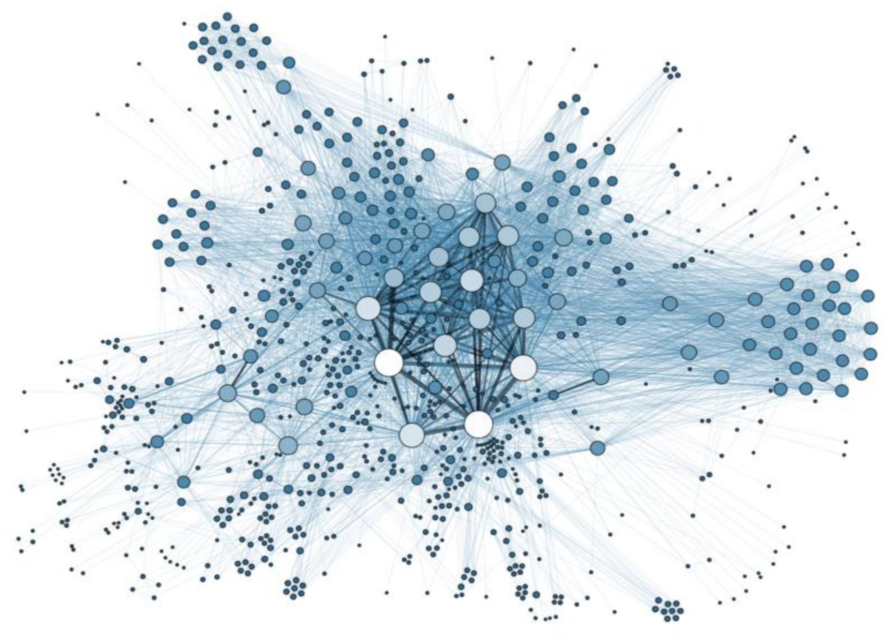
Polysemy in the lexicon within and across languages: Theory and methods

Thanasis Georgakopoulos
(National Research University, Higher School of Economics, Moscow)



28.-30.08.2019

Demines Summer School
“Language Comparison and Typology: German and the Mediterranean languages”

Course description

The main focus of the course will be on polysemy, namely on the phenomenon whereby a linguistic unit exhibits multiple distinct yet related meanings. I will open up by sketching the history of polysemy in the paradigm of cognitive linguistics focusing on lexical semantics. Then I will discuss some of the issues that have sparked debate and critical reflection among researchers. For example, I will try to provide answers for the following questions:

a) Based on which criteria one is able to identify the prototypical meaning of a word? Is it possible to constrain the arbitrariness in the choice of this primary meaning?
b) How to decide when two usages of a word are clearly different or similar enough to be considered as a single meaning?

Within the cognitive-linguistics tradition, there have been several solutions offered for these problematic issues, both top-down and bottom-up. The former are theory-driven which start with a predetermined set of criteria and the latter are data-driven, in which corpus data assist the task of establishing the prototype and of distinguishing between meanings. In this course, I will compare the two approaches and discuss their advantages and shortcomings.

Course description

The second part of the course will be devoted to the cross-linguistic aspect of polysemy and will focus on the method of semantic maps. The semantic map method has been used intensively during the last 20 years. It has proved attractive to typologists because it provides a convenient graphical display of the interrelationships between meanings across languages, while at the same time differentiating what is universal from what is language-specific. In the context of this course, I will give an overview of the model, explain in a step-by-step fashion the principles of the method, present its advantages and disadvantages and highlight the open questions in this field of research.

The course will be both theory- and practice-based. The participants will be introduced to some analytic techniques for identifying protoypes and for distinguishing between meanings within a language as well as to methods for automatically plotting semantic maps.

Course description

Before You Arrive for the Workshop

Please install the following software on the computer which you will bring to the event.

Python: To download go to this page: https://www.python.org/

Note that Python can be used on many operating systems and environments (the same for Anaconda and Gephi).

Find the appropriate one for your own operating system (see the dropdown menu in downloads).

Anaconda: free and open-source distribution of the Python and R programming languages for scientific computing. To download go to this page: https://www.anaconda.com/

Gephi: The Open Graph Viz Platform. To download go to this page: https://gephi.org/

Outline

Part A

1. The notion of Polysemy
   a) Polysemy *vs.* Homonymy
   b) Polysemy *vs.* Vagueness
2. Polysemy in Cognitive Linguistics
   a) *Phase 1*: The full-specification approach
   b) *Phase 2*: Newer developments
   c) *Phase 3*: Corpus linguistics approaches to polysemy

Outline

Part A

3. Polysemy across languages
4. A tool to represent polysemy across languages: Semantic maps
5. Introducing semantic maps
6. Semantic maps: Different types
7. How is a classical semantic map built?
8. Advantages of the semantic map model
9. Problems with classical semantic maps
10. A way out: Inferring semantic maps

Outline

Part B

Workshop activities

The notion of Polysemy

The notion of Polysemy

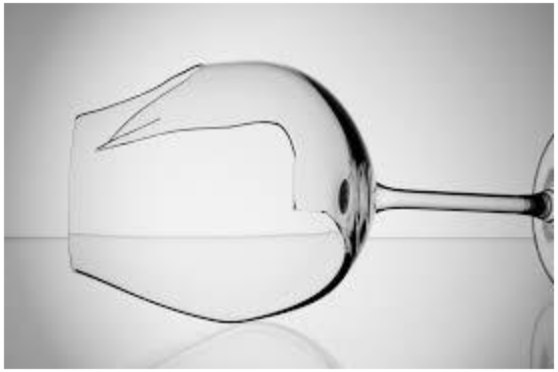
- 2+ related senses associated with the same word

Breal, Michel. 1897. *Essai de Sémantique: Science des Significations*. Paris: Hachette

The notion of Polysemy

Glass

I broke a ***glass***
('Material')



I filled the ***glass***
('Contents of the container')

The notion of Polysemy

Coat

a. ‘an outer garment with sleeves for wearing outdoors’

Put your ***coat*** on.

b. ‘an animal’s covering of fur’

Mountain goats have long shaggy ***coats***.

c. ‘a covering of paint or similar material’

All the door needs is a new ***coat*** of paint.

(Murphy, 2010: 88-89)

The notion of Polysemy

Coat

‘Garment’

‘animal covering’ ‘a covering of paint’

1. The desk has three ***coats*** on it.

2. The desk has three ***coats*** on it, but I still don’t think it looks finished.

zeugma test

3. # The desk has three ***coats*** on it – two parkas and one of walnut varnish.

(Murphy, 2010: 88-89)

Polysemy vs. Homonymy

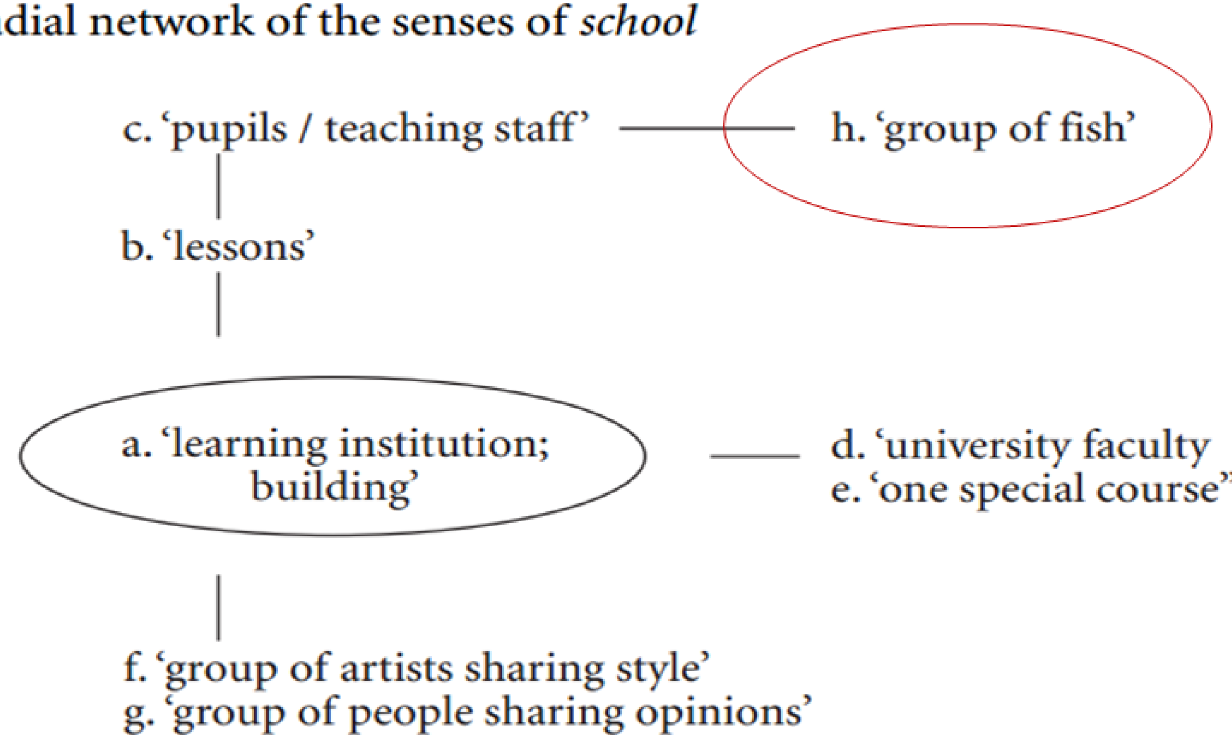
- **Homonymy**: relates to two distinct words that ***happen*** to share the same form in sound (homophones) and/or in writing (homographs)

bank1: 'side of the river'
(from Old Icelandic word for 'hill')

bank2: 'financial institution'
(from Italian banca)

Polysemy vs. Homonymy

Radial network of the senses of *school*



a-g < Latin *schola*
h < Old English *scolu* 'troupe'

(Dirven & Verspoor, 2004: 31-33)

Polysemy vs. Vagueness

Vagueness: If the word has one sense that is general enough that it can be applied to many different things, then the word has a vague, or indeterminate, sense.

Clock

'analog'

'digital'

Polysemy vs. Vagueness

Vagueness: If the word has one sense that is general enough that it can be applied to many different things, then the word has a vague, or indeterminate, sense.

1. Ben is my ***friend***. He's a fine fellow.
2. Georgia is my ***friend***. She's a wonderful woman.

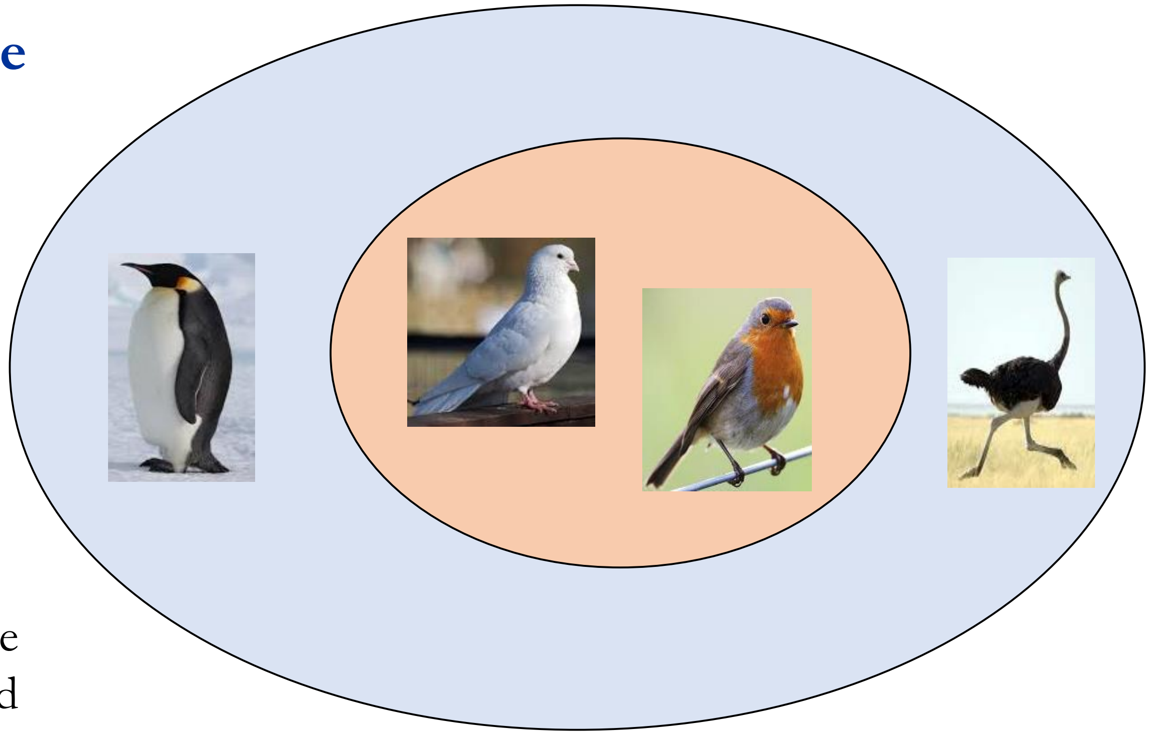
zeugma test

3. One of my ***friends*** is a policewoman and another is a fireman.

Polysemy in Cognitive Linguistics

Polysemy in Cognitive Linguistics

- **Incorporates prototype theory into linguistics**



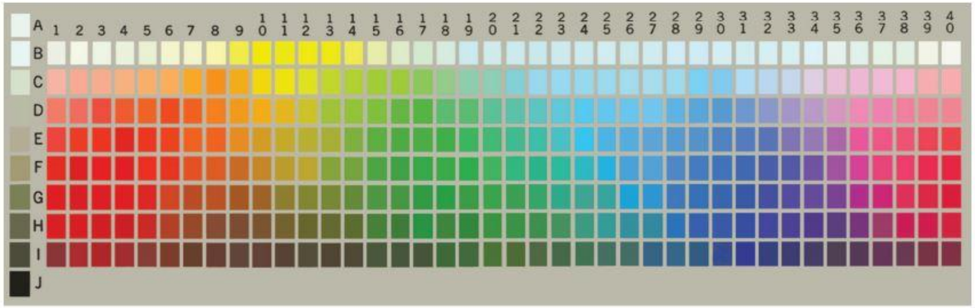
a. Prototypical categories exhibit degrees of typicality

b. Prototypical categories are blurred at the edges

c. Prototypical categories cannot be defined by means of necessary and sufficient attributes

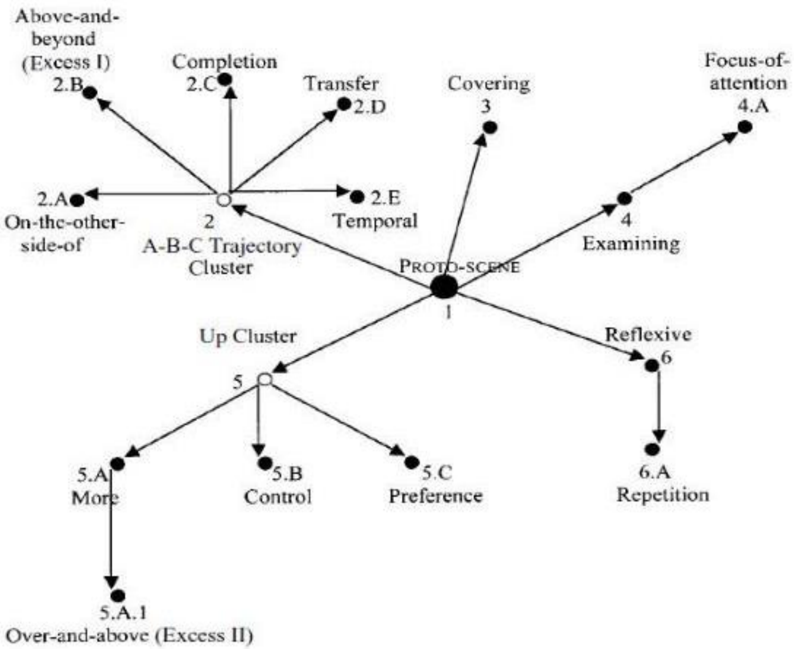
d. Prototypical categories exhibit a family resemblance structure

(Rosch & Mervis, 1975; Rosch, 1978; Geeraerts, 1997)

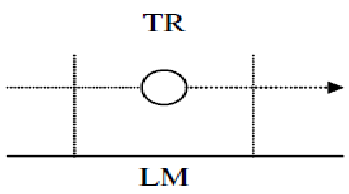


Polysemy in Cognitive Linguistics

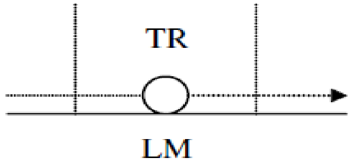
- **Word meanings**: a type of radial category



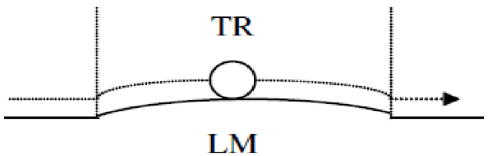
Phase 1: The full-specification approach



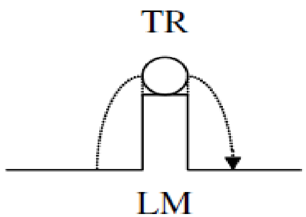
The bird flew over the yard (Schema 1.X.NC) (adapted from Lakoff 1987: 421)



John walked over the bridge (Schema 1.X.C) (adapted from Lakoff 1987: 422)



John walked over the hill (Schema 1.VX.C) (adapted from Lakoff 1987: 422)



Sam climbed over the wall (Schema 1.V.C) (adapted from Lakoff 1987: 422)

- Results in a potentially vast proliferation of senses

- Absence of clear methodological principles for establishing the distinct senses

Phase 2: Newer developments

- **The Principled Polysemy approach**

  - Two major objectives:

    - **Distinguish** between distinct senses and on-line interpretations
    - **Determine** the primary sense based on principled criteria

(Tyler & Evans, 2001; 2003)

Phase 2: Newer developments

- **The Principled Polysemy approach**
  - **Distinguish** between distinct senses and on-line interpretations

- **1st criterion**:

For a sense to count as distinct, it must involve a meaning that is not purely spatial in nature, and/or a spatial configuration holding between the TR and LM that is distinct from the other senses conventionally associated with that preposition

- **2nd criterion**:

There must also be instances of the sense that are context-independent: instances in which the distinct sense could not be inferred from another sense and the context in which it occurs

Phase 2: Newer developments

- **1st criterion**:

For a sense to count as distinct, it must involve a meaning that is not purely spatial in nature, and/or a spatial configuration holding between the TR and LM that is distinct from the other senses conventionally associated with that preposition

- **2nd criterion**:

There must also be instances of the sense that are context-independent: instances in which the distinct sense could not be inferred from another sense and the context in which it occurs

1. The hummingbird is hovering over the flower
2. The helicopter is hovering over the city
3. Joan nailed a board over the hole in the ceiling
4. The tablecloth is over the table

Phase 2: Newer developments

- **The Principled Polysemy approach**

	Sense	Example
1	ABOVE (central sense)	The picture is over the sofa
2.A	ON-THE-OTHER-SIDE-OF	St Paul's is over the river from Southwark
2.B	ABOVE-AND-BEYOND (excess I)	Your article is over the page limit
2.C	COMPLETION	The movie is over
2.D	TRANSFER	The discredited government hand power over to an interim authority
2.E	TEMPORAL	The relationship had altered over the years
3	COVERING	The clouds are over the sun
4	EXAMINING	Mary looked over the document quite carefully
4.A	FOCUS-OF-ATTENTION	The committee agonised over the decision
5.A	MORE	Jerome found over forty kinds of shells on the beach
5.A.1	OVER-AND-ABOVE (excess II)	The heavy rains caused the river to flow over its banks
5.B	CONTROL	She has a strange power over me
5.C	PREFERENCE	I would prefer tea over coffee
6	REFLEXIVE	The fence fell over
6.A	REPETITION	After the false start, they started the race over

TABLE. Distinct senses for *over* identified in Tyler & Evans (2001)

Phase 2: Newer developments

- **The Principled Polysemy approach**
    - **Determine** the primary sense based on principled criteria

1. Earliest attested meaning
2. Predominance in the semantic network
3. Relations to other prepositions
4. Ease of predicting sense extensions

Phase 2: Newer developments

- **Earliest attested meaning**

  - Sanskrit ***upari*** ('above'
  - Ancient Greek ***hupér*** ('over, above')
  - Latin ***super*** ('over, above')
  - Gothic ***ufar*** ('over')
  - Old English ***ofer*** ('over, above')
  - Old Teutonic ***ufa*** ('above')

Phase 2: Newer developments

- **Predominance in the semantic network**

  - The meaning components that are the most frequent in most of the spatial particle's other senses are more likely to be present in the primary sense

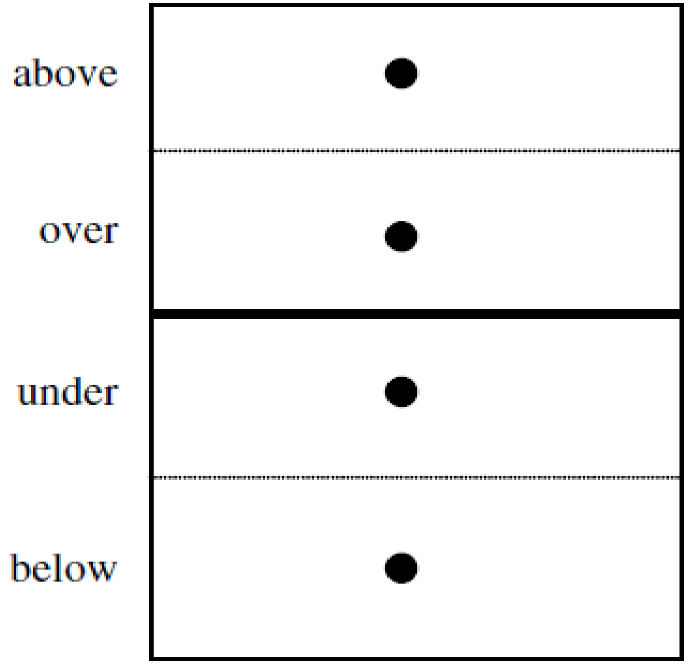
  - The most frequent meaning features of ***over*** involve the features:

    [+ Vertical]
    [- Contact]

Phase 2: Newer developments

- **Relations to other prepositions**

  - It has to form, together with a set of other prepositional primary senses, some sort of paradigm



Phase 2: Newer developments

- **Ease of predicting sense extensions**

- The majority of the senses should be derived from the primary sense

- If a sense is not directly derivable from the primary sense, it should be traceable to the primary one

1. Joan nailed a board over the hole in the ceiling

2. The tablecloth is over the table

Phase 2: Newer developments

- **The Principled Polysemy approach**

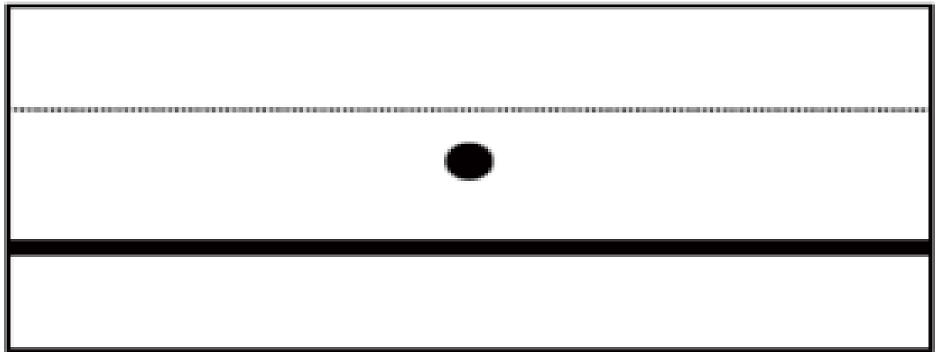


FIGURE. The proto-scene for *over* (based on Tyler & Evans, 2001)

1. She walked ***over*** the bridge
2. She walked ***above*** the bridge

Phase 2: Newer developments

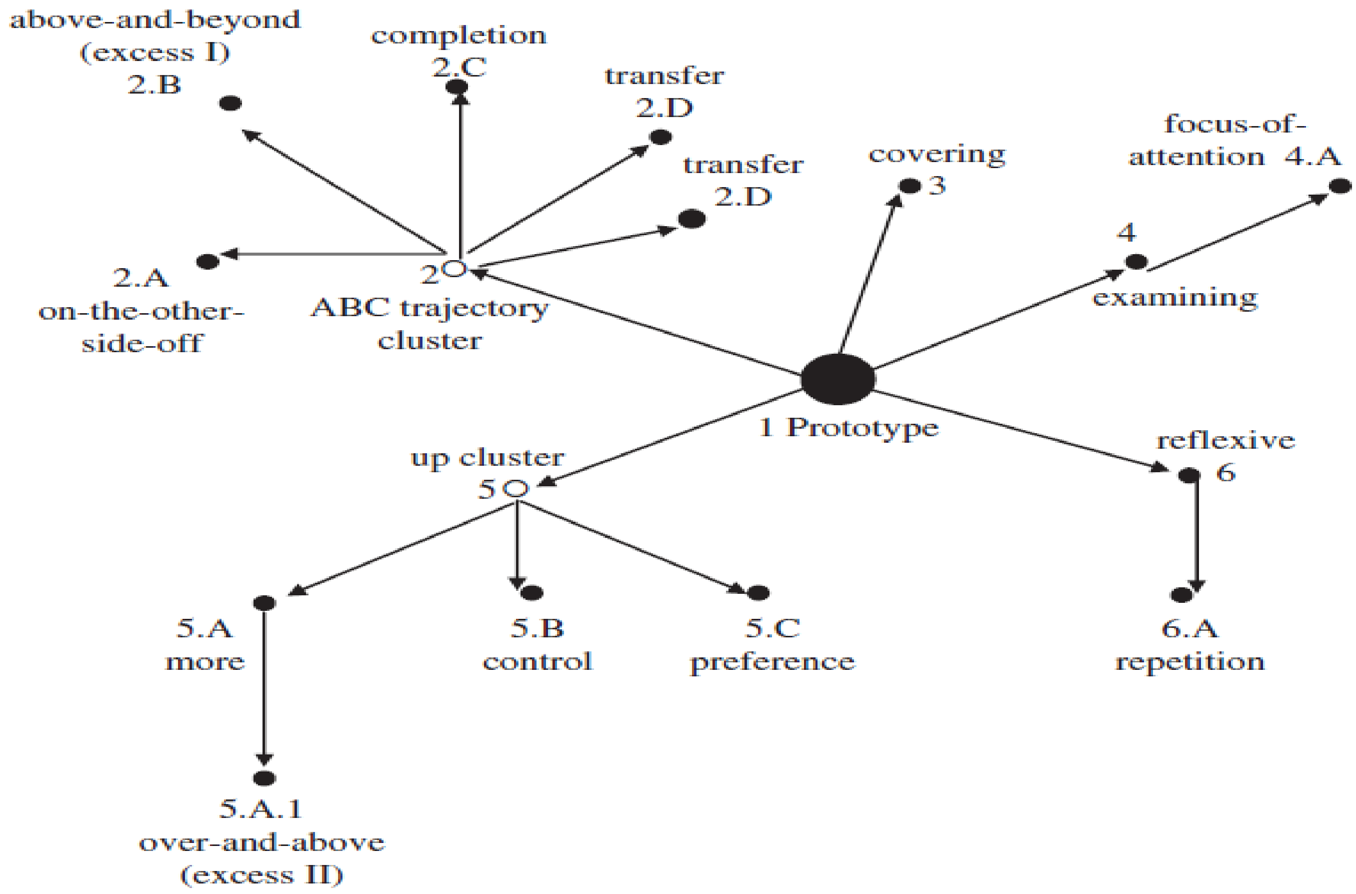


FIGURE. The semantic network for *over* (based on Tyler & Evans, 2003: 80)

Phase 2: Newer developments

- The Principled Polysemy approach
  - More rigorous decision principles
  - Decisions more replicable

Phase 3: Corpus linguistics approaches to polysemy

Basic principle: Lexical polysemy in general inheres in particular contextual features

Phase 3: Corpus linguistics approaches to polysemy

- Atkins (1987): "ID tags": "syntactic or lexical markers in the citations which point to a particular dictionary sense of the word".
- Fillmore & Atkins (1992): analysis of the polysemy of *risk* relies on the distinction of the different types of complements in corpus-derived examples (cf. Hanks 1996)
- Gries (2006) and Berez & Gries (2009): behavioral profiles of the verbs *run* and *get* respectively
- Jansegers and Gries (2017): dynamic behavior profile of the Spanish *sentir*
- Hilpert (2016): highlights the relationship between the collocates of *may* and its different senses
  - Shift in collocational preferences correlates with semantic change and the different meanings of the modal

Phase 3: Corpus linguistics approaches to polysemy

- **Behavioral profile approach**

(Divjak & Gries 2006; Gries 2006; 2010; Berez & Gries 2009; Gries & Divjak 2009; Divjak 2010)

Phase 3: Corpus linguistics approaches to polysemy

- **Behavioral profile approach**
  - **Four Steps**
    - **Retrieve** all instances of the lexeme in context in the form of a concordance
    - **Analyze** and **annotate manually** the concordance lines for a large number of features
    - **Convert** these data into a table of percentage vectors that state how much of the data in percent exhibits a particular feature
    - **Analyze statistically** the data with exploratory tools

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**

  - **Retrieve** all instances of the lexeme in context in the form of a concordance

    - **Data**: International Corpus of English and Brown Corpus of American English

    - 815 instances of ***to run***

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**

- **Retrieve** all instances of the lexeme in context in the form of a concordance

  - **Intransitive**
    - Simons had run down to the villa to get help
    - There are three boats that run from the mainland to the Island
    - On my way to the elevator, I ran into Pete
    - Time is running out
    - We may conceivably run into trouble here
    - When Bush was running for the White House

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**
- **Retrieve** all instances of the lexeme in context in the form of a concordance
  - **Transitive**
    - His brother ran a mile to get the father
    - The club runs regular trips to the cabins
    - The island's newspaper runs a weekly cartoon showing the adventures of 'Vincey'; in its struggle to survive
    - He was running a pizza store

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**
  - **Analyze** and **annotate manually** the concordance lines for a large number of features
    - *N*=252 features (ID tag levels).

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**

- Morphological features of the verb form:
  - Tense, aspect, voice
- Syntactic properties of the clause the verb form occurs in:
  - intransitive vs. transitive vs. complex transitive use of to run,
  - declarative vs. interrogative vs. imperative sentence form,
  - main clause vs. subordinate clause
- Semantic characteristics of the referents of the elements co-occurring with ***to run***:
  - its subjects/heads, objects and complements (which were coded, e.g., as human, animate, concrete countable objects, concrete mass nouns, machines, abstract entities, etc.)
- A paraphrase of ***to run's*** meaning in the citation.

