Referential scales and differential case marking: A study using hierarchical models in Bayesian phylogenetics

Gerhard Jäger

Tübingen University

13th Conference of the Association for Linguistic Typology
Pavia, September 4, 2019

WORDS BONES GENES TOOLS
Tracking Linguistic, Cultural, and Biological Trajectories of the Human Past

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

DFG

Case alignment systems

Universal syntactic-semantic primitives

- three universal core roles
  - **S:** intransitive subject
  - **A:** transitive subject
  - **O:** transitive object

German

Der Junge ist dreckig.
the boy.NOM is dirty
'The boy is dirty.'

Der Junge wirft einen Stein.
DEF boy.NOM throw a.ACC stone
'The boy is throwing a stone.'

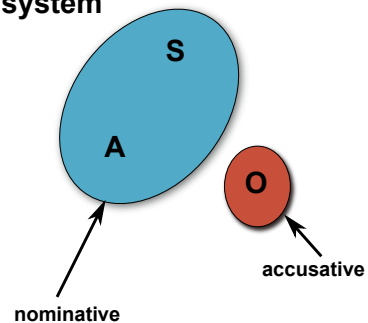
S
A
O

Kalkatungu (Australia)

Kaun muu-yan-ati
dress.ABS dirt-PROP-INCH
'The dress is dirty.'

Kuntu wampa-ngku kaun muu-yan-puni-mi.
not girl-ERG dress.ABS dirty-PROP-CAUS-FUT
'The girl will not dirty the dress.'

Accusative system

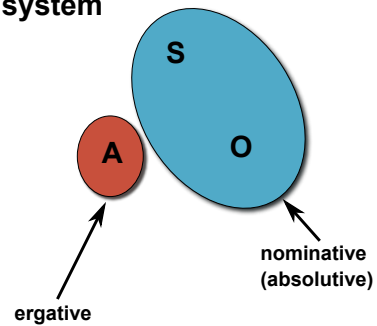


Latin

Puer puellam vidit.
boy.NOM girl.ACC saw *'The boy saw the girl.'*

Puer venit.
boy.NOM came *'The boy came.'*

Ergative system

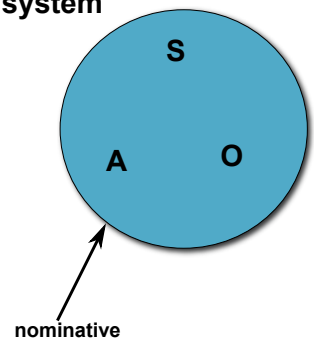


Dyirbal

ŋuma yabu-ŋgu bura-n.
father mother.ERG see-NONFUT
'The mother saw the father.'

ŋuma banaga-nu.
boy.NOM came *'The boy came.'*

Neutral system



Mandarin

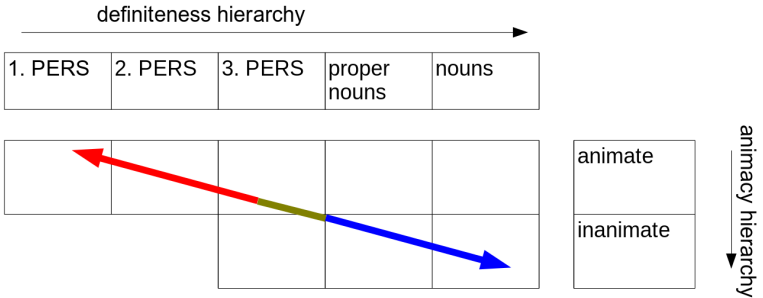
rén lái le.
person come CRS
'The person has come.'

zhāngsān mà lǐsì le ma.
Zhangsan scold Lisi CRS Q
'Did Zhangsan scold Lisi?'

- many languages have mixed systems
- e.g., some NPs have accusative and some have neutral paradigm, such as Hebrew

(1) Ha-seret her?a ?et-ha-milxama
the-movie showed acc-the-war
'The movie showed the war.'

