Synopsis

- spatial reference frames in language, culture, and cognition
- MesoSpace: team, goals, tools
- the Ball & Chair study
- the distribution of the response variables
- the impact of the predictor variables
- discussion and future prospects

Spatial reference frames in language, culture, and cognition (cont.)

- background on reference frames
  - two kinds of place functions (Jackendoff 1983)
    - i.e., functions from reference entities into regions
      - topological (Piaget & Inhelder) – perspective = frame-free
        - means in practice independent of the orientation of the ground,
          the observer, and the figure-ground array (the configuration)

Figure 1. Some configurations that might be described in terms of topological place functions

1.1 The apple is on the skewer
1.2 The band aid is on the shin
1.3 The earring is in the ear (lobe)

Spatial reference frames in language, culture, and cognition (cont.)

- alternative classifications and subtypes

Figure 3. Reference frame types and their classification (A - 'away from'; B - 'back'; D - 'downriver'; F - 'front'; L - 'left'; R - 'right'; T - 'toward'; U - 'upriver'; Bohnemeyer & Levinson ms.)
Spatial reference frames in language, culture, and cognition (cont.)

- Finding: a great deal of crosslinguistic variation
  - in terms of both availability and preferences

- Alignment between language and cognition
  - Preferences for particular frame types in discourse and recall memory covary

- Two competing interpretations
  - Innatist interpretation (Li & Gleitman 2002; Li et al. 2011, inter alia)
    - Innate knowledge of all FO types
    - Variation only in usage preferences
    - Variation caused by adaptation to the environment - topography, population geography, education, literacy
    - Language plays no role in the cultural transmission of practices of spatial reference

- Neo-Whorfian interpretation (Levinson 1996, 2003; Pederson et al. 1998, inter alia)
  - Knowledge of some FO types is culturally transmitted
  - Language plays a key role in the cultural transmission of practices of spatial reference
  - The adaptation to the environment happens at the phylogenetic level, not the ontogenetic level

Spatial reference frames in language, culture, and cognition (cont.)

- But do reference frames diffuse through contact?
  - Languages borrow from one another
    - Phonetic, prosodic, phonotactic patterns; phonemes; morphemes; lexemes; lexical patterns; constructions
  - But reference frames are semantic patterns
    - Which are only indirectly related to particular lexical items

Our test case: the Mesoamerican sprachbund
  - Cf. Campbell 1979; Campbell et al. 1986

The ball is in front of the chair
The ball is left of the chair

Figure 4. Reference frame use in small-scale horizontal space across languages (Bohne et al. and Levinson ms.)

Figure 5. Animals-in-a-Row: design

Figure 6. Results (Levinson 2003: 184)

Figure 7. Innatist vision

Figure 8. Neo-Whorfian vision

Table 1. Animals-in-a-Row in Levinson 2003: the large sample

<table>
<thead>
<tr>
<th>Linguistically Absolute</th>
<th>% of geocentric responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Languages</td>
<td>Predictions</td>
</tr>
<tr>
<td>Amorena</td>
<td>Non-geocentric coding will be absolute</td>
</tr>
<tr>
<td>Pai</td>
<td>Non-geocentric coding will be absolute</td>
</tr>
<tr>
<td>Tsimshian</td>
<td>Non-geocentric coding will be absolute</td>
</tr>
<tr>
<td>Longgu</td>
<td>Non-geocentric coding will be absolute</td>
</tr>
<tr>
<td>Bohor</td>
<td>Non-geocentric coding will be absolute</td>
</tr>
</tbody>
</table>

The role of language contact

-Neo-Whorfians view language as a transmission system for nonlinguistic cognition
-This suggests that not only a person’s L1, but also their L2/L3/…, may affect their cognition
-Experimental support: Boroditsky et al. 2003
  - Learning the grammatical gender system of a made up language influences English speaker’s category associations
  - Counterevidence: Finkbeiner et al. 2003
    - Japanese-English bilinguals behave exactly like monolingual Japanese speakers on a manner/path similarity judgment task
      - Even though Japanese is verb-framed, whereas English is satellite-framed
    - But see Brown & Gullberg 2009

