



Differential object marking: A cross-linguistic study

Kaius Sinnemäki

(kaius.sinnemaki@helsinki.fi)

University of Helsinki

Workshop "Differential object marking: Theoretical
and empirical issues"

Helsinki, August 28, 2009



Background 1/2

- Differential object marking (DOM) has been studied a lot during the last 30 years.
 - The focus on DOM in single languages:
 - **Spanish** (Leonetti 2004), **Guaraní** (Shain 2008), **Mongolian** (Guntsetseg 2008), **Korean** (Kwon & Zribi-Hertz 2008), **Yongren Lolo** (Gerner 2008), **Turkish** (Kornfilt 2009).
 - Or on the factors that affect DOM (e.g. animacy, definiteness, topicality, etc).

Background 2/2

- Little attention has been given to DOM from a **cross-linguistic perspective**.
 - “...we are not aware of a world-wide and sufficiently sampled database on differential case-marking...”
(Bickel & Nichols 2009)
- Yet, Jäger (2007: 102) states: “Almost all accusative languages have DOM.”
 - Compare to Maslova and Nikitina (2007), who sampled 400 languages: **41% of accusative languages had DOM** (cf. McGregor (2009, in press) for recent typological work on ergativity).



Purpose of study 1/2

- A preliminary typological study on DOM in relation to 1) **areas** and 2) **word order types**.
 - Known that case correlates with **verb-final** order and **Eurasia** as a linguistic area (e.g. Bickel 2008).
- Focus on **overt case marking of direct objects** (O; the more patient-like argument of a transitive verb) in transitive clauses. Included marking by affixes, tonal or morphophonological alternations, or isolating formatives. Only **non-pronominal arguments** studied.

Purpose of study 2/2

- No attention given to the **semantic, pragmatic, and syntactic** factors that affect DOM.
 - A cross-linguistic study of these factors would require at least some corpus studies in each language.
 - I simply noted what the sources mentioned about case marking of O and possible restrictions to it.
- I merely contrast **1) differential marking** with **2) consistent/split marking**. (Note: this is not about alignment, but about overt marking.)



Definitions 1/3: DOM

- Differential marking: some objects unmarked while others marked.
 - In Kannada (Lidz 2006), object case is optional for inanimate nouns (1), but obligatory for animates (2).
- (1) Naanu pustaka(-vannu) huDuk-utt-idd-eene.
1SG.NOM book(-ACC) look.for-NPST-be-1SG
'I am looking for a book.' (p. 11)
- (2) Naanu sekretari-yannu/*sekretari huDuk-utt-idd-eene.
1SG.NOM secretary-ACC /*secretary look.for-NPST-be-1SG
'I am looking for a secretary.' (p. 11)



Definitions 2/3: “consistent”

- Consistent marking: all objects are marked with the same marker.
 - Imbabura Quechua (Cole 1985) uses object case regardless of the object being e.g. proper noun (3) or indefinite mass noun (4).

(3) Juzi Marya-ta juya-n-mi.
José María-ACC love-3-VAL
'José loves María.' (p. 98)

(4) Ñuka-ka aycha-ta miku-niñuka wawki tanda-ta miku-j-shna.
1-TOP meat-ACC eat-1 1 brother bread-ACC eat-NMLZ-as
'I eat as much meat as my brother eats bread.' (p. 66)



Definitions 3/3: “split”

- For split marking different case markers are used depending on the context.
 - Evenki uses *-βa* to case-mark (roughly) definite objects (5), while it uses *-jə* to case-mark (roughly) indefinite objects (6) (Bulatova & Grenoble 1999: 8-9).

(5) Bii asii **-βa** əəri-m.
1SG woman-ACC call-1SG
'I call the woman.' (p. 8)

(6) žəβgəə-**jə** gamii, aja bi-mčə.
food -ACC.INDF take good be-COND
'It would be good to get some (any) food.' (p. 9)

Data and sampling 1/2

- Languages classified into **genera** (e.g. Celtic): counts based on counting distinct values in genera to control for the most obvious genealogical bias (Dryer 1989).
 - No particular sampling method: **aiming at maximum genealogical coverage**.
 - Data from language descriptions, sources dealing with DOM, and *The World Atlas of Language Structures* (WALS, Comrie 2005 and Dryer 2005).
- Languages classified into six macro-areas (Dryer 1992).

Data and sampling 2/2

- **678 languages** from **371 genera** (77% coverage).

