Contrasting (mainly) Coreference In Czech And German

From Different Frameworks to Joint Results

GECCo-PDiT

or

PDT-GECCo comparison during the last two years

Anna Nedoluzhko, Katja Lapshinova-Koltunski
What we have:

- two large-scale annotated corpora

GECCO  PDT 3.0
Motivation - general

- compare languages (to know more about each)
- area of discourse phenomena
  - connectives
  - coreference $\rightarrow$ coreference chains
- variation of coreference chains reflects cohesion
Motivation - material

• take advantage of the existing resources reflecting systemic peculiarities of Czech and German
• use Czech and German data annotated on the basis of two different frameworks
comparable and parallel texts in English and German (ca. 80,000 sentences)

written and spoken texts

different genres and registers

morpho-syntactic annotation

texts in Czech (ca. 50,000 sentences)

written texts

journalistic texts with further genre classification

multilayer annotation: morphological, analytical and tectogrammatical

General comparison
Discourse phenomena comparison

<table>
<thead>
<tr>
<th>cohesive devices</th>
<th>textual phenomena</th>
</tr>
</thead>
<tbody>
<tr>
<td>(co)reference</td>
<td>coreference</td>
</tr>
<tr>
<td>lexical cohesion</td>
<td>bridging</td>
</tr>
<tr>
<td>substitution</td>
<td>-</td>
</tr>
<tr>
<td>ellipsis</td>
<td>ellipsis captured in dependency trees</td>
</tr>
<tr>
<td>conjunctives</td>
<td>connectives, arguments, relations</td>
</tr>
</tbody>
</table>
two different theoretical approaches

Halliday and Hasan, 1976

Sgall 1986, Functional grammar,
Zikánová et al., 2015
PDTB-like discourse
He 'd never even bothered to read it. **But** Truman had.
Creating a common scheme

English, German, Czech, other languages

GECCo: English, German

PDT: Czech

various registers and genres

various genres
Discourse Structuring Devices

- DSDs
- Ellipsis
- Bridging/lexical cohesion
- Coreference

Bar chart showing comparisons between PDT and GECCo.
<table>
<thead>
<tr>
<th>Preced. – succ.</th>
<th>reason - result</th>
<th>confrontation</th>
<th>conjunction</th>
</tr>
</thead>
<tbody>
<tr>
<td>synchronous</td>
<td>Pragm. reas. – result</td>
<td>opposition</td>
<td>instantiation</td>
</tr>
<tr>
<td>purpose</td>
<td>pragmatic contrast</td>
<td>specification</td>
<td></td>
</tr>
<tr>
<td>explication</td>
<td>Restr. opp. + except.</td>
<td>equivalence</td>
<td></td>
</tr>
<tr>
<td>condition</td>
<td>concession</td>
<td>generalization</td>
<td></td>
</tr>
<tr>
<td>Pragm. cond.</td>
<td>correction (replacement)</td>
<td>conjunctive alternative</td>
<td></td>
</tr>
</tbody>
</table>

**TEMPORAL**

<table>
<thead>
<tr>
<th>TEMPORAL</th>
<th>CONTINGENCY</th>
<th>COMPARISON (CONTRAST)</th>
<th>EXPANSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEMPORAL</td>
<td>CAUSAL</td>
<td>ADVERSATIVE</td>
<td>ADDITIVE</td>
</tr>
<tr>
<td></td>
<td>relation between events</td>
<td>relation of causality/dependence between</td>
<td>relation of contrast/alternative, for two events which are not true at the same time</td>
</tr>
<tr>
<td>after, afterwards, at the same time..</td>
<td>because, therefore, that's why...</td>
<td>yet, although, by contrast...</td>
<td>and, in addition...</td>
</tr>
<tr>
<td>nachdem, danach, gleichzeitig..</td>
<td>weil, deshalb, aus diesem Grund..</td>
<td>doch, obwohl, im Gegensatz dazu..</td>
<td>und, außerdem..</td>
</tr>
</tbody>
</table>

**discourse markers**

(Attitude markers, modal particles) - in PDT not considered as connectives
Ellipsis

- structural ellipses (of governing verbs and nouns),
- anaphoric zeros (Their reaction was 0 to do nothing and 0 ride it out),
- arguments with control constructions (Peter wants to [Peter] sleep),
- arguments with general arguments (Jane sells at Bata [what] [to whom])

- only in the case of textual relations across grammatical domains
He'd never even bothered to read it. But Truman had [bothered to read id].
Coreference

- Both implicit and explicit relations of coreference are annotated, including indefinite NPs
- Only with explicit marker
John left Mary. He wanted to see his mother.

