

Visualizing the typology of natural classes

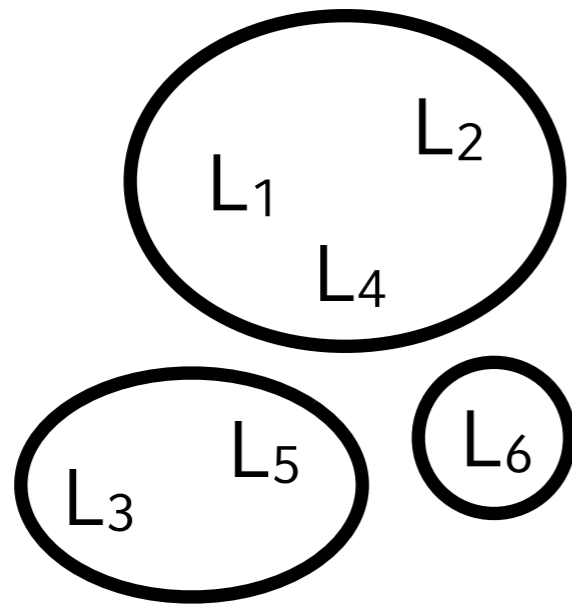
Michael Cysouw

Visualizing the typology of natural (?) classes

Michael Cysouw

Metrics instead of Classes

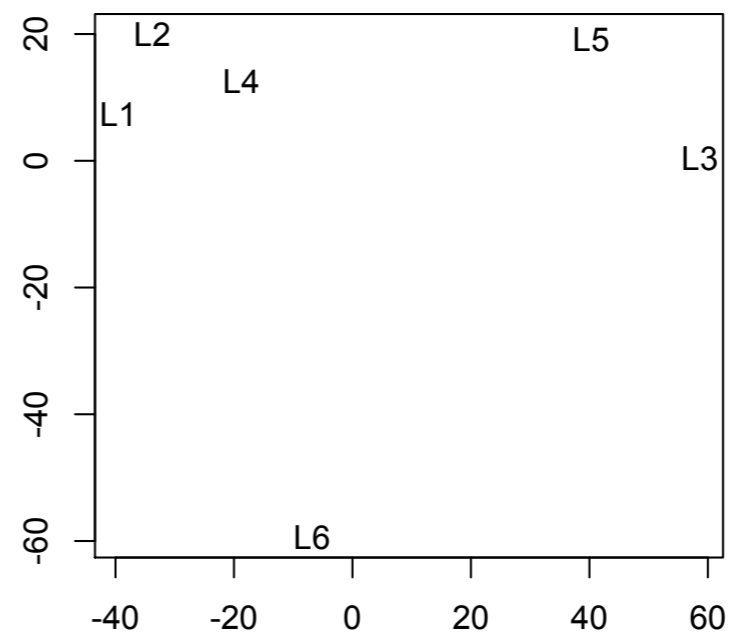
Classes



	L1	L2	L3	L4	L5	L6
L1	0	0	1	0	1	1
L2	0	0	1	0	1	1
L3	1	1	0	1	0	1
L4	0	0	1	0	1	1
L5	1	1	0	1	0	1
L6	1	1	1	1	1	0

Metric

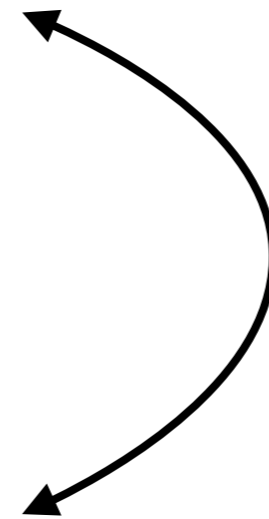
	L1	L2	L3	L4	L5	L6
L1	0	31	95	21	87	79
L2	31	0	98	10	67	83
L3	95	98	0	79	3	89
L4	21	10	79	0	50	71
L5	87	67	3	50	0	90
L6	79	83	89	71	90	0



- **Verb meaning** (V 'verb'): approximation of a cross-linguistic applicable meaning, like BEAT, HIT, GIVE
- **Micro role** (R 'role'): verb-specific roles, like BEATER, HITTEE, GIVING RECIPIENT
- **Coding device** (P 'part'): language-specific form used in coding roles, like case marking or verb inflection
- **Coding set** (C 'coding'): language-specific combination of devices used to code a role, e.g the combination of a specific preposition with a specific case
- **Coding frame** (F 'frame'): combination of coding sets that form a construction, like passives or applicatives
- **Language** (L 'language')

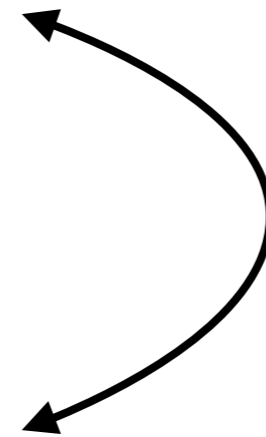
Recurrently similar coding of micro-roles indicates that the micro-roles are similar

micro-role	Icelandic	Hoocak	Chintang
hitter	NP-nom & V.subj	sbj.V	NP-erg & V.subj
hittee	NP-acc	obj.V	NP-abs & V.obj
liker	NP-dat	sbj.V	NP-erg & V.subj
liked entity	NP-nom & V.subj	obj.V	NP-abs & V.obj
throwing goal	inn um+NP-acc	NP+eeja	NP-abs & V.obj
helper	NP-nom & V.subj	sbj.V	NP-erg & V.subj
helpee	NP-dat	obj.V	NP-abs & V.obj



Micro-roles are **similar**, when many languages code them identically

micro-role	Icelandic	Hoocak	Chintang
hitter	NP-nom & V.subj	sbj.V	NP-erg & V.subj
hittee	NP-acc	obj.V	NP-abs & V.obj
liker	NP-dat	sbj.V	NP-erg & V.subj
liked entity	NP-nom & V.subj	obj.V	NP-abs & V.obj
throwing goal	inn um+NP-acc	NP+eeja	NP-abs & V.obj
helper	NP-nom & V.subj	sbj.V	NP-erg & V.subj
helpee	NP-dat	obj.V	NP-abs & V.obj



