Thanasis Georgakopoulos
(University of Liège)
Stéphane Polis
(F.R.S.-FNRS / University of Liège)

ICGL – London – 2017.09.09

Towards a web-based platform for plotting, visualizing and enriching diachronic semantic maps

With a case study on the Greek and Egyptian temporal semantic field

Outline of the talk

- Introduction
  - What are semantic maps?
  - Le Diasema (LExical DIAchronic SEmantic MAps)
- Building semantic maps
  - The network inference problem
  - An algorithm for automatic plotting (Regier et al. 2013)
- Two steps forwards with the automatic plotting
  - The significance of weights
  - Inferring diachronic semantic maps
- Dynamicizing a crosslinguistic semantic map of time-related meanings with Ancient Greek and Egyptian

Introduction

Semantic maps & Le Diasema

What are semantic maps?

'A semantic map is a geometrical representation of functions (…) that are linked by connecting lines and thus constitute a network' (Haspelmath 2003). It constitutes a 'model of attested variation' (Cysouw 2007).

What are semantic maps?

'A semantic map is a geometrical representation of functions (…) that are linked by connecting lines and thus constitute a network' (Haspelmath 2003). It constitutes a 'model of attested variation' (Cysouw 2007).

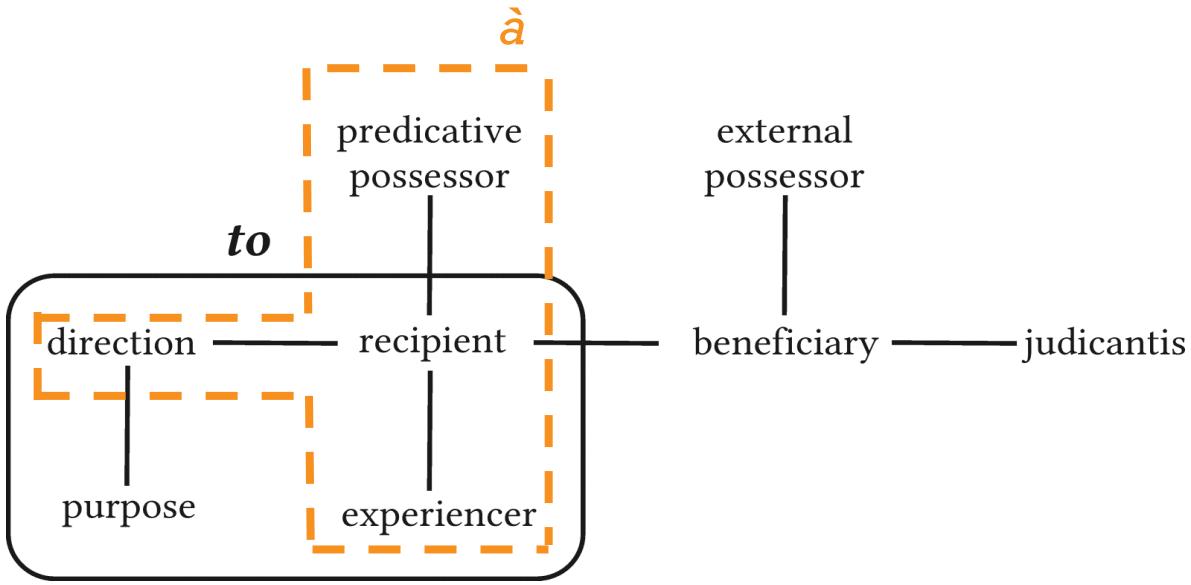


Figure 1. A semantic map of typical dative functions / the boundaries of English *to* and French *à* (based on Haspelmath 2003: 213, 215)

- Sense distinctions are based on **cross-linguistic evidence** and designed to have **cross-linguistic validity**
- They combine the **onomasiological** and the **semasiological** perspective
- **Multifunctionality**. No commitment to a particular claim about conventionalization of senses

What are semantic maps?

Weighted semantic maps

What are semantic maps?

Weighted semantic maps

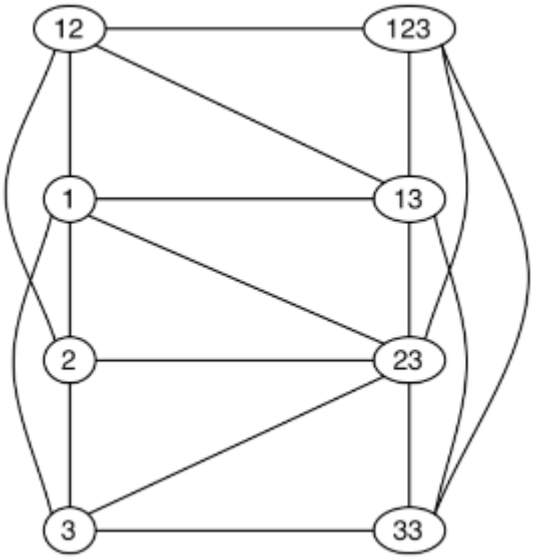


Figure 2a. A simple semantic map of person marking (Cysouw 2007: 231)

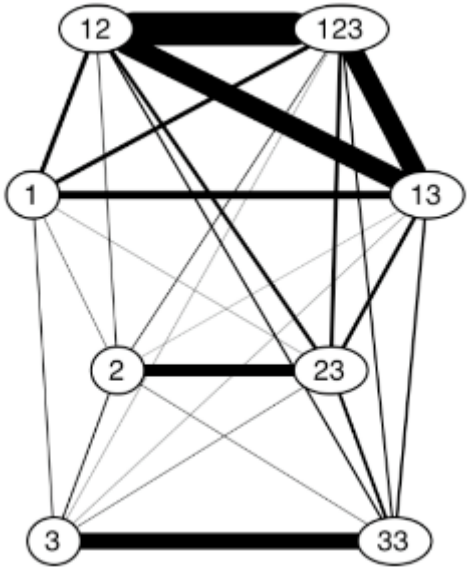


Figure 2a. A weighted semantic map of person marking (Cysouw 2007: 233)

The weighted edges capture the frequencies of each polysemy pattern

What are semantic maps?

Diachronic ('dynamicized') semantic maps

What are semantic maps?

Diachronic ('dynamicized') semantic maps

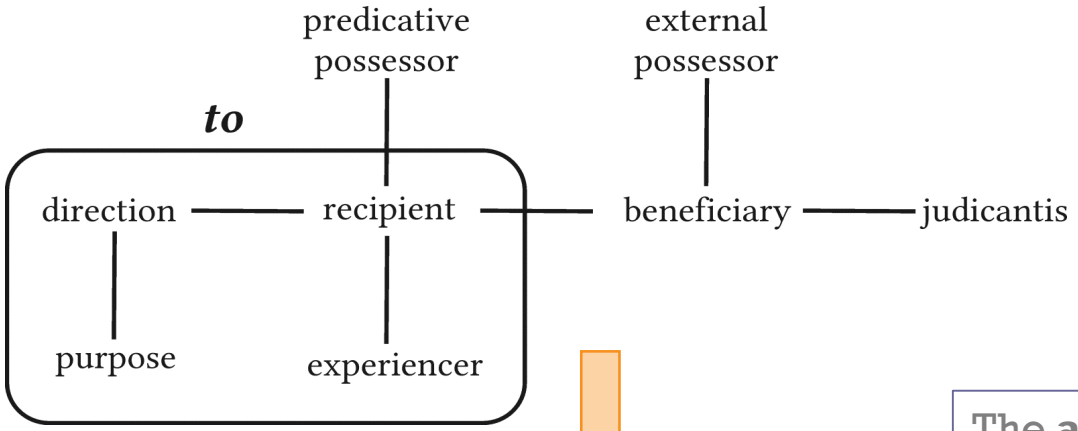


Figure 1. A semantic map of dative functions (Haspelmath 2003: 213)

The **arrows** designate directionality of change

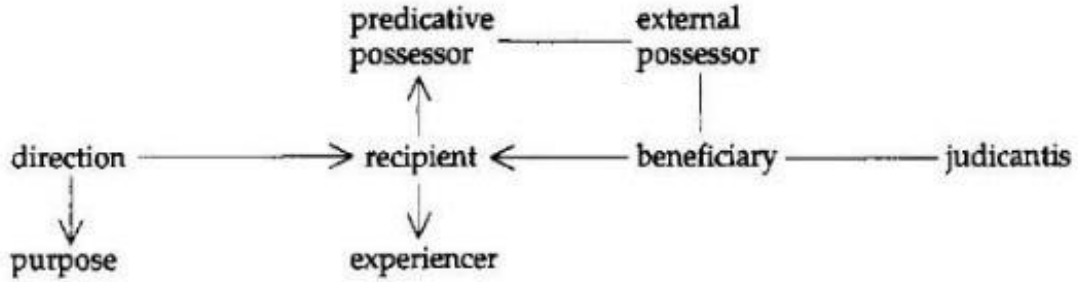


Figure 3. A dynamicized semantic map of dative functions (Haspelmath 2003: 234)

What are semantic maps?

Lexical semantic maps

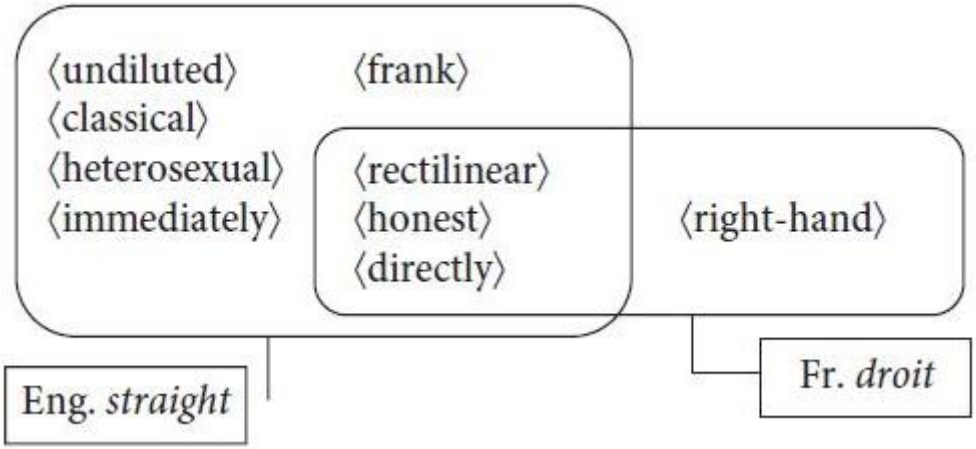


Figure 4. Overlapping polysemies: Eng. straight *vs.* Fr. droit
(François 2008: 167)

Colexification = polyfunctionality

"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)

What are semantic maps?

Lexical semantic maps

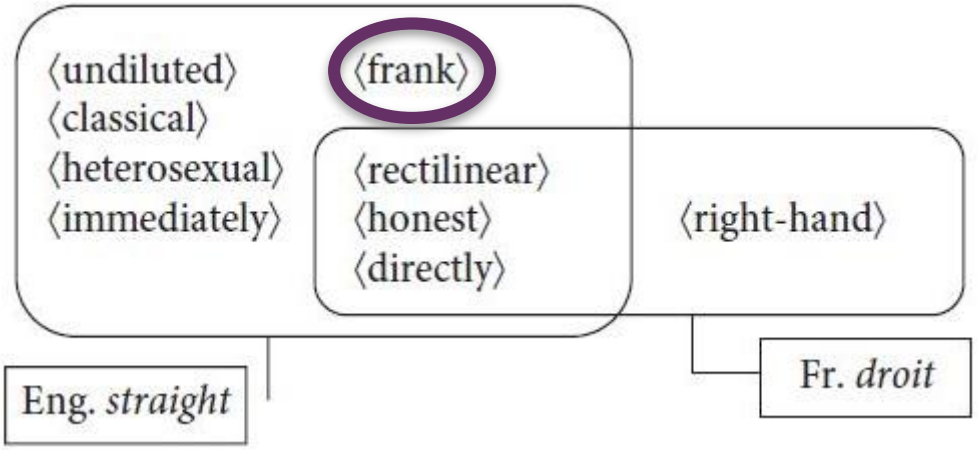


Figure 4. Overlapping polysemies: Eng. straight *vs.* Fr. droit
(François 2008: 167)

Colexification = polyfunctionality

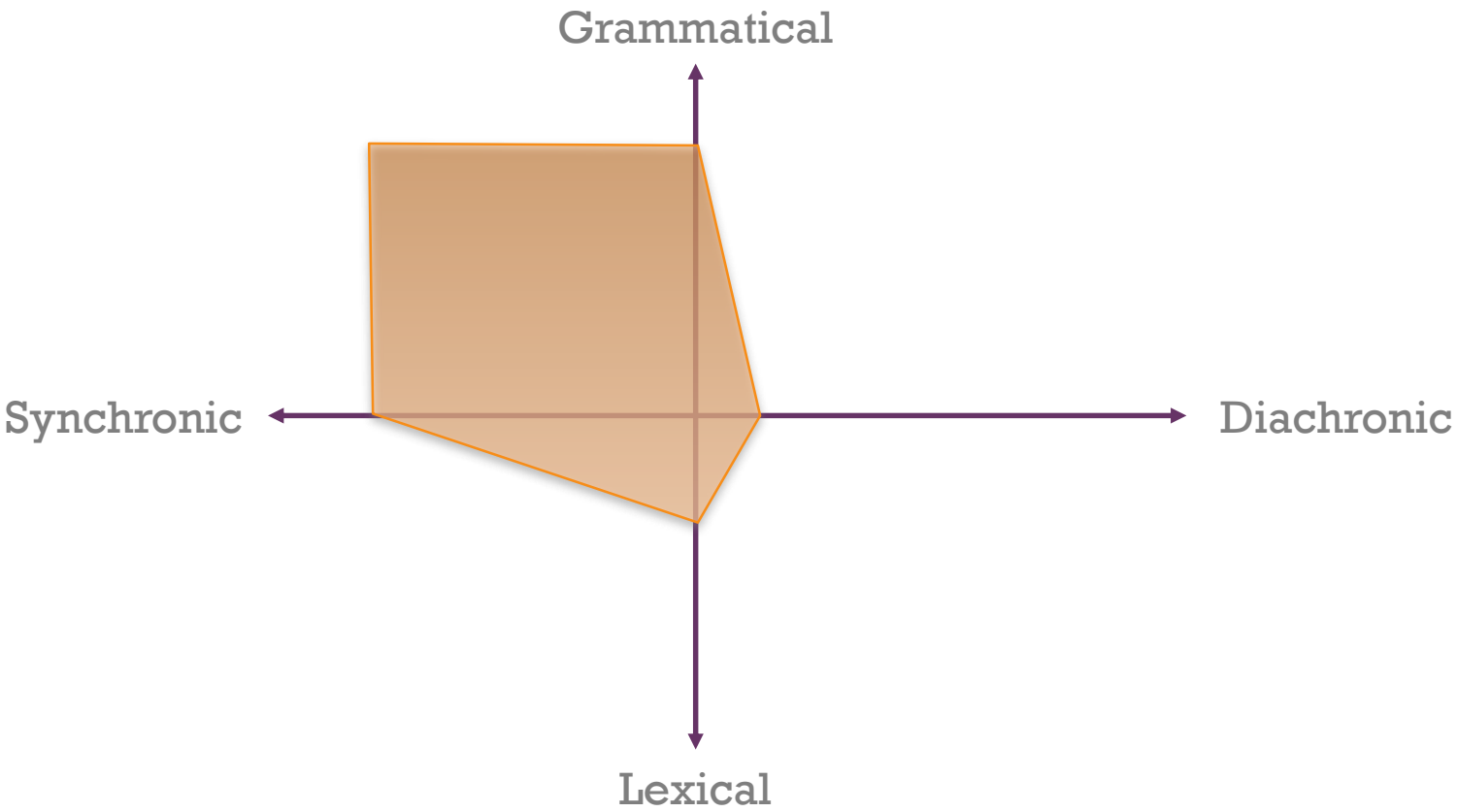
"A given language is said to colexify two functionally distinct senses if, and only if, it can associate them with the same lexical form"
(François 2008: 170)

