

The RST Continuity Corpus

Debopam Das

Åbo Akademi University
Tehtaankatu 2
20500 Turku, Finland
debopam.das@abo.fi

Markus Egg

Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin, Germany
markus.egg@hu-berlin.de

Abstract

We present the RST Continuity Corpus (RST-CC), a corpus of discourse relations annotated for continuity dimensions. Continuity or discontinuity (maintaining or shifting deictic centres across discourse segments) is an important property of discourse relations, but the two are correlated in greatly varying ways. To analyse this correlation, the relations in the RST-CC are annotated using operationalised versions of Givón’s (1993) continuity dimensions. We also report on the inter-annotator agreement, and discuss recurrent annotation issues. First results show substantial variation of continuity dimensions within and across relation types.

1 Introduction

In this paper, we present the RST Continuity Corpus (RST-CC), the first corpus of discourse relations annotated for a wide range of continuity dimensions (e.g., *time*, *space*, *reference*, or *perspective*). These dimensions describe different ways in which a deictic centre can be maintained or updated during a discourse. The corpus contains 1,009 relations from five major relation types, which are a subset of the RST Discourse Treebank (Carlson et al., 2002). In the RST-CC, relations are annotated with respect to Givón’s (1993, ch. 13) seven continuity dimensions. The relations are furthermore annotated for additional features such as *polarity* (positive or negative relation) and *context* (intra- vs. inter-sentential relation).

The paper is structured as follows: Section 2 outlines previous work on continuity and discourse relations. In Section 3, we describe the composition of the RST-CC, its general format, and the selected relations. Section 4 elaborates on the selected continuity dimensions and their operationalisation into distinctive features. Additional features of annotation are described in Section 5. Section 6 reports on the inter-annotator agreement study, and Section 7 discusses recurrent annotation issues. We

present first results in Section 8, and conclude with an outlook on the next steps of our work.

2 Theoretical background

2.1 The notion of continuity

Continuity emerges in multi-segment discourse when the deictic centre remains constant along a situational dimension across segments; e.g., the events or situations described in two segments occur at the same time or share their protagonists. The deictic centre is the point of reference with respect to which context-dependent expressions are evaluated, it is often but not always determined by the speaker.

In contrast, changes along these dimensions, e.g., when a new segment refers to a situation set in an earlier time (like in a flashback) or introduces a new protagonist, result in discontinuity. Continuity is monitored during text processing in that readers maintain or update their frame of reference for dimensions like time, space, character, or causation (Zwaan et al., 1995; Zwaan and Singer, 2003).

We define continuity in terms of thematic coherence (Givón, 1993), which distinguishes seven continuity dimensions or ‘coherence strands’. Maintaining or shifting deictic centres on these dimensions between discourse segments determines the extent of thematic coherence (continuity) or disruption (discontinuity). The seven dimensions are *time*, *space*, *reference*, *action*, *perspective*, *modality*, and *speech act*. The first four dimensions are more concrete and local, the others, more abstract and global, as visualised in Table 1.

The grouping of dimensions is based on effect; consider (1)-(2) from Givón (1993). In (1), a change in the temporal continuity across the two clauses causes a local break, but does not necessarily terminate a larger coherent sequence of clauses in the text. In contrast, a change in one of the global dimensions amounts to a stronger break, which can

local	time
	space
	reference
	action
global	perspective
	modality
	speech act

Table 1: Givón’s coherence strands

terminate such a sequence of clauses. There is such a break in (2), because it exhibits discontinuity in perspective between the two sentences (viewpoint of the author vs. the one of the protagonist).

- (1) She flew in at midnight and left the next day.
- (2) She came in and sat on the bed. She was tired, she thought.

We base our annotation on these continuity dimensions, as they offer a comprehensive range of continuity dimensions. Also, the framework locates continuity at the level of clauses or sequences of clauses and the way they are linked, which is exactly where discourse relations are situated.

2.2 Discourse relations and continuity

Continuity is a crucial feature of discourse (or coherence) relations, which introduce a semantic or pragmatic link (e.g., additive, causal, or adversative) between two discourse segments. The relations exhibit continuity or discontinuity across the discourse segments they link. For example, the clauses in (3)¹ are linked by a CONSEQUENCE relation, the situation presented in the first clause being the consequence of the event in the second clause.

- (3) [The Indian stock markets have been on a five-year high, with dips and corrections,] [since Prime Minister Rajiv Gandhi started liberalizing industry.]

The situation in the first clause of (3) temporally follows the event in the second clause. This backward temporal shift introduces discontinuity along the temporal dimension. Also, the segments have no common discourse referents, which amounts to referential discontinuity. In contrast, there is no spatial shift across the segments, neither is there a perspective change because both segments can be attributed to the same source (the author). So, the relation is continuous for space and perspective.

¹All examples are from the RST Discourse Treebank (Carlson et al., 2002) unless specified otherwise.

