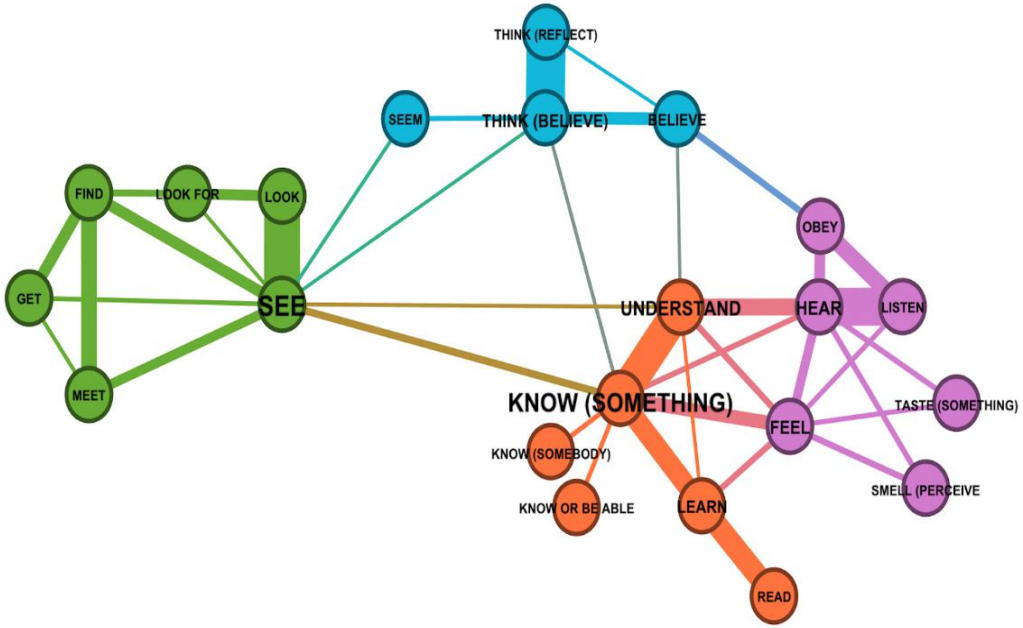
“Πολυσημία στα σημασιολογικά πεδία της αντίληψης και της νόησης: τυπολογικά και τοπολογικά/γεωγραφικά σχήματα”

Θανάσης Γεωργακόπουλος

(Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης/ Ελληνικό Ανοιχτό Πανεπιστήμιο)



Ερευνητικά Σεμινάρια Τομέα Γλωσσολογίας

(Τμήμα Φιλολογίας, ΕΚΠΑ)

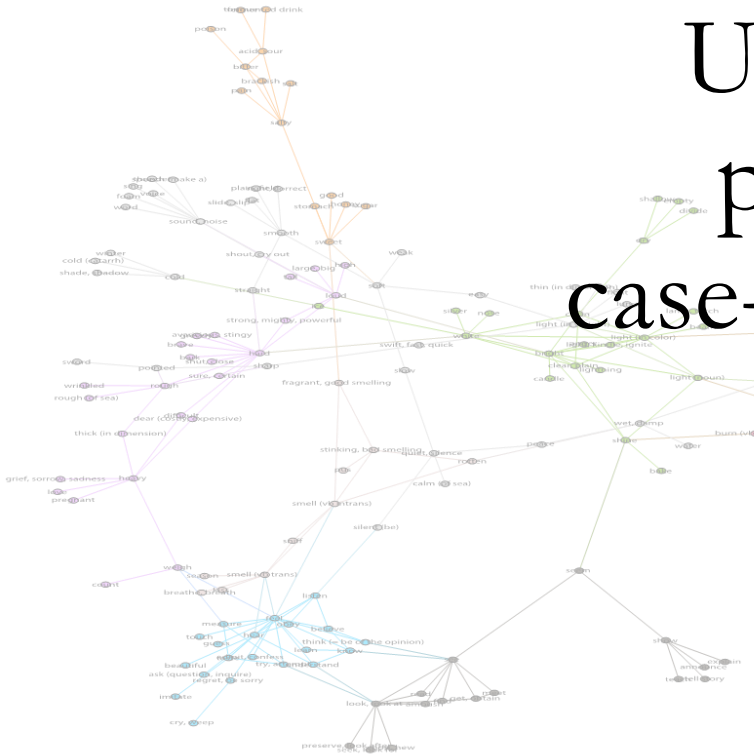
Εργαστήριο Φωνητικής & Υπολογιστικής Γλωσσολογίας

31 Μαρτίου, 2021

ARISTOTLE UNIVERSITY OF THESSALONIKI

ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

Georgakopoulos Thanasis, Eitan Grossman, Dmitry Nikolaev, and Stéphane Polis. 2021-forthcoming. Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain. *Linguistic Typology*.



Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain

Outline of the talk

- Introduction
- The semantic domains: perception and cognition

- Main methods:
  - *Semantic maps*
  - *Colexification networks*

- Associations in the domain of perception and cognition (universal patterns)
- Macro-areal patterns in the domains of perception and cognition

- Discussion

Introduction

Focus:

- On coexpression patterns

(1) Can you ***see*** the bird in that tree?
(2) I just can't ***see*** your point.

(Examples from Princeton WordNet of English)

Colexification

"The capacity, for two senses, to be lexified by the same lexeme in synchrony" (François, 2008: 171).

Introduction

Focus:

- On coexpression patterns

(1) Can you ***see*** the bird in that tree?
(2) I just can't ***see*** your point.

Main objective:

- To investigate universal and areal structures in the lexicon as manifested by coexpression patterns in the semantic domains of perception and cognition

Main questions:

a) To what extent do bottom-up methods using language samples of different sizes match (or challenge) the results of case-studies conducted by experts on individual languages
b) To what extent do these methods reveal new universal or area-specific generalizations about the organization of lexicons?

Introduction

- Generalizations about the cross-linguistic organization of the lexicon are not easily or straightforwardly identified
  - Access to large datasets is needed
- In the past, the availability of such datasets was rather limited
  - Result: the number of languages of a typical lexico-typological study ranged from 10 to 50
    (Koptjevskaja-Tamm, Rakhilina & Vanhove, 2015)
    - Exception: studies that rely on massively parallel texts
      (Wälchli, 2010; Wälchli & Cysouw, 2012; Östling, 2016; Wälchli, 2016)

Introduction

- Nowadays: increasing availability of resources that contain a large amount of lexical information makes large-scale typological studies on the lexicon possible
- Large lexical databases such as CLICS and ASJP have recently been used in order to investigate areal factors in lexical typology

(Gast & Koptjevskaja-Tamm, 2018)

Database of Cross-Linguistic Colexifications

CLICS³ is an online database of colexifications (polysemies or homophonies) in currently 3156 language varieties of the world.