Figure 9. Truth conditions of intrinsic and relative descriptions of Ball & Chair 3.9 (left) and 3.12

Figure 10. Mesoamerican languages map (contemporary distribution)
Source: en.wikipedia.org/wiki/Image:Mesoamericanlanguages.png; lines showing approximate boundaries of Mesoamerican area added by the authors
Synopsis

• spatial reference frames in language, culture, and cognition
• MesoSpace: team, goals, tools
  • the Ball & Chair study
  • the distribution of the response variables
  • the impact of the predictor variables
  • discussion and future prospects

MesoSpace: team, goals, tools

• NSF award #BCS-0723694 Spatial language and cognition in Mesoamerica
• MesoSpace aims to contribute to the debate from two angles
  – we are working on a series of studies that pit linguistic against non-linguistic predictors
    in reference frame use across languages
  – we are also 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

• 13 Mesoamerican (MA) languages
  – Mayan
    • Chol (J.-J. Vázquez)
    • K’ankab’al (E. Mateo)
    • Tzeltal (several variants; G. Polian)
    • Yucatec (J. Bonnemeyer)
  – Mixe-Zoquean
    • Ayutla Mixe (R. Romero)
    • Soteapanec (S. Gutierrez)
    • Tepaptin Zoque (R. Zavala)
  – Oto-Manguean
    • Isthus (Juchitan) Zapotec (G. Pérez)
    • Otomi (N. Hernández, S. Hernández, E. Palancar)
  – Huave (S. Herrera)
  – Purépecha (A. Capistrán)
  – Totonac-Tepehua
    • Huehuetla Tepehua (S. Smythe)
  – Uto-Aztecan
    • Pajapan Nawat (V. Peralta)
  – non-MA “controls”
    – Seri (C. O’Meara)
    – Mayangna (E. Benedicto, A. Eggleston in collaboration with the Mayangna Yulbarangyang Balna)
    – Mexican, Nicaraguan, and Barcelonan Spanish (R. Romero; E. Benedicto, A. Eggleston)

2 (interrelated) domains
  – frames of reference and meronyms (labels for entity parts)

The Ball & Chair study

• 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

The Ball & Chair study (cont.)

– replacement of Men & Tree (M&T) in Pederson et al (1998)
– 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
The Ball & Chair study (cont.)

• the data set of the present study
  – B&C data from 11 varieties
    • 6 Mesoamerican languages
      – Yucatec Maya (J. Bohnemeyer)
      – Ayutla Mixe (R. Romero)
      – San Ildefonso Tuletepec Otomi (N. Hernández, S. Hernández, E. Palancar)
      – Purepecha (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 Yulbarangang Bala)
    • 3 varieties of Spanish
      – from Barcelona (A. Eggleston), Mexico (R. Romero), and Nicaragua (A. Eggleston)

The Ball & Chair study (cont.)

• coding
  – we coded descriptions of the location of the ball
    • distinguishing among eight categories (see Figure 3 above)
      – allocentric intrinsic
      – egocentric intrinsic (=‘direct’; Danziger 2010)
      – egocentric extrinsic = relative
      – intrinsic and relative aligned (Carlson-Radvansky & Irvin 1993)
      – geocentric (= geomorphic, landmark-based, or absolute)
      – vertical absolute
      – vertical absolute and intrinsic aligned (Carlson-Radvansky & Irvin 1993)
      – topological (no reference frame involved; Piaget & Inhelder 1956)

The Ball & Chair study (cont.)

• all of the languages in the sample have the lexical and grammatical resources for using all FoR types
  – in no case does the grammar or lexicon of the language constrain the use of particular frame types
  – reference frames are semantic patterns
    • which are only indirectly related to particular lexical items

The Ball & Chair study (cont.)

– these are all the languages of the MesoSpace sample
  – data from five dyads of participants per variety are included in the analysis
    – except for the case of
      – Mexican Spanish, where up to now only the data from three of the five dyads have been coded
      – Isthmus Zapotec and Barcelona Spanish, where we have data from six dyads
    – responses are accompanied by the researchers’ estimates of the participants’
      • level of education
      • frequency of use of Spanish (as first or second language)
      • frequency of reading and writing

The Ball & Chair study (cont.)

