Distributional Typology

Balthasar Bickel
University of Zürich

The quest for absolute universals

«[W]elcher Gewinn wäre es auch, wenn wir einer Sprache auf den Kopf zusagen dürften: Du hast das und das Einzelmerkmal, folglich hast du die und die weiteren Eigenschaften und den und den Gesamtcharakter! - wenn wir, wie es kühne Botaniker wohl versucht haben, aus dem Lindenblatte den Lindenbaum konstruieren könnten. Dürfte man ein ungeborenes Kind taufen, ich würde den Namen **Typologie** wählen.» (Hans Georg Conon von der Gabelentz 1891:481)

"But what an achievement it would be were we able to confront a language and say to it: 'you have such and such a specific property and hence, also such and such further properties and such and such an overall character' – were we able, as daring botanists have indeed tried, to construct the entire lime tree from its leaf. If one were allowed to baptize an unborn child, I would choose the name **typology**."

The quest for absolute universals

- Implications such as

The quest for absolute universals

- Implications such as
  - if [NP V], then [NP P] and [NP N] (Greenberg Universal #2)

The quest for absolute universals

- Implications such as
  - if [NP V], then [NP P] and [NP N] (Greenberg Universal #2)
  - if Null Subjects, then a specific kind of agreement morphology

The quest for absolute universals

- Implications such as
  - if [NP V], then [NP P] and [NP N] (Greenberg Universal #2)
  - if Null Subjects, then a specific kind of agreement morphology
- But also 'hidden' implications from definitions:

The quest for absolute universals

- Implications such as
  - if [NP V], then [NP P] and [NP N] (Greenberg Universal #2)
  - if Null Subjects, then a specific kind of agreement morphology
- But also 'hidden' implications from definitions:
  - if X is a subordinate clause, then WH is banned inside it

The quest for absolute universals

- Implications such as
  - if [NP V], then [NP P] and [NP N] (Greenberg Universal #2)
  - if Null Subjects, then a specific kind of agreement morphology
- But also 'hidden' implications from definitions:
  - if X is a subordinate clause, then WH is banned inside it
  - if X is a syllable, then it combines into feet (cf. the Prosodic Hierarchy)

The quest for absolute universals

- Implications such as
  - if [NP V], then [NP P] and [NP N] (Greenberg Universal #2)
  - if Null Subjects, then a specific kind of agreement morphology
- But also 'hidden' implications from definitions:
  - if X is a subordinate clause, then WH is banned inside it
  - if X is a syllable, then it combines into feet (cf. the Prosodic Hierarchy)
    - Properties as diagnostics ("*since* WH is banned, it's probably subordinate")

The quest for absolute universals

- Implications such as
  - if [NP V], then [NP P] and [NP N] (Greenberg Universal #2)
  - if Null Subjects, then a specific kind of agreement morphology
- But also 'hidden' implications from definitions:
  - if X is a subordinate clause, then WH is banned inside it
  - if X is a syllable, then it combines into feet (cf. the Prosodic Hierarchy)
    - Properties as diagnostics ("*since* WH is banned, it's probably subordinate")
    - Properties as predictions ("if it's subordinate, *expect* a ban on WH!")

The quest for absolute universals

- Implications such as
  - if [NP V], then [NP P] and [NP N] (Greenberg Universal #2)
  - if Null Subjects, then a specific kind of agreement morphology
- But also 'hidden' implications from definitions:
  - if X is a subordinate clause, then WH is banned inside it
  - if X is a syllable, then it combines into feet (cf. the Prosodic Hierarchy)
    - Properties as diagnostics ("*since* WH is banned, it's probably subordinate")
    - Properties as predictions ("if it's subordinate, *expect* a ban on WH!")

But when is a universal really 'absolute'?

The problem of absolute universals

- An absolute universal is true of all languages, i.e. has no exceptions.

The problem of absolute universals

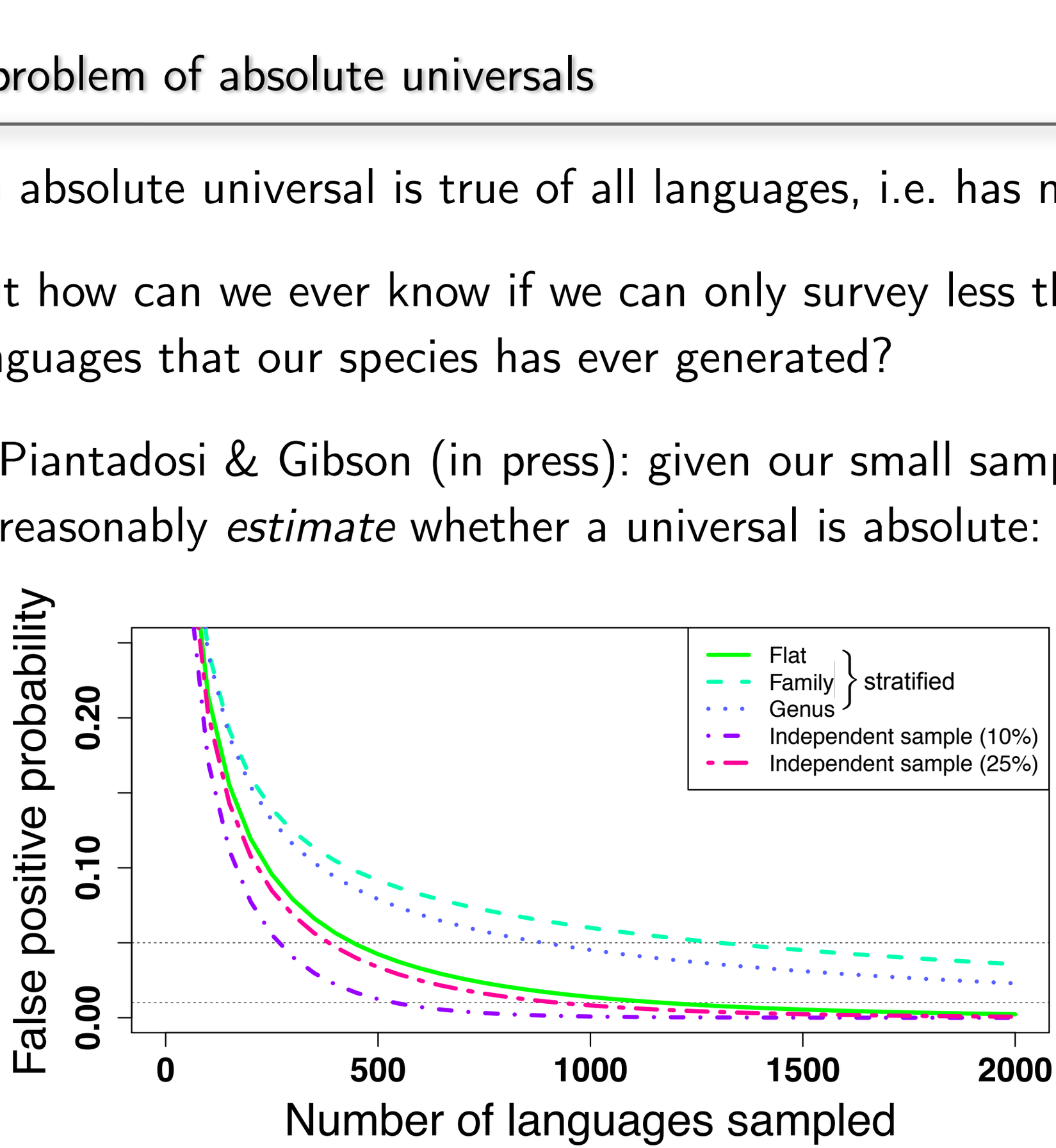
- An absolute universal is true of all languages, i.e. has no exceptions.
- But how can we ever know if we can only survey less than 1% of the languages that our species has ever generated?

The problem of absolute universals

- An absolute universal is true of all languages, i.e. has no exceptions.
- But how can we ever know if we can only survey less than 1% of the languages that our species has ever generated?
  - Piantadosi & Gibson (in press): given our small samples, we can't even reasonably *estimate* whether a universal is absolute:

The problem of absolute universals

- An absolute universal is true of all languages, i.e. has no exceptions.
- But how can we ever know if we can only survey less than 1% of the languages that our species has ever generated?
  - Piantadosi & Gibson (in press): given our small samples, we can't even reasonably *estimate* whether a universal is absolute:



"Independent" means genealogically unrelated.

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.
- Based on the **Pāṇinian Approach** to language:

fl. 5th/6th century BCE

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.
- Based on the **Pāṇinian Approach** to language:
- Rules; then explain away what doesn't fit by positing extra rules:

fl. 5th/6th century BCE

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.
- Based on the **Pāṇinian Approach** to language:
- Rules; then explain away what doesn't fit by positing extra rules:

  - Diachrony: Skt. ***p**ad-* : Lat. ***p**ed-* 'foot'

fl. 5th/6th century BCE

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.
- Based on the **Pāṇinian Approach** to language:
- Rules; then explain away what doesn't fit by positing extra rules:

  - Diachrony: Skt. ***p**ad-* : Lat. ***p**ed-* 'foot'
    Rule 1 *p → p

fl. 5th/6th century BCE

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.
- Based on the **Pāṇinian Approach** to language:
- Rules; then explain away what doesn't fit by positing extra rules:

  - Diachrony: Skt. ***p**ad-* : Lat. ***p**ed-* 'foot'
    Rule 1 *p → p
    but Skt. ***p**ac-* : Lat. ***c**oqu-* 'cooks'

fl. 5th/6th century BCE

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.

- Based on the **Pāṇinian Approach** to language:

- Rules; then explain away what doesn't fit by positing extra rules:

  - Diachrony: Skt. ***p**ad-* : Lat. ***p**ed-* 'foot'
    Rule 1 *p → p
    but Skt. ***p**ac-* : Lat. ***c**oqu-* 'cooks'
    Rule 2: ***pV**k^{w} → **k^{w}V**k^{w}

fl. 5th/6th century BCE

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.

- Based on the **Pāṇinian Approach** to language:

fl. 5th/6th century BCE

- Rules; then explain away what doesn't fit by positing extra rules:

  - Diachrony: Skt. ***p**ad-* : Lat. ***p**ed-* 'foot'
    Rule 1 *p → p
    but Skt. ***p**ac-* : Lat. ***c**oqu-* 'cooks'
    Rule 2: ***pV**k^w → **k^{w}V**k^w

  - Synchrony: Rule 1: All English sentences have a subject: **saw her*

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.

- Based on the **Pāṇinian Approach** to language:

fl. 5th/6th century BCE

- Rules; then explain away what doesn't fit by positing extra rules:

  - Diachrony: Skt. ***p**ad-* : Lat. ***p**ed-* 'foot'
    Rule 1 *p → p
    but Skt. ***p**ac-* : Lat. ***c**oqu-* 'cooks'
    Rule 2: ***pV**k^{w} → **k^{w}V**k^{w}

  - Synchrony: Rule 1: All English sentences have a subject: **saw her*
    but: *he came and _ saw her*

The problem of absolute universals

- **Classical solution:** don't even start to sample languages! Keep the universals anyway, and get rid of any counterexample that may come up.

