Semantic Typology
What is semantic typology?

• categorization

Figure 1. *The spork dilemma*

http://karenjlloyd.com/blog/2009/01/08/extreme-close-up-wall-e/
"...we are parties to an agreement to organize [nature] in this way – an agreement that holds throughout our speech community and is codified in the patterns of our language. This agreement is of course an implicit and unstated one, but its terms are absolutely obligatory; we cannot talk at all except by subscribing to the organization and classification of data which the agreement decrees." (Whorf [1940] 1956: 212-214)
What is semantic typology (cont.)

- A more typical example of linguistic categorization

![Diagram showing basic color terms in the "grue" domain.](image-url)

**Figure 2.** Basic color terms in the "grue" domain
How do languages categorize color?
• collecting color terminologies
  • Berlin & Kay’s (1969) approach
    • collection of basic color terms (BCTs) from 20 languages
    • the BCTs of each language were determined at the outset according to the following criteria
      • monomorphemic – e.g., blue, but not light blue
      • autochthonous – e.g., blue, but not turquoise
      • general currency in the community (not just used by specialists)
      • not referentially restricted to particular objects or non-color properties of those
  • etic grid: the Munsell color chart
    • a continuum of 40 hues evenly spaced in terms of wavelengths cross-tabulated with a continuum of 8 brightness values
    • plus a ten-brightness-levels black-white continuum
    • all colors presented at maximum saturation

Figure 11. The Munsell color chart (http://www.icsi.berkeley.edu/wcs/grid.jpg)
• findings, generalizations, hypotheses
  • Berlin & Kay
    • the extensions of BCTs vary enormously across languages
    • but the focal chips cluster across languages to a highly significant extent – all foci fall into 11 clusters
      • so only between two and eleven foci appear to be lexicalized in any language
    • the make-up of the set of BCTs in any given language is obeys implicational universals which can be interpreted to the effect of an evolutionary scale
      • echoing earlier proposals from the turn of 19th/20th century
      • these implicational generalizations were backed up by a library study on 78 additional languages

Figure 12. Crosslinguistic clusters of focal colors (Berlin & Kay 1969: 9)
Figure 13. Implicational generalizations and evolutionary scale in Berlin & Kay 1969

Figure 14. From Berlin &Kay 1969 to Kay 1975 (MacLaury 2001: 1228)
• the World Color Survey
  • conducted by Berlin, Kay, and William Merrifield 1976-78
    • preliminary findings in Kay, Berlin, & Merrifield 1991; Kay, Berlin, Maffi, & Merrifield 1997
  • analysis of 110 BCT systems, based on data collected from (usually) 25 speakers per language
    • using (broadly) the same procedures as Berlin & Kay 1969

• nature-nurture score, so far
  • evidence points towards constraints on color terminologies that cannot be explained in linguistic terms
    • but require reference to the vision faculty and higher cognitive processing of color
What is semantic typology (cont.)

• semantic typology: distribution

Figure 3. Green and blue terms in WALS (Kay & Maffi 2011)
What is semantic typology (cont.)

• semantic typology: generalizations

<table>
<thead>
<tr>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>[W/R/Y Bk/G/Bu]</td>
<td>[W R Y Bk/G/Bu]</td>
<td>(III Bk/G/Bu)</td>
<td>[W R Y G/Bu Bk]</td>
<td>(IV Bk/Bu)</td>
</tr>
<tr>
<td></td>
<td>[W R Y G/Bu Bk]</td>
<td>(III G/Bu)</td>
<td>[W R Y G/Bu Bk]</td>
<td>(IV G/Bu)</td>
</tr>
<tr>
<td></td>
<td>[W R Y G/Bu Bk]</td>
<td>(III Y/G/Bu)</td>
<td>[W R Y G Bu Bk]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[W R Y G Bu Bk]</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.** Stage model of implicational generalizations, covering 83% (91/110) of the languages of the World Color Survey (Kay & Maffi 1999: 748)
What is semantic typology (cont.)

- **history**
  - **phase I: explicit typological research on semantic categorization starts in the late 19th century**
    - mostly with questionnaire studies such as
      - Morgan 1871 on kinship terminology
      - Darwin 1872 on gesture
    - an early study using non-verbal stimuli: Magnus 1877, 1880 on color naming and discrimination
      - based on a kit of 10 color chips he sent to 61 field investigators
    - much of this research was marred by
      - underdeveloped methods of linguistic analysis
      - racist and social-Darwinist assumptions
  - **phase II: ethnosemantic research**
    - focusing primarily on individual languages
    - often inspired by the Linguistic Relativity Hypothesis
• phase III: resurgence of explicit typology
  • but this time with the benefit of a century of advances in linguistics, cultural anthropology, and cognitive psychology
  • starting with Berlin & Kay’s (1969) work on basic color terms
  • in the 1980s
    • Viberg 1984 on perception verbs
    • Dahl 1985 on tense-mood-aspect system
    • Talmy 1985 on lexicalization patterns in motion descriptions
What is semantic typology (cont.)

- some recent studies
    - spatial frames of reference and spatial categorization in 13 languages
  - Levinson, Meira, & L&C 2003; Khetarpal, Majid, & Regier 2009
    - semantic similarity of ‘topological’ spatial relators in 9 languages
  - Bohnemeyer, Eisenbeiß, & Narasimhan 2006
    - motion event categorization in 17 languages
  - Bohnemeyer 2007
    - argument structure of verbs of cutting and breaking in 17 languages
  - Bohnemeyer *et al.* 2007
    - motion event segmentation in 18 languages
  - Regier, Kay, & Khetarpal 2007
    - semantic similarity of color terms in the 110 languages of the WCS
  - Majid, Boster, & Bowerman 2008
    - semantic similarity of verbs of cutting and breaking in 28 languages
  - Kemp & Regier 2012 (and Levinson’s response)
    - kinship categories across languages
  - Bohnemeyer & O’Meara 2008; O’Meara 2010
    - ethnophysiography: the ethnosemantics of land and water forms
  - Majid 2010; Bohnemeyer 2010
    - body parts and meronymy
  - Majid & Levinson 2011
    - taste and other senses
Why study crosslinguistic semantic variation?

• the answer in a nutshell
  • by studying what is variable across languages in semantic representations of the world
    • we are able to discover boundary conditions on which aspects of cognition may be
      • innate, biologically determined
      • learned and culturally transmitted
Why study crosslinguistic semantic variation? (cont.)

• in effect, we are “mapping the nature-nurture divide in cognition” (Bohnemeyer 2011)

\[ \text{Figure 4. Mapping the nature-nurture divide in cognition} \]
Why study crosslinguistic semantic variation? (cont.)

• in doing so, we clarify the relation between language and non-linguistic cognition
  • and contribute to theories of the syntax-semantics interface

Figure 5. *The relation between language and nonlinguistic cognition*
Why study crosslinguistic semantic variation? (cont.)

• the rationale: sources of knowledge

![Diagram of sources of knowledge](image)

Figure 6. Sources of knowledge
Why study crosslinguistic semantic variation? (cont.)

• the rationale: the relation between variability and cultural transmission
  • the encoding of cognition in the human genome does not appear to be variable
    • there are – fairly superficial – genetic differences across human populations
    • however, there is currently no evidence suggesting that such differences affect cognition
  • it follows that crosslinguistic variation in a given domain of linguistic knowledge is evidence against innateness
Why study crosslinguistic semantic variation? (cont.)

• the rationale: the relation between variability and cultural transmission (cont.)
  • conversely, absence of crosslinguistic variation in a given domain can have a variety of explanations
    • coincidence
    • variability excluded by the fundamental “design features” of language (cf. Hockett 1963)
    • monogenesis and inheritance from the common ancestor (cf. Dunn et al 2011)
    • innateness
  • strong, exceptionless universals are rare among the languages of the world (Evans & Levinson 2009)
    • nevertheless, many general tendencies and implicational generalizations hold and call for explanations
Why study crosslinguistic semantic variation? (cont.)