• the similarity matrix
  – 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 calculated the distances in the “Manhattan” metric
    • where the distance between two points is the sum of the differences of the coordinates
  – we can use this similarity measure to analyze
    • how the responses cluster
    • which factors predict the similarity between participants
Frames of reference in language, culture, and cognition: The Mesoamerican evidence  

The Ball & Chair study (cont.)

- the similarity matrix (cont.)
  - innovation
    - previous multivariate analyses in semantic typology construct similarity matrices over the stimulus items
      - cf. Levinson & Meira 2003; Majid et al 2008
    - in contrast, our approach treats the (dyads of) participants as statistical units
    - this allows us to treat language as a direct predictor variable

Synopsis

- spatial reference frames in language, culture, and cognition
- MesoSpace: team, goals, tools
- the Ball & Chair study
  - the distribution of the response variables
    - the impact of the predictor variables
    - discussion and future prospects

The distribution of the response variables

- the flow of the quantitative analysis
  - step I: identify the response variables that showed the greatest differentiation among participants
    - response variables: the (frequency/probability of) use of each of the eight strategies we coded the data for
  - step II: linear regressions to find the predictor variables significantly contributing to the variance
    - in those response variables identified in step I
    - predictor variables: L1, L2 use, literacy, education, (topography, population geography)

The distribution of the response variables (cont.)

- the similarity matrix
  - 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 calculated the distances in the “Manhattan” metric
    - where the distance between two points is the sum of the differences of the coordinates
  - we can use this similarity measure to analyze
    - how the responses cluster
    - which factors predict the similarity between participants

The distribution of the response variables (cont.)

- how do the participants’ responses cluster?
  - we ran a three-dimensional Multi-Dimensional Scaling (MDS) analysis of the similarity matrix
    - three dimensions produced a better goodness of fit than two
  - cf. Schiffman et al 1981

Figure 15. Plotting the first two dimensions of the MDS analysis
Frames of reference in language, culture, and cognition: The Mesoamerican evidence

MDS analysis (cont.)

- the first dimension of the MDS plot correlates positively with the frequency of geocentric descriptions...
  - Spearman's Rho 0.88
- ... and negatively with the frequency of relative descriptions
  - Spearman's Rho -0.85

Figure 16. Correlations between the first dimension of the MDS plot and the frequency of geocentric (left) and relative (right) descriptions.

MDS analysis (cont.)

- the second dimension shows a very strong negative correlation with the frequency of topological description
  - Spearman's Rho -0.99
- the third dimension exhibits a rather weak correlation with the frequency of intrinsic descriptions
  - Spearman's Rho 0.76

Figure 17. Correlations b/w the 2nd dimension of the MDS plot and the frequency of topological descriptions (left) and b/w the 3rd and the frequency of intrinsic descriptions.

The distribution of the response variables (cont.)

- the effect of relative and geocentric usage can also be visualized in a Neighbor-net of the similarity matrix

Figure 18. The Neighbor-net and its “geography”

• discussion
  - the MDS and Neighbor-net analyses show
    - that the participants differentiated themselves most strongly in their use of relative, geocentric, and topological descriptions
  - the question now: which factors predict which of these strategies a speaker/dyad selects?
    - candidate predictor variables:
      1. L1
      2. L2 (... Ln)
      3. literacy
      4. education
      5. topography
      6. population geography
    - the linear regression we present in the following tests (1) – (4)

Synopsis

- spatial reference frames in language contact
- MesoSpace: team, goals, tools
- the Ball & Chair study
- the distribution of the response variables
- the impact of the predictor variables
- discussion and future prospects

The impact of the predictor variables

- to analyze the role of the predictor variables we conducted several linear regression analyses
- we tested separate models for the strongest differentiating response variables
  - the use of relative and geocentric frames
- we tested these models for two sets of populations
  - on all 11 populations
    - with the predictor variables areal-linguistic affiliation (see below!), literacy, and education
  - on the speakers of the indigenous languages only
    - now including the L2 use of Spanish as a predictor variable
The impact of the predictor variables (cont.)

