Grammatical hybridization and social conditions
Leipzig 16-18 October 2014

Hybridization, social conditions and word order

Martin Haspelmath &
Susanne Maria Michaelis

Max Planck Institute for Evolutionary Anthropology, Leipzig

1. Introduction

"word order is among the linguistic phenomena that are most likely to be affected by language contact" (Heine 2008: 34)
(also Thomason & Kaufman 1988: 88; Dryer 1992: 83; Winford 2003)

But what are the social conditions for word order hybridization?

– do speakers adopt word orders from prestige languages just as often they enhance their language with loanwords?
– is it found as a substrate effect, when a large group of speakers learn another language imperfectly?
– is it due to "intense contact" (Thomason & Kaufman 1988)? What is "intense contact"?

Our claims: – word order borrowing is found under assimilation conditions (**"metatypy"**)

– major word order patterns are (apparently) not borrowed under imposition conditions (**"substrate effects"**)

– large-scale areal patterns cannot be easily understood from modern case studies

2. Cognitive conditions for transfer

adoption

– speakers adopt elements (forms, patterns) from a less familiar language
– identifiable elements are used more or less consciously
for semantic or social reasons (enrichment)

imposition

– speakers impose elements (mostly patterns) from their native language on a second language that they learned imperfectly as adults
– difficult-to-suppress elements from the native language surface unconsciously in the learner version of the recipient language (substrate effect, especially phonology)

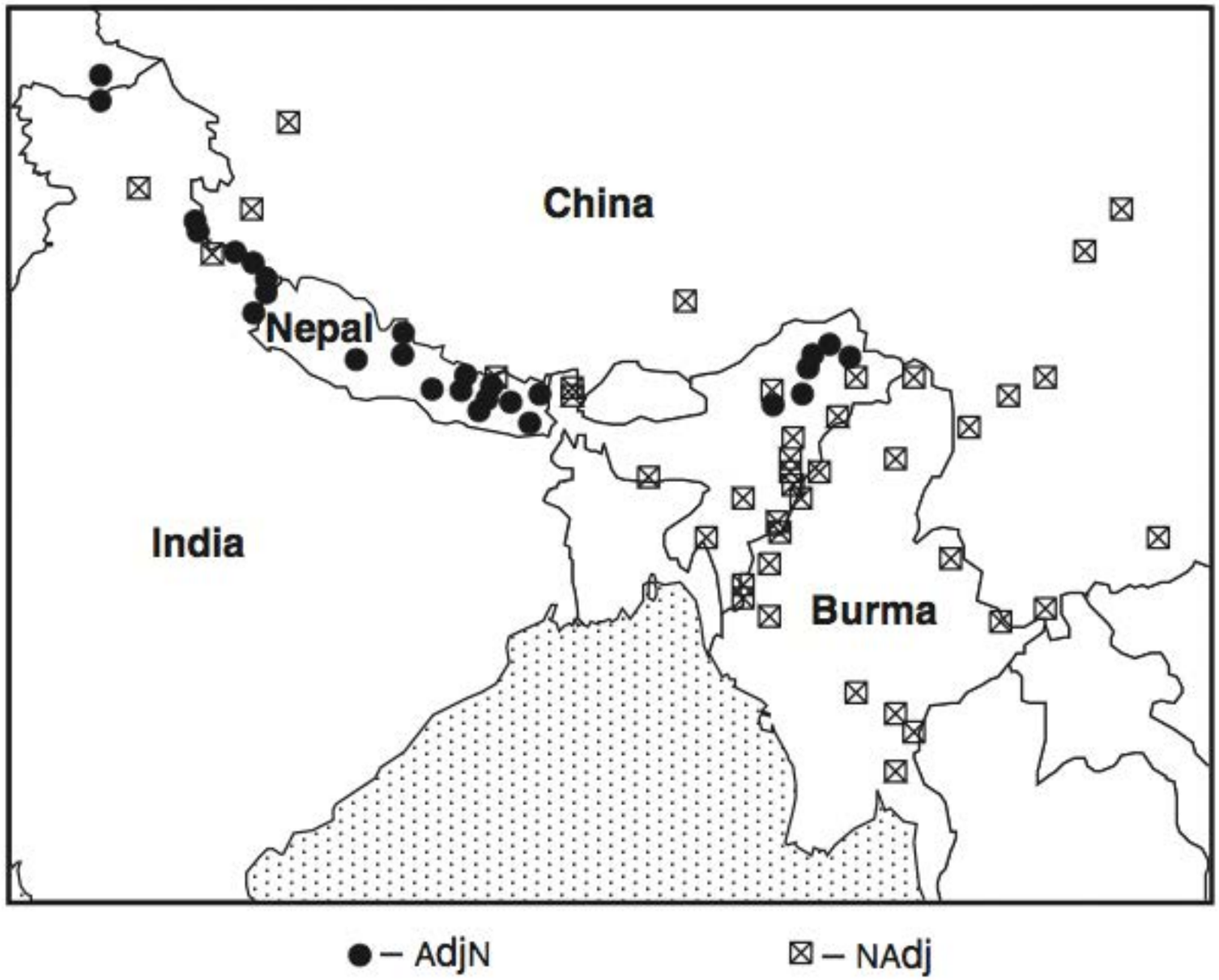
assimilation

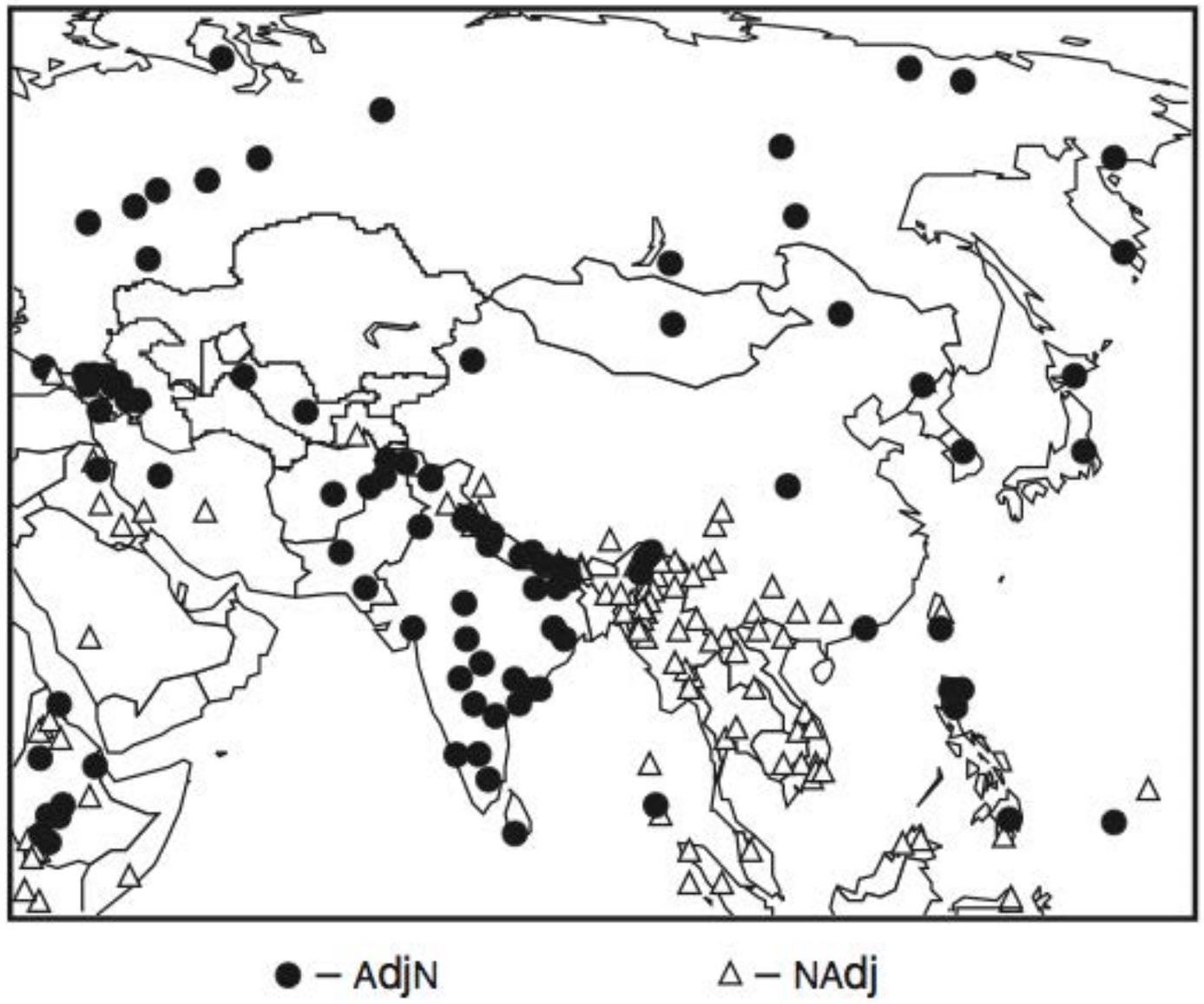
– speakers assimilate patterns in one of their languages to another language that they know just as well (or better)
– difficult-to-suppress elements from a stronger language are used unconsciously in a weaker language (**pattern copying and metatypy**, serving equi-translatability)