• proposed versions of the “big picture”

Figure 7. The big picture according to Whorf

Figure 8. The big picture according to the innatists

Figure 9. The big picture according to Neo-Whorfians

• the goal of the “Neo-Whorfian” program
  • restore culture to its rightful place in the theory of human cognition
How can we study semantic variation?

The Nijmegen approach

- the Nijmegen approach to semantic typology
  - start from a tentative determination of parameters of variation, based on previous studies
  - construct an *etic grid*
    - a possibility space created by a few independent notional dimensions
      - in which every categorized stimulus can be located as a data point
    - e.g., a network of nuclear-family genealogical relations is used as etic grid in studies of kinship terminology
      - following a method pioneered by L. H. Morgan (1871)
    - Berlin & Kay’s (1969) seminal study of color terminologies famously used the Munsell color chart
      - a matrix of 40 hues by eight brightness values, realized in 320 color chips – cf. day 4
      - this approach was pioneered by Brown & Lenneberg 1954
The Nijmegen approach (Cont.)

- encode the “cells” of the grid exhaustively in sets of nonlinguistic stimulus items
- collect preferred descriptions and ranges of possible descriptions
  - in a typologically broadly varied sample of unrelated languages
  - with multiple speakers per language according to a standardized protocol
- try alternative elicitation procedures
  - aimed at exploring the full referential potential of language-particular expressions in the target domain
  - including referential communication tasks
- perform semantic analyses
  - to filter out pragmatically generated meaning components
    - and isolate senses or “intensions”
The Nijmegen approach (Cont.)

— use statistical techniques to analyze correlations
— formulate implicational generalizations
  — e.g., If a language has a basic color term for brown
    » then it also has basic color terms for black, white, red, green,
      yellow, and blue (Berlin & Kay 1969)
  — If a language uses observer-centered ("relative") frames in a given
    domain of spatial reference
    » then it also uses object-centered ("intrinsic") frames in the same
      domain (Pederson et al. 1998)
  — If a language has a pre- or postposition that expresses contact
    ("ON")
    » then it also has a pre- or post-position that expresses
      inclusion/containment ("IN") (Levinson & Meira 2003)

— three large-scale reference studies to date
  • Pederson et al. 1998 on spatial frames of reference
  • Levinson & Meira 2003 on "topological" spatial relations
  • Bohnemeyer et al. 2007 on the segmentation of motion events
Quantitative approaches in typology

• typology as the study of the distribution and co-occurrence
  • of linguistic features/properties
  • is inherently a quantitative project
  • it aims for quantitative generalizations

(1.1) a. All languages have feature X
b. Some languages have feature Y
c. If a language has feature X, it also has feature Y

Figure 1. WALS Feature 81A: Order of Subject, Object and Verb (Dryer 2011)
Quantitative approaches in typology (cont.)

• statistical methods are playing a growing role in typology
  • in the 1980s, Bybee and Dryer pioneered the use of inferential statistics against large language samples in syntactic typology
  • in recent years, there has been a rapid growths in typological applications of multivariate statistics
    • due to the availability of affordable computational power and widely accessible software
  • recent high-profile publications in *Science* (Atkinson 2011, using linear regression) and *Nature* (Dunn et al 2011, using phylogenetic statistics)
    • are bound to further accelerate the trend
  • see Cysouw 2007, 2011; Bickel 2010; Jaeger et al 2011 for discussion on multivariate methods in typology
  • multivariate statistics is first applied to semantic typology in Levinson & Meira 2003 (multi-dimensional scaling)
    • since then: Majid, Gullberg, van Staden, & Bowerman 2007; Majid, Boster, & Bowerman 2008; Regier, Kay, & Khetarpal (2007, 2009); Khetarpal, Majid, & Regier 2009; Bohnemeyer et al 2012; inter alia
Levinson & Meira 2003

• based on BowPed data, but analysis independent of Bowerman & Pederson ms.
  • focus on the expression of topological relations specifically in adpositions

• “orthodox assumptions”
  • topological relations are conceptual primitives
    • or straightforwardly decompose into conceptual primitives (e.g., Jackendoff 1983; Miller & Johnson-Laird 1976)
  • there is a simple homomorphism between internal cognitive representations of topological relations
    • and their linguistic expressions
      • so the former can be straightforwardly read off adpositions and case markers (e.g. Jackendoff & Landau 1992)
specifically

• it is often assumed that there are three “basic” topological place functions AT, ON, and IN
  • AT for *proximity* to a “1-dimensional” ground (point)
  • ON for *contact* to a “2-dimensional” ground (surface)
  • IN for *inclusion* in a “3-dimensional” ground (body)
  • e.g., Clark 1973, Herskovits 1986, Jackendoff & Landau 1992

<table>
<thead>
<tr>
<th>Place function</th>
<th>Ground</th>
<th>Locative (BE)</th>
<th>Goal (TO)</th>
<th>Source (FROM)</th>
<th>Route (VIA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>1D</td>
<td>at</td>
<td>to</td>
<td>from</td>
<td>via</td>
</tr>
<tr>
<td>ON</td>
<td>2D</td>
<td>on</td>
<td>onto</td>
<td>off</td>
<td>across</td>
</tr>
<tr>
<td>IN</td>
<td>3D</td>
<td>in</td>
<td>into</td>
<td>out of</td>
<td>through</td>
</tr>
</tbody>
</table>

Table 1. Place and path functions in English prepositions according to Clark 1973: 41 (adapted to the framework of Jackendoff 1983)
Levinson & Meira 2003 (cont.)

- etic grid and stimuli
  - the 71 BowPed line drawings (Bowerman & Pederson ms.)
    - designed for the crosslinguistic study of linguistic categorization in the domain of topological relations
      - most scenes fall along a constructed conceptual continuum between contact/support ("ON") and containment ("IN")
      - based on observations about the development of topological relations in Dutch, English, and Korean child language

![Continuum Diagram](image)

Figure 2. **Partial superposition-support continuum illustrated by BowPed items 1, 5, 62, 14, and 2** (Levinson & Meira 2003: 488)

- a few scenarios designed to elicit ‘projective’ relations are included as well
• **procedure**
  • for each image
    • find close cultural equivalents of figure and ground
    • negotiate culturally plausible interpretations
  • elicit in a way most appropriate to the target language contextualized locative descriptions

(3.1) A possible elicitation scenario for BowPed: “Imagine you’re talking to somebody who is looking for the [FIGURE]. This person knows where the [GROUND] is, but doesn’t know where the [FIGURE] is. You know the [FIGURE] is; but neither of you can see the [FIGURE] and the [GROUND] right now. The person asks you *Where is the [FIGURE]?* Imagine you want to tell the person where the [FIGURE] is. How do you respond?”

• **scope of analysis**
  • topological information may be expressed in the ground phrase, the locative predicate,…
  • …and in “satellites” and “serial verbs”
Levinson & Meira 2003 (cont.)

- L&M restrict their analysis to the ground phrase
- In the ground phrase, topological information can be expressed
  - by adpositions, case markers, and meronyms

(3.2) **Spatial adposition:** "A spatial adposition is any expression that heads an adverbal phrase of location in ... answers to *where*-questions... This definition is not designed to exclude *spatial nominals*, since they so often gradually develop into ‘true’ adpositions that boundary problems would plague a comparative exercise of this sort." (L&M 2003: 486)

**Figure 3.** Topological information in the predicate: some Tzeltal “dispositional” (adapted from Brown 1994)

**Figure 4.** Meronyms: from body parts via object parts to spatial regions (Levinson 2003: 106, after Heine 1997: 44)
Levinson & Meira 2003 (cont.)

