Developing a Rhetorical Structure Theory Treebank for Czech

Lucie Poláková, Jiří Mírovský, Šárka Zikánová and Eva Hajičová
Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics
Prague, Czech Republic
{polakova, mirovsky, zikanova, hajicova}@ufal.mff.cuni.cz

Abstract

We introduce the first version of the Czech RST Discourse Treebank, a collection of Czech journalistic texts manually annotated using the Rhetorical Structure Theory (RST), a global coherence model proposed by Mann and Thompson (1988). Each document in the corpus is represented as a single tree-like structure, where discourse units are interconnected through hierarchical rhetorical relations and their relative importance for the main purpose of a text is modeled by the nuclearity principle. The treebank is freely available in the Lindat/Clariah-CZ repository under the Creative Commons license; for some documents, it includes two gold annotations representing divergent yet relevant interpretations. The paper outlines the annotation process, provides corpus statistics and evaluation, and discusses the issue of consistency associated with the global level of textual interpretation. In general, good agreement on the structure and labeling could be achieved on the lowest, local tree level and on the identification of the most central (nuclear) elementary discourse units. Disagreements mostly concerned segmentation and, in the structure, differences in the stepwise process of linking the largest text blocks. The project contributes to the advancement of RST research and its application to real-world text analysis challenges.

Keywords: Rhetorical Structure Theory, text coherence, Czech RST Discourse Treebank

1. Introduction

Understanding the entirety of texts, including their meaning, coherence, communicative functions and other aspects, has become increasingly important in both linguistics and NLP. In this respect, human-annotated language corpora provide a unique and invaluable foundation for both fields. In discourse-oriented research, while there are many hand-annotated resources available for local discourse analysis, mostly following the research line of the Penn Discourse Treebank (PDTB, Prasad et al., 2008), there is still much to be desired in terms of global analysis of discourse structure, allowing hierarchical discourse analysis/parsing and a deeper understanding of the underlying relational semantics and the overall meaning of a text. In addition, until recently, much of the research was based on the first available resources (see Section 2) and thus mostly focused on English.

In this paper, we introduce the Czech RST Discourse Treebank (CzRST-DT), a collection of Czech journalistic texts manually annotated using the Rhetorical Structure Theory, an influential global coherence model (Mann and Thompson, 1988; Taboada and Mann, 2006). Our corpus project aims to contribute to the growing body of resources with global coherence analysis, to increase the typological variety of languages analyzed in these resources, and thus to expand the research possibilities in discourse studies. By enlarging the accessible data and sharing our expertise in corpus design, our objective is also to provide valuable insights into the applicability of the RST framework.

The paper is organized as follows. In Section 2, we present relevant related work. In Section 3, we describe the data and tools used, the corpus development is described in Section 4. Section 5 presents basic corpus statistics including the IAA evaluation. Section 6 provides a discussion, Section 7 shows an example RST analysis and Section 8 concludes the paper.

2. Related work

2.1. Rhetorical Structure Theory

The Rhetorical Structure Theory (RST, Mann and Thompson, 1988) is a widely recognized framework for analyzing text structure and coherence. The goal of RST is to provide a formalized analysis of a text that captures a reconstruction of the author's intentions from the reader's point of view. The theory is based on the assumption that coherent texts consist of minimal units (*elementary discourse units*, EDUs) that are recursively linked through *rhetorical relations*. RST treats an entire document as a projective tree structure (compare the example analysis in Example 7). The type of the rhetorical relation linking discourse units is defined in terms of the author's intended effect on the reader together with the application of the principles of nuclearity. The principle of nuclearity suggests two types of rhetorical relations: in mononuclear relations, one text unit entering a rhetorical relation represents more essential information for the text's purpose – the Nucleus (N), while the

LREC-COLING 2024, pages 4802–4810
20-25 May, 2024.

other unit – the Satellite (S) – brings background, supplementary or supporting information, depending on the role of the satellite in different RST relations. In multinuclear relations, the importance of all units is equal: all units are Nuclei. Also, multinuclear relations may consist of more than two units.

