

A faint, light gray world map is visible in the background, showing the outlines of continents and major landmasses. The map is centered on the Atlantic Ocean.

# **Typological Patterns and Hidden Diversity**

**Balthasar Bickel**  
**U. Leipzig**

## A problem

---

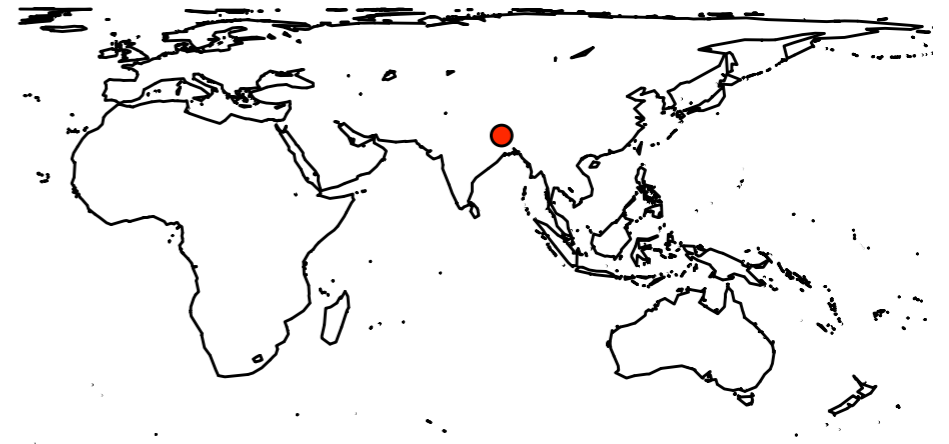
- Typology often destroys its own field: cross-linguistic diversity
- because our analytical notions are systematically gloss over variation and are defined by absolute universals, which often don't hold
  - because typologies are focused on “per language” summaries, glossing over within-language diversity

## Example 1: affixes and clitics

- In many typologies, the notion ‘affix’ implies the following absolute universal: *If something is hosted by a phrase, it is non-selective, and vice-versa.*

Belhare (Kiranti, Sino-Tibetan)

- phendi dabhek=ηa* (vs. *dabhe*)  
axe machete=INS  
‘with an axe and a machete’
- uchoñat phendik=ηa* (vs. *phendi*)  
new axe=INS  
‘with the new axe’
- uchouãt=na*  
new=INS  
‘with the new one’
- ina=ηa*  
DIST.DEM=INS  
‘with that one’

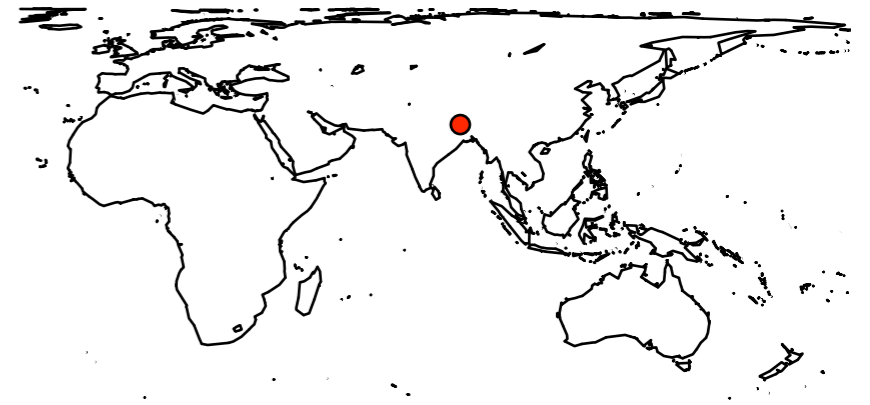


## Example 1: affixes and clitics

- But what about more radically non-selective elements?

Chintang (Sino-Tibetan, Bickel et al. 2007)

- asinda=ta a-ma-im-yokt-e.*  
yesterday=FOC 2-NEG-sleep-NEG-PST  
'You didn't sleep YESTERDAY.'
- asinda a-ma-im-yokt-e=ta.*  
yesterday 2-NEG-sleep-NEG-PST=FOC  
'You DIDN'T SLEEP YESTERDAY.'
- asinda a-ma=ta-im-yokt-e.*  
yesterday 2-NEG=FOC-sleep-NEG-PST  
'You did NOT sleep yesterday.'

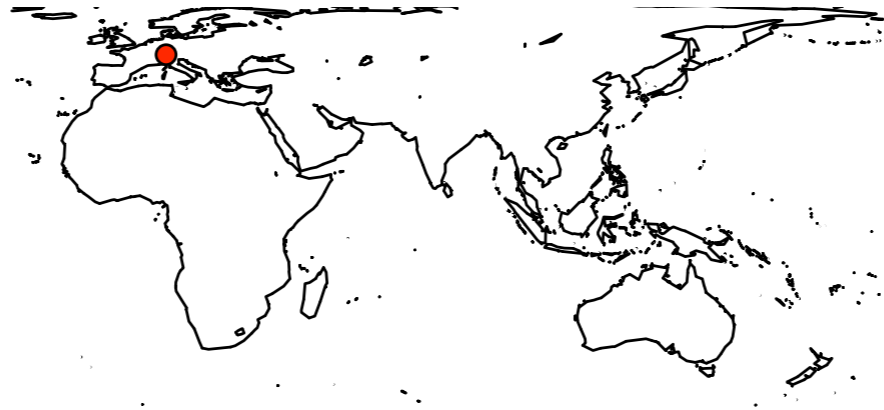


## Example 1: affixes and clitics

- or non-phrasal non-selective elements?

Swiss German

- a. *schlaafe* > *schlööf-l-e* ‘to sleep a bit, take a nap’  
b. *schlaaf* > *schlööf-li* ‘a short sleep, a nap’



## Example 1: affixes and clitics

- In many typologies, the notion ‘clitic’ implies the following absolute universal: *Nonselective elements may appear on the wrong host, selective elements may not.*

Kwak’wala (Anderson 1985)

*nep’id=i=da*      *gənanəm=a* *guk<sup>w</sup>=sa*      *t’isəm.*

throw=SUBJ=DET child=OBJ house=INSTR rock

‘The child threw a rock at the house.’

Belhare (Sino-Tibetan)

a. *n-ta-be*

3pS-come-PST

‘They came.’

b. [*pw unchiŋ*] [*pw tahe*]

unchik      *n-tah-e*

3nsNOM 3pS-come-PST

‘They came.’



## Example 1: affixes and clitics

- In many typologies, the notion ‘affix’ implies the following absolute universal: *If something is selective and attaches to stems (nonphrasal), its position is fixed.*

Chintang (Sino-Tibetan, Bickel et al. 2007)

a. *kha-u-ya-cept-e*

1nsP-3nsA-call-call-PST

b. *kha-ya-u-cept-e*

c. *u-kha-ya-cept-e*

d. *u-ya-kha-cept-e*

e. *ya-kha-u-cept-e*

f. *ya-u-kha-cept-e*

‘They called us.’

a. *u-[kos-a]-[gond-e]*

3nsS-walk-PST-AMBULATIVE-PST

b. *[kos-a]-u-[gond-e]*

walk-PST-3nsS-AMBULATIVE-PST

‘They walked around.’

## Example 1: Conclusion

---

- Notions like ‘affix’ or ‘clitic’ implicitly imply absolute universals of strictly associated properties, e.g.
  - selective  $\leftrightarrow$  fixed position  $\leftrightarrow$  local exponence
  - nonselective  $\leftrightarrow$  phrasal  $\leftrightarrow$  possibly displaced exponence
- But these universals are not absolute: there is much more diversity than the terms allow.
- Typologies based on ‘affixes’ or ‘clitics’ systematically underestimate true diversity.
- unwarranted confidence that “after all, languages are not that different from each other” (cf. Evans & Levinson’s 2009 point)



## Example 2: case alignment

---

- Received knowledge: within languages, case alignments are relative to
  - reference
  - clause type (tense, aspect, periphrasis, dependency, etc.)
- but there is much more diversity beyond this: lexical conditions (Bickel & Nichols 2009 [case handbook])

## Example 2: case alignment — Chintang ditransitives

O=T=G≠A

- a. *akka u-phari pid-a-hã=o!*  
1s[-NOM] 3sPOSS-half[-NOM] give-IMP-1sP.IMP  
'Give me half of it!'

