

# The Analysis of Synonymy and Antonymy in Discourse Relations: An Interpretable Modeling Approach

Asela Reig Alamillo  
Universidad Autónoma del  
Estado de Morelos  
Centro de Investigación en Ciencias  
Cognitivas  
assela.reig@uaem.mx

David Torres Moreno  
Universidad Autónoma del  
Estado de Morelos  
Centro de Investigación en Ciencias  
david.torres@uaem.mx

Eliseo Morales González  
Universidad Autónoma del  
Estado de Morelos  
Centro de Investigación en Ciencias  
eliseo.moralesgon@gmail.com

Mauricio Toledo Acosta  
Universidad Autónoma del  
Estado de Morelos  
Centro de Investigación en Ciencias  
mauricio.toledo@unison.mx

Antoine Taroni  
Universidad Autónoma del  
Estado de Morelos  
Centro de Investigación en Ciencias  
antoine.taroni@phelma.grenoble  
-inp.fr

---

\* Corresponding author.

Action Editor: Ekaterina Shutova. Submission received: 4 August 2022; revised version received: 23 January 2023; accepted for publication: 31 January 2023.

<https://doi.org/10.1162/coli.a.00477>

Jorge Hermsillo Valadez\*

Universidad Autónoma del

Estado de Morelos

Centro de Investigación en Ciencias

jhermsillo@uaem.mx

*The idea that discourse relations are interpreted both by explicit content and by shared knowledge between producer and interpreter is pervasive in discourse and linguistic studies. How much weight should be ascribed in this process to the lexical semantics of the arguments is, however, uncertain. We propose a computational approach to analyze contrast and concession relations in the PDTB corpus. Our work sheds light on the question of how much lexical relations contribute to the signaling of such explicit and implicit relations, as well as on the contribution of different parts of speech to these semantic relations. This study contributes to bridging the gap between corpus and computational linguistics by proposing transparent and explainable computational models of discourse relations based on the synonymy and antonymy of their arguments.*

## 1. Introduction

The interpretation of a discourse as a content unit, instead of as the mere juxtaposition of independent sentences, is possible by virtue of the existence of coherence or discourse relations. Discourse relations (DRs) are an interdisciplinary object of study: On the one hand, using mainly corpus studies, linguistics has recently contributed to our understanding of how these relations are marked in the discourse; on the other hand, research in Natural Language Processing (NLP) and computational linguistics aims chiefly at discourse marker prediction for downstream representation or classification tasks (Atwell, Li, and Alikhani 2021; Bakshi and Sharma 2021; Zeldes et al. 2021; Nie, Bennett, and Goodman 2019).

These disciplinary advances share the theoretical object of study and the use of corpus in the endeavor, yet they very rarely translate into an interdisciplinary dialogue between them that would ultimately broaden our understanding of how discourse coherence is constructed and to what extent the linguistically encoded material contributes to the establishing of different types of discourse relations.

The contributions of this study are twofold. From a theoretical point of view, it is recognized that linguistically encoded material other than connectives contributes to the signaling of DRs, yet it is unclear how much weight is attributed to lexical semantics in this process and whether different types of DRs behave similarly in this respect. We seek to answer the questions of how much lexical relations (specifically, synonymy and antonymy) contribute to the interpretation of explicit and implicit contrast and concession relations and whether different parts of speech (POS) (a.k.a. lexical classes) are of equal importance in signaling these relations, thus contributing to the ongoing debate on the interrelationship between contextual signals and connectives.

At the same time, from a methodological viewpoint, although very valuable proposals have been put forward to manually encode different types of linguistic signals in corpus (Das and Taboada 2019; Crible and Degand 2019; Crible 2022), much less effort is put in developing corpus analysis methodologies that do not rely, or rely

less, on manual analysis and that, instead, take advantage of computational representations of corpus. Hence, we advance toward creating interpretable computational representations, which do not rely on manual coding and which allow us to analyze the lexical–semantic signaling of discourse relations in corpus, thus contributing to provide tools for the benefit of the interface between corpus and computational linguistics.

The rest of the article is organized as follows. In Section 2 we state the problem and formulate the research questions. In Section 3 we propose methods for the computational modeling and analysis of semantic synonymy-antonymy signals in discourse relations of contrast. In Section 4 we present the experimental results after applying our models to the Penn Discourse Treebank-3. We discuss our modeling choices and our results in Section 5, and finally conclude in Section 6.

## 2. Problem Formulation and Research Questions

In this section, we begin by highlighting the need to understand the contribution of lexical semantics in establishing contrast and concession relations. We continue with the motivation to propose a new computational approach to their analysis. We conclude this section by posing the research questions and hypotheses.

### 2.1 Discourse Relations and Lexical Semantics

The processes of establishing suprasentential relations (a.k.a. discourse or coherence relations) between sentences is partially constrained by the linguistic content of the utterance. The most straightforward way in which a specific DR can be linguistically marked or signaled (Taboada 2019) is by means of a connective: Its absence or presence differentiates between implicit and explicit DRs, respectively.

Besides connectives, other elements in the explicit linguistic content are assumed to play a role in the interpretation of discourse relations, thus functioning as cues or signals for the inferring process (Das and Taboada 2018; Crible and Degand 2019; Crible 2022). The conceptual meaning encoded in the discourse segments, and specifically the semantic relations between words, are one of such signals guiding the interpretation of a given discourse relation and interacting with connectives (Das and Taboada 2019; Crible 2022).

It is, for example, intuitively clear that the semantics of *cold* and *warm* is responsible for the contrast reading in (1). In (2) and (3), the semantic content of *rained* and *dry/wet*, together with the interpreter's world knowledge, would lead to a concessive (2) or a causal (3) discourse relation.

- (1) In New York, it's cold today; in Mexico City, it is warm.
- (2) It rained; the streets are dry.
- (3) It rained; the streets are wet.

Negative relations (including contrast and concession causal relations) have been defined as non-basic cognitive relations (Sanders, Spooren, and Noordman 1992). In our examples, although the cognitive ability to establish coherence between sentences would tend to assign a concessive interpretation to (2), the absence of an explicit

connective (*however, but, nevertheless*), guiding the interpreter in this inferring endeavor, makes it more cognitively costly than the same endeavor in (3). This difference is consistently found in the psycholinguistic research and corpus studies find that concessive relations are more often explicitly marked by means of a connective than causal relations (Spooren and Sanders 2008; Xiang and Kuperberg 2015; Xu et al. 2018).

In this study, we are interested in the contribution of lexical semantics and, specifically, synonymy and antonymy, to establishing contrast and concession relations.

The idea that antonymy can signal contrast is present in the literature. Marcu and Echihabi (2002) provided evidence of the role of lexical item pairs as clues in building an unsupervised DR classification system, setting the ground to wonder what the role of lexical patterns is in different relations. Spenader and Stulp (2007) found that antonymy in adjectives seemed to be a source of contrast only in the *but* marked pairs of sentences. Feltracco, Magnini, and Jezek (2018) analyzed the role of conceptual opposition—manually encoded—in contrast relations, and found a low presence of opposites in the arguments of a contrast relation, and a higher occurrence of opposites when the relation is implicit than explicit (16% vs. 5.2%). These results contribute to picturing the role of antonymy in contrast relations, yet the number of occurrences analyzed is low. More recently, in a corpus study with manually annotated discourse signals, Crible (2022) found that semantic relations, mainly antonymy, has a relevant presence in contrast relations (and not in concessive or additive relations), although her study includes only relations with explicit connectives.

The contribution of synonymy to contrast and concession relations is less straightforward, yet its role in constructing discourse coherence is clearly acknowledged: Synonyms are key—together with other anaphoric elements—in establishing topic continuity (Givón 1983), hence contributing to local coherence (Spooren and Sanders 2008; Taboada 2019). We posit that synonymy is a useful feature for representing lexical content in discourse as well as a relevant feature that deserves further attention regarding its contribution to different discourse relations (Lei et al. 2018).

To sum up, the idea that the interpretation of discourse relations is based both on the content explicitly encoded in the discourse and on the common knowledge shared by producer and interpreter, thanks to the human cognitive ability to infer implicit meanings, is pervasive in discourse and linguistic studies. How much weight should be ascribed in this process to the lexical semantics of the arguments is, however, an open question. Likewise, it is unclear to what extent different discourse relations can be characterized and distinguished from other types of discourse relations on the basis of their lexical semantic content, and whether this semantic content also affects the explicit use of a connective.

In order to contribute to answering these questions while advancing in the dialogue between corpus linguistics and computational methodologies, in this article we propose computational representations of concession and contrast discourse relations that capture the contribution of the semantic conceptual content toward the relation.

## 2.2 Constructing Linguistically Informative Computational Representations

Linguistic research using corpora to study DRs has greatly contributed to the identification of a whole range of linguistic signals, covering syntagmatic, morphosyntactic, or semantic features of the sentence (Crible 2022; Taboada 2019). Notably, Prasad, Joshi, and Webber (2010) and Rysová and Rysová (2015) complement the notion of grammaticalized connective, with multiword phrases, also signalling DRs, but syntactically and lexically free, and relying on a broader (pragmatic) context. Nevertheless, these

analyses are based on the systematic manual annotation of linguistic signals in the corpora. Our approach wonders what advances can be made, in the same direction, dispensing with manual annotation. Although many computational approaches have been proposed for the analysis of DRs for predictive and classification purposes (Atwell, Li, and Alikhani 2021; Bakshi and Sharma 2021; Roth and Schulte Im Walde 2014; Lei et al. 2018; Biran and McKeown 2013; Sporleder 2008; Wellner et al. 2006), it is questionable how much these advances have contributed to strengthening the dialogue between NLP and theoretical or corpus linguistics. Indeed, in most cases the features taken into account for the prediction or recognition of a given relation remain unknown to the researcher, and the differences and coincidences between individual types of discourse relations are largely ignored (Lei et al. 2018).

We believe that the dialogue between linguistics and NLP finds a much more fruitful path in the *identification of linguistic patterns* guiding the interpretation of a certain discourse relation. Other studies have advocated a similar approach on DRs (Lei et al. 2018; Taboada and Das 2013) and in NLP in general (Benamara, Taboada, and Mathieu 2017; Boleda 2020). Hence, we propose that lexical semantic relations can be a useful starting point to capture the conceptual meaning of the arguments in contrast and concession DRs, as characterized in the Penn Discourse Treebank (PDTB). By using computational models, our work seeks to understand the occurrence of antonymy and synonymy in contrast and concession relations, and to provide knowledge on the patterns of their co-occurrence with discourse connectives.

### 2.3 Research Questions and Predictions

Our work addresses the following research questions, one methodological and three theoretical:

1. How can discourse relations be computationally modeled in order to capture the contribution of the lexical semantics to the meaning of the discourse relation? We propose to build and analyze interpretable representations of the lexical content of a DR using synonyms and antonyms from the corpus vocabulary for different POS. This should offer answers to the remaining three questions.
2. How much do different POS contribute to the representation of contrast and concession DRs? POS differ in the semantic content that they prototypically represent and the syntactic and discourse functions that they play in the text. Determining what POS, or what combination of them, contributes more to the representation of contrast and concession relations will shed light on how lexical semantics and discourse coherence interact.
3. Are contrast and concession differentiated using these representations? One of the goals of the linguistic analysis and the computational modeling of DRs is being able to set apart different kinds of relations in a corpus. In this sense, it is of interest to determine whether the proposed representations, based on lexical relations, are useful to differentiate contrast and concession relations. Contrast and concession relations are both included under the tag **Comparison** in the PDTB corpus. The tag **contrast** is used when at least two differences between the two

arguments are highlighted, whereas the tag **concession** is used when a causal relation expected from one argument is denied in the other. In principle, this distinction would lead to hypothesize that, in contrast relations, the *differences highlighted between argument 1 and argument 2* would likely be captured by their conceptual meaning (antonyms and/or synonyms in argument 1 and 2), whereas the difference between expected consequences and actual ones would be less closely tied to the conceptual meaning explicitly expressed in the arguments. Based on this idea, we would expect that the representation of discourse relations based on the lexical content of their arguments would set apart contrast and concession relations.

4. Do implicit and explicit discourse relations of contrast and concession behave similarly in terms of these representations? Previous studies on discourse relations exclusively analyze either implicit (Sporleder 2008) or explicit (Crible 2022) DRs, or seem to operate on the implicit idea that the features that characterize explicit discourse relations should be the same features characterizing implicit ones (Biran and McKeown 2013); consequently, in most studies differences are not expected or looked for between the two groups. However, the opposite hypothesis is equally, if not more, plausible taking into account the cognitive process of interpreting coherence: An explicit discourse connective should be expected when the lexical semantic content of the discourse segment contributes less to the discourse relation; in turn, in discourse segments where the lexical semantic content is enough to signal the discourse relation, an implicit connective would be more likely.

### 3. Materials and Methods

In order to find some semantic signaling pattern, we propose to build a representation of each argument of a discourse relation using synonyms and antonyms from the corpus vocabulary, grouping them into 4 lexical classes: nouns, adjectives, verbs, and adverbs. From these argument models, we propose two ways to analyze semantic signaling patterns by connective type in the corpus: First, by constructing a knowledge graph of DRs using these abstractions; and second, by finding synonymy/antonymy relations between these representations. We start by describing the corpus and then the modeling approach.