It means that he doesn’t really love Mary.

My daughter likes to dress herself without my help.

Alex is the boy who kissed Mary.

If you want to prove it, take a look at his outfit.

At last, Mary came to Prague herself and found the Prague atmosphere quite casual.

Peter wants to sleep.

Mary suggested John to go to the theater, but John ignored her wish.

John and Mary kissed Ø.PAT

John asked his mother to advise him how he should behave with Mary, but mother ignored her son’s wish.

John asked mother to come to Mary’s place with him but she decided not to go there.

John saw Mary run around the lake.

John & Mary Ø.ACT run around the lake.
Bridging Relations / Lexical Cohesion

- **PDT**
  - DSDs
  - Ellipsis
  - Bridging/lexical cohesion
  - Coreference

- **GECCo**
  - Bridging/lexical cohesion significantly higher than PDT

- **PDT**
  - 6 specific types of semantic/pragmatic relations

- **GECCo**
  - Many types with explicit marker
**repetition**: same word is repeated (*artist - Artist - artists*)

**synonym**: same meaning (*film - movie*)

**antonym**: opposite meaning (*day - night, husband - lover*)

**hypernym**: more generic word (*street people - vagabond*)

**co-hyponym**: word with same degree of specificity (*movie director - producer - writer*)

**meronym**: a word that names a part of a larger whole (*film - scenes*)

**co-meronym**: a word that names another part of a larger whole (*scenes - characters*)

**holonym**: the larger whole (*film as related to scenes and characters*)

**PART-WHOLE**: (*film - scenes*)

**SET – SUBSET/element of the set**: (*the beds - one of the beds*)

**function – object**: (*prime-minister – government, trainer – football team*)

**semantic/contextual CONTRAST**: (*It is necessary to have lofty aims and not to be satisfied with small (ones)*)

**explicit anaphora without coreference**: (*world war – at that time*)

**REST**: (*family (grandfather - grandson), place – inhabitant, author – work, same denomination to support cohesion of the text (a chance helped – another chance entered the game ...), and event – participant of the event (enterprise - entrepreneur)*)
MUTUAL RESULTING SCHEMA

PDT

coreference

IDENTITY

NON-IDENTITY

bridge

GECCo

coreference

lexical cohesion

discourse connectives

discourse connectives

ELLIPSIS

ellipses

substitution
## MUTUAL RESULTING SCHEMA

<table>
<thead>
<tr>
<th>featID</th>
<th>Czech</th>
<th>German</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDENTITY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>id1</td>
<td>coreference with pronouns</td>
<td>coreference with heads (no extended reference)</td>
</tr>
<tr>
<td>id2</td>
<td>pronouns with arrows to segments and events</td>
<td>reference to verb phrases and longer segments</td>
</tr>
<tr>
<td>id3</td>
<td>NP coreference</td>
<td>coreference with modifiers or def.articles</td>
</tr>
<tr>
<td>id4</td>
<td>coreference with the word <em>same</em></td>
<td>general comp.reference</td>
</tr>
<tr>
<td>id5</td>
<td>coreference with local and temporal adverbs</td>
<td>coreference with local and temporal adverbs</td>
</tr>
<tr>
<td>NON-IDENTITY</td>
<td>relations of MERONYMY</td>
<td>relations of MERONYMY</td>
</tr>
<tr>
<td>nonid1</td>
<td>bridging CONTRAST</td>
<td>particular comparative reference and antonyms</td>
</tr>
<tr>
<td>nonid2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISCOURSE RELATIONS</td>
<td>temporal</td>
<td>temporal</td>
</tr>
<tr>
<td>temp</td>
<td>contingency</td>
<td>causal</td>
</tr>
<tr>
<td>cont</td>
<td>comparison (contrast)</td>
<td>adversative</td>
</tr>
<tr>
<td>comp</td>
<td>expansion</td>
<td>additive</td>
</tr>
<tr>
<td>expan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELLIPSIS</td>
<td>ellipsis</td>
<td>cohesive ellipsis</td>
</tr>
<tr>
<td>ellipsis</td>
<td>textual ellipsis</td>
<td></td>
</tr>
</tbody>
</table>

comp with discourse relations of comparison/contrast

**German**: Arbeiten wie die Polen, **aber** leben wie die Japaner… [Work like the Poles, **but** live like the Japanese…]

**Czech**: Poslední statistické sčítání dopravy proběhlo v roce 1990. Za poslední tři roky se **však** na Českých silnicích zvýšil provoz. [The latest statistical traffic census took place in 1990. Over the past three years, **however**, traffic on Czech roads has increased.]
BUT: could be different

William Gates and Paul Allen in 1975 developed an early language-housekeeper system for PCs, and Gates became an industry billionaire six years after IBM adapted one of these versions in 1981.