2.3 Previous work

Previous work models the relation between continuity and discourse relations in different ways. Sometimes, continuity is treated as a binary feature, and discourse relations, or even whole groups of such relations, are summarily classified as continuous or discontinuous. For instance, Murray (1997) considers CAUSAL relations continuous, and Zufferey and Gygax (2016) regard CONTRASTIVE relations as discontinuous. Asr and Demberg (2012) classify discourse relations in the Penn Discourse Treebank (Prasad et al., 2008) for continuity and discontinuity. They group relations like RESULT, INSTANTIATION, and LIST as continuous and relations like PRAGMATIC CONTRAST, CONTRA-EXPECTATION, or TEMPORAL relations as discontinuous, but leave the CONDITIONAL relations underspecified with respect to continuity.

Other work classifies discourse relations as configurations of individual continuity dimensions (e.g., time, space, or reference). Fetzer (2018) describes relations with a set of continuity dimensions (‘particularized features’), which include temporal and referential continuity, but also continuity of action. Relations are distinguished in terms of the presence or absence of continuity along specific dimensions. For example, CONTINUATION relations are characterised as continuous for dimensions of time, reference, topic, aspect, and lexical coherence, while CONTRAST relations display discontinuity on at least one of these dimensions.

In sum, there is as yet no unanimously accepted classification of discourse relations for continuity. What is more, even individual relations can be continuous and discontinuous on different dimensions simultaneously. For example, CAUSAL relations, which are generally deemed to be continuous, can simultaneously exhibit continuity for the temporal dimension, but discontinuity for the reference dimension, as in (4).

- (4) [As some securities mature and the proceeds are reinvested,] [the problems ought to ease.]

At the same time, CONTRAST relations, usually regarded as discontinuous, can show the same configuration (continuity for time, not for reference):

- (5) [The gasoline picture may improve this quarter,] [but chemicals are likely to remain weak.]

Such cases raise the fundamental question of whether relation types in their entirety can be classified with respect to continuity.

For some dimensions it is even debated whether they introduce continuity or discontinuity in the first place. For example, temporal progression in narration is often cited as indicative of continuity since it represents the expected flow of events (Zwaan, 1996; Zufferey and Gygax, 2016). However, such transitions, particularly when signalled by a temporal connective like *then*, have also been taken to indicate discontinuity (Segal et al., 1991). Asr and Demberg (2012) even regard synchronous temporal relations as discontinuous because they are often used to introduce new events.

The lack of unanimity across approaches and corpus examples like (4) and (5) suggest re-examining the relationship of continuity and discourse relations in detail, i.e., on the level of individual tokens of the relations. For each continuity dimension of a token, continuity must be determined separately.

Since there is as yet no resource for this research question, we compiled the RST-CC, whose format will be described in the next section.

3 The RST Continuity Corpus

The RST-CC comprises relations from the RST Discourse Treebank or RST-DT (Carlson et al., 2002). The RST-DT contains 385 newspaper texts annotated for over 20,000 relations according to Rhetorical Structure Theory or RST (Mann and Thompson, 1988). In RST, relations link a more and a less central discourse unit (nucleus and satellite), or two equally central units (nuclei). Linking is recursive, which models discourse as a tree structure. Elementary discourse units (EDUs) in RST are typically clauses; there may be sub-clausal EDUs units, however (especially in the RST-DT). Fig. 1 illustrates an RST analysis for (6), in which the segments A and B are collectively connected to C by a REASON relation. Text in square brackets represents discourse units; in Fig. 1, arrows go from satellites to nuclei.

- (6) [[The U.S. Coast Guard closed six miles of the Houston Ship Channel,]^A [where about 150 companies have operations,]^B] [because the thick, black smoke obscured the area.]^C

The RST-CC includes five relation types: CAUSAL, CONTRASTIVE, CONDITIONAL, ELABORATION, and TEMPORAL. This selection is moti-

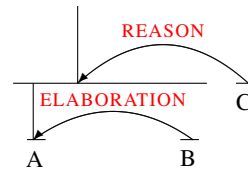


Figure 1: Graphical representation of an RST analysis

vated by previous classifications, which categorise, for example, CAUSAL and ELABORATION relations as continuous (Murray, 1997), CONTRASTIVE relations as discontinuous (Zufferey and Gygax, 2016), TEMPORAL relations as one or the other (Hopper, 1979), and CONDITIONAL relations as underspecified with respect to continuity (Asr and Demberg, 2012).

The relations are also classified in terms of the Cognitive approach to Coherence Relations or CCR (Sanders et al., 1992, 2021), using features such as polarity (positive or negative)² and basic operation (implicational or additive, i.e., causal or non-causal). For instance, ELABORATION relations are positive and additive, whereas CONTRASTIVE relations are negative. Table 2 summarises these classifications.

In the RST-CC, the relation types are subdivided according to the RST-DT relation taxonomy (Carlson and Marcu, 2001); e.g., the CONTRASTIVE type includes the subtypes ANTITHESIS, CONCESSION, and CONTRAST. Table 6 in the Appendix offers a detailed account of the relation types, their member subtypes, and their key features.

