Quantitative Methods in Typology

Balthasar Bickel

University of Leipzig

Rationale of this course

	20th century “Universalist Typology”
nature	method
goal	formulating UG
interest in	absolute implicational universals
part of	classical CogSci
core question	What’s a Possible/ Learnable Human Language?

Anderson 2008: “The object of inquiry in linguistics is the human ability to acquire and use a natural language, and the goal of linguistic theory is an explicit characterization of that ability.”

A paradigm change

Three developments favor a move away from this

- Recognition of ‘universal areality’
- No sample can tell us what is (im)possible
- But samples allow probabilistic inference, using the same techniques as other disciplines (esp. bio-informatics)
- Recognition that answers to the “core question” may better come from empirical language acquisition research and comparative psychology than from linguistics per se

A paradigm change

	20th century “Universalist Typology”	21st century “Distributional Typology”
nature	method	discipline
goal	formulating UG	explaining distributions
interest in	absolute implicational universals	distributional probabilities
part of	classical CogSci	anthropology, including psychology
core question	What’s a Possible/ Learnable Human Language?	What’s Where Why?

What's Where Why?

- *what:* what linguistic structures are there and how can we compare them?
- *where:* how are languages and their structures distributed in space and time, both locally and universally?
- *why:* what factors determine the distribution of languages and their structures in space and time?

What’s Where Why?

Any understanding of What’s Where Why involves quantitative methods:

- *what:* measuring similarities
- *where:* examining the distribution of similarities
- *why:* developing statistical models of the distributions

Quantitative Methods in Typology I:

Comparing structures

Balthasar Bickel

University of Leipzig

Comparing structures

- Bloomfield: “If you want to compare two languages, it helps to know one of them.”
- The Descriptive A Priori = analytical metalanguages or (when named and popular) ‘theoretical frameworks’
- Two necessary ingredients:

1. Absolute Universals = primitive concepts that are required for the analysis of every imagineable language (“Theoretical universal”)
2. Variables = sets (inventories, scales) of primitive concepts that are required for the analysis of at least one language

   Other terms: *tertium comparationis,* comparative concepts, metrics, parameters, features

Too many options to take: how to justify metalanguages?

- One popular answer: claim the metalanguage to be psychologically and biologically real (Chomsky: “UG“, “S_0”) (Descriptive A Priori = Ontological A Priori)

- Problems:

  - if the metalanguage also includes variables (‘has vs. does not have clicks’), all values of the variable must be universally present in the brain, i.e. even in languages that don’t have evidence for them. How can such claims ever be tested? (Nichols 2008)

  - the biological evidence is unclear (gap between genes, expression of genes, and behavior) (e.g. Müller 2009)

  - language acquisition does not seem to rely on any of our linguistic metalanguages but on general learning mechanisms (Tomasello 2003)

A pragmatic approach to metalanguages

- The working linguist’s everyday answer: justify the metalanguage not by appeal to psychology or biology but only by standard principles of scientific inquiry:
  - logical consistency
  - operationalizability (availability of proofs and argumentation)
  - replicability
  - universal descriptive success: all structures can be analyzed in a cross-linguistically consistent way, allowing comparison (compare “subject” and “proto-agent”)
- Compare this to the use of the metrical system as justified by its consistency, operationalizability, replicability, and descriptive and ‘typological’ success — and not by any innateness claim!

A pragmatic approach to metalanguages

- Conceived therefore as analytical instruments, the metalanguage
  - makes no claim on genetic and/or cognitive reality (but may pick up signals that can be explained by cognitively and possibly even genetic patterns, cf. Hawkins 2004, Dediu & Ladd 207)
  - makes no claim on the distribution of its denotata in the world (they could be all over, like “nasals”, or rare like “clicks”)
  - therefore, it provides the right means to study actual distributions, i.e. what’s where why?
- So, let’s work out a good metalanguage!

A key problem for metalanguages:

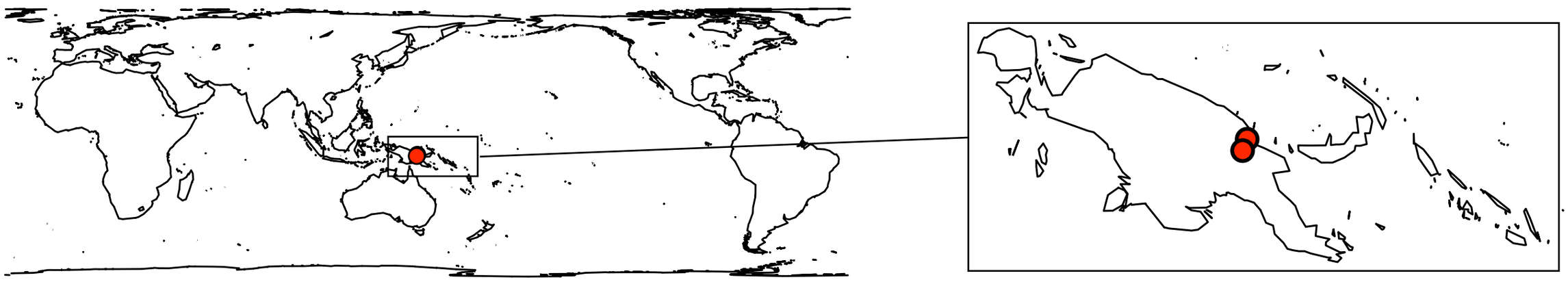
- The structures we find often don’t quite fit with what we know: e.g. is X “really” a passive? “really” an affix? “really” subordinate?
- There are always many similarities but few identities!

‘many similarities, few identities’

- A case study (Bickel in press): “cosubordination”

Amele (Trans-New-Guinea: Madang, Roberts 1987:101)
ija Malolo ugba na ka ji? ana-g na ono nu sum-ud-i bi-bil-igin
1s M. 3s LOC car road mother-3sPOSS LOC there PURP wait-3s-PRED SIM-be-DUR1sDS
ne-?e-b tobo-?o-min belo-w-an.
come.down-DS-3s climb.up-DS-1s go-1d-Y.PST
‘While I waited there at the main road for Malolo’s car he came down. I climbed in and off we went.’

Tauya (Trans-New-Guinea: Madang, MacDonald 1990:101)
wiwi-?ai yate fitau-a-te me nono-ra toute fatemasete-pa ?u-a-?a.
moss-ADESS go TEL-3s-DS DEM child-TOP stick sharpen-SS dig-3s-IND
‘She went away into the moss, and this boy sharpened a stick and dug.’



‘many similarities, few identities’

- Universal definition: clause linkage with conjunct illocutionary scope (operator dependency) (e.g., Foley and Van Valin 1984)

Amele (Trans-New Guinea: Madang; Papua New Guinea; Roberts 1988)
ho busale-ʔe-b dana age gbo-ig-a fo?
pig run.out-DS-3s man 3p hit-3p-T.PST Q
‘Did the pig run out and did the men kill it?’

Tauya (Trans-New Guinea: Madang, Papua New Guinea; MacDonald 1990:226)
tepau-fe-pa yate fitau-a=nae?
break-PRF-SS go throw-2=POLAR.Q
‘Did you break it and go away?’
or ‘Did you go away after breaking it?’ (presupposing either ‘you went away’ or ‘you broke it’)

Classical responses

- Basic principle: the universal definition follows from our metalanguage (the Descriptive A Priori), it cannot have exceptions.
- Therefore, *what looks like an exception, isn't! We can get rid of it, "explain it away".*
- In other words: *Reduce the diversity before you study it!*
- This is characteristic both of theory-of-grammar and typology approaches:

Classical theory-of-grammar approaches

- Search for a higher-ranking principle, e.g. a principle that explains why Tauya *-pa* allows disjunct scope although it "really" establishes cosubordination

  —Problem: no principle known!

- Limit the scope of the definition, e.g. Tauya does not have cosubordination, but "sub-cosubordination"

  —Problem: the structures are so similar to each other that one reading of cosubordination is the sole reading of "sub-cosubordination".

Classical theory-of-grammar approaches (cont'd)

- Assume structural ambiguity: in Tauya, one reading 'really' reflects cosubordination while the other reflects something else (probably subordination, with disjunct scope) (Bickel 1998).

  — Problem: No independent evidence for this

- Revise the definition, e.g. define cosubordination by constraint-free scope (Bickel 1991, Croft 2001)

  — Problem: We can base the definition on any variable we want (e.g. finiteness, tense scope, extraction possibilities, assertedness, etc.) but we may always run into the same problems! "Methodological opportunism" (Croft 2001)

Classical typological approaches

- Typologize exemplars (“basic” cosubordination, like “basic word order“)

  —Problem: which one to pick?

- Define “comparative concepts” that abstract away from language-particular details (Lazard 2006, Haspelmath 2007), perhaps via a “functional” definition (e.g. conjunct scope) —

  —Problem: all language-specific analysis is informed by and benefits from cross-linguistically defined concepts.

- All classical responses convey a sense of “After all, languages are not that different from each other”, and thereby prevent a deeper understanding of the human faculty for diversity (cf. Evans & Levinson 2009)

An alternative: Multivariate Typology

- Diversity means that across languages, things are mostly similar and hardly ever identical.
- But similarity is nothing else but identity in some variables and difference in others.
- For studying similarities, we need large systems of fine-grained variables that fully capture the range of known variation: **Multivariate Typology.**
- And with this, we can **describe and measure the variation, instead of reducing it** — i.e. do what most other disciplines would do when confronted with variation.

