Rhetorical Stucture Theory, RST (retorisen rakenteen teoria)

• Mann & Thompson 1988 (in Finnish, e.g. Komppa 2013)

• Main assumption: Each part of a coherent carefully written text has its significance for the entirety.

• RST describes the relations between text parts.

• These relations are defined to hold between two non-overlapping text spans, a nucleus and a satellite.

• Typically, a clause is considered such a ”text span”.
**Circumstance** (olosuhde)

Probably the most extreme case of Visitors’ Fever I have ever witnessed was a few summers ago... ...when I visited relatives in the Midwest.

The effect: The situation presented in S provides the framework for interpreting N.
**Non-Volitional Result** (ei-tahdonalainen seuraus)

The blast destroyed the plant and most of the surrounding suburbs. Several thousand people were injured, and about 300 are still in hospital.

N

S

The effect: The situation presented in N could have caused the situation presented in S.

---

**Subject-matter relations**

<table>
<thead>
<tr>
<th>Circumstance</th>
<th>Condition</th>
<th>Elaboration</th>
<th>Evaluation</th>
<th>Interpretation</th>
<th>Means</th>
<th>Non-volitional Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Solutionhood</td>
<td>Unconditional</td>
<td>Unless</td>
<td>Volitional Cause</td>
<td>Volitional Result</td>
<td></td>
</tr>
</tbody>
</table>

**Presentational relations**

<table>
<thead>
<tr>
<th>Antithesis</th>
<th>Background</th>
<th>Concession</th>
<th>Enablement</th>
<th>Evidence</th>
<th>Justify</th>
<th>Motivation</th>
<th>Preparation</th>
<th>Restatement</th>
<th>Summary</th>
</tr>
</thead>
</table>

**Multinuclear relations**

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Joint</th>
<th>List</th>
<th>Restatement</th>
<th>Sequence</th>
</tr>
</thead>
</table>
**Concession** (myönnyys)

Although it is toxic to certain animals, … … evidence is lacking that it has any serious long-term effect on human beings.

The effect: Readers positive regard for the situation presented in N is increased by S.

<table>
<thead>
<tr>
<th>Subject-matter relations</th>
<th>Presentational relations</th>
<th>Multinuclear relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circumstance</td>
<td>Antithesis (<em>increases positive regard</em>)</td>
<td>Contrast</td>
</tr>
<tr>
<td>Condition</td>
<td>Background (<em>increases ability</em>)</td>
<td>Joint</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Concession (<em>increases positive regard</em>)</td>
<td>List</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Enablement (<em>increases ability</em>)</td>
<td>Restatement</td>
</tr>
<tr>
<td>Interpretation</td>
<td>Evidence (<em>increases belief</em>)</td>
<td>Sequence</td>
</tr>
<tr>
<td>Means</td>
<td>Justify (<em>increases acceptance</em>)</td>
<td></td>
</tr>
<tr>
<td>Non-volitional Cause</td>
<td>Motivation (<em>increases desire</em>)</td>
<td></td>
</tr>
<tr>
<td>Non-volitional Result</td>
<td>Preparation</td>
<td></td>
</tr>
<tr>
<td>Otherwise</td>
<td>Restatement</td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td>Summary</td>
<td></td>
</tr>
</tbody>
</table>
**Contrast (kontrasti)**

Animals heal, ... ... but trees compartmentalize ('jakaa osiin').

The effect: Readers recognize the comparability and the difference.

<table>
<thead>
<tr>
<th>Subject-matter relations</th>
<th>Presentational relations</th>
<th>Multinuclear relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circumstance</td>
<td>Solutionhood</td>
<td>Contrast</td>
</tr>
<tr>
<td>Condition</td>
<td>Unconditional</td>
<td>Joint</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Unless</td>
<td>List</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Volitional Cause</td>
<td>Restatement</td>
</tr>
<tr>
<td>Interpretation</td>
<td>Volitional Result</td>
<td>Summary</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-volitional Cause</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-volitional Result</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otherwise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Animals heal, ... ... but trees compartmentalize ('jakaa osiin'). They endure a lifetime of injury and infection by setting boundaries that resist the spread of the invading micro-organisms.

• Analysis proceeds hierarchically from smaller units (= clauses) to larger structures (adjacent sentences, paragraphs and so on).
1) **Darwin as a Geologist**

2) He tends to be viewed now as a biologist, 3) but in his five years on the Beagle his main work was geology, 4) and he saw himself as a geologist. 5) His work contributed significantly to the field.
Animals heal, ... but trees compartmentalize (‘jakaa osiin’). They endure a lifetime of injury and infection by setting boundaries that resist the spread of the invading micro-organisms.

Recognition of the relation always rests on functional and semantic judgements alone, not on morphological or syntactic signals.

This could be seen both as
- **cons (for text analysis):** No direct access to either the writer or the readers. Therefore, descriptions cannot be certain but *plausibility judgements*. Prone to differing interpretations.
- **pros (for NLG):** These relations are language independent!
News automation

• Human and machine → a “man-machine marriage.”
• Journalists perform tasks that algorithms cannot, such as in-depth analysis, interviewing, and investigative reporting.
• So, what algorithms can?

(Græffe 2016)

In news automation, it makes sense to focus on a short news article form (2–4 paragraphs, “sähkeuutinen”) with one essential newsworthy fact (uutiskärki).
→ I collected small corpus of such articles.
A HEADING IS PLACED HERE!

Restatement

A leading sentence... (N)

A couple of more sentences (S).

The most-nucleus unit

One box = one paragraph

Elaboration

(N)
(S)

Elaboration

(N)
(S)

Elaboration

(N)
(S)

Other common relations:
Background
Summary
Sequence
Non-Volitional Cause/Result
Elaboration

Nucleus is often an assertion, followed by a Satellite.

The relations is based on **content** BUT is often also marked by **conventional means of cohesion:**
- Pronouns (Party X → It)
- Repetition
- Nominalization (Party X has won → The win...)
- Restrictive relative-clause
- Sentences placed within a paragraph

Komppa 2013

Other common relations:
- Background
- Summary
- Sequence
- Non-Volitional Cause/Result
As far as automated news production is focusing on a simple and succinct news format, the RST won’t offer any breakthrough. However, once we are ready for more complex format, the theory help us clarify the structure and keep it abstract enough for language independency.

Keskusta johtaa, Kokoomus ja Vihreät todennäköisesti nousevat

Keskusta johtaa reilusti toisena tulevaa Kokoomusta. Tällä hetkellä se on saamassa äänistä 23 prosenttia, mikä lisäisi puolueen paikkamäärää Eduskunnassa jopa 7 edustajalla.

On kuitenkin todennäköistä, että tulos tulee äänenlaskennan edetessä vielä muuttumaan. Laskematta ovat nimittäin suurten kaupunkien vaalipäivänä annetut äänet – juuri ne, jotka perinteisesti kartuttavat erityisesti Kokoomuksen ja Vihreiden mutta myös SDP:n ja Perussuomalaisten äänisaalista.