Using quantitative methods for semantic maps

William Croft
University of New Mexico
wcroft@unm.edu

Visualization and explanation

What is the semantic map model?

- It is a tool for visualizing similarity relations between discrete entities
- “Similarity” is defined as: concepts expressed by the same form in one or more languages (***co-expression***; Hartmann et al. 2014)
- This is not the only type of similarity that can be measured by this visualization technique
- But it happens to be the type of similarity that typologists have used the model for

Not just grammatical co-expression

- The “semantic map” model is a model of similarity of any kind, including any kind of co-expression
- It doesn’t have to be co-expression of grammatical elements
- It could be co-expression of lexical elements

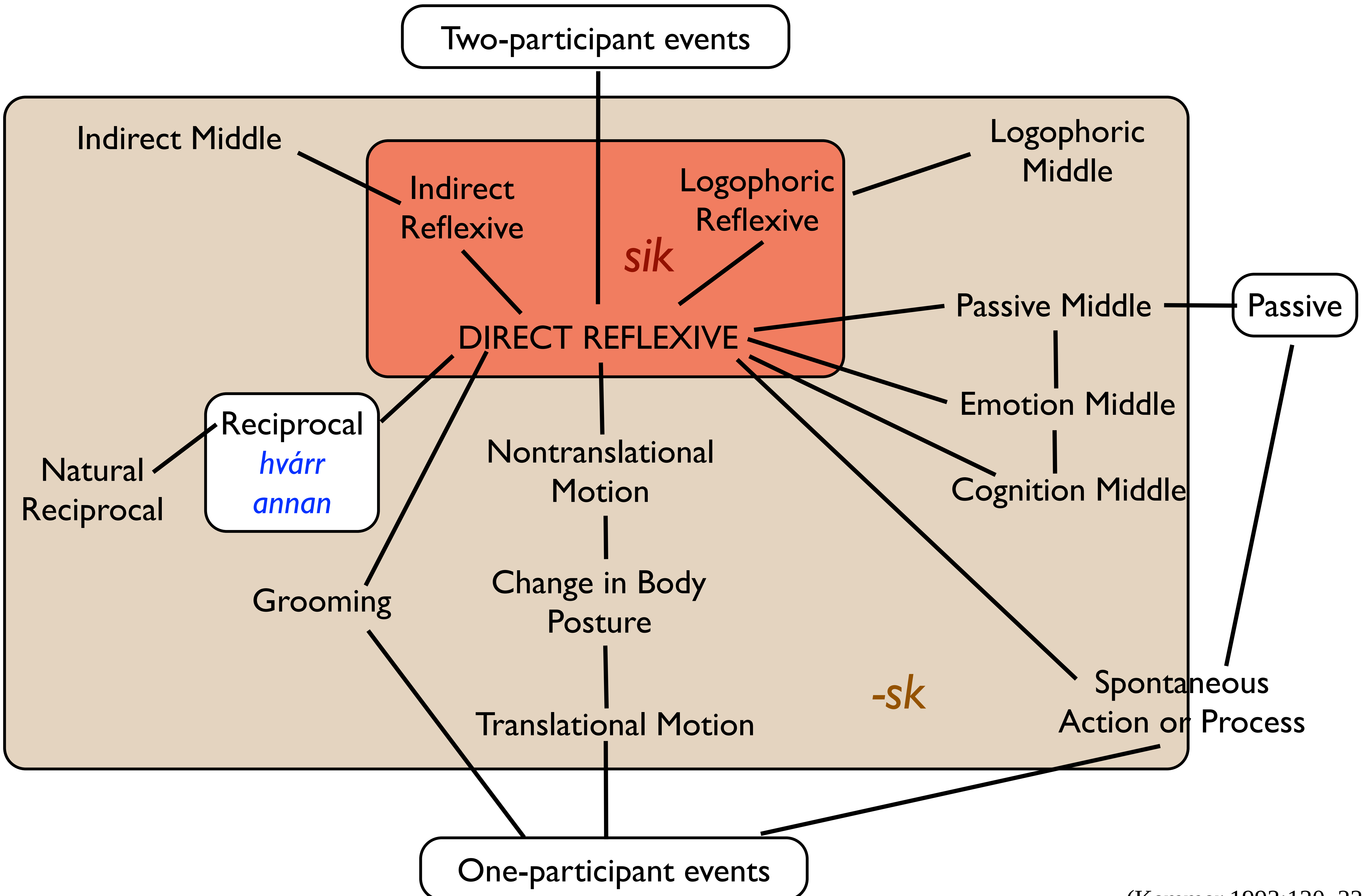
Co-expression and explanation

- Co-expression—similarity defined as two concepts expressed by the same form in at least one language—is a typological generalization (cf. Haiman 1978)
- But many of us would also like an ***explanation*** for co-expression patterns (although some typologists take a nominalist position)

Co-expression and explanation

- Examples of explanations:
  - ✳ conceptual similarity (“mental maps”, “conceptual space”, etc.) of different kinds
  - ✳ diachronic spread (and contraction) of use
  - ✳ phonetic convergence of diachronically unrelated forms (“homonymy”)
- These are not mutually exclusive
- In some cases, their interaction accounts for “anomalies”

The “doughnut”



(Kemmer 1993:120, 226)

A terminological issue

- I use the term ***conceptual space*** for the underlying graph, and ***semantic map*** for language-specific categories mapped onto the space
- It is important to distinguish between ***comparative concepts***, like the conceptual space, and ***language-specific categories***, like the semantic maps (Haspelmath 2010, Croft 2014, inter alia)

Signal and “noise”

- Homonymy introduces “noise” into the conceptual space interpreted as a space of conceptual similarity
- In a different way, diachronic layering of forms like the two Old Norse middle markers also introduces “noise” in the sense that new forms intruding into a conceptual space “break up” similarity networks
- Ideally we would integrate all three explanations, but given a set of synchronic data, we lack the relevant diachronic information

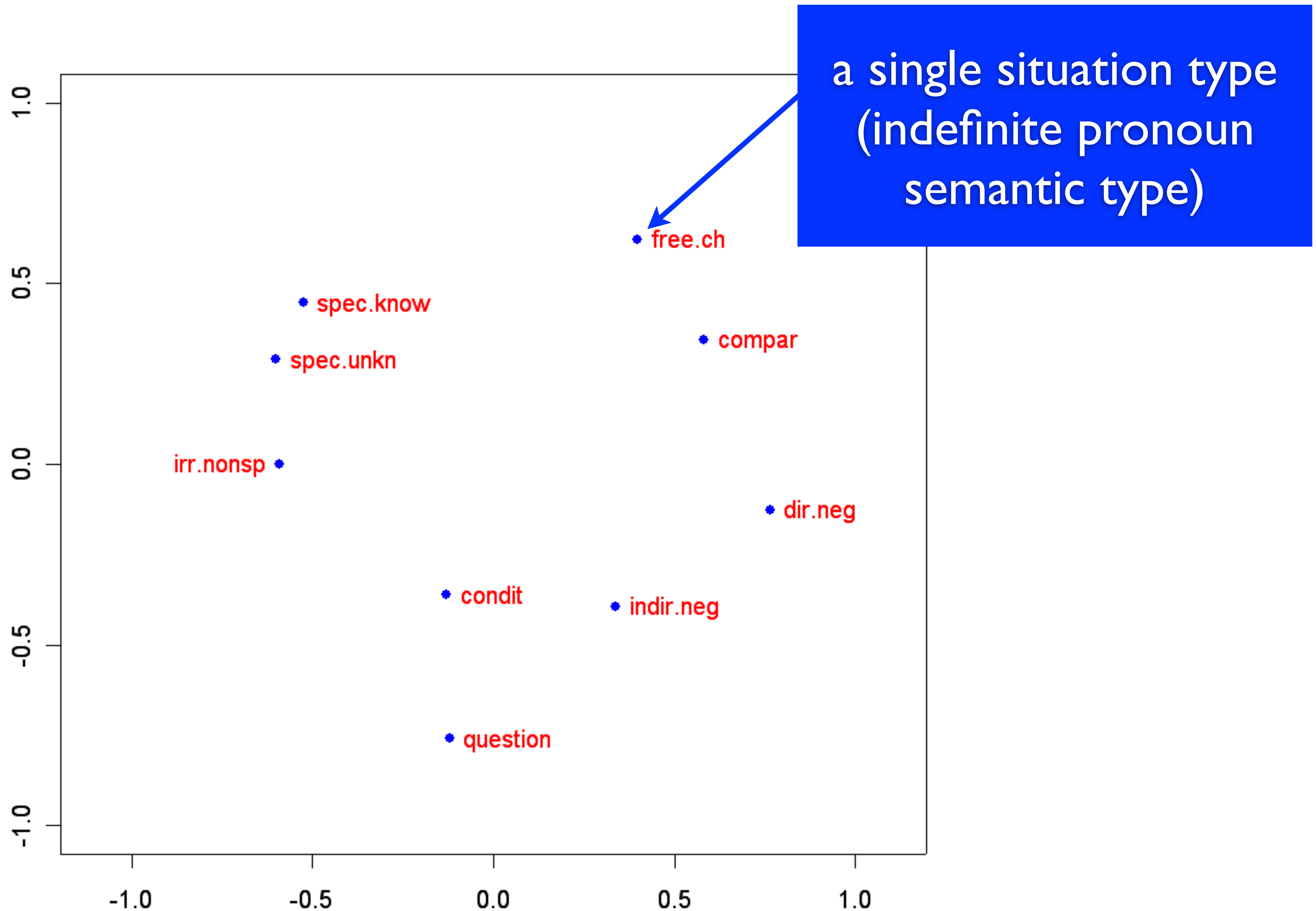
Automated algorithms, 1: MDS and Euclidean models

Multidimensional scaling in analyzing linguistic behavior

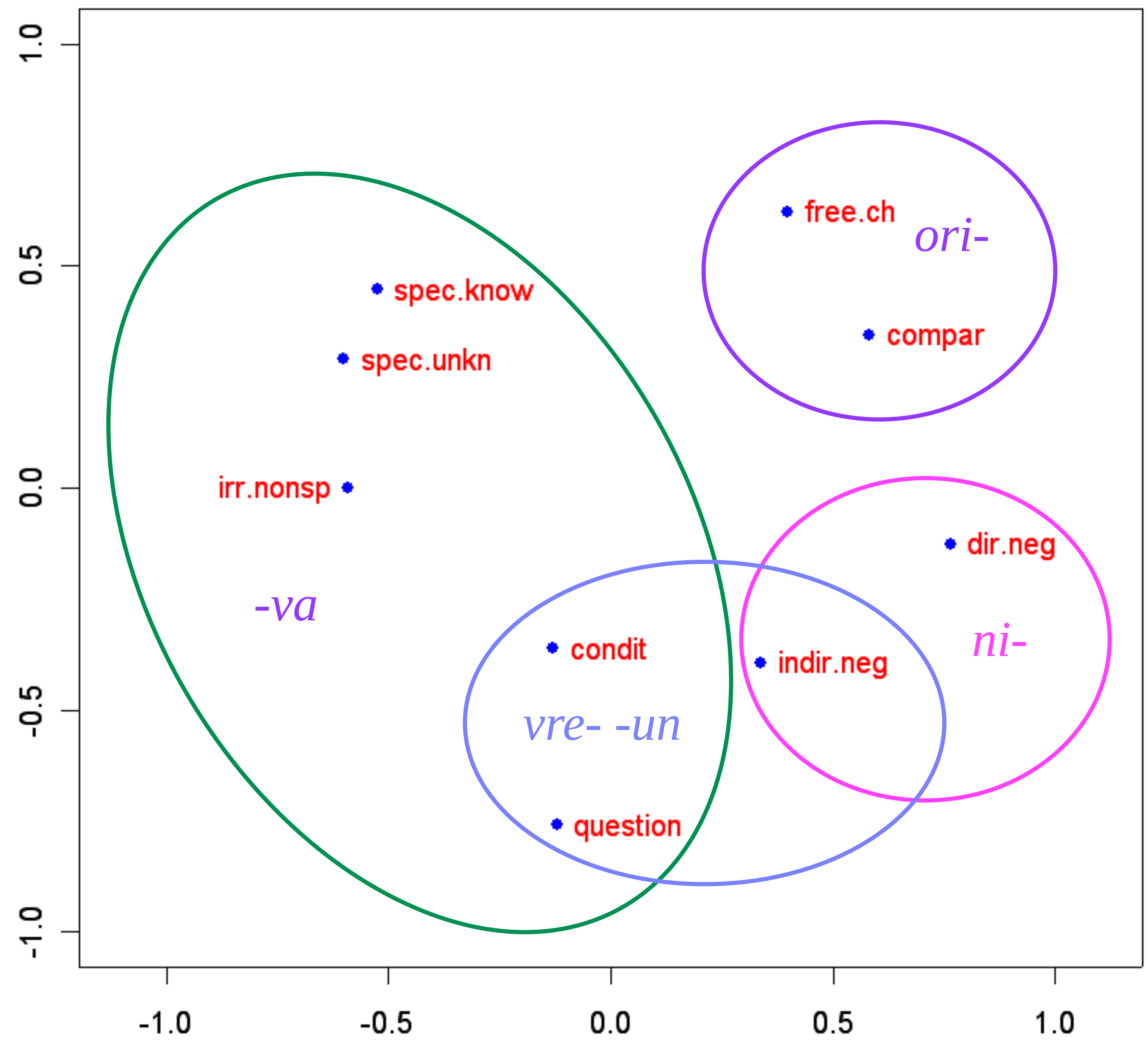
- Linguistic distributional data is similar to voting data: meanings "vote" Y or N on whether they can be expressed by a linguistic form

	Romanian:				Kazakh:			
	va-	*vre- -un*	*ori-*	*ni-*	*älde-*	*bir*	*bolsa da*	*eš*
Specific known	Y	N	N	N	Y	Y	N	N
Specific unknown	Y	N	N	N	Y	Y	N	N
Irrealis nonspecific	Y	N	N	N	N	Y	N	N
Question	Y	Y	N	N	N	Y	Y	N
Conditional	Y	Y	N	N	N	Y	Y	N
Comparative	N	N	Y	N	N	N	Y	N
Free choice	N	N	Y	N	N	N	Y	N
Indirect negation	N	Y	N	Y	N	Y	N	N
Direct negation	N	N	N	Y	N	N	N	Y