#### 3.1 The Corpus: PDTB 3.0

The Penn Discourse Treebank 3.0 (PDTB3) is a large-scale corpus annotated with information related to discourse structure and discourse semantics (cf. Webber et al. 2019, for details). While there are many aspects of discourse that are crucial to a complete understanding of natural language, the PDTB3 focuses on encoding discourse relations. The PDTB3 adopts the predicate-argument view of discourse relations, where a discourse connective (e.g., *because*) is treated as a predicate that takes two text spans as its arguments. The argument that the discourse connective structurally attaches to is called **arg<sub>2</sub>**, and the other argument is called **arg<sub>1</sub>**. The PDTB3 provides annotations for explicit and implicit discourse relations, where an explicit relation contains an explicit discourse connective.

We will only consider DRs of type contrast and concession, both explicit and implicit. The PDTB3 has a total of 26 discourse connectives that have been used to signal contrast, and 47 connectives that have been used to signal concession, each being the core of many discourse relations between arguments  $arg_1$  and  $arg_2$ .

### 3.2 Mathematical Notation

In this subsection we briefly review some mathematical notation used in the following subsections. For a set of objects  $A$ , the cardinality of  $A$ , written  $|A|$ , is the number of elements of  $A$ . We write  $B \subset A$  if all elements of  $B$  are elements of  $A$ ; in this case we say that  $B$  is a subset of  $A$ .

The union of two sets  $A$  and  $B$ , denoted by  $A \cup B$ , is the set of elements that are in  $A$ , in  $B$ , or in both  $A$  and  $B$ . The intersection of two sets  $A$  and  $B$ , denoted by  $A \cap B$ , is the set containing all elements of  $A$  that also belong to  $B$ . The relative complement of  $A$  in  $B$ , denoted by  $B \setminus A$ , is the set of elements in  $B$  but not in  $A$ .

A function  $f : A \rightarrow B$  between two sets  $A$  and  $B$  is a rule assigning to each element of  $A$  exactly one element of  $B$ .

On the other hand, we denote by  $\mathbb{Z}$  the set of integers, that is,  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ . The set  $\mathbb{Z}^n$  is the set of elements with the form  $(z_1, \dots, z_n)$ , where  $z_j \in \mathbb{Z}$ . The element of  $\mathbb{Z}^n$  with a one in the  $j$ -th position and zeros everywhere else is written  $e_j$ . For example,  $e_2 \in \mathbb{Z}^4$  is given by  $e_2 = (0, 1, 0, 0)$ . The canonical basis for  $\mathbb{Z}^m$  is the set  $\{e_1, \dots, e_m\}$ . Thus, suppose  $m = 3$ , and given the canonical basis for  $\mathbb{Z}^3, \{e_1, e_2, e_3\}$ , then for all numbers  $\{a, b, c\}$ ,  $v := a \cdot e_1 + b \cdot e_2 + c \cdot e_3$  is the vector  $(a, b, c) \in \mathbb{Z}^3$ .

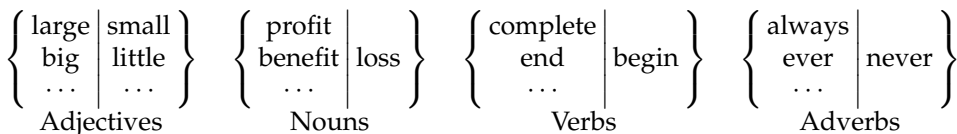
### 3.3 Modeling Arguments as Bags-of-Synonyms/Antonyms

For our purposes, a discourse relation is a **triplet**  $(arg_1, r, arg_2)$ , where  $arg_k$  represents either  $arg_1$  or  $arg_2$  related by the connective  $r$ . For example (taken from file wsj.0617):

	The Manhattan U.S. attorney’s office stressed criminal cases from 1980 to 1987, averaging 43 for every 10,000 adults.	but	the New Jersey U.S. attorney averaged 16.
(4)	$arg_1$	$r$	$arg_2$

In order to build the representations of arguments, we will consider *two-sided* sets of words. In each of these sets, words in the same side are synonyms, and words in opposite sides are antonyms. We build these sets using Wordnet (Fellbaum 1998); in other words, we adopt the WordNet model in the decision of what words stand in a synonymy or an antonymy relation. In the next sections, we provide definitions and lay out the procedures we follow to construct representations of discourse arguments.

*3.3.1 Synonym/Antonym Retrieval Function.* Let  $\mathcal{D}$  be the corpus of documents; in our setting, each document is a discourse relation from the PDTB3 as in Example (4). Prior to the construction of the sets, POS-tagging and Name-Entity-Recognition were carried out to search only for synonyms and antonyms of the words whose POS is of interest to the present study. Let  $\mathcal{V}$  be the vocabulary of  $\mathcal{D}$ . Given a word  $w \in \mathcal{V}$ , we consider



**Figure 1**  
Actual examples of sets  $C_i$  by POS.

a function  $f$  that performs a query to WordNet and returns a set of synonyms of  $w$ :  $\text{syn}_V(w)$ , and a set of antonyms of  $w$ :  $\text{ant}_V(w)$ . Once retrieved, each element  $z$  of  $\text{syn}_V(w)$  or  $\text{ant}_V(w)$ , is tagged along with its part-of-speech  $\text{POS}_z$ . Thus, these subsets are such that  $(w, \text{POS}_w) \in \text{syn}_V(w)$ , and only words pertaining to  $\mathcal{V}$  are actually included in either subset; i.e.  $\text{syn}_V(w) \subset \mathcal{V}$ , and  $\text{ant}_V(w) \subset \mathcal{V}$ . Notice that any of these subsets may be the empty set. Hence, such a function may be formalized as follows:

$$f(w) = (\text{syn}_V(w) \subset \mathcal{V}, \text{ant}_V(w) \subset \mathcal{V})$$

3.3.2 *Bags of Synonyms/Antonyms.* Let us consider collections of the form:

$$C_i = \{C_i^L \subset \mathcal{V}, C_i^R \subset \mathcal{V}\}, \text{ for some } i = 1, 2, \dots, N \tag{1}$$

We will refer to Expression (1) as a “bag-of-synonyms/antonyms”, reminiscent of the bag-of-words model (Harris 1954). In what follows, we describe the procedure to build these collections. Intuitively, we intend each collection to look like this:

$$C_i = \{C_i^L, C_i^R\}$$

where words in  $C_i^L \subset \mathcal{V}$  are synonymous with each other, words in  $C_i^R \subset \mathcal{V}$  are synonymous with each other, and two words  $z \in C_i^L, w \in C_i^R$  are antonymous with each other. Hence, if two words are on the same side, they are synonymous with each other, and antonymous if they are on opposite sides in the same set.

In fact, we aim at constructing these sets in such a way that they only contain words with the same grammatical form (i.e., adjectives, nouns, verbs, or adverbs). In Figure 1, we show some actual examples of these sets, obtained from the PDTB3.

For the sake of clarity, we will make the following abuse of notation  $w \in C_i \cup C_j$  to mean that  $w \in C_i^L \cup C_i^R \cup C_j^L \cup C_j^R$ . Let  $w_0$  be the first word in  $\mathcal{D}$ . The construction of these collections is as follows:

**Step 1.**  $w_0 \in \mathcal{D}$ ;

**Step 2.**  $C_1 \leftarrow f(w_0) = (\text{syn}_V(w_0), \text{ant}_V(w_0))$ ;

**Step 3.**  $N \leftarrow 1$ ;

**Step 4.** FOR EACH  $w_k \in \mathcal{D}$  REPEAT **Step 5** to **16** UNTIL no more words in  $\mathcal{D}$  are found:

**Step 5.** IF  $w_k \notin \bigcup_{i=1}^N C_i$  THEN:CONTINUE, ELSE: GO TO **Step 9**;

**Step 6.**  $C_{N+1} \leftarrow f(w_k) = (\text{syn}_V(w_k), \text{ant}_V(w_k))$ ;

**Step 7.**  $N \leftarrow N + 1$ ;



**Step 8.** GO TO **Step 4**;

**Step 9.** either  $w_k \in C_j^L$  or  $w_k \in C_j^R$ , for some  $j = 1, 2, \dots, m$ ;

**Step 10.** IF  $w_k \in C_j^L$  THEN: CONTINUE, ELSE: GO TO **Step 14**;

**Step 11.**  $C_j^L \leftarrow C_j^L \cup \text{syn}_v(w_k)$ ;

**Step 12.**  $C_j^R \leftarrow C_j^R \cup \text{ant}_v(w_k)$ ;

**Step 13.** GO TO **Step 4**;

**Step 14.**  $C_j^L \leftarrow C_j^L \cup \text{ant}_v(w_k)$ ;

**Step 15.**  $C_j^R \leftarrow C_j^R \cup \text{syn}_v(w_k)$ ;

**Step 16.** GO TO **Step 4**;

After parsing the corpus, the resulting sets are manually curated in order to reduce redundancy. Curation consisted only in eliminating instances of the same word on both sides of the same set. At the end of the construction of the sets, it follows that any two words on the *same side* of each set satisfy one of the following conditions:

- Both words are synonyms between them, according to WordNet.
- Both words are synonyms to a third common word, according to WordNet. This third word is also in the same set.
- There is a third word in the same set such that each of the two words is antonym to this third word, according to WordNet. This third word is also in the same set.

Accordingly, there are words that appear on the same side in several sets, but with different antonyms. Hence, even if there are multiple copies of the same word in different sets, their antonyms will be different and therefore the sets are not redundant.

In this way,  $m$  (527 in our case) sets are obtained, which we order according to their POS. Thus, we define a new set  $S_{all}$  in the following way:

$$S_{all} := \{ \underbrace{C_1, \dots, C_{n1}}_{\text{Adjectives}(aj)}, \underbrace{C_{n1+1}, \dots, C_{n2}}_{\text{Nouns}(n)}, \underbrace{C_{n2+1}, \dots, C_{n3}}_{\text{Verbs}(v)}, \underbrace{C_{n3+1}, \dots, C_m}_{\text{Adverbs}(av)} \}$$

Hence,  $S_{all}$  contains the total number of collections  $m$ , and it should be clear now that we can define disjoint subsets  $S_{aj}$ ,  $S_n$ ,  $S_v$ ,  $S_{av}$  by taking the collections that correspond with the appropriate range of indices in  $S_{all}$  for each subset. Thus, now we are able to define 4 subsets as follows:

$$\begin{aligned} S_{all-aj} &:= S_{all} \setminus S_{aj} & S_{all-n} &:= S_{all} \setminus S_n \\ S_{all-v} &:= S_{all} \setminus S_v & S_{all-av} &:= S_{all} \setminus S_{av} \end{aligned} \quad (2)$$

**3.3.3 Modeling Arguments as Bags-of-Synonyms/Antonyms.** We now describe the method to build the representations. In this article, we only consider contrast and concession DRs.

We start by mapping every document in  $\mathcal{D}$  to a triplet  $(\text{arg}_1^i, r^i, \text{arg}_2^i)$  (cf. Example (4)) for some  $i = 1, 2, \dots, \rho$ , where  $\rho$  is the number of discourse relations under study. Let  $\mathcal{A}_1$  be the set of all arguments  $\text{arg}_1$ , and  $\mathcal{A}_2$  the set of all arguments  $\text{arg}_2$  under consideration. Consider a function  $R$ :

$$R: \mathcal{A}_1 \cup \mathcal{A}_2 \rightarrow \mathbb{Z}^m \quad (3)$$

assigning a vector representation to each argument, and let  $\{e_1, e_2, \dots, e_m\}$  be the canonical basis for  $\mathbb{Z}^m$ . Now, the procedure to construct the representations of arguments in discourse relations is:

**Step 1:**

$$\alpha_j(\text{arg}_k^i) = |\text{arg}_k^i \cap C_j^R| - |\text{arg}_k^i \cap C_j^L| \text{ for } j = 1, 2, \dots, m \quad (4)$$

**Step 2:**

$$R(\text{arg}_k^i) = \alpha_1(\text{arg}_k^i)e_1 + \alpha_2(\text{arg}_k^i)e_2 + \dots + \alpha_m(\text{arg}_k^i)e_m \quad (5)$$

where  $|\text{arg}_k^i \cap C_j^*|$  stands for the number of words of  $\text{arg}_k^i$  contained in  $C_j^*$ .

Thus, instead of representing arguments by their word content, we represent them based on our bags-of-synonyms/antonyms, sorted by POS. Note that Equation (4) may yield 0 because there are two words in the argument that are antonymous to each other or because the argument does not contain any words from the corresponding set. In the first case, we will say that a set  $C_j$  is nullified for that argument. If this situation occurs in all arguments where  $C_j$  appears, we will say that the set is *nullified by algorithm*.

