The Syntax of Adverbials in Early New High German*
Evidence from Syntactically Annotated Corpora

Ulrike Demske (Universität Potsdam)
and
Marianna Patak (Universität Potsdam)

SLE 2016, 1st September 2016, Naples

Contents

1 Introduction 2

2 The Case in Modern German 2

3 Adverbials in Early New High German 4
   3.1 Quantitative Evaluation ................................. 4
      3.1.1 Data ........................................... 4
      3.1.2 Databases ........................................ 7
   3.2 Qualitative Evaluation ................................. 8
      3.2.1 Data ........................................... 8
      3.2.2 Databases ........................................ 11

4 Conclusion 11

*The project is funded by the German Research Foundation (DFG). We would like to thank our student assistants Elisabeth Hoffmann, Malika Reetz, Simon Titze and Henrik Willun for their contribution to our research work on adverbials in Early New High German.
1 Introduction

(1) Wer weiß, ob vielleicht auch vonn wegn vnser sunde Luthers vorgifte lere durch schweigen der gelarten und vorseumlgkeit poisoned doctrine by silence of the scholars and failure der prelaten ßo tieff gewortzelt hat. of the prelates so deeply rooted

‘Who knows whether Luthers poisoned doctrine has been so deeply rooted by the silence of the scholars and by the failure of the prelates, maybe also because of our sins.’

(PENHG, 2016, Martinus Luther: Wy es eyn man sey..., 1522, s. 18)

• Example (1): 6 adverbials1, e.g. vielleicht ‘maybe’ = sentence adverbial, vonn wegn vnser sunde ‘because of our sin’ = causal adverbial, ßo tieff ‘so deeply’ = manner adverbial

• Ebert (1986) worked on syntax of Early New High German

• Work on adverbials at the left periphery in earlier stages of German has been done by Axel (2002) and Donhauser & Petrova (2009)

• Syntax of adverbials in Early New High German?

2 The Case in Modern German

Adverbials in Modern German according to Frey (2003) and Pittner (2004):

• Assumption: syntactic base position of adverbials is reflected by their c-command relations to other adverbials and verbal arguments.

• Regarding their base position adverbials are subsumed in five classes: sentence adverbials > frame adverbials > event-related adverbials > event-internal adverbials > process adverbials

• Distribution of adverbial classes is regulated by syntax.

• Ordering of adverbial types within one adverbial class is regulated by semantics.

1Adverbials are indicated in bold.
Adverbial classes (Pittner, 2004, p. 282):

1. Sentence adverbials: they c-command the finite verb and the base positions of all adverbial adjuncts and verbal arguments.

2. Frame adverbials: they c-command the base positions of event-related adverbials.

3. Event-related adverbials: they c-command the base position of the highest argument and the base positions of event-internal adverbials.

4. Event-internal adverbials: they are minimally c-commanded by the argument they are related to, i.e. no other argument can intervene.

5. Process adverbials: they c-command minimally the verb or ‘verbal phrase’.

(2) a. \[CP \text{ weil } \text{vielleicht}_{SADV} \text{ in keiner Stadt}_{FADV} \text{ heute}_{ERADV} \]
   because maybe in no city today
   \[alle \text{ gerne}_ {EIADV} \text{ (nicht) schnell}_{PADV} \] \[VP \text{ schwimmen gehen} \]
   everybody willingly not fast swim go
   \[wollen] \]
   want
   ‘because maybe today in no city everybody willingly wants to go swimming as fast as possible’

b. \# \[CP \text{ weil } \text{in keiner Stadt}_{FADV} \text{ vielleicht}_{SADV} \text{ heute}_{ERADV} \]
   because in no city maybe today
   \[alle \text{ gerne}_ {EIADV} \text{ (nicht) schnell}_{PADV} \] \[VP \text{ schwimmen gehen} \]
   everybody willingly not fast swim go
   \[wollen] \]
   want
   ‘because maybe today in no city everybody willingly wants to go swimming as fast as possible’

c. \# \[CP \text{ weil } \text{vielleicht}_{SADV} \text{ in keiner Stadt}_{FADV} \text{ heute}_{ERADV} \]
   because maybe in no city today
   \[alle \text{ gerne}_ {EIADV} \text{ schnell}_{PADV} \text{ (nicht)} \] \[VP \text{ schwimmen gehen} \]
   everybody willingly fast not swim go
   \[wollen] \]
   want
   ‘because maybe today in no city everybody willingly wants to go swimming as fast as possible’
3 Adverbials in Early New High German