A spatial model of conceptual similarity among indefinite pronouns



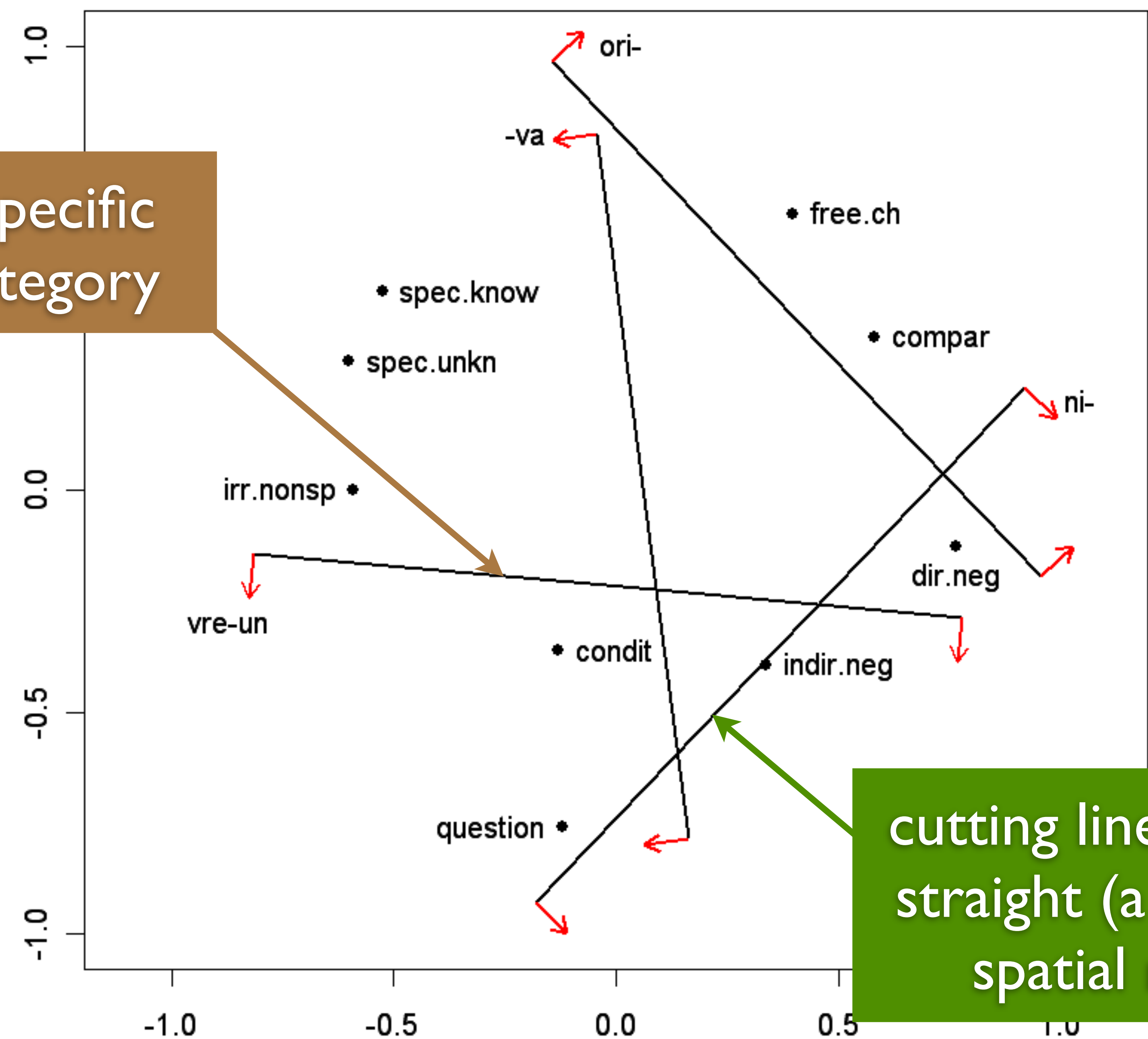
Romanian indefinite pronouns in an MDS spatial model—the wrong way



Conceptual spaces and semantic maps

- In a Euclidean model, the language-specific categories are bisections of the space (cutting lines)
- It is not correct to draw any shape around points/concepts to depict a language-specific category, unlike the “classical” graph structure model

Romanian indefinite pronouns in an MDS spatial model—the right way



a language-specific indefinite category

cutting lines must be straight (a Euclidean spatial model)

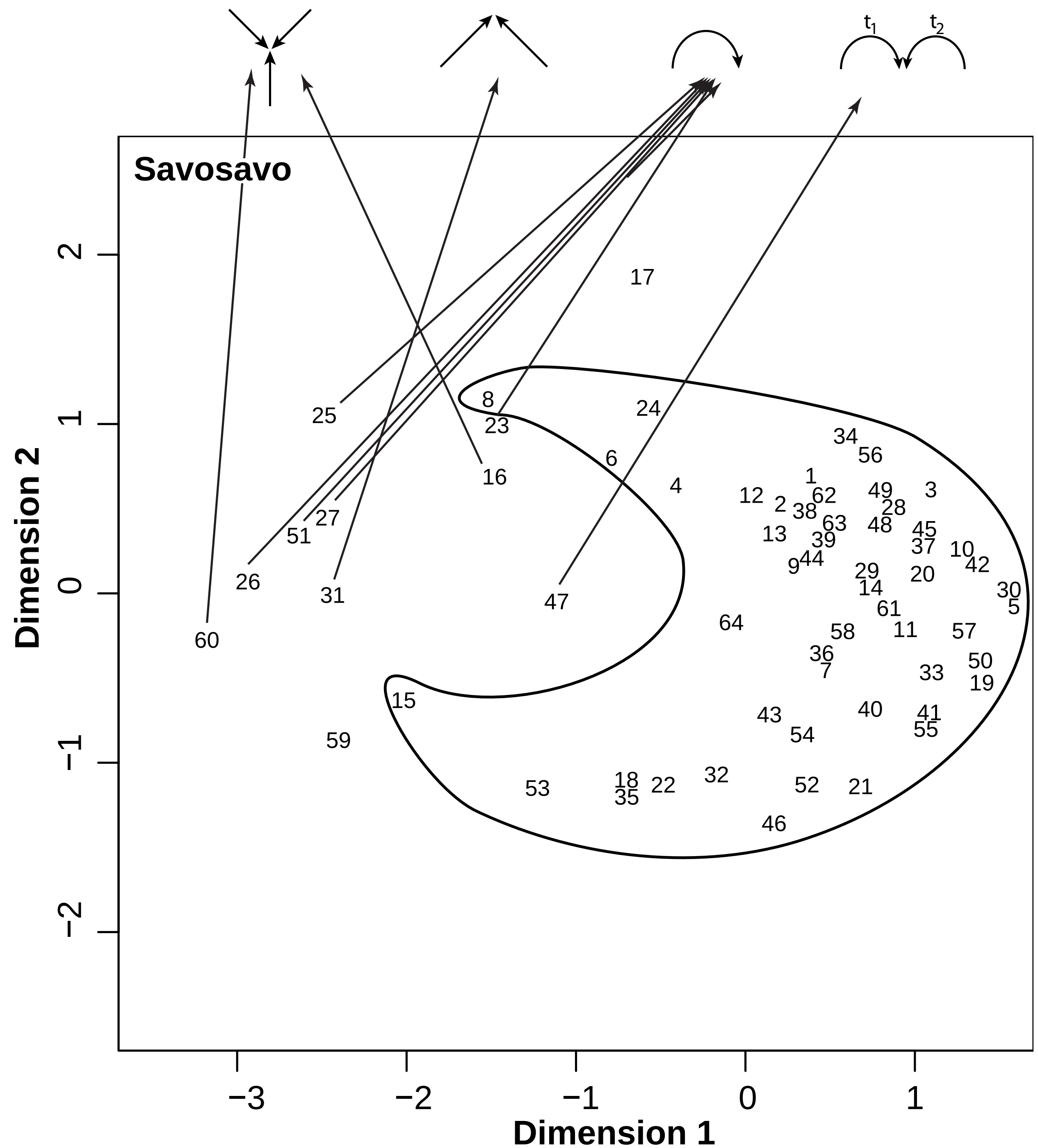


FIGURE 7 | Extensional range for general reciprocal constructions in Jahai (top left), Savosavo (top right), English (bottom left) and Khoekhoe (bottom right).

Majid et al. 2011, *Frontiers in Psychology*

Figure 4. Distribution of the three coding elements in Zenzontepec Chatino

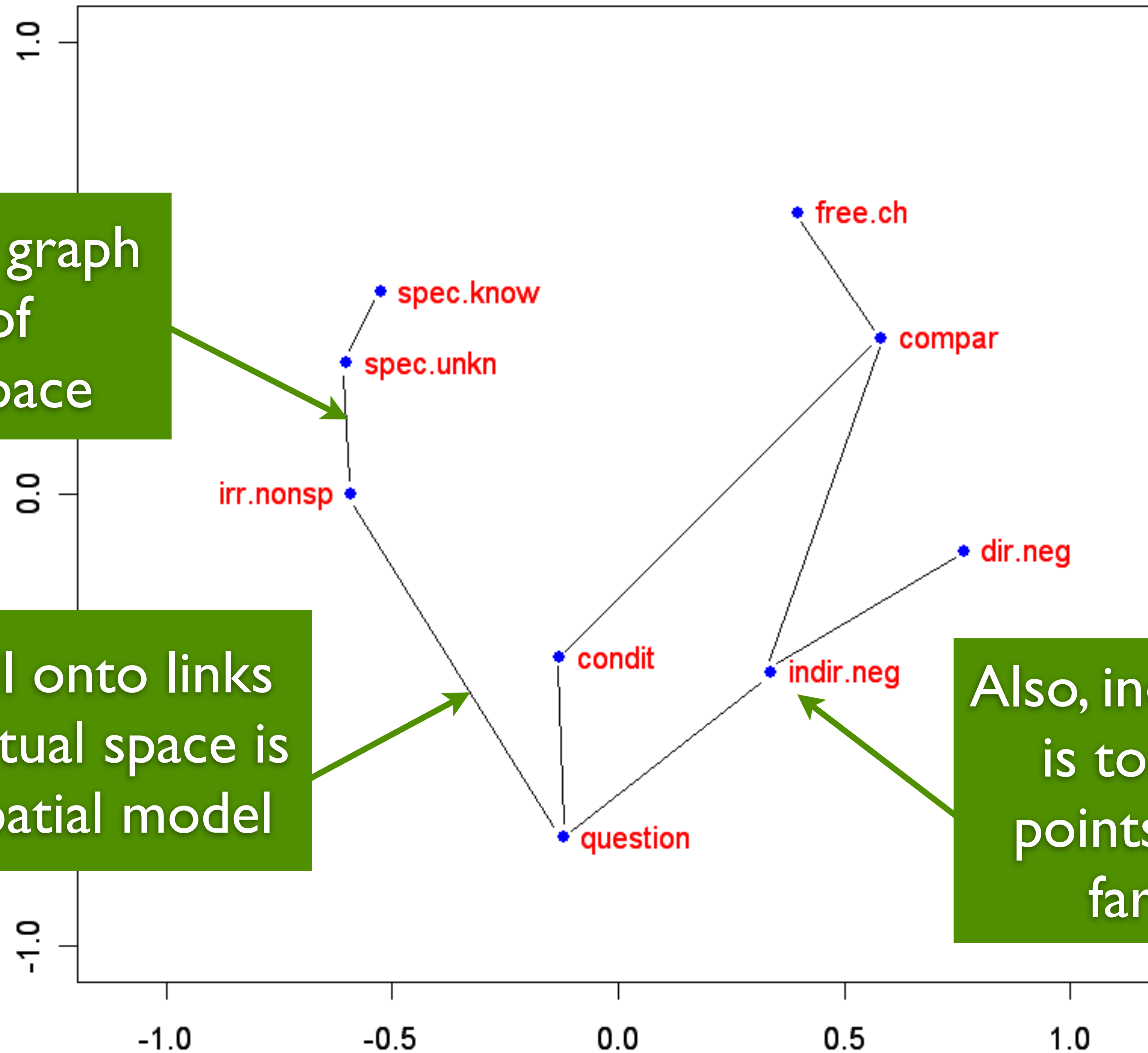
Hartmann et al. 2014, *Studies in Language*

Semantic maps are cutting lines

- A semantic map is a cutting line
- Hence, if one can position the meanings so that a straight cutting line includes all and only the meanings the form stands for*, then the conceptual space is universal
- A conceptual space is only interesting if it is low-dimensional (adding dimensions weakens the constraints on possible cutting lines)

