SeCoRel: Multilingual Discourse Analysis in DISRPT 2025

Sobha Lalitha Devi and **Pattabhi RK Rao** and **Vijay Sundar Ram R**
AU-KBC Research Centre
MIT Campus of Anna University
Chennai, India
`sobha@au-kbc.org`

Abstract

The work presented here describes our participation in DISRPT 2025 shared task in three tasks, Task1: Discourse Unit Segmentation across Formalisms, Task 2: Discourse Connective Identification across Languages and Task 3: Discourse Relation Classification across Formalisms. We have fine-tuned XLM-RoBERTa, a language model to address these three tasks. We have come up with one single multilingual language model for each task. Our system handles data in both the formats .conllu and .tok and different discourse formalisms. We have obtained encouraging results. The performance on test data in the three tasks is similar to the results obtained for the development data.

1 Introduction

This paper describes our system, used in DISRPT2025 shared task “Discourse Relation Parsing and Treebanking (DISRPT)”. This is a Shared Task on Discourse Segmentation, Connective and Relation Identification across Formalisms. The shared task has the following three tasks: a) Task 1- Discourse Segment Identification, b) Task 2 – Discourse Connective Identification and c) Task 3 – Relation Identification. The organizers have provided data from different languages and annotations on these data follow different discourse formalisms. One of the main goals is that only one language model has to be developed which will apply to all languages and formalisms.

Discourse relations are the coherence relations between two discourse segments or also called as Elementary Discourse Units (EDUs) that can be realized explicitly or implicitly in a text. Discourse connectives play a role in signaling the relations in a discourse. They connect two discourse units, which may be a sentence, clause or multiple sentences. The relations can be intra sentential or inter sentential i.e. within a sentence or across sentences.

Thus the main objective of the work presented here is to develop a single language model for each of the task which will work for all languages and formalisms. The pre-trained XLM-RoBERTa language model was adapted through fine-tuning. In the following sections, we give a detailed description of our system.

2 Data

We train and evaluate our models using all the datasets provided by the shared task organizers.[1] In total, the benchmark is composed of 39 datasets, covering 16 languages and 6 frameworks. These datasets were obtained from the following corpora: the Czech RST Discourse Treebank 1.0 (Poláková et al., 2023), the Potsdam Commentary Corpus (Stede and Neumann, 2014; Bourgonje and Stede, 2020), the COVID-19 Discourse Dependency Treebank (Nishida and Matsumoto, 2022), the Discourse Dependency TreeBank for Scientific Abstracts (Yang and Li, 2018; Yi et al., 2021; Cheng and Li, 2019), the Genre Tests for Linguistic Evaluation corpus (Aoyama et al., 2023), the Georgetown University Multilayer corpus (Zeldes, 2017), the RST Discourse Treebank (Carlson et al., 2001), the Science, Technology, and Society corpus (Potter, 2008), the University of Potsdam Multilayer UNSC Corpus (Zaczynska and Stede, 2024), the Minecraft Structured Dialogue Corpus (Thompson et al., 2024), the Strategic Conversations corpus (Asher et al., 2016), the Basque RST Treebank (Iruskieta et al., 2013), the Persian RST Corpus (Shahmohammadi et al., 2021), the ANNOtation DIScursive corpus (Afantenos et al., 2012), the SUMM-RE corpus (Hunter et al., 2024; Prévot et al., 2025), the Dutch Discourse Treebank (Redeker et al., 2012), the Polish Discourse Corpus (Ogrodniczuk et al., 2024;

---

[1] GitHub: https://github.com/disrpt/sharedtask2025, and HuggingFace: https://huggingface.co/multilingual-discourse-hub.

Proceedings of the 4th Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2025), pages 79–86
November 9, 2025

Calzolari et al., 2024), the Cross-document Structure Theory News Corpus (Cardoso et al., 2011), the Russian RST Treebank (Toldova et al., 2017), the RST Spanish Treebank (da Cunha et al., 2011), the RST Spanish-Chinese Treebank (Cao et al., 2018), the Georgetown Chinese Discourse Treebank (Peng et al., 2022b,a), the DiscoNaija corpus (Scholman et al., 2025), the Penn Discourse Treebank (Prasad et al., 2014; Webber et al., 2019), the TED-Multilingual Discourse Bank (English) (Zeyrek et al., 2018, 2019), the LUNA Corpus Discourse Data Set (Tonelli et al., 2010; Riccardi et al., 2016), the Portuguese Discourse Bank (Mendes and Lejeune, 2022; Généreux et al., 2012), the Thai Discourse Treebank (Prasertsom et al., 2024), the Turkish Discourse Bank (Zeyrek and Webber, 2008; Zeyrek and Kurfalı, 2017), and the Chinese Discourse Treebank (Zhou et al., 2014).