- Based on the **Pāṇinian Approach** to language:

fl. 5th/6th century BCE

- Rules; then explain away what doesn't fit by positing extra rules:

  - Diachrony: Skt. ***p**ad-* : Lat. ***p**ed-* 'foot'
    Rule 1 *p → p
    but Skt. ***p**ac-* : Lat. ***c**oqu-* 'cooks'
    Rule 2: ***pV**k^w → **k^{w}V**k^w

  - Synchrony: Rule 1: All English sentences have a subject: **saw her*
    but: *he came and _ saw her*
    Rule 2: Delete a coreferential subject in coordination

The problem of absolute universals

- The Pāṇinian Approach applied to typology:

The problem of absolute universals

- The Pāṇinian Approach applied to typology:
- Example 1: exceptions to Greenberg Universal #2 [PP **P** [NP **N** [NP]]]

Harar Oromo (Kushitic; Owens 1985)

[PP [NP ***maná*** [**NP** *obbolesá xiyyá*]] ***=tt***]

house	brother	my	in
N			**P**

The problem of absolute universals

- The Pāṇinian Approach applied to typology:
- Example 1: exceptions to Greenberg Universal #2 [PP **P** [NP **N** [NP]]]

Harar Oromo (Kushitic; Owens 1985)

[PP [NP ***maná*** [**NP** *obbolesá xiyyá*]] ***=tti***]

house	brother	my	in
N			**P**

- Solution: explain away the counterexamples by restricting the claim:

  *[PP [NP **N** [NP]] **P$_N$**] vs. ✓[PP [NP **N** [NP]] **P$_V$**] and argue that counterexamples don't have nominal but verbal postpositions, or don't have 'real' postpositions anyway (Biberauer et al 2008)

The problem of absolute universals

- Example 2: exceptions to the Prosodic Hierarchy

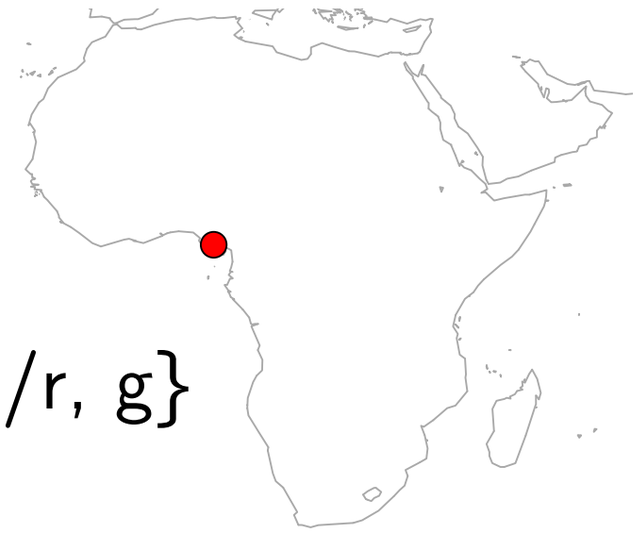
  Gokana (Benue-Congo, Hyman 1983, 2011)

The problem of absolute universals

- Example 2: exceptions to the Prosodic Hierarchy

  Gokana (Benue-Congo, Hyman 1983, 2011)

  - Phonotactics: C_2 a small subset of C_1: C_2 only {b/v, l/r, g}



The problem of absolute universals

- Example 2: exceptions to the Prosodic Hierarchy

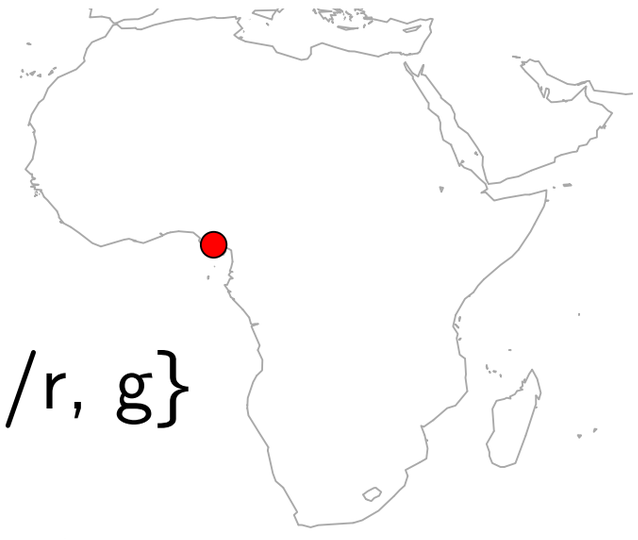
  Gokana (Benue-Congo, Hyman 1983, 2011)

  - Phonotactics: C_2 a small subset of C_1: C_2 only {b/v, l/r, g}
  - But the difference is **never defined by syllables:**

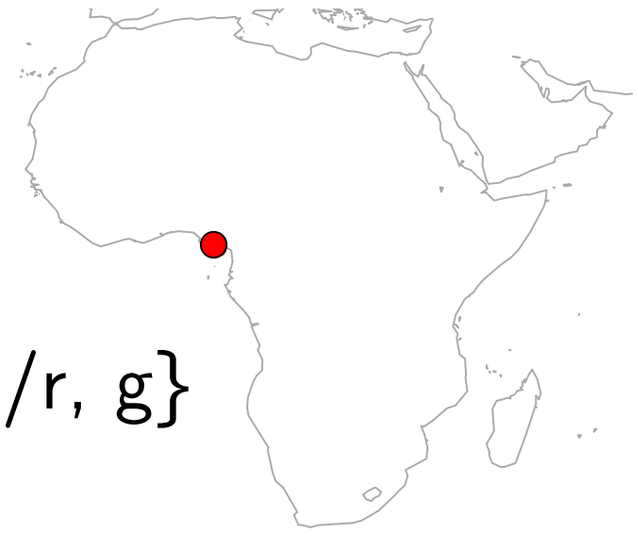
    C_1V, C_1VV, C_1VC_2, C_1VVV, C_1VVVV

    C_1VC_2V, C_1VVC_2V, C_1VC_2VV, C_1VVC_2VV

    unlike in most other languages, e.g. English: $(CCC)V(CCC_{[+coronal]})$



The problem of absolute universals



- Example 2: exceptions to the Prosodic Hierarchy

  Gokana (Benue-Congo, Hyman 1983, 2011)

  - Phonotactics: C_2 a small subset of C_1: C_2 only {b/v, l/r, g}
  - But the difference is **never defined by syllables:**

    C_1V, C_1VV, C_1VC_2, C_1VVV, C_1VVVV

    C_1VC_2V, C_1VVC_2V, C_1VC_2VV, C_1VVC_2VV

    unlike in most other languages, e.g. English: (CCC)V($CCC_{[+coronal]}$)

Analysis without syllables:

(1) ^C_1VV?C_2?V?V?$

(2) disallow *C_1VVC_2

The problem of absolute universals

- Example 2: exceptions to the Prosodic Hierarchy

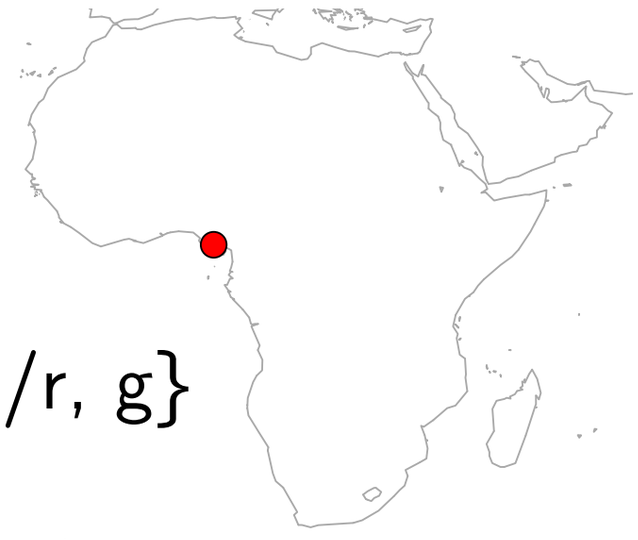
  Gokana (Benue-Congo, Hyman 1983, 2011)

  - Phonotactics: C_2 a small subset of C_1: C_2 only {b/v, l/r, g}
  - But the difference is **never defined by syllables:**

    C_1V, C_1VV, C_1VC_2, C_1VVV, C_1VVVV

    C_1VC_2V, C_1VVC_2V, C_1VC_2VV, C_1VVC_2VV

    unlike in most other languages, e.g. English: (CCC)V($CCC_{[+coronal]}$)



Analysis without syllables:

(1) ^C_1VV?C_2?V?V?$

(2) disallow *C_1VVC_2

Analysis with syllables:

(1) σ: CV(C), (C)V(V)

(2) $C)_\sigma$ only in mono-syllabic words

(3) ϕ: (σ_s σ_w), C_1 in σ_s, C_2 in σ_w

The problem of absolute universals

The problem of absolute universals

- and so on:

The problem of absolute universals

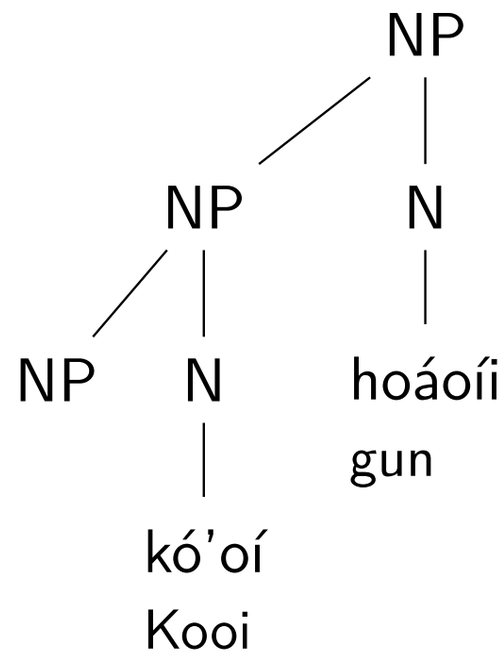
- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

    vs. recursion is not universal

The problem of absolute universals

- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

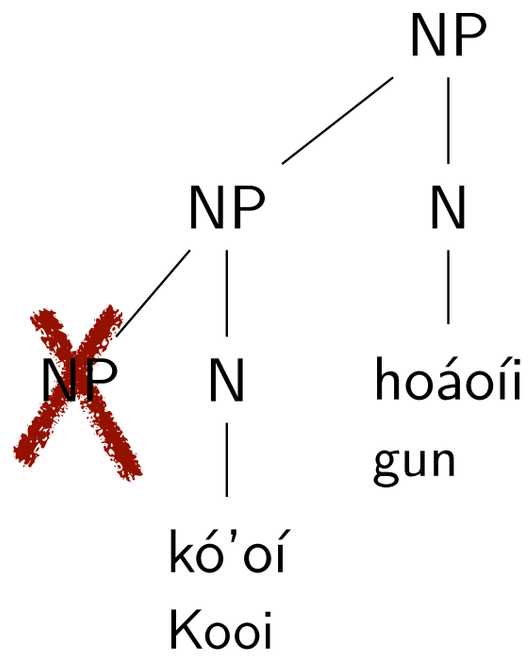
    vs. recursion is not universal



The problem of absolute universals

- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

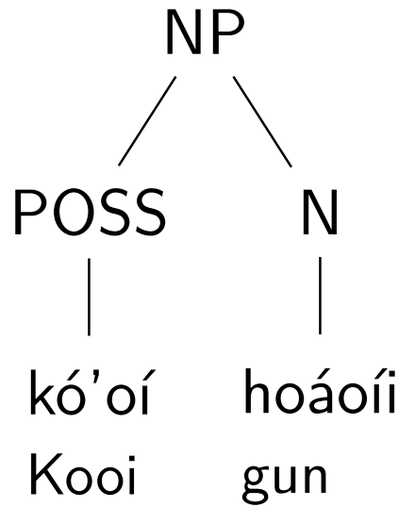
    vs. recursion is not universal



The problem of absolute universals

- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

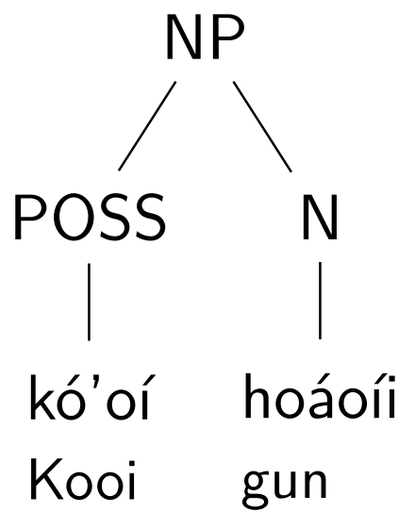
    vs. recursion is not universal



The problem of absolute universals

- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

    vs. recursion is not universal



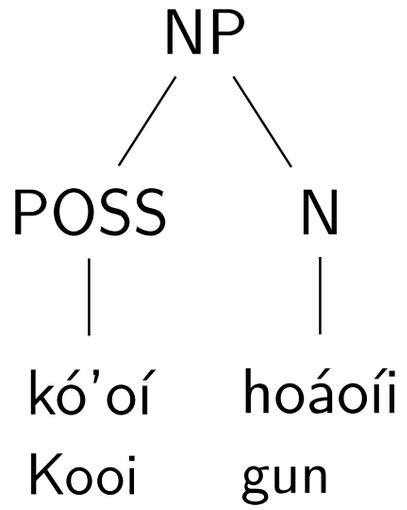
  - words are universal, but interruptable and without a phonological effect in Vietnamese

    vs. the word is not universal

The problem of absolute universals

- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

    vs. recursion is not universal



  - words are universal, but interruptable and without a phonological effect in Vietnamese