*given the presence of noise, i.e. up to a high ***goodness of fit***

Comparing MDS and semantic maps



links represent graph structure of conceptual space

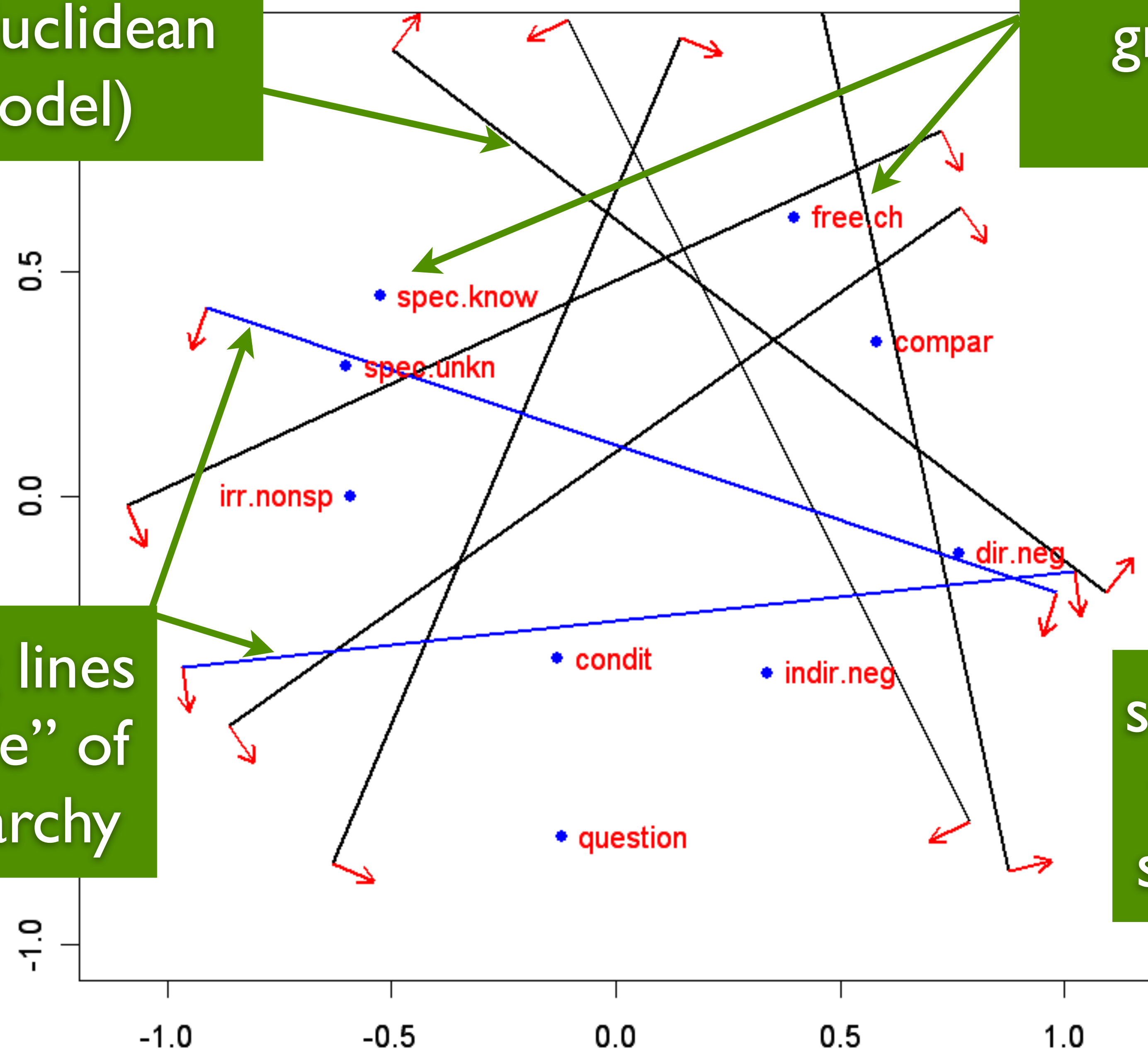
distance maps well onto links except the conceptual space is "curved" in the spatial model

Also, indirect negation is too "close" to points it should be farther from

Comparing MDS and semantic maps

cutting lines must be straight (a Euclidean spatial model)

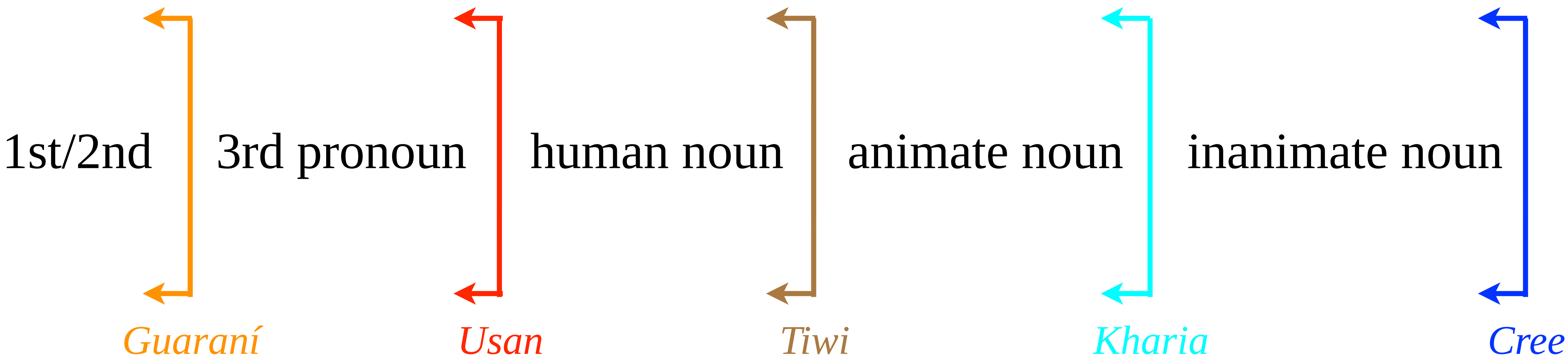
the "ends", *spec.know* and *free.ch*, are never grouped under a cutting line



but some cutting lines cut in the "middle" of the sort-of hierarchy

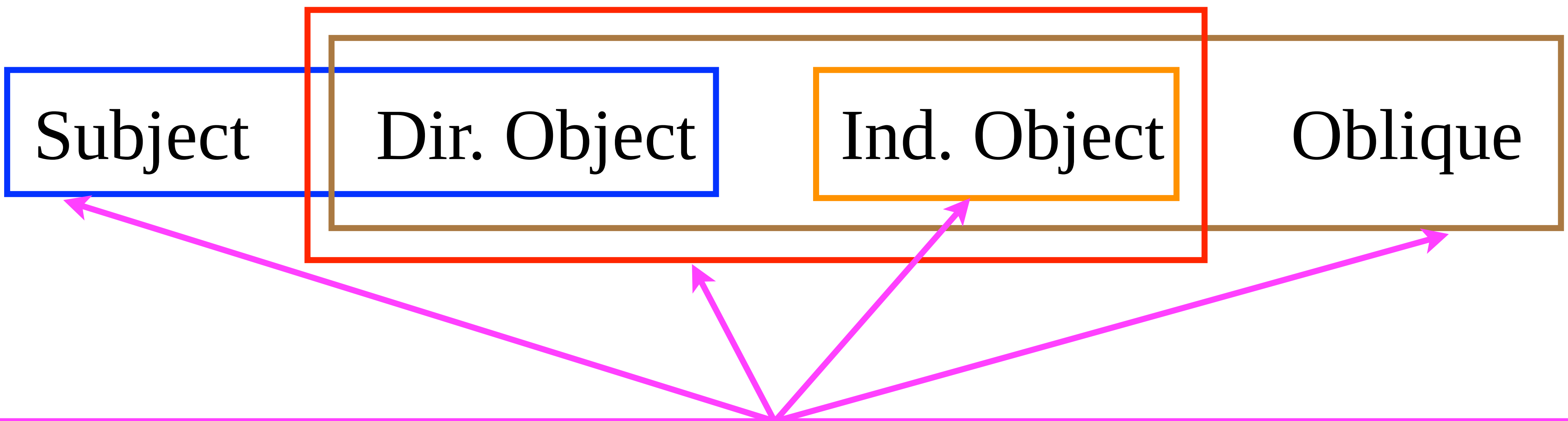
straight cutting lines in the middle mean the space must be curved

Semantic maps and MDS: a one-dimensional spatial model



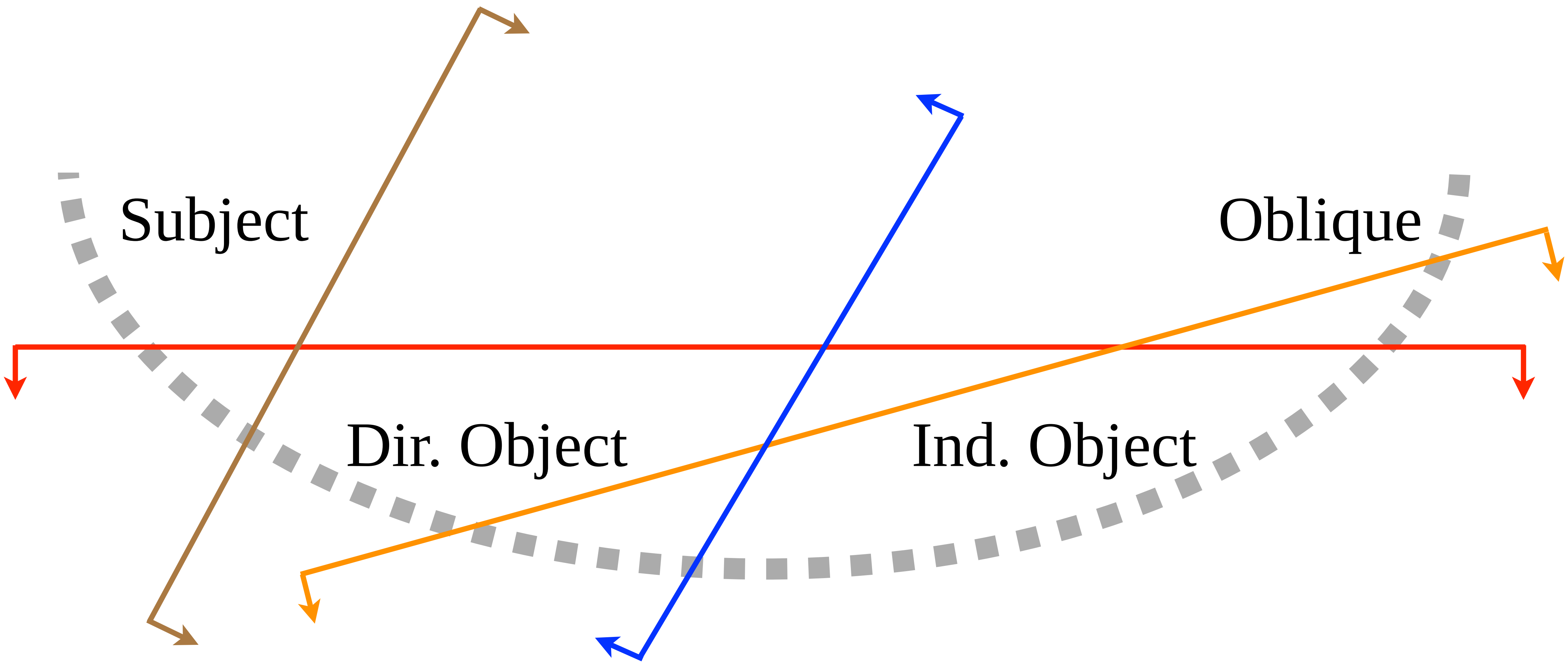
Since all semantic maps include the leftmost end of the Animacy Hierarchy, the Hierarchy can be represented in one dimension

Semantic maps and MDS: a curvilinear model



NP Accessibility Hierarchy: Keenan and Comrie argue that a relative clause construction covers a continuous segment of the Accessibility Hierarchy

Semantic maps and MDS: a curvilinear model



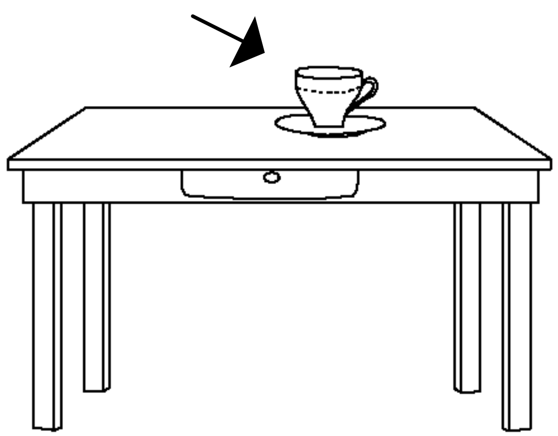
Since cutting lines must be straight, the Hierarchy must be represented as curved in an MDS spatial model

Spatial adpositions

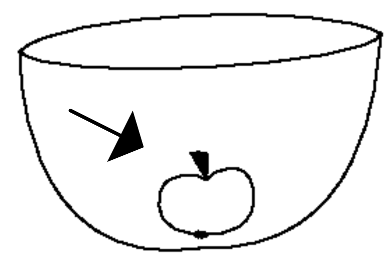
- A set of pictures of spatial situations was constructed to represent situations commonly expressed by English *on* and *in*
- The situations were described by speakers of nine diverse languages (Tiriyó, Trumai, Yukatek, Basque, Dutch, Lao, Ewe, Lavukaleve and Yélîdnye)
- Spatial adpositions only were coded
- An MDS analysis was performed on the data (refined by Croft & Poole)

(Levinson et al., *Language* vol. 79, 2003)

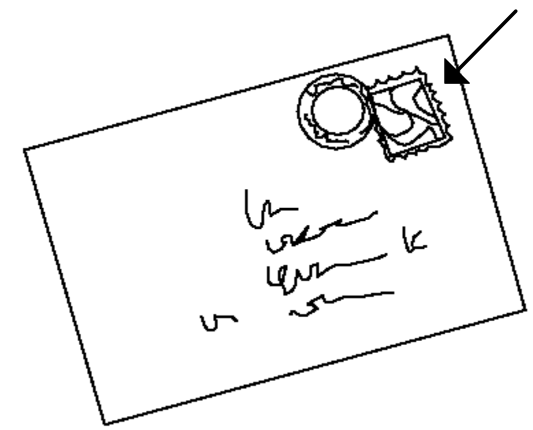
Sample stimuli (Bowerman-Pederson)