3. The macro-level picture: Word order in Trans-Himalayan ("Sino-Tibetan") languages

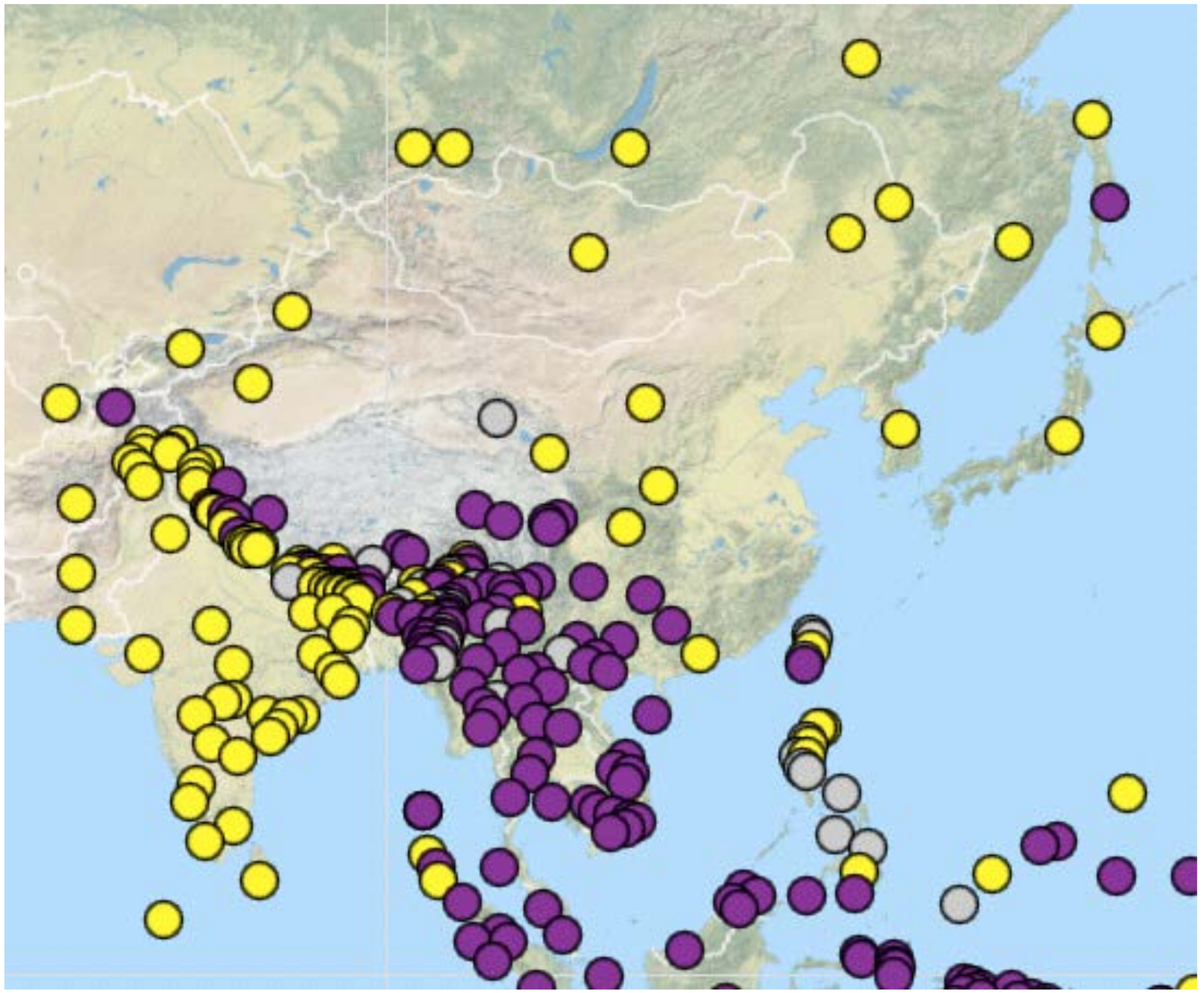
Dryer (2003):

– adjective-noun order is split between western and eastern languages, but in accordance with surrounding languages