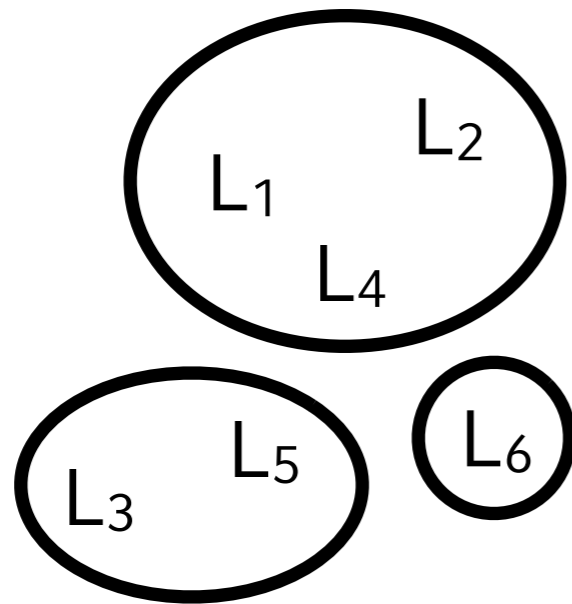
Micro-roles are **different**, when many languages code them differently

Only comparison within languages!

cutting instrument
killing instrument building material
breaking instrument taking source
beating instrument
putting goal
touching instrument
hitting instrument
sitting place recipient
living place carrying
sitting down place
peeler
putter dresser wiper
leader shaver bringer
mangler hairer stealer
toilet paper singer
hider ear teacher
builder shaver faller
shaver follower appearer
talker grin screamer
cookery holder downer pain locus
leaver knower
digger washer
feeder plater
laughter rain
hunter want
crier sad person thing
sick person pain-feeler
dry thing freezing person
peeled object
climbing goal
leaving source
poured substance loaded thing
thought content
told content wanted thing giving recipient
ground thing
carried thing
received thing
built thing
sees thing
cooked food person
shaved person
hidden thing
peel
filling material
going goal
saying addressee
hiding affectee
throwing goal
taught content
talked to person
showing addressee
asking surface
tellee
shoutee

Metrics instead of Classes

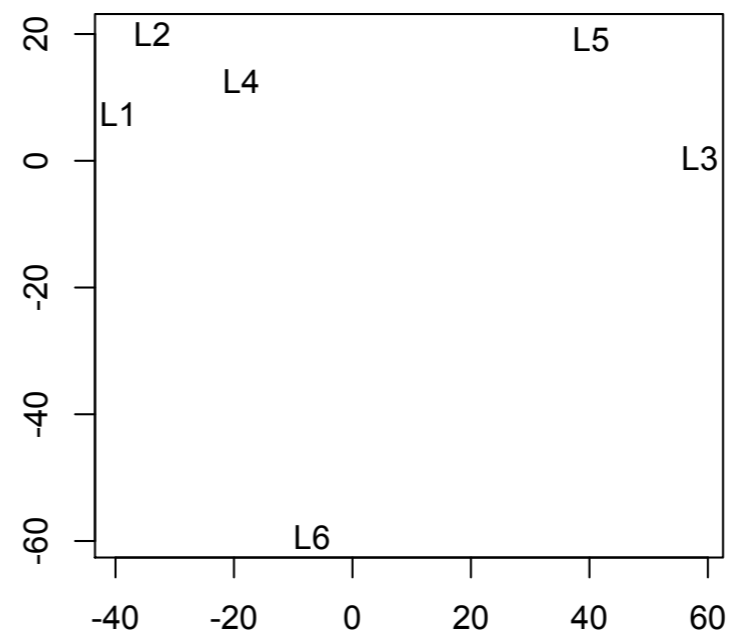
Classes



	L1	L2	L3	L4	L5	L6
L1	0	0	1	0	1	1
L2	0	0	1	0	1	1
L3	1	1	0	1	0	1
L4	0	0	1	0	1	1
L5	1	1	0	1	0	1
L6	1	1	1	1	1	0

Metric

	L1	L2	L3	L4	L5	L6
L1	0	31	95	21	87	79
L2	31	0	98	10	67	83
L3	95	98	0	79	3	89
L4	21	10	79	0	50	71
L5	87	67	3	50	0	90
L6	79	83	89	71	90	0



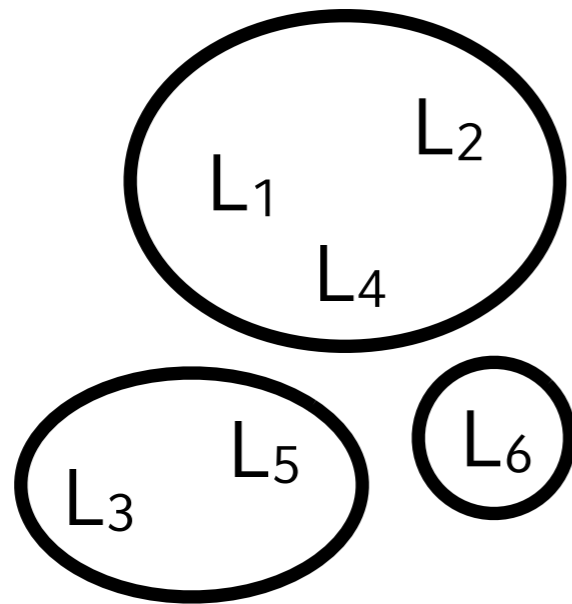
	appearer	askee	asker	beatee	beater	beating instrum.	blinker	boiled thing	breaker	breaking instrum.
Ainu: all
Ainu: dat
Ainu: instr	
Ainu: loc/all
Ainu: loc/all/abl
Ainu: obj.V
Ainu: subj.V	
Alaskan Yupik: NP-abm
Alaskan Yupik: NP-abs V.obj
Alaskan Yupik: NP-abs V.subj