An alternative: Multivariate Typology

- How many structures? — As many as are distinct in V_1 ... V_k
- How many variables? — As many as are of interest to the research question, e.g. all variables needed to capture cross-linguistic differences in the syntax of clause linkage; or in morphological coding.
- Which variables? — Developed as needed for distinguishing structures during data collection (Autotypologizing Method: Bickel & Nichols 2002), or pre-determined by the research question.
  - structural variables: morphosyntactic or semantic properties in which structures are alike or differ.
  - denotation variables: denotations (stimuli, contexts, functions) in which structures are alike or differ

‘many similarities, few identities’: Multivariate typology

clause_linkage (xserve2.rz.uni-leipzig.de)

R...	LID	::Language	II...	::Scope	III...	::Marking	T...	::Scope	T...	::Marking	Foc...	::Marking	Cat...	::categorical_sym...	Cla...	::clause_position	Emb...	::embedding	Jun...	::juncture	Interpr...	::interproposition...
1	14	Alamblak	4	disjunct	1	ok	1	constraint-free	1	ok			1	symmetrical	4	flexible-relational	2	adjoined	2	clause	2	conditional
2	20	Amele	2	blocked	2	banned	1	constraint-free	1	ok			1	symmetrical	1	fixed-final	2	adjoined	2	clause	4	inconsequential
3	20	Amele	3	conjunct	2	banned	3	conjunct	3	harmonic	2	banned	2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
4	20	Amele	3	conjunct			1	constraint-free	1	ok	2	banned	1	symmetrical	1	fixed-final	2	adjoined	2	clause	3	disjunction
5	20	Amele	4	disjunct	1	ok	1	constraint-free	1	ok	1	ok	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	2	conditional
52	20	Amele	4	disjunct	1	ok			1	ok			1	symmetrical	1	fixed-final	2	adjoined	2	clause	6	purposive
7	35	Belhare	1	constraint-free			1	constraint-free	1	ok	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
8	35	Belhare	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
9	35	Belhare			1	ok	1	constraint-free	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	4	inconsequential
10	35	Belhare	4	disjunct			1	constraint-free	1	ok	2	banned	2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	5	concessive
11	35	Belhare	4	disjunct	1	ok	1	constraint-free	1	ok	1	ok	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
12	40	Burushaski	1	constraint-free			1	constraint-free	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
13	629	Chantyal					1	constraint-free	2	banned			2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
51	2862	Chintang	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	3	core	1	narrative
54	71	Dyirbal					1	constraint-free	1	ok			1	symmetrical	5	fixed-initial	2	adjoined				
55	71	Dyirbal					3	conjunct	2	banned			1	symmetrical	5	fixed-initial	2	adjoined			6	purposive
57	71	Dyirbal							2	banned							2	adjoined	4	NP		
14	74	English	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause		
15	74	English	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	6	purposive
16	1494	Fore	1	constraint-free	1	ok	3	conjunct	3	harmonic			2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
17	87	German	4	disjunct	2	banned	1	constraint-free	1	ok	1	ok	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	7	causal
18	87	German	4	disjunct	1	ok	1	constraint-free	1	ok	2	banned	1	symmetrical	4	flexible-relational	2	adjoined	1	sentence	7	causal
19	1556	Godié	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause		
20	1556	Godié	4	disjunct	2	banned	1	constraint-free	1	ok			3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
6	2831	Greek (Ancient)	4	disjunct	2	banned	1	constraint-free	2	banned			2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	9	topic
21	103	Hua	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	8	
22	103	Hua	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	8	
23	103	Hua	1	constraint-free			3	conjunct	3	harmonic			2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
24	103	Hua	4	disjunct	2	banned	1	constraint-free	1	ok	2	banned	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
25	103	Hua	4	disjunct	2	banned	1	constraint-free	1	ok	2	banned	2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	4	inconsequential
26	103	Hua	4	disjunct	2	banned	1	constraint-free	1	ok			2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	2	conditional
27	126	Kâte	1	constraint-free	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
28	126	Kâte	3	conjunct	1	ok	1	constraint-free	2	banned			2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
29	1806	Korafe	1	constraint-free			1	constraint-free	1	ok			2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
30	483	Lenakel					1	constraint-free	1	ok			2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
31	206	Nepali	1	constraint-free	2	banned	3	conjunct	1	ok	1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
32	206	Nepali	4	disjunct	2	banned	1	constraint-free	1	ok	1	ok	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
33	468	Newar (Dolakha)	1	constraint-free					2	banned	1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
34	2866	Pali	1	constraint-free							1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
56	340	Russian	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	1	symmetrical	2	flexible-adjacent	2	adjoined				
35	361	Swahili	3	conjunct			3	conjunct	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
36	361	Swahili	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause		
37	361	Swahili	3	conjunct	1	ok	1	constraint-free	1	ok	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause		
38	361	Swahili	4	disjunct							1	ok	2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	9	topic
39	367	Tauya	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	8	
40	367	Tauya	1	constraint-free	1	ok	3	conjunct	2	banned			2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
41	367	Tauya	2	blocked									2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	4	inconsequential
42	367	Tauya	2	blocked	1	ok							3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
43	2864	Toura	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
44	502	Turkish	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
45	502	Turkish	3	conjunct	1	ok					2	banned	2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
46	393	Usan	1	constraint-free									2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative

100 Browse

Example: variables of clause linkage

ILL-scope: The scope of illocutionary operators in the main clause is

CONJUNCT: extends to the main clause and the dependent clause

DISJUNCT: extends to either the main or the dependent clause but never to both

LOCAL: is limited to the main clause

EXTENSIBLE: extends to either the main clause alone or to both the main clause and the dependent clause, but never to the dependent clause alone

CONSTRAINT-FREE: is not regulated by the clause linkage type

T-scope: The scope of tense or status operators in the main clause is

CONJUNCT: extends to the main clause and the dependent clause

LOCAL: is limited to the main clause

EXTENSIBLE: extends to either the main clause alone or to both the main clause and the dependent clause, but never to the dependent clause alone

Example: variables of clause linkage

Finiteness: The dependent clause is headed by a verb form that is

FINITE: at least as many categories must be marked as in main clauses

NONFINITE: only fewer categories are allowed

ANY: either the same range or less categories can be marked

ILL-mark: Marking of illocutionary force operators in the dependent clause is

OK: allowed

BANNED: not allowed

HARMONIC: allowed but only if it matches the marking on the main clause

T-mark: Marking of tense or status operators in the dependent clause is

OK: allowed

BANNED: not allowed

HARMONIC: allowed but subject to constraints based on the tense or status choice in the main clause

Example: variables of clause linkage

Fore (Trans New Guinea: Kainantu-Gorokan; Scott 1978)

a. *kana-a:-kí-tá* *a-ka-us=ó.*
come-3sPRES.DS-DEP-1dAS 3sP-see-1dA=Q
'Is he coming and we see it?'

b. *na-m-e-g-ánt=ó* *na-ku-w-e.*
1sP-give-2sFUT.DS-DEP-1sAS=IMP eat-FUT-1s-DECL
'Give me something and then I will eat it.' (*i.e.*, 'Give me something to eat!')

Hua (Trans-New Guinea: Kainantu-Gorokan; Haiman 1980:421)

a. *fu=mo d-mi-sa-ga-da* *u-gu-e.* ('DS-CHAIN')
pig=TOP 1sP-give-FUT-3pDS-1sAS go-FUT-1sDECL
'They will give me pork and then I will go.'

b. * *fu=mo d-mi-sa-ga-da* *u-e.* ('DS-CHAIN')
pig=TOP 1sP-give-FUT-3pDS-1sAS go[NFUT]-1sDECL
Intended: 'They will give me pork and so I went', *i.e.*, 'I went because they will give me pork.'

Example: variables of clause linkage

Symmetry: The range of categories that can be expressed on linked clauses is

SYMMETRICAL: must match

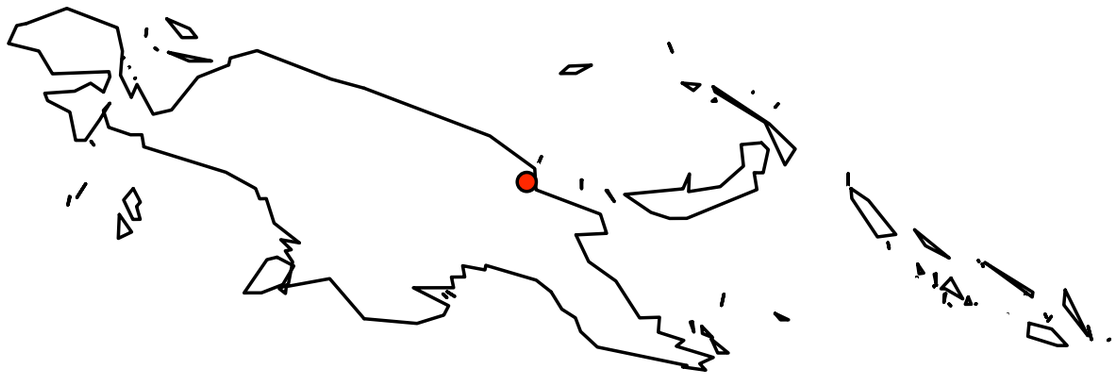
ASYMMETRICAL: can be different

FREE: can be different and can even include elements of different type (different parts of speech, clauses and NPs, etc.)

Amele (Roberts 1987, 1988)

a. * *ija ja hudo-ʔo-min gba ugba sab mane-i-a.*
1s fire open-DS-1s but 3s food cook-1s-T.PST

b. *ija ja hud-ig-a gba ugba sab mane-i-a.*
1s fire open-1s-T.PST but 3s food cook-1s-T.PST
'I lit the fire but she cooked the food.'



Example: variables of clause linkage

Usan (Reesink 1987: 283ff)

a. *munon iyau wârâm-or eng um-orei ʔiyo?*
man dog 3sP.hit-3sF.PST TOP die-3sF.PST Q
'Given that the man hit the dog, did it die?'

b. *yâr-âb eng ye-nipat ʔur big-âr.*
come-SS TOP 1sP-step.over.SS money put-pIMP
'If you come, step over me and put your money (in the basket).'

c. *munon eng, wonou man soau is-orei.*
man TOP 3sPOSS garden landslide go.down-3sF.PST
'As for the man, his garden went down in the landslide.'

Example: variables of clause linkage

WH: Question words and constituent focus inside dependent clauses are

OK: allowed

BANNED: not allowed

Extraction: Extraction of elements of dependent clauses is

OK: allowed

BANNED: not allowed

These are independent and additional variables for at least two reasons...