Relation type	Predicted continuity	Polarity	Basic operation
CAUSAL	continuous	positive	implicational
CONTRASTIVE	discontinuous	negative	additive implicational
CONDITIONAL	not specified	positive negative	implicational
ELABORATION	continuous	positive	additive
TEMPORAL	continuous discontinuous	positive	additive

Table 2: Relation types and their features

In our continuity corpus, we strove to strike a balance between the distribution of the different relation types and that of their subtypes, which turned out to be challenging at times. First, some subtypes have only very few relation tokens, such as ELABORATION-PROCESS-STEP (3 tokens) and INVERTED-SEQUENCE (12 tokens). Second, for

²Negative relations introduce a negation operator in their definition, e.g., OTHERWISE (see the Appendix, Table 6).

certain relation types the distribution of the subtypes in the RST-DT corpus was extremely uneven. For example, in the CONDITIONAL relation type, the subtype CONDITION ranges over 200 tokens, whereas the other subtypes such as CONTINGENCY and OTHERWISE have fewer than 30 tokens.

For an optimal representation of the relation variety, we collected all instances of the infrequent subtypes, further balancing out their low counts by including a higher number of tokens of the more frequent subtypes of the same type³. In the end, we collected 1,009 relations with 276 CAUSAL, 156 CONTRASTIVE, 172 CONDITIONAL, 179 ELABORATION, and 226 TEMPORAL relations. (For the distribution of the subtypes in our corpus, see Table 7 in the Appendix).

Each relation was independently annotated by the two authors for the seven continuity dimensions. Any differences were subsequently adjudicated before including the relation in the corpus.

4 Operationalising continuity dimensions

To annotate the relations in the RST-CC according to Givón’s (1993) seven continuity dimensions, we operationalised them into distinctive features⁴.

4.1 Time

We model temporal continuity using Evers-Vermeul et al.’s (2017) classification of temporality. For a sequence of discourse segments, they distinguish non-temporal, synchronous, and sequential constellations, and divide the latter into chronological and anti-chronological. (7) and (8) exhibit synchronous and anti-chronological constellations, respectively.

(7) [The Ministry of Education is nothing but a cartel for licensed teachers,] [and certainly does not act on behalf of students.]
[relation: CAUSE-RESULT; time: synchronous]

(8) [Monsanto Co., too, is expected to continue reporting higher profit,] [even though its sales of crop chemicals were hurt in the latest quarter by drought in northern Europe and the western U.S.]

³We found some potentially misclassified relation tokens in the RST-DT, especially within the CONDITIONAL relation type. Our corpus does not contain such tokens, however.

⁴The features are summarised in Table 8 in the Appendix.

[relation: CONCESSION; time: anti-chronological]

Evers-Vermeul et al.			Our features
Non-temporal [-TIME]			continuity
Temporal [+TIME]	Synchronous [+SIMULTANEOUS]		
	Sequential [-SIMULTANEOUS]	Chronological [+PRIOR]	
		Anti-chronological [-PRIOR]	discontinuity

Table 3: Operationalisation of the temporal dimension

We assume that anti-chronological pairs of discourse segments introduce temporal discontinuity. All other constellations are classified as continuous; see Table 3. According to this classification, (7) emerges as continuous, and (8) as discontinuous.

4.2 Space

We consider a relation spatially continuous if the events or situations in the discourse segments are non-spatial, as in (9), or situated in the same place. In spatially discontinuous relations, location shifts in between segments, as in (10).

(9) [Passenger car prices jumped 3.8% in September,] [after climbing 0.5% in August and declining in the late spring and summer.]
[relation: TEMPORAL-AFTER; space: no change]

(10) [investment will be more likely to flow toward the other European economies] [and “the U.K. will be less prepared for the single market.”]
[relation: CONSEQUENCE; space: change]

4.3 Reference

We express referential continuity in terms of Centering Theory or CT (Grosz et al., 1995). CT determines for each segment a central discourse referent (‘backward-looking centre’), which can be continued or updated between segments, and occupy different positions on a salience hierarchy for all referents of a segment. This gives rise to four types of transition between segments: *continue*, *retain*, *smooth shift*, and *rough shift*. Poesio et al. (2004) add the types *establishment*, *zero*, and *null*, for the initialisation, termination, or lack of anaphoric reference across segments.

We classify a discourse relation as referentially continuous if the transition between its segments involves some kind of shared referent, like *the Soviets* in (11). Thus, *continue*, *retain*, *smooth shift*, *rough shift*, and *establish* transitions are considered as continuous. In contrast, *zero* and *null* transitions emerge as discontinuous, as in (12), where reference to the Aetna company is discontinued in the second segment.

- (11) It's not enough! [If the Soviets want to be believed,] [they need to start telling the truth about more than the totally obvious.]
[relation: CONDITION; reference: *establish*]
- (12) In a few instances, Aetna knew [it would probably be shelling out big bucks] [even before a client called or faxed in a claim]
[relation: TEMPORAL-BEFORE; reference: *zero*]

4.4 Action

We operationalise action continuity in terms of *script theory* (Schank and Abelson, 1975; Modi et al., 2016), which postulates that part of our knowledge is organised in ‘scripts’ or stereotypical descriptions of routine activities like having a meal in a restaurant or visiting a doctor. This operationalisation makes it possible to support inter-annotator agreement by falling back on existing script data collections like the one of Regneri et al. (2010) or InScript (Modi et al., 2016).