Phase 3: Corpus linguistics approaches to polysemy

	Text	SENSE	USE	TR	INAN_TR	INAN_Tr_Con	TR_BARE	TR_NUM	Pred_Nou
7	θηκε η χαρη ν' ατενισω το θαυμα της σε ολη του την εκταση. Φοβουμαι να πω περισσοτερα, μην πεσω στην κοινοτυπία και λερώσω αυτή τη λευκή –όχι μόνο από το χιόνι– σελίδα. Ιδού ποιός	act	metaphor	animate	NA	NA	NA	singular	no
8	ειωση της τροφης, το γονατισμα, η ορθοστασια και η δημοσια διαπομπευση του μαθητη ο οποιος επεφτε σε παράπτωμα προβλέπονταν από το κανονιστικό πλαίσιο του σχολείου, ενώ, συνηθισμέν	act	metaphor	animate	NA	NA	no	singular	no
9	γάπη μου! Μέχρι τώρα το βούλωνα, ήμουν διακριτικός, αλλά αφού επιμένεις, Όσκαρ, να μου την πέφτεις έτσι, απόψε, με το που θα έρθει η Σαντρίτα, θα της τα πω όλα με το νι και με το σ	attack	metaphor	animate	NA	NA	NA	singular	no
10	ς, αλλα ακομη και παραπατωντας. Πανω που η Τζουλια αρχιζει το τοκ-σοου της στο Κ-5 της την επεσε η Εφορία; Μπας και της την «έστησαν» οι βαρόνοι των Μίντια; Πάνε οι άλλοι και ξοδεύ	attack	metaphor	inanimate	concrete	group.org	no	singular	no
11	ΠΥΡ και ΜΑΝΙΑ ο Μαλκοβιτς με το συμπατριωτη του, επειδη ο ΠΑΟ επεσε σε δυσκολο όμιλο! Τον καταλαβαίνουμε τον Παναθηναϊκό, που φόρεσε...μαύρες πλερέζες μ	be_matched	metaphor	inanimate	concrete	group.org	no	singular	no

LM	LM_TYPE	LM_BARE	LM_NUM	SENTENCE	CLAUSE	TENSE	MODE	ASPECT	WORD_ORDER	ADVERBIAL	TYPE_ADVERBIAL	PP_SEMANT
inanimate	abstract	yes	singular	declarative	dependent	non.past	indicative	perfective	VAdjun	yes	PP	goal
inanimate	abstract	yes	singular	declarative	dependent	past	indicative	imperfecti	SVAdjun	yes	PP	goal
animate	NA	NA	singular	declarative	dependent	non.past	subjuncti	imperfecti	OV	no	NA	NA
animate	NA	NA	singular	declarative	main	past	indicative	perfective	OV	no	NA	NA
inanimate	concrete	yes	singular	declarative	dependent	past	indicative	perfective	SVAdjun	yes	PP	goal

TABLE. Example showing how to annotate the data in Excel (*data from a study on the Modern Greek verb *fall*; Georgakopoulos *under review*)

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**

Variable (ID tag)	ID tag level
Tense	past
Aspect	perfective
Voice	active
Verb form	intransitive
Sentence form	declarative
Clause	main
Animacy of subject	animate
Animacy of complement	animate
Paraphrase	meet

TABLE. Example of annotated ID tags for one citation

Example:

- On my way to the elevator, I ***ran*** into Pete

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**
  - **Convert** these data into a table of percentage vectors that state how much of the data in percent exhibits a particular feature

Variable (ID tag)	ID tag level	meet	escape	copy	…
subject	animate	0.87	0.95	0.08	…
	inanimate	0.13	0.05	0.92	…
tense	non.past	0.41	0.25	0.17	…
	past	0.59	0.75	0.83	…
…					…

TABLE. Examples of Behavioral Profile vectors (*hypothetical numbers)

Phase 3: Corpus linguistics approaches to polysemy

TABLE. BP vector of the sense ‘vert_down’ of the Modern Greek verb *pefto* (Georgakopoulos *under review*)

```
> vert_down.bp <- bp(vert_down)
> vert_down.bp
          USE.literal          USE.metaphor            TR.animate          TR.inanimate      INAN_TR.abstract 
          1.000000000           0.000000000           0.372549020           0.627450980           0.000000000 
     INAN_TR.concrete             INAN_TR.N   INAN_TR.non.concrete INAN_Tr_Concr.action INAN_Tr_Concr.body.part 
          0.781250000           0.015625000           0.203125000           0.019607843           0.000000000 
   INAN_Tr_Concr.food INAN_Tr_Concr.group.org INAN_Tr_Concr.man.made INAN_Tr_Concr.physical INAN_Tr_Concr.substance 
          0.019607843           0.019607843           0.529411765           0.156862745           0.137254902 
 INAN_Tr_Concr.vegetal        TR_NUM.plural       TR_NUM.singular     Pred_Noun_Adj..no    Pred_Noun_Adj..yes 
          0.117647059           0.480000000           0.520000000           1.000000000           0.000000000 
           LM.animate          LM.inanimate      LM_TYPE.abstract      LM_TYPE.concrete LM_TYPE.representation 
          0.042857143           0.957142857           0.000000000           1.000000000           0.000000000 
        LM_NUM.plural      LM_NUM.singular  SENTENCE.declarative   SENTENCE.imperative SENTENCE.interrogative 
          0.185714286           0.814285714           0.980392157           0.000000000           0.019607843 
     CLAUSE.dependent           CLAUSE.main        TENSE.non.past            TENSE.past       MODE.indicative 
          0.431372549           0.568627451           0.380000000           0.620000000           0.872549020 
     MODE.participle      MODE.subjunctive    ASPECT.imperfective     ASPECT.perfective    WORD_ORDER.AdjunVS 
          0.019607843           0.107843137           0.510000000           0.490000000           0.019607843 
        WORD_ORDER.OV  WORD_ORDER.PredNounVS         WORD_ORDER.SV     WORD_ORDER.SVAdjun   WORD_ORDER.SVPredAdj 
          0.009803922           0.000000000           0.147058824           0.401960784           0.000000000 
 WORD_ORDER.SVPredNoun          WORD_ORDER.V     WORD_ORDER.VAdjun    WORD_ORDER.VAdjunS    WORD_ORDER.VPredAdj 
          0.000000000           0.019607843           0.205882353           0.029411765           0.000000000 
   WORD_ORDER.VPredAdS   WORD_ORDER.VPredNoun         WORD_ORDER.VS     WORD_ORDER.VSAdjun  WORD_ORDER.VSPredNoun 
          0.000000000           0.000000000           0.078431373           0.088235294           0.000000000 
        ADVERBIAL.no         ADVERBIAL.yes  TYPE_ADVERBIAL.adverb     TYPE_ADVERBIAL.PP         PP_SEMANT.goal 
          0.264705882           0.735294118           0.066666667           0.933333333           0.700000000 
      PP_SEMANT.other      PP_SEMANT.source  PP_SEMANT.source.goal 
          0.014285714           0.257142857           0.028571429 
```

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**

- **Analyze statistically** the data with exploratory tools
  - For the statistical procedure use

    Levshina, 2015: ch. 15

    (Behavioural profiles, distance metrics and cluster analysis)

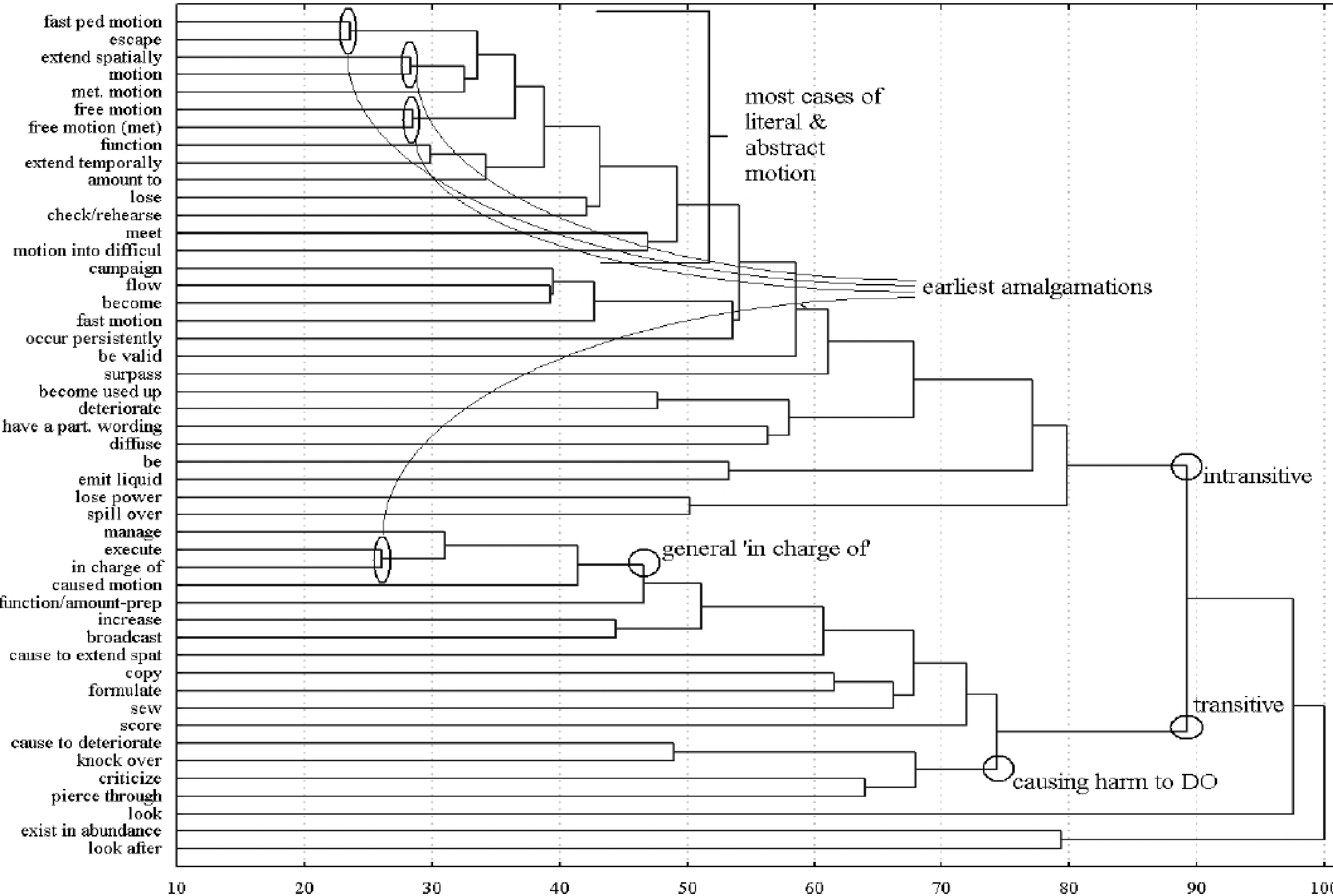


FIGURE. Dendrogram resulting from a hierarchical cluster analysis (Gries, 2006: 82)

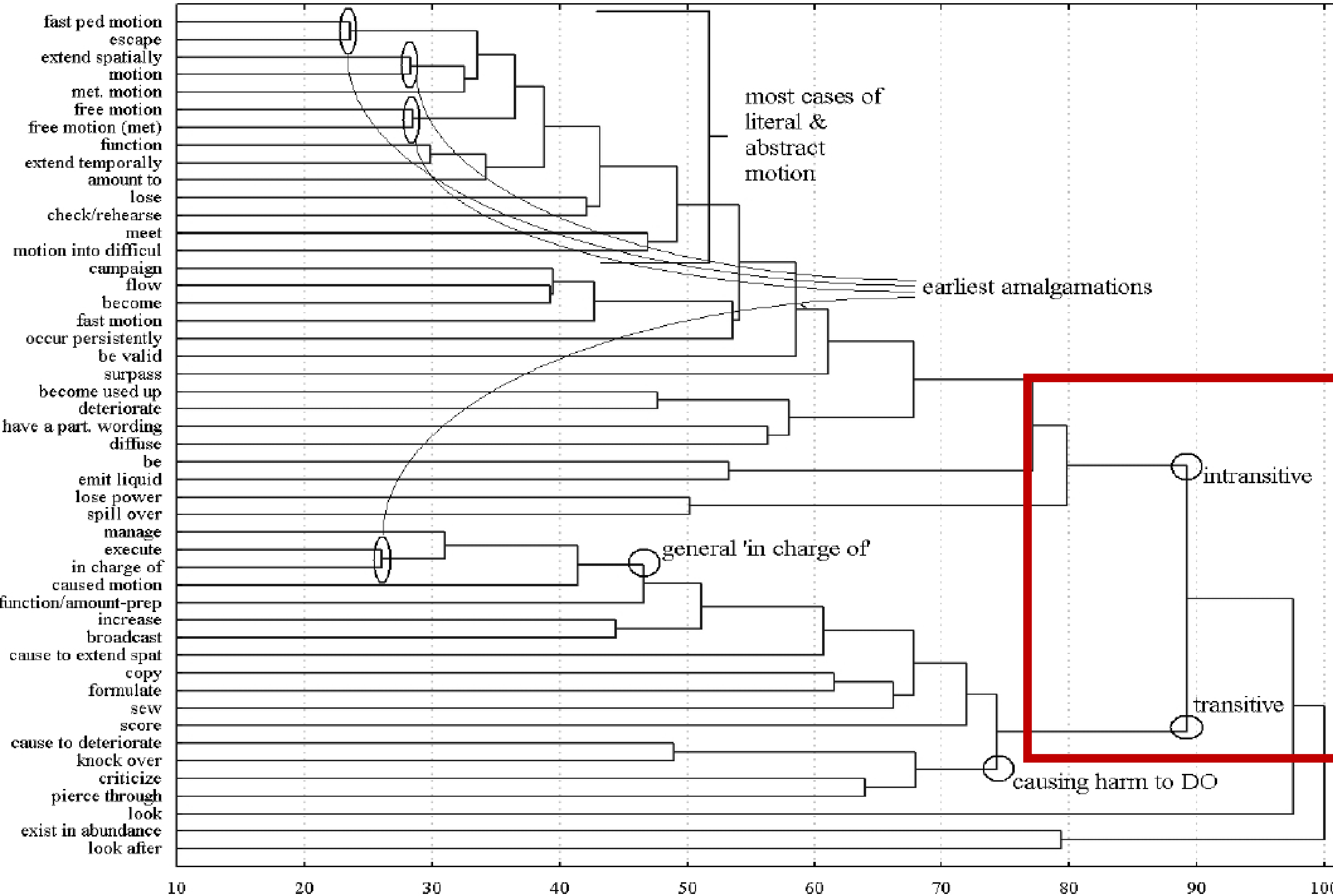


FIGURE. Dendrogram resulting from a hierarchical cluster analysis (Gries, 2006: 82)

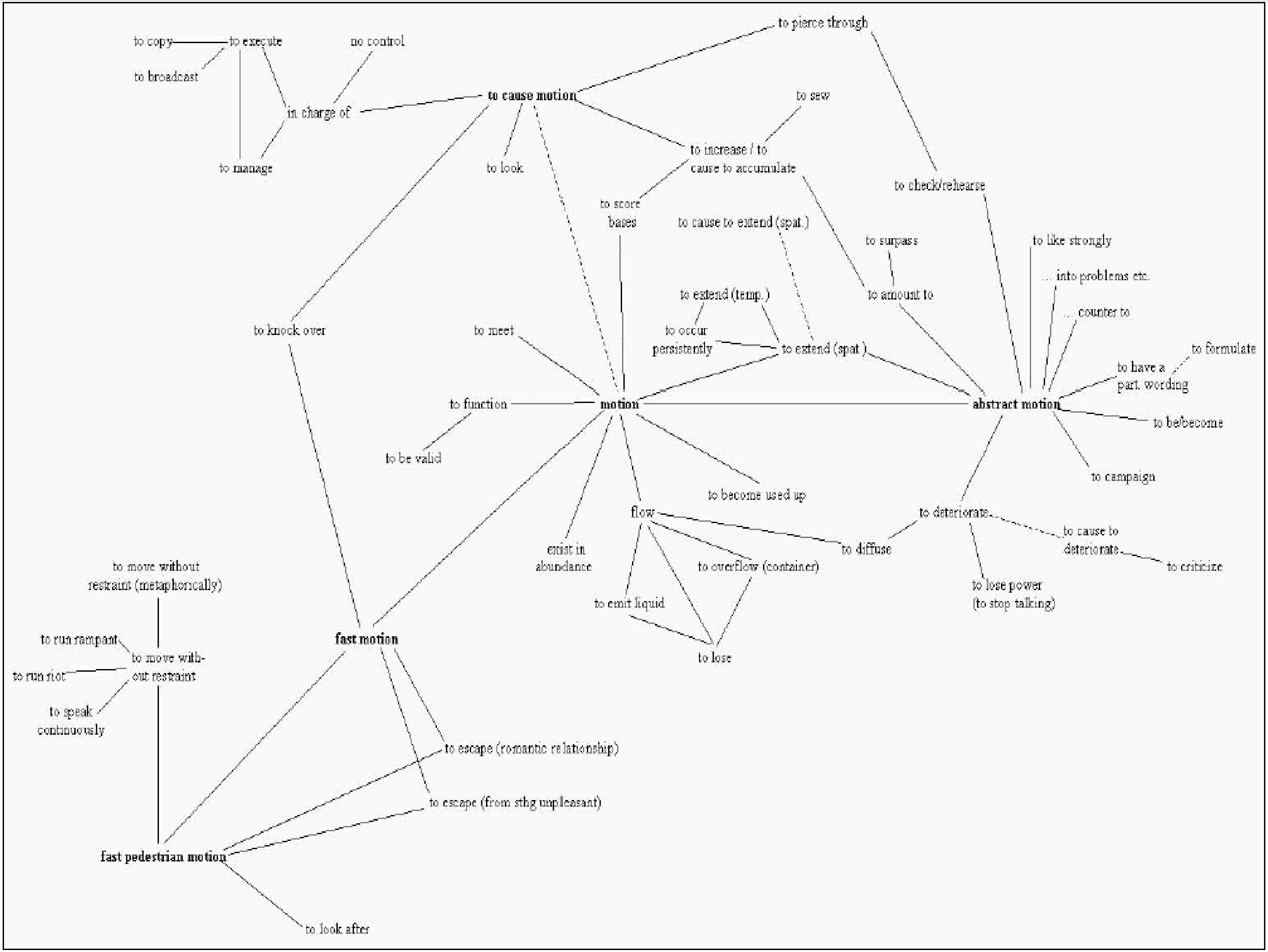
FIGURE. Dendrogram resulting from a hierarchical cluster analysis (Gries, 2006: 82)

Phase 3: Corpus linguistics approaches to polysemy

	TermAll.bp	Limit.bp	GoalPurp.bp	TempBound.bp	Perc.bp	Purpose.bp	On.bp	Moment.bp	Duration.bp	DirAll.bp
Limit.bp	53.12									
GoalPurp.bp	40.46	48.29								
TempBound.bp	49.84	49.88	41.00							
Perc.bp	43.62	44.27	44.79	49.93						
Purpose.bp	45.57	48.56	32.80	40.95	47.43					
On.bp	44.82	44.14	41.73	48.14	39.47	46.97				
Moment.bp	54.77	39.29	43.04	45.04	48.84	46.94	44.68			
Duration.bp	52.84	43.47	44.22	36.09	46.97	45.14	51.09	36.90		
DirAll.bp	33.40	48.43	38.85	42.59	38.89	41.48	43.50	50.09	47.54	
AbstrMot.bp	50.27	41.35	49.81	52.74	42.57	53.54	42.36	47.94	49.00	47.56

```
>
> max(eis_Homer.dist)
[1] 54.76993
> min(eis_Homer.dist)
[1] 32.79767
```

TABLE. Distance matrix showing the pairwise distances between the BP vectors of the preposition *eis* in Homer (Georgakopoulos *in prep*.)



to copy
to execute
no control
to broadcast
to pierce through
to cause motion
in charge of
to sew
to manage
to look
to increase / to cause to accumulate
to check/rehearse
to score bases
to cause to extend (spat.)
to surpass
to like strongly
… into problems etc.
to extend (temp.)
to amount to
… counter to
to knock over
to meet
to occur persistently
to extend (spat.)
to formulate
to have a part. wording
to function
motion
abstract motion
to be/become
to be valid
to campaign
to become used up
flow
to deteriorate
to cause to deteriorate
exist in abundance
to diffuse
to criticize
to move without restraint (metaphorically)
to overflow (container)
to lose power (to stop talking)
to emit liquid
to run rampant
fast motion
to move without restraint
to lose
to run riot
to speak continuously
to escape (romantic relationship)
to escape (from sthg unpleasant)
fast pedestrian motion
to look after

Phase 3: Corpus linguistics approaches to polysemy

- **Case study: *to run* (Gries 2006)**
  - **What about prototypicality?**
    - Information about the most frequent sense
    - Information about the sense with the highest number of different ID tags

Phase 3: Corpus linguistics approaches to polysemy

TABLE. BP vector of the sense 'vert_down' of the Modern Greek verb *pefto* (Georgakopoulos *under review*)

```
> vert_down.bp <- bp(vert_down)
> vert_down.bp
          USE.literal           USE.metaphor              TR.animate            TR.inanimate        INAN_TR.abstract
          1.000000000            0.000000000             0.372549020             0.627450980             0.000000000
     INAN_TR.concrete              INAN_TR.N    INAN_TR.non.concrete  INAN_Tr_Concr.action  INAN_Tr_Concr.body.part
          0.781250000            0.015625000             0.203125000             0.019607843             0.000000000
   INAN_Tr_Concr.food INAN_Tr_Concr.group.org INAN_Tr_Concr.man.made INAN_Tr_Concr.physical INAN_Tr_Concr.substance
          0.019607843            0.019607843             0.529411765             0.156862745             0.137254902
 INAN_Tr_Concr.vegetal          TR_NUM.plural         TR_NUM.singular      Pred_Noun_Adj..no      Pred_Noun_Adj..yes
          0.117647059            0.480000000             0.520000000             1.000000000             0.000000000
           LM.animate            LM.inanimate        LM_TYPE.abstract        LM_TYPE.concrete  LM_TYPE.representation
          0.042857143            0.957142857             0.000000000             1.000000000             0.000000000
        LM_NUM.plural        LM_NUM.singular    SENTENCE.declarative     SENTENCE.imperative  SENTENCE.interrogative
          0.185714286            0.814285714             0.980392157             0.000000000             0.019607843
     CLAUSE.dependent            CLAUSE.main          TENSE.non.past              TENSE.past         MODE.indicative
          0.431372549            0.568627451             0.380000000             0.620000000             0.872549020
     MODE.participle        MODE.subjunctive      ASPECT.imperfective      ASPECT.perfective      WORD_ORDER.AdjunVS
          0.019607843            0.107843137             0.510000000             0.490000000             0.019607843
        WORD_ORDER.OV     WORD_ORDER.PredNounVS           WORD_ORDER.SV       WORD_ORDER.SVAdjun      WORD_ORDER.SVPredAdj
          0.009803922            0.000000000             0.147058824             0.401960784             0.000000000
 WORD_ORDER.SVPredNoun            WORD_ORDER.V       WORD_ORDER.VAdjun       WORD_ORDER.VAdjunS       WORD_ORDER.VPredAdj
          0.000000000            0.019607843             0.205882353             0.029411765             0.000000000
   WORD_ORDER.VPredAdS    WORD_ORDER.VPredNoun           WORD_ORDER.VS       WORD_ORDER.VSAdjun    WORD_ORDER.VSPredNoun
          0.000000000            0.000000000             0.078431373             0.088235294             0.000000000
        ADVERBIAL.no           ADVERBIAL.yes    TYPE_ADVERBIAL.adverb       TYPE_ADVERBIAL.PP            PP_SEMANT.goal
          0.264705882            0.735294118             0.066666667             0.933333333             0.700000000
      PP_SEMANT.other        PP_SEMANT.source   PP_SEMANT.source.goal
          0.014285714            0.257142857             0.028571429
```

Phase 3: Corpus linguistics approaches to polysemy

TABLE. BP vector of the sense 'notice' of the Modern Greek verb *pefto* (Georgakopoulos *under review*)

```
> notice.bp <- bp(notice)
> notice.bp
             USE.literal             USE.metaphor              TR.animate             TR.inanimate        INAN_TR.abstract
              0.00000000               1.00000000              0.00000000               1.00000000              0.07692308
        INAN_TR.concrete                INAN_TR.N     INAN_TR.non.concrete   INAN_Tr_Concr.action  INAN_Tr_Concr.body.part
              0.84615385               0.00000000              0.07692308               0.00000000              1.00000000
      INAN_Tr_Concr.food INAN_Tr_Concr.group.org  INAN_Tr_Concr.man.made  INAN_Tr_Concr.physical  INAN_Tr_Concr.substance
              0.00000000               0.00000000              0.00000000               0.00000000              0.00000000
    INAN_Tr_Concr.vegetal           TR_NUM.plural           TR_NUM.singular       Pred_Noun_Adj..no       Pred_Noun_Adj..yes
              0.00000000               0.00000000              1.00000000               1.00000000              0.00000000
               LM.animate             LM.inanimate         LM_TYPE.abstract         LM_TYPE.concrete  LM_TYPE.representation
              0.15384615               0.84615385              0.09090909               0.90909091              0.00000000
           LM_NUM.plural          LM_NUM.singular     SENTENCE.declarative      SENTENCE.imperative    SENTENCE.interrogative
              0.00000000               1.00000000              1.00000000               0.00000000              0.00000000
        CLAUSE.dependent              CLAUSE.main           TENSE.non.past               TENSE.past         MODE.indicative
              0.38461538               0.61538462              0.23076923               0.76923077              1.00000000
        MODE.participle         MODE.subjunctive      ASPECT.imperfective       ASPECT.perfective       WORD_ORDER.AdjunVS
              0.00000000               0.00000000              0.30769231               0.69230769              0.00000000
           WORD_ORDER.OV     WORD_ORDER.PredNounVS            WORD_ORDER.SV         WORD_ORDER.SVAdjun      WORD_ORDER.SVPredAdj
              0.00000000               0.00000000              0.00000000               0.61538462              0.00000000
   WORD_ORDER.SVPredNoun             WORD_ORDER.V         WORD_ORDER.VAdjun        WORD_ORDER.VAdjuns       WORD_ORDER.VPredAdj
              0.00000000               0.00000000              0.07692308               0.00000000              0.00000000
     WORD_ORDER.VPredAds     WORD_ORDER.VPredNoun            WORD_ORDER.VS        WORD_ORDER.VSAdjun     WORD_ORDER.VSPredNoun
              0.00000000               0.00000000              0.00000000               0.30769231              0.00000000
            ADVERBIAL.no            ADVERBIAL.yes    TYPE_ADVERBIAL.adverb        TYPE_ADVERBIAL.PP           PP_SEMANT.goal
              0.00000000               1.00000000              0.00000000               1.00000000              1.00000000
         PP_SEMANT.other         PP_SEMANT.source     PP_SEMANT.source.goal
              0.00000000               0.00000000              0.00000000
```

Polysemy across languages

(resorting to joint work with Stéphane Polis, F.R.S.-FNRS / ULiège)

Polysemy across languages: Introduction

From language-specific to cross-linguistic regularities

Polysemy across languages: Introduction

From language-specific to cross-linguistic regularities

From polysemy (i.e. radial) networks to semantic maps

Polysemy across languages: Introduction

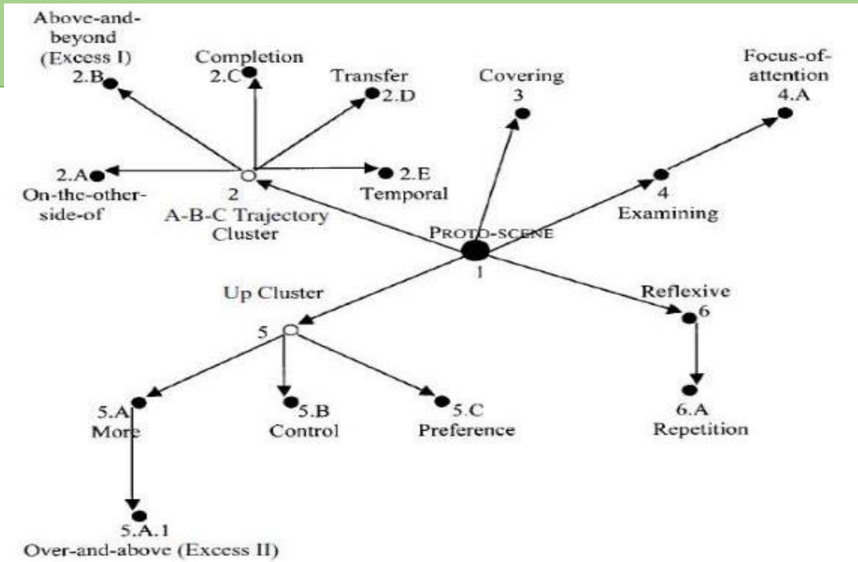
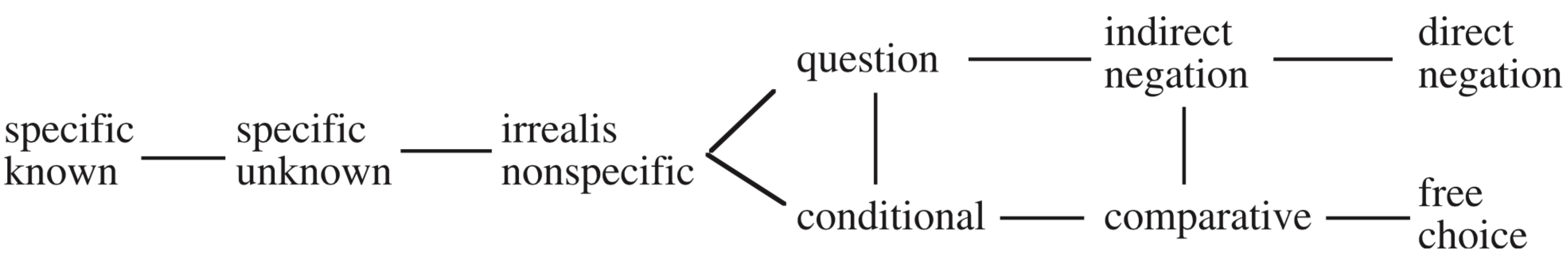


FIGURE. The polysemic radial network for *over* (Tyler & Evans, 2003: 80)

FIGURE. Semantic map of the indefinite pronouns functions (Haspelmath, 1997: 4)



Polysemy across languages: Introduction

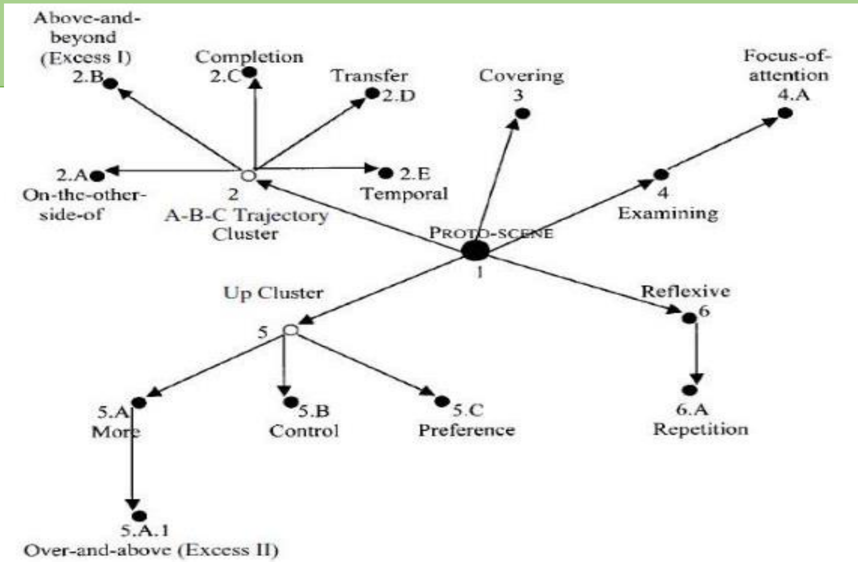


FIGURE. The polysemic radial network for *over* (Tyler & Evans, 2003: 80)

Language specific

- Certain criteria are applied to distinguish between distinct senses and context dependent usages.

(e.g. Tyler & Evans, 2003; cf. Lakoff, 1987)

Polysemy

- Sense distinctions are attributed to speakers' mental representations.

Polysemy across languages: Introduction

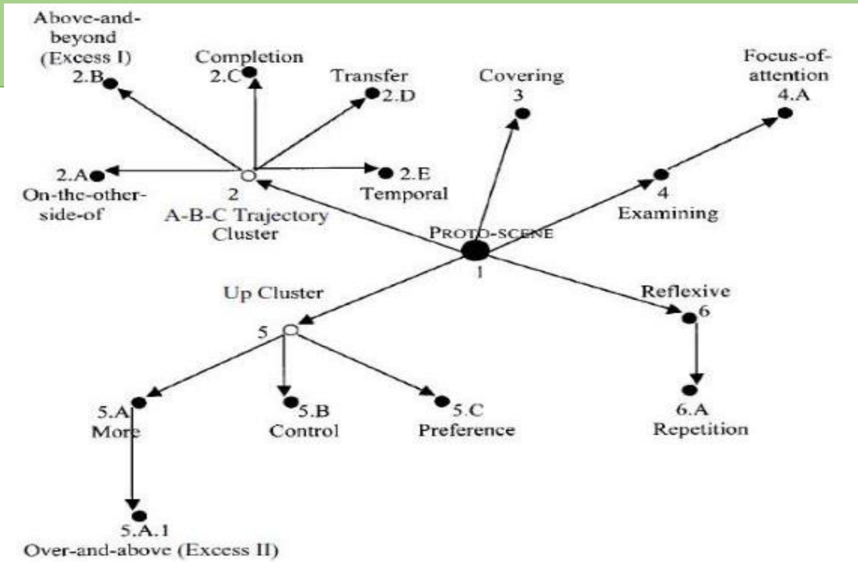


FIGURE. The polysemic radial network for *over* (Tyler & Evans, 2003: 80)

- There is a **prototypical meaning** from which other senses are derived in radial fashion.

- The **semasiological** approach is mainly adopted

- The **arrow** shows derivation of a sense from an another sense (not directionality)

Polysemy across languages: Introduction

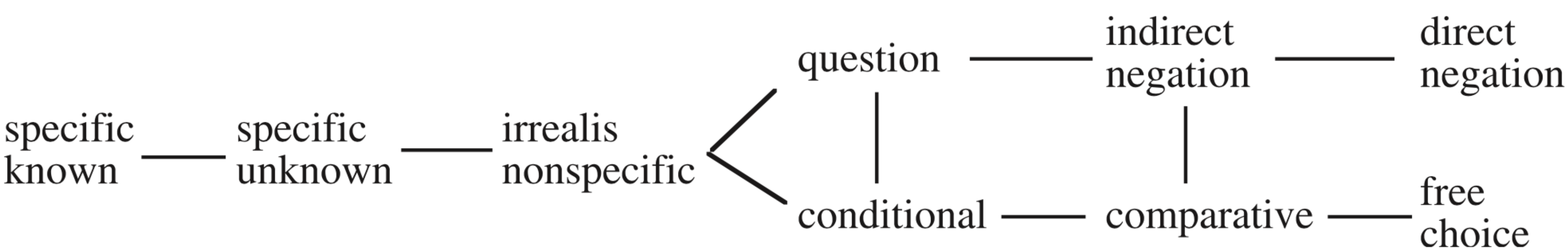
Multifunctionality (Haspelmath, 2003)
No commitment to a particular claim about conventionalization of senses.