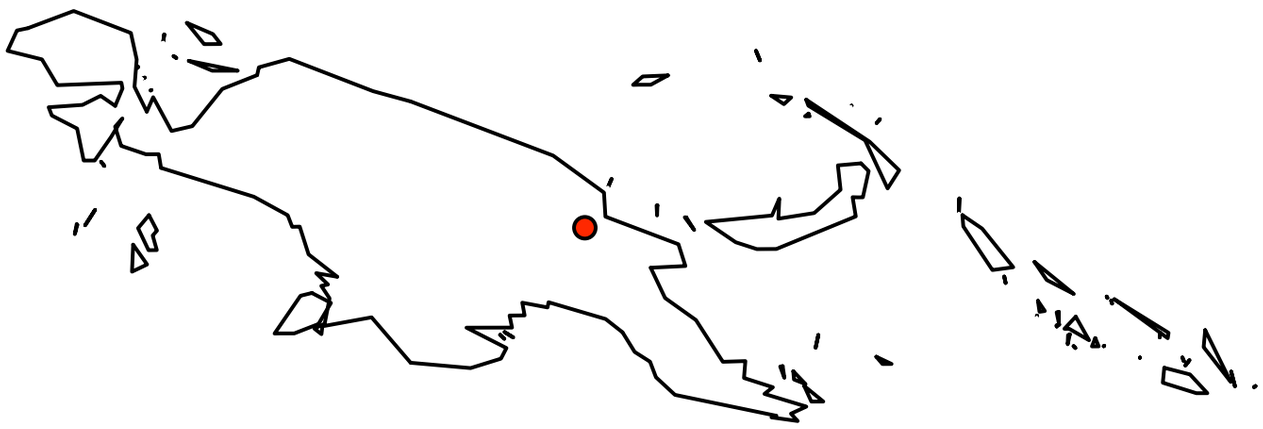
Example: variables of clause linkage

1. Constraints on WH may also hold in the absence of extraction (Foley & Van Valin 1984):

Tauya (MacDonald 1990)

a. *ne-ni we tu-a-te yau-i=ne?* (‘CHAIN’)
3s-ERG who[NOM] [3sP]give-3s-DS [3sP]see-3p=PARAMETRICAL.Q
‘Who did he give it to? and they saw him’ (‘Who did he give it to when they saw him?’)

b. * *we mei fofe-a-nani=ra …* (‘TOPIC’)
who[NOM] here come-3s-ASS=TOP
Intended: ‘*if who did come here…’



Example: variables of clause linkage

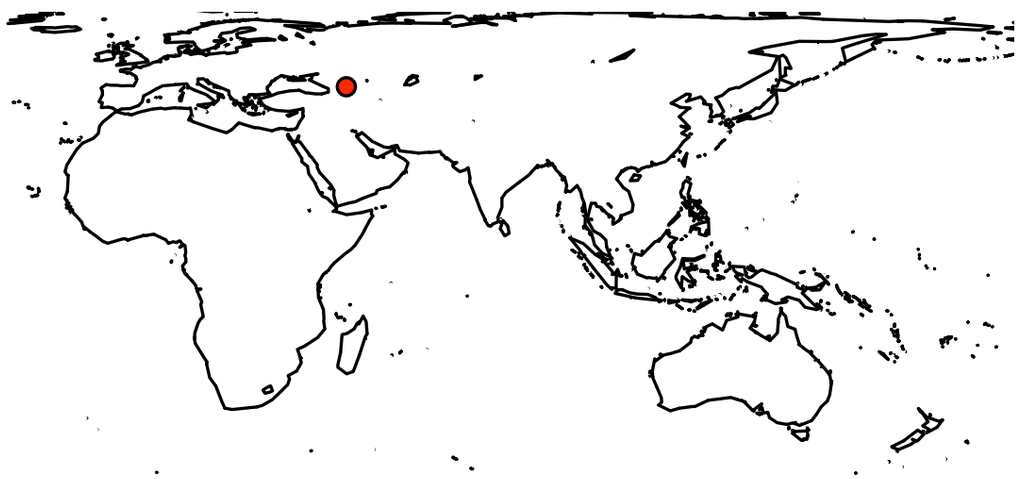
2. The domain relevant for extraction may differ from the domain relevant for WH

Chechen (Nakh-Daghestanian; Good 2003, Molochieva 2008)

a. *Maliika hu iec-na c'a j-e'a-ra?*
M.(J).NOM what buy-CVB house J-come-W.PST
'What did Malika buy and came home?'

b. *Zaara koch ec-na c'a j-e'a-ra.*
Z.(J).NOM dress.NOM buy-CVB home J-come-W.PST
'Zara bought a dress and came home.'

c. * *Zaara(-s) ec-na c'a j-e'a-cha j-olu koch*
Z.(J).(-ERG) buy-CVB home J-come-CVB J-AUX.PTCP dress
Intended: 'the dress that Zara bought and came home.'



Example: variables of clause linkage

FOC: Focus marking on dependent clauses is

OK: allowed

BANNED: not allowed

Chechen (Zarina Molochieva, p.c.)

a. *Maliika tyka-na='a j-agh-na(*='a) c'a j-e'a-ra.*
M.NOM store-DAT=SS J-go-CVB(=FOC) home J-come-W.PST
'Malika went to the store and then came back home.'

b. *Maliika hu iec-na(*='a) c'a j-e'a-ra?*
M.(J).NOM what buy-CVB(=FOC) house J-come-W.PST
'What did Malika buy and came home?'

c. *Maliika hu iec-cha='a c'a j-e'a-ra?*
M.(J).NOM what buy-WHEN=FOC house J-come-W.PST
'What did Malika buy and came home?'

Example: variables of clause linkage

Position: The position of the dependent clause vis-à-vis the main clause with which it enters a dependency relation is

FIXED:POST-MAIN: is fixed and is always after the main clause

FIXED:PRE-MAIN: is fixed and is always before the main clause

FLEXIBLE-ADJACENT: can be before or after the main clause but must be adjacent to it

FLEXIBLE-RELATIONAL: can be before or after the main clause and can be separated from the main clause by other dependent clauses

Chechen (Good 2003)

a. *Maliika tyka-na='a j-agh-na zhejna='a iec-na c'a j-e'a-ra.*
M.NOM store-DAT=SS J-go-CVB book.NOM=SS buy-CVB home J-come-W.PST

b. *Maliika tyka-na='a j-agh-na c'a j-e'a-ra zhejna='a iec-na.*
M.NOM store-DAT=SS J-go-CVB home J-come-W.PST book.NOM=SS buy-CVB

c. *Maliika c'a j-e'a-ra tyka-na='a j-agh-na zhejna='a iec-na.*
M.NOM home J-come-W.PST store-DAT=SS J-go-CVB book.NOM=SS buy-CVB
'Malika went to the store, bought a book, and came back home.'

Example: variables of clause linkage

Layer: The dependent clause adjoins

AD-V: to the predicate and can be center-embedded

AD-S: to the clause and cannot be center-embedded

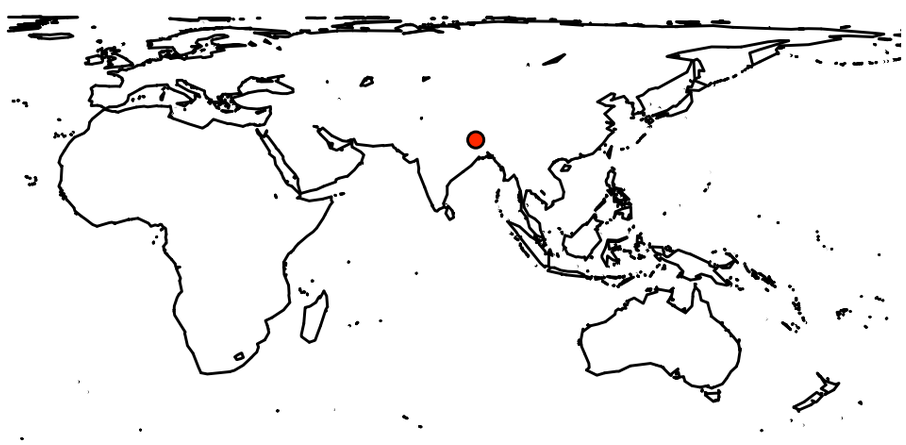
DETACHED: to the clause but is separated syntactically and intonationally

Belhare

a. *Dhankuta him-yakt-a-lok* *khar-e.*
D.[LOC] [3sS-]stumble-IPFV-PST.SBJV-COM [3sS]go-PST
'He went to Dhankuta stumbling.'

b. *u-chom* *pok=naa* *Dhankuta khaʔ-yu.*
3sPOSS-desire [3sS]rise[SBJV]=TOP D.[LOC] [3sS]go-NPST

c. * *Dhankuta u-chom* *pok=naa* *khaʔ-yu.*
D.[LOC] 3sPOSS-desire [3sS]come.up[SBJV]=TOP [3sS]go-NPST
'If he wants, he will go to Dhankuta.'



Example: variables of clause linkage

- Further variables
  - semantic relation between propositions
  - syndesis and prosody
  - cross-clausal coreference constraints (backward anaphora, switch-reference, logophorics etc)

Example: variables of clause linkage

clause_linkage (xserve2.rz.uni-leipzig.de)

