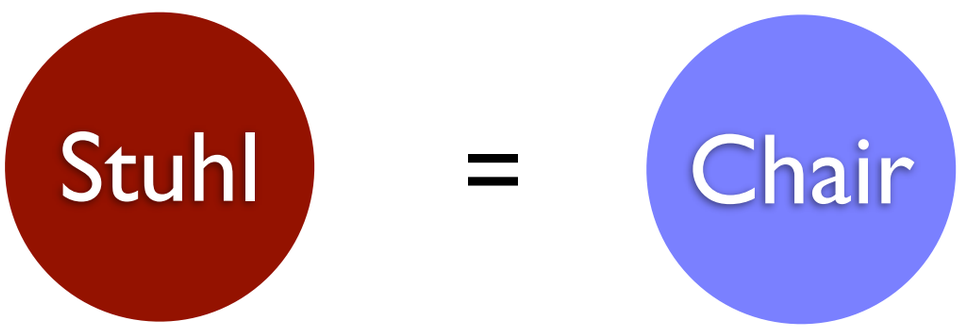
On the principle problem of comparing languages

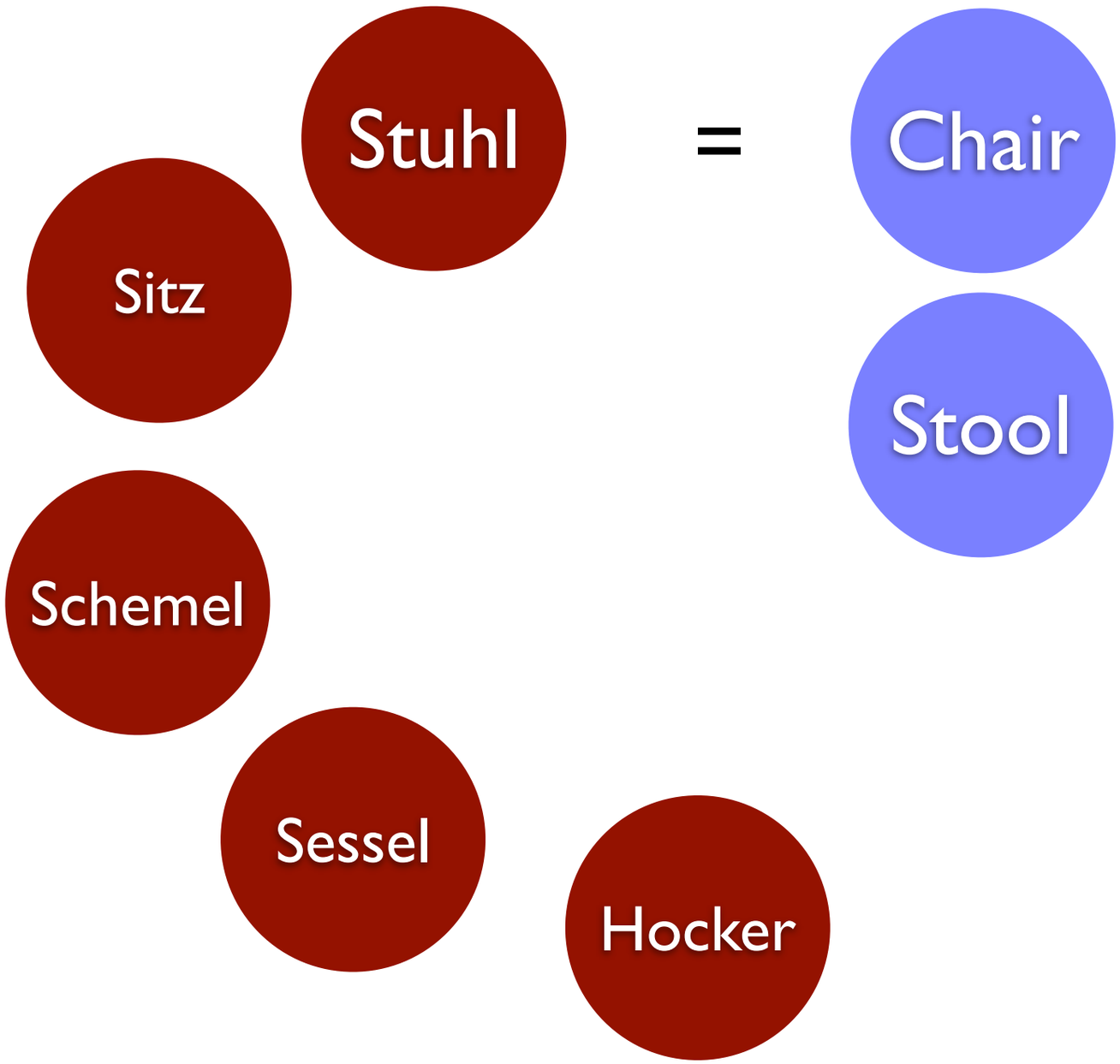
Michael Cysouw
MPI-EVA Leipzig

Stuhl


Stuhl
=
Chair

Stuhl = Chair

Stool


Stuhl
=
Chair
Sitz
Stool
Schemel
Sessel
Hocker

Vorsitz

Präsidium

Lehrstuhl

Stuhl

=

Chair

Sitz

Stool

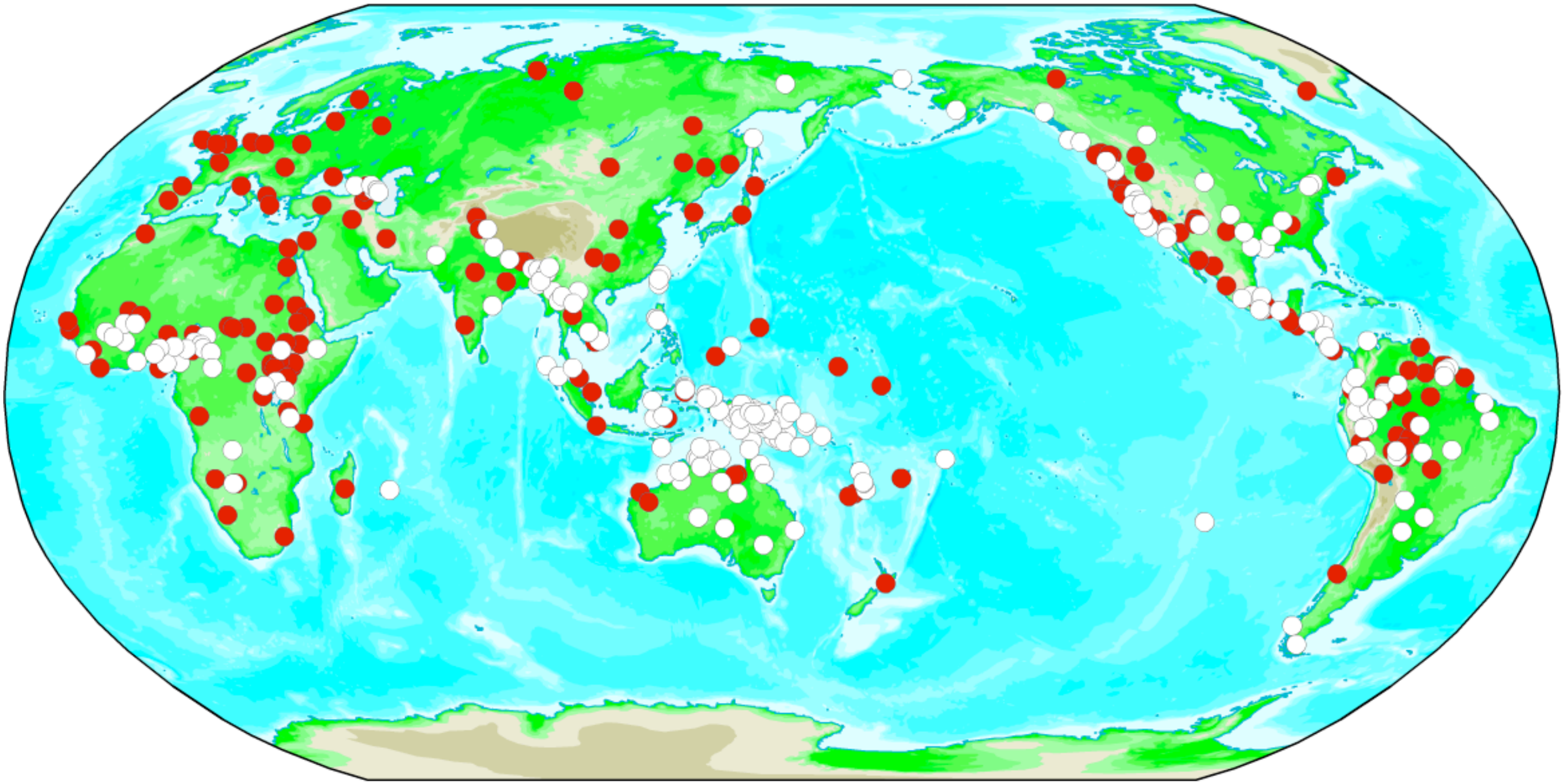
Schemel

Sessel

Hocker

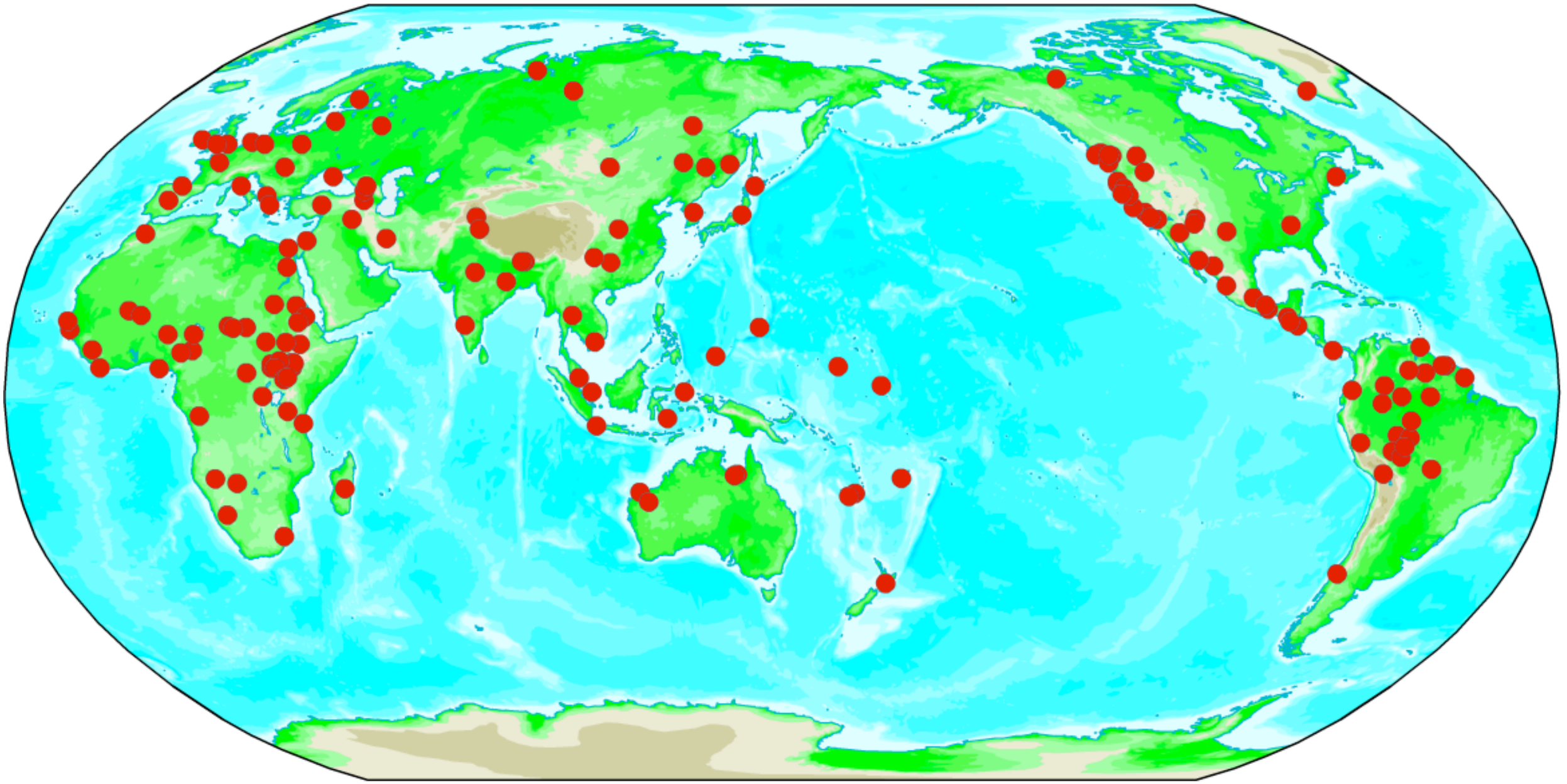
Passive Constructions

Passive Constructions



Siewierska, Anna. "Passive Constructions." World Atlas of Language Structures. Eds. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie. Oxford: Oxford University Press, 2005. 434-37.

Passive Constructions



Siewierska, Anna. "Passive Constructions." World Atlas of Language Structures. Eds. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie. Oxford: Oxford University Press, 2005. 434-37.

Radical Relativism

Radical Relativism

- Constructions (including lexicon) are always language-specific

Radical Relativism

- Constructions (including lexicon) are always language-specific
- In principle, each construction in each language should be uniquely named

Radical Relativism

- Constructions (including lexicon) are always language-specific
- In principle, each construction in each language should be uniquely named
- In practice, the same names are used again and again for reasons of readability

Radical Relativism

- Constructions (including lexicon) are always language-specific
- In principle, each construction in each language should be uniquely named
- In practice, the same names are used again and again for reasons of readability
- This is currently confusing most readers (and authors!)

2.1.1. Emic independent clause classes