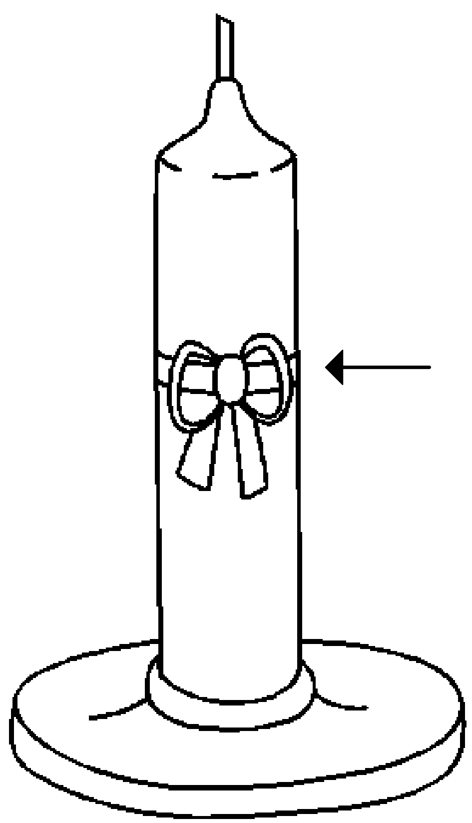
1



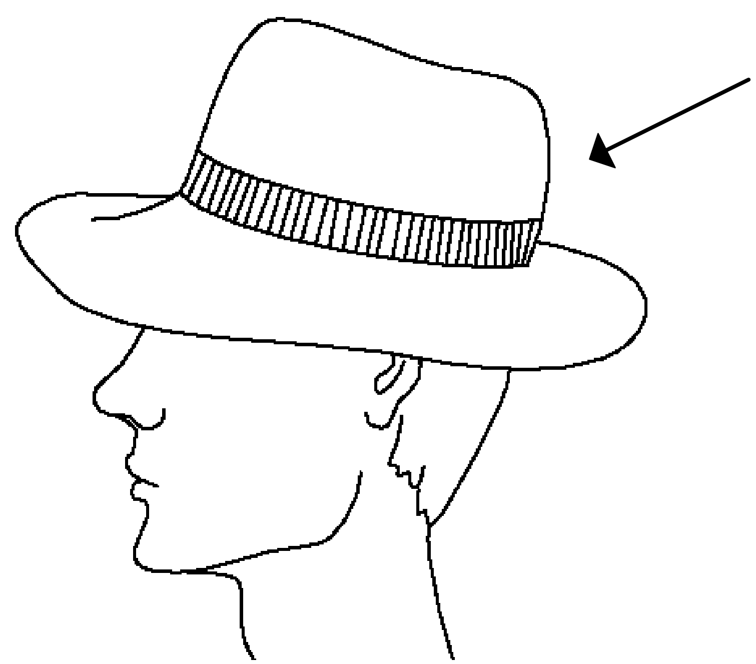
2



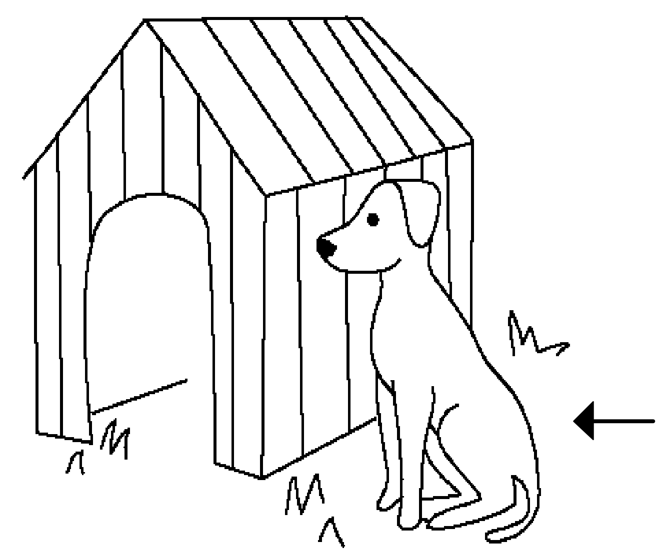
3



4



5



6

Raw data for spatial adpositions: Tiriyó, pictures 11-16

Data is very lopsided; most adpositions are used for only one or a few pictures

	tao	awë	hkao	juuwë	po	rehtë	epoe	epinë
11	0	0	1	1	0	0	0	0
12	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	1	0
14	1	1	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	1

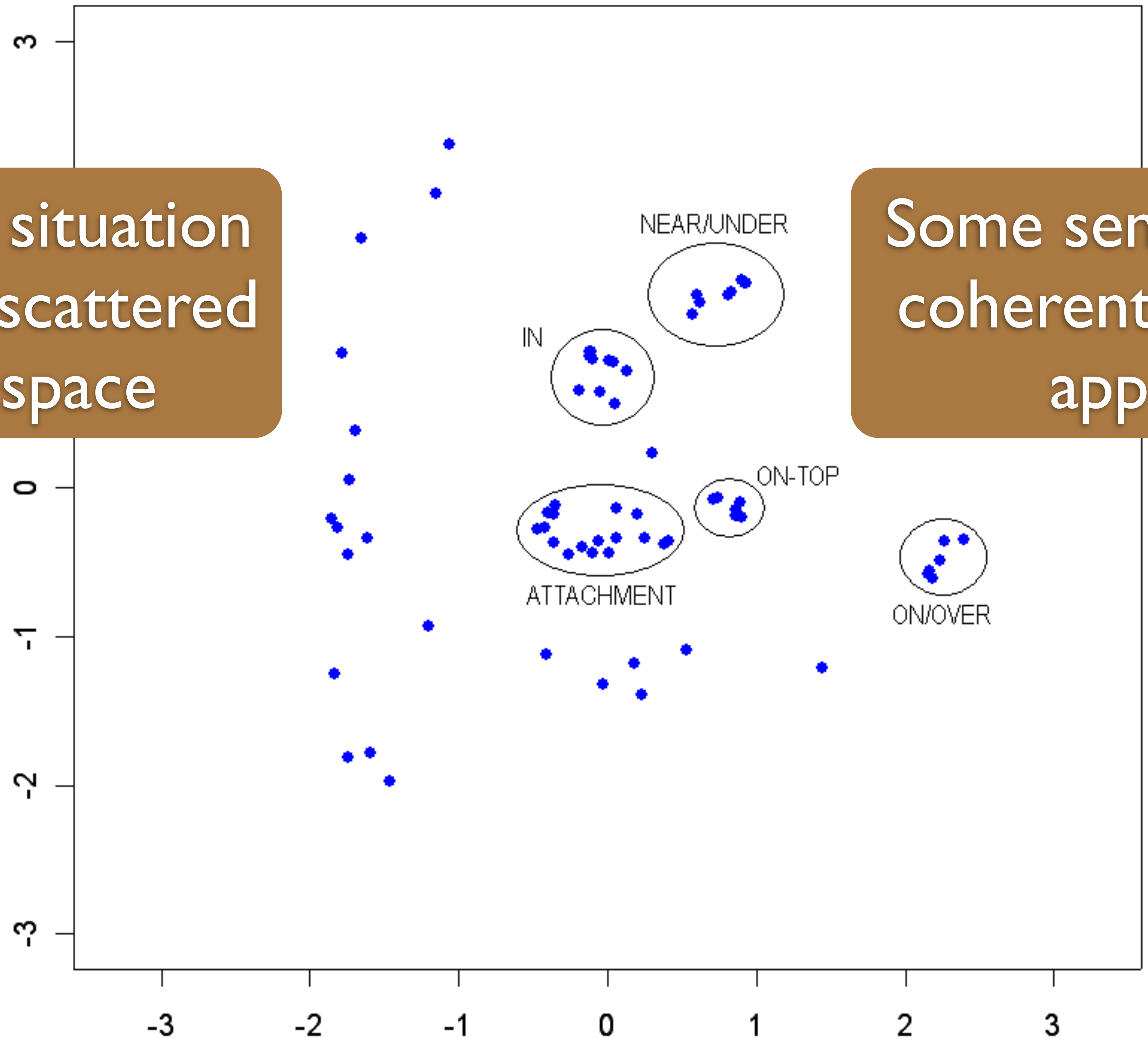
Spatial adpositions by dissimilarity

- Levinson et al. (2003) used a dissimilarity algorithm to analyze the spatial adposition data
- A dissimilarity algorithm cannot use raw crosslinguistic distributional data
- Instead, one must construct a matrix of (dis)similarity, i.e. for each pair of situation types, how often they are/aren't expressed by the same forms

Dissimilarity matrix for adposition data, pictures 1-9

Pictures	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***
1	0.000	8.890	8.610	9.000	8.590	8.890	8.740	8.330	8.890
2	8.890	0.000	8.680	8.860	8.750	8.810	8.750	8.750	8.830
3	8.610	8.680	0.000	8.720	8.680	8.810	8.540	8.470	8.740
4	9.000	8.860	8.720	0.000	8.860	8.920	8.720	8.860	8.790
5	8.590	8.750	8.680	8.860	0.000	8.710	8.600	8.450	8.730
6	8.890	8.810	8.810	8.920	8.710	0.000	8.740	8.810	8.790
7	8.740	8.750	8.540	8.720	8.600	8.740	0.000	8.600	8.640
8	8.330	8.750	8.470	8.860	8.450	8.810	8.600	0.000	8.830
9	8.890	8.830	8.740	8.790	8.730	8.790	8.640	8.830	0.000

Two-dimensional MDS model of adpositions by dissimilarity



But many situation types are scattered in the space

Some semantically coherent clusters appear

Unfolding algorithm

- The unfolding algorithm (Poole 2000, 2005) takes the distribution data directly
- It can therefore handle lopsided data better than the dissimilarity algorithm (dissimilarity compresses the range)
- The result of applying unfolding to the adposition data are much more coherent semantic clusters

Spatial adpositions: Goodness of fit

Dimensions	Classification	APRE
1	94.1%	0.300
2	95.8%	0.501
3	97.1%	0.661

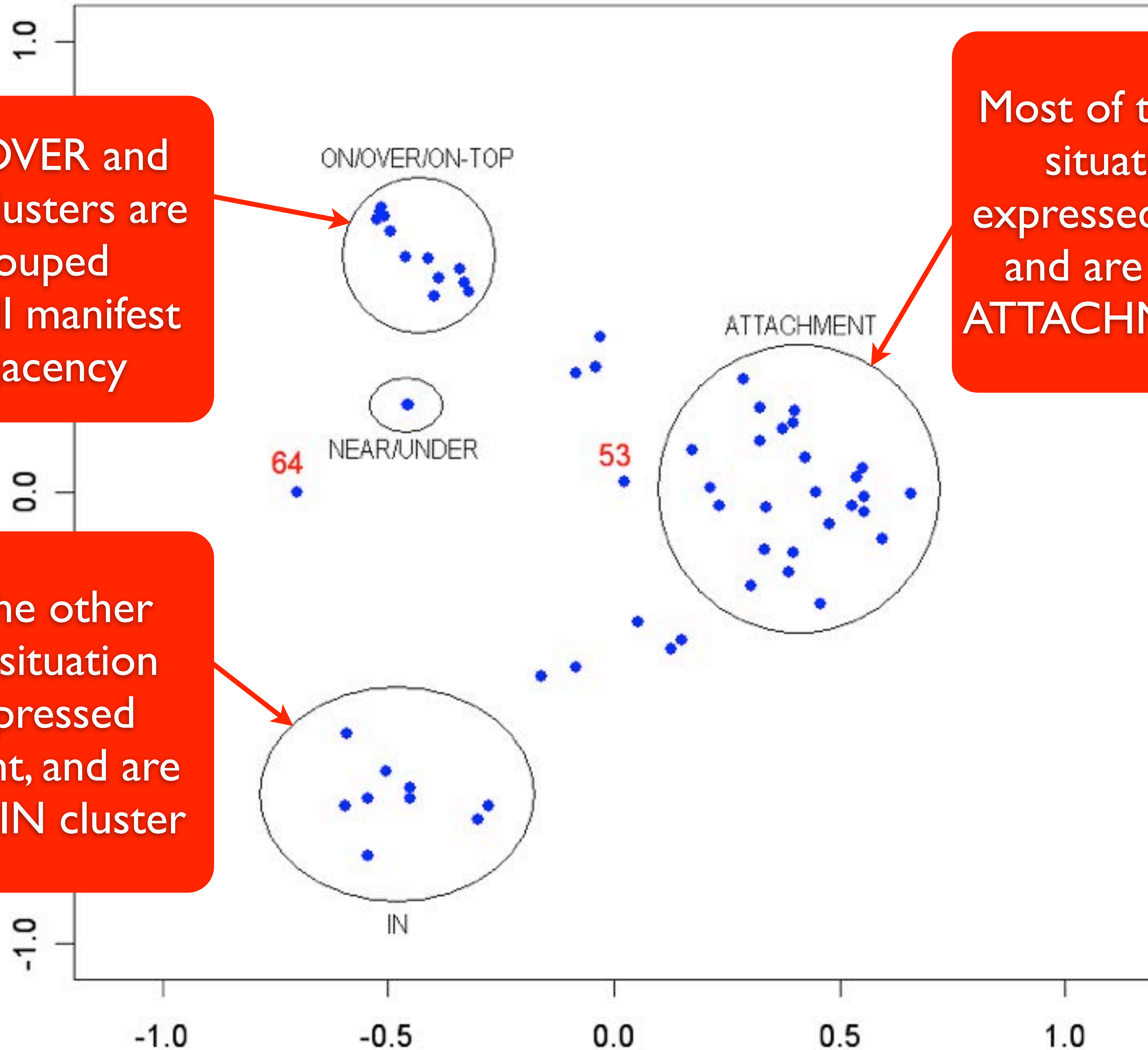
Fitness statistics indicate a two dimensional model is best

Two dimensional MDS model of adpositions by unfolding

The ON/OVER and ON-TOP clusters are now grouped together; all manifest superadjacency

Most of the scattered situation types expressed attachment, and are now in the ATTACHMENT cluster

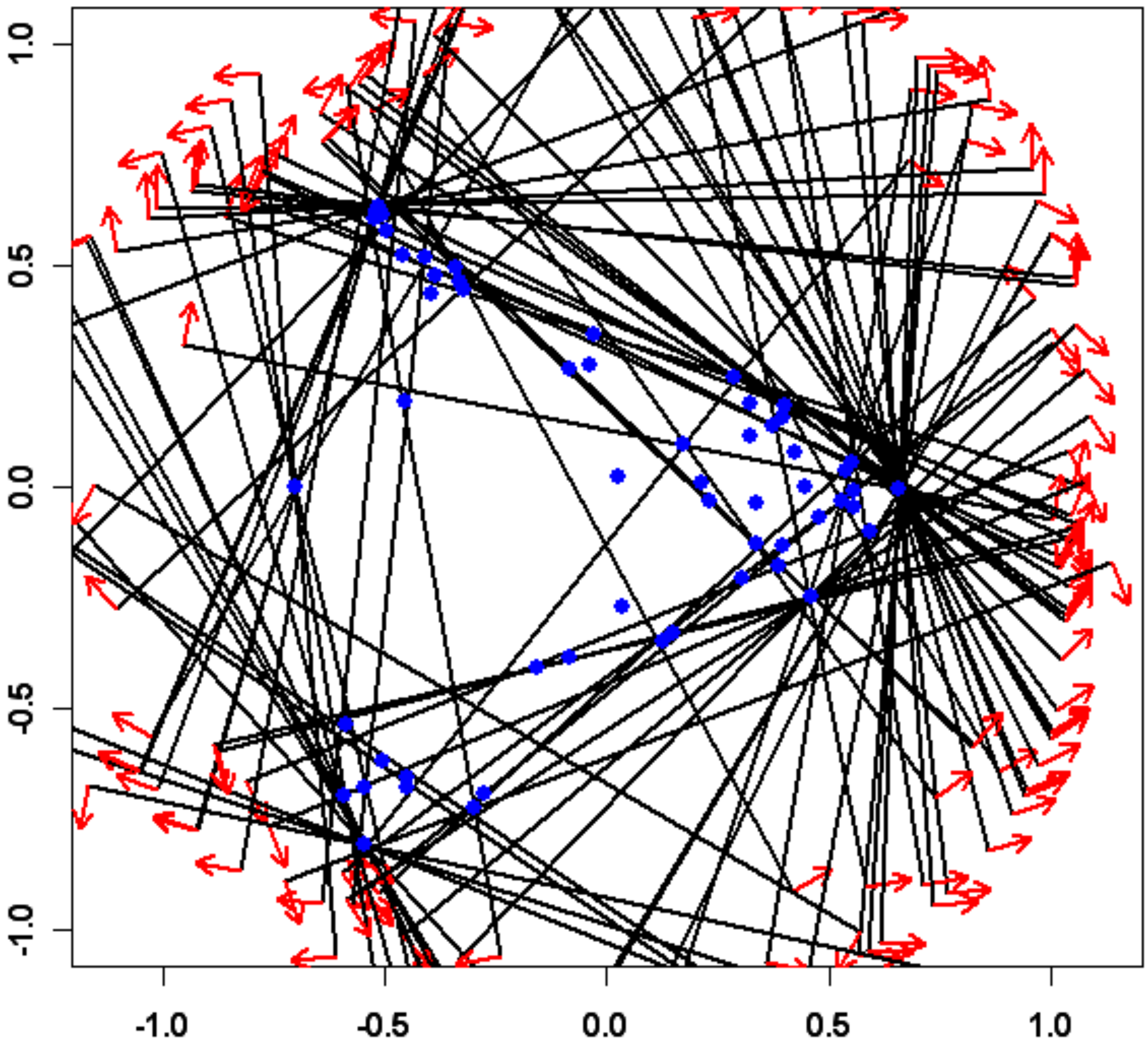
Most of the other scattered situation types expressed containment, and are now in the IN cluster



