

An Assessment of Explicit Inter- and Intra-sentential Discourse Connectives in Turkish Discourse Bank

Deniz Zeyrek, Murathan Kurfalı

Informatics Institute, Middle East Technical University, Ankara,
dezeyrek@metu.edu.tr, kurfali@metu.edu.tr

Abstract

The paper offers a quantitative and qualitative analysis of explicit inter- and intra-sentential discourse connectives in Turkish Discourse Bank, or TDB version 1.1, a multi-genre resource of written Turkish manually annotated at the discourse level following the goals and principles of Penn Discourse TreeBank. TDB 1.1 is a 40K-word corpus involving all major discourse relation types (explicit discourse relations at intra- and inter-sentential positions, implicit discourse relations, alternative lexicalizations and entity relations) along with their senses and the text spans they relate. The paper focuses on the addition of a new set of explicit intra-sentential connectives to TDB 1.1, namely converbs (a subset of subordinators), which are suffixal connectives mostly corresponding to subordinating conjunctions in European languages. An evaluation of the converb sense annotations is provided. Then, with corpus statistics, explicit intra- and inter-sentential connectives are compared in terms of their frequency of occurrence and with respect to the senses they convey. The results suggest that the subordinators tend to select certain senses not selected by explicit inter-sentential discourse connectives in the data. Overall, our findings offer a promising direction for future NLP tasks in Turkish.

Keywords: Turkish, discourse relations, intra-/inter-sentential connective

1. Introduction

Discourse parsing is a challenging task for NLP. It involves various subtasks, such as discourse connective detection, argument detection and sense prediction. Since the release of discourse treebanks, particularly PDTB (Prasad et al., 2014), work on discourse parsing has gained an impetus. But most of this work is on English or European languages; work on other languages is scarce. Turkish Discourse Bank, or TDB, a multi-genre corpus of written Turkish, has been created with this motivation. It has the goal of providing researchers with a clearly defined level of discourse structure and semantics and support NLP and LT research in Turkish (Zeyrek and Webber, 2008).¹ Following the rules and principles of PDTB, TDB (version 1.0) has annotated 8483 relations made salient by explicit discourse connectives, involving the discourse connective itself as well as the two text spans it relates (Demirşahin and Zeyrek, 2017).

Discourse relations not only hold between adjacent sentences (inter-sententially) but can also exist between clauses within a single sentence (intra-sententially). While TDB 1.0 is a richly annotated resource of explicit discourse connectives, it does not annotate all types of intra-sentential discourse connectives and leaves out senses as well as other major discourse relation types that PDTB annotates. Given the need to create a more complete version of TDB and considering budgetary constraints, we have decided to enhance a modest portion of TDB 1.0 with new annotations. Thus, TDB 1.1, a 40K-word subcorpus has been created as described in Zeyrek and Kurfalı (2017). The subcorpus includes a collection of 20 text files (each with 2.000 words) distributed according to the genres covered by TDB 1.0 with the following frequencies: fiction (novel and short

story) (35%), news (30%), research monograph (2%), magazine article (2%), memoir (2%), interview (1%).²

Previous research has shown that parsing decisions conditioned on whether the relation is intra- or inter-sentential yields more effective parsing than decisions based on a single model approach (Stepanov and Riccardi, 2013). In this paper, our focus is what we can gain by considering the distribution of intra- and inter-sentential discourse relations across the corpus and by considering their senses. To this end, we have added converbs, a group of explicit intra-sentential discourse connectives, which were missing in TDB 1.1. Briefly, converbs are suffixal connectives and are a subset of subordinators, which also include postpositions. We first describe this enhancement on the corpus. Then, we assess the distribution of explicit inter- and intra-sentential discourse connectives and evaluate the converb annotations with respect to their senses. The corpus statistics suggest that (a) explicit intra-sentential connectives quantitatively overweigh explicit inter-sentential connectives, (b) explicit intra-sentential discourse connectives belonging to the subordinator type tend to select certain senses not selected by explicit inter-sentential connectives. We suggest that these findings are promising for future discourse parsing efforts of Turkish.

