Attraction through formal resemblance.

Five case studies on constructional contamination

Dirk Pijpops, Isabeau De Smet & Freek Van de Velde

Lilith RU, University of Liège
QLVL RU, University of Leuven
WHAT IS CONSTRUCTIONAL CONTAMINATION?

WHY SHOULD I CARE?
WHAT IS CONSTRUCTIONAL CONTAMINATION?

A usage-based mechanism that creates lexical biases in morphosyntactic variation
WHY SHOULD I CARE?

You're interested in language variation and want to score a quick publication

You're interested in language processing and its effects on the language system
THANKS TO

Gert De Sutter

Tom Ruette
1. How does constructional contamination work?

2. Case studies

3. Recap: how to apply it to your case study?
How does constructional contamination work?

1. An alternation: partitive genitive with -∅ vs. -s

<table>
<thead>
<tr>
<th></th>
<th>-∅</th>
<th>vs.</th>
<th>-s</th>
</tr>
</thead>
<tbody>
<tr>
<td>iets leuk</td>
<td>'something fun'</td>
<td>iets leuks</td>
<td>'something fun-s'</td>
</tr>
<tr>
<td>iets verkeerd</td>
<td>'something wrong'</td>
<td>iets verkeerds</td>
<td>'something wrong-s'</td>
</tr>
</tbody>
</table>
How does constructional contamination work?

1. An alternation: partitive genitive with -∅ vs. -s

2. Another construction that often looks like one of the variants

Partitive genitive: \(ik\ heb\ {\underline{jets\ verkeerd}}\ gegeten.\)

‘I have eaten \underline{something wrong}.’

Construction with adverb: \(ik\ heb\ iets\ {\underline{verkeerd}}\ verstaan.\)

‘I have \underline{misunderstood} something.’
How does constructional contamination work?

1. An alternation: partitive genitive with -Ø vs. -s

2. Another construction that often looks like one of the variants

   Partitive genitive: \[ i k \ heb \ jets \ verkeerd s \ gegeten. \]
   
   ‘I have eaten something wrong.’

   Construction with adverb: \[ i k \ heb \ iets \ verkeerd \ verstaan. \]
   
   ‘I have misunderstood something.’
How does constructional contamination work?

1. An alternation: partitive genitive with -∅ vs. -s
2. Another construction that often looks like one of the variants, viz. -∅
3. That other construction is particularly frequent among some lexical items

Highly frequent among verkeerd:  
- *ik heb iets verkeerd verstaan.*  
  ‘I have misunderstood something’.

Less frequent among leuk:  
- *Je hebt iets leuk ingepakt.*  
  ‘You have wrapped something up in a fun way.’
How does constructional contamination work?

1. An alternation: partitive genitive with -∅ vs. -s

2. Another construction that often looks like one of the variants, viz. -∅

3. That other construction is particularly frequent among some lexical items:
   language users very often hear the string *iets verkeerd* 'something wrong(ly)', without -s, and less often *iets leuk* 'something fun/in a fun way', without -s
How does constructional contamination work?

1. An alternation: partitive genitive with -∅ vs. -s
2. Another construction that often looks like one of the variants, viz. -∅
3. That other construction is particularly frequent among some lexical items:
   language users very often hear the string *iets verkeerd* 'something wrong(ly)', without -s, and less often *iets leuk* 'something fun/in a fun way', without -s
4. Language users store unanalyzed chunks as ready made(s) (Dąbrowska 2014)

⇒ Even among unambiguous partitive genitives, language users will be more inclined to drop the -s for *iets verkeerd(s)* 'something wrong'
How does constructional contamination work?

1. An alternation, e.g. variant X and variant Y
2. Another (unrelated) construction that often looks like one of the variants, e.g. variant X
3. That other construction is highly frequent for word A
4. Language users store unanalyzed chunks as ready mades (Dąbrowska 2014)

⇒ Even among strictly unambiguous instances of the alternating construction, word A will be biased towards variant X
**CASE STUDIES**

<table>
<thead>
<tr>
<th>Dutch</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case study 1: the partitive genitive</td>
<td>Case study 6: Adverb placement</td>
</tr>
<tr>
<td>Case study 2: verbal clusters</td>
<td>(Hilpert 2019: 65-68, Hilpert &amp; Flach forthc.)</td>
</tr>
<tr>
<td>Case study 3: past tense formation</td>
<td></td>
</tr>
<tr>
<td>Case study 4: long and bare infinitives</td>
<td></td>
</tr>
<tr>
<td>Case study 5: comparative placement</td>
<td></td>
</tr>
</tbody>
</table>
## CASE STUDIES

<table>
<thead>
<tr>
<th>Dutch</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Case study 1:</strong> the partitive genitive</td>
<td>Case study 6: Adverb placement</td>
</tr>
<tr>
<td>Case study 2: verbal clusters</td>
<td>(Hilpert 2019: 65-68, Hilpert &amp; Flach forthc.)</td>
</tr>
<tr>
<td>Case study 3: past tense formation</td>
<td></td>
</tr>
<tr>
<td>Case study 4: long and bare infinitives</td>
<td></td>
</tr>
<tr>
<td>Case study 5: comparative placement</td>
<td></td>
</tr>
</tbody>
</table>
1. An alternation: partitive genitive with -∅ vs. -s

-∅ vs. -s

*iets leuk*  
'something fun'

*iets leuks*  
'something fun-s'

*iets verkeerd*  
'something wrong'