The reader may note that the method above considers all collections within  $S_{all}$ . We will denote by  $R_{all}$  the representation considering all POS functions. Thus, after having processed each word  $w$  in the argument  $\text{arg}_k^i$ , its representation is a vector of positive or negative integers, which may look like this:

$$R_{all} = R(\text{arg}_k^i) = (0, -1, \dots, 3) \in \mathbb{Z}^m \quad (6)$$

On the basis of the same principle, we can use the subsets  $S_{all-nj}$ ,  $S_{all-n}$ ,  $S_{all-v}$ ,  $S_{all-av}$  to obtain other representations made up of only bags-of-synonyms/antonyms belonging to specific POS. In this way, we define the additional representations in Table 1.

### 3.4 A Walk-Through Example

In this section, we provide a walk-through example of the modeling steps of the arguments. The following is an actual example from the PDTB3 corpus (file wsj\_1120).

**arg<sub>1</sub>:** Japan has climbed up from the ashes of World War II and a gross national product of about \$800 per capita to reach the heavyweight class among industrialized nations.

In this sentence, the token “climbed” appears in the following bags of synonyms/antonyms: in n° 351 made out of verbs, on the left-hand side; in n° 356 made out of verbs, on the right-hand side; and in n° 494 made out of verbs, on the left-hand side. The fact that a token may be found in more than one set is due to the fact that the bags of



propose to make this analysis seeks to capture these signals in some of the POS we chose, namely: nouns, verbs, adjectives, and adverbs. Any word with any of these grammatical functions, which appears in the corpus and has a synonym or antonym also appearing in the corpus, is present in one of our sets. Moreover, there is no word in any set that does not appear in the corpus. Words that do not have any of these grammatical functions, or do not have synonyms or antonyms, are not included in any set. Therefore, the representation of an argument, which is part of some discourse relation, is an abstraction of the lexical content of the argument: The syntactic structure disappears. Thus, the sentence corresponding to the argument is reduced to a string of integers, each integer corresponding to one of the 527 sets.

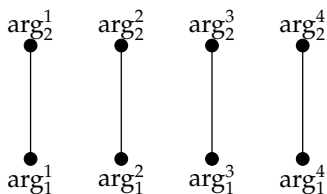
### 3.5 Discourse Relations as Knowledge Graphs

A knowledge graph is an abstract data structure that represents a network of real-world entities and illustrates the relationship between them. It consists of labeled nodes—also called vertices—which represent the entities, interconnected by links—also called edges—which represent the relationships between the entities. In this way, we can visualize in a relatively simple way the interaction between the entities as a graph structure. In what follows, we will refer to this data structure as a knowledge graph or simply as a graph.

In our setting, we will use graphs to describe the relationship between the arguments of a discourse relation by means of a connective. Our goal is to represent each connective as a graph. A first approach is to let each node represent an argument expressed in natural language; that is, without making any abstraction of its lexical content. For each discourse relation kind and each connective type, we consider that each  $\text{arg}_1$  is a node joined to its respective  $\text{arg}_2$  by an edge; we call this configuration a **stick**. As it is very unlikely that one argument (expressed in natural language) appears exactly in more than one discourse relation, every one of these graphs should look like the one depicted in Figure 2.

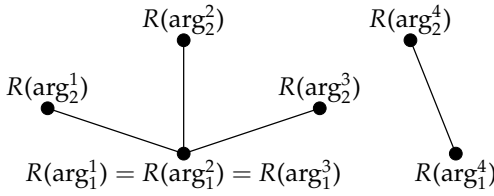
However, this type of graph is not useful for analyzing the behavior of DRs in terms of regularities, or “patterns”, of synonymy/antonymy relations, since each entity is related to a single distinct entity. In order to be able to find lexical–semantic patterns between the arguments of DRs, we will use the representations proposed in the previous section. In this case, each node in a network represents the content of each argument in terms of the bags-of-synonyms/antonyms we have defined.

After obtaining the representation of each argument, we expect to observe a variation in the configuration of the original graph depicted in Figure 2. This variation could



**Figure 2**

Graph depicting 4 discourse relations for some marker as a collection of sticks. The subscript corresponds to either argument 1 or 2 of each relation, and the superscript enumerates the relations.



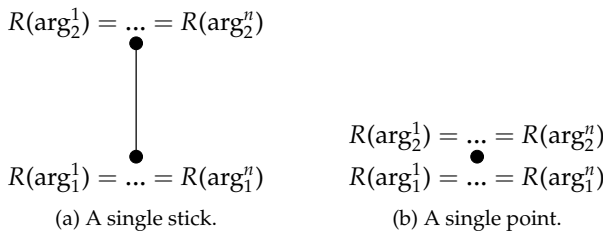
**Figure 3**  
 A rich knowledge graph for some discourse marker after applying the bag-of-synonyms/antonyms representation to the arguments of 4 relations. Left: The first argument of relations 1, 2, and 3 become identical, reducing to a single node; i.e., they share the same lexical-semantic contents. Right: Relationship 4 remains different from the others for the same marker.

result in the appearance of *central* nodes, which represent arguments that share the same lexical-semantic contents as illustrated in Figure 3. We say that we may obtain a “richer”, or “more complex” graph, because ramifications from a node may appear, as it is now possible that two, or more, different argument representations share the same bag-of-synonyms/antonyms: *lexical-semantic similarity patterns appear*.

The reconfiguration of the original graph may lead to two possible extreme cases. The first one is where there are no similarity patterns introduced by our representation. In other words, no two entities share any synonymy/antonymy pattern. The resulting graph is similar to Figure 2. This situation implies that every entity is different under the representation, that is, they have no common information. The second case is depicted in Figure 4. This situation implies that every entity is the same under the representation. That is, they share all the information in terms of synonymy/antonymy patterns.

**3.6 Measuring Inter-Relation Synonymy/Antonymy Patterns Between Arguments**

In terms of graphs, one difference between the three situations described above is the presence of nodes with branching. Branching implies connectivity, and therefore we want to know how strong this connectivity is. This can be characterized in terms of high values of **centrality** measures (Bonacich 2007). In graph analysis, centrality is a very important concept for identifying relevant nodes in a graph; it addresses the question: “What characterizes an important vertex?”. The main thesis is that “a node is important if it is linked to by other important nodes”. One important centrality measure is the **eigenvector centrality**—also called eigencentality—(Golbeck 2013), which is a measure



**Figure 4**  
 Two representations for which many entities share the same bag-of-synonyms/antonyms.

of the influence of a node in a graph. In what follows, we will use the term “centrality” to refer to this measure.

In our setting, centrality measures allow us to know if there is any node with greater relevance than others, in terms of *concentrating a lexical–semantic pattern*, thus becoming an argument that links several others. In this way, measuring the importance of a node is equivalent to assessing the relevance of an argument in terms of capturing a lexical–semantic pattern repeated across the corpus in a given type of discourse relation (contrast or concession, in this article).

For our purposes of analyzing the behavior of our argument representations, we propose to compute the *maximum centrality* and *mean centrality* values for each graph representing all relations established by a discourse marker.

In our context, both metrics are relevant. On the one hand, a low mean centrality value, for example, would tell us that there are only sticks in the graph; that is, all discourse relations are established between a pair  $R(\text{arg}_1)$ ,  $R(\text{arg}_2)$  and no argument is connected to more than a single distinct argument (Figure 2). On the other hand, a high maximum centrality value would tell us that there is one argument (i.e., one  $R(\text{arg}_x)$ ) that connects to two or more arguments that in turn may be connected to other arguments; in this case we say that this node is strongly connected, it is a **concentrating node**. In other words, if we remove this node from the graph, we alter its structure significantly. The most simple case is when given two or more distinct discourse relations, the representation of one of the arguments is the same for all of them (Figure 3).

In the context of our analysis, each discourse marker produces a graph. Therefore, there will be as many graphs in the contrast class as there are connectives associated with this type of discourse relation. The same is true for the concession class. In order to quantify the variations in the configuration of graphs, *attributable to lexical–semantic similarity patterns*, we define two centrality-based metrics, namely,  $\phi_m$  and  $\phi_a$ , as follows. Given a graph  $G(V, E)$ , let  $x_v$  be the eigenvector centrality of node  $v$ . Thus, for all graphs  $G_i(V_i, E_i)$ ,  $i = 1, 2, \dots, n$ , we have:

$$\begin{aligned}\phi_m &= \frac{1}{n} \sum_i^n \max_i \{x_v | \forall v \in V_i\} \\ \phi_a &= \frac{1}{n} \sum_i^n \text{mean}_i \{x_v | \forall v \in V_i\}\end{aligned}\tag{7}$$

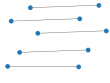



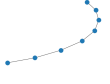
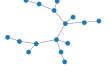
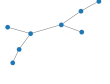

Given a set of graphs, these metrics compute the average of the maximum centrality values and the average of the mean centrality values respectively for the set. The dynamics of these metrics will give us information about the dynamics of the phenomenon of argument concentration or dispersion.

In Table 2, we illustrate the connectivity phenomenon using synthetic graphs. The table shows two sets of graphs, where for each graph the corresponding maximum centrality and mean centrality values are computed. For each set of graphs the values of  $\phi_m$  and  $\phi_a$  are shown.

In our work, we analyze the graphs of explicit and implicit discourse connectives, using the representation  $R_{\text{all}}$ , and observe how these graphs change when we replace  $R_{\text{all}}$  by the representations  $R_{\text{all-aj}}$ ,  $R_{\text{all-av}}$ ,  $R_{\text{all-v}}$ , and  $R_{\text{all-n}}$ . The intention is to analyze the effect of a missing POS on the reconfiguration of the graph of  $R_{\text{all}}$ . We will use both metrics  $\phi_m$  and  $\phi_a$  defined in Equation 7 to perform the analysis.

**Table 2**

Two sets of synthetic graphs with their corresponding maximum eigencentrality and mean eigencentrality values. In each set the graphs are ordered in increasing order of maximum centrality. The values of  $\phi_m$  and  $\phi_a$  are shown for each set. As can be seen, the second set has higher values of  $\phi_m$  and  $\phi_a$ . Higher values of  $\phi_m$  report that the graphs in set *b*) have more important nodes on average, meaning that if these nodes are removed, the structure of the graphs would change significantly.

Set	Item	Graph	$\max(x_v)$	$\text{mean}(x_v)$	$\phi_m$ and $\phi_a$ values per set
a)	1		0.316	0.316	
	2		0.371	0.042	$\phi_m = 0.373$ $\phi_a = 0.278$
	3		0.377	0.377	
	4		0.428	0.377	
b)	5		0.464	0.334	
	6		0.503	0.210	$\phi_m = 0.554$ $\phi_a = 0.296$
	7		0.542	0.324	
	8		0.707	0.322	

**3.7 Finding Intra-Relation Synonymy/Antonymy Patterns**

In this subsection, we describe how we use the representations defined in Subsection 3.3 to discover the presence of synonymy and antonymy between arguments in a discourse relation. Consider the following example, extracted from the manual of the corpus PDTB3 (file wsj\_0359):

(5) After all, gold prices usually **soar** when **inflation** is high. on the other hand, Utility stocks, **thrive** on **disinflation** . . .  
 (arg<sub>1</sub>) (marker) (arg<sub>2</sub>)

In this case, we can observe the presence of synonymy and antonymy between arg<sub>1</sub> and arg<sub>2</sub> (*soar* is synonymous with *thrive*, and *inflation* is antonymous with *disinflation*).

We can detect these *intra-relation synonymy/antonymy matching patterns* using the representations previously defined.

Consider a discourse relation  $(arg_1, r, arg_2)$  for some explicit or implicit connective  $r$ . Now, consider the representation of each argument,  $R(arg_1)$  and  $R(arg_2)$ , given by

$$R(arg_1) = (a_1, \dots, a_K), \quad R(arg_2) = (b_1, \dots, b_K)$$

These representations can be any of the 5 representations previously defined. The element-wise product of two representations is:

$$R(arg_1) * R(arg_2) = (a_1 b_1, \dots, a_K b_K)$$

Observe that a positive value ( $a_i b_i > 0$ ) means that, both  $a_i$  and  $b_i$  have the same sign, that is, *both  $arg_1$  and  $arg_2$  have synonyms from the same set*. On the other hand,  $a_i b_i < 0$  means that *either  $arg_1$  or  $arg_2$  have antonyms from the same set*. Hence, after computing  $R(arg_1) * R(arg_2)$ , we count the number of positive and negative components, and denote them by  $n_{syn}$  and  $n_{ant}$ , respectively. When two or more synonyms of the same set appear in  $arg_1$  and  $arg_2$ , we say that there is a synonymy match. The same is true for antonyms. The count  $n_{syn}$  is the number of synonymy matches, while  $n_{ant}$  is the number of antonymy matches. For example, if  $R(arg_1) = (-2, 0, 1, 0)$ , and  $R(arg_2) = (1, -1, 1, 0)$ , then  $R(arg_1) * R(arg_2) = (-2, 0, 1, 0)$ , and therefore  $n_{syn} = 1$  and  $n_{ant} = 1$ .

The pair  $(n_{syn}, n_{ant})$  gives a 2-dimensional representation of the discourse relation  $(arg_1, r, arg_2)$  in terms of the synonymy/antonymy between the arguments. In the example (5) we would have  $n_{syn} = 1$  and  $n_{ant} = 1$ . Therefore, these counts yield a representation for this discourse relation, which is the point  $(1, 1)$ .

