

Information- and discourse-structure analysis with questions under discussion

Session 3: Discourse segments, adjuncts and rhetorical relations

Kordula De Kuthy, Eberhard-Karls-Universität Tübingen
Arndt Riester, Universität Bielefeld

September 2022

Summer School *Corpus Annotation and Data Analysis (CANDA)*
Georg-August-Universität Göttingen

Rhetorical relations

- Yesterday, we talked about how to construe discourse structure on the basis of QUDs.
- But how are these related to structures using rhetorical relations, such as Rhetorical Structure Theory (RST) or Segmented Discourse Structure Theory (SDRT)?

Mann and Thompson (1988a), Taboada and Mann (2006), and Stede et al. (2017)

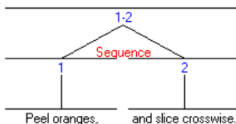
- Coherent texts consist of minimal units, which are connected recursively via rhetorical relations (discourse relations / coherence relations).
- Coherent texts do not contain gaps or unlogical sequences.
- Therefore, texts must be completely connected via rhetorical relations.

- **Nuclearity**
 - According to RST, the relation between two discourse units is often asymmetric: the **Nucleus** is more central, while the **Satellite** is less important.
 - This may but need not coincide with syntactic subordination (hypotaxis).
- **Discourse relations** connect discourse units
- **Hierarchy/recursion**: Connected (complex) discourse units may enter new relations to other units.

Coordination

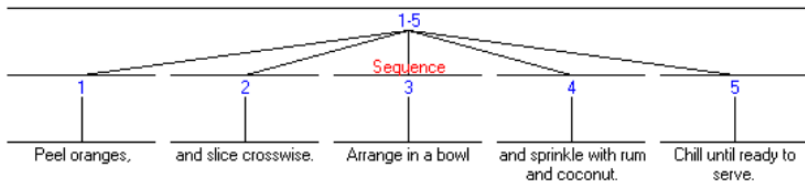
Sentence-internal COORDINATION:

- (1) a. 1. Peel oranges, 2. and slice crosswise.

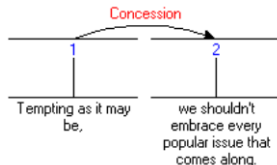


Sentence-external COORDINATION:

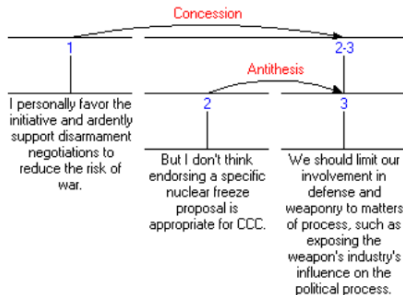
- b. 1. Peel oranges, 2. and slice crosswise. 3. Arrange in a bowl 4. and sprinkle with rum and coconut. 5. Chill until ready to serve.



Sentence-internal CONCESSION:



Sentence-external CONCESSION:
Nucleus itself consisting of two units (2-3)



Example: Relation EVIDENCE

- Nucleus (N): a subjective statement/claim, which the addressee might not accept or might not regard as sufficiently important or positive.
- Satellite (S): a statement that the addressee is likely to accept; usually an “objective” description of a fact.
- Effect: the addressee’s belief in N is increased

(1) [Obama:]

N: let’s face it, my presence on this stage is pretty unlikely.

S: My father was a foreign student, born and raised in a small village in Kenya.

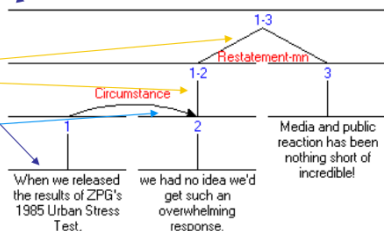
Classification of discourse relations

- **Semantic** relations are used to describe a (possibly complex) state of affairs in the world.
 - e.g. CIRCUMSTANCE, CONDITION, ELABORATION, INTERPRETATION, CAUSE, RESULT, PURPOSE
- **Pragmatic** relations describe the argumentation of the author: What are the claims, and how are they being supported by observations or by other claims?
 - e.g. BACKGROUND, ANTITHESIS, EVIDENCE, CONCESSION, REASON, JUSTIFY, EVALUATION, MOTIVATION, ENABLEMENT
- **Textual** relations work to organize the text and make its understanding easier by providing orienting information, or repetition.
 - e.g. PREPARATION, RESTATEMENT, SUMMARY