The rest of the paper is organized as follows: §2. starts with a brief overview of the TDB project and the underlying approach to discourse is provided. It continues with §2.1., a section on the major annotation categories and a description of intra- and inter-sentential connections. In §2.2., the addition of converbs is explained and an evaluation with respect to their sense tags is presented. §3. introduces the results of a quantitative analysis of explicit inter-sentential and explicit intra-sentential discourse connectives in TDB 1.1, providing frequency lists and examples from the cor-

¹In creating TDB, we are concerned with the local level of discourse, a term we use for low-level relations such as discourse relations (Hobbs, 1985). Thus, as in PDTB, we do not commit ourselves to a description of discourse at the global level.

²Information regarding TDB can be accessed at: <http://www.textlink.ii.metu.edu.tr/turkish-discourse-bank>.

pus. §4. includes an analysis of intra- vs. inter-sentential connections based on senses and discusses the possible implications of the current analysis on Turkish discourse parsing research. Finally, §5. summarizes and concludes the study.

2. Goals and Principles of TDB

In this section, the major annotation categories of TDB are explained as the basis of our work in the current paper. The annotation scheme reflects our approach to discourse, which we adopt from years of research in discourse as well as the PDTB principles.

Our annotation scheme aims to capture discourse relations such as contrast, expansion, contingency, etc. that hold between two text spans. As in the PDTB framework, we take discourse relations as a lexically-grounded phenomenon, where each discourse relation is anchored to a discourse connective. The annotations are created by a tool specifically designed for the TDB project. The TDB tool is a Java-based infrastructure with output representations in XML. It uses standoff annotation methodology, where the beginning and end offsets of the annotated content are kept (Aktaş et al., 2010).

We refer to the lexical anchors of discourse relations as discourse connectives (DCs), which are often (but not limited to) syntactic classes, e.g. conjunctions (*ve* 'and', *ya da* 'or', *çünkü* 'because') and adverbs (*ne var ki* 'nevertheless', *ayrıca* 'in addition'). DCs always relate text spans with an abstract object interpretation, i.e. eventualities, facts, propositions (Asher, 2012), which are referred to as the arguments of a discourse connective and tagged as Arg1 and Arg2. These tags do not indicate any kind of ordering in texts; rather, Arg2 is the text span that is syntactically related to the DC; Arg1 is the other argument. In Turkish, conjunctions and adverbials typically have the Arg1-Arg2 argument ordering.

An important problem to solve in the PDTB-style discourse annotation is to tease apart the discourse and non-discourse role of connectives. For example, in *apples and pears, and* is not a discourse connective since the text spans it relates do not have abstract object interpretations. In TDB, only the discourse use of connectives are annotated, eliminating the non-discourse uses by eye and leaving them unannotated.

In Turkish, discourse connectives are not only represented lexically but also morphologically. Morphology is important particularly for intra-sentential connections. Thus, we distinguish a set of intra-sentential DCs, referred to as subordinators subsuming postpositions (*için*, 'in order to, for the purpose of', *gibi* 'as, like') and converbs (e.g. *-sa* 'if', *-ken* 'meanwhile'). Converbs are a unique aspect of Turkish and other Turkic languages reflecting the role of morphology in clause combining. They usually correspond to subordinate conjunctions typical of European languages (Johanson, 1995). The clause combining role of converbs and postpositions is subordination and they have the same ordering of arguments, i.e. Arg2-Arg1.³

³Thus, postpositions are named as complex subordinators,

Examples (1) and (2) illustrate two converbs -*sa* 'if' and -*CA* 'as'.⁴ Here and in the rest of the paper, we show discourse connectives by underlining. Arg1 is rendered in italic fonts, Arg2 in bold fonts. Each connective is annotated with senses from the PDTB 2.0 sense hierarchy, presented in Fig.1. Where relevant, the sense tag of the relation is presented in square parentheses.

- (1) **Izini bulursanız**, *bu numaraya haber verin.*
If you find his trail, *call this number.* [Contingency:Condition]
- (2) **İndikçe**, *mahzende beliren rafları... gördü.*
As he descended, *he saw the shelves ... in the cellar.* [Temporal:Synchronous]

2.1. Relation Types Annotated in TDB 1.1

TDB 1.1 annotates four types of discourse relations: (i) relations with an overt lexical or morphological signal (conjunctions, adverbials, subordinators), (ii) implicit relations, where the relations lack an overt signal, (iii) alternative lexicalizations and (iv) entity relations.⁵ The annotations are created by determining and selecting the discourse connective span together with its binary arguments and senses.