(from experimental English excerption)

PDT: reason-result

GECCCo: additive relation
German: …praktisch wird es dazu nicht kommen – dafür ist in Deutschland die Bereitschaft zur Solidarität, der Glaube an das “für alle” zu groß. Eine andere Gefahr ist da weit realer: daß die Deutschen… [In practice… it will not come to that— the readiness to practice solidarity, and people’s belief in the “for all” is too pronounced in Germany. Another dangers much more real, however: that the Germans….]

Czech: Jednotlivá studia v apartmánech jsou vybavena kuchyní, takže je možná individuální příprava stravy. [Studio apartments are equipped with kitchens, so everyone can prepare their own food.]
nonid2 with contrast relations

**German:** Dazu gehören zum Beispiel die Halbierung der Energie- und Rohstoffintensität bis 2020 gegenüber 1990 und die Verdoppelung des Anteils erneuerbarer Energien am Energieverbrauch bis 2010. [For example, halving the amount of power and raw material consumption by 2020 compared to 1990 levels and doubling the percentage of renewable energy used as part of total energy consumption by 2010.]

**Czech:** Saldo běžného účtu platební bilance podle odhadu dosáhlo vloni cca 600 USD... I když letos a příští rok je nutné počítat se zpomalením růstu vývozu, prognózujeme, že saldo přesto zůstane kladné. [The balance of the current account deficit is estimated to reach $600 last year ... Although this and the next year we expect the slowdown in export growth, we forecast that the deficit will still remain positive.]
Nominal Ellipsis

German: *All das ist eine kleine Revolution. Die grössere [] ist diese: ... [But there is also a bigger [revolution], and it is this]*

Czech: *Klienti pojišťoven, které ukončí svou činnost, se automaticky vrátí k Všeobecné []. [Clients of insurance companies which shut down will automatically return to the General [one]].*
Applications of the schema

- To calculate all kinds of textual/cohesive phenomena on an excerpt of texts in Czech and German (Katja-Anja-Kerstin, Budapest & LREC 2016)
- To consider coreference chains in Czech and German (Anja & Katja, Dialog 2016)
- To consider abstract anaphora in Czech and German (Anja & Katja, CORBON, San Diego 2016)
- Reduced morphology
- More isolating
- Less flexible word order
- Definite article

- Morphologically rich
- Less isolating
- Flexible word order
- No definite article
Data Excerption

• written discourse
• essays
• comparable topics (economic, political and social issues)
Data Excretion - 1

For general textual phenomena (Budapest, LREC)

<table>
<thead>
<tr>
<th>id</th>
<th>topics</th>
<th>sent</th>
<th>tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Germany and social market economy</td>
<td>121</td>
<td>2035</td>
</tr>
<tr>
<td>2</td>
<td>Optimistic remarks on globalisation</td>
<td>47</td>
<td>971</td>
</tr>
<tr>
<td>3</td>
<td>Politics and globalisation</td>
<td>103</td>
<td>1871</td>
</tr>
<tr>
<td>4</td>
<td>Globalisation and new challenges</td>
<td>27</td>
<td>478</td>
</tr>
<tr>
<td>5</td>
<td>The biggest currency changeover</td>
<td>85</td>
<td>1460</td>
</tr>
<tr>
<td>6</td>
<td>Globalisation and market economy</td>
<td>80</td>
<td>1782</td>
</tr>
<tr>
<td>7</td>
<td>Global market and technical progress</td>
<td>108</td>
<td>1851</td>
</tr>
<tr>
<td>8</td>
<td>Economic and technological changes</td>
<td>73</td>
<td>1795</td>
</tr>
<tr>
<td>9</td>
<td>Doctors and medical system</td>
<td>92</td>
<td>2687</td>
</tr>
<tr>
<td></td>
<td><strong>Total: ALL GERMAN</strong></td>
<td><strong>736</strong></td>
<td><strong>14930</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>id</th>
<th>topics</th>
<th>sent</th>
<th>tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>Germany, politics and history</td>
<td>170</td>
<td>687</td>
</tr>
<tr>
<td>6</td>
<td>Housing</td>
<td>83</td>
<td>1644</td>
</tr>
<tr>
<td>7-8</td>
<td>Technological changes</td>
<td>73</td>
<td>1795</td>
</tr>
<tr>
<td>9-12</td>
<td>Politics</td>
<td>121</td>
<td>1854</td>
</tr>
<tr>
<td>13-14</td>
<td>Economics</td>
<td>149</td>
<td>2568</td>
</tr>
<tr>
<td>15-16</td>
<td>Unemployment</td>
<td>112</td>
<td>2252</td>
</tr>
<tr>
<td>17</td>
<td>Television</td>
<td>55</td>
<td>969</td>
</tr>
<tr>
<td></td>
<td><strong>Total: ALL CZECH</strong></td>
<td><strong>763</strong></td>
<td><strong>11769</strong></td>
</tr>
</tbody>
</table>
For coreference *(Moscow, USA)*
written essays with comparable topics on economic, political and social issues