(2) Ha-seret her?a (*?et-)milxama
the-movie showed (*acc-)war
'The movie showed a war'
(from Aissen, 2003)

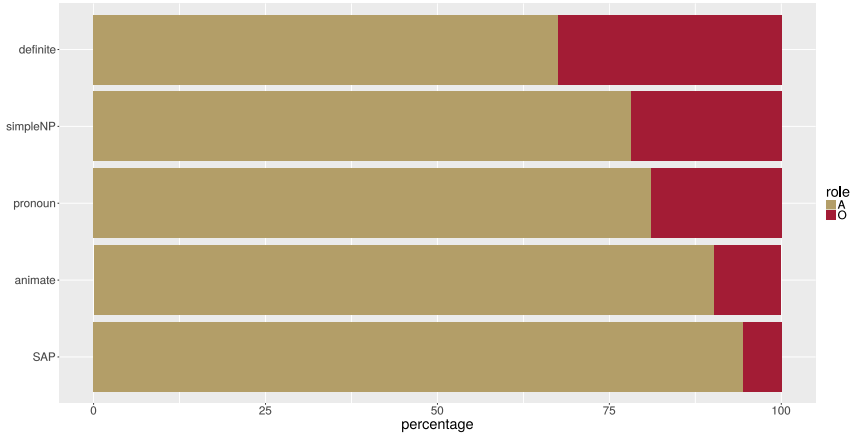


accusative

neutral or tripartite

ergative

probability P(syntactic role|prominence of NP)



A is prominent	A is non-prominent	O is prominent	O is non-prominent
e(rgative)	e(rgative)	a(ccusative)	a(ccusative)
e	e	a	z(ero)
e	e	z	a
e	e	z	z
e	z	a	a
...	...	...	...
z	e	z	z
z	z	a	a
z	z	a	z
z	z	z	a
z	z	z	z

actually attested:

1. **zzzz**: no case marking
2. **zzaa**: non-differential object marking
3. **zzaz**: harmonic differential object marking
4. **ezzz**: non-differential subject marking
5. **zeaz**: split ergative
6. **eeaz**: non-differential subject marking plus differential object marking
7. **ezzz**: dis-harmonic differential subject marking
8. **zezz**: harmonic differential subject marking
9. **zeaa**: harmonic differential subject marking plus non-differential object marking
10. **zzza**: dis-harmonic differential object marking

Differential case marking and referential scales

- received wisdom (Silverstein, 1976; Comrie, 1981; Aissen, 2003, , *inter alia*):
  - if object-marking is differential, upper segments of a referential hierarchy receive accusative marking
  - if object-marking is differential, lower segments of a referential hierarchy receive accusative marking
- Bickel et al. (2015):
  - large differences between macro-areas
  - no universal effects of referential scales on differential case marking

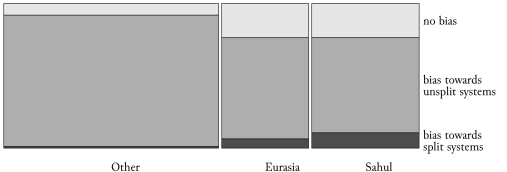


Figure 1: Estimated biases of families having split case markig for A across macro-areas. (The sizes of the individual tiles in the plot are proportional to frequencies, using the 'mosaic' plot technique provided by Meyer et al. 2006)

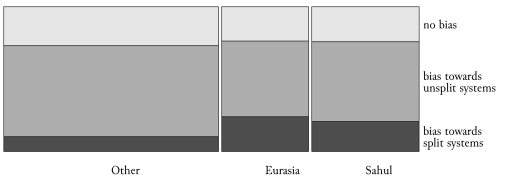
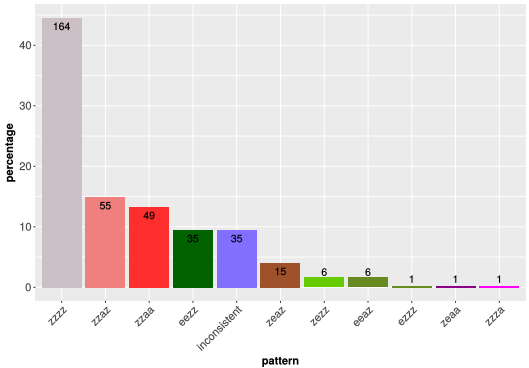


Figure 2: Estimated biases of families having split case markig for P across macro-areas (using the same mosaic plot techniques as in Figure 1)

