

Unifying Discourse Resources with Dependency Framework

Yi Cheng¹ Sujian Li¹ Yueyuan Li²

¹MOE Key Lab of Computational Linguistics, School of EECS, Peking University

²School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University
{yicheng, lisujian}@pku.edu.cn, rowena.lee@sjtu.edu.cn

Abstract

For text-level discourse analysis, there are various discourse schemes but relatively few labeled data, because discourse research is still immature and it is labor-intensive to annotate the inner logic of a text. In this paper, we attempt to unify multiple Chinese discourse corpora under different annotation schemes with discourse dependency framework by designing semi-automatic methods to convert them into dependency structures. We also implement several benchmark dependency parsers and research on how they can leverage the unified data to improve performance.¹

1 Introduction

Discourse parsing aims to construct the logical structure of a text and label relations between discourse units, as shown in Fig. 1 and Fig. 2. Various discourse corpora have been developed to promote the discourse parsing technique. There exist multiple discourse schemes such as Rhetorical Structure Theory (RST) (Carlson et al., 2001) and PDTB (Prasad et al., 2008), which act as guidelines of various discourse corpora. At the same time, text-level discourse annotation is complicated and laborious, so the scale of a single discourse corpora is often much smaller compared with other NLP tasks.

One way to conquer data sparsity is to use different discourse corpora through multi-task learning. One way to conquer the data-sparsity problem is to multi-task learning multiple corpora simultaneously. Prior efforts on that mainly focus on discourse relation classification between two text spans, without considering the whole discourse structure (Liu et al., 2016; Li et al., 2017), since one principle of multi-task learning is that the tasks should be closely related, but the discourse structures of different corpora vary a lot, e.g. shallow predicate-arguments structure (e.g. the relations in Fig. 1) and deep tree structure (e.g. the CDTB tree in Fig. 2).

In this work, we explore another possible way to simultaneously leverage different corpora: we unify the existing discourse corpora under one same framework to form a much larger dataset. Our choice for this unified framework is dependency discourse structure (DDS), where elementary discourse units (EDUs) are directly related to each other without consideration of intermediate text spans. Because as Morey et al. (2018) have argued, DDS is a very general discourse representation framework, and they adopt a dependency perspective to evaluate the English discourse corpora RST-DT (Carlson et al., 2001) and the related parsing techniques.

Our work explore the feasibility of unifying discourse resources under dependency framework with consideration of the following two questions: (i) How to convert other discourse structures into DDS? (ii) How to make the best use of the unified data to improve discourse parsing techniques?

Oriented by the questions above, we unify three Chinese discourse corpora under dependency framework: HIT-CDTB (Zhang et al., 2014), SU-CDTB (Li et al., 2014b) and Sci-CDTB (Cheng and Li, 2019). HIT-CDTB adopts the predicate-argument structure similar to PDTB, with a connective as predicate and two text spans as arguments. Following rhetorical structure theory (RST), SU-CDTB uses a hierarchical tree to represent the inner structure of each text, with EDUs as its leaves and connectives as intermediate

¹The data is available at <https://github.com/PKU-TANGENT/UnifiedDep>.

nodes. Sci-CDTB is a small-scale DDS corpus composed of 108 scientific abstracts. It is the only Chinese DDS corpus as far as we know.

The primary obstacle of unifying these corpora is inconsistency of the representation schemes, such as granularity of EDU and definition of relation types. Besides, the predicate-argument structure of HIT-CDTB leads to the problem that some discourse relations between adjacent text spans are absent, so that the information provided by the original annotation might be insufficient to form a complete dependency structure. To tackle these problems, we redefine granularity of EDU, conduct mapping among different relation sets, and design semi-automatic methods to convert other discourse structures into DDS. As the automatic part, we design the dependency tree transformation method for each corpora. As the manual part, we proofread all the segmentation of EDUs to follow the same definition, complement necessary information, and correct the transformed dependency trees.

Different from [Morey et al. \(2018\)](#) who only consider conversion between RST and DDS, we attempt to unify more discourse schemes into dependency framework and explore whether discourse parsing techniques can be promoted by the unified dataset. Here we implement several discourse dependency parsers and research through experiment on how they can leverage the unified data to improve performance. Then we give out our findings about how to make better use of the unified discourse data.

Contributions of this paper are summarized as follows:

- we propose to integrate the existing discourse resources under a unified framework;
- we design a unified DDS framework and convert three Chinese discourse corpora into dependency structure in a semi-automatic way and get a unified large-scale dataset;
- we implement several discourse dependency parsers and explore how the unified data can be leveraged to improve parsing performance.

