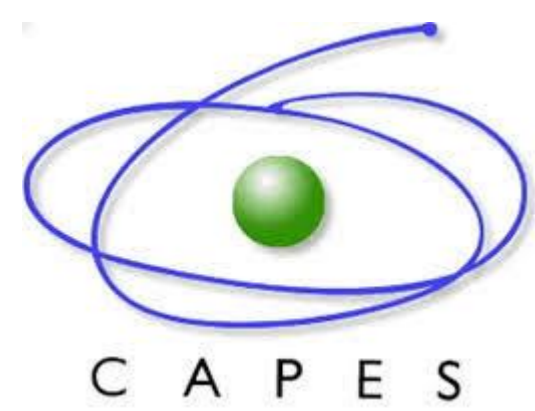


“Sentimento de quê?”

uma lista de sentimentos para a Análise de Sentimentos



Barbara C. Ramos (PUC-Rio)

Cláudia Freitas (PUC-Rio)



Introdução

- **Objetivo:** criar subsídio para a detecção automática de emoções e sentimentos em grandes coleções de texto
- **Desafio:** o que conta como sentimento? O que é um sentimento? Disputa teórica: universalidade vs dependência linguística/cultural
- **Nossa proposta:** investigar o que os falantes de português materializam como sentimento. Deslocamos para os falantes de português a responsabilidade sobre a identificação de algo como sendo uma emoção ou sentimento, e evitamos também a baixa concordância entre anotadores.

Método (1): Exploração em grandes corpora (AC/DC)

Buscas com o padrão léxico-gramatical:

“sentimento de N”: [lema="sentimento"] [lema="de"]
[pos!="V"]* @[pos="N" & func="P<"] within s

Resultado: lista com **2060** lemas diferentes; análise manual de **853** lemas (>3)

Método (2): Pré-organização dos lemas em grupos

- **Grupo 1:** Sentimentos que consideramos convencionais, como *culpa*, *medo* etc:
“*Luís Garcia não pôde furtar-se a um sentimento de pena, ao vê-lo entrar fardado e prestes a seguir para o Sul*”
- **Grupo 2:** Maneiras não convencionais de falar sobre sentimentos:
“*Um sentimento de nacionalidade surgiu nos corações, vago e confuso, mas enérgico*”
- **Grupo 3:** Não se refere a um sentimento, mas ao possuidor do sentimento:
“*Tais eram as idéias e os sentimentos de d. Damiana*”
- **Grupo 4:** Introduce um modificador do sentimento; o “N” não nomeia um sentimento, apenas atribui a ele qualidades:
“*À intuição está relacionado um determinado sentimento de um estado emocional*”

Referências bibliográficas

- Balage Filho, P.P., Aluisio, S.M. e Pardo, T.A.S. (2013) An Evaluation of the Brazilian Portuguese LIWC Dictionary for Sentiment Analysis. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology – STIL*, pages 215–219.
- Carvalho, P., Sarmento, L., Teixeira, J. e Silva, M. J. (2011) Liars and Saviors in a Sentiment Annotated Corpus of Comments to Political Debates. In *ACL (Short Papers)*, pages 564–568.
- Carvalho, P., Silva, M. (2015) Sentlex-pt: principais características e potencialidades. In *Linguística, Informática e Tradução: Mundos que se Cruzam*, Organizado por Simões, Barreiro, Santos, Sousa-Silva & Tagnin, *Oslo Studies in Language* 7(1):425–438.
- Chen, Y. e Skiena, S. (2014) Building Sentiment Lexicons for All Major Languages. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, Maryland, USA, pages 383–389.
- David Le Breton (2019) *Experiências da dor: uma antropologia*, Palestra ministrada por David Le Breton, da Universidade de Estrasburgo, em 13 de março de 2019, Instituto de Estudos Avançados em Humanidades, PUC-Rio.
- Ekman, P. (1999) *Basic Emotions, Handbook of Cognition and Emotion*. Editado por Tim Dalgleish e Mick Power, John Wiley & Sons, Sussex, UK.
- Freitas, C., Motta, E., Miliú, R. L. e César, J. (2014) Sparking Vampires... IJLI Annotating Opinions in a Book Review Corpus. In *New Language Technologies and Linguistic Research: A Two-Way Road*. Sandra Aluisio e Stella E. O. Tagnin, Cambridge Scholars Publishing, 2014, pages 128–146.
- Hearst, M. (1992) Automatic acquisition of hyponyms from large text corpora. In *Proceedings of the 14th International Conference on Computational Linguistics*, pages 539–545.
- Heuser, R.; Moretti, F.; Steiner, E. (2016) The emotions of London. In *Literary Lab Pamphlets*, <https://litlab.stanford.edu/LiteraryLabPamphlet13.pdf>, Março.
- Klinger, R., Samat, S. S. e Reiter, N. (2016) Automatic Emotion Detection for Quantitative Literary Studies: A case study based on Franz Kafka's “Das Schloss” and “Amerika”. In *Digital Humanities 2016: Conference Abstracts*, pages 826–828, Cracóvia, Polónia.
- Maia, B.; Santos, D. (2018) Language, emotion and the emotions: The multidisciplinary and linguistic background. In *Lang Linguist Compass*.
- Mohammad, S. M., e Turney, P. D. (2013) Crowdsourcing a word-emotion association lexicon. In *Computational Intelligence* 29(3):436–465.
- Mota, C. e Santos, D. (2015) Emotions in natural language: a broad-coverage perspective. In *Linguatca*, <https://www.linguatca.pt/aceeso/EmotionsBC.pdf>, Maio.
- Ortony, A., Clore, G. L., e Collins, A. (1988) *The cognitive structure of emotions*. Cambridge and New York: Cambridge University Press.
- Plutchik, R. (1962) *The Emotions: Facts, Theories, and a New Model*. Random House Inc, EUA.
- Plutchik, R. (2001) “The Nature of Emotions”, *American Scientist* 89:344–350.
- Rezende, C.; Coelho, M. (2010) *Antropologia das Emoções*. Editora FGV, Rio de Janeiro.
- Santos, A. G. L., Becker, K. e Moreira, V. (2014) “Um estudo de caso de mineração de emoções em textos multilíngues”, *III Brazilian Workshop on Social Network Analysis and Mining (BraSNAM 2014)*.

