Verb-cluster variations: A Harmonic Grammar analysis

Markus Bader
Goethe-Universität Frankfurt

New Ways of Analyzing Syntactic Variation
Radboud University Nijmegen – November 15.-17, 2012
In German, verbs normally select their dependent elements to the left. This is true for objects . . .

(1) ... dass Peter [ein Buch ← schreibt].
    that P. a book writes

... and also for verbs selected by another verb.

(2) a. ... dass er [es ← geschrieben ← hat].
    that he it written has
    ‘...that he might have written it.’

b. ... dass er [es ← geschrieben ← haben ← könnte].
    that he it written have could
    ‘...that he might have written it.’

c. ... dass [es ← geschrieben ← worden ← sein ← könnte].
    that it written been be could
    ‘...that it might have been written.’
The general pattern thus looks as in (3):

(3) a. $V_2 \leftarrow V_1$
   b. $V_3 \leftarrow V_2 \leftarrow V_1$
   c. $V_4 \leftarrow V_3 \leftarrow V_2 \leftarrow V_1$

There are certain well-known exceptions to (3):
For $V_1 = \text{Aux}_{\text{perfective}}$ and $V_2 = \text{Modal}$, the auxiliary must be fronted to the cluster initial position according to normative grammars of Standard German:

(4) $\text{Aux}_1 \rightarrow V_3 \leftarrow \text{Mod}_2$

\[
\text{dass er es } [\text{hat} \rightarrow [\text{schreiben} \leftarrow \text{wollen}]].
\]
\[
\text{that he it has write want}
\]
\[
\text{‘that he wanted to write it.’}
\]
However . . .

. . . we find a lot of variation across German dialects and varieties:

(5)  a. Certain variants of Austrian and Bavarian:
    dass er es \([\text{schreiben} \leftarrow \text{wollen}] \leftarrow \text{hat}\]. \(V-\text{Mod-Aux}\)

    b. Pattern typical for Austrian and Bavarian:
    dass er es \([\text{schreiben} \leftarrow [\text{hat} \rightarrow \text{wollen}]\]. \(V-\text{Aux-Mod}\)

    c. Standard German:
    dass er es \([\text{hat} \rightarrow [\text{schreiben} \leftarrow \text{wollen}]\]. \(\text{Aux-V-Mod}\)

    d. Pattern typical for Swiss German:
    dass er es \([\text{hat} \rightarrow [\text{wollen} \rightarrow \text{schreiben}]\]. \(\text{Aux-Mod-V}\)

Furthermore, it is reported that dialects often allow for more than one order.
In fact, variation is also found for non-dialect speakers:

- A series of experimental investigations of verb cluster formation has shown that non-dialect speakers of German do not adhere strictly to the Standard German pattern:
- Native speakers of ‘Colloquial German’ are more liberal than prescriptive grammars of ‘Standard German’ in a precisely defined way.

Question addressed in this talk:

- How are variation as observed in acceptability judgments and variation as observed in corpus frequencies related to each other.
3-verb clusters: Introduction

Topics

- Order among verbs within 3-verb clusters
- Comparison of acceptability and frequency

Two different methods to assess the acceptability/grammaticality of sentences:

- **Speeded Grammaticality Judgments (SGJ):**
  Participants judge sentences as either grammatical or ungrammatical under controlled and timed conditions.

- **Magnitude Estimation (ME):**
  Participants evaluate sentences relative to a reference sentence on a continuous scale.

The experimental results discussed in this section are taken from Bader & Schmid (2009) and Bader & Häussler (2010).
3-verb clusters: Expectation

(6) dass Peter ein Buch (hat) lesen (hat) müssen (hat).
that P. a book has read has must has

(7) dass Peter ein Buch (hat) müssen (hat) lesen (hat).
that P. a book has read has must has

<table>
<thead>
<tr>
<th></th>
<th>$V &lt; Mod$</th>
<th>$Aux = 1$</th>
<th>$V = Aux$</th>
<th>$Aux = 2$</th>
<th>$V = Mod$</th>
<th>$Aux = 3$</th>
<th>$Mod = Aux$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Mod &lt; V$</td>
<td>Aux-Mod-V</td>
<td>Mod-Aux-V</td>
<td>Mod-V-Aux</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Expectation based on normative grammar:**
  - high percentages of judgments ‘grammatical’ for order $Aux-V-Mod$
  - low percentages for the remaining five orders
3-verb clusters: Procedures