2 Background

In this section we mainly introduce the three Chinese discourse corpora: HIT-CDTB, SU-CDTB and Sci-CDTB, which we use in this work.

HIT-CDTB Borrowing the discourse scheme of PDTB, HIT-CDTB adopts one-predicate two-arguments structure, where a connective serves as the predicate and two text spans as two arguments, as shown in Fig. 1. The connective can be either explicitly identified if it already exists in the original text, or implicitly added by annotators, while the arguments can be phrases, clauses, sentences, or sentence groups. Each connective corresponds to a relation type. In total, there are 4 coarse-grained types (i.e., *temporal causal*, *comparative* and *extension*) and 22 fine-grained types (e.g., *temporal* is further divided as *synchronous* and *asynchronous*). The documents cover multiple domains, such as news, editorials and popular science articles. When labeling each document, the discourse relations are labeled locally by only considering adjacent text spans, so a complete logical structure of the text may not be obtained.

SU-CDTB Similar to RST, SU-CDTB represents the inner structure of a text with a hierarchical tree with EDUs as leaves, and uses connectives as intermediate nodes to indicate rhetorical relations. The whole text is first divided into several text spans which are recursively divided until getting EDUs. In SU-CDTB, EDUs are segmented according to the punctuation marks. The nucleus-satellite relation structure in RST is retained through arrows in the trees. Connectives in SU-CDTB are also given some relation attributes. 500 news documents from the Chinese Treebank ([Xue et al., 2005](#)) are annotated. A discourse tree is constructed for each paragraph rather than the entire document, with an average of 4.5 EDUs per tree, which are relatively shallow. Besides, the top-down constructed tree cannot cover some particular discourse structures, such as non-adjacent relations. So parsers trained only with this corpus are not very likely to analyze more complex discourse logic.

Sci-CDTB Sci-CDTB is a small Chinese discourse dependency corpus, where EDUs are directly connected with discourse relations without intermediate levels. Its definition of EDU refers to RST-DT, with some modifications based on the linguistic characteristics of Chinese. The head of each EDU and the relations between them are annotated. Sci-CDTB is a small-scale corpus, only composed of 108 annotated

Inner-sentence relation:

- S₁: [一渔船昨晚出海时]u₁ [冷动机突然发生故障。]u₂
 When a fishing boat was sailing out at sea last night, its cold motor broke down.
- S₂: [结果造成机长和3名渔工眼睛受伤]u₃ [还合并有呕吐现象]u₄
 As a result, the captain and three fishermen suffered eye injuries, and kept vomiting.
- S₃: [不过当时渔船已经失去了动力]u₅ (所以^{causality}) [受伤的人员只能在渔船上等待救援]u₆
 But by then the boat had lost power, (so) the injured can only wait on the boat to be rescued.

Inter-sentence relation: S₁ 结果^{causality} S₂

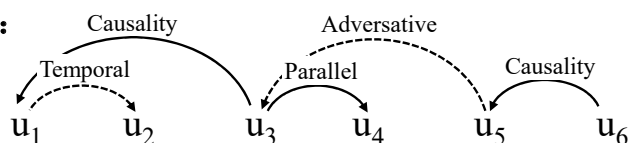
Conversion result:

Figure 1: An example text in HIT-CDTB and its conversion result of DDS. Underlined texts are arguments of labeled HIT-CDTB discourse relations. The dependency relations derived from the original corpus are represented with solid lines in the conversion result, while the ones complemented during conversion are with dotted lines.

scientific abstracts, so it is hard to support the training of a competitive discourse parser. It is the only Chinese DDS corpus to the best of our knowledge.

3 Unifying Discourse Corpora

In this section, we introduce how to convert the three discourse corpora into one unified framework, which mainly involves two aspects, i.e., EDU segmentation and dependency tree construction. Here, we adopt the EDU segmentation guideline of Sci-CDTB, which is similar to RST-DT (Carlson et al., 2001). Basic discourse units of SU-CDTB and HIT-CDTB are divided mainly based on punctuation marks, which is inconsistent with our EDU definition, so we manually re-segment their EDUs to ensure the same segmentation rules.