#### 4. Corpus Analysis Using the Proposed Representations

In this section we present the analysis of contrast and concession DRs in the PDTB3 corpus using our representations. First, we provide quantitative data on the PDTB3 corpus for the present study. Then we show an actual example of the graphs  $R_{all}$  and  $R_{all-aj}$  for the “but” connective in the contrast class. We continue with the results concerning the distributions of the maximum centrality values for both classes, along with a comparative analysis of the  $\phi_m$  and  $\phi_a$  values for our different representations for each class. This analysis provides information on the inter-relation lexical patterns found in the whole set of contrast and concession relations in the corpus. We conclude this section by visualizing a comparative analysis of the counts of intra-relation synonymy/antonymy patterns between arguments for contrast and concession DRs.

##### 4.1 Quantitative Data on the PDTB3 Corpus for the Current Study

In the PDTB3, there are connectives that are used to indicate both contrast and concession. Only two connectives associated with concession were left out of our representation, namely, *despite being* and *or*, because none of our sets had words coming from the arguments associated with these connectives. Consequently, for the purposes of our  $R_{all}$  representation, 26 connectives that were annotated in contrast discourse relations and 45 connectives associated with concession were considered. Recall that each of these connectives ( $r$ ) can be the core of many discourse relations—that is, triplets  $(arg_1, r, arg_2)$ . Table 3 summarizes the amounts and proportions of data we used for

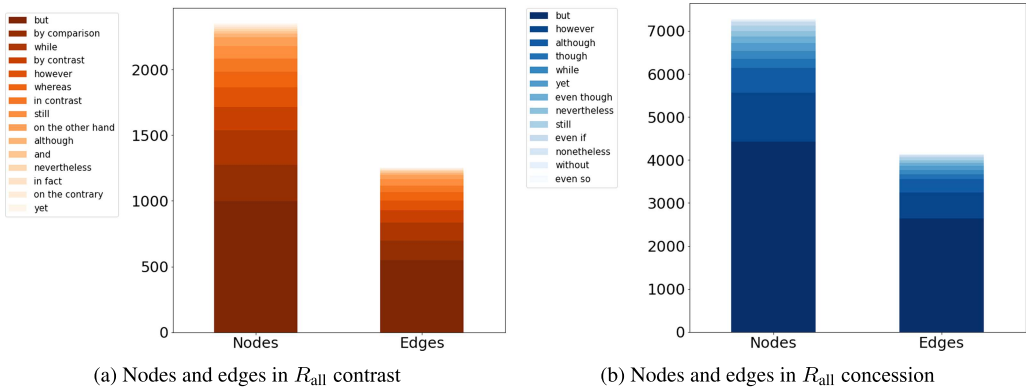


**Table 3**

Data used for the present study.

PDTB3 triplets ( $\text{arg}_1, r, \text{arg}_2$ ) available for representation	
Total number of triplets:	7,217
Contrast triplets:	1,702 (23.6%)
Concession triplets:	5,515 (76.4%)
Explicit triplets:	5,064 (70.2%)
Implicit triplets:	2,153 (29.8%)
Actual number of triplets used for analysis	
$R_{\text{all}}$ :	6,163 (85% of total)
$R_{\text{all-aj}}$	5,502 (76% of total)
$R_{\text{all-n}}$	5,646 (78% of total)
$R_{\text{all-v}}$	3,624 (50% of total)
$R_{\text{all-av}}$	5,945 (82% of total)
Sets usage and coverage	
Total number of sets:	527 of which: 231 (ADJ); 118 (NOUN); 144 (VERB); 34 (ADV).
Nullified by algorithm:	1 (0.2% of total)
Used once:	12 (2.3% of total)
Used more than once:	514 (97.5% of total)
Mean coverage of $\text{arg}_1$ in $R_{\text{all}}$ :	ADJ (18.3%); NOUN (12.9%); VERB (66.0%); ADV (2.8%).
Mean coverage of $\text{arg}_2$ in $R_{\text{all}}$ :	ADJ (18.5%); NOUN (11.9%); VERB (65.8%); ADV (3.8%)
Average cardinality of sets on each side for each POS	
ADJ:	Left: 5; Right: 3
NOUN:	Left: 6; Right: 3
VERB:	Left: 27; Right: 9
ADV:	Left: 3; Right: 2
Statistics on nodes and edges in representation graphs	
Nodes and edges per connective (Contrast)	Figure 5a
Nodes and edges per connective (Concession)	Figure 5b

our experiments and analysis. As mentioned before, to each connective corresponds a graph, which contains both nodes and edges. However, it should be kept in mind that there is not a one-to-one correspondence between a textual argument and a node of the graph, since the latter may be representing two or more arguments (a concentration phenomenon that can be observed in Figure 3). The same is true for edges, in the sense that a relation marked by a particular connective does not necessarily have an edge in



**Figure 5** Distribution of nodes and edges in the  $R_{all}$  graphs of contrast (a) and concession (b) having more than 10 nodes.

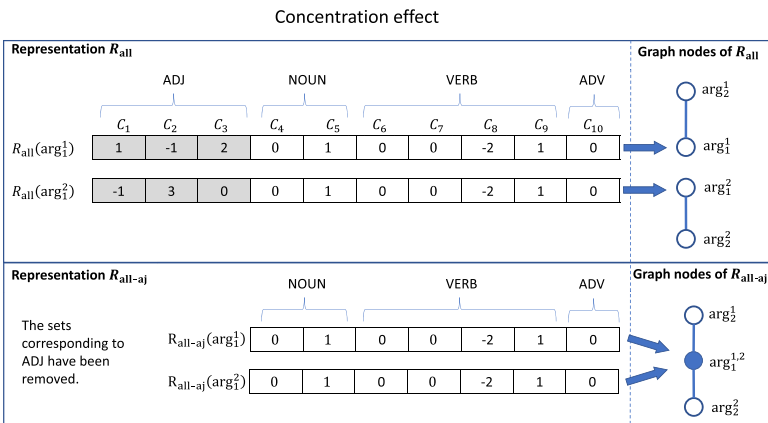
one-to-one correspondence (see Figure 4). The purpose of these bar charts is to show the relative proportion of nodes for each connective and edges for each connective. In the framework of our representation, the proportion of nodes accounts for the number of arguments of each connective and the proportion of edges accounts for the number of relationships between the arguments. Despite not having a one-to-one correspondence, the data yielded by our representation graphs are congruent with the data provided in the PDTB3 Annotation Manual (Appendices A and C).

Figure 5 provides information regarding how connectives that typically encode contrast and concession take on different weights between these two close discourse relations. For example, we see that “by comparison” is specialized in contrast; “although” and “however” are present in both but are proportionally more frequent in concession. The opposite happens with “while”, which is more frequent in contrast. These observations are consistent with the data provided in the PDTB3 Annotation Manual. Looking at the mean coverage of one or the other argument in Table 3, we see that verb sets are used in the highest proportion (66%) in each representation, followed by adjectives with about 18%. Although the number of verb sets is almost half that of adjectives, many verbs are present in several sets (which is reflected in the average cardinality of the sets of this POS), so the weight given to these sets in each representation increases proportionally with this multiplicity.

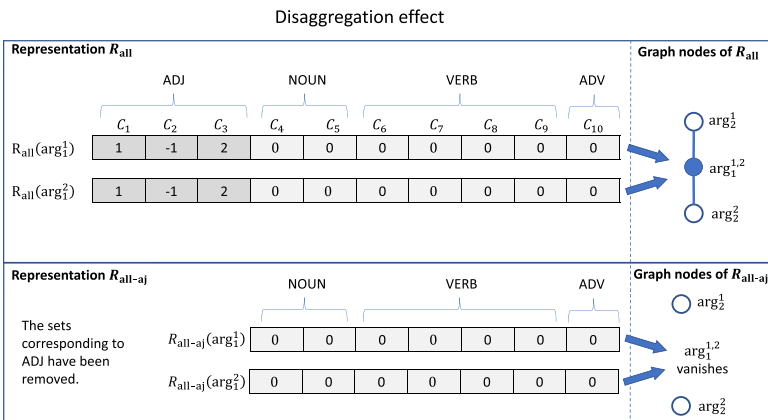
### 4.2 Representative Knowledge Graphs

As we have said, we are interested in analyzing the behavior of the arguments as a function of the POS we have chosen. One aspect we wish to observe and measure is the influence that each POS has when it is removed from the vector representation. Recall that this representation is a string of integers, where each position in the string represents a set of synonyms/antonyms relative to one of these POS. We measure this influence as the ability of a POS to differentiate one argument from another. The differences occur both in a sense of concentration (merging) of nodes and in a sense of disaggregation (separation) of nodes.

Figure 6a illustrates the concentration phenomenon. Assuming two discourse relations for a hypothetical connective, the  $R_{all}$  representation vectors of the two  $arg_1$ 's are shown, using 10 sets that are also hypothetical. In the upper frame we observe the vectors corresponding to each  $arg_1$  and how each of them produces a distinct node in the graph. Each of these  $arg_1$  nodes in turn connects to a node corresponding to  $arg_2$ . Note that the vectors of  $arg_1^1$  and  $arg_1^2$  are distinguished only by the sub-string of adjective (ADJ) integers that are different. The lower frame shows the  $R_{all-aj}$  representation, which results from removing these substrings of adjectives from the  $R_{all}$  representations. Since what remains are identical vectors, a single node is produced in the new graph that synthesizes (*concentrates*) both representations. Thus, both  $arg_2^1$  and  $arg_2^2$  are linked by means of the concentrating node.



(a) Concentration effect: After removal of the sets corresponding to adjectives (ADJ) the vector representations become one and the same: the corresponding graph nodes are fused in a single node.

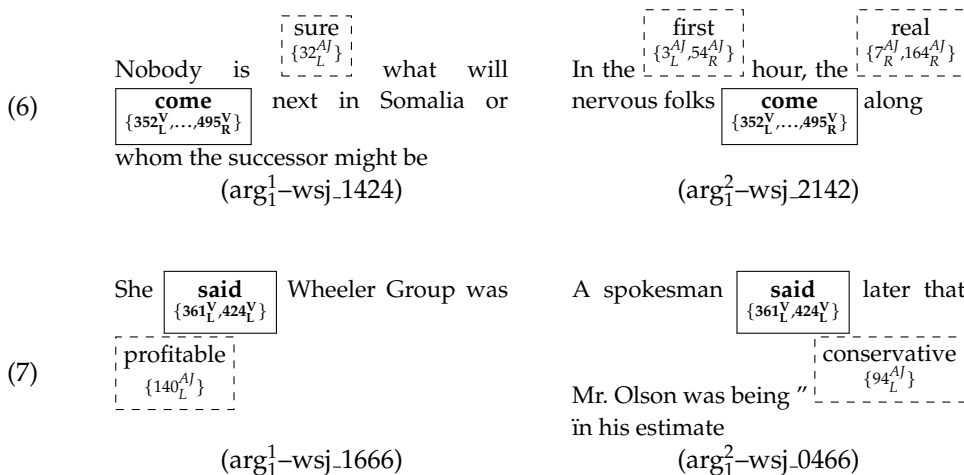


(b) Disaggregation effect: After removal of the sets corresponding to adjectives (ADJ) the vector representations have zeroes everywhere: the corresponding graph nodes vanish.

**Figure 6** Simplified representation of two discourse relations of one hypothetical connective, using hypothetical 10 sets in total; only the representation of  $arg_1^k$  is shown,  $k = 1, 2$ .

Figure 6b illustrates the disaggregation or separation phenomenon. This time we assume that the vectors are identical as a starting condition. By eliminating the adjectives, we remove the non-zero sub-strings. With only null vectors remaining, the node representing both arguments vanishes, and thus, both  $arg_1^1$  and  $arg_2^2$  get separated.

We provide two actual examples from the corpus for the connective “but” coming from files wsj\_1424, wsj\_2142, wsj\_1666, and wsj\_0466. Only  $arg_1^1$  and  $arg_2^2$  are shown as given by the annotation. Text in dashed line boxes represent the words that will be removed from  $R_{all}$  to obtain  $R_{all-aj}$ . The solid-line boxes contain words that will remain in the  $R_{all-aj}$  representation. Note that these words are considered synonymous and are present in the same sets (and on the same side) corresponding to the verbs. Thus, both arguments will be fused into a single node in the corresponding graph.