The database of the Automated Similarity Judgment Program (ASJP) aims to contain 40-item word lists of all the world's languages. A lexical distance can be obtained by comparing the word lists, which is useful, for instance, for classifying a language group and for inferring its age of divergence.

Introduction

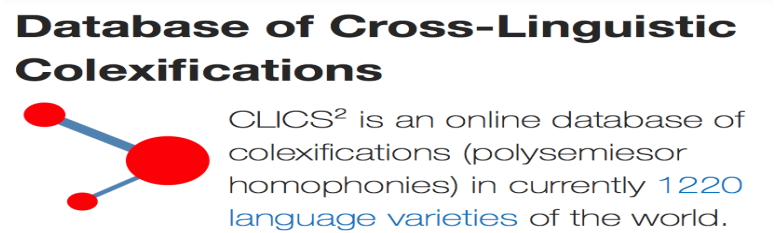
Through different exploratory strategies

- Weighted semantic maps
- Colexification networks
- Correlation plots
- Dimensionality reduction techniques **(! in the paper but not today)**

Data from two different datasets

a) Vanhove's (2008) study of verbs of perception and cognition

b) Database of Cross-Linguistic Colexifications (List et al. 2018)



https://clics.clld.org/

Introduction

Steps

1. Consider a lexical field in its entirety, first showing how this field is structured and unveiling recurrent cross-linguistic colexifications

   - Tools: (a) weighted semantic maps

2. Consider patterns that are specific to particular macro-areas and to smaller regions within macro-areas.

   - Tools: (a) colexification networks
     (b) correlation plots
     (c) dimensionality reduction techniques
     (! in the paper but not today)

The semantic domains:

- Perception & Cognition

Why perception & cognition?

- Both domains are central to human experience

  "Every language has a way of talking about seeing, hearing, smelling, tasting and touching" and "[e]very language has a way of speaking about how one knows"

  (Aikhenvald & Storch, 2013: 1)

- Meanings belonging to both domains appear in the main collections of basic concepts (e.g., Swadesh list, the Leipzig-Jakarta list, the IDS)

- The relevant literature reports on:
  - Within-domain meaning extensions
  - Between-domain semantic connections
  - Both universal and culture-specific patterns

Verbs of perception & cognition

- **Starting point:** SIGHT and HEARING
  - Universally more prominent within the domain of sensory modalities

    (Viberg, 1984; Evans & Wilkins, 2000; Vanhove, 2008; Levinson & Majid, 2014; San Roque et al., 2015; Winter et al., 2018)
  - Cross-linguistically more closely connected to mental perception than smell, taste, and touch

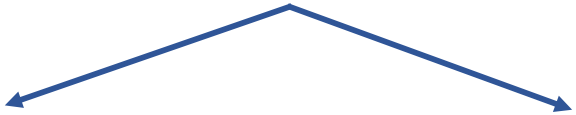
    (Evans & Wilkins, 2000; Sweetser, 1990)

	Controlled activities	Non-controlled experiences
SIGHT	look (at)	see
HEARING	listen (to)	hear

Table 1. Sample English expressions for visual and auditory situations

Verbs of perception & cognition

Semantic extensions



Intrafield *(= Intradomain)*
(senses: same semantic field)

Interfield *(= Interdomain/ Transfield)*
(senses: different semantic field)

(based on Wilkins 1996: 274; cf. Matisoff 1978)

Verbs of perception & cognition

Semantic extensions

Intrafield *(= Intradomain)*
(senses: same semantic field)

Interfield *(= Interdomain/ Transfield)*
(senses: different semantic field)

sight > hearing > touch > { smell / taste }

Figure 1. The sense-modality hierarchy for perception verbs (Viberg 1984: 136)

Markedness criteria (selection)

(a) structural coding
(b) textual frequency & cross-linguistic frequency
(see Croft, 2003: 91ff):

Verbs of perception & cognition

Semantic extensions

Intrafield *(= Intradomain)*
(senses: same semantic field)

Interfield *(= Interdomain/ Transfield)*
(senses: different semantic field)

sight > hearing > touch > { smell / taste }

SIGHT	HEARING
ɲaŋ	ɲaŋ-an

Table 2. SEE and HEAR lexemes in Djaru (Pama-Nyungan language in Australia) (Viberg 2001: 1297)

Verbs of perception & cognition

Semantic extensions

Intrafield *(= Intradomain)*
(senses: same semantic field)

Interfield *(= Interdomain/ Transfield)*
(senses: different semantic field)

sight > hearing > touch > { smell / taste }

Figure 1. The sense-modality hierarchy for perception verbs (Viberg 1984: 136)

- cf. Roque et al. (2015) and Winter et al. 2018:
  - the visual modality dominates; **but**
  - the ranking of the other modalities varies

Verbs of perception & cognition

Semantic extensions

Intrafield *(= Intradomain)*
(senses: same semantic field)

Interfield *(= Interdomain/ Transfield)*
(senses: different semantic field)

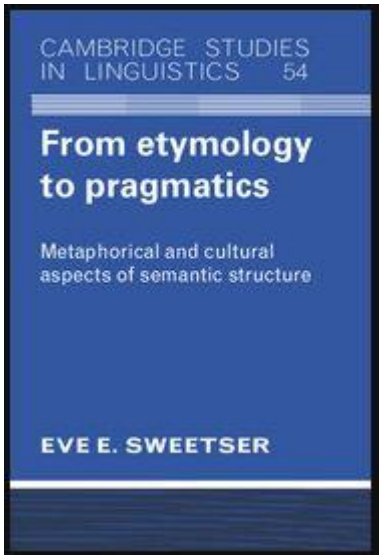
Interfield extensions

Mind-as-body-Metaphor:

The internal self is understood in terms of the bodily external self
(Sweetser 1990: 45)

- **Common cross-linguistically (if not universal)**: the connection between VISION and KNOWLEDGE
(Sweetser 1990: 45)

  - SEEING IS KNOWING
  - HEARING IS UNDERSTANDING



"It would be a novelty for a verb meaning 'hear' to develop a usage meaning 'know' rather than 'understand,' whereas such a usage is common for verbs meanings 'see'"
(Sweetser, 1990: 43)