• does L&M’s decision compare to Berlin & Kay’s to restrict the analysis of color terms to BCTs???
  • a potential problem: languages differ *typologically* in whether they have adpositions and case markers at all
  • and in the extent to which they use them in the ground phrase

• sample
  • 9 mutually unrelated languages all of which have spatial adpositions and at most 1 or 2 general locative cases

*Table 2. Language sample of Levinson & Meira 2003*

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>AFFILIATION</th>
<th>LOCATION</th>
<th>DEMOGRAPHY</th>
<th>CONSULTANTS</th>
<th>RESEARCHER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basque</td>
<td>Isolate</td>
<td>Europe</td>
<td>660,000</td>
<td>26</td>
<td>I. Ibarretxe</td>
</tr>
<tr>
<td>Dutch</td>
<td>Indo-European</td>
<td>Europe</td>
<td>20,000,000</td>
<td>10</td>
<td>D. Wilkins,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. de Witte</td>
</tr>
<tr>
<td>Ewe</td>
<td>Niger-Congo</td>
<td>West Africa</td>
<td>3,000,000</td>
<td>5</td>
<td>F. Ameka</td>
</tr>
<tr>
<td>Lao</td>
<td>Tai-Kadai</td>
<td>Southeast Asia</td>
<td>3,000,000</td>
<td>3</td>
<td>N. Enfield</td>
</tr>
<tr>
<td>Lavukaleve</td>
<td>Isolate</td>
<td>Solomon Islands</td>
<td>1.150</td>
<td>1</td>
<td>A. Terrill</td>
</tr>
<tr>
<td>Tiriyó</td>
<td>Cariban</td>
<td>South America</td>
<td>2,000</td>
<td>10</td>
<td>S. Meira</td>
</tr>
<tr>
<td>Trumai</td>
<td>Isolate</td>
<td>South America</td>
<td>50</td>
<td>3</td>
<td>R. Guirardello</td>
</tr>
<tr>
<td>Yéfi Dnye</td>
<td>Isolate</td>
<td>Papua New Guinea</td>
<td>3,750</td>
<td>4</td>
<td>S. Levinson</td>
</tr>
<tr>
<td>Yukatek</td>
<td>Mayan</td>
<td>Mesoamerica</td>
<td>700,000</td>
<td>5</td>
<td>J. Bohnemeyer,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C. Stolz</td>
</tr>
</tbody>
</table>
the comparability problem

- how does one compare 1 Lavukaleve speaker to 26 Basque speakers?

the fix in Levinson & Meira 2003:
“average” responses within each population

- a picture was assigned to a particular adposition if
  - Tiriyó, Basque, Dutch:
    more than half of the consultants converged on the adposition for this pic
  - languages w/ 4-5 consultants:
    two or more consultants converged on the adposition for this pic
  - languages w/ fewer than 4 consultants:
    any consultant used the adposition for this pic

- this doesn’t solve the problem that the categorization is vastly better evidenced for some languages than for others
Levinson & Meira 2003 (cont.)

- a complication
- many responses combine adpositions and “meronyms” in the ground phrase

“Some languages may, in addition to adpositions, deploy a series of spatial nominals... Here we abstract away from these additional codings ... We do not think that this is wholly legitimate ... However, the idealization is warranted as a response to the orthodox thesis outlined above, namely the claim that the closed-class adpositions yield a specific kind of abstract, universal spatial semantics.” (L&M 2003: 488)
• e.g., L&M would code (3.3) as falling into the domain of the generic preposition *ti’*
  
  • even though the description clearly does not mean the same without the meronym *pàach* ‘back’

(3.3) Te’l kul-ukbal u=pèek’-il
YUC there sit-DIS(B3) A3=dog-REL
  
tu=pàach le=nah=o’
PREP:A3=back DET=house=D2
  
‘There the dog is sitting outside the house’

• their justification: the outcome – the failure to confirm the “orthodox assumptions”
  
  • couldn’t have been remedied had adposition-meronym combinations been treated separately
  
  • this is probably true: L&M’s analysis artificially *inflates* the extension of adpositional categories
    
    • and yet they still turn out to be *narrower* than predicted on the “orthodox assumptions”
  
  • still...
patterns in the data

unsurprisingly, there is a great deal of variation

in the extensions of the adpositional categories

e.g., there is no single 2D arrangement of the 71 pix

that allows to represent all categories as contiguous areas

but how similar are the categories?

i.e., how likely are any two pix to fall in the same category across languages?

towards an answer: multivariate statistics

techniques of quantitative data analysis that permit the discovery of clusters in large data sets – see appendix

Figure 7. Extensional map of adpositional categories in 4 of the languages on 2D “minimal stress” grid (L&M p.500-501)
Levinson & Meira 2003 (cont.)

• L&M use *Multi-Dimensional Scaling* (MDS)
  • an algorithm which translates a *similarity matrix* into a spatial model by interpreting similarity as distance
    • the similarity matrix assigns to each pair of stimulus items their relative similarity in the data set

• background: cluster analysis and MDS
  • two families of techniques specifically designed to analyze similarity data
    • some examples
      • participants rank a set of pictures of facial expressions on a scale of friendliness
        • or simply perform a pile sort on them into an arbitrary number of groups on the basis of similarity
      • participants rank their preferences for a set of goods
        • or a set of ads produced to sell them
• in either case, as the researcher consolidates the responses
  • the question is, what are the criteria underlying the choices?
  • the researcher seeks to recover these criteria by trying to determine what the choices have in common

• the rationale
  • compare every data point to every other data point
  • find some way of ranking or weighting the relative similarity of each pair of data points
    • i.e., the degree to which the participants’ choice was the same in both cases
  • construct a similarity matrix that assigns to each pair of data points their relative (dis)similarity
  • apply an algorithm which translates the similarity matrix into a spatial model
    • on the basis of interpreting dissimilarity as distance
  • interpret that model
Levinson & Meira 2003 (cont.)

- **the output: MDS**
  - MDS interprets dissimilarity between data points as spatial distance (typically on a 2D plot)
    - there are various measures in use, but the most common measure is Euclidean distance
      - distance\((x, y) = \sqrt{\sum_i (x_i - y_i)^2}\) where \(x_i\) and \(y_i\) are the coordinates of \(x\) and \(Y\) in dimension \(i\)
  - a neat comparison, courtesy of Jim Boster
    - imagine the task of constructing a map of a region
      - without any information of where the towns are
      - but with complete information about the relative distances between the towns
    - the result is a map that by necessity gets all the relative locations of the towns right
      - but where the absolute locations are still off by a factor determined in terms of rotation of the map as a whole
  - the desirable number of dimensions of the model is assessed in much the same way as the number of factors
    - in a factor analysis
Levinson & Meira 2003 (cont.)

- **calculating the similarity matrix**
  - first a matrix is calculated for each language
  - every pair of pix was assigned a dissimilarity value b/w 0 (perfectly similar) and 1 (perfectly dissimilar)
    - every pix is perfectly similar to itself
    - the value for pix i and j is calculated according to (3.4)

\[
D_{ij} = \frac{N_{\text{total adpositions}} - N_{\text{adpositions shared b/w i & j}}}{N_{\text{total adpositions}}}
\]

- so similarity increases with the number of shared adpositions by a factor that depends on the size of the overall adposition set
- pix that don’t share any adposition are maximally dissimilar

- then a composite matrix was created for the sample
  - by summing the D coefficients for each pair of pix across the nine languages

- **creating the plot**
  - the composite matrix was fed into the ALSCAL algorithm of SPSS 7.5
  - a plot was obtained interpreting the D coefficients as Euclidian distances
Levinson & Meira 2003 (cont.)

