Multiple Fronting vs. VP Fronting in German

4th Conference on Quantitative Investigations in Theoretical Linguistics
March 28–31, Humboldt Universität Berlin

Felix Bildhauer and Philippa Cook
Freie Universität Berlin

March 29, 2011
Outline

Background

Multiple fronting vs. VP fronting

Corpus study
German as a V2 language

- Exactly one constituent before the finite verb in declarative main clauses
- Can be a verb phrase

(1) [Ganz ins Privatleben zurückziehen] will Höfler sich nicht fully into.the private.life withdraw wants Höfler self not ‘Höfler doesn’t want to completely withdraw from public’ M08/JAN.06537
Multiple fronting

- Sometimes, more than one constituent before the finite verb

(2) \[[\text{Dem Saft}] [\text{eine kräftigere Farbe}] \text{geben}\ \text{Blutorangen}.

to.the juice a more.vivid colour give blood.oranges

‘What gives the juice a more vivid colour is blood oranges.’ R99/JAN.01605
Multiple fronting II

- Can be analyzed as *apparent* multiple fronting
- Involves VP headed by a verb trace
- But: licensing conditions still unclear

![Diagram of VP fronting and (Apparent) Multiple fronting](image)
Multiple fronting III

MF is possibly an instance of pars-pro-toto fronting (Fanselow 2004):

- Entire VP has discourse function (focus/topic)
- Only part of it is fronted

(5)  
  a. Was ist mit dem Buch passiert?  
     ‘What happened to the book?’
  b. [Meiner Freundin] hab ich’s geschenkt
     my.dat girlfriend have I it given  
     ‘I gave it to my girlfriend as a present’
  c. [Meiner Freundin geschenkt] hab ich’s.
     my.dat girlfriend given have I it
     ‘I gave it to my girlfriend as a present’
Contextual licensing of multiple fronting

Two predominant patterns:

1. *Presentation*:
   The verb’s “designated topic” is introduced into the discourse; typically the subject.

   \[ \text{[Dem Saft] [eine kräftigere Farbe] geben Blutorangen.} \]
   to.the juice a more.vivid colour give blood.oranges
   ‘What gives the juice a more vivid colour is blood oranges.’ R99/JAN.01605

   ➤ In line with findings from other SFB-projects on subject focus (Fiedler et al. 2009)
Contextual licensing of multiple fronting II

2. *Propositional Assessment*:
Prominent adverbial spells out the degree to which the main predication holds of a specific discourse referent, realized in the middlefield; typically the subject.

[Nach Brüssel] [zum Demonstrieren] ist Gerd Knecht *nicht* gefahren
to Brussels to demonstrate is G. K. not gone
‘G. K. did not go to Brussels for the demo’ M99/FEB.12802
Multiple fronting vs. VP fronting

- Sometimes, alternation between “exceptional” multiple fronting and “canonical” VP-fronting
Multiple fronting vs. VP fronting

- Sometimes, alternation between “exceptional” multiple fronting and “canonical” VP-fronting

(6) a. [Positiv auf das Ergebnis ausgewirkt] habe sich auch die 1998 positiv on the result affected has self also the 1998 eingeführte teilweise Globalbudgetierung, introduced partial global.budgeting ‘The partially global budgeting which was introduced in 1998 also had a positive effect on the result.’ P98/OKT.39922
Multiple fronting vs. VP fronting

- Sometimes, alternation between “exceptional” multiple fronting and “canonical” VP-fronting

(6)  

   ‘The partially global budgeting which was introduced in 1998 also had a positive effect on the result.’ P98/OKT.39922

   b. [Positiv] [auf das Ergebnis] habe sich der starke Dollar ausgewirkt positiv auf das Ergebnis haben selbst den starke Dollar ausgewirkt
   ‘The strong Dollar had a positive effect on the result.’ N98/NOV.43897
MF-VPF Alternation II

(7) a. [Ganz vom Tisch wischen] kann man das Thema freilich nicht. fully from.the table wipe can one the issue of.course not ‘Of course, one can’t brush the issue aside completely’ RHZ00/MAI.05627

b. [Ganz] [vom Tisch] will er es aber nicht wischen fully from.the table will he it however not wipe ‘He doesn’t want to brush it aside completely though’ NON08/NOV.01483
MF-VPF Alternation III