Interfield extensions

In Australian languages:
(Evans & Wilkins 2000)

Cognitive verbs > 'hear'

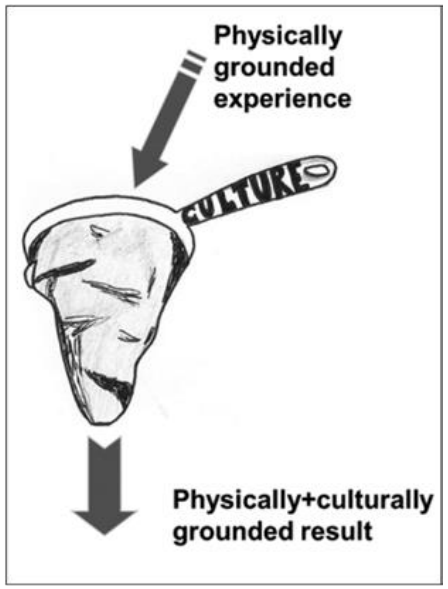


Figure 2. The culture sieve
(Ibarretxe-Antuñano 2013)

A general metaphor:

COGNITION IS PERCEPTION

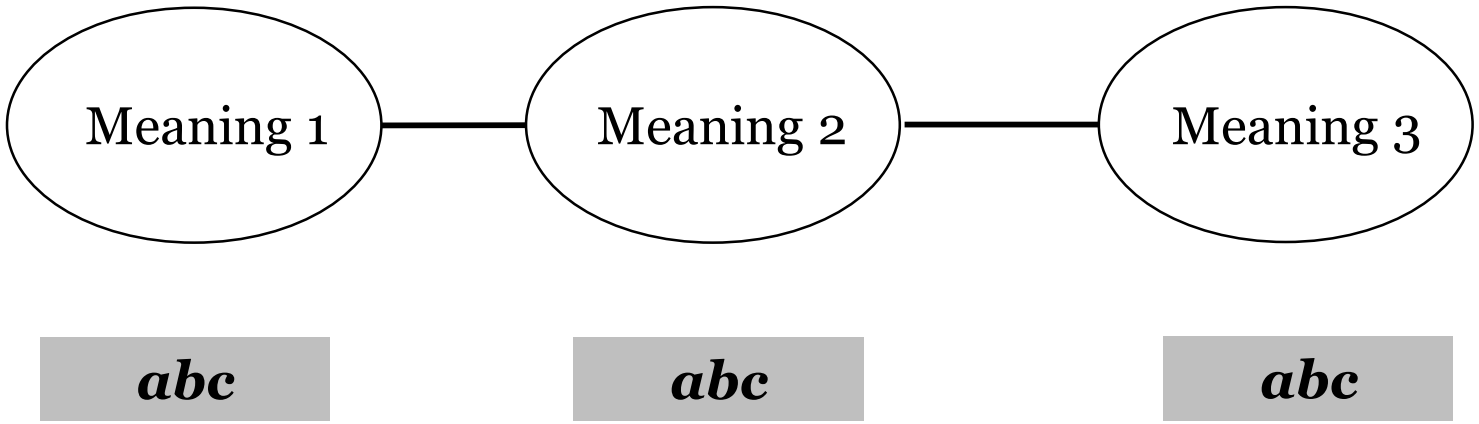
COGNITION IS SEEING
in English

COGNITION IS SMELLING
in Jahai

COGNITION IS HEARING
in Warluwarra

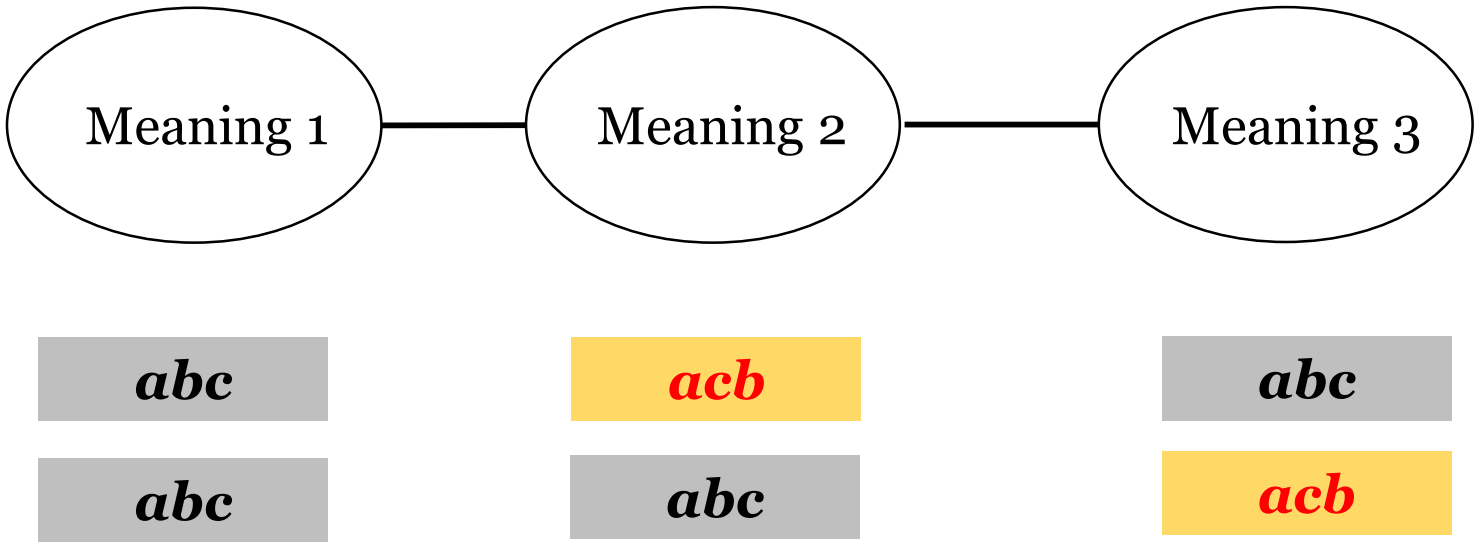
Semantic maps
vs.
Colexification networks