Colexification
"A given language is said to colexify two functionally distinct senses if, and only if, it can associate them with the same lexical form" (François, 2008: 170)

Nodes on the map: meanings/ functions employed by the typologist to express generalizations across languages

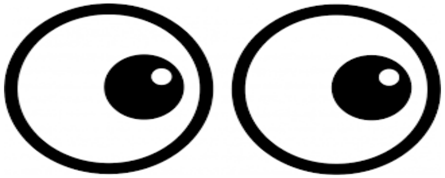
Approach: Both semasiological and onomasiological

FIGURE. Semantic map of the indefinite pronouns functions (Haspelmath, 1997: 4)



A tool to represent polysemy across languages: Semantic maps

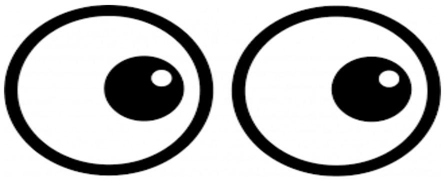
Some background



Some background

VISION

COGNITION



Some background

Some background

Colexifications for "SEE" and "KNOW (SOMETHING)"

Search:

Language	Family	Form for SEE	Gloss for SEE	Form for KNOW (SOMETHING)	Gloss for KNOW (SOMETHING)
guambiano	Barbacoan	*aʃip*	see	*aʃip*	know
orejón	Tucanoan	*jĩãji*	see	*jĩãji*	know
tunebo central	Chibchan	*istínro*	see	*istínro*	know
Old High German	Indo-European	*ir-kennen*	see	*ir-kennen*	know
Hawaiian	Austronesian	*ʔike*	see	*ʔike*	know
Maori	Austronesian	*kite-a*	see	*kite-a*	know
Araona	Pano-Tacanan	*ba*	see	*ba*	know
Ese Ejja	Pano-Tacanan	*ɓa-nahe*	see	*ɓa-nahe*	know
Pacaas Novos	Chapacuran	*kerek*	see	*kerek*	know
Ayoreo	Zamucoan	*i'moʔ*	see	*i'moʔ*	know
Telugu	Dravidian	*arayu*	see	*arayu*	know
Telugu	Dravidian	*aarayu*	see	*aarayu*	know
Mlabri	Austroasiatic	*mʌc*	see	*mʌc*	know
Lame (peve)	Afro-Asiatic	*we*	SEE	*we*	KNOW
Kaera	Timor-Alor-Pantar	*lalo*	to see	*lalo*	to know
Sulung [Puroik]	Sino-Tibetan	*daɬ*^{33}	see	*daɬ*^{33}	know
Hawaiian	Austronesian	*ʻike*	to see	*ʻike*	to know

(CLICS[2] https://clics.clld.org/edges/1409-1410; List et al. 2018)

Some background

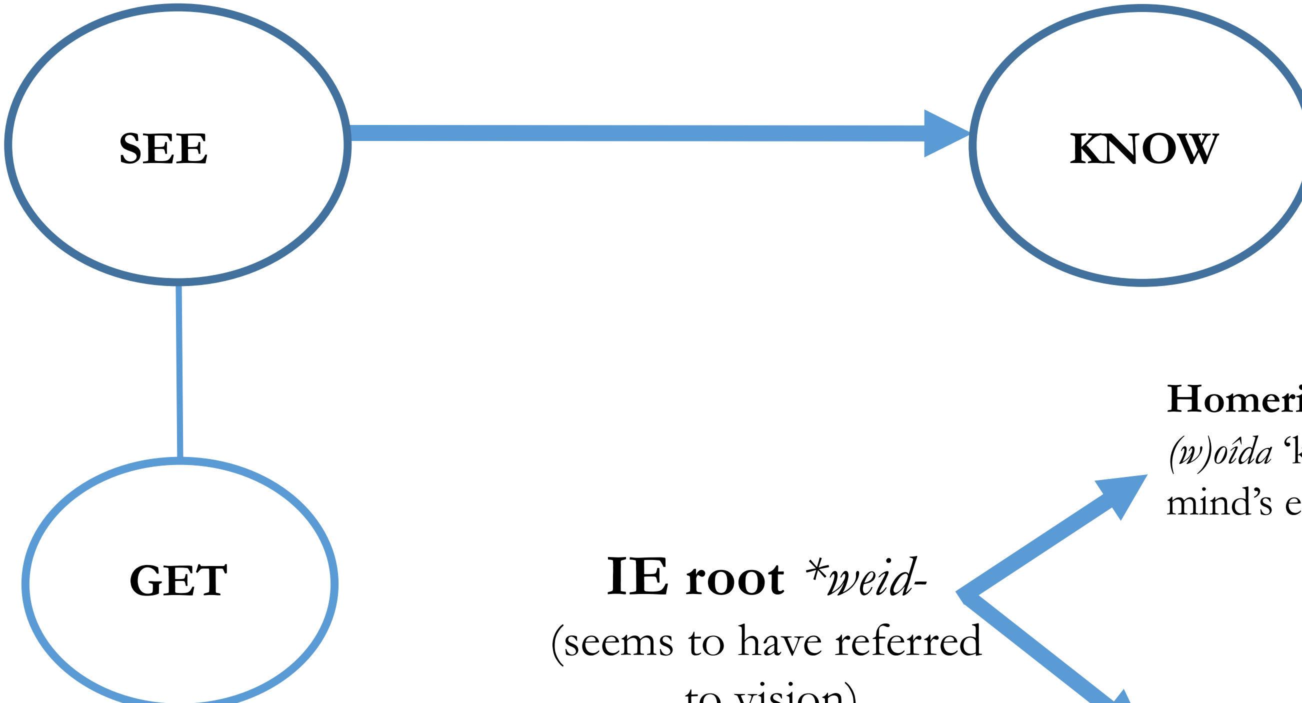
Colexifications for "GET" and "SEE"

Search:

Language	Family	Form for GET	Gloss for GET	Form for SEE	Gloss for SEE
Tswana S.30	Atlantic-Congo	*bóná*	get, obtain	*bóná*	see
Macushi	Cariban	*eraʔma*	get, obtain	*eraʔma*	see
Páez	Páez	*uy-*	get, obtain	*uy-*	see
Sirionó	Tupian	*tea*	get, obtain	*tea*	see
Kaingáng	Nuclear-Macro-Je	*we*	get, obtain	*we*	see
Bura	Afro-Asiatic	*wúta*	OBTAIN	*wutà*	SEE
Burma	Afro-Asiatic	*kwe*	OBTAIN	*kwe*	SEE
Pero	Afro-Asiatic	*wɛyɔ̀*	OBTAIN	*wɛyɔ̀*	SEE

(CLICS[2]; https://clics.clld.org/edges/694-1409; List et al., 2018)

Some background



IE root **weid-*
(seems to have referred to vision)

Homeric Greek perfect: *(w)oîda* ‘know, see with the mind’s eye’

Homeric Greek Aorist: *(w)eîdon* ‘saw’

Introducing semantic maps

ARTICLE

WILEY

The semantic map model: State of the art and future avenues for linguistic research

Thanasis Georgakopoulos[1] | Stéphane Polis[2]

Not a member? Sign up for My OBO. Already a member? My OBO Sign in
Advanced Search

Oxford Bibliographies
Your Best Research Starts Here

Browse by Subject

My Searches (0)

Back to Results

OUP-Reference Gratis Access

Search your library's resources

Sign out

Sign in to an additional subscriber account

Semantic Maps
Thanasis Georgakopoulos

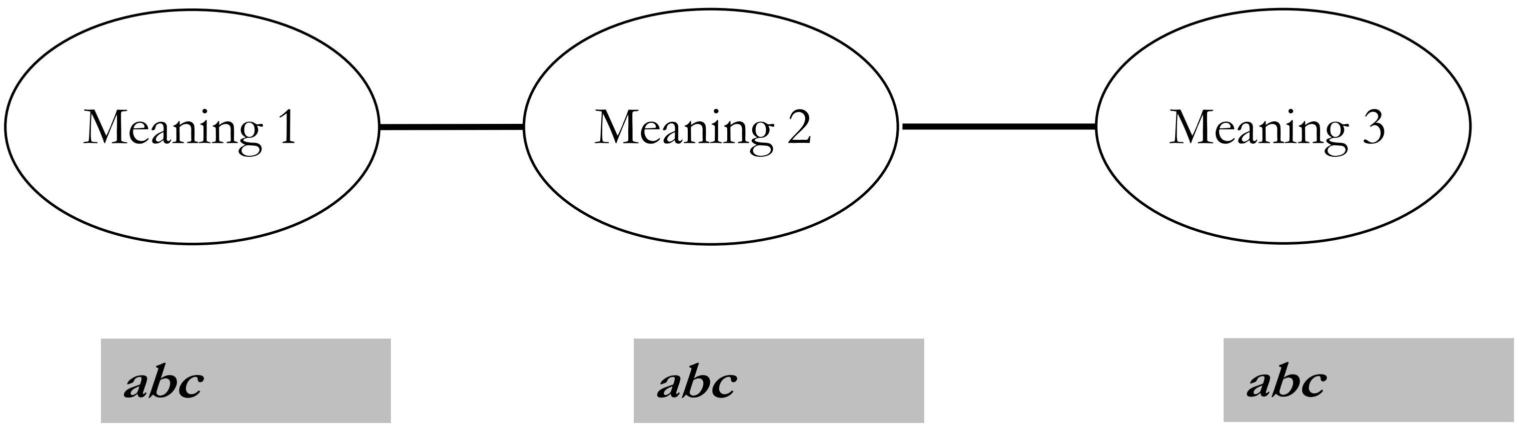
LAST MODIFIED: 15 JANUARY 2019
DOI: 10.1093/OBO/9780199772810-0229

LINGUISTICS
About Linguistics »
Meet the Editorial Board »

JUMP TO OTHER ARTICLES:

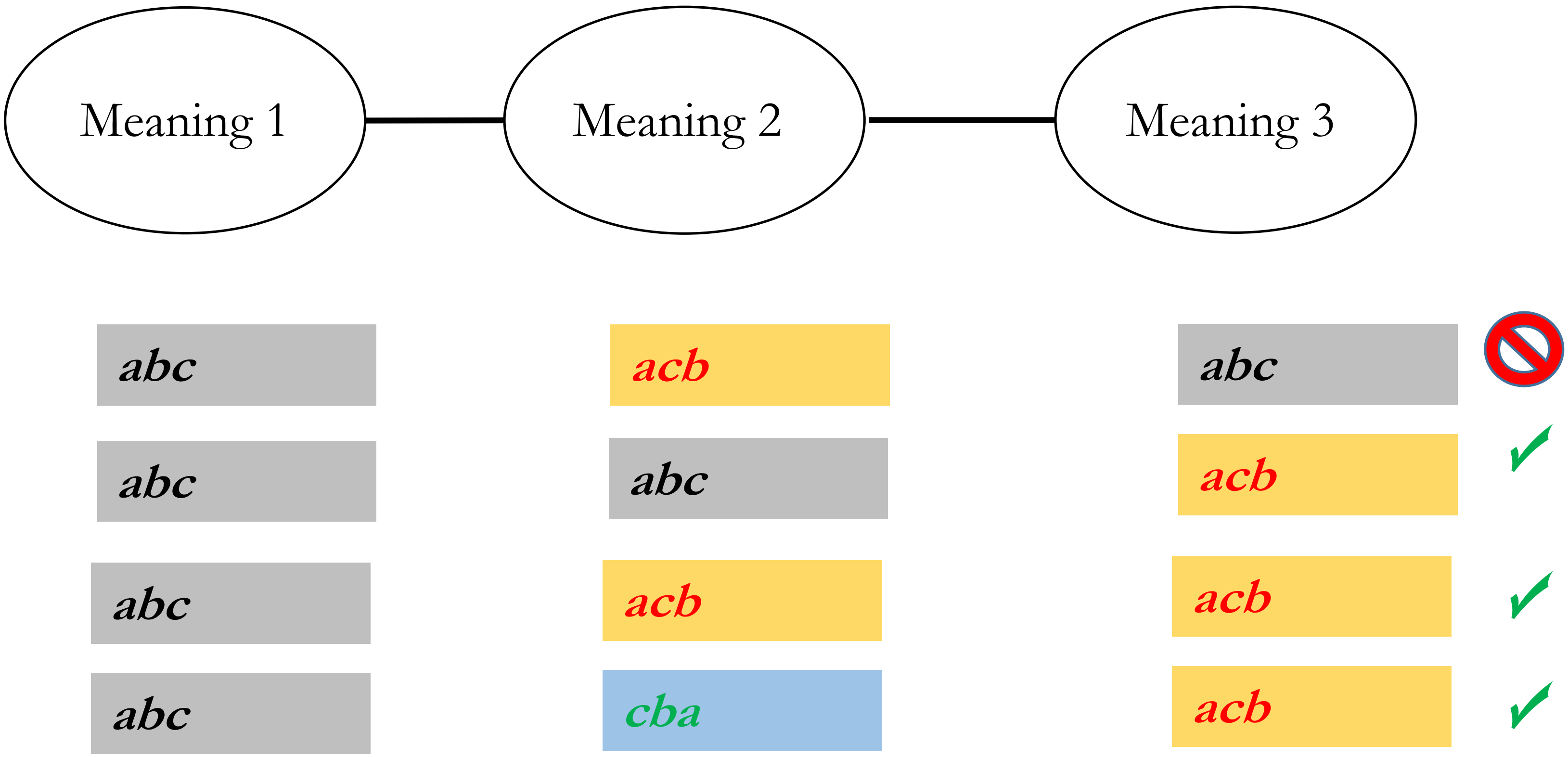
GO

Introducing semantic maps



abc colexifies/ co-expresses **Meaning 1**, **Meaning 2**, and **Meaning 3**

Introducing semantic maps



Introducing semantic maps

- **The semantic map connectivity hypothesis**:

"Any relevant language-specific and construction-specific category should map onto a CONNECTED REGION in conceptual space"

(Croft, 2001: 96)

  - If two given meanings are expressed by one form in at least one language, the corresponding meanings should be connected

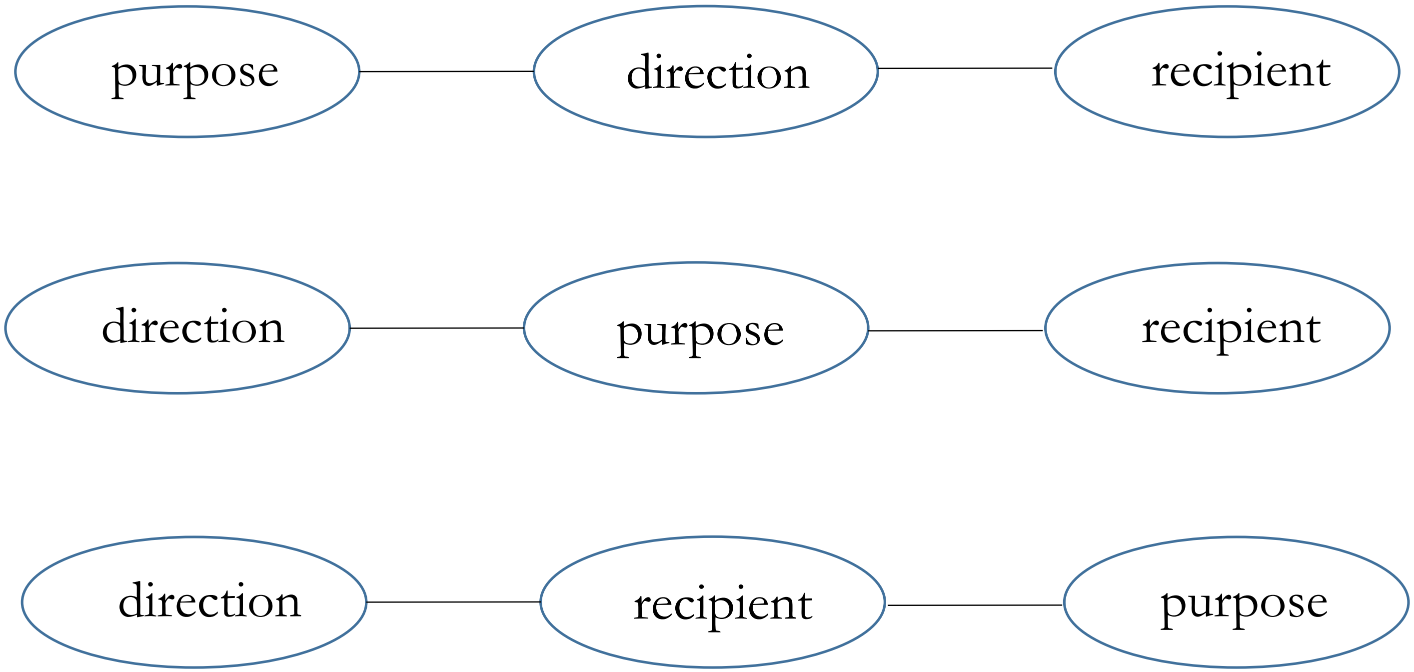
Introducing semantic maps

The English preposition *to*:

- '**Direction**': The teacher is going ***to*** the school
- '**Purpose**': The lifeguard ran ***to*** the rescue of the child
- '**Recipient**': The teacher gave the book ***to*** the student

Introducing semantic maps

How to arrange the different functions?



(Haspelmath, 2003)

Introducing semantic maps

German:

- ‘**Direction**’: Ich gehe ***zu*** Anna
- ‘**Purpose**’: Anna ging ***zum*** Spielen in den Garten

$\neq$

- ‘**Recipient**’: Ich gebe ***dir*** das Buch

direction = purpose ≠ recipient

(Haspelmath, 2003)

Introducing semantic maps

French:

direction = **recipient** ≠ **purpose**

Je vais ***à*** Poseidi

Je donne le livre ***à*** Athina

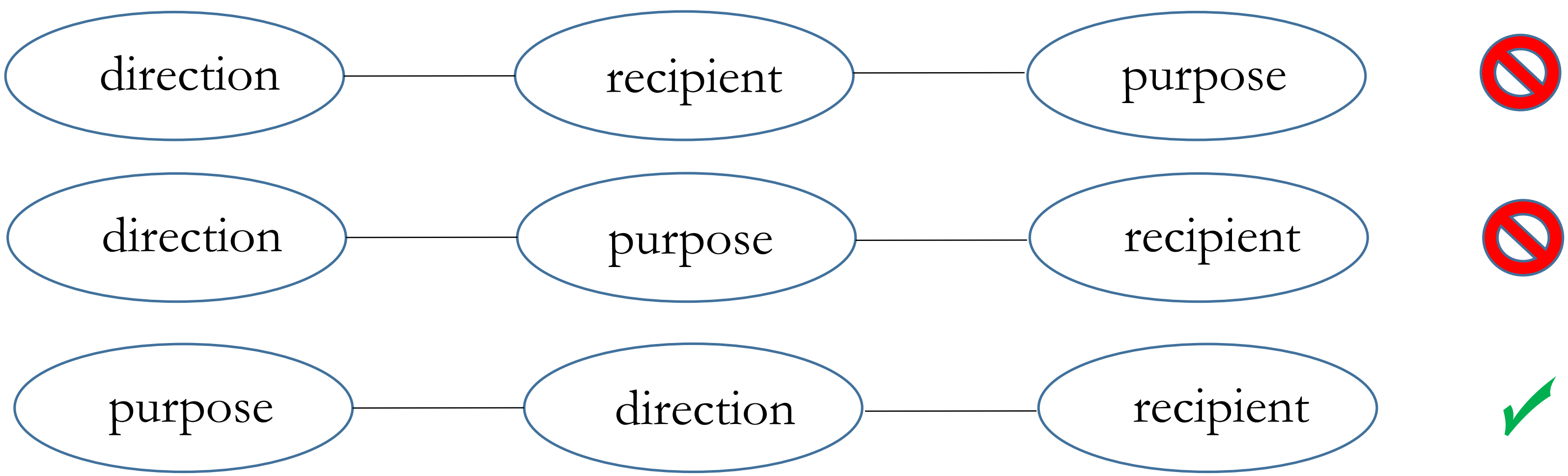
Je donne la balle ***pour*** jouer dans le jardin

(*J' ai quitté la fête tôt **à** arriver à la maison en bon temps
"I left the party early to get home in time"

(cf. Haspelmath, 2003)

Introducing semantic maps

a. direction = recipient = purpose
b. direction = purpose ≠ recipient
c. purpose ≠ direction = recipient



(Haspelmath, 2003)

Introducing semantic maps

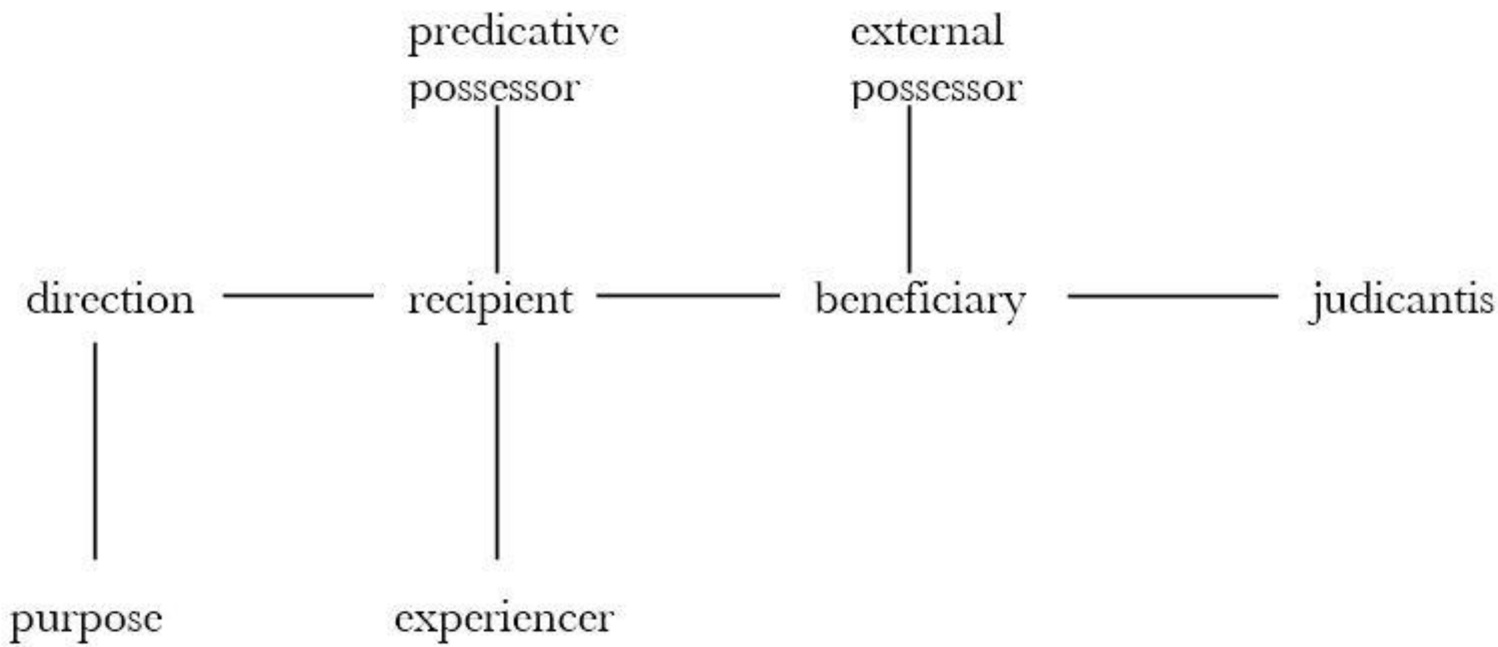


FIGURE. A semantic map of typical dative functions /
(based on Haspelmath, 2003: 213)

Introducing semantic maps

- 'A semantic map is a geometrical representation of functions (…) that are linked by connecting lines and thus constitute a network'

  (Haspelmath, 2003)

- 'A semantic map is a method for visually representing cross-linguistic regularity or universality in semantic structure'

  (Georgakopoulos & Polis, 2018; Georgakopoulos, 2019)

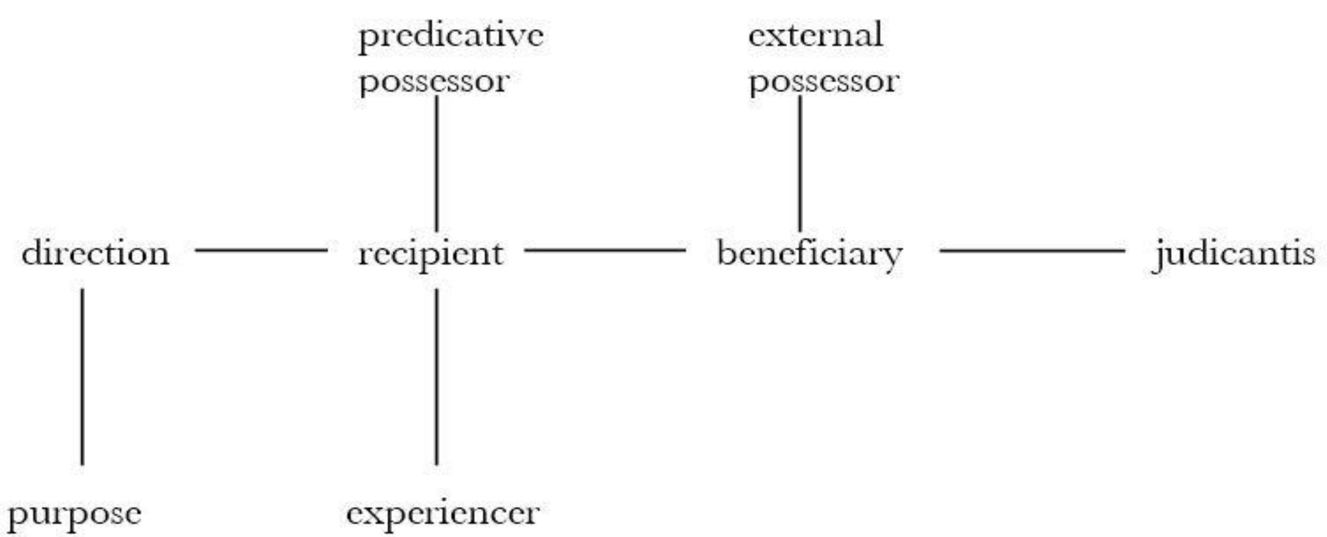


FIGURE. A semantic map of typical dative functions / (based on Haspelmath, 2003: 213)

Introducing semantic maps

- ‘A semantic map is a geometrical representation of functions (…) that are linked by connecting lines and thus constitute a network’

  (Haspelmath, 2003)

- ‘A semantic map is a method for visually representing cross-linguistic regularity or universality in semantic structure’

  (Georgakopoulos & Polis, 2018; Georgakopoulos, 2019)

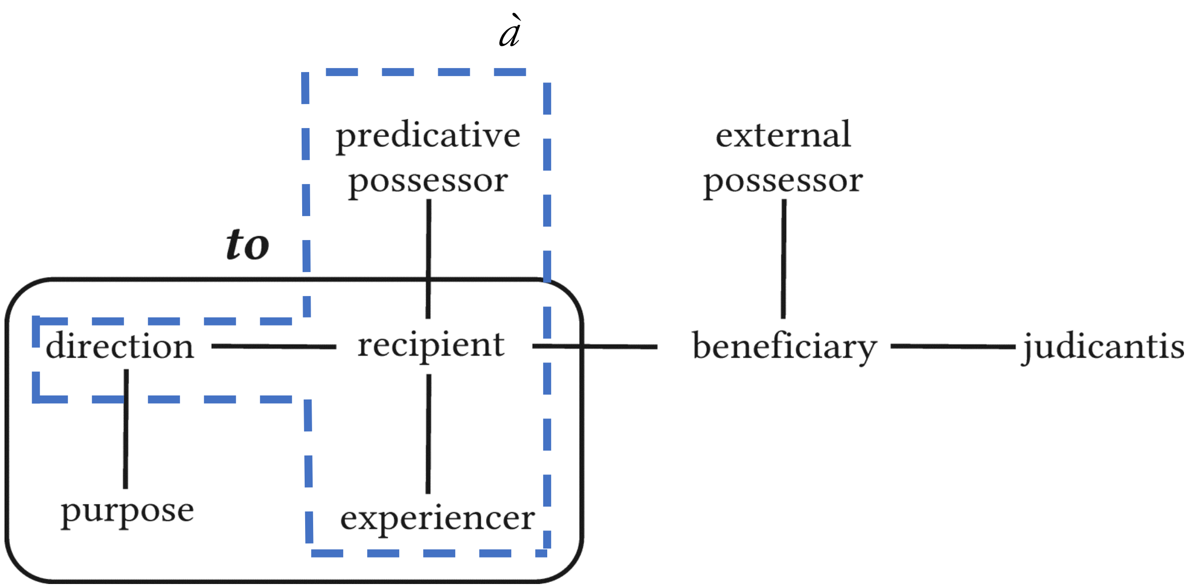


FIGURE. A semantic map of typical dative functions / the boundaries of English *to* and French *à*
(based on Haspelmath, 2003: 213, 215)

Semantic maps: Different types

- Two main types
  - Connectivity maps
  - Proximity maps

ARTICLE

WILEY

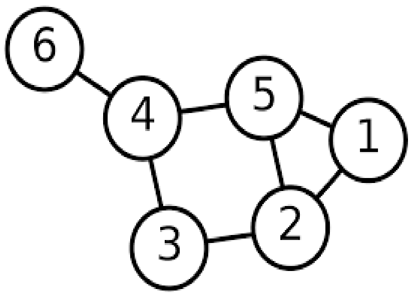
The semantic map model: State of the art and future avenues for linguistic research

Thanasis Georgakopoulos[1] | Stéphane Polis[2]

Semantic maps: Different types

- Two main types
  - **Connectivity maps**
  - Proximity maps

- Graph
  - **Nodes** = meanings
  - **Edges** = relationships between meanings



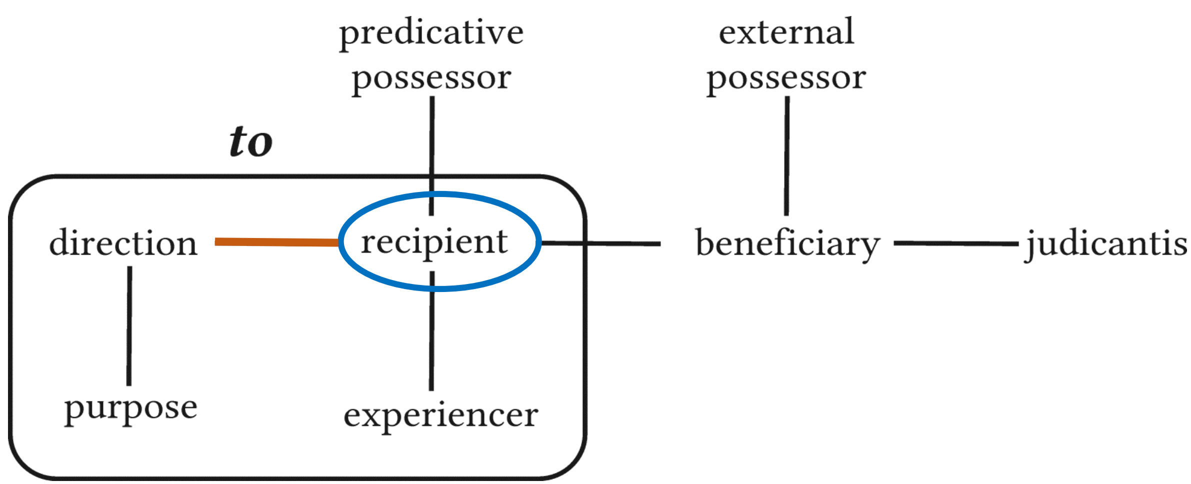


FIGURE. A semantic map of typical dative functions /
the boundaries of English *to* (based on Haspelmath, 2003: 213)

Semantic maps: Different types

- Two main types
  - Connectivity maps
  - **Proximity maps**

- Two-dimensional space
  - **Points** = meanings (or contexts)
  - **Proximity** = similarity between meanings (or contexts)

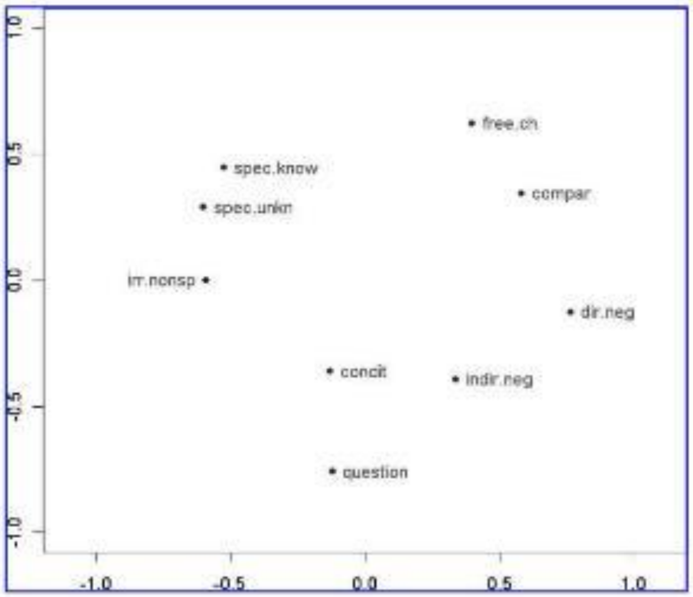


FIGURE. MDS analysis of Haspelmath's (1997) data on indefinite pronouns
(Croft & Poole, 2008: 15)

Semantic maps: Different types

- Two main types
  - Connectivity maps
  - **Proximity maps**

- Two-dimensional space
  - **Points** = meanings (or contexts)
  - **Proximity** = similarity between meanings (or contexts)

1. **Specific known**
   Somebody called you, guess who

2. **Specific unknown:**
   Somebody called you, but I don't know who

3. **Indirect negation:**
   I don't think that **anybody** called

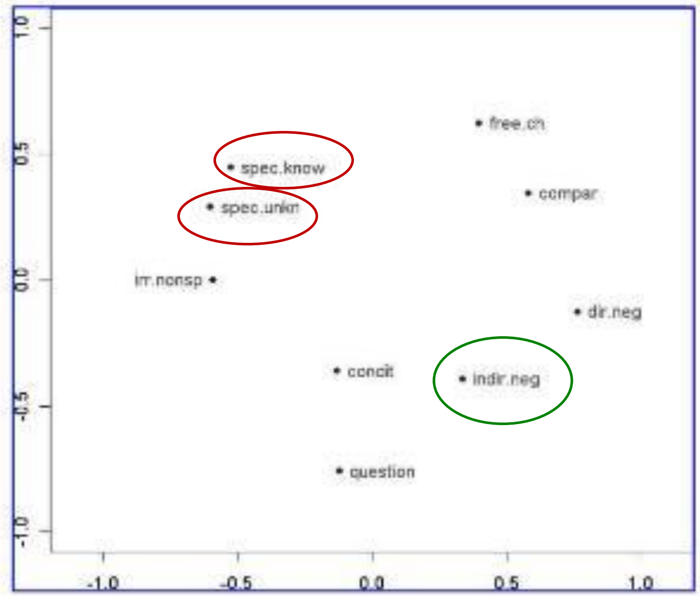


FIGURE. MDS analysis of Haspelmath's (1997) data on indefinite pronouns
(Croft & Poole, 2008: 15)

Classical semantic maps

How is a classical semantic map built?

TABLE Lexemes for tree/wood/forest in four languages

<table>
<tr><th colspan="4">Lexical items</th></tr>
<tr><th>Danish</th><th>French</th><th>German</th><th>Spanish</th></tr>
<tr><td rowspan="3">træ</td><td>arbre</td><td>Baum</td><td>árbol</td></tr>
<tr><td rowspan="3">bois</td><td rowspan="2">Holz</td><td>madera</td></tr>
<tr><td>leña</td></tr>
<tr><td rowspan="2">skov</td><td rowspan="2">Wald</td><td>bosque</td></tr>
<tr><td>forêt</td><td>selva</td></tr>
</table>

ARTICLE WILEY

The semantic map model: State of the art and future avenues for linguistic research

Thanasis Georgakopoulos[1] | Stéphane Polis[2]

How is a classical semantic map built?

- A new node can be added to the map if and only if there is at least one language with a dedicated linguistic form for this node

(Haspelmath, 2003; François, 2008; also Cysouw, 2007; 2010)

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="4">Lexical items</th></tr>
<tr><th>Danish</th><th>French</th><th>German</th><th>Spanish</th></tr>
<tr><td rowspan="5">ANALYTICAL PRIMITIVES</td><td>TREE</td><td rowspan="3">træ</td><td>arbre</td><td>Baum</td><td>árbol</td></tr>
<tr><td>WOOD (mat.)</td><td rowspan="3">bois</td><td rowspan="2">Holz</td><td>madera</td></tr>
<tr><td>FIREWOOD</td><td>leña</td></tr>
<tr><td>FOREST (small)</td><td rowspan="2">skov</td><td rowspan="2">Wald</td><td>bosque</td></tr>
<tr><td>FOREST (large)</td><td>forêt</td><td>selva</td></tr>
</table>

How is a classical semantic map built?

