Tangled: Interaction of linguistic and extralinguistic factors in the use of Dutch causative constructions

Natalia Levshina
Dirk Geeraerts
Dirk Speelman

RU Quantitative Lexicology and Variational Linguistics
University of Leuven
We need to take into account the various levels of the language variation model, and the way they interact.

Sjef Barbiers, ICLaVE 6
Outline

1. A 3D model of variation
2. Dutch causative constructions
3. Data and variables
4. 3D models of Dutch causative constructions
5. Discussion
Background

- third-wave Sociolinguistics: variation is meaningful
- usage-based grammar and lexicology: meaning varies
- lectally enriched studies of semantically/functionally overlapping units: Grondelaers et al. 2002; Glynn 2007; Bresnan & Hay 2008; Speelman & Geeraerts 2009; Bresnan & Ford 2010; Szmrecsanyi 2010, etc.
- so far, no integrative account of relationships between form, function and context
A 3D Model of variation

linguistic form

function

context
Independence

linguistic form

function  context
Independence

- not many examples for near-synonyms or alternations (Bresnan et al. 2007: a small independent effect of channel on the prepositional vs. double-object dative)
- interpretation: no such thing as 'free variation' of meaningful units (cf. Goldberg's principle of No Synonymy)
Interaction

linguistic form

function
context
Interaction

- difference in strength of semantic factors in geographic varieties, e.g. Grondelaers et al. 2002; Bresnan & Hay 2008; Bresnan & Ford 2010
- moderating effects of situational features are less frequently reported (channel: Szmrecsanyi 2010)
- cognitive interpretation: variation in cue validity is more natural in different linguistic systems than within one system
'Differences in input'

linguistic form

function context
'Differences in input'

- e.g. differences in frequency of recipient pronouns in spoken and written corpora explain the difference in proportions of double object datives (Bresnan et al. 2007; Kendall et al. 2010)
- not studied systematically: a part of language?
- interpretation: socioconceptual variation or corpus bias?
Case study

How do geographic and conceptual factors interact in variation of Dutch causative constructions with *doen* and *laten*?
Outline

1. A 3D model of variation
2. Dutch causative constructions
3. Data and variables
4. 3D models of Dutch causative constructions
5. Discussion
Dutch Causative Constructions

De politie deed de auto stoppen.

The police made made/let the car stop

CAUSER

made

AUX

CAUSEE

stoppen

EFFECTED

PREDICATE
## Conceptual variation

<table>
<thead>
<tr>
<th>-done-</th>
<th><strong>laten</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct causation</td>
<td>Indirect causation</td>
</tr>
<tr>
<td>«The initiator produces the effect event directly; there is no intervening energy source ‘downstream’»</td>
<td>«Some other force besides the initiator is the most immediate source of energy in the effect event»</td>
</tr>
</tbody>
</table>

Verhagen & Kemmer (1997)
Outline

1. A 3D model of variation
2. Dutch causative constructions
3. Data and variables
4. 3D models of Dutch causative constructions
5. Discussion
Data

- newspaper corpora TwNC (the Netherlands) and LeNC (Belgium), syntactically parsed
- automatically extracted *doen*/*laten* + Infinitive, manually cleaned up
- 6853 observations = 1168 *doen* + 5685 *laten*
Conceptual variables

CrSem=Anim
EPTtrans=Tr
Coref=Yes
Possess=Yes
CdEventSem=Non-Ment
Polarity=Neg
GeSynt=Periph

CrSem=Inanim
EPTtrans=Intr
Coref=No
Possess=No
CdEventSem=Ment
Polarity=Pos
GeSynt=Central

probability of doen
Distinctive Prototypes

**laten**

Hij liet zich niet pakken door de politie in zijn eigen huis.

He didn't allow the police to arrest him in his own house.

**doen**

Zijn kapsel doet me denken aan een vogelnest.

His hairstyle reminds me of a bird's nest.
Two national 'Dutches' in Europe
Geographic variation of doen/laten

OR = 2.15
Outline

1. A 3D model of variation
2. Dutch causative constructions
3. Data and variables
4. 3D models of Dutch causative constructions
5. Discussion
Model A

- logistic regression with BOTH conceptual and geographic variables, NO 3D interactions tested
- OR Country does not decrease: 2.15 (solo) vs. 2.20 (controlled for semantics)
Model A

- logistic regression with BOTH conceptual and geographic variables, NO 3D interactions tested

- OR Country does not decrease: 2.15 (solo) vs. 2.20 (controlled for semantics)

→ the national difference in doen/laten ratio is NOT due to difference in input
Model B

- tested all two-way 3D interactions
- found significant 3D interactions
Model B

- tested all two-way 3D interactions
- found significant 3D interactions
  → interaction,
  no independence
Interactions

Transitivity * Country

Coreferentiality * Country
Adding Random Effects

- collocations Auxiliary + Effected Predicate: *doen denken (aan)* 'remind (of)', *laten zien* 'show', *laten weten* 'inform', etc.

- Model C: Effected Predicates as random effects in a mixed-effect logistic regression model:
  - adjustments to the intercept
  - adjustments to the effect of Country
Results of ME Modelling

- most effects remain stable
- Coreferentiality*Country is even less outspoken
- Transitivity*Country is no longer significant: was the effect due to specific verbs?

  *laten zien 'see', weten 'know' and horen “hear”:
  28.3% of all Netherlandic observations
  only 3.4% of all Belgian observations

→ interaction is due to difference in input (highly specific semantic patterns)
Random effects
Outline

1. A 3D model of variation
2. Dutch causative constructions
3. Data and variables
4. 3D models of Dutch causative constructions
5. Discussion
Discussion

• we found interactions at the level of general features
• however, some interactions at the more schematic level can be explained by difference in input at the more specific level (due to corpus bias or true socioconceptual variation?)
• thus, studying lower-level schemas is very important for interpreting constructional variation
Tangled