		Tr	Intr	Eq	Quot	
					Tr	Intr
		10	20	30	40	50
Decl	01	11	21	31	41	51
Ex	02	12	22	32	42	52
Q	03	13	23		43	53
Q	04	14	24	34	44	54
Impv	05	15	25		45	55
Impv	06	16	26		46	56

Chart II. Emic independent clause classes

2.1.2. Tagmemic independent clause formula

Cl = {±Mar: Cl mar +Nuc: Cl nuc 10-50}

The independent clause is subdivided into emic distribution classes 11-56 on the basis of the filler of the nucleus slot and of the distribution. Since the distribution classes do not otherwise differ in composition, they are not shown in separate formulas.

2.1.3. Independent clause citation

Ind cl = noy teč in neríyow--noropików to neč ka?ánoneb
'there that water they-drink-where--they-just-now-come the those animals' (those animals were just now going there where they drink that water).

kopi íkomórikon to mónči 'why-you killer the child' (why did you kill the child?).

ne soratíye--ónka to ka kí?inon 'here town-in--not the that-which caring-person' (here in the town there is no one who cares).

nihín--ónka ímokon 'my-daughter--not sleeper' (my daughter, don't sleep).

How to compare unique constructions across languages?

How to compare unique constructions across languages?

- Similarity Semantics: no identity, only similarity

How to compare unique constructions across languages?

- Similarity Semantics: no identity, only similarity
- Meaning is a continuous space, without universal semantic meta-language

How to compare unique constructions across languages?

- Similarity Semantics: no identity, only similarity
- Meaning is a continuous space, without universal semantic meta-language
- The meaning-space can be sampled

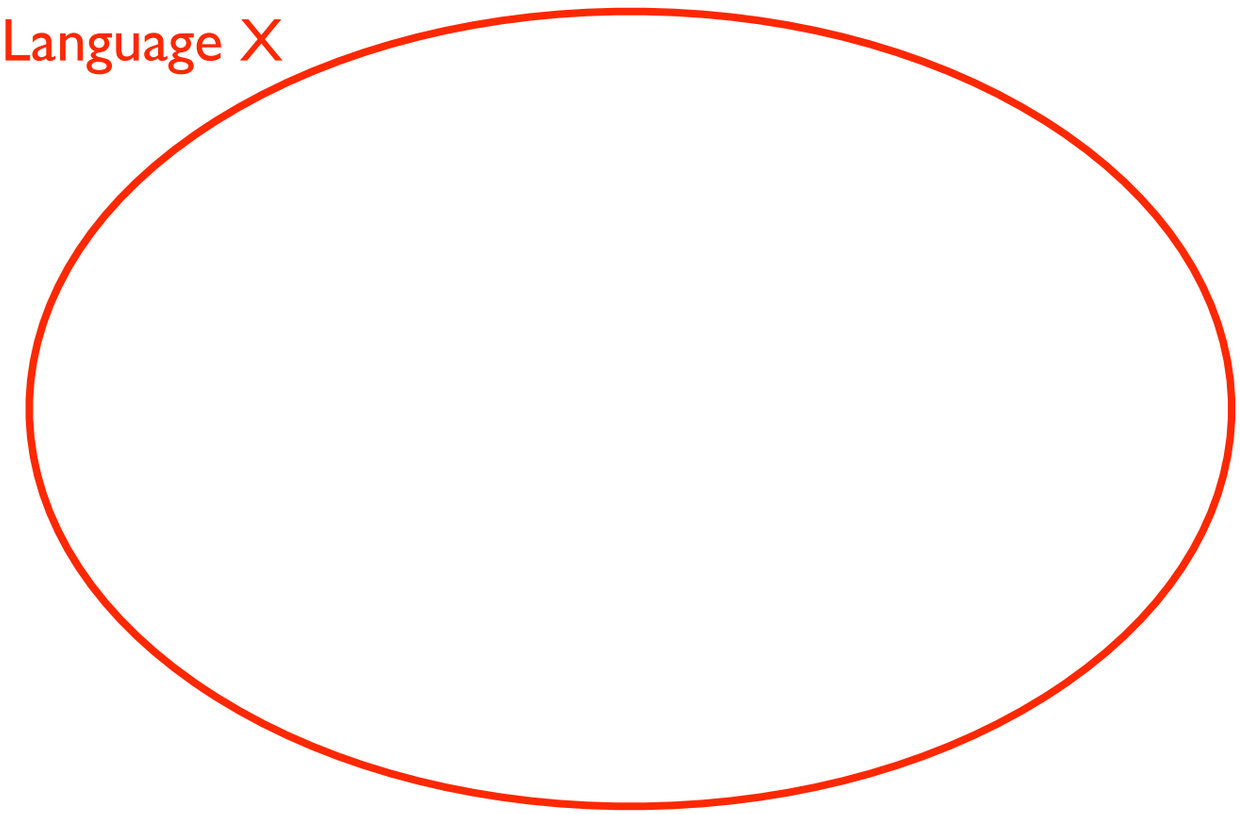
How to compare unique constructions across languages?

