

Elena Cabrio, Alessandro Mazzei and Fabio Tamburini (dir.)

Proceedings of the Fifth Italian Conference on Computational Linguistics CLiC-it 2018
10-12 December 2018, Torino

Accademia University Press

Lexical Opposition in Discourse Contrast

Anna Feltracco, Bernardo Magnini and Elisabetta Jezek

DOI: 10.4000/books.aaccademia.3334
Publisher: Accademia University Press
Place of publication: Torino
Year of publication: 2018
Published on OpenEdition Books: April 8, 2019
Series: Collana dell'Associazione Italiana di Linguistica Computazionale
Electronic EAN: 9788831978682

OpenEdition Books

http://books.openedition.org

Electronic reference
FELTRACCO, Anna ; MAGNINI, Bernardo ; and JEZEK, Elisabetta. *Lexical Opposition in Discourse Contrast* In: *Proceedings of the Fifth Italian Conference on Computational Linguistics CLiC-it 2018: 10-12 December 2018, Torino* [online]. Torino: Accademia University Press, 2018 (generated 07 octobre 2023). Available on the Internet: <http://books.openedition.org/aaccademia/3334>. ISBN: 9788831978682. DOI: https://doi.org/10.4000/books.aaccademia.3334.

The text only may be used under licence . All other elements (illustrations, imported files) are “All rights reserved”, unless otherwise stated.

Lexical Opposition in Discourse Contrast

Anna Feltracco
Fondazione Bruno Kessler
University of Pavia, Italy
University of Bergamo, Italy
anna.feltracco@gmail.com

Bernardo Magnini
Fondazione Bruno Kessler
Trento, Italy
magnini@fbk.eu

Elisabetta Jezek
University of Pavia
Pavia, Italy
jezek@unipv.it

Abstract

English. We investigate the connection between lexical opposition and discourse relations, with a focus on the relation of contrast, in order to evaluate whether opposition participates in discourse relations.[1] Through a corpus-based analysis of Italian documents, we show that the relation between opposition and contrast is not crucial, although not insignificant in the case of implicit relation. The correlation is even weaker when other discourse relations are taken into account.

Italiano. *Studiamo la connessione tra l'opposizione lessicale e le relazioni del discorso, con attenzione alla relazione di contrasto, per verificare se l'opposizione partecipa alle relazioni del discorso. Attraverso un'analisi basata su un corpus di documenti in italiano, mostriamo che la relazione tra opposizione e contrasto non è cruciale, anche se non priva di importanza soprattutto per i casi di contrasto implicito. La correlazione sembra più debole se consideriamo le altre relazioni del discorso.*

1 Introduction

This paper focuses on lexical opposition and discourse contrast. We define opposition as the relation between two lexical units that contrast with each other with respect to one key aspect of their meaning and that are similar for all the other aspects (e.g. to increase / to decrease, up / down). On the other end, we consider discourse contrast as the relation between two parts of a coherent sequence of sentences or propositions (i.e., discourse arguments) that are in conflict. Both opposition and contrast hold between contrasting elements: the first at the lexical level, the other at the discourse level.

In the following example, a contrast relation is identified between the two arguments in square brackets; two opposite terms are found in the arguments of the relation and are underlined.

(1) [The price of this book increased], while [the price of that one decreased.]

Despite the two relations are *per se* independent, the example shows how opposition can participate in contrast; in fact, the opposites *to increase / to decrease* convey the difference based on which the two mentioned entities (i.e., the books) are compared, leading to a contrast.

Indeed, opposition can be found in the context of other discourse relations (e.g. in the temporal relation "Before the decrease of the demand, an increase of the prices was registered"), and discourse contrast can be conveyed through other strategies (e.g. negation and synonyms "Although the price decreased; the demand did not fall" or incompatibility "She has blue eyes, he has green eyes").

However, our analysis focuses on opposition and contrast, and starts with the observation that both linguistic phenomena involve two elements that are similar in many aspects, but that differ in others (Section 2). This similarity have already been considered by works in the computational field, in which opposition is used as a feature for identifying contrast, and viceversa (Section 3). In this paper, we investigate the behaviour of opposition in the context of a contrast relation adopting a corpus-based approach (Section 4). In particular, we study the opposition-contrast intersection by observing how frequently opposites are found in the arguments of a contrast relation in Contrast-Ita Bank (Feltracco et al., 2017), a corpus anno-

[1] Part of this research has already been published in the first author Ph.D. thesis (Feltracco, 2018).

tated with the discourse contrast relation. We analyze the cases in which the two phenomena co-occur, in order to understand the contribution of opposition to discourse contrast (Section 5). The investigation lead us to enrich Contrast-Ita Bank with lexical opposition. Enlarging our focus, we also investigate the behaviour of opposition in the context of other discourse relations in the corpus, by examining which are the relations that involve pairs of opposites in their arguments (Section 6). Finally, we report our concluding observations and our hint for further work (Section 7).

