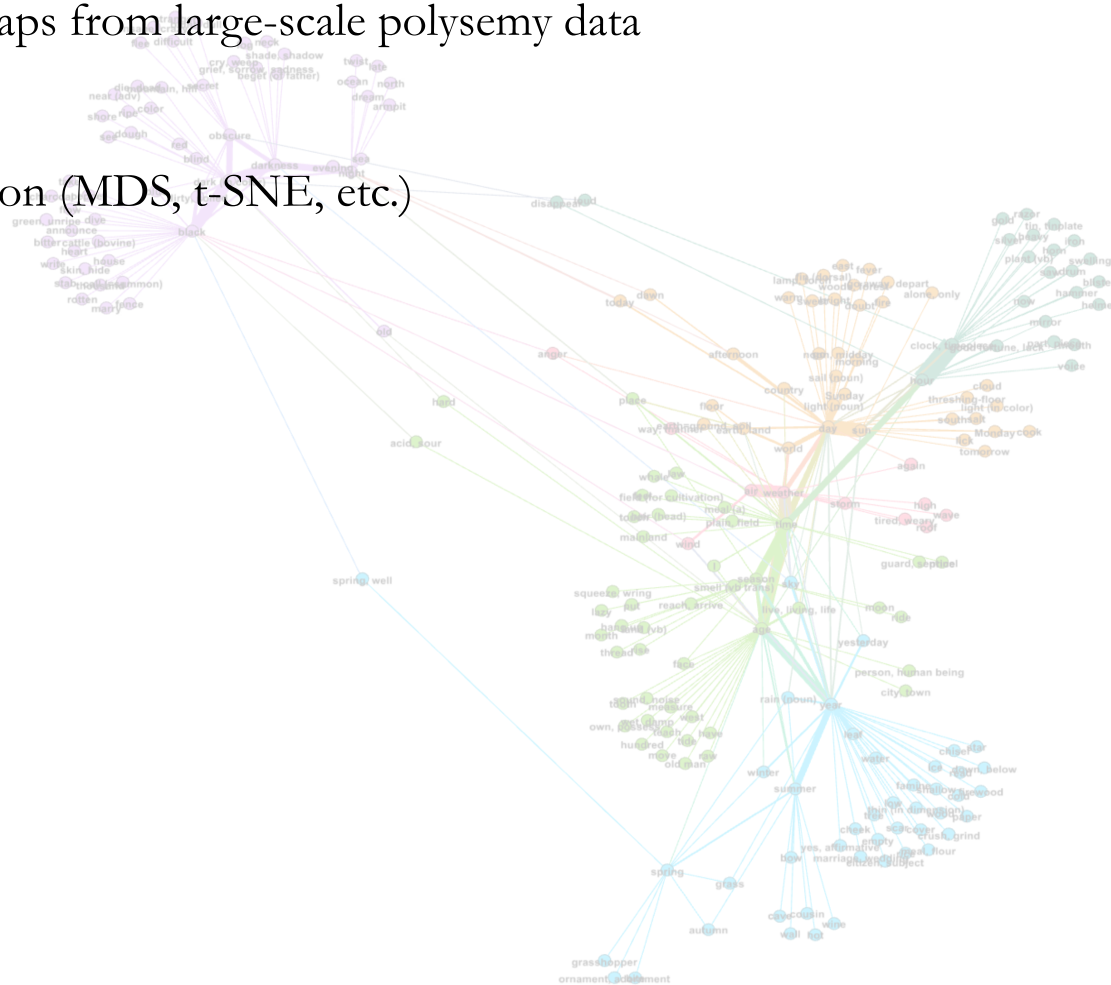
# LExical DIAchronic SEmantic MAps (Le Diasema)
## From simple networks to mixed multi-edge graphs

Thanasis Georgakopoulos & Stéphane Polis
(ULiège / F.R.S.-FNRS)

28.06.2018
University of Liège

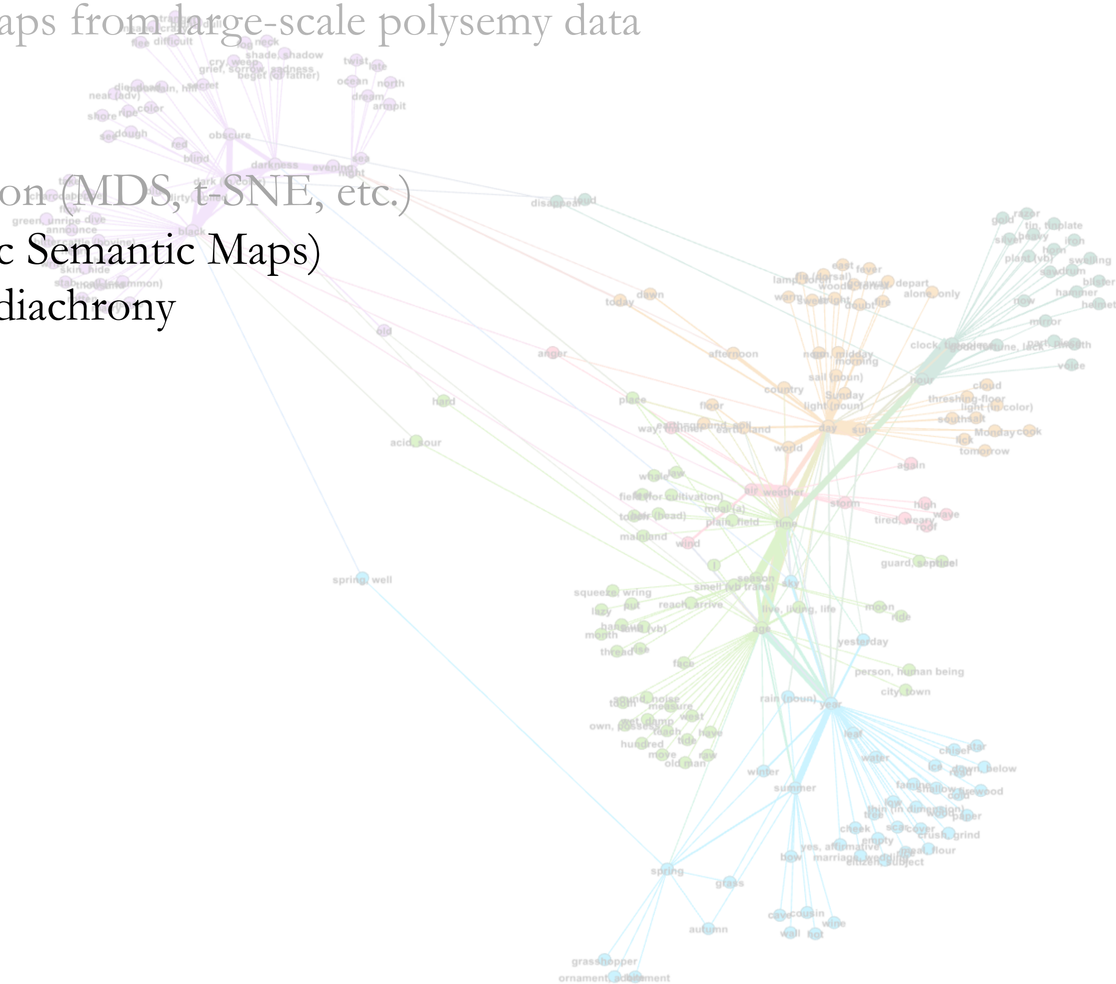
# Outline of the talk

- Inferring (classical) semantic maps from large-scale polysemy data
  - Basic principle
  - Plotting weighted maps
  - Graphs vs feature projection (MDS, t-SNE, etc.)



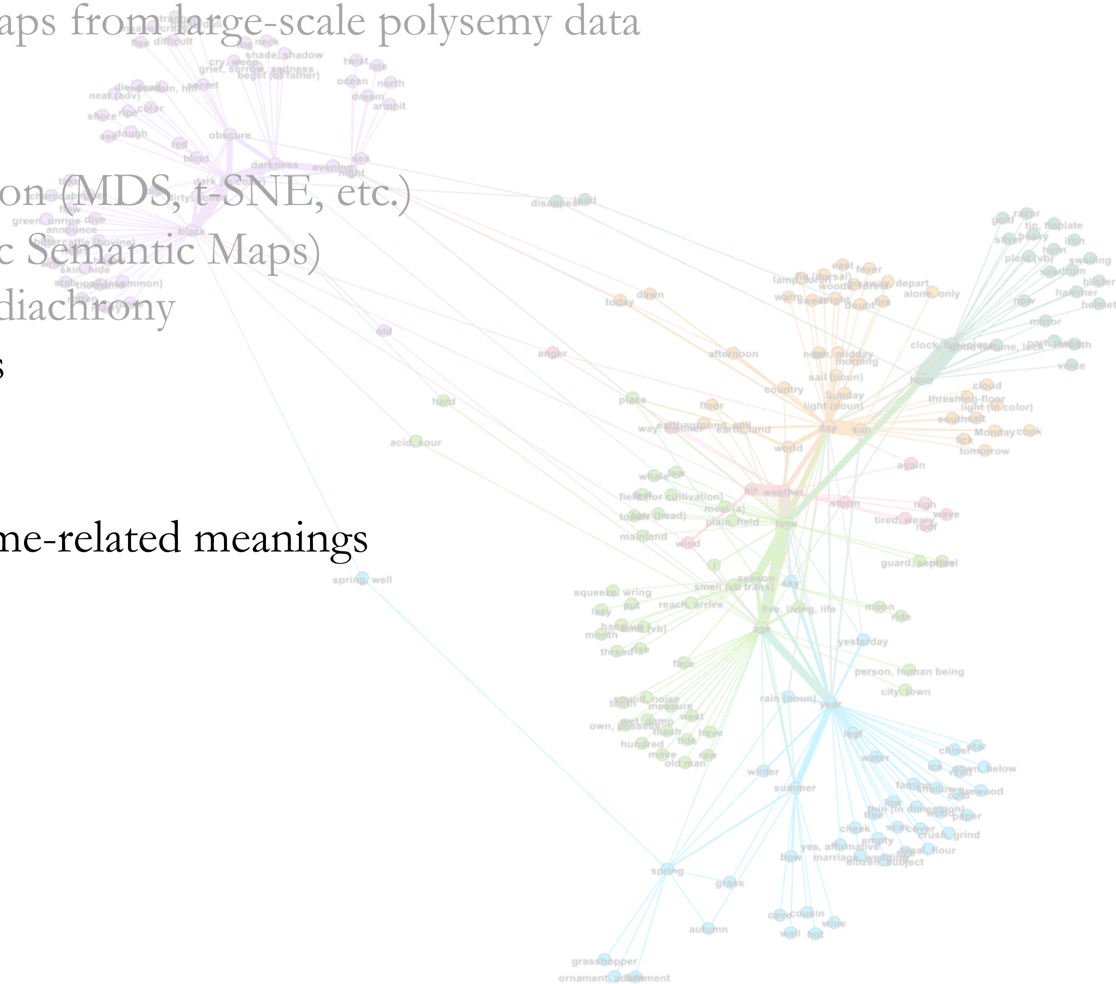
# Outline of the talk

- Inferring (classical) semantic maps from large-scale polysemy data
  - Basic principle
  - Plotting weighted maps
  - Graphs vs feature projection (MDS, t-SNE, etc.)
- Le Diasema (Lexical Diachronic Semantic Maps)
  - Focus on the lexicon and diachrony



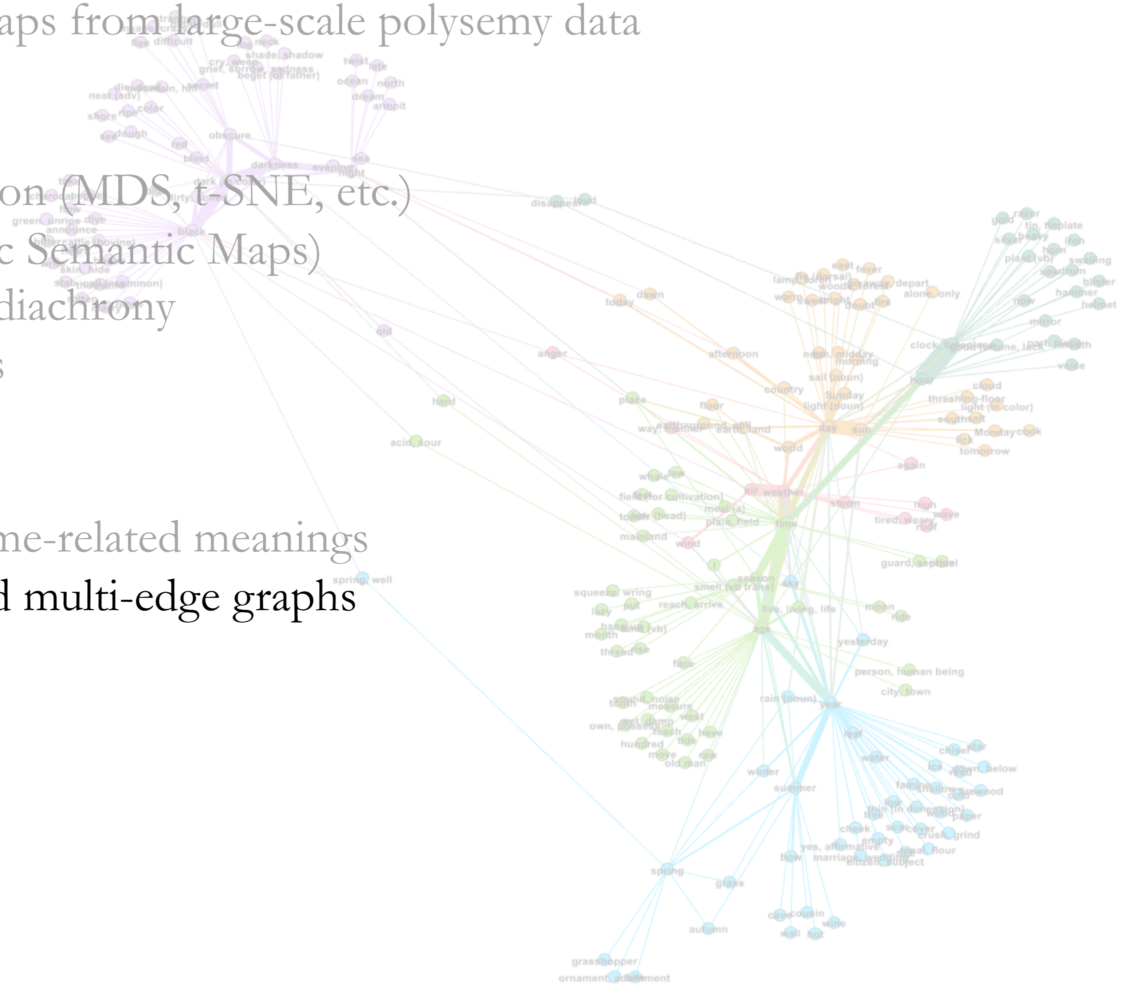
# Outline of the talk

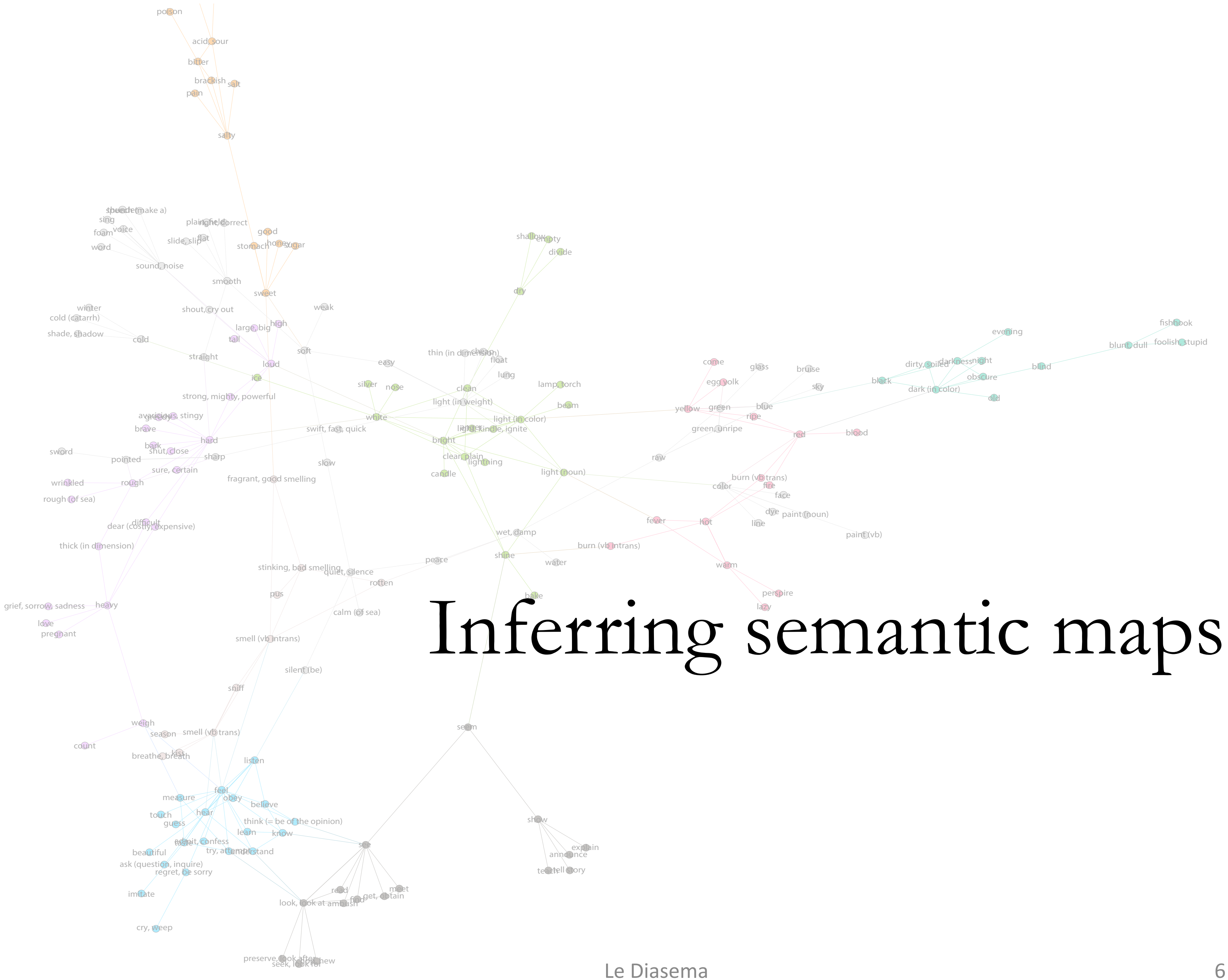
- Inferring (classical) semantic maps from large-scale polysemy data
  - Basic principle
  - Plotting weighted maps
  - Graphs vs feature projection (MDS, t-SNE, etc.)
- Le Diasema (Lexical Diachronic Semantic Maps)
  - Focus on the lexicon and diachrony
- A protocol for diachronic maps
  - Plotting,
  - dynamicizing,
  - and enriching a map of time-related meanings



# Outline of the talk

- Inferring (classical) semantic maps from large-scale polysemy data
  - Basic principle
  - Plotting weighted maps
  - Graphs vs feature projection (MDS, t-SNE, etc.)
- Le Diasema (Lexical Diachronic Semantic Maps)
  - Focus on the lexicon and diachrony
- A protocol for diachronic maps
  - Plotting,
  - dynamicizing,
  - and enriching a map of time-related meanings
- From simple networks to mixed multi-edge graphs






# Inferring semantic maps

# Inferring semantic maps

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

# Inferring semantic maps

Limitation of the semantic map method: practically, it is 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"

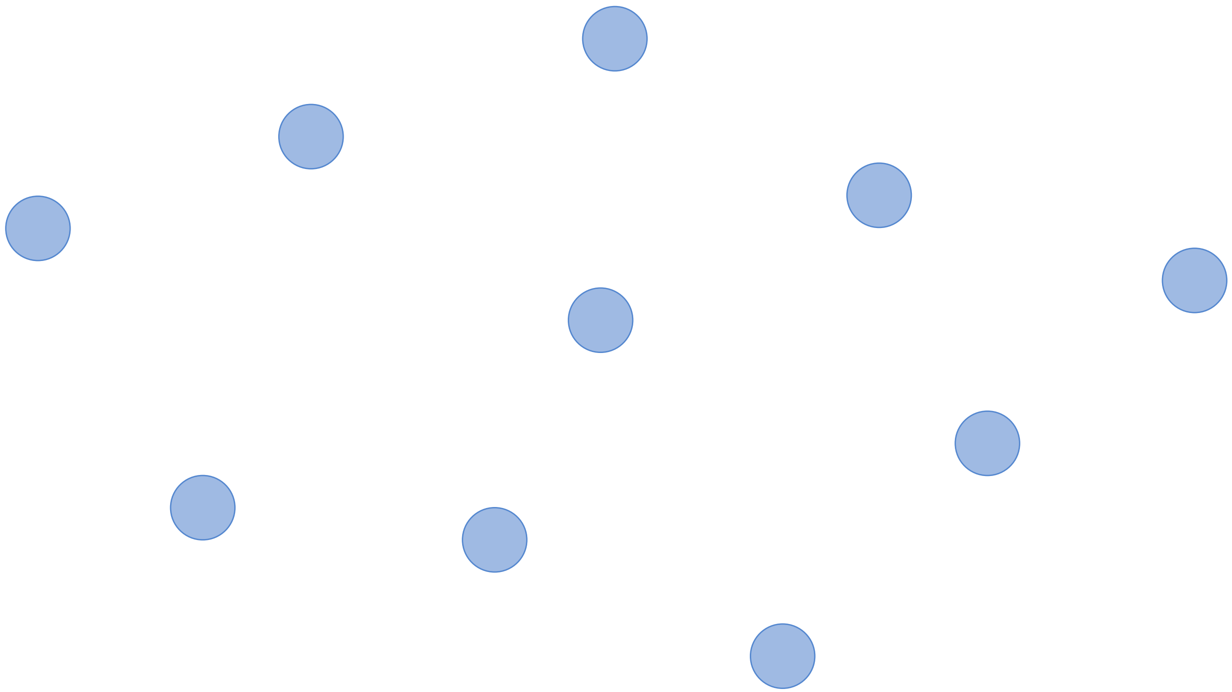
(Croft & Poole 2008: 1)

# 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)

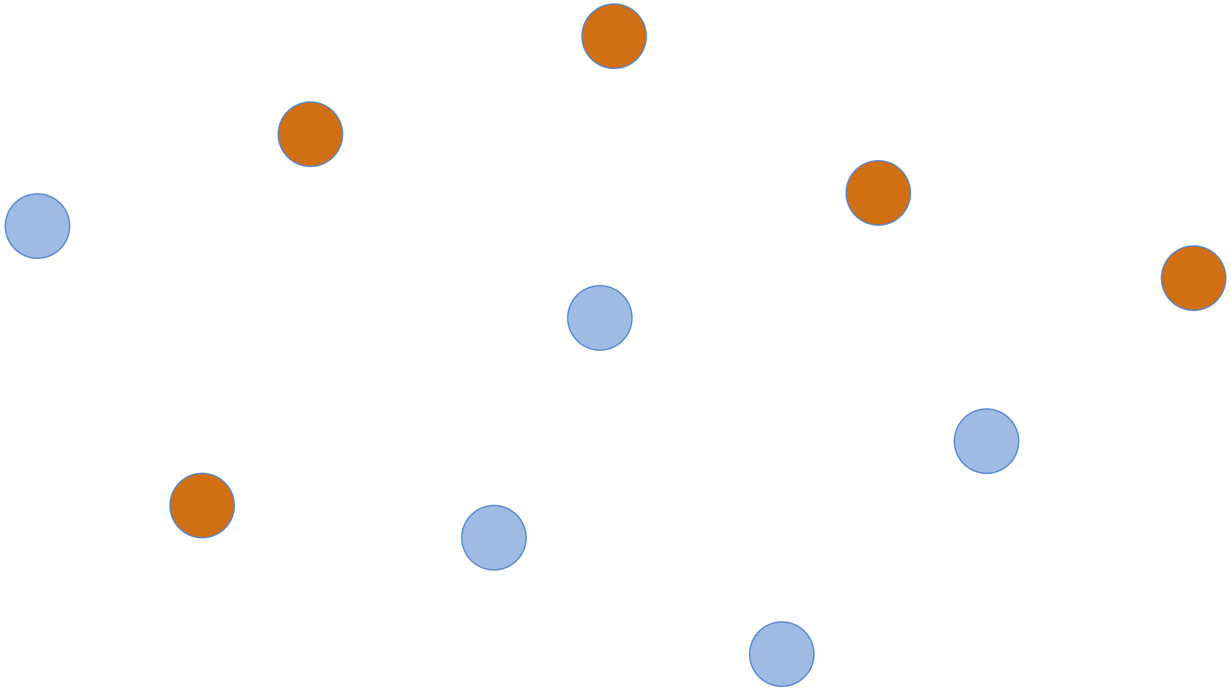
# Inferring semantic maps

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



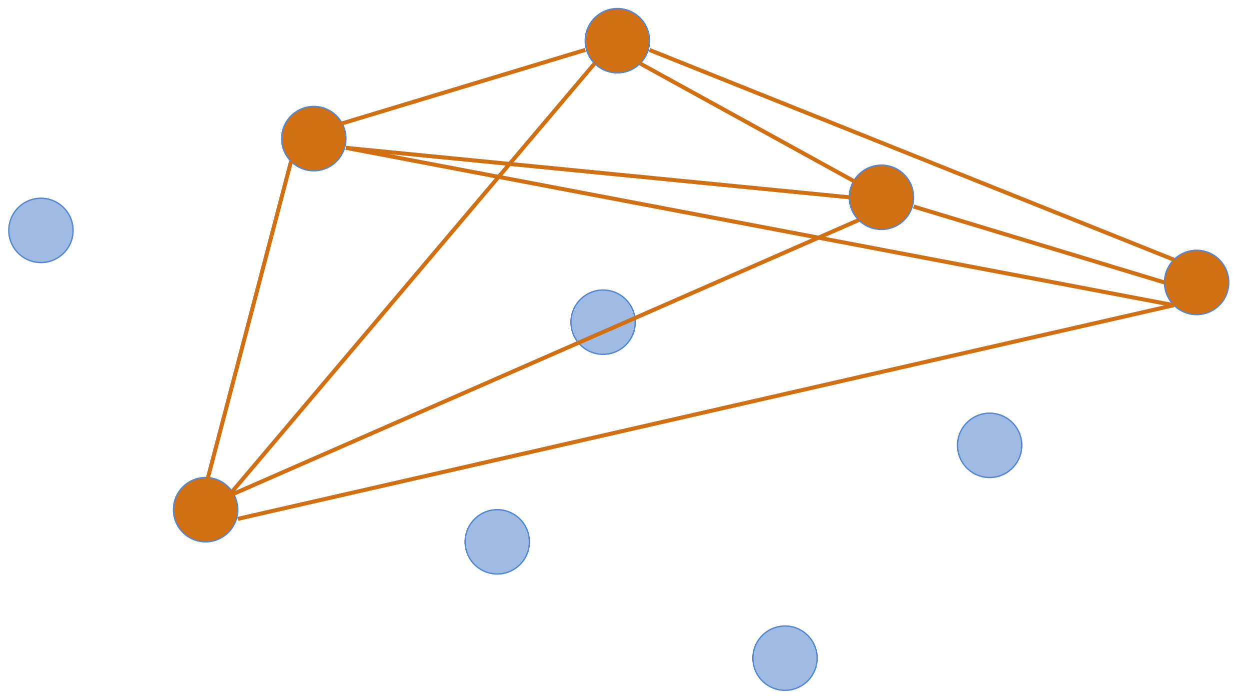
# Inferring semantic maps

- 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)