- Research hypothesis: adverbials in Early New High German show same c-command relations as adverbials in Modern German.
- Test: sentence adverbials and process adverbials (more precisely manner adverbials) - high vs. low position - in the middle field of subordinate and relative clauses\(^2\)
- Databases
  - Two syntactically annotated treebanks (Logačev et al., 2014; Demske; 2007):
    * Mercurius Treebank (two newspapers, 1597 and 1667; 160.000 words) (Rosén et al., 2012)
    * Potsdam Treebank of Early New High German (1350-1650; 320.000 word forms as of August 2016) (PENHG, 2016)
  - Two text corpora:
    * GerManC (1650-1800; 676.508 word forms) (DeReKo, 2016; Scheible et al., 2011)
    * 'Deutsches Textarchiv' (for data from the 17th century; > 150 texts) (DTA, 2016)

3.1 Quantitative Evaluation

3.1.1 Data

Assumptions regarding process adverbials:
- Process adverbials are adjacent to the verbal phrase.
- Sentence negation has to be above process adverbials.
- Base position of process adverbials is relatively low.
- \([CP \text{ subjunction/relative pronoun} > \text{syntactic elements} > \text{sentence negation} > \text{process adverbial} > \text{verbal phrase}]\)

\(^2\)In subordinate/relative clauses the middle field is between the subjunction/relative pronoun and the verbal phrase.
Results regarding process adverbials in ENHG:

<table>
<thead>
<tr>
<th></th>
<th>All Corpora</th>
<th>Without GerManC</th>
<th>Treebanks</th>
<th>Text Corpora</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjacent</td>
<td>81%</td>
<td>81%</td>
<td>82%</td>
<td>79%</td>
</tr>
<tr>
<td>Non-Adjacent</td>
<td>19%</td>
<td>19%</td>
<td>18%</td>
<td>21%</td>
</tr>
<tr>
<td>NEG-PADV</td>
<td>93%</td>
<td>100%</td>
<td>100%</td>
<td>88%</td>
</tr>
<tr>
<td>PADV-NEG</td>
<td>7%</td>
<td>0%</td>
<td>0%</td>
<td>13%</td>
</tr>
</tbody>
</table>

Table 1: Process adverbials in ENHG: adjacency and sentence negation

- Adjacency: in 81% of the data manner adverbials are adjacent to the verbal phrase.
- Sentence negation: in 93% of the data sentence negation is above manner adverbials.

(3) a. 
\[ dz \text{ er mich wol } [VP \text{ verwaren solte}] \]
that he me good detain should
‘that he should detain me well’

(PENHG, 2016, Staden: Wahrhaftig Historia, 1557, s. 878)

b. 
\[ daß alle seine Freunde sich \text{ schleunig} \text{ wiederum bey jhm} [VP \text{ versamleten}] \]
that all his friends themselves immediately again at him come.together
‘that all his friends immediately came together with him again’

(Rosén et al., 2012, Mercurius, #3573 (s2497))

c. 
\[ daz \text{ ir nicht wol } [VP \text{ geslaffen mögl}] \]
that you not good slept could
‘that you did not sleep well’

(PENHG, 2016, Kottanerin: Denkwürdigkeiten, ca. 1442, s. 161)

d. 
\[ daß dieselbigen ohne \text{ sonderliche Laub} \text{ von} \text{ Bergwerck}en \]
that they without special permission from mines
\[ nicht [VP \text{ reisen sol}] \]
not travel should
‘that they should not travel from the mines without special permission’

(DeReKo, 2016, NEWS_P1_OMD_1666_leipzig2)
Table 2: Process adverbials in ENHG: c-command relations

- In 81% of the data process adverbials in ENHG show same c-command relations as process adverbials in Modern German.
- Regarding process adverbials we might accept our research hypothesis.