As for dependency tree construction, we should ensure correctness of both the tree structure and the relations between them. In a dependency tree, each relation connects a *head* EDU to a *dependent* EDU. Each EDU should have one and only one *head*. As the three corpora adopt different relation sets (22 relation types in HIT-CDTB_{dep}, 18 in SU-CDTB_{dep}² and 26 in Sci-CDTB), during conversion we keep the relations unchanged for each corpus, and then map them into 17 predefined relation types, which are basically the same as the ones of SU-CDTB, except that relation *example illustration* is merged into *explanation*. It requires to be further investigated whether there is a more appropriate mapping, and there exists the possibility that these relation sets have inherent incompatibility.

Experiments are conducted both before and after relation mapping.

3.1 Conversion of HIT-CDTB

Formally, a discourse relation in HIT-CDTB can be denoted as $R = \langle \overline{rel}_{HIT}, con, ARG1, ARG2 \rangle$, where \overline{rel}_{HIT} is a relation type, *con* a connective, and *ARG1*, *ARG2* are two arguments. The arguments can be phrases, clauses, sentences or sentence groups. We will use the example in Fig. 1 to elaborate the conversion process of HIT-CDTB.

For the example, HIT-CDTB labels two inner-sentence relations and one inter-sentence relation, which can not form a complete tree for the text. As preparation for DDS construction, EDU segmentation is conducted with an pre-trained segmenter (Yang and Li, 2018b), which are then manually checked to

²HIT-CDTB_{dep} and SU-CDTB_{dep} refer to the converted HIT-CDTB and SU-CDTB before relation mapping.

Text: [张三三十出头,]_a [既没有什么学历,]_b [又没有多少工作经验,]_c [但无论是什么的工作,]_d [他都尽全力做好,]_e [所以只要有重要的任务_{f1} 处长总是会交给他。 _{f2}]_f
 [Zhang is in his early thirties.]_a [He has neither competitive educational background]_b [nor rich experience.]_c [But whatever task assigned to him,]_d [he always tries his best to accomplish it.]_e [So as long as there is an important job _{f1} the director will delegate it to him. _{f2}]_f

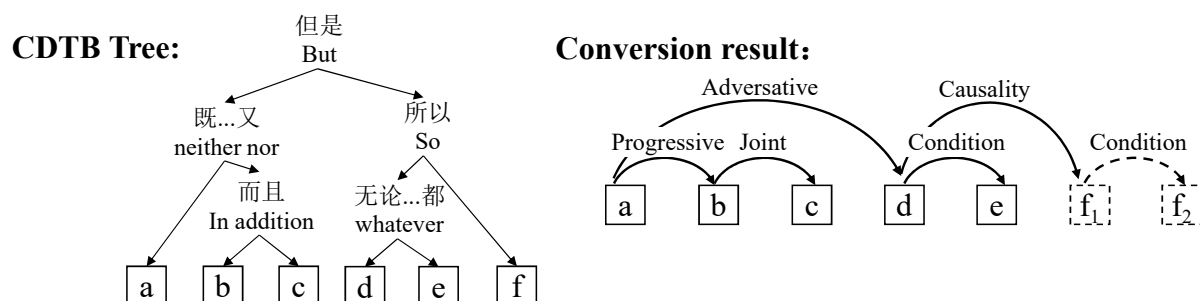


Figure 2: An example SU-CDTB tree and its conversion result of DDS

ensure quality. With EDUs as basic units, we utilize the original relations labeled in HIT-CDTB and complement some relations to construct a complete dependency tree. For instance, the relations derived from the original corpus and the ones complemented during conversion are respectively represented with solid lines and dotted lines in the conversion result of Fig. 1.

Specifically, assume that U_i is the EDU set covered by Arg_i ($i = 1, 2$), the absent relations are **first** complemented to construct a complete dependency subtree. We generate these complemented relations in a rule-based way, by summarizing common discourse markers for each relation type. For example, discourse markers for *Temporal* relation include “时”(when), “前”(before), “后”(after). **Then**, with two subtrees for U_1 and U_2 , we replace U_1 and U_2 with their root EDUs as the arguments of R' (e.g., *Causality* relation between u_1 and u_3 in Fig. 1). **Finally**, all the automatically-added relations are manually checked by two annotators.

3.2 Conversion of SU-CDTB

Like RST, SU-CDTB represents the inner structure of a text with a hierarchical tree, with EDUs as its leaves and connectives as its intermediate nodes. The “nucleus-satellite” structure in RST is retained through the arrows on the tree. In SU-CDTB, EDU segmentation is conducted according to punctuation marks. Fig. 2 shows an example discourse tree in SU-CDTB and its conversion result of DDS.