• the result
  • there’s an IN cluster as predicted
    • but many containment/inclusion items are not in it
      • but instead wind up on the left margin together with protrusion and circumference configuration
  • there are up to three ON clusters
    • an ON-TOP cluster which contains support items
    • an ON/OVER cluster which contains support and superposition items
    • and a large ATTACHMENT cluster
  • there is a NEAR/UNDER cluster
    • which doesn’t clearly correspond to any of the hypothetical primitive topological relations AT, ON, and IN

Figure 8. ALSCAL plot based on the composite dissimilarity matrix of the nine languages of the sample (L&M p.505)
Levinson & Meira 2003 (cont.)

• interpretation – the best crosslinguistic predictors of linguistic categorization in the topological domain
  • are not the hypothetical primitives AT, ON, and IN
  • but are more fine-grained notions such as support, superposition, attachment, etc.
    • which (a) typically do not map one-to-one into adpositional meanings and (b) arguably aren’t purely spatial notions

• implicational generalizations over adposition systems

![Diagram](image)

**Figure 9.** One-dimensional implicational scale à la Berlin & Kay 1969 (L&M p.510)

But Spanish *en* corresponds to Stage-I “AT”, yet coexists with all sorts more specific prepositions!

**Figure 10.** Evolutionary stages à la Kay 1975 (L&M p.512)
• the stimulus artifact problem
  • each point in the MDS plot represents a BowPed pic
  • the distance of each point from each other point
    • represents the similarity of the two items in how they are categorized across the sample
    • with some “stress” resulting from forcing the whole model into two dimensions
• these similarities are a function of the language-specific categorizations
  • but they are also simply a function of the content of the pictures
    • the density of the clusters is a function of how frequently those particular types of scenes occur in the set
    • the relative isolation of other scenes is the result of the particular type occurring only once or twice in the set
• the problem: BowPed was not designed for this
  • there was no attempt to ensure
    • that every logically possible type of topological relation on some etic grid is represented with equal frequency in the set
• how might this affect the analysis?
  • the frequency of a particular type of scene has no effect on distances
    • as long as it is above 1 – at which point the issue is no longer one of the composition of the stimulus set
      • but rather becomes a question of the etic grid
  • to the extent that MDS plots are interpreted visually, cluster density may make a difference
    • the problem may compound that of outliers to be addressed in the next section
• an example from a very different style of analysis
  • the frequency of meronyms (“RSN”)
    • in Yucatec, but not Tseltal, spatial descriptions
      • happens to depend on whether the scene features containment
      • again, the BowPed scenes were not designed for this

• the general point
  • ST stimulus sets are generally not designed according to the standards of quantitative research
    • but it seems to me that they probably should be

Figure 11. Yucatec BowPed responses ground phrase type frequencies by predicate construction type (Bohnemeyer & Brown 2007: 1138)
Majid et al 2008

- Bohnemeyer, Bowerman, & Brown 2001
  - 61 short digital video clips
  - featuring C&B scenes varied in terms of
    - presence of a discernible cause
    - type of theme (fabric, rope, carrots, sticks, ...)
    - type of instrument used (bare hands, hammer, scissors, saw, ...)
    - manner of action (controlled vs. “frenzied”)
    - degree of change (complete vs. partial)

Figure 12. Still pix from four C&B clips
• plus, some clips featured events of opening objects
  • to see whether these are ever described with the same verb as any of the C&B scenes

• protocol
  • participants watched each clip several times
  • then answered two questions asked in their native language
    a) “What did the [actor] do in this clip?”
      • if appropriate, i.e., with the exception of “spontaneous breaking” clips
    b) “What happened to the [theme] in this clip?”

• further elicitation
  • if still necessary, the applicability of three types of descriptions was subsequently tested:
    • active transitive, intransitive activity, and intransitive state change descriptions
• aims
  • study universals and crosslinguistic variation in lexicalization and a-structure classes
  • examine the acquisition of language-specific a-structure patterns

• the sample
  • adult language C&B data has been collected from speakers of about 30 languages so far
    • cf., e.g., Majid & Bowerman (eds.), Bohnemeyer 2007, Majid, van Staden, Boster, & Bowerman (ms.)
Table 4. Language sample used in Majid et al. 2008 (asterisks indicate systems whose semantics was extensively discussed in contributions to Bowerman & Majid eds. 2007)

<table>
<thead>
<tr>
<th>Language*</th>
<th>Language affiliation</th>
<th>Country</th>
<th>Researcher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biak</td>
<td>Austroasiatic</td>
<td>Indonesia</td>
<td>W. van de Heuvel</td>
</tr>
<tr>
<td>Chontal*</td>
<td>isolate</td>
<td>Mexico</td>
<td>L. O’Connor</td>
</tr>
<tr>
<td>Dutch*</td>
<td>Indo-European</td>
<td>Netherlands</td>
<td>M. van Staden</td>
</tr>
<tr>
<td>English*</td>
<td>Indo-European</td>
<td>UK, USA</td>
<td>M. Bowerman, A. Majid, C. Wortmann</td>
</tr>
<tr>
<td>Ewe*</td>
<td>Niger-Congo</td>
<td>Ghana</td>
<td>F. Ameke</td>
</tr>
<tr>
<td>German*</td>
<td>Indo-European</td>
<td>Germany</td>
<td>M. van Staden</td>
</tr>
<tr>
<td>Hindi*</td>
<td>Indo-European</td>
<td>India</td>
<td>B. Narasimhan</td>
</tr>
<tr>
<td>Jalonke*</td>
<td>Niger-Congo</td>
<td>Guinea</td>
<td>F. Lüpke</td>
</tr>
<tr>
<td>Japanese</td>
<td>isolate</td>
<td>Japan</td>
<td>S. Kita</td>
</tr>
<tr>
<td>Kilivila</td>
<td>Austroasiatic</td>
<td>Papua New Guinea</td>
<td>G. Senft</td>
</tr>
<tr>
<td>Kuuk Thaayorre*</td>
<td>Pama-Nyungan</td>
<td>Australia</td>
<td>A. Gaby</td>
</tr>
<tr>
<td>Lao*</td>
<td>Tai</td>
<td>Laos</td>
<td>N. Enfield</td>
</tr>
<tr>
<td>Lijkoe</td>
<td>Niger-Congo</td>
<td>Ghana</td>
<td>F. Ameke</td>
</tr>
<tr>
<td>Mandarin*</td>
<td>Sino-Tibetan</td>
<td>China</td>
<td>J. Chen</td>
</tr>
<tr>
<td>Minaña</td>
<td>Witoroan</td>
<td>Colombia</td>
<td>F. Seifart</td>
</tr>
<tr>
<td>Otoni*</td>
<td>Otoranguean</td>
<td>Mexico</td>
<td>E. Palancar</td>
</tr>
<tr>
<td>Punjabi</td>
<td>Indo-European</td>
<td>Pakistan</td>
<td>A. Majid</td>
</tr>
<tr>
<td>Spanish</td>
<td>Indo-European</td>
<td>Spain, Mexico</td>
<td>M. Bowerman, E. Palancar</td>
</tr>
<tr>
<td>Sranan*</td>
<td>Creole</td>
<td>Surinam</td>
<td>J. Essebey</td>
</tr>
<tr>
<td>Swedish*</td>
<td>Indo-European</td>
<td>Sweden</td>
<td>M. Gullberg</td>
</tr>
<tr>
<td>Tamil*</td>
<td>Dravidan</td>
<td>India</td>
<td>B. Narasimhan</td>
</tr>
<tr>
<td>Tidore*</td>
<td>West Papuan Phylum</td>
<td>Indonesia</td>
<td>M. van Staden</td>
</tr>
<tr>
<td>Tirió*</td>
<td>Cariban</td>
<td>Brazil</td>
<td>S. Meira</td>
</tr>
<tr>
<td>Touro</td>
<td>Papuan isolate</td>
<td>Solomon Islands</td>
<td>M. Dunn, A. Terrill</td>
</tr>
<tr>
<td>Turkish</td>
<td>Altai</td>
<td>Turkey</td>
<td>A. Özyürek</td>
</tr>
<tr>
<td>Tzeltal*</td>
<td>Mayan</td>
<td>Mexico</td>
<td>P. Brown</td>
</tr>
<tr>
<td>Yélî Dnye*</td>
<td>Papuan isolate</td>
<td>Papua New Guinea</td>
<td>S. Levinson</td>
</tr>
<tr>
<td>Yukatek</td>
<td>Mayan</td>
<td>Mexico</td>
<td>J. Bohnemeyer</td>
</tr>
</tbody>
</table>
• **coding and analysis**
  • the question MBB ask: *How much similarity is there across the sample in the categorization of the scenes?*
  • the basic idea behind the answer – same as in Levinson & Meira 2003
    • although MBB’s implementation differs in important respects
    • calculate for each pair of stimulus items how often these wind up in the same linguistic category
    • treat this as a measure of how (dis)similar their categorization is across the languages of the sample