"A function is put on the map if there is at least one pair of languages that differ with respect to this function"
(Haspelmath 2003: 217; cf. François 2008: 168-169)

Le Diasema

Filling a gap

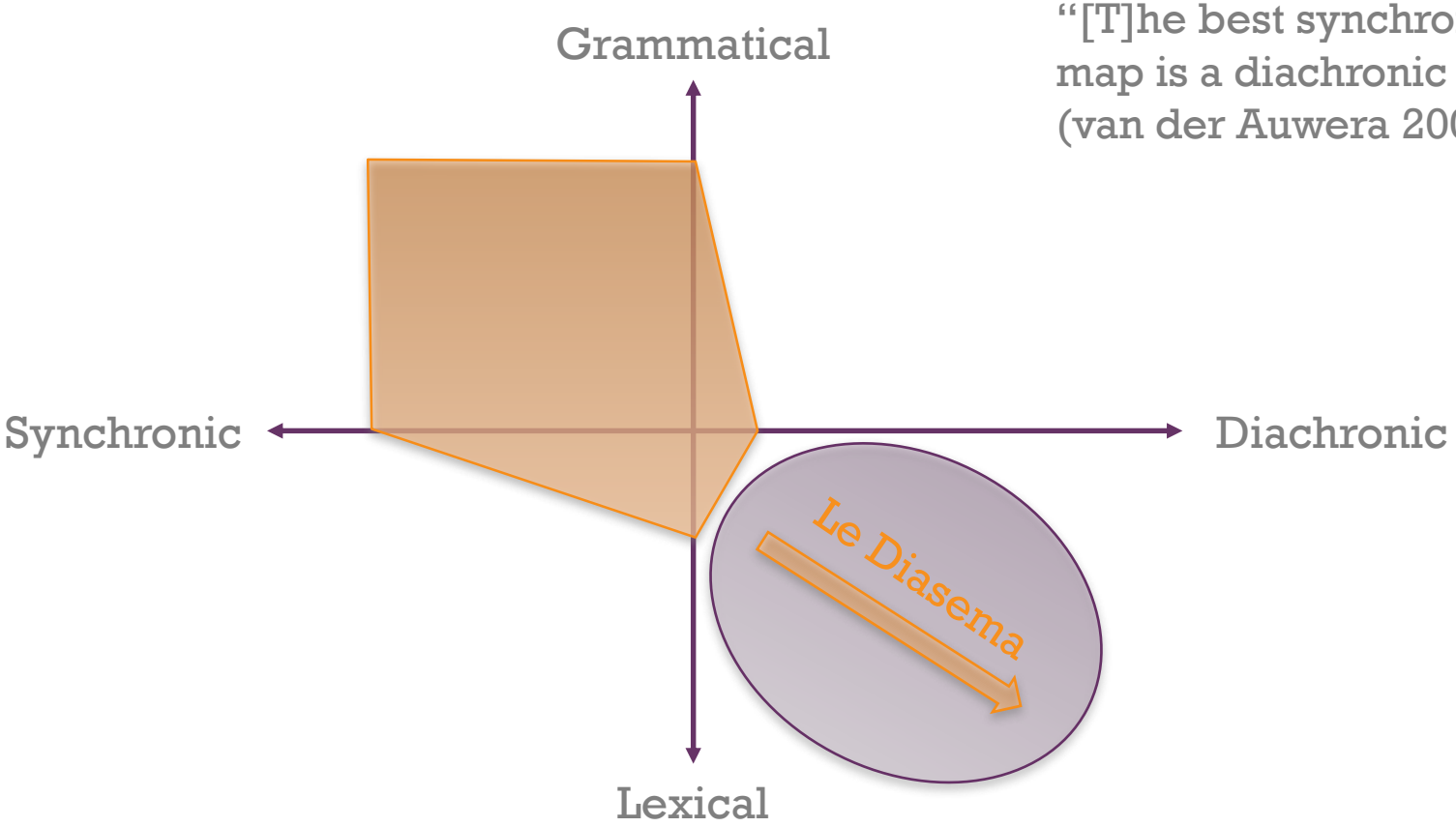
- Adding a diachronic dimension to semantic maps of content words



Le Diasema

Filling a gap

- Adding a diachronic dimension to semantic maps of content words



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

Le Diasema

- Duration
  - December 2016 – December 2018
- Main research question
  - How semantic maps make significant predictions about language change at the lexical level?
- Funding schemes

Université de Liège

BeIPD-COFUND

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

Le Diasema

- To incorporate the diachronic dimension into semantic maps of content words
- To extend the method so as to also include information about the cognitive and cultural factors behind the development of the various meanings
- To create an online platform for automatically plotting diachronic semantic maps based on polysemy data from the languages of the world

Le Diasema

Specific objective for today

- To incorporate the diachronic dimension into semantic maps of content words
- To extend the method so as to also include information about the cognitive and cultural factors behind the development of the various meanings
- To create an online platform for automatically plotting diachronic semantic maps based on polysemy data from the languages of the world

+

Inferring semantic maps

A network inference problem

Inferring semantic maps

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

Inferring semantic maps

- Indeed, a significant limitation of the semantic map method is that it is practically impossible to handle large-scale crosslinguistic datasets manually
  - Up until recently, they were considered "not mathematically well-defined or computationally tractable, making it impossible to use with large and highly variable crosslinguistic datasets" (Croft & Poole, 2008: 1)
  - This led to the development of other approaches using statistical scaling techniques, like Multi-Dimensional Scaling (MDS)

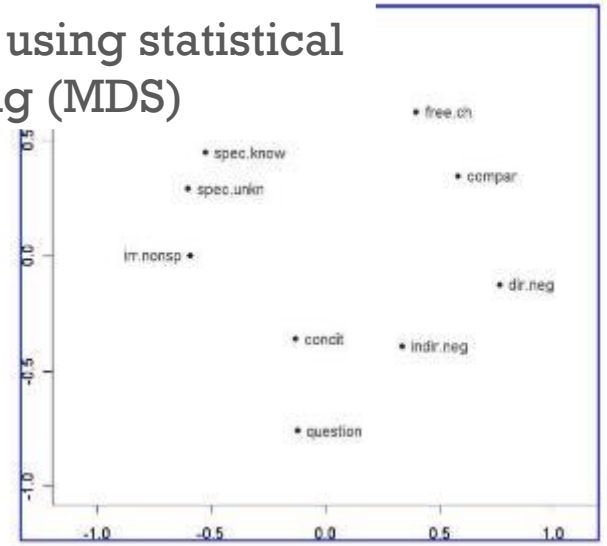


Figure 5. MDS analysis of Haspelmath's 1997 data on indefinite pronouns (Croft & Poole 2008: 15)

Inferring semantic maps

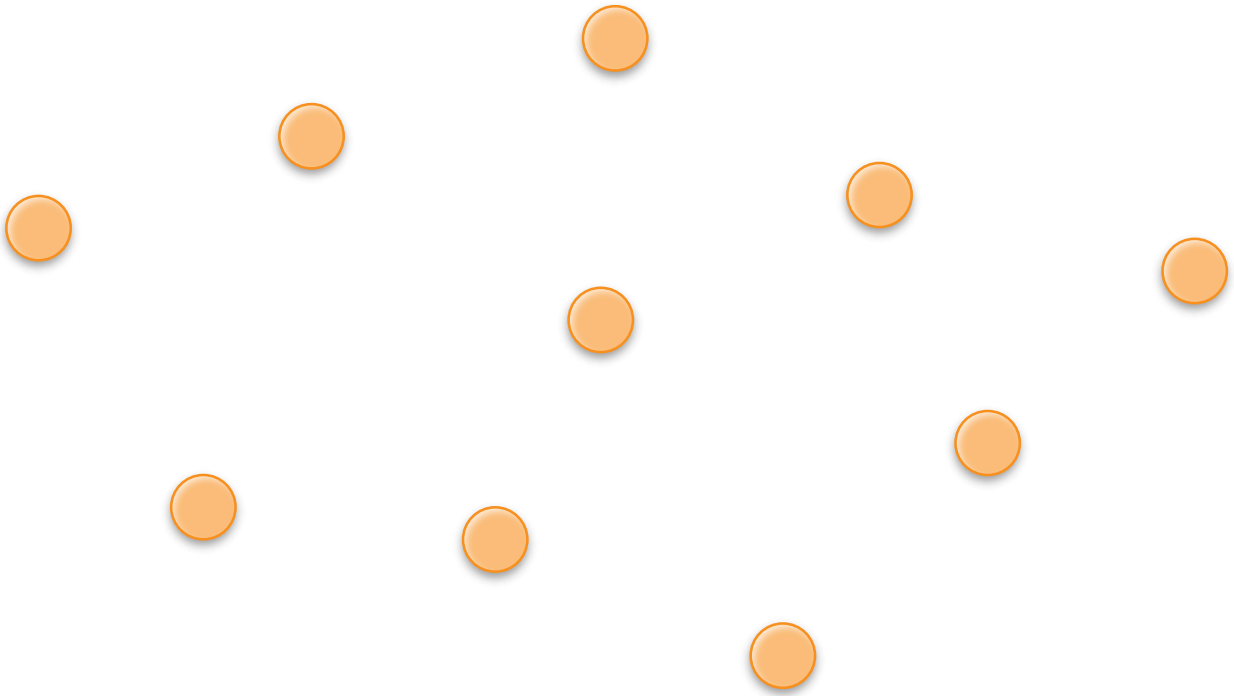
- However, Regier, Khetarpal, and Majid showed that the semantic map inference problem is "formally identical to another problem that superficially appears unrelated: inferring a social network from outbreaks of disease in a population" (Regier et al., 2013: 91)

Inferring semantic maps

- What's the idea?

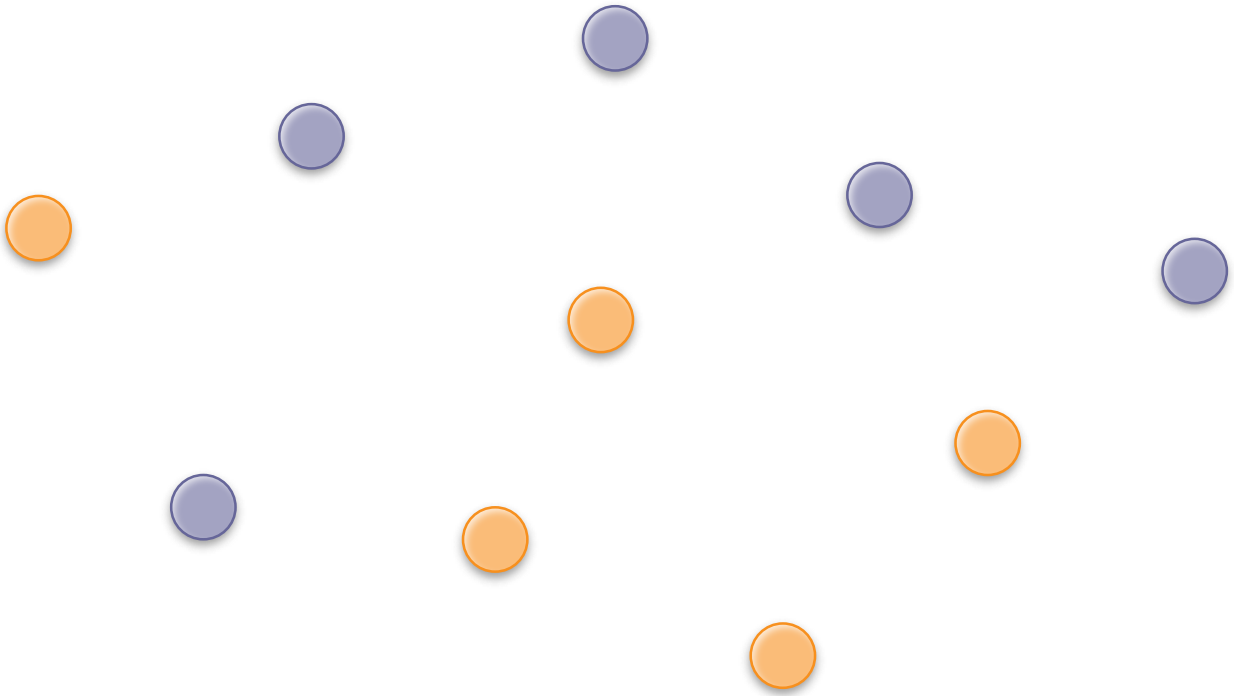
Inferring semantic maps

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



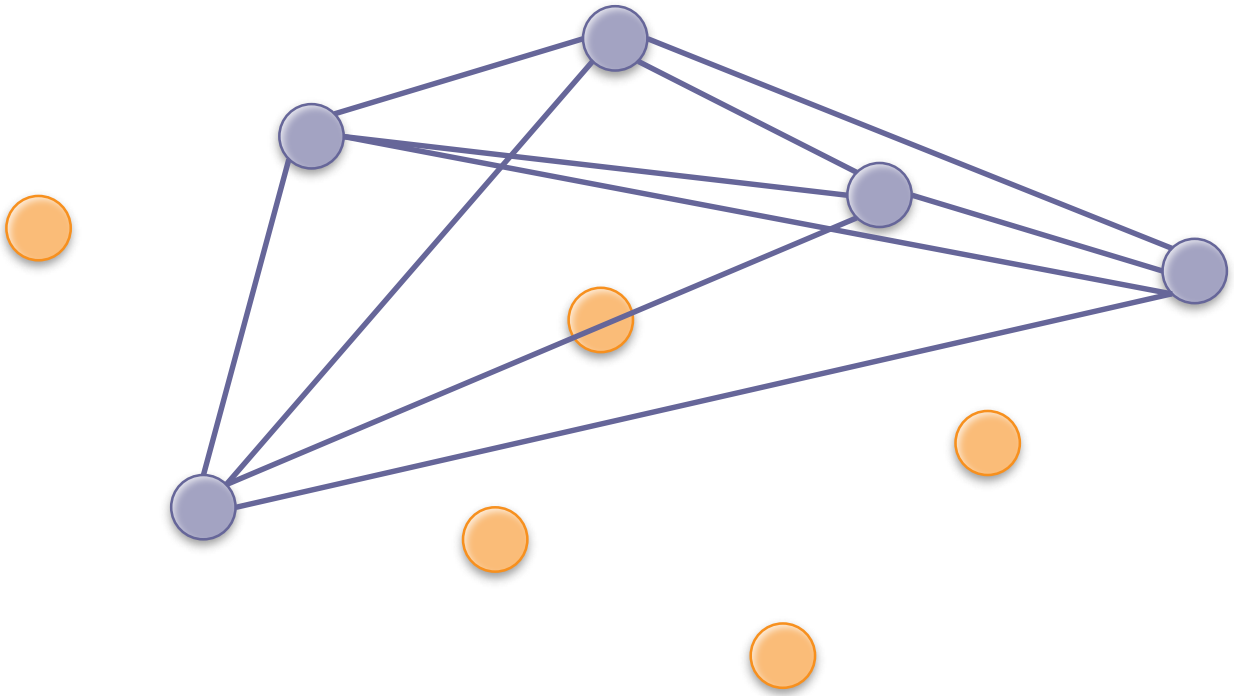
Inferring semantic maps

- 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)