We examine whether the actions or events in the discourse segments can be considered part of a script, so that there is a logical *flow* from one action or event to another. If yes, the relation is considered continuous, as in (13); otherwise, we classify it as discontinuous, as in (14).

- (13) [A substantial warming would melt some of the Earth's polar ice caps,] [raising the level of the oceans]
[relation: SEQUENCE; action: flow]
- (14) [Mercedes officials said they expect flat sales next year] [even though they see the U.S. luxury-car market expanding slightly.]
[relation: CONCESSION; action: no flow]

4.5 Perspective

We distinguish three types of perspective (Pander Maat, 1998): *objective*, *author* (in the form of comments), and *other* (quotations). We consider a discourse relation continuous on the perspective dimension if its two segments share the same perspective, as in (15), otherwise, we classify the relation as discontinuous, as in (16).

- (15) [“Climate varies drastically due to natural causes,” said Mr. Thompson.] [But he said ice samples from Peru, Greenland and Antarctica all show substantial signs of warming.]
[relation: CONTRAST; perspective: no change]
- (16) [“The earnings were fine and above expectations,” said Michael W. Blumstein, an analyst at First Boston Corp.] [Nevertheless, Salomon's stock fell \$1.125 yesterday to close at \$23.25 a share in New York Stock Exchange composite trading.]
[relation: CONTRAST; perspective: change]

4.6 Modality

Modality is predominantly introduced by modal verbs, but also by modal adverbials and verbs like *probably* and *doubt*, respectively. Modal expressions describe what the world would be like according to a ‘modal source’, e.g., wishes, obligations (including laws), or expectations (for a formalisation, see Kratzer 2001).

Discontinuity in modality amounts to a shift of the reality or possible world dimension of the deictic centre. For instance, in (18) below, the modal dimension shifts from the real world (in which Temple-Inland is not expanding) to the world according to Mr. Palmero, in which the company is capable of future debt reduction. If both arguments of a discourse relation are non-modal or if they are modal with respect to the same modal source, we classify the relation as continuous, as in (17); otherwise, as discontinuous, like in (18).

- (17) [Cineplex traded on the New York Stock Exchange at \$11.25 a share, up \$1.125,] [before trading was halted.]
[relation: TEMPORAL-BEFORE; modality: no change]

- (18) Mr. Palmero recommends Temple-Inland, explaining [that it is “virtually the sole major paper company not undergoing a major capacity expansion”] [and thus should be able to lower long-term debt substantially next year.]
[relation: CAUSE-RESULT; modality: change]

4.7 Speech act

Discourse segments can be declarative clauses, questions, or imperatives. When sentence mood changes in between segments, Givón (1993) assumes discontinuity along the speech act dimension.⁵ Thus, relations count as discontinuous if only one of the segments is declarative, as in (19).

- (19) [The next time you hear a Member of Congress moan about the deficit,] [consider what Congress did Friday.]
[relation: CONTINGENCY; speech act: change]

The only exception are rhetorical questions, which we classified as statements (declaratives) in our analysis in spite of their syntactic guise, because they are interpreted as statements. For instance, the second discourse unit of (20) introduces the claim that no one will pay high prices for racehorse anymore:

- (20) [If bluebloods won’t pay high prices for racehorses anymore,] [who will?]
[relation: CONDITION; speech act: no change]

5 Additional features

Features that potentially influence the relationship between discourse relations and continuity are also annotated in the RST-CC. We include the CCR features *polarity* (see Section 3) and *order of segments*. The latter applies to implicational (CAUSAL and CONDITIONAL) relations only: The order is basic if the cause or antecedent segment precedes the result or consequent segment (Sanders et al., 1992); the reverse order indicates a non-basic relation.

⁵This overlaps with but is not identical to speech act relations (Sweetser, 1990), a subset of pragmatic relations, which link one argument to the speech act expressed in the other one.

The relations are annotated for two more features: *nuclearity* (which specifies the segment pair as nucleus-satellite, nucleus-nucleus, or satellite-nucleus, according to RST) and *context* (whether the relation occurs intra- or inter-sententially). The annotation scheme for the additional features is summarised in Table 9 in the Appendix.

For illustration, we provide an example of the RST-CC annotation (for seven continuity dimensions and also for four additional features) in Table 10 in the Appendix.

6 Reliability of annotation

To assess the quality of our annotation, we conducted an annotation experiment. For the seven continuity dimensions, we independently annotated a selection of 240 relations, which are not part of the RST-CC, but represent the five relation types of the corpus. Agreement was substantial according to Cohen’s kappa (Landis and Koch, 1977) for the four dimensions *time*, *reference*, *perspective*, and *modality*, as shown in Table 4. For the remaining dimensions, prevalence prevented the calculation of meaningful κ -values. The agreement scores are 97.07% for *space*, 95.82% for *action*, and 98.74% for *speech act*.⁶

time	reference	perspective	modality
0.72	0.69	0.70	0.76

Table 4: Inter-annotator agreement on four dimensions

To annotate the *action* dimension, we had to consult external encyclopaedic sources, since there were no script data available for the events described in the corpus data. However, our results show that for specialised domains like the economic topics featured in many articles of the RST-CC, external sources can greatly contribute to safeguarding inter-rater agreement.