10%

- b. *huĩsa-ŋa hana chatta na-bopt-e.*  
DEMs-ERG 2s[-NOM] umbrella[-NOM] 3>2-cover-PST  
'She covered you with an umbrella.'

O=G≠T=A

- a. *a-ma-ŋa hana munjei-ŋa na-bhukt-e.*  
1sPOSS-mother-ERG 2s[-NOM] shawl-INS 3>2-cover-PST  
'Mother covered you with a shawl.'

20%

- b. *athomba gol-ŋa rame or-o-ŋs-e.*  
before ball-INS R.[-NOM] throw.at-3sP-PERF-PST  
'He has hit Rame with a ball before.'

O=T≠G≠A

- a. *huĩsa-ŋa dabai u-narek-be yokt-e.*  
DEMs-ERG medicine[-NOM] 3sPOSS-ear-LOC [3sA-]apply-PST[-3sP]  
'S/he put some medicine onto his/her ear.'

70%

- b. *huĩsa-ŋa cuwa gagri-be phatt-e.*  
DEMs-ERG water[-NOM] large.container-LOC [3sA-]fill-PST[-3sP]  
'S/he filled the *gāgrī* with water.'

## Example 2: case alignment

- Exactly the same problem with split-S systems or oblique S or A marking, e.g. German *mich friert* vs. *ich arbeite*
  - German S-ACC:  $S=O=T \neq A \neq G$  (for 1/2sg/3sgMASC)
  - German S-NOM:  $S=A \neq O=T \neq G$  (for 1/2sg/3sgMASC)
- Full set of case alignments in Chintang:
  - $S=A=O=T=G$  class1 in 1; class3 (experiential) in all persons
  - $S=A=O=G \neq T$  class2 in 1
  - $S=A=O=T \neq G$  default in 1
  - $S=O=T=G \neq A$  class1 in 2/3/N
  - $S=O=G \neq A=T$  class2 in 2/3/N
  - $S=O=T \neq A \neq G$  default in 2/3/N
  - $S \neq O=T \neq A \neq G$  class4 (sensations, GEN-S) in 2/3/N
  - $S \neq O=T=A \neq G$  class4 in 1
  - $S=A \neq O=T \neq G$  class5 (NOM-experiencer/A, ERG/INS-stimulus/O)

## Example 2: Conclusion

---

- Great diversity in alignment systems within languages
  - and I haven't even mentioned languages where  $A_{tr} \neq A_{ditr}$ !
  - and not at all alignments in other constructions (e.g. agreement, or raising)!
- Again, as in Example 1, typologies of alignment tend to systematically underestimate the true diversity

# Classical responses

---

- Reduce the diversity before you study it!
- ▶ Theory-of-grammar/framework-centered approaches:
  - search for higher-ranking principles that explain why X behaves like a clitic although it's really an affix
  - revise the analysis or the theory (of alignment, of morphology etc.)
- ▶ Classical typological approaches:
  - typologize exemplars (“basic” alignment; “prototypical” affix)
  - define “comparative concepts” that abstract away from language-particular details (Lazard 2006, Haspelmath 2007)

## An alternative: Multivariate Typology

---

- The cause of variation is that across languages things are mostly similar and hardly ever identical.
- But similarity is nothing else but identity in some variables and difference in others.
- ▶ For studying similarities we need large systems of fine-grained variables that fully capture the range of known variation: **Multivariate Typology.**

(This is similar in spirit to Canonical Typology (Corbett 2005), except that it does not assume canons.)
- ▶ And with this, we can **measure the variation, instead of reducing it** — i.e. do what most other disciplines would do when confronted with variation.

# A Multivariate Typology of grammatical markers

---

1. syntactic host type: *phrasal, terminal*
  2. behavior: *inert, distributive* (stacking or spreading)
  3. selectivity: *restricted, unrestricted*
  4. phonological fusion: *isolating, concatenative, nonconcat.*
  5. flexivity: *flexive, nonflexive* (allomorphy)
  6. phrase position: *final, initial, on head*
  7. syntactic placement: *preceding, following*
  8. phonological position: *prae, post, simul, split*
  9. phonological host: *C, V,  $\sigma$ ,  $\phi$ , pw-Chintang1, pw-Chintang2...*
  10. phonological level: *lexical, postlexical* etc.
  11. licensing: *agreement, government, free choice*
- etc.**

# A Multivariate Typology of case alignment

---

1. Referential category: *1sg, TOP, 3sgMASC.German* etc.
  2. Clause dependency: *main, dependent* etc.
  3. TAM form: *synthetic, periphrastic* etc.
  4. TAM content: *Nepali-Perfective, Chechen-Continuous* etc.
  5. Predicate Class: language-specific, but possibly coded for semantics
- etc.



# Multivariate Typology

---

- Each multivariate typology
  - consists of sets of variables on a level of resolution that is virtually identical to the tools we need for analysing primary data (e.g. in fieldwork):
    - instead of asking: “is X an affix?”, we ask: “does X attach to a phrase or a stem?”, “does X select the category of what it attaches to?” etc.
    - instead of asking: “what is the basic ditransitive alignment?”, we ask “which predicate class shows which alignment under which conditions?”
  - allows for thousands of level combinations: the true diversity

## But what about the good old typological generalizations?

---

- ... can't just get them as well, but better!
- For this we need...
  - A. Heuristics: data mining techniques, e.g.
    - distance-based techniques for finding clusters of similar structures (split graphs, multidimensional scaling etc.) ('prototypes', bottom-up 'canons')
    - entropy-based techniques for finding associations between variables (possibly weighted)

**NB:** if some levels of variables are language-specific, this is OK — it just won't increase similarities or strengthen correlations!

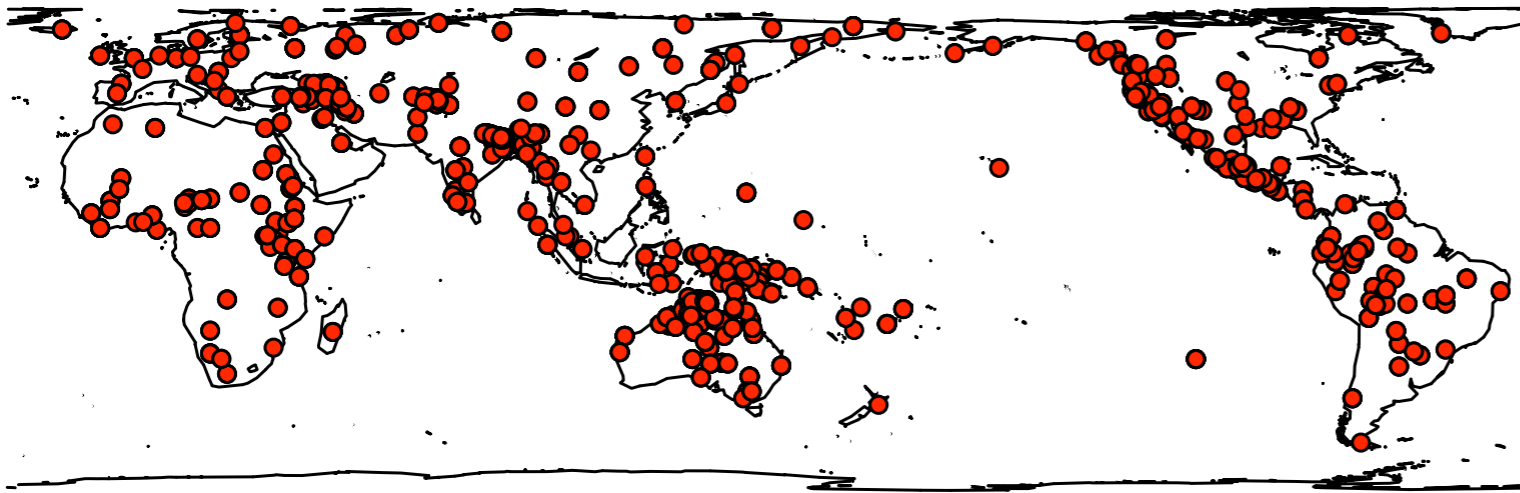
## But what about the good old typological generalizations?

