

Towards discourse annotation and sentiment analysis of the Basque Opinion Corpus

Workshop on Discourse Relation Parsing and Treebanking
(NAACL-HLT 2019)

Jon Alkorta, Koldo Gojenola & Mikel Iruskieta

IXA NLP group. University of the Basque Country (UPV/EHU)

Minneapolis, Minnessota (USA), 6th June, 2019



University of the Basque Country

IXA GROUP



Outline

- 1 Introduction and Related Works
- 2 Theoretical framework and methodology
- 3 Results and discussion
- 4 Conclusion and Future Work

Introduction

- Aims of sentiment analysis:
 - i) **Document level sentiment classification.** A positive or negative evaluation [Pang et al., 2002, Turney, 2002].
 - ii) **Subjectivity classification at sentence level.** A subjective or objective (factual) sentence [Wiebe et al., 1999].
 - iii) **Aspect and entity level.** Identification of the target of one positive or negative opinion [Hu and Liu, 2004].

- Apart from basic resources, **a corpus with subjective information** for sentiment analysis is indispensable.

Examples:

- **Linguistic knowledge**: analysis different linguistic phenomena related to sentiment analysis.
- **Statistic analysis**: extraction of patterns of different linguistic phenomena.

The aim of this work

- Annotate the rhetorical structure of an opinionated corpus in Basque to check out the semantic orientation of rhetorical relations.

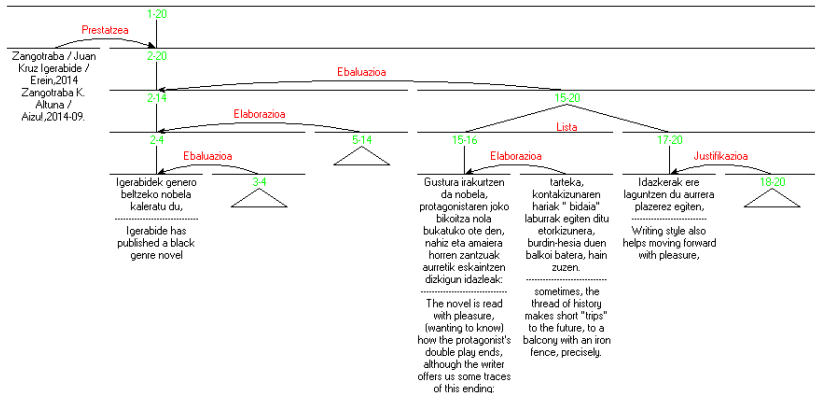
Related works

Author	Theory	Corpus	Annotation	Results
[Refaee and Rieser, 2014]	-	8,868 tweets in Arabic	Semantic orientation Grammatical features	Kappa: 0.84
[Chardon et al., 2013]	SDRT	211 texts (movie reviews, news reactions)	EDUs: subjectivity. Documents: subjectivity and discourse relations	Kappa. EDUs: 0.69, 0.44 Documents: 0.73, 0.58
[Asher et al., 2009]	SDRT	+300 texts (movies, letters, reports)	Discourse and subjectivity annotation	Categorization: 95% Segmentation: 82%
[Mittal et al., 2013]	-	662 reviews in Hindi	Violating expectation conjunctions. Negation.	Discourse + negation, the accuracy: 50.45 to 80.21.

Outline

- 1 Introduction and Related Works
- 2 Theoretical framework and methodology
- 3 Results and discussion
- 4 Conclusion and Future Work

Theoretical framework: Rhetorical Structure Theory (RST)



The Basque Opinion Corpus

- 240 opinion texts collected from different websites.
- Opinion texts of six different domains: sports, politics, music, movies, literature books and weather.
- Usefulness for sentiment analysis:
 - The first person: 1.21% in a Basque objective corpus (Basque Wikipedia) vs. 8.37% in the Basque Opinion Corpus.
 - 8.50% of the words correspond to adjectives in Basque Wikipedia and 9.82% in the corpus for study.
 - Negation, *irrealis blocking* and discourse markers also are in the corpus.

Methodology steps

1- Set the stage for the annotating work.

	A1	A2	Total
Movie	21 + 9	9	30
Weather	10 + 5	5	15
Literature	5	20 + 5	25
Total	50	39	70

2- Annotation procedure and process.

