From picture to preposition: Modeling lexical-semantic variation in the spatial domain

Anja Goldschmidt & Joost Zwarts
Utrecht University

Workshop on Semantic Variation, University of Chicago, October 25-26, 2013
Spatial semantic variation

• Path and motion, ‘to’, ‘away’, ‘into’ (Talmy 1985, Beavers et al. 2010, Pantcheva 2011, ...)
• Axes and frame of reference, ‘behind’, ‘above’, ‘right’ (Levinson 2003, Svenonius 2006, ...)

Lexical-semantic variation in the topological domain

• English on versus Dutch op ‘support’ and aan ‘attachment’
• Japanese ue versus English on ‘contact’ and above ‘no contact’
• Tiriyó tao versus English in and through
• English on versus Yélî Dnye (u)mbêmê ‘on top’ and p:uu ‘attachment’

Levinson & Wilkins (2006)
Two perspectives on spatial variation

• *Intensional* perspective: How do spatial meanings vary across languages in terms of spatial *features* (like attachment and support)?

• *Extensional* perspective: How do spatial meanings vary across languages in terms of set *extensions* consisting of concrete spatial situations (like 1 and 2)

![Diagram](image)
The extensional perspective

- A set of visual stimuli (scenes, pictures, video clips) with spatial configurations
- Systematic elicitation of descriptions of those spatial configurations
- The *extension* of a spatial term as a set of stimuli
- Topological Relations Picture Series (TRPS, Bowerman & Pederson 1992)
- Bowerman (1996), Levinson & Meira (2003), Khetarpal et al. (2009)
Extension of *pëke* (Tiriyó, Meira 2006)

Arrow points to the *figure* of the spatial relation, the other object is the *ground*.
The intensional perspective

• Lexical-semantic variation in the spatial domain is understood in terms of universal spatial features
• Either extracted ‘bottom-up’ from the variational data (Bowerman & Pederson 1996, Levinson & Meira 2003, Regier et al. 2013)
• Or postulated ‘top-down’ for the data (Feist 2000, Xu & Kemp 2010)
Our approach

- Spatial relations are named in concrete situations, like pictures from the TRPS (extensional perspective)
- The spatial relations are analyzed in terms of features based on models of space and force-dynamics (intensional perspective)
- The variation in naming is accounted for by Optimality Theory, the mechanism for dealing with feature-based linguistic variation
Remainder of this talk

• Modeling spatial meanings
• Optimizing spatial namings
• Sources of variation and universals
• Conclusions and questions
Four types of Figure – Ground relations

• Topological: Region Connection Calculus (Randell, Cui & Cohn 1992)
• Projective: vector relations between objects (e.g. Zwarts & Winter 2000, Regier & Carlson 2001)
• Path-based: linear spatial trajectories (e.g. Jackendoff 1983, Nam 1995, Zwarts 2005)
Region Connection Calculus

• Topological logic with primitive regions, relation of *connection* (sharing at least one point) and operation of convex hull.

• Other relations can be defined:

  ![Diagram](image)

  - Disconnected
  - Part ‘inclusion’
  - Overlap
  - Externally connected ‘contact’
Topological relations in RCC

- **Contact**: Eigenspace of F externally connected to eigenspace of G
- **Partial inclusion**: Eigenspace of F overlaps with convex hull of G
- **Full inclusion**: Eigenspace of F part of convex hull of G
Force-dynamic relations

- **Control**: Every movement of the ground causes a movement of the figure.
- **Support**: Every translation of the ground takes the figure along with it.
- **Attachment**: Every rotation of the ground takes the figure along with it.
Projective relations

- **Superiority**: The vector pointing from ground to figure is upward.

- **Inferiority**: The vector pointing from ground to figure is downward.
Paths

• A path is a sequence of pairwise connected regions.
• Can be used to represent a path of motion
• Can also be used to represent elongated objects

arrow through apple, bandana around head
Annotating stimuli

• All pictures of the TRPS have been annotated with these features.