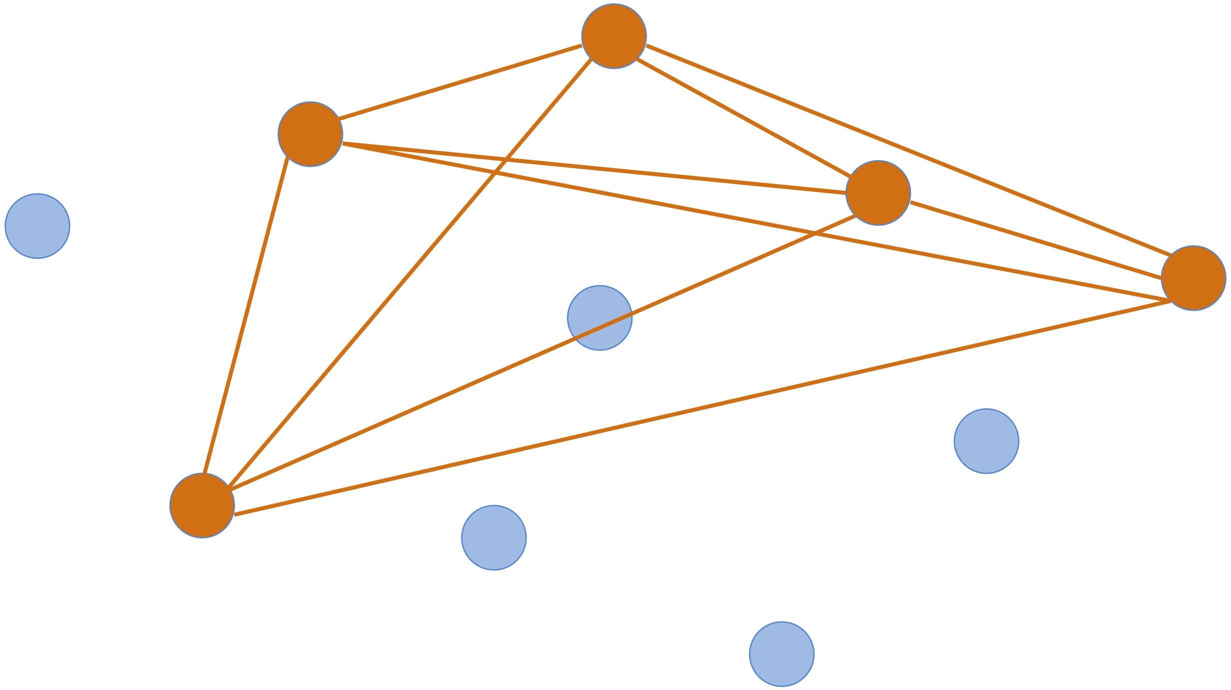
# Inferring semantic maps

- 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)



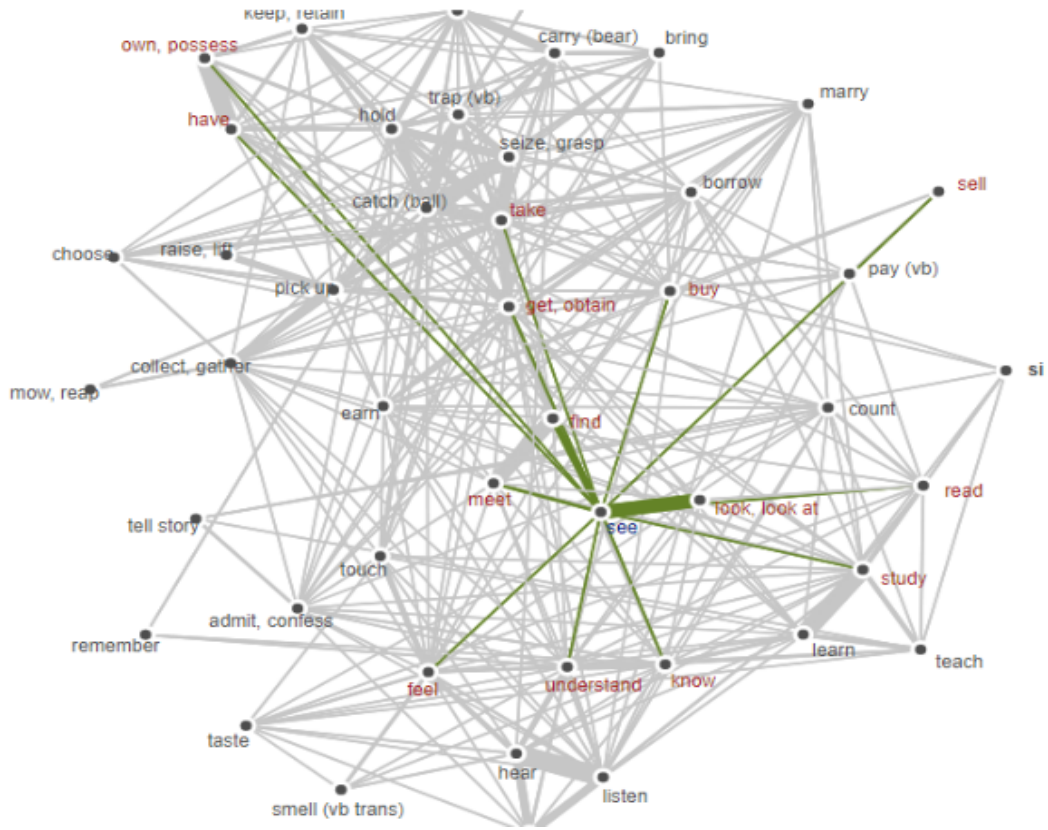
# Inferring semantic maps

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



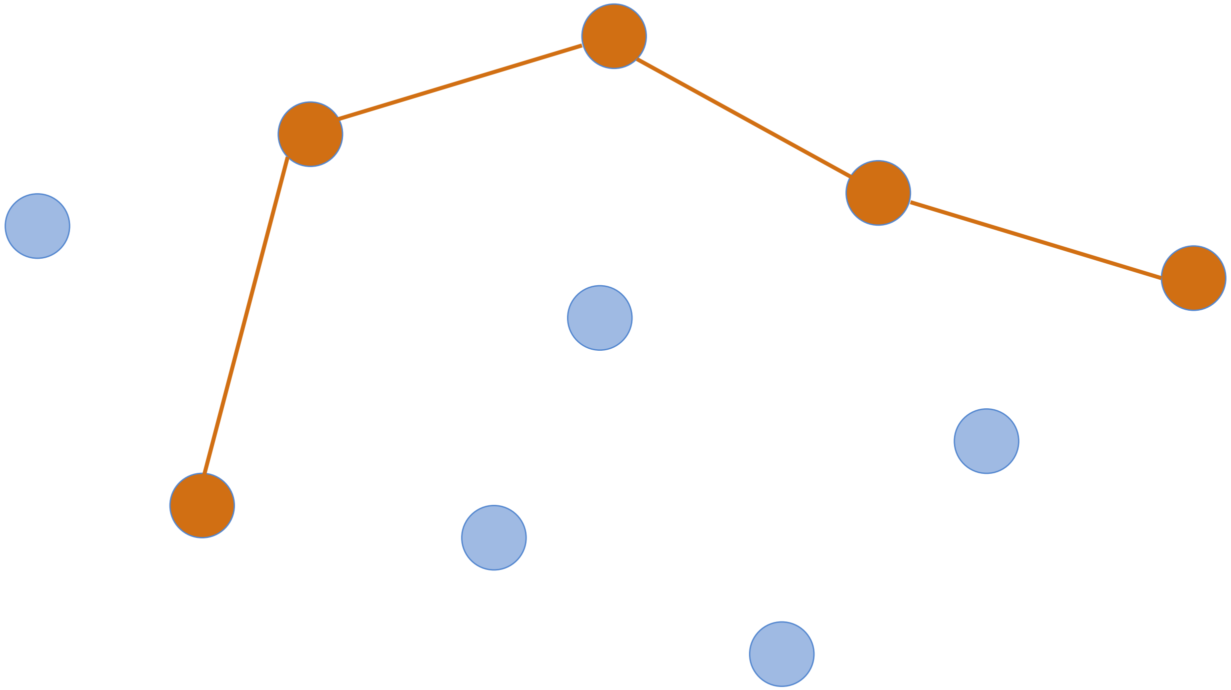
# Inferring semantic maps

- What's the idea?
  - But this is precisely what a colexification network does



# Inferring semantic maps

- What's the idea?
  - The goal would be to find a more economical solution and to have all the social agents connected with as few edges as possible accounting for all the observations

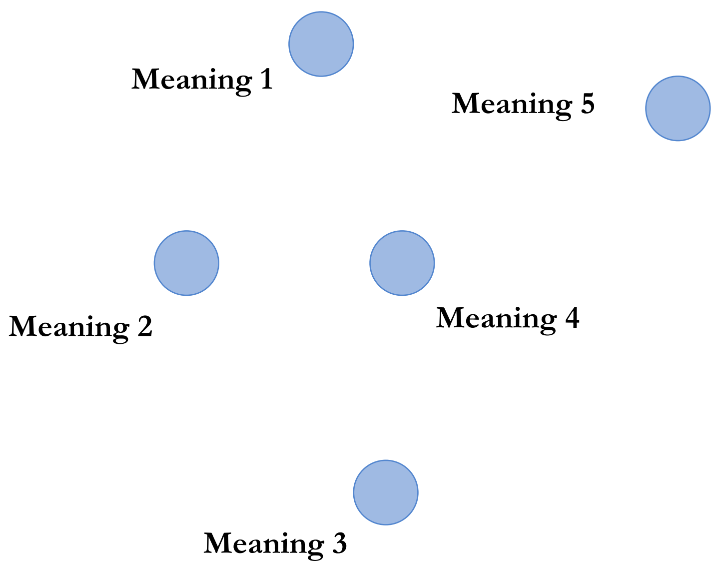


# Inferring semantic maps

- How does it transfer to semantic maps?

# Inferring semantic maps

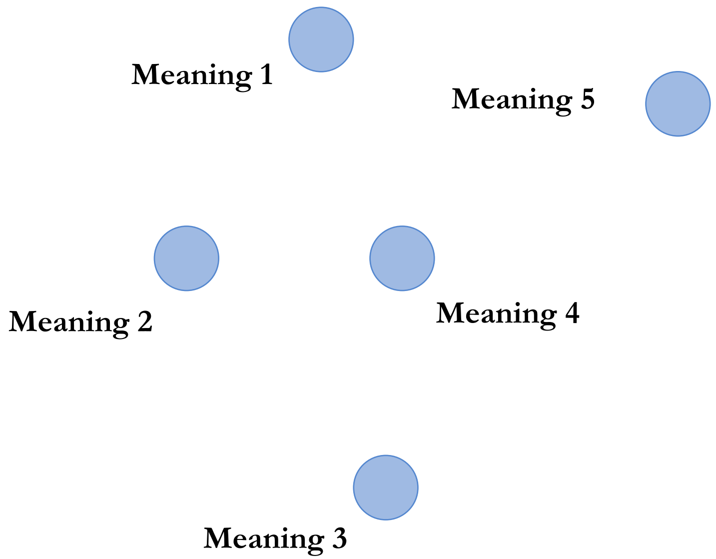
- How does it transfer to semantic maps?
  - Nodes are meanings




| Meaning | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|

# Inferring semantic maps

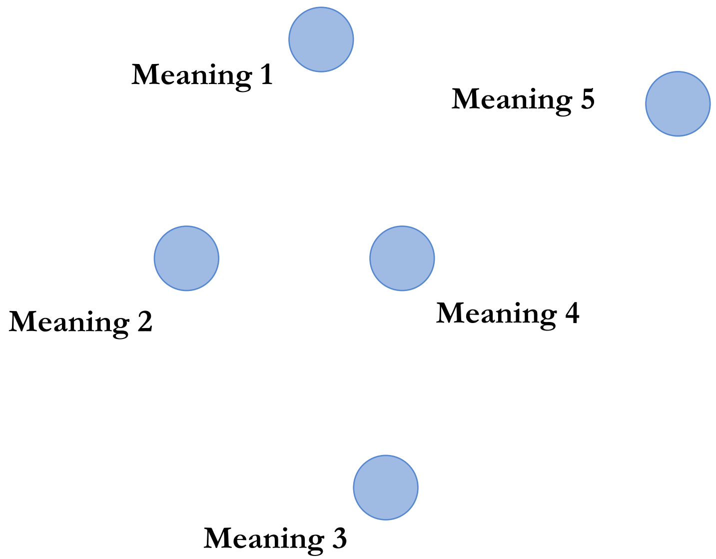
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are polysemic items (connectivity hypothesis)




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

# Inferring semantic maps

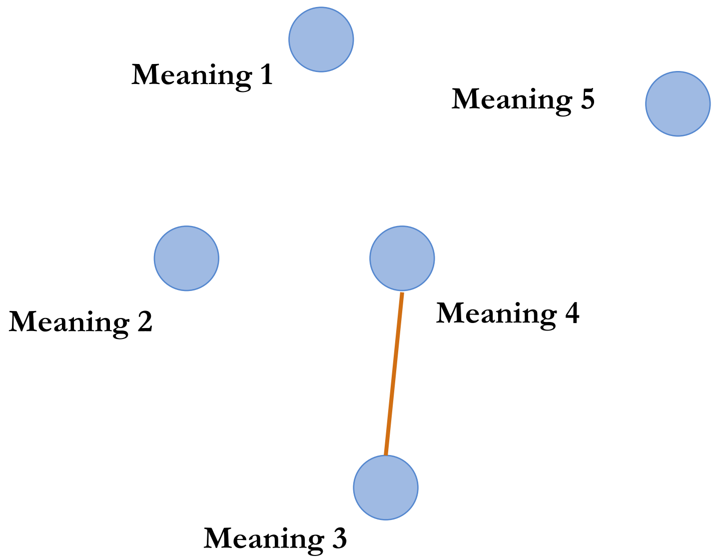
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are polysemic items (connectivity hypothesis)
  - One connects the nodes economically based on these constraints




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

# Inferring semantic maps

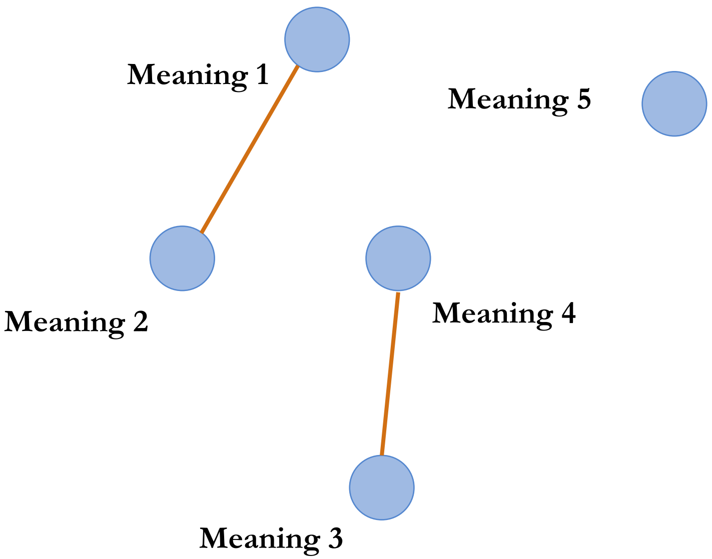
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are polysemic items (connectivity hypothesis)
  - One connects the nodes economically based on these constraints, starting with the edge(s) that accounts for the most frequent constraint(s)




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

# Inferring semantic maps

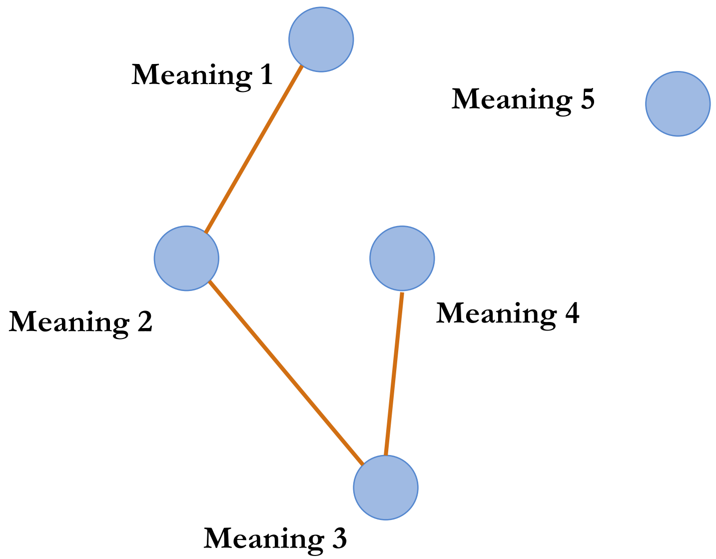
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are polysemic items (connectivity hypothesis)
  - One connects the nodes economically based on these constraints, starting with the edge(s) that accounts for the most frequent constraint(s), and then going down the scale




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

# Inferring semantic maps

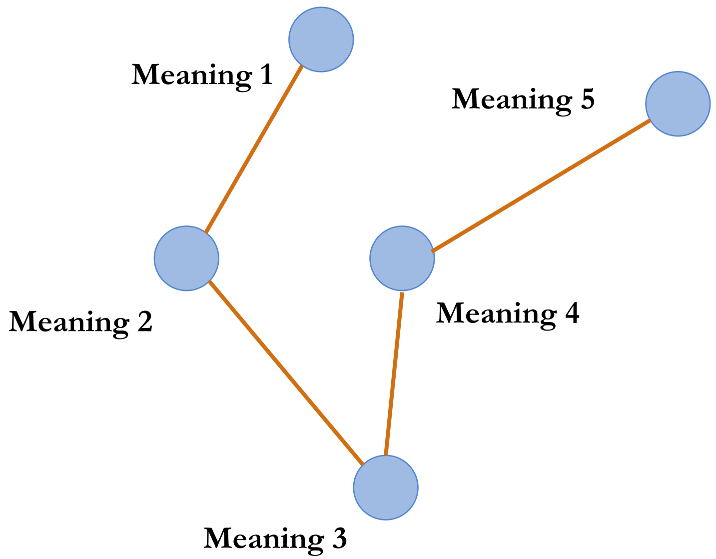
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are polysemic items (connectivity hypothesis)
  - One connects the nodes economically based on these constraints, starting with the edge(s) that accounts for the most frequent constraints(s), and then going down the scale




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

# Inferring semantic maps

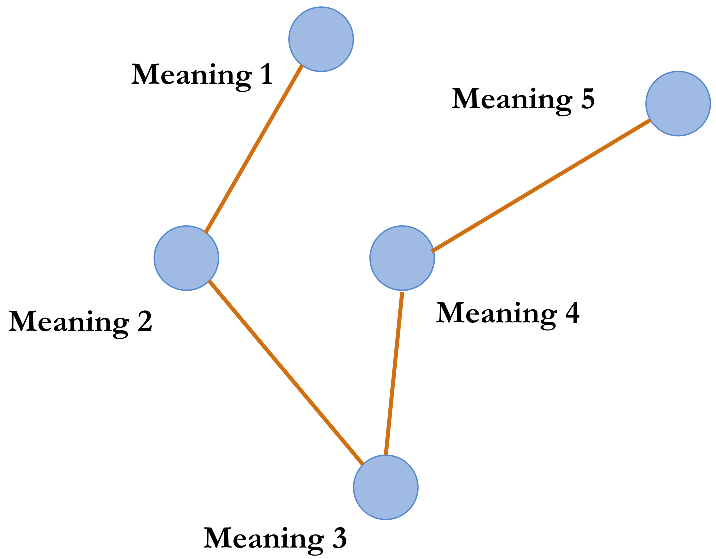
- How does it transfer to semantic maps?
  - Nodes are meanings
  - Constraints are polysemic items (connectivity hypothesis)
  - One connects the nodes economically based on these constraints, starting with the edge(s) that accounts for the most frequent constraint(s), and then going down the scale




| 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 (connectivity hypothesis)
  - One connects the nodes economically based on these constraints, starting with the edge(s) that accounts for the most frequent constraint(s), and then going down the scale




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

# 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)

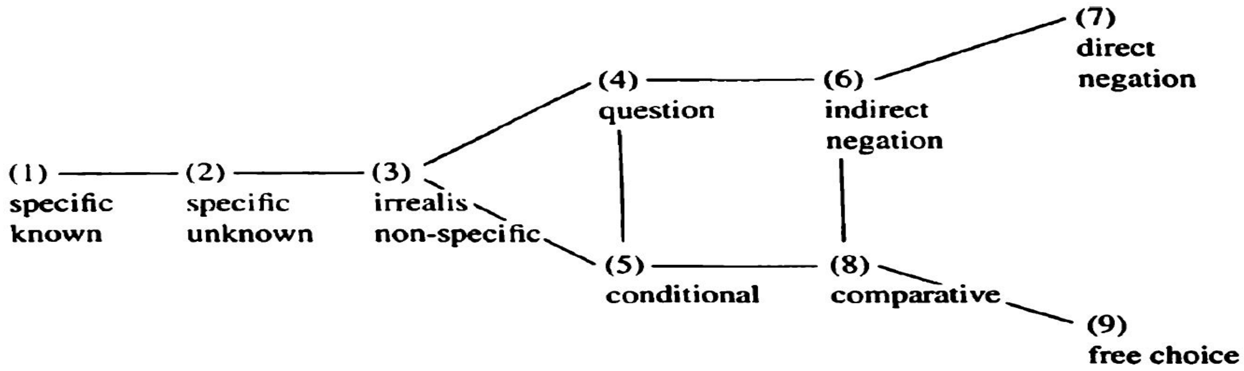



Figure. Haspelmath's (1997: 4) original semantic map of the indefinite pronouns functions

# 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 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

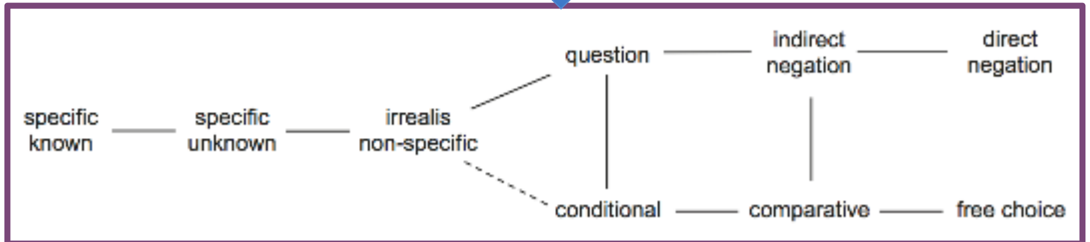
INPUT
(lexical matrix)

| *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)