• **coding the categories**
  • the target categories are those expressed by *verbs*
  • only verbs counted that described
    • “the change in an object from a state of integrity to a state of separation or material destruction” (p. 238)
Majid et al 2008 (cont.)

- a conundrum: what to count as one verb?
  - consider English resultatives and particle verbs
    - are smash and smash to pieces one verb or two?
    - what about break, break in half, break in two, break up...?
    - or should these be treated as different but related verbs?
    - in actual fact, for coding, we tried to treat different complex predicates involving the same roots as different “verbs”
      - however, it was ultimately left to each researcher to decide what counted as a complex predicate in “their” language
  - the coding was done in such a fashion that the analysis could be conducted both on the complex predicates and on the roots
  - the analysis MBB present was performed on the roots
    - so break, break in half, break in two, etc., are all treated as the same verb by this analysis
Majid et al 2008 (cont.)

- **basic similarity scores**
  - for each language, a given pair of scenes was assigned the score 1 if they were described w/ the same “verb”
    - by any speaker of that language
    - and 0 otherwise
  - so the basic similarity score is not affected by how many speakers used the same “verb” for that pair
    - since that in turn would be influenced by the number of speakers recorded for each language
  - in this way, MBB avoid the problem of Levinson & Meira’s (L&M’s) averages
    - but OTOH they fail to distinguish a case in which 100% of speakers of a language treat two scenes under the same “verb”
      - from one where only 5% of the speakers of a language group these two scenes together linguistically
Majid et al 2008 (cont.)

• composite scores
  • MBB computed similarity matrices for each individual language and a composite matrix for the entire sample
  • MBB do not discuss how they computed the cumulative scores for the composite matrix
  • I’m assuming they just added the individual language scores up for each pair of pix

• another factor accounted for in L&M, but apparently not in MBB
  • the number of different categories in the same language that group a pair of scenes together
Majid et al 2008 (cont.)

• multivariate statistics
  • L&M used multi-dimensional scaling (MDS)
  • MBB use three consecutive analyses
    • correspondence analysis, factor analysis, and cluster analysis

• correspondence analysis
  • suppose all stimuli are exactly equidistant from one another in similarity space
    • the languages in the sample would be maximally dissimilar from one another
      • no categorization in one language would be predictive of the categorization in any other language
  • in this case, an (n-1)-dimensional space would be needed to model the similarity matrix for n stimuli
• a 2-D space is needed to represent three equidistant points (an equilateral triangle)
  • a 3-D space to represent four equidistant points (an equilateral tetrahedron)
  • etc.
• OTOH if all stimuli fall into three clusters across the sample, it takes only a 2-D space to model the matrix
• so the number of dimensions needed to create a spatial model of the similarity matrix is a measure
  • of the overall similarity in how the stimuli are categorized across languages
  • the fewer dimensions, the simpler the spatial model, the simpler the matrix, the more uniform the categorization
  • conversely, the less uniform the categorization, the more complex the matrix, the more dimensions needed to model
Majid et al 2008 (cont.)

• results: Figure 13 shows the *eigenvalue* of each dimension
  • the percentage of total variance it accounts for
    • divided by the number of variables = stimulus items
  • the first seven dimensions account for 62%
    • of the variance
  • “the eigenvalues of the remaining dimensions
    • form a scree-slope
    • indicating uninterpretable noise”
Majid et al 2008 (cont.)

• **MBB’s interpretation of the first two dimensions**
  • the first dimension distinguishes reversible from non-reversible events
    • reversible are clips featuring various kinds of opening events rather than cutting/breaking
  • the second dimension distinguishes one clip from all others
    • that clip features pushing away a chair, likewise a reversible event

• **second correspondence analysis**
  • this time focusing only on the clips featuring irreversible events
  • see Figure 14 – now the first four dimensions account for 47% of the variance
Majid et al 2008 (cont.)

- interpretations of the dimensions of the second analysis
  - Figure 4 shows a plot of the clips in a space projected from the first and third dimensions

**Figure 14.** Correspondence analysis of MBB’s C&B data – clips featuring irreversible events only (p. 241)

**Figure 15.** Plot of the first against the third dimension of the similarity space of the “irreversible” clips (p. 241)
Majid et al 2008 (cont.)

- MBB argue that the first dimension represents the “predictability of the locus of separation”
  - *hmmmm...
- the second dimension separates two clips from the rest
  - both of these feature tearing of fabric
- the third dimension separates clips into a “snapping” and a “smashing” cluster
- the fourth dimension singles out one clip that shows an event of punching a hole into a piece of cloth
- MBB’s interim conclusion

“Although the precise categories recognized by the languages in our sample differ, they are highly constrained by the four dimensions we have described. These dimensions delineate a semantic space in which the categories recognized by individual languages, as variable as they are, encompass adjacent clips.” (243)

- correlations of these cross-sample dimensions with dimensions extracted for each language – see Table 5
  - unclear how they computed these
the first two dimensions appear to be highly relevant in most languages
  • although they are not nearly as relevant in some

Table 5. Correlations of MBB’s cross-sample dimensions with dimensions extracted for each language individually (243)