• ‘Ambiguous’ cases:

  11
  either visible part externally connected to water, or complete eigenspace partially included in water

  45
  either externally connected to eigenspace of tree, or inside its convex hull

  50
  either externally connected to wall, or partially included in it
Optimality Theory

- Optimality Theory (OT) developed in Phonology (McCarthy & Prince 1993, ...)
- Uses in Syntax and Semantics (Blutner 2000, Hendriks & de Hoop 2001, de Swart & Zwarts 2009, ...)
- Adapted for the analysis of kinship terminology by Jones (2003)
- Attractive because it generates language variation from a set of universal “building blocks”
Optimality Theory

• OT models an optimization process: it evaluates an input through hierarchically ordered constraints, and generates the optimal output

• Spatial expressions: types of spatial relations as the input, feature bundles representing the categories as the output

• Evaluative constraints are based on features (topological, force-dynamic, ...)

• Language variation is modeled by different orderings of the same constraints
Two types of constraints in OT

- **Faithfulness constraints** require “faith” to a feature. Attachment, for instance, is binary: attachment/no attachment. A corresponding faithfulness constraint would require two terms in a target language, one expressing ‘attachment’, the other expressing ‘no attachment’.

- **Markedness constraints** prevent the expression of a feature distinction. In the case of attachment, a corresponding markedness constraint would prevent the separate expression in two terms of the distinction attachment/no attachment (i.e. distinctions are not marked linguistically).
Analysis of Dutch *aan, op* and English *on*

- taking three pictures of three different types as representative for all tokens of that type gives a simplified category structure:
Modeling the optimization process

<table>
<thead>
<tr>
<th></th>
<th>DUTCH</th>
<th>ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Apple on Branch (SIAt)</strong></td>
<td><em>I</em></td>
<td>DProRel</td>
</tr>
<tr>
<td><strong>Phone on Wall (SLAt)</strong></td>
<td><em>I</em></td>
<td>DProRel</td>
</tr>
<tr>
<td><strong>Cup on Table (SSNAt)</strong></td>
<td><em>I</em></td>
<td>DProRel</td>
</tr>
</tbody>
</table>

SIAt: support, inferiority, attachment  
SLAt: support, lateral, attachment  
SSNAt: support, superiority, no attachment
Analysis of Dutch *in, door, om* and Tiriyó *tae, tao*

- taking three pictures of three different types as representative for all tokens of that type gives a simplified category structure:
The optimization process for Dutch and Tiriyó

<table>
<thead>
<tr>
<th>DUTCH</th>
<th>DInclObj</th>
<th>DPath</th>
<th>*FiP</th>
<th>*FiNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple in Bowl (FiNP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple in Bowl (FiNP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrow through Apple (FiP)</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bandana around Head (GrP)</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TIRIYÓ</th>
<th>DInclObj</th>
<th>*FiP</th>
<th>*FiNP</th>
<th>DPath</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple in Bowl (FiNP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple in Bowl (FiNP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrow through Apple (FiP)</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bandana around Head (GrP)</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FiNP: figure included, no path
FiP: figure included, path
GrP: ground included, path
Two sources of variation

• Variation of conceptualizations:

11


45


50


• This is sometimes mirrored in language:
  Apples can be *in* or *on* a tree, a ship can be *in* or *on* the water, hooks can be *in* or *on* the wall
Two sources of variation

• Variation can be restricted by a universal hierarchy of constraints: Some features are always more important than others

• Probably no language with distinction for contact/no contact in inferiority relations (under), but not superiority relations (on/above)
Conclusions

• Formal(izable) semantic concepts can be used to model the spatial properties of concrete situations,

• which allows for a characterization of how languages optimize the lexicalization of those situations in different ways,

• on the basis of a small number of universally shared constraints.
Bibliography

THANK YOU FOR YOUR ATTENTION!

Contact:
• Anja Goldschmidt: A.Goldschmidt@uu.nl
• Joost Zwarts: J.Zwarts@uu.nl