2 Lexical Opposition and Discourse Contrast

Our definition of opposition in mainly based on the study of Cruse (1986): according to the author, opposition indicates a relation between two terms that *differ along only one dimension of meaning: in respect to all other features, they are identical* (Cruse, 1986, p.197). Examples of opposition are: *to pass / to fail* or *up - down*. In fact, both *to pass / to fail* refer to the result of an examination, but they describe two possible opposite results. Similarly, both *up / down* potentially describe positions with respect to a reference point, the first refers to a higher position, the latter to a lower position.

This definition has some overlap with those proposed for discourse contrast in two of the most important frameworks focused on the study of discourse relations: Rhetorical Structure Theory (Mann and Thompson, 1988) and Segmented Discourse Representation Theory (Asher and Lascarides, 2003). In these theories, the relation of contrast captures cases in which the arguments in the relation *have some aspects in common* (Mann and Thompson, 1988; Carlson and Marcu, 2001), or *have a similar structure* (Asher, 1993), but they *differ in some respect* (i.e., *contrasting themes* (Asher, 1993)) and *are compared with respect to these differences* (Mann and Thompson, 1988). These definitions are consistent with the Penn Discourse Treebank (PDTB) (Prasad et al., 2007) for the sense tag CONTRAST, which is assigned when the arguments of a relation "share a predicate or a property and the difference between the two situations described in the arguments is highlighted with respect to the values assigned to this property" (Prasad et al., 2007, p. 32).

Both opposition and discourse contrast thus involve comparing two elements that are similar in many aspects, but that differ in others; this holds at the lexical level for opposition and at the discourse level for contrast.

3 Opposition and Contrast in NLP

In the area of NLP, the co-occurrence of the opposition and contrast has been considered, for instance, by Roth and Schulte Im Walde (2014), who use what they call *discourse markers* that typically signal a discourse relation, e.g.*but*, for distinguishing paradigmatic relations, including opposition.

Other contributions in the same area use lexical opposition as feature for detecting contrast. As an example, Harabagiu et al. (2006) base the identification of contrast on the opposition relation, given that in some examples "[..] the presence of opposing information contributes more to the assessment of a CONTRAST than the presence of a cue phrase", such as *but* or *although* (Harabagiu et al., 2006).

Marcu and Echihabi (2002) create a system to identify relations of contrast under the hypothesis that some lexical item pairs can "provide clues about the discourse relations that hold between the text span in which the lexical items occur". In a cross-lingual evaluation for English and Swedish, Murphy et al.(2009) show that opposites (*antonyms* in their terminology) are used for different functions: the most common is the one of "creat[ing] or highlight[ing] a secondary contrast within the sentence/discourse".

On the contrary, Spenader and Stulp (2007) give evidence that opposition is not a strong feature for contrast. In particular, they calculate the co-occurrence of opposite adjectives in the contrast relations marked or non-marked by *but* in a corpus. The authors show that opposition is not common in cases of explicit contrast conveyed by *but*, and it is also not very frequent in cases of non-*but* marked contrast. In a similar way, we intend to evaluate whether opposition is a key feature for contrast, or for other discourse relations.

4 Annotating Opposites in Contrast Relations

We carry on our investigation in Contrast-Ita Bank (CIB) (Feltracco et al., 2017)[2], a corpus of 169 Italian documents manually annotated with 372 contrast relations, following the schema proposed

[2] https://hlt-nlp.fbk.eu/technologies/contrast-ita-bank

in the Penn Discourse Treebank. As in the PDTB, the schema in CIB accounts for the identification of *Arg1* and *Arg2*, the two arguments that are compared in a contrast relations. In CIB, two types of contrast are annotated: i) CONTRAST (138 relations), when one the two arguments is similar to the other in many aspects but different in one aspect for which they are compared, and ii) CONCESSION (272 relations), when one argument is denying an expectation that is triggered from the other.[3] CIB accounts for both explicit relations (341) marked by a lexical element (i.e. *connective*, e.g. *but*, *however*) and implicit relations (31).