- the areal-linguistic affiliation variable
  - our dataset includes too many individual languages for a parsimonious model
  - therefore, we grouped the languages according to areal-linguistic affiliation
    - yielding a three-level variable for the 11 populations
      - languages of the Mesoamerican sprachbund, Spanish, and the two non-Mesoamerican indigenous languages
    - and a two-level variable for the models that include the responses from the speakers of the indigenous languages only
      - Mesoamerican sprachbund languages vs. non-Mesoamerican indigenous languages (Seri and Sumu)

implementation
- we used generalized linear mixed-effects models (GLMM; cf. Gelman & Hill 2007, Jaeger 2008)
- implemented using the ARM package in R (Gelman et al 2012)
  - ‘mixed-effects’ models b/c they include random nested intercepts for individual languages and dyads
    - in addition to the ‘fixed’ effects of the predictor variables and an invariable intercept
  - to avoid over-fitting or lack of independence
  - the probability of a given dyad using any of the eight response categories to describe a particular picture
    - is independent of the probability of them using any other type of frame to describe the same picture

The impact of the predictor variables (cont.)

- findings I: GEO, L1-Spanish speakers incl.
  - the fitted geocentric model revealed linguistic affiliation and literacy, but not education, as significant factors
    - there was no effect from membership in the MA sprachbund

The most favored model of the probability of geocentric usage by AIC ranking includes all of these variables

The impact of the predictor variables (cont.)

- findings II: REL, L1-Spanish speakers incl.
  - the fitted relative model revealed linguistic affiliation as the sole significant factor
    - there was no evidence of an areal effect

The most favored model of the probability of relative usage by AIC ranking excludes the nonlinguistic variables

The impact of the predictor variables (cont.)

- findings III: GEO, L1-Spanish speakers excl.
  - the fitted geocentric model showed literacy as the sole significant factor

The most favored model of the probability of geocentric usage by AIC ranking includes only L2-Spanish use and literacy as variables

The impact of the predictor variables (cont.)

- findings IV: REL, L1-Spanish speakers excl.
  - the fitted relative model showed the use of L2 Spanish as the sole significant factor

The most favored model of the probability of relative usage by AIC ranking excludes the nonlinguistic variables
The impact of the predictor variables (cont.)

**discussion: the role of the first language**

- the L1-Spanish speakers differed significantly from the speakers of the indigenous languages
  - using relative frames overall much more frequently and geocentric frames overall much less frequently
  - this contribution of L1 cannot be reduced to a combination of any of the other factors
    - to this extent contra Li & Gleitman 2002

The impact of the predictor variables (cont.)

**discussion: the role of the second language**

- the speakers of the indigenous languages use relative frames in their native languages more frequently
  - the more frequently they use Spanish as an L2
  - this suggests that habituation to the use of relative frames diffuses through contact with Spanish
    - consistent with the Neo-Whorfians view of language as a transmission system for nonlinguistic cognition

The impact of the predictor variables (cont.)

**discussion: the role of the nonlinguistic factors**

- literacy, assessed in terms of the frequency of reading and writing, is a significant predictor of frame use
  - this variable makes a significant independent contribution to the use of geocentric frames, but not to that of relative ones
    - presumably, speakers who read and write more frequently are less likely to use geocentric frames
  - in contrast, we did not find any effect of education
  - overall, this picture is consistent with the varying role of education and literacy across our sample
    - some of the indigenous populations have high education scores across the board
      - and nevertheless use geocentric frames more frequently than relative ones
      - especially the Isthmus Zapotec and Sumu-Mayangna communities

The impact of the predictor variables (cont.)

**discussion: the role of the Mesoamerican area**

- our GLMMs found significant differences b/w the speakers of Spanish and the indigenous languages...
  - ... but not b/w the Mesoamerican and the non-Mesoamerican indigenous languages
  - we thus did not find any evidence of an areal effect
  - given that we did find evidence of contact diffusion of the use of relative frames
    - we decided to probe this lack of evidence of a sprachbund effect further

The impact of the predictor variables (cont.)