The scores reported for *time* and *reference* measure agreement on the binary distinction between continuous and discontinuous relations, as described in Section 4. However, we also calculated scores for more fine-grained classifications.

For *reference*, our annotation of the entire seven-fold classification of Centering Theory also yielded

⁶Prevalence refers to the ratio between the cardinalities of the classes that emerged in the classification. High prevalence leads to high chance agreement. And, since the idea of the kappa statistic is to abstract away from chance agreement, it returns very low kappa values for highly unbalanced samples, even if inter-rater agreement is very high.

substantial agreement ($\kappa = 0.62$). The confusion matrices reveal that agreement is especially high for the preservation, the termination, and the lack of reference continuity (*continue*, *zero*, and *null*, respectively). We interpret this result as confirming the usefulness of Centering Theory for practical annotation initiatives.

For *time*, we annotated a more fine-grained classification into non-temporal/synchronous, chronological, and anti-chronological constellations. Agreement on this classification was only moderate ($\kappa = 0.49$). Subsequent evaluation showed that the problematic distinction was the one between synchronous and chronological, in particular, for implicational relations. The choice of the values for temporal continuity varied over whether the consequent (or result) starts simultaneously with the antecedent (cause) or whether the latter follows the former. The issue is illustrated in (21), for which one annotator assumed that the junk market getting its biggest jolt (cause) is synchronous with it going into a tailspin (consequence), whereas the other one understood a chronological order in that the tailspin began after the jolt.

- (21) [The fragile market received its biggest jolt last month from Campeau Corp...] [At that point, the junk market went into a tailspin...] [Relation: CONSEQUENCE; time: ?]

Subsequent discussion of these decisions revealed that the forced choice between the two possible temporal constellations introduced considerable arbitrariness, which was reflected in low agreement. Consequently, one should avoid forcing a choice in these cases by subsuming the two constellations in the underspecified statement that the consequence does not precede the antecedent. We conclude that such examples pose a severe challenge for approaches to temporal continuity that, unlike ours, regard chronological (as opposed to synchronous) order as non-continuous.

For *perspective*, the high agreement was supported by the fact that newspaper text indicates the sources of direct or indirect quotes very clearly. The disagreements mainly involved distinguishing reported facts from any kind of comment or conclusion drawn from them. For other text types, we envisage that the identification of perspectives must take into account additional linguistic evidence, e.g., in the case of free indirect discourse (Eckardt, 2014).

7 Recurrent annotation issues

This section presents recurring issues for our annotation which make choosing the correct label for a specific continuity dimension challenging.

7.1 *Perspective annotation for implicit attribution*

In newspaper texts, quotes and reported speech are not always indicated (or attributed to their sources) explicitly. This typically happens when a whole series of statements of one single speaker is reported: Some of the statements are presented as a direct quote (*X said, "..."*) or as reported speech (*X said that...*), while the others are not marked explicitly. This is illustrated by (22) [= (18)], where the first segment is a direct quote with attribution to the speaker, while the second one is unmarked, although they both belong to the same statement (made by Mr. Palmero). Accordingly, there is no change of perspective for the relation.

- (22) Mr. Palmero recommends Temple-Inland, explaining [that it is “virtually the sole major paper company not undergoing a major capacity expansion,”] [and thus should be able to lower long-term debt substantially next year.]
[relation: CAUSE-RESULT, perspective: no change]

However, in certain instances it is unclear whether a segment is attributed to a source or not, e.g., in (23), the last segment might be due to Guy Witman or to the author of the article. In this case, even the context of the whole article does not provide a definitive clue to answer this question:

- (23) [[Still, today’s highest-yielding money funds may beat CDs over the next year even if rates fall,] says Guy Witman, an editor of the Bond Market Advisor newsletter in Atlanta.] [That’s because top-yielding funds currently offer yields almost $1\frac{1}{2}$ percentage points above the average CD yield.]
[relation: EXPLANATION-ARGUMENTATIVE, perspective: ?]

7.2 *Annotating modality*

In annotating *modality*, we encountered the problem of indirect speech transforming future-tense

auxiliaries into conditional forms, without introducing modality. For instance, *would* in (24) merely expresses the future tense as it is part of the indirect speech introduced by the matrix clause *he said*:

- (24) He said [construction wouldn't resume] [until market conditions warrant it.]
(relation: CONDITION, modality: no change)

This disambiguation is especially difficult when the scope of the indirect speech is not clear or if the context does not suffice to distinguish between equally plausible readings, as in (25):

- (25) [Sears expected] [that the pricing program wouldn't have any effect on revenue].
(relation: CONTRIBUTION, modality: ?)

Another issue is the scope of modality. The scope of a modal expression might extend over both segments, which entails continuity along the modal dimension, e.g., in (26):

- (26) ... a quarterly dividend of 76 cents, [which would be received] [before the February option expires]
(relation: TEMPORAL-BEFORE, modality: no change)

(26) involves no change of modality, because *would* scopes over both segments. This is reflected in its interpretation as the possibility of receiving a dividend before the expiration of an option.