2.2. RST-annotated Corpora

The RST framework has attracted significant interest and has undergone extensive development and testing. RST-style discourse corpora have been constructed for many languages. The first one, the RST Discourse Treebank (RST-DT, Carlson et al., 2003) annotated for rhetorical relations on 385 English articles from the Wall Street Journal, has become the main resource for experimenting in discourse parsing and similar tasks. Further RST annotation projects followed, both for English (Taboada and Renkema, 2008), (Zeldes, 2017) and for other languages, e.g.: Spanish (Da Cunha et al., 2011), Brazilian Portuguese (Cardoso et al., 2011), Dutch (Redeker et al., 2012), Basque (Iruskieta et al., 2013), German (Stede and Neumann, 2014), Russian (Toldova et al., 2017), Bangla (Das and Stede, 2018), or are in preparation: Persian (Shahmohammadi et al., 2021). Some projects are multilingual: Spanish – Chinese (Cao et al., 2018) or English – Spanish – Basque (Iruskieta et al., 2014).

The usefulness of Rhetorical Structure Theory for various tasks is wide. Apart from discourse parsing (Li et al., 2022), it has been used in many different NLP subfields, such as natural language generation (input: communicative goals and semantic representation, output: text), extractive and abstractive summarization (using the (strong) nuclearity principle[1]), sentiment analysis (aspect-based and hierarchical SA), argument mining, and in writing research (RST as a training tool for writing effective texts, coherence evaluation). A comprehensive overview of RST applications can be found in Hou et al. (2020).

3. Data and Tools

For the annotation of rhetorical structures, we have selected 54 Czech journalistic texts with a length between 6 and 42 sentences from different sub-genres, see Table 1.

The original texts are part of the richly annotated Prague Dependency Treebank 3.5 (Hajič et al., 2018), i.e. they have been independently annotated for local, PDTB-like discourse coherence, including implicit relations and secondary connectives. The treebank introduced here is annotated in a stand-off way on the plain text of the documents (see below), but will allow nevertheless for a future comparison of local and global coherence annotations on the same data.

Text genre	Count	Text genre	Count
comment	14	survey	3
news	10	weather	2
cultural review	7	overview	2
sports	4	letter	2
essay	5	description	1
advice	3	invitation	1

Table 1: Genre division of the annotated documents

The annotations were performed in a locally installed RSTWeb annotation tool (Zeldes, 2016),[2] a freely available and easily configurable client-server tool with a web browser-based interface. The tool stores the texts and their annotations in an internal database and exports them in an XML format with the extension .rs3.

The very same format can be used as an input to the RST-Tace tool (Wan et al., 2019),[3] a freely available tool that we used to measure the inter-annotator agreement (IAA, see Section 5).

4. Corpus Development Process

4.1. Annotation Procedure

Primary practical considerations regarding the development of the treebank concerned the trade-off between comparability with other RST projects and the novelty of the approach based on our previous research on higher-level relations and connectives in shallow discourse annotation (Poláková et al., 2021) and on other relevant work, e.g. Stede (2008), which would suggest some major adjustments to the theory. Finally, for the sake of comparability and usefulness for the discourse community, we adhere most closely to the recent version of RST and to the annotation guidelines applied in the German Potsdam Commentary Corpus (Stede et al., 2016, 2017). Differences in the relation taxonomy are described below in 4.2.

The annotation procedure consists of: (i) segmenting a text into elementary discourse units (EDUs), typically represented by clauses with a predicate verb, (ii) building the tree structure by progressively linking the EDUs and larger units through

[1] The *strong nuclearity principle* was formulated by Marcu (2000). It implies, in short, that the nucleus status of a segment is propagated throughout the tree. When all satellites in a tree are removed, we get a path to the most central segment(s) to the text's purpose (strongly nuclear segment(s)) .

[2] https://gucorpling.org/rstweb/info/

[3] https://github.com/tkutschbach/RST-Tace

rhetorical relations to form a hierarchically connected structure. At the point of creating a rhetorical link, its type (a label from the taxonomy) is also assigned.

Two annotators with linguistic background were trained on the same texts for pilot annotations, followed by two rounds of full-fledged annotations. After the IAA measurements, each round was concluded with qualitative consistency checks and discussions and subsequent updates of the guidelines. Five texts of different genres from the full annotations were double-annotated to measure inter-annotator agreement using the RST-Tace tool (see Section 5). The annotators did not know which documents were selected for the IAA measurement. Finally, the data were cleaned before publication. This included e.g. checking the connectivity of the whole structure, detecting and removing redundant "empty" levels in the tree caused by annotator errors, checking and correcting cases of interrupted segments (use of *Same-unit* and its structuring), consistent handling of headings and subheadings, etc. In this way, the published data is carefully checked, both automatically and manually, and the analyses can be considered as gold.