---

- B. Test strategies: permutation-based statistics, e.g. of
- simple contingency tables (e.g. exact tests)
  - generalized linear models (e.g. Likelihood Ratio)

## Case Study A: mining a database on grammatical markers

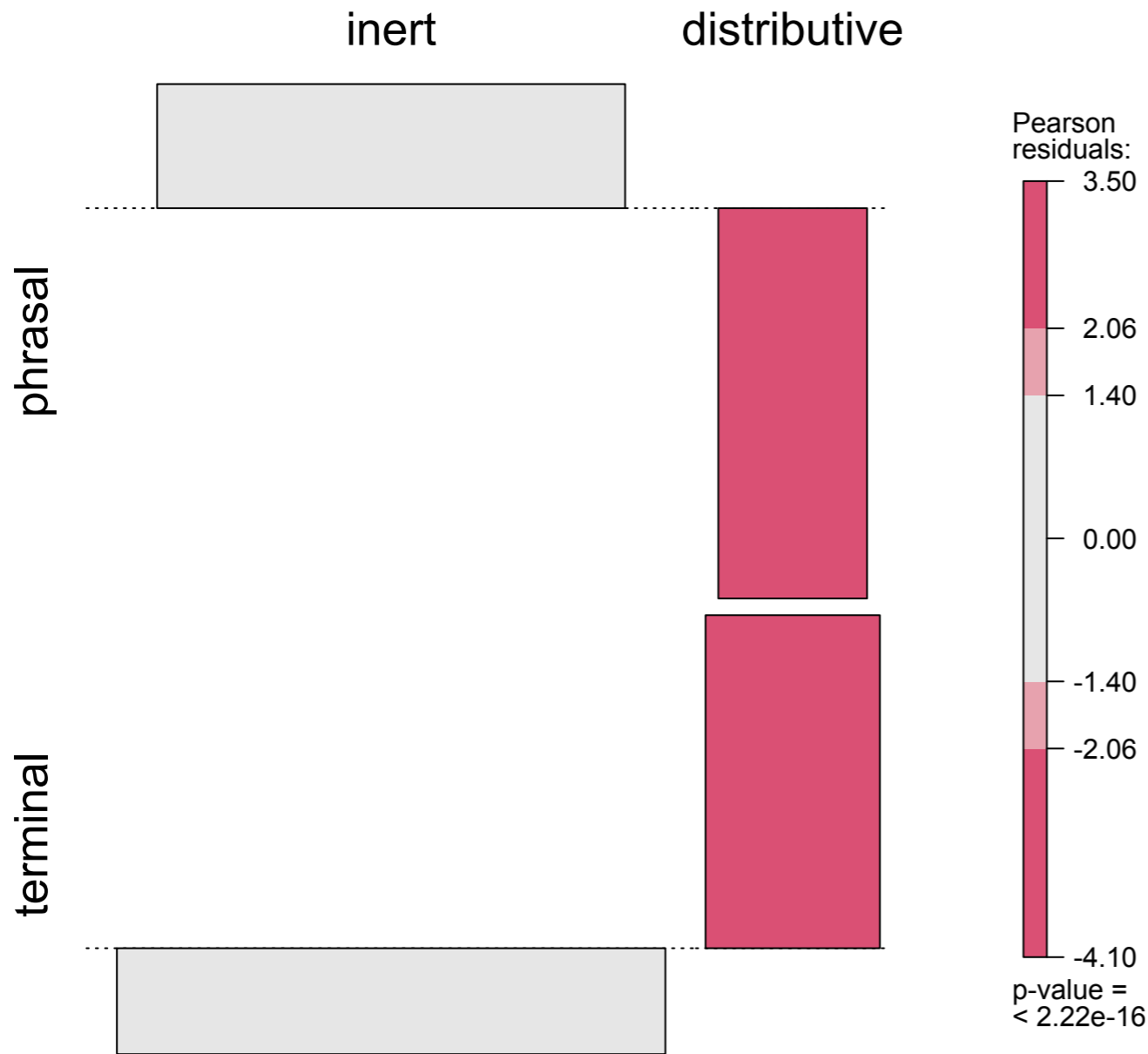
- Multivariate database on grammatical markers, mostly case, number, agreement, TAM, negation, dependent verb forms (data collected in various projects together with Johanna Nichols, Michael Riessler, and Lena Witzlack-Makarevich)
- 1572 markers from 466 languages from 188 stocks



- Computing the relative Mutual Information of all pairs of variables and remove the weakest associations in triplets (methods originally developed for detecting gene associations; Hausser & Strimmer 2009, Margolin et al. 2006)

# Exploring correlations: some findings (= possible universals)

Host type → Behavior ( $\hat{\pi}=0.10$ , reverse  $\hat{\pi}=.08$ )



Warlpiri (Pama-Nyungan; Hale et al. 1995)

a. [*maliki wiri*]=*ngki* =*ji* *yarlku-rnu*.  
 dog big=ERG =[PERF-]1SG.P bite-PT  
 ‘A big dog bit me.’

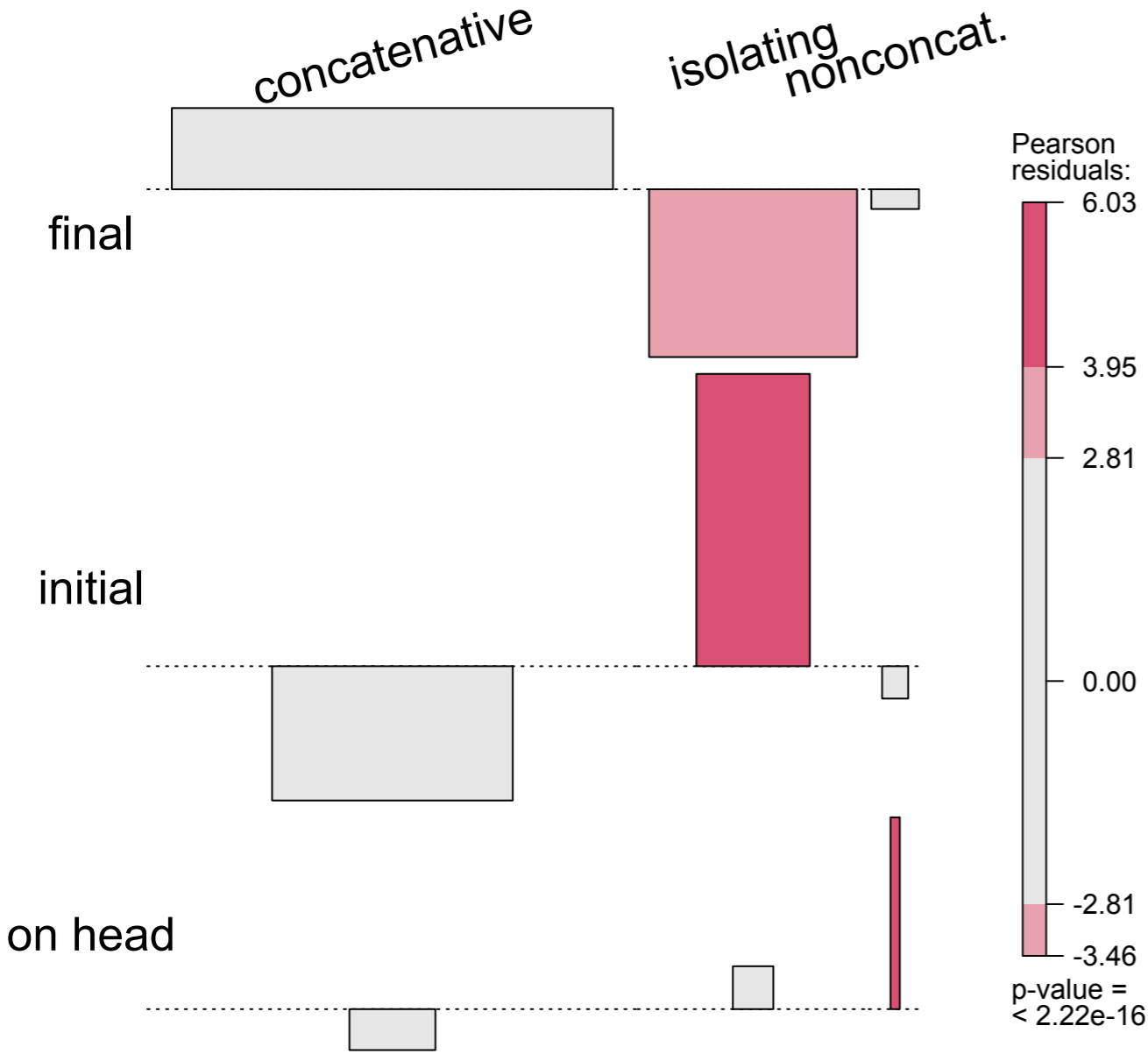
b. [*maliki*]=*rli* =*ji* *yarlku-rnu* [*wiri=ngki*].  
 dog=ERG =[PERF-]1SG.P bite-PT big=ERG  
 ‘A big dog bit me.’