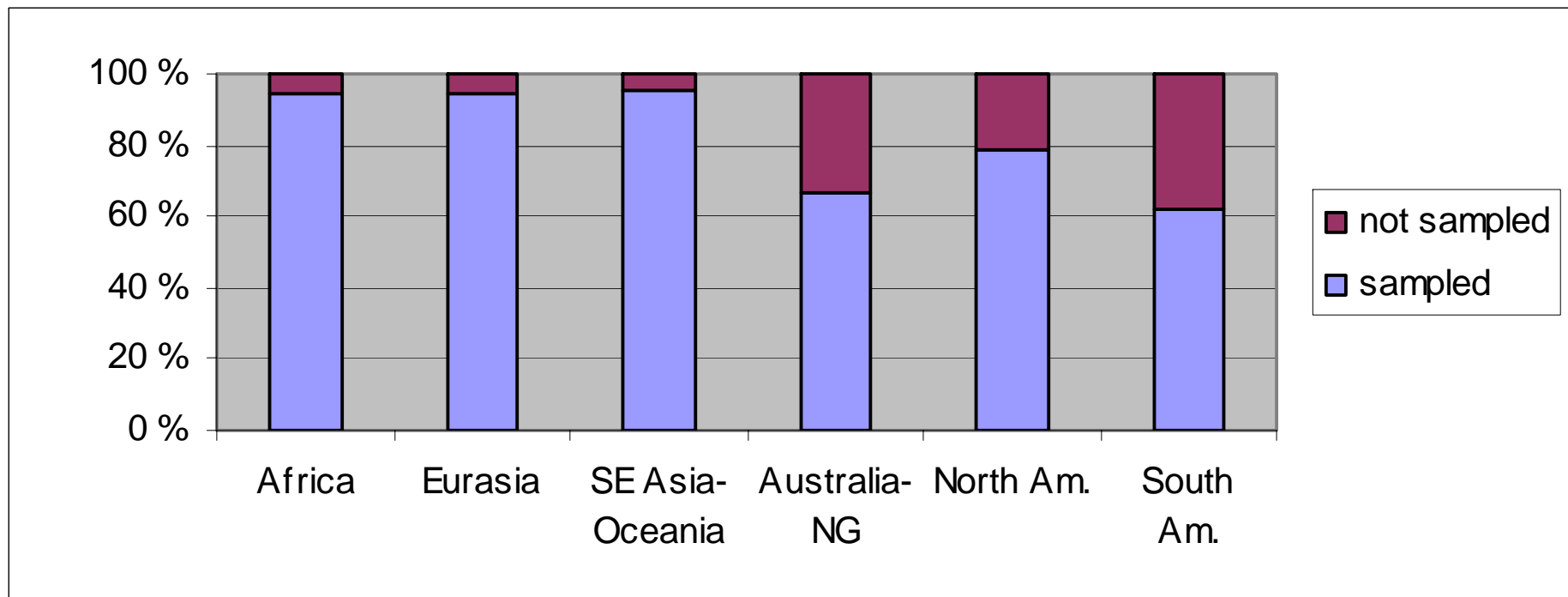


Table: Genealogical coverage of the sample by macro-area.