# Plotting weighted maps

## 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)
```

# Plotting weighted maps

## Weighted semantic maps

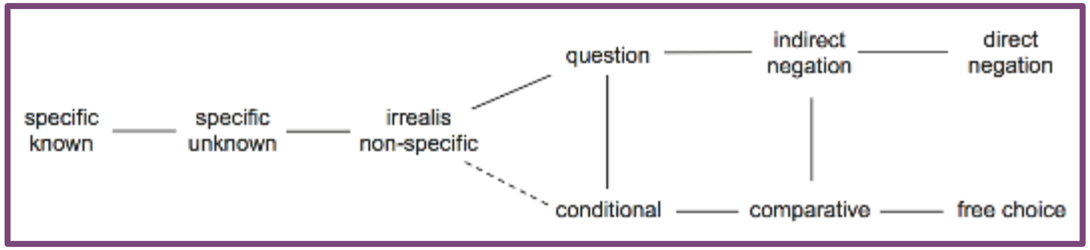
- Generate the map with a modified version of the algorithm of Regier et al. (2013)
  - PRINCIPLE: for each edge that is being added between two meanings of the map by the algorithm, check in the lexical matrix how many times this specific polysemy pattern is attested, and increase the weight of the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

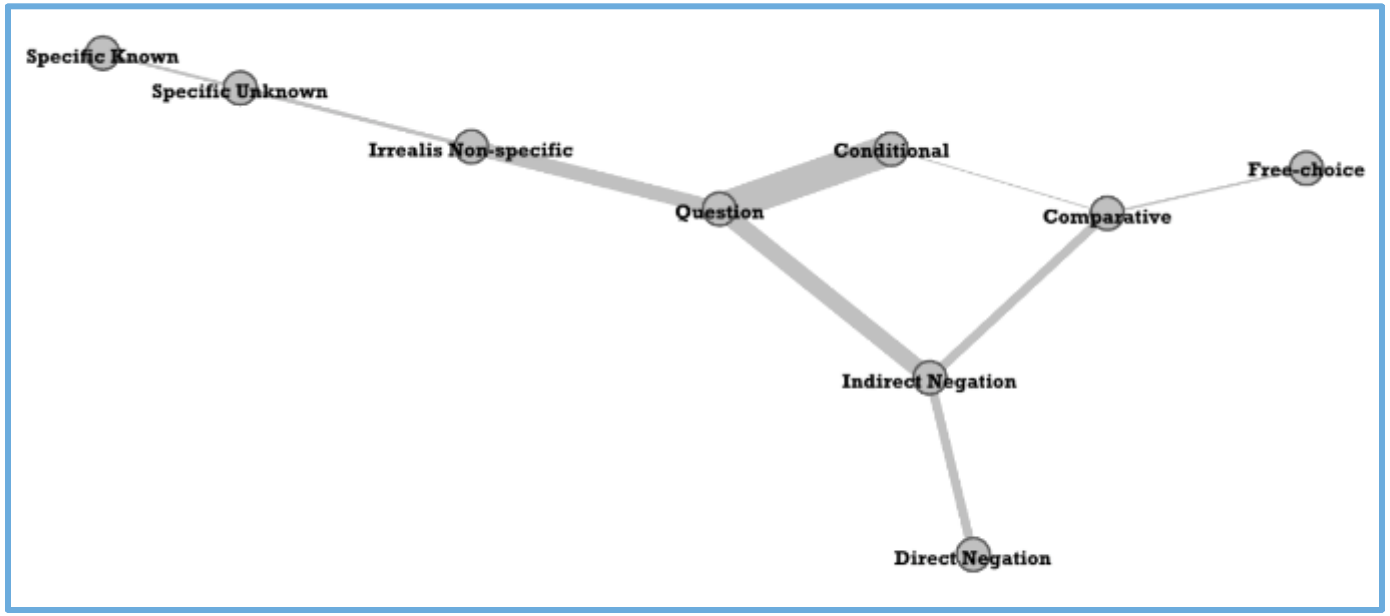
  - Based on the data of Haspelmath (1997), kindly provided by the author, the result between a non-weighted and a weighted semantic map are markedly different

# Plotting weighted maps

## Weighted semantic maps



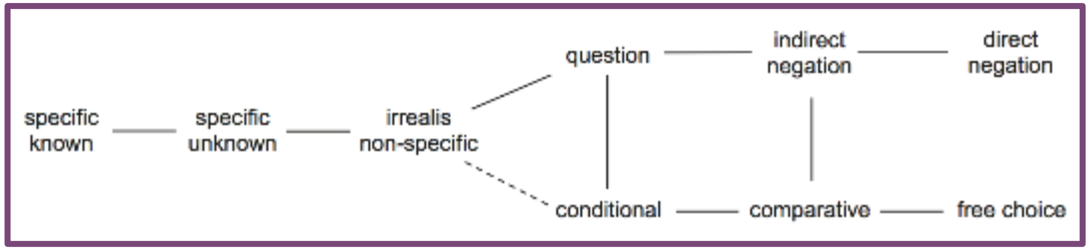

Automatically plotted semantic maps:
non-weighted vs. weighted
(data from Haspelmath 1997)



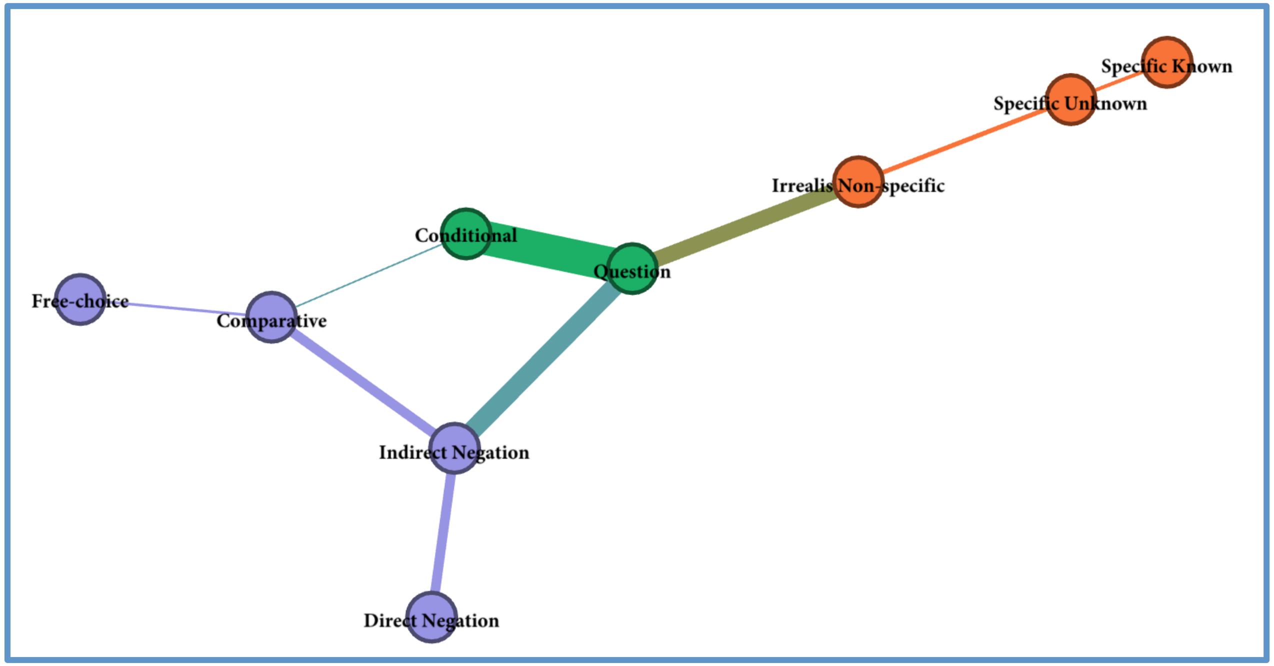

The graph is visualized in Gephi® with the *Force Atlas* algorithm

# Plotting weighted maps

## Weighted semantic maps




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

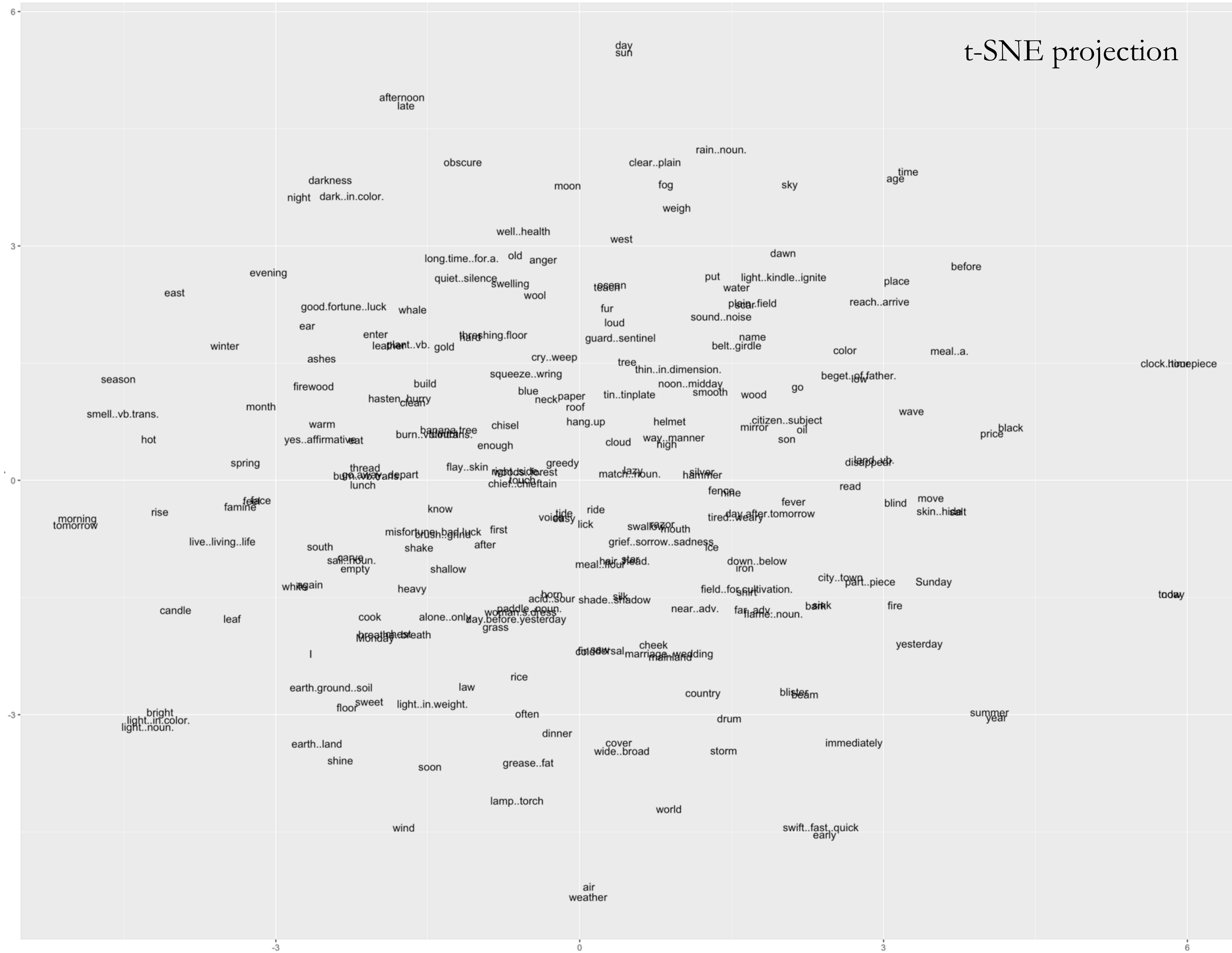



The graph is visualized in Gephi® with the *Force Atlas* algorithm and modularity analysis (Lambiotte et al. 2009)

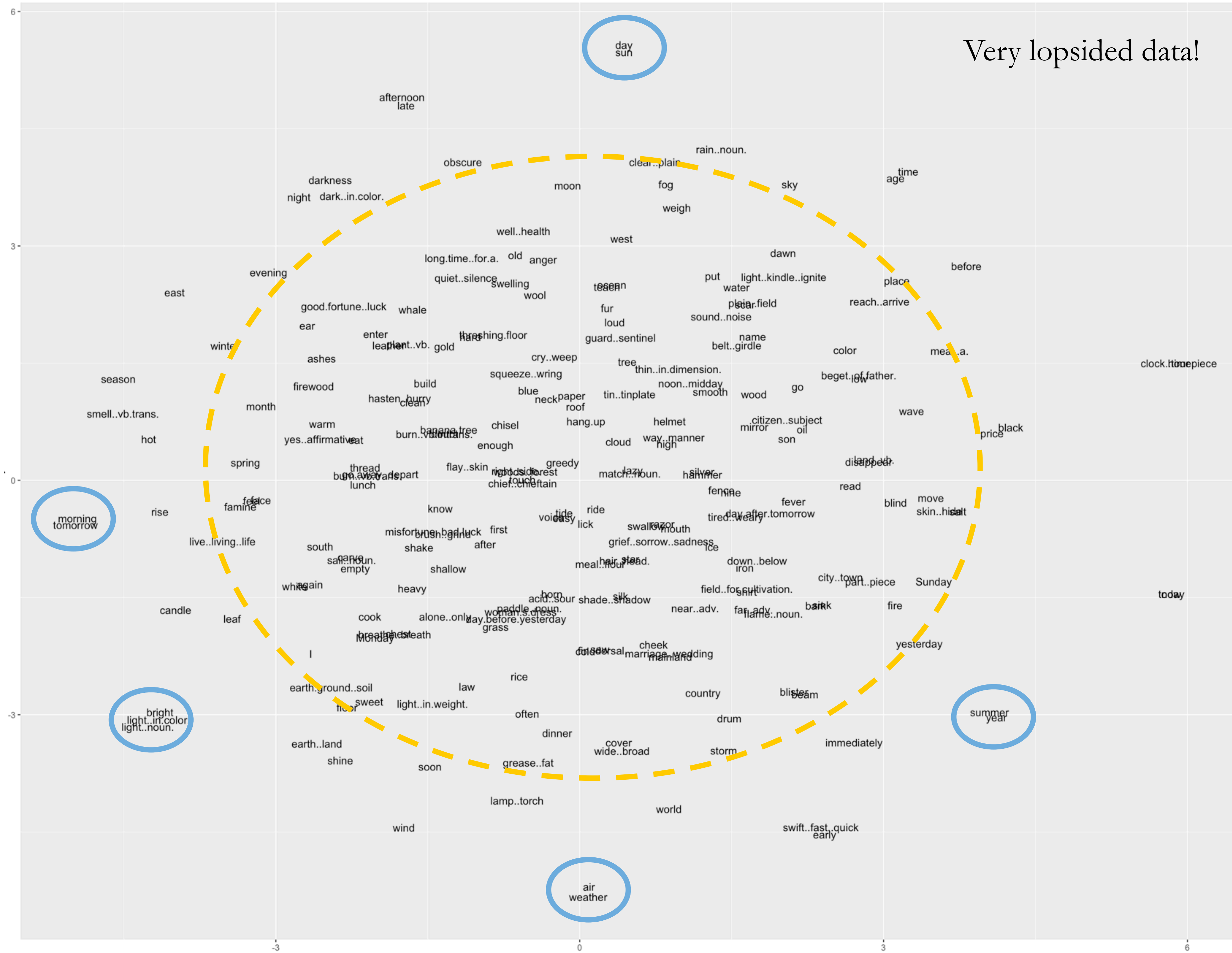
# Graphs vs feature projections

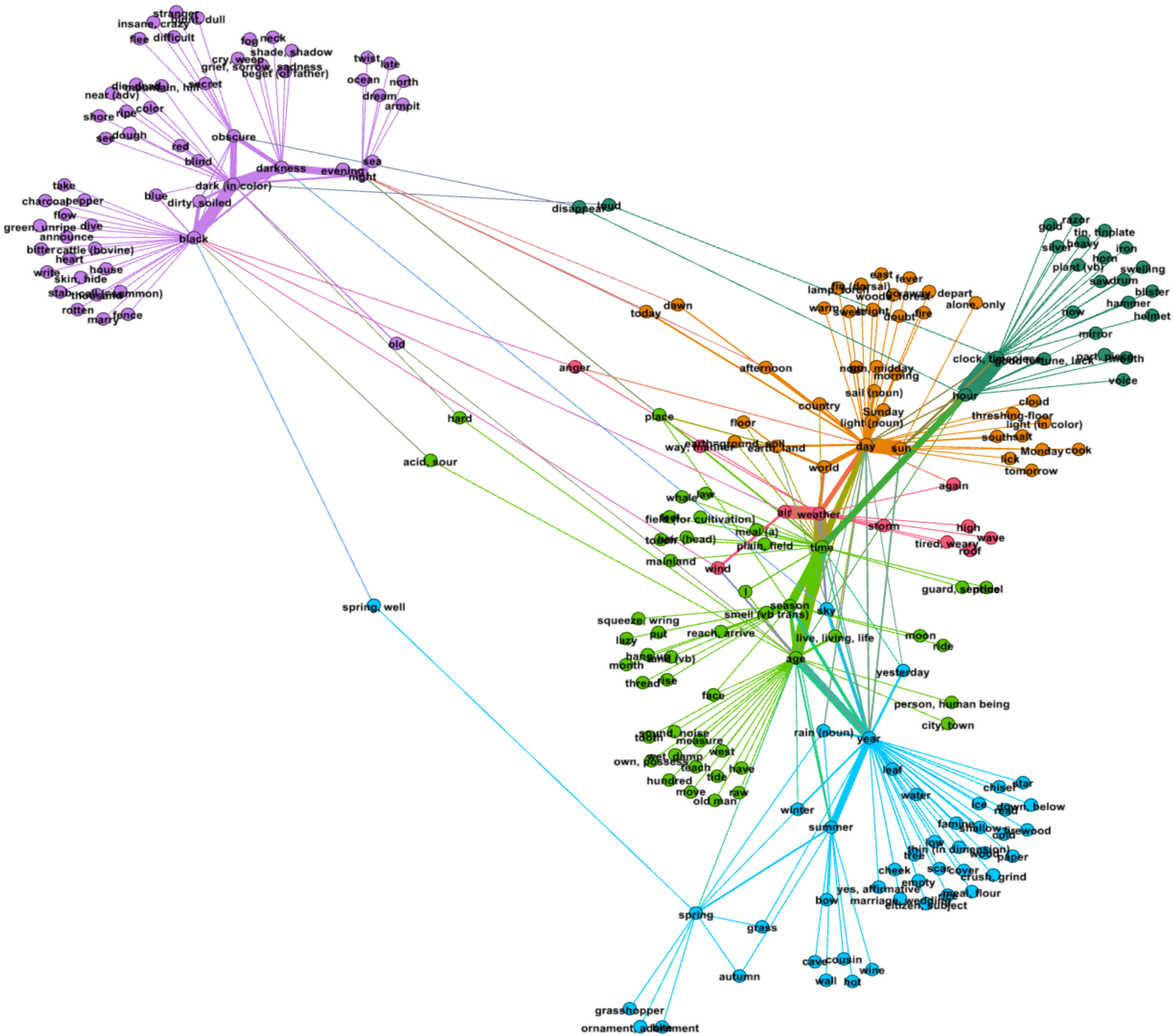
- If the data can be turned into a polysemy matrix, the graphs have some advantages over feature projections based on techniques of dimensionality reduction

  1. Readability and interpretability


t-SNE projection
6
3
0
-3
-3
0
3
6
day
sun
afternoon
late
obscure
clear..plain
rain..noun.
darkness
night
dark..in.color.
moon
fog
sky
age
time
weigh
well..health
west
dawn
before
long.time..for.a.
old
anger
evening
quiet..silence
swelling
put
light..kindle..ignite
place
east
water
wool
good.fortune..luck
whale
fur
reach..arrive
loud
sound..noise
ear
enter
threshing.floor
guard..sentinel
name
winter
gold
belt..girdle
color
meal..a.
ashes
cry..weep
tree
thin..in.dimension.
season
squeeze..wring
beget..of.father.
firewood
build
noon..midday
low
blue
smooth
wood
go
month
tin..tinplate
neck
paper
roof
smell..vb.trans.
wave
warm
chisel
hang.up
helmet
citizen..subject
mirror
oil
black
price
hot
yes..affirmative
enough
cloud
way..manner
high
son
spring
greedy
disappear
thread
flay..skin
depart
lunch
chief..chieftain
match..noun.
silver
hammer
read
fence
fever
blind
move
morning
tomorrow
rise
famine
know
ride
tired..weary
lick
mouth
first
after
live..living..life
south
shake
grief..sorrow..sadness
ice
empty
shallow
down..below
iron
city..town
Sunday
heavy
acid..sour
shade..shadow
field..for.cultivation.
candle
near..adv.
fire
leaf
cook
alone..only
day.before.yesterday
grass
cheek
yesterday
I
earth.ground..soil
law
rice
country
sweet
floor
light..in.weight.
often
drum
bright
light..in.color.
light..noun.
summer
year
dinner
cover
wide..broad
storm
immediately
earth..land
shine
soon
grease..fat
lamp..torch
world
wind
swift..fast..quick
early
air
weather

# Very lopsided data!





stranger
dull
insane, crazy
flee
difficult
fog
neck
shade, shadow
cry, weep
grief, sorrow, sadness
beget (of father)
twist
late
ocean
north
dream
armpit
near (adv)
secret
shore
ripe
color
dough
red
blind
obscure
darkness
evening
night
sea
dark (in color)
blue
dirty, soiled
take
charcoal
pepper
flow
green, unripe
dive
announce
black
bitter
cattle (bovine)
heart
write
house
skin, hide
stab
rotten
fence
marry
disappear
loud
gold
razor
tin, tinplate
silver
heavy
iron
horn
plant (vb)
swelling
sawdust
blister
hammer
helmet
now
mirror
clock, timepiece
good fortune, luck
voice
east
fever
lamp
depart
warm
alone, only
fire
dawn
today
afternoon
noon, midday
morning
sail (noun)
country
Sunday
light (noun)
cloud
threshing-floor
light (in color)
south
salt
Monday
cook
lick
tomorrow
hour
old
anger
place
floor
hard
acid, sour
way, manner
earth, land
day
sun
world
again
air
weather
storm
high
wave
tired, weary
roof
whale
field (for cultivation)
meal (a)
plain, field
time
mainland
wind
guard, sentinel
spring, well
season
sky
squeeze, wring
smell (vb trans)
put
lazy
reach, arrive
live, living, life
moon
ride
month
thread
age
yesterday
face
person, human being
city, town
sound, noise
tooth
measure
west
wet, damp
own, possess
teach
have
hundred
tide
move
raw
old man
rain (noun)
year
leaf
chisel
star
water
ice
down, below
winter
famine
shallow
cold
firewood
summer
low
thin (in dimension)
tree
paper
cheek
scar
cover
crush, grind
empty
yes, affirmative
meal, flour
bow
citizen, subject
spring
grass
cave
cousin
wine
wall
hot
autumn
grasshopper
ornament, adornment

# Graphs vs feature projections

- If the data can be turned into a polysemy matrix, then the graphs have significant advantages over feature projections based on techniques of dimensionality reduction

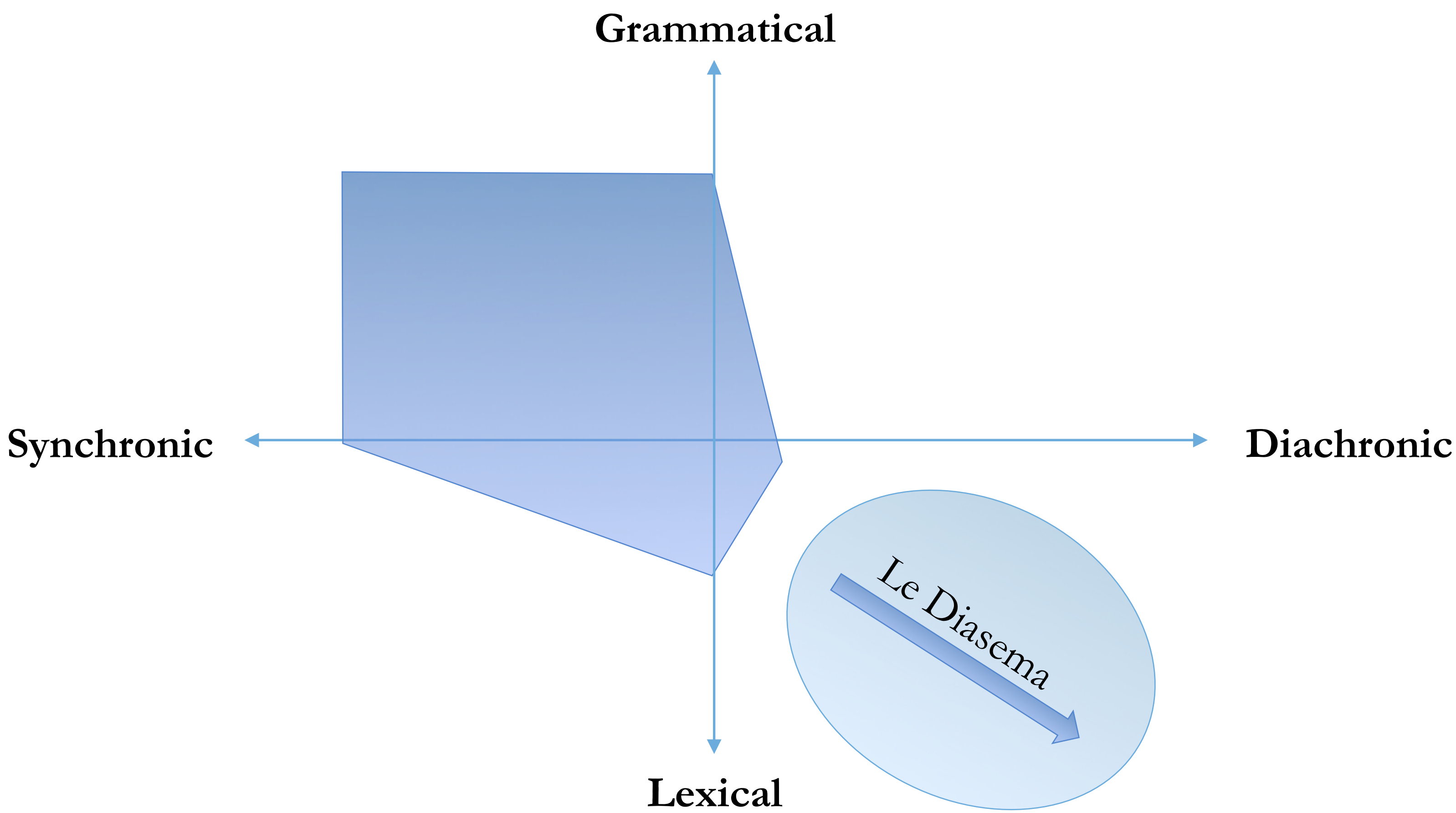
  1. Readability and interpretability
  2. Visual expressiveness
  3. Modifiability
  4. Statistical tools
     - Filtering
     - Clustering
     - Visualizing


poison
acid, sour
bitter
brackish
salt
pain
salty
sing
foam
voice
word
sound, noise
slide, slip
good
stomach
honey
sugar
smooth
sweet
shallow
empty
divide
dry
winter
cold (catarrh)
shade, shadow
cold
shout, cry out
weak
large, big
high
tall
soft
straight
loud
easy
float
lung
ice
silver
nose
clean
lamp, torch
strong, mighty, powerful
light (in weight)
beam
white
light (in color)
swift, fast, quick
brave
hard
bark
shut, close
sharp
sword
pointed
bright
clear, plain
lightning
candle
light (noun)
slow
sure, certain
wrinkled
rough
fragrant, good smelling
rough (of sea)
dear (costly, expensive)
thick (in dimension)
wet, damp
shine
peace
water
burn (vb intrans)
stinking, bad smelling
quiet, silence
rotten
pus
bake
calm (of sea)
grief, sorrow, sadness
heavy
love
pregnant
smell (vb intrans)
silent (be)
sniff
weigh
season
smell (vb trans)
count
breathe, breath
listen
feel
measure
obey
believe
touch
hear
guess
think (= be of the opinion)
learn
know
beautiful
ask (question, inquire)
regret, be sorry
imitate
cry, weep
look, look at
read
get, obtain
meet
seek, look for
seem
show
explain
announce
tell story
come
glass
bruise
egg yolk
sky
yellow
green
blue
ripe
green, unripe
red
blood
raw
burn (vb trans)
fire
color
face
dye
paint (noun)
line
paint (vb)
fever
hot
warm
perspire
lazy
evening
fishhook
blunt, dull
foolish, stupid
dirty, soiled
night
blind
black
obscure
dark (in color)
old

# Le Diasema

# Le Diasema

**Objectives:** 1) Add directionality to semantic maps of content words



# Le Diasema

**Objectives:** 2) Plot *diachronic* and *weighted* semantic maps automatically

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

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

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

# Le Diasema

**Objectives:** 3) Provide information about the *cognitive* and *cultural* factors behind the development of the various meanings

# Le Diasema

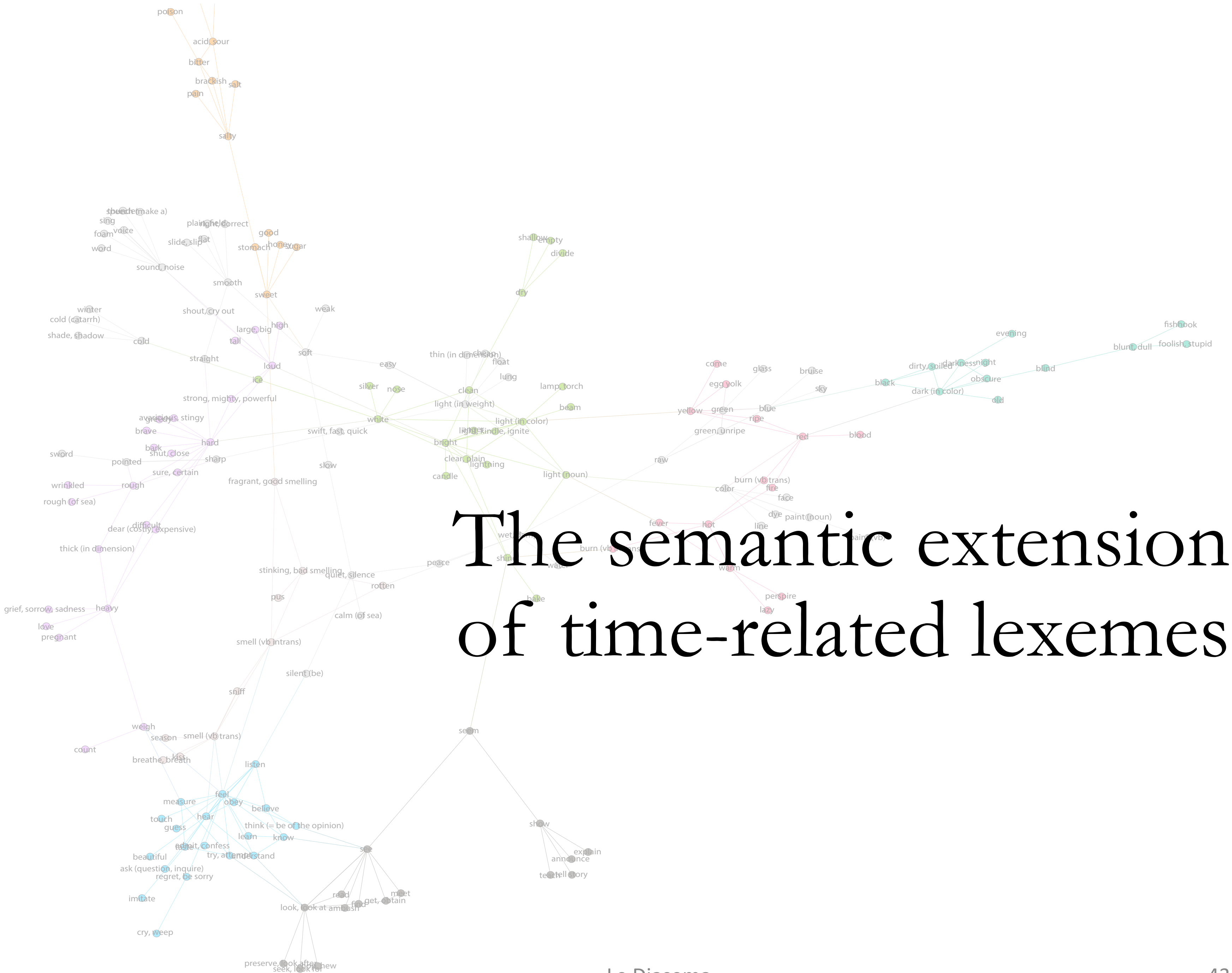
## Today's talk

1. Add directionality to semantic maps of content words
2. Plot diachronic and weighted semantic maps
3. Provide information about the cognitive and cultural factors behind the development of the various meanings

## We present:

- A protocol to construct lexical diachronic semantic maps based on a case-study:

  The semantic extension of **time-related lexeme**
- A way to visualize and analyze the results




# The semantic extension of time-related lexemes

# Protocol to construct a (lexical) diachronic semantic map

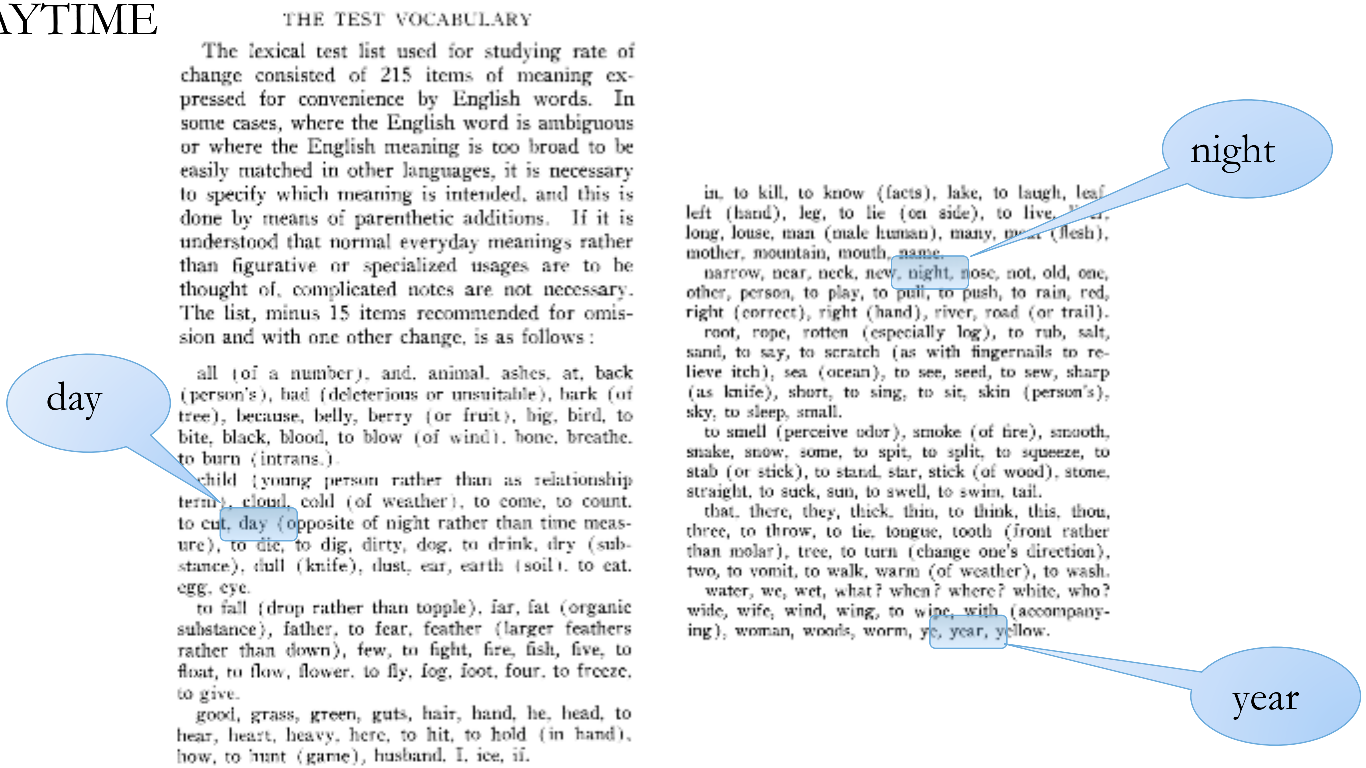
1. Choose the concepts / domains
2. Identify cross-linguistic polysemy patterns
3. Build a lexical matrix
4. Plot a weighted semantic map
5. Remove infrequent polysemy patterns
6. Select languages with diachronic data
7. Ensure comparability
8. Add diachronic information
9. Visualize the complete result

# Protocol to construct a (lexical) diachronic semantic map

## Choice of concepts

- 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 TEST VOCABULARY

The lexical test list used for studying rate of change consisted of 215 items of meaning expressed for convenience by English words. In some cases, where the English word is ambiguous or where the English meaning is too broad to be easily matched in other languages, it is necessary to specify which meaning is intended, and this is done by means of parenthetic additions. If it is understood that normal everyday meanings rather than figurative or specialized usages are to be thought of, complicated notes are not necessary. The list, minus 15 items recommended for omission and with one other change, is as follows:

all (of a number), and, animal, ashes, at, back (person's), bad (deleterious or unsuitable), bark (of tree), because, belly, berry (or fruit), big, bird, to bite, black, blood, to blow (of wind), bone, breathe, to burn (intrans.).

child (young person rather than as relationship term), cloud, cold (of weather), to come, to count, to cut, day (opposite of night rather than time measure), to die, to dig, dirty, dog, to drink, dry (substance), dull (knife), dust, ear, earth (soil), to eat, egg, eye.

to fall (drop rather than topple), far, fat (organic substance), father, to fear, feather (larger feathers rather than down), few, to fight, fire, fish, five, to float, to flow, flower, to fly, fog, foot, four, to freeze, to give.

good, grass, green, guts, hair, hand, he, head, to hear, heart, heavy, here, to hit, to hold (in hand), how, to hunt (game), husband, I, ice, if, in, to kill, to know (facts), lake, to laugh, leaf, left (hand), leg, to lie (on side), to live, [illegible], long, louse, man (male human), many, [illegible] (flesh), mother, mountain, mouth, name.

narrow, near, neck, new, night, nose, not, old, one, other, person, to play, to pull, to push, to rain, red, right (correct), right (hand), river, road (or trail).

root, rope, rotten (especially log), to rub, salt, sand, to say, to scratch (as with fingernails to relieve itch), sea (ocean), to see, seed, to sew, sharp (as knife), short, to sing, to sit, skin (person's), sky, to sleep, small.

to smell (perceive odor), smoke (of fire), smooth, snake, snow, some, to spit, to split, to squeeze, to stab (or stick), to stand, star, stick (of wood), stone, straight, to suck, sun, to swell, to swim, tail.

that, there, they, thick, thin, to think, this, thou, three, to throw, to tie, tongue, tooth (front rather than molar), tree, to turn (change one's direction), two, to vomit, to walk, warm (of weather), to wash.

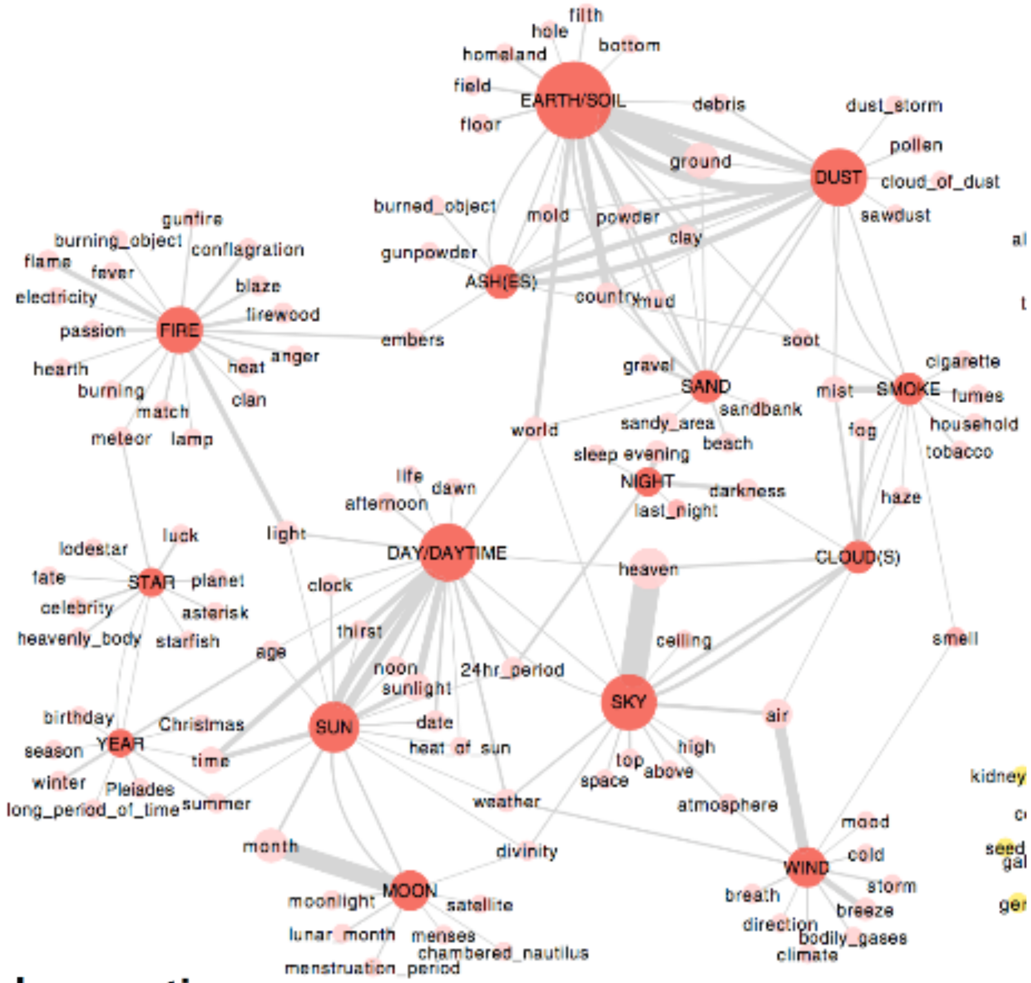
water, we, wet, what? when? where? white, who? wide, wife, wind, wing, to wipe, with (accompanying), woman, woods, worm, ye, year, yellow.

# Protocol to construct a (lexical) diachronic semantic map

## Choice of concepts

- We chose the entries for time-related concepts also for the sake of comparability

  (see, e.g. , Youn et al. 2016)

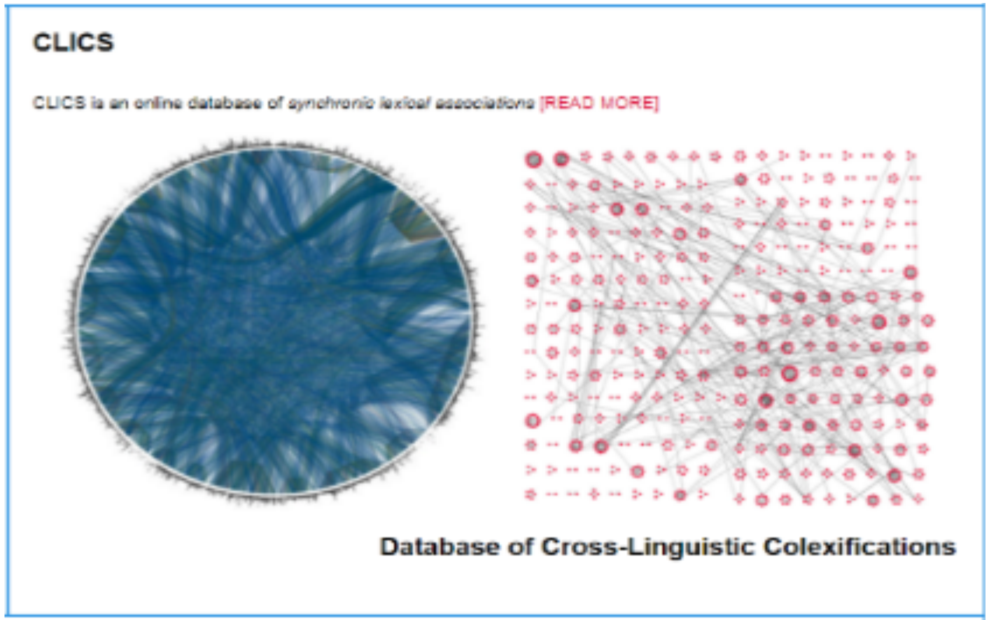



**On the universal structure of human lexical semantics**

Hyejin Youn[a,b,c,1], Logan Sutton[d], Eric Smith[c,e], Cristopher Moore[c], Jon F. Wilkins[c,f], Ian Maddieson[g,h], William Croft[g], and Tanmoy Bhattacharya[c,i,1]

# Protocol to construct a (lexical) diachronic semantic map

## Identify cross-linguistic polysemy patterns



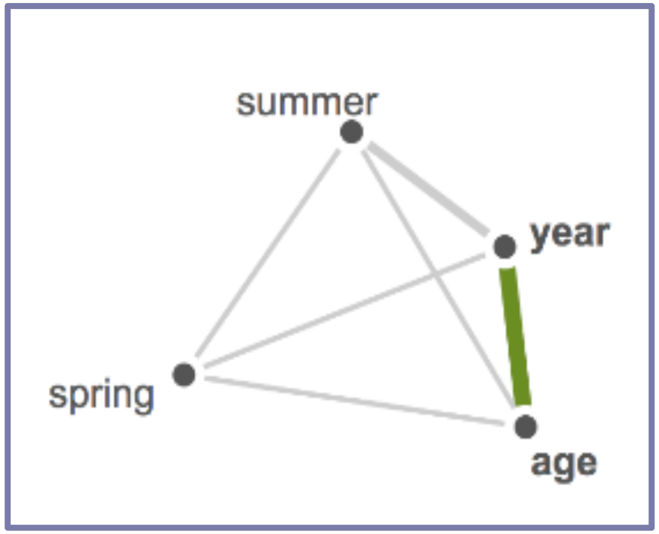

- *N* of lgs: 221
- *N* of lg families: 64
- *N* of concepts: 1280

- Identify in CLICS (List et al. 2014) the main polysemy patterns attested for these three meanings [16 meanings]

# Protocol to construct a (lexical) diachronic semantic map

## Identify cross-linguistic polysemy patterns

- Identify in CLICS (List et al. 2014) the main polysemy patterns attested for these three meanings [16 meanings]
  - **DAY/DAYTIME**: CLOCK/TIMEPIECE, HOUR, SEASON, SUN, TIME, WEATHER
  - **NIGHT**: DARK (in color), DARKNESS, BLACK, OBSCURE
  - **YEAR**: AGE, SPRING, SUMMER



# Protocol to construct a (lexical) diachronic semantic map

## Identify cross-linguistic polysemy patterns

- All the colexification patterns attested for these 16 meanings were gathered from the CLICs source files ([http://clics.lingpy.org/download.php](http://clics.lingpy.org/download.php)):

➡ **381 colexification patterns**

| | A | B | C |
|---|---|---|---|
| 119 | day | afternoon | hau_std:rana//ket_std:iʔ//plj_std:piidɨ//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:šãʔ |
| 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:nału//haw_std:ao//mri_std:ao//tzz_std:osil |
| 130 | day | east | tob_std:naʔaʔk |
| 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:thəm |
| 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:aʔta//crt_std:xloma//haw_std:ao//hdn_Northern:ˁkatʰkáa//ito_std:uwayo//mzh |
| 141 | day | live, living, life | shp_std:niti |

# Protocol to construct a (lexical) diachronic semantic map

## Convert the polysemy patterns 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)
```