To evaluate the role of opposition in the context of a contrast, we manually annotated two opposites *opposite1* and *opposite2*, when the former is part of *Arg1* and the latter is part of *Arg2*. For instance, in Example 1 "The price of this book increased" is *Arg1* and "the price of that one decreased" is *Arg2*, and we marked 'increased" as *opposite1* and "decreased" as *opposite2*.

In this manual exercises, we did not limit our annotation to prototypical opposites (Cruse, 1986, p. 262) or to pairs of mono-token words (typically entries of lexical resources), but we manually marked also larger expressions, including cases similar to Example 2.

(2) [Andrew Smith ha rassegnato le dimissioni ieri], nonostante [i tentativi del premier Tony Blair di convincerlo a rimanere]. [4]

In the example, the light-verb construction *rassegnare le dimissioni* (Eng.'to resign') is considered as the opposite of *rimanere* (Eng.'to remain') and the two are found respectively in the two arguments of the contrast relation, conventionally reported in square brackets.

Furthermore, we include in the annotation also 'opposites in context', that is, pairs of terms that are not intuitively considered opposite but are in an opposition relation in the specific context in which they appear, as it happens in Example 3.

(3) [Sul Nuovo Mercato, Tiscali perde lo 0.05% a 2,23], [E. Biscom sale dell'1,09% a 41,44]. [5]

The two terms *perdere* and *salire* (Eng. 'to lose x', 'to fall by x') are semantically opposite in the specific context of Example 3: they are used in their sense of 'loosing (some value)' and 'increasing (of some value)'.

5 Results of the Annotation

We study the connection between opposition and contrast observing the co-occurrence of the two linguistic phenomena and analyzing whether opposition participates in creating contrast.

5.1 Co-occurrence of the two relations

Out of the 372 contrast relations annotated in CIB, we identified a total of 23 cases in which opposites are present in the arguments of a contrast relation[6].

Table 1 shows that opposition is present both when contrast is conveyed explicitly by mean of a connective (as by *nonostante* in Example 2), and when there is no such element (Example 3); however, there is a higher occurrence when the relation is implicit (16% vs 5.2%). With respect to the types of opposition, it occurs both when CONTRAST or CONCESSION have been marked (Examples 3 and 2 respectively), but it is more frequent with the type CONTRAST (9.2% vs 2.5%).

Senses	Types		tot	% over tot
	Explicit	Implicit		
Contrast	7	4	11	9.2%(/102)
Concession	6	0	6	2.5% (/234)
Both	5	1	6	16.6% (/36)
tot	18	5	**23**	
% over tot	5.2%(/341)	16% (/31)		

Table 1: Opposition in discourse contrast in CIB.

5.2 The role of opposition

We conducted a deeper investigation in order to evaluate whether the opposites in the arguments of a contrast relation actually contribute to it.

In Example 4 opposition triggers the contrast relation.

(4) [uno dei due è ricco di cellule staminali], [l' altro ne è povero].[7]

In this case (and in Examples 2 and 3), the contrast relation holds because two entities (e.g. 'one', 'the other') that share a property (i.e. 'to

[3]The presence of one type of relation does not exclude the other.

[4]Eng.:[Andrew Smith resigned yesterday,] despite [Prime Minister Tony Blair's attempts to persuade him to stay.]

[5]Eng.:[On the New Market, Tiscali looses 0.05% to 2.23], [E. Biscom rises by 1.09% to 41.44].

[6]We manually recognized 20 relations; other 3 were identified *ad posteriori* applying the methodology described in Section 6.

[7]Eng.:[one is rich in stem cells],[the other is poor of them.]

have stem cell') are compared with respect to different values that this property takes (i.e. 'to be rich of them', 'to be poor of them'): the different values can be expressed through opposites (i.e. *ricco*/*povero*).

Other examples includes case in which the contrast relation stem from a comparison between two values of a property assigned to the same entity, as happens for the example in Example 5.

(5) Il commercialista [doveva essere il cavaliere bianco chiamato a salvare la Chini] e, invece, [è stato quello che l' ha affossata].[8]

In the example, the contrast arises from the comparison between the opposite roles of the participant: *to save (something) / to ruin (something)*.

Opposition is central for the discourse contrast in these examples. This is not the case for Example 6, for which the opposition does not act as a source for the discourse contrast relation.

(6) [A dispetto degli sforzi della pubblica amministrazione..], [gli investimenti privati in termini di istruzione sono ancora bassi.][9]

In the example, the opposite adjectives *pubblico / privato* (Eng. 'private / public') are attributes of two entities involved: one can say that the participants do have opposite characteristics. However, the contrast relation does not stem from this opposition; rather, it is based on the comparison between the 'positive efforts' on the one hand and the 'low investments' on the other hand.