Implicit relations: Where a discourse relation is not made explicit by a connective, readers can still infer a relation; these have been referred to as implicit discourse relations. In PDTB, when an implicit relation is inferred between text spans, the annotator is asked to insert an explicit connective that makes the relation salient. A sample annotation following the PDTB guidelines is provided in (3).

- (3) *Nihayet o da satıldı.* (Implicit=sonuçta) **Mülk olarak elde Süleymaniye'deki konak ile Küçük Çamlıca'daki köşkten başka bir şey kalmadı.**
Eventually, it was sold too. (Implicit=as a result) **There was nothing in their hand as property except the villa in Süleymaniye and the manor house in Çamlıca.** [Contingency:Cause:Result]

Alternative lexicalizations: Discourse relations can also be expressed by other means of lexicalizing a relation (AltLex), such as *because of this* (Prasad et al., 2010). In TDB, we annotate such cases as "phrasal expressions" (Zeyrek et

converbs are referred to as simplex subordinators (Zeyrek and Webber, 2008).

⁴Throughout the paper, all examples are rendered in Turkish orthography. We use capital letters to capture vowel or consonant harmony, which are operative in stems as well as suffixes. For example, the converb 'if' in example (1) has two surface forms due to vowel harmony: *-se* and *-sa*. We use H to indicate any high vowel, A to indicate e or a (see Table 5). Similarly, the converb 'as' in example (2) has 4 surface forms due to vowel as well as consonant harmony: *-ca*, *-ce*, *ça*, *çe*. To capture consonant harmony, i.e. the harmony of a voiced consonant with its voiceless counterpart as in c and ç (the letters for voiced and voiceless affricates, respectively), we use C. To indicate the harmony of d and t, we use D. (see Table 7).

⁵PDTB also annotates no relations. We leave annotation of no relations for further work.

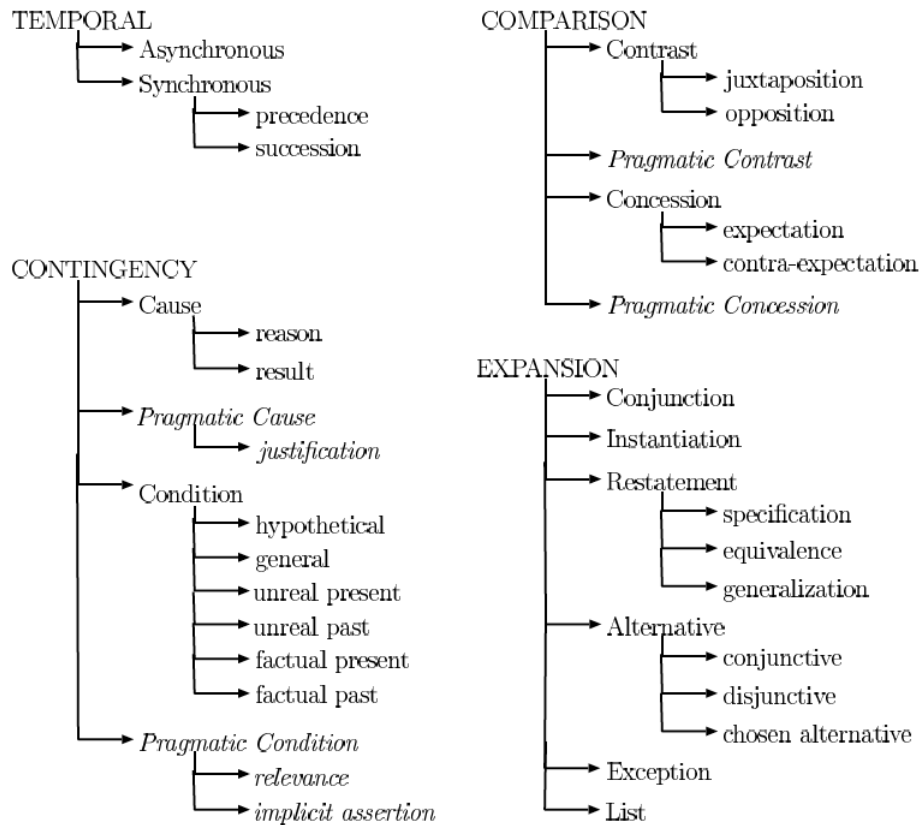


Figure 1: PDTB 2.0 sense hierarchy (Prasad et al., 2014)

al., 2013). These are devices that are often compositionally formed on the basis of a core element, such as a post-position *rağmen* 'despite' resulting in expressions as *buna rağmen* 'despite this'. A sample annotation is provided in (11) below.⁶ Phrasal expressions form a subset of alternative lexicalizations in the PDTB framework (Prasad et al., 2010).

