Melanie Andresen & Heike Zinsmeister
melanie.andresen@uni-hamburg.de

STYLISTIC DIFFERENCES BETWEEN CLOSELY RELATED DISCIPLINES:
METATEXT IN LINGUISTICS AND LITERARY STUDIES
Introduction

Style in related disciplines

- German Studies = Linguistics + Literary Studies
- But: Scholars perceive the two disciplines as (very!) distinct.
- Students are expected to adapt to writing conventions of both disciplines.
- How do academic texts of Literary Studies and Linguistics differ stylistically?
Introduction

Data

- 60 PhD theses, 30 each from Linguistics and Literary Studies
- in German
- plain text
- citations, examples etc. extracted
Introduction

N-gram analysis

- data-driven approach
- n-gram = sequence of $n$ elements, e.g. words

<s> I will go hiking. </s>

unigrams:  <s>, I, will, go, hiking, </s>
bigrams:   <s> I, I will, will go, go hiking, hiking </s>
trigrams:  <s> I will, I will go, will go hiking, go hiking </s>
... ...
Introduction

N-gram analysis

- data-driven approach
- n-gram = sequence of $n$ elements, e.g. words
- Which n-grams differ in frequency between the corpora?
- measure for comparison: Log Likelihood (Dunning 1993)
N-gram analysis

Example: Trigrams that are more frequent in Linguistics:

<table>
<thead>
<tr>
<th>rank</th>
<th>LLR</th>
<th>German Trigram</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>261.46</td>
<td>in Bezug auf</td>
<td>with regard to</td>
</tr>
<tr>
<td>2</td>
<td>236.23</td>
<td>der vorliegenden Arbeit</td>
<td>the present text</td>
</tr>
<tr>
<td>3</td>
<td>233.30</td>
<td>&lt;s&gt; bei der</td>
<td>&lt;s&gt; At the</td>
</tr>
<tr>
<td>4</td>
<td>203.39</td>
<td>in der Regel</td>
<td>generally speaking</td>
</tr>
<tr>
<td>5</td>
<td>157.66</td>
<td>in der vorliegenden</td>
<td>in the present</td>
</tr>
<tr>
<td>6</td>
<td>156.06</td>
<td>Rahmen der vorliegenden</td>
<td>course of the present</td>
</tr>
<tr>
<td>7</td>
<td>155.72</td>
<td>Bezug auf die</td>
<td>regard to the</td>
</tr>
<tr>
<td>8</td>
<td>153.20</td>
<td>im Hinblick auf</td>
<td>with regard to</td>
</tr>
<tr>
<td>9</td>
<td>150.47</td>
<td>die Ergebnisse der</td>
<td>the results of</td>
</tr>
<tr>
<td>10</td>
<td>144.92</td>
<td>&lt;s&gt; bei den</td>
<td>&lt;s&gt; At the</td>
</tr>
</tbody>
</table>
**N-gram analysis**

N-grams related to metatext that are more frequent in Linguistics:

<table>
<thead>
<tr>
<th>size</th>
<th>rank</th>
<th>German n-gram</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10</td>
<td>der vorliegenden</td>
<td>the present</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>in Kapitel</td>
<td>in chapter</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>der vorliegenden Arbeit</td>
<td>the present text</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>im Folgenden</td>
<td>in the following</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>im Rahmen der vorliegenden</td>
<td>in the present</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>&lt;s&gt; zusammenfassend lässt sich</td>
<td>summarizing it can be [...]</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>&lt;s&gt; im folgenden werden die</td>
<td>in the following [...] will be</td>
</tr>
</tbody>
</table>
N-gram analysis

Text comments

- are more frequent in Linguistics than in Literary Studies.
- emerged as relevant from a data-driven analysis:
  - are an important difference between the disciplines.
  - are realized in a formulaic way.
Typological classification

- Fandrych and Graefen (2002): text comments
- Hyland (2005, p. 49): interactive metadiscourse > frame markers
- Ädel (2006, p. 20): metatext > text-oriented metadiscourse
Afros and Schryer (2009) compare promotional (meta)discourse:
- more pathos appeals in Literary Studies
- ‘transcending borders with literary genres’ (ibid., p. 63)

Hyland (2005, p. 57): ‘the more discursive “soft” fields employed more metadiscourse overall’
- but no clear tendency for frame markers
Are there significant differences between the use of metadiscourse of Literary Studies and Linguistics?