The shared task was held in 2019 (Zeldes et al., 2019), 2021 (Zeldes et al., 2021), 2023 (Braud et al., 2023) and 2025 (), with more information on the data format in (Braud et al., 2024).

3 System Description

Motivated by the works presented in the previous DISRPT 2021 and 2023 workshops, a fine-tuning strategy was chosen. XLM-RoBERTa architecture is considered suitable for multilingual tasks, like question answering, discourse parsing because it employs self-attention mechanisms to effectively capture contextual dependencies within the text.

For discourse relation identification (task 3), the problem is framed as a classification task, in which it will learn to categorize discourse relations between different parts of the text.

Tasks 1 and 2 are handled as a sequence labeling task. This approach aims to identify and label the boundaries of discourse units within a given sequence of text.

XLM-RoBERTa is a transformer network-based model framework which relies on a strong self-attention mechanism to understand and interpret context effectively. This self-attention mechanism allows the model to weigh the significance of different parts of the input sequence, irrespective of their position, leading to a more nuanced understanding of the input data (Conneau et al., 2020).

XLM-RoBERTa-base (XLM-R-B) is a multilingual language model, well-suited for this shared task. XLM-R-B has a relatively smaller parameter size of 2.55B compared to XLMR Large, which translates to fewer computational resources required for processing. This efficiency makes it a practical choice for the present shared task.

3.1 Hyper-parameter Fine Tuning

In our approach, for fine-tuning XLM-RoBERTa we follow on the work of (Wolf et al., 2019), who offered a thorough framework for training for text classification models with Hugging Face's Transformers library. Although their configuration provided a strong basis for training the model, we modified it to better fit the discourse datasets provided in the shared task. Increasing the number of epochs from the initial setting to 10 was a crucial change that enabled the model to go through more thorough training and better absorb the subtleties of the data. In order to achieve effective gradient descent during training and maximize the trade-off between stability and quick convergence, we also changed the learning rate. Refining the batch sizes was another important modification. We set the evaluation batch size at 16 and the training batch size at 8. These modifications were designed to ensure adequate data flow for model learning while managing memory limitations on our hardware. In order to avoid over fitting, we also adjusted regularization parameters like the weight decay. The model's efficiency and generalization were enhanced by these adjusted parameters in conjunction with the monitoring of training and evaluation performance. All these optimizations were same for all three tasks. We were not able to get access to licensed datasets such as pdtb, thus these datasets were trained without words.

4 Results

Evaluation was done on the outputs produced by the system using the evaluation script provided by the organizers. The results are tabulated in the Tables 1, 2, 3, 4 and 5, for each of the tasks on different file formats and languages. Table 1 and 2 display the results obtained for task 1. Table 3 and 4 display the results obtained for task 2, And Table 5 displays the results obtained for task 3.

In task1 and task 2, the major challenge was tokenizing the input data into sentences. The data being multilingual we had to employ a multilingual sentence splitter. We had developed a basic sentence splitter using heuristic rules which handles different language texts. The results for (*.tok) files evidently shows the impact of sentence splitting.