Empirical distribution

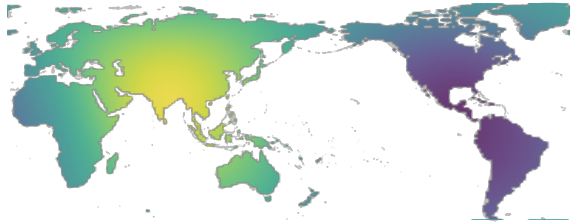
Bickel et al.'s (2015) sample

- genetically diverse sample of 460 case marking systems
- used here: 368 systems
  - one system per language
  - only languages with ISO code
  - only languages present in ASJP
- 2 out of 333 systems (99.4%) are obey the Silverstein hierarchy (not counting inconsistent states)

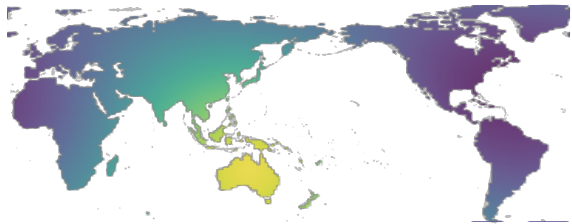


- differential object marking concentrated in Eurasia
- diffential subject marking concentrated in Sahul
- only cases of anti-DOM and anti-DSM (one instance of each) in North America

Differential object marking

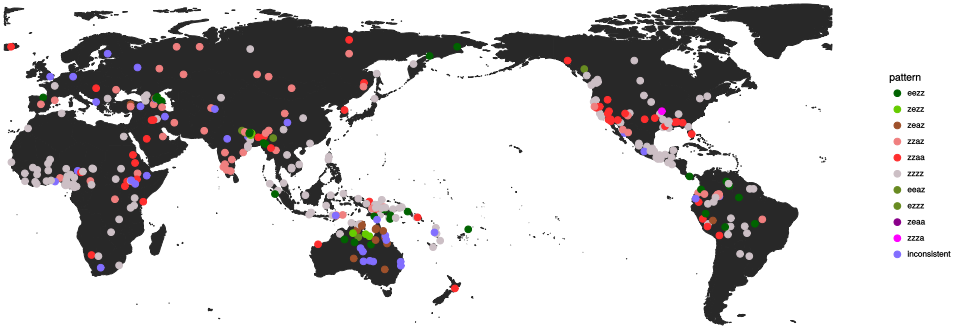


Differential subject marking

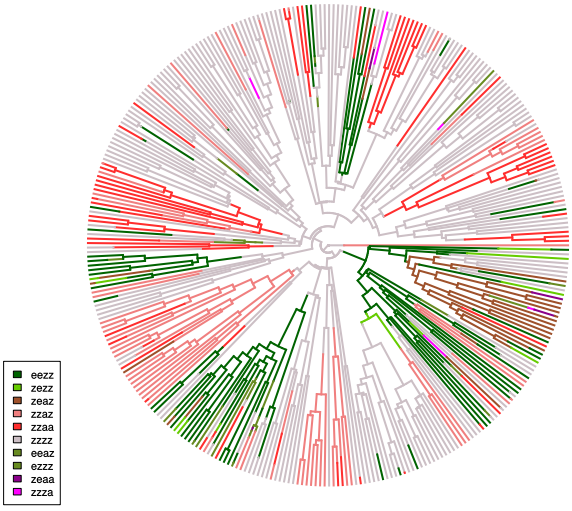


- languages are phylogenetically structured
- if two closely related languages display the same pattern, these are not two independent data points

$\Rightarrow$ we need to control for phylogenetic dependencies



Phylogenetic non-independence



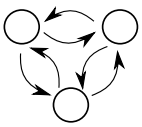
Maslova (2000):

> *"If the A-distribution for a given typology cannot be assumed to be stationary, a distributional universal cannot be discovered on the basis of purely synchronic statistical data."*

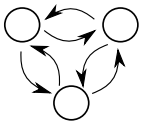
> *"In this case, the only way to discover a distributional universal is to* ***estimate transition probabilities*** *and as it were to 'predict' the stationary distribution on the basis of the equations in (1)."*