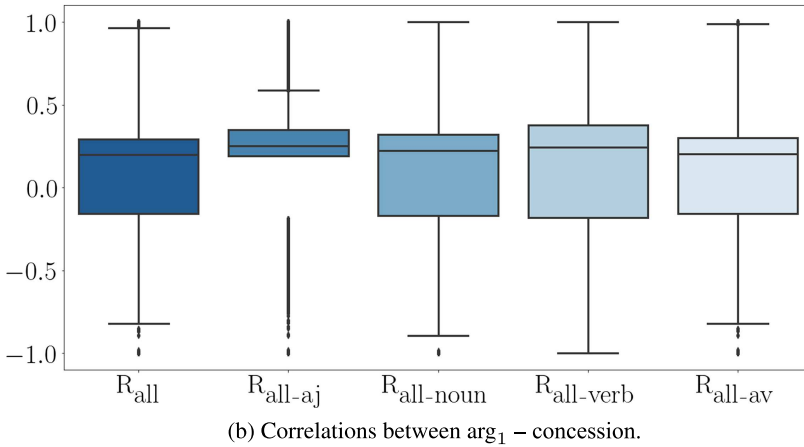
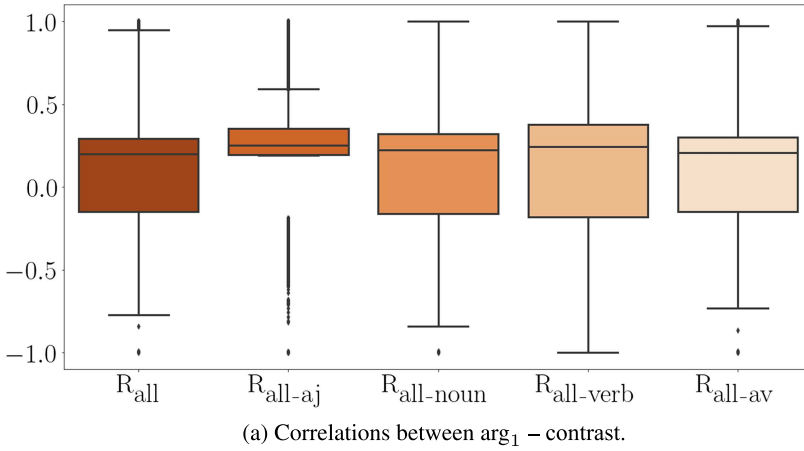


To quantitatively validate the hypothesis that POS removal contributes to node concentration or node disaggregation, we propose to compare the vectors of  $arg_1$  and  $arg_2$  with each other using a Pearson correlation measure. This measure gives 1 if the vectors are identical or  $-1$  if they have coordinates with opposite signs. A correlation in absolute value between 0 and 1 would tell us that the vectors are similar to a lesser or greater degree.

We performed this evaluation as follows. For each connective, we take all arguments 1 and calculate the Pearson correlation between them, keeping only the significant correlations (p-value less than 1%). Figures 7a and 7b show boxplots of these correlation values for the four representations considered in this study for the contrast and concession connectives, respectively. We perform the same procedure for arguments 2. Figures 8a and 8b show the boxplots of the correlations between arguments 2 corresponding to the contrast and concession connectives, respectively.

To compare the samples of each representation against  $R_{all}$ , we conducted Mann-Whitney U tests to evaluate whether the central values were different. We obtained p-values (not shown) below 1% of significance, which tells us that the distributions are indeed quantitatively different.

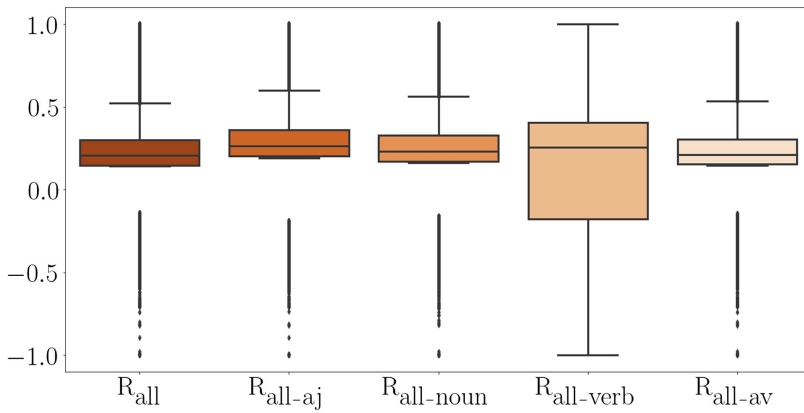
What can be said from these results is that the removal of adjectives from  $R_{all}$  makes a notable difference with respect to the other POS, especially in the representation of  $arg_1$ , in the sense that adjectives make a difference toward argument similarity (i.e.,



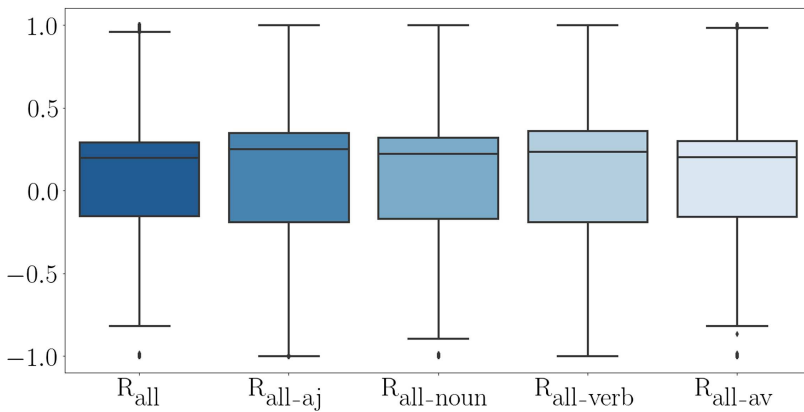
**Figure 7** Pearson correlations between  $arg_1$  vectors in function of the representation. Mann-Whitney U-tests between  $R_{all}$  and each of the other representations are significant below 1%.

concentration) both for contrast and concession relations. A similar effect occurs in nouns for contrast relations with representations of  $arg_2$ . Taking into account that the average weight of adjectives in each representation is 18% (see Table 3), similar to nouns but lower than verbs, we can say that adjectives contribute significantly to make a difference in the similarity of the arguments under our representation  $R_{all}$ .

The effect of verbs is also noticeable (particularly obvious in  $arg_2$  of both contrast and concession) but in the opposite direction to that of adjectives. In this case, the spectrum of correlations tends to open up to encompass more negative correlations. The increase in negative correlations indicates that the remnant vectors have coordinates of similar magnitude with opposite signs. In our representation framework this indicates that in what remains after removing the verbs there are words that belong to the same sets but on opposite sides; that is, in what remains there is a strong synonymy/antonymy relationship. In other words, the verbs homogenize the representations, making them more similar to each other. One might think that this phenomenon is rooted in the fact that the coverage of verbs is much higher than that



(a) Correlations between  $arg_2$  – contrast.



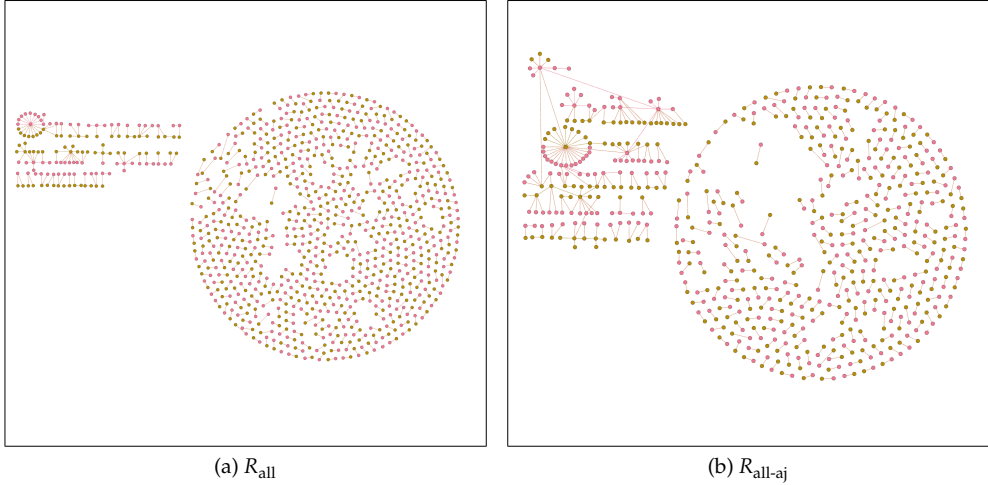
(b) Correlations between  $arg_2$  – concession.

**Figure 8**

Pearson correlations between  $arg_2$  vectors in function of the representation. Mann-Whitney U-tests between  $R_{all}$  and each of the other representations are significant below 1%.

of the other POS (see Table 3). However, the fact that when verbs are present the correlations tend to be positive is an obvious sign that the content of the verb substring tends to be more similar; that is, they contain sets with words on the same side: a more homogeneous relationship of synonymy. Thus, unlike adjectives, verbs seem to play a cohesive role in the representation  $R_{all}$ . Nevertheless, it can be observed that the phenomenon of increased correlation, both positive and negative, is greater in contrast than in concession.

To visually illustrate the concentrating effect on actual graph representations, we show in Figure 9 the graphs of  $R_{all}$  (Figure 9a) and  $R_{all-aj}$  (Figure 9b) for the connective “but” in the contrast class. Each frame of Figure 9 is a graph. We have sectioned each graph in two parts in order to observe two types of interconnections between nodes. In each frame, one can observe a collection of sticks, that is, a pair of nodes connected by an edge, distributed in a disk; and, on the left side of the disk, one can observe nodes with branches. Figure 9b shows that after removing the adjectives the concentration



**Figure 9** Representations of all contrast discourse relations with the connective *but*. We use the representations  $R_{all}$  and  $R_{all-aj}$  to show the influence of adjectives. A phenomenon of concentration of some nodes is observed in  $R_{all-aj}$  when adjectives are removed from the  $R_{all}$  representation. We aim at measuring these concentration or dispersion phenomena using our metrics  $\phi_m$  and  $\phi_a$  introduced in Section 3.6.

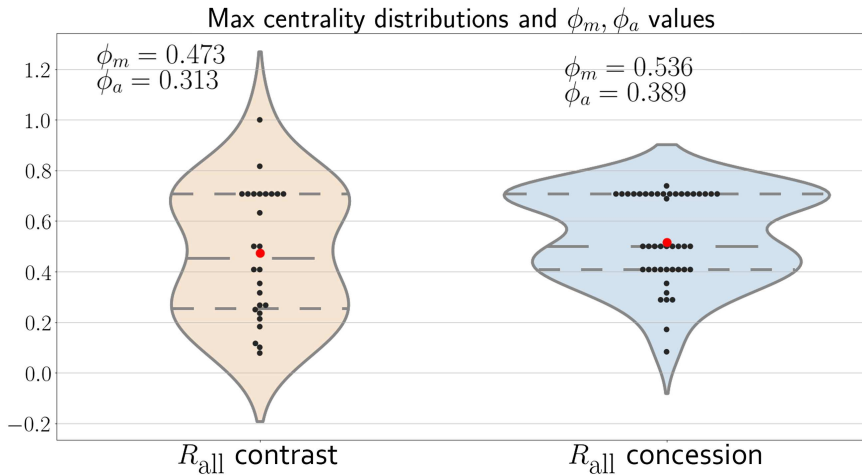
phenomenon seems to increase, to the extent that many sticks have disappeared from the disc, and new nodes with branches appear on the left side.

The branching phenomenon gives rise to different degrees of connectivity strength, as explained in Section 3. In the graphs above, as we move from representation  $R_{all}$  to representation  $R_{all-aj}$  (i.e., without adjectives), we observe the presence of more nodes with branches. This means that, by not considering adjectives in the representation, more arguments share the same synonymy/antonymy patterns. These changes in the configuration of the nodes can be quantified for either the contrast or concession class by our metrics  $\phi_m$  and  $\phi_a$ , as will be shown in the next subsection.

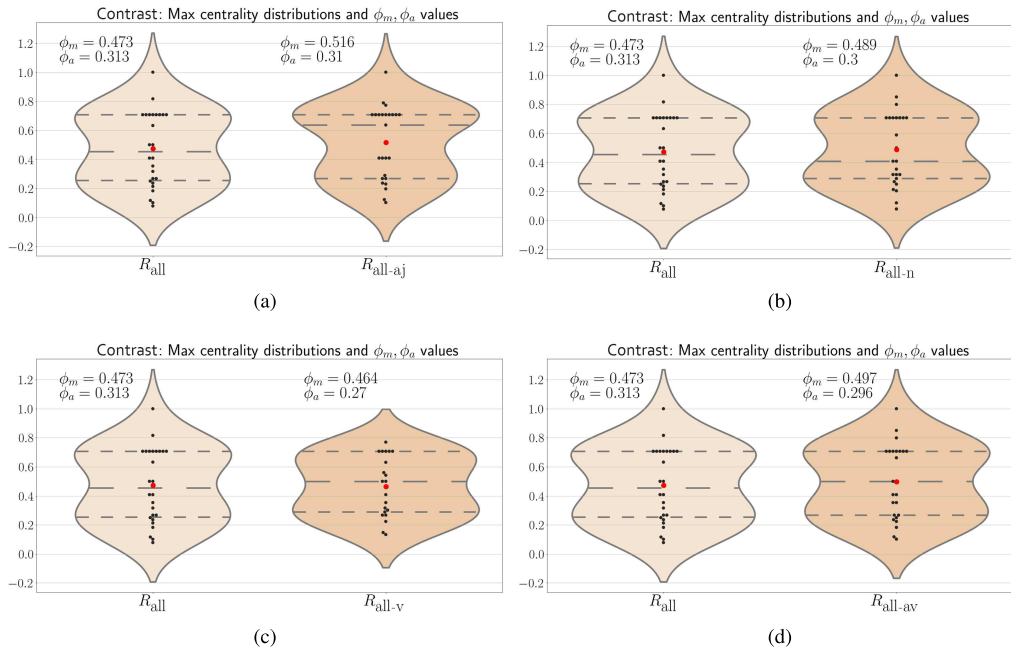
**4.3 Measuring Inter-Relation Synonymy/Antonymy Patterns Between Arguments: Assessing the Connectivity of Graphs Per Class**

Now, we turn our attention to the general situation, by observing in Figures 10 to 12 the distributions of the maximum centrality values along with the metrics  $\phi_m$  and  $\phi_a$  for the set of graphs corresponding to  $R_{all}$  within each class, and as we move from representation  $R_{all}$  to representations  $R_{all-aj}$ ,  $R_{all-av}$ ,  $R_{all-v}$ , and  $R_{all-n}$ . In the figures, each black dot is the maximum centrality of the graph corresponding to some connective. The striped lines represent (from bottom to top) the first quartile, the median, and the third quartile. The red dot shows the mean value.