**Speeded Grammaticality Judgments**
- Word-by-word presentation in the middle of the screen
- Presentation time for each word: 225 ms plus an additional 25 ms per character
- End-of-sentence judgments with a deadline of 2000 ms

**Magnitude Estimation**
- First, a reference item is presented to which the participant assigns an arbitrary numeric value (> 0).
- All further items are judged in proportion to the reference item on a continuous numerical scale.
- Each individual data point is divided by the reference value and the resulting ratio is log-transformed.
### 3-verb clusters: Representative results

**Left:** Speeded grammaticality judgments

**Right:** Magnitude estimation

<table>
<thead>
<tr>
<th></th>
<th>V&lt;Mod</th>
<th>Mod&lt;V</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aux=1</strong></td>
<td>86</td>
<td>12</td>
</tr>
<tr>
<td><strong>Aux=2</strong></td>
<td>69</td>
<td>3</td>
</tr>
<tr>
<td><strong>Aux=3</strong></td>
<td>26</td>
<td>2</td>
</tr>
</tbody>
</table>

**SGJ (%)**
- Aux=1: 86
- Aux=2: 69
- Aux=3: 26

**ME (log)**
- Aux=1: 0.51
- Aux=2: 0.20
- Aux=3: -0.27

- Aux=1: -0.29
- Aux=2: -0.58
- Aux=3: -0.66
In all experiments, the Standard German order Aux-V-Mod received the best judgments.

Independently of regional background, the partially inverted order V-Aux-Mod was judged only somewhat worse, despite being ungrammatical according to normative grammar.

The remaining orders were judged as unacceptable, but with differing degrees:

- When the modal verb preceded the lexical verb, the aux-initial order Aux-MOD-V was judged better than the other two orders (MOD-Aux-V and MOD-V-Aux).
3-verb clusters: Expectation

(8) dass Peter ein Buch hat lesen müssen.  \text{Aux-V-Mod}
that P. a book has read must

(9) dass Peter ein Buch lesen hat müssen.  \text{V-Aux-Mod}
that P. a book read has must

\begin{center}
\begin{tabular}{l|c|c|c}
  & \text{Aux = 1} & \text{Aux = 2} & \text{Aux = 3} \\
\hline
\text{V < Mod} & \text{Aux-V-Mod} & \text{V-Aux-Mod} & \text{V-Mod-Aux} \\
\hline
\text{Mod < V} & \text{Aux-Mod-V} & \text{Mod-Aux-V} & \text{Mod-V-Aux} \\
\end{tabular}
\end{center}

(10) a. \textit{The Mod-V Constraint}
The complement of a modal verb precedes the modal verb.
b. \textit{The Aux-Mod Restriction Restriction}
When the perfect auxiliary selects a modal verb, it must precede it.
**On Gradience and Frequency**

Table: Experimental data for word order in 3-verb clusters.

<table>
<thead>
<tr>
<th></th>
<th>V&gt;Mod Constr.</th>
<th>Aux&gt;Mod Constr.</th>
<th>ME Exp.</th>
<th>SGJ Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>V&lt;Mod</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aux = 1</td>
<td>.22</td>
<td></td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Aux = 2</td>
<td></td>
<td>.09</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Aux = 3</td>
<td></td>
<td></td>
<td>-.12</td>
<td>28</td>
</tr>
<tr>
<td>Mod&lt;V</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aux = 1</td>
<td>*</td>
<td>-.13</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Aux = 2</td>
<td>*</td>
<td></td>
<td>-.25</td>
<td>5</td>
</tr>
<tr>
<td>Aux = 3</td>
<td>*</td>
<td></td>
<td>-.29</td>
<td>4</td>
</tr>
</tbody>
</table>

- The two orders which violate neither of the two constraints are judged best.
- The two orders which violate only a single constraint are judged worse, but are not rejected completely.
- The two orders which violate both constraints are rejected completely.
Frequency of verb-final clusters in complementizer-initial embedded clauses in the Tiger Corpus (Version 2), a treebank containing about 50,000 sentences (ca. 888,000 tokens) from the German newspaper *Frankfurter Rundschau*.