- (Typically) **marked** relations
 - CONCESSION (*despite, however, yet, although*), CONDITION (*if*)
- vs. (typically) **unmarked** relations
 - BACKGROUND, RESTATEMENT
- Sequence of units: **Nucleus before Satellite**
 - ELABORATION
- vs. **Satellite before Nucleus**
 - CONCESSION (*despite, although*)

RST trees (Figure by M. Stede)

- A **horizontal line** covers a span of text (possibly made up of further spans)
- A **vertical line** signals the nucleus or nuclei
- A **curve** represents a relation, and the direction of the arrow, the direction of satellite towards nucleus



Stede et al. (2017: 4f.)

- Elementary discourse units correspond to clauses.
- They include adjunct clauses, but not complement / argument clauses.
- All discourse units are required to contain a verb.
- Coordinated clauses and coordinated VPs are segmented.

- Text segmentation: Does a particular sentence or constituent form its own discourse segment?
- Identification of discourse relations: which one is the most plausible. Two analysts may disagree, which relation is the correct one.
- Inventory of discourse relations:
 - Many different classifications have been proposed.
 - Compromise between richness of detail (as many relations as possible) and clarity / inter-annotator agreement (as few relations as possible).

Discourse segmentation in the QUD framework

Segmentation principles in the QUD-tree framework

- Segmentation is oriented along the lines of **information units** (each containing a focus).
- An information unit is a segment of discourse that
 - can function as the **answer to a question**.
 - is semantically **optional** / independent from the preceding or upcoming information.
- Segmentation is even less constrained by syntax than assumed in RST.
- In particular, we argue that **adjuncts** and **conjuncts** form separate discourse units.
- NP- and VP-coordinations are treated as elliptical statements.

Texts are split into elementary units (mainly assertions) at **conjunctions**, i.e. isolate single assertions (A). If necessary, reconstruct ellipses.

A_{28'}: I will always be grateful for everything that you've done

A_{28''}: and I ~~will always be grateful for~~ all the incredible work that you put in.

Text segmentation

- * separate main clauses
- * separate adjunct clause
- * separate coordination

Tonight,*more than 200 years after a former colony won the right to determine its own destiny,*the task of perfecting our union moves forward.*It moves forward because of you.*It moves forward because you reaffirmed the spirit that has triumphed over war*and depression;*the spirit that has lifted this country from the depths of despair to the great heights of hope –*the belief that*while each of us will pursue our own individual dreams,*we are an American family,

Non-at-issue material and adjuncts

Non-at-issue material

“not at issue”: Potts (cf. 2005), Tonhauser (2012), and Horn (2014)

Non-at-issueness (relative to Q):

An expression X in an utterance U is not at issue with respect to the current Question under Discussion Q iff the deletion of X has no effect on the truth-conditions of the main proposition denoted by U – in other words, if X is **optional** with respect to Q .

- Typical NAI material: supplements such as parentheticals, non-restrictive modifiers and other adjunct-like material.

Q₄: {Tonight the American people reminded us that we did WHAT?}

> A₄: [[Tonight,]_T [in this election,]_{NAI} [you,]_T [the American people,]_{NAI} reminded [us]_T that, [while our road has been hard,]_{NAI} [we]_T have [picked ourselves up]_F]~

As argued in Brunetti et al. (2021) adjuncts (e.g. certain PPs or adverbials) can take on different information-structural roles:

Focus

- (2) Q₀: With whom will you have dinner?
> A₀: We will have dinner [with some friends]_F.

Background

- (3) A₁₄: [...] My father got a scholarship to study in [...] America.
Q₁₅: {What happened while he was studying in America?}
> A₁₅: **While studying here**, my father [met my mother]_F.

As argued in Brunetti et al. (2021) adjuncts (e.g. certain PPs or adverbials) can take on different information-structural roles:

Contrastive topic

- (4) Q₁: What do you prefer to eat?
> Q_{1.1}: What do you prefer to eat in the morning?
>> A_{1.1}: [**In the morning**]_{CT} I prefer [cereals]_F.
> Q_{1.2}: What do you prefer to eat during lunch?
>> A_{1.2}: [**For lunch**]_{CT} I normally have [a sandwich]_F.

In all these cases, adjuncts form an integral part of the current discourse unit.