<table>
<thead>
<tr>
<th>Topics</th>
<th>sent</th>
<th>tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Germany and social market economy</td>
<td>121</td>
<td>2035</td>
</tr>
<tr>
<td>2 Optimistic remarks on globalisation</td>
<td>47</td>
<td>971</td>
</tr>
<tr>
<td>3 Politics and globalisation</td>
<td>103</td>
<td>1871</td>
</tr>
<tr>
<td>4 Globalisation and new challenges</td>
<td>27</td>
<td>478</td>
</tr>
<tr>
<td>5 The biggest currency changeover</td>
<td>85</td>
<td>1460</td>
</tr>
<tr>
<td>6 Globalisation and market economy</td>
<td>80</td>
<td>1782</td>
</tr>
<tr>
<td>7 Global market and technical progress</td>
<td>108</td>
<td>1851</td>
</tr>
<tr>
<td>8 Economic and technological changes</td>
<td>73</td>
<td>1795</td>
</tr>
<tr>
<td>Total: ALL GERMAN</td>
<td>645</td>
<td>12243</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>id</th>
<th>topics</th>
<th>sent</th>
<th>tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>Germany, politics and history</td>
<td>170</td>
<td>687</td>
</tr>
<tr>
<td>6</td>
<td>Housing</td>
<td>83</td>
<td>1644</td>
</tr>
<tr>
<td>7</td>
<td>Technological changes</td>
<td>73</td>
<td>1795</td>
</tr>
<tr>
<td>8-9</td>
<td>Politics</td>
<td>121</td>
<td>1854</td>
</tr>
<tr>
<td>10-13</td>
<td>Economics</td>
<td>149</td>
<td>2568</td>
</tr>
<tr>
<td>14</td>
<td>Unemployment</td>
<td>112</td>
<td>2252</td>
</tr>
<tr>
<td>15</td>
<td>Television</td>
<td>55</td>
<td>969</td>
</tr>
<tr>
<td>Total: ALL CZECH</td>
<td>628</td>
<td>11399</td>
<td></td>
</tr>
</tbody>
</table>
## Analysis on Data Excerpton – 1
Subcategories (*norm. per 1000*)

<table>
<thead>
<tr>
<th>featID</th>
<th>id1</th>
<th>id2</th>
<th>id3</th>
<th>id4</th>
<th>id5</th>
<th>nonid1</th>
<th>nonid2</th>
<th>temp</th>
<th>cont</th>
<th>comp</th>
<th>expan</th>
<th>ellipsis</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>88.4</td>
<td>38.2</td>
<td>144.7</td>
<td>3.3</td>
<td>12.0</td>
<td>52.9</td>
<td>28.8</td>
<td>106.5</td>
<td>52.2</td>
<td>79.0</td>
<td>181.5</td>
<td>14.1</td>
</tr>
<tr>
<td>Czech</td>
<td>97.7</td>
<td>64.6</td>
<td>597.3</td>
<td>0.0</td>
<td>10.2</td>
<td>88.4</td>
<td>37.4</td>
<td>14.4</td>
<td>66.3</td>
<td>86.7</td>
<td>136.8</td>
<td>50.1</td>
</tr>
</tbody>
</table>

### Identity

- **id1**: coreference with pronouns
- **id2**: pronouns with arrows to segments and events
- **id3**: NP coreference
- **id4**: coreference with the word *same*
- **id5**: coreference with local and temporal adverbs

### Non-Identity

- **nonid1**: relations of MERONYMY
- **nonid2**: bridging CONTRAST

### Discourse Relations

- **temp**: temporal
- **cont**: contingency
- **comp**: comparison (contrast)
- **expan**: expansion