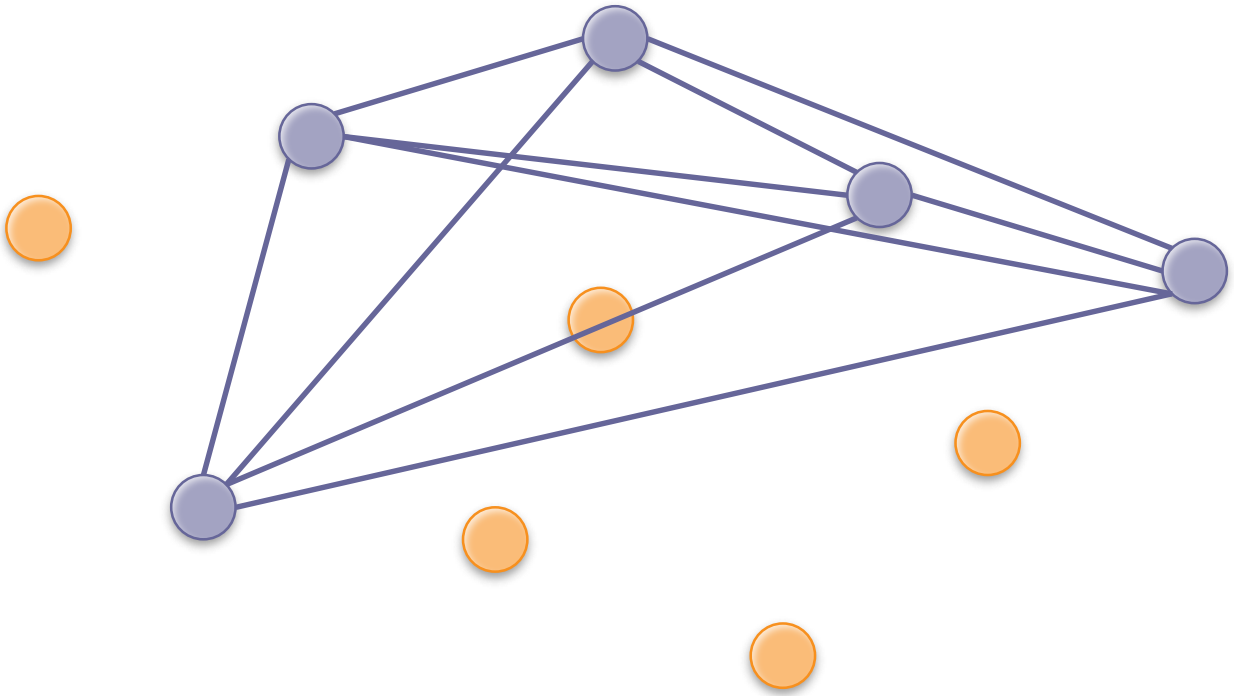
Inferring semantic maps

- 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)



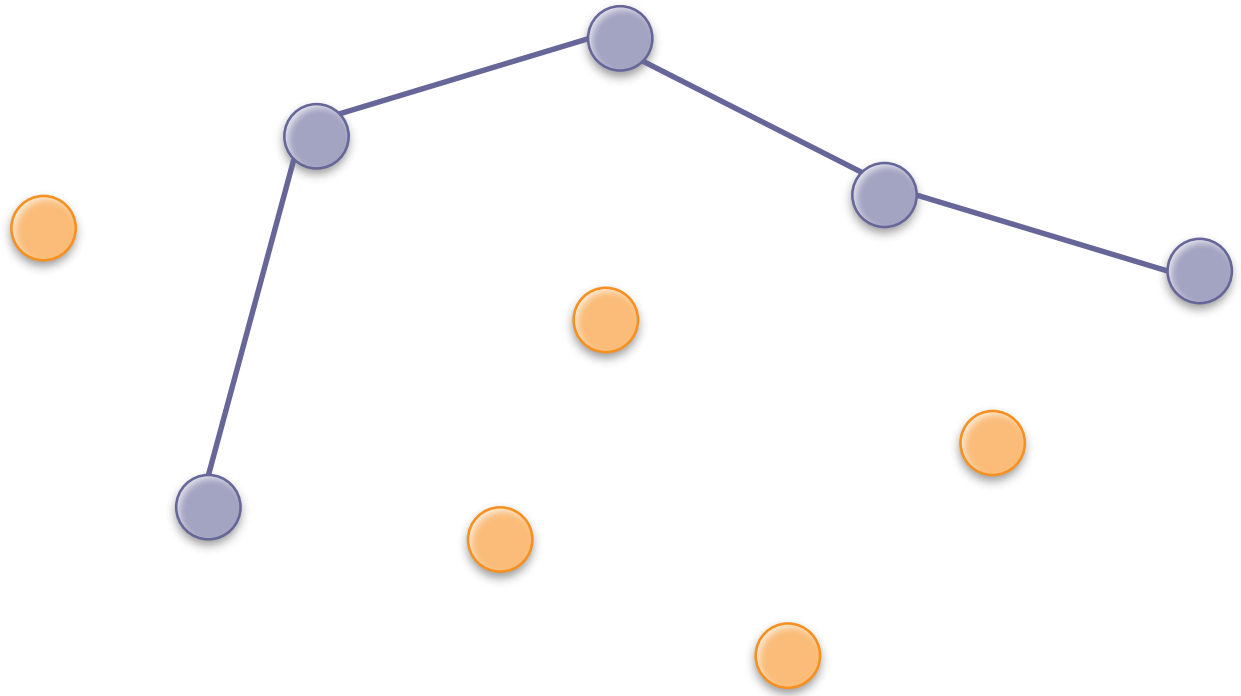
Inferring semantic maps

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



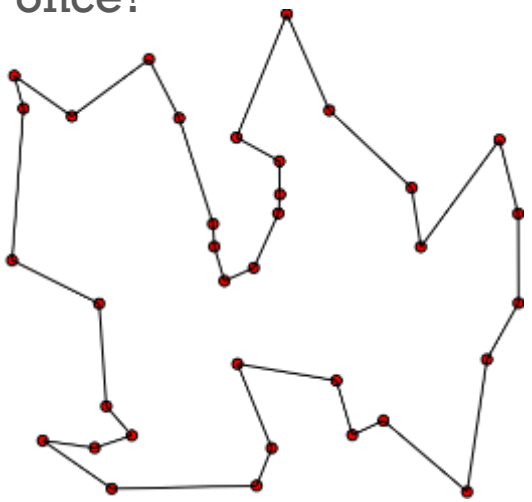
Inferring semantic maps

- 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?
  - The goal would be to have all the social agents connected with as few edges as possible
  - Such a **Network Inference** problem looks intuitively simple, but is computationally hard to solve
    - Cf. the travelling salesman problem [TSP]: "Given a list of cities and the distance between each pair of cities, what is the shortest possible route that visits each city exactly once?"

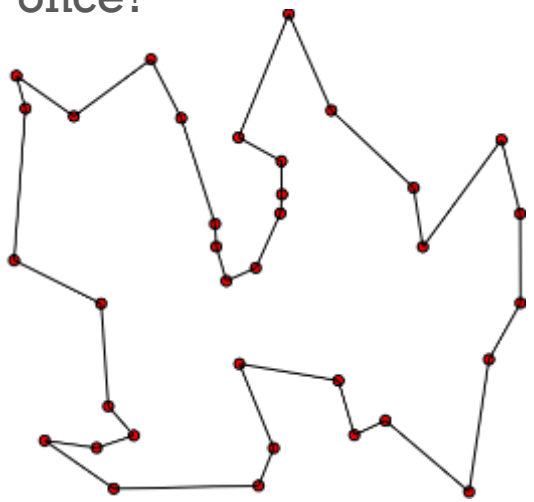


Inferring semantic maps

- What's the idea?
  - The goal would be to have all the social agents connected with as few edges as possible
  - Such a **Network Inference** problem looks intuitively simple, but is computationally hard to solve
    - Cf. the travelling salesman problem [TSP]: "Given a list of cities and the distance between each pair of cities, what is the shortest possible route that visits each city exactly once?"
  - 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,***]

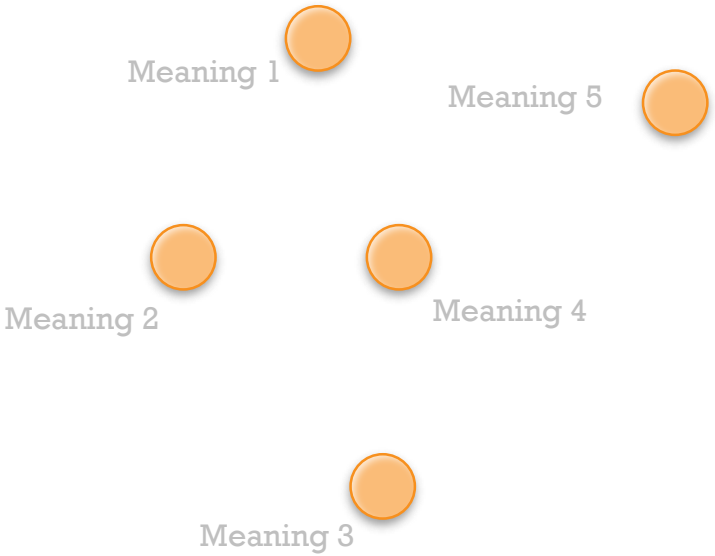


Inferring semantic maps

- How does it transfer to semantic maps?

Inferring semantic maps

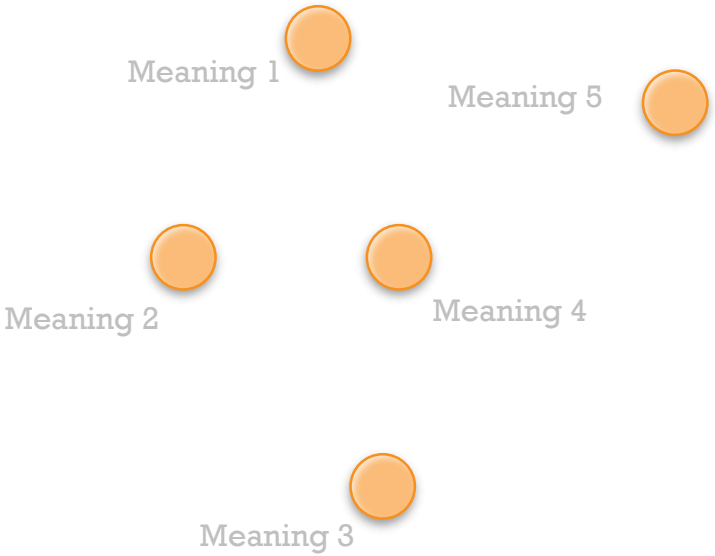
- How does it transfer to semantic maps?
  - Nodes are meanings



Meaning	1	2	3	4	5

Inferring semantic maps

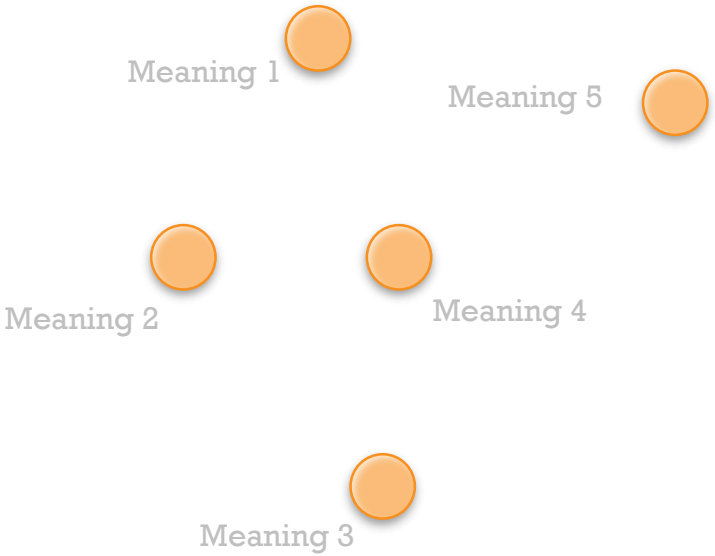
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are Polysemic items



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

Inferring semantic maps

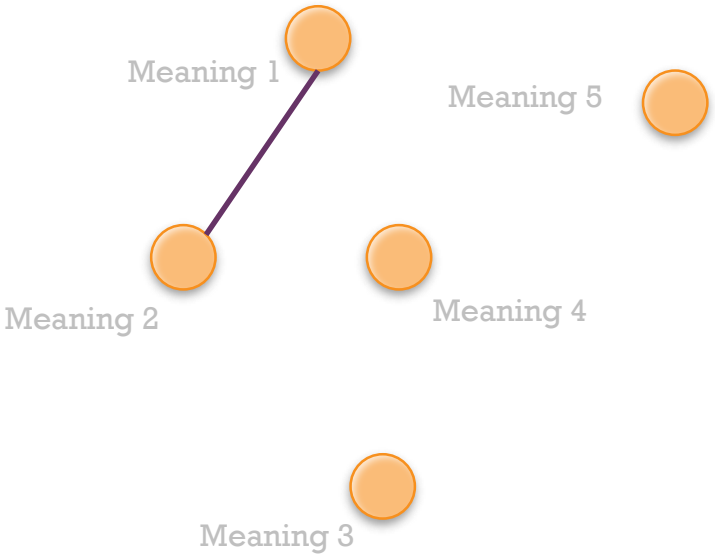
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are Polysemic items
  - One connects the nodes economically based on these constraints



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

Inferring semantic maps

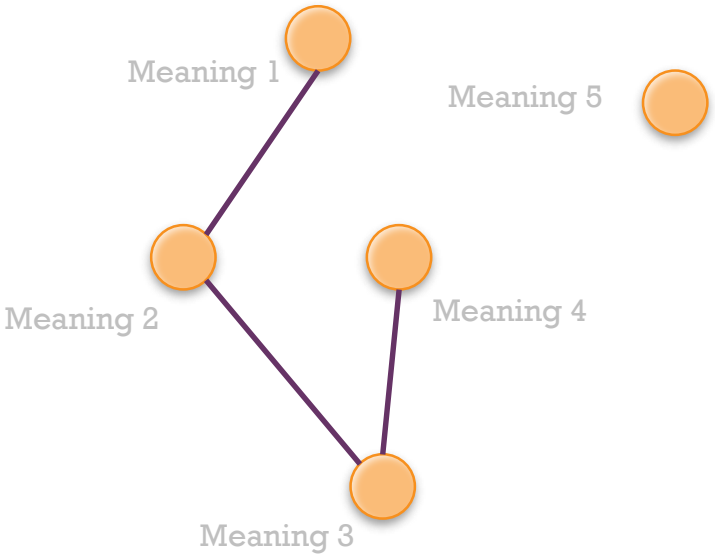
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are Polysemic items
  - One connects the nodes economically based on these constraints