(11)  

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>V[fin] - final</td>
<td>3959</td>
<td>2635</td>
<td>319</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>V[fin] - initial</td>
<td>-</td>
<td>-</td>
<td>11</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

- **Newspaper Corpus**: the newspaper corpus made available by the Institute for German Language (IDS) in Mannheim/Germany ([http://www.ids-mannheim.de](http://www.ids-mannheim.de))
- **Web Corpus**: the deWaC corpus of German internet texts made available by the University of Bologna Baroni et al. (2009)
**Table**: Corpus and experimental data for word order in 3-verb clusters.

<table>
<thead>
<tr>
<th>V&lt;Mod</th>
<th>Aux = 1</th>
<th>Aux = 2</th>
<th>Aux = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>V&gt;Mod</td>
<td>Newspaper</td>
<td>Web</td>
<td>ME</td>
</tr>
<tr>
<td>V&lt;Mod</td>
<td>Aux = 1</td>
<td>*</td>
<td>0</td>
</tr>
<tr>
<td>Aux = 2</td>
<td>99.5</td>
<td>96.2</td>
<td>.22</td>
</tr>
<tr>
<td>Aux = 3</td>
<td>0.5</td>
<td>3.8</td>
<td>.09</td>
</tr>
<tr>
<td>Mod&lt;V</td>
<td>Aux = 1</td>
<td>*</td>
<td>0</td>
</tr>
<tr>
<td>Aux = 2</td>
<td>0</td>
<td>0</td>
<td>-.25</td>
</tr>
<tr>
<td>Aux = 3</td>
<td>0</td>
<td>0</td>
<td>-.29</td>
</tr>
</tbody>
</table>

- The frequency distribution is much more skewed than the acceptability distribution: Small variations in acceptability go hand in hand with large variations in frequency.
- Orders with zero or near-zero frequency can still differ with regard to acceptability.
Orders with zero or near-zero frequency can still differ with regard to acceptability.

- This argues against theories like Stochastic OT (Boersma & Hayes, 2001) which tie acceptability to frequency on the level of whole structures.

- What is needed instead is a way to relate frequency to acceptability on the level of individual constraints.

The frequency distribution is much more skewed than the acceptability distribution.

- Frequency and acceptability must be related to each other in a non-linear way.
On Gradience and Frequency

Relating frequency to acceptability:

- Recent versions of *Harmonic Grammar* (Pater, 2009; Boersma & Pater, 2008) have the desired properties (see also Goldwater & Johnson, 2003; Jäger, 2003; Jäger & Rosenbach, 2006):
  - The weight of individual constraints is learned by frequency.
  - The relationship between frequency and constraint weight is non-linear.

- Constraint weights can be related to acceptability in the way of *Linear OT*: Keller (2000, 2006)
On Gradience and Frequency

Constraints for modelling:

(12) a. Align(V, L) (V > Mod)
    b. Align(Aux_{Mod}, L) (Aux_{Mod} > Mod)

Experiments with the praat program indicate that constraint weights accounting for acceptability can be learned from frequency.

Open question: How to model the advantage of Aux-V-Mod over V-Aux-Mod?

- Align(Aux_{Mod}, Initial)?
- Making Align(Aux_{Mod}, L) a gradable constraint?
The final part of this talk presents new experimental evidence as well as corpus data on 4-verb clusters.

3-verb cluster

(13) dass jemand die Schubkarre hätte reparieren müssen.  
that someone the wheelbarrow had repair must  
‘...that someone should have repaired the wheelbarrow.’

4-verb clusters

(14) dass die Schubkarre hätte repariert werden müssen.  
that the wheelbarrow had repaired be must  
‘...that the wheelbarrow should have been repaired.’

(15) dass Max die Schubk. hätte reparieren lassen müssen.  
that M. the wheelb. had repaired let must  
‘...that Werner should have let repair the wheelbarrow.’
4-verb clusters

(16) dass die Schubkarre hätte repaired werden must
     that the wheelbarrow had repaired be must
     ‘... that the wheelbarrow should have been repaired.’