Introducing semantic maps

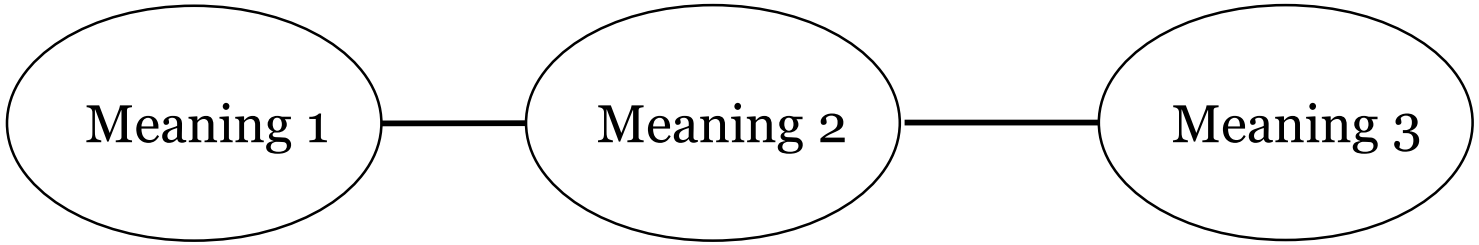


abc colexifies/ co-expresses **Meaning 1**, **Meaning 2**, and **Meaning 3**

Introducing semantic maps



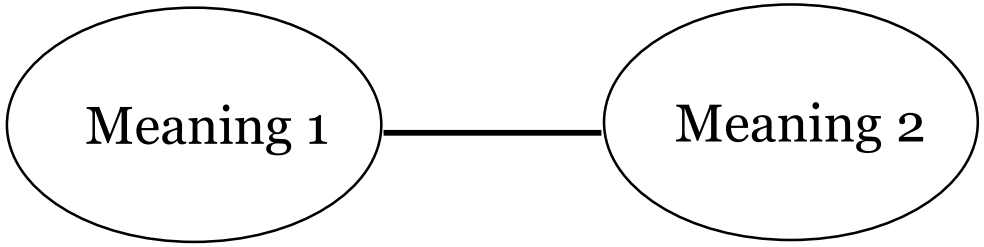
Introducing semantic maps



abc	*acb*	*abc*	⊘
abc	*abc*	*acb*	✓
abc	*acb*	*acb*	✓
abc	*cba*	*acb*	✓

Introducing semantic maps

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



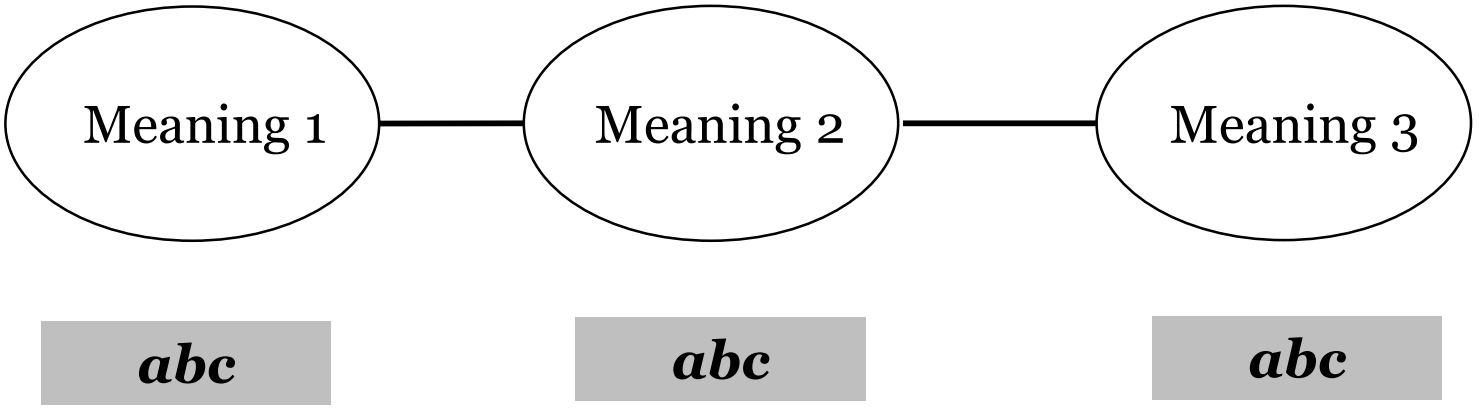
Form ***abc*** ***abc***

Introducing semantic maps

The semantic map connectivity hypothesis:

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

(Croft, 2001: 96)

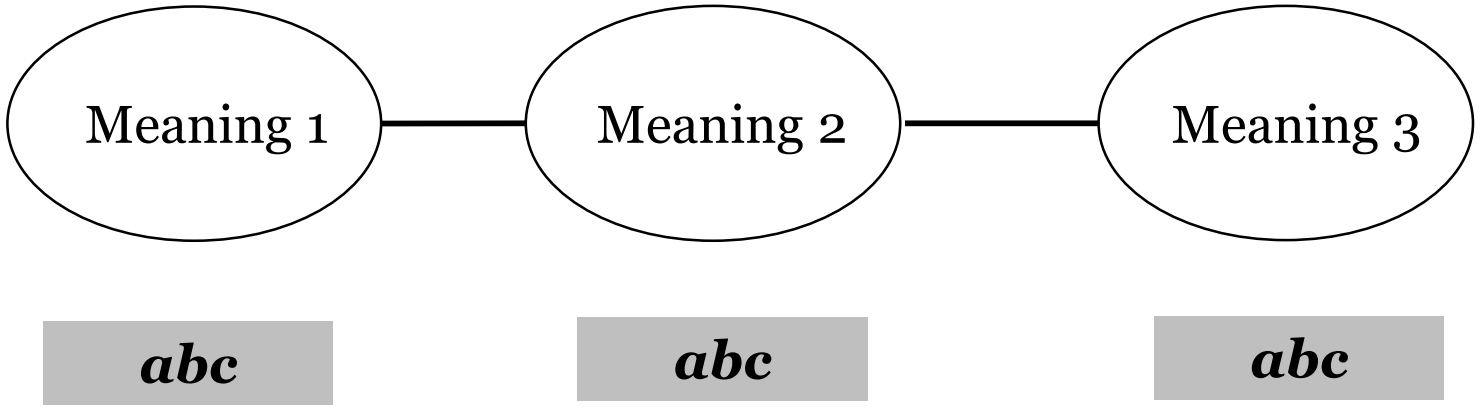


Introducing semantic maps

The economy principle:

- Two meanings are connected by an edge if and only if they are not already part of a subgraph of meanings expressed by a single polysemic item in a given language of the sample

(Georgakopoulos & Polis, 2018)