- **Onomasiology:**
  - Pick a domain
  - Identify the core meanings in this domain
  - Search for the individual forms that express these meanings in different languages

- **Semasiology:**
  - List in a lexical matrix all the meanings attested for each form of the language sample

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="4">Lexical items</th></tr>
<tr><th>Danish</th><th>French</th><th>German</th><th>Spanish</th></tr>
<tr><td rowspan="5">ANALYTICAL PRIMITIVES</td><td>TREE</td><td rowspan="3">træ</td><td>arbre</td><td>Baum</td><td>árbol</td></tr>
<tr><td>WOOD (mat.)</td><td rowspan="3">bois</td><td rowspan="2">Holz</td><td>madera</td></tr>
<tr><td>FIREWOOD</td><td>leña</td></tr>
<tr><td>FOREST (small)</td><td rowspan="2">skov</td><td rowspan="2">Wald</td><td>bosque</td></tr>
<tr><td>FOREST (large)</td><td>forêt</td><td>selva</td></tr>
</table>

How is a classical semantic map built?

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

ANALYTICAL PRIMITIVES	Lexical items			
	Danish	**French**	**German**	**Spanish**
TREE	*træ*	*arbre*	*Baum*	*árbol*
WOOD (mat.)		*bois*	*Holz*	*madera*
FIREWOOD				*leña*
FOREST (small)	*skov*		*Wald*	*bosque*
FOREST (large)		*forêt*		*selva*

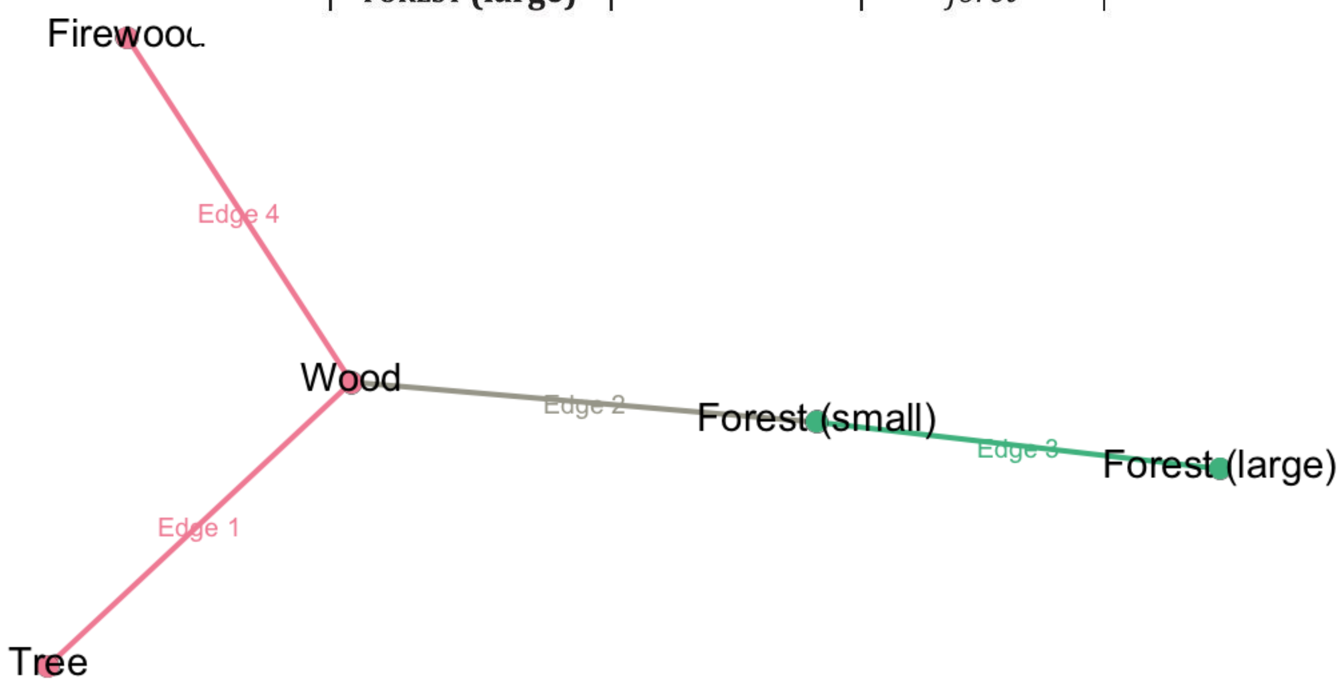


FIGURE. A semantic map inferred from the data in the Table

How is a classical semantic map built?

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

		Lexical items			
		Danish	**French**	**German**	**Spanish**
ANALYTICAL PRIMITIVES	TREE	*træ*	*arbre*	*Baum*	*árbol*
	WOOD (mat.)		*bois*	*Holz*	*madera*
	FIREWOOD				*leña*
	FOREST (small)	*skov*		*Wald*	*bosque*
	FOREST (large)		*forêt*		*selva*

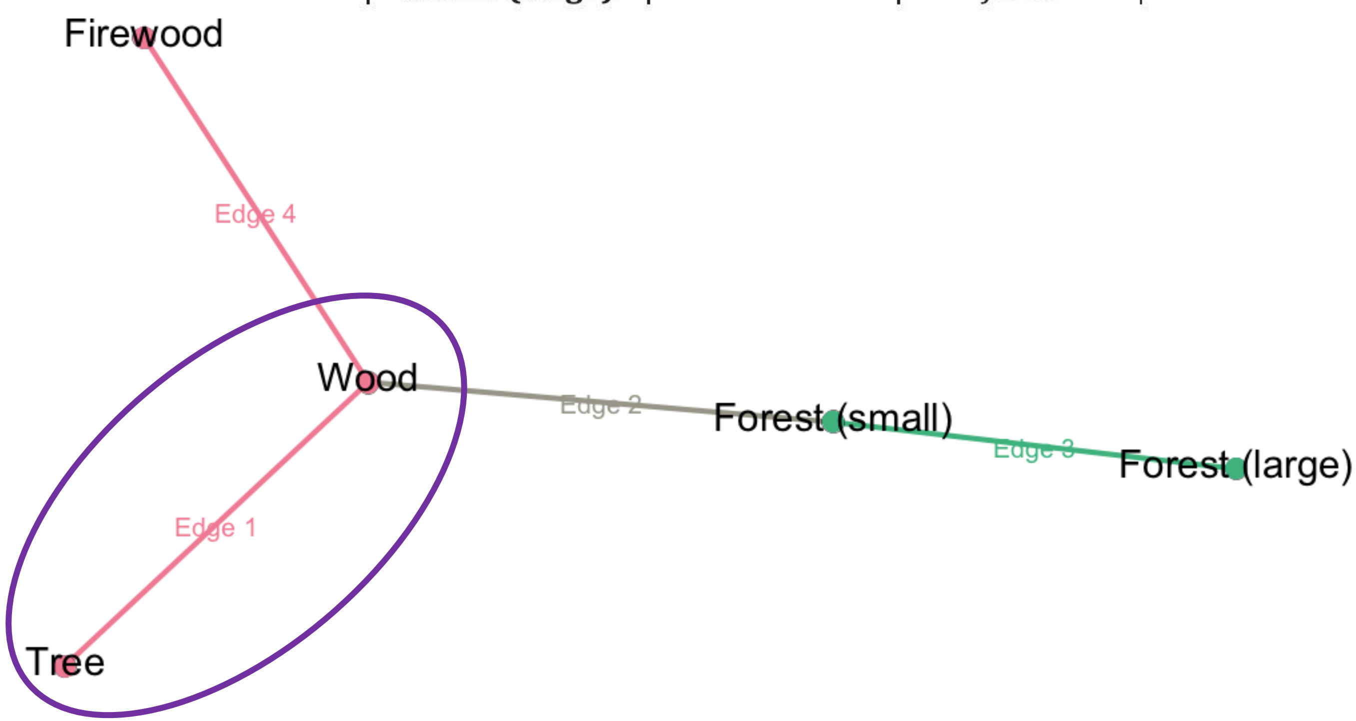


FIGURE. A semantic map inferred from the data in the Table

How is a classical semantic map built?

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

ANALYTICAL PRIMITIVES	Lexical items			
	Danish	**French**	**German**	**Spanish**
TREE	*træ*	*arbre*	*Baum*	*árbol*
WOOD (mat.)			*Holz*	*madera*
FIREWOOD		*bois*		*leña*
FOREST (small)	*skov*		*Wald*	*bosque*
FOREST (large)		*forêt*		*selva*

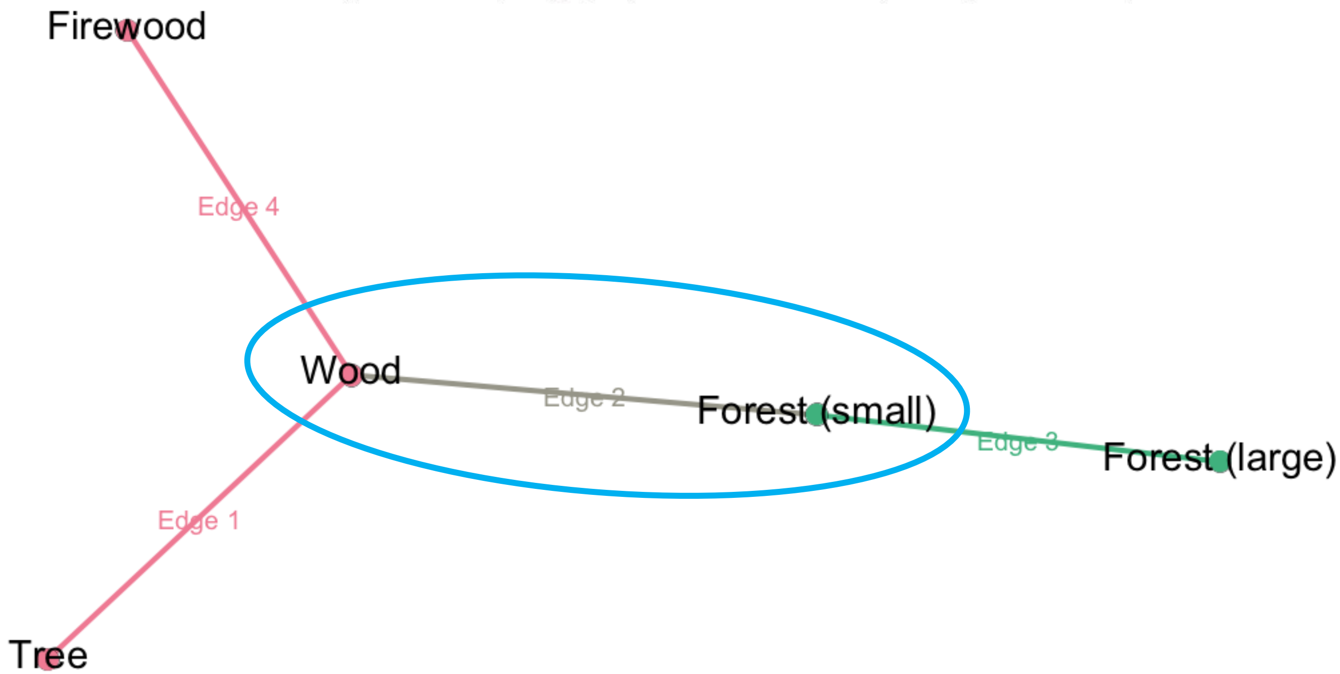


FIGURE. A semantic map inferred from the data in the Table

How is a classical semantic map built?

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

<table>
<tr><th colspan="2"></th><th colspan="4">Lexical items</th></tr>
<tr><th colspan="2"></th><th>Danish</th><th>French</th><th>German</th><th>Spanish</th></tr>
<tr><td rowspan="5">ANALYTICAL PRIMITIVES</td><td>TREE</td><td rowspan="3">træ</td><td>arbre</td><td>Baum</td><td>árbol</td></tr>
<tr><td>WOOD (mat.)</td><td rowspan="3">bois</td><td rowspan="2">Holz</td><td>madera</td></tr>
<tr><td>FIREWOOD</td><td>leña</td></tr>
<tr><td>FOREST (small)</td><td rowspan="2">skov</td><td rowspan="2">Wald</td><td>bosque</td></tr>
<tr><td>FOREST (large)</td><td>forêt</td><td>selva</td></tr>
</table>

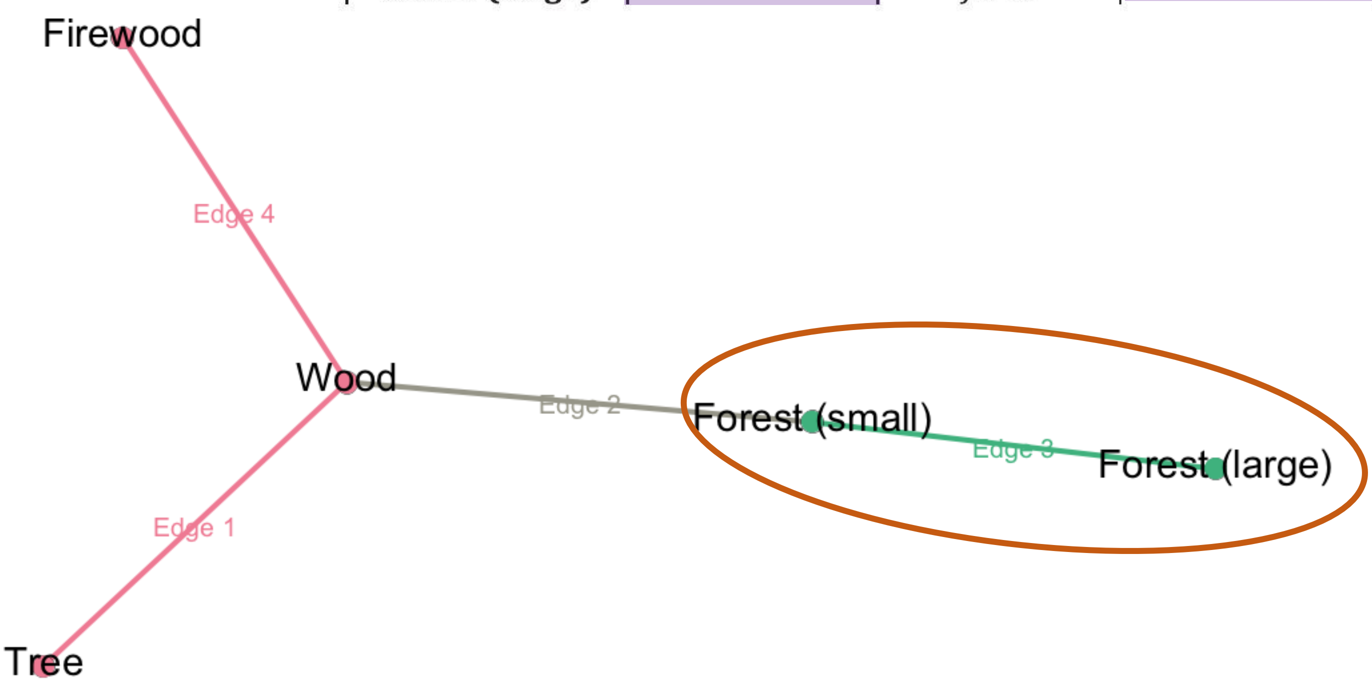


FIGURE. A semantic map inferred from the data in the Table

How is a classical semantic map built?

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

		Lexical items			
		Danish	**French**	**German**	**Spanish**
ANALYTICAL PRIMITIVES	**TREE**		*arbre*	*Baum*	*árbol*
	WOOD (mat.)	*træ*	*bois*	*Holz*	*madera*
	FIREWOOD				*leña*
	FOREST (small)	*skov*		*Wald*	*bosque*
	FOREST (large)		*forêt*		*selva*

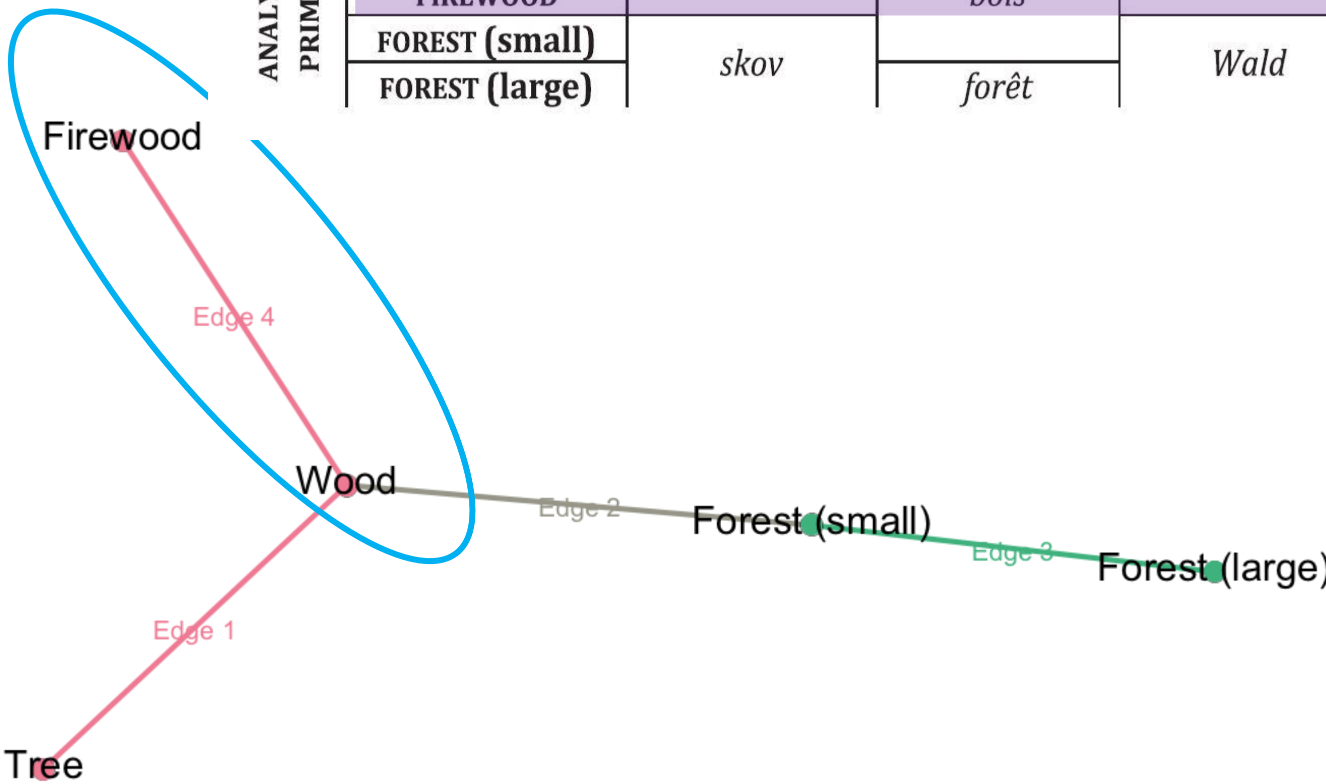


FIGURE. A semantic map inferred from the data in the Table

How is a classical semantic map built?

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

		Lexical items			
		Danish	**French**	**German**	**Spanish**
ANALYTICAL PRIMITIVES	**TREE**	*træ*	*arbre*	*Baum*	*árbol*
	WOOD (mat.)		*bois*	*Holz*	*madera*
	FIREWOOD				*leña*
	FOREST (small)	*skov*		*Wald*	*bosque*
	FOREST (large)		*forêt*		*selva*

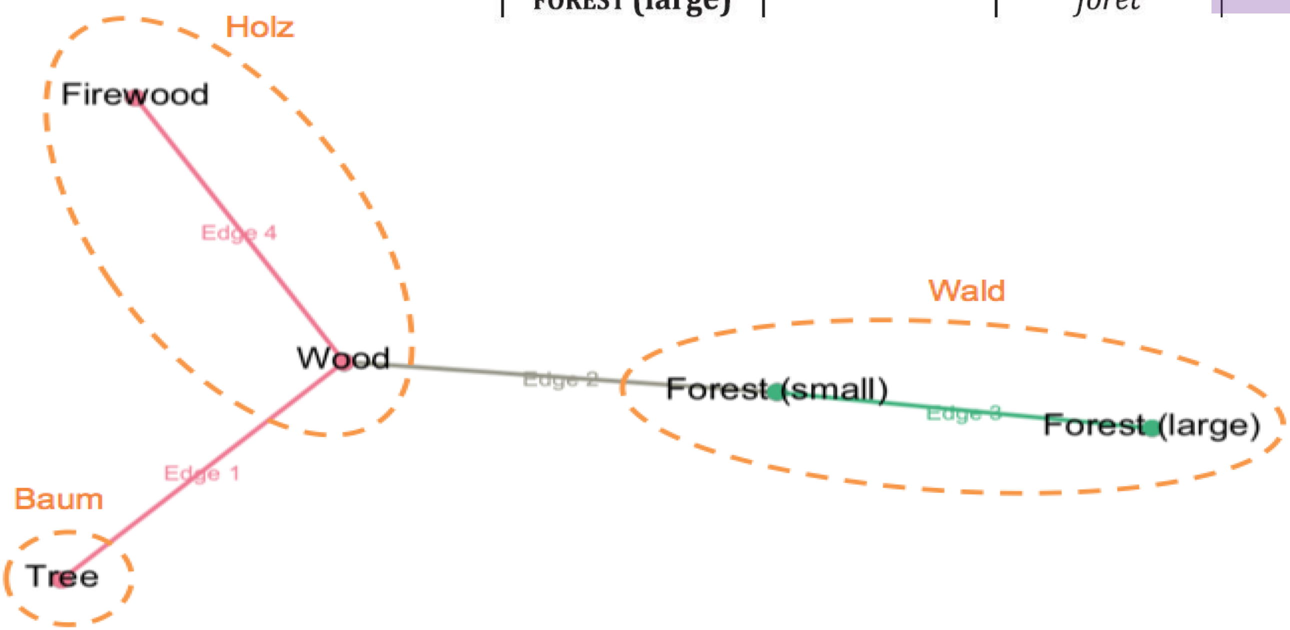


FIGURE. A semantic map inferred from the data in Table 3, with the German lexemes mapped onto the nodes

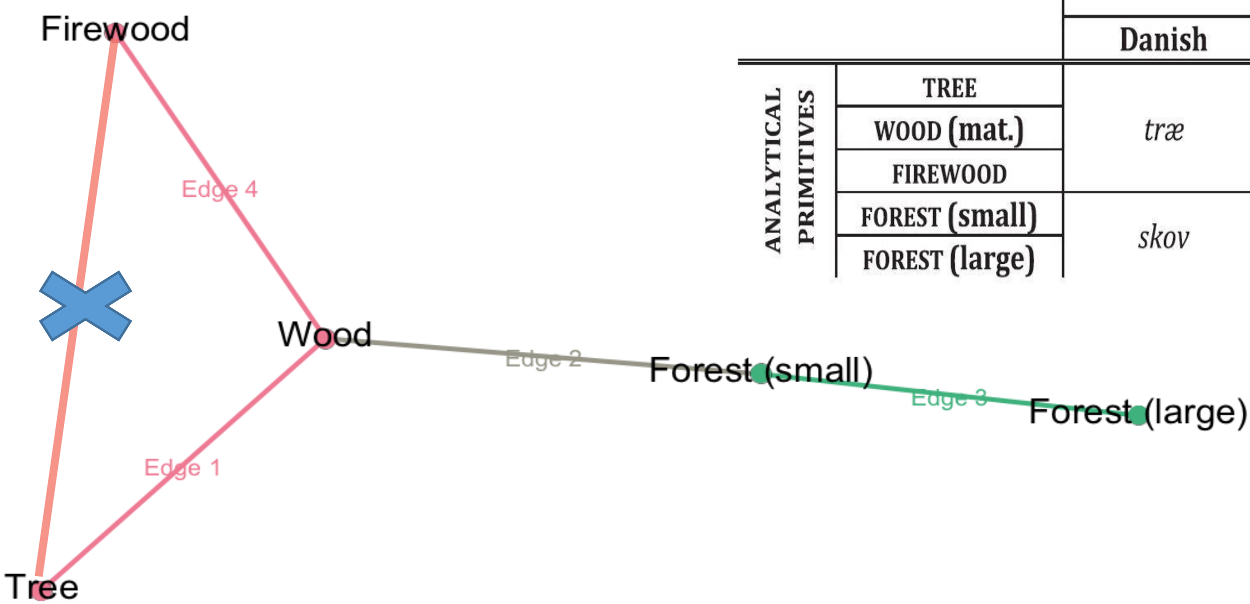
How is a classical semantic map built?

Economy principle:
Given three meanings (**Meaning1**, **Meaning2**, **Meaning3**), if the linguistic items expressing **Meaning1** and **Meaning3** always express **Meaning2**, there is no need to draw an edge between **Meaning1** and **Meaning3**

(Regier et al., 2013; Georgakopoulos & Polis, 2018)

TABLE Partitioning of the TREE–WOOD–FOREST semantic domain

		Lexical items			
		Danish	French	German	Spanish
ANALYTICAL PRIMITIVES	TREE		*arbre*	*Baum*	*árbol*
	WOOD (mat.)	*træ*		*Holz*	*madera*
	FIREWOOD		*bois*		*leña*
	FOREST (small)	*skov*		*Wald*	*bosque*
	FOREST (large)		*forêt*		*selva*



Colexification networks vs. semantic maps

FIGURE. Colexification network of the domains of perception and cognition (data from CLICS[2])

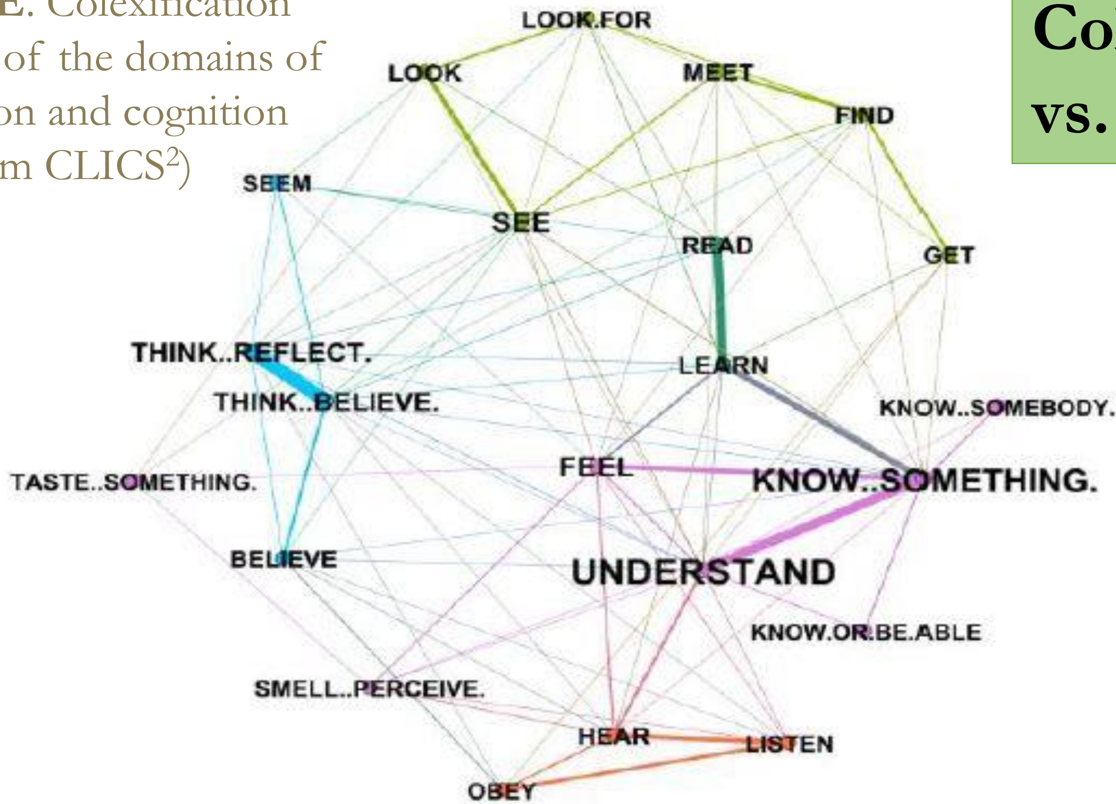


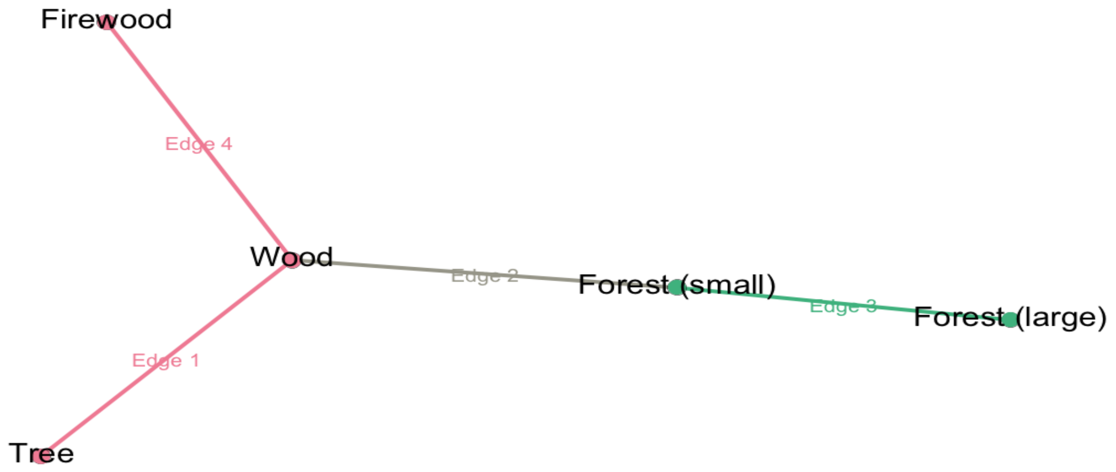
FIGURE. Semantic map of the domains of perception and cognition (data from CLICS[2])

(See Georgakopoulos et al. *under revision*. Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain. *Linguistic Typology*)

Advantages of the semantic map model

- It is neutral with respect to the monosemy/vagueness–polysemy–homonymy distinction
  (Haspelmath, 2003)

- They are at the same time implicational (Haspelmath, 1997) and falsifiable (Cysouw et al., 2010, p. 1).