The conversion process consists of three steps. **First**, as the annotation scheme of SU-CDTB is similar to that of RST-DT, we conduct conversion in a way similar to the algorithm of transforming RST-style discourse trees into DDS (Li et al., 2014a; Muller et al., 2012; Hirao et al., 2013). **Then**, we further subdivide some of the EDUs because EDU granularity in SU-CDTB is larger than our definition. **Finally**, we manually label the relations between newly-segmented EDUs, such as f_1 and f_2 in Fig. 2. As SU-CDTB constructs a complete tree for each text, the conversion process is relatively labor-saving.

3.3 Corpus Statistics

Size of HIT-CDTB_{dep}, SU-CDTB_{dep} and Sci-CDTB are presented in Table 3.3. We also list the statistics of RST-DT and SciDTB (Yang and Li, 2018a), an English dependency corpus, for comparison. We can see that Sci-CDTB has a much smaller scale than HIT-CDTB_{dep} and SU-CDTB_{dep}. HIT-CDTB_{dep} and SU-CDTB_{dep} have similar number of relations, but the average length of each document (531.7 characters, 28.6 EDUs) in HIT-CDTB_{dep} is larger than that in SU-CDTB_{dep} (94.8 characters, 4.8 EDUs). Size of the unified DDS corpus Unified_{dep} is comparable to the two English discourse corpora.

Corpus	HIT-CDTB _{dep}	SU-CDTB _{dep}	Sci-CDTB	Unified _{dep}	SciDTB	RST-DT
#Doc	353	2332	108	2793	798	385
#Rel	9796	8181	1392	19369	18978	21787

Table 1: Corpus Size Comparison

Corpus	HIT-CDTB _{dep}		SU-CDTB _{dep}		Sci-CDTB		Unified _{dep}	
	Rel.	%	Rel.	%	Rel.	%	Rel.	%
Relation Distribution	joint	46.4	joint	53.2	elaboration	29.3	joint	52.7
	explanation	8.7	explanation	10.9	joint	17.0	explanation	16.7
	progressive	8.5	causality	8.3	enablement	9.9	causality	6.9
	expression	8.1	continuation	6.8	bg-general	9.7	continuation	4.3
	causality	6.2	goal	4.1	evaluation	6.1	progressive	4.1

Table 2: Distribution of the 5 most frequently-used relations in HIT-CDTB_{dep}, SU-CDTB_{dep} and Sci-CDTB

Two annotators have participated in the manual labeling work. It takes them about 3 months and 3 weeks respectively to do the manual annotation and checking for HIT-CDTB and SU-CDTB. In comparison, it takes two annotators 3 months to build Sci-CDTB from scratch, which has a much smaller scale than the converted data, showing that converting existing resources is a relatively efficient way to generate dependency discourse dataset. Conversion of SU-CDTB is less time-consuming because its discourse structure is more complete than HIT-CDTB. For HIT-CDTB, annotators take much more time on EDU segmentation and tree completion, since this corpus is constructed bottom up and its scheme is more different from our discourse dependency framework. Table 2 shows distribution of the five most frequent relations in each corpus. We can see that the relation distribution in Sci-CDTB is quite different from the other corpora.

4 Dependency Discourse Parsing

Baselines Following the work of Yang and Li (2018a), we implement four Chinese discourse dependency parsers:

- **Graph-based Parser** adopts Eisner algorithm to predict the most possible dependency tree structure (Li et al., 2014b). It refers to the graph-based method of syntax parsing, which adopts Eisner algorithm and MST algorithm to predict the most possible dependency tree structure. For simplicity, an averaged perceptron is implemented to train weights.
- **Vanilla Transition-based Parser** adopts the transition method of dependency parsing proposed by Nivre (2003), employing the action set of arc-standard system (Nivre et al., 2004). An SVM classifier is trained to predict transition action for a given configuration.
- **Two-stage Transition-based Parser** (Wang et al., 2017) first adopts the transition-based method to construct a naked dependency tree, and then uses another SVM to predict relations, which can take the tree-structure as features.
- **Bert-based Parser** also conducts parsing in a two-stage transition way, but in the second stage it uses a bert-based model, which incorporates BERT (Devlin et al., 2018) with one additional output layer, to identify relation types. It keeps the pre-trained parameters³ and is fine-tuned on our task using Adam.