Mnemonically let's call this table: **CR**

Then: **$RR = CR^T \cdot CR$** (matrix multiplication)

Metrics instead of Classes

Classes

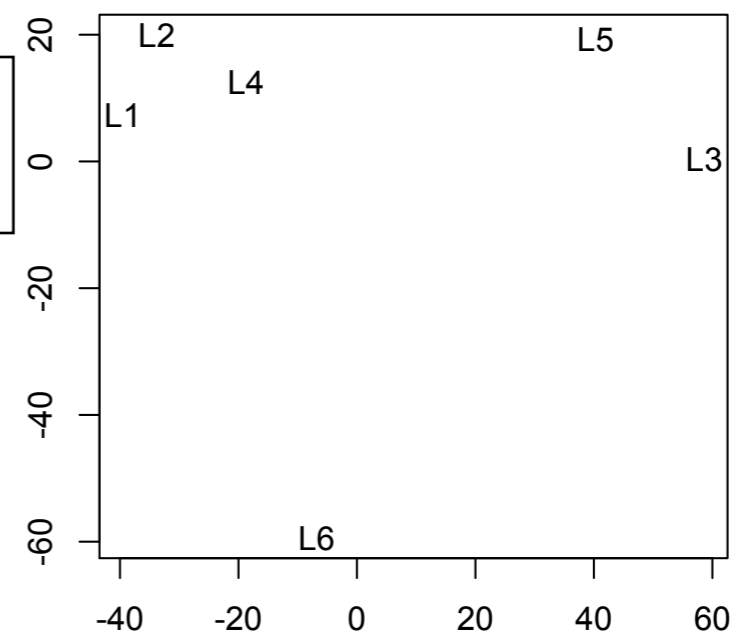


	L1	L2	L3	L4	L5	L6
L1	0	0	1	0	1	1
L2	0	0	1	0	1	1
L3	1	1	0	1	0	1
L4	0	0	1	0	1	1
L5	1	1	0	1	0	1
L6	1	1	1	1	1	0

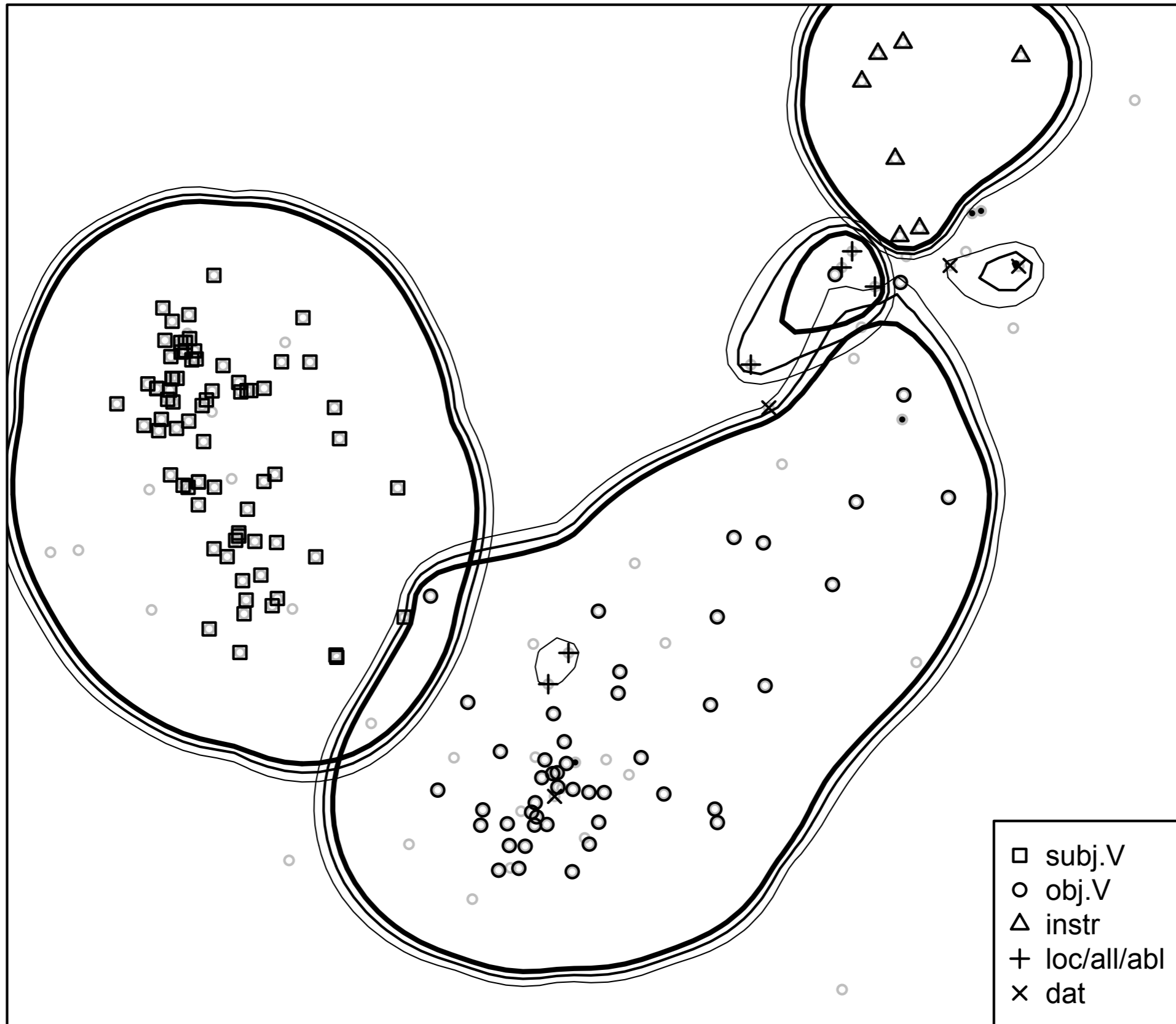
Metric

	L1	L2	L3	L4	L5	L6
L1	0	31	95	21	87	79
L2	31	0	98	10	67	83
L3	95	98	0	79	3	89
L4	21	10	79	0	50	71
L5	87	67	3	50	0	90
L6	79	83	89	71	90	0