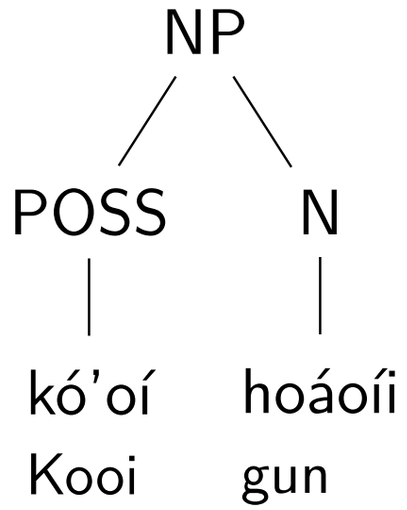
    vs. the word is not universal

(1) *reddish*, ***not*** *reddish*, **red-**not**-ish*

The problem of absolute universals

- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

    vs. recursion is not universal



  - words are universal, but interruptable and without a phonological effect in Vietnamese

    vs. the word is not universal

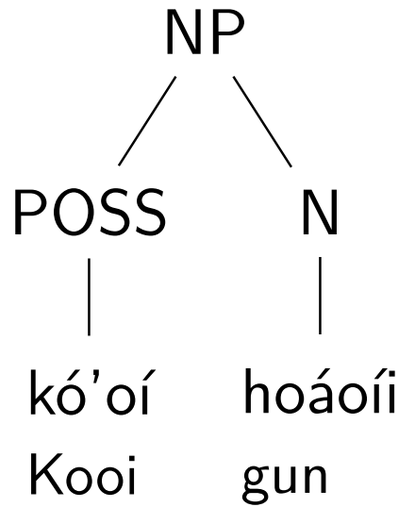
(1) *reddish*, ***not*** *reddish*, **red-**not**-ish*

(2) *đo-đỏ*, ***không*** *đo-đỏ*, *đo-**không**-đỏ*

The problem of absolute universals

- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

    vs. recursion is not universal



  - words are universal, but interruptable and without a phonological effect in Vietnamese

    vs. the word is not universal

(1) *reddish*, ***not*** *reddish*, **red-**not**-ish*

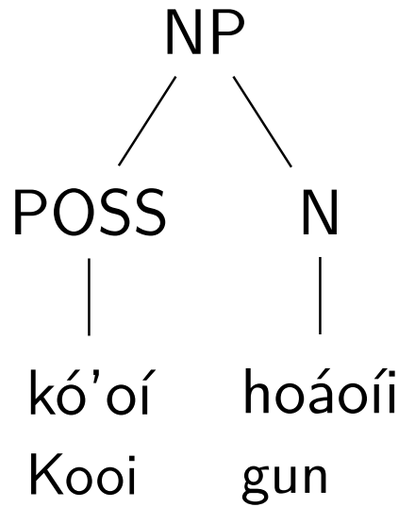
(2) *đo-đỏ*, ***không*** *đo-đỏ*, *đo-**không**-đỏ*

(3) *hoa hồng* 'rose; pink flower'

The problem of absolute universals

- and so on:
  - recursive phrase structure is universal, but limited to 1 level in Pirahã

    vs. recursion is not universal



  - words are universal, but interruptable and without a phonological effect in Vietnamese

    vs. the word is not universal

(1) *reddish*, ***not*** *reddish*, **red-**not**-ish*

(2) *đo-đỏ*, ***không*** *đo-đỏ*, *đo-**không**-đỏ*

(3) *hoa hồng* 'rose; pink flower'

How can we know what's right?

The classical answer (from a typological perspective)

The classical answer (from a typological perspective)

- Criteria of adequacy allow one to choose between analyses (Chomsky 1964) and thereby between universals: an adequate analysis

The classical answer (from a typological perspective)

- Criteria of adequacy allow one to choose between analyses (Chomsky 1964) and thereby between universals: an adequate analysis

  (1) captures all patterns and generalizations ('descriptive adequacy')

The classical answer (from a typological perspective)

- Criteria of adequacy allow one to choose between analyses (Chomsky 1964) and thereby between universals: an adequate analysis

  (1) captures all patterns and generalizations ('descriptive adequacy')

  (2) is stated in a universal theory (Chomsky 1955), i.e. in cross-linguistically viable terms

The classical answer (from a typological perspective)

- Criteria of adequacy allow one to choose between analyses (Chomsky 1964) and thereby between universals: an adequate analysis

  (1) captures all patterns and generalizations ('descriptive adequacy')

  (2) is stated in a universal theory (Chomsky 1955), i.e. in cross-linguistically viable terms

$\rightarrow$ no help here!

The classical answer (from a typological perspective)

The classical answer (from a typological perspective)

- More criteria: an adequate analysis

The classical answer (from a typological perspective)

- More criteria: an adequate analysis

  (1) yields a learnable grammar ('explanatory adequacy', Chomsky 1964)

The classical answer (from a typological perspective)

- More criteria: an adequate analysis

  (1) yields a learnable grammar ('explanatory adequacy', Chomsky 1964)

  (2) is psychologically realistic (Bresnan & Kaplan 1982)

The classical answer (from a typological perspective)

- More criteria: an adequate analysis

  (1) yields a learnable grammar ('explanatory adequacy', Chomsky 1964)

  (2) is psychologically realistic (Bresnan & Kaplan 1982)

  (3) is biologically and evolutionarily realistic (Chomsky 2004)

The classical answer (from a typological perspective)

- More criteria: an adequate analysis

  (1) yields a learnable grammar ('explanatory adequacy', Chomsky 1964)

  (2) is psychologically realistic (Bresnan & Kaplan 1982)

  (3) is biologically and evolutionarily realistic (Chomsky 2004)

$\longrightarrow$ mostly of little help since the evidence is rarely unambigous

The classical answer (from a typological perspective)

- More criteria: an adequate analysis

  (1) yields a learnable grammar ('explanatory adequacy', Chomsky 1964)

  (2) is psychologically realistic (Bresnan & Kaplan 1982)

  (3) is biologically and evolutionarily realistic (Chomsky 2004)

$\longrightarrow$ mostly of little help since the evidence is rarely unambigous

- Standard everyday solution: let's "just" assume a universal anyway (e.g. a ban on *[$_{PP}$ [$_{NP}$ **N** [$_{NP}$]] **P**] or a requirement of syllables in all languages), as "Working Hypothesis", and then see how convincingly we can explain away the counterexamples

The classical answer (from a typological perspective)

- More criteria: an adequate analysis

  (1) yields a learnable grammar ('explanatory adequacy', Chomsky 1964)

  (2) is psychologically realistic (Bresnan & Kaplan 1982)

  (3) is biologically and evolutionarily realistic (Chomsky 2004)

$\longrightarrow$ mostly of little help since the evidence is rarely unambigous

- Standard everyday solution: let's "just" assume a universal anyway (e.g. a ban on *[$_{PP}$ [$_{NP}$ **N** [$_{NP}$]] **P**] or a requirement of syllables in all languages), as "Working Hypothesis", and then see how convincingly we can explain away the counterexamples

- Back to square 1!

The alternative

The alternative

- Forget absolute universals and stop asking "What's possible?", instead ask "**What's Where Why?**"

The alternative

- Forget absolute universals and stop asking "What's possible?", instead ask "**What's Where Why?**"

- But, how exactly? My goals here:

The alternative

- Forget absolute universals and stop asking "What's possible?", instead ask "**What's Where Why?**"
- But, how exactly? My goals here:
  - Show one specific approach to the "What's Where Why?" question: **Distributional Typology**

The alternative

- Forget absolute universals and stop asking "What's possible?", instead ask "**What's Where Why?**"
- But, how exactly? My goals here:
  - Show one specific approach to the "What's Where Why?" question: **Distributional Typology**
  - Show that this turns typology from a Pāṇinian into a normal science

WHAT

WHAT

- Given the choice, assume variables not universals

WHAT

- Given the choice, assume variables not universals
- *Example:* do not assume the Prosodic Hierarchy as an absolute universal (Vogel 2009), but take its nodes as variables

WHAT

- Given the choice, assume variables not universals
- *Example:* do not assume the Prosodic Hierarchy as an absolute universal (Vogel 2009), but take its nodes as variables

WHAT

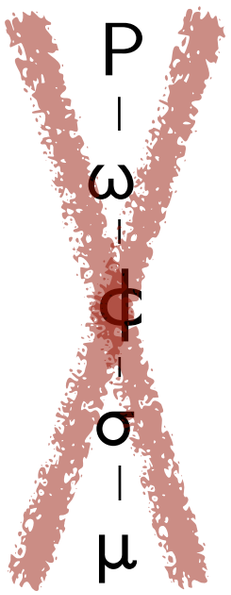
- Given the choice, assume variables not universals
- *Example:* do not assume the Prosodic Hierarchy as an absolute universal (Vogel 2009), but take its nodes as variables

But no rule combining syllables into words in Vietnamese:

hoa hồng ‘rose; pink flower’

WHAT

- Given the choice, assume variables not universals
- *Example:* do not assume the Prosodic Hierarchy as an absolute universal (Vogel 2009), but take its nodes as variables

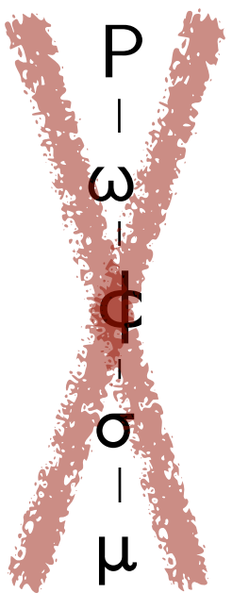


But no rule combining syllables into words in Vietnamese:

hoa hồng ‘rose; pink flower’

WHAT

- Given the choice, assume variables not universals
- *Example:* do not assume the Prosodic Hierarchy as an absolute universal (Vogel 2009), but take its nodes as variables



Alternative (Schiering, Hildebrandt & Bickel 2010):

Each phonological pattern targets a sequence of

- phonological units, creating μ, σ, ϕ
- morphological units, creating w-domains (ω, P)

Some variables per language:

1. number of phonological patterns targeting w-domains
2. size of w-domains
3. number of non-isomorphic w-domains

etc.

WHAT: Case study on words

1. Number of (non-item-specific) phonological patterns targeting w-domains

WHAT: Case study on words

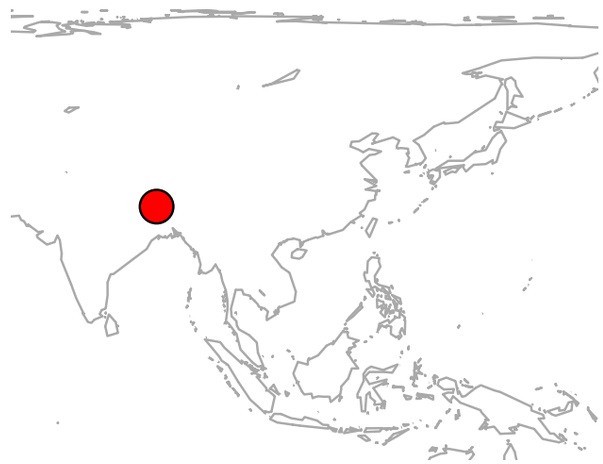
1. Number of (non-item-specific) phonological patterns targeting w-domains

   - Vietnamese (Austro-Asiatic): N=0

WHAT: Case study on words

1. Number of (non-item-specific) phonological patterns targeting w-domains

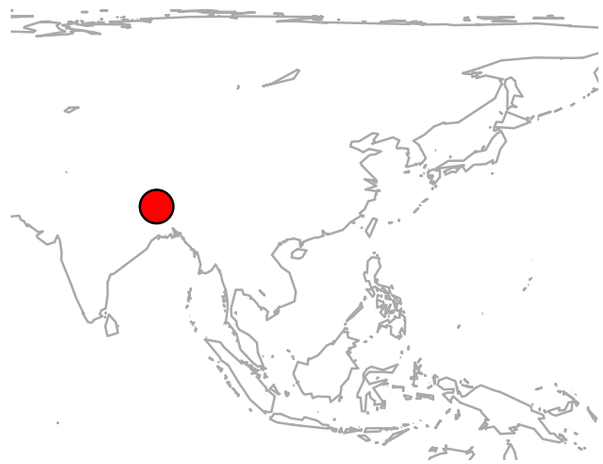
   - Vietnamese (Austro-Asiatic): N=0
   - Limbu (Sino-Tibetan): N=10



WHAT: Case study on words

1. Number of (non-item-specific) phonological patterns targeting w-domains

   - Vietnamese (Austro-Asiatic): *N*=0
   - Limbu (Sino-Tibetan): *N*=10



Liquid Alternation etc

gm-[**stem-gm-ptcl**]

kɛ-[**L**i'-**L**e-**L**oː] > kɛ[**l**i'**r**e**r**oː]

2sPOSS-bow-GEN-PTCL

'of your bow'

mɛ-[**L**uːg-ɛ-**L**oː] > mɛ[**l**uːgɛ**r**oː]

[3]nsS-fall-PST-PTCL

'they fell down'

Coronal Assimilation etc.