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

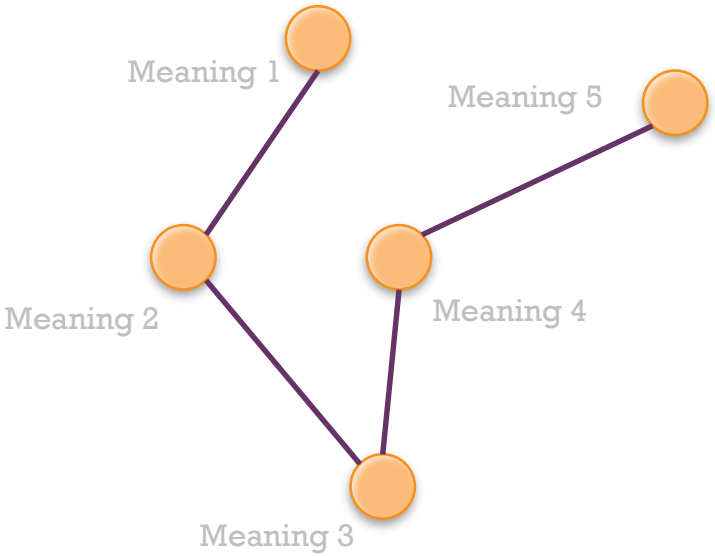


Meaning	1	2	3	4	5
Polysemic item A	√	√			
Polysemic item B		√	√	√	

(but no edge needed between meaning 2 and meaning 4, as they are connected through meaning 3)

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

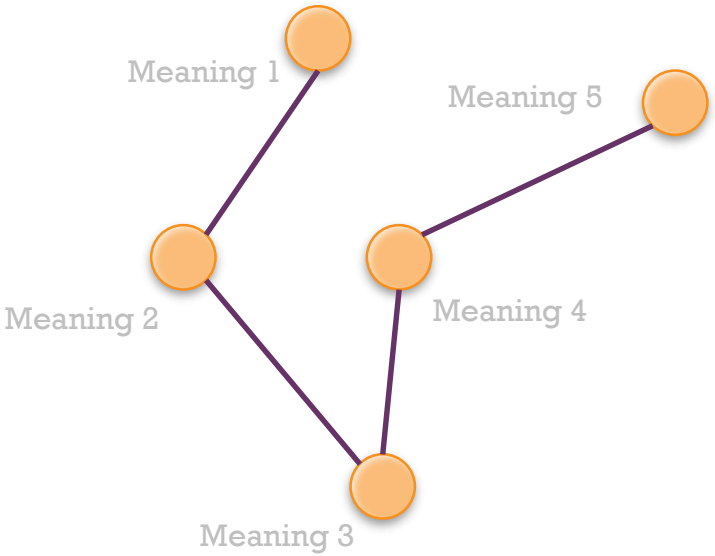


Meaning	1	2	3	4	5
Polysemic item A	√	√			
Polysemic item B		√	√	√	
Polysemic item C			√	√	√

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

- Having tested the algorithm on the crosslinguistic data of Haspelmath (1997) and Levinson et al. (2003), Regier et al. (2013) conclude that the approximations produced by the algorithm are of *high quality*, which means that they produce equal or better results than the manually plotted maps

Inferring semantic maps

INPUT
(lexical matrix)

Language	*Word*	Specific Known	Specific Unknown	Irrealis Non-specific	Question	Conditional	Indirect Negation
		SK	**SU**	**IR**	**QN**	**CD**	**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	Specific Unknown	Irrealis Non-specific	Question	Conditional	Indirect Negation
		SK	**SU**	**IR**	**QN**	**CD**	**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

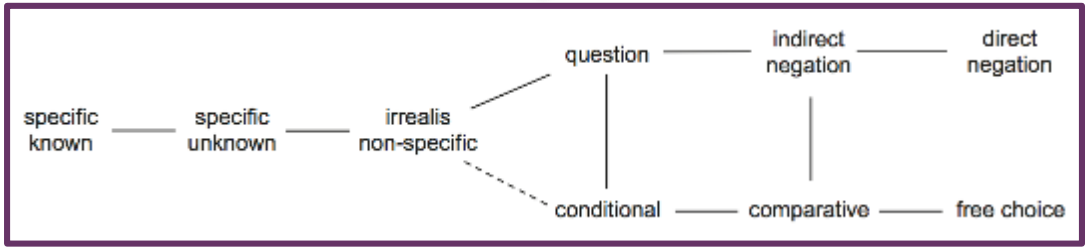
INPUT
(lexical matrix)

Language	*Word*	Specific Known	Specific Unknown	Irrealis Non-specific	Question	Conditional	Indirect Negation
		SK	**SU**	**IR**	**QN**	**CD**	**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)



Automatic plotting: Two steps forwards

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

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

Can weighted and diachronic semantic map be plotted automatically?

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

Can weighted and diachronic semantic map be plotted automatically?

YES!

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

Can weighted and diachronic semantic map be plotted automatically?

YES!

HOW?

Automatic plotting: Two steps forward

Weighted semantic maps

Automatic plotting: Two steps forward

Weighted semantic maps

- Generate the map with a modified version of the algorithm of Regier et al. (2013)

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)
```

Automatic plotting: Two steps forward

Weighted semantic maps

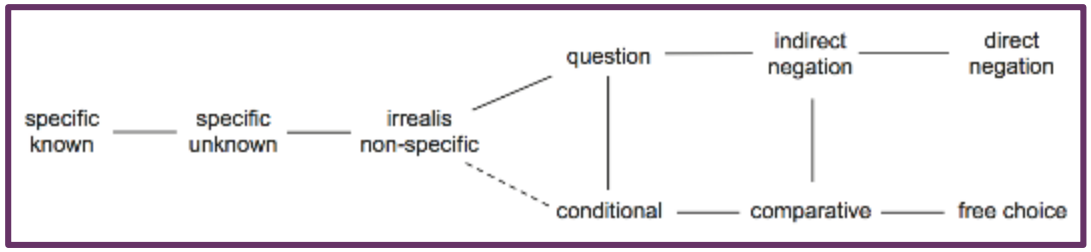
- Generate the map with a modified version of the algorithm of Regier et al. (2013)
  - PRINCIPLE: for each edge that is being added between two meanings of the map by the algorithm, check in the lexical matrix how many times this specific polysemy pattern is attested, and increase the weight of the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

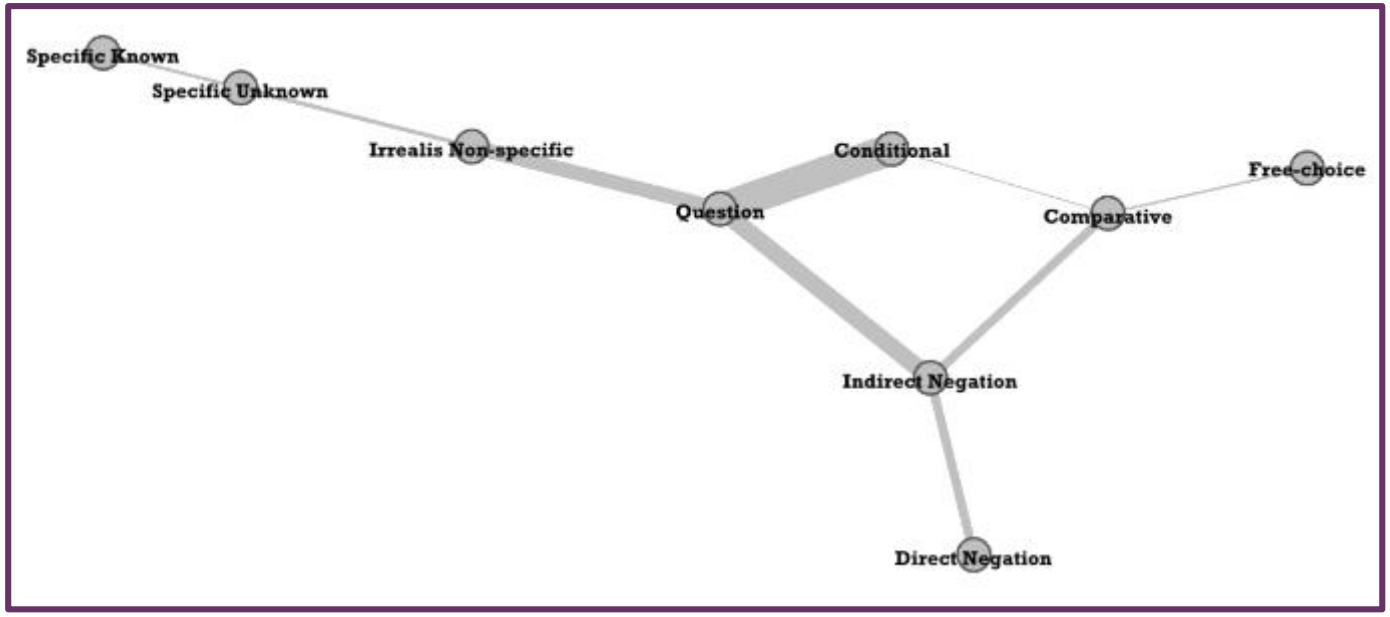
  - Based on the data of Haspelmath (1997), kindly provided by the author, the result between a non-weighted and a weighted semantic map are markedly different

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

Diachronic semantic maps

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

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

- Expand the lexical matrix so as to include information about diachrony
- Generate the graph with the algorithm of Regier et al. (2013)

Automatic plotting: Two steps forward

Diachronic semantic maps

- Expand the lexical matrix so as to include information about diachrony
- Generate the graph with the algorithm of Regier et al. (2013)
- Enrich the graph with oriented edges (where relevant)
  - PRINCIPLE: 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.

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

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
```

Automatic plotting: Two steps forward

Diachronic semantic maps

INPUT (diachronic lexical matrix)

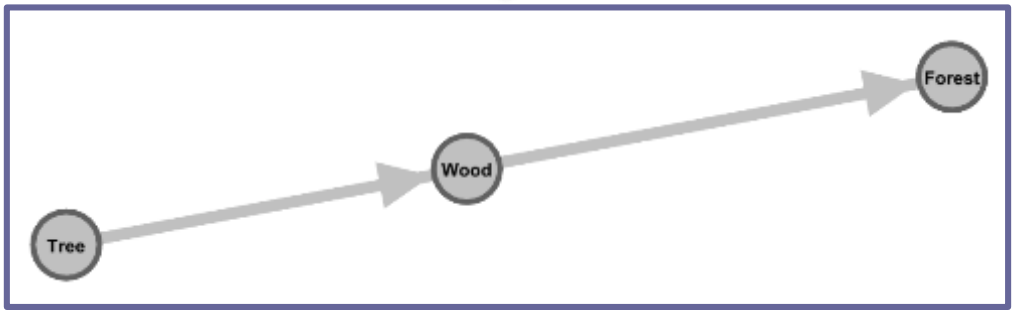
Source of constraint	*Constraint name*	*Constraint Time*	Sense_1 Tree	Sense_2 Wood	Sense_3 Forest
Language_1	Word_1	0	1	0	0
Language_1	Word_1	1	1	1	0
Language_2	Word_1	0	1	0	0
Language_2	Word_2	0	0	1	0
Language_2	Word_2	1	0	1	1
Language_3	Word_1	0	1	1	0
Language_3	Word_2	0	0	0	1

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)



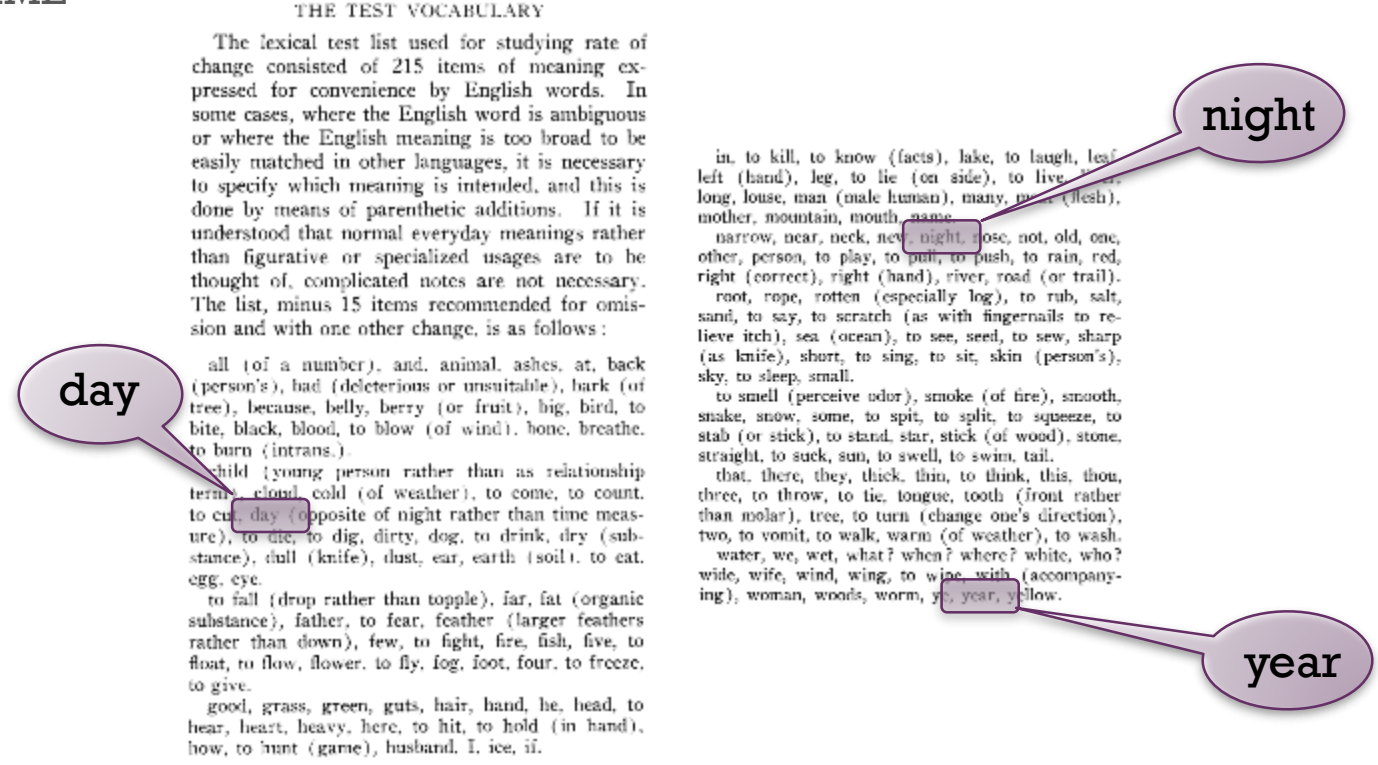
The semantic extension of time-related lexemes

Inferring a semantic map based
on cross-linguistic colexification patterns
and enriching it with diachronic data