(8) a. [Gänzlich auf Derbys verzichten] will Nix aber auch künftig nicht fully on Derbys renounce will Nix but also in the future not ‘Nix doesn’t want to completely give up Derbys in the future either though’ RHZ99/MAI.16980

b. [Gänzlich] [auf seine Erfahrung] will “Py” aber nicht verzichten. fully on his experience will “Py” however not renounce “’Py” doesn’t want to completely do without his experience though’ M03/MAR.14862
MF-VPF Alternation IV

(9) a. [Ganz ins Privatleben zurückziehen] will Höfler sich nicht fully into.the private.life withdraw wants Höfler self not ‘Höfler doesn’t want to completely withdraw from public’ M08/JAN.06537

b. [Ganz] [ins Privatleben] wird sich Karl Popp (…) freilich fully into.the private.life will self Karl Popp of.course nicht zurückziehen. not withdraw not withdraw ‘Of course, Karl Popp doesn’t want to completely withdraw from public’ N00/DEZ.59676
Explaining the MF-VPF alternation

Can we isolate factors that determine the MF-VPF alternation?

- Information-structural factors?
- Other factors?
Explaining the MF-VPF alternation

Can we isolate factors that determine the MF-VPF alternation?

- Information-structural factors?
- Other factors?

Examine via corpus study
Corpus studies of MF constructions

Quantitative corpus studies of MF constructions are generally difficult:

- Large amount of data whose status is not entirely clear (Instance of MF or not?)
Corpus studies of MF constructions

Quantitative corpus studies of MF constructions are generally difficult:

- Large amount of data whose status is not entirely clear (Instance of MF or not?)
- Rare phenomenon: Requires very large corpora for meaningful statistical analysis (we are using DeReKo, 2.5 billion words)
Corpus studies of MF constructions

Quantitative corpus studies of MF constructions are generally difficult:

- Large amount of data whose status is not entirely clear (Instance of MF or not?)
- Rare phenomenon: Requires very large corpora for meaningful statistical analysis (we are using DeReKo, 2.5 billion words)
- Working with non-tagged, non-parsed corpora: No general way of sampling instances of MF
Corpus studies of MF constructions

Quantitative corpus studies of MF constructions are generally difficult:

- Large amount of data whose status is not entirely clear (Instance of MF or not?)
- Rare phenomenon: Requires very large corpora for meaningful statistical analysis (we are using DeReKo, 2.5 billion words)
- Working with non-tagged, non-parsed corpora: No general way of sampling instances of MF
- Workaround: Search by relatively concrete patterns (e.g. lexically provide one or both of the fronted constituents)
Corpus study: Data

\[ \begin{array}{cc}
1 & 2 \\
\end{array} \nonfin \quad \begin{array}{c}
\text{V} \\
\text{fin} \\
\text{......} \\
\end{array} \quad = \text{VPF} \\
\begin{array}{cc}
1 & 2 \\
\end{array} \quad \begin{array}{c}
\text{V} \\
\text{fin} \\
\text{......} \nonfin \\
\end{array} \quad = \text{MF} \\

- \nonfin \text{ is a modal/auxiliary verb: } \{ \text{haben, sein, wollen, sollen, können, dürfen, müssen, werden, mögen} \} \\
- \begin{array}{c}
2 \\
\end{array} \text{ is an NP or PP} \\
- \begin{array}{c}
1 \\
\end{array} \text{ is an Adjective or Adverb}
Synsets: Selected from Germanet semantic network

\[
\begin{align*}
\text{Set} & \\
\text{GANZ} & \text{POSITIV} \\
\text{\textquote{completely}} & \text{\textquote{positively}} \\
\text{ganz} & \text{positiv} \\
gänzlich & \text{günstig} \\
ganz und gar & \text{vorteilhaft} \\
völlig & \text{glücklich} \\
komplett & \text{negativ} \\
vollständig & \text{ungünstig} \\
\end{align*}
\]
Example

(10) VP fronting
   a. \([\text{Adv PP } V_{\text{nonfin}}] \ V_{\text{fin}} (XP_3 \ldots XP_n)\)
   b. \([\text{Ganz aus den Augen verloren}] \ \text{habe er die Kunst ohnehin nie.}\)
      fully out the eyes lost has he the art anyway never
      ‘He has never fully lost sight of art anyway.’ RHZ08/NOV.19605