Assumptions regarding sentence adverbials:
- Base position of sentence adverbials is high in the clause.
- No other syntactic element may occur above sentence adverbials.
- \([CP \text{ subjunction/relative pronoun} > \text{(topic)} > \text{sentence adverbial} > \text{(sentence negation)} > \text{syntactic elements} > \text{verbal phrase}]\]

Results regarding sentence adverbials (SADV) in Early New High German:

<table>
<thead>
<tr>
<th></th>
<th>All Corpora</th>
<th>Without GerManC</th>
<th>Treebanks</th>
<th>Text Corpora</th>
</tr>
</thead>
<tbody>
<tr>
<td>SADV-element</td>
<td>17%</td>
<td>16%</td>
<td>18%</td>
<td>16%</td>
</tr>
<tr>
<td>element-SADV</td>
<td>83%</td>
<td>84%</td>
<td>82%</td>
<td>84%</td>
</tr>
</tbody>
</table>

Table 3: Sentence adverbials in ENHG: highest position

- In 17% of the data sentence adverbials are the highest element in the clause.
- But: in 83% of the data one (or more) syntactic element appears in front of a sentence adverbial.

(4) a. *dz warlich vil kinder Abrahe ausz den stemen gehauen vnder*  
that truly many children A. of the tribe made among  
den Heyden seind  
the Gentiles are  
‘that truly many children made of Abrahams’ tribe are among the Gentiles’  
(PENHG, 2016, Franck: Chronik, 1534, s. 90)

b. *das zween winter zwen Sommer in einem jar gewißlich da*  
that two winter two summer in one year certainly there  
vorhanden sein.  
present are  
‘that there are certainly two winter and two summer present in one year’
In ENHG sentence adverbials tend not to occupy the highest position.

Research hypothesis? Continues under section 3.2.1

3.1.2 Databases

Assumption: historical treebanks supply search tools which make it particularly easy to locate constituents such as adverbials (Dipper, 2015).

Research hypothesis: in the two treebanks of ENHG we find most data with process and sentence adverbials.

Test: compare the two treebanks to the two text corpora according to their amount of data.

Problematic: unknown size of selected texts of the text corpora

<table>
<thead>
<tr>
<th>Word Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercurius</td>
</tr>
<tr>
<td>PENHG</td>
</tr>
<tr>
<td>GerManC</td>
</tr>
<tr>
<td>DTA</td>
</tr>
</tbody>
</table>

Table 4: Treebanks vs. Text Corpora: size of selected texts

GerManC: size of individual texts is not recorded in word forms.

DTA: size of individual texts is not recorded in word forms.

But: DTA and Potsdam Treebank of Early New High German have the same text ‘Staden: Wahrhaftig Historia, 1557’.

Solution: compare DTA to Potsdam Treebank of Early New High German regarding the ‘Staden’-Text.

---

The ‘Staden’-Text is longer in DTA than in Potsdam Treebank of Early New High German. All hits of DTA which appear obviously after that part the ‘Staden’-text of Potsdam Treebank of Early New High German ends were excluded from the calculation.
Results regarding the Text ‘Staden: Wahrhaftig Historia, 1557’

<table>
<thead>
<tr>
<th></th>
<th>Word forms</th>
<th>Data</th>
<th>Data/WF</th>
</tr>
</thead>
<tbody>
<tr>
<td>PENHG</td>
<td>20,280</td>
<td>13</td>
<td>0,001</td>
</tr>
<tr>
<td>DTA</td>
<td>20,280</td>
<td>17</td>
<td>0,001</td>
</tr>
</tbody>
</table>

Table 5: Potsdam Treebank of Early New High German vs. DTA

- No tendency
- Corpora to small (20,280 word forms)?

3.2 Qualitative Evaluation

3.2.1 Data

- Process adverbials in ENHG tend to maintain same c-command relations as process adverbials in Modern German.
- But sentence adverbials?
  - In 17% of the data sentence adverbials occupy the highest position in the clause.
  - In 83% of the data one (or more) syntactic element appears in front of a sentence adverbial.

Middle field in Modern German according to Frey (2004):

- Discourse configurational with respect to topics
- Topics occur only in front of sentence adverbials.
- Elements which occur in front of other adverbial classes than sentence adverbials are not topics\(^4\).

Deeper insight into the ‘element-SADV’ data of sentence adverbials:

- In 83% of the data syntactic elements occur in front of sentence adverbials.
- Are these elements topical phrases?
- Test: no other adverbial class may intervene between the syntactic element and the sentence adverbial\(^5\).