German

[*Ein grosser Hund*] *biss* *mich*.  
 a.NOM big.NOM dog.NOM bite-PST 1sACC  
 ‘A big dog bit me.’

# Exploring correlations: some findings (= possible universals)

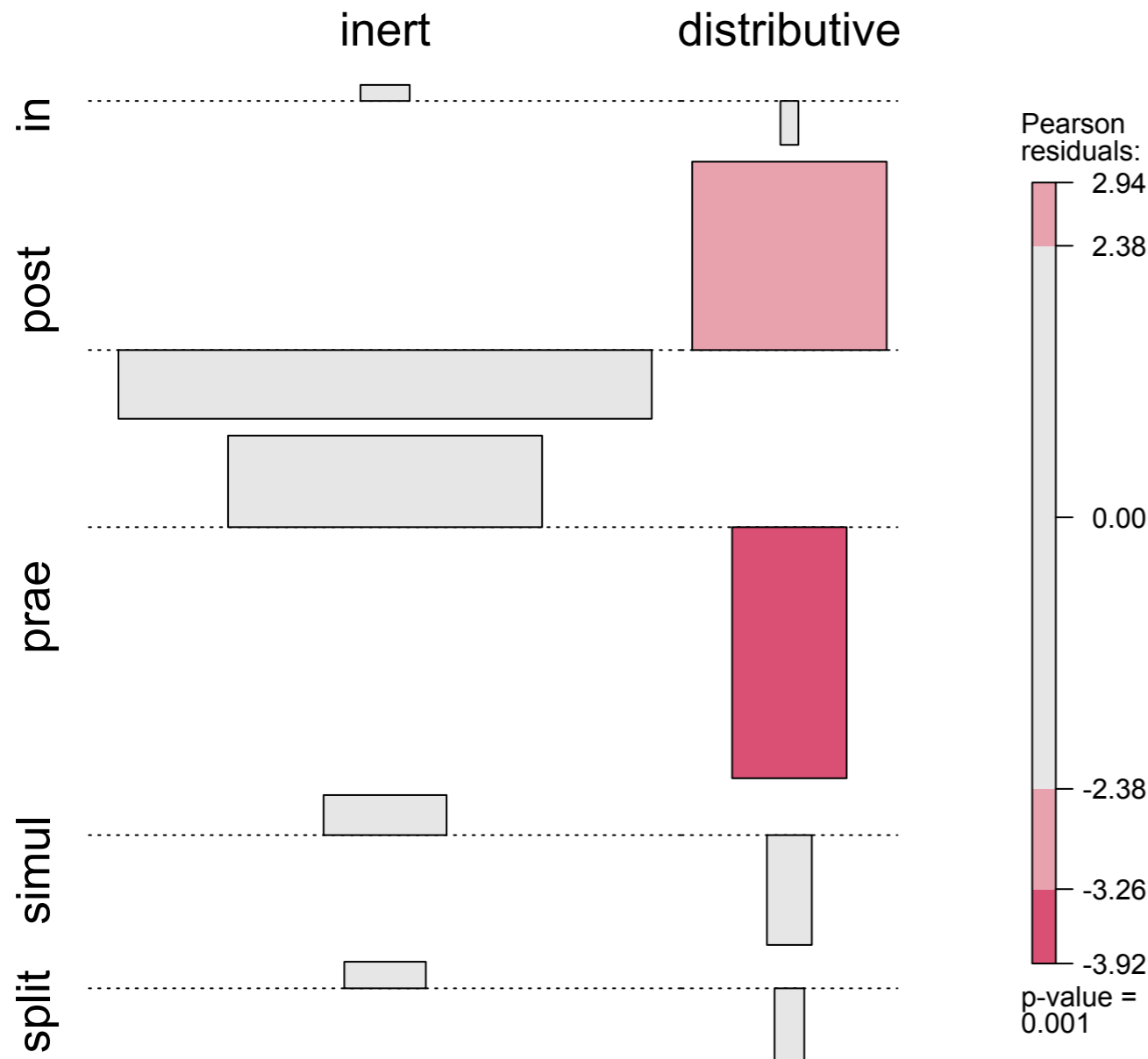
Phrase position → Phon. Fusion ( $\hat{\pi}=0.26$ , reverse  $\hat{\pi}=.21$ )