File name	**Prec Dev**	**Prec Test**	**Rec Dev**	**Rec Test**	**F1 Dev**	**F1 Test**
ces. rst.crdt	93.40	89.63	91.40	91.30	92.40	90.46
eng. erst.gum	95.60	91.80	83.00	79.84	88.80	85.40
eng. rst.sts	60.30	71.18	89.30	73.51	72.00	72.32
eng. sdrt.stac	88.60	85.02	91.20	90.00	89.80	87.45
fra. sdrt.annodis	92.13	90.65	82.19	80.00	86.88	85.05
rus. rst.rrt	90.90	91.99	92.25	92.59	91.60	92.29
zho. dep.scidtb	93.72	86.69	94.35	97.02	94.03	91.56
deu. rst.pcc	95.63	97.13	93.26	91.86	94.43	94.42
eng. rst.oll	80.00	85.76	97.14	89.99	87.74	87.79
eng. rst.umuc	87.28	89.73	91.15	85.27	89.17	87.45
eus. rst.ert	92.05	90.71	90.69	89.72	91.36	90.21
nld. rst.nldt	96.00	96.75	97.95	97.04	96.96	96.89
spa. rst.rststb	93.67	90.02	95.46	92.17	94.56	91.08
zho. rst.gcdt	91.57	93.58	84.52	82.87	87.90	87.90
eng. dep.scidtb	96.52	96.49	92.48	92.39	94.46	94.39
eng. rst.rstdt	83.37	82.66	83.41	82.73	83.39	82.70
eng. sdrt.msdc	97.00	96.36	93.97	93.55	95.46	94.94
fas. rst.prstc	92.79	93.15	93.92	93.43	93.35	93.29
por. rst.cstn	91.66	90.96	92.53	95.42	92.10	93.14
spa. rst.sctb	78.04	85.20	93.20	85.71	84.95	85.45
zho. rst.sctb	49.20	55.71	93.20	89.88	64.42	68.72
Mean	**88.00**	**88.00**	**91.00**	**89.00**	**89.00**	**88.00**

Table 1: Evaluation Results for Task 1: Discourse Segmentation (for *.tok files)

For some language files such as ita.pdtb.luna, the sentence splitting was not efficient. In general it is observed that the results obtained for *.tok files are better. One probable reason is that for these files our sentence splitting algorithm worked better.

We observe that there are many false positives in zho.rst.sctb and spa.rst.sctb which has led to high recall and low precision. In the dataset eng.rst.umuc, the system has failed to learn segment start which is with-in the sentence. This has affected both recall and precision. Similar problem is observed in eng.rst.rstdt dataset also.

In task 2, the system has identified single word connectives with high precision and recall. It has poorly identified connectives with multiple words and apostrophe such as 'the same way', 'it would be same thing if', 'because of that', 'years have passed' etc. Improving the tokenization and contextual learning will boost the accuracy of connective identification.

In the relation identification task (task 3), we observed that the major errors are in the identification of 'elaboration' and 'conjunction' relation types. 'Elaboration' relation type is confused with relation types such as 'conjunction', 'organization', 'temporal' and 'frame'. Similarly 'conjunction is confused with 'temporal', 'explanation', 'frame' and 'causal'. We need to address these two relation types for improving the accuracy of the relation identification system. We need to train the system with syntactic features.

5 Conclusion

We have submitted our test runs for all the three tasks of the DISRPT 2025 shared task. We have fine-tuned the XLM-Roberta to handle multilingual and multi-formalism data. The three models and the system runs are available in the following link: `https://drive.google.com/drive/folders/1g3Rcve50OvlEWuqDzr8twiFipP8YLGhC?usp=sharing`

Acknowledgments

We thank the organizers of DISRPT 2025 for providing us the datasets and giving us opportunity to participate in this task.

References

Stergos Afantenos, Nicholas Asher, Farah Benamara, Myriam Bras, Cécile Fabre, Mai Ho-dac,

File name	Prec Dev	Prec Test	Rec Dev	Rec Test	F1 Dev	F1 Test
ces.rst.crdt	93.00	91.93	91.00	91.93	92.00	91.93
deu.rst.pcc	96.35	97.13	93.62	91.86	94.96	94.43
eng.dep.scidtb	96.52	96.50	92.44	92.39	94.44	94.40
eng.erst.gum	95.85	96.00	83.60	83.89	89.31	89.54
eng.rst.oll	81.52	85.62	96.07	88.89	88.20	87.22
eng.rst.rstdt	54.17	54.49	24.59	23.53	33.83	32.87
eng.rst.sts	63.40	86.09	93.47	88.39	75.56	87.22
eng.rst.umuc	87.29	90.08	91.15	85.09	89.18	87.51
eng.sdrt.msdc	97.47	96.68	93.12	92.98	95.24	94.80
eng.sdrt.stac	90.26	88.36	92.74	95.41	91.48	91.75
eus.rst.ert	92.46	90.60	90.55	89.86	91.49	90.23
fas.rst.prstc	92.80	93.16	93.92	93.58	93.36	93.37
fra.sdrt.annodis	91.02	87.35	82.01	80.42	86.28	83.74
fra.sdrt.summre	61.06	57.22	87.85	89.17	72.05	69.71
nld.rst.nldt	95.73	96.75	97.96	96.75	96.83	96.75
por.rst.cstn	91.67	90.97	92.54	95.42	92.10	93.14
rus.rst.rrt	91.26	92.28	92.72	92.65	91.98	92.47
spa.rst.rststb	93.43	90.59	94.99	92.17	94.20	91.38
spa.rst.sctb	77.42	86.14	93.20	85.12	84.58	85.63
zho.dep.scidtb	93.73	86.69	94.35	97.02	94.04	91.57
zho.rst.gcdt	93.04	93.88	85.31	84.34	89.01	88.86
zho.rst.sctb	49.75	57.14	96.12	95.24	65.56	71.43
Mean	**85.42**	**86.62**	**88.79**	**87.55**	**86.17**	**86.36**