- Following the annotation guidelines proposed by [Das and Taboada, 2018].
- Weather texts were annotated in 20 minutes while movie and literature texts were annotated in one hour.

3- Measurement of inter-annotator agreement.

Inter-annotator agreement was measured in two ways:

- The qualitative evaluation method [Iruskieta et al., 2015] using F-measure.
- In contrast with the qualitative evaluation, the manual evaluation did not take the central subconstituent factor into account.

4- Semantic orientation extraction.

- Use of the Basque version of the SO-CAL tool [Taboada et al., 2011].
- Extraction of the sentiment valence of 75 instances of CONCESSION and EVALUATION relations.

5- Results.

- Percentage of rhetorical relations with the same label annotated by two persons.
- Accumulated values of sentiment valences in nuclei and satellites in texts of different domains.

Outline

- 1 Introduction and Related Works
- 2 Theoretical framework and methodology
- 3 Results and discussion**
- 4 Conclusion and Future Work

RST annotation: inter-annotator agreement

- Type of rhetorical relation.

Domain	Agreement (%)	Agreement (RR)
Weather	43.59	17 of 39
Literature	41.67	70 of 168
Movies	37.73	83 of 220
Total	39.81	170 of 427

Sentiment analysis: sentiment valence of rhetorical relations

- We sum all the sentiment valence of words of CONCESSION and EVALUATION rhetorical relations.
- The results of the sum are given based on nuclearity.

Sum of sentiment valences	CONCESSION		EVALUATION	
	Nucleus	Satellite	Nucleus	Satellite
Weather	39.41	39.75	49.86	33.35
Literature	61.02	68.73	53.13	80.30
Movies	13.98	19.45	26.01	45.58
Total	114.41 (47.21 %)	127.93 (52.79 %)	128.99 (45.00%)	159.23 (55.00%)

- CONCESSION.

- [S[Puntu ahulak izan arren,]_{-1.5} N[filma erakargarri eta berezia da Victoria.]]₊₆]_{+4.5} (ZIN19)
[S[Although it has weak points,]_{-1.5} N[Victoria is an entertaining and special movie.]]₊₆]_{+4.5}

- EVALUATION.

- [N[Bada, erraz ikusten den filma da "The danish girl"].]₊₁ S[Atsegina da, hunkigarria, entretenigarria]₊₆]₊₇ (ZIN15).
[N[So, "The danish girl" is a film easy to watch.]]₊₁ S[It is nice, touching, entertaining.]]₊₆]₊₇

RST annotation: inter-annotator agreement

- Automatic evaluation in a more strict scenario (if and only if the central subconstituent is the same) following [Iruskieta et al., 2015]
 - **Constituent (C)**. All the EDUs that compose each discourse unit or span.
 - **Attachment point**. The node in the RS-tree to which the relation is attached.
 - **N-S or nuclearity** Specification of the compared relations regarding direction (NS, NS or NN).
 - **Relation**. The same type of rhetorical relation to the attachment point of two or more EDUs in order to get the same effect.

- Results according to automatic evaluation concerning discourse annotation.

Domain	Constituent		Attachment		N-S		Relation	
	Match	F1	Match	F1	Match	F1	Match	F1
Weather	20/37	0.54	9/37	0.24	22/37	0.59	15/37	0.41
Literature	84/155	0.54	67/155	0.43	105/155	0.68	48/155	0.31
Movies	112/221	0.56	88/221	0.40	147/221	0.67	68/221	0.31
Total	216/413	0.52	164/413	0.40	274/413	0.66	131/413	0.32

Discussion: relevant RR disagreement

A1	A2		
RRs		#	Total
ELABORATION	MOTIVATION	9	19
ELABORATION	INTERPRETATION	6	
RESULT	ELABORATION	4	
INTERPRETATION	JUSTIFICATION	4	4
CONCESSION	CONTRAST	6	14
EVALUATION	CONTRAST	4	
LIST	CONJUNCTION	4	

Usefulness of the corpus for sentiment analysis

- We can combine the subjectivity information with features of type of rhetorical relations to make a better sentiment analysis and classification.
- 1) Subjectivity extraction: words with sentiment valence tend to appear more in satellites than in nuclei.