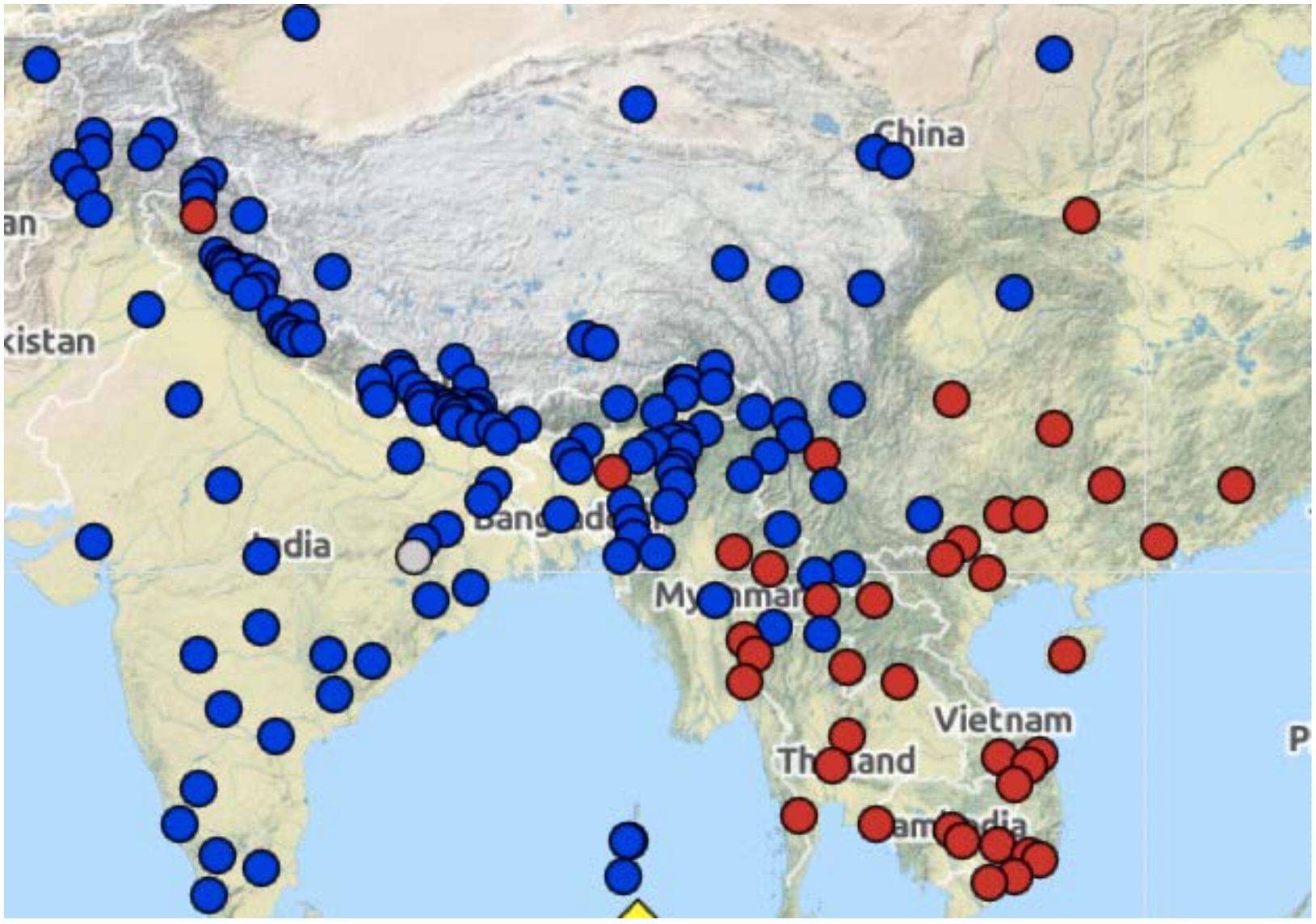




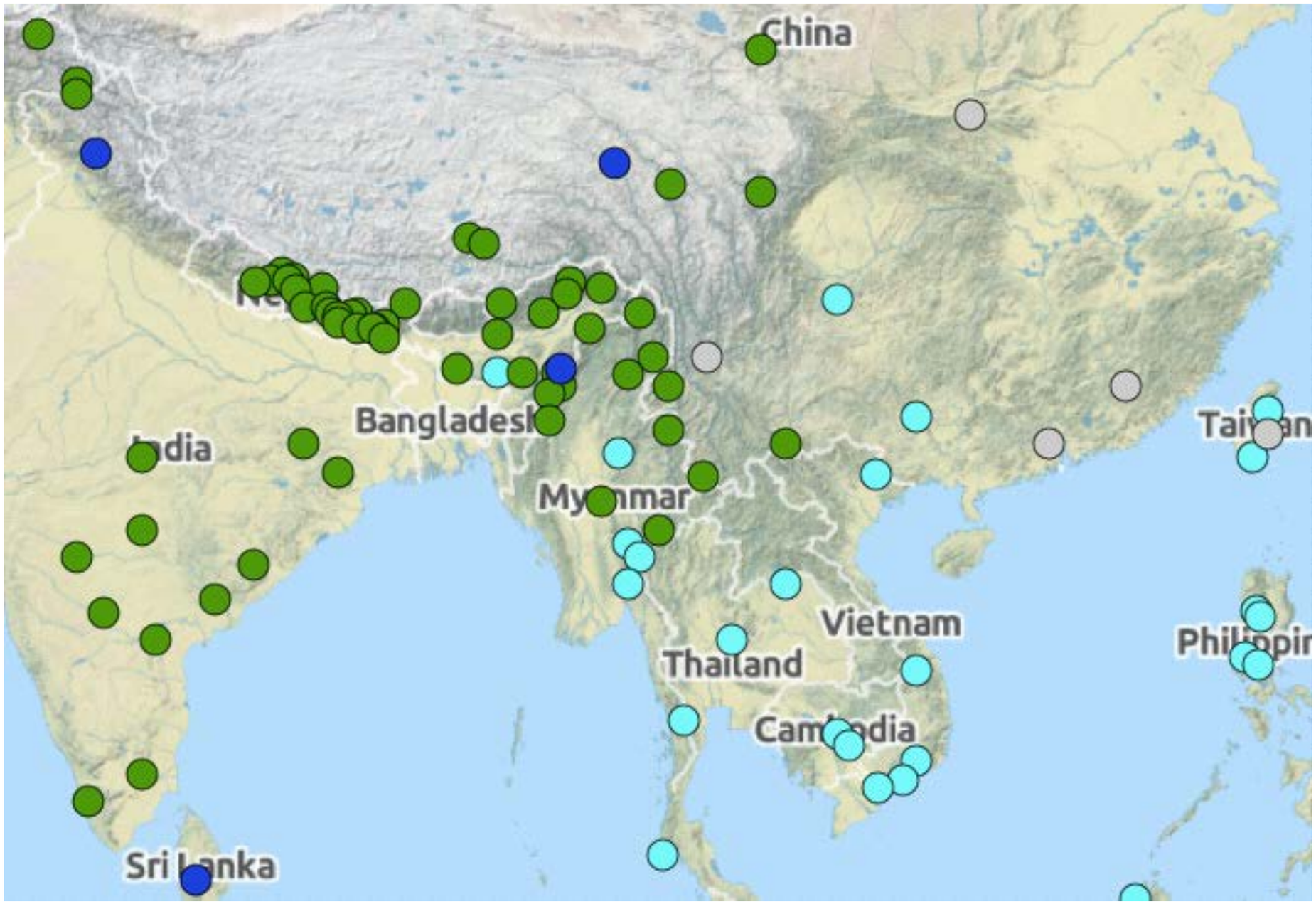
adjective-noun order in Asia (Dryer 2003)



adjective-noun order in WALS (Dryer 2005a/2013a)



order of S, V, O in WALS (Dryer 2005b/2013b)



order of SVO and rel. clause in WALS (Dryer 2005bc)

3. The macro-level picture: Word order in Trans-Himalayan ("Sino-Tibetan") languages

Dryer (2003):

– adjective-noun order is split between western and eastern languages, but in accordance with surrounding languages

– OV/VO order is split in a similar way

– the unusual combination of RelN with VO in Sinitic can be explained if RelN (etc) is due to Tungusic/Mongolic influence, and if VO is due to Tai-Kadai and Hmong-Mien influence

3. The macro-level picture: Word order in Trans-Himalayan ("Sino-Tibetan") languages

Dryer (2003):

– adjective-noun order is split between western and eastern languages, but in accordance with surrounding languages

– OV/VO order is split in a similar way

– the unusual combination of RelN with VO in Sinitic can be explained if RelN (etc) is due to Tungusic/Mongolic influence, and if VO is due to Tai-Kadai and Hmong-Mien influence

This is a particularly striking case, but there are quite a few other similar cases, e.g.

– western Uralic SVO order vs. general Uralic SOV
– Munda SOV order vs. general Austroasiatic SVO
– Bel OV order vs. general Oceanic VO order (Ross 2007)
– Romansch verb-second vs. general Romance SVO

In fact, according to Dryer (2011: 371),
"the majority of word order changes occur due to contact with unrelated languages".

4. The micro-level picture

examination of some situations from two recent edited volumes (Aikhenvald & Dixon 2006, Matras & Sakel 2007):

RECIPIENT < DONOR	WORD-ORDER CHANGE	SOURCE
Likpe < Ewe	none	Ameka 2006
Basque < French/Spanish	postnominal RCs	Jendraschek 2006
Tetun Dili < Portuguese	none	Hajek 2006
Pennsylvania German < English	none	Burridge 2006
Cantonese < Tai-Kadai (?)	property-standard	Matthews 2006
Hup < Tucanoan	SOV (< earlier SVO)	Epps 2006
Amuesha < Quechua	none	Adelaar 2006
Mawayana < Cariban	none	Carlin 2006
Tasawaq (Songhay) < Tamajeq (Tuareg)	none	Kossmann 2007
Khuzistani Arabic < Persian	N-Adj-Poss, V-Aux	Matras & Shabibi 2007
Domari < Arabic	SOV, NAdj, NPoss	Matras 2007
Macedonian Turkish < Macedonian	NPoss	Matras & Tufan 2007
Kildin Saami < Russian	none	Rießler 2007
Manange < Nepali	none	Hildebrandt 2007
Vietnamese < Chinese	none	Alves 2007
Yaqui < Spanish	none	Estrada & Guerrero 2007
Imbabura Quechua < Spanish	SVO, N-RelC	Estrada & Guerrero 2007

Overall, not very much word-order change, but some word orders that a minor language adopts from a major (or national) language.

Three examples:

(1) Imbabura Quechua (Gómez-Rendón 2007: 512)
Kallar-naku-nchik shuk mushuk semana-ta.
begin-RECP-1PL one new week-ACC
'We all start a new week.'

(2) a. Standard Turkish
damad-ın eşya-lar-ı
groom-GEN clothes-PL-3SG.POSS
'the groom's clothes'

b. Macedonian Turkish (Gostivar)
ruba-lar-i damad-ın
clothes-PL-3SG.POSS groom-GEN

c. Macedonian
ališt-ta na zet-ot
clothes-DEF of groom-DEF (Matras & Tufan 2007: 219-220)

(3) a. Cantonese
gou1 gwo3 keoi5
tall pass 3SG
'taller than him' (Matthews 2006: 229)

b. Thai
suung kwaa khaw
tall pass 3SG
'taller than him'

c. Mandarin
bǐ tā gāo
than 3SG tall
'taller than him'

At most the Cantonese example might be explained as a substrate effect, but this is very uncertain, because the contact influence happened long ago.

More recent word-order borrowings all seem to involve a minor language influenced by a major language (Basque, Hup, Khuzistani Arabic, Domari, Macedonian Turkish, Imbabura Quechua), apparently via **assimilation/metatypy**.

Can this mechanism explain the macro-level picture?
Maybe, but only if it can be made plausible that it also works when the languages are in a symmetric social relationship.

5. Word order patterns in pidgin and creole languages

(i) subject – verb – object

(ii) numeral and noun

(iii) wh-phrases in content questions

- data from *APiCS* (2013)

• Guiding assumptions:

(a) In pidginization/creolization processes substrate adult speakers interpret structural/phonological patterns of the lexifier language in terms of their dominant/native language(s).

(b) We assume second language acquisition to be relevant in pidginization/creolization.

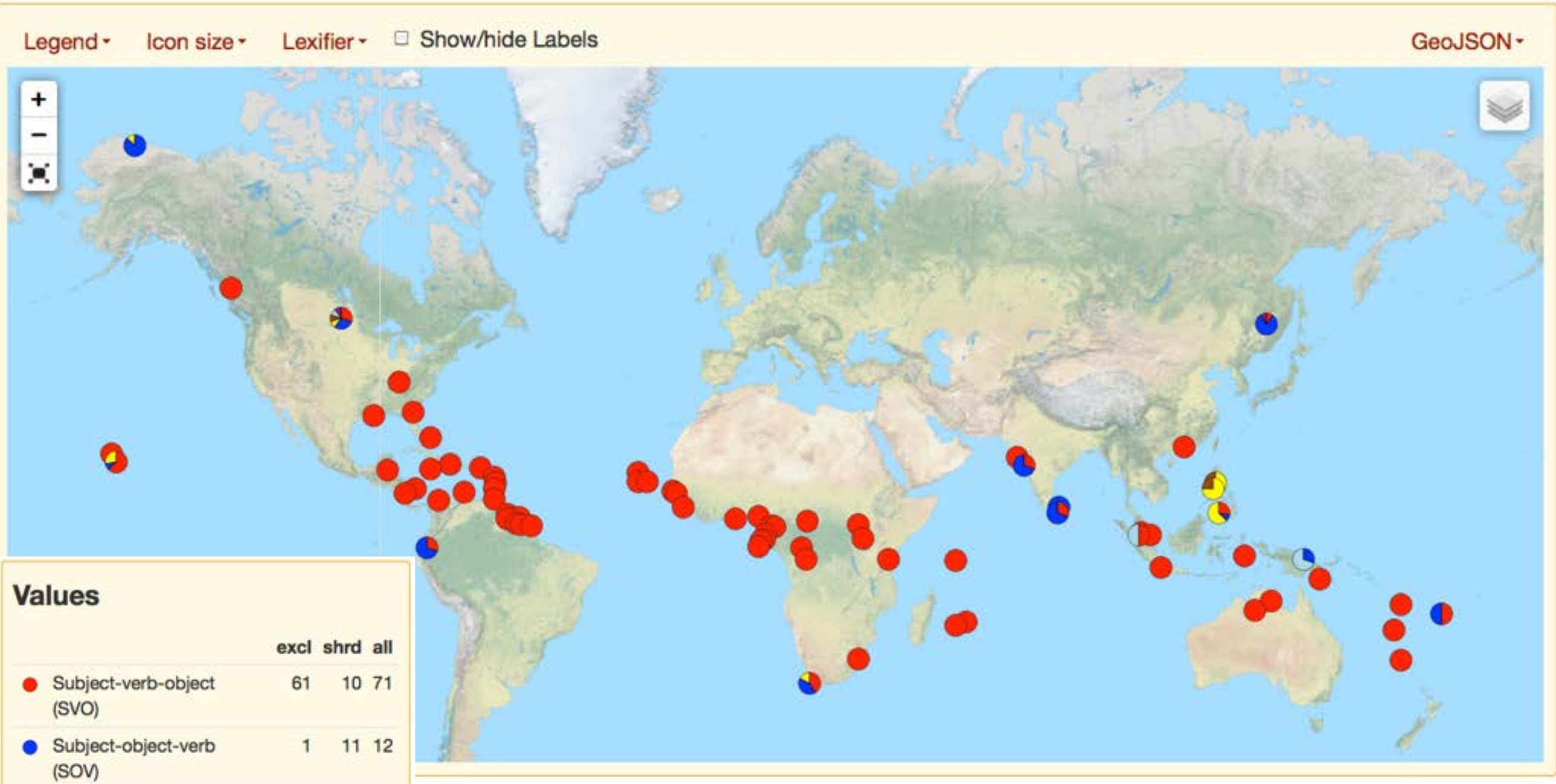
(c) But: data from second language acquisition suggest that **basic word order is usually not imposed** on the target language/lexifier, i.e. SLA speakers acquire the correct word order of the TL/lexifier within a reasonable short time (Siegel 2008: 204-6).