Less phonological interaction with phrase-initial than with other formatives

# Exploring correlations: some findings (= possible universals)

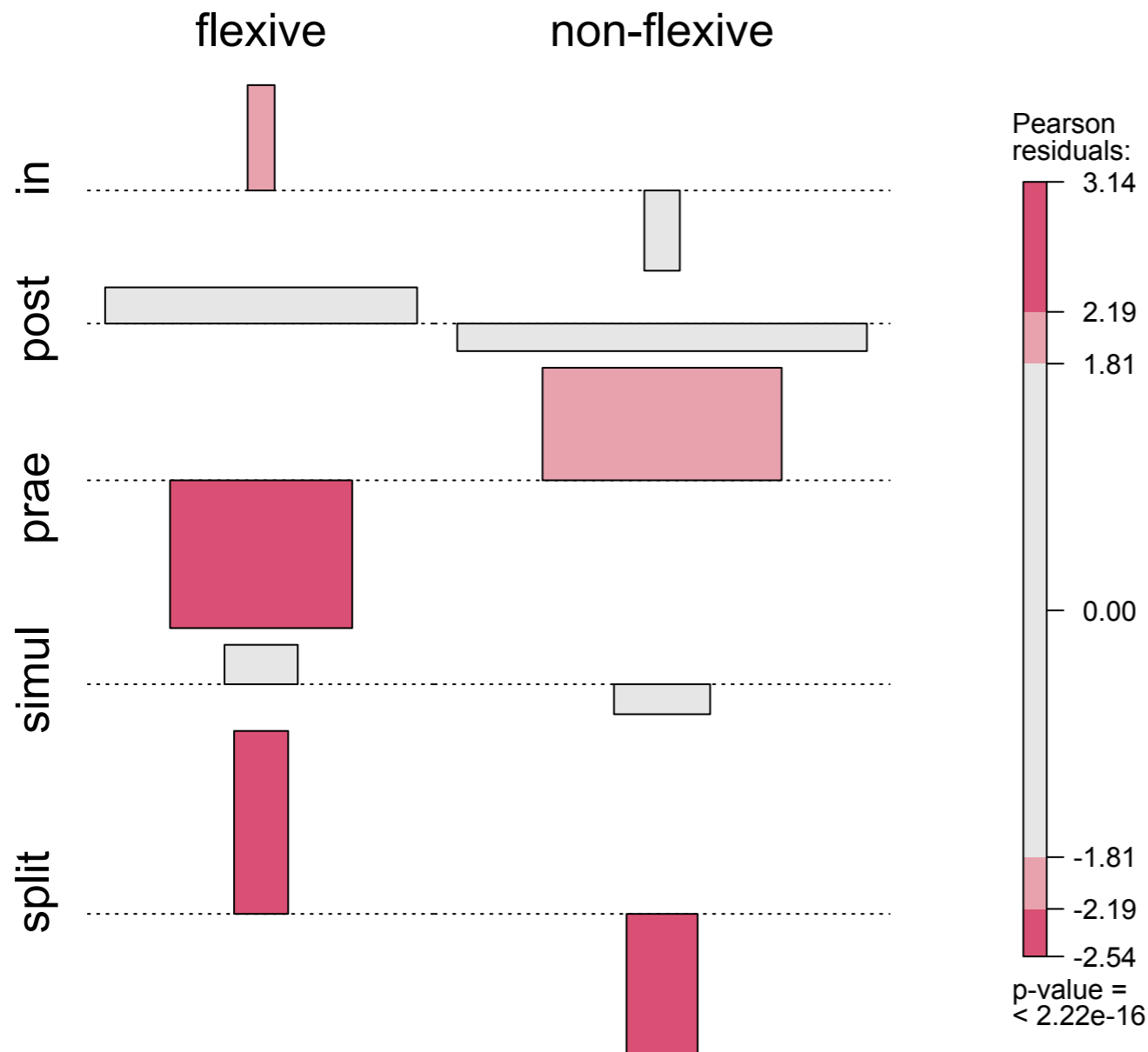
Phon. Position  $\leftrightarrow$  Behavior ( $\hat{\pi}=0.08$ , reverse  $\hat{\pi}=.05$ )



Spreading and stacking almost always with final elements (cf. Plank 1995: hardly any Präfixaufnahme)

# Exploring correlations: some findings (= possible universals)

Phon. Position  $\leftrightarrow$  Flexivity/Allomorphy ( $\hat{\pi}=0.04$ , reverse  $\hat{\pi}=.04$ )



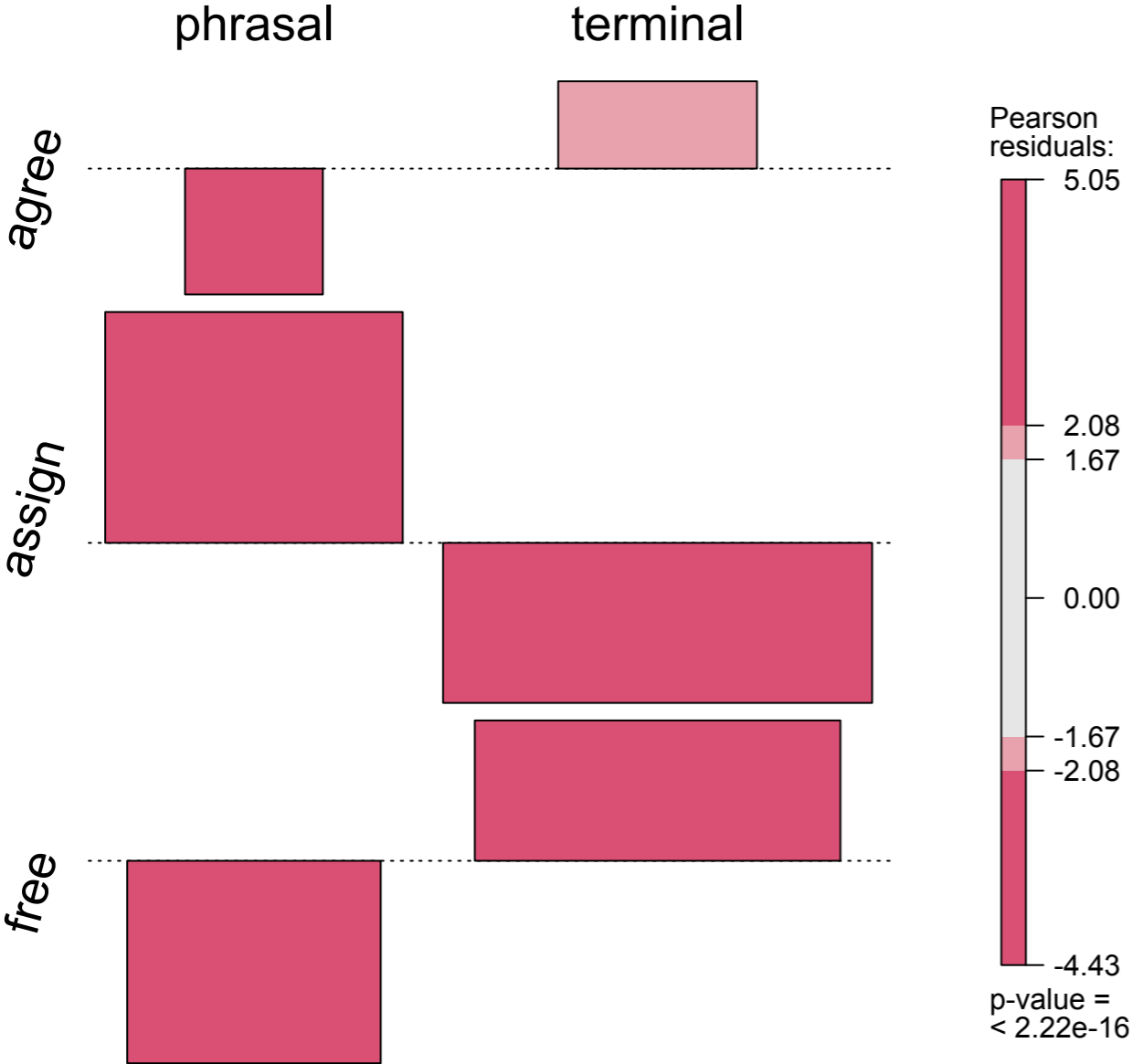
Less flexivity/allomorphy with initial than with other formatives;

Splits often associated with allomorphy (not just with position)



