Spatial reference frames in language contact

- two central questions
  - are practices of language use contact-diffused?
  - can such practices constitute areal features?
- a domain in which to look for answers: spatial frames of reference

Spatial reference frames in language contact (cont.)

- projective – framework-dependent
  - the place function returns a region defined in a coordinate system centered on the reference entity
  - the axes of the coordinate system are derived from an anchor
    - in intrinsic frames, the anchor is the reference entity
    - in relative frames, it is the body of an observer
    - in absolute frames, it is some environmental entity/feature

Spatial reference frames in language contact (cont.)

- background on reference frames
  - two kinds of place functions (Jackendoff 1983)
    - I.e., functions from reference entities into regions
      - topological (Piaget & Inhelder 1956) – 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 contact (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 contact (cont.)

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

Spatial reference frames in language contact (cont.)

- alignment between language and cognition
  - preferences for particular frame types in discourse and recall memory covary

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

<table>
<thead>
<tr>
<th>Linguistically</th>
<th>Relative</th>
<th>Prediction: Non-verbal coding will be absolute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linguistically</td>
<td>Relative</td>
<td>Prediction: Non-verbal coding will be relative</td>
</tr>
</tbody>
</table>

Spatial reference frames in language contact (cont.)

- the role of language contact
  - the Neo-Whorfians view language as a transmission system for nonlinguistic cognition
  - this suggests that not only a person’s L1, but also their L2/3/…, 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
  - counter-evidence: 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

Spatial reference frames in language contact (cont.)

- our test case: the Mesoamerican *sprachbund*

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

Spatial reference frames in language contact (cont.)

- two competing interpretations
  - Innatist interpretation (Li & Gleitman 2002; Li et al 2011, inter alia)
    - innate knowledge of all FoR 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 FoR 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 at the ontogenetic level

Spatial reference frames in language contact (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

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

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

<table>
<thead>
<tr>
<th>relative</th>
<th>intrinsic</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>in which type of FoR?</td>
</tr>
</tbody>
</table>

December 14th, 2012
In search of areal effects in semantic typology: Reference frames in Mesoamerica

December 14th, 2012

Spatial reference frames in language contact (cont.)

- possible areal features according to Campbell 1979, Campbell, Kaufman, & Smith-Stark 1986
  - lack of phonemic voicing contrasts in stops and fricatives
    - shared throughout MA, with only a few exceptions
      - including Tequistlatec, Huave, and some OM
    - neighbors to the north (e.g., O’odham; Tarahumara) and south (e.g., Sumu, Miskito; Chibchan) do have them
  - no V-final constituent orders except in Mixean
    - Yuman and most Uto-Aztecan languages to the north and Chibchan and Misumalpan to the south are V-final
    - correlated with the absence of V-final order, adnominal possessors predominantly follow the possessum
      - not so in Sumu and Miskito to the south of MA; not in most UA languages to the north of MA

Spatial reference frames in language contact (cont.)

- possible areal features (cont.)
  - few or no adpositions
  - relational nouns and applicatives used instead
  - semantic calques
    - this includes the vigesimal numeral system and a rich set of meronymic metaphors

Spatial reference frames in language contact (cont.)

- particularly interesting for our purposes
  - the vigesimal system and the meronym calques at least suggest contact-diffused usage practices
    - although the effect is in this case “set” in the meanings of lexical items
  - the role of Spanish as the dominant contact language of the area
    - relative frames of reference play only a minor role in many Mesoamerican languages
      - cf. the contributions to O’Meara & Pérez Báez (eds.) 2011 and references therein
    - in contrast, Spanish as a European language favors relative frames in small-scale space
      - cf. Eggleston 2012 on Peninsular (Barcelonan) and Nicaraguan Spanish in comparison to Sumu-Mayangna

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

Spatial reference frames in language contact (cont.)

Table 1.
Some pan-MA calques (CK&S p. 553)

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

MesoSpace: team, goals, tools (cont.)