- guiding questions:

  (a) Do pidgins/creoles continue lexifier patterns as we would predict from a second language use scenario in pidginization/creolization? Assuming that minimal exposure to the lexifier would cause rapid restructuring (Siegel 2008: 205).

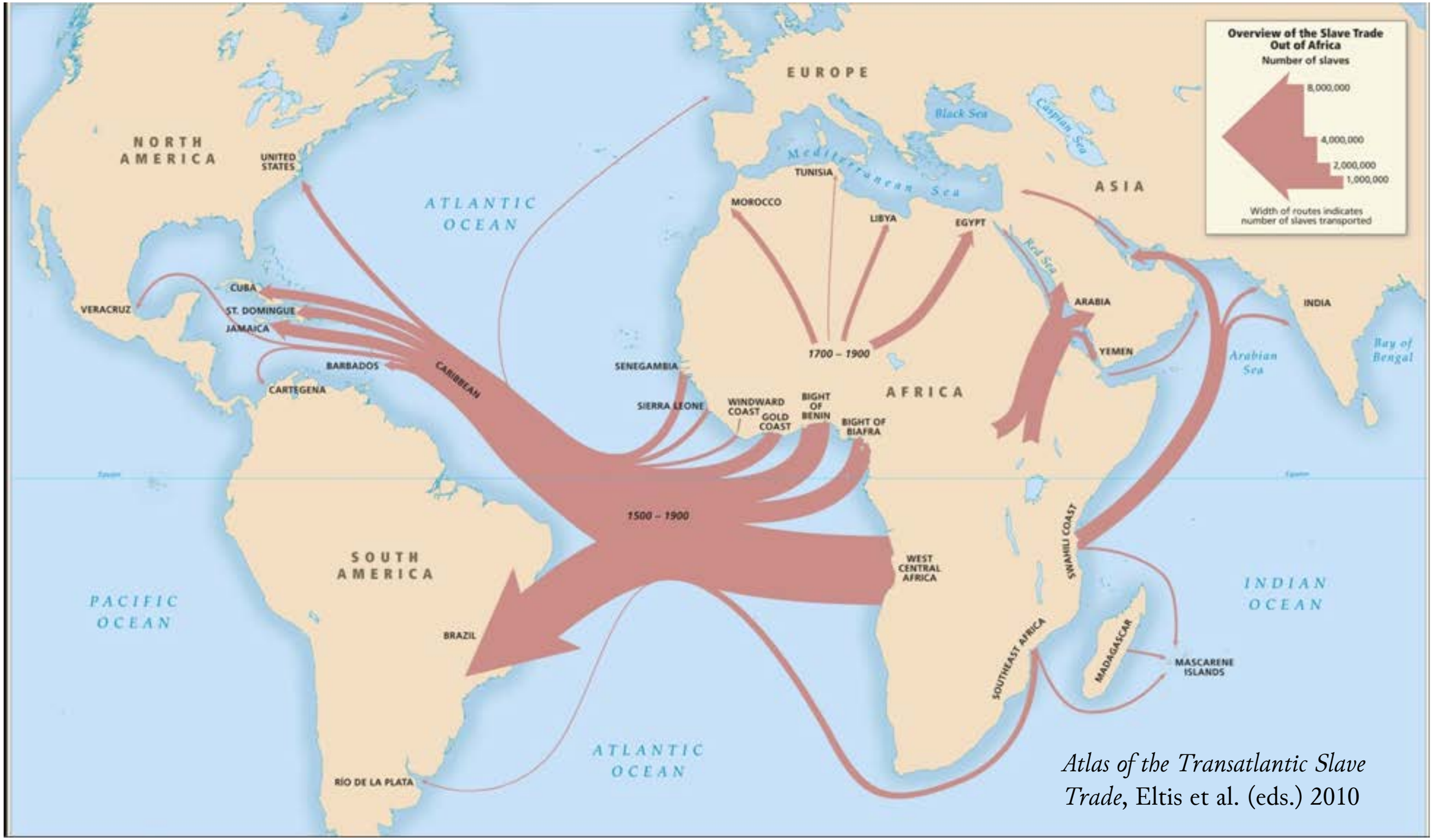
  (b) If we find contradicting data, are there social conditions which make an assimilation/metatypy scenario more likely?

5.1. Order of subject, object and verb (Huber & APiCS Consortium 2013)



- SVO extremely widespread within pidgins/creoles
- South Asian and Philippine Ibero-Romance-based creoles show deviant patterns

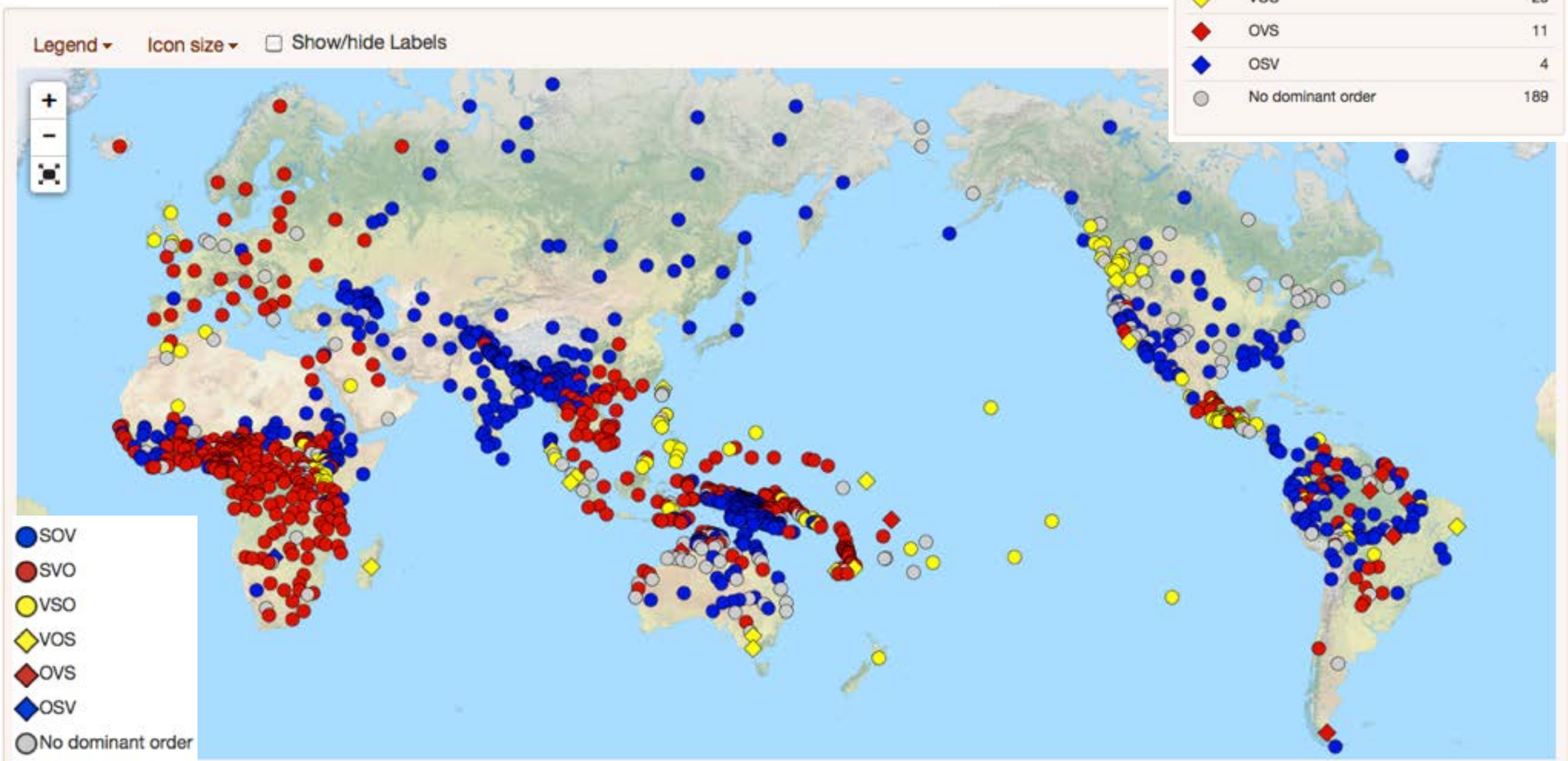
Slave trade out of Africa



Atlas of the Transatlantic Slave Trade, Eltis et al. (eds.) 2010

Feature 81A: Order of Subject, Object and Verb

In chapter 81: Order of Subject, Object and Verb by Matthew S. Dryer cite



- lexifiers overwhelmingly have SVO
- substrates with SVO: West African, Bantu, insular Southeast Asia, parts of Melanesia
- sub-/adstrates with SOV: South Asia, Ijo (West Africa)

- interesting cases:

Berbice Dutch (Dutch-based, Guyana): SVO
  – Dutch (lexifier): no dominant order, (SVO in simple clauses)
  – Ijo (substrate, West Africa): SOV

South Asian Portuguese-based creoles: SVO/SOV
  – Portuguese (lexifier): SVO
  – sub-/adstrates: SOV

- case study: **Korlai** (Portuguese-based, India; Clements 2001, Smith 2012)

Korlai used to have SVO in early stages of creolization, but has been shifting to the SOV pattern of its adstrate Marathi —> metatypy

- Chabacano varieties (Philippines): VSO
  - Portuguese (lexifier): SVO
  - sub-/adstrates (Tagalog): VSO

BUT:

– difficult to reconstruct word order right after creolization,

– no reliable documents,

– more frequent SVO patterns in older written Chabacano texts may only reflect potential Spanish influence on (Sippola, p.c.)