Table 2: Evaluation Results for Task 1: Discourse Segmentation (for *.conllu files)

File name	Prec Dev	Prec Test	Rec Dev	Rec Test	F1 Dev	F1 Test
deu.pdtb.pcc	80.85	81.31	86.36	78.72	83.51	79.99
eng.pdtb.gum	87.64	87.89	92.69	86.00	90.10	86.93
eng.pdtb.tedm	79.00	83.41	71.81	76.19	75.23	79.63
ita.pdtb.luna	81.37	68.29	59.71	53.63	68.87	60.08
pcm.pdtb.disconaija	64.48	74.09	56.09	73.71	60.00	73.90
pol.iso.pdc	71.20	70.20	57.66	63.61	63.72	66.74
por.pdtb.crpc	82.95	81.62	76.81	71.87	79.76	76.44
por.pdtb.tedm	75.49	77.40	75.49	79.31	75.49	78.34
tha.pdtb.tdtb	-	-	-	-	-	-
tur.pdtb.tedm	78.33	78.43	34.81	32.38	48.20	45.84
Zho.pdtb.ted	78.54	-	82.68	-	80.56	-
Mean	**77.98**	**78.08**	**69.41**	**68.38**	**72.54**	**71.98**

Table 3: : Evaluation Results for Task 2: Discourse Connective Identification (for .tok files)

File name	Prec Dev	Prec Test	Rec Dev	Rec Test	F1 Dev	F1 Test
deu.pdtb.pcc	80.00	81.32	86.36	78.72	83.06	79.99
eng.pdtb.gum	86.55	90.65	83.43	82.23	84.96	86.23
eng.pdtb.tedm	79.00	83.89	71.82	76.62	75.24	80.09
ita.pdtb.luna	78.49	64.11	52.51	51.34	62.93	57.02
pcm.pdtb.disconaija	61.89	70.47	47.97	65.21	54.04	67.74
pol.iso.pdc	71.28	71.39	57.88	65.02	63.88	68.06
por.pdtb.crpc	83.24	81.42	76.81	71.69	79.89	76.25
por.pdtb.tedm	74.76	76.92	75.49	78.82	75.12	77.86
tha.pdtb.tdtb	75.27	77.15	84.23	86.16	79.49	81.40
tur.pdtb.tedm	76.27	78.43	33.33	32.39	46.39	45.84
zho.pdtb.ted	78.67	69.10	82.68	82.03	80.62	75.02
Mean	**76.86**	**76.80**	**68.41**	**70.02**	**71.42**	**72.32**

Table 4: : Evaluation Results for Task 2: Discourse Connective Identification (for .conllu files)