- Similarity Semantics: no identity, only similarity
- Meaning is a continuous space, without universal semantic meta-language
- The meaning-space can be sampled
- Similarity of constructions can be established based on this sample

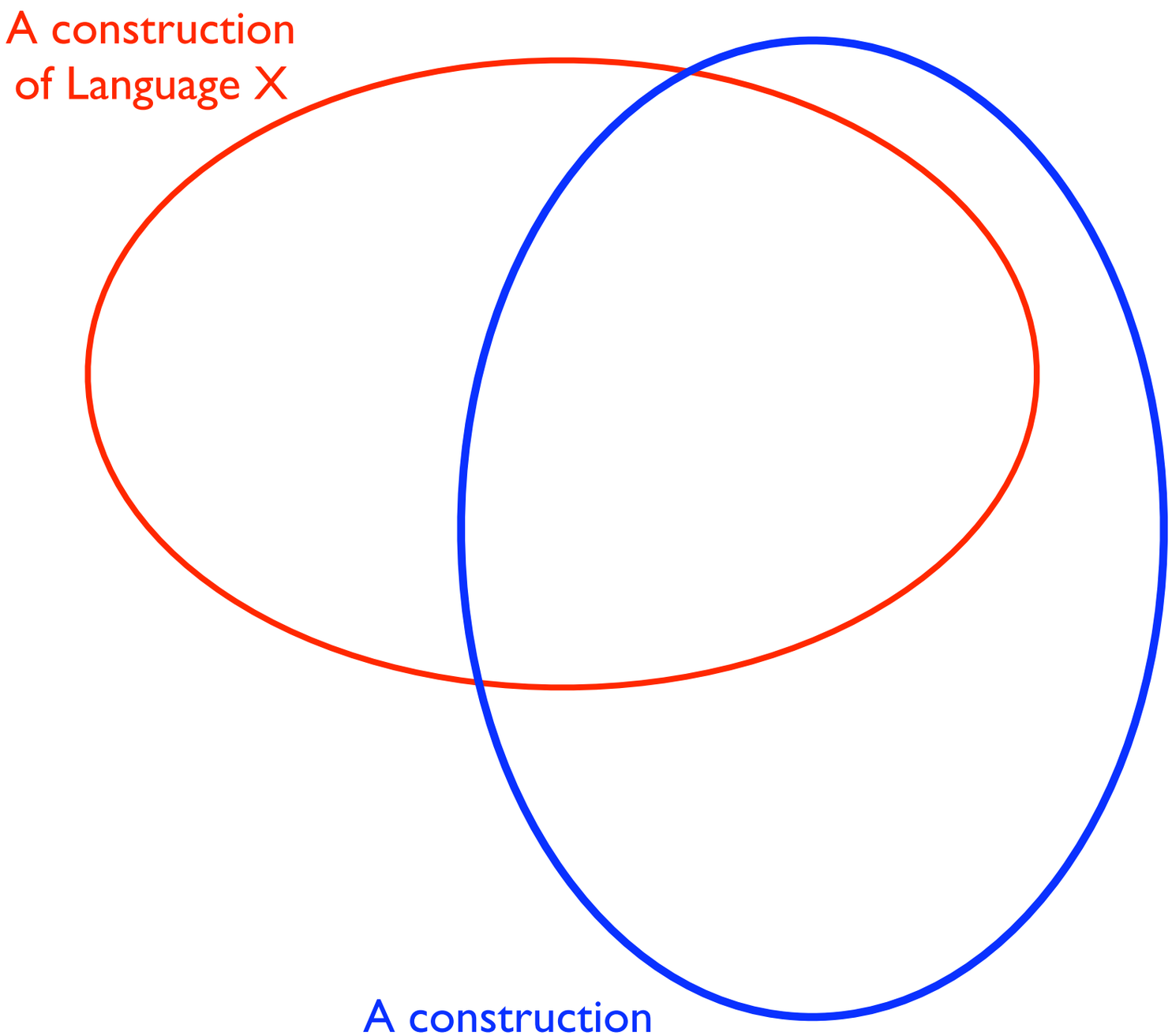
Meaning-space

Meaning-space

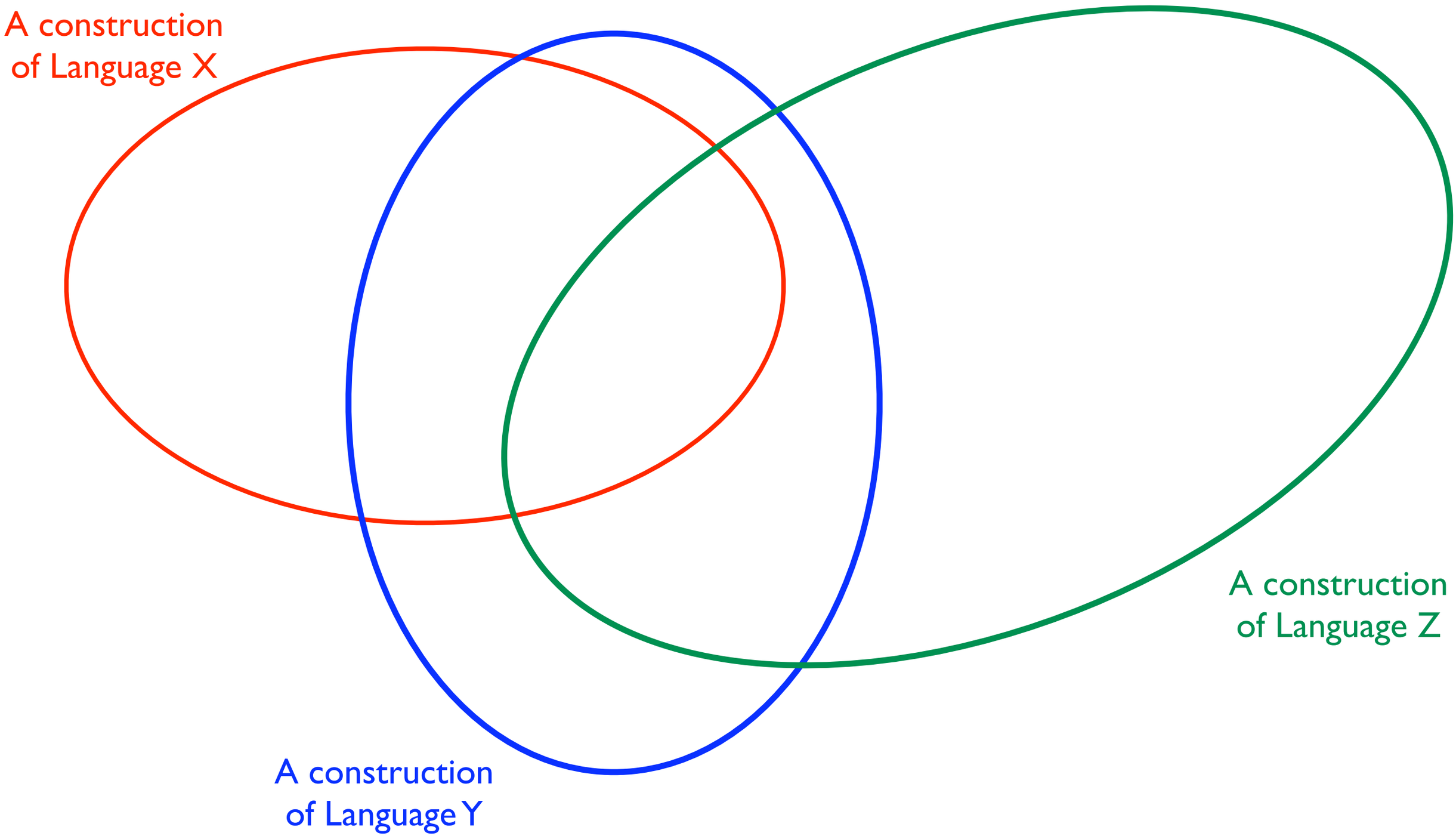
A construction
of Language X