Python script $\alpha$

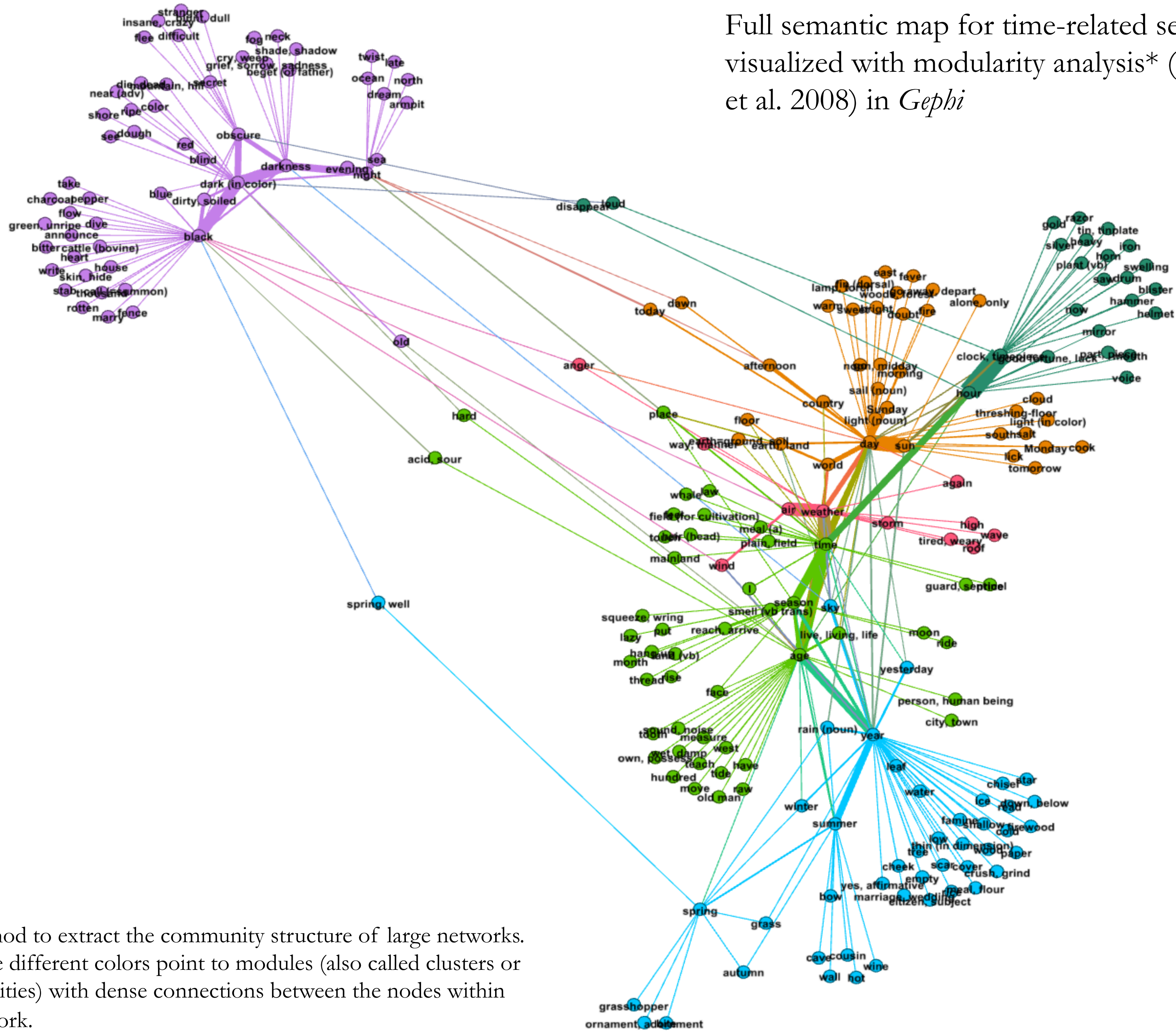
Languages Forms Meanings

| | 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 | ʿara | 1 | 0 | 0 | 1 |
| 6 | ncg_std | йуз | 1 | 0 | 0 | 0 |
| 7 | mri_std | pakeke | 1 | 0 | 0 | 0 |
| 8 | pbb_std | hi?ph | 1 | 0 | 0 | 0 |
| 9 | khv_Khvarshi | замана | 1 | 0 | 0 | 0 |

1 when a meaning is attested for one form

Lexical matrix

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



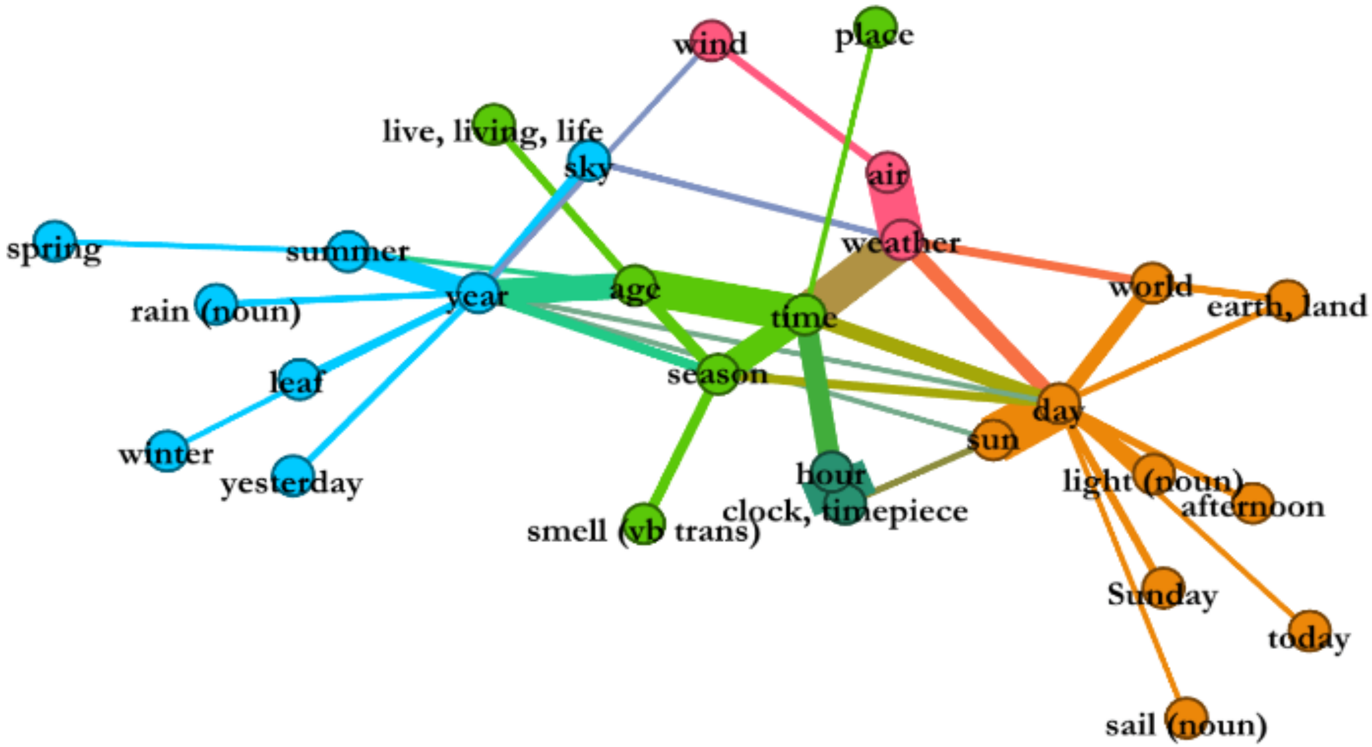

* A method to extract the community structure of large networks. Here, the different colors point to modules (also called clusters or communities) with dense connections between the nodes within the network.

# Remove infrequent polysemy patterns


**Semantic map** of time-related senses (colexification patterns attested in 2+ languages)

Two connected sub-networks

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

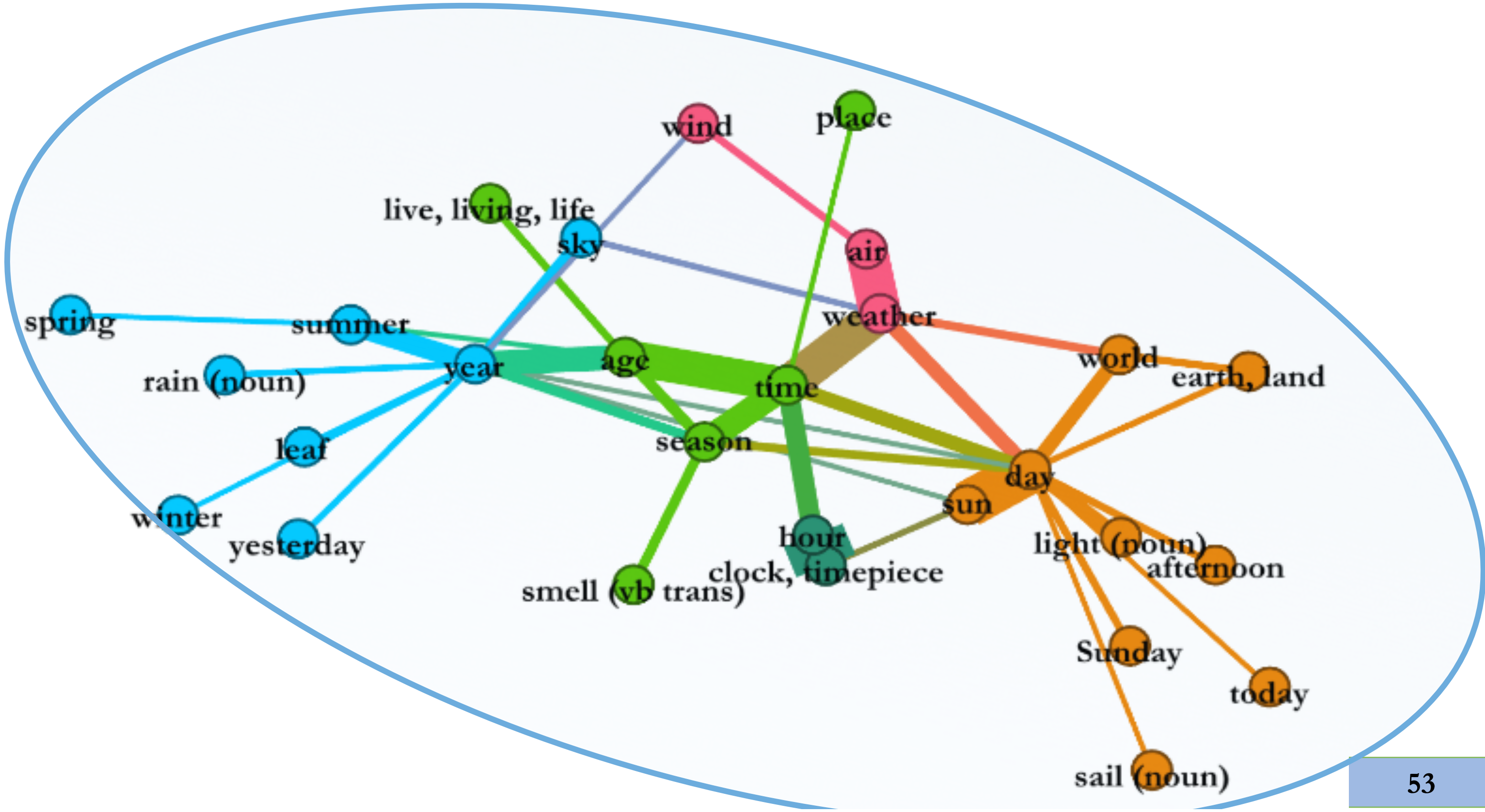


# Remove infrequent polysemy patterns


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

Two connected sub-networks

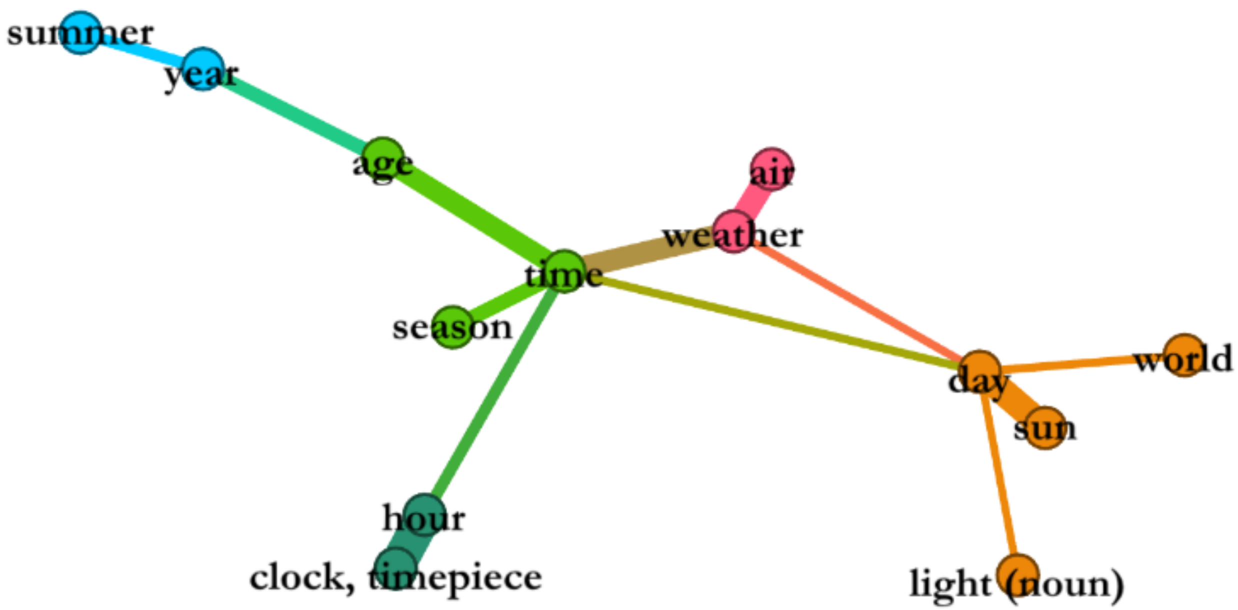
- NIGHT/DARKNESS/DARK
- **DAY/TIME/AGE/YEAR**



# Protocol to construct a (lexical) diachronic semantic map

## Remove infrequent polysemy patterns

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



# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

(1) *Meanings*: tree *(source)*—forest (*target*) (ID: 600); *Form*: dar; *Language*: Aghul; *Realization Type*: **synchronic polysemy**

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

(1) *Meanings*: tree *(source)*—forest (*target*) (ID: 600); *Form*: dar; *Language*: Aghul; *Realization Type*: **synchronic polysemy**

(2) *Meanings*: doll *(source)*—nymph, chrysalis (*target*) (ID: 927); *Form*: kukla; *Language pair*: Russian —Czech; *Realization Type*: **Cognate**

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

(1) *Meanings*: tree *(source)*—forest (*target*) (ID: 600); *Form*: dar; *Language*: Aghul; *Realization Type*: **synchronic polysemy**

(2) *Meanings*: doll *(source)*—nymph, chrysalis (*target*) (ID: 927); *Form*: kukla; *Language pair*: Russian —Czech; *Realization Type*: **Cognate**

(3) *Meanings*: arc *(source)* → rainbow (*target*) (ID: 393); *Form*: Bogen → Regenbogen; *Language*: German; *Realization Type:* **Morphological derivation**

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

(1) *Meanings*: tree *(source)*—forest (*target*) (ID: 600); *Form*: dar; *Language*: Aghul; *Realization Type*: **synchronic polysemy**

(2) *Meanings*: doll *(source)*—nymph, chrysalis (*target*) (ID: 927); *Form*: kukla; *Language pair*: Russian —Czech; *Realization Type*: **Cognate**

(3) *Meanings*: arc *(source)* → rainbow (*target*) (ID: 393); *Form*: Bogen → Regenbogen; *Language*: German; *Realization Type:* **Morphological derivation**

(4) *Meanings*: to count *(source)* → speech (*target*) (ID: 11); *Forms*: ratio → Rede; *Languages*: Latin (*donor*) → German (*target*); *Realization Type:* **Borrowing**

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

(1) *Meanings*: tree *(source)*—forest (*target*) (ID: 600); *Form*: dar; *Language*: Aghul; *Realization Type*: **synchronic polysemy**

(2) *Meanings*: doll *(source)*—nymph, chrysalis (*target*) (ID: 927); *Form*: kukla; *Language pair*: Russian —Czech; *Realization Type*: **Cognate**

(3) *Meanings*: arc *(source)* → rainbow (*target*) (ID: 393); *Form*: Bogen → Regenbogen; *Language*: German; *Realization Type:* **Morphological derivation**

(4) *Meanings*: to count *(source)* → speech (*target*) (ID: 11); *Forms*: ratio → Rede; *Languages*: Latin (*donor*) → German (*target*); *Realization Type:* **Borrowing**

(5) *Meanings*: to catch *(source)* → to hunt (*target*) (ID: 415); *Forms*: capto → cacciare; *Languages*: Latin → Italian; *Realization Type:* **Diachronic semantic evolution**

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

DatSemShifts

Home | Semantic shifts | Meanings | Languages | Participants | Publications | Contact us | Log in

| ID | Source | Direction | Target | Status | Contributed by | Accepted realization | Show |
|---|---|---|---|---|---|---|---|
| 53 | time | — | weather | Accepted | DG | 4 | Show |
| 109 | time | — | opportunity | Accepted | IG | 2 | Show |
| 395 | time | — | hour | Accepted | DG | 2 | Show |
| 406 | time | — | 24 hours | Suspended | DG | 0 | Show |
| 795 | time | → | one time, once | New | MB | 0 | Show |
| 1446 | time | → | journal, magazine | Accepted | IG | 3 | Show |

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

DatSemShifts

Home | Semantic shifts | Meanings | Languages | Participants | Publications | Contact us | Log in

| ID | Source | Direction | Target | Status | Contributed by | Accepted realization | Show |
|---|---|---|---|---|---|---|---|
| 53 | time | — | weather | Accepted | DG | 4 | Show |
| 109 | time | — | opportunity | Accepted | IG | 2 | Show |
| 395 | time | — | hour | Accepted | DG | 2 | Show |
| 406 | time | — | 24 hours | Suspended | DG | 0 | Show |
| 795 | time | → | one time, once | New | MB | 0 | Show |
| 1446 | time | → | journal, magazine | Accepted | IG | 3 | Show |

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)

| ID | Source | Direction | Target | Status | Contributed by |
|---|---|---|---|---|---|
| 1446 | time | → | journal, magazine | Accepted | IG |

Comments:
Ср. греч. хронограф, откуда могут быть кальки.