Advantages of the semantic map model

- They have proven to be an efficient tool in historical linguistics, and especially in grammaticalization studies

(e.g., Narrog & van der Auwera, 2011)

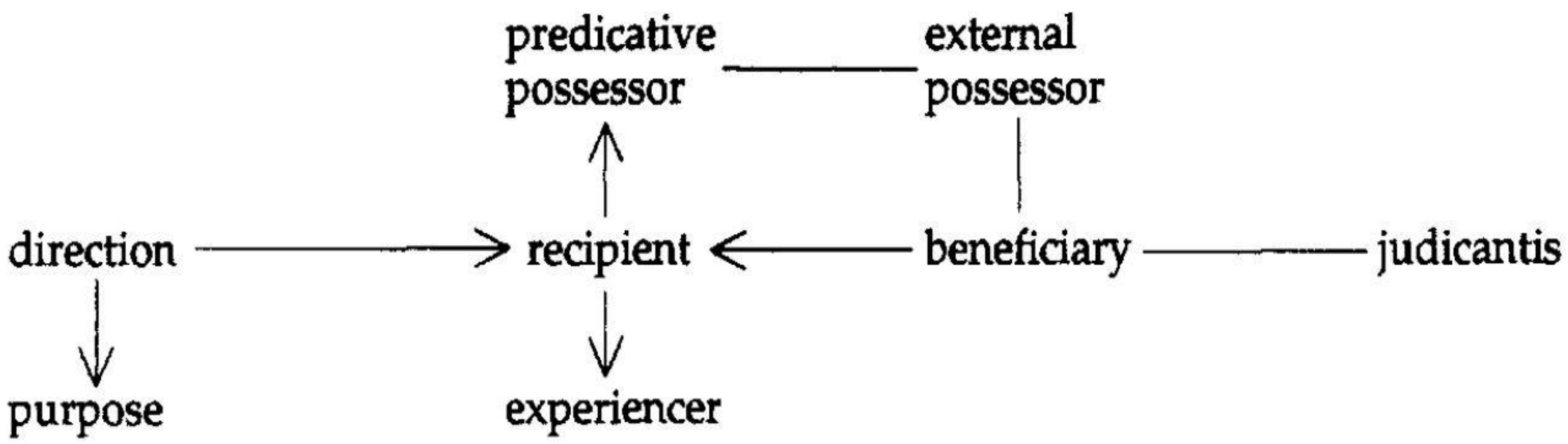


FIGURE. A dynamicized semantic map of dative functions (Haspelmath, 2003: 234)

Advantages of the semantic map model

- They can integrate information about the frequency of polysemy patterns

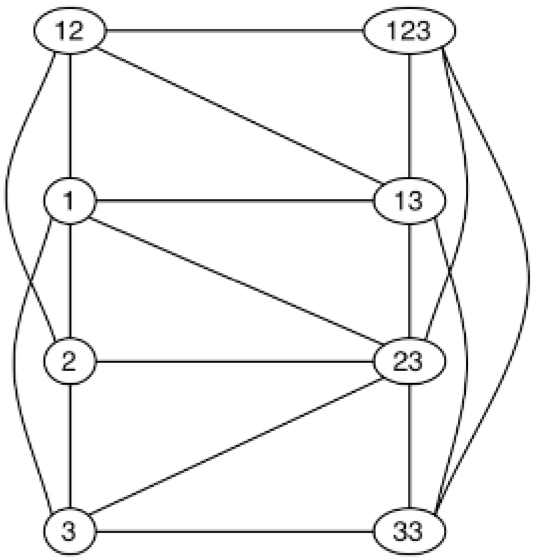


FIGURE a. A simple semantic map of person marking (Cysouw, 2007: 231)

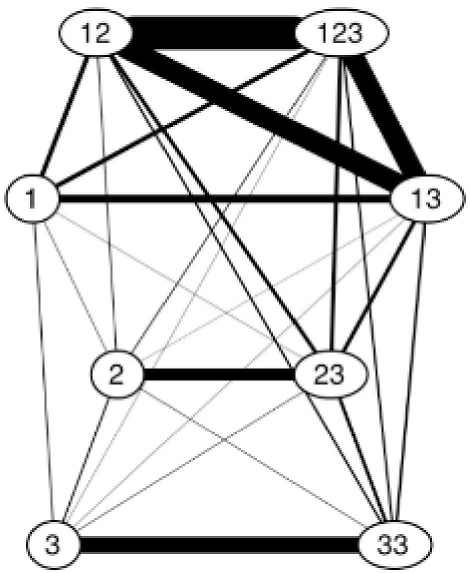


FIGURE b. A weighted semantic map of person marking (Cysouw, 2007: 233)

Advantages of the semantic map model

- The meanings of semantic maps can be of any kind: "grammatical," "lexical," or "constructional"

Lexical semantic maps

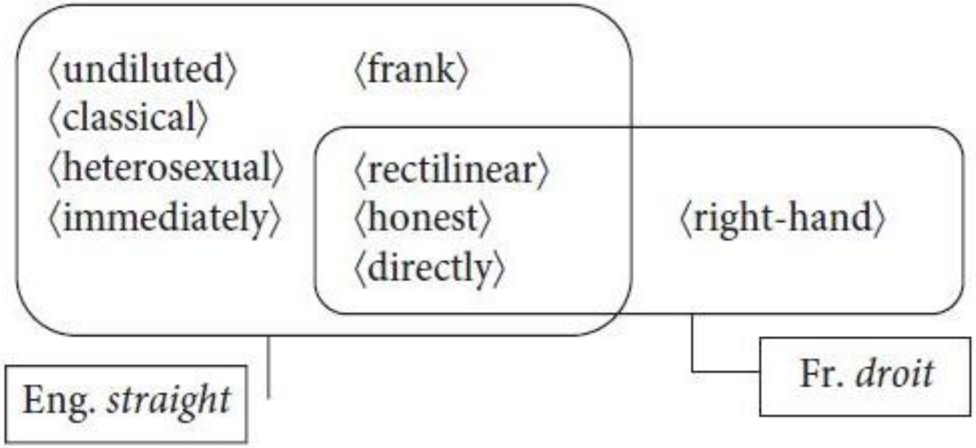


FIGURE. Overlapping polysemies:
Eng. straight *vs.* Fr. droit
(François, 2008: 167)

<u>Terms</u>

colexification = polyfunctionality = multifunctionality = polysemy

"A given language is said to colexify two functionally distinct senses if, and only if, it can associate them with the same lexical form"
(François, 2008: 170)

Advantages of the semantic map model

- The meanings of semantic maps can be of any kind: “grammatical,” “lexical,” or “constructional”

Lexical semantic maps

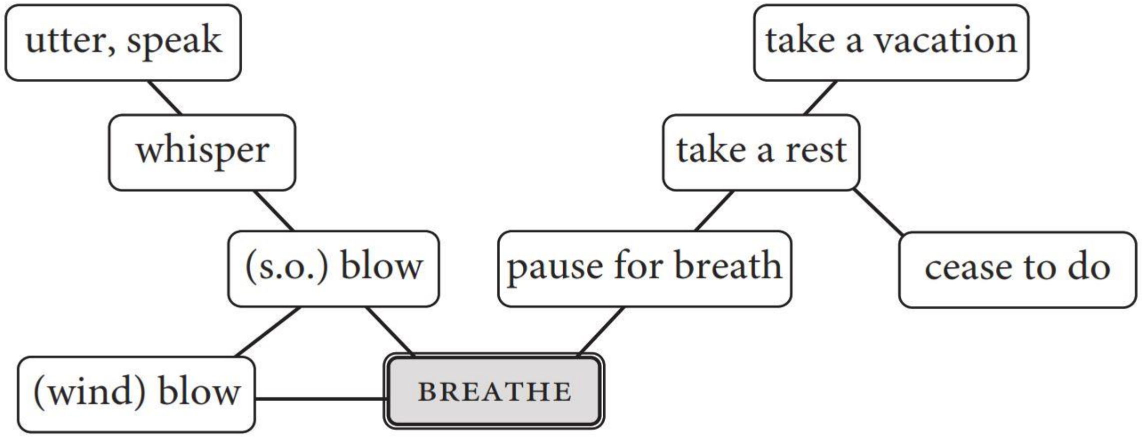
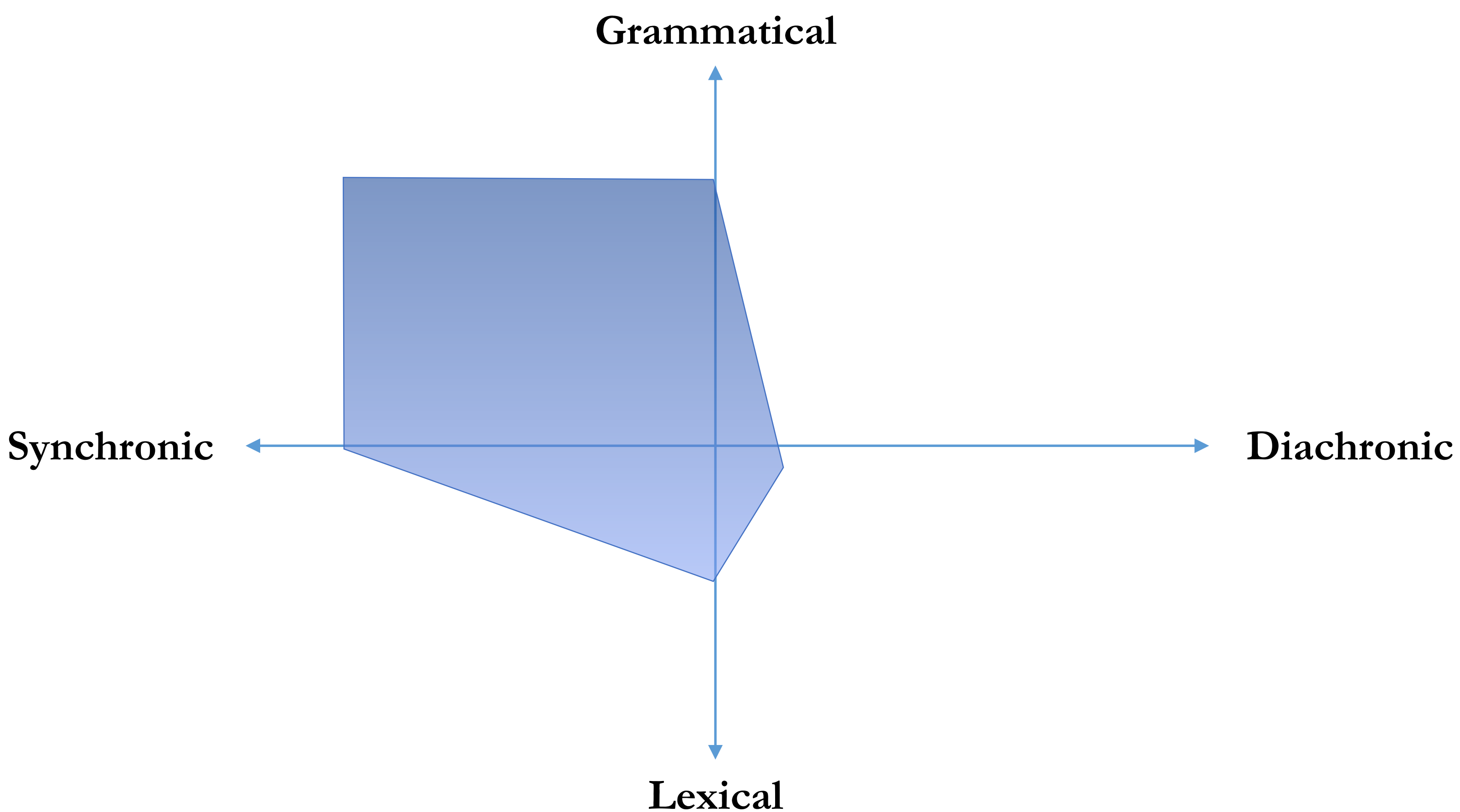


FIGURE. A (partial) semantic map for BREATHE
(François, 2008: 185)

Terms
colexification = polyfunctionality = multifunctionality = polysemy

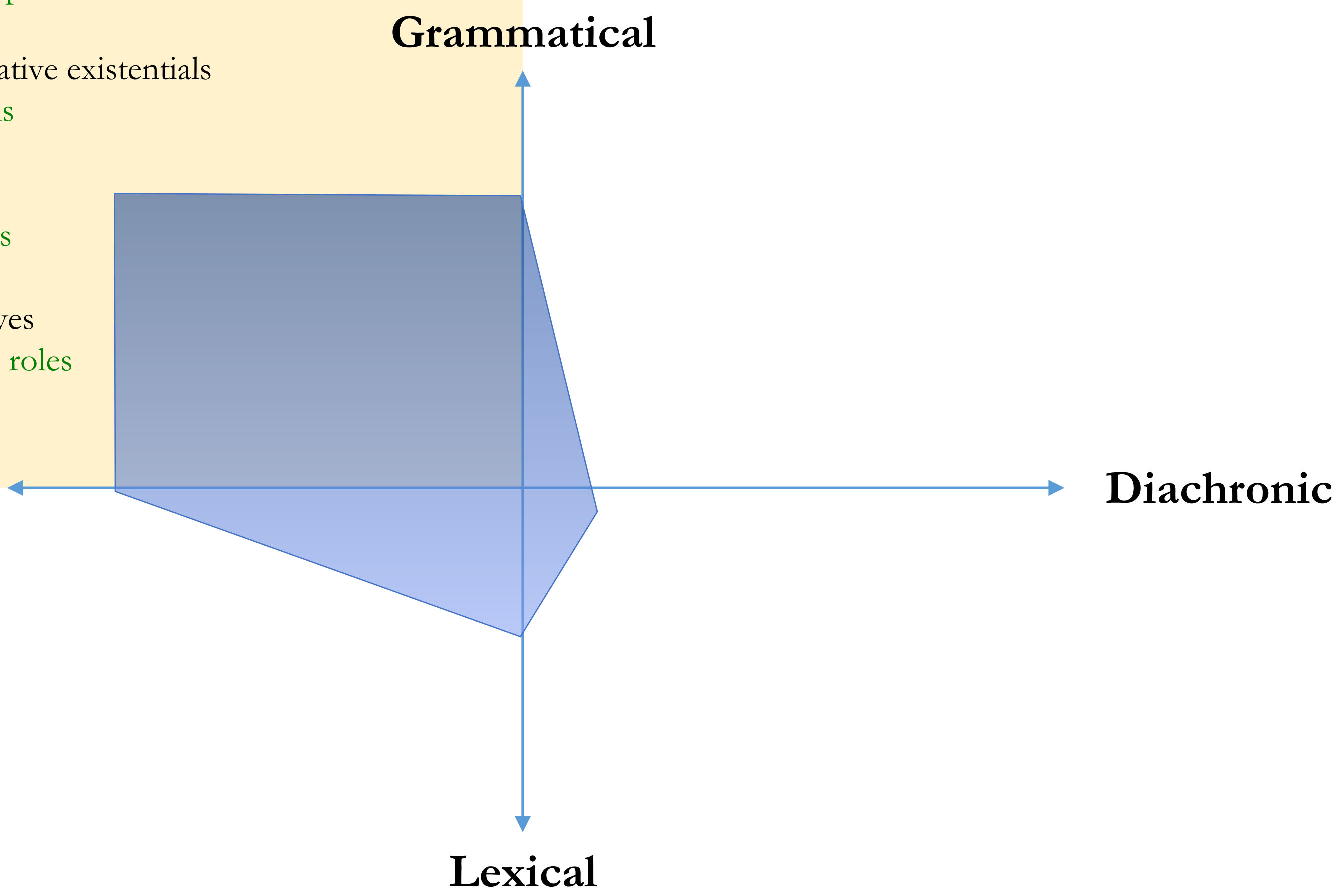
Different domains



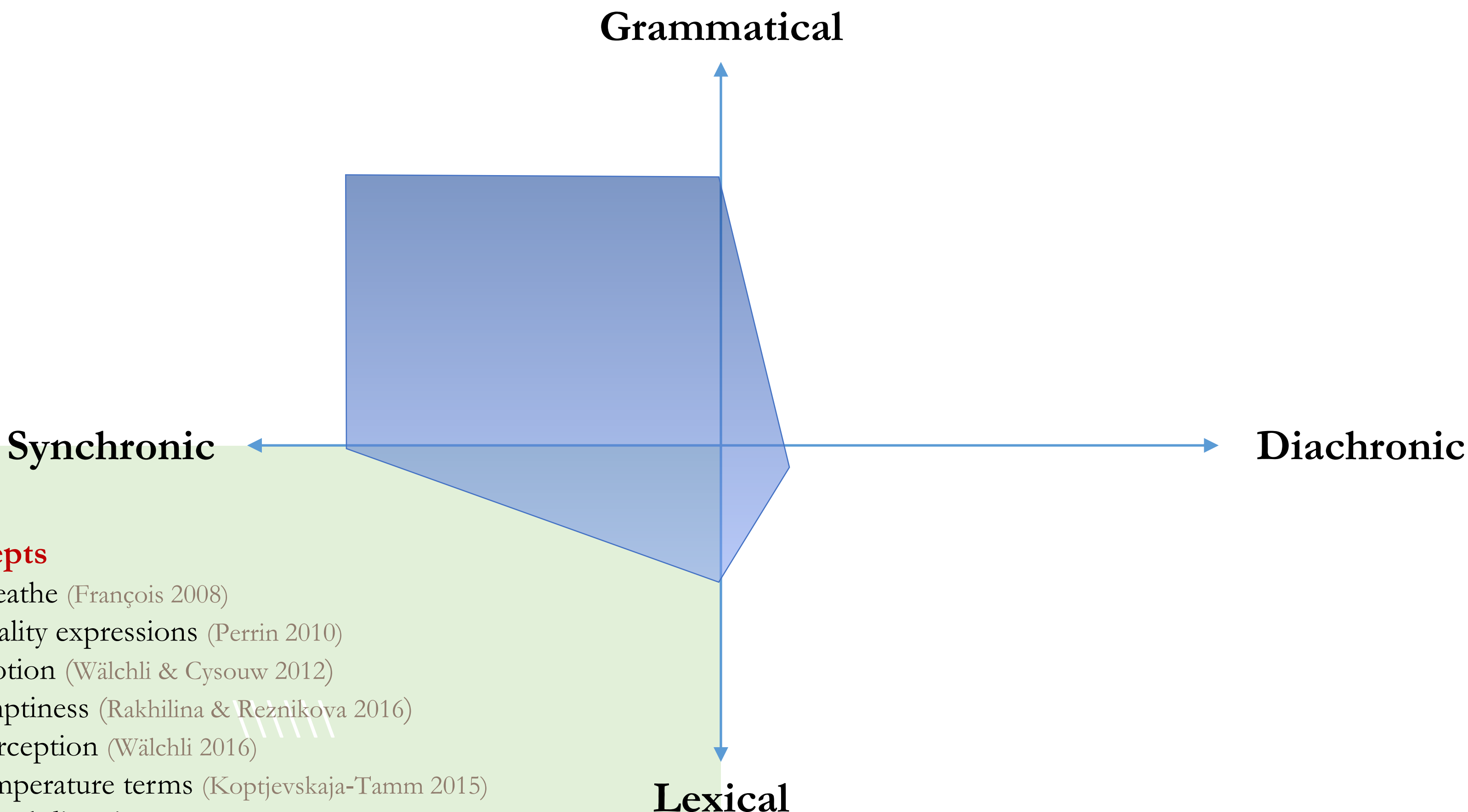
Different domains

Domains (not exhaustive)

• additives • adversatives • temporal markers
• conditionals • coordination
• negative polarity items • negative existentials
• core arguments • impersonals
• modality • middle voice
• secondary predication
• tense and aspect • intensifiers
• imperative-hortatives
• indefinite pronouns • partitives
• complementation • semantic roles



Different domains

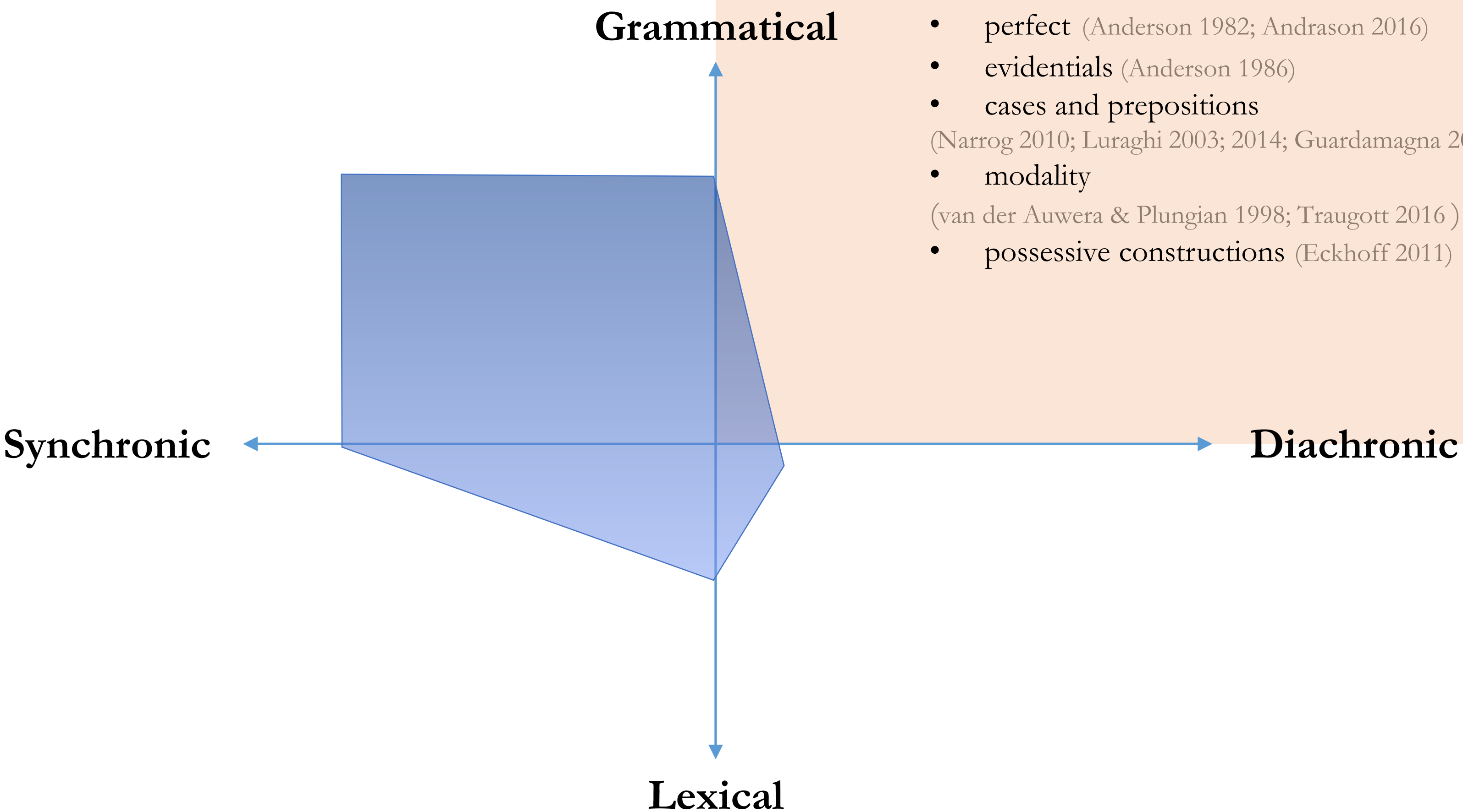


Concepts

- breathe (François 2008)
- quality expressions (Perrin 2010)
- motion (Wälchli & Cysouw 2012)
- emptiness (Rakhilina & Reznikova 2016)
- perception (Wälchli 2016)
- temperature terms (Koptjevskaja-Tamm 2015)
- visual direction (Rakhilina *et al.* 2017)
- natural and spatial features

(Georgakopoulos *et al.* 2016; Youn *et al.* 2016)

Different domains



Concepts

- perfect (Anderson 1982; Andrason 2016)
- evidentials (Anderson 1986)
- cases and prepositions

(Narrog 2010; Luraghi 2003; 2014; Guardamagna 2016)

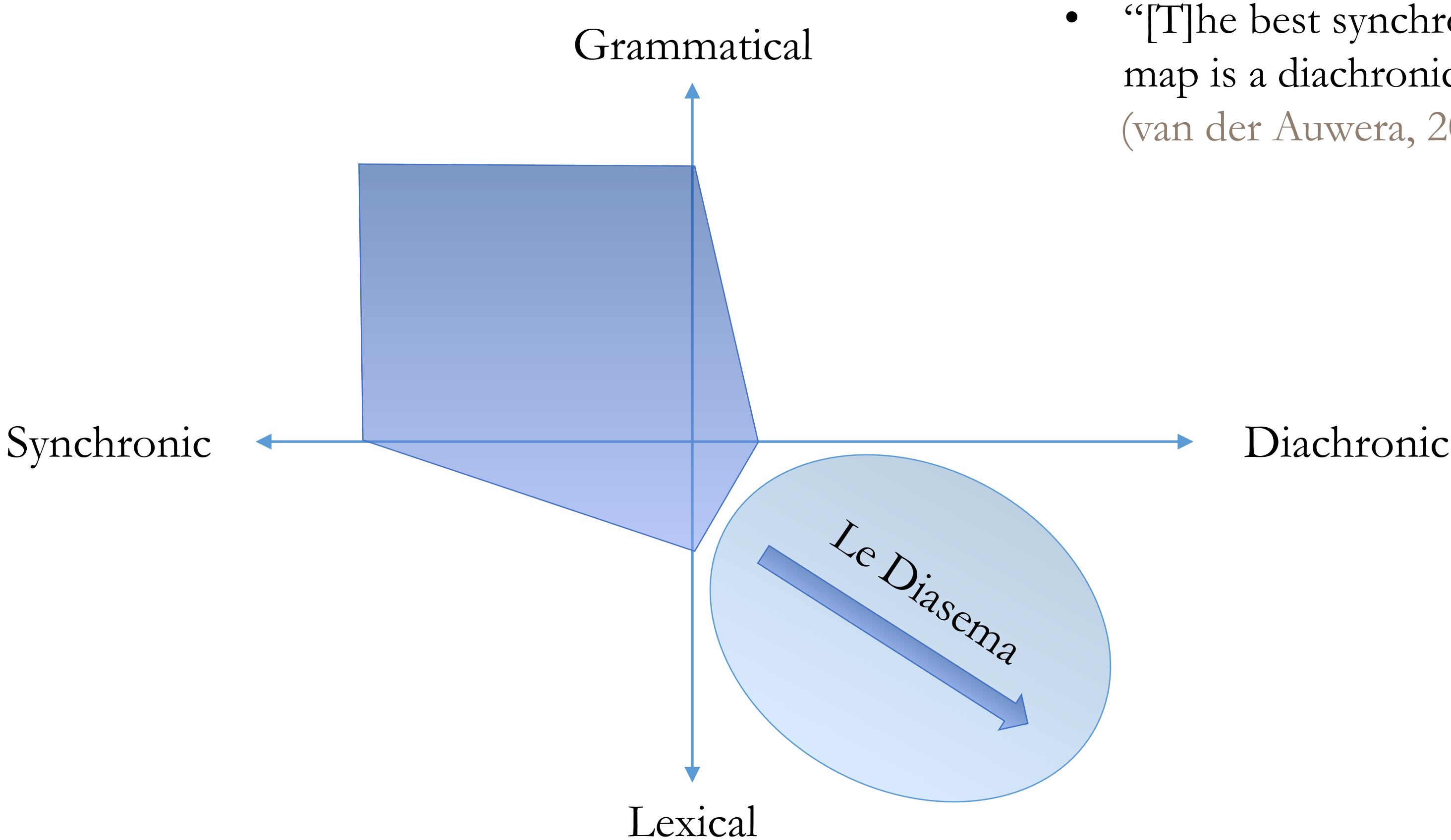
- modality

(van der Auwera & Plungian 1998; Traugott 2016)

- possessive constructions (Eckhoff 2011)

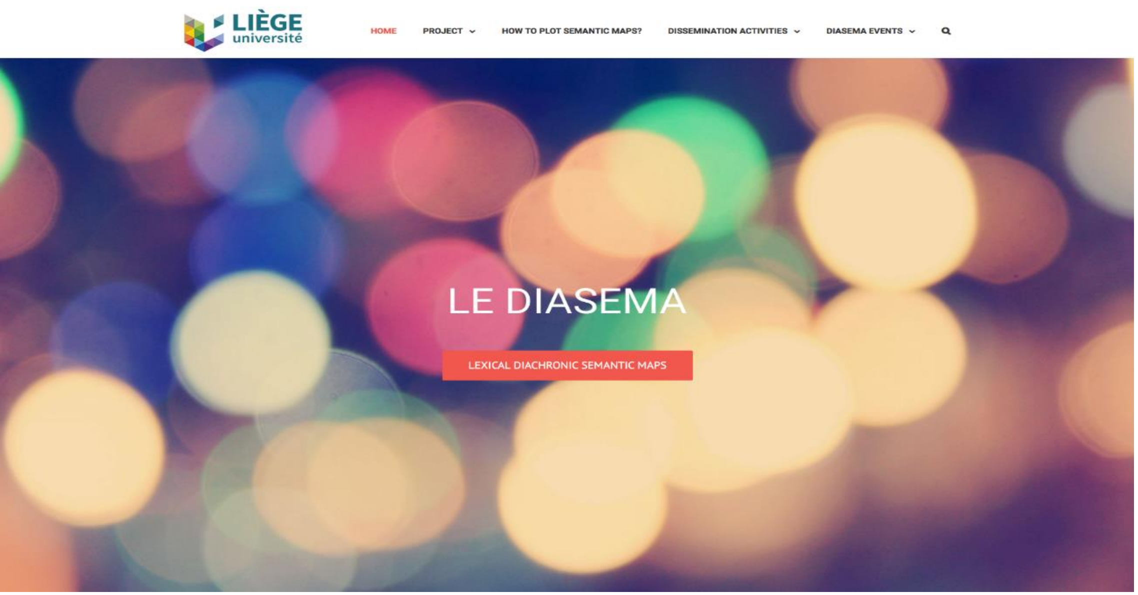
Different domains

- Adding the diachronic dimension to semantic maps of content words



- "[T]he best synchronic semantic map is a diachronic one" (van der Auwera, 2008: 43)

http://web.philo.ulg.ac.be/lediasema/



http://web.philo.ulg.ac.be/lediasema/

Thanasis GEORGAKOPOULOS

Stéphane POLIS

Problems with classical semantic maps

- They predict much more than is actually found

  (Cysouw, 2001, pp. 609–610)

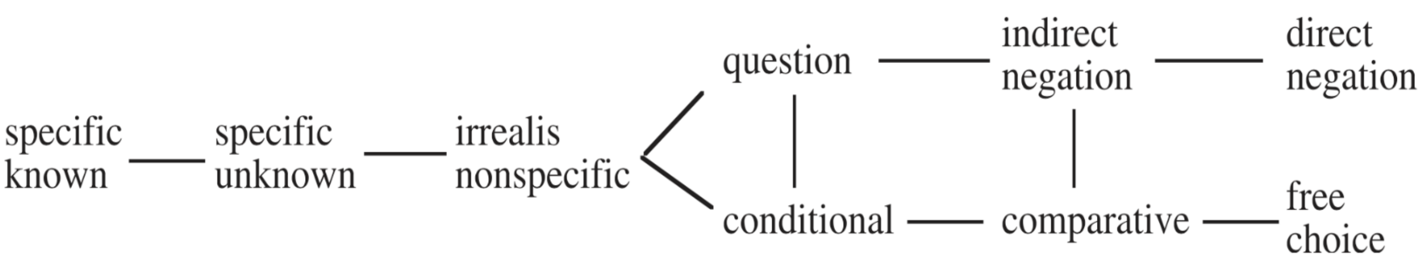


FIGURE. Semantic map of the indefinite pronouns functions (Haspelmath, 1997, p. 4)

- There are **105** different possibilities for mapping a linguistic form
  - Haspelmath (2003, p. 76): only **39** different kinds of mapping are actually found in his dataset

Problems with classical semantic maps

- Practically impossible to handle large-scale crosslinguistic datasets manually

  - “not mathematically well-defined or computationally tractable, making it impossible to use with large and highly variable crosslinguistic datasets”

  (Croft & Poole 2008: 1)

> “ideally (…) it should be possible to generate semantic maps automatically on the basis of a given set of data”
> (Narrog & Ito 2007: 280)

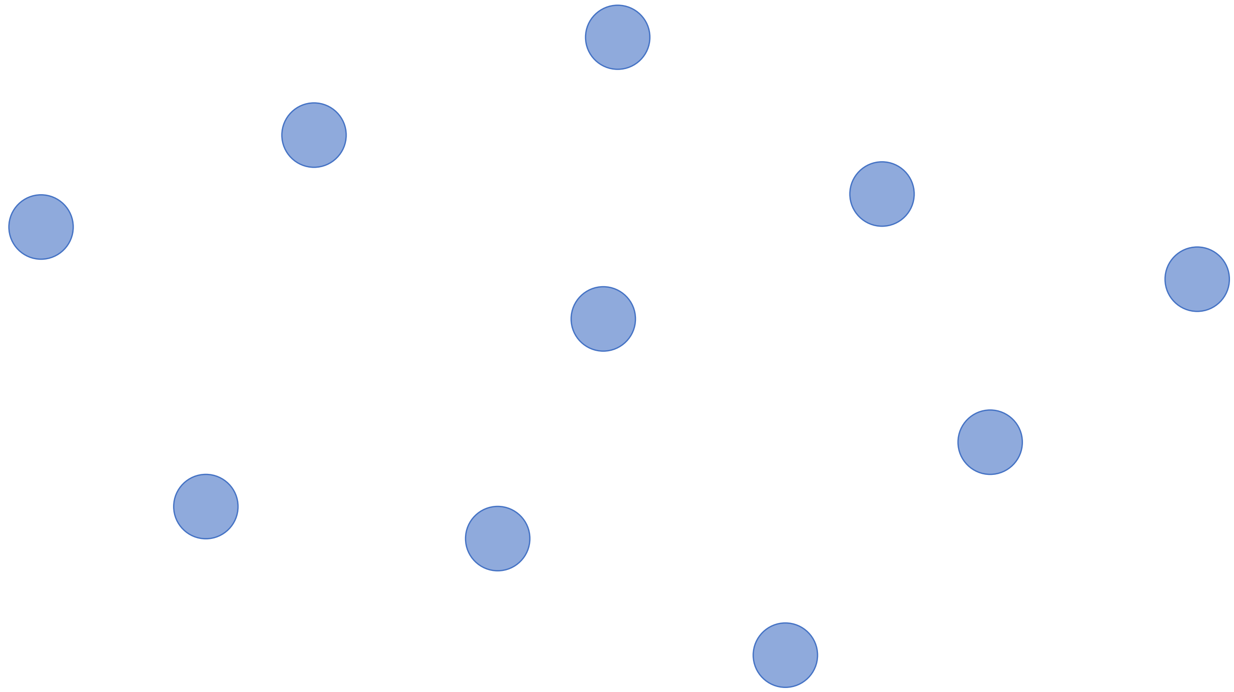
A way out: Inferring semantic maps

Inferring semantic maps

- Regier, Khetarpal, and Majid showed that the semantic map inference problem is "formally identical to another problem that superficially appears unrelated: inferring a social network from outbreaks of disease in a population" (Regier et al., 2013: 91)

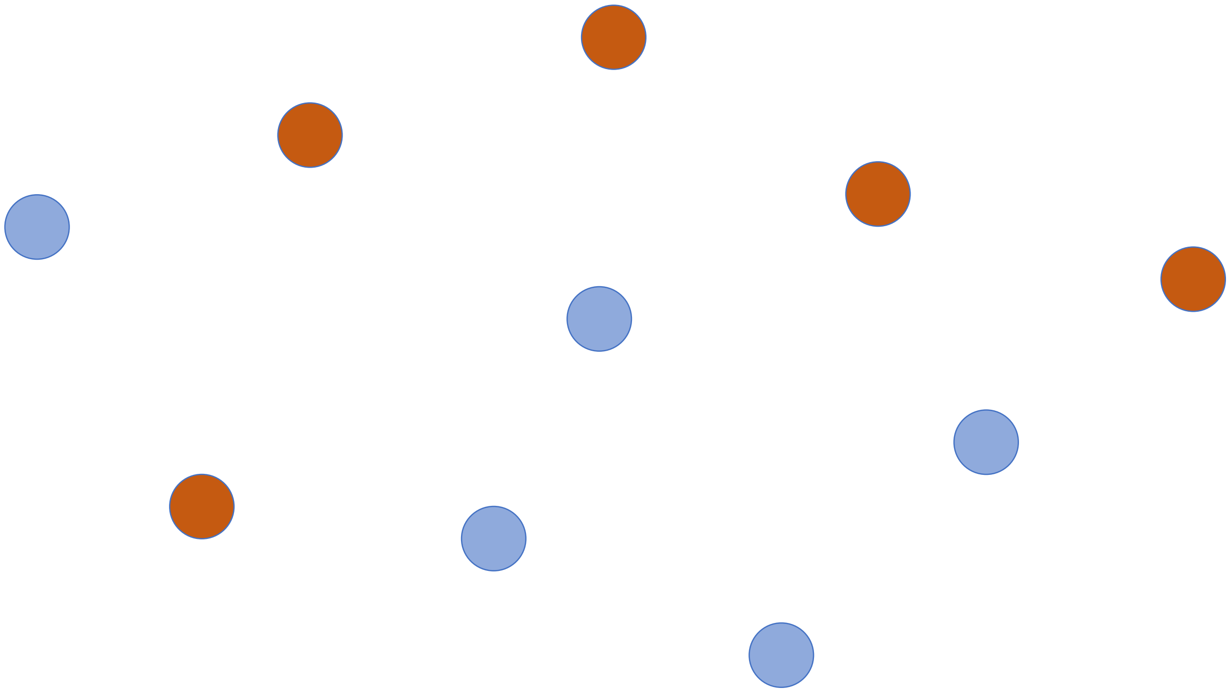
Inferring semantic maps

- What's the idea?
  - Let's consider a group of social agents (represented by the nodes of a potential graph)