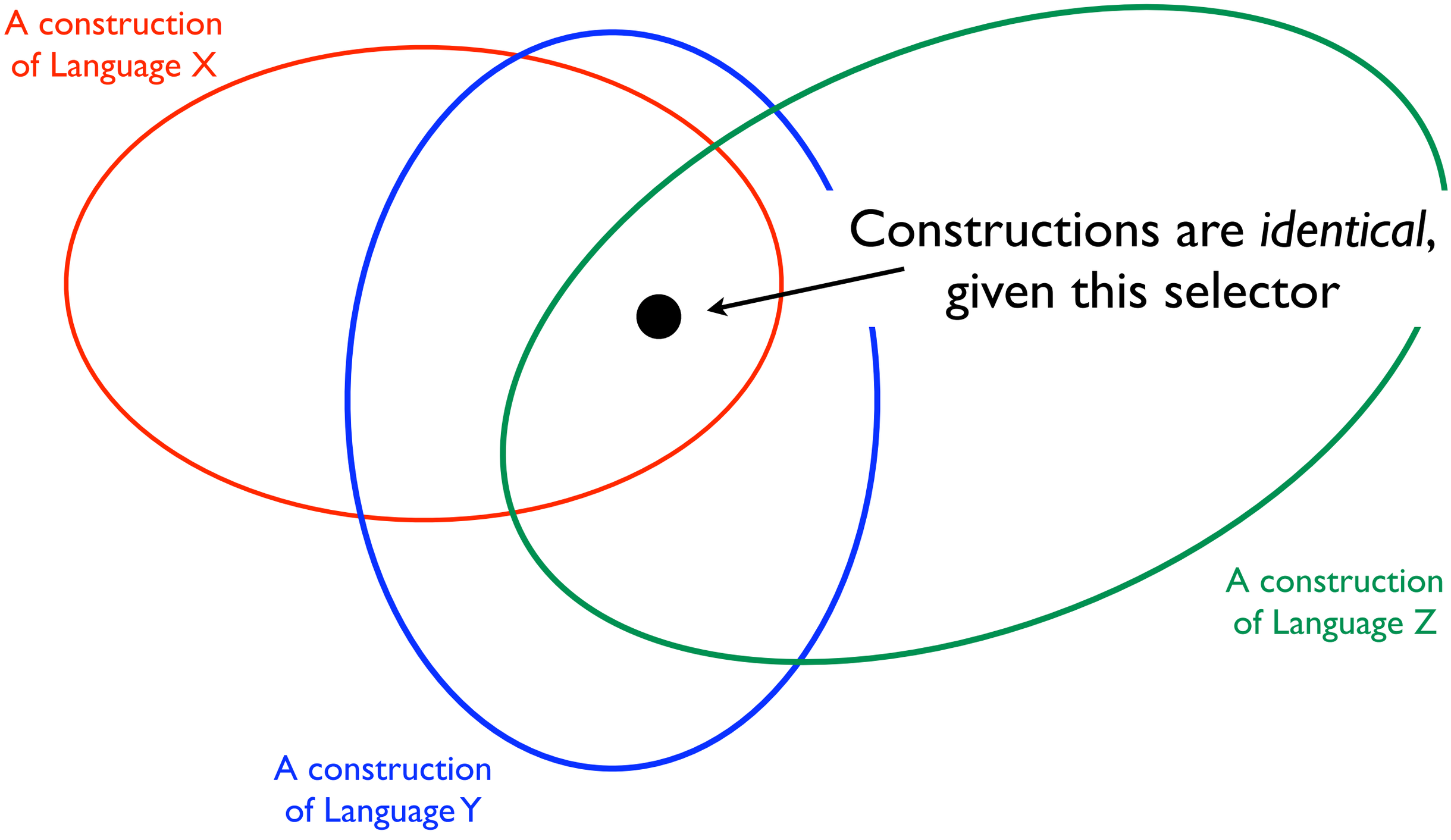
Meaning-space



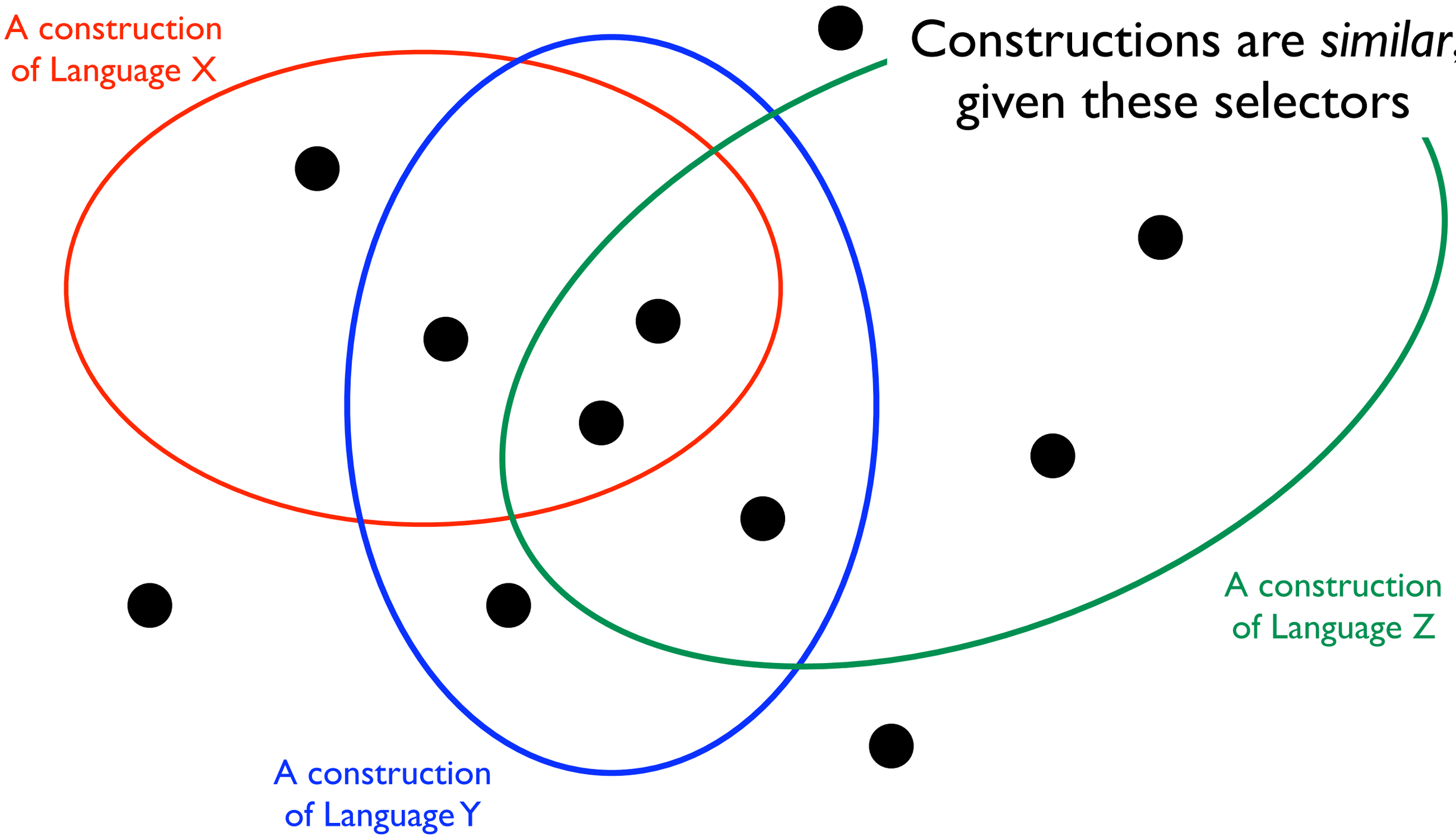
Meaning-space



Meaning-space



Meaning-space



Selectors

Selectors

- The meaning-space can be sampled by choosing (arbitrary) 'selectors'
  - items in parallel texts
  - pictures, videos, translational questionnaires
  - (more abstract) functions
  - formal restrictions