### Ellipsis

- **ellipsis**: textual ellipsis

- **German**: coreference with heads (no extended reference), reference to verb phrases and longer segments, coreference with modifiers or def.articles, general comp.reference, coreference with local and temporal adverbs

- **Czech**: relations of MERONYMY, particular comparative reference and antonyms

- **Czech**: temporal, causal, adversative, additive

- **German**: cohesive ellipsis
Textual phenomena on Data Excerption - 1
Textual phenomena on Data Excerption - 1
RESULTS

- interoperable schemes permit multilingual analyses of resources from different approaches, saves time and effort as no additional annotation is required
- insights into differences between German and Czech
- limitations of the dataset: although comparable, still different sources, authors, size
- numbers show more differences in conceptualizations and annotation details, not in languages so far
MUTUAL RESULTING SCHEMA

- PDT
  - coreference
  - bridging
  - discourse connectives
  - ellipsis

- GECCo
  - coreference
  - lexical cohesion
  - discourse connectives
  - ellipsis

- IDENTIFY
  - DISCOURSE
  - ELLIPSIS

- NON-IDENTITY
  - substitution
Coreference chains and abstract anaphora

• Are there any language contrasts with respect to coreference chain features between German and Czech?
• Are there any differences between German and Czech if abstract coreference is concerned?
Features to answer:

- longest chains (chains with the greatest number of elements in a text),
- number of coreference pairs (= number of anaphorically referring expressions),
- types of anaphors (e.g. pronouns, nominal modifiers, temporal and local adverbs and so on),
- number of anaphora referring to antecedents other than NP/pronoun,
- types of anaphora referring to antecedents other than NP/pronoun
### Chain properties in German and Czech

<table>
<thead>
<tr>
<th></th>
<th>German</th>
<th>Czech</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of coreference chains</td>
<td>225</td>
<td>550</td>
</tr>
<tr>
<td>length of coreference chains (mean value per text)</td>
<td>2.45</td>
<td>3.2</td>
</tr>
<tr>
<td>longest chain</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>number of coreference relations</td>
<td>327</td>
<td>1231</td>
</tr>
</tbody>
</table>
### Chain properties in German and Czech

<table>
<thead>
<tr>
<th></th>
<th>German</th>
<th>Czech</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of coreference chains</td>
<td>225</td>
<td>550</td>
</tr>
<tr>
<td>length of coreference chains (mean value per text)</td>
<td>2.45</td>
<td>3.2</td>
</tr>
<tr>
<td>longest chain</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>number of coreference relations</td>
<td>327</td>
<td>1231</td>
</tr>
<tr>
<td></td>
<td>996</td>
<td>1027</td>
</tr>
</tbody>
</table>

Note: German numbers are adjusted by adding repetitions and removing zeros.
### Chain properties in German and Czech

<table>
<thead>
<tr>
<th></th>
<th>German</th>
<th>Czech</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of coreference chains</td>
<td>225</td>
<td>550</td>
</tr>
<tr>
<td>length of coreference chains (mean value per text)</td>
<td>2.45</td>
<td>3.2</td>
</tr>
<tr>
<td>longest chain</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>number of coreference relations</td>
<td>327</td>
<td>1231</td>
</tr>
</tbody>
</table>
Longest chains in German and Czech

German
mostly contain
• personal and demonstrative pronouns
• definite nominal phrases

Czech
mostly contain
• hyperthematic named entities

Lexical cohesion in GECCO
Úvahy padesát let od SNP

Povstání a vůbec celou antifašistickou rezistenci považuji za významnou událost, která podala nefalšovaný obraz Slovenska, řekl v srpnu 1994 v rozhovoru pro slovenský deník Pravda předseda Slovenského svazu protifašistických bojovníků generál Ján Husák.

Také nejvýznamnější slovenští politici od prezidenta přes premiéra až po předsedu parlamentu považují Slovenské národní povstání za pozitivní jev ve slovenských dějinách.

Dnes je už známé, že historie SNP byla překroucena ve prospěch komunistické ideologie.

Po listopadu 1989 sice tento černobílý pohled zmizel, ale vakuum využili evážně exiloví historici a spojili obhajobu Slovenského státu s odsouzením SNP.

Oprostit SNP od nánosu lží a mýtů si kladla za cíl první vědecká konference o SNP, která se uskutečnila letos v dubnu v Donovalech.

S blížícím se 50. výročí SNP dochází na politické scéně ve slovenském tisku k souboji dvou diametrálně odlišných koncepcí výkladu SNP: na straně jedné se povstání označuje za puč proti vlastní státnosti a na straně druhé je SNP chápáno jako vědomé odmítnutí totality a fašismu.