# Exploring correlations: some findings (= possible universals)

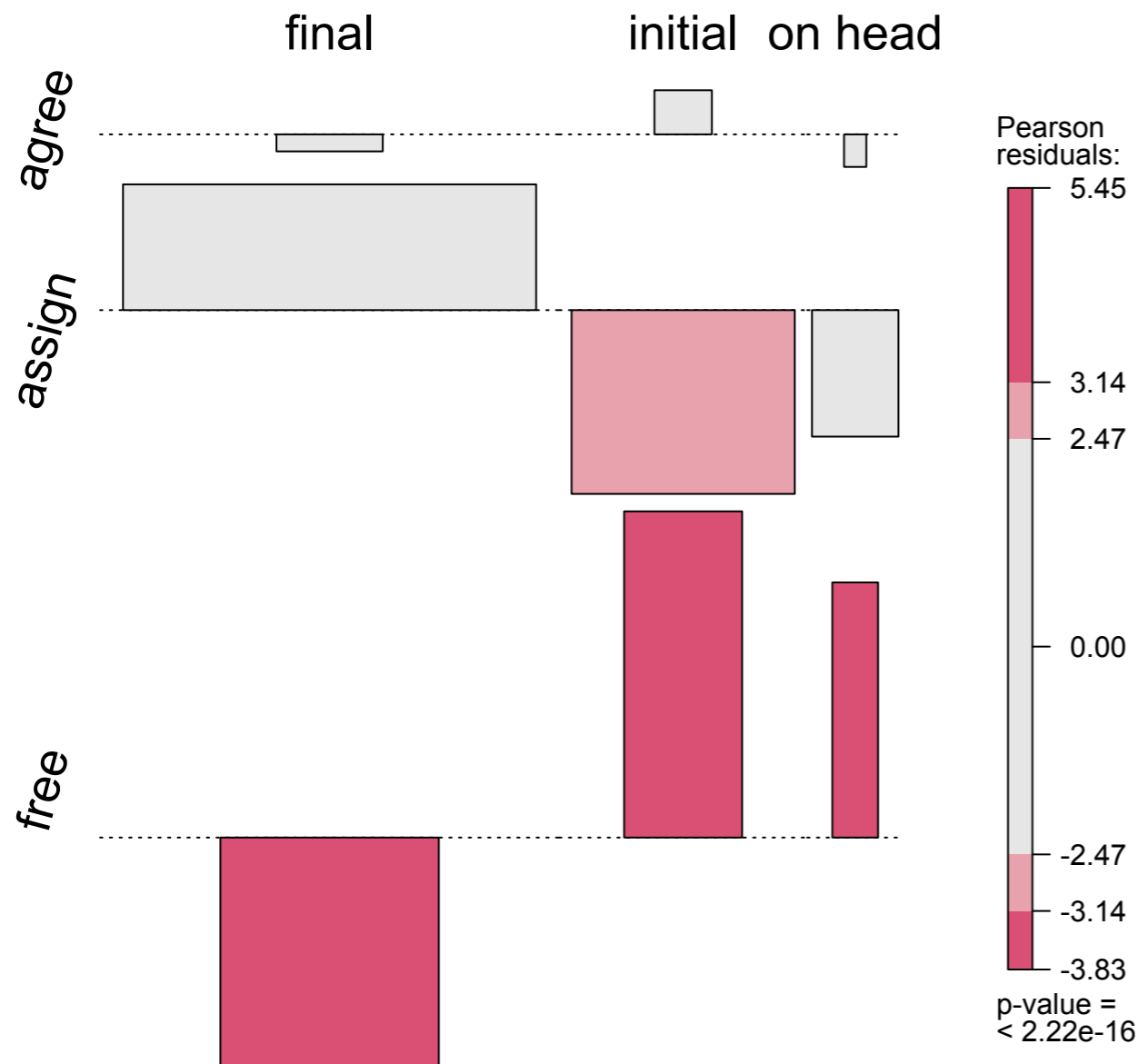
Licensing  $\leftrightarrow$  Host type ( $\hat{\pi}=0.10$ , reverse  $\hat{\pi}=.06$ )



Cases tend to be phrasal, all else terminal

# Exploring correlations: results (= possible universals)

Content type → Phrase position ( $\hat{\pi}=0.31$ , reverse  $\hat{\pi}=.21$ )



Free choice formatives (NEG, TAM) prefer initial or on head positions more than syntactically triggered formatives.

Saliency of initial positions?

## Two likely prototypes

---

1. Formatives that occur spontaneously (e.g. negation) or by agreement and that tend to be
  - initial (if phrasal)
  - prefixal/proclitic phonologically
  - preceding their syntactic host
  - isolating (phonologically)
  - nonflexive
  - inert (if terminal)
2. Formatives that are assigned (e.g. case) and that
  - final (if phrasal)
  - suffixal/enclitic phonologically
  - following their syntactic host
  - concatenative (phonologically)
  - flexive
  - distributive (if terminal)

# Methodological implication of Case Study A

---

- Multivariate typologies allow
  - detection of possible universals without first reducing the variation
  - by standard statistical mining techniques

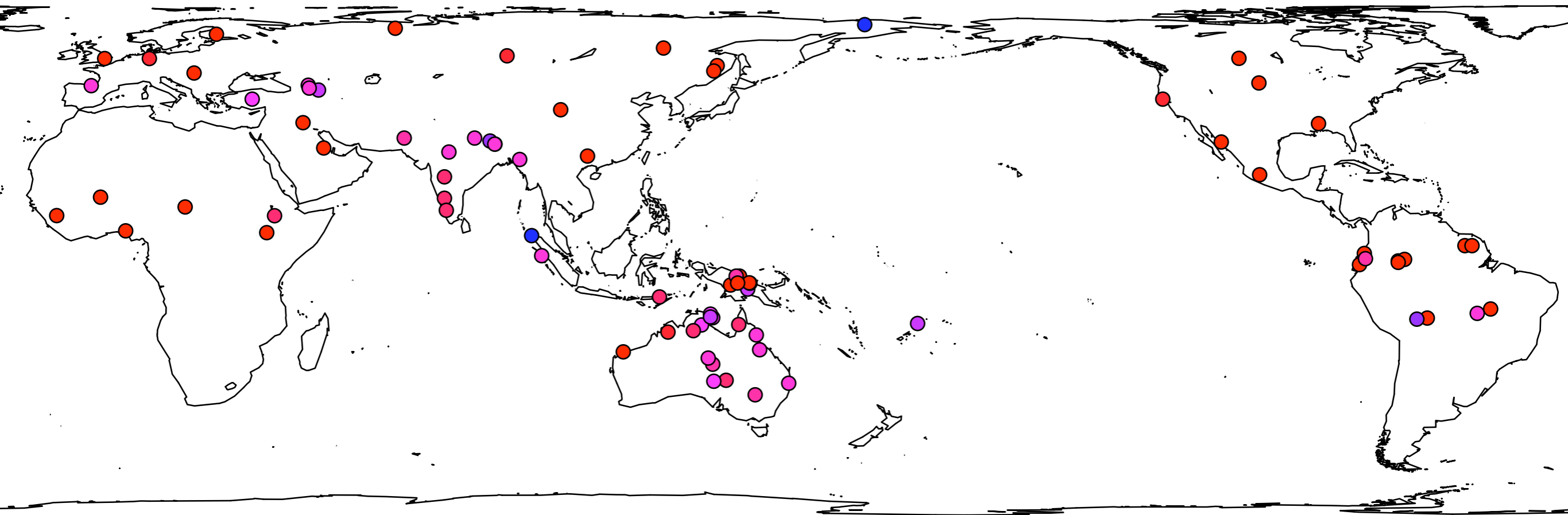
## Case Study B: testing a hypothesis on case alignment

- Multivariate database on case/adposition alignment (data collected together with Johanna Nichols, Lennart Bierkandt and especially Alena Witzlack-Makarevich)
- Alignments computed from lists of the arguments covered by each case, relative to properties of the arguments (reference, lexical class etc.) and the context (main vs dep. clause etc.)
- Some sample entries from Hindi (thanks to Alena Witzlack-Makarevich and Kamal Choudhary):

Alignment	Reference	Predicate Class	Clause	Dependency
$S=A_{ditr}=O=T \neq A=G$		A-DAT	non-PTCP_based	main
$S \neq A=A_{ditr} \neq O=G \neq T$	N-high	S-GEN	PTCP_based	main
$S=A=A_{ditr}=O=T \neq G$	N-low	<default>	non-PTCP_based	main
$S=A_{tr}=A_{ditr}=T \neq O=G$	N-high	<default>	non-PTCP_based	main
$S=G \neq A_{tr}=A_{ditr}=O=T$	N-low	S-DAT	non-PTCP_based	main
$S=O=G \neq A_{tr}=A_{ditr}=T$	N-high	S-DAT	non-PTCP_based	main
$S=O=T \neq A_{tr}=A_{ditr} \neq G$	N-low	<default>	PTCP_based	main