- 14 Mesoamerican (MA) languages
  - Mayan
    - Ch'ol (J.-J. Vázquez)
    - K’anjob’al (E. Mateo)
    - Tzutul (several variants; G. Pollan)
    - Tz'utujil (J. Bohmeyer)
  - Mixe-Zoquean
    - Ayutla Mixe (R. Romero)
    - Soctapanc (S. Gutierrez)
    - Tecpatán Zoque (R. Zavala)
  - Oto-Manguean
    - Isthmus (S. Smythe)
    - Zapotec (G. Pérez)
    - Otomi (N. Hernández, S. Hernández, E. Palán)
  - Huave (S. Herrera)
  - Pupepecha (A. Capistrán)
  - Totonac-Tepehuan
    - Huehueta Tepehua (S. Smythe)
  - Uto-Aztecan
    - Pajapana (V. Peralta)
In search of areal effects in semantic typology: Reference frames in Mesoamerica

December 14th, 2012

MesoSpace: team, goals, tools (cont.)

- non-MA "controls"
  - Seri (C. O'Meara)
  - Cora (Uto-Aztecan; V. Vázquez)
  - Mayangna (E. Benedicto, A. Eggleston in collaboration with the Mayangna Yulbarangyang Balna)
  - Mexican, Nicaraguan, and Barcelonian Spanish (R. Romero; E. Benedicto, A. Eggleston)

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

Figure 12. The MesoSpace team (minus V. Peralta and R. Tucker)

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 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

Figure 13. Meronyms in Ayoquesco Zapotec (left) and Tenejapa Tseltal (adapted from MacLaury 1989 and Levinson 1994)

Figure 14. Design of the Men and Tree task (Pederson et al. 1998: 562)

Figure 15. Two of the Ball & Chair fotos, featuring an intrinsic contrast

The Ball & Chair study (cont.)

- these are all the languages of the MesoSpace sample the data from which have been coded so far

- 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, 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 data set of the present study
  - B&C data from 11 varieties
    - 6 Mesoamerican languages
      - Yucatec Maya (J. Bohnemeyer)
      - Ayutla Mixe (R. Romero)
      - San Idelfonso Tultepec Otomi (N. Hernández, S. Hernández, E. Palancar)
      - Purépecha (or Tarascan; A. Capistrán)
      - Chacema Tzeltal (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)

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)
In search of areal effects in semantic typology: Reference frames in Mesoamerica

December 14th, 2012

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
  - a given speech community’s preferences for using particular frame types are strictly a matter of usage
    - they are a part of the community’s practices of language use
  - the question the studies reported here address is this:
    - does the frame use of individual speakers reflect not only the practices of their L1 speech community
      - but also those of communities whose languages they use as L2 speakers?

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

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

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 distribution of the response variables
- how do the participants’ responses cluster?
  - MDS analysis shows two broad groups
    - cf. Schiffman et al 1981

Figure 16. MDS plot

MDS analysis (cont.)
- a strong correlation emerges b/w the dimension of the MDS plot and the use of geocentric frames
  - Spearman’s Rho 0.95
  - and weaker negative correlation between the first dimension and the use of relative frames
    - Spearman’s Rho -0.8
- the second dimension shows a weak correlation with the frequency of topological descriptions
  - Spearman’s Rho 0.79

Figure 17. Correlations between the dimensions of the MDS plot and the frequency of geocentric (left), relative (center), and topological (right) descriptions.
In search of areal effects in semantic typology: Reference frames in Mesoamerica

December 14th, 2012

Enter MesoSpace (cont.)

– the effect of relative and geocentric usage can also be visualized in a Neighbor-net of the similarity matrix
  • using Splitstree4 (cf. Hudson & Bryant 2006)

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

MDS analysis (cont.)

– 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 a linear regression analysis

• we tested separate models for the strongest differentiating response variables
  • the use of relative and geocentric frames

• we modeled the geocentric and relative FoR scores
  – of just the speakers of the indigenous languages

  – as a function of education level, literacy level, L2-Spanish usage level
  • and areal-linguistic affiliation: Mesoamerican vs. non-Mesoamerican

Linear regression analysis (cont.)