**probing the lack of evidence for an areal effect**

- we ran a cluster analysis of the similarity matrix
  - including again the data from the L1-Spanish speakers
- we applied an agglomerative algorithm using the ‘cluster’ and ‘MASS’ packages in R

**findings**

- with three exceptions from three different varieties, the speakers of the three Spanish varieties cluster together
  - due to their unifying high relative and low geocentric scores
- in contrast, no clear differentiation between the Mesoamerican (MA) and non-MA indigenous languages emerged

Figure 19. Cluster analysis dendrogram of the similarity matrix of the Ball & Chair data.

Figure 20. Color-coding the clusters: red – Mesoamerican; green – non-Mesoamerican
Synopsis

- spatial reference frames in language contact
- MesoSpace: team, goals, tools
- the Ball & Chair study
- the distribution of the response variables
- the impact of the predictor variables
- discussion and future prospects

Discussion and future prospects

- language as an influence on frame use
  - linear regressions of data from speakers of 11 varieties suggest that L1 is an irreducible factor in frame selection
  - a speaker’s first language is a powerful predictor of their probability of using relative and geocentric frames
    - more specifically, speaking any variety of Spanish predicts a very different usage profile from speaking any indigenous language
  - this effect of first language cannot be reduced to effects of literacy and education
    - this finding conforms to the Neo-Whorfian predictions

Discussion and future prospects (cont.)

- estimated frequency of L2 Spanish use is also a significant predictor of the use of relative frames
  - by speakers of the indigenous languages in the sample
    - so is literacy, but not education
  - this finding supports the hypothesis that reference frame types diffuse through language contact
    - this likewise accords with the Neo-Whorfian view
    - in contrast, we did not find evidence for an areal effect
      - the speakers of the Mesoamerican languages distinguish themselves from the speakers of the Spanish varieties
        - but not clearly from the speakers of the two non-Mesoamerican indigenous languages Seri and Sumu

Discussion and future prospects (cont.)

- the Mesoamerican linguistic area as a fossilized sprachbund
  - the contact that caused the convergence of linguistic features in MA unfolded mostly in pre-Columbian times
  - at present, contact among indigenous languages is mostly restricted to certain hotspots
    - chiefly, to parts of Oaxaca, Chiapas, and Guatemala
  - none of the indigenous languages of the MesoSpace subsample are currently in contact with one another
  - given the evidence for intra-variety mutability of reference frame use...
    - ... it stands to reason that areal effects in frame use that may have existed prior to the Conquest are no longer visible

Discussion and future prospects (cont.)

- by hypothesis, any feature that can be contact-diffused should also be able to be areally shared
  - so our failure to find an areal effect seems to call for an explanation

- possible factors
  - sampling artifact
    - the use of reference frames in Seri and/or Sumu could be accidentally so similar to that in some MA languages
      - as to mask a possible areal effect
  - effects of current vs. historic contact
    - whereas the effect of Spanish on the use of reference frames may be ongoing, a sprachbund effect likely not
      - since the MA sprachbund is no longer "active" in many regions

Discussion and future prospects (cont.)

- what’s next?
  - include data from additional Mesoamerican languages in the analysis
  - run a second analysis based on speakers’ self-estimations of Spanish use, literacy, and education
  - run similar analyses on the recall memory data
  - extend all of the above to languages from other parts of the world
    - as part of the new project Spatial Language and Cognition Beyond Mesoamerica
      - NSF Award No. BCS-1053123
      - http://www.acsu.buffalo.edu/~jb77/Mesospace1b.html
Acknowledgements

• I would like to thank
  — ... our teachers and consultants, the speakers of the languages the MesoSpace team has been studying
  — ... my colleagues, the members of the MesoSpace team

Acknowledgements (cont.)

• I would like to thank (cont.)
  — ... the National Science Foundation, for the necessary resources to realize these studies
  — ... the institutions who have partnered with MesoSpace to lend us support, CIESAS and the MPI for Psycholinguistics
  — ... Matthew Dryer, Jeff Good, Marianne Gullberg, Florian Jaeger, Jean-Pierre Koenig, Steve Levinson, David Mark, Wolfgang Wölk
    • and the members of the UB Semantic Typology Lab, for advice
  — ... audiences at
    • CILSA V, the International Conference on Yucatecan Linguistics, and the Workshop on Quantitative Methods in Areal Typology
    • for comments on previous presentations of some of the material
  — ... you!

References


References (cont.)