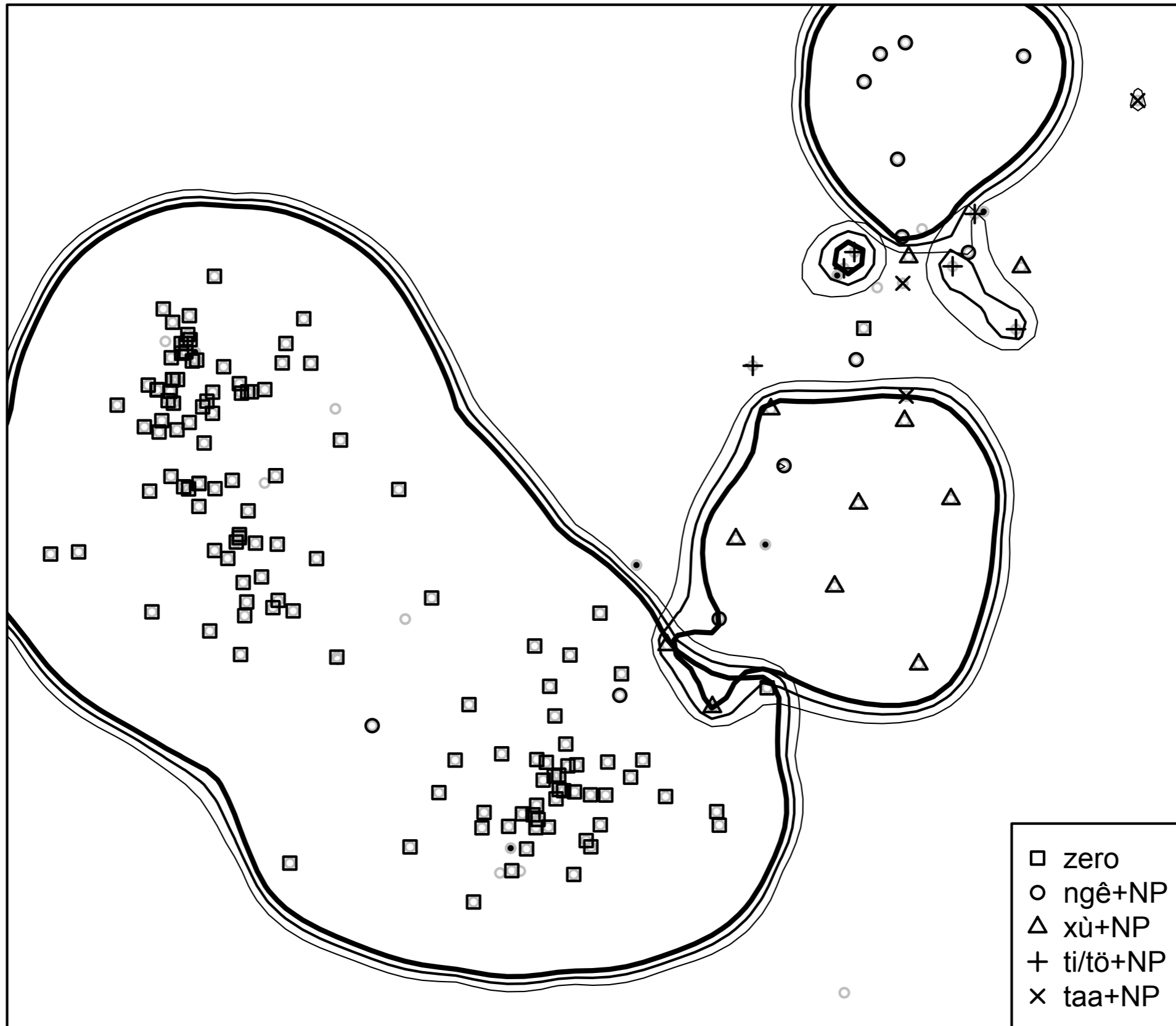
**Multidimensional
Scaling**



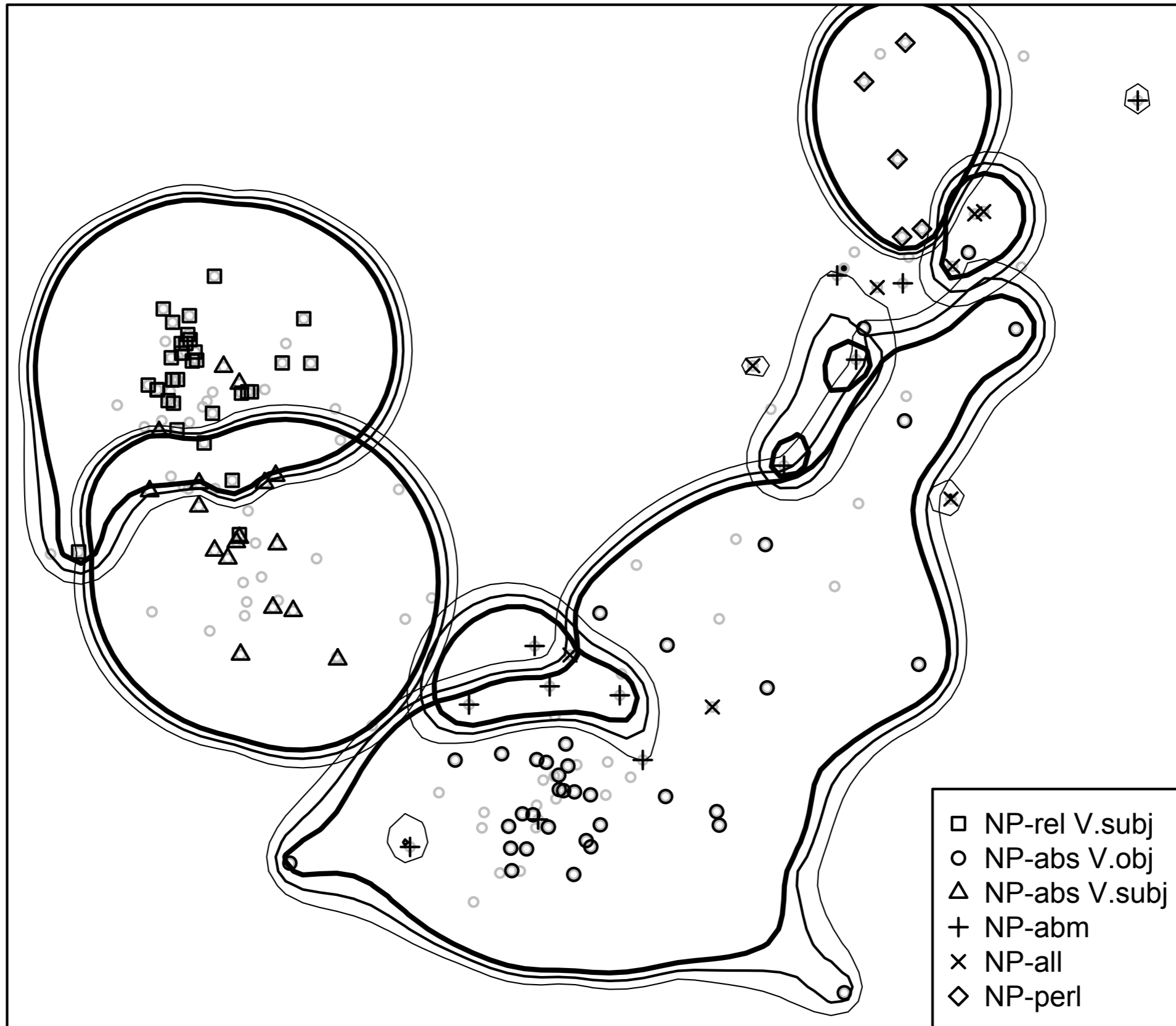
Ainu



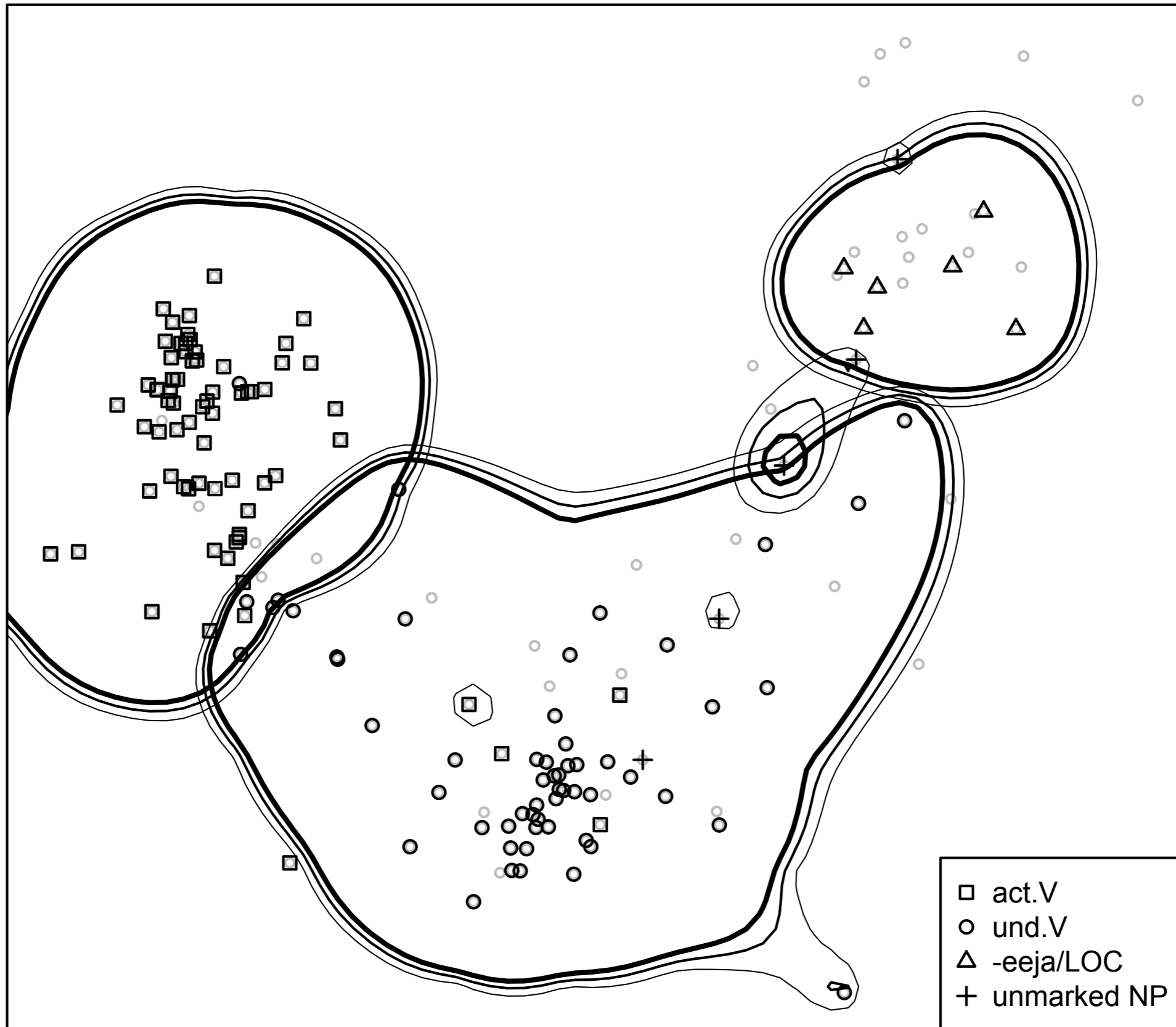
Xârâcùù



Alaskan.Yupik

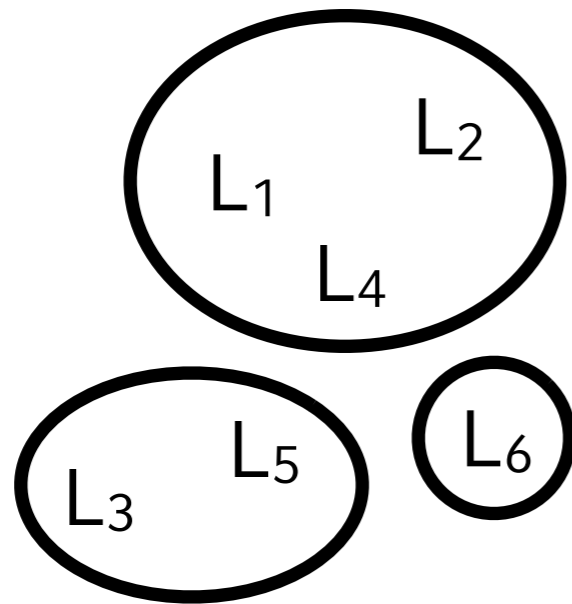


Hoocaq



Metrics instead of Classes

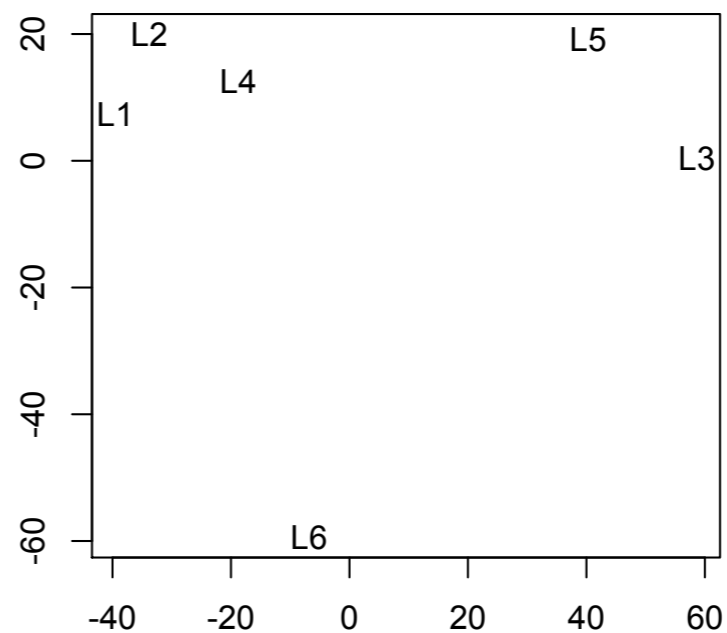
Classes