Conceptual categories (clusters)

- Does the crosslinguistic MDS analysis reveal linguistically relevant universal conceptual categories?
- What is universal are the individual situation types and their conceptual relations to each other
- That is, it is the dimensions of the spatial model that describe the linguistically relevant semantic properties

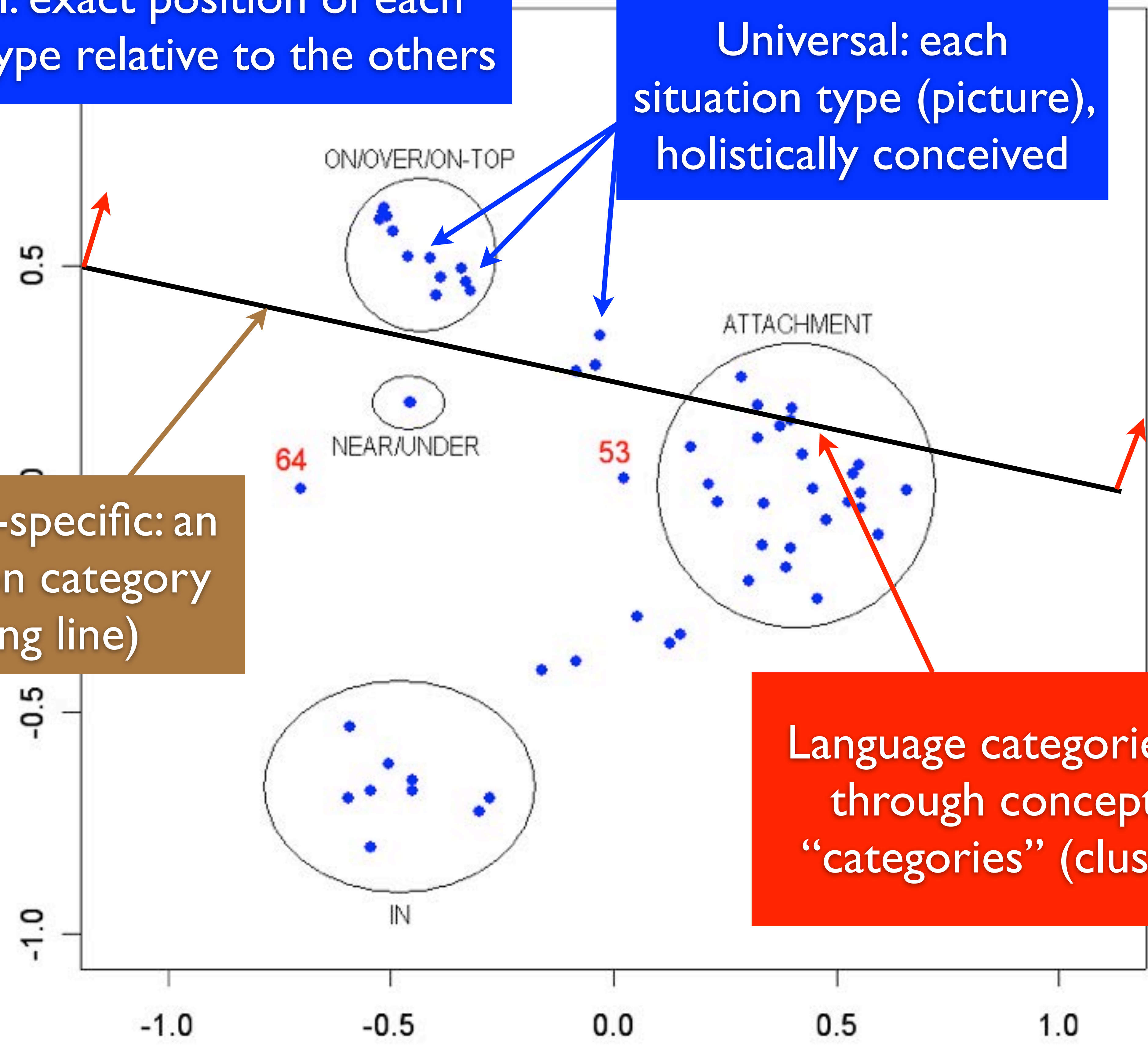
All adposition categories



Language universal and language-specific

Universal: exact position of each situation type relative to the others

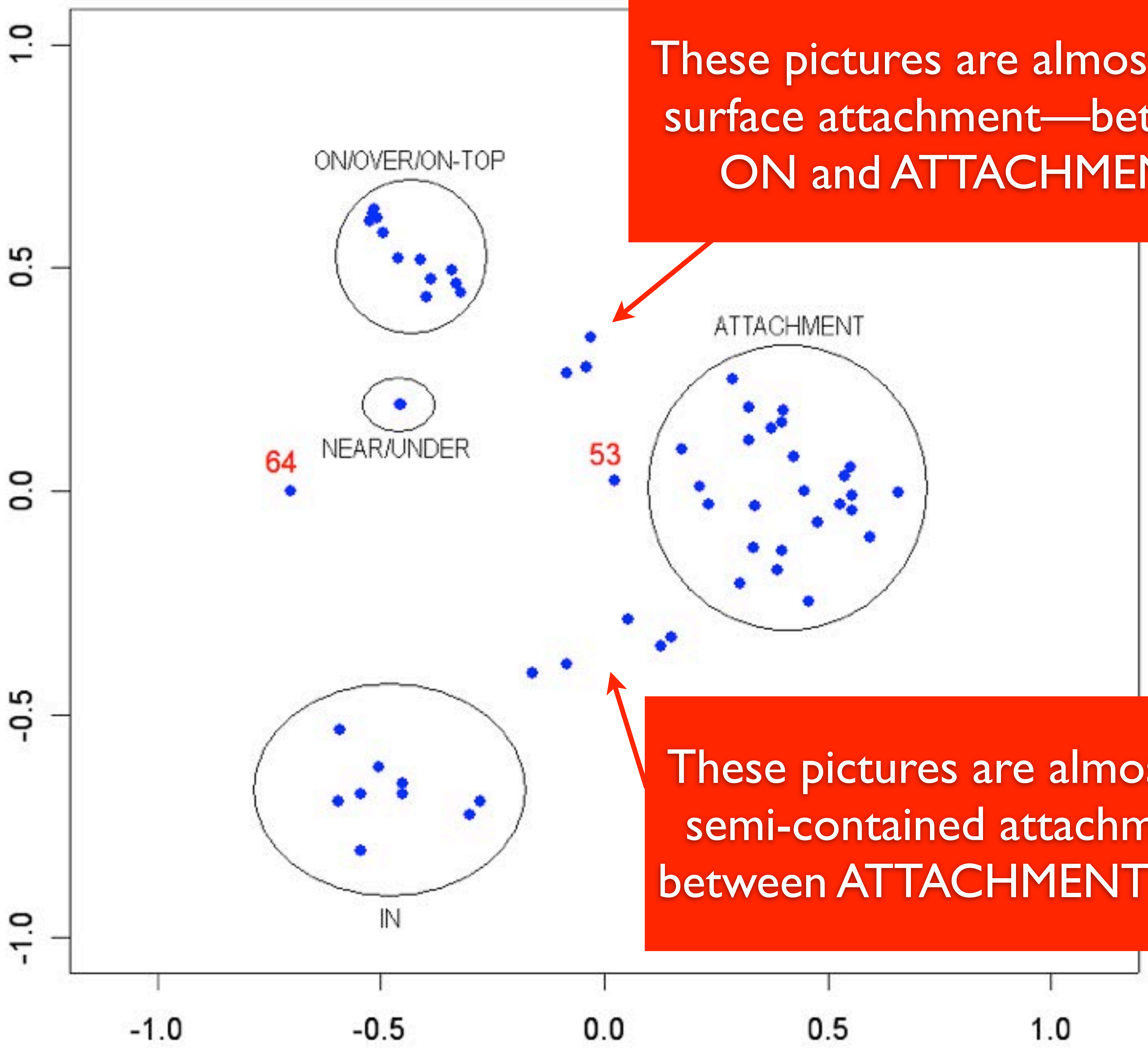
Universal: each situation type (picture), holistically conceived



Language-specific: an adposition category (cutting line)

Language categories cut through conceptual “categories” (clusters)

The importance of relations between situation types

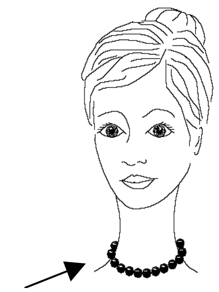


These pictures are almost all of surface attachment—between ON and ATTACHMENT

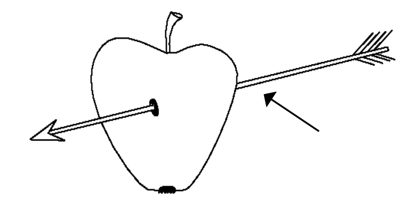
These pictures are almost all of semi-contained attachment—between ATTACHMENT and IN

Between IN and ATTACHMENT

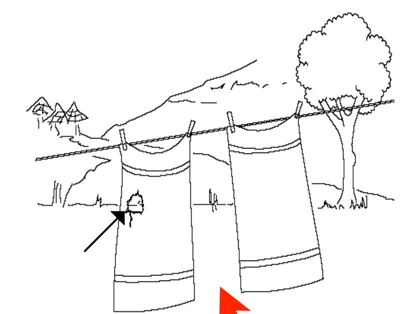
ATTACHMENT



51

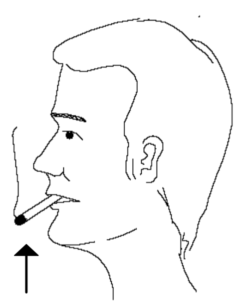


30

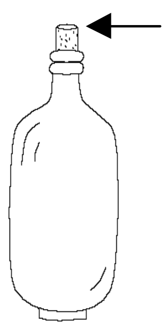


18

Closer to ATTACHMENT, the figure is or creates a hole in the ground, but can extend beyond the ground



39

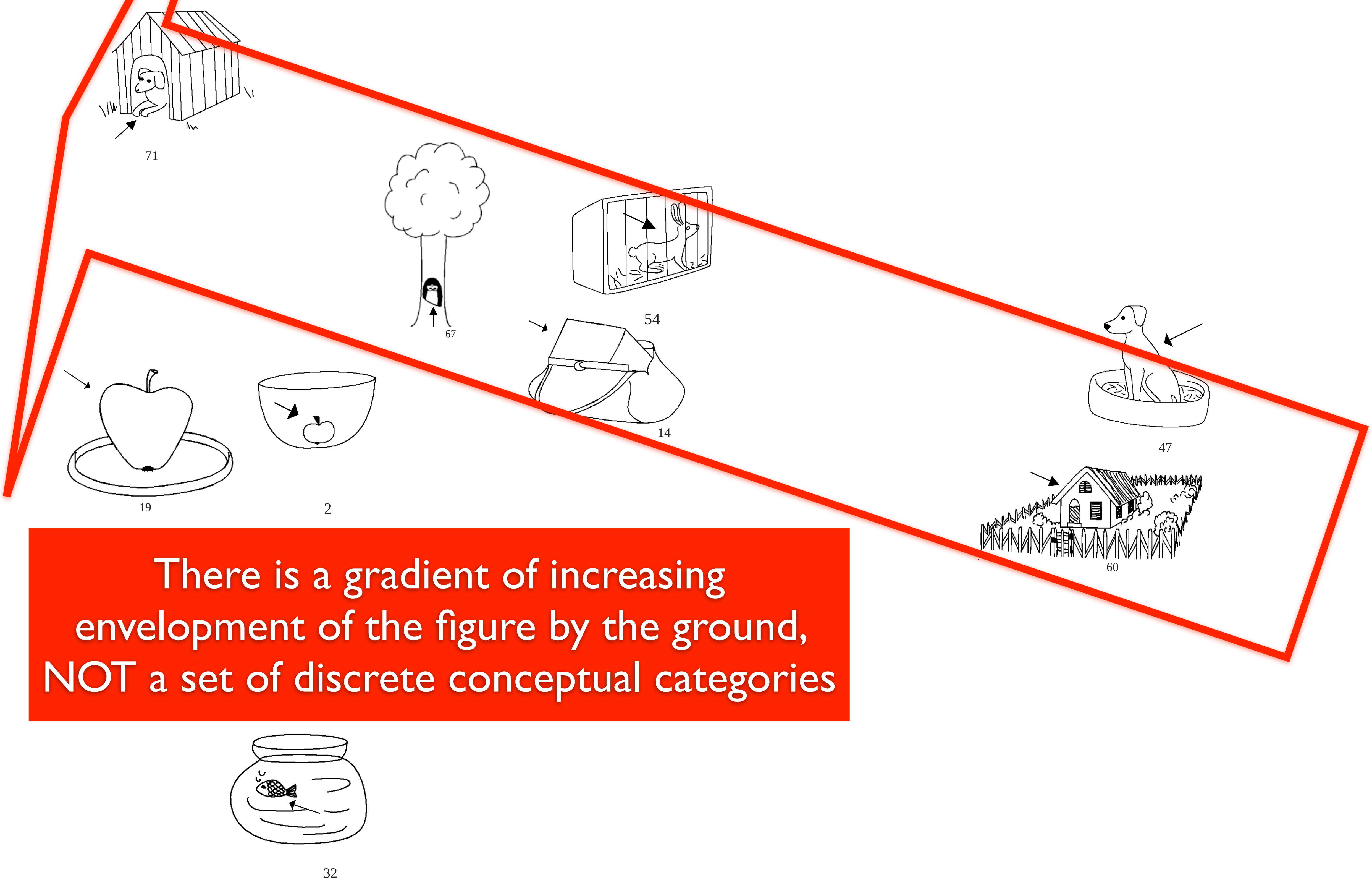


62

Somewhat closer to IN, the figure is partly contained in the ground, which has an opening, not a hole