The phylogenetic comparative method

Markov process

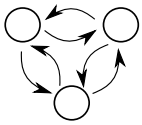


Markov process



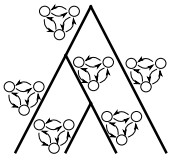
Phylogeny

Markov process

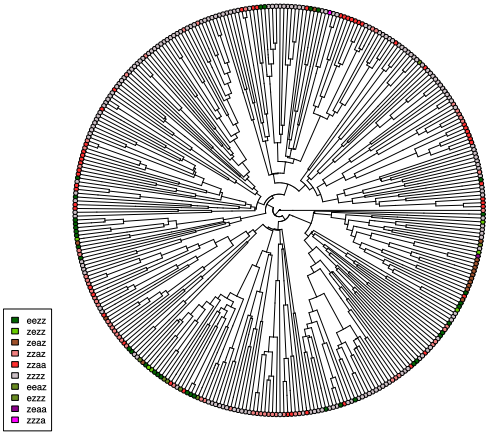


Phylogeny

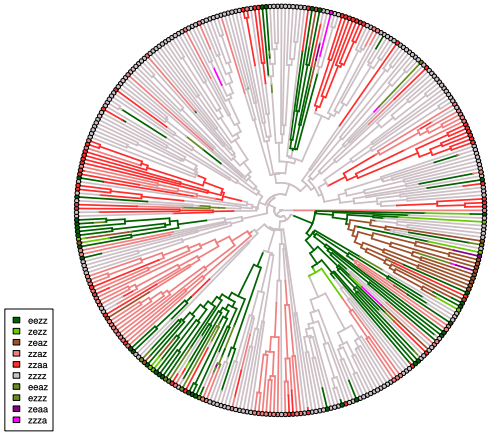
Branching process



- if phylogeny and states of extant languages are known...



- if phylogeny and states of extant languages are known...
- ... transition rates and ancestral states can be estimated based on Markov model

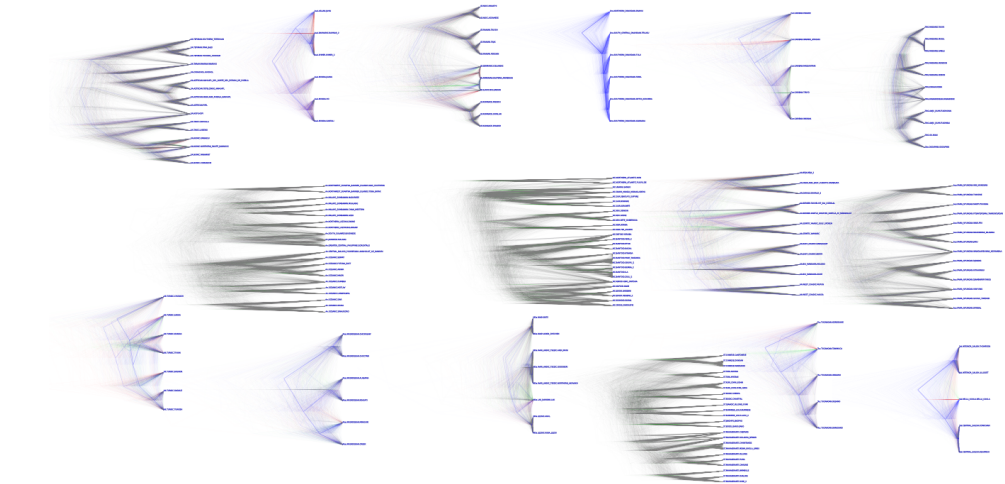


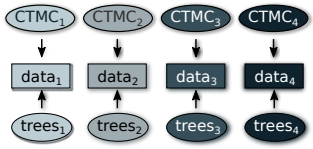
Cases in equilibrium

Phylogenetic trees for the case data

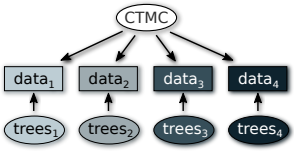
- 39 families and 63 isolates in the intersection of the Autotyp data and ASJP (Wichmann et al., 2018)
- for each of these families, I inferred a posterior distribution of 1,000 trees (using lexical data from ASJP) to reflect uncertainty in tree structure and branch length
- Glottolog tree was used as constraint tree