Confirmed by 3 Guru(s)

Derivation: German *Zeit* → *Zeitung, Zeitschrift* 'newspaper, journal'

Derivation: Karaim *вахт* 'time' → *вахтлых* 'journal'

Polysemy: Polish *czas* 'time' — 'journal'

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- The Catalogue of Semantic Shifts in the Languages of the World (Zalizniak, 2006; Zalizniak et al., 2012; http://semshifts.iling-ran.ru/)
  - Relies predominantly on synchronic polysemy
  - Mirrors the polysemous view of semantic change

    (see Sweetser 1990: 9; Evans, 1992: 476; Geerarts, 1997: 6; Wilkins, 1996: 269-271; Evans & Wilkins, 2000: 549ff)

- We share this view, but:
  - **Our focus**: diachronic semantic developments of individual lexemes in the course of their semantic history
  - **Advantage:** theorize about semantic change based on actual data
  - **Disadvantage:** not many languages with significant diachronic data!

# Protocol to construct a (lexical) diachronic semantic map

## Select languages with diachronic data

- **Ancient Greek** ($8^{th}$ c. BC – $4^{th}$ c. AD)
  - Perseus digital library (http://www.perseus.tufts.edu/hopper/), TLG (http://stephanus.tlg.uci.edu)
  - Cunliffe (*A lexicon of the Homeric Dialect*), LSJ
  - Dictionary of Selected Synonyms in the Principal Indo-European Languages (Buck, 1949)
  - Etymological dictionaries (e.g., Beekes, 2010)

- **Ancient Egyptian** ($26^{th}$ c. BC – $10^{th}$ c. AD)
  - Corpora
    - Thesaurus Linguae Aegyptiae (http://aaew.bbaw.de/tla/)
    - Ramses Online (http://ramses.ulg.ac.be)
  - Lexical resources (Dictionaries and Coptic etymological dictionaries)

# Protocol to construct a (lexical) diachronic semantic map

## Ensure comparability

- Provide definitions for the 13 concepts
  - Use Concepticon (http://concepticon.clld.org)
    (a) the concept sets are given a unique definition
    (b) CLICS is one of the lists included in Concepticon*

| CONCEPT | DEFINITION IN CONCEPTICON | ADJUSTED DEFINITION |
|---|---|---|
| AGE | The period of time that a person, animal or plant has lived or is expected to live. | |
| DAY/DAYTIME | The period between sunrise and sunset where one enjoys daylight. | |
| SUN | The particular star at the centre of our solar system, from which the Earth gets light and heat. | The star that is the source of light and heat for the planets in the solar system (Wordnet) |

Table. Definitions of concepts (incomplete)

* The elicitation of the data in CLICS precedes the addition of the definitions in Concepticon

# Protocol to construct a (lexical) diachronic semantic map

## Ensure comparability

- Proceed onomasiologically

| CONCEPT | DEFINITION IN CONCEPTICON | ADJUSTED DEFINITION | LEXEME IN *AEG* | LEXEME IN *AG* |
|---|---|---|---|---|
| AGE | The period of time that a person, animal or plant has lived or is expected to live | | *jꜣk.t*<br>*šms*<br>*snḫy.t*<br>*jꜣw* (old age)<br>*jz.t* (age, decline) | *hēlikía* |
| DAY/DAYTIME | The period between sunrise and sunset where one enjoys daylight. | | *hrw*<br>*ḥḏ.t*<br>*rʕw* (*nb*)<br>*sw* (calenderic) | *êmar*<br>*ēṓs* |
| SUN | The particular star at the centre of our solar system, from which the Earth gets light and heat. | The star that is the source of light and heat for the planets in the solar system (Wordnet) | *rʕw*<br>*šw*<br>*jtn*<br>etc. | *hḗlios or ēélios* |

Table. Definitions of concepts and lexemes expressing the concepts

# Protocol to construct a (lexical) diachronic semantic map

## Ensure comparability

- Proceed onomasiologically
- Proceed semasiologically
  - List the different meanings of the lexemes identified
    - Dictionary-based
    - Other available resources
    - Corpus queries
  - Collect at least two text examples of each of the meanings

# Protocol to construct a (lexical) diachronic semantic map

## Add diachronic information

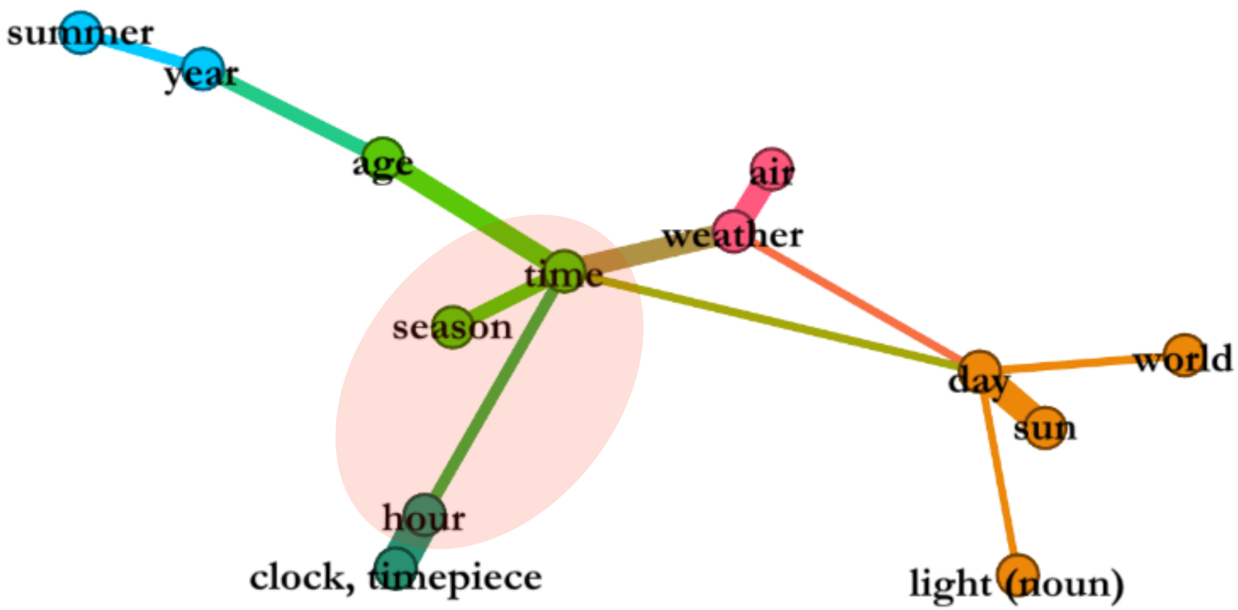
- Include directionality of change
  - Add arrows to the existing synchronic map
  - Enrich the existing synchronic map with additional nodes and add arrows (if necessary)
- Visualize the (complex) results

**+ Semantic analysis**

# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

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



# The semantic extension of time-related lexemes

## Ancient Greek: *hṓra* ‘season’

- Proto Indo-European root **Hieh,-r-, Hioh,-r-* ‘year’ (Beekes, 2010: 1681)

(1) *hóssá* *te* *phúlla* *kaì* *ánthea*
REL.NOM.PL.N PTC leaf: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)

Approx.
**8th c. BC**

**hôrai (seasons):**

- **spring** (*éaros hṓrē* ‘spring season’; *Iliad* 6.148),
- **winter** (*hṓrēi kheimeríēi* ‘in wintry season’; *Odyssey* 5.485)
- **summer/autumn** (*hṓrai epibríseian* ‘in rainy seasons’; *Odyssey* 24.344)

# The semantic extension of time-related lexemes

## Ancient Greek: *hṓra* 'time/moment'

Approx. **8th c. BC**

| (2) | *óphra* | *Poseidā́ōni* | *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)

# The semantic extension of time-related lexemes

Ancient Greek: *hṓra* 'time/moment'

5th c. BC

| | | | | | |
|---|---|---|---|---|---|
| (3) | *makrá* | *moi* | *neîsthai* | *kat'* | *amaksitón:* |
| | long:NOM.SG.F | 1SG.DAT | go:PRS.INF.M/P | DIR.INFR | highway:ACC.SG.M |
| | ***hṓra*** | *gàr* | *sunáptei* | | |
| | **time:NOM.SG.F** | PTC | join.together:PRS.3SG | | |

'Returning home by highway is too long; for **time** is approaching'
(Pindar, *Pythian* 4.247)

# The semantic extension of time-related lexemes

Ancient Greek: *hṓra* ‘time/moment’ ⇒ ‘hour’

Approx.
**1st c. AD**

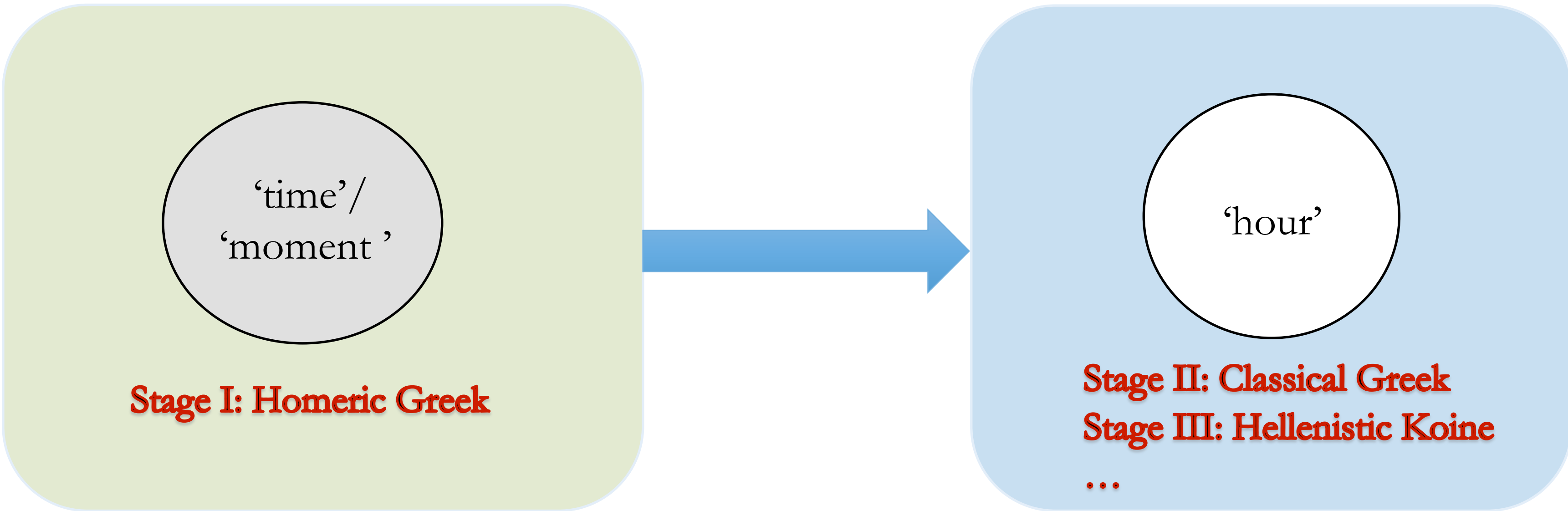
| | | | | | | |
|---|---|---|---|---|---|---|
| (4) | *oukhì* | *dṓdeka* | ***hôraì*** | *eisin* | *tês* | *hēméras;* |
| | NEG | twelve | hour:NOM.PL.F | be.PRS.3PL | ART.GEN.SG.F | day:GEN.SG.F |

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

- *hôra* conveyed the meaning ‘hour’ as early as the 4th c. BC

# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

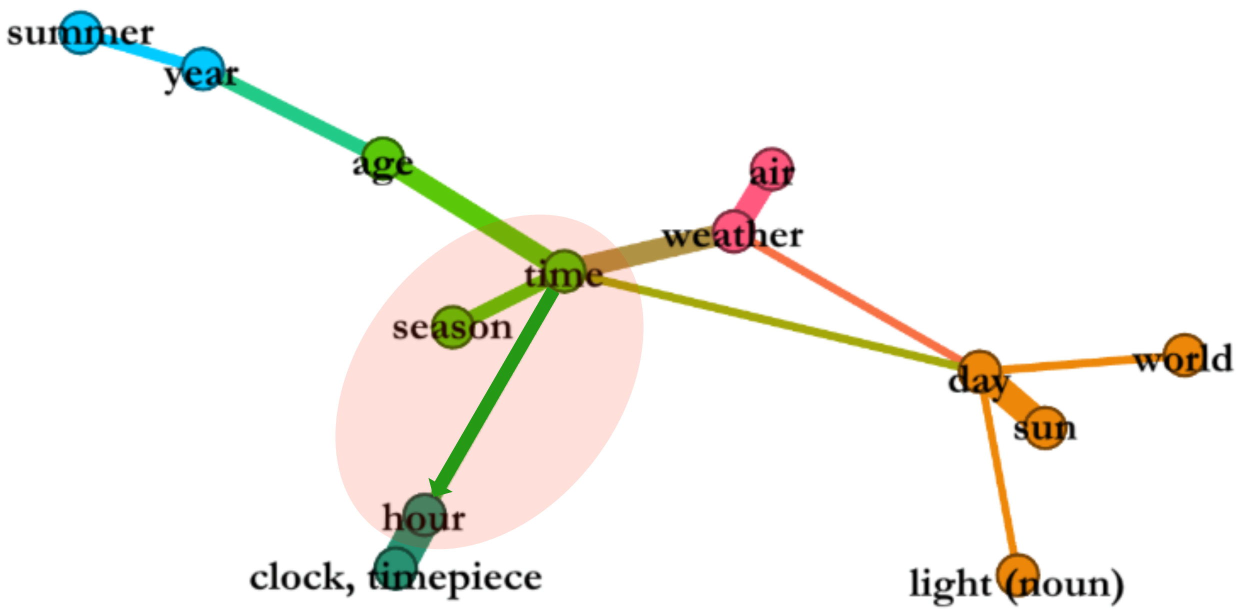



**Metonymy**: due to the correlation between the canonical time periods and the time these take to unfold

# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

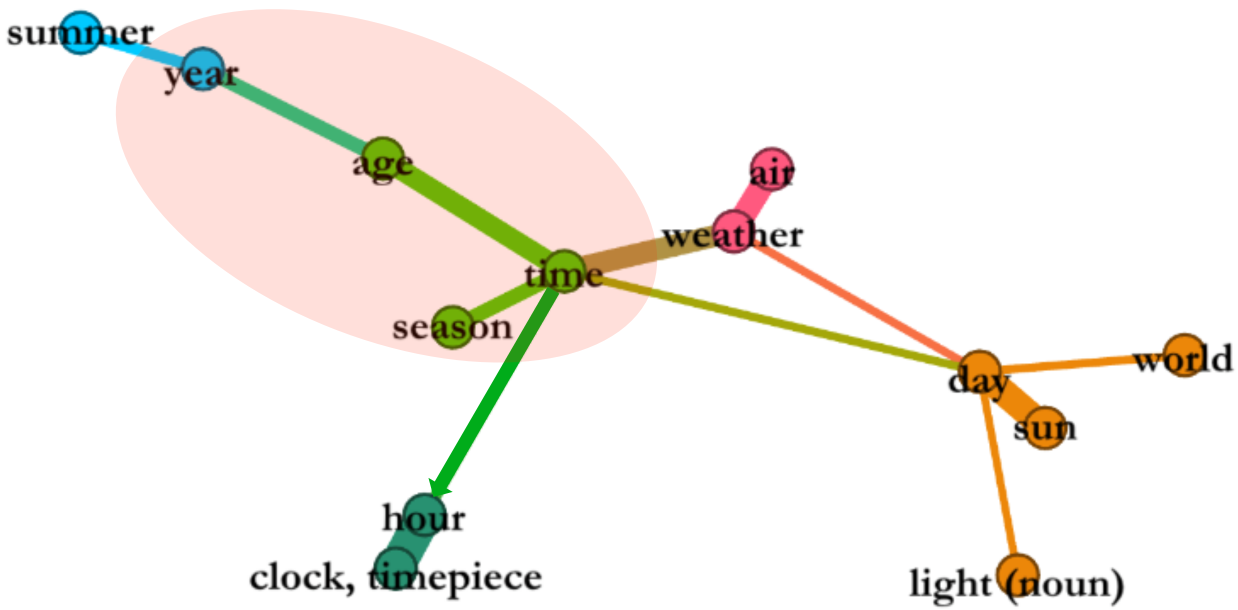
- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns
  - **SEASON-TIME-HOUR**



# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns
  - **TIME-AGE-YEAR-SEASON:** Reintegrating edges



# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns
  - **TIME-AGE-YEAR-SEASON:** Reintegrating edges

| | | | | | |
|---|---|---|---|---|---|
| (5) | *dôke* | *dé* | *m'* | *ekdeíras* | *askòn* |
| | give:AOR.3SG | PTC | 1SG.ACC | strip.off:PTCP.AOR.NOM.SG.M | skin:ACC.SG.M |
| | *boòs* | ***enneṓroio,*** | | | |
| | ox:GEN.SG.M | **nine.years.old:GEN.SG.M** | | | |

'He gave me a wallet, made of the hide of an ox **nine years old**, which he flayed' (Homer, *Odyssey* 10.19)

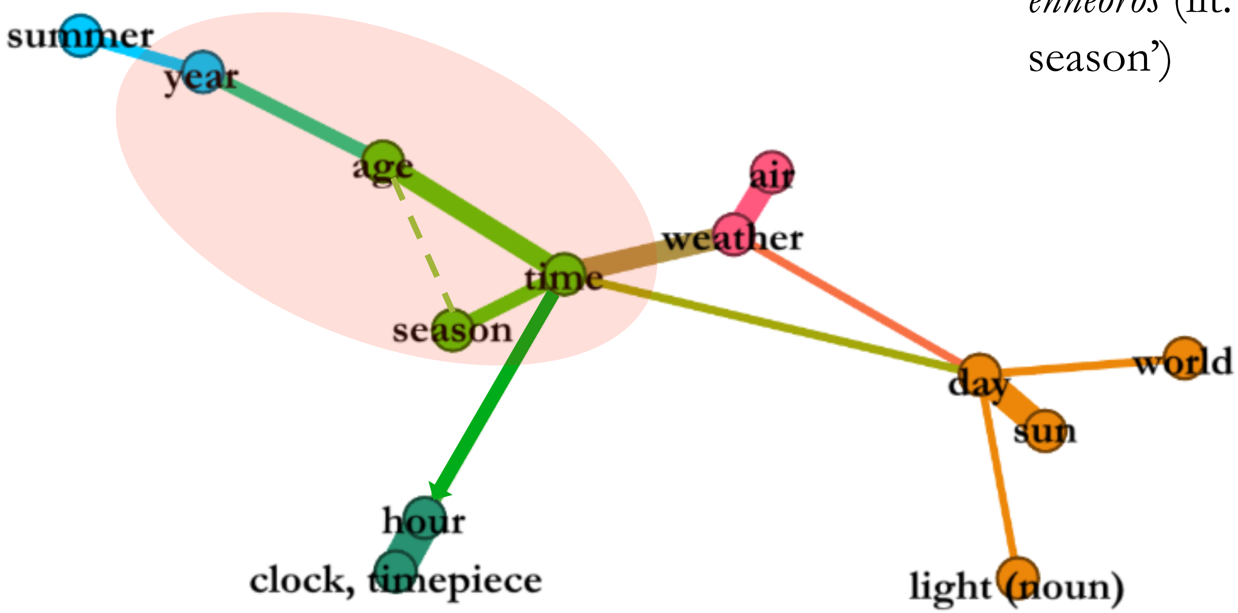
Approx.
**8th c. BC**

# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

- The diachronic material allows us to add diachronic information (graphically, oriented edges) between frequent colexification patterns
  - **TIME-AGE-YEAR-SEASON:** Reintegrating edges

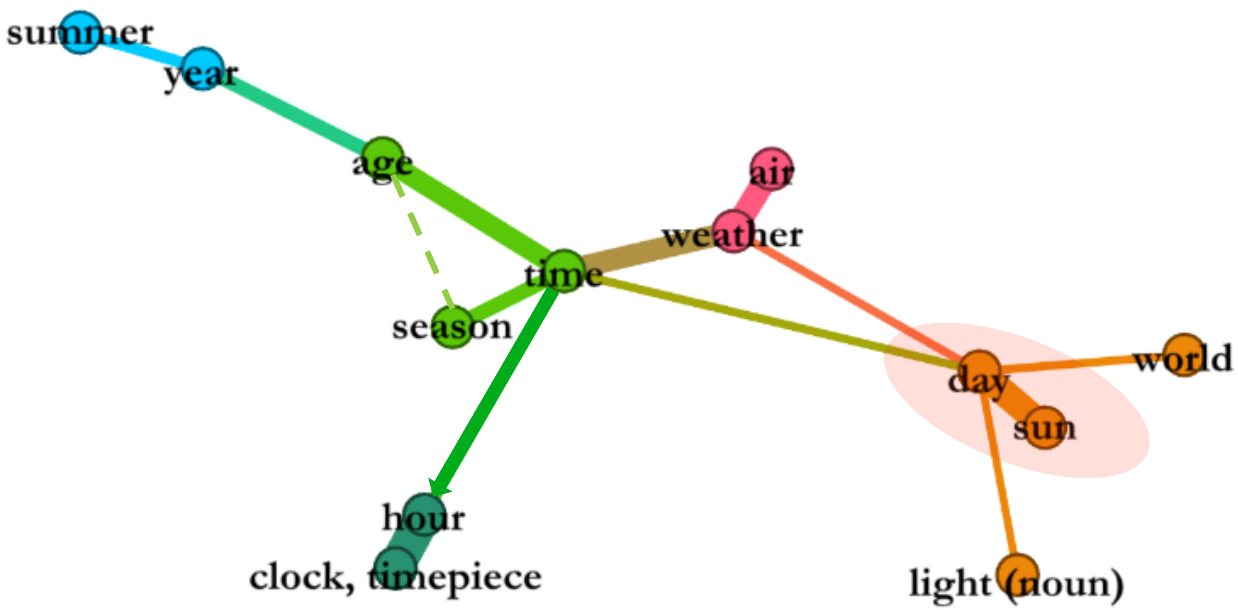
**Loose colexifications:**
*ennéoros* (lit. 'in the ninth season')



# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

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



# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

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

| | | | | | | |
|---|---|---|---|---|---|---|
| (6) | *pân* | *d'* | *êmar* | *pherómēn,* | *háma* | *d'* |
| | whole:ACC.SG.N | PTC | day:ACC.SG.N | carry:IMPF.1PL.M/P | ADV | PTC |

| | | | | |
|---|---|---|---|---|
| ***ēelíōi*** | *katadúnti* | *káppeson* | *en* | *Lḗmnōi* |
| **sun:DAT.SG.M** | set:PTCP.AOR.DAT.SG.M | fall:AOR.1PL | in | Lemnos:DAT.SG |

'the whole day long I was carried headlong, and at **sun**set I fell in Lemnos'
(Homer, *Iliad* 1.592-593)

Approx.
**8th c. BC**

# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

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

**5th c. BC**

(7)

| | | | | | |
|---|---|---|---|---|---|
| *ékheis,* | *egṓ* | *te* | *sé:* | ***hēlíous*** | *dè* |
| have:PRS.2SG | 1SG.NOM | PTC | 2SG.ACC | **sun:ACC.PL.M** | PTC |

| | | | |
|---|---|---|---|
| *muríous* | *mólis* | *dielthṑn* | *ēisthomēn* |
| infinite:ACC.PL.M | ADV | pass:PTCP.AOR.NOM.SG.M | perceive:AOR.1SG.MID |

| | | |
|---|---|---|
| *tà* | *tês* | *theoû* |
| ART.ACC.PL.N | ART.GEN.SG.F | god:GEN.SG |

'You have me, and I have you; although it was hard to live through so many **days**, I now understand the actions of the goddess.' (Euripides, *Helen* 652-653)

# The semantic extension of time-related lexemes

## Add arrows to the existing synchronic map

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

(8)

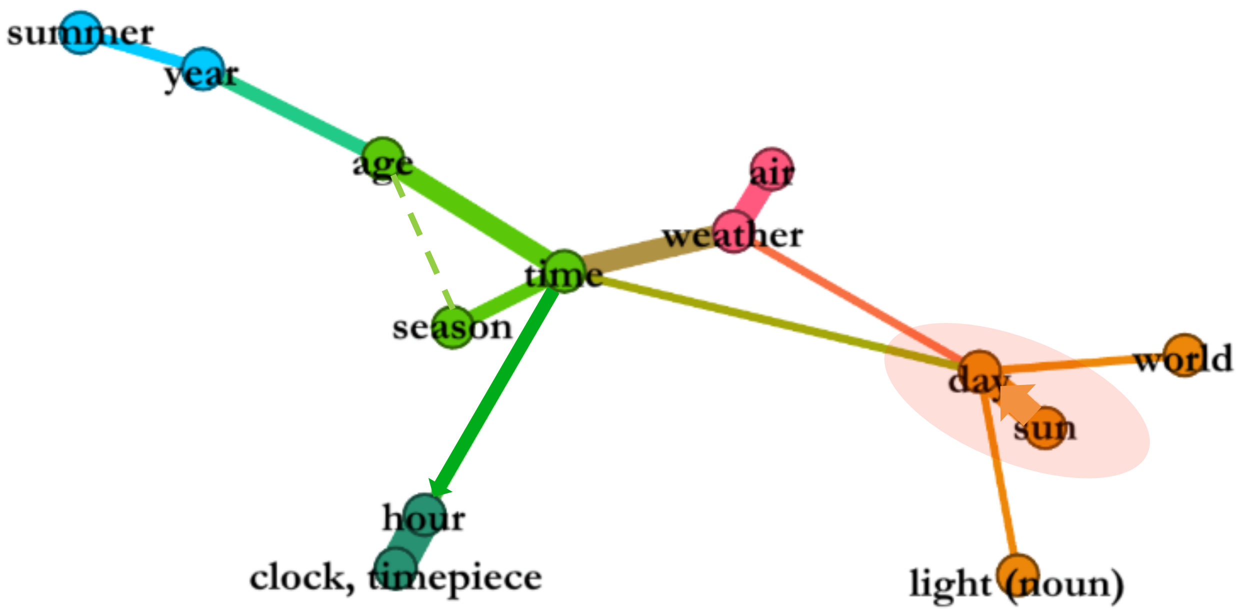
| | | | |
|---|---|---|---|
|  |  |  |  |
| *ḥs(y)-f* | *tw* | *rꜤw* | *nb* |
| PRAISE:SBJV-3SG.M | 2SG.M | sun | QUANT |

'(You should pray god non-stop,) so that he praise you every day .'
(P. Chester Beatty IV, v° 4,10)