Observing Figure 10 we see that the distribution of the maximum centrality values are more scattered in the contrast relations than in the concession ones. Figures 11a to 11d show the effects on the distribution of maximum centrality values of  $R_{all}$  when adjectives ( $R_{all-aj}$ ), nouns ( $R_{all-n}$ ), verbs ( $R_{all-v}$ ), and adverbs ( $R_{all-av}$ ) are removed, respectively. The corresponding effects on the concession relations are shown in Figures 12a to 12d.



**Figure 10** The distribution of maximum centrality values of the representations  $R_{all}$  of connectives in the class of contrast (left) and concession (right).



**Figure 11** Distributions of maximum centrality values in graphs by type of representation in discourse contrast relations. The distribution of  $R_{all}$  is always shown as a reference. Variations in the values of  $\phi_m$  and  $\phi_a$  indicate a phenomenon of concentration or dispersion of connections between nodes.

In Figure 11a, we observe that the median has shifted toward higher values of maximum centrality, approaching the third quartile: Concentrations of a larger number of points are observed at the top of the plot. The value of  $\phi_m$  increases while  $\phi_a$  remains practically the same for the  $R_{all-aj}$  representations. In Figure 11b, we observe another

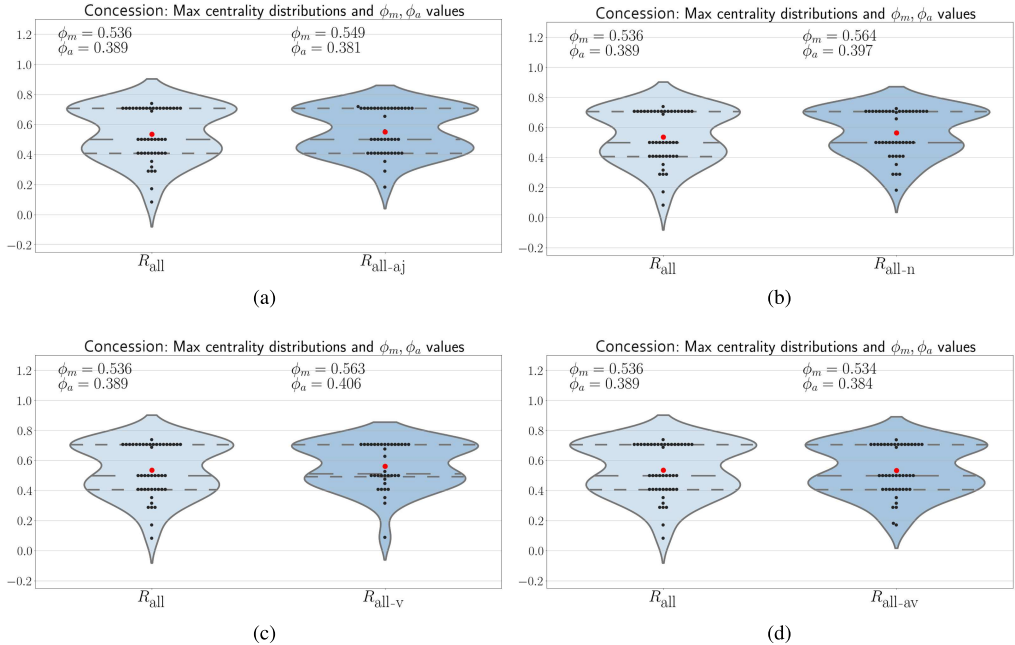


concentration effect. This time, however, the median has decreased slightly and the gap between the median and the first quartile has narrowed: concentrations of a larger number of points are observed at the bottom of the plot. Again the value of  $\phi_m$  grows while  $\phi_a$  remains practically the same for the  $R_{all-n}$  representations. In summary, in both cases we observe an increase in the average maximum centrality. Hence, we infer the existence of a greater number of arguments unified by a single node; that is, the number of concentrated arguments rises when adjectives and nouns are removed.

Figure 11c shows an opposite phenomenon this time—namely, a spread instead of a concentration. In this case, a decrease of both  $\phi_m$  and  $\phi_a$  is present. Therefore, when verbs are removed, fewer arguments are clustered around a single node, and there are also fewer nodes concentrating arguments. In other words, arguments become more spread out when verbs are removed, resulting in a more scattered distribution as can be seen in the figure.

Finally, in Figure 11d, we observe yet another phenomenon, this time resulting from an increase in the value of  $\phi_m$  and a decrease in the value of  $\phi_a$ . In this case, something that could be described as scattered concentrations is observed. In other words, when adverbs are removed, concentrating nodes are lost, but the number of concentrated arguments rises.

Looking at the concession connectives, we observe the following. In Figure 12a the phenomenon observed for the  $R_{all-aj}$  representations of concessive relations is similar to that of adverbs in contrast relations, in the sense that a similar dynamic occurs in the arguments of concession relations when adjectives are removed as revealed by the



**Figure 12** Distributions of maximum centrality values in graphs by type of representation in discourse concession relations. The distribution of  $R_{all}$  is always shown as a reference. Variations in the values of  $\phi_m$  and  $\phi_a$  indicate a phenomenon of concentration or dispersion of connections between nodes.

values of  $\phi_m$  and  $\phi_a$ . That is, in concession relations, when adjectives are removed, concentrating nodes are lost, but the number of concentrated arguments rises.

Figures 12b and 12c show the same phenomenon, in turn similar to that of nouns in contrast relations. That is, in concession relations, when nouns or verbs are removed, more concentrating nodes appear, and the number of concentrated arguments rises.

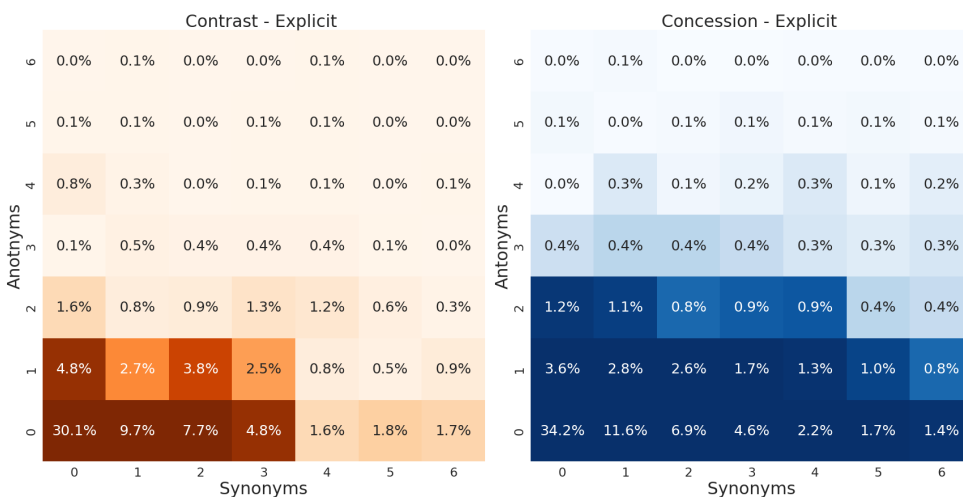
Finally, in Figure 12d the observed phenomenon is similar to that of verbs in contrast relations. That is, in concession relations, when adverbs are removed, in a very slight way, concentrating nodes are lost, and the number of concentrated arguments goes down.

#### 4.4 Quantifying Synonymy/Antonymy Relationships Between Arguments Within a Discourse Relation

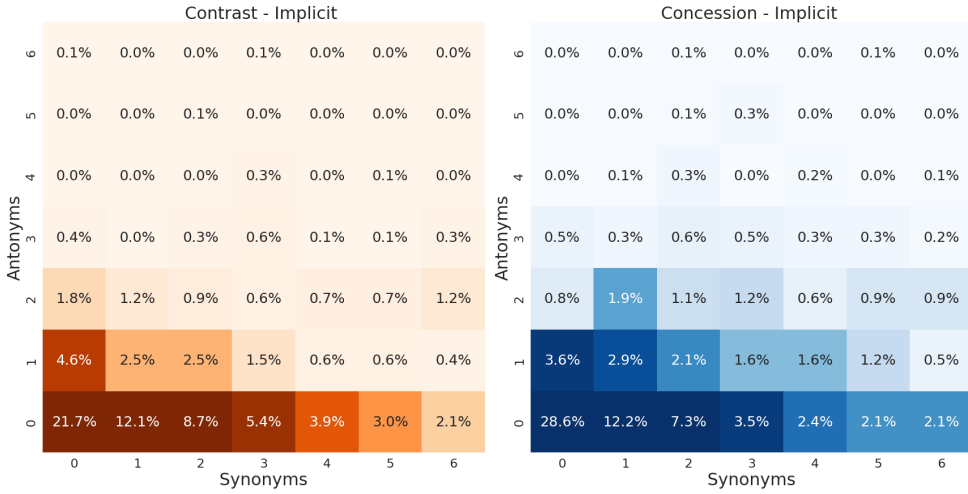
In order to tackle our last two research questions, we now quantify the number of intra-relation synonymy-antonymy matching correspondences between arguments ( $arg_1$  and  $arg_2$  in a given DR), using the 2-dimensional representations described in Subsection 3.7. In order to visualize these points, we use heat maps. Figure 13 shows the counts of synonymy matches ( $n_{syn}$ ) and antonymy matches ( $n_{ant}$ ) in explicit DRs of type contrast and concession, while Figure 14 shows the corresponding counts in implicit contrast and concession DRs. In all cases we used the representation  $R_{all}$ . Each heat map shows proportions of the number of triplets where there is neither synonymy nor antonymy, as well as the proportions of the number of triplets where there are one or more synonymy or antonymy matches. For ease of reading, we will say that the maps show proportions of synonymy and antonymy, keeping in mind that it refers to the former.

In order to quantify the differences between the heatmaps, we conducted a non-parametric Mann-Whitney-Wilcoxon test (Neuhäuser 2011). Table 4 summarizes the results.

At first sight, the higher concentration of triplets in the (0,0) cell indicates that the most common situation is one in which no synonymy or antonymy relation is found



**Figure 13** Proportion of antonymy and synonymy between arguments in explicit DRs of contrast and concession.



**Figure 14** Proportion of antonymy and synonymy between arguments in implicit DRs of contrast and concession.

**Table 4** Statistical significance tests using Mann-Whitney-Wilcoxon to measure differences between the heat maps of synonymy–antonymy correspondences within a discourse relation.

Test	Data ( $U_1, U_2$ )	$p$ -value
1	Contrast–Explicit, Concession–Explicit	0.00012
2	Contrast–Implicit, Concession–Implicit	0.35807
3	Contrast–Explicit, Contrast–Implicit	0.43556
4	Concession–Explicit, Concession–Implicit	4.7934E-6

between argument 1 and 2. The proportions in the maps, nevertheless, indicate that the majority of the represented triplets—spread across all the other cells but the (0,0)—contain intra-relation matches of synonyms and antonyms. The figures also graphically show that, as expected, pairs of synonyms ( $x$  axis) are more frequent than pairs of antonyms ( $y$  axis). In turn, the proportion of antonymy matches drops much faster than the proportion of synonymy matches, indicating that the occurrence of more than one pair of antonyms is extremely infrequent, while more than one pair of synonyms in the same triplet does occur in the analyzed sentences. These results and, specifically, the observed differences between contrast and concession and implicit and explicit relations, answer our research questions and will be discussed in Section 5.3.

**5. Discussion of the Analysis**

The present study was motivated by theoretical and methodological questions. We divide the discussion in subsections according to our research questions.

## 5.1 Computational Models

The first question addressed in this article is a methodological one: How can discourse relations be computationally modeled in order to capture the contribution of the lexical semantics to the discourse relation meaning? We proposed representations of DRs based on POS-bags-of-synonyms/antonyms and were able to computationally capture both inter-relation semantic patterns (i.e., patterns of synonyms/antonyms found in arguments of the whole set of contrast or concession relations occurrences), and intra-relation patterns (i.e., matches of synonymy/antonymy between  $\text{arg}_1$  and  $\text{arg}_2$  of a discourse relation).