8 First results

We provide the distribution of continuous relations (proportions in percentages) for five relation types with respect to the seven continuity dimensions in Table 5, with the highest and lowest scores for a dimension in bold font.

We found that some continuity dimensions show uniformity across relation types. Relations of all types are found to be overwhelmingly continuous (> 98%) for the dimensions *space* and *speech act*, and almost never continuous (< 2%) for *action*⁷. We believe that this is due to our data: In particular, the non-narrative character of our data is responsible for the low degree of *action* continuity and

⁷However, even these dimensions exhibit 100% continuity or discontinuity for a specific relation type only rarely: ELABORATION is 100% continuous for *space*, CONTRASTIVE, 0.00% for *action*, and 100% for *speech act*.

for the high degree of *space* continuity. In addition, there are very few questions and imperatives in our newspaper data, which explains the overall continuity for *speech act*. Due to these limitations of our corpus, we believe that the uniformity we found for the *space*, *speech act*, and *action* dimensions does not suggest that these dimensions are less important for continuity in discourse relations; instead, they might become distinctive if material from other registers is investigated.

For the dimensions *time*, *reference*, *perspective*, and *modality*, however, there is considerable difference between the relation types, as summarised in Table 5. In addition, we found that the relation types are not homogeneously continuous or discontinuous, but can be simultaneously more continuous for some dimensions but less continuous or even predominantly discontinuous for other dimensions. In particular, CONTRASTIVE relations are the least continuous for *reference* and *perspective*, but highly continuous for *time*. CONDITIONAL relations are the most continuous for *perspective*, and the least continuous for *modality*. TEMPORAL relations are the least continuous for *time*, but the most continuous for *reference* and *modality*. What is more, continuity is not uniform even for a single dimension of one of these relations; e.g., only 82.61% (and not 100%) of the CAUSAL relations are continuous for *time*.

Continuity scores for *reference* are consistently lower for two reasons: There are many small discourse segments in the RST-DT, which reduces the chance of finding a shared referent across the segments. This is illustrated by (27), where the target relation, CONSEQUENCE-N, holds between segments A (the single word 'lost') and B.

- (27) [Mr. Lagnado said] [that] [although retailers probably won't ever recover sales] [lost]^A [because of the California quake and Hurricane Hugo,]^B [they could see some benefits later on.]

Since neither of the two segments has a background-looking centre (Cb), referential continuity of the relation is calculated as *null*, which amounts to discontinuity. Moreover, we did not consider eventive and propositional referents in the analysis. However, as long as we compare only reference scores across the relation types (or subtypes), this will not affect our results.

Relation	Time	Reference	Perspective	Modality	Space	Action	Speech act
CAUSAL	82.61	30.79	85.87	80.79	97.46	2.54	99.64
CONDITIONAL	81.98	35.47	93.61	61.63	98.84	5.81	98.26
CONTRASTIVE	91.67	23.72	67.31	77.56	98.08	0.00	100
ELABORATION	93.85	34.64	78.21	85.47	100	0.56	99.44
TEMPORAL	74.34	38.50	90.27	92.92	97.35	0.88	98.67
mean	84.04	32.90	83.94	80.57	98.23	1.98	99.21

Table 5: Continuity scores across relation types

The correlations between relation types and continuity along a specific dimension are significant at $p < .001$ ($p < .05$ for *reference*) for all dimensions except *space* and *speech act*.

9 Conclusions and outlook

We presented the RST Continuity Corpus (RST-CC), which comprises five major types of discourse relations annotated for a wide array of continuity dimensions and additional features. We envisage two applications of the corpus. First, the RST-CC will contribute to a more precise characterisation of discourse relations, providing a systematic, detailed, and reliable resource for examining the relationship between continuity (dimensions) and discourse relations. In addition, the corpus can also be used to test hypotheses about correlations between continuity dimensions and discourse relations. For example, CONTRASTIVE relations often present information about different (though comparable) items or information from different sources, and one can test whether this would lead to low scores for *reference* and *perspective* continuity.

Second, the corpus, in conjunction with parallel resources like the RST Signalling Corpus (Das et al., 2015), will contribute to the study of discourse signalling, e.g., to explore the *continuity hypothesis* (Murray, 1997), which entails that discontinuous discourse relations are harder to process, and hence, their processing should be facilitated by more explicit signalling.

Furthermore, it is an important research question whether continuity in discourse relations patterns uniformly or differently across genres or languages. For further work in this field, the development of the RST-CC could be a model for similar resources for different genres and different languages (other than news texts in English, as in the RST-CC).

Finally, we believe that our decompositional approach towards continuity would support further in-depth analyses of discourse relations. The varying effect of different continuity dimensions on discourse relations, for instance, would help resolv-

ing incongruities found in the study of discourse processing (why certain discourse relations are processed quicker and remembered better than others).

For a broader empirical basis for such investigations, we will extend the RST-CC, adding more instances of the relation types covered so far, but also including additional relation types like BACKGROUND, COMPARISON, EVALUATION, and EXPLANATION. The final version of the RST-CC will be published via the Linguistic Data Consortium.