Introducing semantic maps

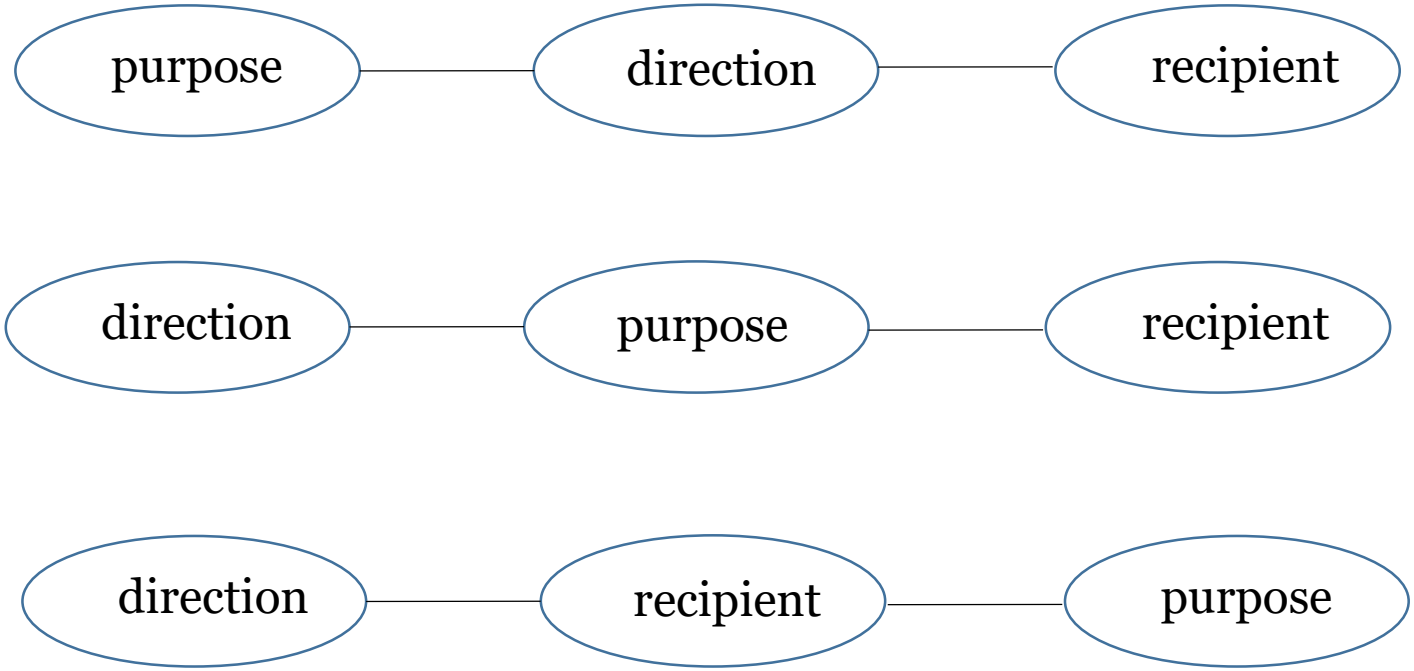
The English preposition *to*:

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

(Haspelmath, 2003)

Introducing semantic maps

How to arrange the different meanings?



(Haspelmath, 2003)

Introducing semantic maps

German:

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

  ≠
- '**Recipient**': Ich gebe ***dir*** das Buch

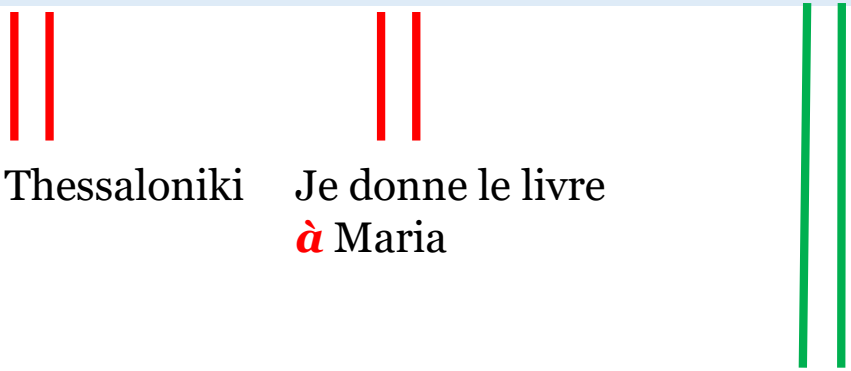
direction = purpose ≠ recipient

(Haspelmath, 2003)

Introducing semantic maps

French:

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



Je vais ***à*** Thessaloniki

Je donne le livre ***à*** Maria

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

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

(Haspelmath, 2003)

Introducing semantic maps

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

Introducing semantic maps

How to arrange the different meanings?

purpose ——— direction ——— recipient

direction ——— purpose ——— recipient

direction ——— recipient ——— purpose

Introducing semantic maps

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

Introducing semantic maps

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

direction ——— recipient ——— purpose 

Introducing semantic maps

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

direction ——— recipient ——— purpose 

direction ——— purpose ——— recipient

Introducing semantic maps

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

direction ——— recipient ——— purpose 

direction ——— purpose ——— recipient 

Introducing semantic maps

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

direction ——— recipient ——— purpose 

direction ——— purpose ——— recipient 

purpose ——— direction ——— recipient ✓

Introducing semantic maps

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

(Haspelmath, 2003)

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

(Georgakopoulos & Polis, 2018)

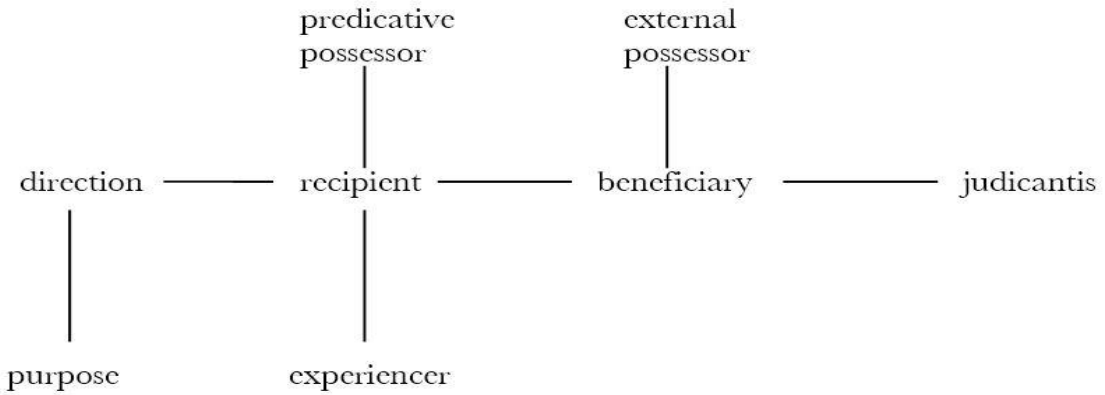


Figure 3. A semantic map of typical dative functions
(based on Haspelmath 2003: 213)