File name	Dev Data Accuracy	Test Data Accuracy
ces.rst.crdt	39.02	42.57
deu.pdtb.pcc	54.17	59.28
deu.rst.pcc	42.31	45.05
eng.dep.covdtb	66.28	68.10
eng.dep.scidtb	81.27	79.84
eng.erst.gentle	-	44.95
eng.erst.gum	43.61	46.89
eng.pdtb.gentle	-	46.31
eng.pdtb.gum	49.23	49.58
eng.pdtb.pdtb	28.29	26.92
eng.pdtb.tedm	49.44	53.28
eng.rst.oll	53.23	41.33
eng.rst.rstdt	10.12	10.90
eng.rst.sts	40.80	35.06
eng.rst.umuc	58.10	59.09
eng.sdrt.msdc	85.80	84.86
eng.sdrt.stac	65.88	67.64
eus.rst.ert	51.30	54.64
fas.rst.prstc	52.10	51.52
fra.sdrt.annodis	59.27	51.85
ita.pdtb.luna	60.68	65.07
nld.rst.nldt	51.06	55.69
pcm.pdtb.discon	54.47	56.54
pol.iso.pdc	47.89	50.07
por.pdtb.crpc	69.11	73.88
por.pdtb.tedm	58.42	64.29
por.rst.cstn	61.78	61.40
rus.rst.rrt	60.11	62.70
spa.rst.rststb	69.19	57.98
spa.rst.sctb	65.96	65.41
tha.pdtb.tdtb	95.66	96.21
tur.pdtb.tdb	25.40	24.94
tur.pdtb.tedm	50.71	49.04
zho.dep.scidtb	65.48	67.44
zho.pdtb.cdtb	60.81	58.92
zho.pdtb.ted	59.74	59.92
zho.rst.gcdt	60.44	55.93
zho.rst.sctb	52.13	55.97
Average	**52.61**	**55.29**

Table 5: Evaluation for Task 3 Discourse Relation Classification

Anne Le Draoulec, Philippe Muller, Marie-Paule Péry-Woodley, Laurent Prévot, Josette Rebeyrolles, Ludovic Tanguy, Marianne Vergez-Couret, and Laure Vieu. 2012. An empirical resource for discovering cognitive principles of discourse organisation: the ANNODIS corpus. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, pages 2727–2734, Istanbul, Turkey. European Language Resources Association (ELRA).

Tatsuya Aoyama, Shabnam Behzad, Luke Gessler, Lauren Levine, Jessica Lin, Yang Janet Liu, Siyao Peng, Yilun Zhu, and Amir Zeldes. 2023. GENTLE: A genre-diverse multilayer challenge set for English NLP and linguistic evaluation. In *Proceedings of the 17th Linguistic Annotation Workshop (LAW-XVII)*, pages 166–178, Toronto, Canada.

Nicholas Asher, Julie Hunter, Mathieu Morey, Benamara Farah, and Stergos Afantenos. 2016. Discourse structure and dialogue acts in multiparty dialogue: the STAC corpus. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pages 2721–2727, Portorož, Slovenia. European Language Resources Association (ELRA).

Peter Bourgonje and Manfred Stede. 2020. The Potsdam Commentary Corpus 2.2: Extending Annotations for Shallow Discourse Parsing. In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*, Paris, France. European Language Resources Association (ELRA).

Chloé Braud, Yang Janet Liu, Eleni Metheniti, Philippe Muller, Laura Rivière, Attapol Rutherford, and Amir Zeldes. 2023. The DISRPT 2023 shared task on elementary discourse unit segmentation, connective detection, and relation classification. In *Proceedings of the 3rd Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2023)*, pages 1–21, Toronto, Canada. The Association for Computational Linguistics.

Chloé Braud, Amir Zeldes, Laura Rivière, Yang Janet Liu, Philippe Muller, Damien Sileo, and Tatsuya Aoyama. 2024. DISRPT: A multilingual, multi-domain, cross-framework benchmark for discourse processing. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 4990–5005, Torino, Italia. ELRA and ICCL.

Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors. 2024. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*. ELRA and ICCL, Torino, Italy.

Shuyuan Cao, Iria da Cunha, and Mikel Iruskieta. 2018. The RST Spanish-Chinese treebank. In *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*, pages 156–166, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Paula Christina Figueira Cardoso, Erick Galani Maziero, Maria Lucía del Rosario Castro Jorge, M. Eloize, R. Kibar Aji Seno, Ariani Di Felippo, Lucia Helena Machado Rino, Maria das Graças Volpe Nunes, and Thiago Alexandre Salgueiro Pardo. 2011. CST-News - a discourse-annotated corpus for single and multi-document summarization of news texts in Brazilian Portuguese. In *Proceedings of the 3rd RST Brazilian Meeting*, pages 88–105, Cuiabá, Brazil.

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 Second SIGdial Workshop on Discourse and Dialogue*, 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. Association for Computational Linguistics.