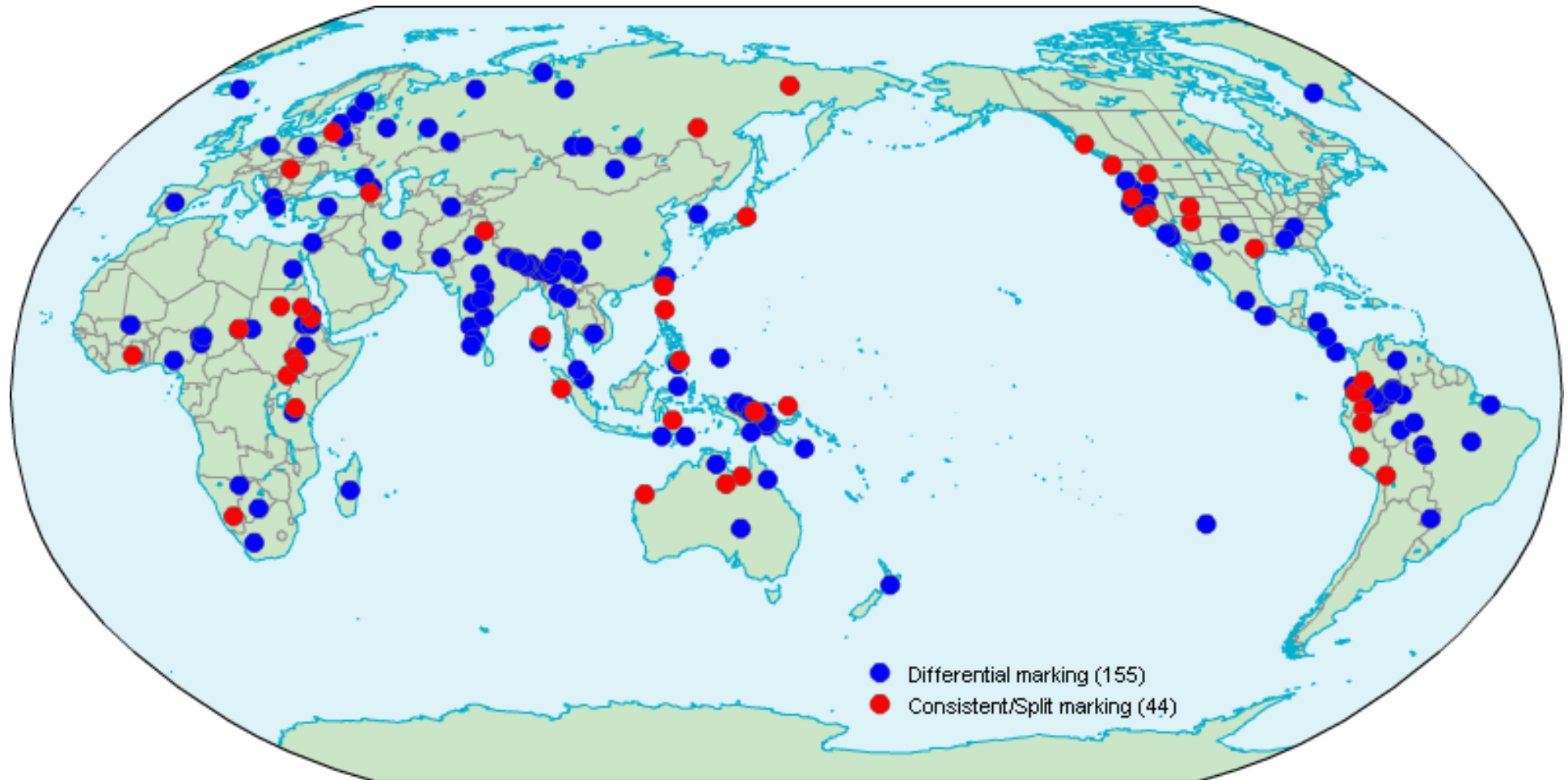
Results 1/5: descriptive

Case marking of O	N	%	% of O-case
None	272	65	
Differential	104	25	72
Consistent/Split	41	10	28
Total	417	100	100