R...	LID	::Language	II...	::Scope	III...	::Marking	T...	::Scope	T...	::Marking	Foc...	::Marking	Cat...	::categorical_sym...	Cla...	::clause_position	Emb...	::embedding	Jun...	::juncture	Interpr...	::interproposition...
1	14	Alamblak	4	disjunct	1	ok	1	constraint-free	1	ok			1	symmetrical	4	flexible-relational	2	adjoined	2	clause	2	conditional
2	20	Amele	2	blocked	2	banned	1	constraint-free	1	ok			1	symmetrical	1	fixed-final	2	adjoined	2	clause	4	inconsequential
3	20	Amele	3	conjunct	2	banned	3	conjunct	3	harmonic	2	banned	2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
4	20	Amele	3	conjunct			1	constraint-free	1	ok	2	banned	1	symmetrical	1	fixed-final	2	adjoined	2	clause	3	disjunction
5	20	Amele	4	disjunct	1	ok	1	constraint-free	1	ok	1	ok	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	2	conditional
52	20	Amele	4	disjunct	1	ok			1	ok			1	symmetrical	1	fixed-final	2	adjoined	2	clause	6	purposive
7	35	Belhare	1	constraint-free			1	constraint-free	1	ok	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
8	35	Belhare	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
9	35	Belhare			1	ok	1	constraint-free	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	4	inconsequential
10	35	Belhare	4	disjunct			1	constraint-free	1	ok	2	banned	2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	5	concessive
11	35	Belhare	4	disjunct	1	ok	1	constraint-free	1	ok	1	ok	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
12	40	Burushaski	1	constraint-free			1	constraint-free	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
13	629	Chantyal					1	constraint-free	2	banned			2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
51	2862	Chintang	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	3	core	1	narrative
54	71	Dyirbal					1	constraint-free	1	ok			1	symmetrical	5	fixed-initial	2	adjoined				
55	71	Dyirbal					3	conjunct	2	banned			1	symmetrical	5	fixed-initial	2	adjoined			6	purposive
57	71	Dyirbal							2	banned							2	adjoined	4	NP		
14	74	English	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause		
15	74	English	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	6	purposive
16	1494	Fore	1	constraint-free	1	ok	3	conjunct	3	harmonic			2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
17	87	German	4	disjunct	2	banned	1	constraint-free	1	ok	1	ok	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	7	causal
18	87	German	4	disjunct	1	ok	1	constraint-free	1	ok	2	banned	1	symmetrical	4	flexible-relational	2	adjoined	1	sentence	7	causal
19	1556	Godié	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause		
20	1556	Godié	4	disjunct	2	banned	1	constraint-free	1	ok			3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
6	2831	Greek (Ancient)	4	disjunct	2	banned	1	constraint-free	2	banned			2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	9	topic
21	103	Hua	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	8	
22	103	Hua	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	8	
23	103	Hua	1	constraint-free			3	conjunct	3	harmonic			2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
24	103	Hua	4	disjunct	2	banned	1	constraint-free	1	ok	2	banned	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
25	103	Hua	4	disjunct	2	banned	1	constraint-free	1	ok	2	banned	2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	4	inconsequential
26	103	Hua	4	disjunct	2	banned	1	constraint-free	1	ok			2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	2	conditional
27	126	Kâte	1	constraint-free	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
28	126	Kâte	3	conjunct	1	ok	1	constraint-free	2	banned			2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
29	1806	Korafe	1	constraint-free			1	constraint-free	1	ok			2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
30	483	Lenakel					1	constraint-free	1	ok			2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
31	206	Nepali	1	constraint-free	2	banned	3	conjunct	1	ok	1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
32	206	Nepali	4	disjunct	2	banned	1	constraint-free	1	ok	1	ok	3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
33	468	Newar (Dolakha)	1	constraint-free					2	banned	1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
34	2866	Pali	1	constraint-free							1	ok	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
56	340	Russian	1	constraint-free	1	ok	1	constraint-free	1	ok	1	ok	1	symmetrical	2	flexible-adjacent	2	adjoined				
35	361	Swahili	3	conjunct			3	conjunct	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
36	361	Swahili	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause		
37	361	Swahili	3	conjunct	1	ok	1	constraint-free	1	ok	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause		
38	361	Swahili	4	disjunct							1	ok	2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	9	topic
39	367	Tauya	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	1	symmetrical	2	flexible-adjacent	2	adjoined	2	clause	8	
40	367	Tauya	1	constraint-free	1	ok	3	conjunct	2	banned			2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
41	367	Tauya	2	blocked									2	asymmetrical	4	flexible-relational	2	adjoined	2	clause	4	inconsequential
42	367	Tauya	2	blocked	1	ok							3	constraint-free	4	flexible-relational	2	adjoined	2	clause	9	topic
43	2864	Toura	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative
44	502	Turkish	3	conjunct	2	banned	3	conjunct	2	banned	2	banned	2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
45	502	Turkish	3	conjunct	1	ok					2	banned	2	asymmetrical	1	fixed-final	2	adjoined	2	clause	1	narrative
46	393	Usan	1	constraint-free									2	asymmetrical	2	flexible-adjacent	2	adjoined	2	clause	1	narrative

100 Browse

But what about the good old typological generalizations?

- ... can’t just get them as well, but better!
- For this we need...

  A. Heuristics: data mining techniques, e.g.

    - distance-based techniques for finding clusters of similar structures (split graphs, multidimensional scaling etc.) (‘prototypes’, bottom-up ‘canons’)
    - entropy-based techniques for finding associations between variables (possibly weighted)

But what about the good old typological generalizations?

B. Test strategies: permutation-based statistics, e.g. of

- simple contingency tables
- generalized linear models

= next topic. For now, let’s explore some data mining techniques useful for comparing structures

Mining clause linkage data

- Are there any structures across languages that are more similar to each other than to others?
- A standard answer: compute the relative Hamming distance (proportion of different values in a set of variables)

Mining clause linkage data: measuring distances

Language	Marker	ILL-scope	T-scope	Finiteness	ILL-mark	T-mark	Symmetry	WH	Extraction	FOC	Position	Layer
Amele:but	*gba*	local	local	fin.	*	✔	symm.	NA	NA	NA	fix.:pre	ad-S
Amele:chain	*-me, -ʔV*	conj.	conj.	nonfin.	*	*	asymm.	NA	NA	NA	fix.:pre	ad-S
Amele:or	*fo ~o*	conj.	local	fin.	✔	✔	symm.	NA	NA	NA	fix.:pre	ad-S
Amele:cond	*fi*	disj.	extens.	any	✔	✔	flex	✔	NA	NA	flex-rel.	ad-S
Greek:p.c.	*-men/-nt*	disj.	extens.	nonfin.	*	*	asymm.	NA	NA	NA	flex-rel.	ad-S
Belhare:and	*=cha...=cha*	flex	local	fin.	✔	✔	symm.	NA	*	*	flex-adj.	ad-S
Belhare:chain	*ki(na)(huŋ)*	flex	extens.	fin.	harm.	✔	asymm.	✔	possible	✔	flex-adj.	ad-S
Belhare:inc	*-kone*	local	extens.	nonfin.	✔	*	asymm.	✔	*	✔	flex-adj.	ad-S
Belhare:sub	*-naa...*	disj.	extens.	fin.	*	✔	flex	✔	*	✔	flex-rel.	ad-S
Burúshaski:chain	*n(V)-´-Σ-(i)n*	flex	extens.	nonfin.	NA	*	asymm.	NA	NA	*	flex-adj.	ad-S
Chantyal:chain	*-si, -rə*	NA	extens.	nonfin.	NA	*	asymm.	NA	NA	NA	flex-adj.	ad-S

	Amele:but	Amele:chain	Amele:or	Amele:cond	Greek:p.c.	Belh.:and	Belh.:chain	Belh.:incons	Belh.:sub	Burush.:chain
Amele:chain	0.62									
Amele:or	0.25	0.62								
Amele:cond	0.62	0.88	0.50							
Greek:part.coni.	0.62	0.38	0.75	0.50						
Belhare:and	0.38	0.88	0.25	0.50	0.75					
Belhare:chain	0.50	0.75	0.38	0.44	0.62	0.30				
Belhare:inconseq	0.62	0.50	0.62	0.56	0.38	0.50	0.36			
Belhare:sub	0.38	0.75	0.50	0.22	0.38	0.50	0.45	0.55		
Burushaski:chain	0.71	0.43	0.71	0.71	0.29	0.38	0.38	0.25	0.75	
Chantyal:chain	0.67	0.33	0.67	0.67	0.17	0.50	0.33	0.00	0.67	0

Mining clause linkage data: measuring distances

Language	Marker	ILL-scope	T-scope	Finiteness	ILL-mark	T-mark	Symmetry	WH	Extraction	FOC	Position	Layer
Amele:but	*gba*	local	local	fin.	*	✔	symm.	NA	NA	NA	fix.:pre	ad-S
Amele:chain	*-me, -ʔV*	conj.	conj.	nonfin.	*	*	asymm.	NA	NA	NA	fix.:pre	ad-S
Amele:or	*fo ~o*	conj.	local	fin.	✔	✔	symm.	NA	NA	NA	fix.:pre	ad-S
Amele:cond	*fi*	disj.	extens.	any	✔	✔	flex	✔	NA	NA	flex-rel.	ad-S
Greek:p.c.	*-men/-nt*	disj.	extens.	nonfin.	*	*	asymm.	NA	NA	NA	flex-rel.	ad-S
Belhare:and	*=cha...=cha*	flex	local	fin.	✔	✔	symm.	NA	*	*	flex-adj.	ad-S
Belhare:chain	*ki(na)(huŋ)*	flex	extens.	fin.	harm.	✔	asymm.	✔	possible	✔	flex-adj.	ad-S
Belhare:inc	*-kone*	local	extens.	nonfin.	✔	*	asymm.	✔	*	✔	flex-adj.	ad-S
Belhare:sub	*-naa...*	disj.	extens.	fin.	*	✔	flex	✔	*	✔	flex-rel.	ad-S
Burúshaski:chain	*n(V)-ˊ-Σ-(i)n*	flex	extens.	nonfin.	NA	*	asymm.	NA	NA	*	flex-adj.	ad-S
Chantyal:chain	*-si, -rə*	NA	extens.	nonfin.	NA	*	asymm.	NA	NA	NA	flex-adj.	ad-S

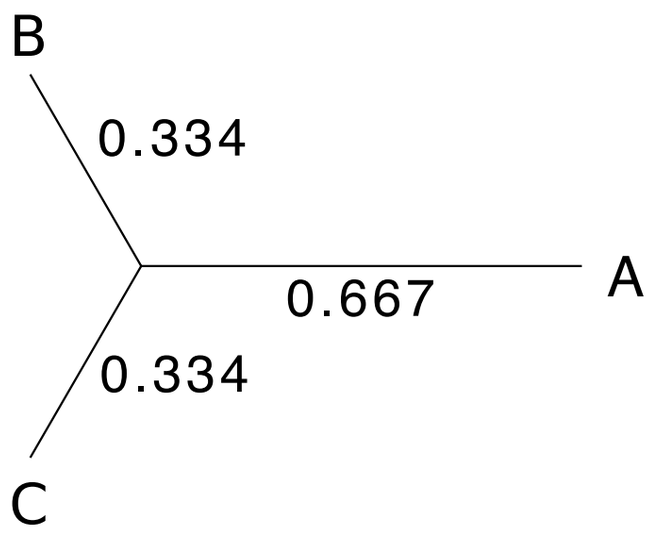
Two issues

1. NAs reduce comparability (but perhaps we can extrapolate).
2. Variables with language-specific content (e.g. “Marker”, “Language”) are not interesting for this, they just increase the baseline distance