# The semantic extension of time-related lexemes

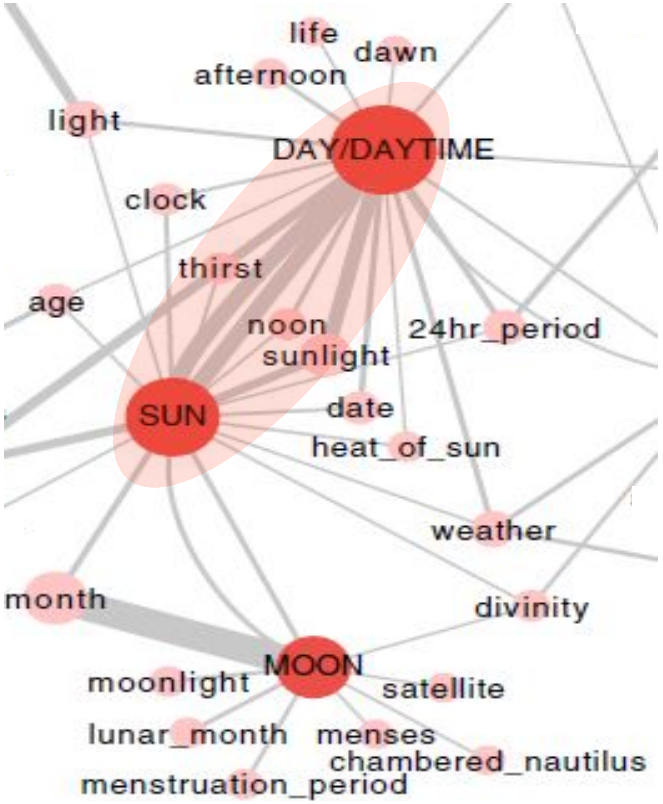
## Add arrows to the existing synchronic map

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



# The semantic extension of time-related lexemes

- **SUN-DAY**

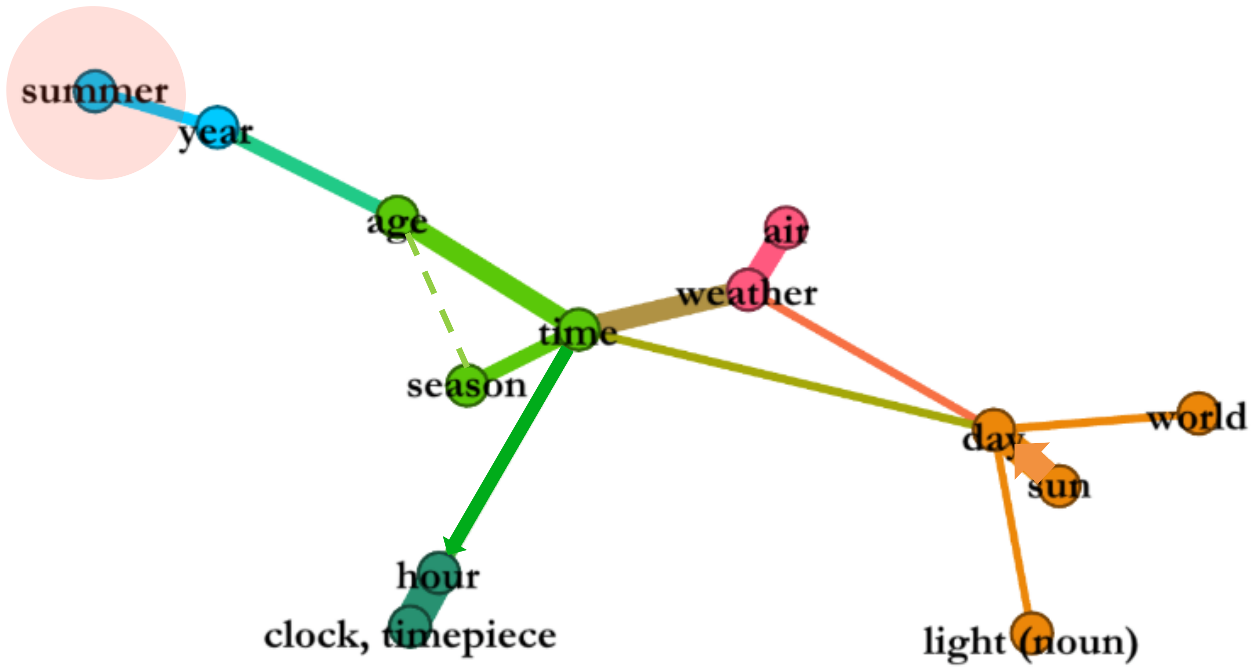



Snapshot from Youn *et al.* (2016) showing polysemy patterns of celestial objects and natural settings

# The semantic extension of time-related lexemes

## Enriching the map

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



# The semantic extension of time-related lexemes

## Enriching the map

- **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 |

(http://clics.lingpy.org/all.php?gloss=summer)

# The semantic extension of time-related lexemes

## Enriching the map

*In Homer, *karpós* is used for 'harvest'

- Ancient Greek: *théros* 'summer' ⇒ 'harvest'*

(9)

| | | | | |
|---|---|---|---|---|
| *autàr* | *epḕn* | *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 |

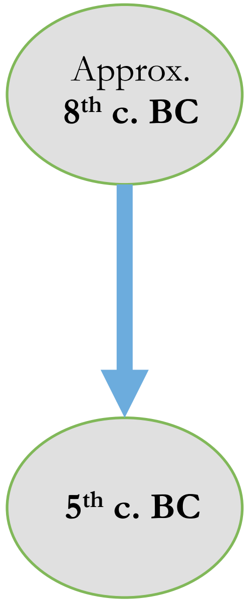
'But when **summer** comes and rich autumn' (Homer, *Odyssey* 11.192)

(10)

| | | | | |
|---|---|---|---|---|
| *kâit'* | *anḕr* | *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** |

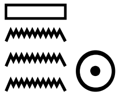
'he has only made himself a name by reaping another's **harvest**'
(Aristophanes, *Knights* 392)

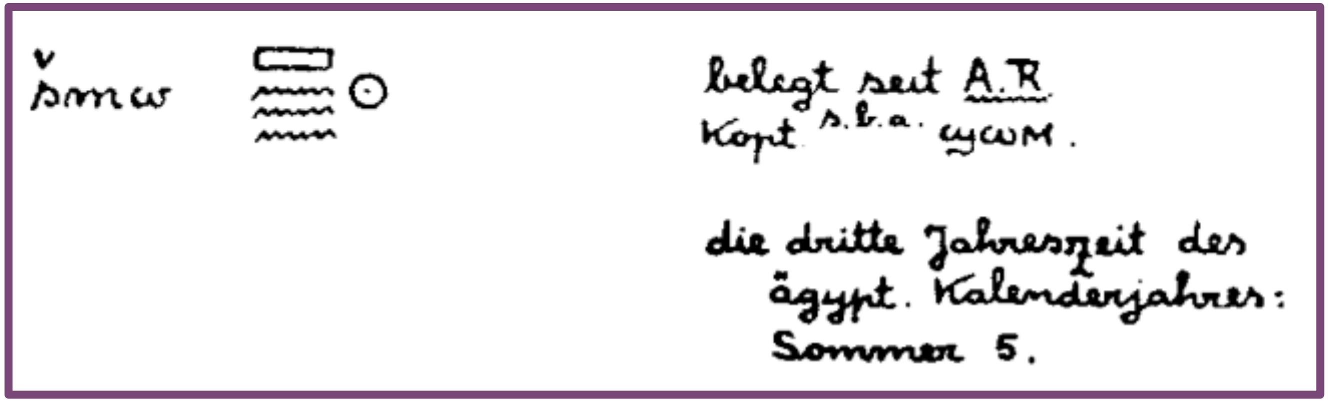


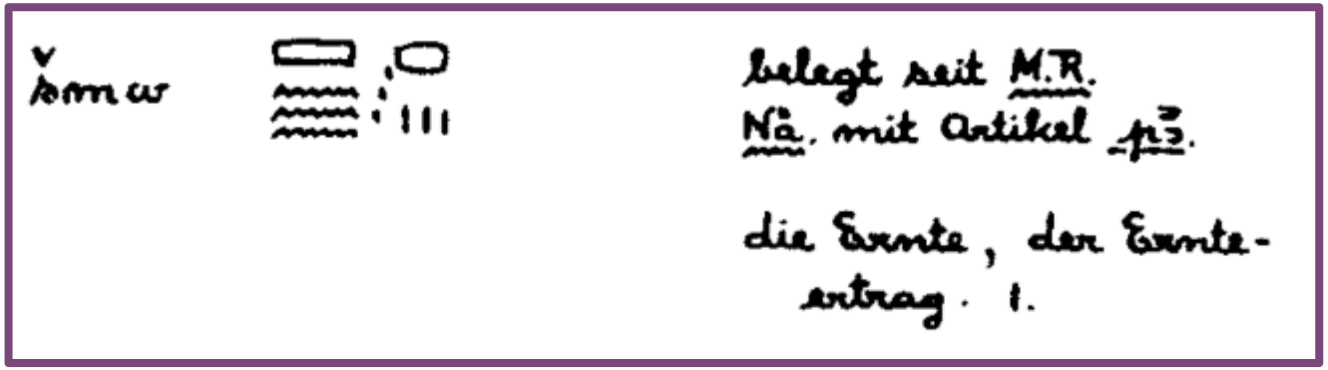
# The semantic extension of time-related lexemes

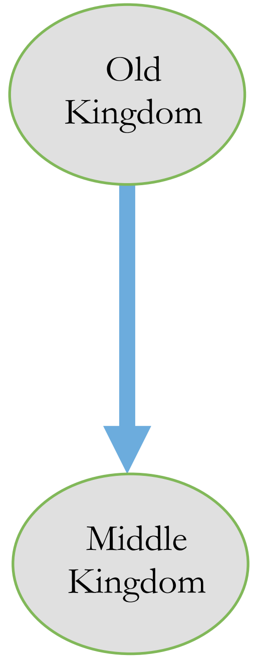
## Enriching the map

*In OEg, another lemma is used for 'harvest' (*ꜣzḫ*)

**Ancient Egyptian:**  *šmw* 'summer' ⇒ *šmw* 'harvest'*


šmw
belegt seit A.R.
Kopt. s.b.a. ϣⲱⲙ.
die dritte Jahreszeit des ägypt. Kalenderjahres: Sommer 5.


šmw
belegt seit M.R.
Nä. mit Artikel pꜣ.
die Ernte, der Ernte-ertrag. 1.

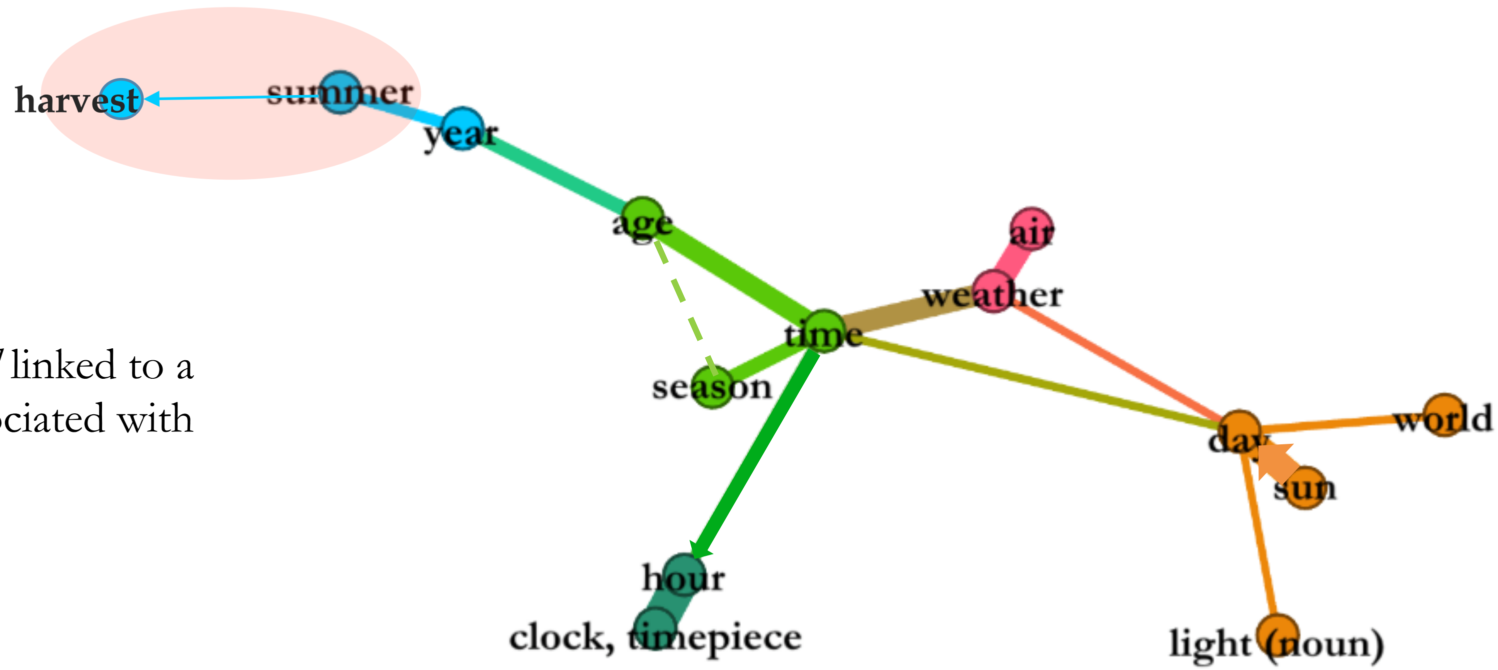


# The semantic extension of time-related lexemes

## Enriching the map

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

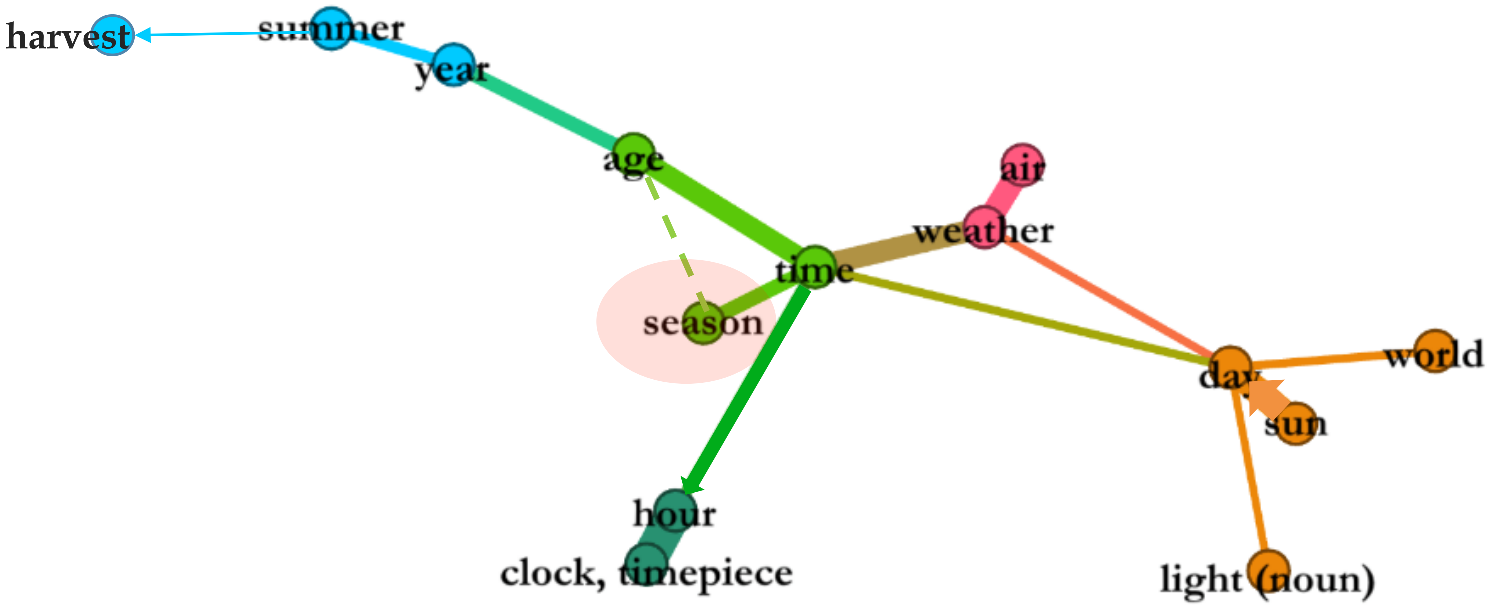
**Metonymy:**
A particular *period* linked to a salient *activity* associated with the period



# The semantic extension of time-related lexemes

## Enriching the map

- The material allows us to add new polysemy patterns and to provide a diachronic account
  - **SEASON**



# The semantic extension of time-related lexemes

## Enriching the map

5th c. BC

- **SEASON → YOUTH**

(11)

| | | | | | |
|---|---|---|---|---|---|
| *eph'* | *hoîs* | *prosḗkei* | *semnúsesthai* | *tḕn* | *pólin,* |
| SUPR | REL.DAT.PL.M | belong:PRS.3SG | exalt:INF.M/P | ART.ACC.SG.F | city:ACC.SG.F |

| | | | | |
|---|---|---|---|---|
| *eàn* | *kállei* | *kaì* | ***hṓrai*** | *dienegkóntes* |
| CONJ | beauty:DAT.SG.N | CONJ | **youth:DAT.SG.F** | differ:AOR.PTCP.NOM.PL.M |

| | | | | |
|---|---|---|---|---|
| *ekplḗksōsí* | *tinas* | *kaì* | *perimákhētoi* | *eks* |
| amaze:3AOR.SUBJ | INDEF.ACC.PL.F | CONJ | fought.for:NOM.PL.M | ELAT |

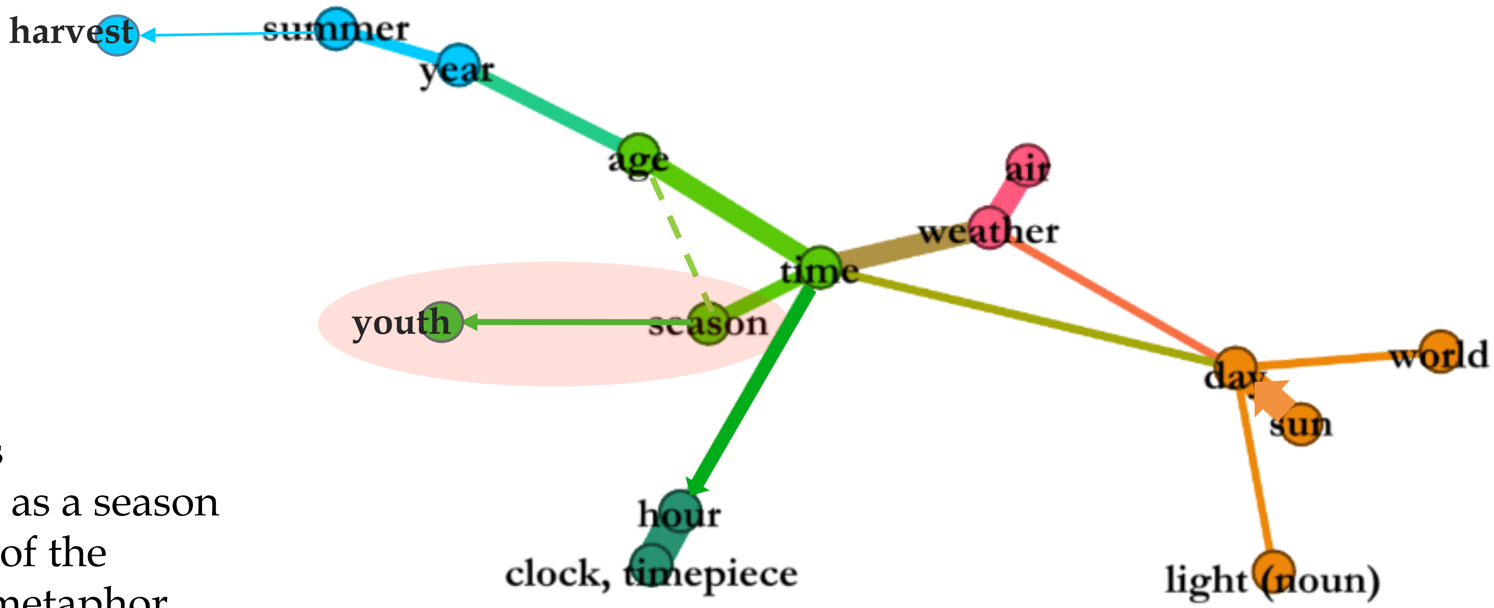
| | |
|---|---|
| *erōtos* | *génōntai* |
| love:GEN.SG.M | be.AOR.SUBJ.3PL.MID |

'of whom the city may well be proud, if by their surpassing beauty and **youthful charm** they infatuate one person or another' (Aeschines, *Against Timarchus* 1.134)

# The semantic extension of time-related lexemes

## Enriching the map

- The material allows us to add new polysemy patterns and to provide a diachronic account
  - **SEASON**




- The life cycle is conceptualized as a season along the lines of the LIFE IS A YEAR metaphor

(see Lakoff & Turner, 1989: 18; Sullivan, 2017: 387).

# The semantic extension of time-related lexemes

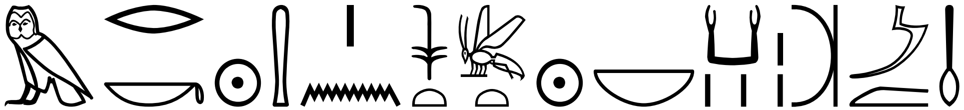
## Language-specific colexification patterns

- The material allows us to highlight unexpected pathways of change:
  - **From temporal proximity to spatial proximity**

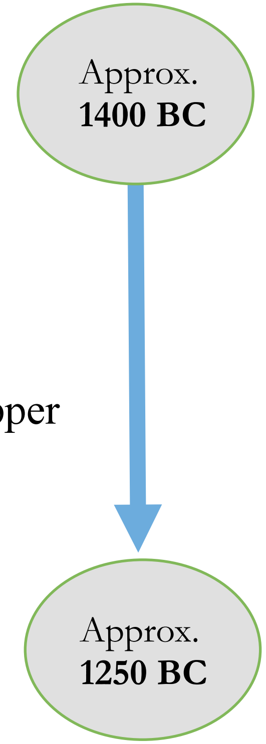
- **What about the TIME IS SPACE Metaphor?**
  - (Cross-linguistically Time to Space transfers are extremely rare; cf. French depuis; Haspelmath 1997)

# The semantic extension of time-related lexemes

## Ancient Egyptian



*Peasant*, B1, 103-104




| | | | | | |
|---|---|---|---|---|---|
| (12) | ***m*** | ***rk*** | *ḥm-f* | *nswt-bity* | *nb-kꜣw-rᶜ* |
| | **in** | **time** | Majesty-3SG.M | King of U. and L. Egypt | Nebkaure |

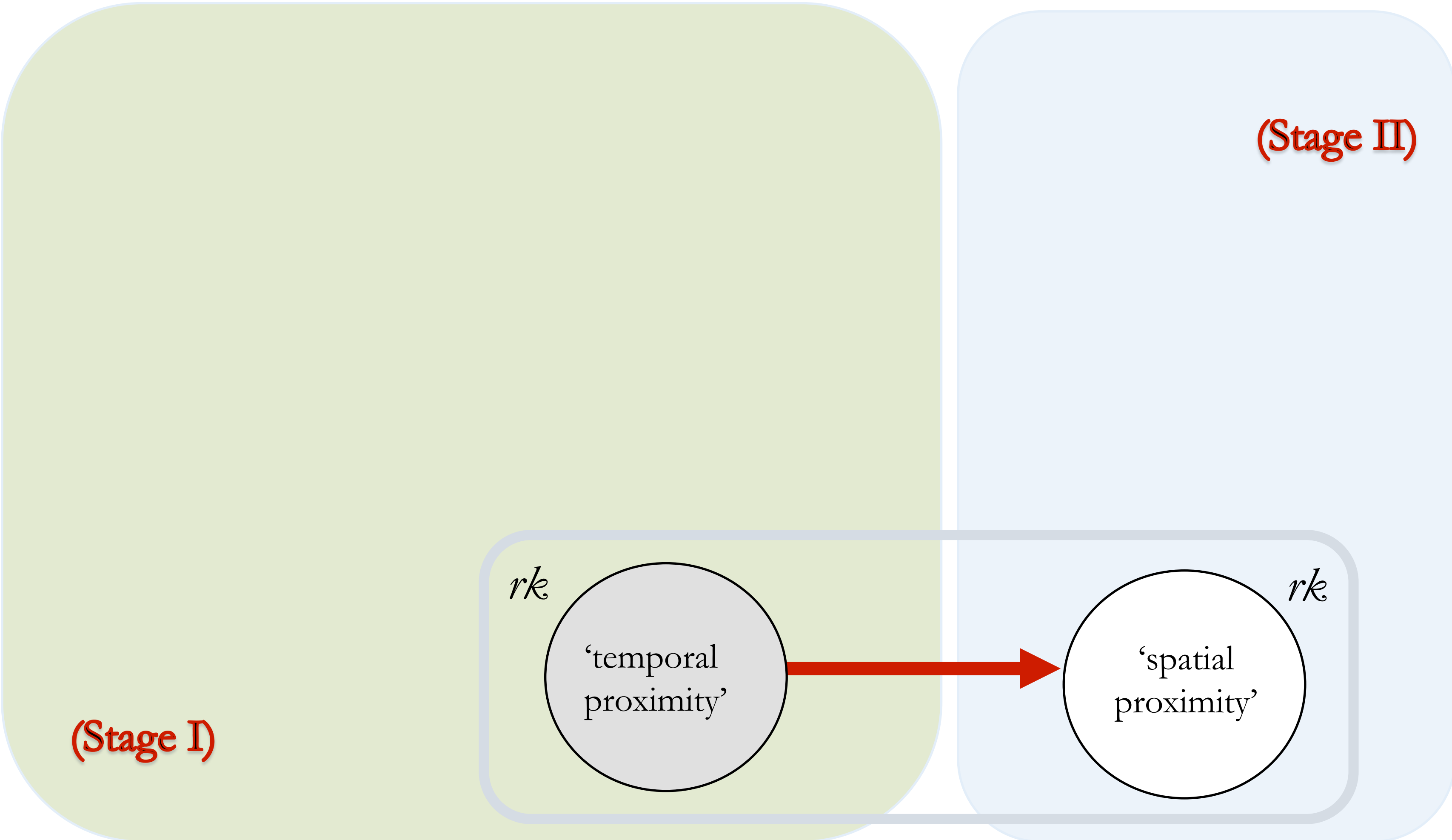
‘(Now, the peasant spoke these word) **during the time** of his Majesty, the King of Upper and Lower Egypt, Nebkaure (the justified)’ (= Parkinson 1991: 19)

| | | | | | | |
|---|---|---|---|---|---|---|
| (13) | *sbty* | *ḏr* | ***m*** | ***rk*** | *mšᶜ-f* | (= K*RI* II, 6,8) |
| | rampart | strong | **in** | **time** | army-3SG.M | |

(talking about the King, who is described as)
‘A strong rampart around his army, (their shield in the day of fighting)’

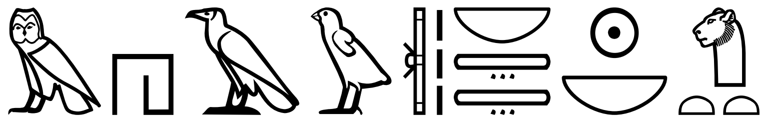
# The semantic extension of time-related lexemes

## Ancient Egyptian



# The semantic extension of time-related lexemes

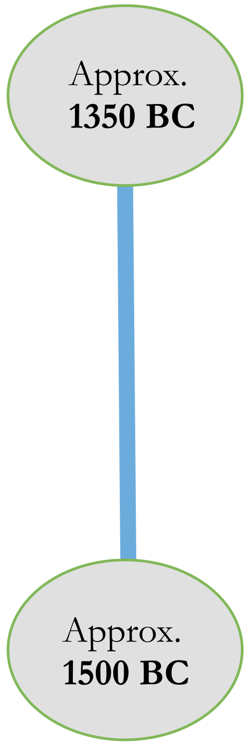
## Ancient Egyptian

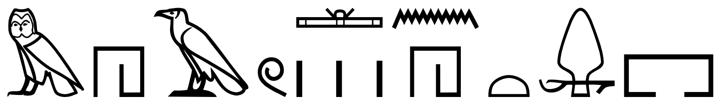


*Biography of Ahmose*, 5

| (14) | *m* | *hꜣw* | *nb* | *tꜣ-wj* | *nb-pḥ.tj-r*$^{c}$ |
|---|---|---|---|---|---|
| | 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)