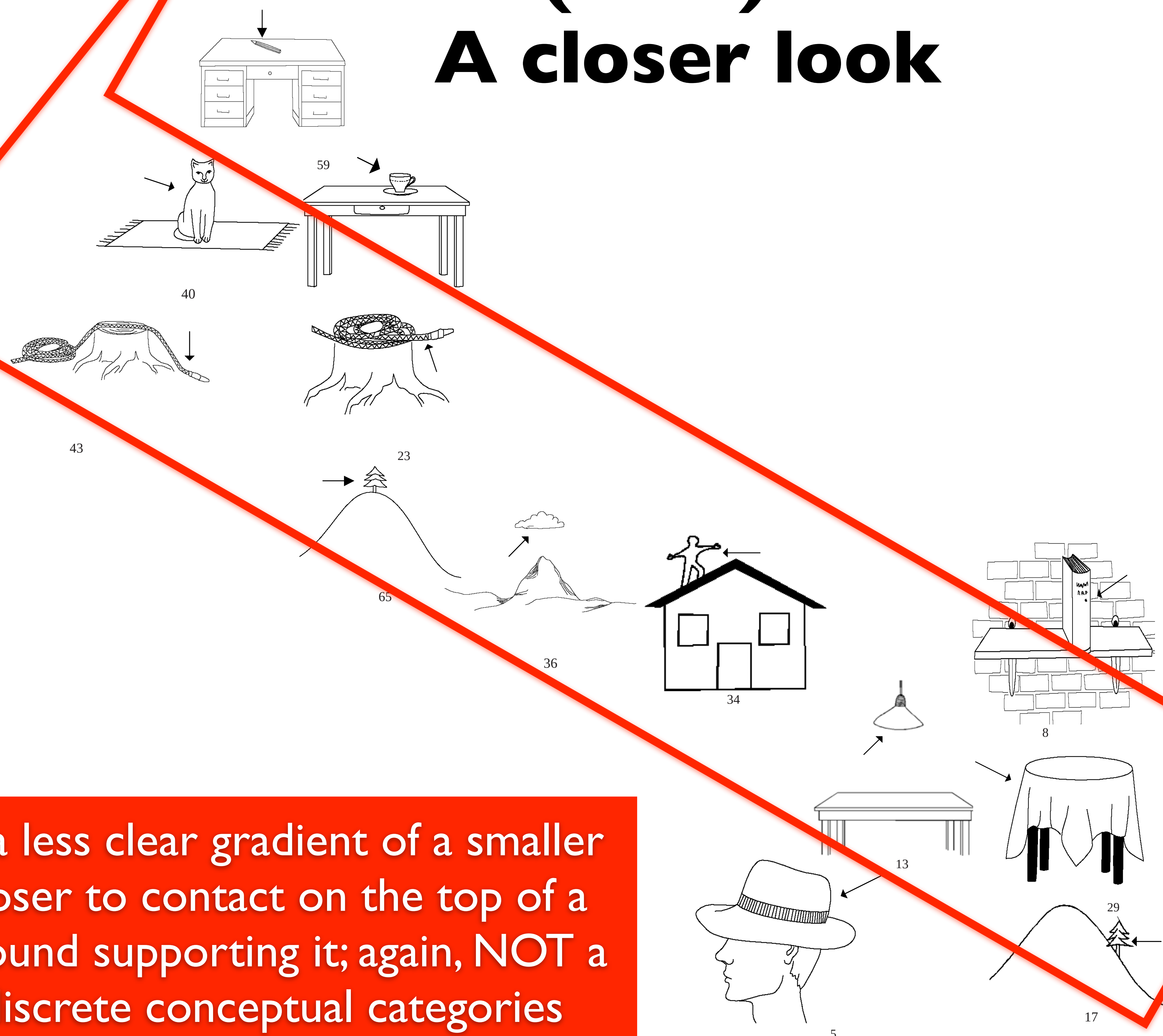
IN

The IN “cluster”: A closer look



There is a gradient of increasing envelopment of the figure by the ground, NOT a set of discrete conceptual categories

The ON (TOP) “cluster”: A closer look



There is a less clear gradient of a smaller figure closer to contact on the top of a flatter ground supporting it; again, NOT a set of discrete conceptual categories

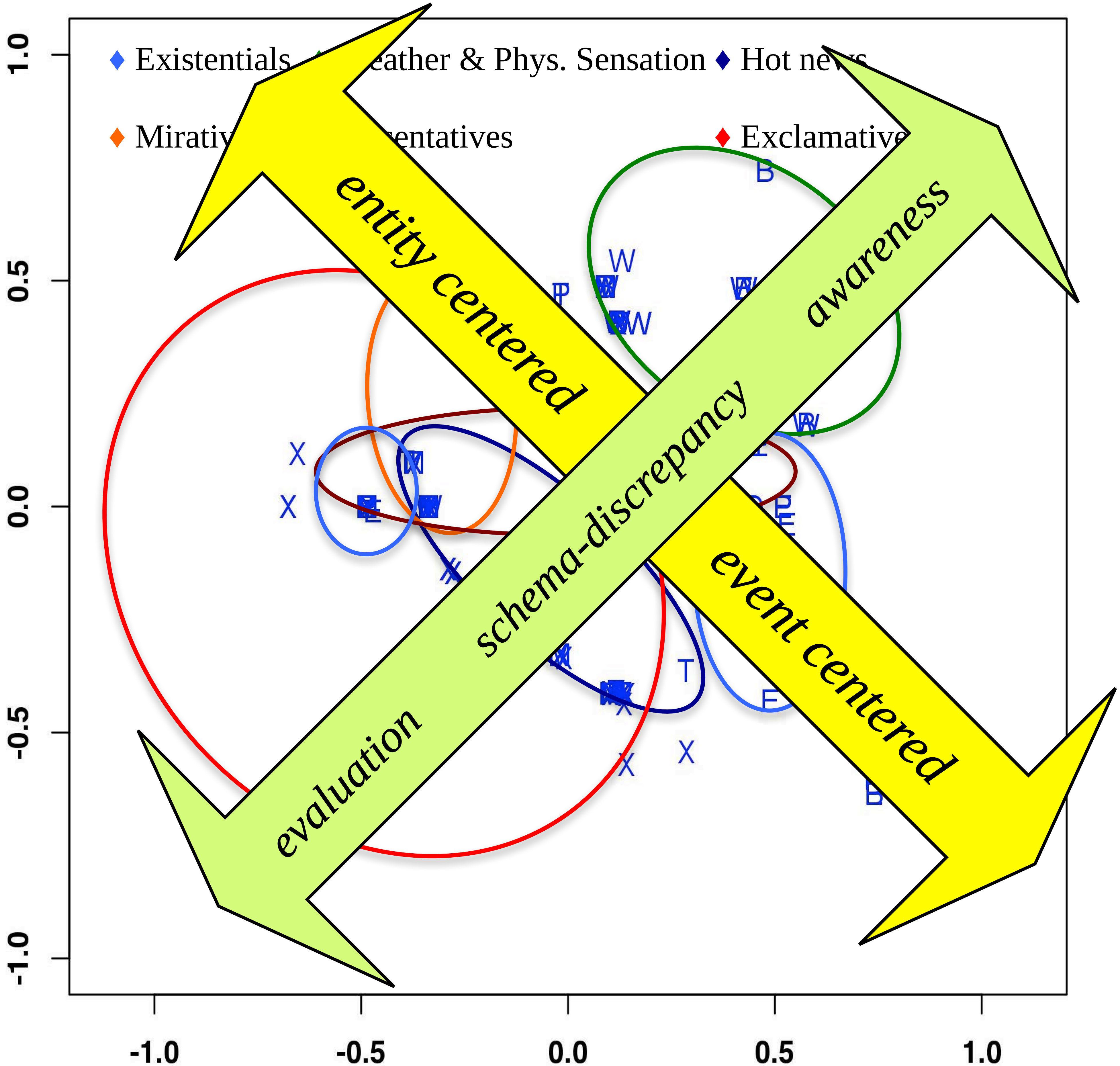
Beyond co-expression

Not just conceptual similarity

- Recall that the "semantic map" model is a model of similarity of any kind
- It doesn't have to be co-expression of meanings by a form
- For example, it could be similarity of the form of constructions in terms of certain structural traits of the constructions

MDS analysis of constructional similarity

- García Macías (2016) selected 360 constructions from 101 languages, expressing thetic meanings of different kinds (existential, presentation, hot news, weather, physical sensation), miratives and exclamatives
- He created a matrix of constructions coded with respect to shared morphosyntactic properties (e.g. defective verb, specially marked subject, overt coding of function, etc.)


Existentials
eather & Phys. Sensation
Hot ne
Mirativ
entatives
Exclamative
entity centered
event centered
awareness
schema-discrepancy
evaluation
1.0
0.5
0.0
-0.5
-1.0
-1.0
-0.5
0.0
0.5
1.0

Automated algorithms, 2: graph models

Automating "classical" semantic maps (graphs)

- "Classical" semantic maps don't have fitness metrics applied to them
- Nor do they normally provide a visualization of frequency of co-expression, like higher-dimensional MDS spaces do
- But they can, and should
- And they do, in the Regier et al. (2013) model, based on an algorithm to derive social networks from epidemiological data (Angluin et al. 2010)

The goal, and the utility function

adding the other edges contributes to capturing only one category; so the first edge has a higher utility

adding this edge contributes to capturing two categories/ semantic maps

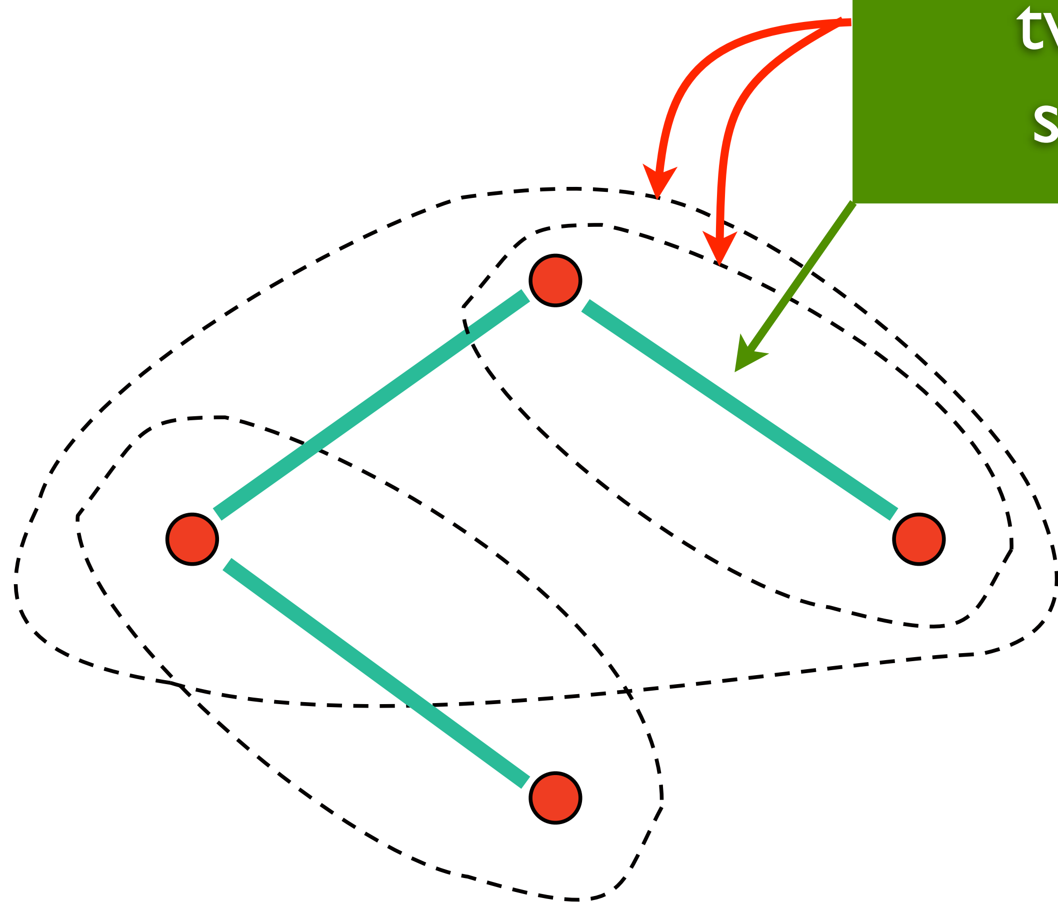
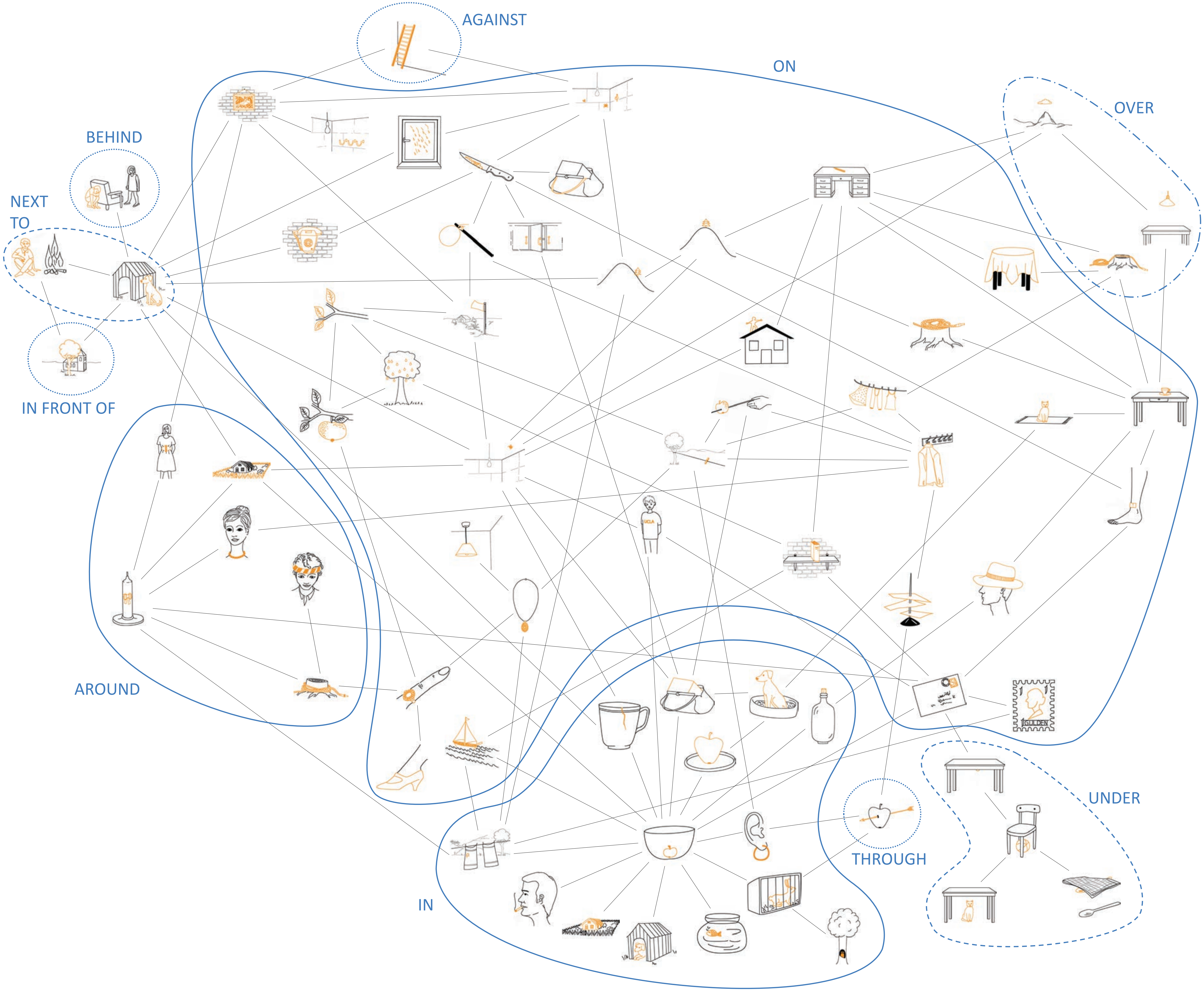