<table>
<thead>
<tr>
<th>Language</th>
<th>Dimension 1</th>
<th>Dimension 2</th>
<th>Dimension 3</th>
<th>Dimension 4</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bislam</td>
<td>.90</td>
<td>.75</td>
<td>.89</td>
<td>.81</td>
<td>.84 (.07)</td>
</tr>
<tr>
<td>Chontal</td>
<td>.81</td>
<td>.92</td>
<td>.04</td>
<td>.19</td>
<td>.47 (.47)</td>
</tr>
<tr>
<td>Dutch</td>
<td>.78</td>
<td>.88</td>
<td>.64</td>
<td>.55</td>
<td>.71 (.15)</td>
</tr>
<tr>
<td>English</td>
<td>.93</td>
<td>.93</td>
<td>.48</td>
<td>.63</td>
<td>.74 (.23)</td>
</tr>
<tr>
<td>Ewe</td>
<td>.80</td>
<td>.70</td>
<td>.89</td>
<td>.70</td>
<td>.77 (.09)</td>
</tr>
<tr>
<td>German</td>
<td>.82</td>
<td>.77</td>
<td>.77</td>
<td>.51</td>
<td>.72 (.14)</td>
</tr>
<tr>
<td>Hindi</td>
<td>.80</td>
<td>.57</td>
<td>.20</td>
<td>.59</td>
<td>.56 (.28)</td>
</tr>
<tr>
<td>Japonke</td>
<td>.88</td>
<td>.81</td>
<td>.87</td>
<td>.83</td>
<td>.85 (.03)</td>
</tr>
<tr>
<td>Japanese</td>
<td>.90</td>
<td>.90</td>
<td>.53</td>
<td>.81</td>
<td>.71 (.17)</td>
</tr>
<tr>
<td>Kihivila</td>
<td>.94</td>
<td>.71</td>
<td>.95</td>
<td>.91</td>
<td>.88 (.11)</td>
</tr>
<tr>
<td>Kuuk Thaayarre</td>
<td>.93</td>
<td>.95</td>
<td>.76</td>
<td>.89</td>
<td>.88 (.09)</td>
</tr>
<tr>
<td>Lao</td>
<td>.60</td>
<td>.53</td>
<td>.73</td>
<td>.48</td>
<td>.59 (.11)</td>
</tr>
<tr>
<td>Likpe</td>
<td>.77</td>
<td>.45</td>
<td>.85</td>
<td>.78</td>
<td>.71 (.18)</td>
</tr>
<tr>
<td>Mandarin</td>
<td>.79</td>
<td>.75</td>
<td>.79</td>
<td>.81</td>
<td>.79 (.03)</td>
</tr>
<tr>
<td>Mirinya</td>
<td>.83</td>
<td>.67</td>
<td>.57</td>
<td>.15</td>
<td>.56 (.29)</td>
</tr>
<tr>
<td>Otori</td>
<td>.93</td>
<td>.68</td>
<td>.91</td>
<td>.66</td>
<td>.80 (.14)</td>
</tr>
<tr>
<td>Punjabi</td>
<td>.86</td>
<td>.94</td>
<td>.04</td>
<td>.21</td>
<td>.52 (.44)</td>
</tr>
<tr>
<td>Spanish</td>
<td>.79</td>
<td>.61</td>
<td>.60</td>
<td>.36</td>
<td>.59 (.18)</td>
</tr>
<tr>
<td>Sranan</td>
<td>.85</td>
<td>.95</td>
<td>.04</td>
<td>.74</td>
<td>.65 (.41)</td>
</tr>
<tr>
<td>Swedish</td>
<td>.89</td>
<td>.91</td>
<td>.70</td>
<td>.73</td>
<td>.81 (.11)</td>
</tr>
<tr>
<td>Tamil</td>
<td>.82</td>
<td>.92</td>
<td>.19</td>
<td>.62</td>
<td>.64 (.32)</td>
</tr>
<tr>
<td>Tidore</td>
<td>.75</td>
<td>.94</td>
<td>.82</td>
<td>.81</td>
<td>.83 (.08)</td>
</tr>
<tr>
<td>Tiriyo</td>
<td>.83</td>
<td>.97</td>
<td>.88</td>
<td>.74</td>
<td>.86 (.10)</td>
</tr>
<tr>
<td>Touou</td>
<td>.83</td>
<td>.87</td>
<td>.70</td>
<td>.50</td>
<td>.73 (.17)</td>
</tr>
<tr>
<td>Turkish</td>
<td>.88</td>
<td>.62</td>
<td>.43</td>
<td>.14</td>
<td>.52 (.31)</td>
</tr>
<tr>
<td>Tzeltal</td>
<td>.84</td>
<td>.88</td>
<td>.89</td>
<td>.74</td>
<td>.84 (.07)</td>
</tr>
<tr>
<td>Yéli Dnye</td>
<td>.81</td>
<td>.62</td>
<td>.20</td>
<td>.68</td>
<td>.56 (.26)</td>
</tr>
<tr>
<td>Yurake</td>
<td>.93</td>
<td>.83</td>
<td>.90</td>
<td>.81</td>
<td>.87 (.06)</td>
</tr>
<tr>
<td>Mean</td>
<td>.83</td>
<td>.79</td>
<td>.62</td>
<td>.62</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.09</td>
<td>.15</td>
<td>.30</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>.60</td>
<td>.45</td>
<td>-.04</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>.94</td>
<td>.97</td>
<td>.95</td>
<td>.91</td>
<td></td>
</tr>
</tbody>
</table>
• **factor analysis**  
  a la Romney, Weller, and Batchelder (1986)
  • this permits a comparison of the 28 individual-language matrices in their entirety
    • whereas the dimensions produced by the correspondence analysis only account for about half of the variance
  • method: correlate the individual matrices, then run a factor analysis over the resulting correlation matrix
  • two FA analyses: principal component analysis (PCA) and minimal residual factor analysis (MRFA)
  • languages that categorize similarly should correlate positively
    • those that do not should correlate negatively

*Figure 16. Correlation of first factor scores and the aggregate for individual languages in PCA and MRFA*  
(p. 244)
Majid et al 2008 (cont.)

- all languages do correlate positively
- Figure 17: plot of the first two factors
  - this is a similarity space - the distance between two languages now represents the (dis)similarity of their categorizations
- MBB interpret the first two dimensions as showing “agreement” and “differentiation” respectively

Figure 17. Factors 1 and 2 from the PCA model: Agreement and differentiation among languages in their overall classification of “cutting and breaking” events. (244)
discussion: the aggregation problem

on aggregating extensional data from different languages to assess similarity of categorization

extension underspecifies sense

(4.1) a. *The glass is half full*
b. *The glass is half empty*

(4.2) a. *The cat is above the aquarium*
b. *The aquarium is below the cat*

(4.3) a. *Kryten is a mechanoid*
b. *Kryten is a mechanoid, and he’s not an anteater*
• the categorizations of the stimuli semantic typologists collect are extensional
  • they simply group the stimuli into sets without regard for the criteria = meanings under which they are grouped
• it is, however, the senses of the categorizations semantic typologists are really after
  • semantic typology is an exploration of the mind
  • nobody claims people represent (“generative”) categories in terms of lists of members
The aggregation problem (cont.)

- extensional categorization may underspecify sense contrasts in crosslinguistic research
  - example: Levinson 2007 – Yélî Dnye extensionally *more or less* distinguishes CUT-type from BREAK-type scenes
  - but in terms of its sense (its conceptual content) construes this contrast in completely different terms

---

**Figure 18.** Intensions of Yélî C&B verbs (left) with some illustrations (Levinson 2007: 210-211)
The aggregation problem (cont.)

- extensionally, this distinction only registers in very few scenes
- these effects may be too small to move the needle
  - when lumped together with extensional data from other languages

**Figure 19.** Plot of dimensions 1 and 3 of correspondence analysis of C&B verbs (Majid, Boster, & Bowerman 2008: 241)

**Figure 20.** Similarity of languages in their C&B classifications (Principal Component Analysis, dimensions 1-2) (MBB p.244)
The aggregation problem (cont.)

• when one aggregates category extensions across languages
  • one loses the one thing that “makes sense” of them conceptually
    • their relations to the other categories of the same language
Regier et al. 2007

- one alternative: assessing the “naturalness” of extensional categorizations intra-linguistically
  - by comparing them to categorizations of equal complexity that are maximally “natural”
    - in the sense that they maximize within-category similarity and minimize across-category dissimilarity
    - this has been attempted in Regier, Kay, & Khetarpal (2007, 2009) and Khetarpal, Majid, & Regier 2009

![Figure 21. Comparing the predicted “optimal” partitioning of CIELAB color space across 3, 4, 5, and 6 categories (left) against actual color term systems of the World Color Survey (right) (Regier, Kay, & Khetarpal 2007a: 1)](image)

![Figure 21. Naturalness of Berinmo color terms in terms of partitioning CIELAB, compared to 19 variants obtained by rotating the Berinmo system by 1, 2, etc. hue columns of the Munsell chart (Regier, Kay, & Khetarpal 2007b: 1440)](image)
The aggregation problem (cont.)