4.2. Taxonomy of RST Relations

The original account of RST proposes a set of 24 rhetorical relations.[4] The RST Discourse Treebank distinguishes 78 types of relations in 16 classes (Carlson and Marcu, 2001). The annotation in the Potsdam Commentary Corpus (PCC), the RST version closest to our approach, uses 31 rhetorical relations (Stede et al., 2017, 2016). The PCC taxonomy has been adapted for the annotation of newspaper editorials, i.e., opinion texts. Therefore, pragmatic relations play a key role in it. Both the classical RST of Mann and Thompson and of the Potsdam group emphasize that the taxonomy of rhetorical markers is open and can be adapted to the nature of the annotated texts.

The set of rhetorical relations used for our annotation contains 36+1 relations, see Table 2 which also reports the distributions of the relations in the corpus. Differences from the PCC guidelines include the addition of the following six (5+1) relations:

Gradation – we lacked a straightforward label for a common situation of escalating the content importance, often represented by connective patterns such as *not only ... but also*.

Disjunction – classifying disjunctive links as *Conjunction* (or *Otherwise*) seemed like a loss of information.

[4] compare also the relation set and definitions at https://www.sfu.ca/rst/01intro/definitions.html

Rhetorical relation	Frequency
Primarily pragmatic, mononuclear	
background	42
concession	64
concession-N	4
antithesis	28
evidence	37
reason	24
reason-N	10
justify	33
evaluation-S	22
evaluation-N	14
motivation	6
enablement	4
Primarily semantic, mononuclear	
circumstance	29
cause	41
result	18
condition	29
otherwise	5
unless	3
means	10
purpose	18
elaboration	135
entity-elaboration	84
interpretation	23
solutionhood	11
gradation	13
Textual, mononuclear	
preparation	55
summary	3
restatement	8
attribution	48
Multinuclear	
conjunction	96
disjunction	4
joint	49
list	102
sequence	31
contrast	47
restatement-M	6
Technical relation (for units split by an embedded content)	
same-unit	36

Table 2: Overview of rhetorical relations in five rhetorical types used for the annotation of the Czech RST Discourse Treebank, with overall frequencies of the relations in the corpus

Concession-N – during annotation, contexts with opposite nuclearity were found where the unit(s) expressing expectation was considered more crucial than the unit(s) expressing the violation of a causal principle, compare Example 1.

(1) *[Nucleus: Šéf bývalé tajné služby byl obžalován z vlastizrady. Hrozí mu až patnáct let vězení (+ 15 EDUs)]* ***Concession-N*** *[Satellite: Datum soudního procesu ale zatím nebylo oznámeno.]*
[*[Nucleus: The head of the former secret service has been charged with treason. He faces up to fifteen years in prison. (+ 15 further EDUs)]* ***Concession-N*** *[Satellite: However, the date of the trial has not yet been announced.]*]

In this example, the whole document, consisting of 18 elementary discourse units, elaborates on a topic of a secret agent accused of treason. The last, 18th unit says that the date of the trial has not yet been set, which is a kind of additional information (satellite), appended to the main text topic (nucleus). At the same time, the satellite shows a denied expectation, which is the opposite direction of nuclearity to that defined as *Concession*.
Restatement-M(ultinuclear) was reintroduced[5] for contents of equal importance.
Attribution was reintroduced as a result of an earlier segmentation decision: to prevent an inconsistent analysis, where an attribution clause merges into one unit with only a part of a multi-unit reported speech. For instance, previously *She said that she would come* only was one elementary unit, and *and that she would bring some food* was another. In our approach, these are three separate EDUs.
For practical reasons, we have reintroduced the technical relation of *Same-unit*, to account for discontinuous units. Such a situation is illustrated by Example 2 and Figure 1. In our proposal, by default, the embedded content is appended to the leftmost part of the discontinuous unit.