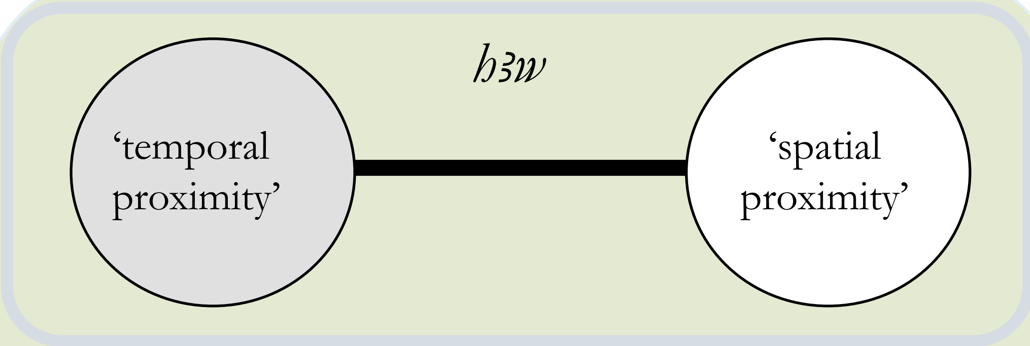
*Sinuhe*, B8

| (15) | *m* | *hꜣw* | *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)

# The semantic extension of time-related lexemes

## Ancient Egyptian

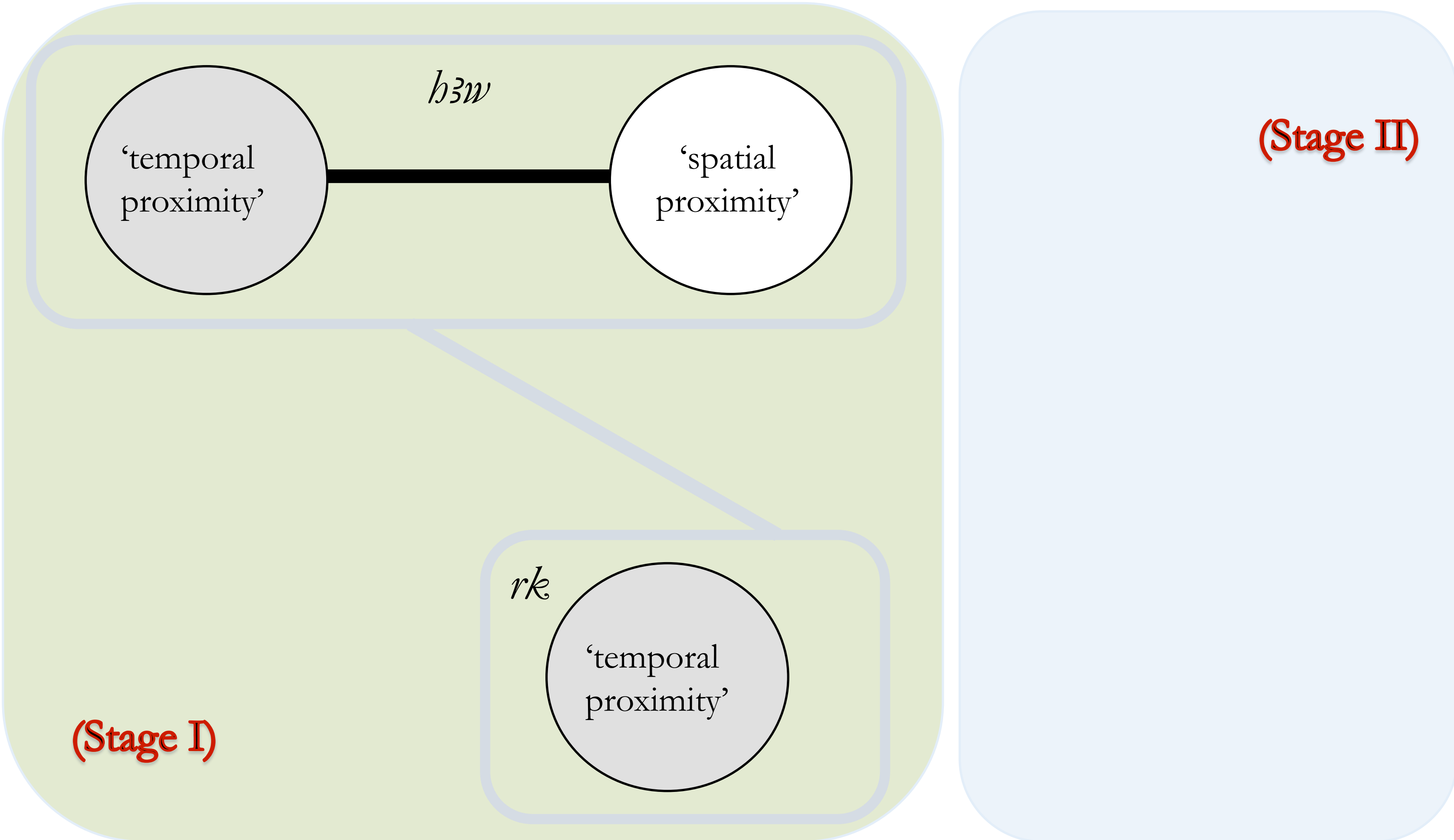



(Stage II)

(Stage I)

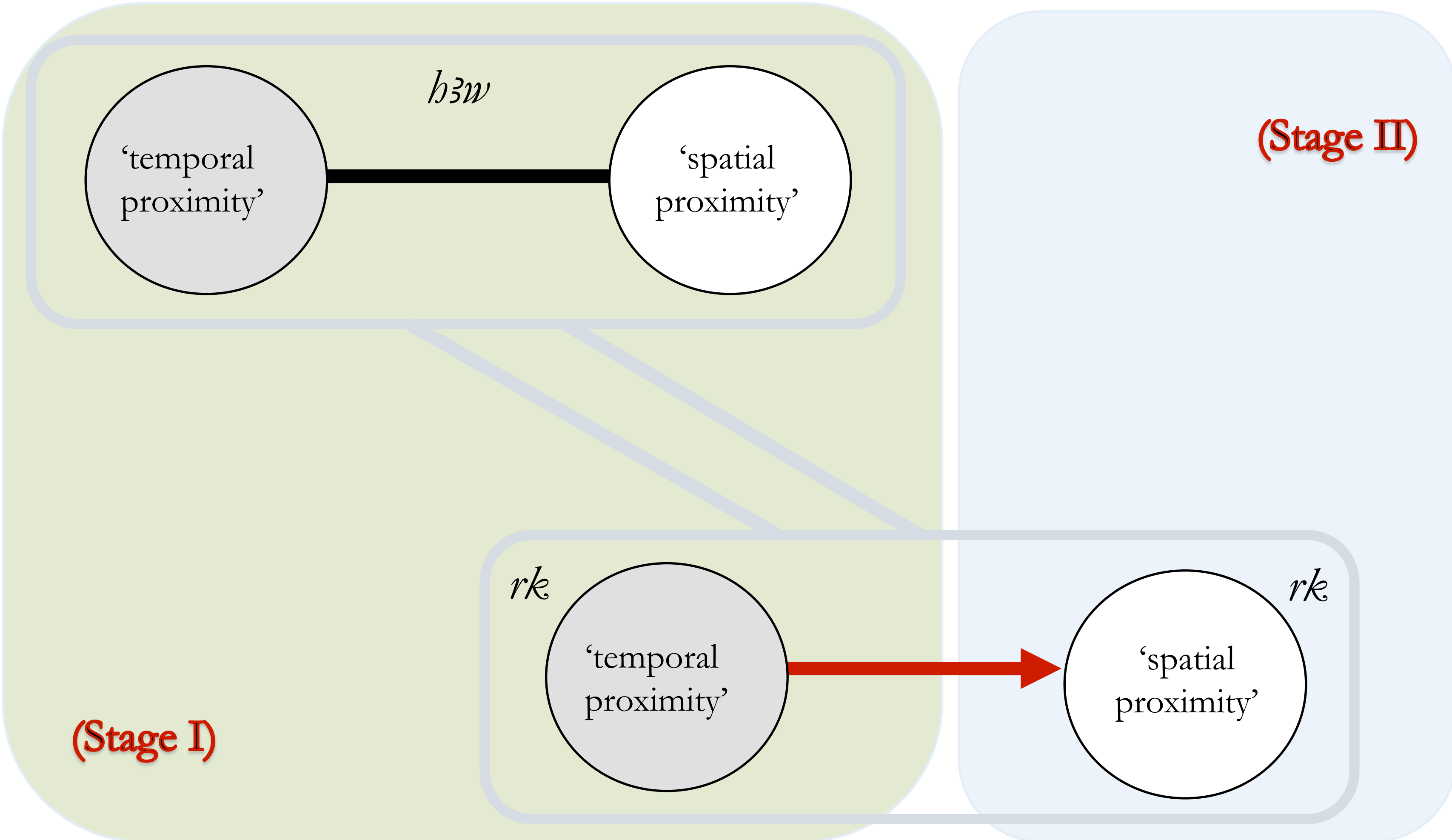
# The semantic extension of time-related lexemes

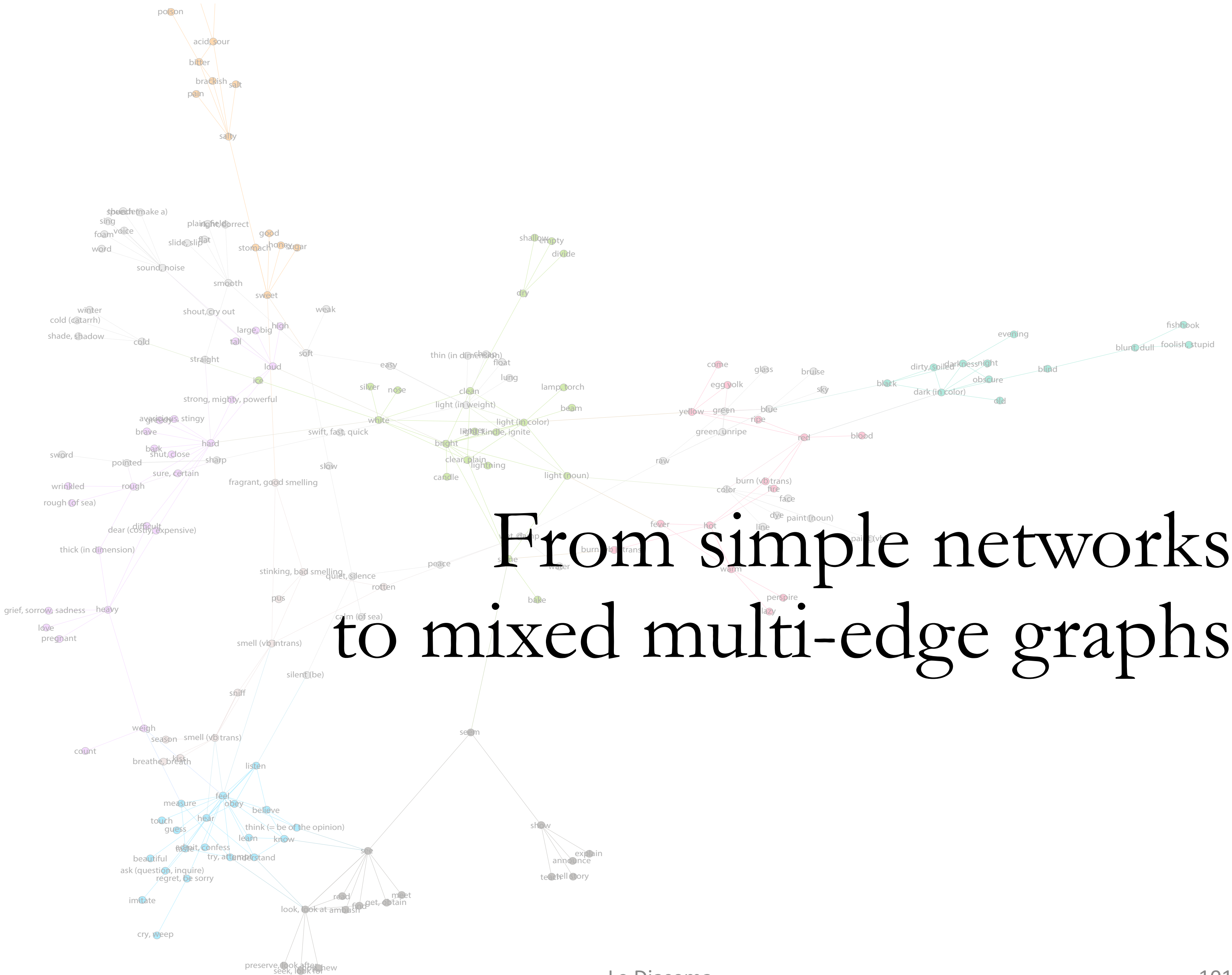
## Ancient Egyptian



# The semantic extension of time-related lexemes

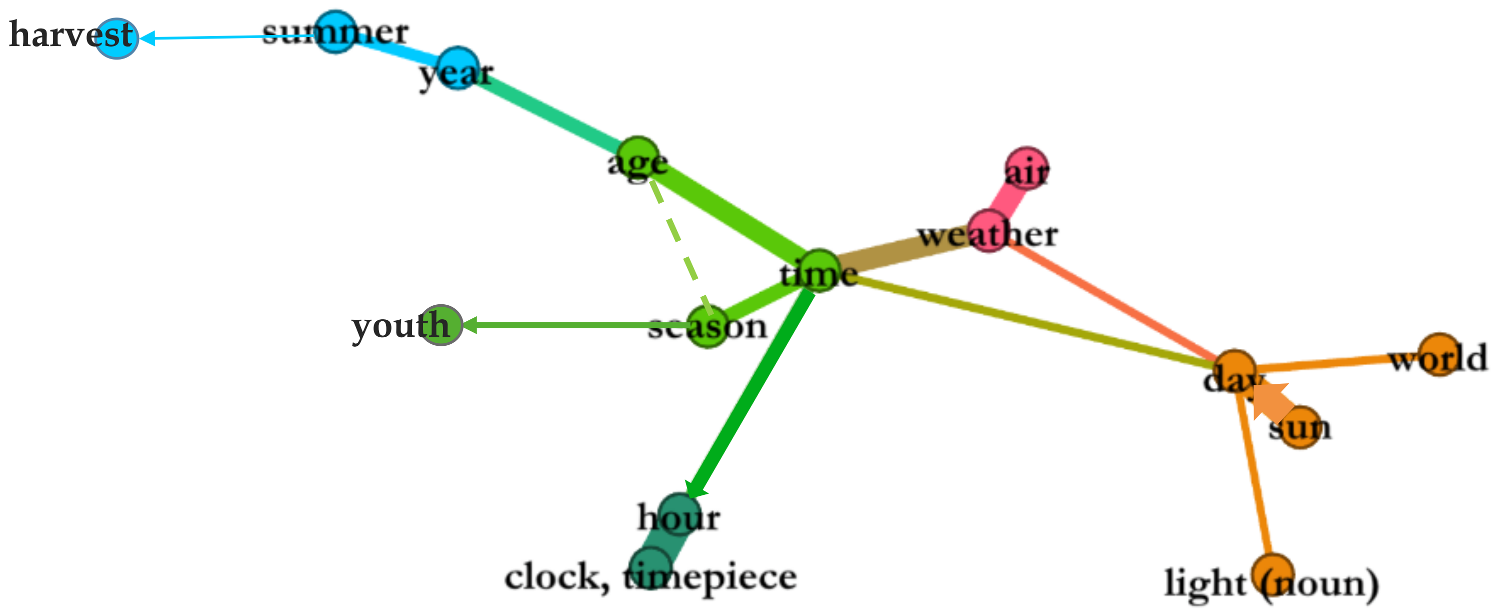
## Ancient Egyptian



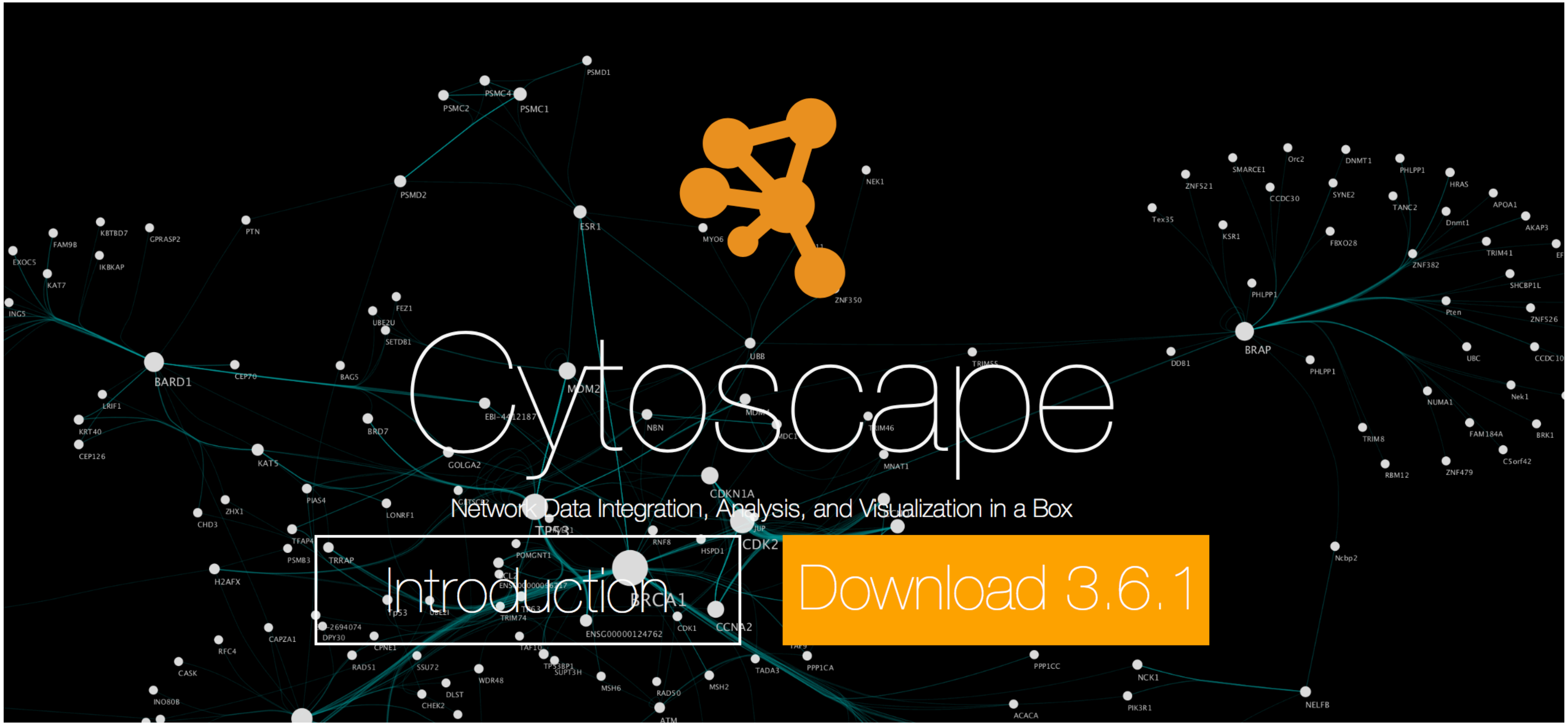


# From simple networks to mixed multi-edge graphs

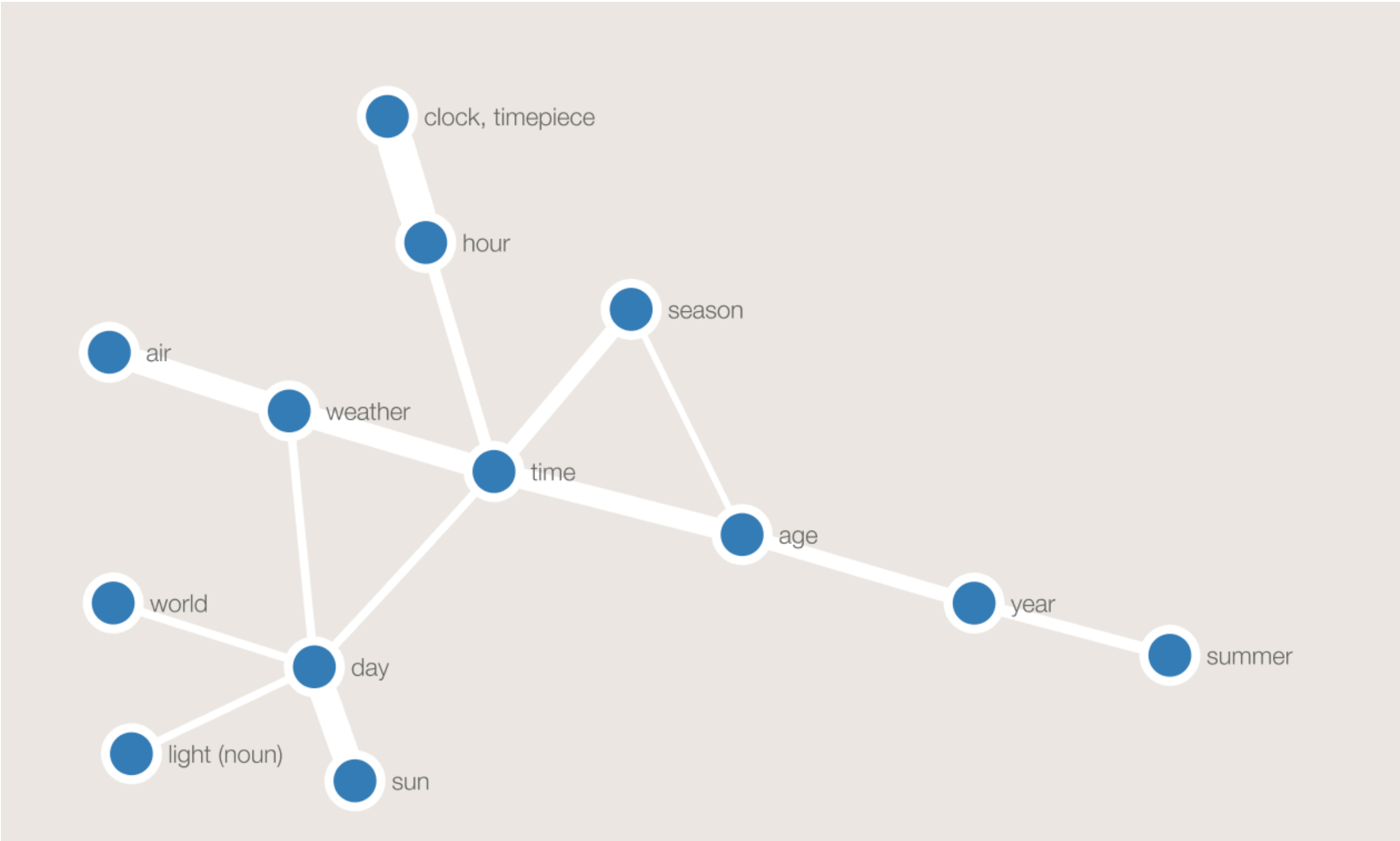
# From simple networks to mixed multi-edge graphs



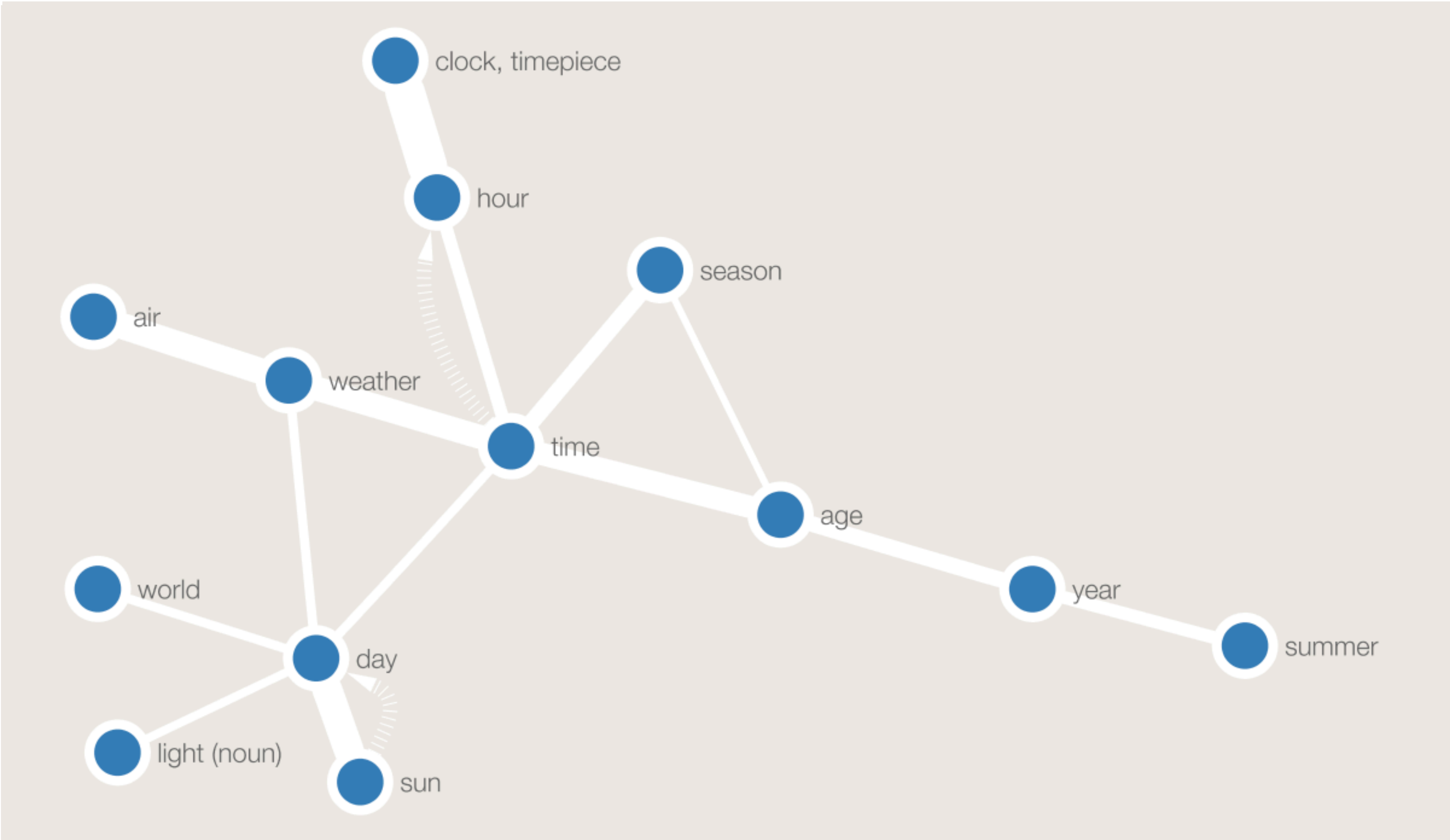
# From simple networks to mixed multi-edge graphs



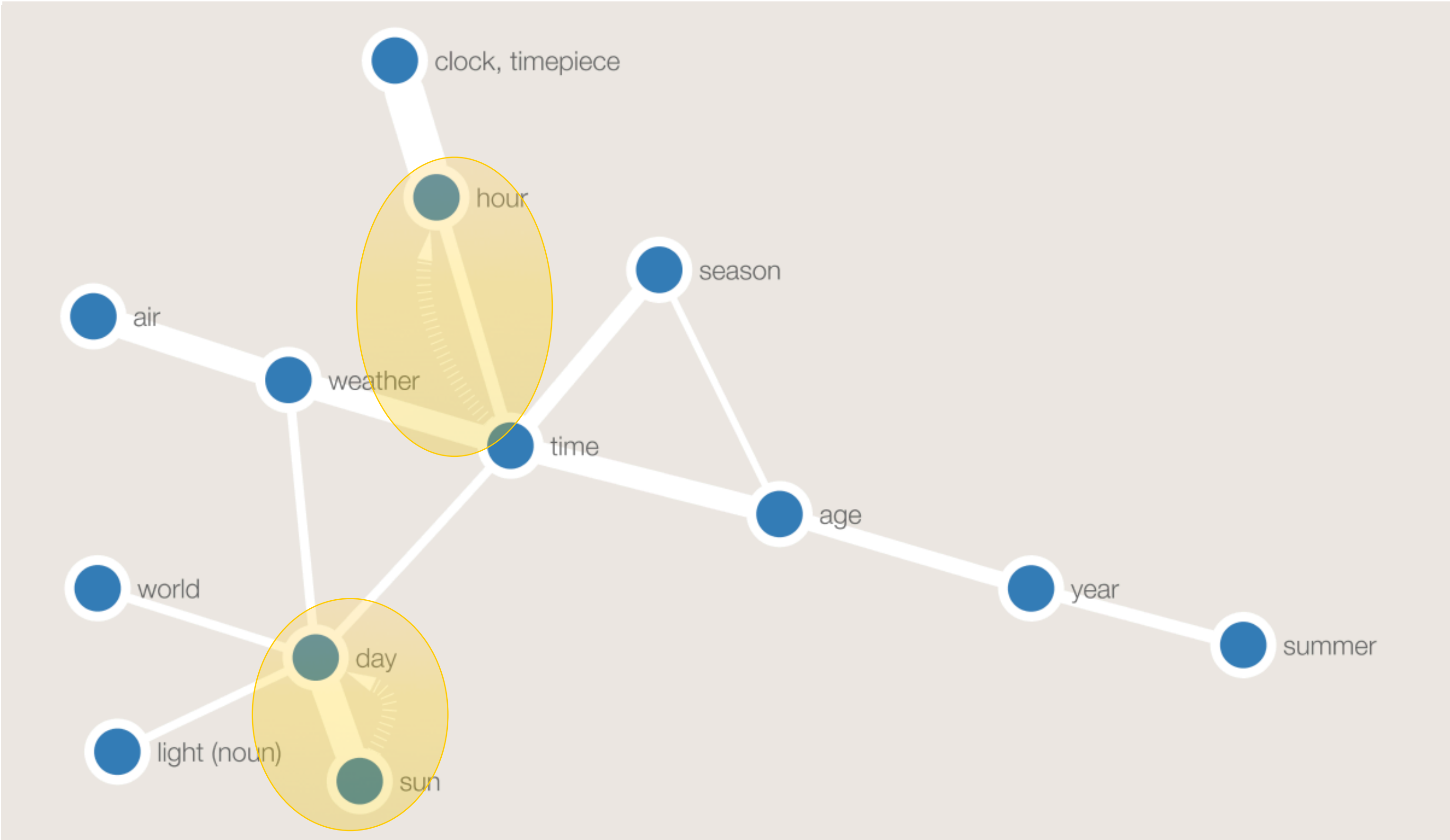
# From simple networks to mixed multi-edge graphs