Phylogenetic trees for the case data

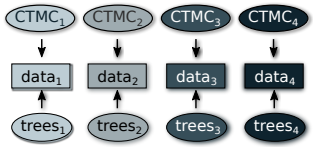




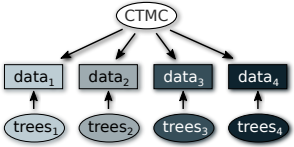
area-specific



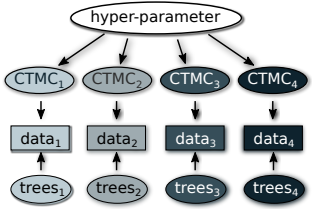
universal



area-specific



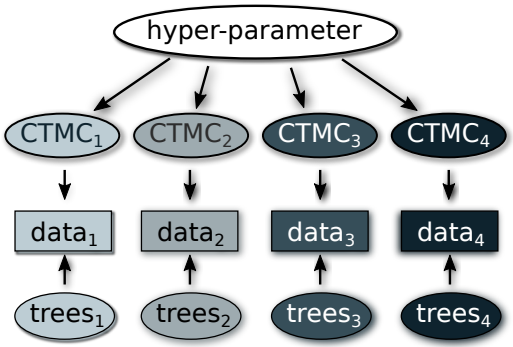
universal



hierarchical

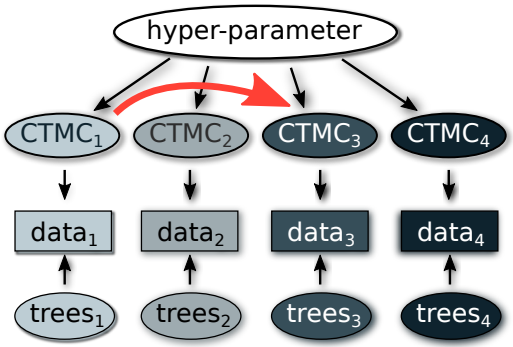
Hierarchical Models to capture areal effects

- each macro-area has its own parameters
- parameters are all drawn from the same distribution f
- shape of f is learned from the data
- prior assumption that there is little cross-area variation $\rightarrow$ can be overwritten by the data



Hierarchical Models to capture areal effects

- each macro-area has its own parameters
- parameters are all drawn from the same distribution f
- shape of f is learned from the data
- prior assumption that there is little cross-area variation $\rightarrow$ can be overwritten by the data
- enables **information flow across areas**

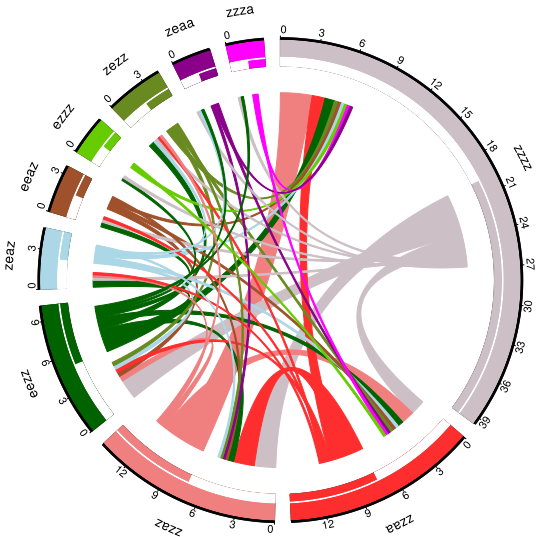


What about isolates?

- Continuous Time Markov Chain defines a unique **equilibrium distribution**
- hierarchical model assumes a different CTMC, and thus a different equilibrium distribution for each lineage
- by modeling assumption, root state of a lineage is drawn from this distribution (Uniformity Principle)
- isolates are treated as families of size 1, i.e., they are drawn from their equilibrium distribution