(2) *Castro prohlásil, že Spojené státy, ačkoli mohly udělit 150 000 víz, poskytly jich pouze 11 000.*
[*Castro said that the USA, even though they could have issued 150,000 visas, only issued 11,000 of them.*]

5. The Treebank in Numbers

Table 3 gives an overview of the basic treebank figures, such as the number of the relations annotated etc. Table 4 shows the inter-annotator agreement by the RST-Tace tool on five texts annotated by two annotators. The annotations of these files had to be slightly modified to meet the input requirements of the tool, i.e. they were changed to form a single tree in cases where a piece of text remained unconnected (a meta information, a title and a subtitle, etc.). The number of sentences after the modification is given in brackets. In addition, the tool only compares document pairs with unified EDU segmentation, thus disregarding the most burning issue of disagreement (see Section 6). Therefore, the IAA documents had to be segmented in a unified way first.[6]
The measurements show that despite the relatively low overall scores (the F-measure varied from 0.5 to 0.71 for nuclearity, from 0.21 to 0.5 for relation assignment, and from 0.41 to 0.66 for average), there are many near-matches (such as opposite nuclearity, close relations, different number of units in an otherwise corresponding multinuclear structure, etc.). This is also confirmed by the manual inspection of the tree structures (see Section 6).

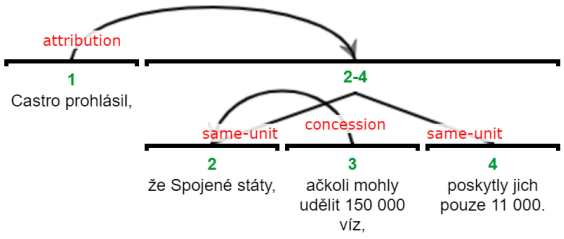

Figure 1: Annotation of an embedded segment in the Example 2 with the technical multinuclear relation *Same-unit*, and an example of *Attribution* relation annotation.

6. Discussion

This section explains some annotation choices and discusses the most common disagreements from the qualitative point of view.
Segmentation. Surprisingly, the most problematic part of the annotation was a uniform segmentation into EDUs. Compared to the annotation principles in PCC, we made the following segmentation decisions, which were subsequently reflected both in the shape of the structure and in the necessary introduction of new rhetorical relations. The main segmentation issues are restrictive and

[5] Reintroduced labels are those that were present in some of the earlier versions of RST (mostly included by Carlson et al., 2003) and that the Potsdam Commentary corpus team decided to drop for some reason. Our motivation to reintroduce them is data-driven.

[6] Measuring inter-annotator agreement on rhetorical structures is a very complex task, we thus provide a detailed explanation of the advanced tool outputs in a Supplementary material on a dedicated webpage `https://ufal.mff.cuni.cz/czrst-dt1.0/supplements`.

Specification	Count
Unique texts in the corpus	54
Number of RST analyses	59
Total no. of sentences	901
Total no. of tokens	14 514
Total elementary discourse units	1 422
Total rhetorical relations	1 192
Average no. of sentences per text	16.7
Average no. of tokens per text	268.8
Average no. of tokens per sentence	16.1

Table 3: Corpus Statistics

non-restrictive clauses and attribution.[7]

Restrictive and non-restrictive relative clauses, where in practice it is very difficult to draw a clear line. We decided to follow an instruction to segment as much as possible, to account for possibly relevant pieces of information (e.g. in contexts like *This is the question that bothers me the most.* we distinguish two EDUs.) This decision might have led to a higher number of *Entity-elaboration* labels in our corpus.

Attribution and reported content: To avoid inconsistent analysis, where an attribution clause merges into one unit with only a part of a multi-unit reported speech, we segment attribution clauses from the content they introduce based on a list of verbs of saying and an additional set of annotation instructions (see the annotation manual, Poláková, 2023) and we reintroduce the *Attribution* label. The annotation of attribution is illustrated in Figure 1 above.

Next type of disagreements concern the tree **structure**: A thorough comparison of disagreements in tree structures by two annotators suggests the following findings: (i) Good agreement on the structure and labeling can be achieved on the lowest, local tree level. This may be attributed to the guidance of syntax and local coherence clues, but partly also to the solid linguistic background of the annotators. (ii) There is a good agreement on the identification of the most central EDUs, i.e. the structures of both annotators point to the same units as being the most important ones (strong nuclearity principle), the overall structures show similar global patterns.