The semantic extension of time-related lexemes

Plotting a synchronic semantic map

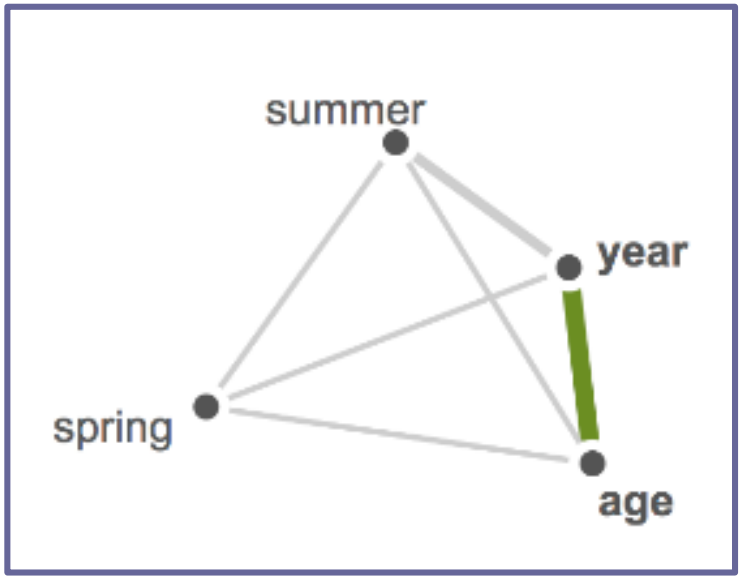
- For the purpose of universality and stability, we chose the entries for time-related concepts in the Swadesh 200-word list (Swadesh 1952: 456-457)
  - DAY/DAYTIME
  - NIGHT
  - YEAR



The semantic extension of time-related lexemes

Plotting a synchronic semantic map

- We identified in the database of Crosslinguistic Colexifications (CLICs; http://clics.lingpy.org/main.php; List et al. 2014) the main polysemy patterns attested for these three meanings (subgraph approach) [16 meanings]
  - **DAY/DAYTIME**: CLOCK/TIMEPIECE, HOUR, SEASON, SUN, TIME, WEATHER
  - **NIGHT**: DARK (in color), DARKNESS, BLACK, OBSCURE
  - **YEAR**: AGE, SPRING, SUMMER



The semantic extension of time-related lexemes

Plotting a synchronic semantic map

- All the colexification patterns attested for these 16 meanings were gathered in the CLICs source files (http://clics.lingpy.org/download.php), ending up with **381 colexification** patterns

	A	B	C
119	day	afternoon	hau_std:rana//ket_std:i?//plj_std:piidi//rus_std:den//tli_std:yakyee
120	day	again	kha_std:sngi
121	day	age	gui_std:ˈara//yad_std:hnda
122	day	anger	tzz_std:k'ak'al
123	day	bright	tzz_std:k'ak'al
124	day	clock, timepiece	gue_std:wurlngarn//sei_std:šā?
125	day	cloud	haw_std:ao
126	day	country	cbr_std:niti//shp_std:niti
127	day	dawn	haw_std:ao//waw_std:enmari
128	day	doubt	haw_std:lā
129	day	earth, land	cag_std:nafu//haw_std:ao//mri_std:ao//tzz_std:osil
130	day	east	tob_std:na?a?k
131	day	fever	tzz_std:k'ak'al
132	day	fin (dorsal)	haw_std:lā
133	day	fire	jpn_std:hi
134	day	go	ote_std:pa//oym_std:aa
135	day	go away, depart	ote_std:pa
136	day	hour	sap_Standard:aknim//shb_std:tham
137	day	lamp, torch	ito_std:uwayo
138	day	lick	cmn_std:tian
139	day	light (in color)	mri_std:ao
140	day	light (noun)	con_std:a?ta//crt_std:xloma//haw_std:ao//hdn_Northern:ʔkatʰkáa//ito_std:uwayo//mzh
141	day	live, living, life	shp_std:niti
142	day	Monday	shp_std:niti
143	day	morning	crt_std:xloma//guq_std:kreibu
144	day	noon, midday	ind_std:siang//plj_std:piidi

The semantic extension of time-related lexemes

Plotting a synchronic semantic map

- All the colexification patterns attested for these 16 meanings were gathered in the CLICs source files (http://clics.lingpy.org/download.php), ending up with **381 colexification** patterns
- These synchronic polysemy patterns were converted into a **lexical matrix**

```
Tmap = [Tsenses]
for t in Tclean:
        split_langWord = t[2].split('//')
        for couple in split_langWord:
                langWord = couple.split(':')
                line = [langWord[0], langWord[1]]
                for i in range (2,len(Tsenses)):
                        line.append('0')
                line[Tsenses.index(t[0])] = '1'
                line[Tsenses.index(t[1])] = '1'
                Tmap.append(line)
```

Languages Forms Meanings

	A	B	C	D	E	F
1			age	acid, sour	city, town	day
2	yad_std	hnda	1	1	0	1
3	vec_std	edat	1	0	0	0
4	jpn_std	toshi	1	0	1	0
5	gui_std	ʻara	1	0	0	1
6	nog_std	йуз	1	0	0	0
7	mri_std	pakeke	1	0	0	0
8	pbb_std	hiʔph	1	0	0	0
9	khv_Khvarshi	замана	1	0	0	0

1 when a meaning is attested for one form

Python script α

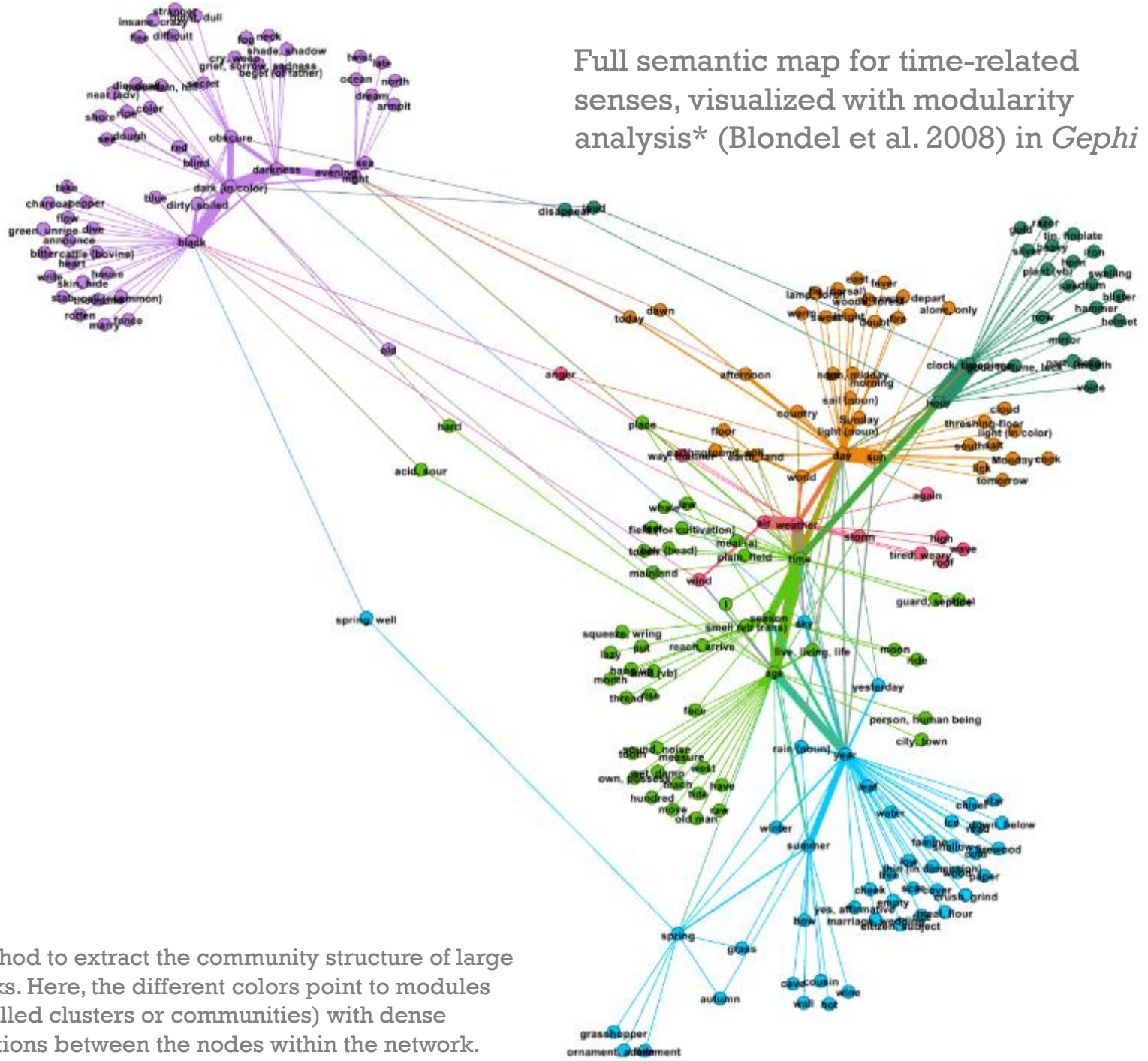
Lexical matrix

The semantic extension of time-related lexemes

Plotting a synchronic semantic map

- All the colexification patterns attested for these 16 meanings were gathered in the CLICs source files (http://clics.lingpy.org/download.php), ending up with **381 colexification** patterns
- These synchronic polysemy patterns were converted into a **lexical matrix**
- From this lexical matrix, we inferred a **weighted semantic map** based on the adapted version of the algorithm suggested by Regier et al. (2013)

Full semantic map for time-related senses, visualized with modularity analysis* (Blondel et al. 2008) in *Gephi*



* A method to extract the community structure of large networks. Here, the different colors point to modules (also called clusters or communities) with dense connections between the nodes within the network.

The semantic extension of time-related lexemes

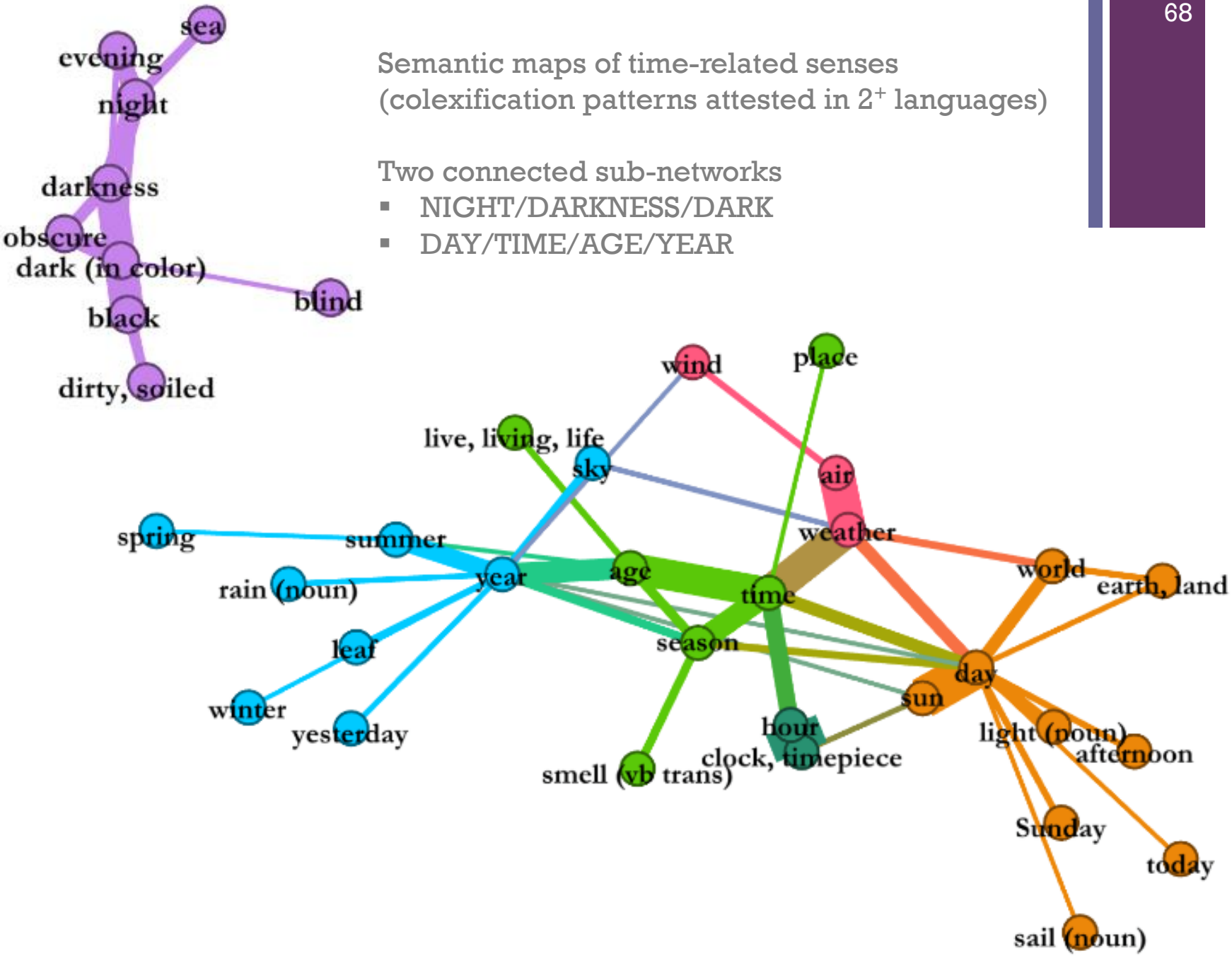
Plotting a synchronic semantic map

- All the colexifications patterns attested for theses 16 meanings were gathered in the CLICs source files (http://clics.lingpy.org/download.php), ending up with **381 colexifications** patterns
- These synchronic polysemy patterns were converted into a **lexical matrix**
- From this lexical matrix, we inferred a **weighted semantic map**, based on an adapted version of the algorithm by Regier et al. (2013)
- The **weighted edges** allow us to get rid of poorly attested patterns of polysemy (keeping only those attested in 2^+ languages),

Semantic maps of time-related senses
(colexification patterns attested in 2^+ languages)