*iets verkeerds*  
'something wrong-s'
2 & 3. Strings that superficially resemble partitive genitives on -∅ are particularly frequent with assessment adjectives *verkeerd* 'wrong', *beter* 'better',... and color adjectives

\[
\begin{array}{ccc}
\text{ iets verkeerd } & \text{ niets verkeerd } & \text{ iets beter } \\
\text{ Something wrong } & \text{ nothing better } & \text{ something better } \\
\text{ wat geel } & \text{ veel blauw } & \\
\text{ Something yellow } & \text{ a lot of blue } & \\
\end{array}
\]

⇒ Prediction: Even among unambiguous partitive genitives, language users will be more inclined to drop the -s for phrases containing these adjectives
• Only strictly unambiguous partitive genities

• Mixed regression model, controlling for other known factors and random lexical preferences
**CASE STUDIES**

<table>
<thead>
<tr>
<th>Dutch</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Case study 1:</strong> the partitive genitive</td>
<td>Case study 6: Adverb placement</td>
</tr>
<tr>
<td><strong>Case study 2:</strong> verbal clusters</td>
<td>(Hilpert 2019: 65-68, Hilpert &amp; Flach forthc.)</td>
</tr>
<tr>
<td>Case study 3: past tense formation</td>
<td></td>
</tr>
<tr>
<td>Case study 4: long and bare infinitives</td>
<td></td>
</tr>
<tr>
<td>Case study 5: comparative placement</td>
<td></td>
</tr>
</tbody>
</table>
Verbal clusters

1. An alternation: VERBAL PARTICIPLE + AUXILIARY VS. AUXILIARY + VERBAL PARTICIPLE

   *dat de deur net door John* **gesloten is**    vs. *dat de deur net door John* **is gesloten**

   That the door just by John **closed is**          that the door just by John **is closed**

2. Another construction that superficially looks like one of the variants: ADJECTIVE + COPULA

   *dat de deur al geruime tijd* **gesloten is**

   That the door already some time **closed is**

⇒ Prediction 1: the more often a participle is used as an adjective, the more often it will appear in the order VERBAL PARTICIPLE + AUXILIARY in unambiguously verbal instances with the auxiliaries *zijn 'be'* or *worden 'become'*
Verbal clusters

Prediction 1: The more often a participle is used as an adjective, the more often it will appear in the PARTICIPLE + AUXILIARY order in unambiguously verbal instances, with the auxiliaries *zijn* 'to be' or *worden* 'become'

   e.g. *gesloten* 'closed' would prefer the PARTICIPLE + AUXILIARY order compared to *geopend* 'opened'

   → From there, the effect might spread:  ... *gisteren gesloten is*  ~  *gisteren gesloten heeft*
   
   closed is  closed has

Prediction 2: The same lexical bias might also appear among verbal clusters with other auxiliaries, but it should be weaker.
Verbal clusters

- Dataset from Gert De Sutter, with auxiliaries *zijn* 'be', *worden* 'become' and *hebben* 'have'
- De Sutter distinguished between ambiguous & unambiguous verbal clusters
- We only looked at unambiguous verbal clusters
- For each participle:

\[
\text{Adjectiveness} = \arcsin \left( \frac{\text{occurrences as an adjective in the Corpus of Spoken Dutch}}{\sqrt{\text{Total number of occurrences in the Corpus of Spoken Dutch}}} \right)
\]
Auxiliaries *zijn* 'be' and *worden* 'become'

Estimated probability of the PARTICIPLE + AUXILIARY order

Odds Ratio = 3.96
p-value = 0.001

Auxiliary *hebben* 'have'

Estimated probability of the PARTICIPLE + AUXILIARY order

Odds Ratio = 2.54
p-value = 0.132
CASE STUDIES

Dutch

Case study 1: the partitive genitive
Case study 2: verbal clusters
Case study 3: past tense formation
Case study 4: long and bare infinitives
Case study 5: comparative placement

English

Case study 6: Adverb placement
(Hilpert 2019: 65-68,
Hilpert & Flach forthc.)
Other case studies

3. Paste tense formation (Vosters 2012, De Smet 2021)

Gisteren *graafde* ik een put \[\text{vs.}\] Gisteren *groef* ik een put

Yesterday digged I a hole \[\text{vs.}\] Yesterday dug I a hole

4. Long and bare infinitives among the posture verbs (Pijpops, De Smet & Van de Velde 2018)

*Terwijl* we zaten *te wachten*,... \[\text{vs.}\] *Terwijl* we zaten *wachten*,...

While we sat to wait \[\text{vs.}\] while we sat wait

5. Comparative placement (Van de Velde & Pijpops 2018)

*Een* lager *dan verwachte* opkomst \[\text{vs.}\] *een* lagere opkomst *dan verwacht*

A lower than expected attendance \[\text{vs.}\] a lower attendance than expected


*The* driver was *instantly killed* \[\text{vs.}\] *the* driver was *killed instantly*
Recap

1. An alternation, e.g. variant X and variant Y
2. Another (unrelated) construction that often looks like one of the variants, e.g. variant X
3. That other construction is more frequent for word A than word B
4. Language users store unanalyzed chunks as ready mades (Dąbrowska 2014)

⇒ Test for constructional contamination:

   Even among strictly unambiguous instances of the alternating construction, word A will be biased towards variant X
Want to know more?


dirk.pijpops@uliege.be

Don't like Dutch?


References