Disagreements, on the other hand, mostly concern differences in the stepwise process of linking the largest blocks of texts. There is an observed tendency to make the highest tree levels multinuclear, and thus the analysis becomes more technical. At least in our type of data (newspaper texts of various types), there sometimes seems to be no reliable (surface or semantic) cue as to how to connect the largest blocks (three or four of them). The more arbitrary the solution seems, the more annotators tend to create a multinuclear relation in which all large blocks are on the same level. In this way, the purpose of the analysis – to build a hierarchical structure – gives way to a technical solution (to build a connected graph). This conveys the message that rhetorical relations at this highest level may be vague and, in most cases, open to different yet relevant interpretations. A typical pattern of disagreement at the highest tree level in our data is the interpretation as [*Background/Preparation* (S) → Main message (N)], as opposed to [Main message (N) ← *Elaboration* (S)]. This is also visible in the analysis of Example 3 in Section 7 below, which corresponds to the former pattern, with an alternative analysis (the first segment – the title as the strongly nuclear segment) that would correspond to the latter pattern. In our opinion, this is a matter of a specific text type (here, journalism) and can only be attributed to personal preferences of the annotators. It is an example of the interpretative subjectivity of any text recipient, including skilled text researchers, but it also reveals where the RST framework is likely to reach its limits.

Recognizing subjectivity is an unavoidable aspect of text interpretation. The literature and our own previous research in this direction show (Daneš, 1988, Poláková and Synková, 2021) that in most cases this is related to the different accessibility of different types of inferences to the recipients. For a reliable global coherence modeling, we see the following mitigation strategies: agreement of the annotators after joint discussion, agreement after adjudication by a third judge, or the acceptance of the "good", relevant disagreement (Das et al., 2017). In line with the insights of Plank (2022), we incline to refer to this phenomenon as human label variation.

7. Example RST Analysis

Figure 2 shows a rhetorical structure analysis for the Czech text in Example 3 below. The example text contains 5 paragraphs, 10 sentences, 13 elementary discourse units (EDUs) and 12 rhetorical relations (among them 2 multinuclear). According to the analysis, the segment most central to the text's purpose (the strongly nuclear EDU)[8] is the segment 5: *což bude působit na vzestup cen* [*which will drive up the prices*]. This segment is also semantically very close to the title of the text, which could be an alternative central segment.

[7] We also encountered some language-specific, syntactic issues regarding e.g. participial constructions, different types of ellipses etc., but we do not elaborate on these within the scope of this paper.

[8] Stede (2008); Iruskieta et al. (2015).

Text genre	Sentences	EDUs	Nuclearity	Relation	Constituent	Attach. point	Average
Comment	12 (9)	15	0.71	0.50	0.64	0.79	0.66
			0.54	0.43	0.70	0.77	0.61
Advice	21 (19)	31	0.54	0.34	0.51	0.49	0.47
			0.36	0.30	0.55	0.47	0.42
Survey	18 (15)	16	0.65	0.29	0.65	0.71	0.57
			0.41	0.21	0.81	0.69	0.53
News	10 (8)	14	0.57	0.21	0.43	0.43	0.41
			0.39	0.13	0.47	0.39	0.34
Review	13 (12)	27	0.50	0.30	0.50	0.60	0.48
			0.31	0.26	0.62	0.58	0.44

Table 4: Inter-annotator agreement as measured by the RST-Tace tool on five double-annotated documents from five different genres; for each measurement, there are two lines of output given by the tool, representing (according to the description of the tool) the F-measure and the inter-annotator agreement, respectively. The agreement is composed of four properties: Nuclearity: the direction of the relation; Relation: the name of the relation; Constituent: the unit(s) where the satellite (or one of the nuclei in the case of multinuclear relations) is located; Attachment point: the unit(s) where the constituent is linked, compare (Wan et al., 2019). The numbers in brackets indicate the number of sentences after the annotations have been modified to form a single tree (as required by the tool).