Two connected sub-networks

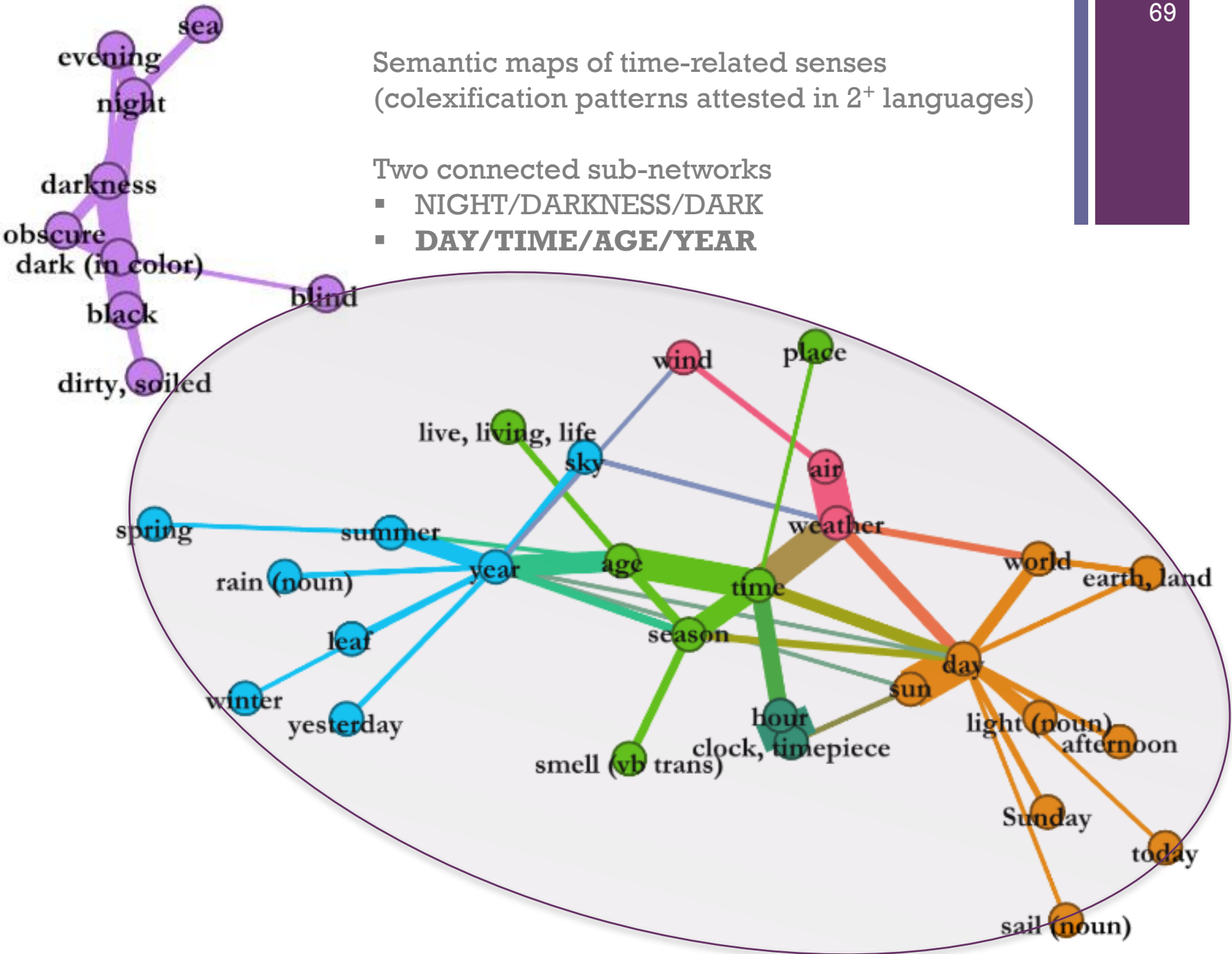
- NIGHT/DARKNESS/DARK
- DAY/TIME/AGE/YEAR



Semantic maps of time-related senses
(colexification patterns attested in 2^+ languages)

Two connected sub-networks

- NIGHT/DARKNESS/DARK
- **DAY/TIME/AGE/YEAR**



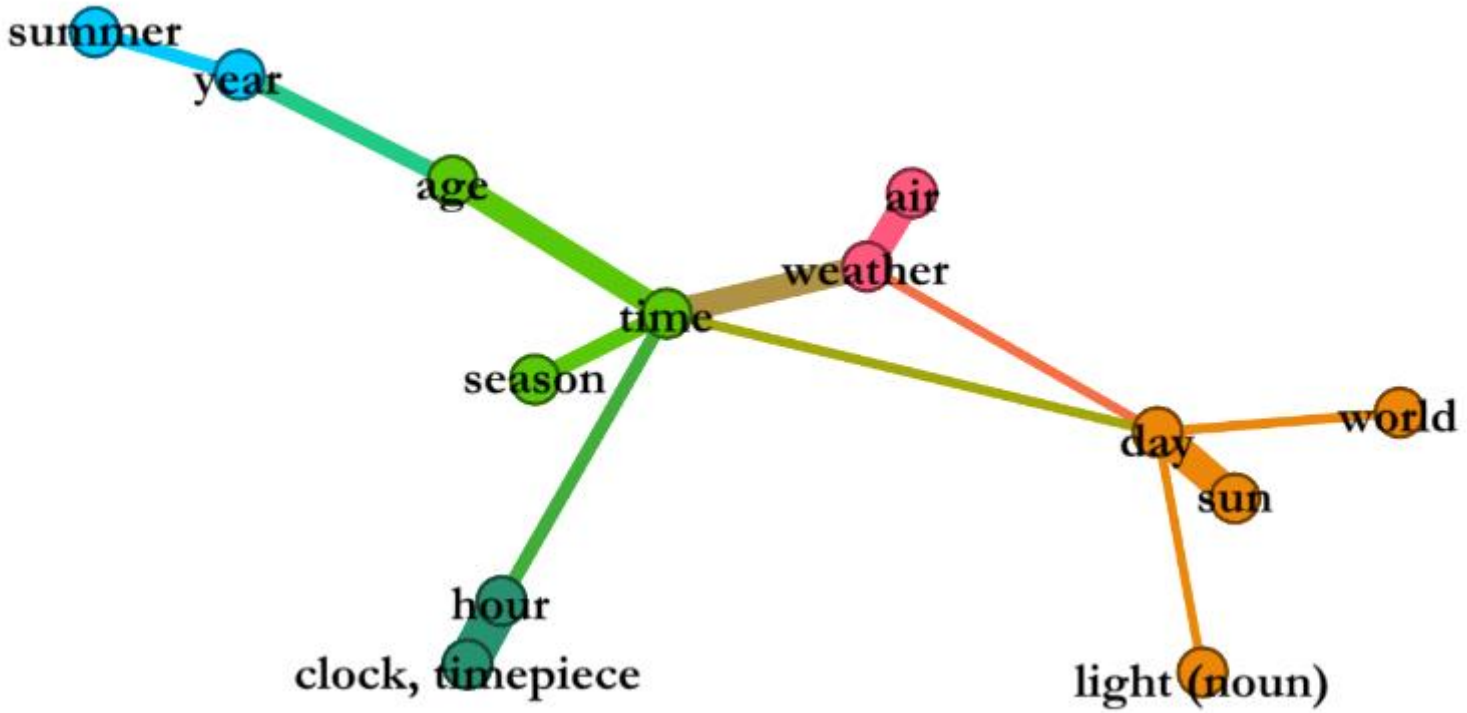
The semantic extension of time-related lexemes

Towards a dynamicized semantic map

The semantic extension of time-related lexemes

Towards a dynamicized semantic map

- In order to investigate directionality of change, 13 meanings that are connected on this map in at least 8 different languages were kept as a basis for diachronic investigation



The semantic extension of time-related lexemes

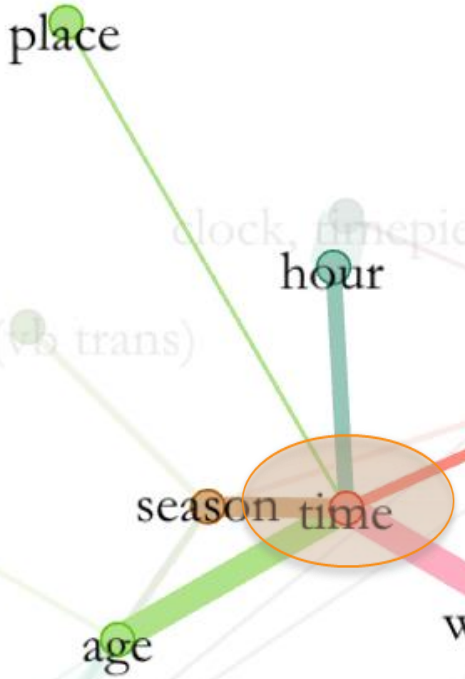
Towards a dynamicized semantic map

- **Diachronic data**
  - Ancient Greek (8th – 4th c. BC; in a few cases till 1st c. BC)
    - Perseus digital library (http://www.perseus.tufts.edu/hopper/), TLG (http://stephanus.tlg.uci.edu)
    - Cunliffe (*A lexicon of the Homeric Dialect*), LSJ
  - Ancient Egyptian (26th c. BC – 10th c. AD)
    - Thesaurus Linguae Aegyptiae (http://aaew.bbaw.de/tla/)
    - The Ramses corpus (http://ramses.ulg.ac.be),
    - Lexical resources (Coptic etymological dictionaries)

The semantic extension of time-related lexemes

Towards a dynamicized semantic map

- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns
  - **TIME?**



Ancient Greek

hṓra 'season/time/moment'

Approx. **8th c. BC**

(1) *hóssá* *te* *phúlla* *kaì* *ánthea*
REL.NOM.PL.N PTC leave:ACC.PL.N CONJ flower:ACC.PL.N

gígnetai ***hṓrēi***
become:PRS.3SG **season:DAT.SG.F**
'as are the leaves and the flowers in their **season**' (Homer, *Iliad* 2.468)

(2) *óphra* *Poseidáōni* *kaì* *állois* *athanátoisin*
CONJ Poseidon:DAT.SG.M CONJ other:DAT.PL immortal:DAT.PL

speísantes *koítoio* *medṓmetha:*
pour.libation:PART.AOR.NOM.PL.M bed:GEN.SG.M think.of:PRS.1PL.SUBJ.M/P

toîo *gàr* ***hṓrē***
dem.GEN.SG PTC **time:NOM.SG.F**
'that when we have poured libations to Poseidon and the other immortals, we may bethink us of sleep; for it is the **time** thereto' (Homer, *Odyssey* 3.333-334)

Ancient Greek

hṓra 'season/time/moment' ⇒ 'hour'

Approx. **5th c. BC**

(3) *anastàs* *dè* *prỳi* *pseustheìs*
raise.up:PTCP.AOR.NOM.SG.M PTC early deceive:PTCP.AOR.PASS.NOM.SG.M

tês ***hṓras*** *badízein*
ART.GEN.SG.F **time:GEN.SG.F** walk:PRS.INF

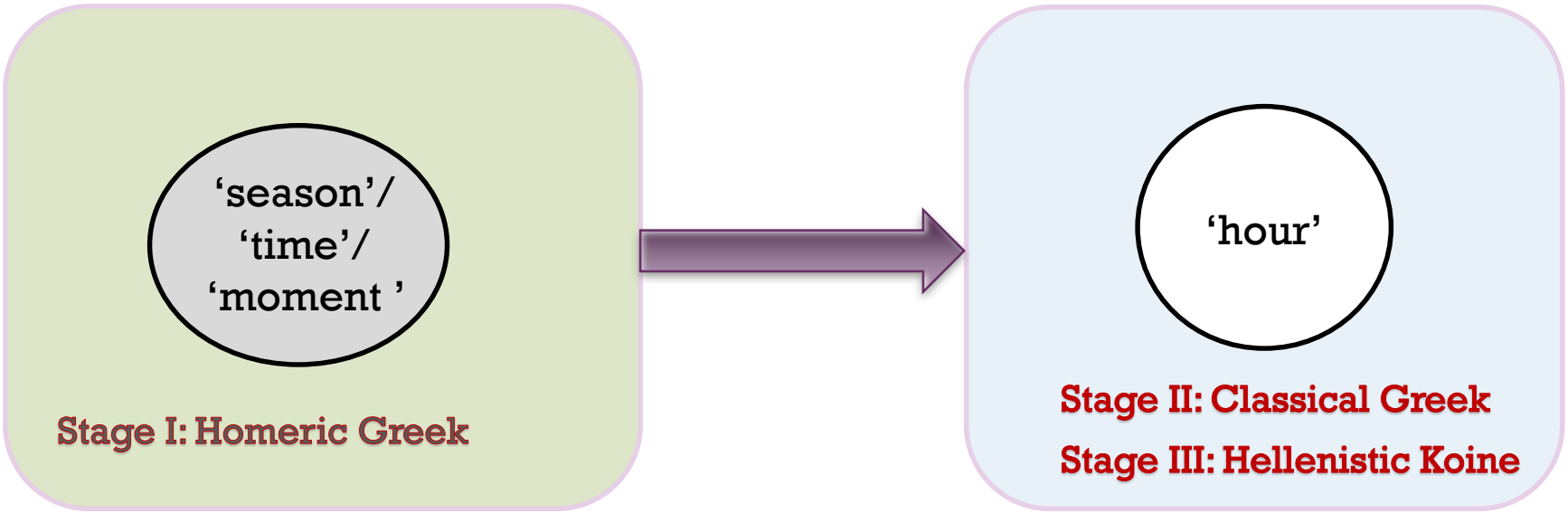
'He arose early, mistaking the **time/hour**, and started off on his walk' (Andocides, *On the Mysteries* 1.38)

Approx. **1st c. AD**

(4) *oukhì* ***dṓdeka*** ***hôrai*** *eisin* *tês* *hēméras;*
NEG **twelve** **hour:NOM.PL.F** be.PRS.3PL ART.GEN.SG.F day:GEN.SG.F

'Aren't there twelve hours of daylight?' (New Testament, John 11.9.2)

Ancient Greek

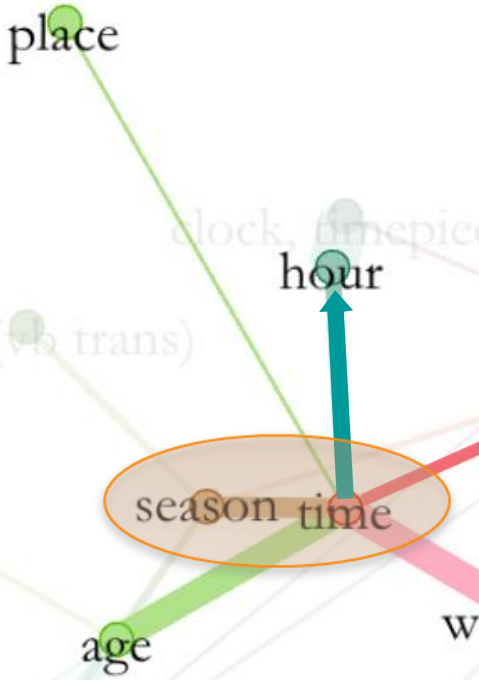


Metonymy: due to the correlation between the canonical time periods and the time these take to unfold

The semantic extension of time-related lexemes

Towards a dynamicized semantic map

- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns
  - **TIME** ⇒ HOUR



The semantic extension of time-related lexemes

Towards a dynamicized semantic map

- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns
- The material allows us to to add new polysemy patterns, and to provide a diachronic account
  - SUMMER?

(Ancient) culture-specific colexification patterns

- **Summer?**

There are 17 links involving the concept "summer": ?