Mining clause linkage data: aggregating distances

3. With 69 constructions we get $\frac{69 \cdot (69-1)}{2} = 2,346$ pairs!

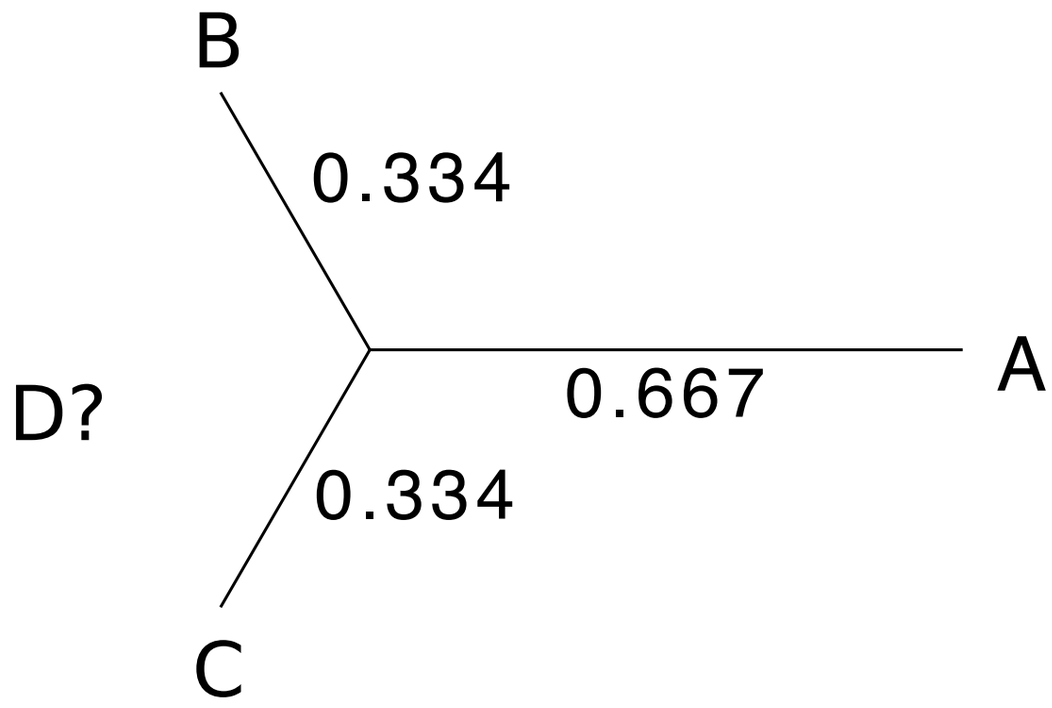
- How can we detect general patterns about what's more similar to each other (e.g., having "low average distances")?
- Aggregate and plot it!
- But there is a geometrical problem:
  - DIST(A,B)=1, DIST(A,C)=1, DIST(B,C)=2/3



Mining clause linkage data: aggregating distances

- now add D: dist(A,D)=1, dist(B,D)=1/3, dist(C,D)=1/2

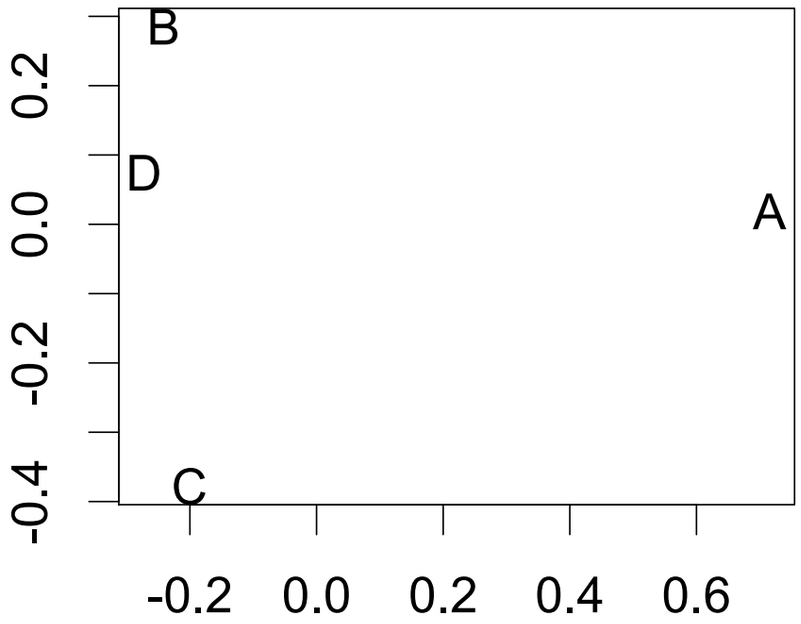
	A	B	C
B	1.000		
C	1.000	0.667	
D	1.000	0.333	0.500



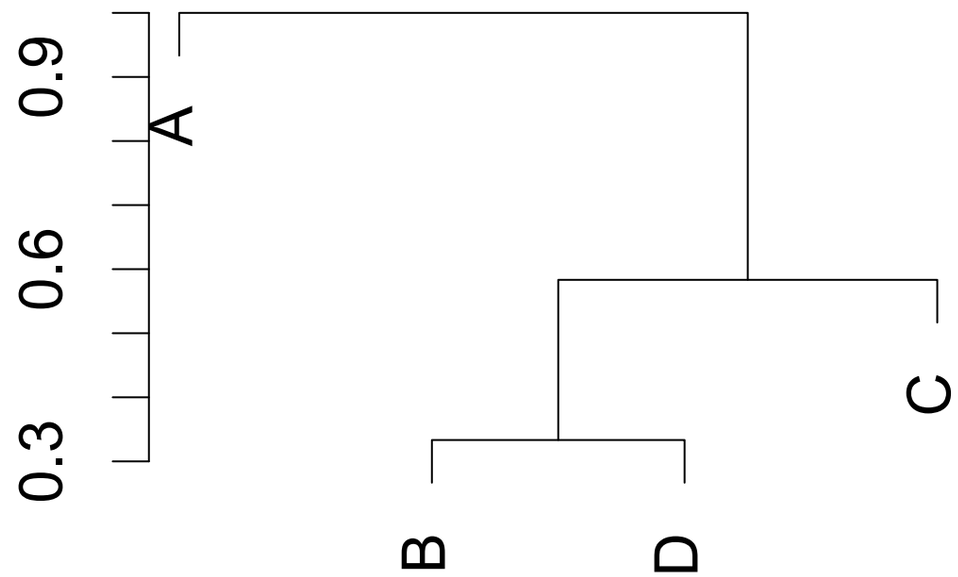
Mining clause linkage data: aggregating distances

	A	B	C
B	1.000		
C	1.000	0.667	
D	1.000	0.333	0.500

Multidimensional Scaling: find (algorithmically) lowest dimensionality with least distortion

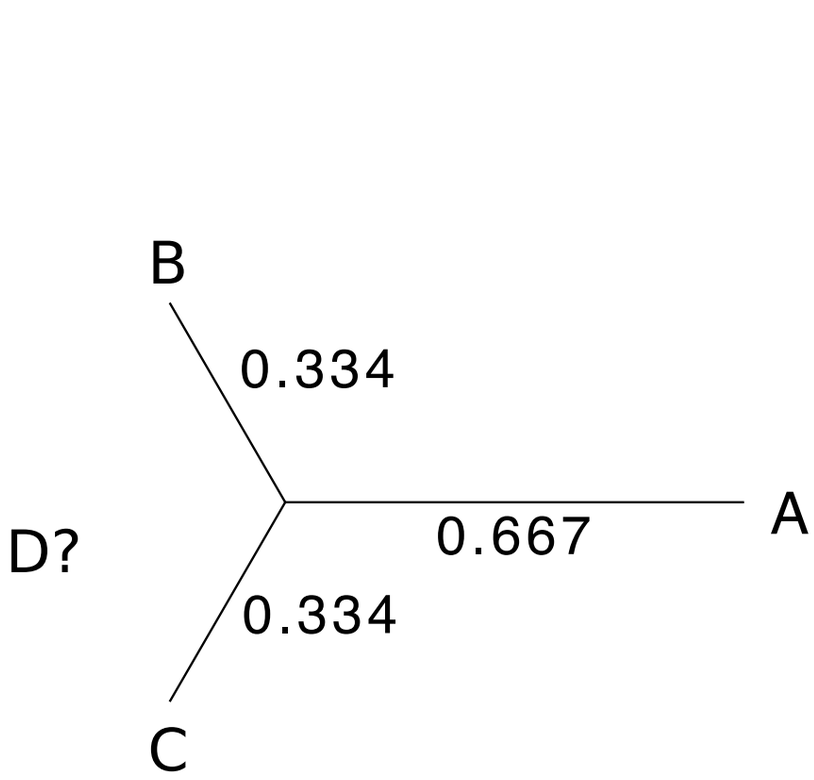


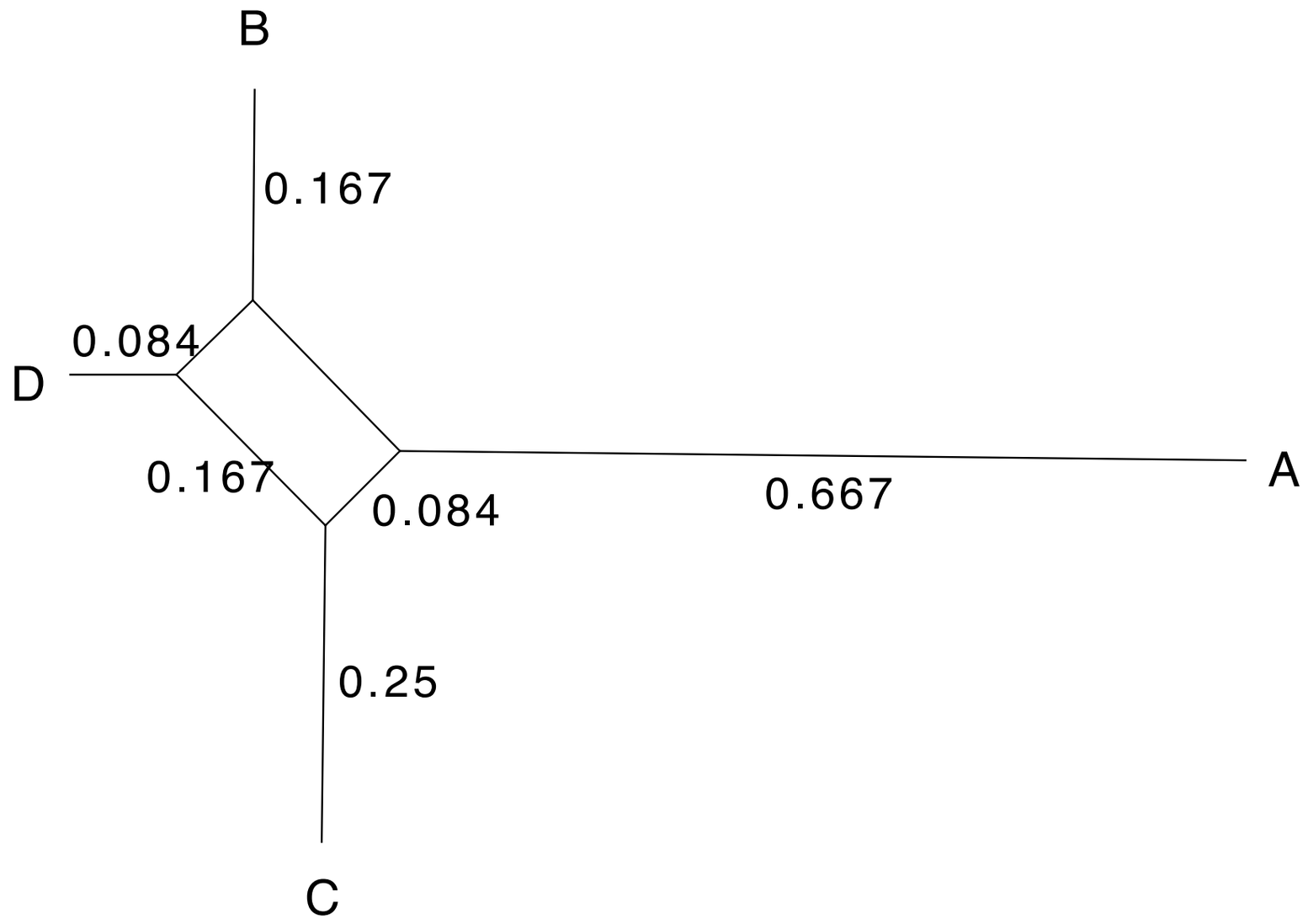
Hierarchical Clustering: stepwise grouping, minimizing distances between groups (e.g. via averages)