Out of 23 cases, in 17 opposites are crucial for the contrast relation while in 6 they do not affect the contrast relation. It seems that when opposites appear in the context of a contrast relation they frequently contribute to the phenomena.

We also performed an inter annotator agreement exercises among two annotators to understand whether to distinguish cases in which opposition contributes in conveying the discourse relation (and cases in which they do not) is an easy operation.[10] We register disagreement in 3 cases out of 20, that corresponds to a Dice's coefficient of 85%. After a reconciliation step, in which annotators compared their annotations, and could revise their decisions, two cases were solved, while a third, reported in Example 7 remained.

(7) [A decorrere da domenica 12 entra in vigore il nuovo orario invernale per il servizio extraurbano e la Trento - Malè.] [Da lunedì 13 entra invece in vigore il nuovo orario invernale 2004 / 2005 per il servizio urbano di Trento e Rovereto.] [11]

In this case, one annotator considered that the contrast among the two situations described in the arguments of the discourse relation originates from the opposites *suburban / urban*. Conversely, the other annotator recognized the different dates of entering into force of the two service (i.e. Sunday 12 vs Monday 13) as the source of the resulting discourse contrast.

6 Opposition and Other Discourse Relations

We performed a further analysis evaluating cases of opposites in other discourse relations. We carried on this investigation inspecting the entire CIB corpus and adopting an external resource in which opposites are registered[12]. We automatically retrieved from the corpus pairs of opposites in a windows of 25 token[13]. We retrieved 152 cases that we manually analyzed considering:

- whether the two opposites appear in their opposite sense (e.g. the verbs *andare* / *tornare* are opposite as far as the first verb is not consider as a modal) - data are reported in the second column of Table 2-, and if so:
- whether they are somehow related in the text or not (e.g. in *è subentrato un fatto nuovo, determinato dal fatto che i vincitori del vecchio regime non..* the two opposites properties are of two unrelated entities while in *proposte ufficiali o ufficiose*, the two opposites are in a coordinating relation) - data are

[8] Eng.: The accountant [was supposed to be the white knight designated to save the Chini] and, on the contrary, [he has been the one that ruined it.]

[9] Eng.: [Despite public administration efforts.], [private investments in terms of education are still low.]

[10] One annotator is an author of this paper, the second annotator, who has some familiarity with linguistic tasks, was provided with simple oral instructions through which we ask her to judge the contribution of the opposites when in the context of a contrast relation. We acknowledge Enrica Troiano for collaborating as second annotator.

[11] Eng.: [Starting from Sunday 12 the new winter timetable for the suburban service and for the Trento - Mal enters into force.][From Monday 13 instead the new winter timetable 2004 / 2005 for the urban service of Trento and Rovereto enters into force.]

[12] Dizionario dei Sinonimi e dei Contrari - Rizzoli Editore, http://dizionari.corriere.it/dizionario_sinonimi_contrari

[13] The number was set observing that opposites were found at a maximum distance of 24 tokens in contrast relations.

reported in the third column of Table 2. If the opposites are related:

- whether they are in the arguments of a discourse relation, as in Example 4 - fourth column of the table.

Total	Opposite sense	Related	In Discourse relation
152	100	72	19

Table 2: Opposition in discourse relations.

Results show that in a large number of pairs the two opposites are not actually used in their opposite sense (52 cases = 152 - 100) or are not related in the text (28 cases = 100 - 72). The opposites are found in the arguments of a discourse relation just in 18 cases (11.8 % of the total), suggesting that lexical opposition is not an indicator for the presence of a discourse relation.

A further analysis brought us to investigate also in which discourse relations opposites are involved, following the PDTB classification.[14] We also investigated if opposition is central for these relations. Data are reported in Table 3.

# opp. per relation	discourse relation	# opp. central per relation
7	Comparison.*Contrast*	1
1	Comparison.*Concession*	1
6	Expansion.*Conjunction*	3
3	Expansion.*Level-of-detail*	1
1	Contingency.*Cause*	1
1	Contingency.*Condition*	1
19		8

Table 3: Number of opposition relations in different discourse relations, and their centrality.

From Table 3, we see that opposition co-occurs with different discourse relations, especially Conjunction, but in a more limited number of cases with respect to contrast.[15]

Moreover, comparing the first and the third column of the table, it can be noticed that, as it happens for discourse contrast (see Section 5.2), opposition is not always contributing to the discourse relation itself, meaning that it does not play central role in conveying the relation. As an example, compare Example 8 in which opposition is judged as central, with Example 9 in which it is not.