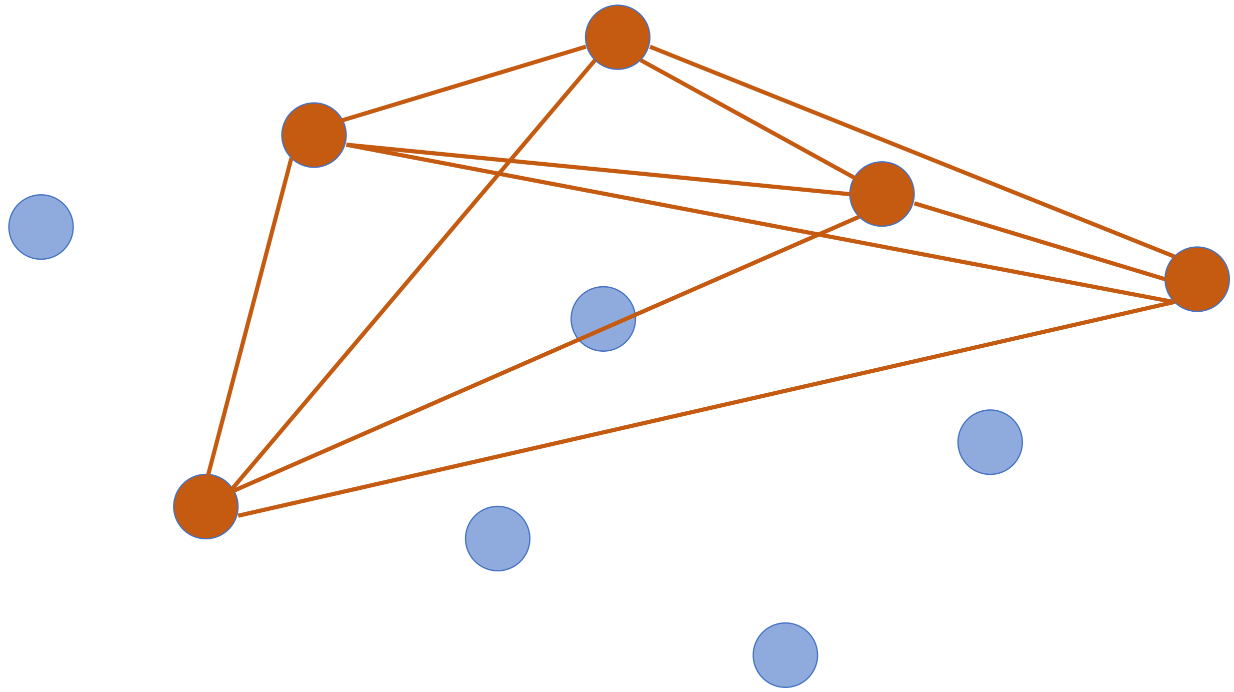
Inferring semantic maps

- What's the idea?
  - If one observes the same disease for five of these agents (technically called a constraint on the nodes of the graph)



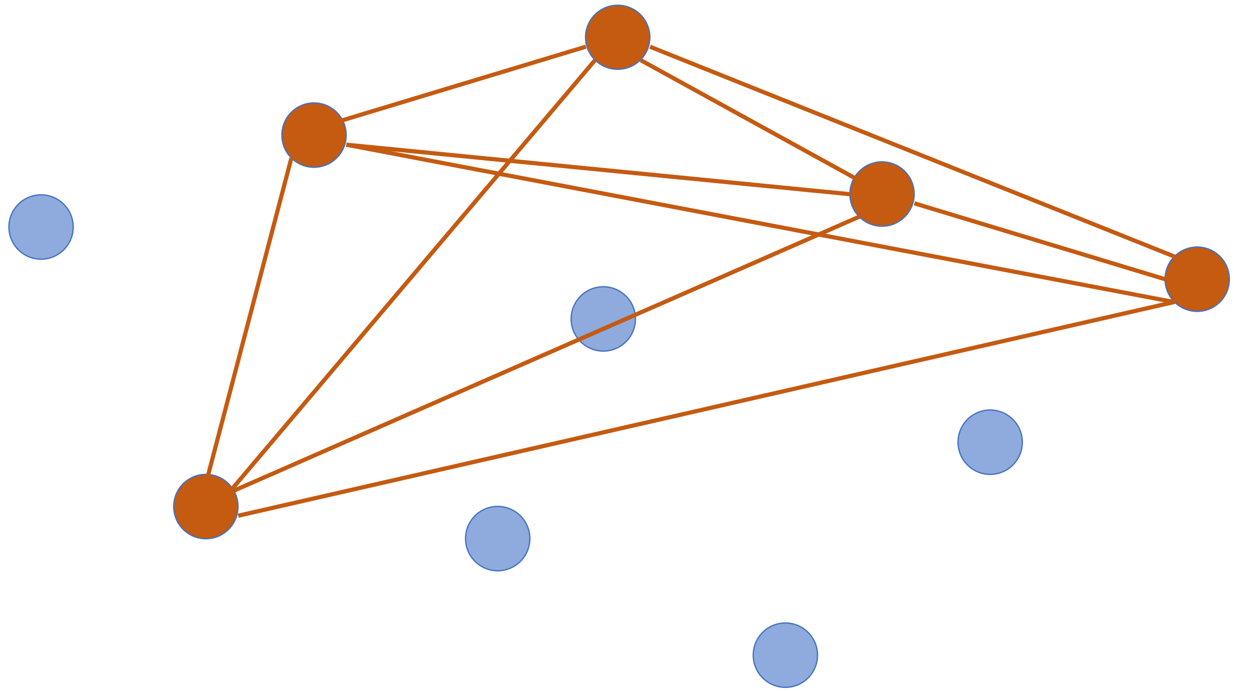
Inferring semantic maps

- What's the idea?
  - One can postulate that all the agents met, so that all the nodes of the graph are connected (10 edges between the 5 nodes)



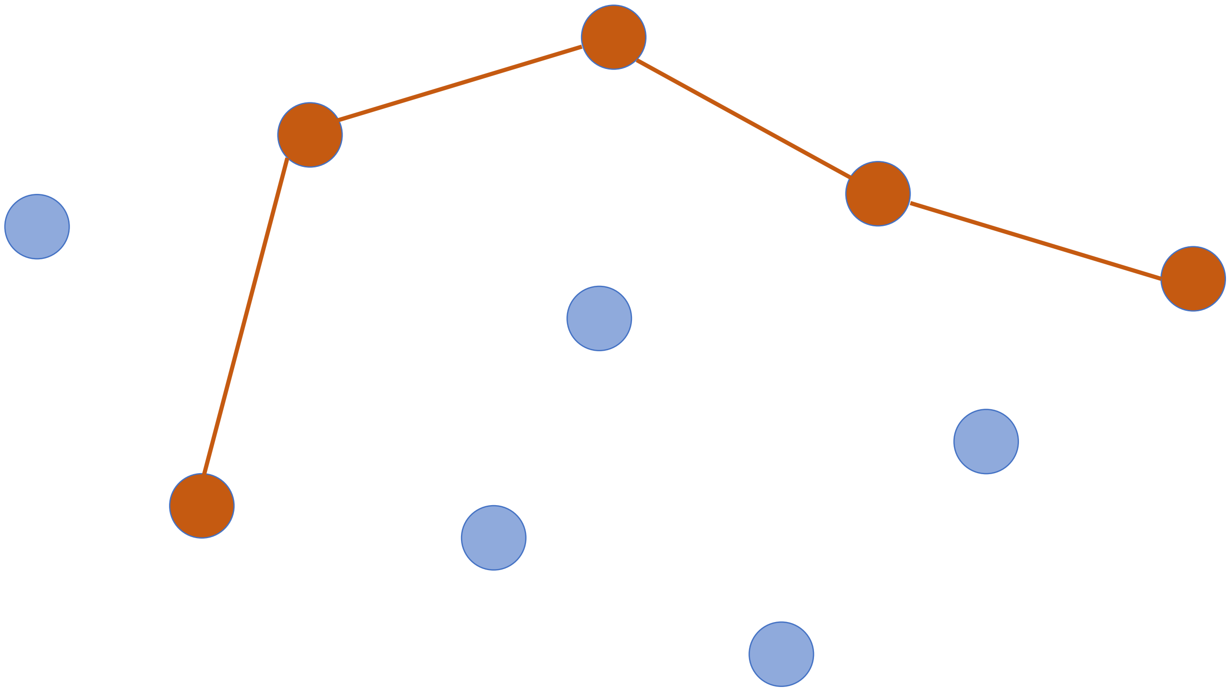
Inferring semantic maps

- What's the idea?
  - This is neither a very likely, nor a very economic explanation



Inferring semantic maps

- What's the idea?
  - The goal would be to have all the social agents connected with as few edges as possible



Inferring semantic maps

- What's the idea?
    - Such a Network Inference problem looks intuitively simple, but is computationally hard to solve

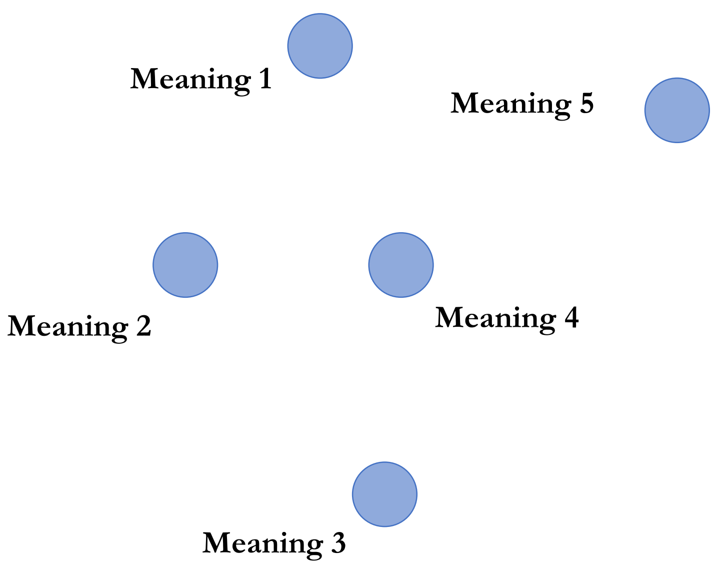
    - Angluin et al. (2010) concluded that the problem is indeed computationally intractable, but proposed an algorithm that approximates the optimal solution nearly as well as is theoretically possible

Inferring Social Networks from Outbreaks

Dana Angluin[1,*], James Aspnes[1,**], and Lev Reyzin[2,***]

Inferring semantic maps

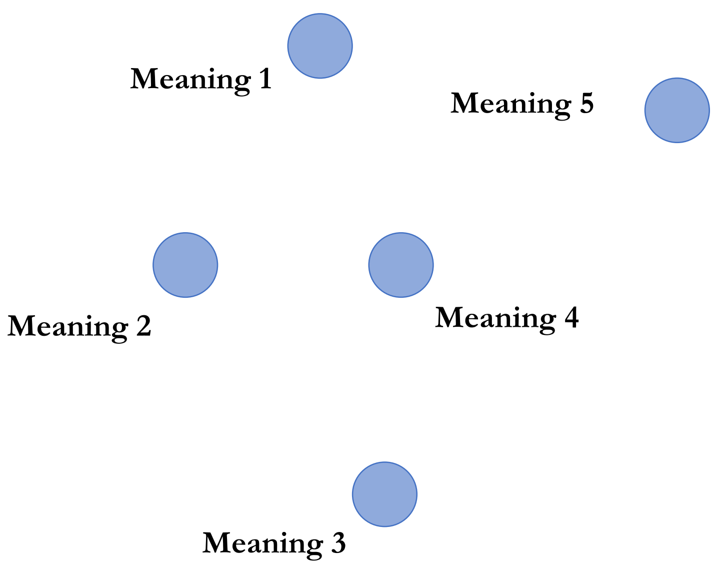
- How does it transfer to semantic maps?
  - Nodes are meanings



Meaning	1	2	3	4	5

Inferring semantic maps

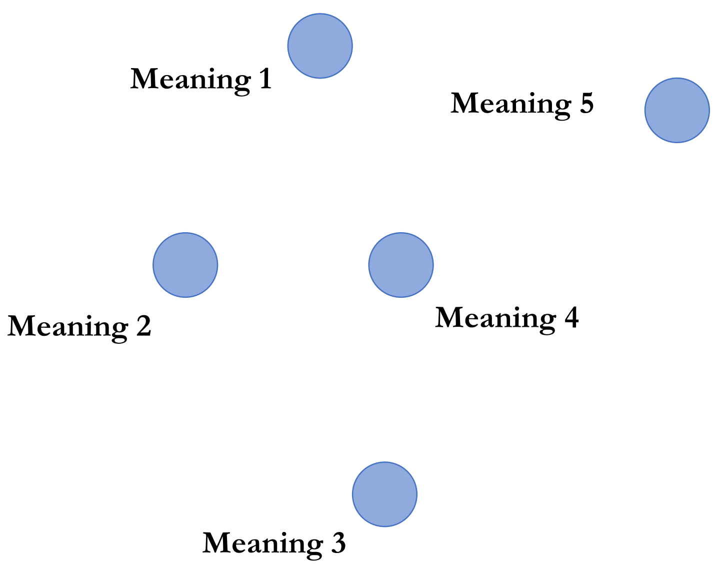
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are Polysemic items



Meaning	1	2	3	4	5
Polysemic item A	√	√			
Polysemic item B		√	√	√	
Polysemic item C			√	√	√

Inferring semantic maps

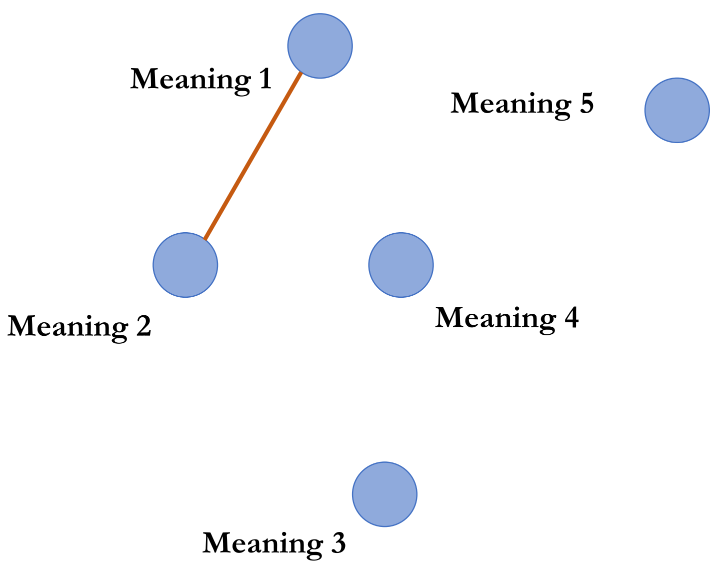
- How does it transfer to semantic maps?
    - Nodes are meanings
    - Constraints are Polysemic items
    - One connects the nodes economically based on these constraints



Meaning	1	2	3	4	5
Polysemic item A	√	√			
Polysemic item B		√	√	√	
Polysemic item C			√	√	√

Inferring semantic maps

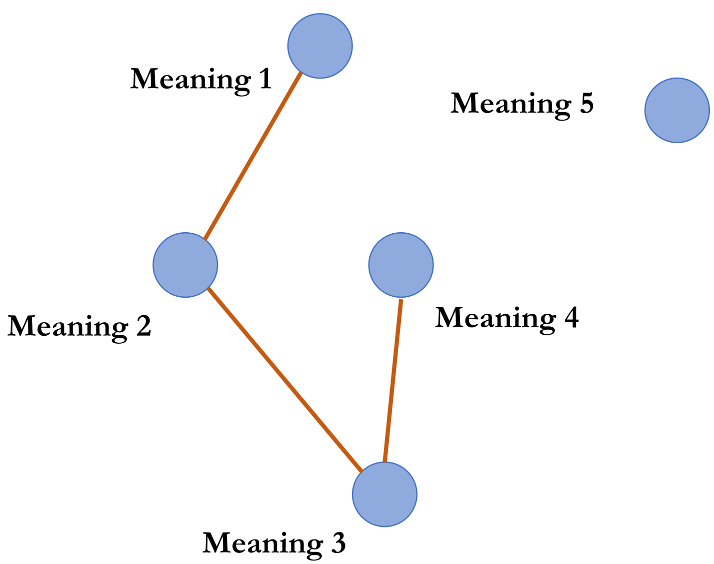
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are Polysemic items
  - One connects the nodes economically based on these constraints



Meaning	1	2	3	4	5
Polysemic item A	√	√			

Inferring semantic maps

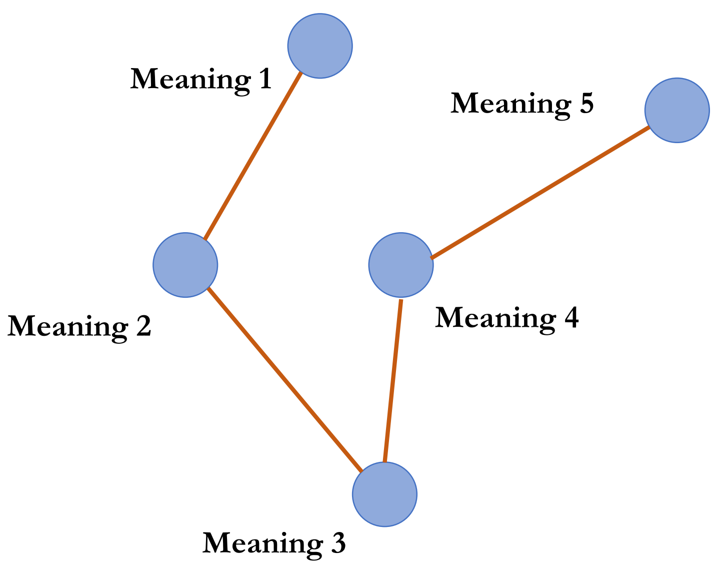
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are Polysemic items
  - One connects the nodes economically based on these constraints



Meaning	1	2	3	4	5
Polysemic item A	√	√			
Polysemic item B		√	√	√	

Inferring semantic maps

- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are Polysemic items
  - One connects the nodes economically based on these constraints

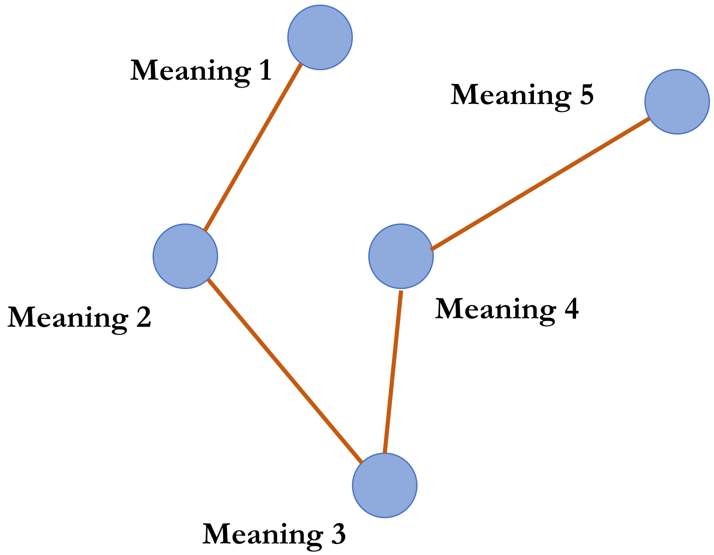


Meaning	1	2	3	4	5
Polysemic item A	√	√			
Polysemic item B		√	√	√	
Polysemic item C			√	√	√

Inferring semantic maps

- How does it transfer to semantic maps?

The result is a map that accounts for all the polysemy patterns, while remaining as economic as possible



Meaning	1	2	3	4	5
Polysemic item A	√	√			
Polysemic item B		√	√	√	
Polysemic item C			√	√	√

Inferring semantic maps

- Regier et al. (2013): the approximations produced by the Angluin et al. algorithm are of high quality
  - Tested on the crosslinguistic data of Haspelmath (1997) and Levinson et al. (2003)

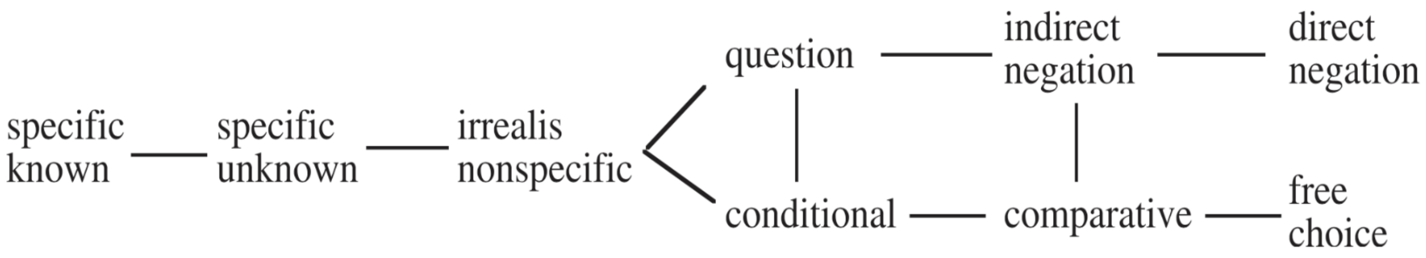


FIGURE. Semantic map of the indefinite pronouns functions (Haspelmath, 1997: 4)

Inferring semantic maps

- **How to prepare the material?**

Source of constraint	Constraint name	Meaning_1	Meaning_2	Meaning_3	Meaning_4	Meaning_5
		"xxx"	"xxx"	"xxx"	"xxx"	"xxx"
Language 1	Form	1	1	0	1	0
Language 2	Form	1	1	1	1	1
Language 3	Form	1	1	1	0	0
Language 4	Form 1	1	1	0	0	0
Language 4	Form 2	1	0	1	1	1

TABLE. A hypothetical lexical matrix for four languages

Inferring semantic maps

INPUT
(lexical matrix)

Language	*Word*	Specific Known **SK**	Specific Unknown **SU**	Irrealis Non-specific **IR**	Question **QN**	Conditional **CD**	Indirect Negation **IN**
German	"etwas"	1	1	1	1	1	1
German	"irgend"	0	1	1	1	1	1
German	"je"	0	0	0	1	1	1
German	"jeder"	0	0	0	0	0	1
German	"n-"	0	0	0	0	0	0
Dutch	"dan ook"	0	0	1	1	1	1
Dutch	"enig"	0	0	0	1	1	1
Dutch	"iets"	1	1	1	1	1	1
Dutch	"niets"	0	0	0	0	0	0
English	"any"	0	0	0	1	1	1
English	"ever"	0	0	0	1	1	1
English	"no"	0	0	0	0	0	0
English	"some"	1	1	1	1	1	0

Inferring semantic maps

INPUT
(lexical matrix)

Language	*Word*	Specific Known **SK**	Specific Unknown **SU**	Irrealis Non-specific **IR**	Question **QN**	Conditional **CD**	Indirect Negation **IN**
German	"etwas"	1	1	1	1	1	1
German	"irgend"	0	1	1	1	1	1
German	"je"	0	0	0	1	1	1
German	"jeder"	0	0	0	0	0	1
German	"n-"	0	0	0	0	0	0
Dutch	"dan ook"	0	0	1	1	1	1
Dutch	"enig"	0	0	0	1	1	1
Dutch	"iets"	1	1	1	1	1	1
Dutch	"niets"	0	0	0	0	0	0
English	"any"	0	0	0	1	1	1
English	"ever"	0	0	0	1	1	1
English	"no"	0	0	0	0	0	0
English	"some"	1	1	1	1	1	0

Inferring semantic maps

INPUT
(lexical matrix)

Language	Word	Specific Known SK	Specific Unknown SU	Irrealis Non-specific IR	Question QN	Conditional CD	Indirect Negation IN
German	"etwas"	1	1	1	1	1	1
German	"irgend"	0	1	1	1	1	1
German	"je"	0	0	0	1	1	1
German	"jeder"	0	0	0	0	0	1
German	"n-"	0	0	0	0	0	0
Dutch	"dan ook"	0	0	1	1	1	1
Dutch	"enig"	0	0	0	1	1	1
Dutch	"iets"	1	1	1	1	1	1
Dutch	"niets"	0	0	0	0	0	0
English	"any"	0	0	0	1	1	1
English	"ever"	0	0	0	1	1	1
English	"no"	0	0	0	0	0	0
English	"some"	1	1	1	1	1	0

Inferring semantic maps

INPUT
(lexical matrix)

Language	*Word*	Specific Known SK	Specific Unknown SU	Irrealis Non-specific IR	Question QN	Conditional CD	Indirect Negation IN
German	"etwas"	1	1	1	1	1	1
German	"irgend"	0	1	1	1	1	1
German	"je"	0	0	0	1	1	1
German	"jeder"	0	0	0	0	0	1
German	"n-"	0	0	0	0	0	0
Dutch	"dan ook"	0	0	1	1	1	1
Dutch	"enig"	0	0	0	1	1	1
Dutch	"iets"	1	1	1	1	1	1
Dutch	"niets"	0	0	0	0	0	0
English	"any"	0	0	0	1	1	1
English	"ever"	0	0	0	1	1	1
English	"no"	0	0	0	0	0	0
English	"some"	1	1	1	1	1	0

FORM-MEANING CORRESPONDENCES

1 => form has meaning

0 => form does not have meaning

Inferring semantic maps

INPUT
(lexical matrix)

Language	*Word*	Specific Known **SK**	Specific Unknown **SU**	Irrealis Non-specific **IR**	Question **QN**	Conditional **CD**	Indirect Negation **IN**
German	"etwas"	1	1	1	1	1	1
German	"irgend"	0	1	1	1	1	1
German	"je"	0	0	0	1	1	1
German	"jeder"	0	0	0	0	0	1
German	"n-"	0	0	0	0	0	0
Dutch	"dan ook"	0	0	1	1	1	1
Dutch	"enig"	0	0	0	1	1	1
Dutch	"iets"	1	1	1	1	1	1
Dutch	"niets"	0	0	0	0	0	0
English	"any"	0	0	0	1	1	1
English	"ever"	0	0	0	1	1	1
English	"no"	0	0	0	0	0	0
English	"some"	1	1	1	1	1	0

ALGORITHM
(python script)

```
# MAIN LOOP
objfn = C(G,T)
while (objfn < 0):
    print ("objective fn is currently", objfn,)
    max_score = 0
    # choose next edge greedily: the one that increases objfn the most
    for e in PossE:
        # temporarily add e to graph G
        G.add_edge(*e)
        score = C(G,T) - objfn
        G.remove_edge(*e)
        if (score > max_score):
            max_score = score
            max_edge = e
```

Inferring semantic maps

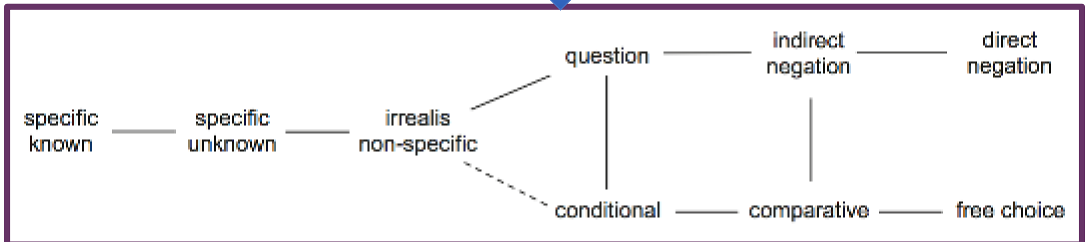
INPUT
(lexical matrix)

Language	*Word*	Specific Known SK	Specific Unknown SU	Irrealis Non-specific IR	Question QN	Conditional CD	Indirect Negation IN
German	"etwas"	1	1	1	1	1	1
German	"irgend"	0	1	1	1	1	1
German	"je"	0	0	0	1	1	1
German	"jeder"	0	0	0	0	0	1
German	"n-"	0	0	0	0	0	0
Dutch	"dan ook"	0	0	1	1	1	1
Dutch	"enig"	0	0	0	1	1	1
Dutch	"iets"	1	1	1	1	1	1
Dutch	"niets"	0	0	0	0	0	0
English	"any"	0	0	0	1	1	1
English	"ever"	0	0	0	1	1	1
English	"no"	0	0	0	0	0	0
English	"some"	1	1	1	1	1	0

ALGORITHM
(python script)

```
# MAIN LOOP
objfn = C(G,T)
while (objfn < 0):
    print ("objective fn is currently", objfn,)
    max_score = 0
    # choose next edge greedily: the one that increases objfn the most
    for e in PossE:
        # temporarily add e to graph G
        G.add_edge(*e)
        score = C(G,T) - objfn
        G.remove_edge(*e)
        if (score > max_score):
            max_score = score
            max_edge = e
```

RESULT
(semantic map)



Inferring semantic maps

FIGURE a. Semantic map of the indefinite pronouns functions (Haspelmath, 1997, p. 4)

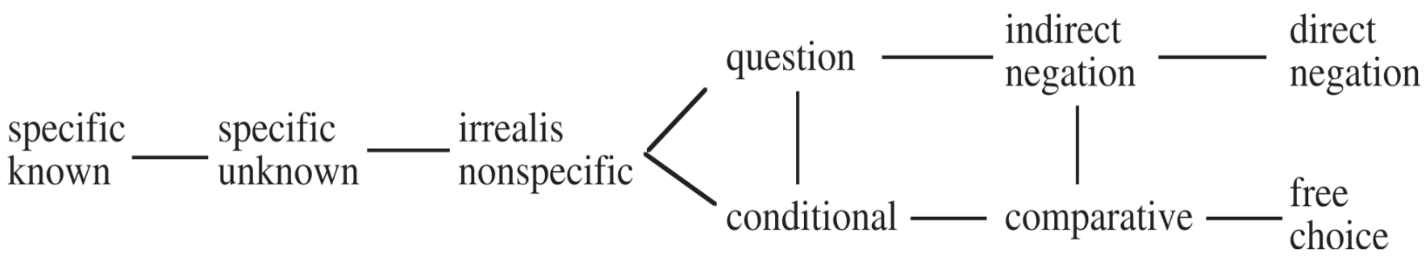
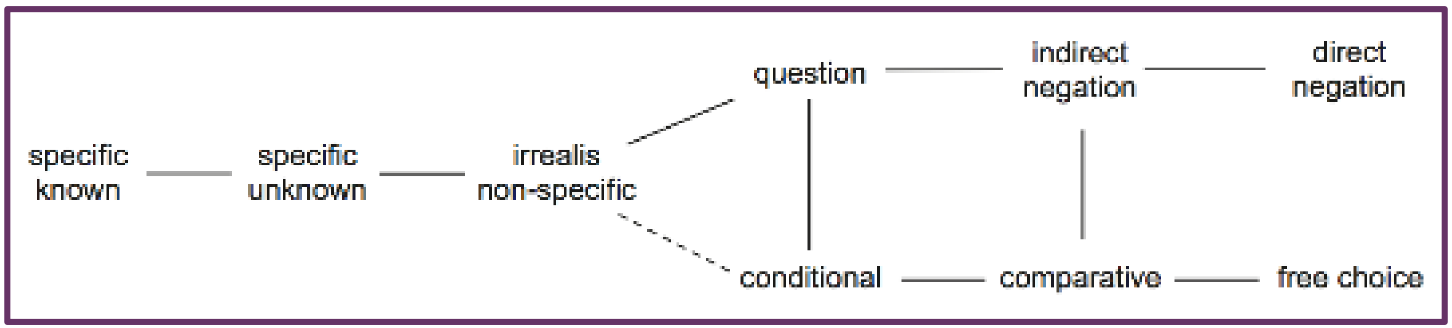


FIGURE b. Semantic map of the indefinite pronouns functions in Regier et al. (2013) using the same dataset with Haspelmath (1997)



Workshop activities

Workshop activity 1

TABLE. Polysemy data for lexemes expressing the meanings "thread, string" (based on the resources provided by William S. Annis; https://lingweenie.org/conlang/maps/thread.py)

Languages	Forms	Meanings
Malayalam	*tantu*	filament, tendril, fiber, thread, sinew
Thai	*yaawngM*	fiber, filament, string, thread, cord, vein
Turkish: iplik	*iplik*	thread, yarn, fiber, filament
Maori	*tui*	thread, lashing
Russian	*верёвка*	cord, rope, string
Zulu	*intambo*	cord, rope, string, thread, wire
Buang	*aggis*	rope, vine, string
Kagwahiva	*-enimboa*	thread, cord
Nootkah	*diqi:ba*	net, twine, thread
Rotokas	*viripato*	thread, string, fishing line, cord
Tlingit	*tás*	thread, sinew
Inupiaq	*ivalu*	thread, sinew

Workshop activity 1

Task

1. Convert the polysemy data provided into a **lexical matrix** (with language-form as rows and meanings as columns)

EXAMPLE

	A	B	C	D	E	F	G	H	I	J	K
1	***Language***	***Word***	Sense_1	Sense_2	Sense_3	Sense_4	Sense_5	Sense_6	Sense_7	Sense_8	Sense_9
2			**Specific Kr**	**Specific Ur**	**Irrealis Noı**	**Question**	**Conditiona**	**Indirect Ne**	**Comparatiı**	**Direct Neg:**	**Free-choice**
3	Portugues	qualquer	0	1	1	1	1	1	1	1	1
4	Hindi/Urd	bhii	0	0	1	1	1	1	1	1	1
5	Turkish	herhangi	0	0	1	1	1	1	1	1	1
6	Ancash Qu	-pis	0	0	1	1	1	1	1	1	1
7	English	any	0	0	0	1	1	1	1	1	1
8	French	que ce soi	0	0	0	1	1	1	1	1	1
9	Irish	ar bith /ac	0	0	0	1	1	1	1	1	1
10	Swahili	(CL-o Cl-o	0	0	0	1	1	1	1	1	1
11	Yakut	-da	0	0	0	0	0	1	1	1	1

TABLE. A lexical matrix based on the dataset for the implicational map of indefinite pronouns in Haspelmath 1997

Workshop activity 1

Task

1. **Convert** the polysemy data provided into a **lexical matrix** (with language-form as rows and meanings as columns)

2. **Build** manually a **classical semantic map**—based on this lexical matrix—that respects both the connectivity hypothesis and the economy principle

Connectivity hypothesis:
"Any relevant language-specific and construction-specific category should map onto a CONNECTED REGION in conceptual space"

- If two given meanings are expressed by one form in at least one language, the corresponding meanings should be connected

Economy principle:
Given three meanings (**Meaning1**, **Meaning2**, **Meaning3**), if the linguistic items expressing **Meaning1** and **Meaning3** always express **Meaning2**, there is no need to draw an edge between **Meaning1** and **Meaning3**

Workshop activity 1

Task

1. **Convert** the polysemy data provided into a **lexical matrix** (with language-form as rows and meanings as columns)

2. **Build** manually a **classical semantic map**—based on this lexical matrix—that respects both the connectivity hypothesis and the economy principle

3. **Use** the **algorithm** provided by Regier et al. (2013) to automatically infer a graph and then comparing the manually and automatically plotted maps.*

http://lclab.berkeley.edu/regier/semantic-maps/SemanticMap.py

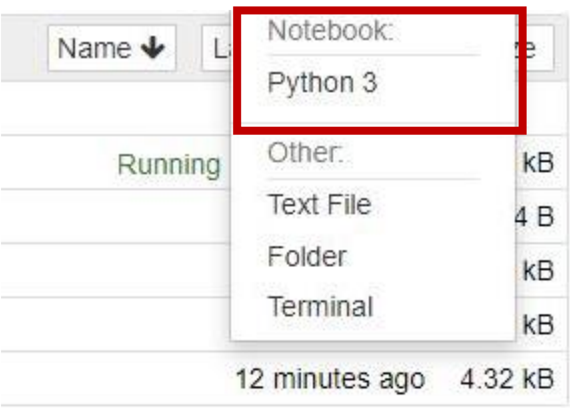
*** The algorithm needed to be amended; check the seminar's Dropbox folder for the amended script**