[**gm-stem-gm-ptcl**]

[kɛ-**n**-pa] > [kɛ**m**ba]

2sPOSS-KIN-father

'your father'

[mɛ-**n**-mɛ**t**-paŋ] > [mɛ**m**mɛ**p**paŋ]

[1]nsA-NEG-tell-1>3[s].PST

'we did not tell him'

WHAT: Case study on words

2. Size of w-domains: proportion of morpheme types inside the domain

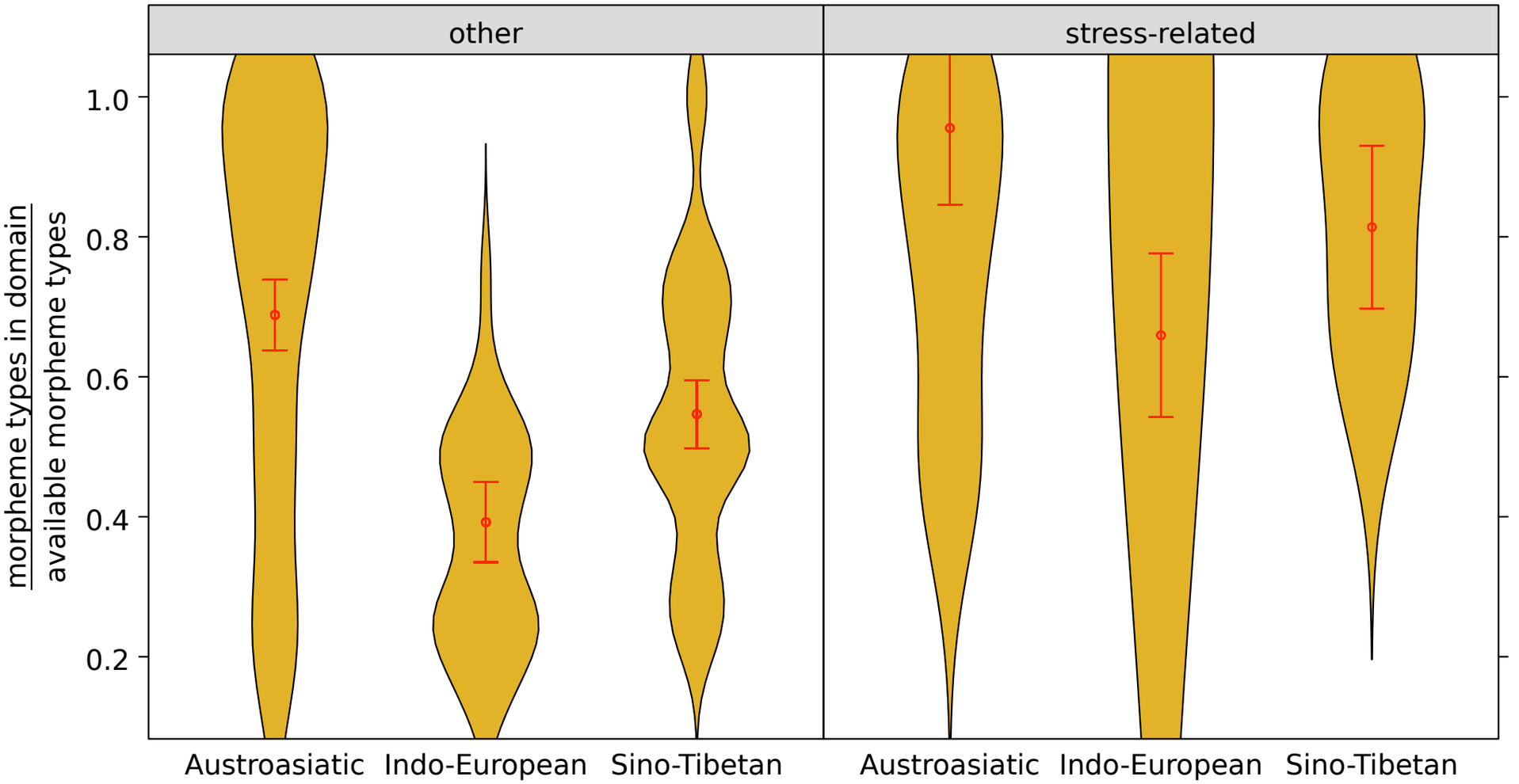
2. Size of w-domains: proportion of morpheme types inside the domain
   - [gm-stem-gm-ptcl]: $s = 1$

2. Size of w-domains: proportion of morpheme types inside the domain
   - [gm-stem-gm-ptcl]: $s = 1$
   - gm-[stem-gm-ptcl]: $s = \frac{3}{4}$

WHAT: Case study on words

2. Size of w-domains: proportion of morpheme types inside the domain
   - [gm-stem-gm-ptcl]: $s = 1$
   - gm-[stem-gm-ptcl]: $s = \frac{3}{4}$

Survey of 40 languages from three families:



Stress-related:
e.g. Czech
['*do=Prah-y*]

WHAT: Multivariate typology

WHAT: Multivariate typology

- Don't categorize languages into gross types, use fine-grained matrices.

WHAT: Multivariate typology

- Don't categorize languages into gross types, use fine-grained matrices.
- Example "ergative" in Hindi

WHAT: Multivariate typology

- Don't categorize languages into gross types, use fine-grained matrices.
- Example "ergative" in Hindi

Nīnā *hãs-ī.*
Nina.**NOM** laugh-PP.SG.FEM
S

'Nina laughed.'

Nīnā-ne *film* *dekh-ī.*
Nina-**ERG** movie.NOM see-PP.SG.FEM
A P

'Nina saw a movie.'

WHAT: Multivariate typology

- Don't categorize languages into gross types, use fine-grained matrices.
- Example "ergative" in Hindi

Nīnā *hãs-ī.*
Nina.**NOM** laugh-PP.SG.FEM
S

'Nina laughed.'

Nīnā-ne *film* *dekh-ī.*
Nina-**ERG** movie.NOM see-PP.SG.FEM
A P

'Nina saw a movie.'

Nīnā-ne *nahā-yā.*
Nina-**ERG** bathe-PP.SG.MASC
S

'Nina bathed.'

WHAT: Multivariate typology

- Don't categorize languages into gross types, use fine-grained matrices.
- Example "ergative" in Hindi

Nīnā *hãs-ī.*
Nina.**NOM** laugh-PP.SG.FEM
S

'Nina laughed.'

Nīnā-ne *film* *dekh-ī.*
Nina-**ERG** movie.NOM see-PP.SG.FEM
A P

'Nina saw a movie.'

Nīnā-ne *nahā-yā.*
Nina-**ERG** bathe-PP.SG.MASC
S

'Nina bathed.'

Nīnā-ko *ṭhaṇḍī* *lag-ā.*
Nina-**ACC** cold feel-PP.SG.MASC
S

'Nina felt cold.'

Nīnā-ko *sikṣak* *dikh-ā.*
Nina-**ACC** teacher[NOM] visible-PP.SG.MASC
A P

'Nina saw the/a teacher.'

WHAT: Multivariate typology

- Don't categorize languages into gross types, use fine-grained matrices.
- Example "ergative" in Hindi

Nīnā *hãs-ī.*
Nina.**NOM** laugh-PP.SG.FEM
S

'Nina laughed.'

Nīnā-ne *film* *dekh-ī.*
Nina-**ERG** movie.NOM see-PP.SG.FEM
A P

'Nina saw a movie.'

Nīnā-ne *nahā-yā.*
Nina-**ERG** bathe-PP.SG.MASC
S

'Nina bathed.'

Nīnā-ko *ṭhaṇḍī* *lag-ā.*
Nina-**ACC** cold feel-PP.SG.MASC
S

'Nina felt cold.'

Nīnā-ko *sikṣak* *dikh-ā.*
Nina-**ACC** teacher[NOM] visible-PP.SG.MASC
A P

'Nina saw the/a teacher.'

Nīnā-ne *film-ko* *dekh-ā.*
Nina-ERG movie-**ACC** see-PP.SG.MASC
A P

'Nina saw the movie.'

WHAT: Multivariate typology

ISO639.3	ID	Role	PoS	co.Role	co.PoS	PredCat	Clause	Predicate Class
ctn	1327	A_d	non-excl	ANY	ANY	ANY	ANY	Primary object verbs: some verbs denoting covering events, events of destructive impact like throwing, kicking, hitting, or cutting
ctn	1327	A_d	N	ANY	ANY	ANY	ANY	Primary object verbs ...
ctn	1327	A_d	non-excl	ANY	ANY	ANY	ANY	Double object verbs: physical and mental transfer events (translated as 'send, bring, take, move to, give, pass to, infect, feed, tell, ask for, show' etc.), also verbs like *yukt-* 'to keep for someone', which represent a kind of 'intended transfer'; verbs of covering ('cover, bury, pour, throw, spray at, soil, stain,' etc.)
ctn	1327	A_d	N	ANY	ANY	ANY	ANY	Double object verbs ...
ctn	1327	A_d	non-excl	ANY	ANY	ANY	ANY	the default ditransitive predicate class
ctn	1327	A_d	N	ANY	ANY	ANY	ANY	the default ditransitive predicate class
ctn	1327	A	non-excl	ANY	ANY	ANY	ANY	the default transitive predicate class
ctn	1327	A	N	ANY	ANY	ANY	ANY	the default transitive predicate class
ctn	1327	T	non-excl	ANY	ANY	ANY	ANY	Primary object verbs ...
ctn	1327	T	N	ANY	ANY	ANY	ANY	Primary object verbs ...
hin	92	A_d	ANY	ANY	ANY	ANY	PP-`hin`	main	the default ditransitive predicate class
hin	92	A	ANY	ANY	ANY	ANY	PP-`hin`	main	predicates with ERG depending on 'conscious choice' or volitioANYlity (alterANYtion possible only in perfective): *samajh* 'understand, suppose', *bhul* 'forget', *jan* 'give birth (to)', *phãd* 'leap over', *bak* 'to talk nonsense', *har* 'lose, be defeated' (Butt 2001: 127)
hin	92	A	ANY	ANY	ANY	ANY	PP-`hin`	main	the default transitive predicate class
hin	92	S	ANY	ANY	ANY	ANY	PP-`hin`	main	predicates with ERG/NOM-Sintr/Atr depending on 'conscious choice' or volitioANYlity (alterANYtion possible only in perfective): intr. verbs: *bhõk* 'bark', *jhãk* 'peep, look into/ through', *khãs* 'cough', *chĩk* 'sneeze', *muskara* 'smile', *thuk* 'spit', *mut* 'uriANYte', *hag* 'defecate', *naha* 'bathe', *ro* 'cry', *hãs* 'laugh', *so* 'sleep' (Butt 2001: 127)

WHERE AND WHY

WHERE AND WHY

- Variables have distributions (in time and space)

WHERE AND WHY

- Variables have distributions (in time and space)
- Goal: predict these distributions?

WHERE AND WHY

- Variables have distributions (in time and space)
- Goal: predict these distributions?
- A normal science, non-Pāṇinian approach:

WHERE AND WHY

- Variables have distributions (in time and space)
- Goal: predict these distributions?
- A normal science, non-Pāṇinian approach:

  1. A causal theory on what determines the distribution

WHERE AND WHY

- Variables have distributions (in time and space)
- Goal: predict these distributions?
- A normal science, non-Pāṇinian approach:

  1. A causal theory on what determines the distribution

  2. A statistical model based on this

WHERE AND WHY

- Variables have distributions (in time and space)
- Goal: predict these distributions?
- A normal science, non-Pāṇinian approach:

  1. A causal theory on what determines the distribution

  2. A statistical model based on this

  3. A method for testing the model against data

WHERE AND WHY

- Variables have distributions (in time and space)
- Goal: predict these distributions?
- A normal science, non-Pāṇinian approach:

  1. A causal theory on what determines the distribution
  2. A statistical model based on this
  3. A method for testing the model against data

- Same here! (No Working Hypothesis, no explaining away of counter-examples etc.)

WHERE AND WHY

- Variables have distributions (in time and space)
- Goal: predict these distributions?
- A normal science, non-Pāṇinian approach:

  1. A causal theory on what determines the distribution
  2. A statistical model based on this
  3. A method for testing the model against data

- Same here! (No Working Hypothesis, no explaining away of counter-examples etc.)
- Illustrate by way of a case study

Causal theories

1. A functional theory: presence of A≠P case is driven by V-final word order (Greenberg 1963, Siewierska 1996, Dryer 2002, Hawkins 2004 etc.)