Introducing semantic maps

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

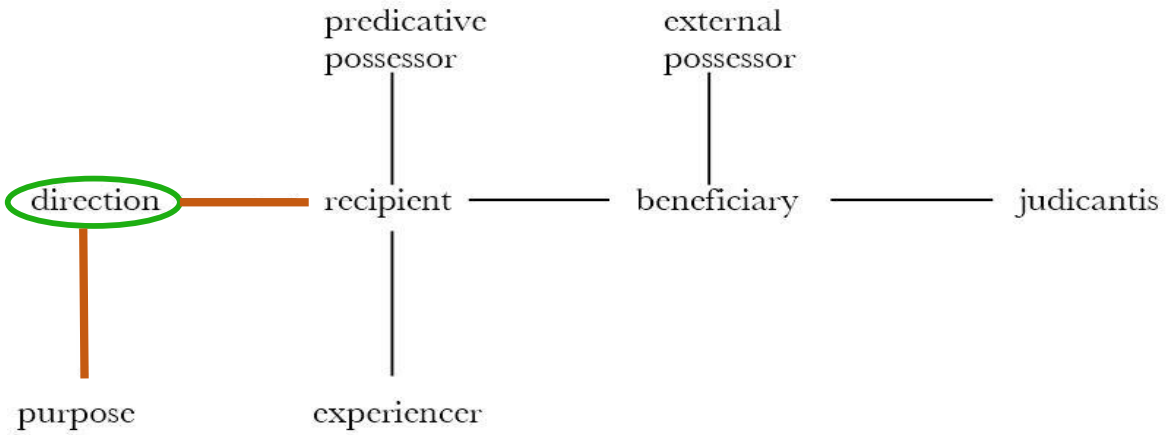


Figure 3. A semantic map of typical dative functions
(based on Haspelmath 2003: 213)

Colexification networks

Form: *abc*

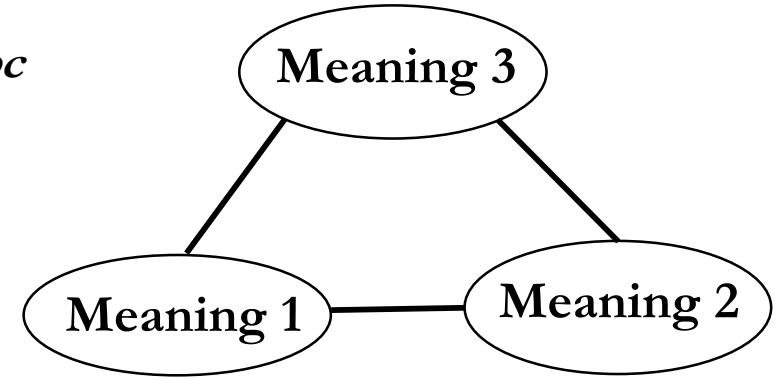


Figure 4a. Abstract colexification network

- Colexification networks are restricted to pairwise associations
- No implicational universals can be inferred from the graph

abc colexifies **Meaning 1**, **Meaning 2**, and **Meaning 3**

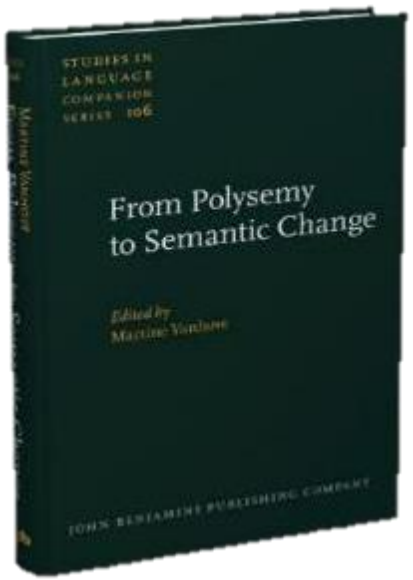
Form: *abc*

Figure 4b. Abstract classical semantic map

Associations in the domain of perception and cognition

Universal patterns

Associations in the domain of perception

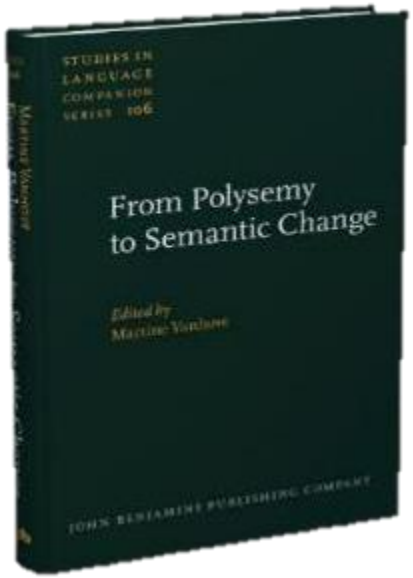


Vanhove 2008

- Sample of 25 languages (8 phyla); mostly African

- **Transfield**: the auditory modality prevails over the visual modality
  - Stronger semantic association of hearing and mental perception

Associations in the domain of perception



Vanhove 2008

- Sample of 25 languages (8 phyla); mostly African

A	B	C	D	E	F	G	H	I	J	K
		See	Heed	Understar	Know	Learn	Think	Hear	Obey	Remember
English	see	1	0	1	1	1	0	0	0	0
German	sehen	1	0	1	1	0	0	0	0	0
French	voir	1	0	1	1	0	0	0	0	0
Italian	vedere	1	0	1	0	0	0	0	0	0
Russian	*IE weid-\	1	0	0	1	0	0	0	0	0
Arabic	ra_a	1	0	0	1	0	1	0	0	0
Beja	rha	1	0	1	0	0	0	0	0	0
Sar	_	1	1	1	0	0	0	0	0	0
Yulu	*e_ __.ga_	1	0	0	1	0	0	0	0	0
Kasem	(na_)	1	0	1	1	0	0	0	0	0
Swahili	kuona	1	0	1	0	0	0	0	0	0
Wolof	gis	1	0	1	0	0	0	0	0	0
English	listen	0	1	0	0	0	0	1	1	0
English	hear	0	0	1	1	1	0	1	0	0
German	hren	0	1	1	1	1	0	1	1	0