Workshop activity 1

1. Everything needs to be in the same folder:
    - **xls. file, python script**

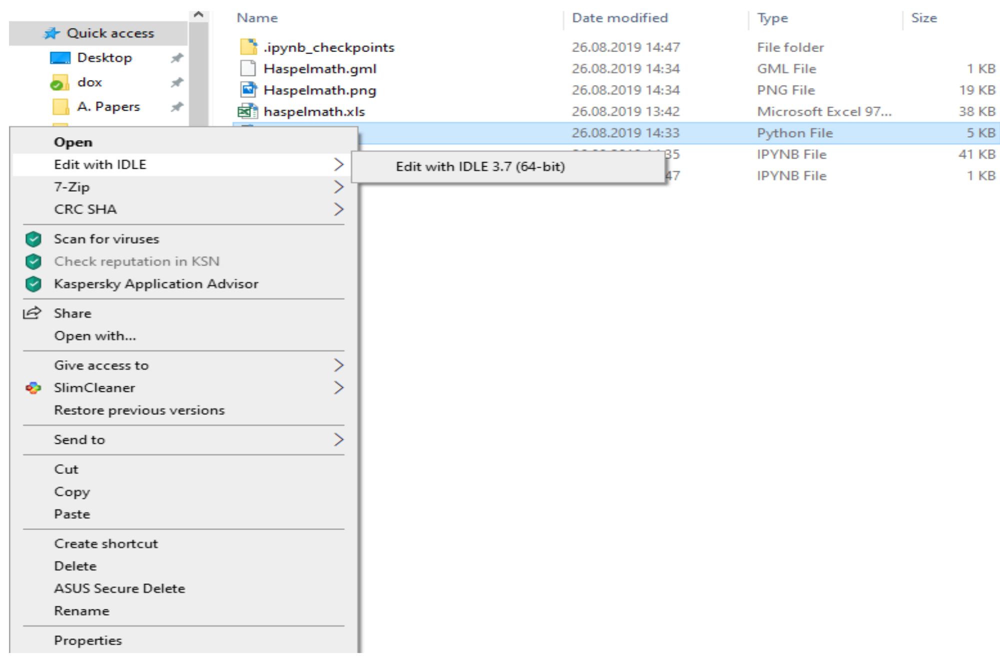
2. **Run Anaconda**
    - Run Jupyter Notebook

3. **Locate** the folder where the files are



Workshop activity 1

1. Go to the folder where you saved your data
    - Right click on the **python file**
    - Edit with **Idle**

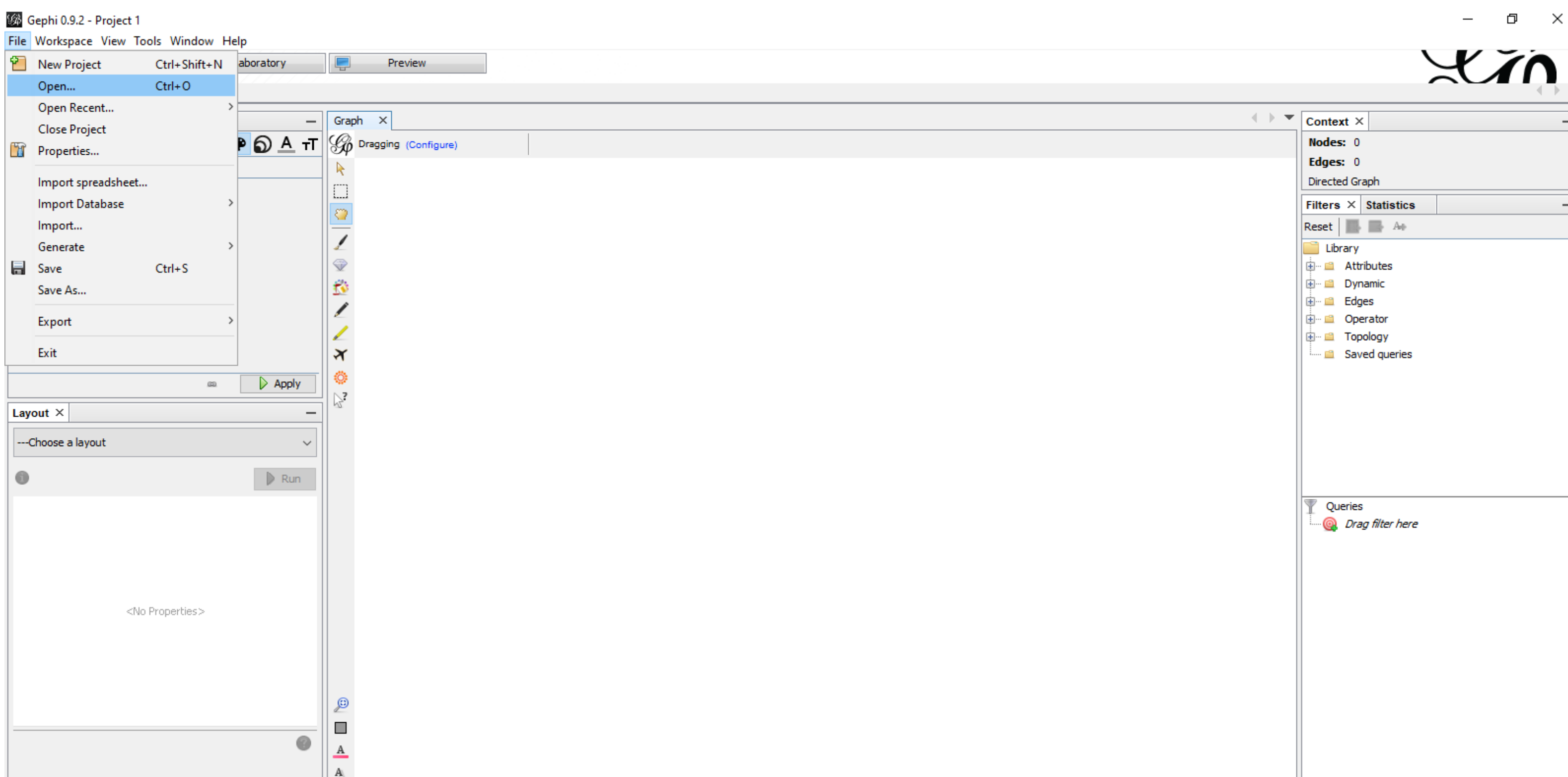


2. Copy the code and paste it to the Anaconda environment
    - Click **Run**

Workshop activity 1

1. Run Gephi

2. File → Open

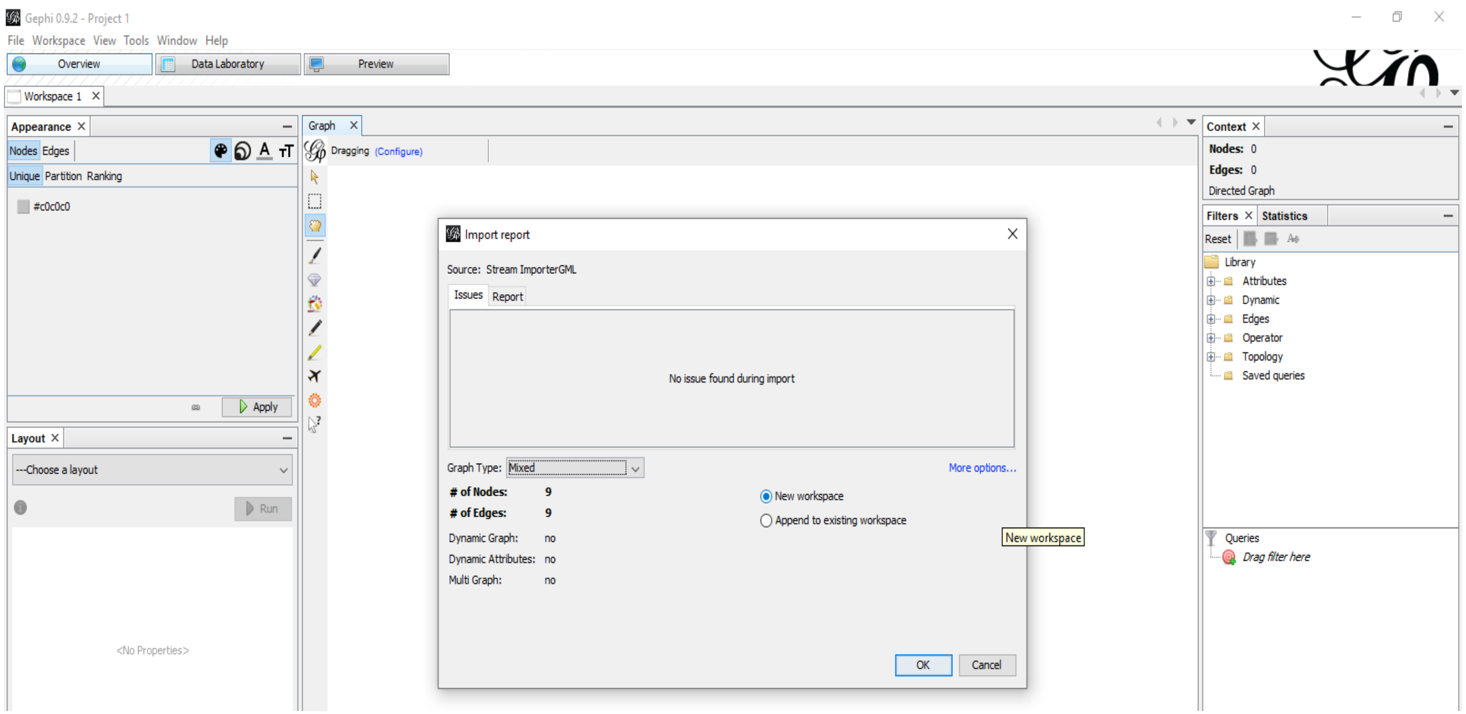


Workshop activity 1

1. Locate the folder where you saved your data

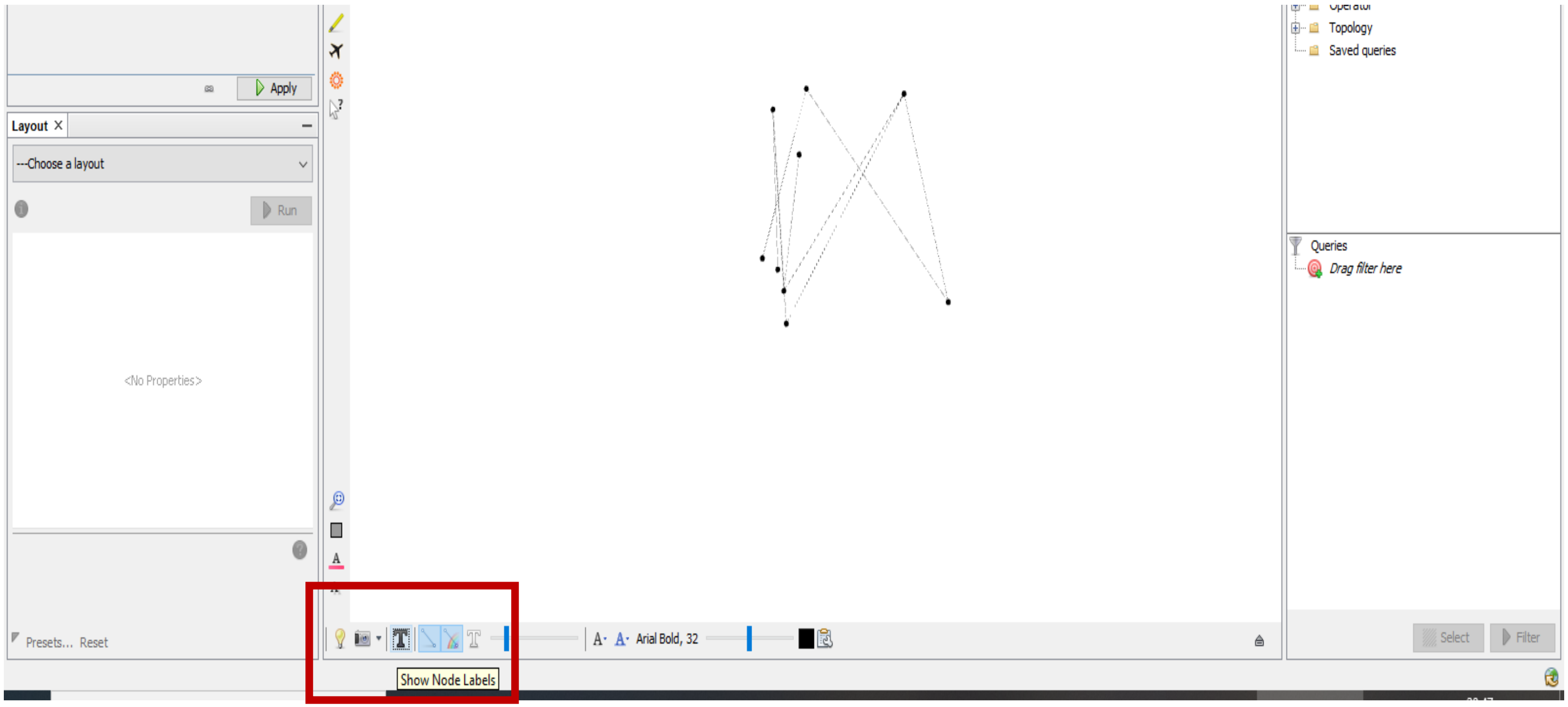
2. Open the **.gml file**

2. Dropdown menu: Graph: choose **undirected**



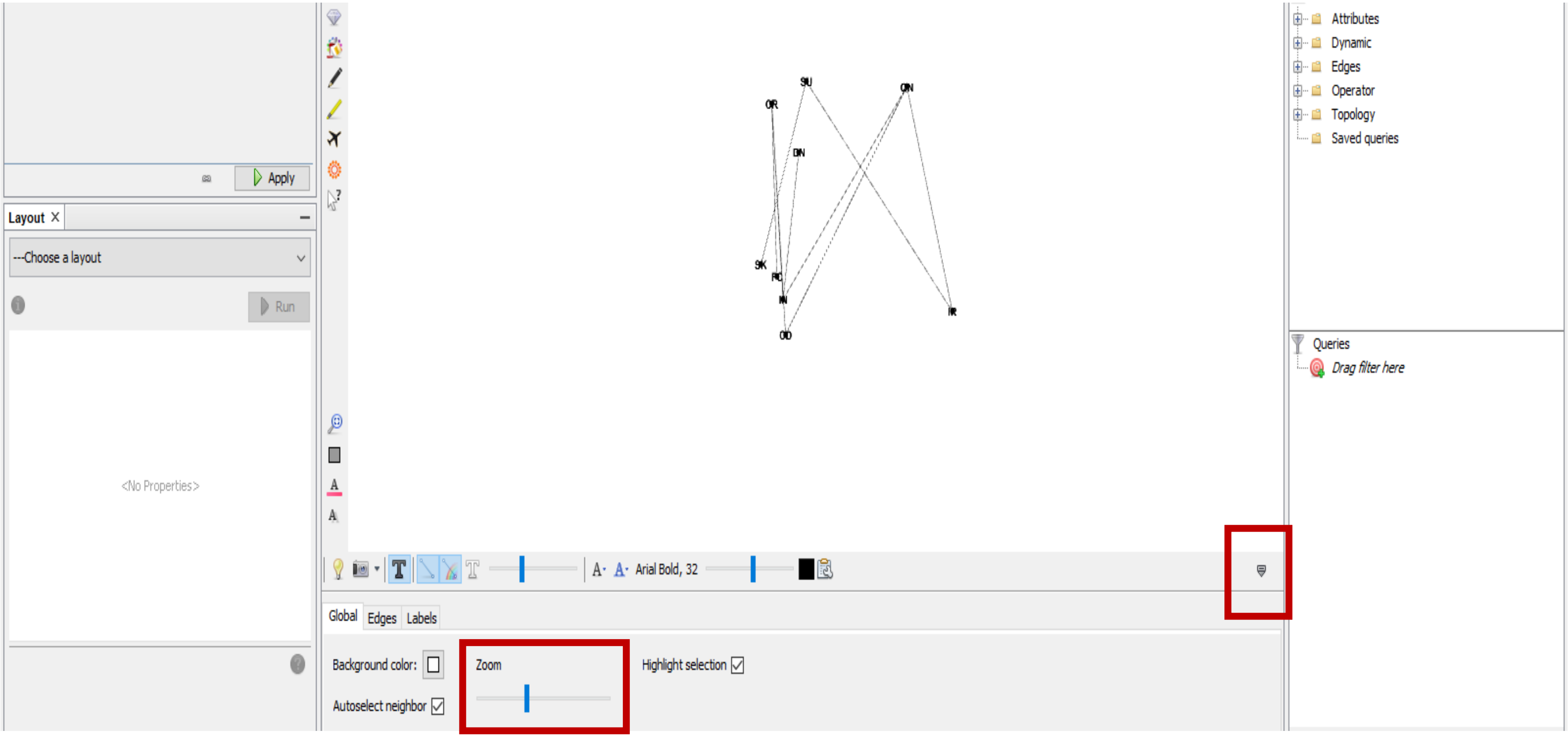
Workshop activity 1

1. Click on the labels icon to show node (i.e. meaning) labels



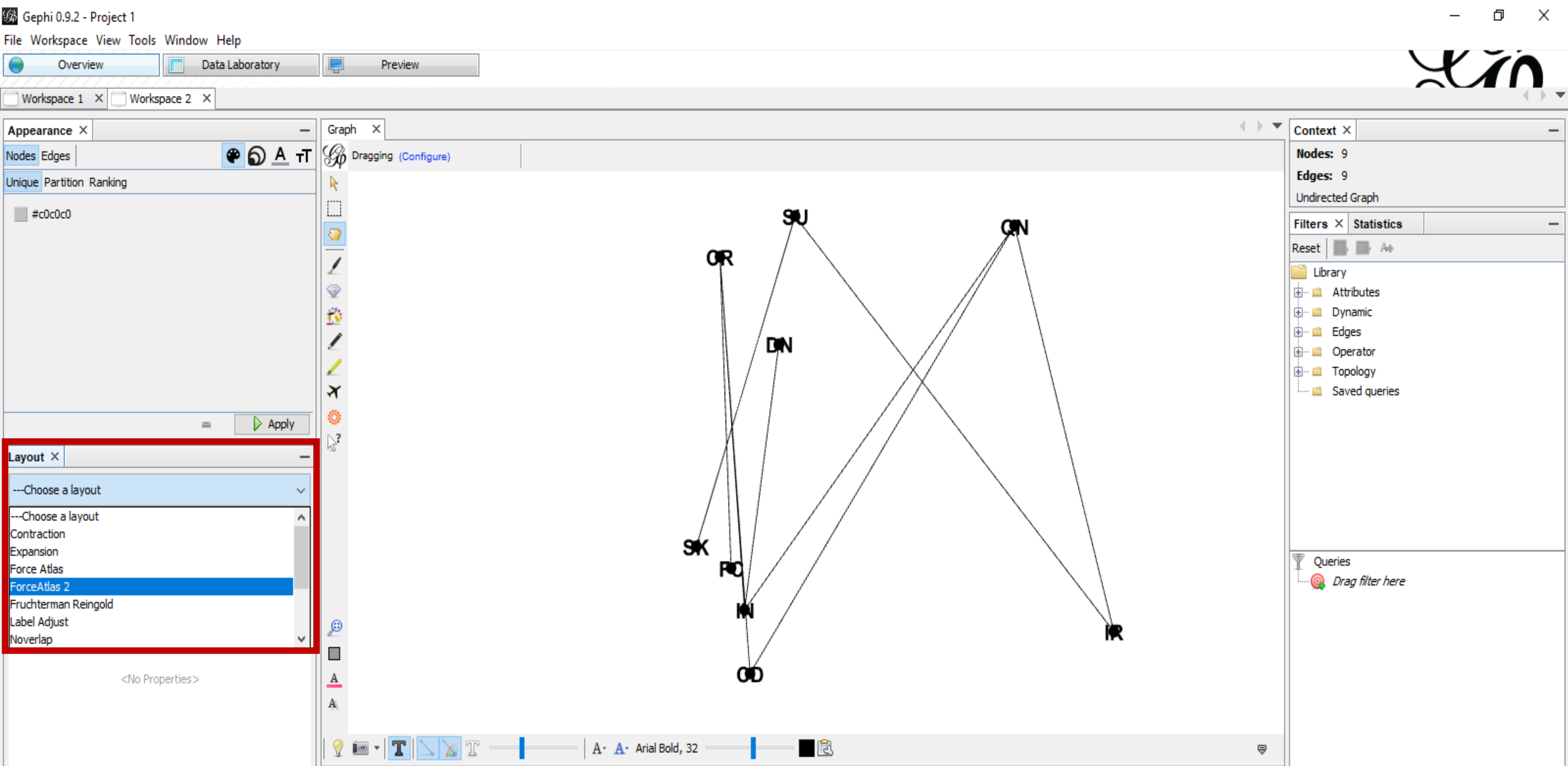
Workshop activity 1

1. Zoom in or out



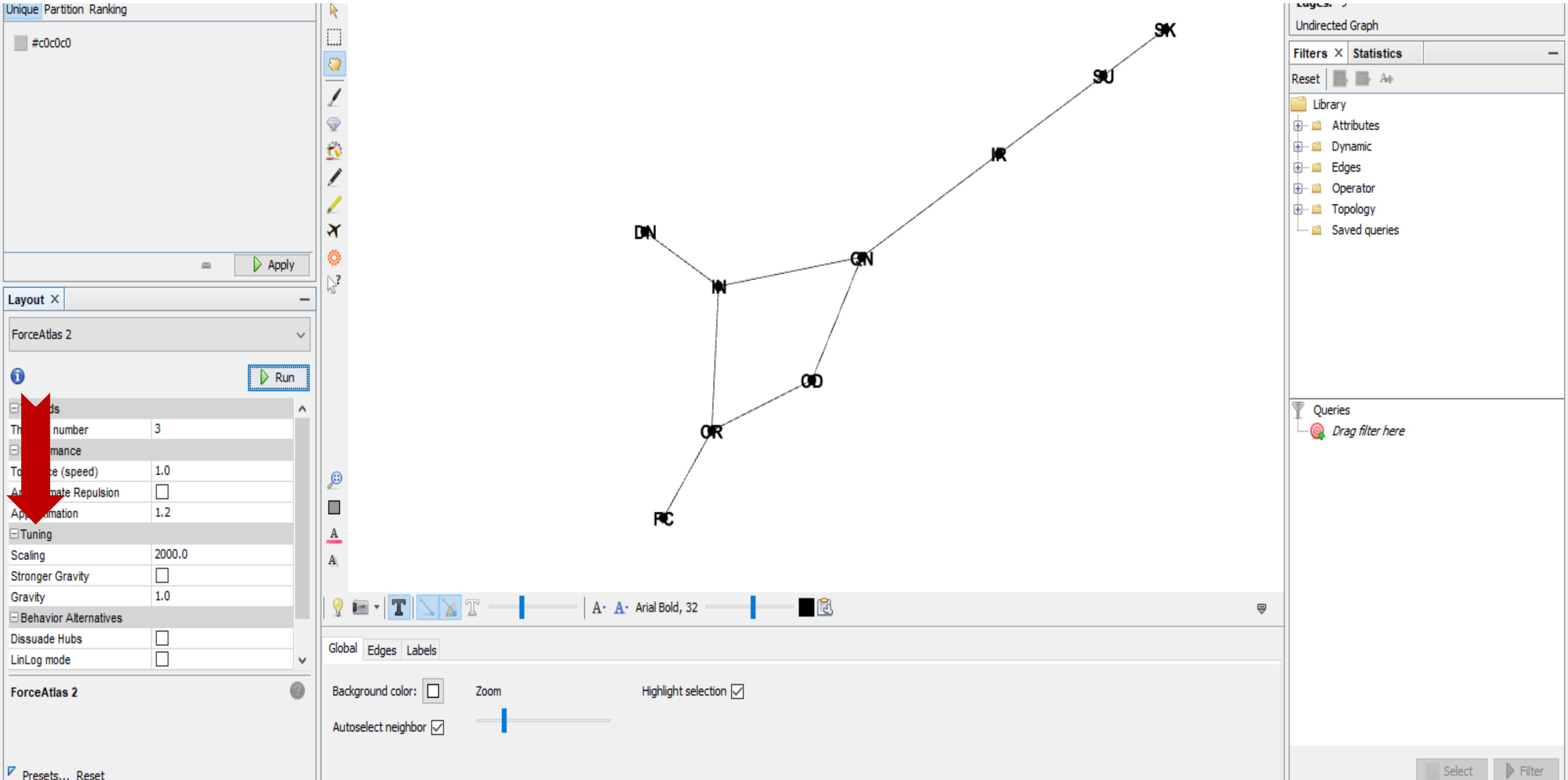
Workshop activity 1

1. Choose **a layout**
2. Adjust **the scaling** (e.g. 1000, or 2000, etc.)
3. **Run**



Workshop activity 1

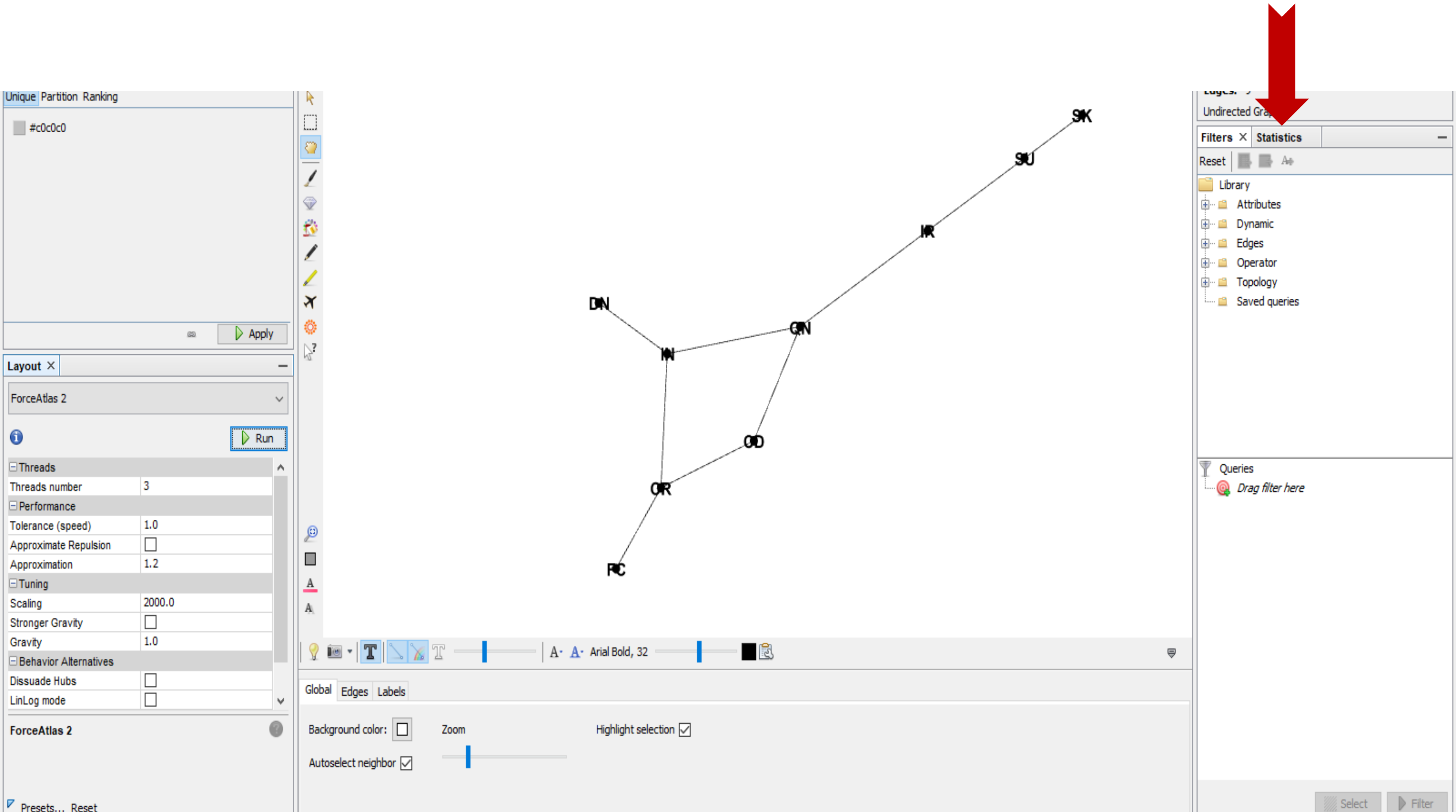
1. Choose **a layout**
2. Adjust **the scaling**
3. **Run**