	L1	L2	L3	L4	L5	L6
L1	0	0	1	0	1	1
L2	0	0	1	0	1	1
L3	1	1	0	1	0	1
L4	0	0	1	0	1	1
L5	1	1	0	1	0	1
L6	1	1	1	1	1	0

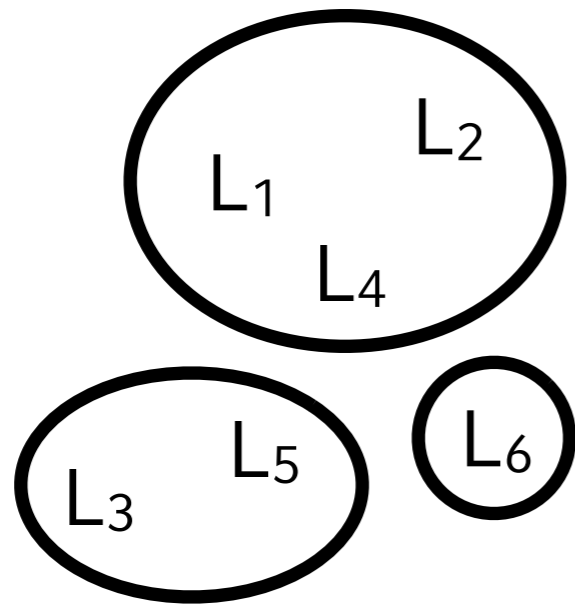
Metric

	L1	L2	L3	L4	L5	L6
L1	0	31	95	21	87	79
L2	31	0	98	10	67	83
L3	95	98	0	79	3	89
L4	21	10	79	0	50	71
L5	87	67	3	50	0	90
L6	79	83	89	71	90	0



Metrics instead of Classes