Figure 2. *Formalization of the semantic map inference problem. We are given a set of semantic functions (vertices V, shown as small circles), and groupings of these functions into language-specific categories (constraints $S_i \subseteq V$, each shown by a dashed outline). We seek the minimum set of edges E (shown as links between vertices) such that each grouping picks out a connected region of the overall graph $G = (V, E)$.*

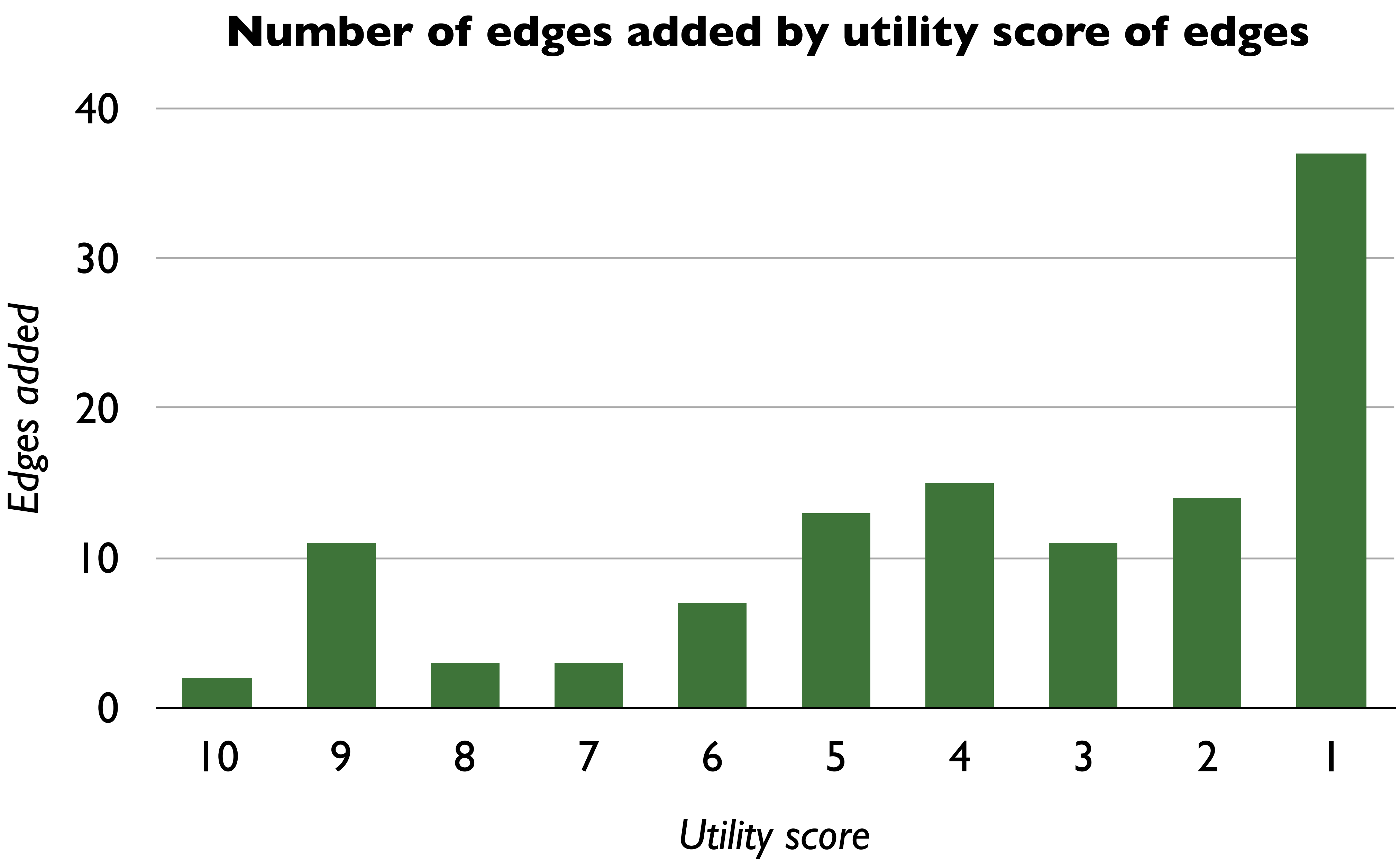

AGAINST
ON
OVER
BEHIND
NEXT
TO
IN FRONT OF
AROUND
IN
THROUGH
UNDER

The utility function: goodness of fit

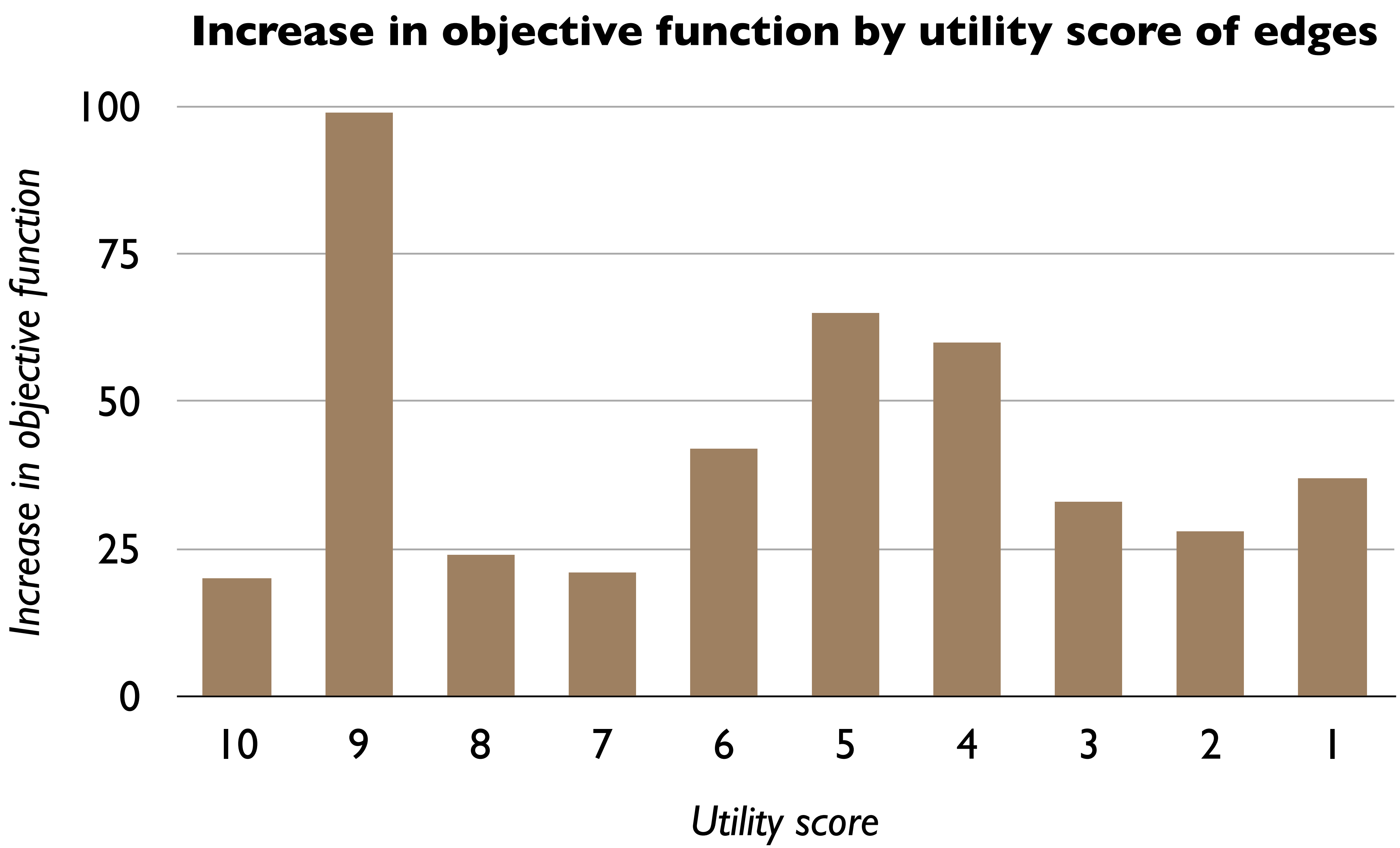
objective fn is currently -436 adding ('R2', 'R32') with score 10
objective fn is currently -426 adding ('R1', 'R40') with score 10
objective fn is currently -416 adding ('R16', 'R24') with score 9
objective fn is currently -407 adding ('R16', 'R31') with score 9
objective fn is currently -398 adding ('R2', 'R14') with score 9
objective fn is currently -389 adding ('R2', 'R19') with score 9
objective fn is currently -380 adding ('R2', 'R54') with score 9
objective fn is currently -371 adding ('R19', 'R47') with score 9
objective fn is currently -362 adding ('R1', 'R59') with score 9
objective fn is currently -353 adding ('R29', 'R59') with score 9
objective fn is currently -344 adding ('R6', 'R38') with score 9
objective fn is currently -335 adding ('R8', 'R59') with score 9
objective fn is currently -326 adding ('R59', 'R65') with score 9
objective fn is currently -317 adding ('R1', 'R23') with score 8
objective fn is currently -309 adding ('R13', 'R36') with score 8
objective fn is currently -301 adding ('R6', 'R49') with score 8
objective fn is currently -293 adding ('R2', 'R60') with score 7
objective fn is currently -286 adding ('R54', 'R67') with score 7
objective fn is currently -279 adding ('R34', 'R59') with score 7
objective fn is currently -272 adding ('R2', 'R71') with score 6
objective fn is currently -266 adding ('R12', 'R52') with score 6
objective fn is currently -260 adding ('R20', 'R56') with score 6
objective fn is currently -254 adding ('R33', 'R57') with score 6
objective fn is currently -248 adding ('R16', 'R53') with score 6
objective fn is currently -242 adding ('R3', 'R28') with score 6
objective fn is currently -236 adding ('R6', 'R64') with score 6
objective fn is currently -230 adding ('R12', 'R20') with score 5
objective fn is currently -225 adding ('R12', 'R35') with score 5
objective fn is currently -220 adding ('R27', 'R41') with score 5
objective fn is currently -215 adding ('R27', 'R45') with score 5
objective fn is currently -210 adding ('R33', 'R37') with score 5
objective fn is currently -205 adding ('R33', 'R41') with score 5
objective fn is currently -200 adding ('R41', 'R56') with score 5
objective fn is currently -195 adding ('R44', 'R52') with score 5
objective fn is currently -190 adding ('R3', 'R35') with score 5
objective fn is currently -185 adding ('R3', 'R68') with score 5
objective fn is currently -180 adding ('R30', 'R54') with score 5
objective fn is currently -175 adding ('R7', 'R56') with score 5
objective fn is currently -170 adding ('R43', 'R59') with score 5
objective fn is currently -160 adding ('R36', 'R59') with score 4
objective fn is currently -156 adding ('R9', 'R22') with score 4
objective fn is currently -152 adding ('R9', 'R33') with score 4
objective fn is currently -148 adding ('R9', 'R70') with score 4
objective fn is currently -144 adding ('R10', 'R21') with score 4
objective fn is currently -140 adding ('R10', 'R57') with score 4
objective fn is currently -136 adding ('R12', 'R25') with score 4
objective fn is currently -132 adding ('R12', 'R61') with score 4
objective fn is currently -128 adding ('R44', 'R50') with score 4
objective fn is currently -124 adding ('R57', 'R63') with score 4
objective fn is currently -120 adding ('R4', 'R55') with score 4
objective fn is currently -116 adding ('R3', 'R8') with score 4
objective fn is currently -112 adding ('R8', 'R11') with score 4
objective fn is currently -108 adding ('R48', 'R52') with score 4
objective fn is currently -104 adding ('R8', 'R45') with score 4
objective fn is currently -100 adding ('R12', 'R66') with score 3
objective fn is currently -97 adding ('R2', 'R18') with score 3
objective fn is currently -94 adding ('R2', 'R62') with score 3
objective fn is currently -91 adding ('R10', 'R55') with score 3
objective fn is currently -88 adding ('R46', 'R55') with score 3
objective fn is currently -85 adding ('R1', 'R5') with score 3
objective fn is currently -82 adding ('R52', 'R58') with score 3
objective fn is currently -79 adding ('R18', 'R28') with score 3
objective fn is currently -76 adding ('R14', 'R70') with score 3
objective fn is currently -73 adding ('R17', 'R65') with score 3
objective fn is currently -70 adding ('R33', 'R69') with score 3
objective fn is currently -67 adding ('R14', 'R47') with score 2
objective fn is currently -65 adding ('R2', 'R11') with score 2
objective fn is currently -63 adding ('R2', 'R39') with score 2
objective fn is currently -61 adding ('R4', 'R15') with score 2
objective fn is currently -59 adding ('R4', 'R42') with score 2
objective fn is currently -57 adding ('R4', 'R51') with score 2
objective fn is currently -55 adding ('R7', 'R34') with score 2
objective fn is currently -53 adding ('R17', 'R52') with score 2
objective fn is currently -51 adding ('R4', 'R18') with score 2
objective fn is currently -49 adding ('R17', 'R18') with score 2
objective fn is currently -47 adding ('R18', 'R57') with score 2
objective fn is currently -45 adding ('R22', 'R30') with score 2
objective fn is currently -43 adding ('R6', 'R7') with score 2
objective fn is currently -41 adding ('R37', 'R43') with score 2
objective fn is currently -37 adding ('R1', 'R13') with score 1
objective fn is currently -36 adding ('R9', 'R20') with score 1
objective fn is currently -35 adding ('R9', 'R37') with score 1
objective fn is currently -34 adding ('R10', 'R27') with score 1
objective fn is currently -33 adding ('R33', 'R70') with score 1
objective fn is currently -32 adding ('R41', 'R45') with score 1
objective fn is currently -31 adding ('R44', 'R56') with score 1
objective fn is currently -30 adding ('R30', 'R69') with score 1
objective fn is currently -29 adding ('R2', 'R26') with score 1
objective fn is currently -28 adding ('R2', 'R69') with score 1
objective fn is currently -27 adding ('R11', 'R18') with score 1
objective fn is currently -26 adding ('R1', 'R35') with score 1
objective fn is currently -25 adding ('R7', 'R65') with score 1
objective fn is currently -24 adding ('R7', 'R68') with score 1
objective fn is currently -23 adding ('R23', 'R65') with score 1
objective fn is currently -22 adding ('R29', 'R43') with score 1
objective fn is currently -21 adding ('R38', 'R49') with score 1
objective fn is currently -20 adding ('R2', 'R68') with score 1
objective fn is currently -19 adding ('R7', 'R15') with score 1
objective fn is currently -18 adding ('R7', 'R26') with score 1
objective fn is currently -17 adding ('R44', 'R58') with score 1
objective fn is currently -16 adding ('R6', 'R17') with score 1
objective fn is currently -15 adding ('R14', 'R61') with score 1
objective fn is currently -14 adding ('R6', 'R26') with score 1
objective fn is currently -13 adding ('R2', 'R3') with score 1
objective fn is currently -12 adding ('R6', 'R48') with score 1
objective fn is currently -11 adding ('R6', 'R25') with score 1
objective fn is currently -10 adding ('R1', 'R43') with score 1
objective fn is currently -9 adding ('R7', 'R14') with score 1
objective fn is currently -8 adding ('R7', 'R36') with score 1
objective fn is currently -7 adding ('R2', 'R15') with score 1
objective fn is currently -6 adding ('R9', 'R51') with score 1
objective fn is currently -5 adding ('R6', 'R15') with score 1
objective fn is currently -4 adding ('R3', 'R4') with score 1
objective fn is currently -3 adding ('R3', 'R53') with score 1
objective fn is currently -2 adding ('R42', 'R44') with score 1
objective fn is currently -1 adding ('R6', 'R44') with score 1