Entity relations: It is possible for a discourse relation to hold between two entities rather than eventualities. Such cases are referred to as Entity Relations (EntRel) and are not assigned a sense tag in PDTB-style annotation, as shown in example (4).

- (4) *Aşıklı Höyük bu yerleşimlerden biri. Aksaray ilinin Kızılkaya Köyü'nün yakınında, Melendiz Nehri'nin kıyısında yer alıyor.*
The Aşıklı Tumulus is one of these settlements. It is located by the Melendiz River near the Kızılkaya village of Aksaray province. [EntRel]

Implicit relations, alternative lexicalizations and entity relations are out of scope of this study but have been overviewed here for the sake of completeness. In the rest of the paper, the term *discourse connective* refers to *explicit* DCs unless otherwise stated.

⁶In TDB 1.0, around 5% of the annotated relations consists of phrasal expressions (Demirşahin and Zeyrek, 2017).

As mentioned above, discourse relations may hold both inter-sententially and intra-sententially. Thus, a DC can be categorized as inter-sentential (inter-S) or intra-sentential (intra-S) according to where its arguments are situated in the text. Intra-S connectives relate two text spans situated in the same sentence. We use orthographic conventions in postprocessing the data in terms of inter- and intra-S connections. Typically, in Turkish, arguments of an intra-S connective are not separated by terminal punctuation marks, e.g. '!', '!', '?' (example 5). Arguments of an inter-S connective are in different sentences (example 6).

- (5) *Müthiş soğuk bir gündü ve tipi şeklinde kar yağıyordu.*
It was an extremely cold day and there was a blizzard. [Expansion:Conjunction]
- (6) *Kızınca bir çocuk kadar bile olamazdım. Bir teneye tekme atamazdım, mesela.*
Even when I was angry, I would not be able to act like a child. I could not even kick the bin, for instance. [Expansion:Instantiation]

2.2. Annotation Procedure of Converbs and Evaluation

As with other connectives, converbs have a DC and a non-DC use and to distinguish these cases is important. The abstract object criterion is helpful, showing us when to take a converb as a DC. For example, sentence (7)

can only be interpreted as having a single abstract object interpretation, blocking the interpretation of an *entering event* separate from the *running event*; thus in this case, the converb -ArAk 'by (means of)' fulfills a non-DC role. As opposed to this, in example (8), the converb has a DC role since it relates two abstract objects (a *questioning event* and a *referring event*) expressed in the arguments.

- (7) Ayşe eve koşarak girdi.
Ayşe walked into the house (by) running.
- (8) .. **bir bakıma kendine de gönderme yaparak**,
yazılış mantığını sorgular.
.. *he questions the wording (by) referring to himself in a way.*

In Turkish, when a converb has a DC role, it is always an explicit intra-S connective. Thus, with the addition of converbs, we extend the coverage of explicit intra-S DCs in TDB 1.1. This allows us to provide an assessment of explicit intra-S and inter-S connectives, which is the focus of our study.

Converb annotations are realized in two phases. First, connective-argument annotations are created by two independent annotators who go over the whole corpus determining the DC use, eliminating the NDC use of converbs. The resulting annotations are discussed in meetings, where agreed versions are produced by a unanimous decision and recorded. In the second phase, the converb-argument annotations are checked for their correctness by two independent annotators and sense tags are assigned; then the cycle is repeated. The inter-annotator agreement for the senses at all levels of the sense-hierarchy is strong, i.e. ≥ 0.8 (see Table 1).⁷

Sense	IAA
Level-1	89.5%
Level-2	81.9%
Level-3	80.0%

Table 1: IAA results of converb Level-1 senses in TDB 1.1

Overall, 12 unique tokens of converbs realized by 38 different surface forms amounting to 105 converb tokens are added to the data. Hence, the number of explicit connectives increased in the data. The current coverage of TDB 1.1 is shown in Table 2 (Zeyrek and Kurfalı, 2017).