Classes

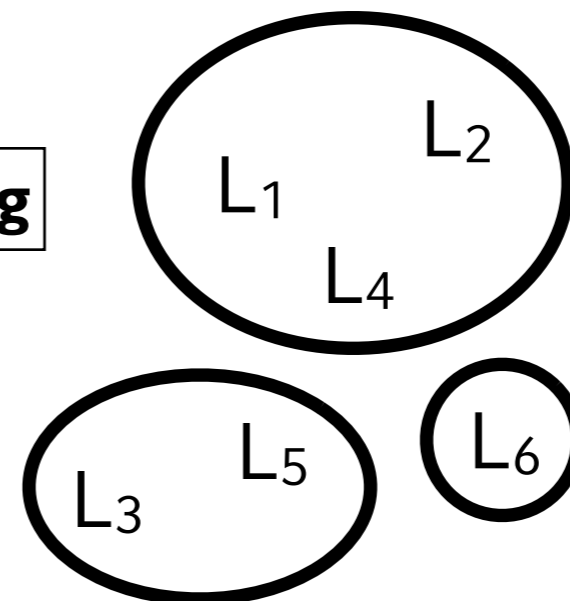


	L1	L2	L3	L4	L5	L6
L1	0	0	1	0	1	1
L2	0	0	1	0	1	1
L3	1	1	0	1	0	1
L4	0	0	1	0	1	1
L5	1	1	0	1	0	1
L6	1	1	1	1	1	0

Metric

	L1	L2	L3	L4	L5	L6
L1	0	31	95	21	87	79
L2	31	0	98	10	67	83
L3	95	98	0	79	3	89
L4	21	10	79	0	50	71
L5	87	67	3	50	0	90
L6	79	83	89	71	90	0