Podle názoru místopředsedkyně Slovenské národní strany (SNS) Evy Slavkovské přineslo SNP slovenskému národu jen to, že povstal se zbraní v ruce proti fašismu a nepřijal ho trpělivě jako český národ.

ΩVzalo mu však mnohem více: vlastní samostatnost.

Podle komentáře slovenského levicového listu Pravda povstání bránilo hodnoty, na kterých Evropa stála a stojí.

SNS a další národně orientované strany (např. Křesťanská sociální unie, Společnost dr. Jozefa Tisa, Slovenská národní jednota) se přiklánějí k názoru, že šlo o puč proti vlastní státnosti.

Podle představitelů Slovenské národní jednoty Češi v SNP obrali Slováky o vlastní stát a na 40 let jim vnutili čechokomunistický režim.

Strana demokratické levice po přehodnocení faktu, že vedoucí silou SNP byla komunistická strana, se jednoznačně staví za antifašistický a demokratický odzak povstání.

O postoji Hnutí za demokratické Slovensko (HZDS) k SNP se dá říci, že je přinejmenším neujasněný.

Na loňských oslavách 49. výročí SNP předseda HZDS a bývalý premiér Vladimír Mečiar sice zvůli.

Ani postoj Křesťansko-demokratického hnutí (KDH) nebyl vždy jasný.

Představitelé KDH však neodmítli dialog a dnes považují povstání za významnou kapitolu slovenské historie.

Slovenské národní povstání politicky organizovala Slovenská národní rada, která vznikla již v září 1943 a zastupovala komunistický a občanský odboj.

Podle Vánoční dohody bylo cílem ozbrojeného boje obnovení Československé republiky.

Během SNP působily na území Slovenska vojenské mise Spojenců, a to sovětská, britská a americká.

Povstání zkomplikovalo plány nacistů na obranu Karpat, připoutalo nemalé nacistické jednotky na Slovensko, rozrušilo komunikační systém v týlu fronty a vytvořilo nebezpečné ohnisko nepokojů.

Ω Posilovalo také odpor v sousedních zemích.

O Slovensko O mělo význam také politicky: O zlikvidovalo chápání Slovenska na mezinárodní scéně jako německého vazala.
Úvahy padesát let od **SNP**

(3) **Povstání** a vůbec celou antifašistickou rezistenci považuji za významnou událost, která podala nefalšovaný obraz Slovence, řekl v srpnu 1994 v rozhovoru pro slovenský deník Pravda předseda Slovenského svazu protifašistických bojovníků generál Ján Husák.

(4) Také nejvýznamnější slovenská politika od prezidenta přes premiéra až po předsedu parlamentu považuji **Slovenské národní povstání** za pozitivní jev ve slovenských dějinách.

(5) Dnes je už známe, že historie **SNP** byla překroucena ve prospěch komunistické ideologie.

(6) Po listopadu 1989 sice tento černobilný pohled zmizel, ale vacuum využili převážně exiloví historici a spojili obhajobu Slovenského státu s odsouzením **SNP**.

(7) Oprostit **SNP** od nánosu lží a mýtů si klidla za cíl první vědeckou konference o **SNP**, která se uskutečnila letos v dubnu v Donovalech.

(8) S blížícím se 50. výročím **SNP** dochází na politické scéně i ve slovenském tisku k souboji dvou diametrně odlišných koncepcí výkladu **SNP**: na straně jedné se povstání označuje za **puč** proti vlastní státnosti a na straně druhé je **SNP** chápáno jako vědomé odmítnutí totality a fašismu.

(9) Podle názoru místopředsedkyně **Slovenské národní strany** (SNS) Evy Slavkovské přineslo **SNP** slovenskému národu jen to, že povstal se zbraní v ruce proti fašismu a nepřijal ho trpě jen jako český národ.

(10) O Vzalo mu však mnohem víc: vlastní samostatnost.

(11) Podle komentáře slovenského levicového listu Pravda povstání bránilo hodnoty, na kterých Evropa stála a stojí.

(12) SNS a další národně orientované strany (např. Křesťanská sociální unie, Společnost dr. Jozefa Tisa, Slovenská národní jednota) se přiklání k názoru, že sho o **puč** proti vlastní státnosti.

(13) Podle představitelů **Slovenské národní jednoty Češi** v **SNP** obrali Slováky o vlastní stát a na 40 let jim vznutil čechokomunistický režim.