\(^5\)Unfortunately non-topical phrases like quantificational phrases does not pop up in our data.
Test:

- \([\text{CP} \text{ subjunction/relative pronoun} > \text{topic} > \!\text{no other adverbial class!} > \text{sentence adverbial} > \ldots]\)

Results regarding topics:

<table>
<thead>
<tr>
<th></th>
<th>All Corpora</th>
<th>Without GerManC</th>
<th>Treebanks</th>
<th>Text Corpora</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element-SADV</td>
<td>79%</td>
<td>76%</td>
<td>78%</td>
<td>76%</td>
</tr>
<tr>
<td>Element-ADV-SADV</td>
<td>21%</td>
<td>24%</td>
<td>22%</td>
<td>19%</td>
</tr>
</tbody>
</table>

Table 6: Elements above sentence adverbials

- In 79% of the data syntactic elements are followed directly by sentence adverbials.
- In 21% of the data an adverbial class other than sentence adverbial intervenes between the sentence adverbial and the syntactic element.
- What is the syntactic function of the elements which occur in front of sentence adverbials?

Results regarding syntactic function:

<table>
<thead>
<tr>
<th></th>
<th>All Corpora</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>57%</td>
</tr>
<tr>
<td>Object</td>
<td>14%</td>
</tr>
<tr>
<td>Adjunct</td>
<td>29%</td>
</tr>
</tbody>
</table>

Table 7: Syntactic function of the elements above sentence adverbials

- In 57% of the data subjects appear above sentence adverbials.
- In 14% of the data objects appear above sentence adverbials.
- 71% support: syntactic elements above sentence adverbials are topics.
- Research hypothesis regarding c-command relations of sentence adverbials?

(5) a. \(\text{welche Sie sonder Zweifel zur Conservation der allgemeinen Ruhe}\)
which she without doubts for conservation the general calmness
\(\text{hat}\)
has
\(\text{‘which without doubts she has for conservation of the general calmness’}\)

(Rosén et al., 2012, Mercurius, #3880 (s2804))
b. \textit{das in \textit{villeicht} die Heyden \textit{vmb bringen moechten}}
that him maybe \textit{the Gentiles kill want}
‘that maybe the Gentiles want to kill him’
\hfill (PENHG, 2016, Fierrebras, 1532, s. 482)

c. \textit{das zween winter zwen Sommer in einem jar \textit{gewislich} da}
that two \textit{winter two summer in one year certainly there}
\textit{vorhanden sein}.
present are
‘that two winter and two summer certainly are present there in one year’
\hfill (DTA, 2016, Staden: Wahrhaftig Historia, 1557, p. 5)

\begin{tabular}{|c|c|}
\hline
All Corpora & \\
Yes & 83\% \\
No & 17\% \\
\hline
\end{tabular}

Table 8: Sentence adverbials in ENHG: c-command relations

- In 83\% of the data sentence adverbials in ENHG maintain same c-command relations as sentence adverbials in Modern German.
- Regarding sentence adverbials we might accept our research hypothesis.

Further consequences: Middle field in Early New High German

- According to Speyer (2011):
  - Middle field in ENHG is structurally determined.
  - No continuity from Indoeuropean to Modern German
    * Indoeuropean: information structure determines syntax.
    * ENHG: grammar determines syntax.
    * Modern German: information structure determines syntax.
- We argue against Speyer (2011):
  - Adverbials in ENHG maintain same c-command relations as adverbials in Modern German.
  - Process adverbials tend to occur adjacent to the verbal phrase.
  - Sentence adverbials tend to occupy the highest position in ENHG.
  - Historical evidence for Cinque’s (1999) proposed hierarchy of functional projections
– If an element is above a sentence adverbial it is likely that it is a topical phrase.
– Middle field of ENHG is determined by information structure.
– Middle field of ENHG is as discourse configurational as the middle field of Modern German.

3.2.2 Databases

Historical treebanks have advantages for linguistic research (Haug, 2015). But do they have any advantages according to our topic?

• Contra: Quantitative evaluation (corpus too small?)
• Pro: it is easy to specify syntactic categories like clause type and constituents
• But: semantics is not annotated
• Therefore: too many adverbials which has to be manually checked

4 Conclusion

Data:

• A very first insight into (relative) syntactic ordering of adverbials in the middle field of Early New High German was given.
• In Early New High German process and sentence adverbials maintain same c-command relations as in Modern German.
• In Early New High German the middle field is determined by information structure.

Databases:

• Treebanks have advantages compared to text corpora.
• They are obviously on the side of syntax.
• Research on syntax-semantics interface still needs to be half-manually made.

References


Corpora