(17)

<table>
<thead>
<tr>
<th>Aux = 1</th>
<th>V-Pass &lt; Mod</th>
<th>Mod &lt; V-Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aux V4 Pass Mod</td>
<td>AUX1 V4 Pass3 Mod2</td>
<td>AUX1 Mod2 Pass3 V4</td>
</tr>
<tr>
<td>Aux = 2</td>
<td>V4 Aux Pass Mod</td>
<td>Mod2 Aux V4 Pass3</td>
</tr>
<tr>
<td>Aux = 3</td>
<td>V4 Pass Aux Mod</td>
<td>Mod2 V4 Aux Pass3</td>
</tr>
<tr>
<td>Aux = 4</td>
<td>V4 Pass Mod Aux</td>
<td>Mod2 V4 Pass Aux1</td>
</tr>
</tbody>
</table>
Experiment 1: 4-verb clusters
Method: Magnitude estimation

Ich glaube, dass die Schubkarre schon letzte Woche . . .
I think that the wheelbarrow already last week

V < Mod

<table>
<thead>
<tr>
<th>Aux</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HÄTTE repariert werden müssen (had repaired be must)</td>
</tr>
<tr>
<td>2</td>
<td>repariert HÄTTE werden müssen (repaired had to be must)</td>
</tr>
<tr>
<td>3</td>
<td>repariert werden HÄTTE müssen (repaired must have been)</td>
</tr>
<tr>
<td>4</td>
<td>repariert werden müssen HÄTTE (repaired must have been)</td>
</tr>
</tbody>
</table>

Mod < V

<table>
<thead>
<tr>
<th>Aux</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HÄTTE müssen repariert werden (must have been repaired)</td>
</tr>
<tr>
<td>2</td>
<td>müssen HÄTTE repariert werden (must have been repaired)</td>
</tr>
<tr>
<td>3</td>
<td>müssen repariert HÄTTE werden (must have been repaired)</td>
</tr>
<tr>
<td>4</td>
<td>müssen repariert werden HÄTTE (must have been repaired)</td>
</tr>
</tbody>
</table>

Translation for all conditions:
‘I think that the wheelbarrow had to be repaired already last week.’
Experiment 2: 4-verb clusters
Method: Magnitude estimation

Ich glaube, dass Max die Schubkarre sofort . . .
I think that M. the wheelbarrow immediately

| V < Mod | Aux=1 | . . . HÄTTE reparieren lassen müssen |
| Aux=2  | . . . reparieren HÄTTE lassen müssen |
| Aux=3  | . . . reparieren lassen HÄTTE müssen |
| Aux=4  | . . . reparieren lassen müssen HÄTTE |

| Mod < V | Aux=1 | . . . HÄTTE müssen reparieren lassen |
| Aux=2  | . . . müssen HÄTTE reparieren lassen |
| Aux=3  | . . . müssen reparieren HÄTTE lassen |
| Aux=4  | . . . müssen reparieren lassen HÄTTE |

Translation for all conditions:
‘I think that the wheelbarrow had to be repaired already last week.’
Figure: Results of Experiment 1 and 2
4-verb clusters

Figure: Results of Experiment 4 of Bader & Schmid (2009) and results of Experiment 1
4-verb clusters

Table: Corpus and experimental data for word order in 3-verb clusters.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>V&lt;Mod</td>
<td>Aux = 1</td>
<td>100</td>
<td>95.1</td>
<td>.26</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td>Aux = 2</td>
<td>0</td>
<td>4.0</td>
<td>.17</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td>Aux = 3</td>
<td>0</td>
<td>0.8</td>
<td>.04</td>
<td>.11</td>
</tr>
<tr>
<td></td>
<td>Aux = 4</td>
<td>*</td>
<td>0</td>
<td>0</td>
<td>-.31</td>
</tr>
<tr>
<td>Mod&lt;V</td>
<td>Aux = 1</td>
<td>*</td>
<td>0</td>
<td>0</td>
<td>-.08</td>
</tr>
<tr>
<td></td>
<td>Aux = 2</td>
<td>*</td>
<td>0</td>
<td>0</td>
<td>-.34</td>
</tr>
<tr>
<td></td>
<td>Aux = 3</td>
<td>*</td>
<td>0</td>
<td>0</td>
<td>-.43</td>
</tr>
<tr>
<td></td>
<td>Aux = 4</td>
<td>*</td>
<td>0</td>
<td>0</td>
<td>-.45</td>
</tr>
</tbody>
</table>

The results are similar to those for 3-verb clusters:

- All orders that violate neither of the two constraints receive positive acceptability scores, all other orders receive negative acceptability constraints.
- Only orders that violate no constraint are found in the corpus.
- There is some evidence that the position of the auxiliary should be modeled by a gradable constraint.
Variation as seen in German complex verb clusters presents a complex relationship between frequency and acceptability.

- Variation in acceptability despite zero frequency.
- Small variations in acceptability go hand in hand with large variations in frequency.

Recent developments in weight-based OT allow us to capture the frequency-relationship observed for verb-clusters.
References