- However, sometime adjuncts can express new, but at the same time **additional, parenthetical, or peripheral information**
- Despite being new information, they do not constitute the main point of the utterance

A: "The greatest fear I have", and I quote you, "regarding the disclosures is nothing will change."

A: That was one of your greatest concerns at the time

A: but in the meantime, there is a vivid discussion about the situation with the NSA, **not only in America but also in Germany and in Brazil.**

IS-peripheral adjuncts

- IS-peripheral adjuncts are part of a broader focus and provide some peripheral or parenthetical information.

- > A₁₅: “The greatest fear I have”, and I quote you, “regarding the disclosures is nothing will change.”
- > Q₁₆: *{What about this fear?}*
- >> Q_{16.1}: *{What about this fear at the time you said that?}*
- >>> A_{16.1}: [[That]_T was [one of your greatest concerns]_F [at the time]_{CT}]~
- >> Q_{16.2}: *{What about this fear in the meantime?}*
- >>> A_{16.2}: but [[in the meantime]_{CT} [there is a vivid discussion about the situation with the NSA, **not only in America but also in Germany and in Brazil**]_F]~

Adjuncts in sentence internal position

- > A₇: [[c'est une expérience ça [que je n'ai jamais oubliée]_F]~
'It's an experience that I've never forgotten'
- > Q₈: *{What about the harshness of this experience?}*
- >> A_{8'}: [ce qui est dur [ce n'est pas surtout quand on est très jeune ce n'est vraiment pas d'être pauvre]_F]~
'What is hard is not, especially when one is very young, it is not just to be poor' [...]
- >> A_{8''}: [[c'est de se dire je n'en sortirai jamais]_F]~
'it is to say "I will never get out of it"'

(Ramm 2011)

(5) a. When did you arrive?

b. I arrived yesterday evening **with some friends**.

- The manner adjunct *with some friends* "encodes **information that is new**, i.e. not part of the background, but **does not contribute to answering the relevant question** and thus cannot be part of the focus in the strict sense either."
- "[...] the adjunct, in a way, **answers a question that has not been asked**" (Ramm 2011: 146).

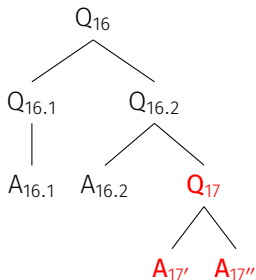
IS-peripheral adjuncts as independent discourse segments

- IS-peripheral adjuncts answer an independent QUD,
- reconstructed on the basis of the semantic content of the host

- > Q₁₆: {*What about this fear?*}
- >> Q_{16.1}: {*What about this fear at the time you said that?*}
- >>> A_{16.1}: [[That]_T was [one of your greatest concerns]_F [at the time]_{CT}]
- >> Q_{16.2}: {*What about his fear in the meantime?*}
- >>> A_{16.2}: but [[in the meantime]_{CT} [there is a vivid discussion about the situation with the NSA,]_F]~
- >>> Q₁₇: {*Where is this vivid discussion?*}
- >>>> A_{17'}: [[not only in America]_F]~
- >>>> A_{17''}: [[but also in Germany and in Brazil]_F]~

IS-peripheral adjuncts as independent discourse segments

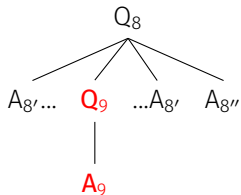
QUD-tree



Sentence internal adjunct answering an independent QUD

- > Q₈: *{What about the harshness of this experience ?}*
- >> A_{8'}...: *ce qui est dur ce n'est pas*
'What is hard is not...'
- >> Q₉: *{When is it not hard to be poor?}*
- >>> A₉: *[surtout quand on est très jeune]_F*
'especially when one is very young'
- >> ...A_{8'}: *ce n'est vraiment pas [d'être pauvre (...)]_F ~*
'...it is not really to be poor'
- >> A_{8''}: *(...) c'est [de se dire je n'en sortirai jamais]_F*
'it is to say "I will never get out of it"'

QUD-tree corresponding to sentence with internal adjunct



- As we said earlier, discourse segments stand in some rhetorical relation with another segment.
- Elementary Discourse Units are defined as:
 - "clauses or clause-like units that are unequivocally the NUCLEUS or SATELLITE of a rhetorical relation that adds some significant information to the text." (Marcu et al. 1999: p. 50).
- In the case of IS-peripheral adjuncts, the rhetorical relation is established with the host sentence.