References

- Fatemeh Asr and Vera Demberg. 2012. Measuring the strength of linguistic cues for discourse relations. In *Proc. Workshop on Advances in Discourse Analysis and its Computational Aspects*, pages 33–42.
- Lynn Carlson and Daniel Marcu. 2001. [Discourse Tagging Manual](#). ISI Technical Report ISI-TR-545, USC.
- Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2002. [RST Discourse Treebank, LDC2002T07](#).
- Debopam Das, Maite Taboada, and Paul McFetridge. 2015. [RST Signalling Corpus, LDC2015T10](#).
- Regine Eckardt. 2014. *The semantics of free indirect discourse: How texts allow us to mind-read and eavesdrop*. Brill, Leiden.
- Jacqueline Evers-Vermeul, Jet Hoek, and Merel Scholman. 2017. On Temporality in Discourse Annotation. *Dialogue & Discourse*, 8(2):1–20.
- Anita Fetzer. 2018. The encoding and signalling of discourse relations in argumentative discourse. In M. de los Ángeles Gómez González and J. L. Mackenzie, editors, *The Construction of Discourse as Verbal Interaction*, page 13–44. John Benjamins.
- Talmy Givón. 1993. *English Grammar: A function-based introduction*, volume 2. John Benjamins.
- Barbara Grosz, Aravind Joshi, and Scott Weinstein. 1995. Centering: A Framework for Modelling the Local Coherence of Discourse. *Computational Linguistics*, 21(2):203–225.

- Paul Hopper. 1979. Aspect and foregrounding in discourse. In Talmy Givón, editor, *Syntax and semantics, Vol. 12. Discourse and syntax*, page 213–241. Academic Press, New York.
- Angelika Kratzer. 2001. Modality. In Arnim v. Stechow and Dieter Wunderlich, editors, *Semantics*, page 639–650. Mouton, Berlin.
- Richard Landis and Gary Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33:159–174.
- William Mann and Sandra Thompson. 1988. Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. *Text*, 8:243–281.
- Ashutosh Modi, Tatjana Anikina, Simon Ostermann, and Manfred Pinkal. 2016. InScript: Narrative texts annotated with script information. In *Proceedings of LREC 2016*.
- John Murray. 1997. Connectives and narrative text: The role of continuity. *Memory and Cognition*, 25(2):227–236.
- Henk Pander Maat. 1998. Classifying negative coherence relations on the basis of linguistic evidence. *Journal of Pragmatics*, 30:177–204.
- Massimo Poesio, Rosemary Stevenson, Barbara Di Eugenio, and Janet Hitzeman. 2004. Centering: A Parametric Theory and Its Instantiations. *Computational Linguistics*, 30(3):309–363.
- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltasakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. 2008. The Penn Discourse Treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)*, pages 2961–2968, Marrakech, Morocco.
- Michaela Regneri, Alexander Koller, and Manfred Pinkal. 2010. Learning script knowledge with web experiments. In *Proceedings of the 48th Annual Meeting of the ACL*, page 979–988.
- Ted Sanders, Wilbert Spooren, and Leo Noordman. 1992. Toward a taxonomy of coherence relations. *Discourse Processes*, 15:1–35.
- Ted J.M. Sanders, Vera Demberg, Jet Hoek, C.J. Scholman, Merel, Fatemeh Torabi Asr, Sandrine Zufferey, and Jacqueline Evers-Vermeul. 2021. Unifying dimensions in coherence relations: How various annotation frameworks are related. *Corpus Linguistics and Linguistic Theory*, 17(1):1–71.
- Roger Schank and Robert Abelson. 1975. Scripts, Plans, and Knowledge. In *Proceedings of the 4th IJCAI*, page 151–157. Morgan Kaufmann Publishers Inc.
- Erwin Segal, Judith Duchan, and Paula Scott. 1991. The role of interclausal connectives in narrative structuring: Evidence from adults’ interpretations of simple stories. *Discourse Processes*, 14(1):27–54.
- Eve Sweetser. 1990. *From Etymology to Pragmatics: The Mind-Body Metaphor in Semantic Structure and Semantic Change*. Cambridge University Press, Cambridge.
- Sandrine Zufferey and Pascal Gygax. 2016. The Role of Perspective Shifts for Processing and Translating Discourse Relations. *Discourse Processes*, 53(7):532–555.
- Rolf Zwaan. 1996. Processing Narrative Time Shifts. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22(5):1196–1207.
- Rolf Zwaan, Mark Langston, and Arthur Graesser. 1995. The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6(5):292–297.
- Rolf Zwaan and Murray Singer. 2003. Text Comprehension. In A. Graesser, M. Gernsbacher, and S. Goldman, editors, *Handbook of Discourse Processes*, pages 83–121. Erlbaum.