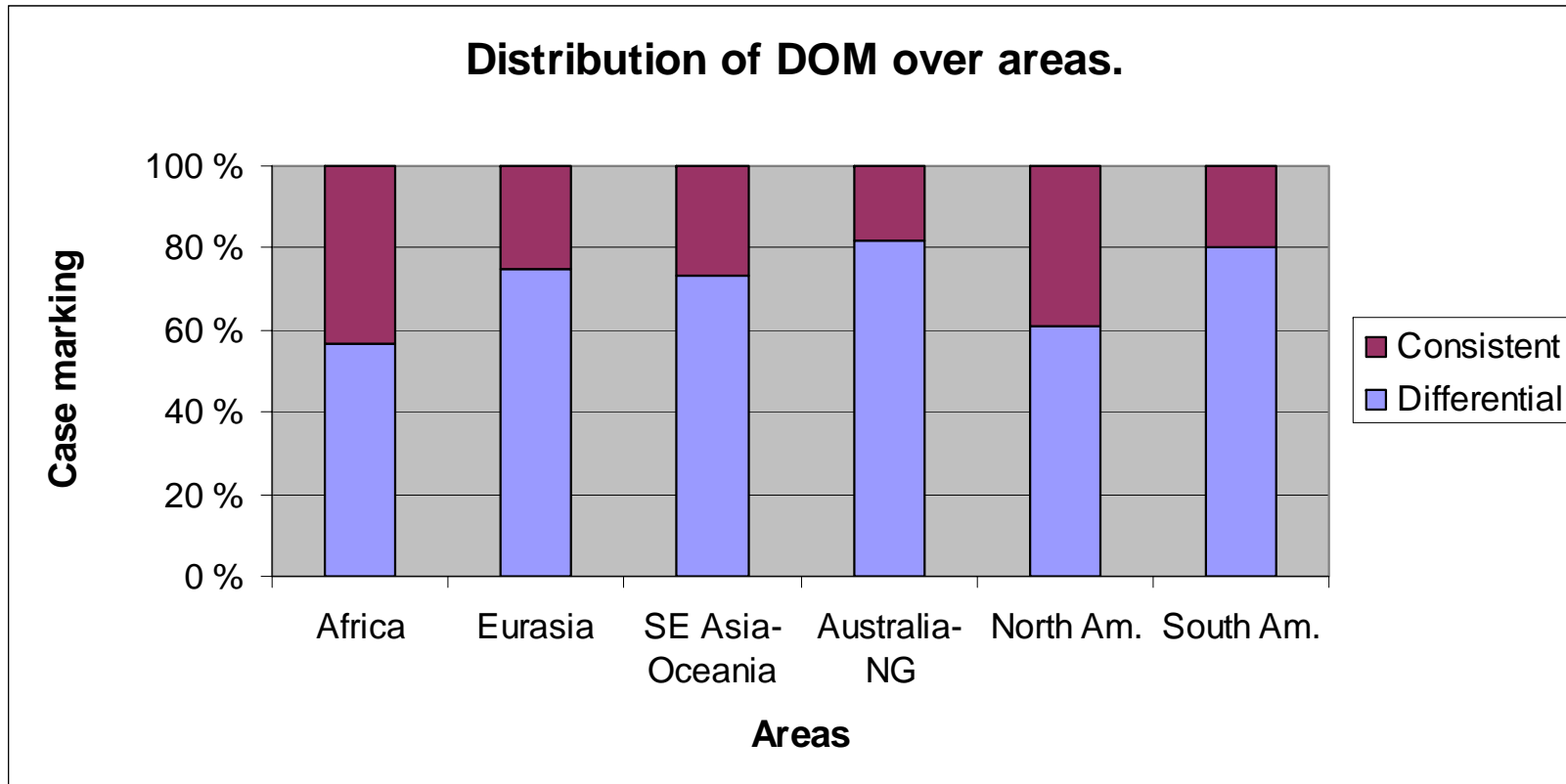
- **35% of languages had case marking of O** (cf. 36% in Maslova & Nikitina 2007), **>70% of these had DOM** and less than 1/3 had consistent/split marking.