- Nonsentential phrases, and adjuncts in particular, can be independent discourse units (Schauer 2000, Schauer and Hahn 2000)
 - Events are characterized by typical properties such as **agent, patient, location, instrument, time frames**.
 - These typical properties have traditionally not been treated by use of coherence relations between discourse units.
 - "whenever nontypical, unpredictable information pieces have to be accounted for, coherence relations **may capture their value-adding semantics**."

- The criteria under which adjuncts should be analyzed as independent discourse units are the following:
 - **they refer to non-typical properties, or**
 - their semantic interpretation partially refers to typical properties, but the intended meaning is not fully covered by them;

IS-peripheral adjuncts are in a rhetorical relation with their host

- **ELABORATION relation** (Mann and Thompson 1988b)
 - *"S [= Satellite] presents additional detail about the situation or some element of subject matter which is presented in N [= Nucleus]"*
 - The adjunct does not simply inform about the place of the discussion, but emphasizes on how the discussion has become vivid.

> A₁₅: "The greatest fear I have", and I quote you, "regarding the disclosures is nothing will change."

>>> A_{16.2}: but *[[in the meantime]_{CT} [there is a vivid discussion about the situation with the NSA,]_F~*

>>> Q₁₇: *{Where is this vivid discussion?}*

>>>> A_{17'}: *[[not only in America]_F~*

>>>> A_{17''}: *[[but also in Germany and in Brazil]_F~*

- CIRCUMSTANCE relation

- The temporal clause provides some parenthetical information specifying in what circumstances the speaker's claim (that being poor is not a problem) mostly holds.

> Q₈: *{What about the harshness of this experience ?}*

>> A_{8'}: *ce qui est dur ce n'est pas*

'What is hard is not...'

>> Q₉: *{When is it not hard to be poor?}*

>>> A₉: *[surtout quand on est très jeune]_F*

'especially when one is very young'

>> A_{8'}: *ce n'est vraiment pas [d'être pauvre (...)]_F ~*

'...it is not really to be poor'

>> A_{8''}: *(...) c'est [de se dire je n'en sortirai jamais]_F]~*

'it is to say "I will never get out of it"'

IS-peripheral adjuncts are in a rhetorical relation with their host

- If the adjunct is a separated segment, then it is in a rhetorical relation with its host, but not the other way around.
- In order to be independent segments, adjuncts have to meet *two* criteria:
 - they have to provide new information
 - they have to be in a rhetorical relation with their host.

Combining QUDs and rhetorical relations

(6) [They would give me an African name, Barack, or “blessed”.]₂₅

(7) Q_{25a}: {What would Obama’s parents do with him?}

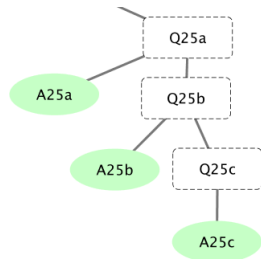
> A_{25a}: [[They]_T would [give]_F me [an African name]_F,]~

> Q_{25b}: {What name would they give to him?}

> > A_{25b}: [[Barack,]_F]~

> > Q_{25c}: {What does *Barack* mean?}

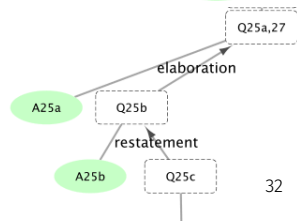
> > > A_{25c}: or [[“blessed”]_F]~.



Coherence relation between sub-clausal units:

A_{25a} ← A_{25b}: ELABORATION

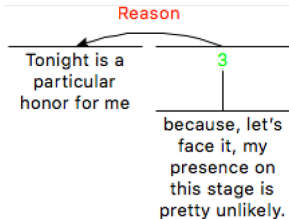
A_{25b} ← A_{25c}: RESTATEMENT



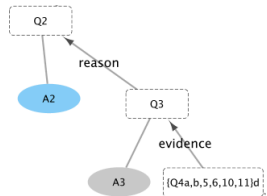
Subordinating Relations

All relations can be translated into question nodes (e.g. Q₃ expresses the same information as the discourse relation)

- (8) A₂: Tonight is a particular honor for me
Q₃: {Why is it a particular honor for Obama to speak on this stage?}
> A₃: because, [let's face it,]_{NAI} [[my presence on this stage]_T [is pretty unlikely]_F]~.