A Appendix

Dimension	Value	(Dis)continuous?
time	non-temporal	continuous
	synchronous	
	chronological	
	anti-chronological	discontinuous
space	no change	continuous
	change	discontinuous
reference	continue	continuous
	retain	
	smooth shift	
	rough shift	
	establish	
	zero	discontinuous
null		
action	flow	continuous
	no flow	discontinuous
perspective	no change	continuous
	change	discontinuous
modality	no change	continuous
	change	discontinuous
speech act	no change	continuous
	change	discontinuous

Table 8: Continuity dimensions and their values

add. feature	value
polarity	positive
	negative
order of segments	basic
	non-basic
nuclearity	S-N
	N-S
	N-N
context	intra-sentential
	inter-sentential

Table 9: Additional features and their values

Relation type	Relation subtype	Definition: key feature(s)
CAUSAL	CAUSE	Nucleus (N) is the cause; Satellite (S) is the result.
	RESULT	N is the result; S is the cause.
	CAUSE-RESULT (multinuclear)	Cause and result are equally important.
	CONSEQUENCE-S	Weaker version of CAUSE; N is the cause; S is the consequence.
	CONSEQUENCE-N	Weaker version of RESULT; N is the consequence; S is the cause.
	CONSEQUENCE (multinuclear)	Weaker version of CAUSE-RESULT; cause and consequence are equally important.
CONTRASTIVE	ANTITHESIS	N and S stand in contrast with each other.
	CONTRAST (multinuclear)	Two equally important units stand in contrast with each other.
	CONCESSION	The contrast arises due to a violated expectation between N and S.
CONDITIONAL	CONDITION	The consequent holds if the antecedent holds.
	CONTINGENCY	In any context, the consequent holds if the antecedent holds.
	HYPOTHETICAL	Like CONDITION, in addition, the antecedent is assumed to be true.
	OTHERWISE (mostly multinuclear)	The consequent does not hold if the antecedent does.
ELABORATION	ELABORATION-ADDITIONAL	S provides additional information about N.
	ELABORATION-GENERAL-SPECIFIC	S provides specific information about N.
	ELABORATION-OBJECT-ATTRIBUTE	S is an embedded clause/NP modifying an object/entity representing N.
	ELABORATION-PART-WHOLE	S specifies or elaborates on a part of N.
	ELABORATION-PROCESS-STEP	S enumerates the steps for carrying out a process introduced by N.
	ELABORATION-SET-MEMBER	N introduces a set/list of information; S elaborates on one (or more) member of the set/list
	EXAMPLE	S provides an example for the information in N.
	DEFINITION	S provides a definition of N.
TEMPORAL	TEMPORAL-BEFORE	The situation in N occurs before or leading up to the situation in S.
	TEMPORAL-AFTER	The situation in N occurs after the situation in S.
	TEMPORAL-SAME-TIME	The situations in N and S occur at approximately the same time.
	SEQUENCE	A multinuclear list of events presented in chronological order.
	INVERTED-SEQUENCE	A multinuclear list of events presented in reverse chronological order.

Table 6: Relation types, relation subtypes, and their key features

Relation type	Relation subtype	#	#
CAUSAL	CAUSE	43	276
	RESULT	52	
	CAUSE-RESULT (multinuclear)	52	
	CONSEQUENCE-S	52	
	CONSEQUENCE-N	52	
	CONSEQUENCE (multinuclear)	25	
CONTRASTIVE	ANTITHESIS	52	156
	CONCESSION	52	
	CONTRAST (multinuclear)	52	
CONDITIONAL	CONDITION	108	172
	CONTINGENCY	27	
	HYPOTHETICAL	22	
	OTHERWISE (predominantly multinuclear)	15	
ELABORATION	ELABORATION-ADDITIONAL	44	179
	ELABORATION-GENERAL-SPECIFIC	22	
	ELABORATION-OBJECT-ATTRIBUTE	22	
	ELABORATION-PART-WHOLE	22	
	ELABORATION-PROCESS-STEP	3	
	ELABORATION-SET-MEMBER	22	
	EXAMPLE	22	
DEFINITION	22		
TEMPORAL	TEMPORAL-BEFORE	35	226
	TEMPORAL-AFTER	57	
	TEMPORAL-SAME-TIME	56	
	SEQUENCE	66	
	INVERTED-SEQUENCE	12	
total			1009

Table 7: Distribution of relations types and subtypes

Relation to be annotated:			
To be sure, [big investors might put away their checkbooks in a hurry] [if stocks open sharply lower today]			
[relation: CONDITION]			
Dimension	Value	Explanation	Continuity
time	change	The consequent or protasis (first segment) precedes the antecedent or apodosis (second segment).	discontinuous
space	no change	The segments have no spatial markers; hence, the relation is non-spatial.	continuous
reference	null	None of the segments has a backward-looking centre (Cb).	discontinuous
action	no flow	The transition of the segments does not represent part of a script (a stereotypical situation or routine activity).	discontinuous
perspective	no change	Both segments bear the perspective of the writer.	continuous
modality	change	The first segment uses the modal verb 'might' while the second one uses none.	discontinuous
speech act	no change	Both segments are declarative sentences.	continuous
add. feature	value		
polarity	positive		
order of segments	non-basic (consequent-antecedent)		
nuclearity	N-S		
context	intra-sentential		

Table 10: Example of RST-CC annotation