What additional insights can be gained by inspecting instances in detail?

selected examples:

- (im) Folgenden (in the following)
- zusammenfassend (summarizing)
im Folgenden (in the following)
Frequency of *im Folgenden*

Total: 569

- Linguistics: 235
- Literary Studies: 234

n = 30 each
Frequency of *im Folgenden*

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linguistics</td>
<td>18.97</td>
<td>20.17</td>
</tr>
<tr>
<td>Literary Studies</td>
<td>9.80</td>
<td>8.95</td>
</tr>
</tbody>
</table>

- **t-test**: $p = 0.028$
- **Wilcoxon rank-sum test**: $p = 0.080$
- **Cohen's $d$ (effect size)**: 0.59

→ big difference in means but high variance
Metatextual vs. intertextual

(e.g. Ädel 2006, p. 28)
Metatextual vs. intertextual

**Metatextual**

In the following, possible reasons for these differences will be addressed.

Im Folgenden wird auf mögliche Gründe für diese Unterschiede eingegangen.

(Lin_Dui_13)

**intertextual**

In the following d'Holbach explains [...]

Im Folgenden führt d’Holbach aus, dass [...] (Lit_Kob_25)
Metatextual vs. intertextual

\[ X^2: p < 0.001 \]
frequent in German text comments (in comparison to English, see Fandrych and Graefen 2002, pp. 28-33)

most frequent: sollen, which ‘indicates that the impetus for an action is external, i.e. an agent is required to carry out the will of another person or an institution.’ (ibid., p. 32)

→ hedging strategy
**Modal verbs**

*in the following without modal verb*

In the following, only Shakespeare's *The Tempest* will be picked out.

Im Folgenden wird lediglich Shakespeares *The Tempest* herausgegriffen. (Lit_Jen_19)

*in the following with modal verb*

This shall be determined more closely in the following.

Diese **soll** im Folgenden genauer eruiert werden. (Lit_Kie_23)
None vs. other: $X^2$: $p = 0.003$

sollen vs. other: $X^2$: $p = 0.002$

Sample, $n = 161$
### Most frequent verbs

<table>
<thead>
<tr>
<th>Verb</th>
<th>Translation</th>
<th>Ling.</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>eingehen auf</td>
<td>go into sth.</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>untersuchen</td>
<td>investigate</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>darstellen</td>
<td>depict</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>vorstellen</td>
<td>present</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>zeigen</td>
<td>show</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>erläutern</td>
<td>explain</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>betrachten</td>
<td>consider</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

→ hypothesis: more speech act verbs in Linguistics
Speech act verbs (SAV)

- hypothesis: more speech act verbs in Linguistics
- theoretical argument: clearer distinction between investigation and text in Linguistics
- definition of SAV: listed in either Harras et al. (2007) or GermaNet (Hamp and Feldweg 1997)
im Folgenden (in the following)

Speech act verbs (SAV)

\[ X^2: p = 0.001 \]

sample, n = 161
zusammenfassend (summarizing)
Frequency of *zusammenfassend*

- Total: 167
- Linguistics: 49
- Literary Studies: 49

*n = 30 each*
### Frequency of *zusammenfassend*

<table>
<thead>
<tr>
<th>Field</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linguistics</td>
<td>5.60</td>
<td>5.80</td>
</tr>
<tr>
<td>Literary Studies</td>
<td>1.94</td>
<td>8.95</td>
</tr>
</tbody>
</table>

- **t-test:** $p = 0.0011$
- **Wilcoxon rank sum test:** $p = 0.0002$
- **Cohen's $d$ (effect size):** 0.91

→ significant difference and large effect size
zusammenfassend (summarizing)

Metatextual vs. intertextual

Fisher's test: \( p = 0.029 \)

\( n = 216 \)


**Most frequent verbs**

<table>
<thead>
<tr>
<th>Verb</th>
<th>Translation</th>
<th>Ling.</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>festhalten</td>
<td>record</td>
<td>56</td>
<td>10</td>
</tr>
<tr>
<td>sagen</td>
<td>say</td>
<td>19</td>
<td>11</td>
</tr>
<tr>
<td>feststellen</td>
<td>determine</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>darstellen</td>
<td>depict</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td></td>
<td>154</td>
<td>41</td>
</tr>
</tbody>
</table>

→ very stable pattern, esp. in Linguistics
Conclusions
Conclusions

Summary

- Candidates for metatext can be identified automatically, but should be inspected in detail.
- Linguistics make more use of metatext than Literary Studies.
- Literary Studies use more modal verbs in metatext than Linguistics.
- Linguistics uses more speech act verbs (in metatext).
Conclusions

Possible Explanations

- different aesthetic demands (see also Afros and Schryer 2009; and results related to article titles by Haggan 2004)

- less analytical research process in Literary Studies
  - less subdividable into distinct steps
  - less universal research steps

- stronger influence of the English writing culture on Linguistics (see discussion in Fandrych and Graefen 2002)
Thank you for your attention!


Melanie Andresen & Heike Zinsmeister
melanie.andresen@uni-hamburg.de

STYLISTIC DIFFERENCES BETWEEN CLOSELY RELATED DISCIPLINES: METATEXT IN LINGUISTICS AND LITERARY STUDIES