- **the rub**: this presupposes a language-independent standard of comparison
  - some measure of similarity that allows one to determine what a maximally “natural” system would look like
    - in Regier, Kay, & Khetarpal (2007, 2009): the optimal partitioning of the CIELAB color space with n categories
      - the distance between two points/colors in CIELAB is presumed to represent their psychological (dis)similarity
    - in Khetarpal, Majid, & Regier 2009: the output of a pile sort of the BowPed pix by native speakers of Dutch and English
  - the problem: semantic typologists can’t just assume that there *are* universally valid “naturalness” standards
  - so now we need to go and test conceptual categorization for each language community
    - then assess how optimally that culture-specific similarity space is partitioned in the language
    - and finally compare the results across populations

➢ a new Whorfian project!

**Figure 22.** The chips of the World Color Survey stimuli plotted in CIELAB space (Regier, Kay, & Khetarpal 2007: 1440)
Enter MesoSpace

- MesoSpace (I-Ib) aims to contribute to this debate from two angles
  - we are investigating a possible lexico-syntactic factor that may bias speakers against relative FoRs
    - namely the productive use of shape-based meronyms in the representation of space
  - in addition, we are working on a series of studies that pit linguistic against non-linguistic predictors
    - in reference frame use across languages
  - the following is a presentation of preliminary findings from the second line of inquiry
Enter MesoSpace (cont.)

• our research questions
  – to what extent is FoR use in a given community correlated with
    • features of the local topography
    • properties of population geography (density, urbanization, infrastructure)
    • level of education
    • literacy
  – is there a correlation between (first and second) language and FoR use
    – that cannot be reduced to come combination of the above non-linguistic variables?
Enter MesoSpace (cont.)

- our tool for studying the use of FoRs in discourse
- a referential communication task: Ball & Chair (B&C)
  - replacing Men & Tree (M&T) in Pederson et al (1998) etc.
  - B&C allows us to discover selection preferences for any of the FoR types
    - at the in-door scale
    - M&T may for various reasons depress the use of intrinsic FoRs

Figure 23. Design of the Men and Tree task (Pederson et al. 1998: 562)

Figure 24. Two of the Ball & Chair fotos, featuring an intrinsic contrast
Enter MesoSpace (cont.)

• recall memory test: New Animals
  – an almost identical replication of the Animals In A Row (AIAR) design
  • Of Levinson 1996 and Pederson et al. 1998

  step I: memorize array  
  step II: 180° rotation and move to recall table  
  step III: reconstruction of the stimulus array

Figure 25. Animals-in-a-Row: design

  – differences: the toy animals used; the number of trials; ...
  – disadvantage: no distinct intrinsic response type
Enter MesoSpace (cont.)

• the pilot study presented in Bohnemeyer et al 2012
  • an analysis of the B&C data from 11 varieties
    • 6 Mesoamerican languages
      • Yucatec Maya (J. Bohnemeyer)
      • Ayutla Mixe (R. Romero)
      • San Ildefonso Tultepec Otomí (N. Hernández, S. Hernández, E. Palancar)
      • Purépecha (or Tarascan; A. Capistrán)
      • Chacoma Tseltal (G. Polian)
      • Juchitán (Isthmus) Zapotec (G. Pérez)
    • 2 non-Mesoamerican indigenous languages
      • Seri (C. O’Meara)
      • Sumu-Mayangna (E. Benedicto, A. Eggleston, Mayangna Yulbarangyang Balna)
    • 3 varieties of Spanish
      • from Barcelona (A. Eggleston), Mexico (R. Romero), and Nicaragua (A. Eggleston)
Enter MesoSpace (cont.)

• based on five dyads of participants per variety
  – except for the case of Mexican Spanish, where up to now
    only the data from three of the five dyads have been coded

– accompanied by the participants’ demographic data
  • age
  • sex
  • level of education
  • frequency of use of Spanish (as first or second language)
  • frequency of reading and writing
Enter MesoSpace (cont.)

• predictions for the use of FoRs

• the innatist camp predicts that
  • literacy and education level are the most powerful factors ✔
  • the native language and the use of Spanish play no independent role ❌

• the neo-whorfean camp predicts that
  • the native language and the use of Spanish are powerful irreducible factors ✔
  • literacy level of education merely play secondary roles ❌
Enter MesoSpace (cont.)

• a phylogenetic analysis: *Neighbor-net*
  – we restrict ourselves to games 2 and 4 of B&C
  – for every participate and every type of frame
    • we calculated the number of photos described using the particular frame type
  – we classified the frames or reference strategies
    • distinguishing among eight categories
      – intrinsic object-centered
      – intrinsic ‘direct’, egocentric
      – relative
      – ambiguous between intrinsic and relative
      – geocentric
      – vertical absolute
      – ambiguous between vertical absolute and vertical intrinsic
      – topological (no FoR)
• for each participant, we calculated a set of eight frequencies
• these sets can be interpreted as points in an octodimensional space
• the distances between the points represent the similarity across the participants’ responses
• we can use this similarity measure to test a simplified version of the predictions
  • the innatist camp: similarity in literacy and education predicts similarity in FoR use
  • the neo-whorfian camp: similarity in first/second language predicts similarity in FoR use
• we calculated the distances in the so-called “Manhattan” metric
  • where the distance between two points is the sum of the differences of the coordinates
• we entered the distance matrix representing the similarity b/w the participants in the Splitstree4 program
  • cf Huson & Bryant 2006
• we applied the Neighbor-net algorithm
  • cf Bryant & Moulton 2004
• Neighbor-net represents the distance b/w two points
  • as a set of alternative routes between two terminales nodes of the graph
    • permitting a two-dimensional representation of the distances with minimal distortion
Enter MesoSpace (cont.)

- Neighbor-nets has bee used in syntactic typology (Bickel 2010; Cysouw 2007)
  - but as far as we know ours is the first application in semantic typology
    - and the first based on experimental data
- ours is also the first application of multivariate statistics to the analysis of inter-participant variation
  - by treating language as a predictor variable
    - we are avoiding the problem of how to aggregate the data from the speakers of each language
  - in principle, this approach can be applied to the data of Levinson & Meira 2003 and Majid et al. 2008 as well
    - here, too, it makes sense to ask to what extent participants cluster by language in their linguistic categorization of the stimuli
    - a reanalysis is currently in preparation
Enter MesoSpace (cont.)

- results

Figure 26. The Neighbor-net
Enter MesoSpace (cont.)

- the resulting Neighbor-net has two salient poles
  - one pole characterized by high relative values, the other by high geocentric values

**Figure 27. The Neighbor-net and its “geography”**
Enter MesoSpace (cont.)

- this interpretation is confirmed by an MDS analysis of the same similarity matrix

Figure 28. MDS plot
Enter MesoSpace (cont.)

- a strong correlation emerges b/w the dimension of the MDS plot and the use of geocentric frames
  - Spearman’s Rho 0.9545343
  - and weaker negative correlation between the first dimension and the use of negative frames
    - Spearman’s Rho -0.7995417
- the second dimension shows a weak correlation with the frequency of topological descriptions
  - Spearman’s Rho 0.791383

![Figure 29. Correlations between the dimensions of the MDS plot and the frequency of geocentric (left), relative (center), and topological (right) descriptions.](image)
Enter MesoSpace (cont.)

- the participants cluster strongly by their native language
  - all speakers of the three varieties of Spanish are found in the relative zone with only one exception
  - the geocentric zone is inhabited by speakers of indigenous language, above all Mesoamerican ones, with one exception

Figure 30. The Neighbor-net with the terminal nodes (speakers) color-coded for group of native language
Enter MesoSpace (cont.)