## Case Study B: testing a hypothesis on case alignment

- Hypothesis: there is a universal preference for S=A alignment mirroring an S=A preference in incremental processing (Bornkessel-Schlesewsky, Choudhary, Witzlack-Makarevich & Bickel 2008; also cf. Nichols 1993, Maslova & Nikitina 2007)
- Mapping the proportion of S=A case alignments per language (languages with exhaustive coding of predicate classes only,  $N = 80$ )



## Case Study B: testing a hypothesis on case alignment

---

- If the hypothesis is true, there must be a principle that pushes languages towards  $S=A$  over time, either by maintaining  $S=A$  or by innovating  $S=A$ .
- For this, we need to look at all conditions because the trend could show up in any of these, e.g. through
  - changing or keeping the types or sizes of predicate classes
  - changing or keeping alignments in some clause types or under some referential conditions (e.g. loosing ERG in Ns).
- And so, if the hypothesis holds, we expect that, across all conditions, there are significantly more families skewed towards  $S=A$  than there are families skewed towards  $S \neq A$  or not skewed at all,  
*independent of the geographical location.*

## Case Study B: testing a hypothesis on case alignment

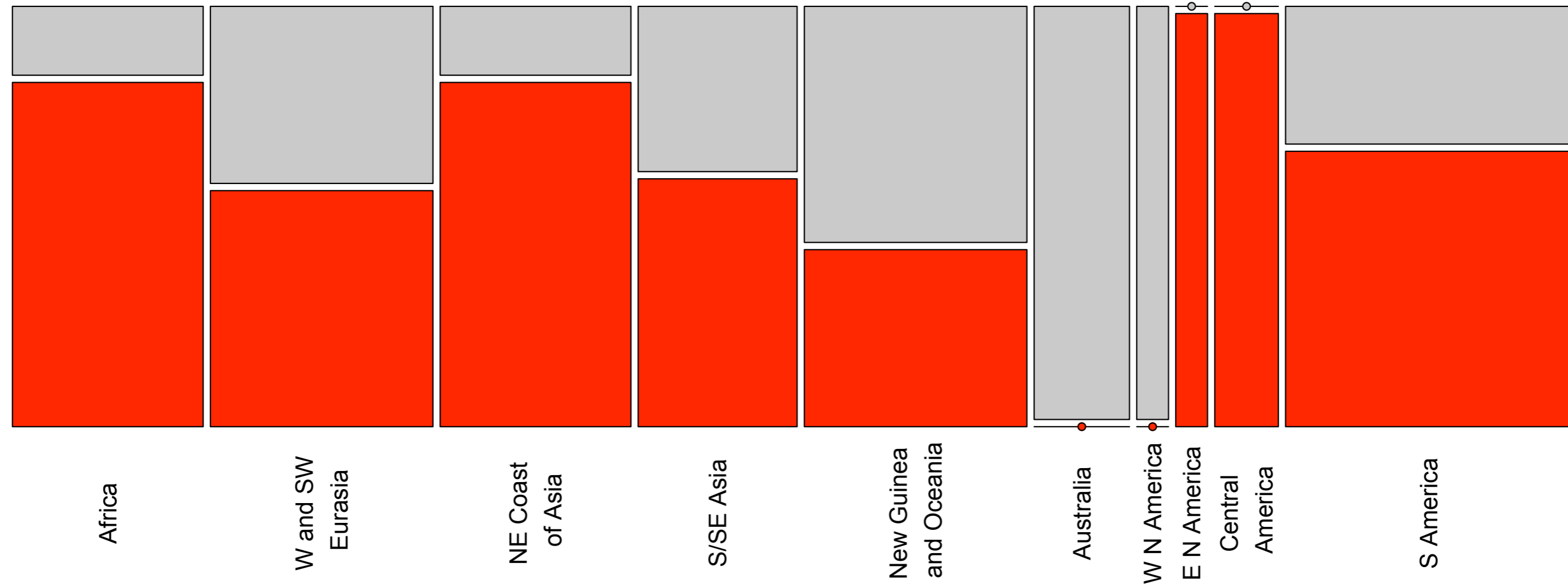
---

- Find the highest taxa that are not split across known areas and compute the proportions of S=A alignments within these taxa, giving variable weights to the lexically largest ('default') predicate class.
- Estimate the skewing of a family by Binomial Tests on whether the proportion of S=A exceeds what can be expected on the basis of all possible permutations
- and cross-tabulate the families against the areas.



# Case Study B: testing a hypothesis on case alignment

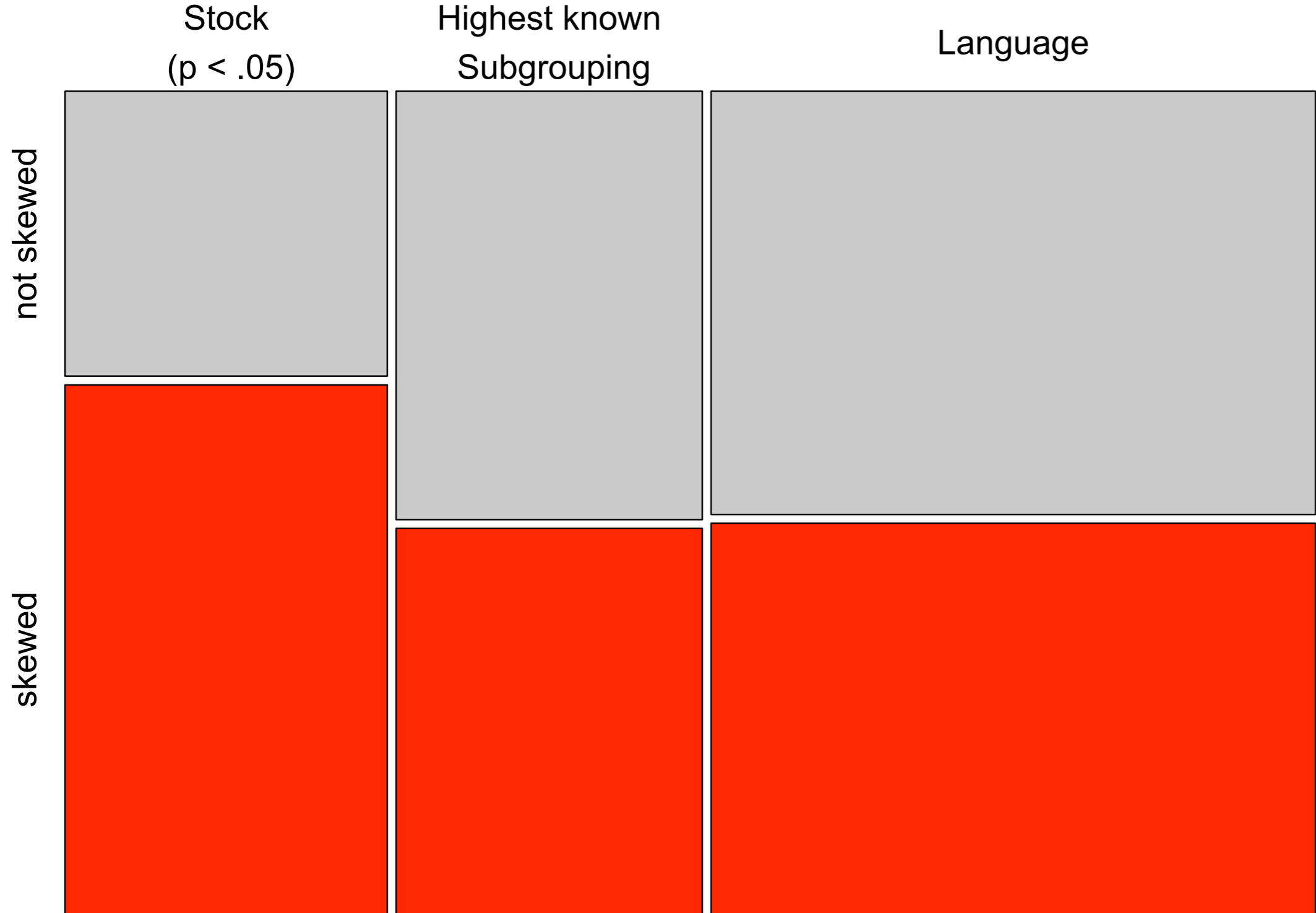
## Proportion of highest taxa with S=A skewing per area



