1. Theories of discourse structure

Early discourse resources were task-oriented

- For example, an engineering explaining to an apprentice how to repair a pump

And the structure of task-oriented discourse often mirrored the structure of the task

Pre-computational theories had focused on narrative structures

- Story grammars, so-called, basically taxonomic and flat

These gave way to structurally rich generative models

- Grosz and Sidner's Discourse Theory
- Mann and Thompson's Rhetorical Structure Theory (RST)

Both were expressed in terms of *coherence relations*

- Also sometimes called discourse relations
- Between the interpretations of sentences/utterances
  - After some amount of abstraction

Still depending on observable phenomena (cohesion) to detect/identify them

2. Grosz and Sidner’s Discourse Theory (GSDT)

GSDT approaches the hierarchical nature of discourse at three levels

- Linguistic structure
  - What is actually said/written
- Intentional structure
  - Speaker’s goals and purposes
  - Organised into a relational structure
  - I.e. this is where discourse relations come in
- Attentional structure
  - Speaker’s focus of attention
3. GSDT: Intentional structure

There is an overall *discourse purpose* (DP)

- The basic purpose of the whole discourse

A discourse consists of *discourse segments* (DS)

- Each segment has one or more *discourse segment purposes* (DSPs)
  - How a segment contributes to the DP

As well as *segment relations*

- **satisfaction-precedence**  DSP₁ must be satisfied before DSP₂
- **dominance**  DSP₁ dominates DSP₂ if fulfilling DSP₂ constitutes part of fulfilling DSP₁

4. Attentional state

Attention is represented in terms of a *focus stack*

- A stack of *focus spaces*
  - Each containing objects, properties and relations salient during its corresponding DS
  - As well as its DSPs
- That is, content plus purpose

The discourse focus is always on the focus space at the top of the stack

State changes are modeled by transition rules controlling the addition/deletion of focus spaces

- Information at lower levels may or may not be available at higher levels

Focus spaces are pushed onto the stack when a new DS is detected

- And popped when one is completed

5. Discourse structure influences coreference

Consider our earlier example:

```
Welcome to word processing:
  ◦ That's using a computer to type letters and reports [push]
    ■ Make a typo? [push]
    ■ No problem
    ■ Just back up, type over the mistake, and it's gone
  ◦ And, it eliminates retyping [pop]
```

Or this (more realistic, but still not real):
6. Detecting subtopics/Identifying discourse relations

There are a variety of cues that make these structures easier to recognise:

- **implicit** lexical chains, tense and aspect
- **explicit cue phrases**, including:
  - conjunctions
    - 'because', 'but'
  - conjunctive adverbials
    - 'nevertheless', 'instead'
  - temporal adverbials
    - 'then', 'afterwards'

- **suprasegmental** intonational variation

Different cues, different actions

- push
- pop
- chain
  - That is, pop then push

7. Learning cues

The Penn Discourse TreeBank provides a resource for supervised learning of cues:

- Annotated with connectives, their arguments, and the senses they convey
- Approximately 18,000 explicit connectives (cue phrases) and 16,000 implicit ones
As with many other linguistic phenomena, the distribution of connectives follows Zipf's law

<table>
<thead>
<tr>
<th>Connective</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>but</em></td>
<td>3308</td>
</tr>
<tr>
<td><em>and</em></td>
<td>3000</td>
</tr>
<tr>
<td><em>if</em></td>
<td>1223</td>
</tr>
<tr>
<td><em>because</em></td>
<td>858</td>
</tr>
<tr>
<td><em>while</em></td>
<td>781</td>
</tr>
<tr>
<td><em>however</em></td>
<td>465</td>
</tr>
<tr>
<td><em>therefore</em></td>
<td>26</td>
</tr>
<tr>
<td><em>otherwise</em></td>
<td>24</td>
</tr>
<tr>
<td><em>as soon as</em></td>
<td>20</td>
</tr>
<tr>
<td><em>if and when</em></td>
<td>5</td>
</tr>
<tr>
<td><em>conversely</em></td>
<td>2</td>
</tr>
</tbody>
</table>

The corpus isn't large enough to have all the connectives we might expect.

**8. Discourse relations**

What kinds of things are we looking for?

- Kinds of coherence

How *do* discourse segments "hang together"?

Essentially an **abductive** question

- Abduction is reasoning to the best explanation

Which often *is* explanation

Compare

```
John hid Bill's car keys. He was drunk
John hid Bill's car keys. He likes spinach
```

Here's a list from Mann & Thompson (1988) *Rhetorical Structure Theory*:

- Circumstance
- Solutionhood
- Elaboration
- Background
- Enablement and Motivation
  - Enablement
  - Motivation
- Evidence and Justify
  - Evidence
  - Justify
- Relations of Cause
  - Volitional Cause
  - Non-Volitional Cause
  - Volitional Result
  - Non-Volitional Result
- Purpose
- Antithesis and Concession
  - Antithesis
  - Concession
- Condition and Otherwise
  - Condition
9. RST, cont'd

Relations defined by constraints on the participants

* called **nucleus** (**N**) and **satellite** (**S**)  
* With respect not only to **N** and **S**  
* But also the writer/speaker (**W**) and the reader/listener (**R**)  

For example, the **Evidence** relation

<table>
<thead>
<tr>
<th>Relation Name</th>
<th>Constraints on N</th>
<th>Constraints on S</th>
<th>Constraints on N+S</th>
<th>Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence</td>
<td>R might not believe <strong>N</strong> to a degree satisfactory to <strong>W</strong></td>
<td><strong>R</strong> believes <strong>S</strong> or will find it credible</td>
<td><strong>R</strong>'s comprehending <strong>S</strong> increases <strong>R</strong>'s belief of <strong>N</strong></td>
<td><strong>R</strong>'s belief of <strong>N</strong> is increased</td>
</tr>
</tbody>
</table>

Relations may be either symmetric (e.g. Contrast) or asymmetric (e.g. Purpose)

* There's a set of graphical conventions for diagramming relations and their overall pattern in a discourse
See J&M 21.2.1 for more on RST

- But ...

The exact details of the list of relations and their definitions are not important

- But their length and subjective nature are

I'm sceptical of the ultimate viability of attempting to categorise discourse relations this way

10. Contrasting approaches to coreference in discourse

A motivating antecedant to Grosz's discourse work was Hobbs's 1978 pronoun resolver

- Originally purely syntactic
- Gender, person, number checks
- Carefully-crafted search order
  - Up from the pronoun
  - Across left-to-right, breadth-first, for each S or NP as you go up
  - Back to the previous sentences

A revised version also included selectional restrictions
The use of selectional restrictions is necessary to reject "536" (dates don't move) and "the castle ..." (large buildings don't move either) to arrive at "the residence ..." (allowing for metaphorical motion, I guess?)

Grosz (at SRI) and Sidner (at MIT) moved from this kind of approach

- Where all sentences are treated equally

to one dependent on hierarchical discourse structure

- Called **Centering theory**
- Building on their earlier approach to discourse structure
- Key roles for two kinds of center:
  - One **backward-looking** (the current focus of attention)
  - Multiple **forward-looking** (potential future foci)
- And rules associated with push, pop etc. for updating these

See J&M 21.6.1 and 21.6.2 for detailed expositions of these two approaches