3. Distribution of Intra- and Inter-sentential Connectives in the Corpus

A quantitative analysis of explicit intra-S DCs (including converbs) and inter-S DCs shows that explicit intra-S DCs are more frequent than explicit inter-S DCs, as shown in Table 3.

⁷The converb annotations were created by Işın Demirşahin, Ahmet Faruk Acar, Arzu Burcu Güven and Nihan Soyca, post-graduate students at Cognitive Science Department, Middle East Technical University. The IAA is measured by the exact match method (Miltakaki et al., 2004).

TYPE	Frequency
Explicit	868
Implicit	407
EntRel	541
AltLex	108

Table 2: Distribution of discourse relation types in TDB 1.1 including the converbs added to Explicit relations

TYPE	Frequency	Ratio
Inter-S Explicit DR	688	79.2%
Intra-S Explicit DR	180	20.8%

Table 3: Distribution of inter-S and intra-S explicit discourse relations in TDB 1.1

Analysis of inter-sentential connectives: We find that among explicit inter-S DCs, *ama* ('but, yet') is the most frequent DC as well as the most ambiguous one with 7 different sense tags. Table 4 presents the 10 most frequent explicit inter-S connectives together with their canonical syntactic types, which constitute 82.7% of all the explicit inter-S tokens in the corpus.

DC	Gloss	Syntactic type	Frequency
ama	<i>but</i>	conjunction	48
çünkü	<i>because</i>	conjunction	21
ayrıca	<i>in addition</i>	adverb	17
sonra	<i>then</i>	adverb	13
ancak	<i>however/yet</i>	adverb	12
oysa	<i>however</i>	adverb	12
fakat	<i>but</i>	conjunction	7
aslında	<i>in fact</i>	adverb	7
ve	<i>and</i>	conjunction	7
önce	<i>before</i>	adverb	5

Table 4: 10 most frequent inter-S DCs and their canonical syntactic types

Example (9) presents one of the inter-S uses of *ama*.

- (9) *Birden geldiğini duydum. Ama, göremedim onu. Suddenly, I heard her come. Yet, I could not see her.* [Comparison: Concession: contra-expectation]

Analysis of intra-sentential connectives: Of all the explicit intra-DCs in the corpus, *ve* 'and' is the most frequent connective comprising 28.3% of the explicit intra-S DCs and 23.2% of the explicit DCs in TDB 1.1 (Table 5). Two converbs, namely -HnCA 'when' and -ken 'meanwhile' are among the most frequent explicit intra-S DCs in the corpus. Example (10) below shows the purpose sense of the post-position *için*, the second most frequent explicit intra-S DC in the corpus.

- (10) **Onu görmek için** tüm zamanınızı parkta geçirmeye başlarsınız.

DC	Gloss	Syntactic type	Frequency
ve	<i>and</i>	conjunction	195
için	<i>since/in order to</i>	postposition	91
ama	<i>but</i>	conjunction	71
sonra	<i>after</i>	postposition	58
-HncA	<i>when</i>	converb	22
gibi	<i>as/like</i>	postposition	18
çünkü	<i>because</i>	conjunction	17
-ken	<i>meanwhile</i>	converb	16
ancak	<i>however/yet</i>	adverb	15
kadar	<i>to the degree that</i>	postposition	13

Table 5: 10 most frequent explicit intra-S DCs and their canonical syntactic types

You would start to spend all your time in the park in order to see her. [Contingency:Purpose]

4. An Analysis Based on Senses

Table 6 shows the distribution of explicit intra- and inter-S DCs across Level-1 senses.⁸

Sense	Expl. Inter-S DCs		Expl. Intra-S DCs	
	Count	Ratio	Count	Ratio
Contingency	35	0.19	167	0.24
Temporal	25	0.14	149	0.22
Comparison	82	0.45	126	0.18
Expansion	36	0.22	246	0.36

Table 6: Distribution of Level-1 senses among explicit inter- and explicit intra-S discourse relations

The numbers in bold fonts indicate the most frequently occurring instances. Table 6 also shows that while Comparison is the most frequent Level-1 sense selected for explicit inter-S connectives, Expansion is the most common Level-1 sense chosen for explicit intra-S connectives.