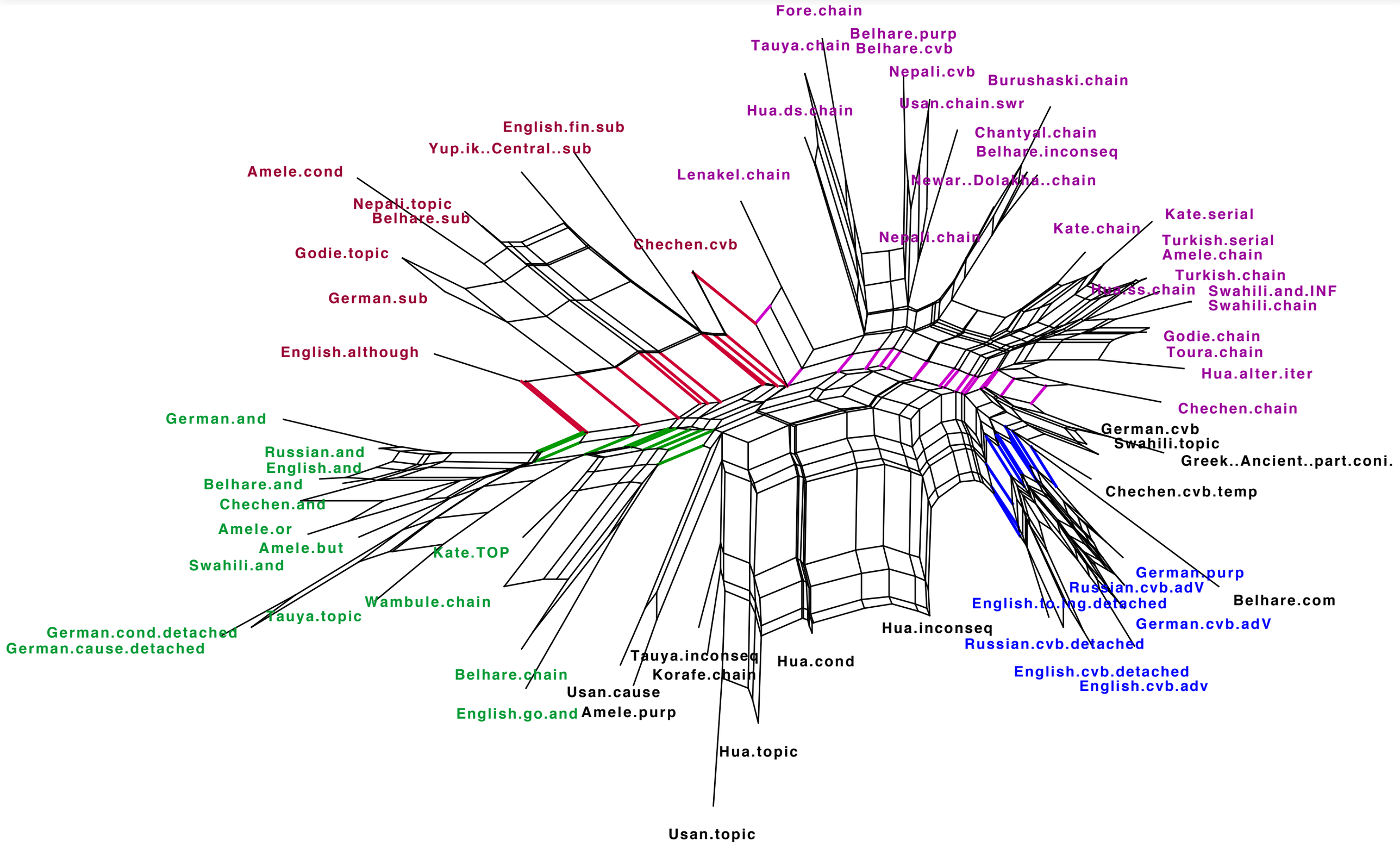
Mining clause linkage data: aggregating distances

- A more informative alternative: **split graph** methods, e.g. NeighborNet:





Mining clause linkage data: type clusters (“prototypes”)



Neighbor-Net of clause linkage types (69 constructions, 24 languages, 11 variables)

Mining clause linkage data: associations

- Many concepts of traditional metalanguages are absolute universals in disguise
- E.g., if S stands for ‘sentence (verb plus all arguments)’ and [S_1 S_2] represents subordination, then (as for example in RRG): **[[S_1] S_2]=ILL**

therefore,

(1) ILL scope is not shared (disjunct)

(2) WH is banned

- (1) and (2) define the structure (“are diagnostic of it”), and
- the structure “predicts” (1) and (2), and (1) ↔ (2)

Mining clause linkage data: associations

- Problem: Perhaps. Perhaps not!
- Unless the structure is “given” (descriptive A priori = ontological A priori!), we need empirical evidence.
- But rather than relying on a single datapoint, we want to estimate the degree to which the structure is supported from many datapoints.
  - We know that single datapoints are prone to error!
  - We will never know the truth, but we can try and separate signal from noise, statistically.
- Quantitative, rather than qualitative thinking.

Mining clause linkage data: associations

- Entropy of X with estimated value probabilities $p_{x1}...p_{xk}$:

$$H(x) = -\sum_{i=1}^{k} p(x_i) \cdot \log(p(x_i))$$

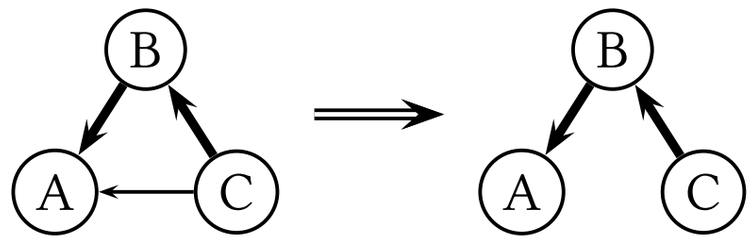
- Joint entropy: $\mathrm{H}(\mathrm{X}, \mathrm{Y}) = -\sum \mathrm{p}_{\mathrm{x_i},\mathrm{y_i}} \log(\mathrm{p}_{\mathrm{x_i},\mathrm{y_i}})$

- Mutual information: $\mathrm{I}(\mathrm{X}; \mathrm{Y}) = \mathrm{H}(\mathrm{X}) + \mathrm{H}(\mathrm{Y}) - \mathrm{H}(\mathrm{X}, \mathrm{Y})$

- Predictability of X given Y ('Y → X'): $\pi(\mathrm{X}|\mathrm{Y}) = \dfrac{\mathrm{I}(\mathrm{X}; \mathrm{Y})}{\mathrm{H}(\mathrm{X})}$

Mining clause linkage data: associations

- Estimate predictabilities between all pairs of variables
- Reduce triplets via weakest link deletion:



- Examine those with non-zero predictability:

Mining clause linkage data: associations

Implication	$\hat{\pi}(X\|Y)$	$\hat{\pi}(Y\|X)$
ILL-scope $\leftrightarrow$ ILL-mark	.09	.05
ILL-scope $\rightarrow$ WH	.18	.08
ILL-scope $\rightarrow$ FOC	.22	.06
ILL-scope $\leftrightarrow$ position	.23	.20
T-scope $\leftrightarrow$ T-mark	.29	.26
T-scope $\leftrightarrow$ symmetry	.21	.17
T-scope $\rightarrow$ FOC	.11	.05
T-scope $\leftrightarrow$ layer	.09	.06
ILL-mark $\rightarrow$ extraction	.22	.12
T-mark $\leftrightarrow$ finiteness	.31	.28
symmetry $\leftrightarrow$ finiteness	.28	.28
symmetry $\leftrightarrow$ ILL-mark	.13	.12
position $\rightarrow$ layer	.12	.06

But which values of these variables (features of these attributes) are associated?