Although discourse relations have been the subject of growing attention in corpus linguistic studies as well as in NLP and computational linguistics research, the gap between linguistic and computational approaches remains wide. In recent years, we have seen the emergence of Transformers, which are deep neural networks based on self-attentional mechanisms, and have been shown to be able to better deal with long-range correlations in text processing (Vaswani et al. 2017). Prominent state-of-the-art models like GPT (Radford et al. 2019) and XLNet (Yang et al. 2019) are pre-trained using autoregressive language models, while BERT (Devlin et al. 2019) uses a denoising approach. The usefulness of these models to discover latent relations between text units or as text analysis tools has been demonstrated in countless contributions for solving complex tasks, such as sentiment analysis (Hoang, Bihorac, and Rouces 2019), semantic textual matching (Xia et al. 2021), semantic role labeling (Larionov et al. 2019), and discourse analysis (Kiyomaru and Kurohashi 2021), among many others.

Still, the question remains as to what exactly these models extract at the linguistic level—or, in other words, what these models tell us about the properties of language. With respect to this question, a recent analysis of BERT (Rogers, Kovaleva, and Rumshisky 2020) shows that this model can be very useful for extracting some linguistic knowledge, especially syntactic, and for producing embedded representations with high quality contextualized distributional properties. Despite this, BERT remains vulnerable to variations in context (Atwell, Li, and Alikhani 2021) or syntactic structure (Rogers, Kovaleva, and Rumshisky 2020). On the other hand, extracting knowledge about the linguistic functions that BERT attention heads manage to classify is an arduous and complex process, and the role of attention remains moot (Rogers, Kovaleva, and Rumshisky 2020). Although we know that these models are capable of establishing long-range correlations in the presence of a very large amount of data, we only have some hints about how the model makes decisions based on how its different processing layers “pay attention” to some words or tokens.

In view of this, we believe that our proposal, although simple in its structure, enables a corpus analysis that can lead to a more detailed, cautious, and eventually basic, or elementary, interpretation of how the relations of synonymy and antonymy, which exist in concession and contrast discourse relations, come into play and determine certain properties attributed to this type of linguistic discourse relations. In this sense, our proposed representations of discourse relations, based on POS-bags-of-synonyms/antonyms, allows the computational study of these linguistic forms in a more transparent and linguistically meaningful way.

## 5.2 Inter-relation Analysis

The quantitative analysis of our representations  $R_{\text{all}}$ ,  $R_{\text{all-aj}}$ ,  $R_{\text{all-av}}$ ,  $R_{\text{all-v}}$ , and  $R_{\text{all-n}}$  provides insights into the inter-relation patterns of synonymy and antonymy and into the

importance of POS in these patterns. On the one hand, correlations between argument representation vectors show that particularly adjectives (and nouns to a lesser degree) make a significant difference toward vector similarity when removed. Also, verbs play a cohesive role between arguments, since by removing them the representations tend to become more differentiated. This indicates that adjectives (and nouns to a lesser degree) are differentiators with more strength than the other POS, and that verbs contribute to the similarity of arguments. On the other hand, after performing a Mann-Whitney U-test, we found that the difference between the distributions of the maximum centrality values of the contrast and concession graphs (Figure 10) is significant to less than 1%. In other words, the similarity patterns between our global representations of the synonymy/antonymy content of contrast discourse relations are in fact distinct from those of concession.

In linguistic terms, this indicates that there are lexical–semantic convergences found across contrast and across concession relations in the corpus. In other words, although these representations are mainly built to provide answers to our theoretical questions (more directly pertaining to intra-relation patterns), the proposed representations, capturing argument concentration or dispersion, also offer a measure of the topic homogeneity of the corpus: The higher the number of arguments concentrating in a single node, the higher the semantic homogeneity of the content included in discourse relation segments in the corpus. In this sense, the results in Sections 4.2 and 4.3 indicate that discourse segments involved in concession relations show higher synonymy/antonymy homogeneity than those included in contrast relations in the corpus. This contributes to a potential measure of topic homogeneity. Both the general increase in vector correlations, and the more scattered distribution of the maximum centrality values in contrast than in concession relations (Figure 10), draw this picture.

Our analysis provides answers to the three theoretical questions posed in the study. We first wondered how much different parts of speech contribute to the representation of contrast and concession discourse relations. Previous studies considering lexical terms expressing opposition either limit themselves to one part of speech (Spencer and Stulp 2007), do not specify the word class of the lexical elements considered, or do not analyze the contribution of the different parts of speech (Feltracco, Magnini, and Jezek 2018; Marcu and Echiabi 2002). Our results show that the inclusion or exclusion of all the parts of speech in the representation plays a relevant role: Removing adjectives, adverbs, verbs, and nouns results in changes in the graph configuration, revealing that all of them contribute to capturing inter-relation lexical patterns between arguments. Their contribution in contrast and concession relations is, nevertheless, different.

Removing adjectives and nouns from contrast representations (Figures 11a and 11b) results in higher argument concentration, which indicates that both adjectives and nouns were serving a discriminating role in the argument representation in contrast relations. The topic homogeneity is captured across contrast segments in spite of the fact that nouns and adjectives are playing a more discriminating role. In turn, the topic homogeneity in contrast relations seems to be more dependent on verbs, since when verbs are removed from the contrast representations, fewer arguments are clustered around a single node and there are fewer nodes concentrating arguments (Figure 11c). The effect of removing adverbs is less straightforward: Concentrating nodes are lost but the number of concentrated arguments rises (Figure 11d), thus indicating that arguments that were collapsed into a single node by virtue of the presence of an adverb are still collapsed with other arguments when adverbs are removed. In concession relations, removing nouns and verbs (Figures 12b and 12c) results in representations with more concentrating nodes and higher number of concentrated arguments. Again, this

indicates that nouns and verbs were discriminating in concession relations. The effect of adjectives is less clear in concession relations, where removing them results in changes in the graph but does not indicate a straightforward concentrating or dispersing effect; removing adverbs in concession relations, finally, results in more argument dispersion. Despite these encouraging results, Mann-Whitney U-tests on these (within class) distributions revealed no statistical significance. This indicates that the differences in the centrality values just discussed are only trends. Even so, these trends are congruent with the significant results that the correlations between vector representations gave us, in the sense that adjectives in particular have a differentiating role and verbs provide cohesion in the representation  $R_{all}$ . The lack of statistical significance regarding the maximum centrality is partly explained by the high degree of abstraction introduced by the centrality values. Indeed, compared to the number of vectors used in the calculation of correlations, the number of centrality values is too small to achieve statistical significance. On the other hand, as there is a strong dominance of sticks in general, the maximum centrality values lose strength, so the relevance of the concentrating effect of each POS in the graphs is diminished.

### 5.3 Intra-relation Analysis

The remaining two questions are answered through the analysis of intra-relation synonymy/antonymy patterns. The first question was whether contrast and concession discourse relations were differentiated using this model of representation and, specifically, whether the presence of antonymy and synonymy patterns between argument 1 and 2 of a given relation differed in contrast and concession DRs, as captured by our representations. First, notice that the presence of synonymy and antonymy between the two arguments in contrast and concession relations is, overall, high (around 70%), in contrast with the findings in previous studies using other corpora (Feltracco, Magnini, and Jezek 2018; Spenader and Stulp 2007).

Overall, the presence of pairs of synonyms and antonyms between arguments is almost parallel in the two types of discourse relations. However, from Table 4 and Figures 13 and 14, we see that explicit concession relations show significantly more antonymy and synonymy counts than explicit contrast relations ( $p < 0.1\%$ ). This difference is not found within implicit relations. Regarding the presence of synonymy and antonymy, data shows that antonymy is, overall, less frequent than synonymy in both contrast and concession relations.

The presence of more intra-relation lexical matches in explicit concession relations than in contrast relations is, in principle, unexpected, since contrast discourse relations were expected to be more dependent on the lexical-semantic content of its arguments than concession relations (see Section 2). On a closer look, however, one can see that antonymy is in fact slightly more frequent in explicit contrast than in explicit concession relations, a tendency in agreement with previous studies (Feltracco, Magnini, and Jezek 2018; Crible 2022). Regarding synonymy, we posit that intra-relation synonymy overall contributes to creating coherence through topic continuity (Lei et al. 2018), a discourse function equally displayed in contrast and concession relations.

The unexpected result just mentioned (more lexical matches in concession than in contrast relations) might be due to the PDTB3 encoding procedure, in which “whenever concession can be taken as holding, it is annotated as such, even if contrast also holds by definition” (cf. Webber et al. 2019). This causes that a discourse relation in which there are at least two differences between  $arg_1$  and  $arg_2$  is not necessarily marked as contrast; instead, it could be tagged as concession. The difference in annotation between

PDTB2 and PDTB3 is most likely affecting our ability to clearly tear apart the contribution of lexical semantics to contrast and concession relations, since an important portion of the concession discourse relations in PDTB3 could, in fact, have been annotated as contrast. The annotation manual specifies that, for explicit relations, 75% of the relations in PDTB2 were labeled as contrast and 25% as concession, while in PDTB3, with the new annotation, 78% of the relations are labeled as concession and only 22% are labeled as contrast (i.e., those in which a concessive reading cannot be found). Although we have not investigated the PDTB2 corpus directly, the “provenance” field in the PDTB3 shows that the annotation of 70% of concession relations in PDTB2 has been changed. However, the manual indicates that the nature of the changes can have more than one origin (not just the class), but this is not specified. The percentages just presented give reasons to suppose that most of these changes correspond to relations that are now labeled as concession **despite** the fact that contrast also holds.

If this is the case, having access to the specific set of discourse relations that changed from contrast to concession in PDTB2 and PDTB3 would allow to analyze the three groups of relations separately (1: “only contrast”, labeled as contrast in PDTB3; 2: “contrast & concession”, labeled as contrast in PDTB2 and changed to concession in PDTB3; and 3: “only concession”). Under our logic, the group “contrast & concession” is expected to display lexical patterns more similar to the “only contrast” group, that is to say, a higher presence of antonymy, as well as the overall presence of synonymy indicating topic continuity in both types of relations. The group “only concession” would be expected to differ more evidently from the contrast and the contrast & concession group, and be more dependent on synonymy (and less so on antonymy, according to previous studies) than the remaining two. Although this final analysis is not provided in detail in this article, we ran our experiments by separating Concession Changed from Concession Not Changed, obtaining the following results.

Table 5 shows that the differences between the changed concession group and the contrast group in PDTB3 are not significant, but are significant with the unchanged concession group. Although there may be other sources of change in this new group, these results strongly suggest that our representations capture a difference in the presence of matches between synonymy and antonymy in contrast and concession, masked in the original analysis by the particularities of PDTB3 annotation.

Finally, we wondered whether implicit and explicit discourse relations behave similarly in terms of this representation. The literature on discourse relations modeling seems to operate under the assumption that explicit and implicit discourse relations follow the same linguistic patterns, thus training their models with explicit discourse relations in order to infer implicit ones. This assumption, nevertheless, is at odds with the more straightforward communicative hypothesis that the speaker takes into account

**Table 5**

Statistical significance tests using Mann-Whitney-Wilcoxon to measure differences between the heat maps of synonymy–antonymy matches between changed and not changed relations.

Test	Data ( $U_1, U_2$ )	$p$ -value
1	Concession–Explicit CHANGED, Concession–Explicit NO CHANGE	0.00734
2	Concession–Implicit CHANGED, Concession–Implicit NO CHANGE	0.004012
3	Concession CHANGED, Contrast	0.42323

the difficulty of inferring the discourse relation in order to decide between the explicit or implicit connective (Asr and Demberg 2012a). This idea has mostly been put forward in order to compare the use of connectives for more basic (additive and positive causal) versus less basic (contrast and concession) discourse relations (Das and Taboada 2019; Hoek and Zufferey 2015). Under the same logic, if a discourse relation can be easily inferred from the explicit semantic content in the two arguments, implicit connectives would be expected, whereas explicit ones would be more frequent when the conceptual semantics contributes to a lesser extent to establish the discourse relation.

With this hypothesis, the analysis of synonymy and antonymy global match counts between first and second arguments allows us to compare implicit and explicit contrast and concession relations. The results indicate that there is a significant difference between the presence of synonymy-antonymy matches in explicit concession DRs vs. implicit concession DRs: Explicit discourse relations show a higher proportion of antonymy-synonymy matches than implicit ones. The difference between explicit and implicit contrast relations is not significant. In our data, therefore, when the lexical conceptual semantics in the arguments contributes with more information (mainly in terms of synonymy) to establishing the contrast or concession discourse relation, it is not more likely for the connective to remain implicit. In fact, the opposite occurs for concession relations. These results, showing that synonymy-antonymy matches are equally present in implicit and explicit contrast relations, are in line with the presence of relevant linguistic cues in both implicit and explicit Substitution relations in Webber (2013) (while less relevant linguistic cues only appear in explicit ones). The higher presence of synonymy-antonymy matches in explicit concession than implicit concession in our data, in turn, suggests that the easiness of inferring a concession discourse relation is not so much dependent on the synonym and antonym matches: It is possible to think that the concession relations that writers in our corpus decided to leave implicit were easily inferred based on contextual or discourse knowledge that antonymy and synonymy—and, therefore, our representations—are not capturing, whereas explicit connectives frequently co-occur with intra-relations lexical matches, specially of synonyms, narrowing the kind of discourse relation holding between arguments.