Partitioning



“Argument Types”

Class 1	Class 3	Class 7	Class 10	Class 12
bringing recipient	giving recipient	loading place	carrying goal	beating instrument
saying addressee	showing addressee	putting goal	going goal	breaking instrument
sending recipient	teachee	throwing goal	living place	cover
talked to person	tellee	tying goal	pouring goal	cutting instrument
			sitting place	filling material
			sitting-down place	hitting instrument
				killing instrument
				touching instrument

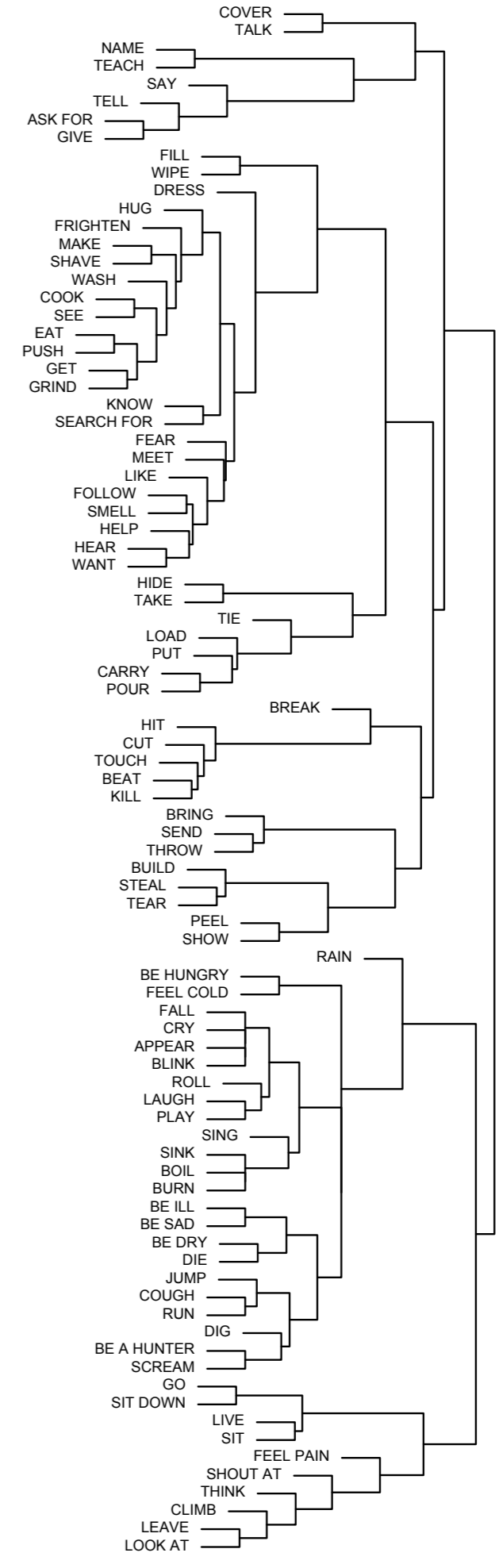
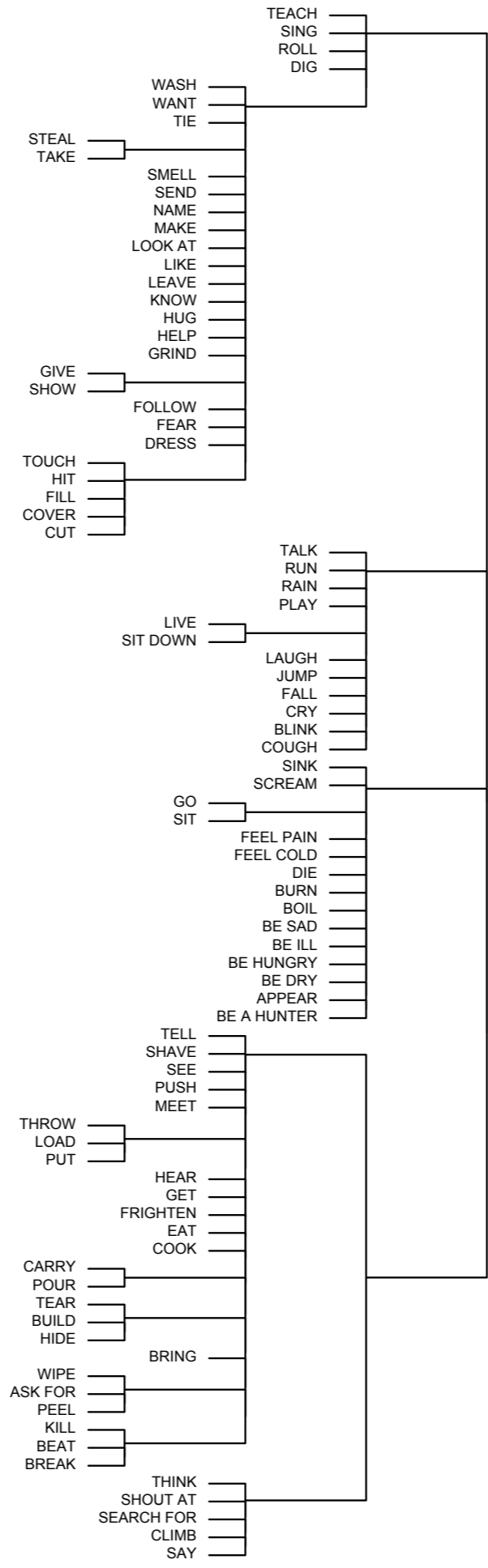
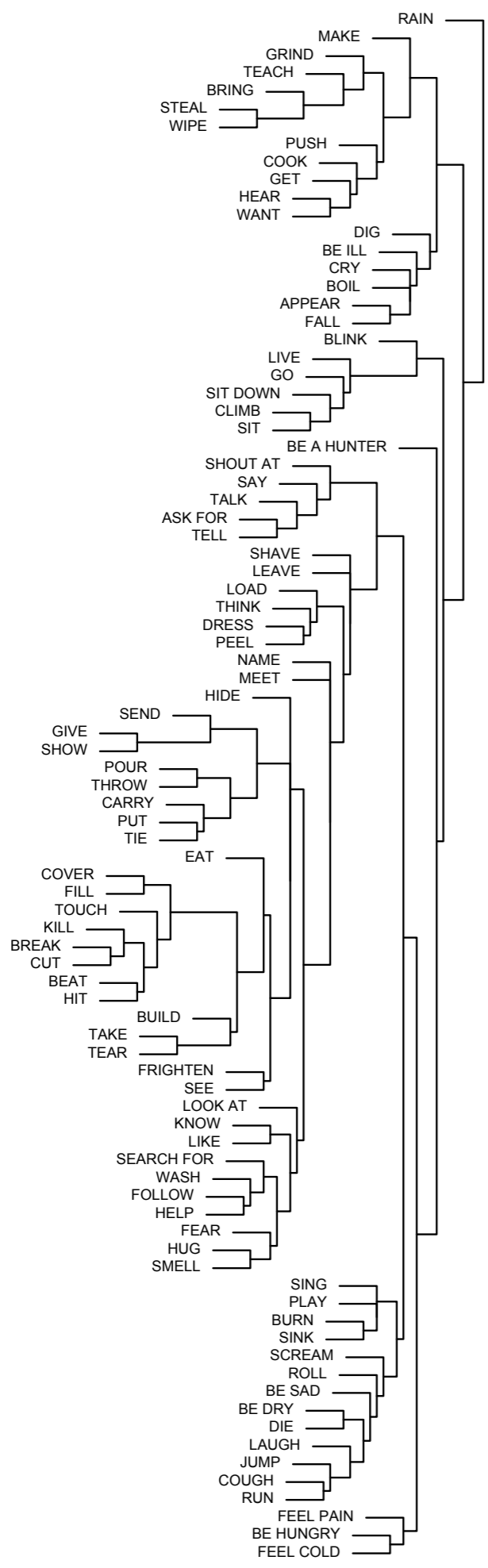
Class 2	Class 4	Class 5	Class 8	Class 9
appearer	boiled thing	broken thing	asker	askee
blinker	burnt thing	dressee	beater	beatee
climber	dieer	eaten food	breaker	brought thing
cougher	dry thing	fear stimulus	bringer	built thing
crier	fallee	filled container	builder	carried thing
laugher	freezing person	followee	carrier	climbing goal
liver	goer	gift	cooker	cooked food
...