- You could prune edges with the lowest utility value(s)

The utility function: goodness of fit



The utility function: goodness of fit



The utility function: visualizing frequency

```
objective fn is currently -436 adding ('R2', 'R32') with score 10
objective fn is currently -426 adding ('R1', 'R40') with score 10
objective fn is currently -416 adding ('R16', 'R24') with score 9
objective fn is currently -407 adding ('R16', 'R31') with score 9
objective fn is currently -398 adding ('R2', 'R14') with score 9
objective fn is currently -389 adding ('R2', 'R19') with score 9
objective fn is currently -380 adding ('R2', 'R54') with score 9
objective fn is currently -371 adding ('R19', 'R47') with score 9
objective fn is currently -362 adding ('R1', 'R59') with score 9
objective fn is currently -353 adding ('R29', 'R59') with score 9
objective fn is currently -344 adding ('R6', 'R38') with score 9
objective fn is currently -335 adding ('R8', 'R59') with score 9
objective fn is currently -326 adding ('R59', 'R65') with score 9
objective fn is currently -317 adding ('R1', 'R23') with score 8
objective fn is currently -309 adding ('R13', 'R36') with score 8
objective fn is currently -301 adding ('R6', 'R49') with score 8
objective fn is currently -293 adding ('R2', 'R60') with score 7
objective fn is currently -286 adding ('R54', 'R67') with score 7
objective fn is currently -279 adding ('R34', 'R59') with score 7
objective fn is currently -272 adding ('R2', 'R71') with score 6
objective fn is currently -266 adding ('R12', 'R52') with score 6
objective fn is currently -260 adding ('R20', 'R56') with score 6
objective fn is currently -254 adding ('R33', 'R57') with score 6
objective fn is currently -248 adding ('R16', 'R53') with score 6
objective fn is currently -242 adding ('R3', 'R28') with score 6
objective fn is currently -236 adding ('R6', 'R64') with score 6
objective fn is currently -230 adding ('R12', 'R20') with score 5
objective fn is currently -225 adding ('R12', 'R35') with score 5
objective fn is currently -220 adding ('R27', 'R41') with score 5
objective fn is currently -215 adding ('R27', 'R45') with score 5
objective fn is currently -210 adding ('R33', 'R37') with score 5
objective fn is currently -205 adding ('R33', 'R41') with score 5
objective fn is currently -200 adding ('R41', 'R56') with score 5
objective fn is currently -195 adding ('R44', 'R52') with score 5
objective fn is currently -190 adding ('R3', 'R35') with score 5
objective fn is currently -185 adding ('R3', 'R68') with score 5
objective fn is currently -180 adding ('R30', 'R54') with score 5
objective fn is currently -175 adding ('R7', 'R56') with score 5
objective fn is currently -170 adding ('R43', 'R59') with score 5
objective fn is currently -160 adding ('R36', 'R59') with score 4
objective fn is currently -156 adding ('R9', 'R22') with score 4
objective fn is currently -152 adding ('R9', 'R33') with score 4
objective fn is currently -148 adding ('R9', 'R70') with score 4
objective fn is currently -144 adding ('R10', 'R21') with score 4
objective fn is currently -140 adding ('R10', 'R57') with score 4
objective fn is currently -136 adding ('R12', 'R25') with score 4
objective fn is currently -132 adding ('R12', 'R61') with score 4
objective fn is currently -128 adding ('R44', 'R50') with score 4
objective fn is currently -124 adding ('R57', 'R63') with score 4
objective fn is currently -120 adding ('R4', 'R55') with score 4
objective fn is currently -116 adding ('R3', 'R8') with score 4
objective fn is currently -112 adding ('R8', 'R11') with score 4
objective fn is currently -108 adding ('R48', 'R52') with score 4
objective fn is currently -104 adding ('R8', 'R45') with score 4
objective fn is currently -100 adding ('R12', 'R66') with score 3
objective fn is currently -97 adding ('R2', 'R18') with score 3
objective fn is currently -94 adding ('R2', 'R62') with score 3
objective fn is currently -91 adding ('R10', 'R55') with score 3
objective fn is currently -88 adding ('R46', 'R55') with score 3
objective fn is currently -85 adding ('R1', 'R5') with score 3
objective fn is currently -82 adding ('R52', 'R58') with score 3
objective fn is currently -79 adding ('R18', 'R28') with score 3
objective fn is currently -76 adding ('R14', 'R70') with score 3
objective fn is currently -73 adding ('R17', 'R65') with score 3
objective fn is currently -70 adding ('R33', 'R69') with score 3
objective fn is currently -67 adding ('R14', 'R47') with score 2
objective fn is currently -65 adding ('R2', 'R11') with score 2
objective fn is currently -63 adding ('R2', 'R39') with score 2
objective fn is currently -61 adding ('R4', 'R15') with score 2
objective fn is currently -59 adding ('R4', 'R42') with score 2
objective fn is currently -57 adding ('R4', 'R51') with score 2
objective fn is currently -55 adding ('R7', 'R34') with score 2
objective fn is currently -53 adding ('R17', 'R52') with score 2
objective fn is currently -51 adding ('R4', 'R18') with score 2
objective fn is currently -49 adding ('R17', 'R18') with score 2
objective fn is currently -47 adding ('R18', 'R57') with score 2
objective fn is currently -45 adding ('R22', 'R30') with score 2
objective fn is currently -43 adding ('R6', 'R7') with score 2
objective fn is currently -41 adding ('R37', 'R43') with score 2
objective fn is currently -37 adding ('R1', 'R13') with score 1
objective fn is currently -36 adding ('R9', 'R20') with score 1
objective fn is currently -35 adding ('R9', 'R37') with score 1
objective fn is currently -34 adding ('R10', 'R27') with score 1
objective fn is currently -33 adding ('R33', 'R70') with score 1
objective fn is currently -32 adding ('R41', 'R45') with score 1
objective fn is currently -31 adding ('R44', 'R56') with score 1
objective fn is currently -30 adding ('R30', 'R69') with score 1
objective fn is currently -29 adding ('R2', 'R26') with score 1
objective fn is currently -28 adding ('R2', 'R69') with score 1
objective fn is currently -27 adding ('R11', 'R18') with score 1
objective fn is currently -26 adding ('R1', 'R35') with score 1
objective fn is currently -25 adding ('R7', 'R65') with score 1
objective fn is currently -24 adding ('R7', 'R68') with score 1
objective fn is currently -23 adding ('R23', 'R65') with score 1
objective fn is currently -22 adding ('R29', 'R43') with score 1
objective fn is currently -21 adding ('R38', 'R49') with score 1
objective fn is currently -20 adding ('R2', 'R68') with score 1
objective fn is currently -19 adding ('R7', 'R15') with score 1
objective fn is currently -18 adding ('R7', 'R26') with score 1
objective fn is currently -17 adding ('R44', 'R58') with score 1
objective fn is currently -16 adding ('R6', 'R17') with score 1
objective fn is currently -15 adding ('R14', 'R61') with score 1
objective fn is currently -14 adding ('R6', 'R26') with score 1
objective fn is currently -13 adding ('R2', 'R3') with score 1
objective fn is currently -12 adding ('R6', 'R48') with score 1
objective fn is currently -11 adding ('R6', 'R25') with score 1
objective fn is currently -10 adding ('R1', 'R43') with score 1
objective fn is currently -9 adding ('R7', 'R14') with score 1
objective fn is currently -8 adding ('R7', 'R36') with score 1
objective fn is currently -7 adding ('R2', 'R15') with score 1
objective fn is currently -6 adding ('R9', 'R51') with score 1
objective fn is currently -5 adding ('R6', 'R15') with score 1
objective fn is currently -4 adding ('R3', 'R4') with score 1
objective fn is currently -3 adding ('R3', 'R53') with score 1
objective fn is currently -2 adding ('R42', 'R44') with score 1
objective fn is currently -1 adding ('R6', 'R44') with score 1
```

- Edges could have thickness based on their utility score

Which models?

Graph models and Euclidean models

- The "classical" graph model and the MDS Euclidean model are both legitimate visualizations
- The graph model is more useful when there are a small number of nodes (concepts) being compared
- The Euclidean model is more useful when there are a medium to large number of nodes being compared

Graph models and Euclidean models

- The graph model assumes a discrete underlying conceptual space, while he Euclidean spatial model represents a continuous underlying conceptual space
- The graph model cannot be interpreted in terms of "dimensions"; the two-dimensional visual display of the graph is just one of convenience (e.g. minimizing the crossing of edges)
- The Euclidean model's dimensions can (and should) be interpreted

MDS and similar models

- MDS is one of a family of multivariate analyses
- It is an unsupervised distance model
- Unsupervised = the categories or groupings are not specified in advance
- Distance = represents similarity directly. In this respect, it differs from eigenanalysis methods (principal components analysis, factor analysis, correspondence analysis)

MDS and similar models

- Eigenanalysis converts the matrix of data to another matrix of the same dimensionality such that
  - ✦ each dimension is uncorrelated with every other dimension
  - ✦ the first dimension accounts for the most variance in the data, the second for the next most variance, and so on
- This has consequences for interpreting the typically two-dimensional visualizations

MDS and similar models

- An MDS spatial model represents all the variance in the data in the displayed dimensions, while an eigenanalysis represents only a subset of the variance
- In an MDS spatial model is a true Euclidean spatial representation; an eigenanalysis is a visual representation of the variance in the first two principal components

MDS and similar models

- In an MDS spatial model, all distances are interpretable. Hence the analysis is invariant under translation and rotation.
- In an eigenanalysis, each dimension must be interpreted separately:

It is customary to summarize the row and column coordinates in a single plot. However, it is important to remember that in such plots, you can only interpret the distances between row points, and the distances between column points, but not the distances between row points and column points.
(http://www.statsoft.com/Textbook/Correspondence-Analysis/, accessed 7 June 2018)

Resources

- Multidimensional scaling: https://github.com/jaytimm/MDS_for_Linguists *(NB: the code at my website now gives wrong results; the user guide that is there is still mostly good but will be updated)*
- Graph structure: http://lclab.berkeley.edu/regier/semantic-maps/