#### 5.4 Usefulness of Our Representations

Lastly, the proposed representations are useful to capture differences among discourse markers. Even though a detailed analysis of the behavior of each connective is out of the scope of this article, in our data, specific connectives show different patterns as shown by the distributions of the maximum centrality values. Previous literature has addressed the idea that discourse markers can be organized in terms of their “cue strength” (Asr and Demberg 2012b), a probabilistic measure of their ambiguity and monosemy, and suggests that the way different discourse markers interact with signals in their context is related to their strength or weakness (Crible 2022). In this line, our proposal opens the possibility to further analyze how different connectives interact with lexical semantic information in the discourse context.

Nevertheless, our approach to represent discourse relations could be useful in contrast and concession pattern recognition tasks using classical classification or regression methods. Machine learning methods allow the selection of suitable features, thus enabling, for example, the analysis to discern the linguistic features that characterize the effectiveness of a discourse (El Baff et al. 2020). In this respect, a Linear Discriminant Analysis could be used with our representations, which could in turn help to evaluate



the ability of each POS to predict the connective or class to which a discourse relation belongs.

Finally, due to the nature of the PDTB3, our analysis is now confined to news texts. Our methodology, however, could be applied to contrast and concession relations in other genres, both written and spoken, contributing to our understanding of the potential differences in discourse relation patterns among discourse genres.

## 6. Conclusion

Our work addresses important questions in corpus linguistics regarding the role of semantic signals in contrast and concession discourse relations. To achieve this, we propose a computational modeling approach to discourse relations in corpus that is transparent about how certain semantic features are automatically analyzed for the sake of linguistic interpretability of the results. In this sense, our approach allows us to abstract lexical–semantic signals of synonymy and antonymy between the arguments of each discourse relation, thus obtaining information regarding the contribution of synonymy and antonymy in the signaling of discourse relations and showing the differences and similarities between types of relations (contrast vs. concession, and implicit vs. explicit), according to the different word classes (POS).

Our model allows us to observe a greater contribution of adjectives, verbs, and nouns in contrast relations than in concession relations. Although we were able to appreciate differences in all POS, our results can be improved. In contrast to what has been reported in the literature so far, our approach allows for a deeper analysis of the role of different word classes as lexical cues.

From the intra-relational analysis, our results show that in general synonymy is more frequent than antonymy, and these lexical signals are more frequent in concession than in contrast in PDTB3. However, antonymy, as expected, is more frequent in contrast, and our results suggest that the differences between contrast and concession are real. When inspecting the concession relations annotated in PDTB3 as changed, our representations also allow us to observe significant differences between relations marked as changed and those that are not. By doing so, we infer that our approach distinguishes between “real” concession relationships and those that are apparently masked by the annotation of these relationships in PDTB3.

Finally, the presence of synonymy and antonymy does not correlate in the expected way with the explicit or implicit occurrence of the connector. There are no more implicit relations in cases where there are more synonymy and antonymy. This could be because one aspect to improve in our approach is to integrate the modeling of negation and other possible cues interacting with synonymy and antonymy.

We point out some weaknesses of the approach. On the one hand, we acknowledge the importance of identifying phrasal verbs, previously to the search for synonyms or antonyms. On the other hand, the approach does not consider the polarity of sentences as given by negation, for example. These factors should be considered in future work.

Although discourse relations have been the subject of growing attention in corpus linguistics studies as well as in NLP and computational linguistics research, the gap between linguistic and computational approaches remains wide, and scarce efforts are being made to deepen the dialogue between these disciplines. Bridging this gap, although challenging, is important for interdisciplinary work and offers a promising landscape toward a more complete understanding of discourse linguistic phenomena. We believe our work is a contribution in this direction. Not the least, we believe our method opens the possibility to extend this kind of analysis to a broader audience, as

the methods employed may be automated. Hence, it is possible to extend this kind of analysis to other corpus that may or may not be annotated, contributing thus to deepen research in corpus linguistics.

### Acknowledgments

Research was partially funded by CONACYT Project A1-S-24213 of Basic Science and CONACYT grants 28268, 29943, and 732458. The authors thank CONACYT for the computer resources provided through the INAOE Supercomputing Laboratory's Deep Learning Platform for Language Technologies.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyzes, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

### References

- Asr, Fatemeh Torabi and Vera Demberg. 2012a. Implicitness of discourse relations. In *Proceedings of COLING 2012*, pages 2669–2684.
- Asr, Fatemeh Torabi and Vera Demberg. 2012b. Measuring the strength of linguistic cues for discourse relations. In *Proceedings of the Workshop on Advances in Discourse Analysis and Its Computational Aspects*, pages 33–42.
- Atwell, Katherine, Junyi Jessy Li, and Malihe Alikhani. 2021. Where are we in discourse relation recognition? In *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 314–325.
- Bakshi, Sahil and Dipti Sharma. 2021. A transformer based approach towards identification of discourse unit segments and connectives. In *Proceedings of the 2nd Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2021)*, pages 13–21.
- Benamara, Farah, Maite Taboada, and Yannick Mathieu. 2017. Evaluative language beyond bags of words: Linguistic insights and computational applications. *Computational Linguistics*, 43(1):201–264. [https://doi.org/10.1162/COLI\\_a.00278](https://doi.org/10.1162/COLI_a.00278)
- Biran, Or and Kathleen McKeown. 2013. Aggregated word pair features for implicit discourse relation disambiguation. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 69–73.
- Boleda, Gemma. 2020. Distributional semantics and linguistic theory. *Annual Review of Linguistics*, 6:213–234.
- Bonacich, Phillip. 2007. Some unique properties of eigenvector centrality. *Social Networks*, 29(4):555–564. <https://doi.org/10.1016/j.socnet.2007.04.002>
- Crible, Ludivine. 2022. The syntax and semantics of coherence relations. *International Journal of Corpus Linguistics*, 27(1):59–92. <https://doi.org/10.1075/ijcl.19109.cri>
- Crible, Ludivine and Liesbeth Degand. 2019. Domains and functions: A two-dimensional account of discourse markers. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, 24. <https://journals.openedition.org/discours/9992>
- Das, Debopam and Maite Taboada. 2018. Signalling of coherence relations in discourse, beyond discourse markers. *Discourse Processes*, 55(8):743–770. <https://doi.org/10.1080/0163853X.2017.1379327>
- Das, Debopam and Maite Taboada. 2019. Multiple signals of coherence relations. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, 24. <https://journals.openedition.org/discours/9992>
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186.
- El Baff, Roxanne, Henning Wachsmuth, Khalid Al Khatib, and Benno Stein. 2020. Analyzing the persuasive effect of style in news editorial argumentation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3154–3160.
- Fellbaum, Christiane. 1998. *WordNet: An Electronic Lexical Database*. Bradford Books. <https://doi.org/10.7551/mitpress/7287.001.0001>

- Feltracco, Anna, Bernardo Magnini, and Elisabetta Jezeq. 2018. Lexical opposition in discourse contrast. In *Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018)*, 6 pages.
- Givón, T. 1983. *Topic Continuity in Discourse: A Quantitative Cross-language Study*. John Benjamins. <https://doi.org/10.1075/ts1.3>
- Golbeck, Jennifer. 2013. Chapter 3 - network structure and measures. In Jennifer Golbeck, editor, *Analyzing the Social Web*. Morgan Kaufmann, Boston, pages 25–44.
- Harris, Zellig S. 1954. Distributional structure. *Word*, 10(2–3):146–162. <https://doi.org/10.1080/00437956.1954.11659520>
- Hoang, Mickel, Oskar Alija Bihorac, and Jacobo Rouces. 2019. Aspect-based sentiment analysis using BERT. In *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, pages 187–196.
- Hoek, Jet and Sandrine Zufferey. 2015. Factors influencing the implicature of discourse relations across languages. In *Proceedings of the 11th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-11)*, 7 pages.
- Kiyomaru, Hirokazu and Sadao Kurohashi. 2021. Contextualized and generalized sentence representations by contrastive self-supervised learning: A case study on discourse relation analysis. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5578–5584. <https://doi.org/10.18653/v1/2021.naacl-main.442>
- Larionov, Daniil, Artem Shelmanov, Elena Chistova, and Ivan Smirnov. 2019. Semantic role labeling with pretrained language models for known and unknown predicates. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*, pages 619–628.
- Lei, Wenqiang, Yuanxin Xiang, Yuwei Wang, Qian Zhong, Meichun Liu, and Min-Yen Kan. 2018. Linguistic properties matter for implicit discourse relation recognition: Combining semantic interaction, topic continuity and attribution. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence*, pages 4848–4855. <https://doi.org/10.1609/aaai.v32i1.11933>
- Marcu, Daniel and Abdessamad Echihabi. 2002. An unsupervised approach to recognizing discourse relations. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 368–375.
- Neuhäuser, Markus. 2011. Wilcoxon–Mann–Whitney test. In Miodrag Lovric, editor, *International Encyclopedia of Statistical Science*. Springer, Berlin Heidelberg, pages 1656–1658.
- Nie, Allen, Erin Bennett, and Noah Goodman. 2019. DisSent: Learning sentence representations from explicit discourse relations. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4497–4510. <https://doi.org/10.18653/v1/P19-1442>
- Prasad, Rashmi, Aravind Joshi, and Bonnie Webber. 2010. Realization of discourse relations by other means: Alternative lexicalizations. In *COLING 2010: Posters*, pages 1023–1031.
- Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.
- Rogers, Anna, Olga Kovaleva, and Anna Rumshisky. 2020. A primer in BERTology: What we know about how BERT works. *Transactions of the Association for Computational Linguistics*, 8:842–866. [https://doi.org/10.1162/tacl\\_a\\_00349](https://doi.org/10.1162/tacl_a_00349)
- Roth, Michael 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)*, pages 524–530. <https://doi.org/10.3115/v1/P14-2086>
- Rysová, Magdaléna and Kateřina Rysová. 2015. Secondary connectives in the Prague Dependency Treebank. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*, pages 291–299.
- Sanders, Ted J. M., Wilbert P. M. Sporeen, and Leo G. M. Noordman. 1992. Toward a taxonomy of coherence relations. *Discourse Processes*, 15(1):1–35. <https://doi.org/10.1080/01638539209544800>
- Spender, Jennifer and Gert Stulp. 2007. Antonymy in contrast relations. In *Proceedings: Seventh International Workshop on Computational Semantics*, 12 pages.
- Sporeen, Wilbert and Ted Sanders. 2008. The acquisition order of coherence relations: On cognitive complexity in discourse.

- Journal of Pragmatics*, 40(12):2003–2026. <https://doi.org/10.1016/j.jpragma.2008.04.021>
- Sporleder, Caroline. 2008. Lexical models to identify unmarked discourse relations: Does WordNet help? *Journal for Language Technology and Computational Linguistics*, 23(2):20–33. <https://doi.org/10.21248/jlcl.23.2008.105>
- Taboada, Maite. 2019. The space of coherence relations and their signalling in discourse. *Language, Context and Text*, 1(2):205–233. <https://doi.org/10.1075/langct.00009.tab>
- Taboada, Maite and Debopam Das. 2013. Annotation upon annotation: Adding signalling information to a corpus of discourse relations. *Dialogue & Discourse*, 4(2):249–281. <https://doi.org/10.5087/dad.2013.211>
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008.
- Webber, Bonnie. 2013. What excludes an alternative in coherence relations? In *Proceedings, 10th International Conference on Computational Semantics*, pages 276–287.
- Webber, Bonnie, Rashmi Prasad, Alan Lee, and Aravind Joshi. 2019. The Penn Discourse Treebank 3.0 annotation manual. Philadelphia, University of Pennsylvania.
- Wellner, Ben, James Pustejovsky, Catherine Havasi, Anna Rumshisky, and Roser Sauri. 2006. Classification of discourse coherence relations: An exploratory study using multiple knowledge sources. In *Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue*, pages 117–125. <https://doi.org/10.3115/1654595.1654618>
- Xia, Tingyu, Yue Wang, Yuan Tian, and Yi Chang. 2021. Using prior knowledge to guide BERT’s attention in semantic textual matching tasks. In *WWW ’21: The Web Conference 2021, Virtual Event*, pages 2466–2475. <https://doi.org/10.1145/3442381.3449988>
- Xiang, Ming and Gina Kuperberg. 2015. Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience*, 30(6):648–672. <https://doi.org/10.1080/23273798.2014.995679>, PubMed: 25914891
- Xu, Xiaodong, Qingrong Chen, Klaus-Uwe Panther, and Yicheng Wu. 2018. Influence of concessive and causal conjunctions on pragmatic processing: Online measures from eye movements and self-paced reading. *Discourse Processes*, 55(4):387–409. <https://doi.org/10.1080/0163853X.2016.1272088>
- Yang, Zhilin, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R. Salakhutdinov, and Quoc V. Le. 2019. XLNet: Generalized autoregressive pretraining for language understanding. In *Advances in Neural Information Processing Systems*, pages 5753–5763.
- Zeldes, Amir, Yang Janet Liu, Mikel Iruskieta, Philippe Muller, Chloé Braud, and Sonia Badene. 2021. The DISRPT 2021 shared task on elementary discourse unit segmentation, connective detection, and relation classification. In *Proceedings of the 2nd Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2021)*, pages 1–12. <https://doi.org/10.18653/v1/2021.disrpt-1.1>