Word order in pidgins/creoles, lexifiers, and sub-/adstrates

	pidgin/ creole	lexifier	substrate/ adstrate	social setting
Haitian Creole	SVO	SVO	SVO	substrates not available anymore
Berbice Dutch	SVO	no dominant/ SVO	SOV	substrates not available anymore
Korlai	SVO--> SOV	SVO	SOV	bilingualism, dominant adstrate
Chabacano	VSO	SVO	VSO	multilingualism, adstrates still available

interim summary

creole data from basic word order feature

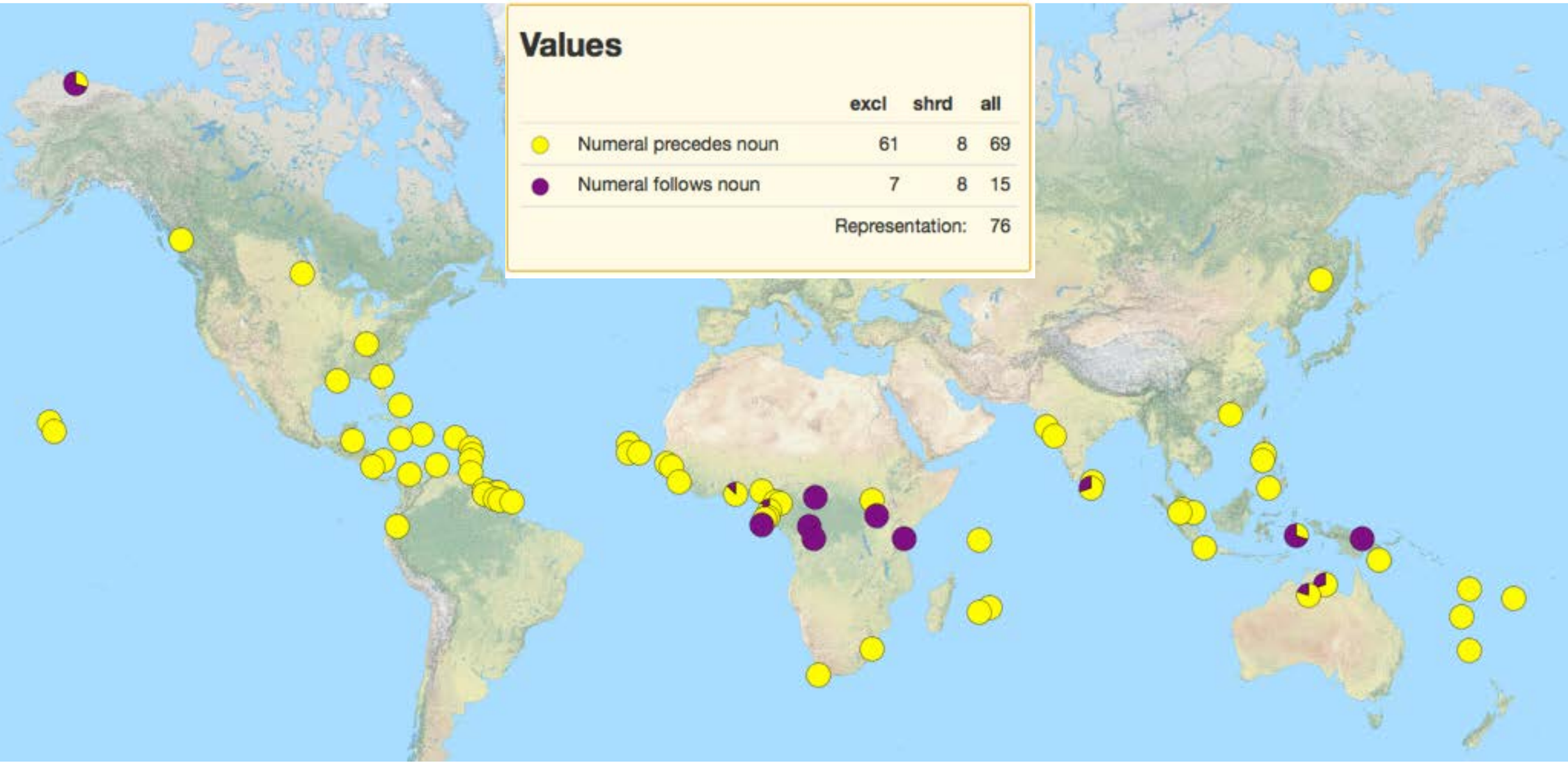
supports idea:

– creoles continue lexifier patterns

– where we find adstrate patterns (e.g. Korlai)

—> metatypic change

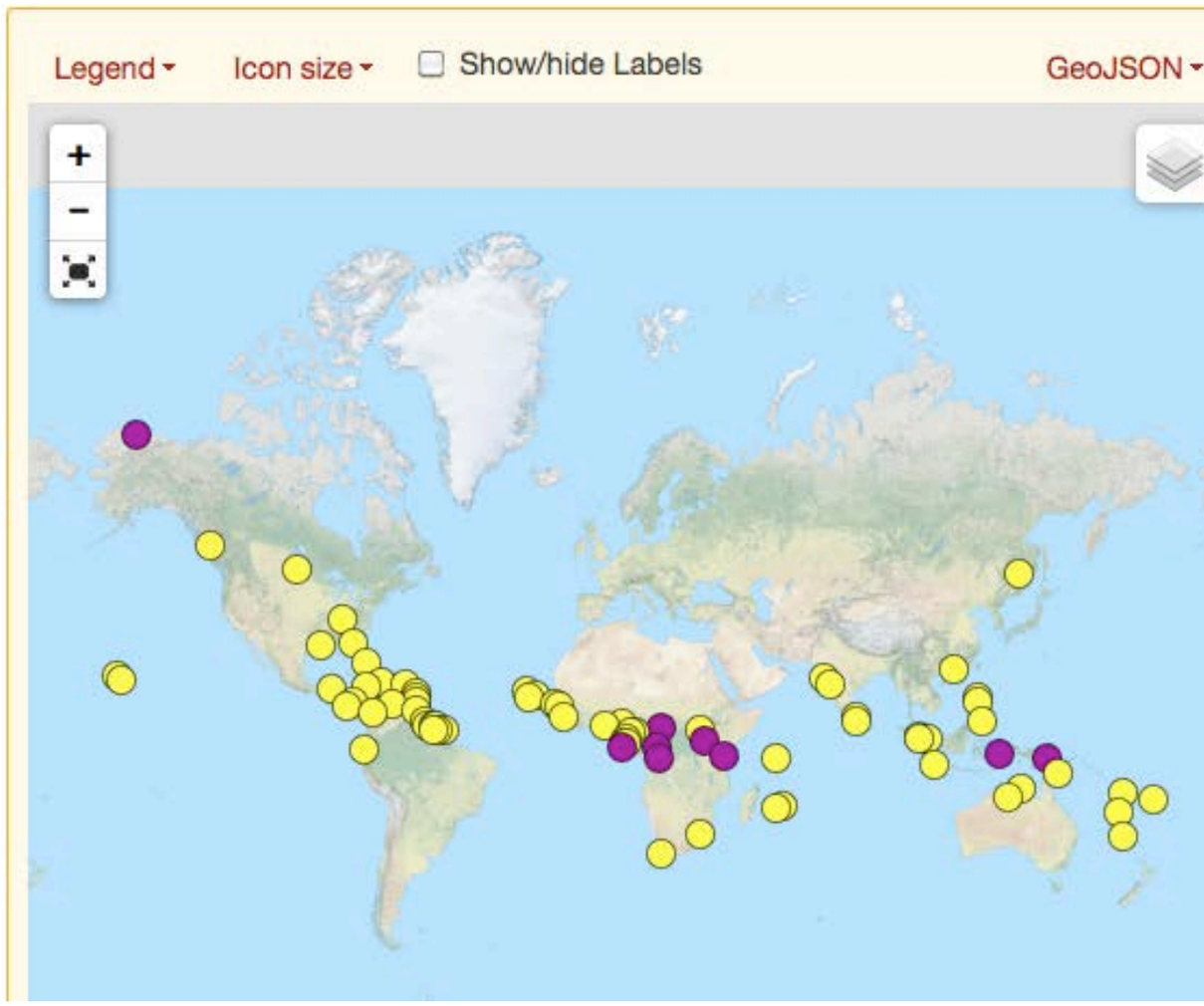
5.2. Order of cardinal numeral and noun: Haspelmath, Michaelis & *APiCS* Consortium 2013a



- Pidgins/creoles show overwhelmingly numeral-noun order.
- Some pidgins/creoles in Africa, and Ambon Malay and Pidgin Yimas-Arafundi have the noun-numeral order.