(3) *Očekává se vzestup cen cukru.*
Podle odhadu britské obchodní firmy E. D. and F. Man bude objem letošní produkce cukru v zemích EU nižší než loni. Přesto by se však jeho ceny ještě v prvním čtvrtletí příštího roku neměly zvyšovat.
Teprve potom se projeví stoupající poptávka ze strany Ruska a Číny, což bude působit na vzestup cen. Jejich pohyb směrem vzhůru navíc podpoří nízká úroveň sklizně na Kubě i v samotných zemích EU, místo očekávaných 17,55 mil. tun pouze 14,8 milionu.
Odborníci odhadují, že Čína ho bude nucena příští rok dovézt až 2 mil. tun. Objem kubánské produkce má klesat i příští rok, protože vzhledem k nedostatku pohonných hmot tam bude plocha pro pěstování cukrové třtiny snížena na 63 %. Oficiální odhad tamní produkce pro sezónu 1993/94 uvádí číslo 4 mil. tun, zatímco neoficiální údaje hovoří pouze o 3,8 mil. tun.
V zemích EU, kde bylo loni dosaženo téměř rekordní produkce, bude letos sklizeň nižší. Hmotnost cukrové řepy je totiž menší asi o 20 až 30 %.

English translation:[9]
[*Sugar prices are expected to rise.*|$_1$

According to the estimate by the British trading firm E. D. and F. Man, EU sugar production this year will be lower than last year.|$_2$ *Nevertheless, prices should not increase in the first quarter of next year.*|$_3$

Only then a rising demand from Russia and China will be reflected,|$_4$ *which will drive up the prices.*|$_5$ *In addition, the low level of harvests in Cuba and in the EU countries themselves, instead of the expected 17.55 million tonnes, will support the upward movement, with only 14.8 million tonnes.*|$_6$

Experts estimate that China will be forced to import up to 2 million tonnes next year.|$_7$ *The volume of Cuban production is also expected to fall next year,*|$_8$ *as the area under sugar cane cultivation there will be reduced to 63% due to fuel shortages.*|$_9$ *The official estimate of local production for the 1993/94 season gives a figure of 4 million tonnes,*|$_{10}$ *while unofficial figures are only 3.8 million tonnes.*|$_{11}$

In EU countries, where near-record production was achieved last year, the harvest will be lower this year.|$_{12}$ *In fact, the weight of sugar beet is about 20 to 30% less.*|$_{13}$]

8. Conclusion

The Czech RST Discourse Treebank was published in June 2023 (Poláková et al., 2023) in the Lindat/Clariah-CZ repository under the Creative Commons licence.[10] It contains original Czech texts and RST annotations of 54 documents selected from the Prague Dependency Treebank. Five of the documents were used for the IAA measurements; for these, also the concurrent annotations done by the other annotator are a part of the package. The release includes corpus metadata, link to a dedicated website, and the annotation manual. Since the same documents were previously independently annotated for local discourse coherence, the presented project will allow future comparison of local and global coherence analyses of the same data.

[9] With marked segmentation to EDUs according to Figure 2.