Results 2/5: map

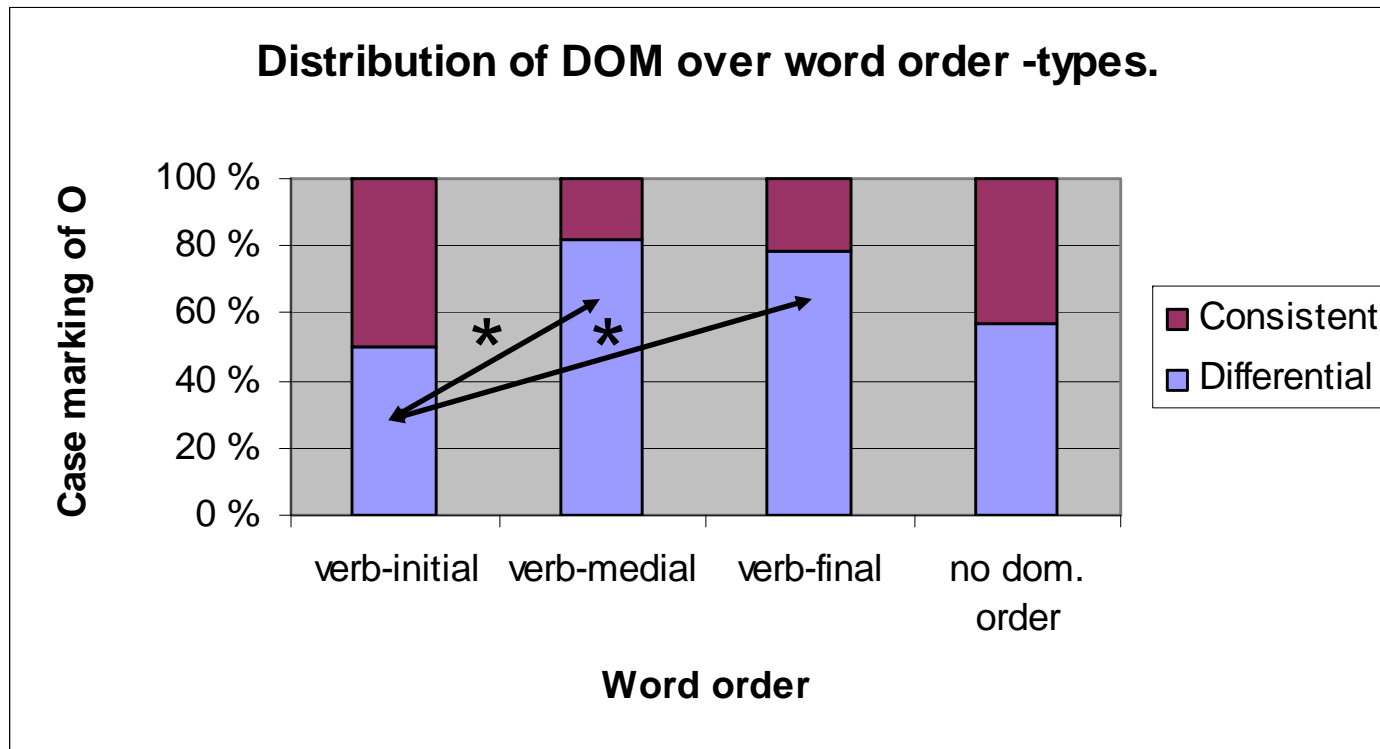


Results 3/5: areas



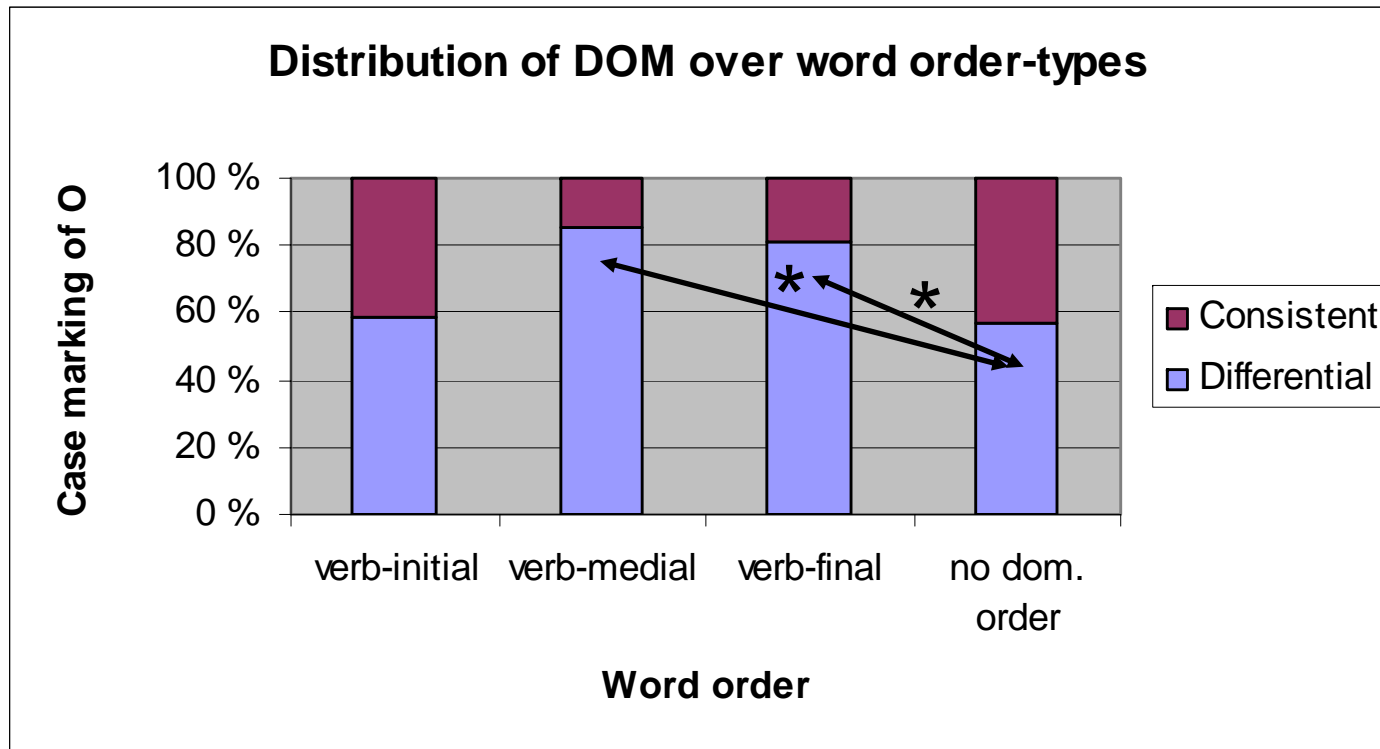
- **No statistical effect for areal distribution** ($p = 0.329$, tested with randomized chi-square).