Selectors

- The meaning-space can be sampled by choosing (arbitrary) 'selectors'
  - items in parallel texts
  - pictures, videos, translational questionnaires
  - (more abstract) functions
  - formal restrictions
- The selectors 'select' the relevant constructions for the comparison

Selectors

- The meaning-space can be sampled by choosing (arbitrary) 'selectors'
  - items in parallel texts
  - pictures, videos, translational questionnaires
  - (more abstract) functions
  - formal restrictions
- The selectors 'select' the relevant constructions for the comparison
- Language comparison has to sum up over all relevant constructions

Verbs of motion in parallel texts
(data from Bernhard Wälchli)

Verbs of motion in parallel texts
(data from Bernhard Wälchli)

Selectors	Acholi	Ainu	Albanian	Armenian	Avar	Aymara	Bambara	Basque
1.05.0	bino	paye	shkoj	erTam	baC'ine	mistu	taa	etorri
1.07.0	bino	ek	vij	gam	baC'ine	juta	na	etorri
1.09.0	bino	ek	vij	gam	baC'ine	juta	na	etorri
1.10.4	a	yan	dal	veranam	Lug'ine	mistu	bO	atera
1.10.5	pye	ran	zbres	iJanem	reSt'ine	jala	jigin	jaitsi
1.11.4	a	–	–	linim	bag'ize	juta	bOra	–
1.12.0	ryamo	omande	shtrEngoj	hanem	ine	nuct'i	ye	bultzatu
1.14.0	bino	ek	shkoj	gam	ine	juta	taa	etorri
1.16.0	kato	shirikush	kaloj	anCanem	ine	sara	taama	joan
1.17.0	bino	araki	vij	gam	biLLine	juta	tugu	etorri
1.18.4	lobo	paye	shkoj	gnam	ine	arka	tugu	jarrai
1.20.0	cito	paye	shkoj	gnam	ine	arka	tugu	abiatu
1.21.0	donyo	paye	arrij	mtanem	Sweze	ma	taa	sar
1.21.1	donyo	ahun	hyj	mtanem	baC'ine	ma	don	joan

Verbs of motion in parallel texts
(data from Bernhard Wälchli)

Selectors	Acholi	Ainu	Albanian	Armenian	Avar	Aymara	Bambara	Basque
1.05.0	bino	paye	shkoj	erTam	baC'ine	mistu	taa	etorri
1.07.0	bino	ek	vij	gam	baC'ine	juta	na	etorri
1.09.0	bino	ek	vij	gam	baC'ine	juta	na	etorri
1.10.4	a	yan	dal	veranam	Lug'ine	mistu	bO	atera
1.10.5	pye	ran	zbres	iJanem	reSt'ine	jala	jigin	jaitsi
1.11.4	a	–	–	linim	bag'ize	juta	bOra	–
1.12.0	ryamo	omande	shtrEngoj	hanem	ine	nuct'i	ye	bultzatu
1.14.0	bino	ek	shkoj	gam	ine	juta	taa	etorri
1.16.0	kato	shirikush	kaloj	anCanem	ine	sara	taama	joan
1.17.0	bino	araki	vij	gam	biLLine	juta	tugu	etorri
1.18.4	lobo	paye	shkoj	gnam	ine	arka	tugu	jarrai
1.20.0	cito	paye	shkoj	gnam	ine	arka	tugu	abiatu
1.21.0	donyo	paye	arrij	mtanem	Sweze	ma	taa	sar
1.21.1	donyo	ahun	hyj	mtanem	baC'ine	ma	don	joan

Verbs of motion in parallel texts
(data from Bernhard Wälchli)

Selectors	Acholi
1.05.0	bino
1.07.0	bino
1.09.0	bino
1.10.4	a
1.10.5	pye
1.11.4	a
1.12.0	ryamo
1.14.0	bino
1.16.0	kato
1.17.0	bino
1.18.4	lobo
1.20.0	cito
1.21.0	donyo
1.21.1	donyo

Verbs of motion in parallel texts
(data from Bernhard Wälchli)

Acholi	bino	a	pye	ryamo	kato	lobo	cito	donyo
1.05.0	+							
1.07.0	+							
1.09.0	+							
1.10.4		+						
1.10.5			+					
1.11.4		+						
1.12.0				+				
1.14.0	+							
1.16.0					+			
1.17.0	+							
1.18.4						+		
1.20.0							+	
1.21.0								+
1.21.1								+

Selectors	1	2	3	4	5	6
Construction 1	+	+				
Construction 2			+	+		
Construction 3					+	+

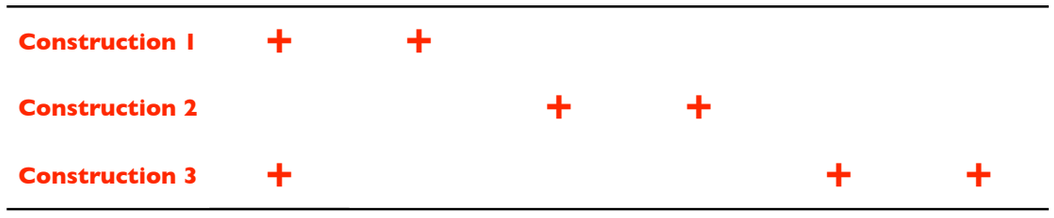
Selectors	1	2	3	4	5	6
Construction 1	+	+				
Construction 2			+	+		
Construction 3					+	+
Construction 1	+		+	+	+	
Construction 2	+					+

Selectors	1	2	3	4	5	6
Construction 1	+	+				
Construction 2			+	+		
Construction 3					+	+
Construction 1	+		+	+	+	
Construction 2		+				+

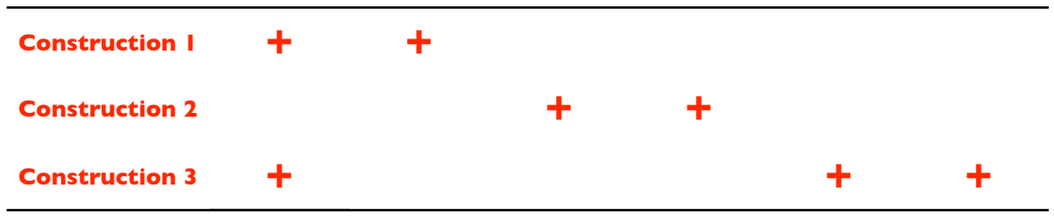
Construction 1	+					
Construction 2		+				
Construction 3			+			
Construction 4				+	+	+

Selectors	1	2	3	4	5	6
Construction 1	+	+				
Construction 2			+	+		
Construction 3	+				+	+
Construction 1	+		+	+	+	
Construction 2		+	+	+		+
Construction 1	+					
Construction 2		+			+	
Construction 3	+		+			
Construction 4	+			+	+	+

Method of Language Comparison, version A



Method of Language Comparison, version A



	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

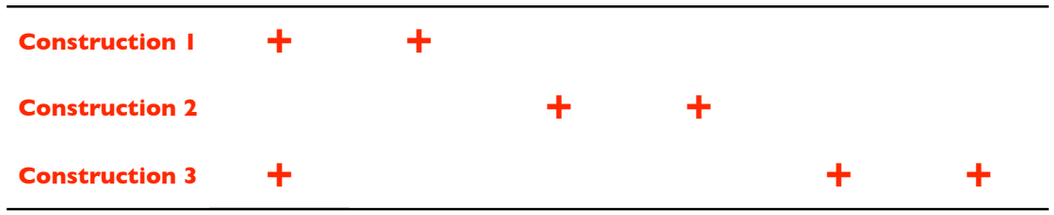
Method of Language Comparison, version A

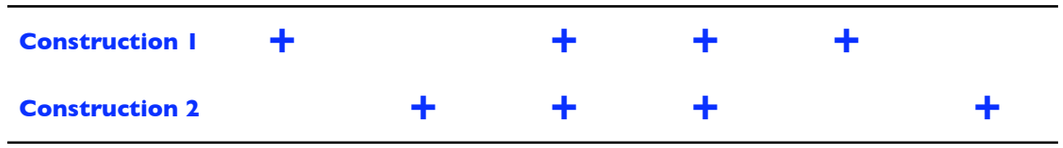
Construction 1	+	+					
Construction 2			+	+			
Construction 3	+				+	+	

Construction 1	+			+	+	+	
Construction 2		+	+	+			+

	1	2	3	4	5	6
1	0	1	3	3	1	1
2	1	0	2	2	2	2
3	3	2	0	0	2	2
4	3	2	0	0	2	2
5	1	2	2	2	0	0
6	1	2	2	2	0	0