(11) Multiple fronting
   a. \([\text{Adv} [\text{PP}] \ V_{\text{fin}} (XP_3 \ldots XP_n) \ V_{\text{nonfin}}] \)
   b. \([\text{Ganz} [\text{aus dem Auge}] \ \text{will man das Thema nicht verlieren}\)
      Fully out the eye wants one the subject not lose
      ‘They don’t want to lose sight of the issue completely.’
      RHZ06/MAI.21197
Procedure: Data extraction

Extract all sentences from the corpus that…
Procedure: Data extraction

Extract all sentences from the corpus that...

- start with an item from the GANZ or POSITIV set
- contain one of the listed modals/auxiliaries
Procedure: Data extraction

Extract all sentences from the corpus that...

- start with an item from the GANZ or POSITIV set
- contain one of the listed modals/auxiliaries
- Manually scan results for target structure
- Discard passives, extraposed relative clauses, copula constructions
Procedure: Data extraction

Extract all sentences from the corpus that...

- start with an item from the GANZ or POSITIV set
- contain one of the listed modals/auxiliaries
- Manually scan results for target structure
- Discard passives, extraposed relative clauses, copula constructions

<table>
<thead>
<tr>
<th>Set</th>
<th>ganz</th>
<th>positiv</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPF</td>
<td>114</td>
<td>93</td>
<td>207</td>
</tr>
<tr>
<td>MF</td>
<td>83</td>
<td>102</td>
<td>185</td>
</tr>
<tr>
<td>TOTAL</td>
<td>197</td>
<td>195</td>
<td>382</td>
</tr>
</tbody>
</table>

Manually annotate the remaining 382 tokens.
Factors related to information structure

**Givenness** and **Focus**

- Relatively “low-level” concepts
- Can be operationalised (more or less) straightforwardly

Annotation based on guidelines by Götze et al. (2007)
Factors related to information structure II

Givenness of the second fronted element (NP or PP):
  ▶ Levels: given, non-given, non-referential

Givenness of the subject (in the middlefield):
  ▶ Levels: given, non-given, non-referential

Main focus (mostly in the middlefield):
  ▶ subject, quantifying-adverbal, other-middlefield, other-prefield
Factors *not* related to information structure

- Length/heaviness of fronted constituents

(12) a. [Ganz auf eine Beschränkung im Herzen Nürnbergs verzichten] fully on a restriction in the heart Nuremberg renounce

will die Stadt aber nicht wants the city but not

‘The city doesn’t want to give up on a regulation for the heart of Nuremberg.’

b. [Ganz] [auf Fleisch] mögen viele Gourmets nicht verzichten.

fully on meat want many gourmets not renounce

‘Many gourmets don’t want to give up meat completely.’

- Measured as number of syllables
Potential factors *not* related to information structure II

- Length/heaviness of the middlefield subject


b. [Positiv] [auf das Ergebnis] *habe* sich der starke Dollar ausgewirkt

- Measured as number of syllables
Model

- Logistic regression
- Multiple Fronting coded as “1”
- Stepwise model selection (optimizing AIC)
Model

- Logistic regression
- Multiple Fronting coded as “1”
- Stepwise model selection (optimizing AIC)

Factors contained in optimal model:

1. Set
2. Givenness of second fronted element
3. Focus
4. Length (i.e., difference in length)
5. Interaction: Set * Foc
6. Interaction: Set * Givenness of second fronted element
<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>S. E.</th>
<th>Wald Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.03111</td>
<td>0.25987</td>
<td>-0.12</td>
<td>0.9047</td>
</tr>
<tr>
<td>Set=positiv</td>
<td>1.06440</td>
<td>0.40895</td>
<td>2.60</td>
<td><strong>0.0092</strong></td>
</tr>
<tr>
<td>XP2.giv=giv</td>
<td>-0.16177</td>
<td>0.36799</td>
<td>-0.44</td>
<td>0.6602</td>
</tr>
<tr>
<td>XP2.giv=non.ref</td>
<td>-0.45948</td>
<td>0.39542</td>
<td>-1.16</td>
<td>0.2452</td>
</tr>
<tr>
<td>Foc=qadv</td>
<td>-0.33804</td>
<td>0.32916</td>
<td>-1.03</td>
<td>0.3044</td>
</tr>
<tr>
<td>Foc=other</td>
<td>-0.19246</td>
<td>0.81288</td>
<td>-0.24</td>
<td>0.8128</td>
</tr>
<tr>
<td>Foc=vf</td>
<td>2.25827</td>
<td>1.08922</td>
<td>2.07</td>
<td><strong>0.0381</strong></td>
</tr>
<tr>
<td>Diff</td>
<td>0.03130</td>
<td>0.01376</td>
<td>2.27</td>
<td><strong>0.0229</strong></td>
</tr>
<tr>
<td>Set=positiv*Foc=qadv</td>
<td>-0.42943</td>
<td>1.01120</td>
<td>-0.42</td>
<td>0.6711</td>
</tr>
<tr>
<td>Set=positiv*Foc=other</td>
<td>-2.07765</td>
<td>1.44120</td>
<td>-1.44</td>
<td>0.1494</td>
</tr>
<tr>
<td>Set=positiv*Foc=vf</td>
<td>-10.40430</td>
<td>36.76462</td>
<td>-0.28</td>
<td>0.7772</td>
</tr>
<tr>
<td>Set=positiv*XP2.giv=giv</td>
<td>0.45825</td>
<td>0.54270</td>
<td>0.84</td>
<td>0.3984</td>
</tr>
<tr>
<td>Set=positiv*XP2.giv=non.ref</td>
<td>-4.00293</td>
<td>1.13074</td>
<td>-3.54</td>
<td><strong>0.0004</strong></td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>S. E.</th>
<th>Wald Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set=positiv</td>
<td>1.06440</td>
<td>0.40895</td>
<td>2.60</td>
<td><strong>0.0092</strong></td>
</tr>
<tr>
<td>Foc=vf</td>
<td>2.25827</td>
<td>1.08922</td>
<td>2.07</td>
<td><strong>0.0381</strong></td>
</tr>
<tr>
<td>Diff</td>
<td>0.03130</td>
<td>0.01376</td>
<td>2.27</td>
<td><strong>0.0229</strong></td>
</tr>
<tr>
<td>Set=positiv*XP2.giv=non.ref</td>
<td>-4.00293</td>
<td>1.13074</td>
<td>-3.54</td>
<td><strong>0.0004</strong></td>
</tr>
</tbody>
</table>
Main effect: Focus

Main focus in prefield: MVB

(13) a. Bereits im Gründungsjahr präsentierte sich das neugegründete Streichorchester in Dortelweil das erste Mal der Öffentlichkeit...

b. [Ganz] [der Barockmusik] hatte man sich in diesen Jahren entirely to the baroque music had one himself in these years verschrieben.
dedicated
‘During these years, they dedicated themselves entirely to baroque music.’
## Results

<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>S. E.</th>
<th>Wald Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set=positiv</td>
<td>1.06440</td>
<td>0.40895</td>
<td>2.60</td>
<td>0.0092</td>
</tr>
<tr>
<td>Foc=vf</td>
<td>2.25827</td>
<td>1.08922</td>
<td>2.07</td>
<td>0.0381</td>
</tr>
<tr>
<td>Diff</td>
<td>0.03130</td>
<td>0.01376</td>
<td>2.27</td>
<td>0.0229</td>
</tr>
<tr>
<td>Set=positiv*XP2.giv=non.ref</td>
<td>-4.00293</td>
<td>1.13074</td>
<td>-3.54</td>
<td>0.0004</td>
</tr>
</tbody>
</table>
Main effect: Length

Long fronted XP, short subject: MF (Difference = 7)

(14) [Ganz] [auf die Tätigkeit mit Büchern] wollte sie jedoch nicht verzichten
Main effect: Length

Long fronted XP, short subject: MF (Difference = 7)

(14) [Ganz] [auf die Tätigkeit mit Büchern] wollte sie jedoch nicht verzichten

Short fronted XP, long subject: VPF (Difference = -21)