Results 4/5: word order



- **A significant effect for word order** ($p = 0.0135$), mostly caused by verb-initial languages.

Results 5/5: word order



- **Split-type removed:** A significant effect for word order ($p = 0.0278$), mostly caused by **no dom. order type**.

Discussion

- Case marking of O is likely to be differential.
 - Why is consistent marking so rare?
 - Why do some languages mark all Os?
- DOM clustered in Eurasia, Meso-America, and the Amazon - no statistical effect.
 - Differential and consistent marking may be equally (un)likely to diffuse areally.
- A small statistical effect for verb-initial order.
 - This was a consequence of including languages with splits. If splits excluded, an effect for no dom. order: free word order well-motivated by consistent marking.



Future research

- Design a typology for DOM with a smaller sample (cf. McGregor (in press) for a typology for ergativity).
- Investigate the cross-linguistic distributions of DOM properties.

Thank you!



References 1/2

- Bickel, B. & J. Nichols 2009. The geography of case. *The Oxford Handbook of Case*, A. Malchukov & A. Spencer (eds.), 479-93. Oxford: OUP.
- Bickel, B. 2008 (ms.). A general method for the statistical evaluation of empirical universals. <<http://www.uni-leipzig.de/~bickel/>>.
- Bulatova, N. & L. Grenoble 1999. Evenki. *Languages of the World/Materials* 141. München: Lincom Europa.
- Cole, P. 1985. Imbabura Quechua. London: Croom Helm.
- Comrie, B. 2005. Alignment of case marking. *The World Atlas of Language Structures*, M. Haspelmath et al. (eds.), 398-405. Oxford: OUP.
- Dryer, M. 1989. Large linguistic areas and language sampling. *Studies in Language* 13: 257-92.
- Dryer, M. 1992. The Greenbergian word order correlations. *Language* 68: 81-138.
- Dryer, M. 2005. Order of subject, object, and verb. *The World Atlas of Language Structures*, M. Haspelmath et al. (eds.), 330-3. Oxford: OUP.
- Gerner, M. 2008. Ambiguity-driven differential object marking in Yongren Lolo. *Lingua* 118: 296-331.
- Guntsetseg, D. 2008. Differential object marking in Mongolian. *Working Papers of the SFB 732, Incremental Specification in Context* 1: 53-69.



References 2/2

- Jäger, G. 2007. Evolutionary game theory and typology: A case study. *Language* 83(1): 74-109.
- Kornfilt, J. 2009. DOM and two types of DSM in Turkish. *Differential Subject Marking*, H. de Hoop and P. de Swart (eds.), 79-111. Dordrecht: Springer.
- Kwon, S.-N. & A. Zribi-Hertz 2008. Differential function marking, case, and information structure: Evidence from Korean. *Language* 84(2): 258-99.
- Leonetti, M. 2004. Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics* 3: 75-114.
- Lidz, J. 2006. The grammar of accusative case in Kannada. *Language* 82(1): 10-32.
- Maslova, E. and T. Nikitina 2007 (ms.). Stochastic universals and dynamics of cross-linguistic distributions: the case of alignment types. <<http://www.stanford.edu/~emaslova/>>.
- McGregor, W.B. (in press). Optional ergative case marking systems in a typological-semiotic perspective. *Lingua*.
- McGregor, W.B. 2009. Typology of ergativity. *Language and Linguistics Compass* 3(1): 480-508.
- Shain, C. 2008. Differential object marking in Paraguayan Guaraní. A Senior Honor's Thesis, The Ohio State University.