Table 3. Snapshot of the polysemy data from Vahnove 2008

Associations in the domain of perception

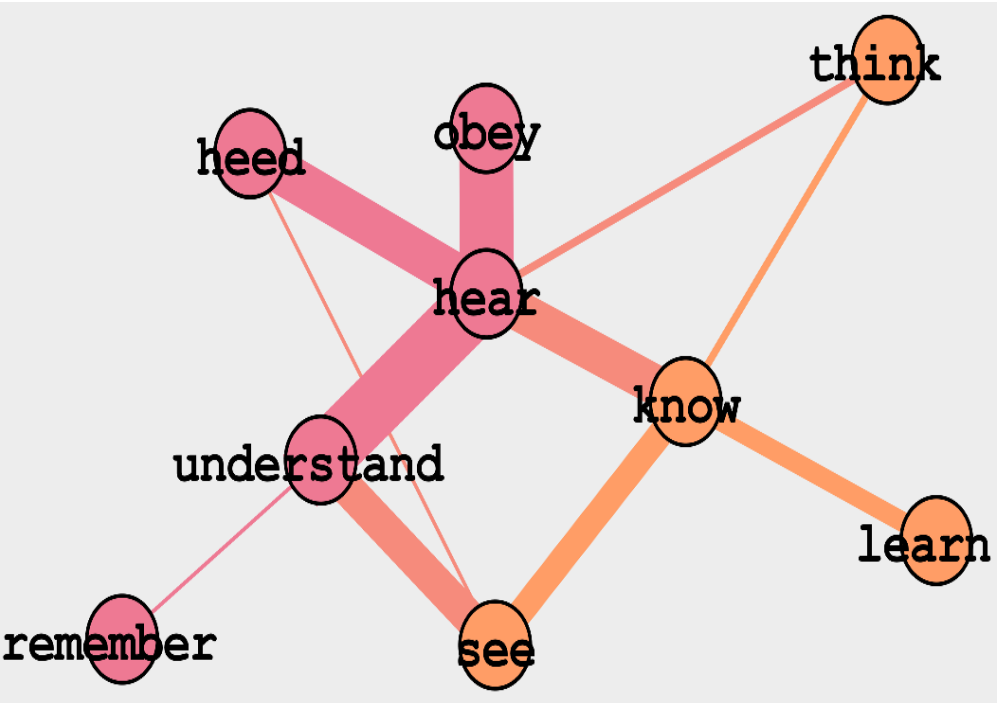


Figure 5. Semantic map of the associations between the verbs of *seeing*, *hearing* and *cognition* (visualized with modularity analysis* in Gephi)

Vanhove 2008

- Sample of 25 languages (8 phyla); mostly African

- **Transfield**: the auditory modality prevails over the visual modality
  - Stronger semantic association of hearing and mental perception

* 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.

Associations in the domain of perception

Database of Cross-Linguistic Colexifications

CLICS² is an online database of colexifications (polysemiesor homophonies) in currently 1220 language varieties of the world.

(List et al., 2018; https://clics.clld.org)

- *N* of Colexification patterns: **2638**
- *N* of concepts: **2487**
- Based on 15 lexical datasets

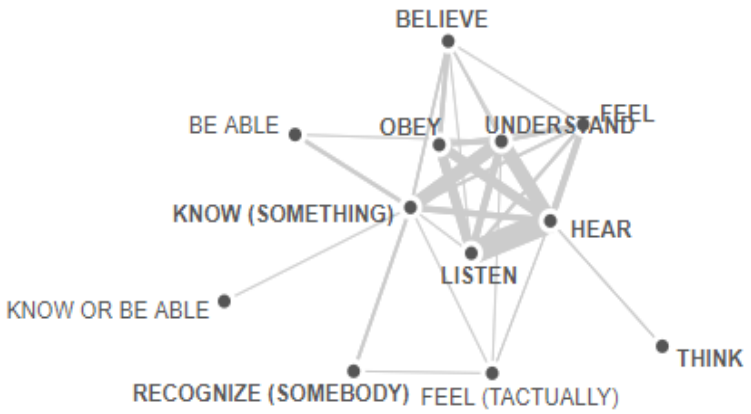
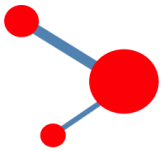


Figure 6. Colexification network of KNOW in CLICS³

Associations in the domain of perception

Database of Cross-Linguistic Colexifications



CLICS² is an online database of colexifications (polysemiesor homophonies) in currently 1220 language varieties of the world.

- In CSV format all the meanings that are attested for lexemes that express at least one of the four concepts SEE, LOOK, HEAR and LISTEN
  - **4045** different word forms
  - **819** word forms colexify at least 2 meanings
  - The forms co-express 362 meanings in total

language:form	Meaning 1	Meaning 2
aro_std:[ba]//ayo_std:[iˈmoʔ]//haw_std:[ʔike]//mcq_std:[ɓanahe]//mri_std:[kitea]//tel_std:[aarayu]//tel_std:[arayu]	see	know
agr_std:[wainat]//arn_std:[pe]//con_std:[ˈatʰeye]//cwg_std:[yow]//emp_std:[uˈnu]//kgp_std:[we]//kpv_std:[addzɪnɪ]//kyh_std:[mah]//mca_std:[wen]//mri_std:[kitea]//oym_std:[ɛsa]//pbb_std:[uy]//plt_std:[mahìta]//pui_std:[duk]//ray_std:[tikeʔa]//rtm_std:[ræe]//sap_Enlhet:[neŋwetayʔ]//sei_std:[aʔo]//shb_std:[taa]//sja_std:[unu]//swh_std:[ona]//tbc_std:[le]//yag_std:[tiki]	see	find
kgp_std:[we]//mbc_std:[eraʔma]//pbb_std:[uy]//sap_Standard:[akwitayi]//srq_std:[tea]//udi_std:[акъсун]	see	get, obtain

Table 4. Snapshot of the polysemy data from CLICS

Associations in the domain of perception

- **From CLICS to a more economic map**

The economy principle:

Given three meanings ($Meaning_A$, $Meaning_B$, $Meaning_C$), if the linguistic items expressing $Meaning_A$ and $Meaning_C$ always express $Meaning_B$, there is no need to draw an edge between $Meaning_A$ and $Meaning_C$ (the resulting map will not be triangular, with all the meanings connected).