Method of Language Comparison, version A

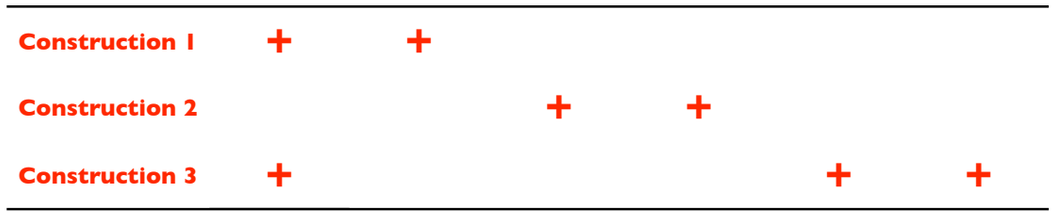


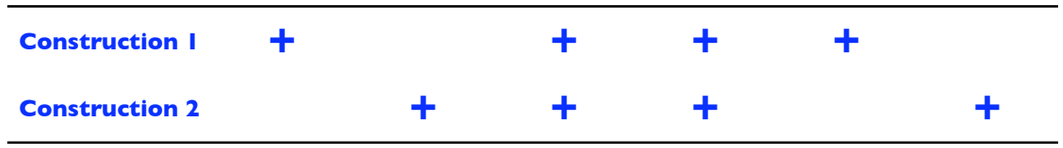


	1	2	3	4	5	6
1	0	1	3	3	1	1
2	1	0	2	2	2	2
3	3	2	0	0	2	2
4	3	2	0	0	2	2
5	1	2	2	2	0	0
6	1	2	2	2	0	0

	1	2	3	4	5	6
1	0	2	1	1	0	2
2	2	0	1	1	2	0
3	1	1	0	0	1	1
4	1	1	0	0	1	1
5	0	2	1	1	0	2
6	2	0	1	1	2	0

Method of Language Comparison, version A

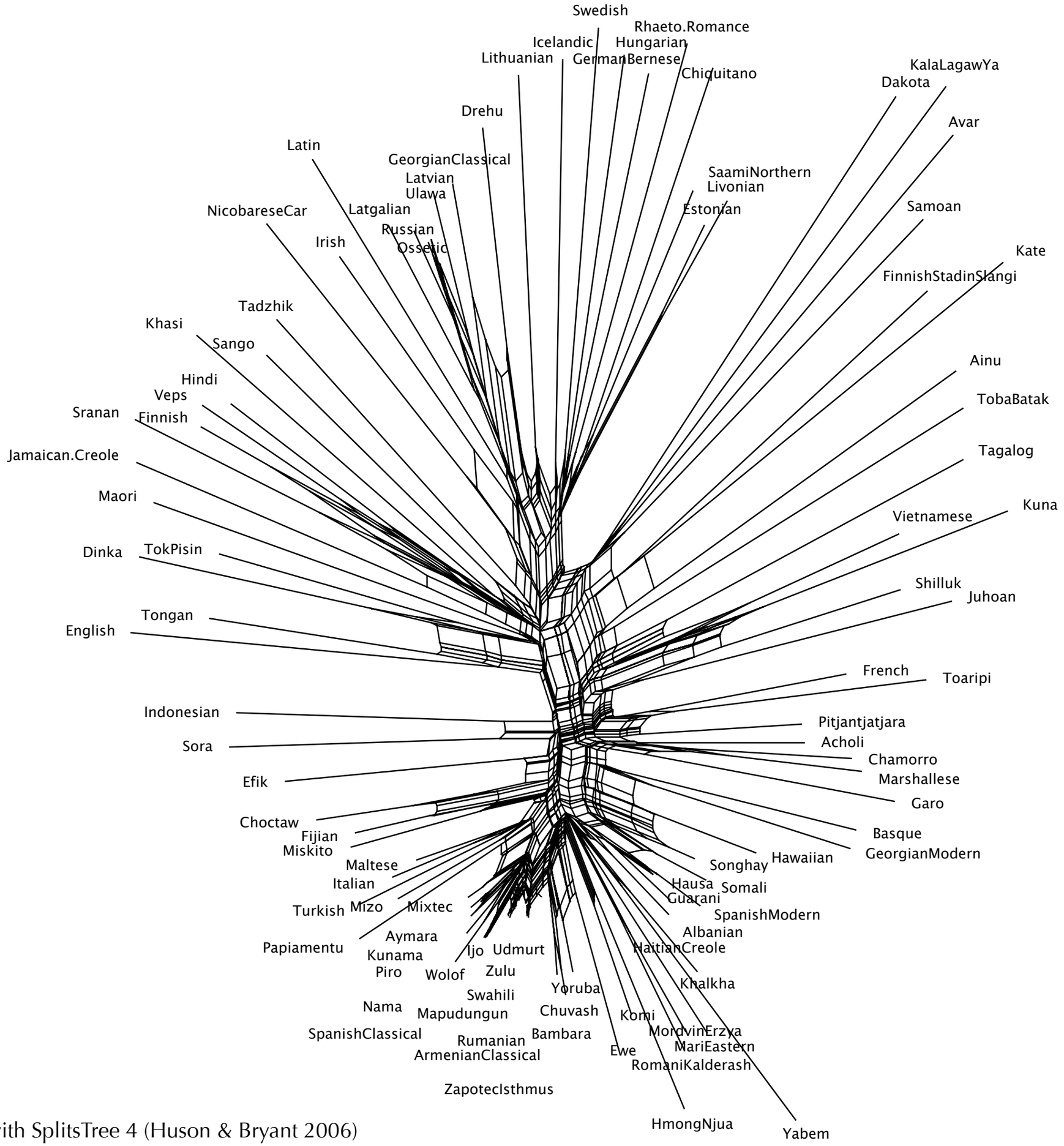




	1	2	3	4	5	6
1	0	1	3	3	1	1
2	1	0	2	2	2	2
3	3	2	0	0	2	2
4	3	2	0	0	2	2
5	1	2	2	2	0	0
6	1	2	2	2	0	0

How similar are they ? (Matrix Correlation)

	1	2	3	4	5	6
1	0	2	1	1	0	2
2	2	0	1	1	2	0
3	1	1	0	0	1	1
4	1	1	0	0	1	1
5	0	2	1	1	0	2
6	2	0	1	1	2	0



NeighborNet, made with SplitsTree 4 (Huson & Bryant 2006)

From construction-types to construction-tokens

Selectors	1	2	3	4	5	6
Construction 1	+	+				
Construction 2			+	+		
Construction 3	+				+	+

From construction-types to construction-tokens

Selectors	1	2	3	4	5	6
Construction 1	+					
Construction 4		+				
Construction 2			+			
Construction 5			+			
Construction 3				+		
Construction 6						+
Construction 7	+					

From construction-types to construction-tokens

Selectors	1	2	3	4	5	6
Construction 1	+					
Construction 4		+				
Construction 2			+			
Construction 5			+			
Construction 3				+		
Construction 6					+	
Construction 7	+					

Similarity between constructions is language-specific!

Comparative notions ?

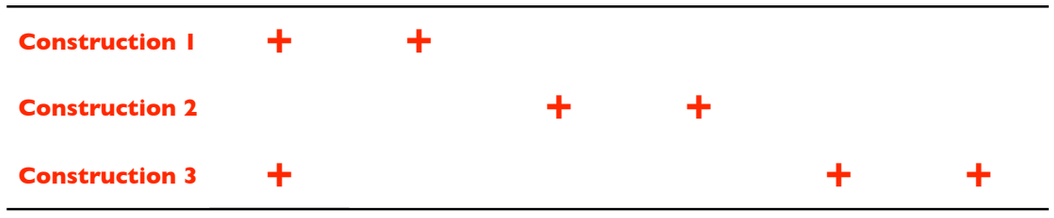
Comparative notions ?

- Constructions are language specific

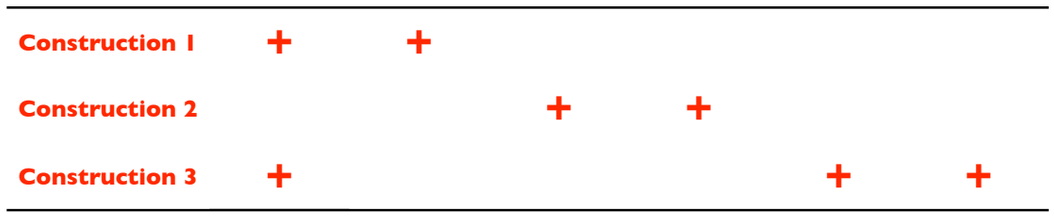
Comparative notions ?