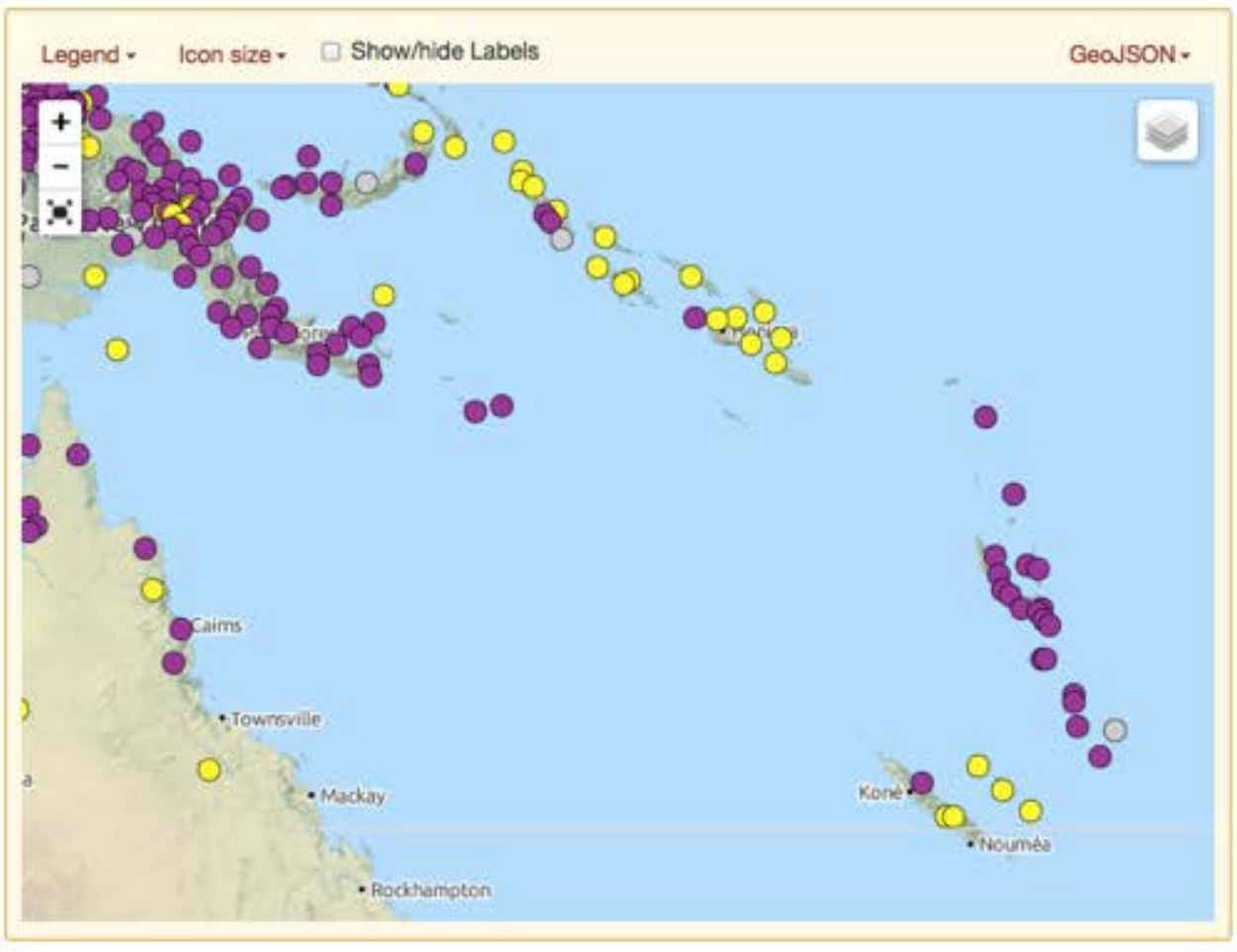
Question: Do the pidgins/creoles continue the lexifier patterns?

Order of cardinal numeral and noun: *WALS* match



- European lexifiers: preposed numeral
- African substrates: postposed numeral (except for Wolof and Ijo)
- South Asian adstrates: preposed numeral

WALS – *APiCS* comparison: Melanesia



- Melanesia has both orders:
  - preposed numeral
  - postposed numeral

Melanesian pidgins/creoles only show

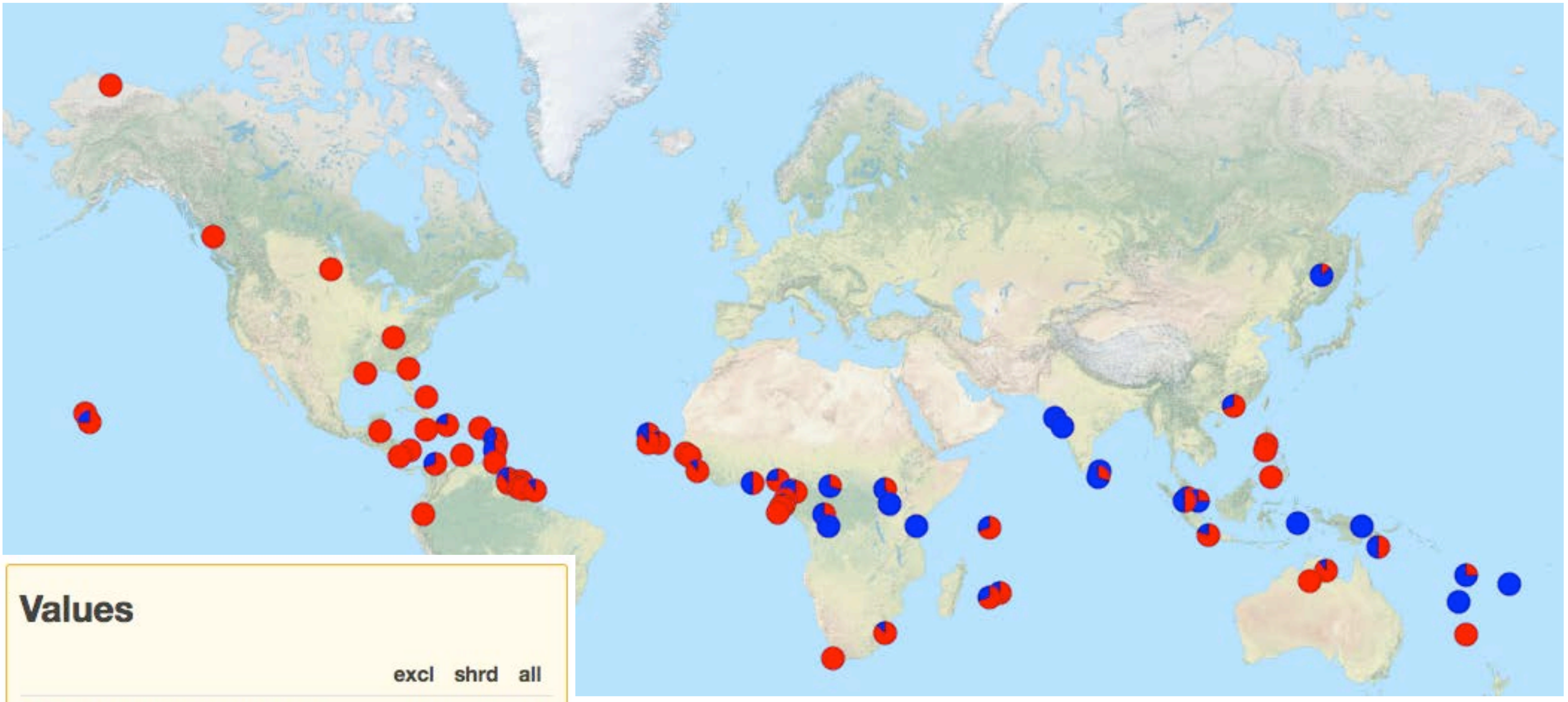
- preposed numeral

Preposed/postposed numerals in pidgins/creoles, lexifiers, and sub-/adstrates

	pidgin/ creole	lexifier	substrate/ adstrate	social setting
Haitian Creole	preposed	preposed	postposed	substrates not available anymore
Saramaccan	preposed	preposed	postposed	substrates not available anymore
Fa d'Ambô	postposed	preposed	postposed	substrates not available anymore
Principense	preposed (except for 'one')[a]	preposed	postposed	substrates not available anymore
Tok Pisin	preposed	preposed	preposed	multilingualism, adstrates still available
Bislama	preposed	preposed	postposed	multilingualism, adstrates still available

[a] In older Principense all numerals could be postposed.

5.3. Position of interrogative phrase: Haspelmath & *APiCS* Consortium 2013b



Values

	excl	shrd	all
Interrogative phrase initial	31	35	66
Interrogative phrase not initial	10	35	45
Representation:			76

Question: Do the pidgins/creoles continue the lexifier patterns?