(14) Strana demokratické levice po přehodnocení faktu, že vedoucí silou **SNP** byla komunistická strana, se jednoznačně staví za antifašistický a demokratický odkaz **povstání**.

(15) O postoji Hnutí za demokratické Slovensko (HZDS) k **SNP** se dá říci, že je přinejmenším neujasněný.

(16) Na loňských oslavách 49. výročí **SNP** předseda HZDS a bývalý premiér Vladimír Mečiar sice zvůlí.

(17) Ani postoj **Křesťanskodemokratického hnutí** (KDH) nebyl vždy jasný.

(19) Představitelé KDH však neodmítli dialog a dnes považují **povstání** za významnou kapitolu slovenské historie.

(20) **Slovenské národní povstání** politicky organizovala Slovenská národní rada, která vznikla již v září 1943 a zastupovala komunistický a občanský odboj.

(22) Podle Vánoční dohody bylo cílem ozbrojeného boje obnovení Československé republiky.

(24) Během **SNP** působily na území Slovenska vojenské mise Spojenců, a to sovětská, britská a americká.

(27) **Povstání** zkomplikovalo plány nacistů na obranu Karpát, připoutalo nemalé nacistické jednotky na Slovensko, rozrušilo komunikační systém v týlu fronty a vytvořilo nebezpečné ohnisko nepokoju.

(28) **O** Posilovalo také odpor v sousedních zemích.

(29) Pro Slovensko **O** mělo význam také politicky: **O** zlikvidovalo chápání Slovence na mezinárodní scéně jako německého vazala.
Frequency of Coreference Chains of Various Length

more than 2-element chains:

Two-element coreference chains:

- **German** – 69% (156 cases)
- **Czech** – 62% (343 cases)

TENDENCY TO SHORTER CHAINS IN GERMAN

adding lexical cohesion changes the situation not very substantially
Results so far...

- comparable results as for the properties of coreference chains
- properties of chains - genre- or register-dependent???
- thematic similarities or differences of the selected texts
Expressive means of coreference

- personal pronoun / zero
- local/temp (here, now)
- bare NP
- demonstrative modifier + NP (this fact)
- demonstrative head (this)
Antecedents of Demonstrative Heads

Czech

German

72% refer to abstract entities

39% refer to abstract entities

demonstrative head (this)
Abstract Anaphors referring to...

<table>
<thead>
<tr>
<th></th>
<th><strong>German</strong></th>
<th>%</th>
<th><strong>Czech</strong></th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>to clauses</td>
<td>5</td>
<td>7.35</td>
<td>13</td>
<td>20.63</td>
</tr>
<tr>
<td>to sentences</td>
<td>45</td>
<td>66.18</td>
<td>20</td>
<td>31.75</td>
</tr>
<tr>
<td>to bigger</td>
<td>18</td>
<td>26.47</td>
<td>30</td>
<td>47.62</td>
</tr>
<tr>
<td>segments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>68</td>
<td>100</td>
<td>63</td>
<td>100</td>
</tr>
</tbody>
</table>
Expressing means of Abstract Anaphors...

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
<th>Percentage</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>demonstrative head (<em>dies, dazu/ ten [this]</em>)</td>
<td>44</td>
<td>65%</td>
<td>28</td>
<td>44%</td>
</tr>
<tr>
<td>demonstrative modifier + NP (<em>diese Frage/ tato otázka [this question]</em>)</td>
<td>16</td>
<td>24%</td>
<td>17</td>
<td>27%</td>
</tr>
<tr>
<td>bare NP</td>
<td>0</td>
<td>0%</td>
<td>10</td>
<td>16%</td>
</tr>
<tr>
<td>temporal/local (<em>hier, da, nun/ tam, tady [here, there, now]</em>)</td>
<td>3</td>
<td>4%</td>
<td>4</td>
<td>6%</td>
</tr>
<tr>
<td>personal pronoun (<em>es [≈ it]</em>)/ zero anaphora</td>
<td>3</td>
<td>4%</td>
<td>2</td>
<td>3%</td>
</tr>
<tr>
<td>comparative</td>
<td>2</td>
<td>3%</td>
<td>2</td>
<td>3%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>68</td>
<td>100%</td>
<td>63</td>
<td>100%</td>
</tr>
</tbody>
</table>
Logically...

- demonstrative heads (this) – often to clauses

Cizinci podstatně přispěli k německému hospodářskému a kulturnímu vývoji, proč jejich počet naopak ve statistikách nezdůrazňovat a tím veřejně uznat jejich zásluhy o německou hospodářskou a politickou demokracii? [Foreigners have contributed significantly to the German economic and cultural development, so why not to emphasize their number in statistics, and to acknowledge their merit of the German economic and political democracy by this?]