[10] http://hdl.handle.net/11234/1-5174

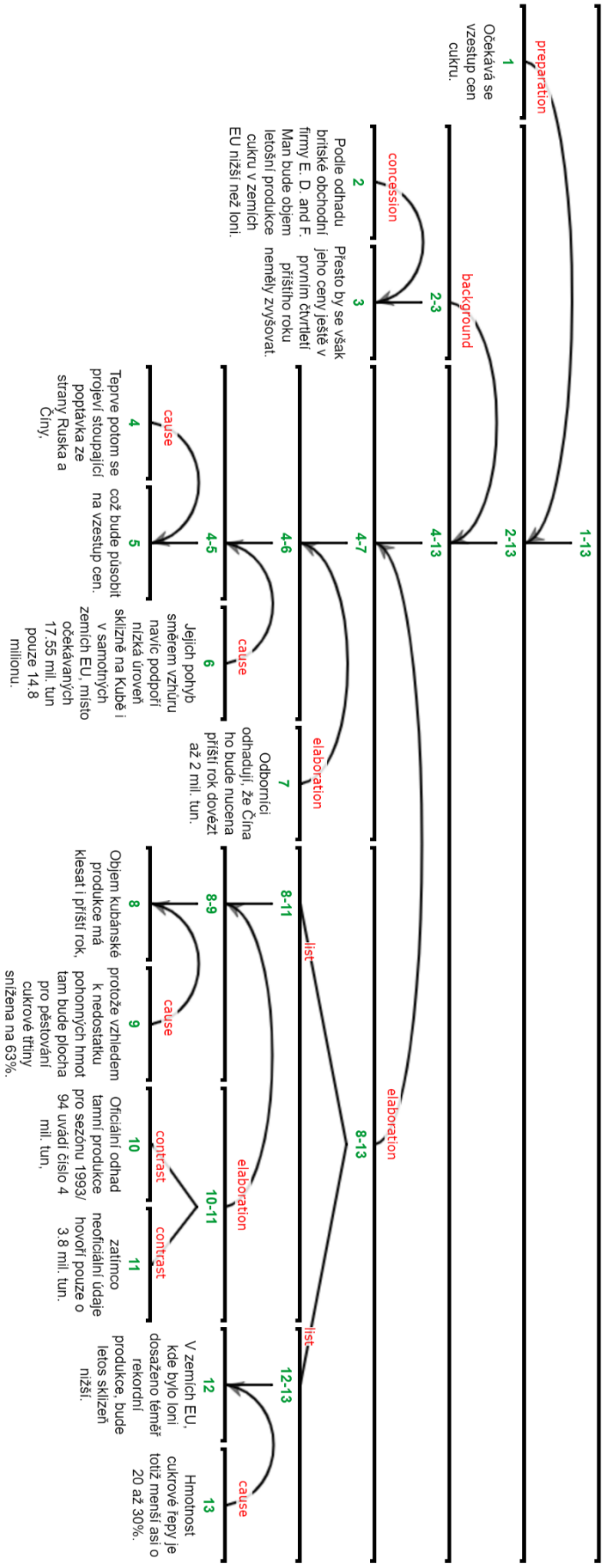

Figure 2: Example RST analysis for the Czech text in Example 3 as provided by the RSTweb tool

Ethics Statement

We honor the ethical code set out in the ACL Code of Ethics. There are no special ethical issues involved in this work.

Acknowledgements

The authors gratefully acknowledge support from the Grant Agency of the Czech Republic (project no. 20-09853S, *Global Coherence of Czech Texts in the Corpus-Based Perspective*) and from the Ministry of Education, Youth and Sports of the Czech Republic, Project No. LM2023062 LINDAT/CLARIAH-CZ.

9. Bibliographical References

Shuyuan Cao, Iria da Cunha, and Mikel Iruskieta. 2018. The RST Spanish-Chinese Treebank. In *Proceedings of the Joint Workshop of Linguistic Annotation, Multiword Expression and Constructions (LAW-MWE-CxG-2018)*, pages 156–166.

Paula Christina Figueira Cardoso, Erick Galani Maziero, Mara Elena Lucia, R. Castro Jorge, Ariani Di Felippo, Lucia Helena Machado Rino, Maria das Graças, Volpe Nunes, Thiago Alexandre Salgueiro Pardo, and Rodovia Washington Luís. 2011. CSTNews - A Discourse-Annotated Corpus for Single and Multi-Document Summarization of News Texts in Brazilian Portuguese. In *Proceedings of the 3rd RST Brazilian Meeting*.

Lynn Carlson and Daniel Marcu. 2001. Discourse Tagging Reference Manual. Technical Report 54, ISI Technical Report ISI-TR-545.

Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2003. Building a Discourse-tagged Corpus in the Framework of Rhetorical Structure Theory. In *Current and new directions in discourse and dialogue*, pages 85–112. Springer.

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. Association for Computational Linguistics.

František Daneš. 1988. Předpoklady a meze interpretace textu [Preconditions and Limits of Text Interpretation]. In *Slavica Pragensia XXXII*, pages 85–109, Praha. AUC – Philologica.

Debopam Das and Manfred Stede. 2018. Developing the Bangla RST Discourse Treebank. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Debopam Das, Manfred Stede, and Maite Taboada. 2017. The Good, the Bad, and the Disagreement: Complex Ground Truth in Rhetorical Structure Analysis. In *Proceedings of the 6th Workshop on Recent Advances in RST and Related Formalisms*, pages 11–19, Santiago de Compostela, Spain. Association for Computational Linguistics.

Shengluan Hou, Shuhan Zhang, and Chaoqun Fei. 2020. Rhetorical Structure Theory: A Comprehensive Review of Theory, Parsing Methods and Applications. *Expert Syst. Appl.*, 157:113421.

Mikel Iruskieta, Marıa J Aranzabe, Arantza Diaz de Ilarraza, Mikel Lersundi, and Oier Lopez de Lacalle. 2013. The RST Basque TreeBank: An Online Search Interface to Check Rhetorical Relations. In *4th Workshop RST and Discourse Studies*, pages 40–49.