- many pidgins/creoles show interrogative phrases fronted:

(6) Saramaccan (English-based; Suriname, Aboh et al. 2013)
Andí *a* *bì* *bái?*
what 3SG TNS buy
'What did he buy?'

andí 'what' imposed from the Gbe substrate

(7) Berbice Dutch (Dutch-based; Guyana, Kouwenberg 2013)
wanɛrɛ *so* *ju* *mu-a* *rita* *anga?*
when FOC 2SG go-IPFV Rita LOC
'When are you going to Rita's?'

Some pidgin/creoles show non-fronted interrogative phrases:

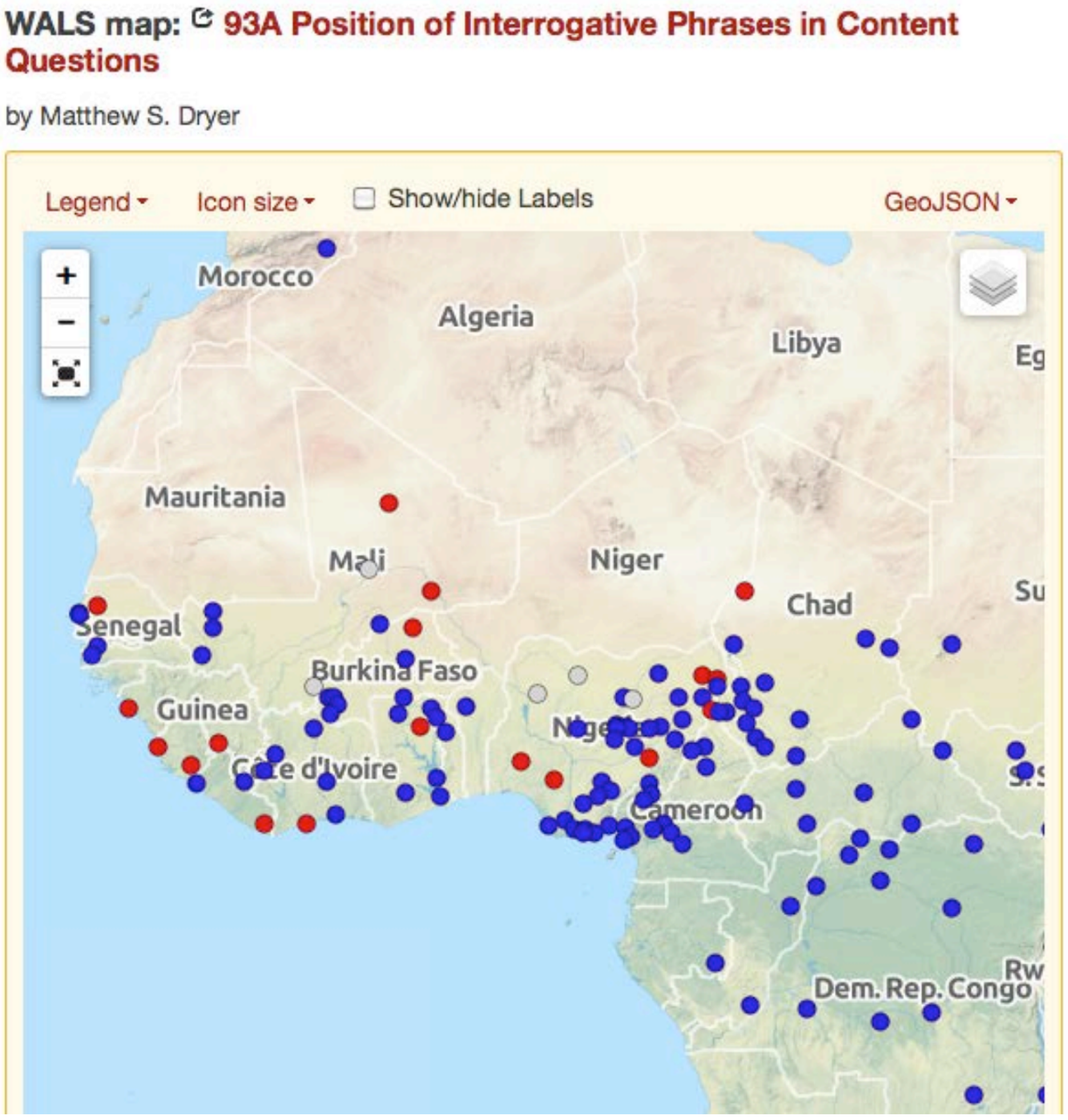
(7) Tayo (French-based; New Caledonia, Ehrhart & Revis 2013)
ta *tape* ***ki?***
you hit whom
'Whom did you hit?'

(8) Korlai (Portuguese-based; India, Clements 2013)
Use ***kɛ̃*** *tɛ?*
2SG.FORMAL who COP.PRS
'Who are you?'

- European **lexifier languages** show fronted wh-phrases; colloquial French has both fronted and in-situ wh-phrases,

Où *est-ce que tu es allée*?

Tu es allée ***où***? 'Where did you go?' (Coveney 1995, 1996)