Mining clause linkage data: associations

- Pearson residuals:

	[,1]	[,2]	sum
[1,]	5	0	5
[2,]	5	10	15
sum	10	10	20

- Expected rows: 50% each:

	[,1]	[,2]	sum
[1,]	2.5	2.5	5
[2,]	7.5	7.5	15
sum	10	10	20

- Absolute residual $(O\text{-}E)^2 = (5\text{-}2.5)^2$ and $(10\text{-}7.5)^2$
- Relative residual:

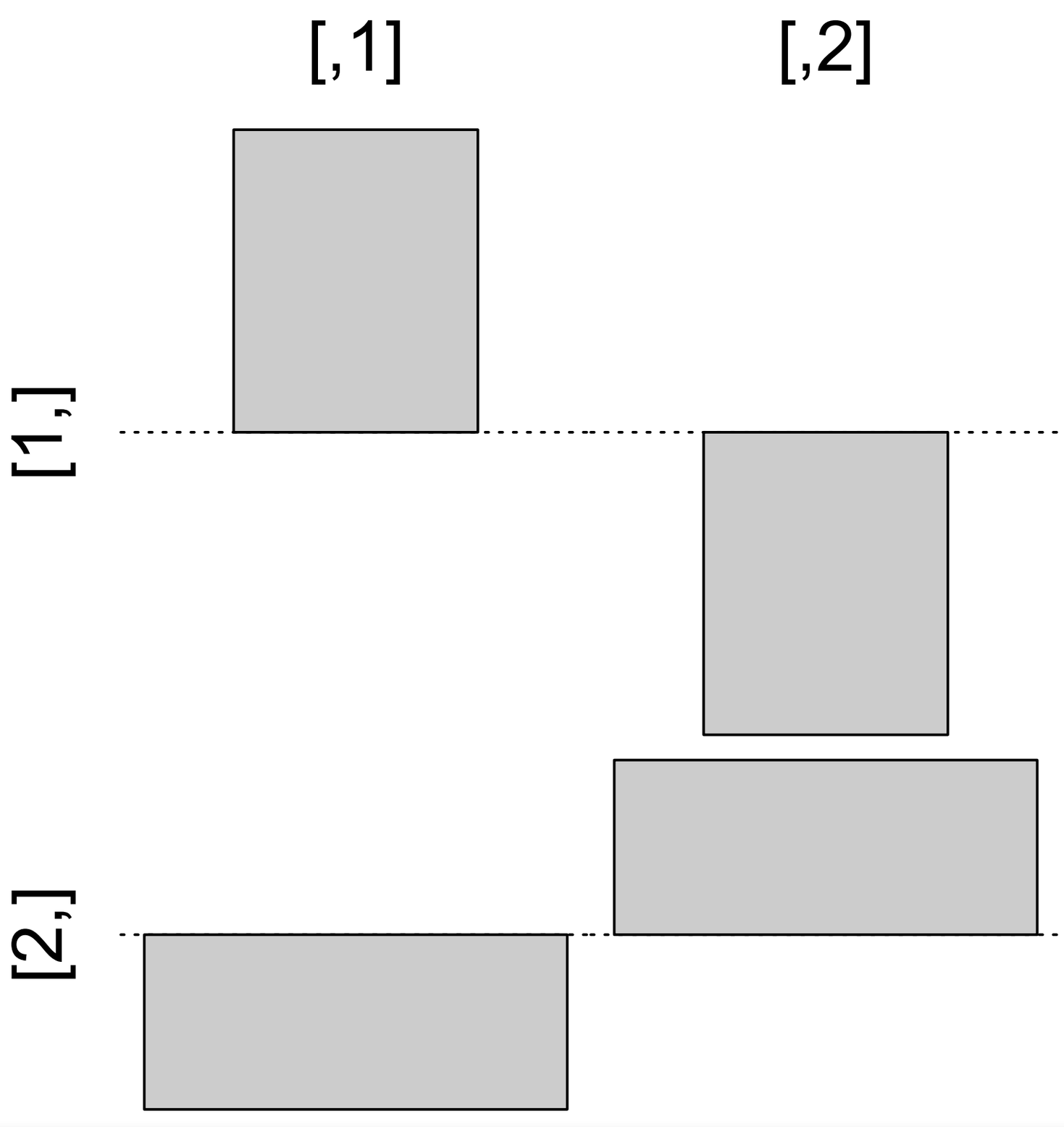
$$R = \frac{(O-E)^2}{E} = \frac{O-E}{\sqrt{E}} = \frac{n-\hat{\mu}}{\sqrt{\hat{\mu}}}$$

	[,1]	[,2]
[1,]	1.58	-1.58
[2,]	-0.91	0.91

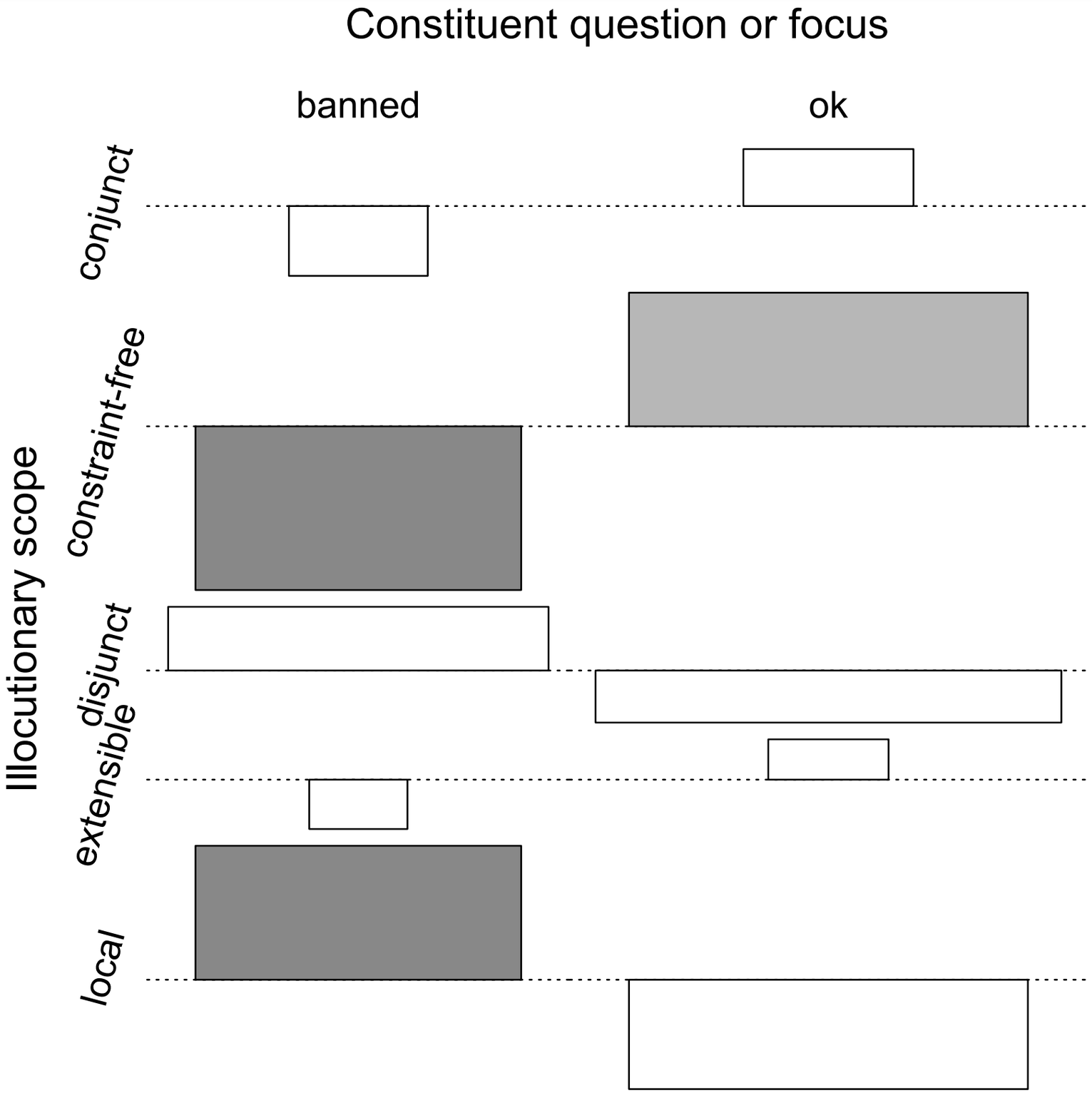
Mining clause linkage data: associations

- Association plots:

	[,1]	[,2]
[1,]	1.58	-1.58
[2,]	-0.91	0.91



Mining clause linkage data: associations

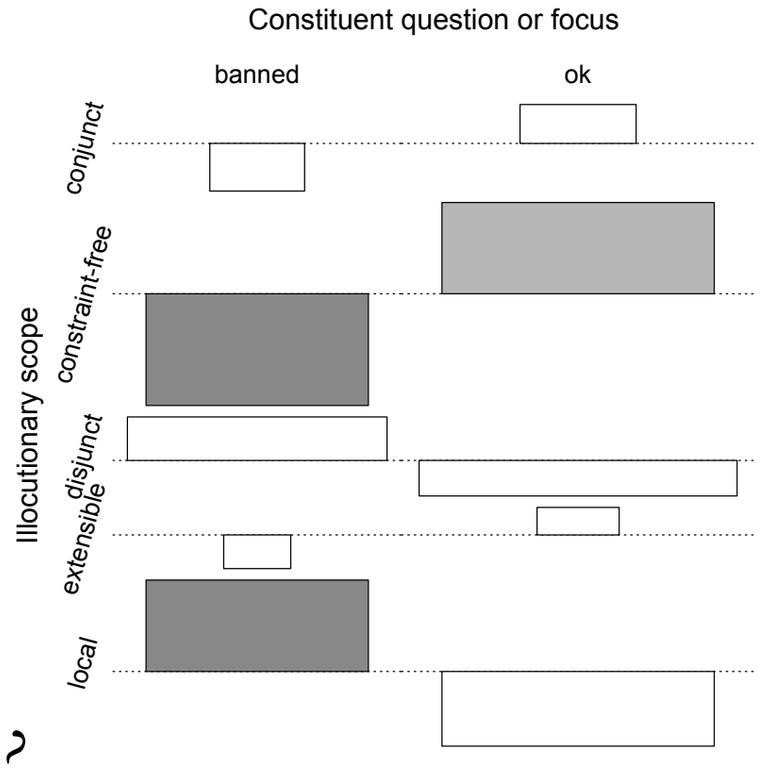


Association plot, showing Pearson residuals

Examples

- Significantly more frequent than expected under H_0:

  No scope constraint, WH ok



Belhare (Bickel 1993)