Metrics As for metrics, we use UAS and LAS to measure the dependency tree labeling accuracy without and with relation labels respectively. LAS_O and LAS_U denote using the original relation set of each corpus

³<https://github.com/huggingface/transformers>

	Train	Dev	Test
HIT-CDTB _{dep}	250	50	53
SU-CDTB _{dep}	1600	400	332
Sci-CDTB	68	20	20
Unified _{dep}	1918	470	405

Table 3: Division of training, validation and test set for the three corpora

	HIT-CDTB _{dep}			SU-CDTB _{dep}			Sci-CDTB		
	UAS	LAS _O	LAS _U	UAS	LAS _O	LAS _U	UAS	LAS _O	LAS _U
Graph-based	0.353	0.237	0.255	0.585	0.415	0.428	0.338	0.175	0.199
Vanilla Trans	0.835	0.551	0.588	0.803	0.580	0.588	0.525	0.276	0.299
Two-stage	0.835	0.565	0.600	0.803	0.587	0.597	0.525	0.276	0.299
Bert-based	0.835	0.564	0.576	0.803	0.767	0.783	0.525	0.358	0.408
Two-stage _{Uni}	0.813	-	0.614	0.802	-	0.591	0.603	-	0.393
Bert-based _{Uni}	0.813	-	0.606	0.802	-	0.637	0.603	-	0.220
Human	0.872	0.723	-	0.897	0.774	-	0.806	0.627	-

Table 4: Performance Comparison of Benchmark Parsers on different dependency discourse corpora

or the predefined unified relation set.

Results Table 4 compares performance of the benchmark parsers. To give a rough upper bound of the parsing performance, the last row lists the consistency of human annotation by comparing two annotators’ labelling results on 30 documents from each corpus respectively. Division of training, validation and test set is shown in Table 3. The first four methods in the first block show the results of the benchmark parsers which are trained and tested respectively on HIT-CDTB_{dep}, SU-CDTB_{dep} and Sci-CDTB.

Graph-based method is less effective than the others, but it could probably be improved by using other training methods like MIRA. Among the transition-based methods, *Bert-based* performs the best on SU-CDTB_{dep} and Sci-CDTB_{dep} with respect to LAS, but is 2.4% lower than *Two-stage* on HIT-CDTB_{dep}, probably because HIT-CDTB_{dep} is a multi-domain corpus, and feature-based methods are more robust to changes in different domains. Comparing the three corpora, parsing results on Sci-CDTB are the worst because Sci-CDTB is too small for supervised learning. As we expect, SU-CDTB performs the best since it has relatively shallow tree structure and its labeling is highly consistent. With original relations mapped to the unified relation set, LAS_U results of all the methods on the three corpora have been improved compared with LAS_O, proving that the unified relation set we use is acceptable. Two-stage transition-based method performs better than the vanilla one with respect to LAS due to the addition of tree structural features in relation type prediction.

Two-stage_{Uni} and Bert-based_{Uni} use Unified_{dep} as training data, and are tested on each corpus respectively. From Table 4, we can see that Sci-CDTB has a significant promotion of 7.8% on UAS, while HIT-CDTB_{dep} and SU-CDTB_{dep} slightly decline. One possible explanation is that documents in HIT-CDTB_{dep} are much longer than in SU-CDTB_{dep} so their structures are not similar enough to improve each other’s performance. For Sci-CDTB, however, it also has a small numbers of EDUs per text so parsing performance can be improved by learning the augmented data of short trees from SU-CDTB.

For LAS_U, both methods increase on HIT-CDTB_{dep}, but drop on SU-CDTB_{dep}. As HIT-CDTB_{dep} covers multiple domains and has an average of 28.6 EDUs per text, its annotation is more difficult and the DDS-conversion process is also more complicated. In comparison, with an average of 4.5 EDUs per tree, SU-CDTB only focuses on news domain and its DDS-conversion is relatively easy. As a result, SU-CDTB_{dep} keeps a better data consistency than HIT-CDTB_{dep}, which may explain why the augmented data promotes the results on HIT-CDTB_{dep}, but introduces noise data for SU-CDTB_{dep}. As Sci-CDTB is from scientific domain and its relations are much different from those of HIT-CDTB and SU-CDTB, the

augmented labeled relations do not bring much improvement on relation labeling.

Overall, Unified_{dep} can serve as a discourse dataset for researching cross-domain or long text discourse parsing. However, our unification method may also introduce noise due to difference in text length and domain, so it remains to be considered how to better leverage this unified large corpus to improve the Chinese discourse parsing techniques.