Alexis Conneau, KartikayKhandelwal, NamanGoyal VishravChaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, MyleOtt, Luke Zettlemoyer, and VeselinStoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451.

Iria da Cunha, Juan-Manuel Torres-Moreno, and Gerardo Sierra. 2011. On the development of the RST Spanish treebank. In *Proceedings of the 5th Linguistic Annotation Workshop*, pages 1–10, Portland, Oregon, USA. Association for Computational Linguistics.

Michel Généreux, Iris Hendrickx, and Amália Mendes. 2012. Introducing the reference corpus of contemporary Portuguese online. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, pages 2237–2244, Istanbul, Turkey. European Language Resources Association (ELRA).

Julie Hunter, Hiroyoshi Yamasaki, Océane Granier, Jérôme Louradour, Roxane Bertrand, Kate Thompson, and Laurent Prévot. 2024. MEETING: A corpus of French meeting-style conversations. In *Actes de la 31ème Conférence sur le Traitement Automatique des Langues Naturelles, volume 1 : articles longs et prises de position*, pages 508–529, Toulouse, France. ATALA and AFPC.

Mikel Iruskieta, María Jesús Aranzabe, Arantza Diaz de Ilarraza, Itziar Gonzalez-Dios, Mikel Lersundi, and Oier Lopez de Lacalle. 2013. The RST Basque TreeBank: An online search interface to check rhetorical relations. In *4th Workshop on RST and Discourse Studies*, pages 40–49, Fortaleza, Brasil.

Amália Mendes and Pierre Lejeune. 2022. CRPC-DB a Discourse Bank for Portuguese. In *Proceedings of the 15th International Conference on Computational Processing of Portuguese (PROPOR 2022)*, pages 79–89, Berlin, Heidelberg. Springer-Verlag.

Noriki Nishida and Yuji Matsumoto. 2022. Out-of-domain discourse dependency parsing via bootstrapping: An empirical analysis on its effectiveness and limitation. *Transactions of the Association for Computational Linguistics*, 10:127–144.

Maciej Ogrodniczuk, Aleksandra Tomaszewska, Daniel Ziembicki, Sebastian Żurowski, Ryszard Tuora, and Aleksandra Zwierzchowska. 2024. Polish Discourse Corpus (PDC): Corpus design, ISO-compliant annotation, data highlights, and parser development. In (Calzolari et al., 2024), pages 12829–12835.

Siyao Peng, Yang Janet Liu, and Amir Zeldes. 2022a. Chinese Discourse Annotation Reference Manual. Research Report, Georgetown University (Washington, D.C.).

Siyao Peng, Yang Janet Liu, and Amir Zeldes. 2022b. GCDT: A Chinese RST Treebank for Multigenre and Multilingual Discourse Parsing. In *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing*, pages 382–391, Online only. Association for Computational Linguistics.

Lucie Poláková, Šárka Zikánová, Jiří Mírovský, and Eva Hajičová. 2023. Czech RST discourse treebank 1.0. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL).

Andrew Potter. 2008. Interactional Coherence in Asynchronous Learning Networks: A Rhetorical Approach. *Internet and Higher Education*, 11(2):87–97.

Rashmi Prasad, Bonnie Webber, and Aravind Joshi. 2014. Reflections on the Penn Discourse TreeBank, comparable corpora, and complementary annotation. *Computational Linguistics*, 40(4):921–950.

Ponrawee Prasertsom, Apiwat Jaroonpol, and Attapol T. Rutherford. 2024. The Thai Discourse Treebank: Annotating and Classifying Thai Discourse Connectives. *Transactions of the Association for Computational Linguistics*, 12:613–629.

Laurent Prévot, Roxane Bertrand, and Julie Hunter. 2025. Segmenting a large French meeting corpus into elementary discourse units. In *Proceedings of the 26th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL 2025)*, Avignon, France.

Gisela Redeker, Ildikó Berzlánovich, Nynke van der Vliet, Gosse Bouma, and Markus Egg. 2012. Multi-layer discourse annotation of a Dutch text corpus. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, pages 2820–2825, Istanbul, Turkey. European Language Resources Association (ELRA).

Giuseppe Riccardi, Evgeny A. Stepanov, and Shammur Absar Chowdhury. 2016. Discourse connective detection in spoken conversations. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6095–6099.