Workshop activity 1

To identify communities in the data

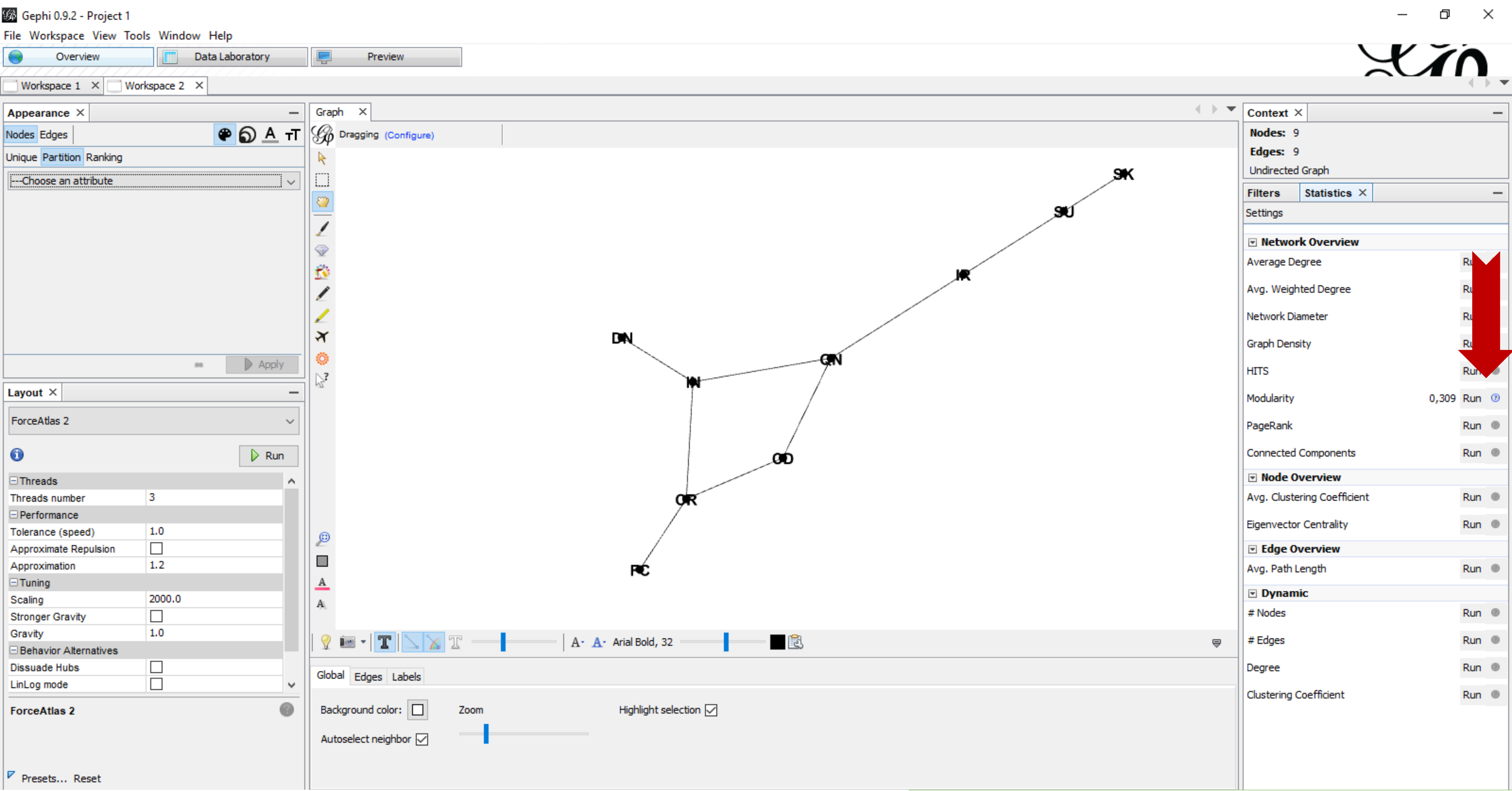
1. Click on **Statistics**



Workshop activity 1

To identify communities in the data

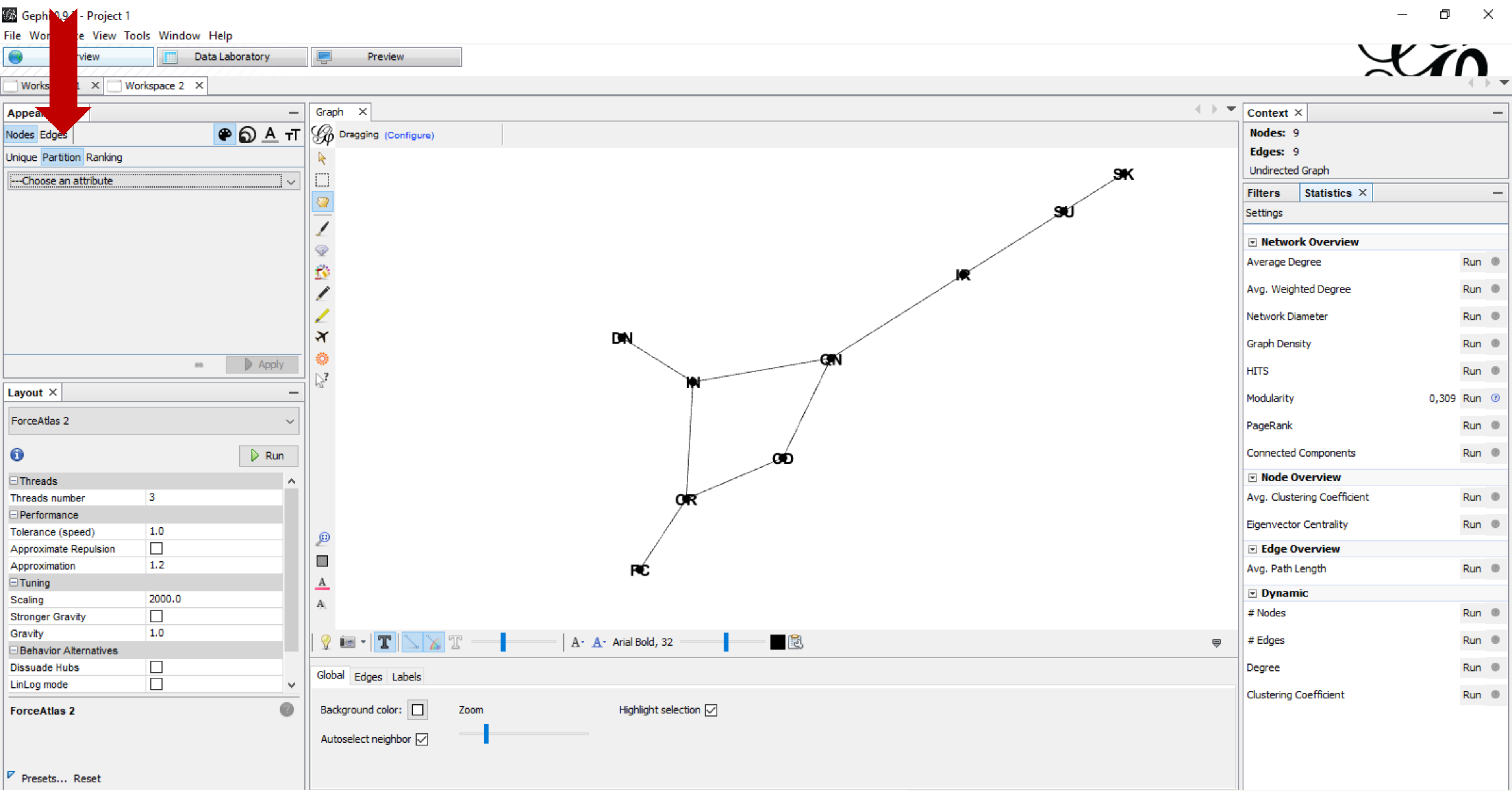
2. Run **Modularity**



Workshop activity 1

To identify communities in the data

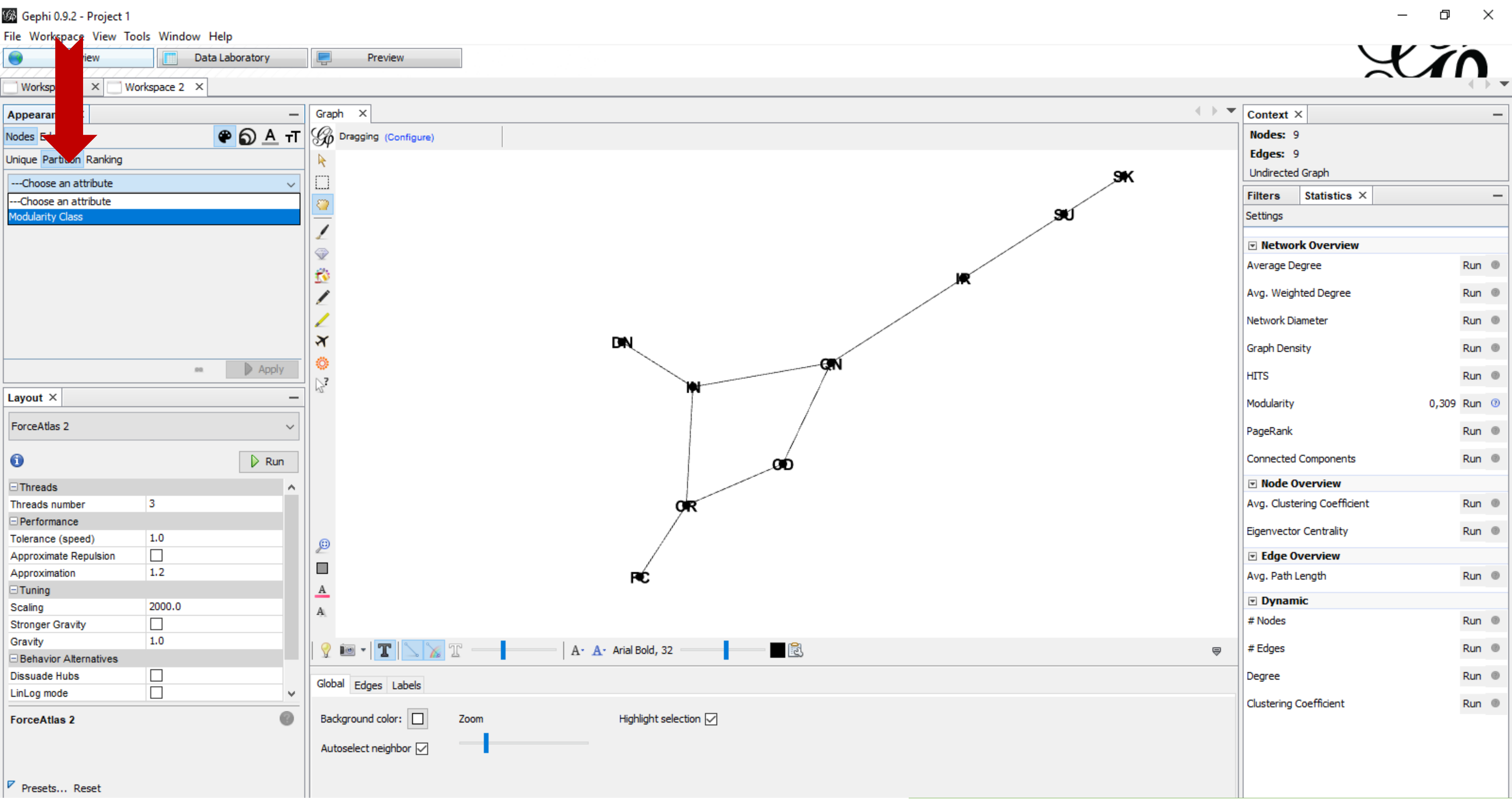
3. Click on **Partition**



Workshop activity 1

To **identify** communities in the data

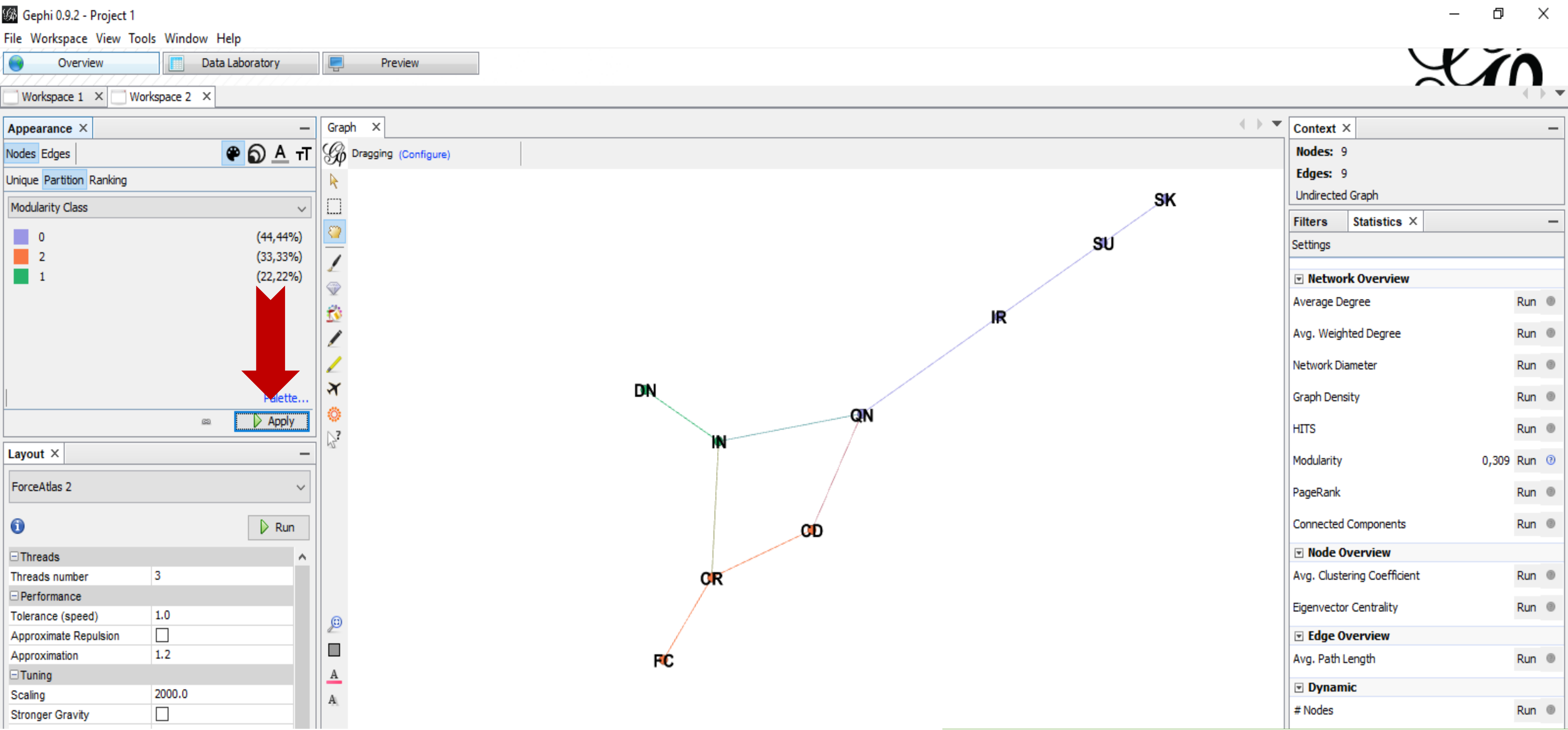
4. Choose an **attribute: Modularity class**



Workshop activity 1

To identify communities in the data

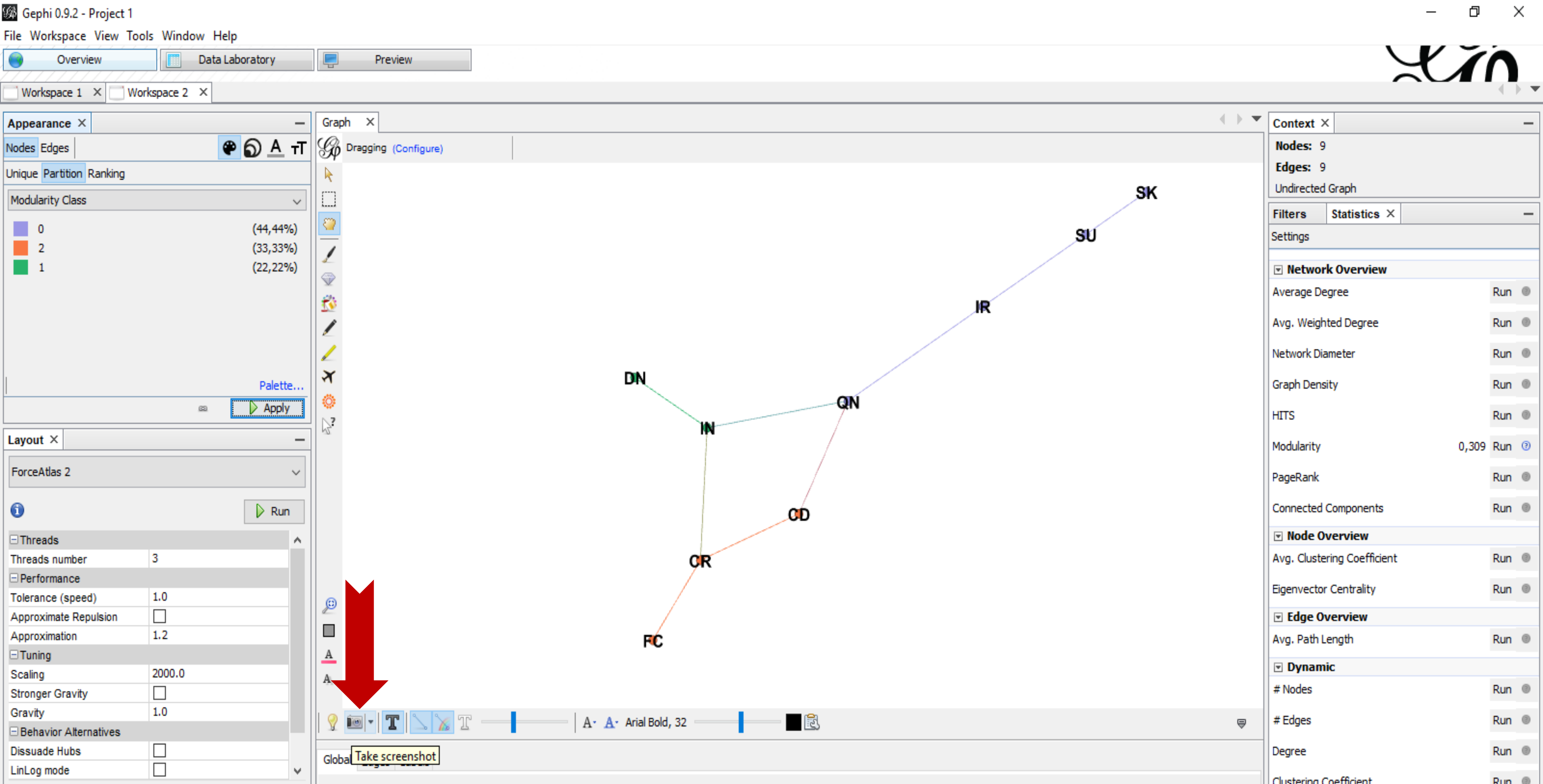
5. Click on **Apply**



Workshop activity 1

To identify communities in the data

6. Take **screenshot** and **save** the graph



Workshop activity 2

Task

1. **Download** the dataset for the implicational map of indefinite pronouns in Haspelmath 1997: https://zenodo.org/record/1157061#.XWOzvOgzaM-

2. **Use** the **algorithm** provided by Regier et al. (2013) to automatically infer a graph*

http://lclab.berkeley.edu/regier/semantic-maps/SemanticMap.py

*** The algorithm needed to be amended; check the seminar's Dropbox folder for the amended script**

Automatic plotting: Two steps forward

Weighted and diachronic semantic maps

Automatic plotting: Two steps forward

- Weighted semantic maps are much more informative than regular semantic maps, because they visually provide information about the frequency of polysemy patterns

    - Georgakopoulos, Eitan Grossman, Dmitry Nikolaev & Stéphane Polis. *Under revision*. Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain. *Linguistic Typology*.

- Diachronic semantic maps are much more informative than regular semantic maps, because they visually provide information about possible pathways of change

    - Georgakopoulos Thanasis & Stéphane Polis. *Under revision*. Dynamic semantic maps of content words. The diachrony of time-related lexemes. *Journal of Historical Linguistics*.

Automatic plotting: Two steps forward

Weighted semantic maps

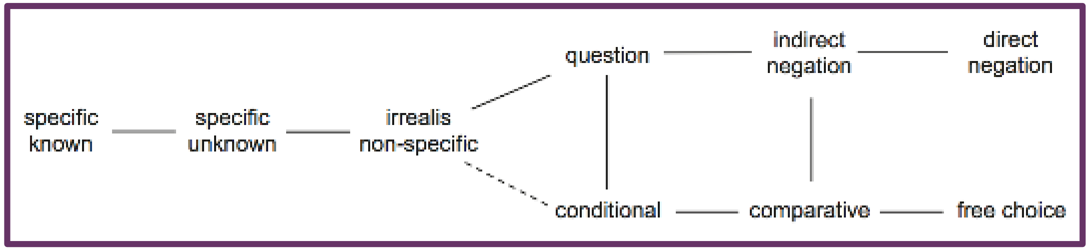
- Generate the map with a modified version of the algorithm of Regier et al. (2013)
  - PRINCIPLE: for each edge that is being added between two meanings of the map by the algorithm, check in the lexical matrix how many times this specific polysemy pattern is attested, and increase the weight of the edge accordingly

```
edgeWeight = 0
for sns in sensesTupleList:
        if (max_edge[0] in sns) and (max_edge[1] in sns):
                edgeWeight += 1
G.add_edge(*max_edge,weight=edgeWeight)
```

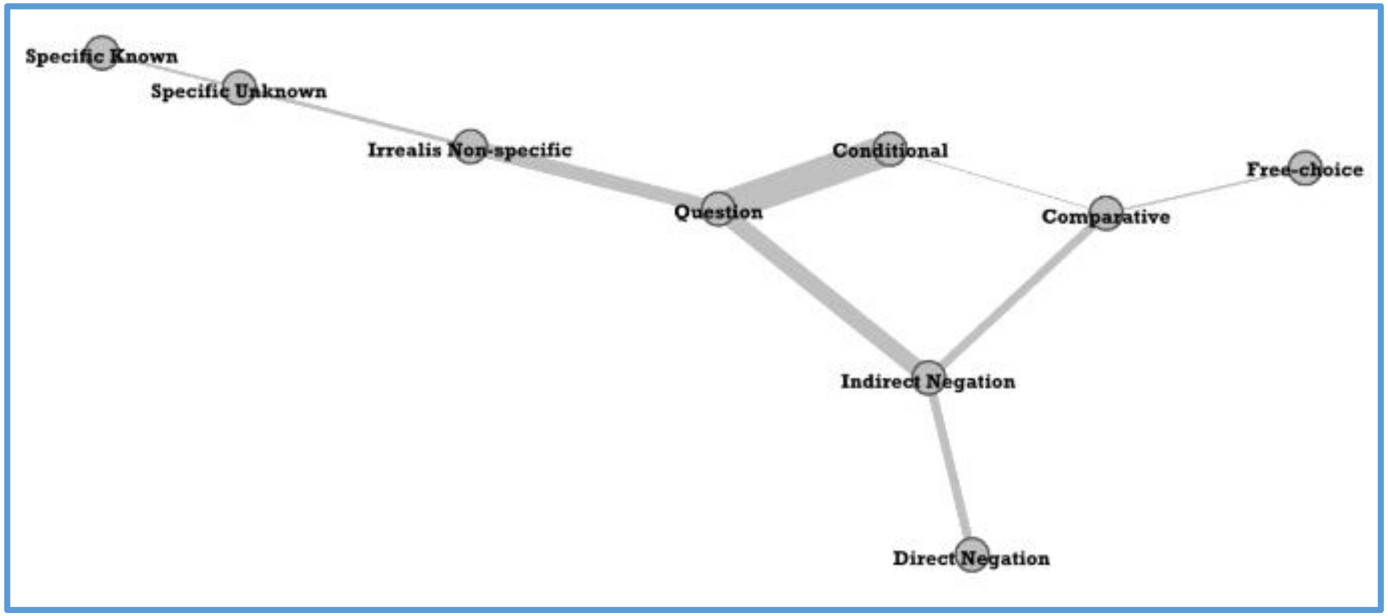
  - Based on the data of Haspelmath (1997), kindly provided by the author, the result between a non-weighted and a weighted semantic map are markedly different

Automatic plotting: Two steps forward

Weighted semantic maps



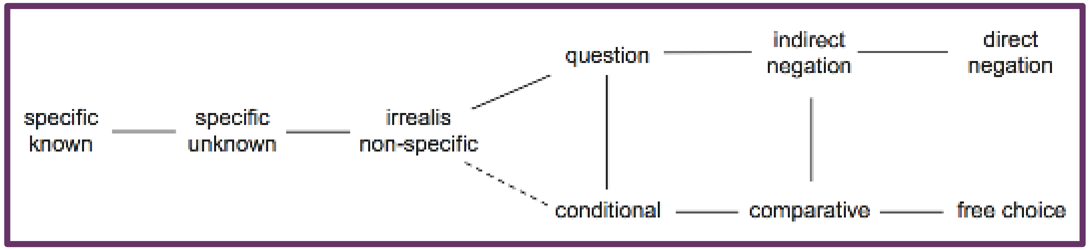
Automatically plotted semantic maps: non-weighted vs. weighted
(data from Haspelmath 1997)



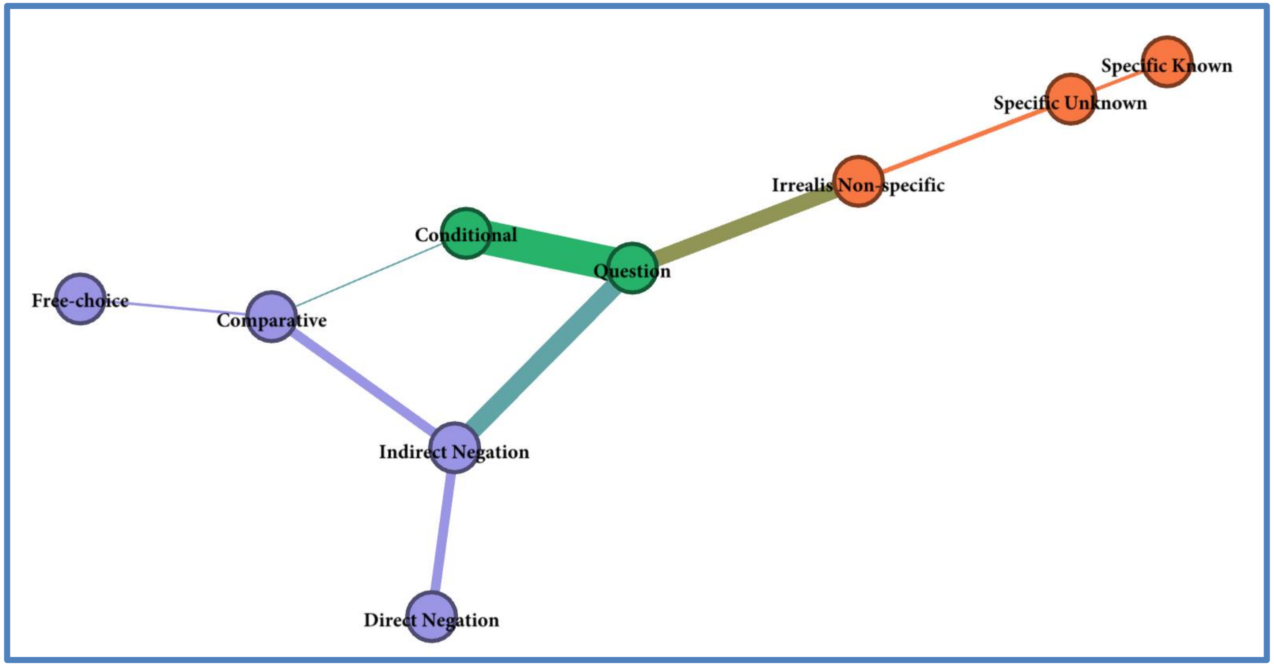
The graph is visualized in Gephi® with the *Force Atlas* algorithm

Automatic plotting: Two steps forward

Weighted semantic maps



Automatically plotted semantic maps: non-weighted vs. weighted
(data from Haspelmath 1997)



The graph is visualized in Gephi® with the *Force Atlas* algorithm and modularity analysis (Lambiotte et al. 2009)

Automatic plotting: Two steps forward

Diachronic semantic maps

- Expand the lexical matrix so as to include information about diachrony

Source of constraint	*Constraint name*	*Constraint Time*	Sense_1 Tree	Sense_2 Wood	Sense_3 Forest
Language_1	Word_1	0	1	0	0
Language_1	Word_1	1	1	1	0
Language_2	Word_1	0	1	0	0
Language_2	Word_2	0	0	1	0
Language_2	Word_2	1	0	1	1
Language_3	Word_1	0	1	1	0
Language_3	Word_2	0	0	0	1

The diachronic stages are
indexed by numbers:
0, 1, 2, etc.

Automatic plotting: Two steps forward

Diachronic semantic maps

- Expand the lexical matrix so as to include information about diachrony

Source of constraint	*Constraint name*	*Constraint Time*	Sense_1 Tree	Sense_2 Wood	Sense_3 Forest
Language_1	Word_1	0	1	0	0
Language_1	Word_1	1	1	1	0
Language_2	Word_1	0	1	0	0
Language_2	Word_2	0	0	1	0
Language_2	Word_2	1	0	1	1
Language_3	Word_1	0	1	1	0
Language_3	Word_2	0	0	0	1

The meaning of a word can change from one stage to another (e.g., Word_2 of Language_2 expresses the meaning Wood during stage 0 and Wood & Forest during stage 1)

Automatic plotting: Two steps forward

Diachronic semantic maps

- Generate the graph with the algorithm of Regier et al. (2013)
- Enrich the graph with oriented edges (where relevant)
    - PRINCIPLE: (1) we convert the undirected graph into a *directed graph*
      (2) for each edge in the graph, if the meaning of node A is attested for one diachronic stage, while the meaning of node B is not, check in the lexical matrix if there is a later diachronic stage of the same language for which this specific word has both meaning A and B (or just meaning B). If this is the case, we can infer a meaning extension from A to B.

```
H = G.to_directed()     # convert the graph 'G' into a directed Graph 'H' in order to explo
                        # all the possibilities as regards the relationship between the noc
                        # (i.e., both A -> B and B -> A for all the connected nodes, crucic
                        # not only A -> B)
nx.set_edge_attributes(H, 'type', 'undirected') # set the default value to "undirected" for

for u,v,e in H.edges(data=True):        # loop over all the edges in the DiGraph 'H'
        for t in T_Full:                # look at the metadata and senses for one line in T
                if t.count(u) == 1 and t.count(v) == 0: # if the meaning of node 'u' in the
                                                        # while the meaning of node 'v' is
```

Automatic plotting: Two steps forward

Diachronic semantic maps

INPUT
(diachronic
lexical matrix)

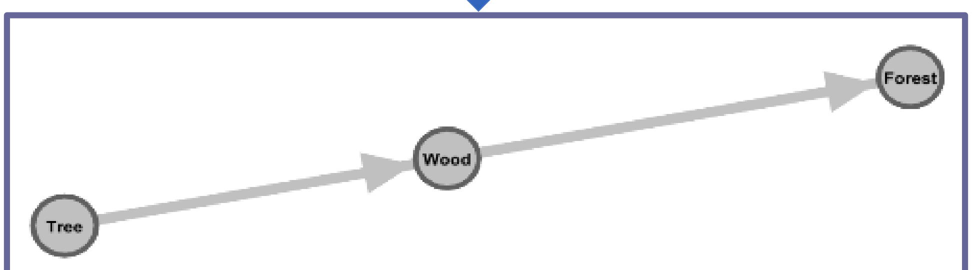
Source of constraint	*Constraint name*	*Constraint Time*	Sense_1 Tree	Sense_2 Wood	Sense_3 Forest
Language_1	Word_1	0	1	0	0
Language_1	Word_1	1	1	1	0
Language_2	Word_1	0	1	0	0
Language_2	Word_2	0	0	1	0
Language_2	Word_2	1	0	1	1
Language_3	Word_1	0	1	1	0
Language_3	Word_2	0	0	0	1

ALGORITHM
(python script
for inferring
oriented edges)

```
H = G.to_directed()     # convert the graph 'G' into a directed Graph 'H' in order to explo
                        # all the possibilities as regards the relationship between the nod
                        # (i.e., both A -> B and B -> A for all the connected nodes, crucia
                        # not only A -> B)
nx.set_edge_attributes(H, 'type', 'undirected') # set the default value to "undirected" for

for u,v,e in H.edges(data=True):        # loop over all the edges in the DiGraph 'H'
        for t in T_Full:                # look at the metadata and senses for one line in T
                if t.count(u) == 1 and t.count(v) == 0: # if the meaning of node 'u' in the
                                                         # while the meaning of node 'v' is
```

RESULT
(dynamic
semantic map)



References (selection)

Angluin, D., Aspnes, J., & Reyzin, L. (2010). Inferring social networks from outbreaks. In M. Hutter, F. Stephan, V. Vovk, & T. Zeugmann (Eds.), *Algorithmic learning theory 2010* (pp. 104–118). Berlin: Springer.

Atkins, B. T. S. (1987). Semantic ID tags: corpus evidence for dictionary senses. *Proceedings of the Third Annual Conference of the UW Centre for the New Oxford English Dictionary* (pp. 17–36).

Berez, A. L., & Gries, S. T. (2009). In defense of corpus-based methods: A behavioral profile analysis of polysemous get in English. In S. Moran, D. S. Tanner, & M. Scanlon (Eds.), *The 24th Northwest Linguistics Conference* (University of Washington Working Papers in Linguistics), vol. 27 (pp. 157-166). Seattle, WA: Department of Linguistics.

Breal, M. (1897). *Essai de Sémantique: Science des Significations*. Paris: Hachette.

Croft, W. (2001). *Radical construction grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, W., & Poole, K. T. (2008). Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. *Theoretical Linguistics, 34*(1), 1–37. https://doi.org/10.1515/thli.2008.001

Cysouw, M. (2001). Martin Haspelmath, indefinite pronouns (Oxford studies in typology and linguistic theory). Oxford: Clarendon press, 1997. Pp. xvi 364. Journal of Linguistics, 37(3), 607–612. Cysouw, M. (2007). Building semantic maps: The case of person marking. In B. Wälchli, & M. Miestamo (Eds.), *New challenges in typology* (pp. 225–248). Berlin/New York: Mouton de Gruyter.

Cysouw, M., Haspelmath, M., & Malchukov, A. L. (2010). Introduction to the special issue "semantic maps: Methods and applications". *Linguistic Discovery, 8*(1), 1–3. https://doi.org/10.1349/ps1.1537-0852.a.358

Dirven, R., & Verspoor, M. (2004). *Cognitive explorations of language and linguistics*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.

References (selection)

Divjak, D., & Gries, S. T. (2006). Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory, 2*(1), 23–60. DOI 10.1515/CLLT.2006.002

Divjak, D. (2010). *Structuring the Lexicon: A Clustered Model for Near-Synonymy*. Berlin: de Gruyter.

Fillmore, C. J., & Atkins, B. T. (1992). Towards a frame-based lexicon: The semantics of RISK and its neighbors. In A. Lehrer & E. F. Kittay (Eds.), *Frames, fields and contrasts: New essays in semantics and lexical organization* (pp. 75-102). Hillsdale, NJ: Lawrence Erlbaum Associates.

François, A. (2008). Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In M. Vanhove (Ed.), *From polysemy to semantic change. Towards a typology of lexical semantic associations* (pp. 163–215). John Benjamins: Amsterdam/Philadelphia.

Geeraerts, D. (1997). *Diachronic Prototype Semantics*. Oxford: Clarendon Press.

Georgakopoulos, T., & Polis, S. (2018). The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass, 12*(2), 1–33.

Georgakopoulos, T. (2019). Semantic Maps. In Mark Aronoff (Ed.), *Oxford Bibliographies in Linguistics*. New York: Oxford University Press. DOI 10.1093/OBO/9780199772810-0229

Georgakopoulos, T. *Under review*. A two-dimensional semantic analysis of falling in Modern Greek: A typological and corpus-based approach. *Acta Linguistica Petropolitana*.

Gries, S. T. (2006). Corpus-based methods and cognitive semantics: The many senses of to run. In S. T. Gries, & A. Stefanowitsch (Eds.), *Corpora in Cognitive Linguistics* (pp. 57–99). Berlin/ New York: Mouton De Gruyter.

Gries, S. T., & Divjak, D. (2009). Behavioral profiles: A corpus-based approach to cognitive semantic analysis. In V. Evans, & S. Pourcel (Eds.), *New directions in cognitive linguistics* (pp. 57–75). Amsterdam: John Benjamins.

References (selection)

Gries, S. T. (2010). Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. The Mental Lexicon, 5(3), 323–346. doi 10.1075/ml.5.3.04gri

Hanks, P. (1996). Contextual dependency and lexical sets. *International Journal of Corpus Linguistics, 1*(1), 75–98.

Haspelmath, M. (1997). *Indefinite pronouns*. Oxford: Oxford University Press

Haspelmath, M. (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (Ed.), *The new psychology of language* (Vol. 2) (pp. 211–243). New York: Erlbaum.

Hilpert, M. (2016). Change in modal meanings: Another look at the shifting collocates of may. *Constructions and Frames, 8*(1), 66-85.

Jansegers, M & S. T. Gries. (2017). Towards a dynamic behavioral profile: A diachronic study of polysemous sentir in Spanish. *Corpus Linguistics and Linguistic Theory*. [online first] DOI 10.1515/cllt-2016-0080

Kay, P., & Fillmore, C. (1999). *Grmatical constructions and linguistic generalizations: The* What's X doing Y? *construction*. *Language, 75*, 1-33.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Levinson, S., Meira, S., & the Language and Cognition Group (2003). 'Natural concepts' in the spatial topological domain—Adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language, 79*(3), 485–516.

Levshina, N. (2015). *How to do Linguistics with R: Data exploration and statistical analysis*. Amsterdam/ Philadelphia: John Benjamins.

Murphy, L. (2010). *Lexical Meaning*. Cambridge: Cambridge University Press.

List, J.-M., Greenhill, S., Anderson, C., Mayer, T., Tresoldi, T, & Forkel, R. (Eds.). 2018. *Database of Cross-Linguistic Colexifications*. Jena: Max Planck Institute for the Science of Human History. http://clics.clld.org.

References (selection)

Narrog, H., & van der Auwera, J. (2011). Grammaticalization and semantic maps. In H. Narrog, & B. Heine (Eds.), *The Oxford handbook of grammaticalization* (pp. 318–327). Oxford: OUP.

Narrog, H., & Ito, S. (2007). Re-constructing semantic maps: The comitative-instrumental area. *STUF - Language Typology and Universals, 60*(4), 273–292. https://doi.org/10.1524/stuf.2007.60.4.273

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 1999. URL http: www.R-project.org/

Regier, T., Khetarpal, N., & Majid, A. (2013). Inferring semantic maps. *Linguistic Typology, 17*(1), 89–105. https://doi.org/10.1515/lity-2013-0003

Rosch, E. ([1978] 1999). Principles of categorization. In B. Lloyd, & E. Rosch (Eds.), *Cognition and Categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum [reprinted in E. Margolis, & S. Laurence (Eds.) (1999) *Concepts: Core Readings*, (pp. 189–206). Cambridge, MA: MIT Press].

Rosch, E., & Mervis, C. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology, 7*, 573–605.

Tyler, A., & Evans, V. (2001). 'Reconsidering prepositional polysemy networks: the case of over'. *Language, 77*(4), 724–65.

Tyler, A., & Evans, V. (2003). *The Semantics of English Prepositions. Spatial Scenes, Embodied Meaning and Cognition*. Cambridge: Cambridge University Press

Thank you!

E-mail:
ageorgakopoulos@hse.ru
athanasphil@gmail.com

Webpage:
https://thanasis-georgakopoulos.weebly.com/
https://scopus.academia.edu/ThanasisGeorgakopoulos