- Constructions are language specific
- However, some (limited) options exist to analyse constructions typologically
  - morpological structure
  - order of elements
  - length
  - …

Method of Language Comparison, version B

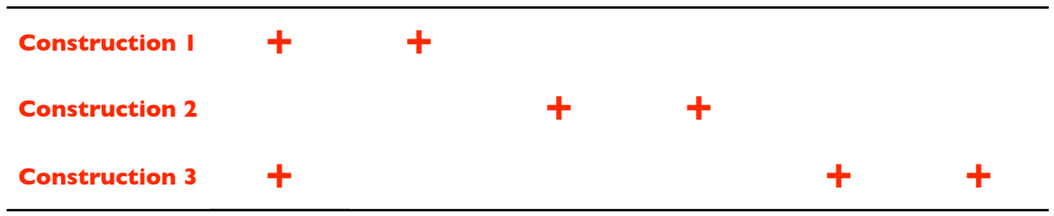


Method of Language Comparison, version B



Differences	Construction 1	Construction 2
Construction 1		
Construction 2		
Construction 3		

Method of Language Comparison, version B



Differences	Construction 1	Construction 2
Construction 1	4	4
Construction 2	2	2
Construction 3	3	5

Method of Language Comparison, version B

	Length
Construction 1	7
Construction 2	3
Construction 3	4
Construction 1	5
Construction 2	2

Method of Language Comparison, version B

	Length
Construction 1	7
Construction 2	3
Construction 3	4
Construction 1	5
Construction 2	2

Differences	Construction 1	Construction 2
Construction 1	| 7 - 5 | = 2	| 7 - 2 | = 5
Construction 2	| 3 - 5 | = 2	| 3 - 2 | = 1
Construction 3	| 4 - 5 | = 1	| 4 - 2 | = 2

Method of Language Comparison, version B

Method of Language Comparison, version B

Functions	Construction 1	Construction 2
Construction 1	4	4
Construction 2	2	2
Construction 3	3	5

Method of Language Comparison, version B

Functions	Construction 1	Construction 2
Construction 1	4	4
Construction 2	2	2
Construction 3	3	5

Length	Construction 1	Construction 2
Construction 1	2	5
Construction 2	2	1
Construction 3	1	2

Method of Language Comparison, version B

Functions	Construction 1	Construction 2
Construction 1	4	4
Construction 2	2	2
Construction 3	3	5

How similar are they ? (Matrix Correlation)

Length	Construction 1	Construction 2
Construction 1	2	5
Construction 2	2	1
Construction 3	1	2

case study:
Inchoative - causative verb pairs

case study: Inchoative - causative verb pairs

- Inchoative

  “The stick broke.”

case study: Inchoative - causative verb pairs

- Inchoative
  "The stick broke."
- Causative
  "The girl broke the stick."

case study: Inchoative - causative verb pairs

- Inchoative
  "The stick broke."

- Causative
  "The girl broke the stick."

Based on data from: Haspelmath, Martin. 1993. "More on the typology of inchoative/causative verb alternations." In: Comrie, B. & Polinsky, M. (eds.) *Causatives and transitivity*. Amsterdam: Benjamins, 87-120.

Some non-identical verb pairs in English

Some non-identical verb pairs in English

- *die - kill*
- *learn - teach*
- *rise - raise*
- *get lost - lose*
- *go out - put out*

Verb pairs investigated

Verb pairs investigated

begin	dry	melt
boil	fill	open
break	finish	rise/raise
burn	freeze	rock
change	gather	roll
close	get lost/	sink
connect	lose	split
destroy	go out/put	spread
develop	out	stop
die/kill	improve	turn
dissolve	learn/teach	wake up

Strategies for encoding inchoative-causative relation

(Terminology of: Haspelmath 1993)

Strategies for encoding inchoative-causative relation

- Different roots:
  Suppletive strategy (e.g. *die* - *kill*)

(Terminology of: Haspelmath 1993)

Strategies for encoding inchoative-causative relation

- Different roots:
  Suppletive strategy (e.g. *die* - *kill*)
- No differentiation:
  Labile strategy (e.g. *break*)

(Terminology of: Haspelmath 1993)

Strategies for encoding inchoative-causative relation

- Different roots:
  Suppletive strategy (e.g. *die* - *kill*)
- No differentiation:
  Labile strategy (e.g. *break*)
- Causative derived from Inchoative:
  Causative strategy (e.g. German *enden* - *beenden)*

(Terminology of: Haspelmath 1993)

Strategies for encoding inchoative-causative relation

- Different roots:
  Suppletive strategy (e.g. *die - kill*)
- No differentiation:
  Labile strategy (e.g. *break*)
- Causative derived from Inchoative:
  Causative strategy (e.g. German *enden - beenden)*
- Inchoative derived from Causative:
  Anticausative strategy (e.g. *be destroyed - destroy*)

(Terminology of: Haspelmath 1993)

Strategies for encoding inchoative-causative relation

- Different roots:
  Suppletive strategy (e.g. *die - kill*)
- No differentiation:
  Labile strategy (e.g. *break*)
- Causative derived from Inchoative:
  Causative strategy (e.g. German *enden - beenden)*
- Inchoative derived from Causative:
  Anticausative strategy (e.g. *be destroyed - destroy*)
- No (synchronic) primacy for one or the other:
  Equipollent strategy (e.g. German *versinken - versenken*)

(Terminology of: Haspelmath 1993)

Arabic

Class A: C / CC
1. saʱaa / saʱʱaa
8. darasa / darrasa
14. damara / dammara
31. waqafa / waqqafa

Class B: in / ø
2. inkasara / kasara
5. infataʱa / fataʱa
6. inqafala / qafala
13. inṣahara / ṣahara
30. inšaqqa / šaqqa

Class C: t / ʔ
3. iʱtaraqa / ʔaʱraqa
22. intahaa / ʔanhaa

Class D: t / ø
9. iltamma / lamma
10. intašara / našara
17. irtabaṭa / rabaṭa
21. irtafaʕa / rafaʕa
27. imtalaʔa / malaʔa

Class E: ø / ʔ
11. ġariqa / ʔaġraqa
18. ġalaa / ʔaġlaa
23. daara / ʔadaara
26. ðaaba / ʔaðaaba

Class F: ta / ø
12. tabaddala / baddala
16. taṭawwara / ṭawwara
19. taʔarjaʱa / ʔarjaʱa
24. tadaʱraja / daʱraja
25. tajammada / jammada
28. taʱassana / ʱassana

Singular cases:
4. maata / qatala
7. badaʔa
15. daaʕa / xasira
20. inṭafaʔa / ʔaṭfaʔa
29. jaffa / jaffafa

English

Class A: Identical
1. wake up
2. break
3. burn
5. open
6. close
7. begin
9. gather
10. spread
11. sink
12. change
13. melt
16. develop
17. connect
18. boil
19. rock
22. finish
23. turn
24. roll
25. freeze
26. dissolve
27. fill
28. improve
29. dry
30. split
31. stop

Singular cases:
4. die / kill
8. learn / teach
14. be destroyed / destroy
15. get lost / lose
20. go out / put out
21. rise / raise

Finnish

Class A: ø / ttA
1. herätä / herättää
3. palaa / polttaa
8. oppia / opettaa
10. levitä / levittää
13. sulaa / sulattaa
18. kiehua / kiehuttaa
19. kiikkua / kiikuttaa
20. sammua / sammuttaa
21. kohota / kohottaa
22. loppua / lopettaa
24. vieriä / vierittää
25. jäätyä / jäädyttää
26. liueta / liuottaa
31. pysähtyä / pysähdyttää

Class B: U / A
2. murtua / murtaa
12. muuttua / muuttaa
16. kehittyä / kehittää
23. vääntyä / vääntää
27. täyttyä / täyttää
28. parantua / parantaa

Class C: UtU / ø
5. avautua / avata
6. sulkeutua / sulkea
14. tuhoutua / tuhota

Singular cases:
4. kuolla / tappaa
7. alkaa / aloitaa
9. kokoontua / koota
11. laskea
15. hukkaantua / hukata
17. yhtyä / yhdistää
29. kuivaa / kuivata
30. haljeta / halkaista