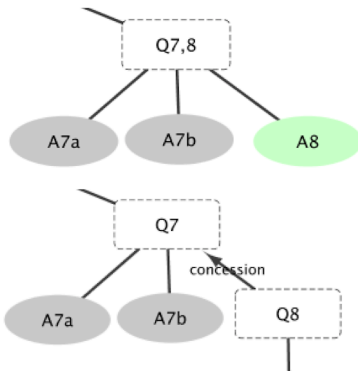
RST representation



QUD representation

Added value of discourse relations

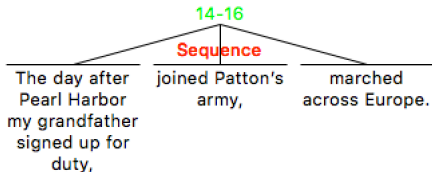
- (9) Q_{7,8}: {What about Obama's paternal grandfather?}
> A_{7a}: [His father,_T [my grandfather,_{NAI}] was [a cook,_F
> A_{7b}: [a domestic servant to the British]_F.
> Q₈: {What did he want despite being a servant?}
> > A₈: But [my grandfather]_T [had larger dreams]_F for [his son]_T.



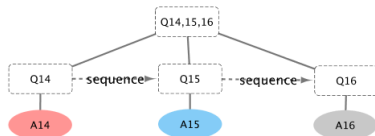
Coordinating relations

(Adversative) CONTRAST or (temporal) SEQUENCE presume subquestions and contrastive topics.

- (10) Q_{14,15,16}: {What did the grandfather do after the Pearl Harbor attack?}
> Q₁₄: {What did the grandfather do on the (exact) day after Pearl Harbor?}
> > A₁₄: [[The day after Pearl Harbor]_{CT} [my grandfather]_T [signed up for duty,]_F]_~
> Q₁₅: {What did the grandfather do at t₁₅?}
> > A₁₅: [[joined Patton's army,]_F]_~
> Q₁₆: {What did the grandfather do at t₁₆?}
> > A₁₆: [[marched across Europe]_F]_~



RST representation












QUD representation



[Click here](#)

[Tree created by Amalia Canes Nápoles]

References

-  Horn, Laurence. 2014. Information structure and the landscape of (non-)at-issue meaning. In Caroline Fery and Shinichiro Ishihara (eds.), *The oxford handbook of information structure*. Oxford University Press.
-  Mann, William and Sandra Thompson. 1988a. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8(3). 243–281.
-  Mann, William and Sandra Thompson. 1988b. Rhetorical structure theory: toward a functional theory of text organization. *Text-interdisciplinary Journal for the Study of Discourse* 8(3). 243–281.
-  Marcu, Daniel, Estibaliz Amorrortu and Magdalena Romera. 1999. Experiments in constructing a corpus of discourse trees. In *Proceedings of the ACL workshop on Standards and Tools for Discourse Tagging*, 48–57. College Park, MD, USA.
-  Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford University Press.

-  Ramm, Wiebke. 2011. Discourse-structural salience from a cross-linguistic perspective: coordination and its contribution to discourse (structure). In Christian Chiarcos, Berry Claus and Michael Grabski (eds.), *Salience: multidisciplinary perspectives on its function in discourse*, vol. 227 (Trends in Linguistics), 143–172. Berlin/New York: De Gruyter Mouton.
-  Schauer, Holger. 2000. From elementary discourse units to complex ones. In *Proceedings of the 1st sigdial workshop on discourse and dialogue - volume 10* (SIGDIAL '00), 46–55. Hong Kong: Association for Computational Linguistics. <https://doi.org/10.3115/1117736.1117742>.
-  Schauer, Holger and Udo Hahn. 2000. Phrases as carriers of coherence relations. In *Proceedings of the annual meeting of the cognitive science society*, 22.
-  Stede, Manfred, Maite Taboada and Debopam Das. 2017. *Annotation Guidelines for Rhetorical Structure*. Potsdam / Burnaby.

-  Taboada, Maite and William Mann. 2006. Rhetorical Structure Theory: Looking back and moving ahead. *Discourse Studies* 8. 423–459.
-  Tonhauser, Judith. 2012. Diagnosing (not)-at-issue content. In *Proceedings of Semantics of Underrepresented Languages in the Americas (SULA)*, 239–254. Amherst, MA, USA.