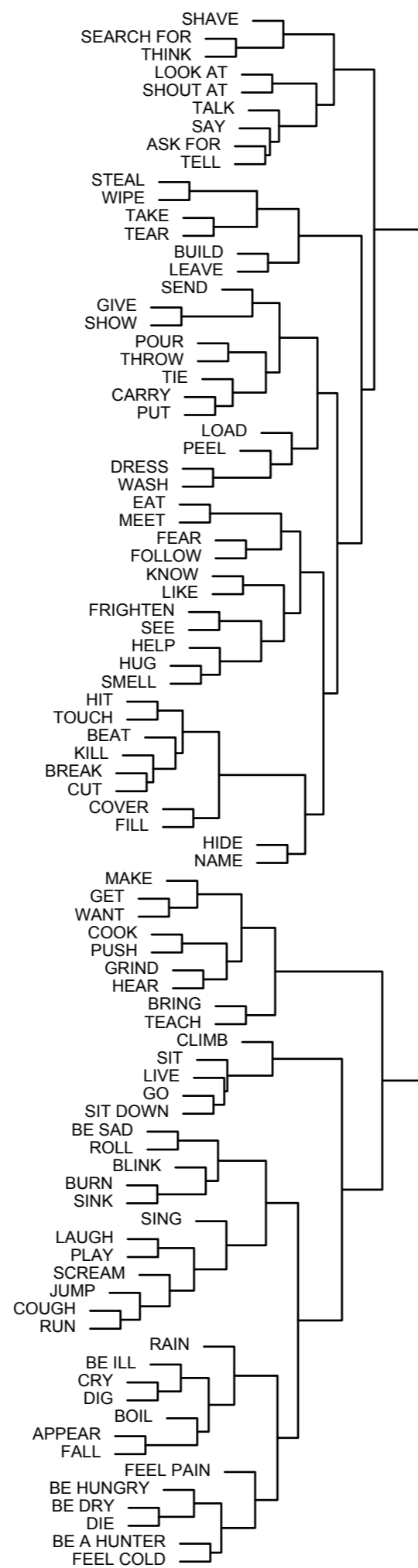
Verb Classes?

- There are different possibilities to compare verbs via their construction frame
- Basic idea: verbs are similar when their microroles are similar
- Using the table **VR** (verbs by roles), then:
$$\mathbf{VV} = \mathbf{VR} \cdot \mathbf{CR}^T \cdot \mathbf{CR} \cdot \mathbf{VR}^T$$
- Or, we need a table with the partitioning of the role **RG** (roles by groups), then:
$$\mathbf{VV} = \mathbf{VR} \cdot \mathbf{RG} \cdot \mathbf{RG}^T \cdot \mathbf{VR}^T$$
- Many more possibilities: normalization, centering...

	hittee	hitter	hitting instrument
beatee	23	6	0
beater	6	23	0
beating instrument	1	0	13

	askee	asker	requested thing
beatee	8	6	11
beater	2	20	5
beating instrument	0	0	2





Conclusion

- A notion of *similarity* is a generalization of a notion of *classification*
- Similarity can be used to derive a partitioning, hierarchical clusters, scaling, etc.
- Similarity between micro roles can be established through cross-linguistic variation
- Verbal classes can be derived from the similarity between coding frames