- A direct analysis confirms this distribution

**Figure 31.** Mean frequency of relative (blue) vs. geocentric use by native language group (left: Spanish; center: Mesoamerican; right: non-Mesoamerican indigenous)
Enter MesoSpace (cont.)

• the second language also appears to play a strong role
  • in the relative zone, there is no participant who does not speak Spanish
  • and there are only two who do not speak it frequently
  • however, no clear tendency emerges in the geocentric zone

Figure 32. The Neighbor-net color-coded for the use of Spanish as first or second language
• again, a direct analysis confirms the distribution

Figure 33. Mean frequency of relative (blue) vs. geocentric frames by frequency of use of Spanish (left: none; center: infrequent; right: frequent)
Enter MesoSpace (cont.)

- the level of education also is a predictor of FoR use
  - among the participants with high relative values, all but two have formal education
    - there is no clear tendency in the geocentric zone

Figure 34. The Neighbor-net color-coded by level of education (blue: no formal education; green: primary education only; red: secondary or higher)
Enter MesoSpace (cont.)

- the same tendency is apparent from a direct analysis

Use of geocentric vs. relative frames by level of education

![Bar chart showing mean frequency of relative (blue) vs. geocentric frames by level of education.]

Figure 35. Mean frequency of relative (blue) vs. geocentric frames by level of education (left: no formal education; center: primary only; right: secondary and higher)
literacy likewise is a predictor

among the participants with high relative values, all but two are literate

- the illiterate participants are the same two who lack formal education
- again no salient tendency emerges in the zona geocentric zone

Figure 36. The Neighbor-net color-coded by level of education
(blue: reads/writes daily; green: reads/writes infrequently; red: does not read or write)
Enter MesoSpace (cont.)

• the same tendency emerges from a direct analysis

Use of relative vs. geocentric frames by frequency of reading and writing

Figure 37. Mean frequency of relative (blue) vs. geocentric frames by level of education (left: does not read or write; center: reads/writes infrequently; right: blue: reads/writes daily)
Enter MesoSpace (cont.)

• however, these factors are interdependent
  • being a speaker of an indigenous language continues to be weakly correlated with lower levels of education and literacy
    • due to the persistent socioeconomic marginalization
    • this holds for many of the populations in our sample – but not for all!
  • counterexamples: speakers of Juchitan Zapotec and Mayangna

**Figure 38.** Education level (blue: no formal; red: primary only) by language group (left: Spanish; center: Mesoamerican)

**Figure 39.** Literacy (blue: no reading/writing; red: infrequently) by language group (left: Spanish; center: Mesoamerican)

• these interdependencies are the very factors motivating a multivariate analysis like Neighbor-net in the first place!
• which factors are more powerful?
  • we calculated for each value of each variable the sum of the distances among the participants with that value
    • meaning the sum of the distances in the similarity matrix
  • the mean distance is a rough measure of the density of the cluster caused by the particular value
  • it turns out that the values with the densest clusters belong to the linguistic variables
    • the absence of formal education is the strongest non-linguistic predictor
      • but it comes in only in fourth place

<table>
<thead>
<tr>
<th>innatist predictions</th>
<th>✔</th>
</tr>
</thead>
<tbody>
<tr>
<td>neo-Whorfean predictions</td>
<td>✓</td>
</tr>
</tbody>
</table>
Enter MesoSpace (cont.)

<table>
<thead>
<tr>
<th>variable</th>
<th>value</th>
<th>mean distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>&lt; 30</td>
<td>35.8421052631579</td>
</tr>
<tr>
<td></td>
<td>≥ 30</td>
<td>39.1714285714286</td>
</tr>
<tr>
<td>sex</td>
<td>male</td>
<td>35.0315789473684</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>40.4063492063492</td>
</tr>
<tr>
<td>level of education</td>
<td>no formal education</td>
<td>28.9285714285714286</td>
</tr>
<tr>
<td></td>
<td>primary education</td>
<td>30.9934640522876</td>
</tr>
<tr>
<td></td>
<td>secondary education</td>
<td>39.3397849462366</td>
</tr>
<tr>
<td>literacy</td>
<td>does not read/write</td>
<td>30.7111111111111</td>
</tr>
<tr>
<td></td>
<td>reads/writes infrequently</td>
<td>34.5263157894737</td>
</tr>
<tr>
<td></td>
<td>reads/writes daily</td>
<td>36.5661375661376</td>
</tr>
<tr>
<td>use of Spanish (as L1 or L2)</td>
<td>none</td>
<td>24.666666666666667</td>
</tr>
<tr>
<td></td>
<td>infrequent</td>
<td>35.8</td>
</tr>
<tr>
<td></td>
<td>frequent/daily</td>
<td>36.2804597701149</td>
</tr>
<tr>
<td>native language group</td>
<td>Mesoamerican</td>
<td>34.7943548387097</td>
</tr>
<tr>
<td></td>
<td>non-Mesoamerican indigenous</td>
<td>26.4</td>
</tr>
<tr>
<td></td>
<td>Spanish</td>
<td>23.5164835164835</td>
</tr>
</tbody>
</table>

- we are currently working on an inferential analysis
- using mixed models linear regression
  - to determine which factors are significant
  - and whether the impact of language can be reduced to non-linguistic factors as predicted by Li & Gleitman 2002

<table>
<thead>
<tr>
<th>innatist predictions</th>
<th>✔️</th>
</tr>
</thead>
<tbody>
<tr>
<td>neo-Whorfian predictions</td>
<td>✔</td>
</tr>
</tbody>
</table>
discussion

- both the linguistic factors (L1, L2) and education and literacy predict the use of FoRs
- however, these variables are interrelated
- a preliminary multivariate analysis suggests that the linguistic factors may be the most powerful ones

- this is evidence supporting the neo-Whorfian view
  - language does seem to be an important factor in the cultural transmission of cognitive practices
Enter MesoSpace (cont.)

• **the future of MesoSpace**

<table>
<thead>
<tr>
<th>project</th>
<th>completed</th>
<th>in progress</th>
<th>planned</th>
</tr>
</thead>
<tbody>
<tr>
<td>descriptive multivariate analysis of linguistic data from additional languages</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>inclusion of topography and population geography in the analysis of the linguistic data</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>linear regression analysis of the linguistic data</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>multivariate analysis of the recall memory data</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>gathering of linguistic data with additional stimuli and tasks</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>gathering of linguistic and nonlinguistic data in other parts of the world (Africa, Asia, South America)</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

*Table 7. No rest for the wicked!*
But are we ready to go quantitative?

• all of the studies presented above suffer from the same fundamental flaw
  • they perform quantitative analyses according to the standards of the social and behavioral sciences
    • on data that was collected according to the customary qualitative protocols of the humanities tradition in linguistics

• how this flaw manifests itself
  • usually, too few speakers per language or linguistically defined population are tested
    • to ensure that the sample is representative of the language community
But are we ready to go quantitative? (cont.)

- the criteria used in recruiting participants are left to the intuition and habits of the participating researchers
  - again, there are no principal safeguards in place to ensure that each sample is representative of the speech community
- approaches to dealing with intra-speaker variation are usually ad hoc
  - sometimes only a participant’s first response is included in the analysis, sometimes their preferred response
    - and sometimes all responses are included or some aggregate measure is computed
- perhaps all the quantitative studies in semantic typology conducted to date will have to be considered pilot studies
  - by the standards of the social and behavioral sciences
HOW COOL IS ALL OF THIS?

• WANT TO KNOW MORE!
  • COME TO SEMANTIC TYPOLGY LAB!
  • TUESDAYS 4-5:20, Spring 2017, Baldy 617
  • Email Erika to be added to the listserve (ebelling@buffalo.edu)
References


References (cont.)