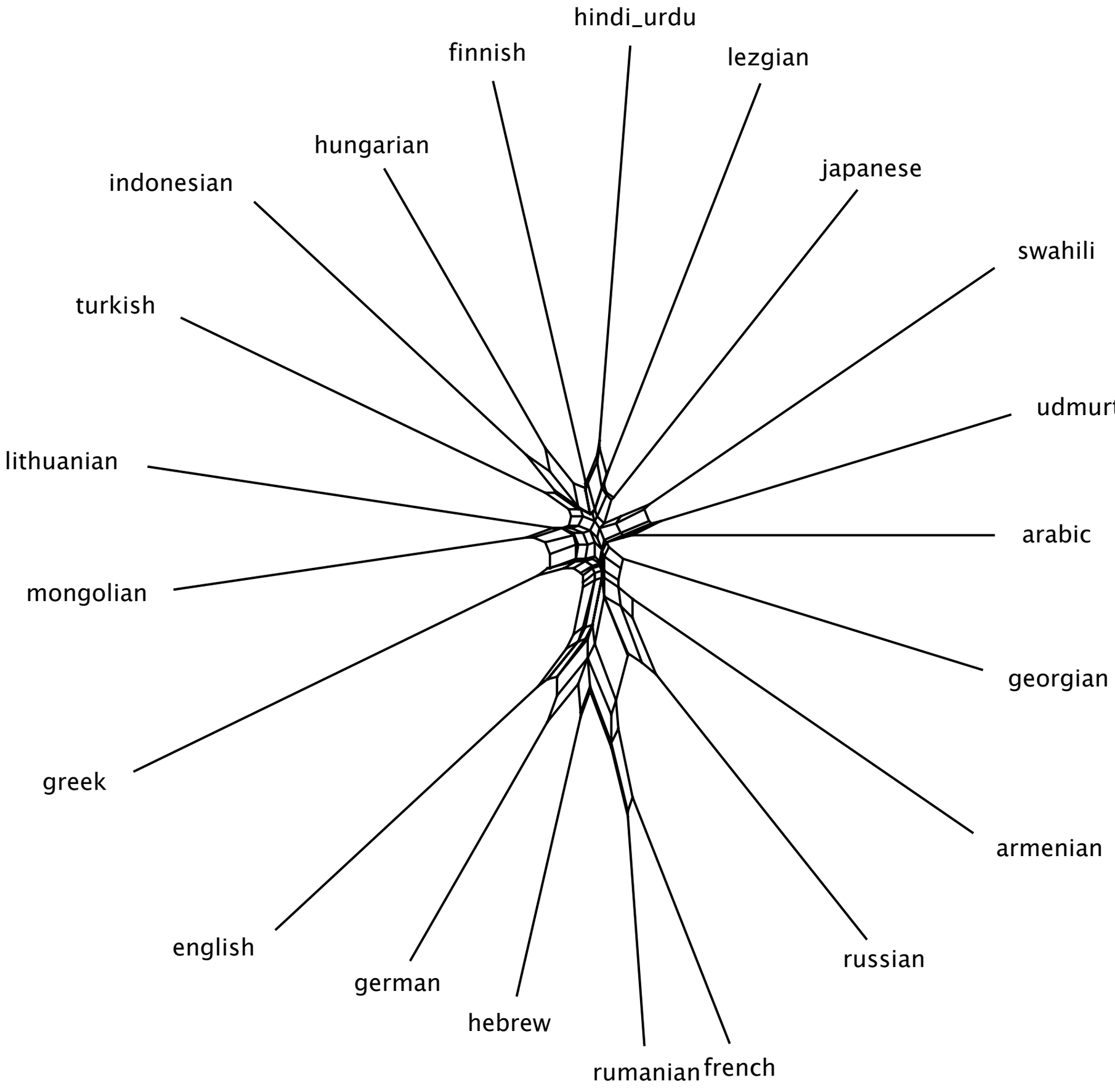
French

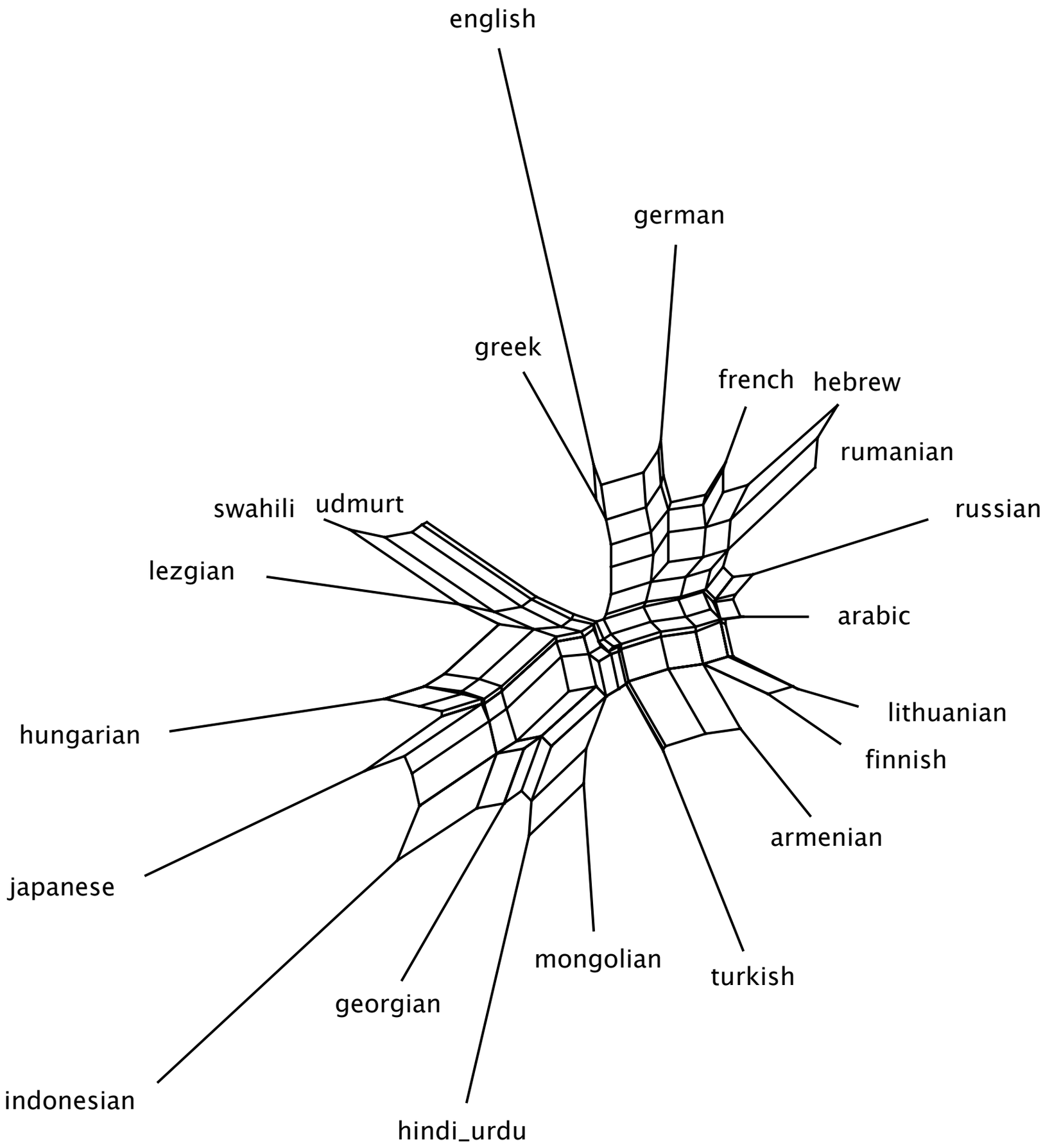
Class A: se / ø
1. se réveiller / réveiller
2. se briser / briser
5. s’ouvrir / ouvrir
6. se fermer / fermer
9. s’assembler / assembler
10. s’étendre / étendre
11. s’enfoncer / enfoncer
15. se perdre / perdre
16. se développer / développer
17. se lier / lier
19. se balancer / balancer
20. s’éteindre / éteindre
21. se lever / lever
23. se tourner / tourner
26. se dissoudre / dissoudre
27. se remplir / remplir
28. s’améliorer / améliorer
30. se fendre / fendre
31. s’arrêter / arrêter

Class B: Identical
3. brûler
7. commencer
8. apprendre
12. changer
22. finir
24. rouler
25. geler
29. sécher

Class C: ø / faire
13. fondre / faire fondre
18. bouillir / faire bouillir

Singular cases:
4. mourir / tuer
14. être détruit / détruir


hindi_urdu
finnish
lezgian
hungarian
indonesian
japanese
swahili
turkish
udmurt
lithuanian
arabic
mongolian
georgian
greek
armenian
english
russian
german
hebrew
rumanian
french


english
german
greek
french
hebrew
rumanian
russian
swahili
udmurt
lezgian
arabic
lithuanian
finnish
armenian
hungarian
japanese
mongolian
turkish
georgian
indonesian
hindi_urdu

Summary

Summary

- Language comparison works on constructions (either types or tokens)

Summary

- Language comparison works on constructions (either types or tokens)
- Relevant constructions are selected specifically for the comparison

Summary

- Language comparison works on constructions (either types or tokens)
- Relevant constructions are selected specifically for the comparison
- Constructions are language specific

Summary

- Language comparison works on constructions (either types or tokens)
- Relevant constructions are selected specifically for the comparison
- Constructions are language specific
- Comparison of (a set of) constructions possible via selector-similarity

Summary

- Language comparison works on constructions (either types or tokens)
- Relevant constructions are selected specifically for the comparison
- Constructions are language specific
- Comparison of (a set of) constructions possible via selector-similarity
- Only limited cross-linguistic comparison of constructions possible