Type of RR	Nucleus	Satellite
CONCESSION	situation affirmed by author	situation which is apparently inconsistent but also affirmed by author
EVALUATION	a situation	an evaluative comment about the situation

2) Discourse information.

- **CONCESSION.**

- **Result:** The semantic orientation of nucleus must be the semantic orientation of all the rhetorical relation.

- **EVALUATION.**

- **Result:** The weight must be assigned to the satellite because that part of the relation is more important.

Outline

- 1 Introduction and Related Works
- 2 Theoretical framework and methodology
- 3 Results and discussion
- 4 Conclusion and Future Work

Conclusions

- Inter-annotator agreement.
 - Annotation of a part of the Basque Opinion Corpus using RST.
 - The inter-annotator agreement: 39.81%.
 - The results of automatic tool regarding constituent and nuclearity are higher than 0.5 (inter-annotator agreement).
- The usefulness of the corpus for sentiment analysis.
 - Useful to extract subjectivity information of different rhetorical relations.
 - CONCESSION: the semantic orientation of the nucleus prevails.
 - EVALUATION: words with sentiment valence concentrate on satellite.

Future Work

- Building of extended annotation guidelines to annotate the corpus with more reliability.
- Annotation of the entire corpus.
- Analysis regarding the distribution of the subjective information in relations.

Any question?

Gracias
Eskerrik
Thanks asko

Towards discourse annotation and sentiment analysis of the Basque Opinion Corpus

Workshop on Discourse Relation Parsing and Treebanking
(NAACL-HLT 2019)

Jon Alkorta, Koldo Gojenola & Mikel Iruskieta

IXA NLP group. University of the Basque Country (UPV/EHU)

Minneapolis, Minnessota (USA), 6th June, 2019







University of the Basque Country

IXA GROUP



References I

-  Asher, N., Benamara, F., and Mathieu, Y. Y. (2009).
Appraisal of opinion expressions in discourse.
[Lingvisticæ Investigationes](#), 32(2):279–292.
-  Chardon, B., Benamara, F., Mathieu, Y., Popescu, V., and Asher, N. (2013).
Measuring the effect of discourse structure on sentiment analysis.
In [International Conference on Intelligent Text Processing and Computational Linguistics](#), pages 25–37. Springer.
-  Das, D. and Taboada, M. (2018).
RST Signalling Corpus: A Corpus of Signals of Coherence Relations.
[Lang. Resour. Eval.](#), 52(1):149–184.
-  Hu, M. and Liu, B. (2004).
Mining and summarizing customer reviews.
In [Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining](#), pages 168–177. ACM.

References II



Iruskieta, M., Da Cunha, I., and Taboada, M. (2015).

A qualitative comparison method for rhetorical structures: identifying different discourse structures in multilingual corpora.

[Language resources and evaluation](#), 49(2):263–309.



Mittal, N., Agarwal, B., Chouhan, G., Bania, N., and Pareek, P. (2013).

Sentiment Analysis of Hindi Reviews based on Negation and Discourse Relation.

In [Proceedings of the 11th Workshop on Asian Language Resources](#), pages 45–50.



Pang, B., Lee, L., and Vaithyanathan, S. (2002).

Thumbs up?: sentiment classification using machine learning techniques.

In [Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10](#), pages 79–86. Association for Computational Linguistics.



Refaee, E. and Rieser, V. (2014).

An Arabic Twitter Corpus for Subjectivity and Sentiment Analysis.

In [LREC](#), pages 2268–2273.

References III



Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M. (2011).
Lexicon-based methods for sentiment analysis.
[Computational Linguistics](#), 37(2):267–307.



Turney, P. D. (2002).
Thumbs up or thumbs down?: semantic orientation applied to unsupervised
classification of reviews.
In [Proceedings of the 40th annual meeting on association for computational
linguistics](#), pages 417–424. Association for Computational Linguistics.



Wiebe, J. M., Bruce, R. F., and O'Hara, T. P. (1999).
Development and use of a gold-standard data set for subjectivity classifications.
In [Proceedings of the 37th annual meeting of the Association for Computational
Linguistics](#), pages 246–253.