4.1. Observations on the Sense of Intra-sentential Discourse connectives

In this section we offer two observations. First, we find that there are several second- or third-level senses which seem to be preferred by subordinators (both postpositions and converbs) as shown in Table 7.⁹

We note that most of the senses in Table 7 can well be conveyed by means of an AltLex containing the core lexeme of the intra-S connective, as in the use of the postposition

⁸TDB allows assigning multiple senses to a discourse relation. Therefore, in Table 6, overall, there are more senses than the number of discourse relations.

⁹*Manner* and *Degree* have been introduced on the basis of Turkish data (Zeyrek and Kurfalı, 2017). *Purpose* has been borrowed from PDTB 3.0 sense hierarchy (Webber et al., 2016).

rağmen 'despite' in the phrasal expression *buna rağmen* 'despite this' (see example (11)). We conjecture that this implies the ubiquity of the senses selected by postpositions or phrasal expressions derived from them.

Sense	Freq.	DC/Gloss	Syntactic type of DC
Contingency: Purpose	79	için (<i>since/in order to</i>)	postposition
Temporal: Synchronous	57	-DHğHndA (<i>when</i>)	converb
Expansion: Manner	28	kadar (<i>as</i>)	postposition
Comparison: Concession: expectation	22	rağmen (<i>despite</i>)	postposition
Comparison: Degree	11	kadar (<i>to the degree that</i>)	postposition

Table 7: The most frequent second- or third-level senses selected only by converbs or postpositions in their discourse connective roles

- (11) *..pistin görülmediği anlaşılmış, buna rağmen radar yardımı istenmemiştir.*
..it was understood that the airfield was not seen; despite this, radar help was not requested.
 [Comparison:Concession:contra-expectation]

Secondly, Table 8 shows that three explicit temporal DCs exhibit a tendency towards Asynchronous:precedence or Asynchronous:succession depending on whether they establish inter- or intra-level connection. By definition, in precedence relations, Arg1 precedes Arg2; in succession relations, Arg2 precedes Arg1. The inter-S *sonra* 'later' (syntactically an adverb) exhibits Arg1-Arg2 ordering and hence prefers the Temporal:Asynchronous:precedence sense. On the other hand, the intra-S *sonra* 'after' (syntactically a postposition) displays Arg2-Arg1 ordering and tends to convey the Temporal:Asynchronous:succession sense. Examples (12) and (13) exemplify argument ordering and the sense of an adverb and a postposition, respectively.

- (12) *Durun, oturun biraz. Anlatın. Gidersiniz sonra.*
Wait, have a sit for a while. Tell us. You can leave later. [Temporal:Asynchronous:precedence]

[adverb]

- (13) *İyice kendine geldikten sonra getirilen kuru elbiseleri giydi.*
After regaining his consciousness, he wore the dry clothes they brought. [Temporal:Asynchronous:succession]

[postposition]

4.2. Implications

The differences between inter- and intra-sentential relations in discourse have long been recognized and employed

DC/Syntactic type	Sense (Temporal:Asynchronous)
ardından (inter-S)/Adv ardından (intra-S)/P	precedence (2/2) succession(2/2)
sonra (inter-S)/Adv sonra (intra-S)/P	precedence (13/13) succession (58/58)
önce (inter-S)/Adv önce (intra-S)/P	succession (5/5) precedence (8/8)

Table 8: Precedence or succession senses selected by 3 explicit Temporal:Asynchronous connectives. The numbers in parentheses refer to the number of times the subsense is selected and the number of all temporal relations selected for the connective, respectively. Adv stands for adverb, P for postposition.

with success in computational analysis of discourse involving automatic argument extraction and sense labeling (Sporleder and Lascarides, 2008; Joty et al., 2013; Liu and Lapata, 2017; Braud and Denis, 2014). In the current study, for the first time, we offered corpus statistics from Turkish regarding the distribution of inter- and intra-sentential relations with respect to their senses to set the basis for future computational analysis and discourse parsing studies. Our findings suggest that whether or not a given explicitly marked relation holds inter- or intra-sententially may carry valuable information for Turkish as well. Therefore, the reported results are promising for future efforts on automatic sense disambiguation of explicit discourse connectives in Turkish.

