THE INFLUENCE OF COHERENCE RELATIONS ON A SENTIMENT OF A DISCOURSE (BASED ON FRENCH NEWS ARTICLES)

Loda Sylvette
National Technical University
"Kharkiv Polytechnic Institute",
Kharkiv, Pushkinskaya str., 79/2, tel. 707–63–60,
e-mail: simplement.sy@gmail.com

Coherence is currently a topic of intense debate in the international linguistic community, as well as automatic discourse analysis. The possibility of automatic semantic and sentiment analysis is very important in modern informational world. Now that we have an access to large amounts of information coming every day, it's becoming more and more complicated to analyze it.

Nowadays computational linguistics can operate various methods of sentiment and semantic discourse analysis, depending on the data they are applied to. But there is something that unites all of the topics and texts, here I am talking about connectives that are used to mark coherence relations between discourse segments. In my current research I want to address a question of a possible influence of connectives on semantic meaning and an overall sentiment of a text, narrowing the scope of research to financial news in French language as an example. Although it is a well-known fact that research into coherence strategy and automatic language processing is considered relevant to all spheres of human communication.

A text is coherent if it is designed around a common topic. In the reading process, the individual units of information enter meaningful relationships to one another. The text coheres and is not just a sequence of sentences to be interpreted in isolation. Those individual units that are united by specific relationships are called elementary discourse units (EDU). An EDU is a span of text, usually a clause, but in general ranging from minimally a Noun Phrase to maximally a sentence, that denotes a single event or type of event, serving as a complete, distinct unit of information that the subsequent discourse may connect to. This connection is performed by a certain coherence relation, which is a specific relationship, holding on the semantic or the pragmatic level of description, between adjacent units of text.

Coherence relations are usually explicitly marked with connectives, which belong to a closed-class non-inflectable lexical items. Their key property is their relational meaning: connectives set two discourse segments into correspondence with each other. From the semantic viewpoint, they therefore denote two-place relations. Syntactically, a connective can be a subordinating or coordinating conjunction, an adverbial, or (arguably) a preposition.

In my research I decided to define main classes of connectives with respect to their influence on a sentiment and semantic meaning of a discourse. The main idea was to make the classification of connectives as meaningful for practical use as possible. So basically we can identify the following groups of connectives according to their properties:

1 - connectives, changing sentiment of a nucleus sub-sentence and an overall sentiment of a phrase

E.g.: [L'abondance des sources alternatives de pétrole et les permis d'exploiter accordés en masse en Irak vont doper la production de pétrole]{S; pos}, mais [son prix restera élevé]{N, neg}. Overall sentiment: negative.

2 - connectives, influencing only satellite sub-sentence

E.g.: [Le dollar US a affiché son mois le plus bas depuis septembre]{N, neg} malgré que [Janet Yellen ait déclaré que la Banque Centrale va probablement maintenir sa politique monétaire]{S; pos}. Overall sentiment: negative.

3 - connectives that do not influence the sentiment of a span of discourse at all

E.g.: [Oddo réitère son opinion 'neutre' sur Aéroports de Paris]{pos} et [il relève l'objectif de cours à 88 euros]{pos}.

4 - connectives, that introduse a EDU giving different kinds of additional information (cause, consequence, result, time, place, example, comparison etc.)

E.g.: [Les ventes de camions légers ont grimpé de 9,7% en février, à 64 579]{pos}, alors qu'[elles avaient été de 58 867 en février 2013]{comparison}.

These groups are created out of a lexical base of French connectives which I observed in my previous course paper, containing 328 connectives. The new classification has a very important practical use in automatic discourse analysis and natural language processing, as it allows to enhance the effectiveness of sentiment and semantic analysis of texts.

Bibliography

1. Stede, Manfred (2011), Discourse Processing, Potsdam : Morgan & Claypool Publishers.

2. Charlotte Roze, Laurence Danlos (2009), Base lexicale des connecteurs discursifs du français, Paris: Université Paris Diderot.

3. Busquets, Joan (2013), Analyse du discours,Relations de cohérence, Bordeaux: Universit´e Bordeaux-3.

4. Bateman, John and Judy Delin (2005), « Rhetorical structure theory », In Encyclopedia of Language and Linguistics (2nd ed., pp. 589-596), Oxford : Elsevier.

5. Mann,William C. and Sandra A. Thompson (1986), Rhetorical Structure Theory : Description and Construction of Text Sructures (Technical Report No. ISI/RS- 86-174), Marina del Rey, CA : Information Sciences Institute.

6. Hobbs, Jerry, (1985). On the Coherence and Structure of Discourse (Research Report 85-37), Stanford, CA : CSLI.

7. Kehler, Andrew (2002), Coherence, Reference, and the Theory of Grammar, Stanford, CA : CSLI.