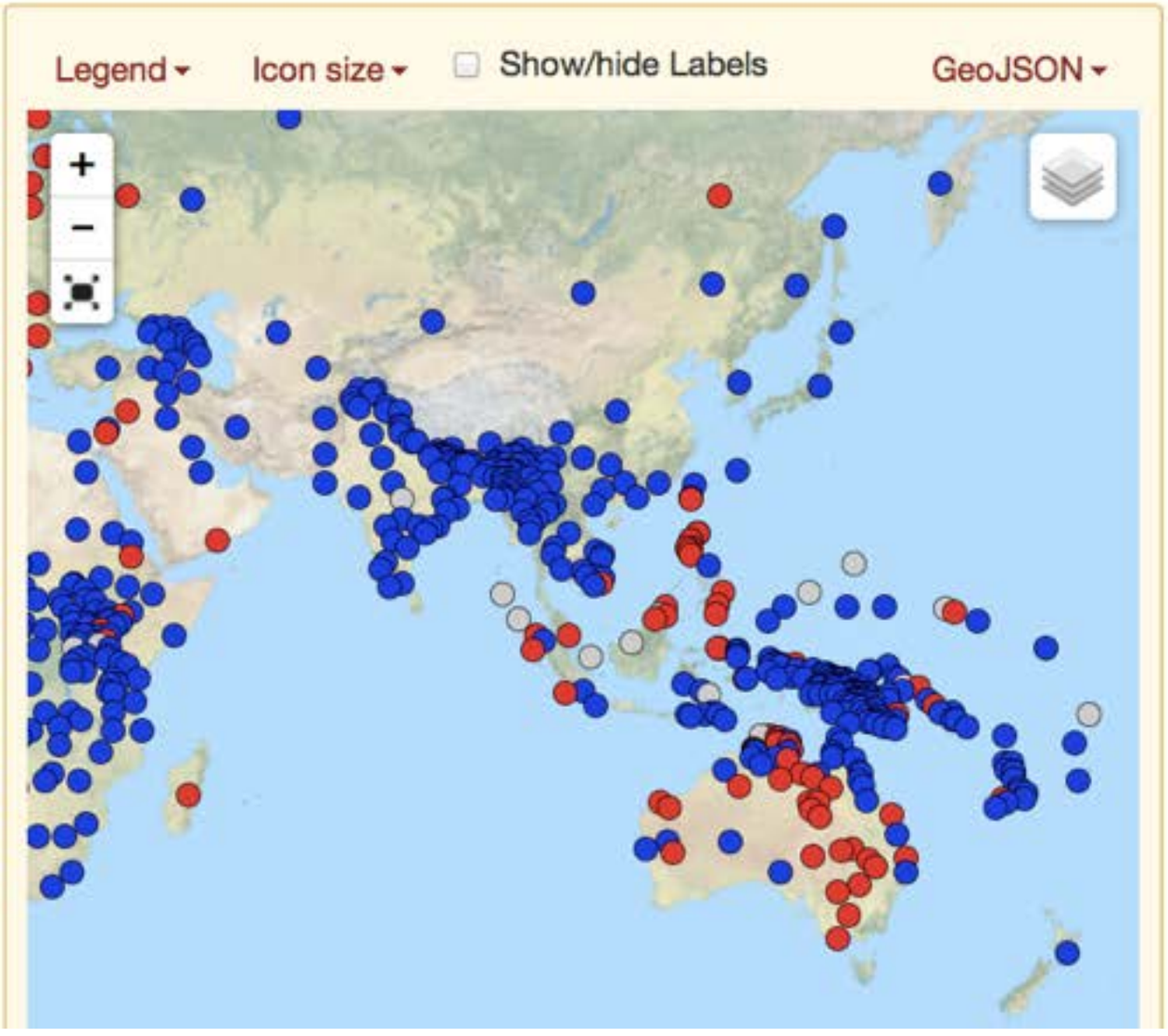


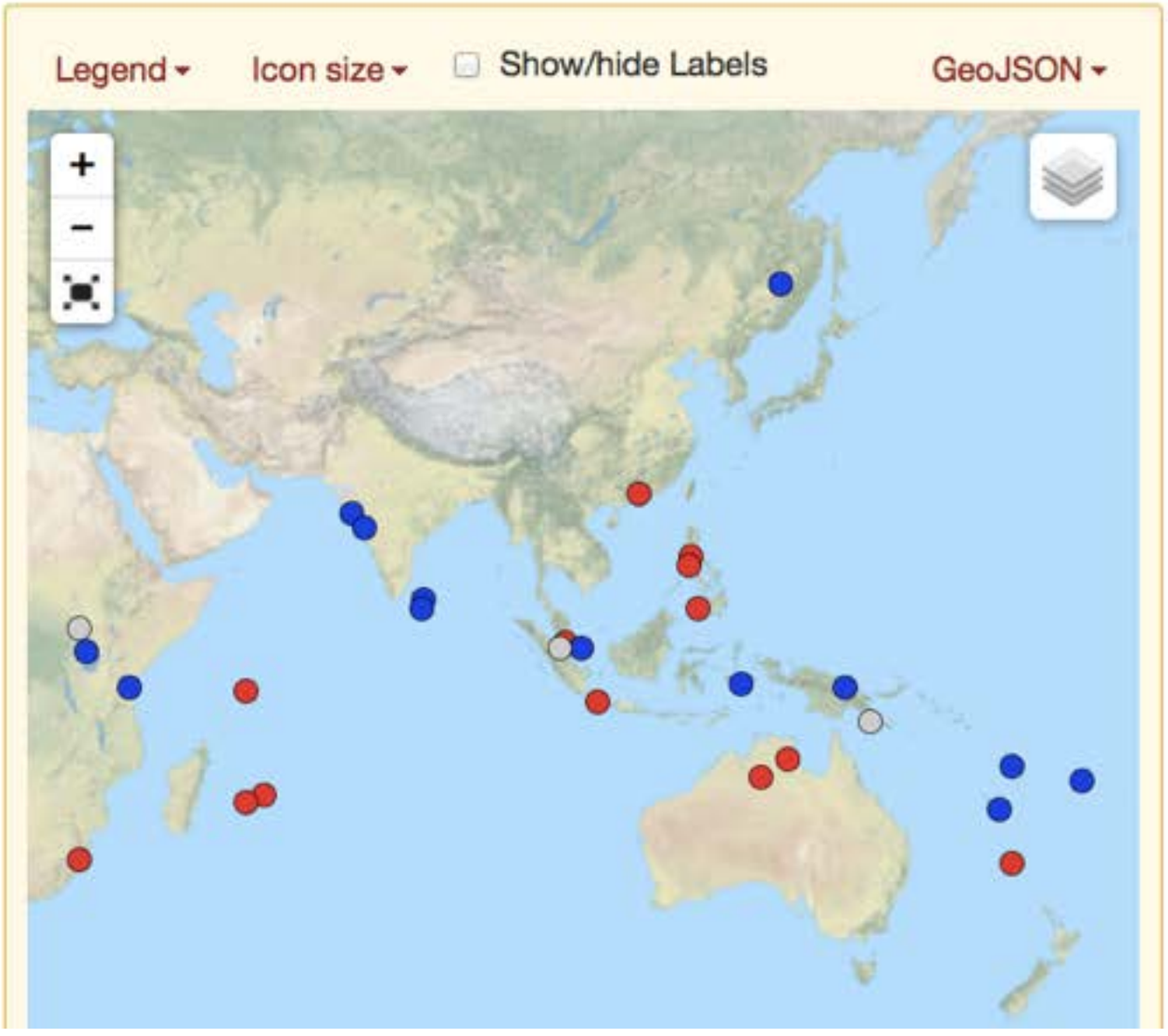
- West African substrates:

Gbe, Yoruba, Wolof: fronted wh-phrases

Ewe, Akan, Ijo: non-fronted wh-phrases

WALS – *APiCS* comparison: South(east) Asia, Oceania





- South India: non-fronted
- Melanesia: non-fronted
- parts of insular Southeast Asia and Australia:

fronted, mixed

Lexifier and substrates coincide

• Saramaccan with mainly Gbe substrate continues the coinciding pattern of its lexifiers (English/Portuguese) and West African substrates, fronted wh-phrases.

• The same is true for the Chabacano varieties: they continue the pattern of the lexifier (Spanish) and of the Philippinic adstrates: fronted wh-phrases.

Lexifier prevails over substrate

• Jamaican and Belizean, for instance, with a major Akan substrate (non-fronted pattern) show the lexifier pattern of English, ---> fronted wh-phrases.

• The same holds for Berbice Dutch: Ijo (its only substrate) has non-fronted wh-phrases, Dutch has fronted, so has Berbice Dutch.

Adstrates prevail over lexifier (metatypy)

• Tayo closely mirrors the patterns of its Oceanic adstrate languages Cèmuhí, Drubéa, Xârâcùù (Corne 1999: 34): non-fronted wh-phrases (lexifier spoken French also allows for non-fronted wh-phrases, but at a much smaller scale, Coveney 1995).

• Bislama mirrors its Oceanic adstrates, e.g. Efate and Paamese which have non-fronted wh-phrases against its lexifier English.

• Tok Pisin, mixed: one of its main sub-/adstrates Tolai also shows mixed behavior, Tigak has non-fronted wh-phrases

• Korlai, Diu Indo-Portuguese, Sri Lanka Portuguese have non-fronted wh-phrases as have their sub-/adstrates, adstrate Marathi (Clements 1996:179)

Position of wh-phrases, f = fronted, non-f = non-fronted

	pidgin/ creole	lexifier	substrate/ adstrate	social setting
Saramaccan	f	f	f	substrates not available anymore
Chabacano	f	f	f	adstrates still available
Jamaican	f	f	non-f	substrates not available anymore
Berbice Dutch	f	f	non-f	substrate not available anymore
Tayo	non-f	f/non-f	non-f	multilingualism, adstrates still available
Bislama	non-f (f)	f	non-f	multilingualism, adstrates still available
Korlai	non-f	f	non-f	bilingualism, dominant adstrate

6. Conclusion

• Even though substrate effects have sometimes been invoked to explain word order change (perhaps most famously for Ethiopic Semitic languages and Akkadian, cf. Leslau 1945, Heine 2008), the creoles and pidgins that we examined do not provide good evidence for this.

• It seems that word order patterns are transferred by assimilation, not by imposition, but in modern contact situations we do not see much evience for this.

• It remains to be seen whether assimilation can explain large-scale areal patterns of word order.

References

Aikhenvald, Alexandra Y. & R.M.W. Dixon (eds.). 2006. *Grammars in contact: A cross-linguistic typology*. Oxford: Oxford University Press.

Clements, J. Clancy. 2001. Word order shift and natural L2 acquisition in a Portuguese creole. In Camps & Wiltshire (eds.), *Romance syntax...* , 73-87.

Clements, J. Clancy. 2001. Word order shift and natural L2 acquisition in a Portuguese creole. In Joaquim Camps & Caroline R. Wiltshire (eds.), *Romance syntax, semantics and L2 acquisition*, 73–87. Amsterdam: Benjamins.

Corne, Chris. 1999. *From French to Creole: The development of new vernaculars in the French colonial world*. London: University of Westminster Press.

Coveney, Aidan. 1995. The use of the QU-final interrogative structure in spoken French. *French Language Studies* 5. 143–171.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–138.

Dryer, Matthew S. 2005a. Order of adjective and noun. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 354–357. Oxford: Oxford University Press.

Dryer, Matthew S. 2005b. Order of subject, object, and verb. In Martin Haspelmath, Matthew S Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 330 – 341. Oxford University Press.

Dryer, Matthew S. 2005c. Order of relative clause and noun. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 366–369. Oxford: Oxford University Press.

Dryer, Matthew S. 2011. The evidence for word order correlations. *Linguistic Typology* 15(2). 335–380.

Enoch O. Aboh, Tonjes Veenstra, Norval S.H. Smith. 2013. Saramaccan structure dataset. In: Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) *Atlas of Pidgin and Creole Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

Gómez-Rendón, Jorge. 2007. Grammatical borrowing in Imbabura Quichua (Ecuador). In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 481–521. Berlin: Mouton de Gruyter.

Haspelmath, Martin & The APiCS Consortium. 2013. Position of interrogative phrases in content questions. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures*, 44–47. Oxford: Oxford University Press.

Haspelmath, Martin & Susanne Maria Michaelis & The APiCS Consortium &. 2013. Order of cardinal numeral and noun. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures*, 22–25. Oxford: Oxford University Press.

Heine, Bernd. 2008. Contact-induced word order change without word order change. In Peter Siemund & Noemi Kintana (eds.), *Language contact and contact languages*, 33ff. Amsterdam: Benjamins.

Huber, Magnus & The APiCS Consortium. 2013. Order of subject, object and verb. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures*, 2–5. Oxford: Oxford University Press.

Lefebvre, Claire. 1998. *Creole genesis and the acquisition of grammar: The case of Haitian creole.* Cambridge: Cambridge University Press.

Leslau, Wolf. 1945. The influence of Cushitic on the Semitic languages of Ethiopia: A problem of substratum. *Word* 1(1): 59–82.

Matras, Yaron & Jeanette Sakel (eds.). 2007. *Grammatical borrowing in cross-linguistic perspective.* Berlin: Mouton de Gruyter.

Matthews, Stephen. 2006. Cantonese grammar in areal perspective. In: Aikhenvald & Dixon (eds.) *Grammars in contact: A cross-linguistic typology*. Oxford: Oxford University Press.

Sabine Ehrhart, Melanie Revis. 2013. Tayo structure dataset. In: Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Siegel, Jeff. 2008. *The emergence of pidgin and creole languages*. Oxford: Oxford University Press.

Silvia Kouwenberg. 2013. Berbice Dutch structure dataset. In: Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) *Atlas of Pidgin and Creole Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

Smith, Ian. 2012. Measuring substrate influence: Word order features. In Cardoso et al. (eds). *Ibero-Asian creoles: comparative perspectives.* Amsterdam: Benjamins, 125-48.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Winford, Donald. 2003. *An introduction to contact linguistics*. Malden, Mass.: Blackwell Pub.