Merel CJ Scholman, Marian Marchal, AriaRay Brown, and Vera Demberg. 2025. DiscoNaija: a discourse-annotated parallel Nigerian Pidgin-English corpus. *Language Resources and Evaluation*, pages 1–37.

Sara Shahmohammadi, Hadi Veisi, and Ali Darzi. 2021. Persian Rhetorical Structure Theory. *arXiv preprint arXiv:2106.13833*.

Manfred Stede and Arne Neumann. 2014. Potsdam Commentary Corpus 2.0: Annotation for discourse research. In *Proceedings of the Language Resources and Evaluation Conference (LREC 2014)*, pages 925–929, Reykjavik, Iceland.

Kate Thompson, Julie Hunter, and Nicholas Asher. 2024. Discourse structure for the Minecraft corpus. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 4957–4967, Torino, Italia. ELRA and ICCL.

Svetlana Toldova, Dina Pisarevskaya, Margarita Ananyeva, Maria Kobozeva, Alexander Nasedkin, Sofia Nikiforova, Irina Pavlova, and Alexey Shelepov. 2017. Rhetorical relations markers in Russian RST treebank. In *Proceedings of the 6th Workshop on Recent Advances in RST and Related Formalisms*, pages 29–33, Santiago de Compostela, Spain. Association for Computational Linguistics.

Sara Tonelli, Giuseppe Riccardi, Rashmi Prasad, and Aravind Joshi. 2010. Annotation of discourse relations for conversational spoken dialogs. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)*, Valletta, Malta. European Language Resources Association (ELRA).

Bonnie Webber, Rashmi Prasad, Alan Lee, and Aravind Joshi. 2019. The Penn Discourse TreeBank 3.0 annotation manual. Technical report, University of Edinburgh, Interactions, LLC, University of Pennsylvania.

T Wolf, L Debut, V Sanh, J Chaumond, C Delangue, A Moi, P Cistac, T Rault, R Louf, and M Funtowicz. 2019. Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.

An Yang and Sujian Li. 2018. 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, Melbourne, Australia. Association for Computational Linguistics.

Cheng Yi, Li Sujian, and Li Yueyuan. 2021. Unifying discourse resources with dependency framework. In *Proceedings of the 20th Chinese National Conference on Computational Linguistics*, pages 1058–1065, Huhhot, China. Chinese Information Processing Society of China.

Karolina Zaczynska and Manfred Stede. 2024. Rhetorical strategies in the UN security council: Rhetorical Structure Theory and conflicts. In *Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 15–28, Kyoto, Japan. Association for Computational Linguistics.

Amir Zeldes. 2017. The GUM Corpus: Creating Multilayer Resources in the Classroom. *Language Resources and Evaluation*, 51(3):581–612.

Amir Zeldes, Debopam Das, Erick Galani Maziero, Juliano Antonio, and Mikel Iruskieta. 2019. The DISRPT 2019 shared task on elementary discourse unit segmentation and connective detection. In *Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019*, pages 97–104, Minneapolis, MN. Association for Computational Linguistics.

Amir Zeldes, 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, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Deniz Zeyrek and Murathan Kurfalı. 2017. TDB 1.1: Extensions on Turkish discourse bank. In *Proceedings of the 11th Linguistic Annotation Workshop*, pages 76–81, Valencia, Spain. Association for Computational Linguistics.

Deniz Zeyrek, Amália Mendes, Yulia Grishina, Murathan Kurfalı, Samuel Gibbon, and Maciej Ogrodniczuk. 2019. TED Multilingual Discourse Bank (TED-MDB): A parallel corpus annotated in the PDTB style. *Language Resources and Evaluation*, pages 1–27.

Deniz Zeyrek, Amália Mendes, and Murathan Kurfalı. 2018. Multilingual extension of PDTB-style annotation: The case of TED multilingual discourse bank. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Deniz Zeyrek and Bonnie Webber. 2008. A discourse resource for Turkish: Annotating discourse connectives in the METU corpus. In *Proceedings of the 6th Workshop on Asian Language Resources*.

Yuping Zhou, Jill Lu, Jennifer Zhang, and Nianwen Xue. 2014. Chinese Discourse Treebank 0.5 LDC2014T21.