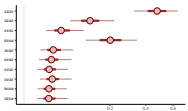
Results

Estimated transitions

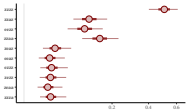


Estimated equilibrium distributions

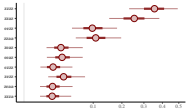
Africa



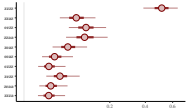
Americas



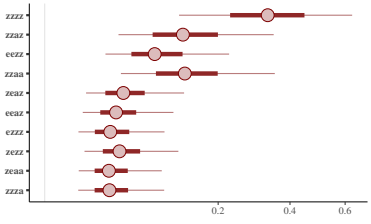
Eurasia



Sahul



posterior prediction



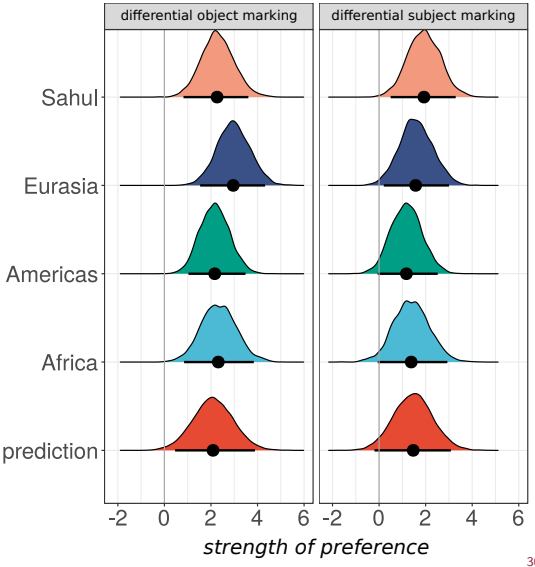
Preference for scale-respecting differential case marking

- **strength of preference** of DOM over anti-DOM:

$$\log \frac{P(..\text{az})}{P(..\text{za})}$$

- DSM over anti-DSM:

$$\log \frac{P(\text{ze}..)}{P(\text{ez}..)}$$



Conclusion

- considerable variation between macroareas concerning the dynamic process governing the diachrony of alignment systems, and the resulting long-term averages
- still, consistent preference for DOM/DSM over anti-DOM/DSM

Judith Aissen. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21(3):435–483, 2003.
Balthasar Bickel, Alena Witzlack-Makarevich, and Taras Zakharko. Typological evidence against universal effects of referential scales on case alignment. In Ina Bornkessel-Schlesewsky, Andrej L. Malchukov, and Marc D. Richards, editors, *Scales and hierarchies: A cross-disciplinary perspective*, pages 7–43. de Gruyter, Berlin/Munich/Boston, 2015.
Georg Bossong. *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Günther Narr Verlag, Tübingen, 1985.
Bernard Comrie. *Language Universals and Linguistic Typology*. Basil Blackwell, Oxford, 1981.
Gerhard Jäger. Phylogenetic inference from word lists using weighted alignment with empirically determined weights. *Language Dynamics and Change*, 3(2):245–291, 2013.
Gerhard Jäger. Support for linguistic macrofamilies from weighted sequence alignment. *Proceedings of the National Academy of Sciences*, 112(41):12752–12757, 2015. doi: 10.1073/pnas.1500331112.
Gerhard Jäger. Global-scale phylogenetic linguistic inference from lexical resources. arXiv:1802.06079, 2018.
Gerhard Jäger and Søren Wichmann. Inferring the world tree of languages from word lists. In S. G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Feher, and T. Verhoef, editors, *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*, 2016. Available online: http://evolang.org/neworleans/papers/147.html.
Elena Maslova. A dynamic approach to the verification of distributional universals. *Linguistic Typology*, 4(3):307–333, 2000.
Mark Pagel and Andrew Meade. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *The American Naturalist*, 167(6): 808–825, 2006.
Mark Pagel and Andrew Meade. BayesTraits 2.0. software distributed by the authors, November 2014.
Hugo Reyes-Centeno, Katerina Harvati, and Gerhard Jäger. Tracking modern human population history from linguistic and cranial phenotype. *Scientific Reports*, 6, 2016.
Frederik Ronquist and John P. Huelsenbeck. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12):1572–1574, 2003.
Michael Silverstein. Hierarchy of features and ergativity. In R. M. W. Dixon, editor, *Grammatical Categories in Australian Languages*, pages 112–171. Australian Institute of Aboriginal Studies, Canberra, 1976.
Søren Wichmann, Eric W. Holman, and Cecil H. Brown. The ASJP database (version 18). http://asjp.clld.org/, 2018.