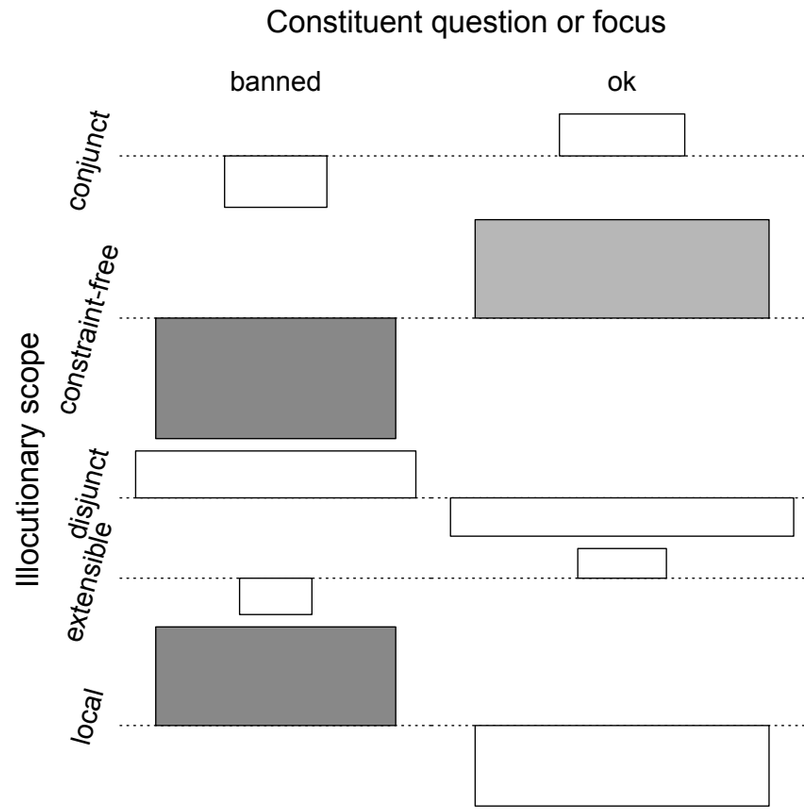
a. *khar-e ki jutta ŋŋ-in-ghutt-he-ga i?*
[3sS]go-PST SEQ shoes[NOM] 3sA-buy-bring.for-PST-2sP Q
'Did she go [there] and buy you shoes?'
or 'Did she buy you shoes when she went [there]?' (presupposing either 'she went' or 'she bought')

b. *laitar hene lept-he-ga ki sʌlai am-t-u-ga?*
lighter where throw-PST-2sA SEQ matches light-NPST-3sP-2sA
'Where did you throw the lighter so that you have to use matches?'

Examples

- Significantly more frequent than expected under H_0:

  Local scope, no WH



Tauya (MacDonald 1990)

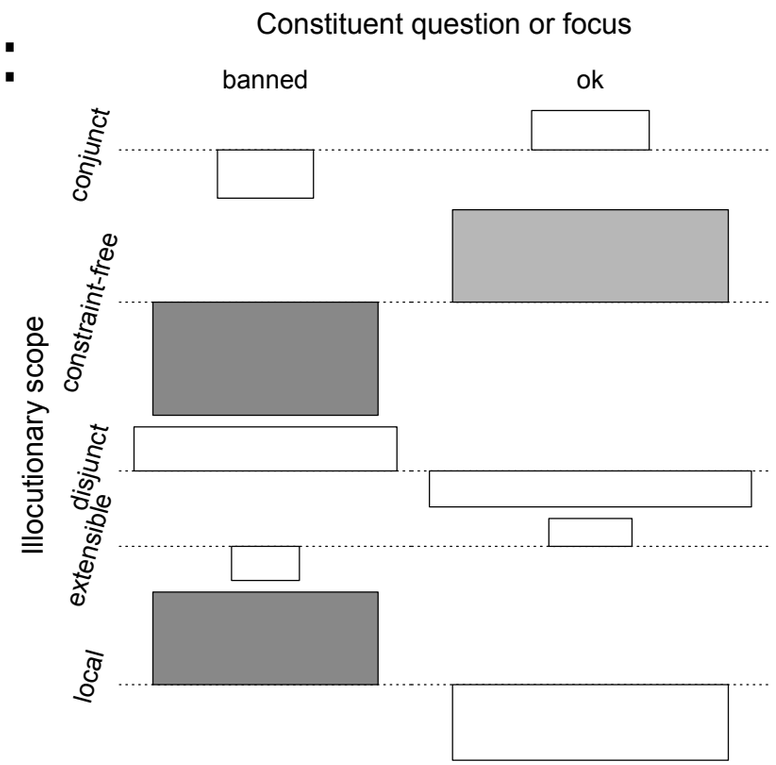
a. *nen mei momune-i-nani=ra pofei-ti nen=tu-e=nae?*
   3p here sit-3p-ASS=TOP talk-CONJ 3p=give-2=POLAR.Q
   'They sat here and/but did you talk to them?'
   or 'Since they sat here, did you talk to them?'

b. * *we mei fofe-a-nani=ra …*
   who[NOM] here come-3s-ASS=TOP
   Intended: '*if who did come here…'

Examples

- A bit more frequent than expected under H_0:

  Disjunct scope, no WH



Lakhota (Siouan, Van Valin 1995)

a. *šúka ki táku yaxtáka he?*
   dog DET INDEF/WH [3s>3s-]bite Q
   'What did the dog bite?' *or:* 'Did the dog bite something?'

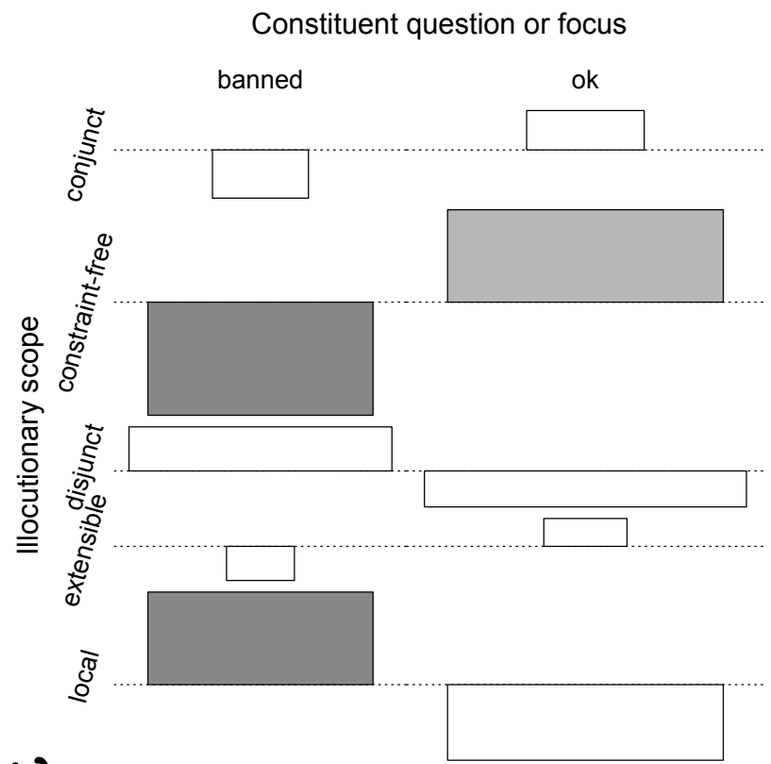
b. *wičháša ki táku yúte ečhúhą, tha-wíču ki mní ikíčiču he?*
   man DET eat INDEF/WH while 3sPOSS-wife DET water get.for Q
   'While the man was eating something, did his wife get water for him?'
   not: '*What did his wife get him water, while the man was eating?'

and many languages with WH extraction (English among them)

Examples

- A bit less frequent than expected under H_0:

  Disjunct scope, WH ok



Belhare

a. *ne-e yuŋ-a=naa mundhupt-he i?*
DEM-LOC [3sS]sit-SBJV.PST=TOP [3sS]chat-PST Q
'When he was here, did he say something?' (or was he silent?)
or 'Did he say something when he was here?' (or later only?)
but not 'Was he here, and did he say something?'

b. *sa-a ya=m-phekt-a-k=naa ŋ-khatd-at-ni-gak=phe?*
who-[s]ERG call=3nsA-call-SBJV.PST-2=TOP NEG-go-PST-NEG-2=IRR
'You hadn't gone if who had called you?'

Summary

- Recall that this was an exploratory study: large-scale multivariate typologies currently under development
  - clause linkage
  - grammatical relations

- Once the data exist and are mined, each associations needs **theoretical interpretation and modeling** so that it can be tested for universal validity

→ the topic of the next part

Compare to traditional approach

- Instead of empirical correlations, universal notions defining bundles of properties:
  - cosubordination: {conjunct ILL scope, WH allowed}
  - subordination: {disjunct or local ILL scope, WH banned}
- These are in fact absolute, exceptionless universals:
  - conjunct ↔ WH ok
  - disjunct/local ↔ WH banned

Compare to traditional approach

- Exceptions (as in Belhare) need to be “explained away”
- Or the absolute univeral needs to be limited in its predictive value, e.g.
    - subordination = structure with disjunct or local ILL-scope.
- But then, no argumentation available of the kind:

“since X is subordinate, WH is not allowed” or “since X is subordinate, there is an intervening S node, and we cannot move WH”

Compare to traditional approach

- This is why in traditional approaches, metalanguages have a strong interest in maintaining absolute universals
- In traditional approaches, there is no way of understanding relations between structures or generalizations across structures without absolute universals
    - Another example: branching direction (head parameter)
- In MULTIVARIATE TYPOLOGY, by contrast, all generalizations are probabilistic, and therefore have expections
    - Instead of “if X bans disjunct scope, it must ban WH formation”, we get “... it is likely to ...”, i.e. statistical instead of absolute universals

Overall summary

- Absolute universals and variables are analytical tools, more or less successful in measuring variation and detecting distributional patterns, but with no claim on any deeper reality (i.e. Descriptive A priori ≠ Ontological A priori)
- Received absolute universals typically need decomposition into multiple variables allowing for much more variation
- Multivariate Typology allows analytical coverage of phenomena ‘in-between’ and captures similarities without forcing structures into pre-conceived universal slots
  - Instead of asking “is X subordinate?”, we better ask questions like “does X enforce disjunct scope?”, “does X allow WH?” etc.
- Hundreds of variables, thousands of levels = the true diversity!

Overall summary

- These are the right questions for fieldwork.
- And, at the same time, these are the right questions for typology
- Answers lead to tables which we can mine for distributional patterns
- And so, the object of inquiry is these patterns, and the goal is to explain them.

  (Contrast this with Anderson's definition of the object and goal of linguistics: *"The object of inquiry in linguistics is the human ability to acquire and use a natural language, and the goal of linguistic theory is an explicit characterization of that ability."*)

- Interestingly, more in line with other disciplines dealing with distributions (e.g. genetics, ecology, economics).