- No significant differences between areas, Fisher Exact Test,  $p > .05$  (two-sided), regardless of the weighting factor for default classes
- But full confirmation will require more data from North America (currently only one sufficiently large family in each)

# Case Study B: testing a hypothesis on case alignment

Taxa skewed towards S=A (all areas, unweighted or weighted)



## Methodological implication of Case Study B

---

- Full datasets allow testing of diachronic trends because we get many datapoints per genealogical unit — from family to language — and can look into trends within these units, *allowing for all relevant conditions* (lexical classes, referential conditions)
- Reduced (‘simplified’) datasets do not allow testing for such trends, and they distort the true diversity.

# Overall conclusions

---

- Multivariate Typology seeks
  - typological variables as close to analytical tools in primary analysis (fieldwork), and vice-versa
  - typological patterns in full, not in artificially reduced diversity

# Overall conclusions

---

- The result is probabilistic patterns of distributional skewings
  - Instead of “if X spreads, it must be attached to terminals”, we get “... it is likely to ...”, i.e. statistical instead of absolute universals
    - Also, instead of traditional notions like ‘affix’ that entail bundles of strictly associated properties (‘terminal host’, ‘selective’ etc.), we get quantifiable degrees of association between properties
- More in line with other disciplines dealing with distributional skewings (e.g. genetics, ecology, economics).

# References

---

- Anderson, Stephen R., 1985. Inflectional morphology. In Shopen, Timothy (ed.) *Language typology and syntactic description*, vol. 3, 150 – 201. Cambridge: Cambridge University Press.
- Bickel, Balthasar, 2007. Typology in the 21st century: major current developments. *Linguistic Typology* 11, 239 – 251.
- Bickel, Balthasar, 2008. A general method for the statistical evaluation of typological distributions. Ms. University of Leipzig, [http://www.uni-leipzig.de/~bickel/research/papers/testing\\_universals\\_bickel2008.pdf](http://www.uni-leipzig.de/~bickel/research/papers/testing_universals_bickel2008.pdf).
- Bickel, Balthasar, in press. Capturing particulars and universals in clause linkage: a multivariate analysis. In Bril, Isabelle (ed.) *Clause-hierarchy and clause-linking: the syntax and pragmatics interface*. Amsterdam: Benjamins [preprint available at [http://www.uni-leipzig.de/~bickel/research/papers/multivariate\\_clauselinkage\\_bickel2009.pdf](http://www.uni-leipzig.de/~bickel/research/papers/multivariate_clauselinkage_bickel2009.pdf)].
- Bickel, Balthasar, Goma Banjade, Martin Gaenszle, Elena Lieven, Netra Paudyal, Ichchha P. Rai, Manoj Rai, Novel K. Rai, & Sabine Stoll, 2007. Free prefix ordering in Chintang. *Language* 83, 43 – 73.
- Bickel, Balthasar & Johanna Nichols, 2002. Autotypologizing databases and their use in fieldwork. In Austin, Peter, Helen Dry, & Peter Wittenburg (eds.) *Proceedings of the International LREC Workshop on Resources and Tools in Field Linguistics, Las Palmas, 26 - 27 May 2002*. Nijmegen: MPI for Psycholinguistics.
- Bickel, Balthasar & Johanna Nichols, 2009. Case marking and alignment. In Malchukov, Andrej & Andrew Spencer (eds.) *The Oxford Handbook of Case*, 304 – 321. Oxford: Oxford University Press.
- Bickel, Balthasar, Manoj Rai, Netra Paudyal, Goma Banjade, Toya Nath Bhatta, Martin Gaenszle, Elena Lieven, Iccha Purna Rai, Novel K. Rai, & Sabine Stoll, in press. The syntax of three-argument verbs in Chintang and Belhare (Southeastern Kiranti). In Malchukov, Andrej, Martin Haspelmath, & Bernard Comrie (eds.) *Studies in ditransitive constructions*. Berlin: Mouton de Gruyter [pre-print available at [http://www.uni-leipzig.de/~bickel/research/papers/ctn\\_ditrans\\_bickeletal.pdf](http://www.uni-leipzig.de/~bickel/research/papers/ctn_ditrans_bickeletal.pdf)].
- Corbett, Greville G., 2005. The canonical approach in typology. In Frajzyngier, Zygmunt, Adam Hodges, & David S. Rood (eds.) *Linguistic diversity and language theories*, 25 – 49. Amsterdam: Benjamins.
- Croft, William & Keith T. Poole, 2008. Inferring universals from grammatical variation: multidimensional scaling for typological analysis. *Theoretical Linguistics* 34, 1–37.
- Cysouw, Michael, 2003. Against implicational universals. *Linguistic Typology* 7, 89–110.

# References

---

- Cysouw, Michael, 2007. New approaches to cluster analysis of typological indices. In Köhler, Reinhard & Peter Grzбек (eds.) *Exact methods in the study of language and text*, 61–76. Berlin: Mouton de Gruyter.
- Evans, Nicholas & Stephen C. Levinson, in press. The Myth of Language Universals: language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*.
- Hale, Kenneth L., Mary Laughren, & Jane Simpson, 1995. Warlpiri. In Jacobs, Joachim, Arnim von Stechow, Wolfgang Sternefeld, & Theo Vennemann (eds.) *Syntax: ein internationales Handbuch zeitgenössischer Forschung, 2. Halbband*, 1430 – 1451. Berlin: Mouton de Gruyter.
- Haspelmath, Martin, 2007. Pre-established categories don't exist: consequences for language description and typology. *Linguistic Typology* 11, 119 – 132.
- Hausser, Jean & Korbinian Strimmer, 2009. Entropy inference and the James-Stein estimator, with application to nonlinear gene association networks. *Journal of Machine Learning Research* 10, xx–xx.
- Janssen, Dirk, Balthasar Bickel, & Fernando Zúñiga, 2006. Randomization tests in language typology. *Linguistic Typology* 10, 419 – 440.
- Lazard, Gilbert, 2006. *La quête des invariants interlangues: la linguistique est-elle une science?* Paris: Champion.
- Levinson, Stephen C., Sergio Meira, & the Language and Cognition Group, 2003. 'Natural concepts' in the spatial topological domain – adpositional meanings in crosslinguistic perspective: an exercise in semantic typology. *Language* 79, 485 – 516.
- Margolin, Adam, Ilya Nemenman, Katia Basso, Chris Wiggins, Gustavo Stolovitzky, Riccardo Favera, & Andrea Califano, 2006. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a Mammalian cellular context. *BMC Bioinformatics* 7, S7.
- Maslova, Elena & Tatiana Nikitina, 2007. Stochastic universals and dynamics of cross-linguistic distributions: the case of alignment types. Ms. Stanford University, <http://www.stanford.edu/~emaslova/Publications/Ergativity.pdf>.
- Nichols, Johanna, 1993. Ergativity and linguistic geography. *Australian Journal of Linguistics* 13, 39 – 89.
- Plank, Frans, 1995. (Re-)introducing Suffixaufnahme. In Plank, Frans (ed.) *Double case: agreement by Suffixaufnahme*, 3 – 110. New York: Oxford University Press.