   [NP V] : [$Ø_A$ NP_P V] or [NP_A $Ø_P$ V]

   [NP-x_P V]: [$Ø_A$ NP_P V]

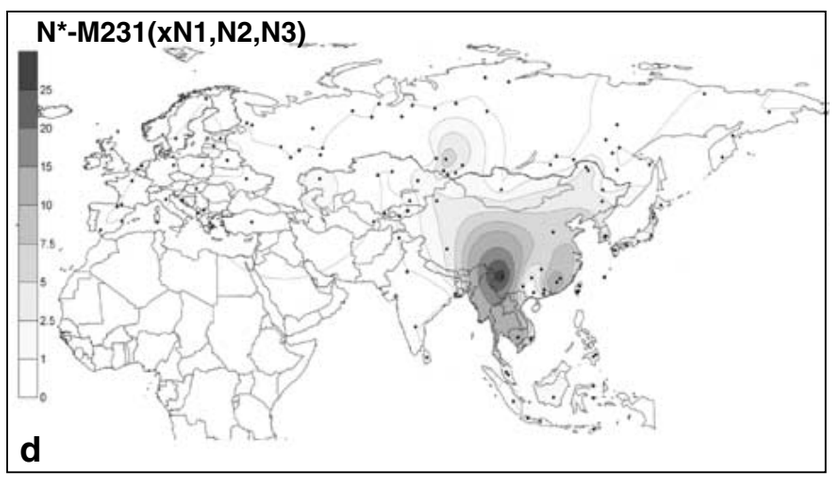
Causal theories

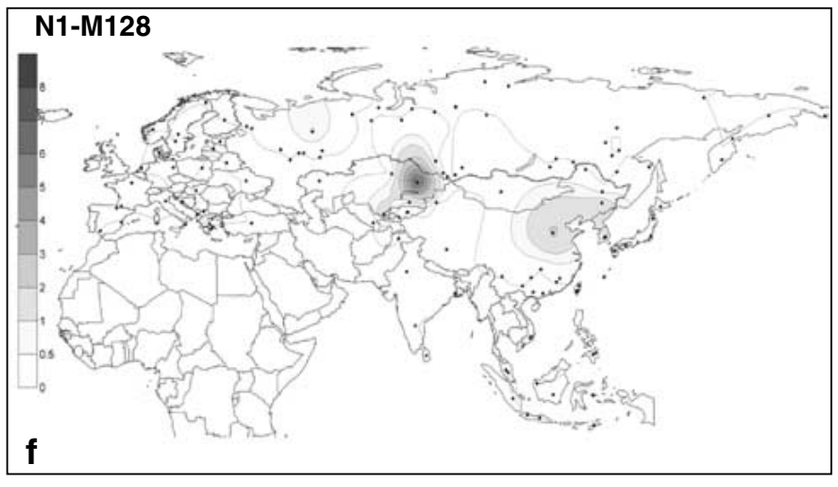
2. Event-based theory: presence of A≠P case is driven by diffusion in the wake of the Eurasian spreads

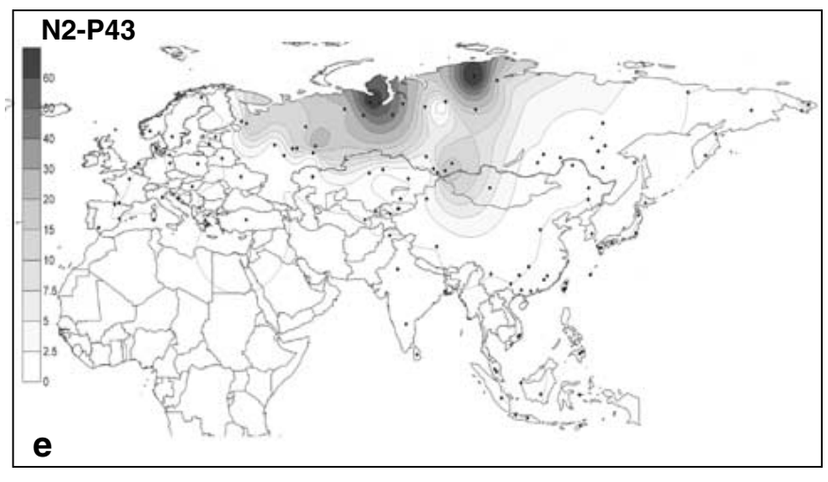
Causal theories

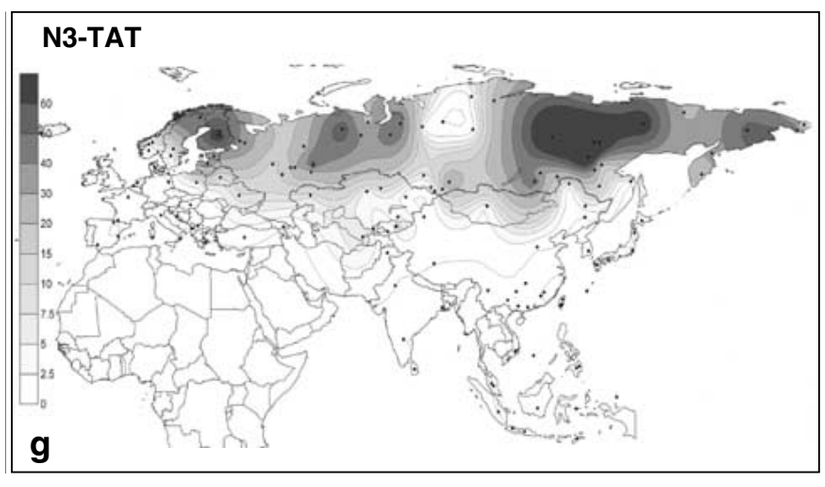
2. Event-based theory: presence of A≠P case is driven by diffusion in the wake of the Eurasian spreads

14 ±4 ky





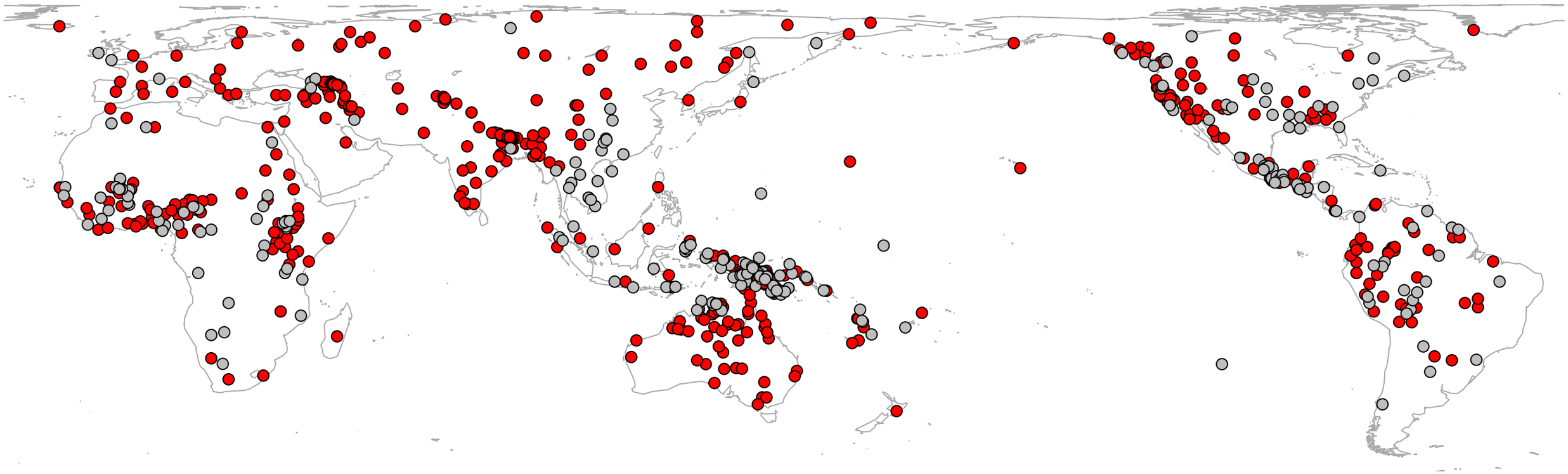


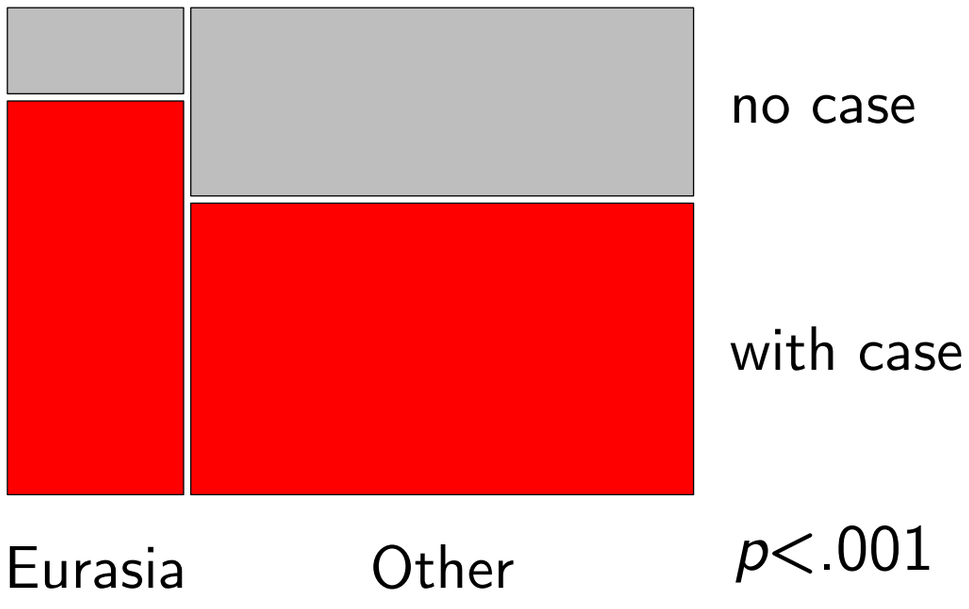


Rootsi et al. 2007 in *Europ. J. Hum. Gen.*

Causal theories

Data from AUTOTYP (Witzlack-Makarevich et al. 2011+) on case and WALS (Dryer 2005) on word order: $\boldsymbol{N = 489}$





'with case' = marking that differentiates between two argumental NPs of at least one kind (e.g. only first and second person pronouns) in at least some bivalent predicates (e.g. perhaps only in some experiencer predicates with an oblique experiencer).

Statistical modeling

Statistical modeling

- Model the possible impact of functional and event-based factors at the same time and watch out for interactions

Statistical modeling

- Model the possible impact of functional and event-based factors at the same time and watch out for interactions

  $\rightarrow$ standard regression models

Statistical modeling

- Model the possible impact of functional and event-based factors at the same time and watch out for interactions

  $\rightarrow$ standard regression models

- Model these impacts over time, since they are the result of diachronic biases of retention or innovation of prefered patterns

Statistical modeling

- Model the possible impact of functional and event-based factors at the same time and watch out for interactions

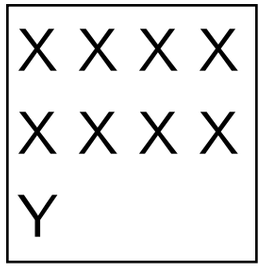
  $\rightarrow$ standard regression models

- Model these impacts over time, since they are the result of diachronic biases of retention or innovation of prefered patterns

  $\rightarrow$ Family Bias Method

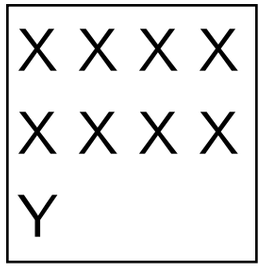
The Family Bias Method

Synchronic observations
on *demonstrably related*
languages:



The Family Bias Method

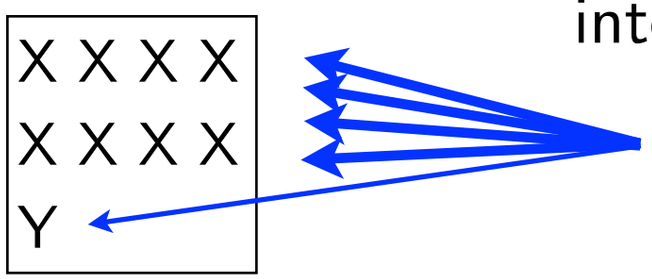
Synchronic observations on *demonstrably related* languages:



Possible diachronic interpretations:

The Family Bias Method

Synchronic observations on *demonstrably related* languages:

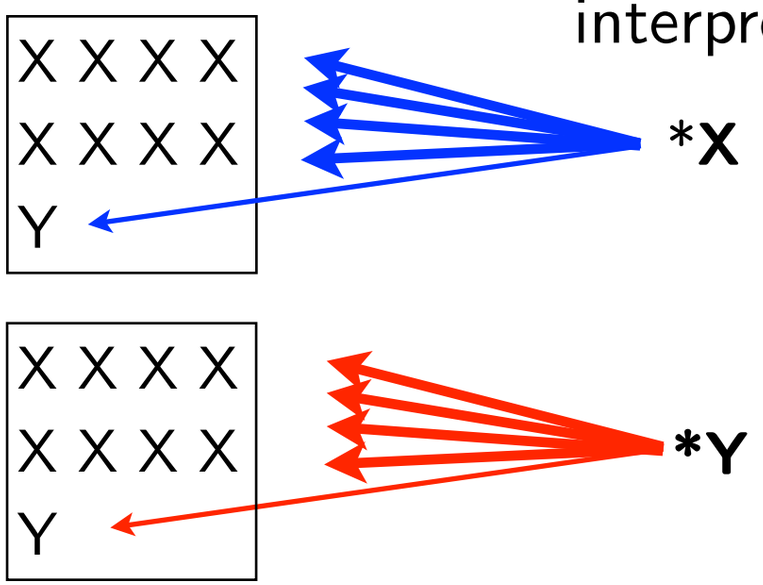


Possible diachronic interpretations:

***X**

The Family Bias Method

Synchronic observations on *demonstrably related* languages:



Possible diachronic interpretations:

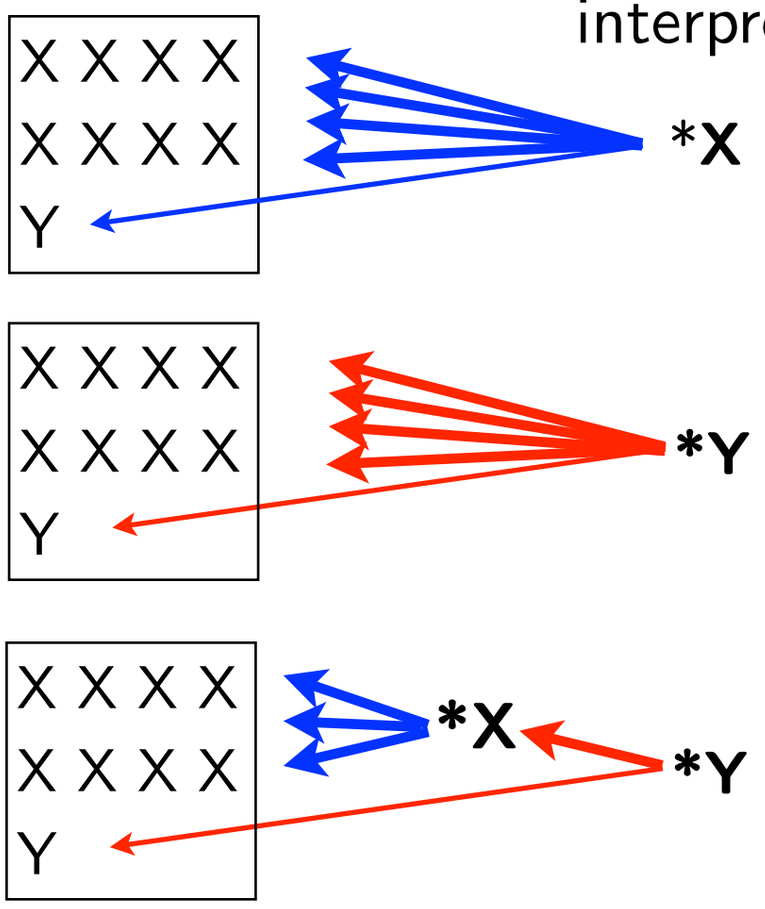
***X**

***Y**

The Family Bias Method

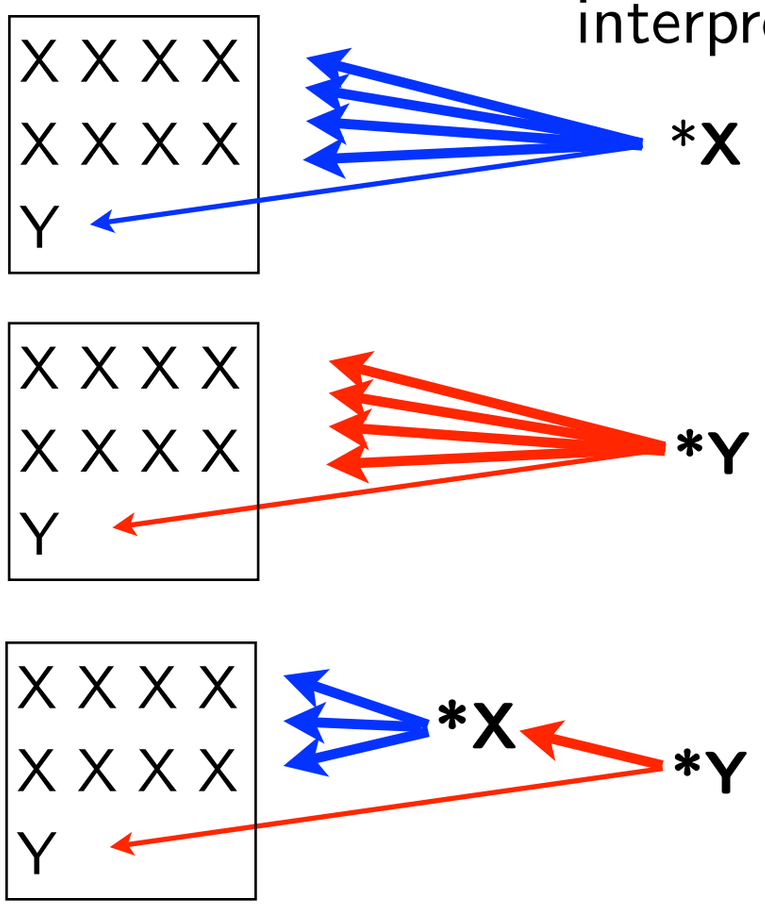
Synchronic observations on *demonstrably related* languages:

Possible diachronic interpretations:



The Family Bias Method

Synchronic observations on *demonstrably related* languages:

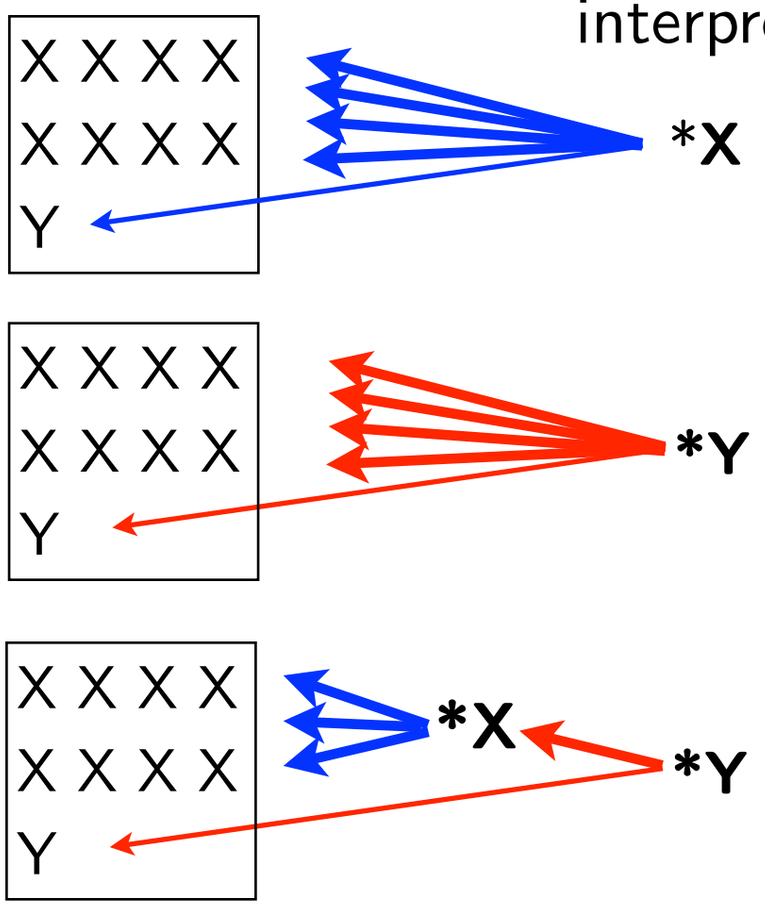


Possible diachronic interpretations:

Conclusion: different probabilities of innovation *and* retention

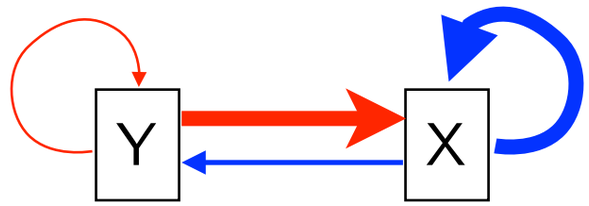
The Family Bias Method

Synchronic observations on *demonstrably related* languages:



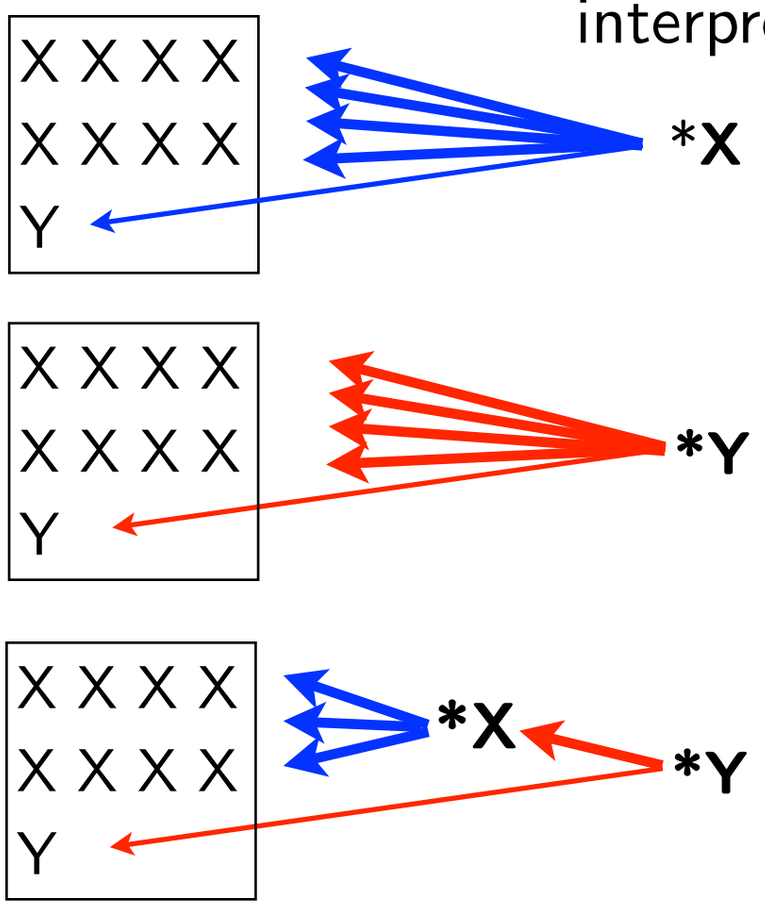
Possible diachronic interpretations:

Conclusion: different probabilities of innovation *and* retention



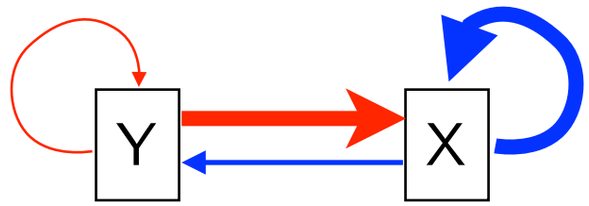
The Family Bias Method

Synchronic observations on *demonstrably related* languages:



Possible diachronic interpretations:

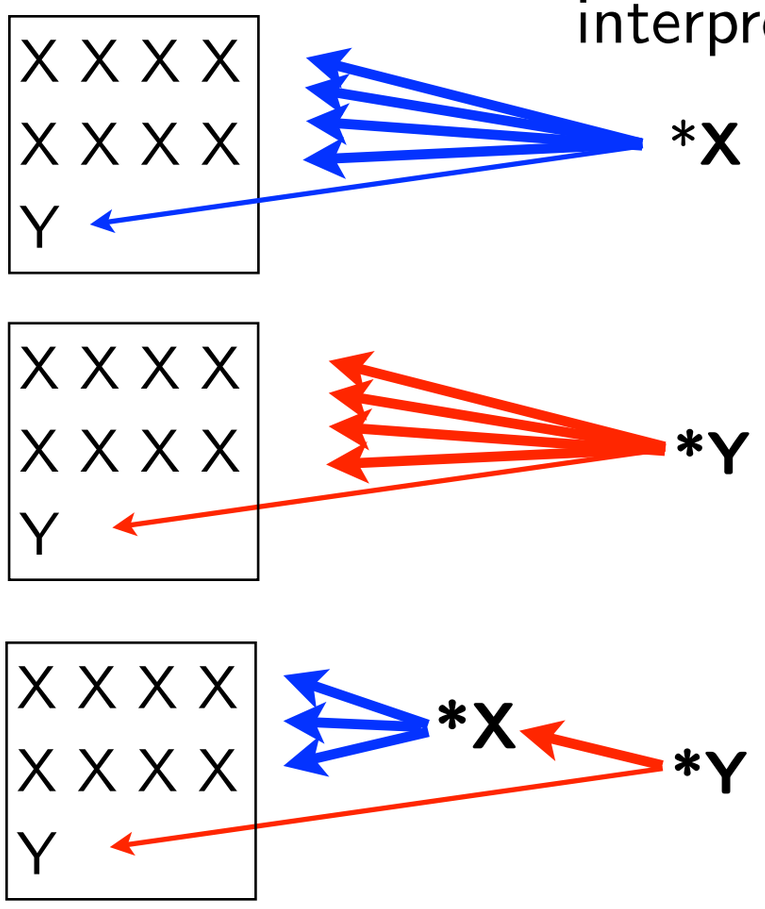
Conclusion: different probabilities of innovation *and* retention



$Pr(Y{>}X) > Pr(X{>}Y)$

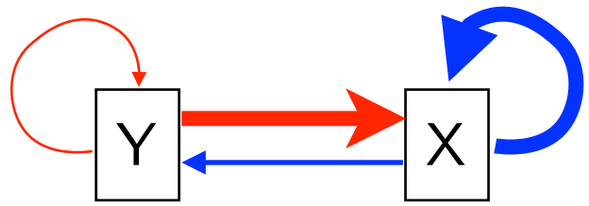
The Family Bias Method

Synchronic observations on *demonstrably related* languages:



Possible diachronic interpretations:

Conclusion: different probabilities of innovation *and* retention

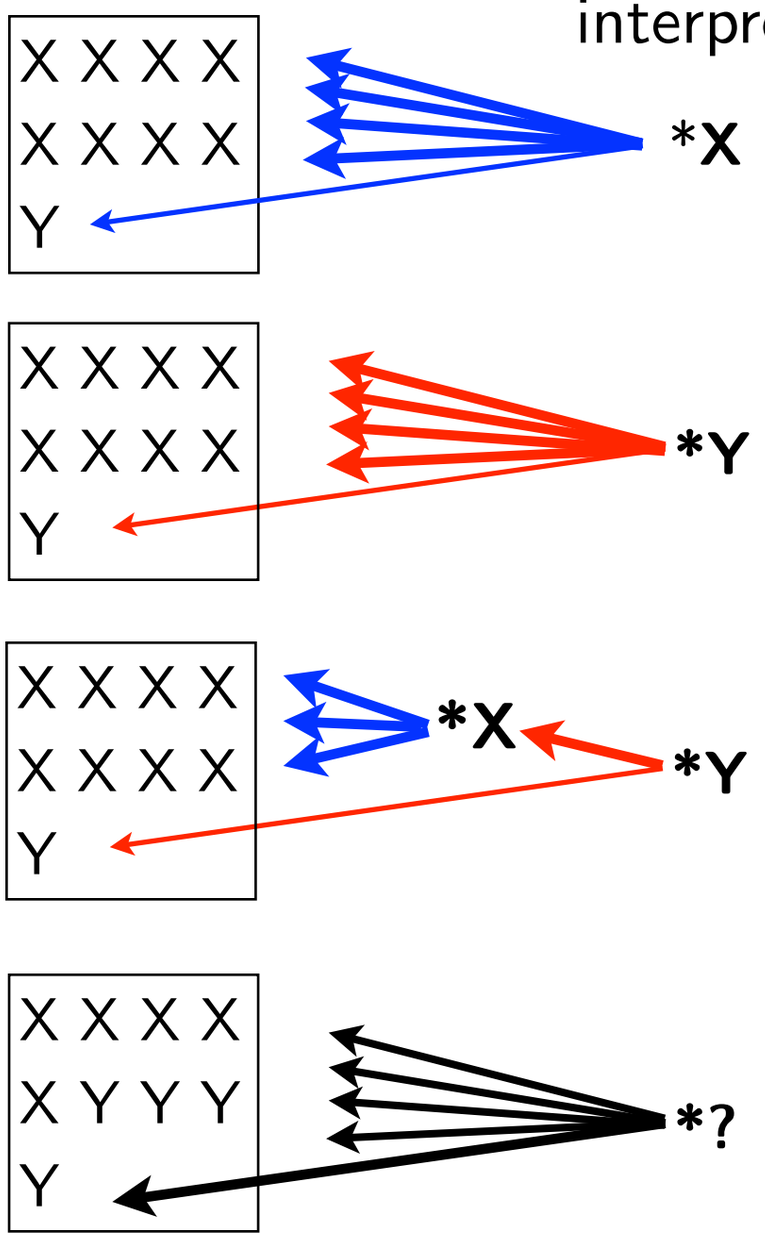


$Pr(Y{>}X) > Pr(X{>}Y)$

("**Family Bias**")

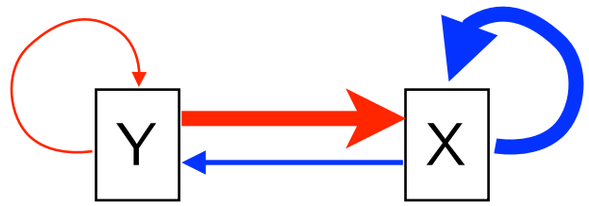
The Family Bias Method

Synchronic observations on *demonstrably related* languages:



Possible diachronic interpretations:

Conclusion: different probabilities of innovation *and* retention

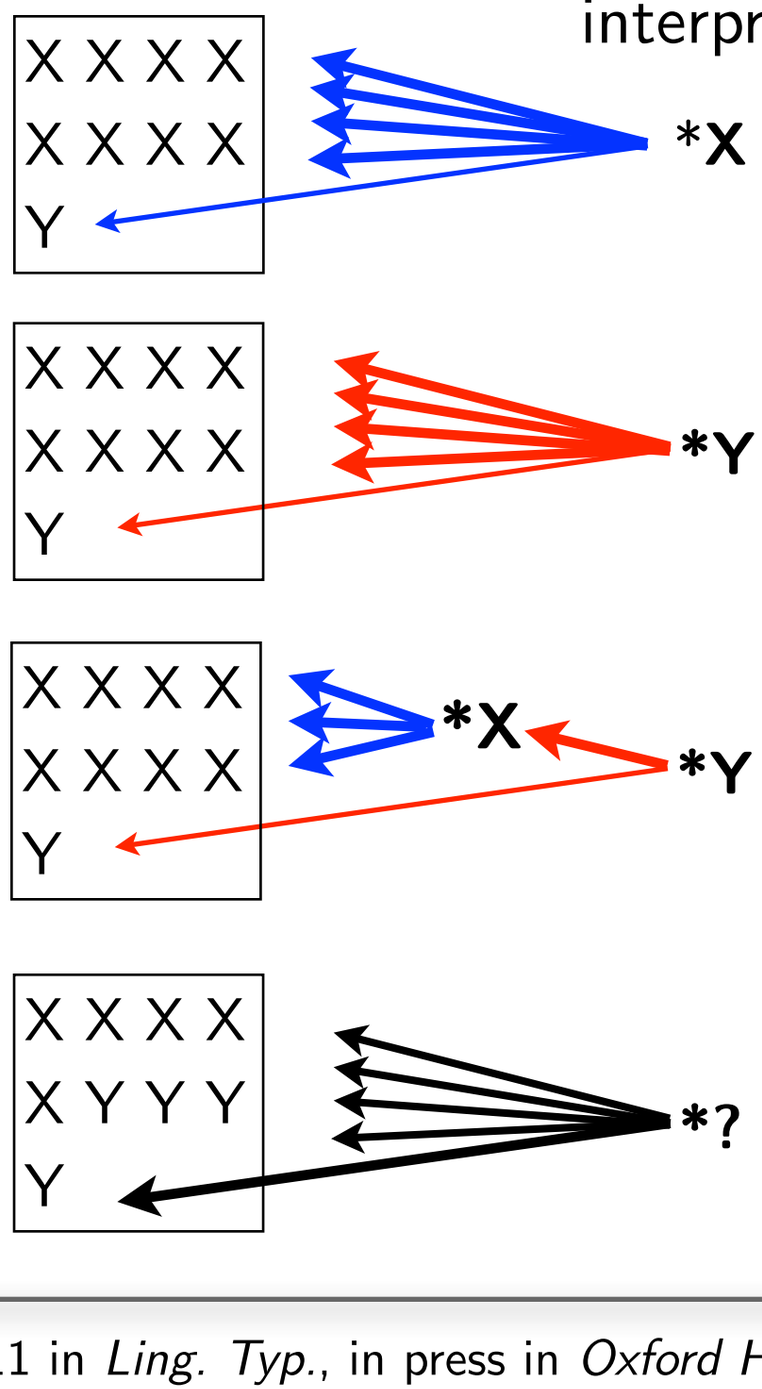


$Pr(Y{>}X) > Pr(X{>}Y)$

(“**Family Bias**”)

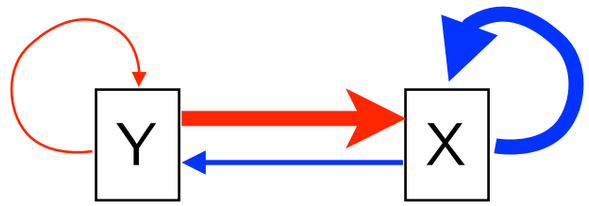
The Family Bias Method

Synchronic observations on *demonstrably related* languages:



Possible diachronic interpretations:

Conclusion: different probabilities of innovation *and* retention



$Pr(Y{>}X) > Pr(X{>}Y)$

(“**Family Bias**”)

$Pr(Y{>}X) \approx Pr(X{>}Y)$

(“no bias”, “diverse”)

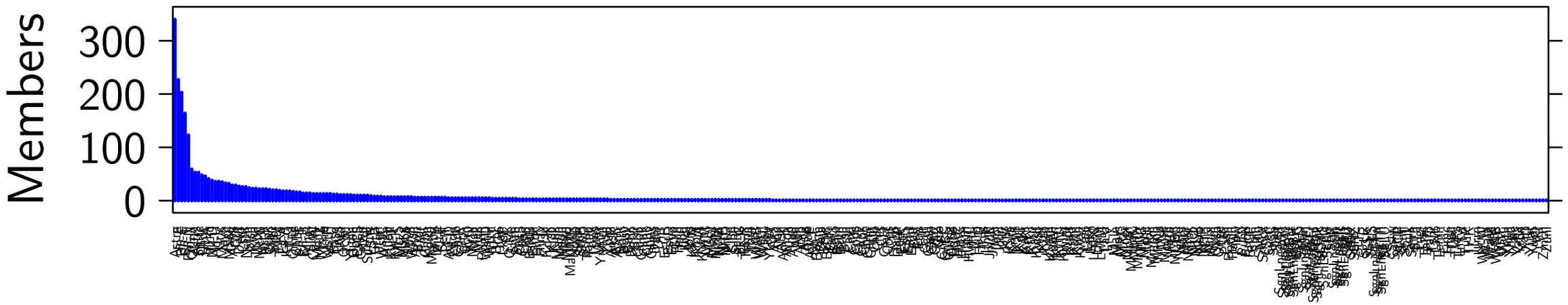
The Family Bias Method

- Estimate biases in large families ($N \geq 5$), using binomial tests

The Family Bias Method

- Estimate biases in large families ($N \geq 5$), using binomial tests
- Extrapolate to small families based on bias probabilities of large families and the data in small families, including single-member families (isolates, or families represented only by one member in a given database)

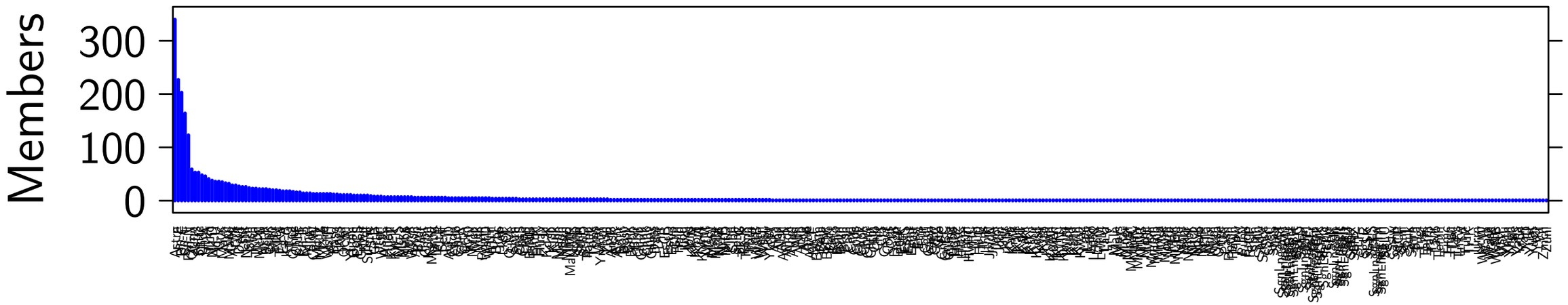
because, after all, this is where the data are:



The Family Bias Method

- Estimate biases in large families ($N \geq 5$), using binomial tests
- Extrapolate to small families based on bias probabilities of large families and the data in small families, including single-member families (isolates, or families represented only by one member in a given database)

  because, after all, this is where the data are:

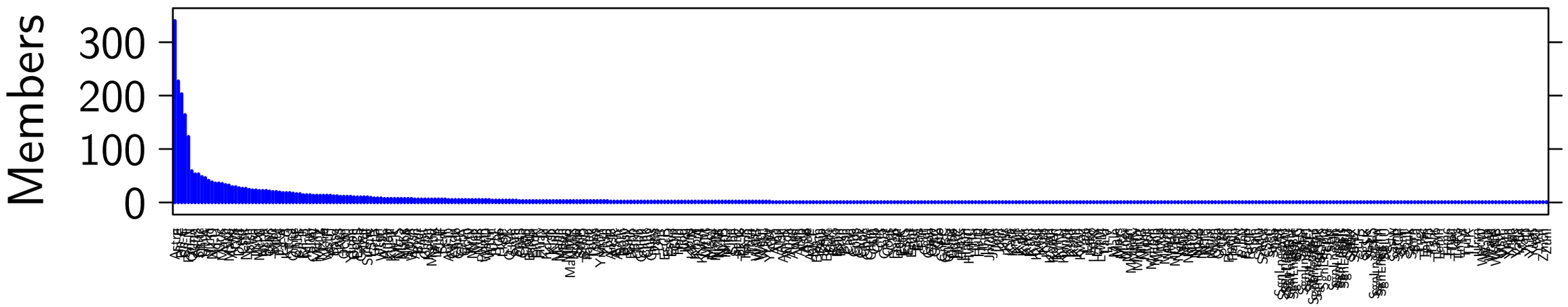


- Sample the world as exhaustively as possible (depart from the tradition!)

The Family Bias Method

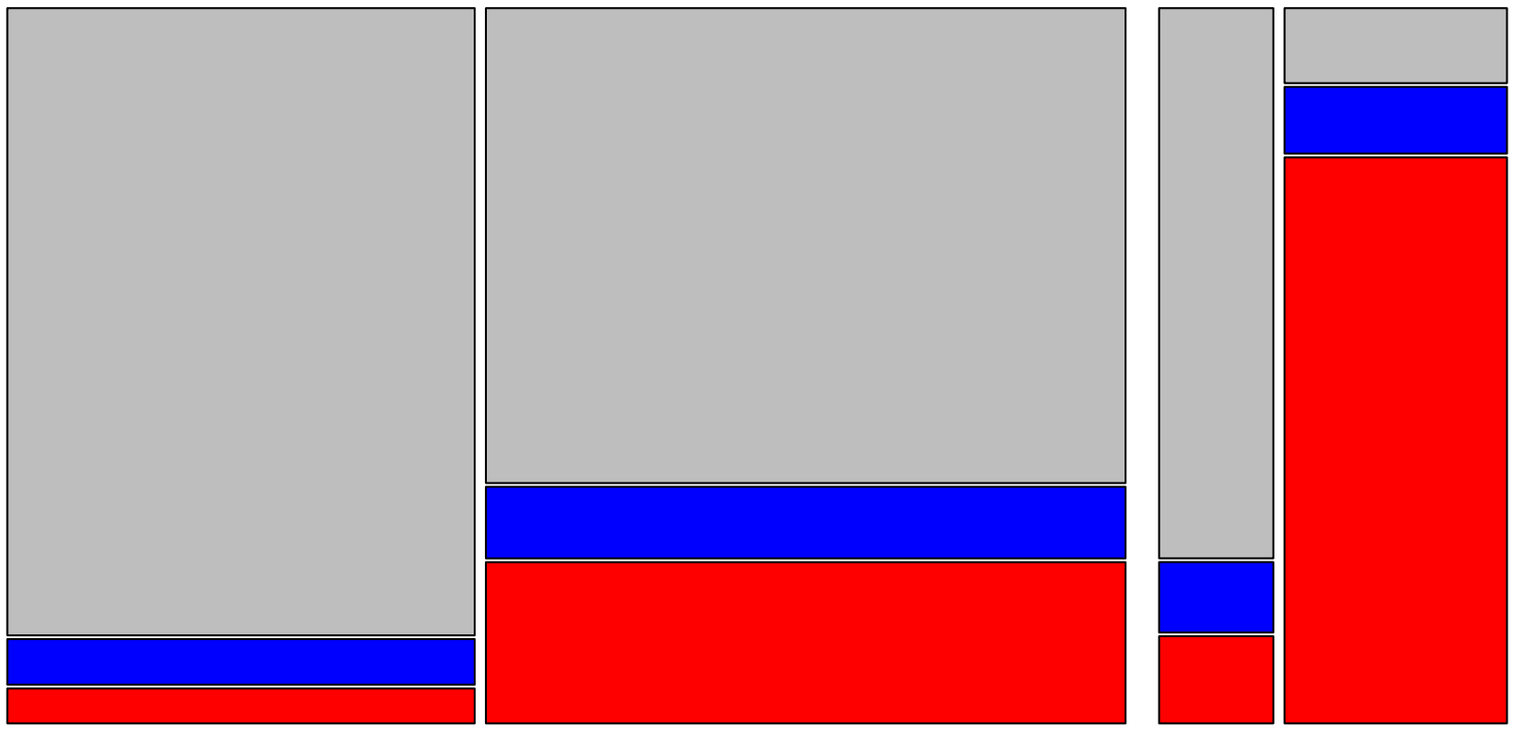
- Estimate biases in large families ($N \geq 5$), using binomial tests
- Extrapolate to small families based on bias probabilities of large families and the data in small families, including single-member families (isolates, or families represented only by one member in a given database)

  because, after all, this is where the data are:



- Sample the world as exhaustively as possible (depart from the tradition!)
- Software available at `http://www.uzh.ch/spw/software`

Testing the effects



V...] ...V] V...] ...V]

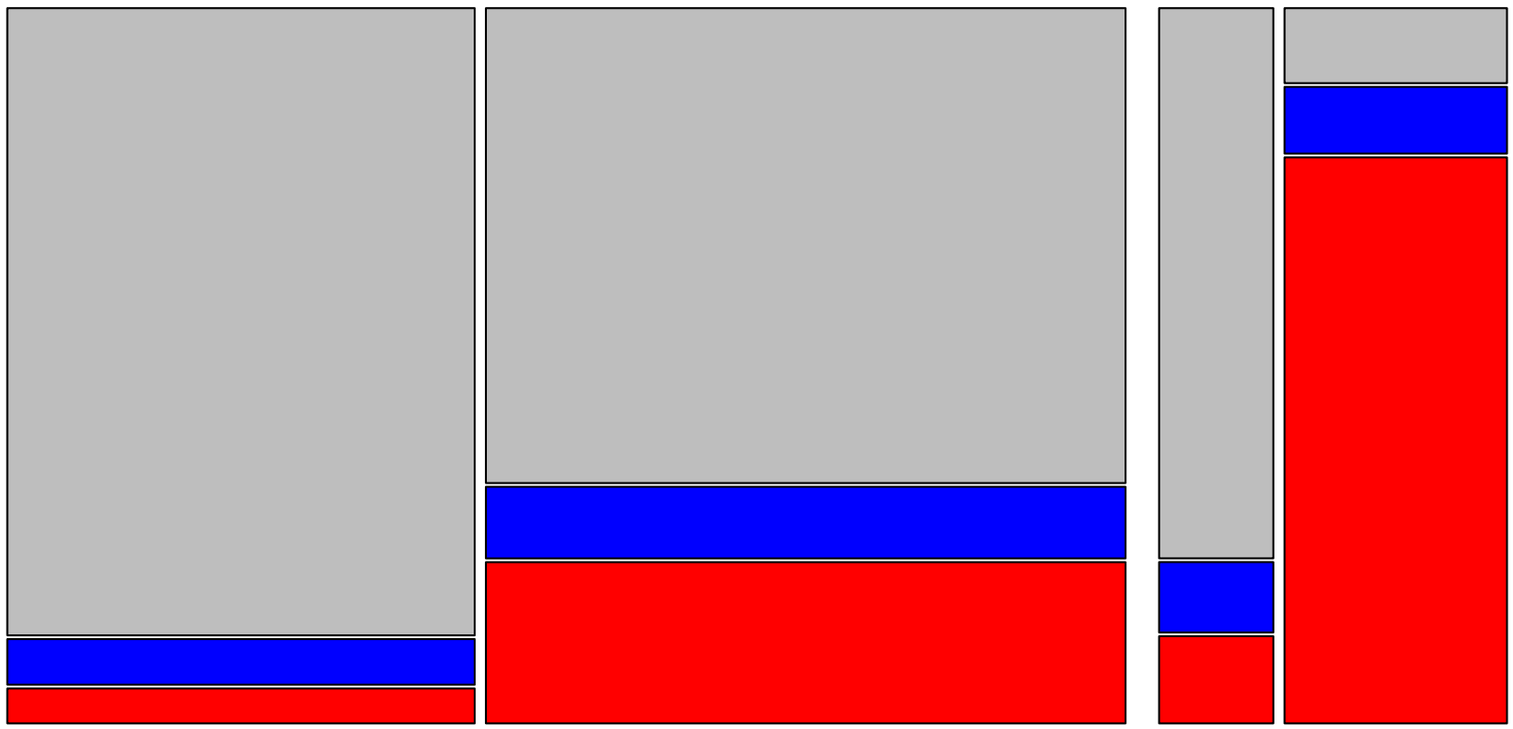
Other Eurasia

Bias for case vs. against case is determined both

- by the contact history of Eurasia: case tends to be better preserved or (re-)created in Eurasia (AREA × BIAS TYPE**,** p=.034)
- by processing principles: case tends to be better preserved or (re-)created in v-final families (ORDER × BIAS TYPE, p=.027)

These effects are independent of each other (three-way interaction is *n.s.*)

Testing the effects



Diversification vs. stability is determined both

- by the contact history of Eurasia, but only in v-final groups (three-way interaction, p=.011): v-final groups diversify less in Eurasia than elsewhere (AREA × DIVERSITY, p<.001), no such effect in non-final groups
- by processing principles: v-final languages diversify less than non-v-final languages (factorial analysis across areas, both p<.001)

Conclusions

Conclusions

- Distributional Typology follows the normal science triad — causal theory, statistical modeling, testing — instead of Pāṇini

  (The details in statistical methods may (and will) change, but the mode of thinking stays.)

Conclusions

- Distributional Typology follows the normal science triad — causal theory, statistical modeling, testing — instead of Pāṇini

  (The details in statistical methods may (and will) change, but the mode of thinking stays.)

- Results do **not** depend on

Conclusions

- Distributional Typology follows the normal science triad — causal theory, statistical modeling, testing — instead of Pāṇini

  (The details in statistical methods may (and will) change, but the mode of thinking stays.)

- Results do **not** depend on
  - individual datapoints ("counterexamples") and fights on what is the "right" analysis, but on general patterns

Conclusions

- Distributional Typology follows the normal science triad — causal theory, statistical modeling, testing — instead of Pāṇini

  (The details in statistical methods may (and will) change, but the mode of thinking stays.)

- Results do **not** depend on
  - individual datapoints ("counterexamples") and fights on what is the "right" analysis, but on general patterns
  - sampling choices since the Family Bias Method uses *exhaustive* samples

Conclusions

- Distributional Typology follows the normal science triad — causal theory, statistical modeling, testing — instead of Pāṇini

  (The details in statistical methods may (and will) change, but the mode of thinking stays.)

- Results do **not** depend on
  - individual datapoints ("counterexamples") and fights on what is the "right" analysis, but on general patterns
  - sampling choices since the Family Bias Method uses *exhaustive* samples

- Distributional Typology fits with the old insight that **nothing in linguistics makes sense expect in the light of history** (cf. Dobzhansky re biology),

  .... as linguists knew all along!