5 Conclusions

In our work, we design semi-automatic methods to unify three Chinese discourse corpora with dependency framework. Our methods of converting PDTB-style and RST-style discourse annotation into DDS can be used to convert more other corpora and further enlarge our unified dataset. At the same time, by implementing several benchmark parsers on the converted data, we find that augmenting training set with the unified data can to some extent improve performance, but may also introduce noise and bring performance loss due to difference in text length and domain. This unified dataset is potentially helpful to research on cross-domain and long text discourse parsing, which will be our future work.

Acknowledgments

We thank the NLP Lab at Soochow University and HIT-SCIR for offering the discourse resources (*i.e.* SU-CDTB and HIT-CDTB) to conduct the research.

References

- Lynn Carlson, Daniel Marcu, and Mary Ellen Okurovsky. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of the SIGDIAL 2001 Workshop, The 2nd Annual Meeting of the Special Interest Group on Discourse and Dialogue, Saturday, September 1, 2001 to Sunday, September 2, 2001, Aalborg, Denmark*.
- Yi Cheng and Sujian Li. 2019. Zero-shot Chinese discourse dependency parsing via cross-lingual mapping. In *Proceedings of the 1st Workshop on Discourse Structure in Neural NLG*, pages 24–29, Tokyo, Japan, November. Association for Computational Linguistics.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805.
- Tsutomu Hiraio, Yasuhisa Yoshida, Masaaki Nishino, Norihito Yasuda, and Masaaki Nagata. 2013. Single-document summarization as a tree knapsack problem. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pages 1515–1520, Seattle, Washington, USA, October*. Association for Computational Linguistics.
- Sujian Li, Liang Wang, Ziqiang Cao, and Wenjie Li. 2014a. Text-level discourse dependency parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014, June 22–27, 2014, Baltimore, MD, USA, Volume 1: Long Papers*, pages 25–35.
- Yancui Li, Wenhe Feng, Jing Sun, Fang Kong, and Guodong Zhou. 2014b. Building chinese discourse corpus with connective-driven dependency tree structure. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25–29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*, pages 2105–2114.
- Haoran Li, Jiajun Zhang, and Chengqing Zong. 2017. Implicit discourse relation recognition for english and chinese with multiview modeling and effective representation learning. *ACM Trans. Asian Low Resour. Lang. Inf. Process.*, 16(3):19:1–19:21.
- Yang Liu, Sujian Li, Xiaodong Zhang, and Zhifang Sui. 2016. Implicit discourse relation classification via multi-task neural networks. In Dale Schuurmans and Michael P. Wellman, editors, *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12–17, 2016, Phoenix, Arizona, USA*, pages 2750–2756. AAAI Press.
- Mathieu Morey, Philippe Muller, and Nicholas Asher. 2018. A dependency perspective on rst discourse parsing and evaluation. *Computational Linguistics*, 44(2):197–235.

- Philippe Muller, Stergos Afantenos, Pascal Denis, and Nicholas Asher. 2012. Constrained decoding for text-level discourse parsing. In *Proceedings of COLING 2012*, pages 1883–1900, Mumbai, India, December. The COLING 2012 Organizing Committee.
- Joakim Nivre, Johan Hall, and Jens Nilsson. 2004. Memory-based dependency parsing. In *Proceedings of the Eighth Conference on Computational Natural Language Learning (CoNLL-2004) at HLT-NAACL 2004*, pages 444–449. Association for Computational Linguistics.
- Joakim Nivre. 2003. An efficient algorithm for projective dependency parsing. *Proceedings of the 8th International Workshop on Parsing Technologies (IWPT)*, 07.
- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind K. Joshi, and Bonnie L. Webber. 2008. The penn discourse treebank 2.0. In *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2008, 26 May - 1 June 2008, Marrakech, Morocco*.
- Yizhong Wang, Sujian Li, and Houfeng Wang. 2017. A two-stage parsing method for text-level discourse analysis. In *Meeting of the Association for Computational Linguistics*.
- Naiwen Xue, Fei Xia, Fu-Dong Chiou, and Martha Palmer. 2005. The penn chinese treebank: Phrase structure annotation of a large corpus. *Nat. Lang. Eng.*, 11(2):207–238.
- An Yang and Sujian Li. 2018a. Scidtb: Discourse dependency treebank for scientific abstracts. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 444–449. Association for Computational Linguistics.
- Jingfeng Yang and Sujian Li. 2018b. Chinese discourse segmentation using bilingual discourse commonality. In <https://arxiv.org/abs/1809.01497>. arXiv.
- Muyu Zhang, Bing Qin, and Ting Liu. 2014. Chinese discourse relation semantic taxonomy and annotation. *Journal of Chinese Information Processing*, 28:26–28.