Mikel Iruskieta, Iria da Cunha, and Maite Taboada. 2014. A Qualitative Comparison Method for Rhetorical Structures: Identifying Different Discourse Structures in Multilingual Corpora. *Languages Resources and Evaluation*.

Mikel Iruskieta, Arantza Diaz de Ilarraza, Gorka Labaka, and Mikel Lersundi. 2015. The detection of central units in Basque scientific abstracts. In *5th Workshop "RST and Discourse Studies" in Actas del XXXI Congreso de la Sociedad Española del Procesamiento del Lenguaje Natural. SEPLN*.

Jiaqi Li, Ming Liu, Bing Qin, and et al. 2022. A Survey of Discourse Parsing. *Frontiers of Computer Science*, 16(165329).

William C. Mann and Sandra A. Thompson. 1988. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text-Interdisciplinary Journal for the Study of Discourse*, 8:243–281.

Daniel Marcu. 2000. *The Theory and Practice of Discourse Parsing and Summarization*. MIT Press, Cambridge, MA.

Barbara Plank. 2022. Is Human Label Variation Really so Bad for AI? Invited Talk. Clarin Annual Conference, Prague, Czech Republic.

Lucie Poláková. 2023. Instrukce pro anotaci Rhetorical Structure Theory (RST) v češtině [Instructions for the annotation of Rhetorical Struc-

ture Theory (RST) in Czech]. Annotation Manual. Institute of Formal and Applied Linguistics, Charles University, Prague, Czech Republic.

Lucie Poláková, Jiří Mírovský, Šárka Zikánová, and Eva Hajičová. 2021. Discourse relations and connectives in higher text structure. *Dialogue & Discourse*, 12(2):1–37.

Lucie Poláková and Pavlína Synková. 2021. Pragmatické aspekty v popisu textové koherence [Pragmatic aspects in the description of discourse coherence]. *Naše řeč*, 104(4):225–242.

Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. 2008. The Penn Discourse TreeBank 2.0. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, pages 2961–2968, Marrakech. European Language Resources Association.

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'12)*, pages 2820–2825, Istanbul, Turkey. European Language Resources Association (ELRA).

Sara Shahmohammadi, Hadi Veisi, and Ali Darzi. 2021. Persian Rhetorical Structure Theory. arXiv. https://doi.org/10.48550/arXiv.2106.13833.

Manfred Stede. 2008. Disambiguating rhetorical structure. *Research on Language and Computation*, 6:311–332.

Manfred Stede and Arne Neumann. 2014. Potsdam Commentary Corpus 2.0: Annotation for Discourse Research. In *Proceedings of LREC 2014*, pages 925–929, Reykjavik, Iceland.

Manfred Stede, Maite Taboada, and Debopam Das. 2017. Annotation Guidelines for Rhetorical Structure. Technical report, University of Potsdam and Simon Fraser University.

Manfred Stede et al. 2016. *Handbuch Textannotation: Potsdamer Kommentarkorpus 2.0*. Number 8 in Potsdam Cognitive Science Series. Universitätsverlag Potsdam.

Maite Taboada and William C. Mann. 2006. Rhetorical Structure Theory: Looking Back and Moving Ahead. *Discourse studies*, 8(3):423–459.

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.

Shujun Wan, Tino Kutschbach, Anke Lüdeling, and Manfred Stede. 2019. RST-Tace A Tool for Automatic Comparison and Evaluation of RST Trees. In *Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019*, pages 88–96.

Amir Zeldes. 2016. rstWeb-A Browser-based Annotation Interface for Rhetorical Structure Theory and Discourse Relations. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, pages 1–5.

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

10. Language Resource References

Jan Hajič and others. 2018. *Prague Dependency Treebank 3.5*. Data/Software, Charles University, MFF, ÚFAL, Prague, Czech Republic. [link].

Lucie Poláková, Šárka Zikánová, Jiří Mírovský, and Eva Hajičová. 2023. Czech RST Discourse Treebank 1.0. Prague, Czech Republic. ÚFAL MFF UK, LINDAT.

Maite Taboada and Jan Renkema. 2008. Discourse Relations Reference Corpus. Simon Fraser University and Tilburg University.