(8) Sabato [partenza alle 7.01] ed [arrivo alle 19.36.][16]

(9) [..il gruppo ha proseguito l'opera di riorganizzazione societaria], [mettendo un po' d' ordine nelle partecipazioni non legate al core business delle singole controllate..][17]

In the Conjunction relation of Example 8, the two opposite terms indicate the (opposite) events that are coordinated via the conjunction *e*. In Example 9 (a case of EXPANSION.Level-of-detail relation), the two opposites are somehow related (i.e. the *group* is operating for the *singles* subsidiaries), but they are not central for the relation, which is determined by the two events: *proseguire l'opera* and *mettendo [..] ordine*.

7 Conclusion and Further Work

Through the annotation of opposites in the arguments of contrast relations in Contrast-Ita Bank, we aim at providing new insights over the role of opposition in discourse contrast. Overall, we register 23 cases of opposition over 372 contrast relations in our dataset. This number is not high and one we can expect the number to be higher in a larger dataset. However, this limited number suggests that the presence of opposites is not an impacting feature for the identification of contrast relation in the Italian language. It is, however, quite frequent for *implicit* relations, suggesting that the use of opposition can be a strategy to convey contrast when there is a lack of a connective (such as *but* or *however*) that lexicalizes the relation. Moreover, we show that also the co-occurrence of opposition and other discourse relations is low. Despite, in related work opposition has been used as a feature for identifying contrast, the result of our investigation suggests that opposition does not appear to be a strong informative feature and this can possibly lead to a decrease in precision in the process of identifying contrast (i.e., many false positives are expected).

Further and symmetrical work includes the classification of the phenomena that can lead to contrast.

[14]The complete list of the PDTB 3.0 relations can be found in (Webber et al., 2016).

[15]The data for CONTRAST and CONCESSION are part of the ones reported in Table 1, which consider also multi-token expressions and 'opposites in context'.

[16]Eng.: On Saturday, [departure at 7.01] and [arrival at 19.36.]

[17]Eng.: [.. the group has continued the work of corporate reorganization], [putting some order in the shareholdings that not tied to the core business of the single subsidiaries..]

References

Nicholas Asher and Alex Lascarides. 2003. *Logics of conversation*. Cambridge University Press.

Nicholas Asher. 1993. *Reference to Abstract Objects in Discourse*. Kluwer Academic Publishers, Dordrecht.

Lynn Carlson and Daniel Marcu. 2001. Discourse tagging reference manual. *ISI Technical Report ISI-TR-545*, 54:56.

D Alan Cruse. 1986. *Lexical semantics*. Cambridge University Press.

Anna Feltracco, Bernardo Magnini, and Elisabetta Ježek. 2017. Contrast-Ita Bank: A corpus for Italian Annotated with Discourse Contrast Relations. In *Proceedings of the Fourth Italian Conference on Computational Linguistic (CLiC-it 2017)*.

Anna Feltracco. 2018. *Lexical Opposition and Discourse Contrast: A Data-driven Investigation*. Ph.D. thesis, University of Bergamo.

Sanda Harabagiu, Andrew Hickl, and Finley Lacatusu. 2006. Negation, contrast and contradiction in text processing. In *Association for the Advancement of Artificial Intelligence (AAAI)*, volume 6, pages 755–762.

William C Mann and Sandra A Thompson. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text-Interdisciplinary Journal for the Study of Discourse*, 8(3):243–281.

Daniel Marcu and Abdessamad Echihabi. 2002. An unsupervised approach to recognizing discourse relations. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 368–375. Association for Computational Linguistics.

M Lynne Murphy, Carita Paradis, Caroline Willners, and Steven Jones. 2009. Discourse functions of antonymy: a cross-linguistic investigation of Swedish and English. *Journal of pragmatics*, 41(11):2159–2184.

Rashmi Prasad, Eleni Miltsakaki, Nikhil Dinesh, Alan Lee, Aravind Joshi, Livio Robaldo, and Bonnie L Webber. 2007. The Penn Discourse Treebank 2.0 Annotation Manual.

Michael Roth and Sabine Schulte Im Walde. 2014. Combining word patterns and discourse markers for paradigmatic relation classification. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, volume 2, pages 524–530.

Jennifer Spenader and Gert Stulp. 2007. Antonymy in contrast relations. In *Seventh International Workshop on Computational Semantics*, volume 3, page 100.

Bonnie Webber, Rashmi Prasad, Alan Lee, and Aravind Joshi. 2016. A discourse-annotated corpus of conjoined vps. *LAW X*, page 22.