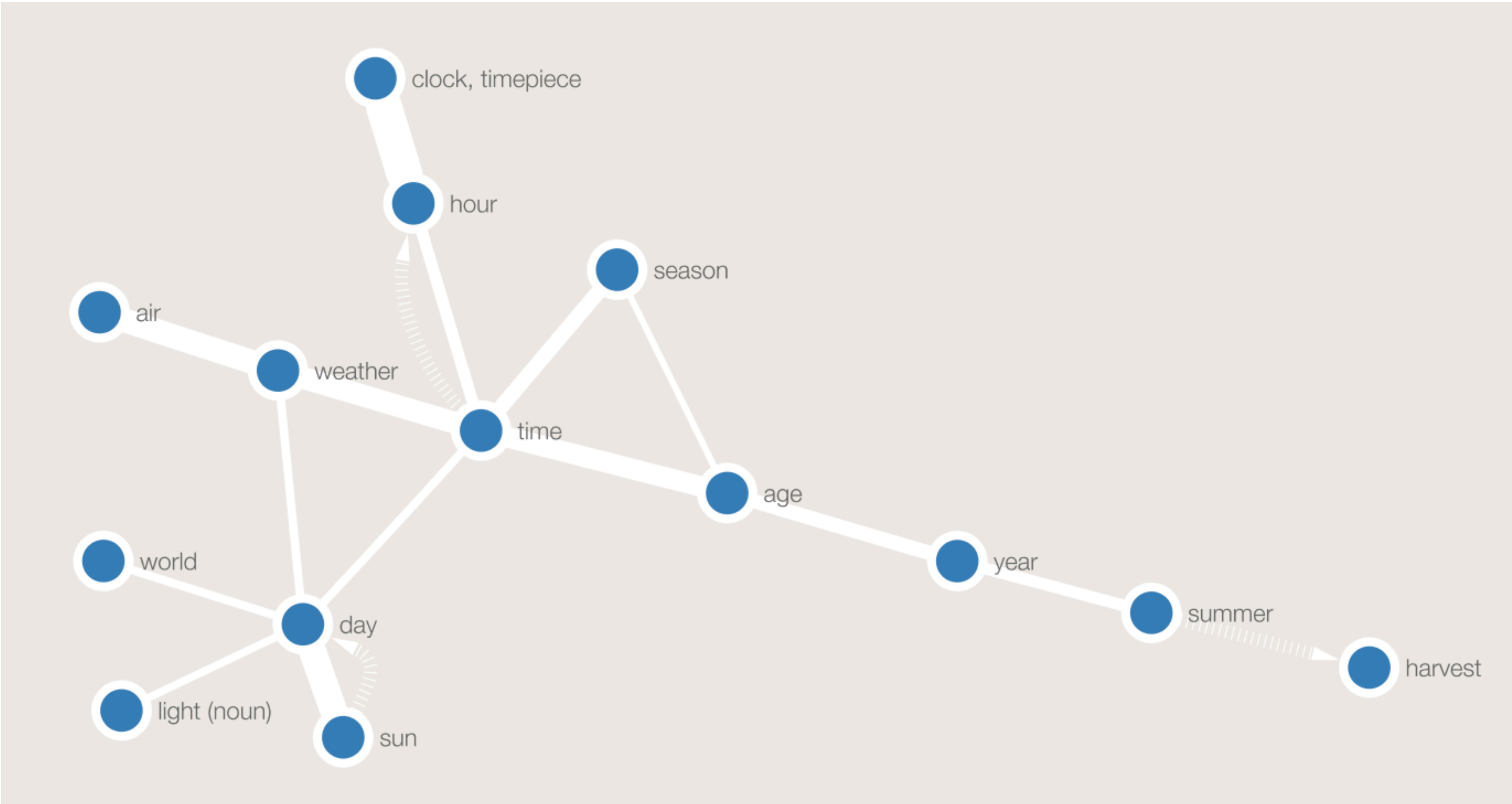
# From simple networks to mixed multi-edge graphs



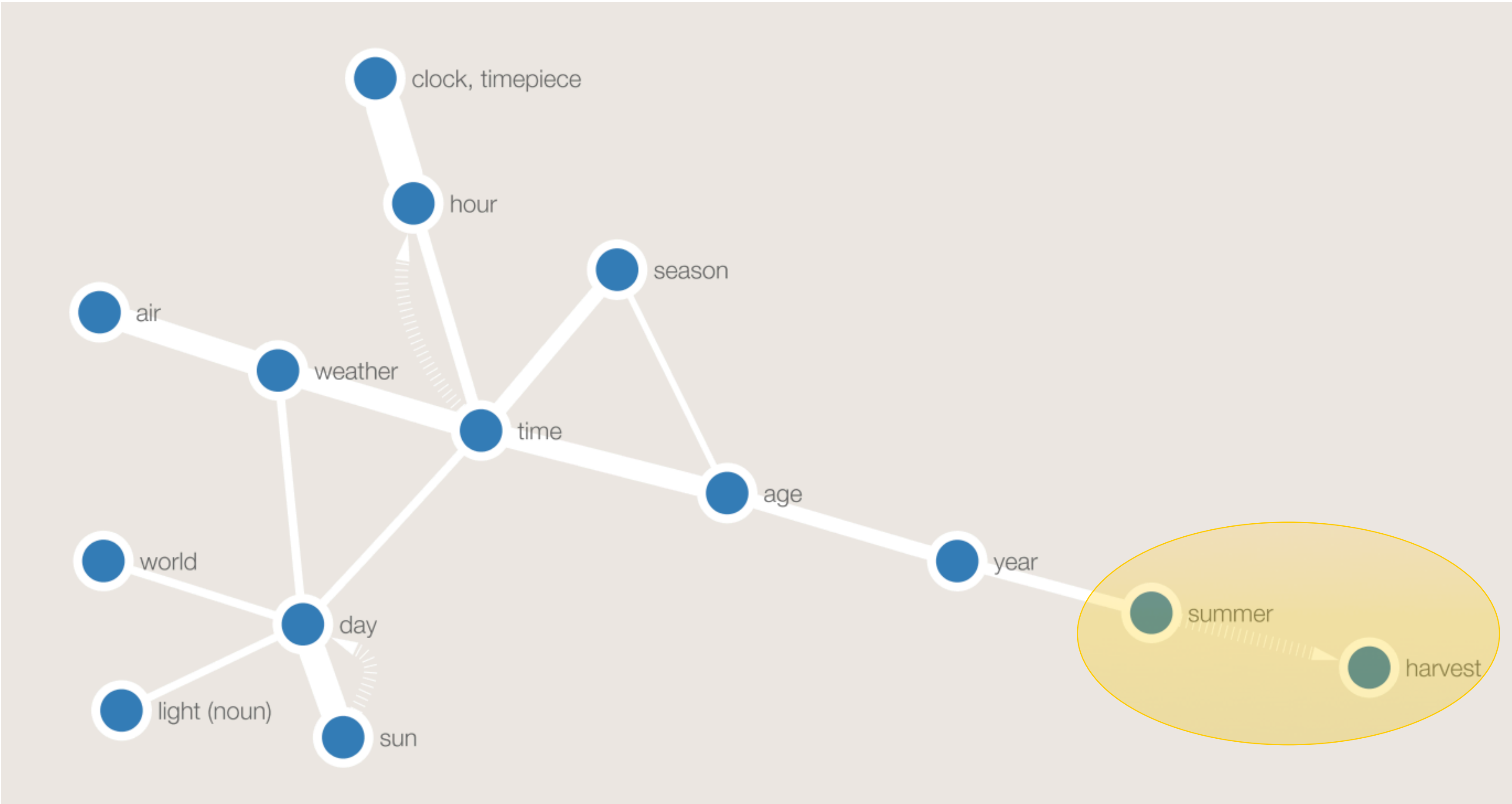
# From simple networks to mixed multi-edge graphs



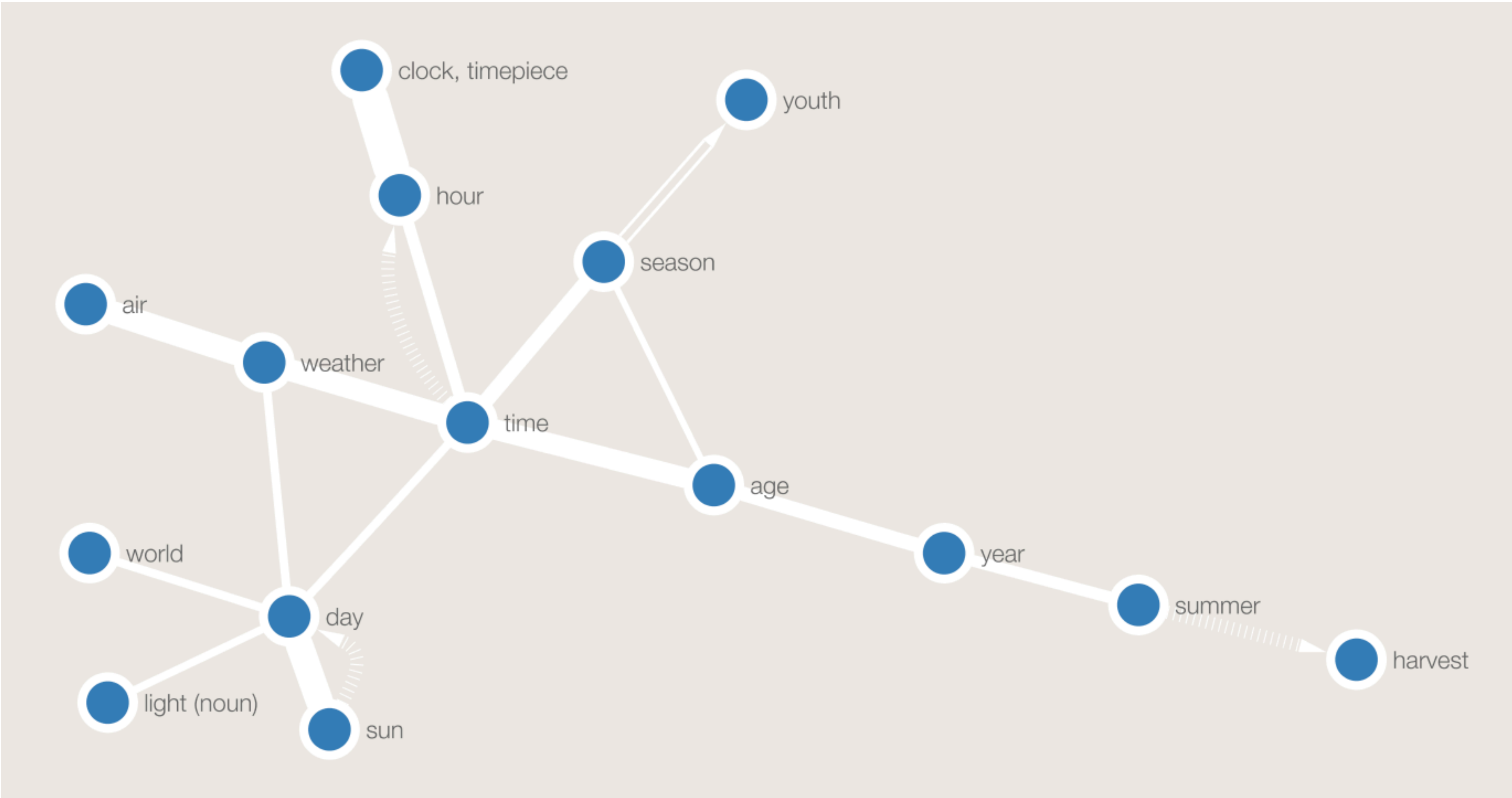
# From simple networks to mixed multi-edge graphs



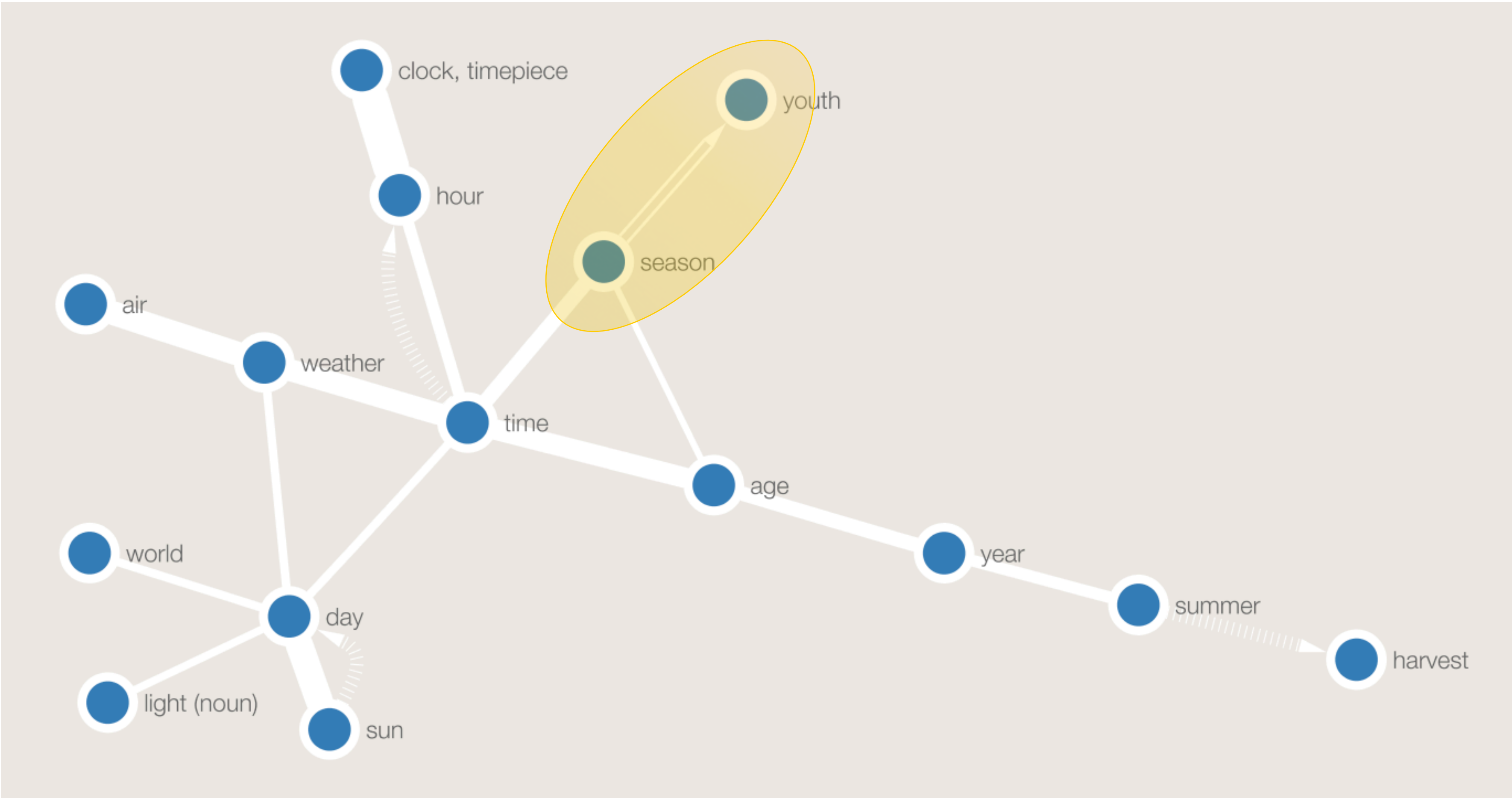
# From simple networks to mixed multi-edge graphs



# From simple networks to mixed multi-edge graphs



# From simple networks to mixed multi-edge graphs



# From simple networks to mixed multi-edge graphs

**Can we infer directionalities automatically?**

# From simple networks to mixed multi-edge graphs

- 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 |

# From simple networks to mixed multi-edge graphs

- 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.

# From simple networks to mixed multi-edge graphs

- 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)

# From simple networks to mixed multi-edge graphs

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

# From simple networks to mixed multi-edge graphs

- Expand the lexical matrix so as to include information about diachrony
- Generate the graph with the weighted version of 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.

# From simple networks to mixed multi-edge graphs

INPUT
(diachronic lexical matrix)

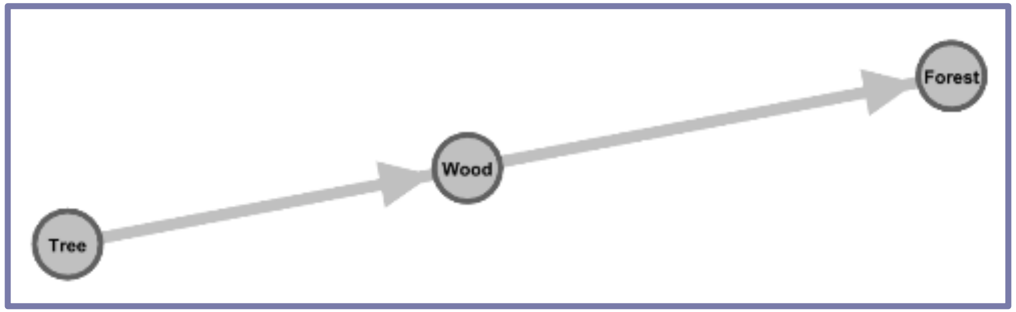
| *Source of constraint* | *Constraint name* | *Constraint Time* | Sense_1 Tree | Sense_2 Wood | Sense_3 Forest |
|---|---|---|---|---|---|
| Language_1 | Word_1 | 0 | 1 | 0 | 0 |
| Language_1 | Word_1 | 1 | 1 | 1 | 0 |
| Language_2 | Word_1 | 0 | 1 | 0 | 0 |
| Language_2 | Word_2 | 0 | 0 | 1 | 0 |
|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

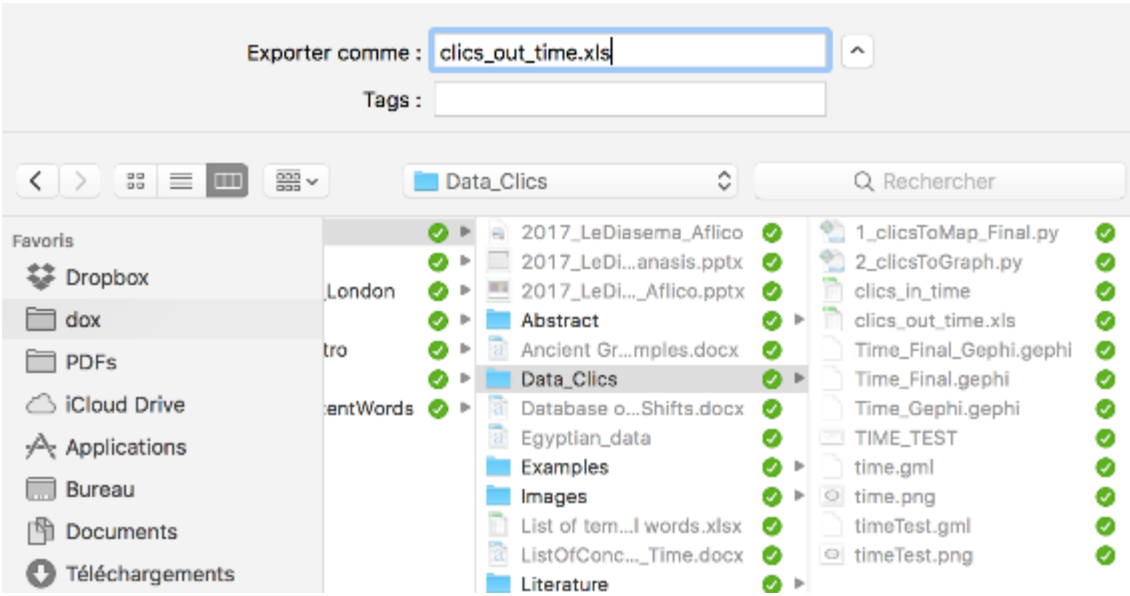


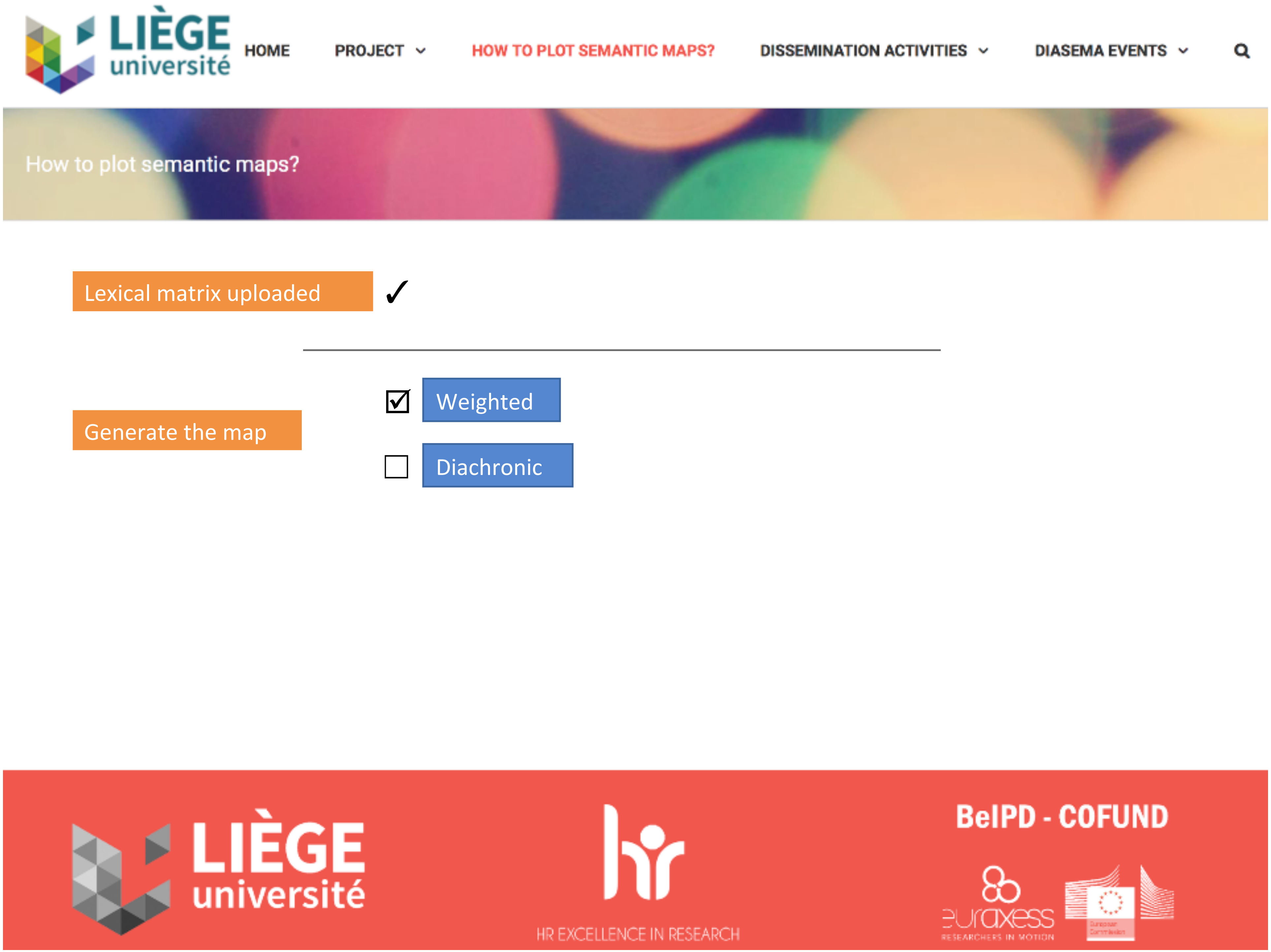
How to plot semantic maps?

Upload the lexical matrix (.xls file)

# How to plot semantic maps?

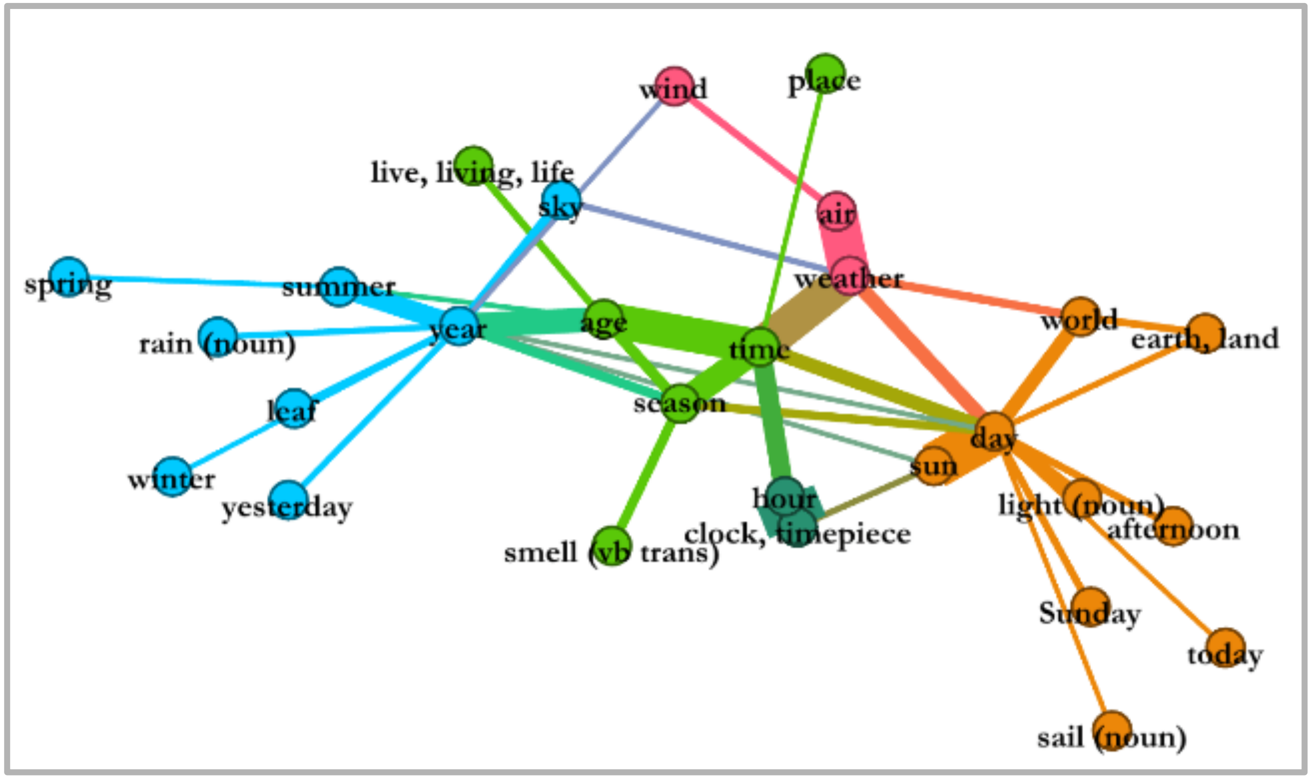
## Upload the lexical matrix (.xls file)





LIÈGE université
HOME
PROJECT
HOW TO PLOT SEMANTIC MAPS?
DISSEMINATION ACTIVITIES
DIASEMA EVENTS
How to plot semantic maps?
Lexical matrix uploaded ✓
Generate the map
☑ Weighted
☐ Diachronic
LIÈGE université
HR EXCELLENCE IN RESEARCH
BelPD - COFUND
euraxess
RESEARCHERS IN MOTION

# How to plot semantic maps?



# Conclusions

Thanks!

s.polis@uliege.be
athanasios.georgakopoulos@uliege.be