- (dem.mod.+)NP (this/the decline) – often to sentences

Dnes se tento počet snížil na asi půl milionu, jenže důvodem <poklesu> je přede vším skutečnost, že ten, kdo není zaměstnán déle než rok, již podporu nedostane. [Today, that number dropped to about half a million, but the reason for the decline is the fact that anyone who is not employed for more than a year, gets no support anymore.]

- dem.mod.+general noun – often to larger segments
Gleichzeitig brauchen wir mindestens eine Verdoppelung des Wohlstands. Wenn wir die Armutseingebungen der Erde anschauen, weiß jeder sofort, dass dies das Mindeste an moralischer Herausforderung ist.

[At the same time, we need to double the current level of prosperity. One look at the poor regions throughout the world is enough to make anyone realize that this is the most urgent moral challenge we face.]
Es ist eine der wichtigsten Aufgaben des Staates, die Erhaltung des freien Wettbewerbs sicherzustellen. Versagt der Staat auf diesem Felde, dann ist es bald um die soziale Marktwirtschaft geschehen. [Protecting free competition is one of the state’s most important tasks. If the state fails in this area, the social market economy will soon be lost].
Nicht multinationale Unternehmen und Großbanken schaffen die meisten Arbeitsplätze, sondern der Mittelstand. Hier zeigt sich auch die Bedeutung der Berufsausbildung. [The majority of jobs are not created by multinational conglomerates, but by the "Mittelstand", the small and medium-sized enterprises. The importance of vocational training also becomes evident here.]
Jeder hatte sich darin auf irgendeine Weise definiert, nicht zuletzt auch die Vertreter des Kapitals in den marktwirtschaftlich organisierten Demokratien. Für sie war es wichtig, den Beweis zu erbringen, dass die demokratisch organisierte Marktwirtschaft auch sozial dem ineffizienten autoritären Kommunismus weit überlegen war. Es war ein voller Erfolg bis zu dem Punkt, als der OstWest-Konflikt zugunsten des Westens entschieden war.

[Everyone defined themselves in some way or another in accordance with this; not least those who represented investments in the market economies underlying the democratic countries. For such people, it was important to produce proof that democratically organized market economies were far superior to inefficient authoritarian communism in social as well as economic terms. It was a "roaring success" until the time came when the East-West conflict was resolved in favour of the West, at which point it was no longer necessary to furnish proof.]
Vždyť firma Bell začala jako první na světě provozovat městskou telefonní síť - a to právě v Québecu. – [Indeed, the Bell company started to operate the urban telephone network first in the world, and it was in Quebec.]
Poskytla mi smích už když jsem četl rukopis. Zdeňkovi však tehdy do smíchu nebylo. [It made me laugh when I was reading the manuscript. However, Zdenek, had no intention to laugh then.]
Czech: demonstrative head ⇝ clause

Cizinci podstatně přispěli k německému hospodářskému a kulturnímu vývoji, proč jejich počet naopak ve statistikách nezdůrazňovat a tím veřejně uznat jejich zásluhy o německou hospodářskou a politickou demokracii? [Foreigners have contributed significantly to the German economic and cultural development, so why not to emphasize their number in statistics, and to acknowledge their merit of the German economic and political democracy by this?]
Dnes se tento počet snížil na asi půl milionu. Důvodem poklesu je především skutečnost, že ten, kdo není zaměstnán déle než rok, již podporu nedostane. [Today, that number dropped to about half a million. However, the reason for the decline is the fact that anyone who is not employed for more than a year, gets no support anymore.]
V současnosti pracuje ve věznicích České republiky asi padesát procent z 9671 odsouzených. Poslední údaje uvádějí, že v půli srpna bylo ve vězení také 8467 obviněných, celkově tedy 18138 osob. Informoval o tom LN Eduard Vacek. [At present, about fifty percent of the 9,671 inmates work in the prisons of Czech Republic. The latest data show that in mid-August there were charged in jail 8467, a total of 18,138 persons. Eduard Vacek informed LN about this.]
Conclusion

- technical issues
- differences of typological character (absence of definiteness or pro-drops) *also* have influence on the properties of coreference chains
- the main differences in the length and number of chains (reflecting the topic structuring) are rather genre- or even domain-dependent
References


Navarretta C., Olsen S. (2009), The annotation of pronominal abstract anaphora in Danish texts and dialogues. DAD report 1. Centre for Language Technology, University of Copenhagen.