5. Summary and Conclusion

This paper started with new enrichments on TDB 1.1 involving a new type of explicit intra-S DCs - converbs, added to the corpus as a further type of subordinator discourse connectives. With the addition of converbs, two goals have been accomplished: (a) a unique aspect of Turkish morphology that has a bearing on clause combining and hence intra-sentential discourse relations has been captured in the corpus, and (b) the overall frequency of intra-sentential explicit discourse connectives increased. This enhancement allowed us to make a comparison between explicit intra- and inter-sentential discourse connectives in TDB 1.1. With various corpus statistics, the paper showed that explicit intra-S DCs occur more frequently than explicit inter-S DCs and it presented evidence for the senses preferred by explicit intra-sentential connectives, namely subordinators but not by explicit inter-S discourse connectives. The paper thus argued that the analysis of explicit intra- and inter-S DCs with respect to the senses they choose is a promising direction for further discourse parsing studies on Turkish. Our aim for the future is to implement the ideas that arise from the current work on various NLP tasks.

6. Acknowledgements

This research was partially supported by METU Project Funds (project ID: BAP-07-04-2017-001). We thank Dilek

Deniz Bilgiç for recording the converb annotations and Tunç Güven Kaya for his help in postprocessing the data. We are grateful to our three anonymous reviewers. All remaining errors are ours.

7. References

- Aktaş, B., Bozsahin, C., and Zeyrek, D. (2010). Discourse relation configurations in Turkish and an annotation environment. In *Proceedings of the fourth linguistic annotation workshop*, pages 202–206. Association for Computational Linguistics.
- Asher, N. (2012). *Reference to abstract objects in discourse*, volume 50. Springer Science & Business Media.
- Braud, C. and Denis, P. (2014). Combining natural and artificial examples to improve implicit discourse relation identification. In *COLING*.
- Demirşahin, I. and Zeyrek, D. (2017). Pair annotation as a novel annotation procedure: The case of Turkish Discourse Bank. In *Handbook of Linguistic Annotation*, pages 1219–1240. Springer.
- Hobbs, J. R. (1985). On the coherence and structure of discourse. *Technical report CSLI-85-37, Center for the Study of Language and Information, Stanford University*.
- Johanson, L. (1995). On Turkic converb clauses. *Haspelmath & König (eds.)*, 1995:313–48.
- Joty, S. R., Carenini, G., Ng, R. T., and Mehdad, Y. (2013). Combining intra- and multi-sentential rhetorical parsing for document-level discourse analysis. In *ACL (1)*, pages 486–496.
- Liu, Y. and Lapata, M. (2017). Learning contextually informed representations for linear-time discourse parsing. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1300–1309.
- Miltsakaki, E., Prasad, R., Joshi, A. K., and Webber, B. L. (2004). The Penn Discourse Treebank. In *LREC*.
- Prasad, R., Joshi, A., and Webber, B. (2010). Realization of discourse relations by other means: alternative lexicalizations. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, pages 1023–1031. Association for Computational Linguistics.
- Prasad, R., Webber, B., and Joshi, A. (2014). Reflections on the Penn Discourse Treebank, comparable corpora, and complementary annotation. *Computational Linguistics*, 40(4):921–950.
- Sporleder, C. and Lascarides, A. (2008). Using automatically labelled examples to classify rhetorical relations: An assessment. *Natural Language Engineering*, 14(3):369–416.
- Stepanov, E. A. and Riccardi, G. (2013). Comparative evaluation of argument extraction algorithms in discourse relation parsing. In *13th International Conference on Parsing Technologies (IWPT 2013)*, volume 36, page 44.
- Webber, B., Prasad, R., Lee, A., and Joshi, A. (2016). A discourse-annotated corpus of conjoined VPs. *LAW X*, page 22.
- Zeyrek, D. and Kurfalı, M. (2017). TDB 1.1: Extensions on Turkish Discourse Bank. *LAW XI 2017*, pages 76–81.

- Zeyrek, D. and Webber, B. L. (2008). A discourse resource for Turkish: Annotating discourse connectives in the METU Corpus. In *IJCNLP*, pages 65–72.
- Zeyrek, D., Demirsahin, I., Sevdik-Callı, A., and Çakıcı, R. (2013). Turkish Discourse Bank: Porting a discourse annotation style to a morphologically rich language. *Dialogue & Discourse*, 4(2):174–184.