RESULTADO E DISCUSSÃO

- Lista de 742 palavras de sentimento:

		Corpus		
		Obras	Todos exceto Obras	Todos
Grupo 1	sema = emo	33 (62%)	76 (35%)	109 (40%)
	sema ≠ emo	20 (38%)	141 (65%)	161 (60%)
Grupo 2	sema = emo	0%	26 (6,1%)	26 (5,5%)
	sema ≠ emo	49 (100%)	398 (93,9%)	446 (94,5%)
Grupo 3		11	59	70
Grupo 4		07	24	31

- A distribuição das palavras de sentimento por campos lexicais pouco mencionados na literatura:

Campo lexical 1: lemas referentes ao nacionalismo.

Sentimento de....

americanismo – anticomunismo – antilhanía – anti-lusitanismo – bairrismo – brasilidade – cidadania – civismo – democracia – germanidade – germanismo – identidade – italianidade – mineiridade – nação – nativismo – nacionalização – nacionalidade – nacionalismo – origem – pátria – patriota – patriotismo – pertença – pertencimento – povo – raça – ufanía – ufanismo – união – xenofobia.

Campo lexical 2: lemas referentes à espiritualidade. Sentimento de...

abstenção – adoração – católico – cobiça – comunhão – devoção – doação – fanatismo – fé – idolatria – perdão – religião – religiosidade

Campo lexical 3: lemas referentes ao parentesco. Sentimento de...

avó – irmandade – irmão – mãe – maternidade – orfandade – pai – paternidade – viúvo

Desafios Futuros

- Desenvolver estratégias para lidar com a cauda longa (frequência <3);
- Organizar a lista de palavras e adicioná-la à anotação do AC/DC;
- Encaixar as palavras encontradas nos clusters existentes e justificar a criação de novos clusters;
- Detectar automaticamente sentimentos e emoções em grandes coleções de textos em língua portuguesa.

- Santos, D. e Bick, E. (2000) “Providing Internet access to Portuguese corpora: the AC/DC project”, *Proceedings of the Second International Conference on Language Resources and Evaluation*, Edited by Maria Gavrilidou, George Carayannis, Stella Markantonatou, Stelios Piperidis e Gregory Stainhauer, LREC 2000, pages 205–210.
- Santos, D., Freitas, C. e Bick, E. (2018) Obras: a fully annotated and partially human-revised corpus of Brazilian literary works in the public domain. In *OpenCor*, Canela, RGS, Brasil.
- Souza, M.; Vieira, R.; Buseti, D.; Chishman, R. E Alves, I. M. (2012) Construction of a Portuguese Opinion Lexicon from multiple resources. In *8th Brazilian Symposium in Information and Human Language Technology*. Tausczik, Y., Pennebaker, J. (2010) The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. In *Journal of Language and Social Psychology*, 29(1), pages 24–54.
- Tomkins, S. S. (1962) *Affect Imagery Consciousness: Volume I, The Positive Affects*, London: Tavistock Publications.
- Tomkins, S. S. (1963) *Affect Imagery Consciousness: Volume II, The Negative Affects*, London: Tavistock Publications.
- Wiebe, J., Wilson, T., e Cardie, C. (2005) Annotating expressions of opinions and emotions in language. In *Language resources and evaluation*, 39(2-3), pages 65–210.
- Wierzbicka, A. (1999) *Emotions Across Languages and Cultures: Diversity and Universals*, pages 1–48, Inglaterra: Cambridge University Press.