Concept	IDS-Key	Occurrences	Families	Languages	Network		Forms
year	14.73	233	10	16	COM	SUB	FORMS
age	14.12	257	2	3	COM	SUB	FORMS
bow	20.24	231	2	2	COM	SUB	FORMS
spring	14.75	174	2	3	COM	SUB	FORMS
autumn	14.77	167	1	1	COM	SUB	FORMS
cave	1.28	256	1	1	COM	SUB	FORMS
cousin	2.55	346	1	1	COM	SUB	FORMS
hang up	9.341	280	1	1	COM	SUB	FORMS
hot	15.85	303	1	1	COM	SUB	FORMS
put	12.12	306	1	1	COM	SUB	FORMS
rain (noun)	1.75	257	1	1	COM	SUB	FORMS
reach, arrive	10.55	329	1	1	COM	SUB	FORMS
rise	10.21	334	1	1	COM	SUB	FORMS
season	14.78	193	1	1	COM	SUB	FORMS
sun	1.52	245	1	1	COM	SUB	FORMS
wall	7.27	239	1	1	COM	SUB	FORMS
wine	5.92	162	1	1	COM	SUB	FORMS

Ancient Greek

théros 'summer' ⇒ 'harvest'

(5) *autàr* *epèn* *élthēisi* ***théros*** *tethaluîá*
PTC when come:AOR.SUBJ.3SG **summer:NOM.SG.M** thrive:PART.PERF.NOM.SG.F
t' *opṓrē*
PTC autumn:NOM.SG.F
'But when **summer** comes and rich autumn' (Homer, *Odyssey* 11.192)

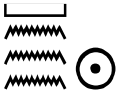
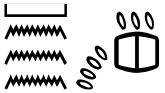
Approx. **8th C. BC**

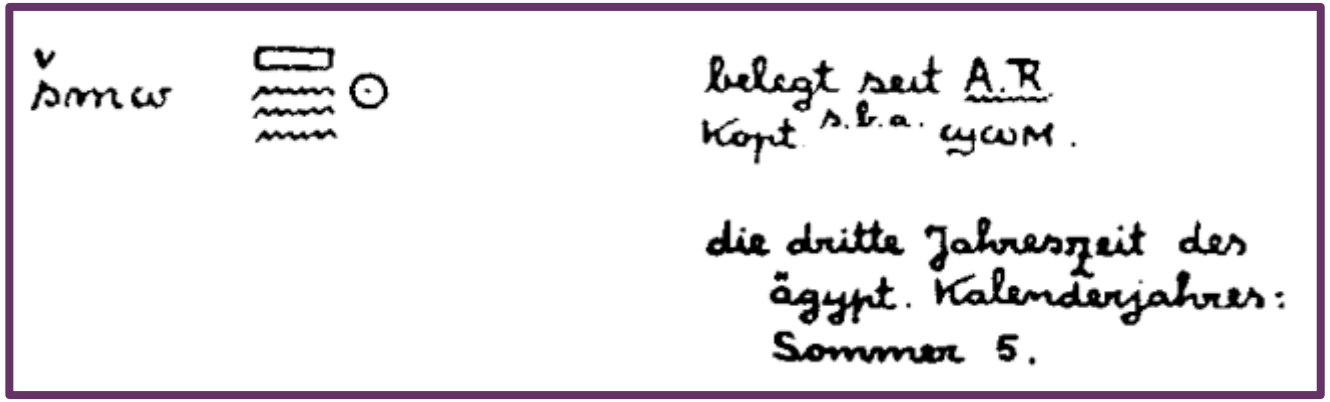
(6) *kâit'* *anèr* *édoksen* *eînai,* *tallótrion*
ADV man:NOM.SG.M seem:AOR.3SG be.INF another:GEN.SG

amôn ***théros***
reap.corn:PTCP.PRS.NOM.SG.M **summer:ACC.SG.N**
'he has only made himself a name by reaping another's **harvest**' (Aristophanes, *Knights* 392)

Approx. **5th C. BC**

Ancient Egyptian

 Smw ‘summer’ ⇒  Smw ‘harvest’



Old Kingdom

Middle Kingdom

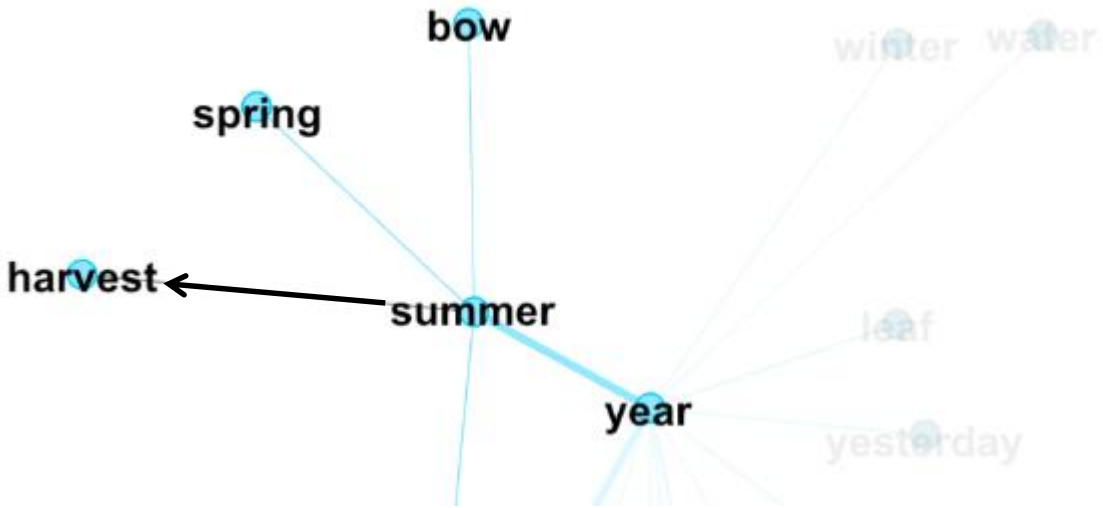
The semantic extension of time-related lexemes

Towards a dynamicized semantic map

- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns

- The material allows us to to add new polysemy patterns, and to provide a diachronic account
  - SUMMER

HARVEST

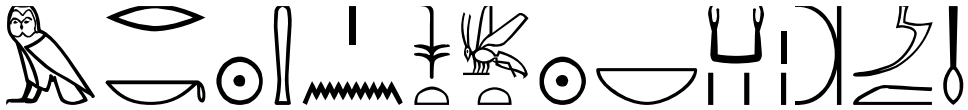


The semantic extension of time-related lexemes

Towards a dynamicized semantic map

- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns
- The material allows us to to add new polysemy patterns, and to provide a diachronic account
- The material allows us to highligh unexpected pathways of change: from temporal proximity to spatial proximity

Language-specific colexification patterns



Peasant, B1, 103-104

(7)

m	**rk**	Hm-f	nswt-bity	nb-kAw-ra
in	**time**	Majesty-3SG.M	King of U. and L. Egypt	Nebkaure

'(Now, the peasant spoke these word) **during the time** of his Majesty, the King of Upper and Lower Egypt, Nebkaure (the justified)' (= Parkinson 1991: 19)

Approx. **1400 BC**

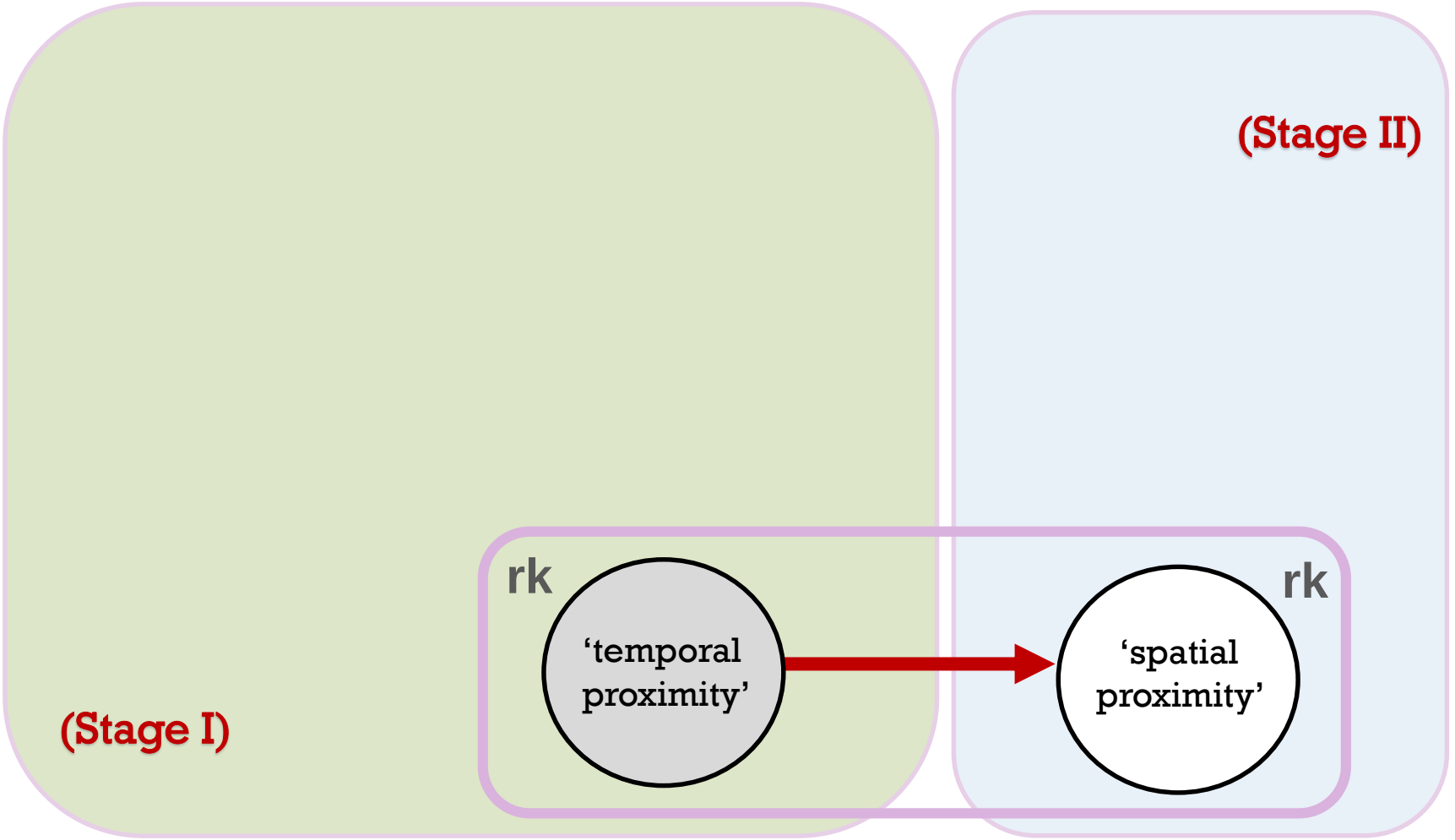
(8)

sbty	Dr	**m**	**rk**	mSa-f	(= K*RI* II, 6,8)
rampart	strong	**in**	**proximity**	army-3SG.M	

(speaking of the King who is)
'A strong rampart around his army, (their shied in the day of fighting)'

Approx. **1250 BC**

Language-specific colexification patterns

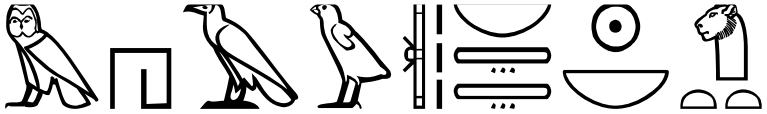


Language-specific colexification patterns

- Counterexample to the TIME IS SPACE metaphor?
  - Cross-linguistically Time to Space transfers are extremely rare

(cf. French *depuis*; Haspelmath 1997)

Language-specific colexification patterns



Biography of Ahmose, 5

(9) **m hAw** nb tA-wj nb-pH.tj-ra

in prox-time lord land-DU Nebphtire

(And then I became a soldier (…),)
'**during the time** of the lord of the Two Lands, Nebpehtire (justified, when I was a young man, not having a wife yet)' (= *Urk*. IV, 2,13)