(15) [Völlig aus den Schlagzeilen verschwunden] ist hingegen die so bitter nötige Diskussion über Deponien und Verbrennungsanlagen für Sondermüll.
## Results

<table>
<thead>
<tr>
<th>Set</th>
<th>Coef</th>
<th>S. E.</th>
<th>Wald Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>positiv</td>
<td>1.06440</td>
<td>0.40895</td>
<td>2.60</td>
<td>0.0092</td>
</tr>
<tr>
<td>Foc=vf</td>
<td>2.25827</td>
<td>1.08922</td>
<td>2.07</td>
<td>0.0381</td>
</tr>
<tr>
<td>Diff</td>
<td>0.03130</td>
<td>0.01376</td>
<td>2.27</td>
<td>0.0229</td>
</tr>
<tr>
<td>positiv*XP2.giv=non.ref</td>
<td>-4.00293</td>
<td>1.13074</td>
<td>-3.54</td>
<td>0.0004</td>
</tr>
</tbody>
</table>
Interaction: Set * Givenness

When Set=positiv and second fronted element=non-referential: VPF

(16) [Negativ zu Buche geschlagen habe vor allem das Saisonende in den Wintersportzentren negatively to book hit has before all the season.end in the winter.sports.centers
‘The end of the winter sports season had a negative financial effect.’
## Accuracy

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPF</td>
<td>157</td>
<td>66</td>
</tr>
<tr>
<td>MF</td>
<td>50</td>
<td>119</td>
</tr>
<tr>
<td>Total</td>
<td>207</td>
<td>185</td>
</tr>
</tbody>
</table>

- Model classifies 70% of data correctly.
- Simply guessing (VPF) yields 52.8% correct classifications.
Summary

- Length and focus play a role across synonym sets
- Influence of Set, Interaction Set*Givenness: point to “local” (more low-level) generalizations (synonym sets)
Summary

- Length and focus play a role across synonym sets
- Influence of Set, Interaction Set*Givenness:
  point to “local” (more low-level) generalizations (synonym sets)

Improve model, other candidate factors:

- More fine-grained information-structural categories:
  fronted XP and non-finite verb part of a single IS-unit?
Summary

- Length and focus play a role across synonym sets
- Influence of Set, Interaction Set*Givenness: point to “local” (more low-level) generalizations (synonym sets)

Improve model, other candidate factors:

- More fine-grained information-structural categories: fronted XP and non-finite verb part of a single IS-unit?
- Collocation of prefíeld material
Summary

- Length and focus play a role across synonym sets
- Influence of Set, Interaction Set*Givenness: point to “local” (more low-level) generalizations (synonym sets)

Improve model, other candidate factors:

- More fine-grained information-structural categories: fronted XP and non-finite verb part of a single IS-unit?
- Collocation of prefield material
- Role of predictability of the verb from the Adv-XP sequence
Summary

- Length and focus play a role across synonym sets
- Influence of Set, Interaction Set*Givenness: point to “local” (more low-level) generalizations (synonym sets)

Improve model, other candidate factors:

- More fine-grained information-structural categories: fronted XP and non-finite verb part of a single IS-unit?
- Collocation of prefield material
- Role of predictability of the verb from the Adv-XP sequence
  - Collocation of the [Adv + XP] string with a particular verb: hearer can anticipate the verb when processing the prefield material
Summary

- Length and focus play a role across synonym sets
- Influence of Set, Interaction Set*Givenness:
  point to “local” (more low-level) generalizations (synonym sets)

Improve model, other candidate factors:

- More fine-grained information-structural categories:
  fronted XP and non-finite verb part of a single IS-unit?
- Collocation of prefield material
- Role of predictability of the verb from the Adv-XP sequence
  - Collocation of the [Adv + XP] string with a particular verb:
    hearer can anticipate the verb when processing the prefield material
Summary

- Length and focus play a role across synonym sets
- Influence of Set, Interaction Set*Givenness: point to “local” (more low-level) generalizations (synonym sets)

Improve model, other candidate factors:

- More fine-grained information-structural categories: fronted XP and non-finite verb part of a single IS-unit?
- Collocation of prefield material
- Role of predictability of the verb from the Adv-XP sequence
  - Collocation of the [Adv + XP] string with a particular verb: hearer can anticipate the verb when processing the prefield material

Thank you.


