

## TextLink: Resource survey compiled by WG1

Jiri Mirovsky  
Manfred Stede

### Inventories

- **Lexconn** 1.0, 2009
- Charlotte Roze, Laurence Danlos, Philippe Muller (Paris, Toulouse)
- freely available XML
- French
- 328 connective words, phrases
- syntactic type, semantic relation (SDRT/RST)

```

<connecteur cat="adv" id="c225" relations="narration" type="coord">
  <forme>
    ensuite
  </forme>
</connecteur>

```

```

<connecteur cat="adv" id="c226" relations="continuation" type="coord"> •
  <forme>
    ensuite
  </forme>
  <exemple>
    Je trouvais mille prétextes pour ne pas descendre de voiture.
    D'abord, il n'y avait pas de place pour se garer. Ensuite, ils devaient être en
    train de dîner quelque part. Des choses comme ça.
  </exemple>
</connecteur>

```

## Inventories

- **DimLex** 2.0, 2015 (1.0, 1998)
- Manfred Stede, Tatjana Scheffler (Potsdam)
- freely available XML
- German
- 290 connective words, multiwords
- syntactic features, ambiguity info, semantic relation (PDTB 3)

```

<entry id="k5" word="allerdings">
  <orth type="cont" canonical="1" onr="k5o1">
    <part type="single">allerdings</part>
  </orth>

  <disambi>
    <conn_d>1</conn_d>
    <sem_d>0</sem_d>
  </disambi>

  <focuspart>0</focuspart>
  <non_conn_reading>
    <example type="ADV" tfreq="0">"Hast du das gehoert?" "Allerdings"!<</example>
  </non_conn_reading>

  <stts> <example type="ADV" tfreq="236">Das neue Cabrio ist auf dem Markt, allerdings ist
  es so teuer, dass man es sich nicht kaufen sollte.</example> </stts>

  <syn> <cat>konnadv</cat>
    <ordering> <ante>0</ante> <post>1</post> <insert>0</insert> </ordering>
  </syn>
</entry>

```

## Inventories

- **Hebrew Discourse Markers 1.0, 2009**
- Yael Maschler (Haifa)
- available in print only
- Hebrew
- textual, interpersonal discourse markers in spoken language
- semantic, pragmatic, prosodic features

## Inventories

- Spanish Discourse Particles
- Antonio Briz, Salvador Pons, José Portolés (Valencia)
- dictionary for human users, available online: <http://www.dpde.es/>
- Spanish
- syntax, semantics, prosody, ..., English/Italian/Portuguese correspondents

## Inventories

- TextLink activities so far:
  - Correspondences between French Lexconn and German DiMLex (STSM, Margot Colinet 2015)
  - ...more ?

## Corpora

### Corpora: two subgroups

- available and some XML
- Connectives (etc), a la PDTB or similar
  - French Discourse Treebank
  - ? CST News Corpus (Brazilian Portugese)
  - Potsdam Commentary Corpus (German)
  - Prague Dependency Treebank (Czech, English)
  - Penn Discourse Treebank (English)
  - Turkish Discourse Bank
  - RST Signaling Corpus (English)
- Coherence relations a la RST, SDRT, SandersEtAl
  - Annodis (French)
  - Potsdam Commentary Corpus (German)
  - DiscAn (Dutch)
  - RST Signaling Corpus (English)
  - Basque RST Treebank
  - Multilingual RST Treebank (English, Spanish, Basque)