Approx. **1350 BCE**

Sinuhe, B8

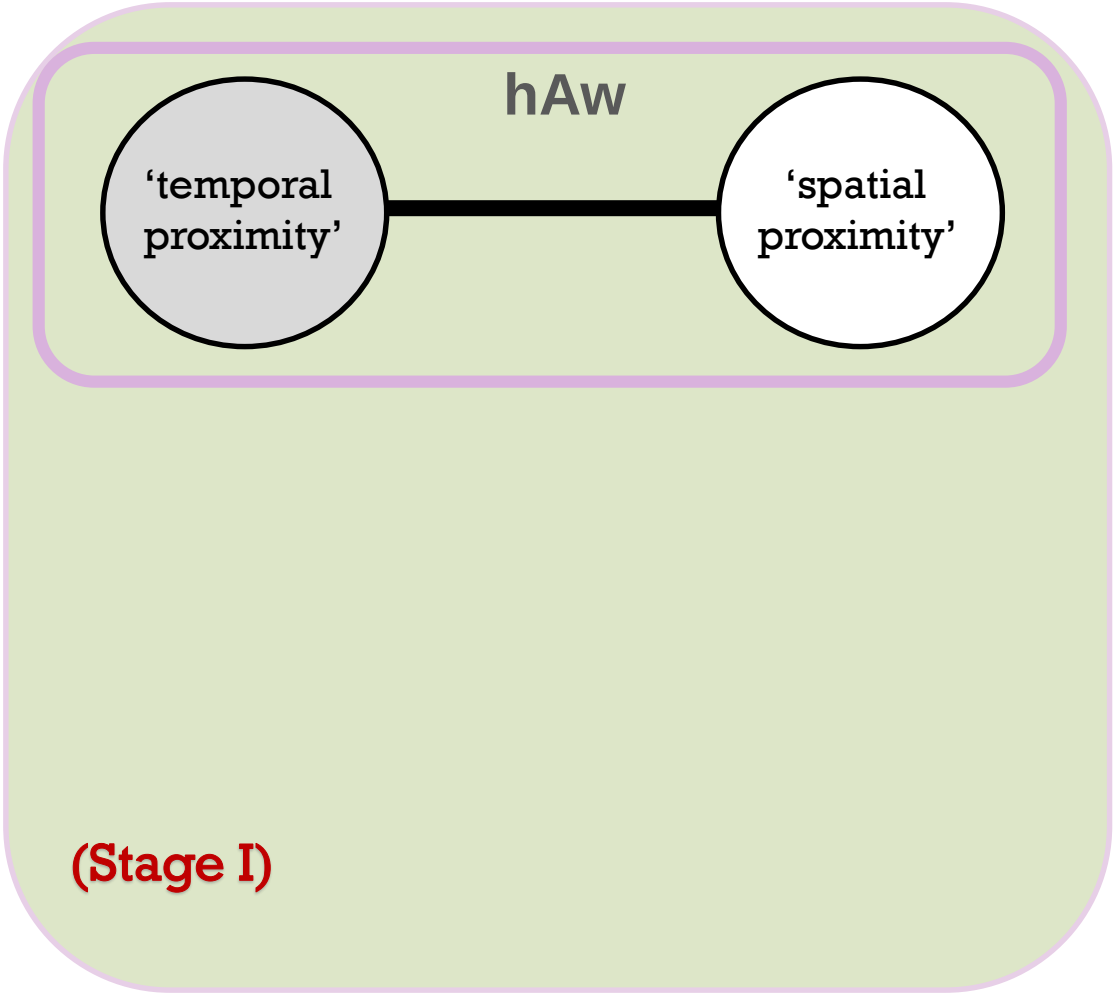
(10) **m hAw** nh.t

in prox-space Sycamore

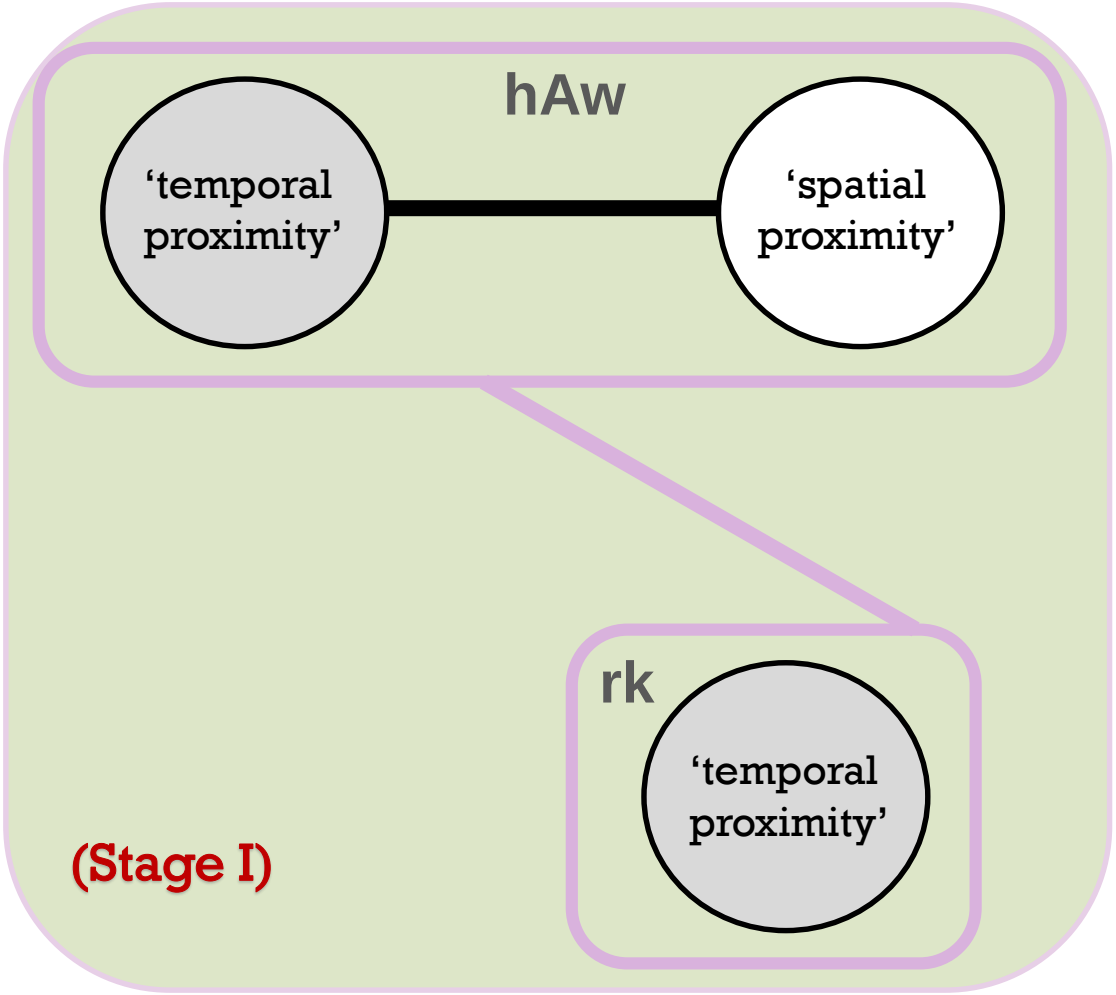
'(I crossed the place called The Two Truths,) **in the vicinity** of The Sycamore" (and I landed at The Island of Snefru)' (= Koch 1990: 14)

Approx. **1500 BCE**

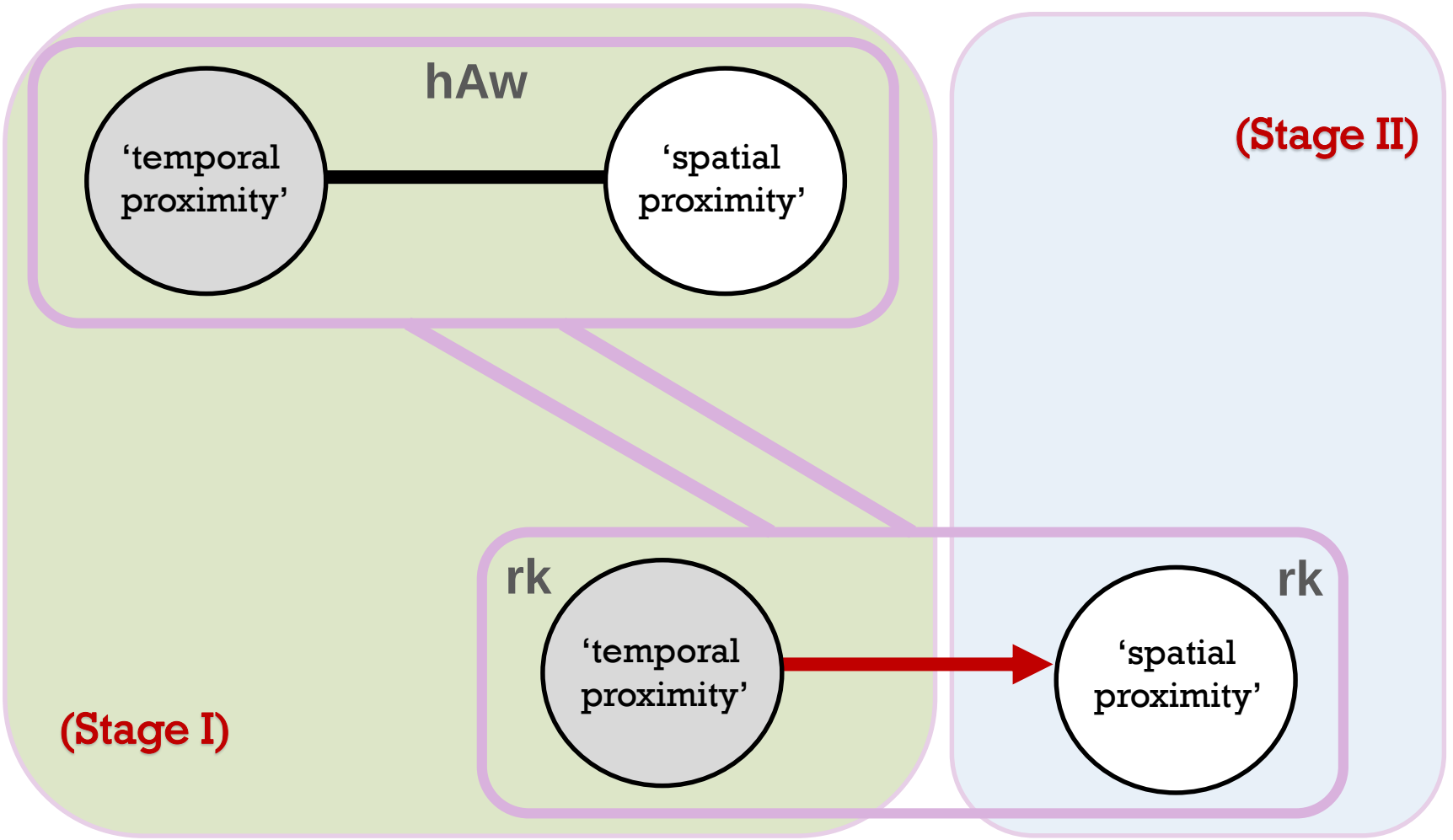
Language-specific colexification patterns



Language-specific colexification patterns



Language-specific colexification patterns



+

Conclusions

The web-based platform

Conclusions

- In this talk, we have shown that weighted diachronic semantic maps can be automatically plotted based on lexical matrix of crosslinguistic colexification patterns
- Furthermore, we demonstrated that language-specific studies reveal interesting polysemy patterns, with unexpected pathways of change

Conclusions

- In this talk, we have shown that weighted diachronic semantic maps can be automatically plotted based on lexical matrix of crosslinguistic colexification patterns
- Furthermore, we demonstrated that language-specific studies reveal interesting polysemy patterns, with unexpected pathways of change
- One can expect the linguists to build solid polysemy matrices
- BUT, one cannot expect them to master the technicalities invloved for plotting semantic maps automatically and exploring the result

Conclusions

- In this talk, we have shown that weighted diachronic semantic maps can be automatically plotted based on lexical matrix of crosslinguistic colexification patterns
- Furthermore, we demonstrated that language-specific studies reveal interesting polysemy patterns, with unexpected pathways of change
- One can expect the linguists to build solid polysemy matrices
- BUT, one cannot expect them to master the technicalities invloved for plotting semantic maps automatically and exploring the result

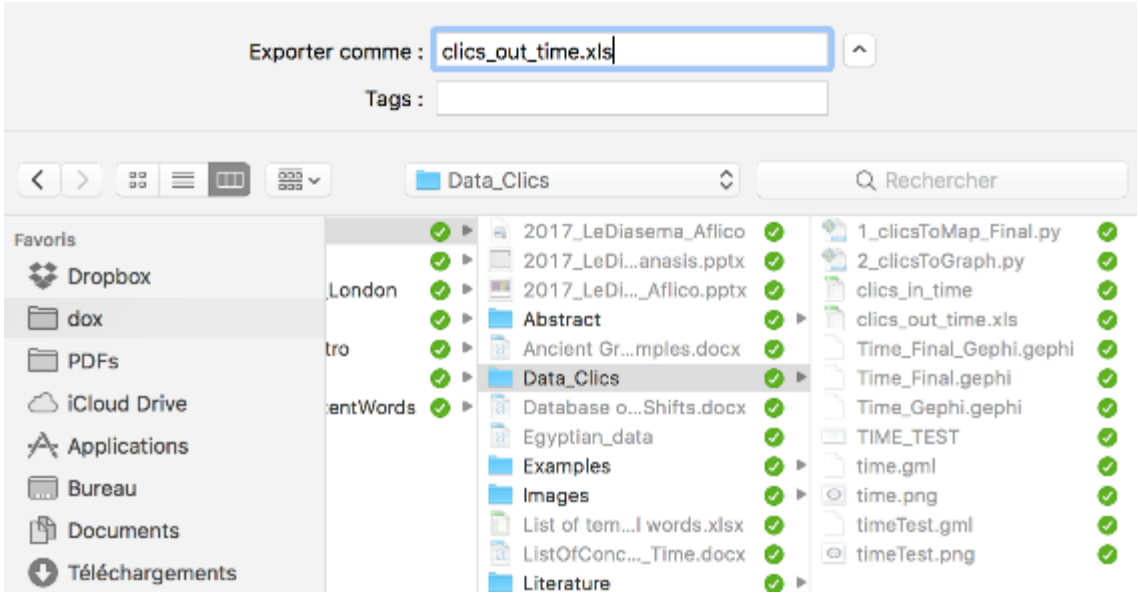
A web-based platform can easily do the work

How to plot semantic maps?

Upload the lexical matrix (.xls file)

How to plot semantic maps?

Upload the lexical matrix (.xls file)



How to plot semantic maps?

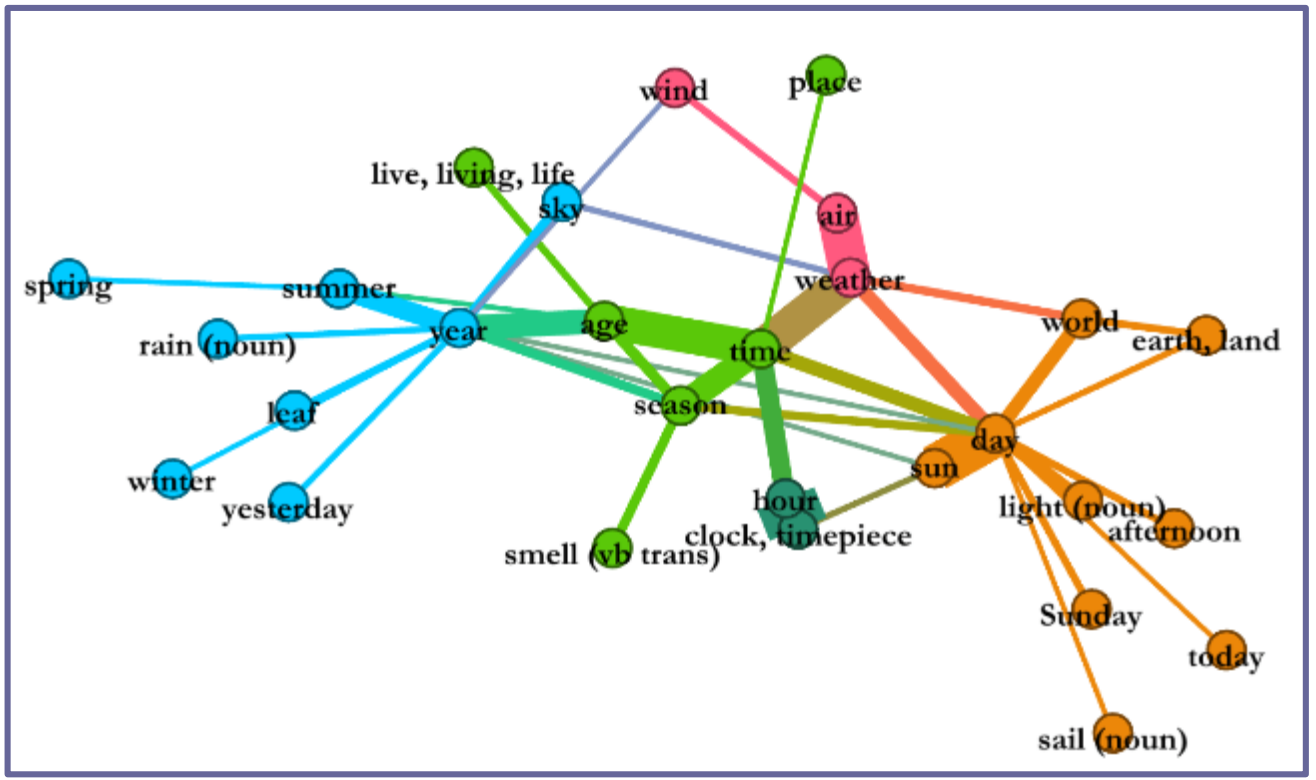
Lexical matrix uploaded ✓

Generate the map

☑ Weighted

☐ Diachronic

How to plot semantic maps?



Selected references

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.

van der Auwera, J. (2008). In defense of classical semantic maps. *Theoretical Linguistics*, *34*(1), 39–46.

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

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.

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). Amsterdam/Philadelphia: John Benjamins.

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. 211–242). Mahwah/ New Jersey: Lawrence Erlbaum Associates.

List, J.-M., Mayer, Th., Terhalle, A., & Urban, M. (2014). *CLICS: Database of Cross-Linguistic Colexifications*. Marburg: Forschungs-zentrum Deutscher Sprachatlas (Version 1.0, online available at http://CLICS.lingpy.org, accessed on 2016-27-10).

Narrog, H., & Ito, S. (2007). Re-constructing semantic maps: the comitative-instrumental area. *Language Typology and Universals, 60*(4), 273–292.

Regier, T., Khetarpal, N., & Majid, A. (2013). Inferring semantic maps*, in: Linguistic Typology*, *17*, 89–105.

Swadesh, M. (1952). Lexicostatistic Dating of Prehistoric Ethnic Contacts. *Proceedings of the American Philosophical Society*, *96*, 452–463.

Wälchli, B., & Cysouw, M. (2012). Lexical typology through similarity semantics: Toward a semantic map of motion verbs. *Linguistics, 50*(3), 671–71.

Thanks!

athanasios.georgakopoulos@ulg.ac.be

s.polis@ulg.ac.be

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