• implementation
  – we used a generalized linear mixed-effects model (GLMM; cf. Gelman & Hill 2007, Jaeger 2008)
    • implemented using the ARM package in R (Gelman et al 2012)
  – it is a ‘mixed-effects’ model in that it includes 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

Linear regression analysis (cont.)

• findings
  – the fitted geocentric model revealed

L2-Spanish use and literacy as significant factors

Generalized linear mixed model fit by the Laplace approximation
Formula: Lgeoc ~ edu + esp + lit + Ltyp + (1 | ID) + (1 | LANG)
Data: ..1

AIC BIC logLik deviance
1559 1598 -772.7 1545

Random effects:
Groups Name Variance Std.Dev
ID (Intercept) 1.6226 1.2738
LANG (Intercept) 2.0463 1.4305

Number of obs: 1787, groups: ID, 82; LANG, 8

Fixed effects:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.0799 1.3550 -0.797 0.4255
edu -0.4788 0.4951 -0.967 0.3335
esp -0.8469 0.3333 -2.526 0.0115 *
lit 1.1892 0.4836 2.459 0.0139 *
LtypMES 0.3375 1.2535 0.269 0.7878

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Generalized linear mixed model fit by the Laplace approximation
Formula: Lgeoc ~ edu + esp + lit + Ltyp + (1 | ID) + (1 | LANG)
Data: ..1

AIC BIC logLik deviance
1559 1598 -772.7 1545

Random effects:
Groups Name Variance Std.Dev
ID (Intercept) 1.6226 1.2738
LANG (Intercept) 2.0463 1.4305

Number of obs: 1787, groups: ID, 82; LANG, 8

Fixed effects:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.0799 1.3550 -0.797 0.4255
edu -0.4788 0.4951 -0.967 0.3335
esp -0.8469 0.3333 -2.526 0.0115 *
lit 1.1892 0.4836 2.459 0.0139 *
LtypMES 0.3375 1.2535 0.269 0.7878

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
In search of areal effects in semantic typology: Reference frames in Mesoamerica

December 14th, 2012

Linear regression analysis (cont.)
• findings (cont.)
  – the fitted relative model revealed only L2-Spanish use as significant

Generalized linear mixed model fit by the Laplace approximation
Formula: Lrel ~ edu + esp + lit + Ltyp + (1 | ID) + (1 | LANG)
Data: ...

Correlation of Fixed Effects:

AIC BIC logLik deviance
data: 1787, groups: ID, 82; LANG, 8

Fixed effects:
Estimate Std. Error z value
(Intercept) -2.43468 0.57355 -4.245 2.19e-05 ***
edu -0.20378 0.28745 -0.709 0.4784
esp 0.45204 0.19250 2.348 0.0189 *
lit 0.05716 0.28627 0.200 0.8417
LtypMES -0.58065 0.40908 -1.419 0.1558

--- Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:

(edu)
edu -0.022 esp -0.630 lit -0.154
esp -0.154 lit -0.755 -0.076
lit -0.076 LtypMES 0.250

Random effects:

Variance Std.Dev
ID (Intercept) 0.44961 0.67053
LANG (Intercept) 0.14426 0.37981

Number of obs: 1787, groups: ID, 82; LANG, 8

Linear regression analysis (cont.)
• findings
  – our GLMMs did not find a significant effect of the areal-linguistic affiliation variable
  • Wald-p = .79 for the geocentric and .16 for the relative model

• discussion
  – 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
  – our failure to find evidence of an areal effect caused us to conduct further analyses

Synopses
• 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
• estimated frequency of L2 Spanish use is 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
• 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.)

- 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

- we would like to thank
  — … our teachers and consultants, the speakers of the languages the MesoSpace team has been studying
  — … our colleagues, the members of the MesoSpace team
  — … 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 Gulberg, Florian Jaeger, Steve Levinson, David Mark, and Wolfgang Wölk
  — … the QAALT organizers
  — … you!

References


References (cont.)