(Georgakopoulos & Polis, 2018)

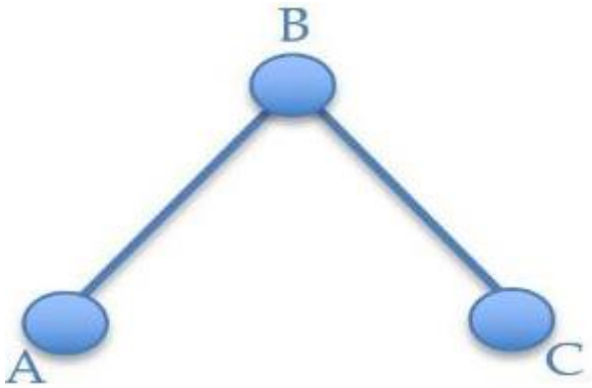


Figure 7. An abstract semantic map

Associations in the domain of perception

- **From CLICS to a more economic map**

```
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 α

- The polysemy patterns were converted into a lexical matrix

Languages

Forms

Source of constraint	Constraint name	Meaning 1 'SEE	Meaning 2 'KNOW'	Meaning 3 'GET, OBTAIN
Araona	ba	1	1	0
Ayoreo	i'moʔ	1	1	0
Hawaiian	ʔike	1	1	0
Ese	ɓanahe	1	1	0
Maori	kitea	1	1	0
Telugu	aarayu	1	1	0
Kaingang	we	1	0	1
Macushi	eraʔma	1	0	1
Páez	uy	1	0	1
Sanapaná (Standard)	akwitayi	1	0	1
Sirionó	tea	1	0	1
Udi	акъсун	1	0	1

Associations in the domain of perception

- **From CLICS to a more economic map**

Languages Forms

Source of constraint	Constraint name	Meaning 1 'SEE	Meaning 2 'KNOW'	Meaning 3 'GET, OBTAIN
Araona	ba	1	1	0
Ayoreo	iˈmoʔ	1	1	0
Hawaiian	ʔike	1	1	0
Ese	ɓanahe	1	1	0
Maori	kitea	1	1	0
Telugu	aarayu	1	1	0
Kaingang	we	1	0	1
Macushi	eraʔma	1	0	1
Páez	uy	1	0	1
Sanapaná (Standard)	akwitayi	1	0	1
Sirionó	tea	1	0	1
Udi	акъсун	1	0	1

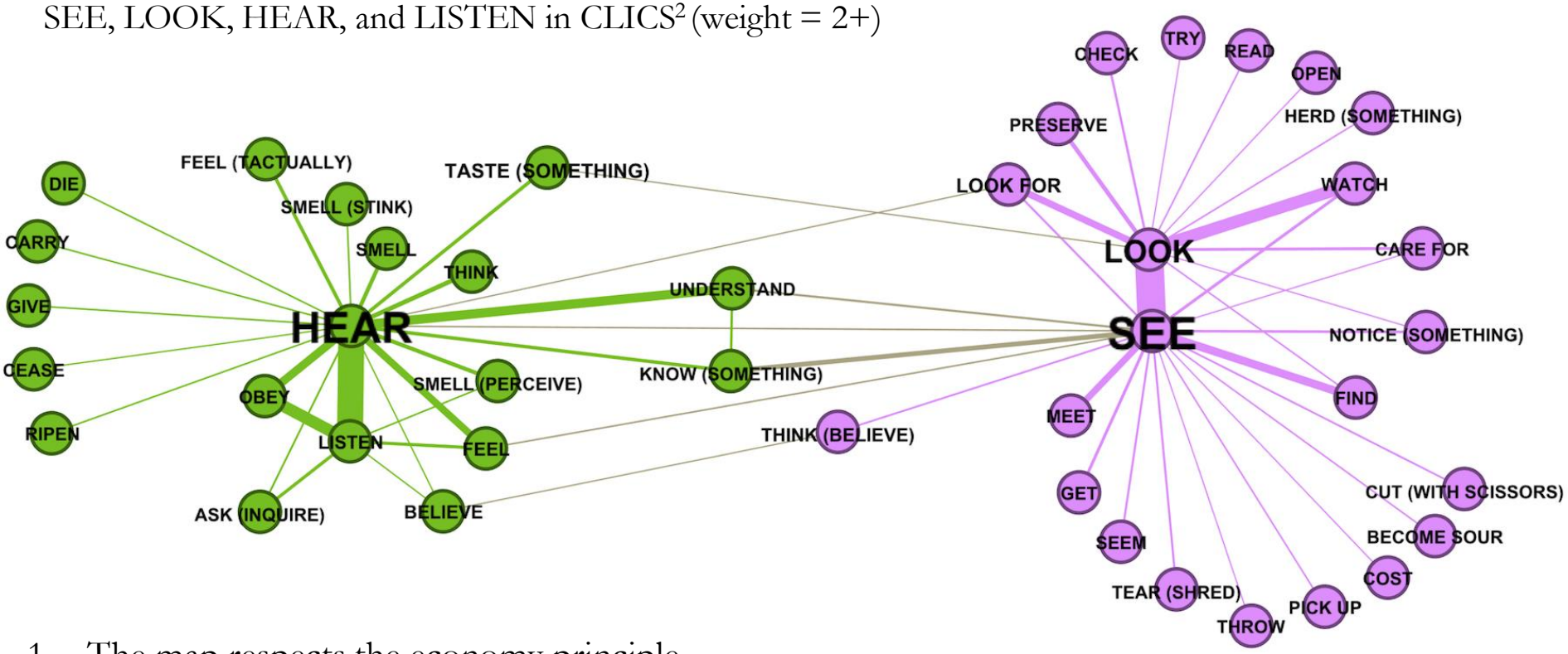
- **Weighted semantic map** based on the adapted version of the algorithm suggested by Regier et al. (2013)

```
# CREATE INITIAL GRAPH
# graph G: add each term's nodes, no edges in graph yet.
G = nx.Graph()  # create empty graph (undirected)
PossE = []      # list of possible edges, filled below
for t in T:
        # add all nodes in t, if not already in graph
        for n in t:
                if (not G.has_node(n)):
                        G.add_node(n)
        # add to PossE a link between each pair of nodes in t
        # adding a link between every node in G is needless and slower
        for pair in allpairs(t):
                u = pair[0]
                v = pair[1]
                if (not (((u,v) in PossE) or ((v,u) in PossE))):
                        PossE.append((u,v))
```

Python script β

Associations in the domain of perception

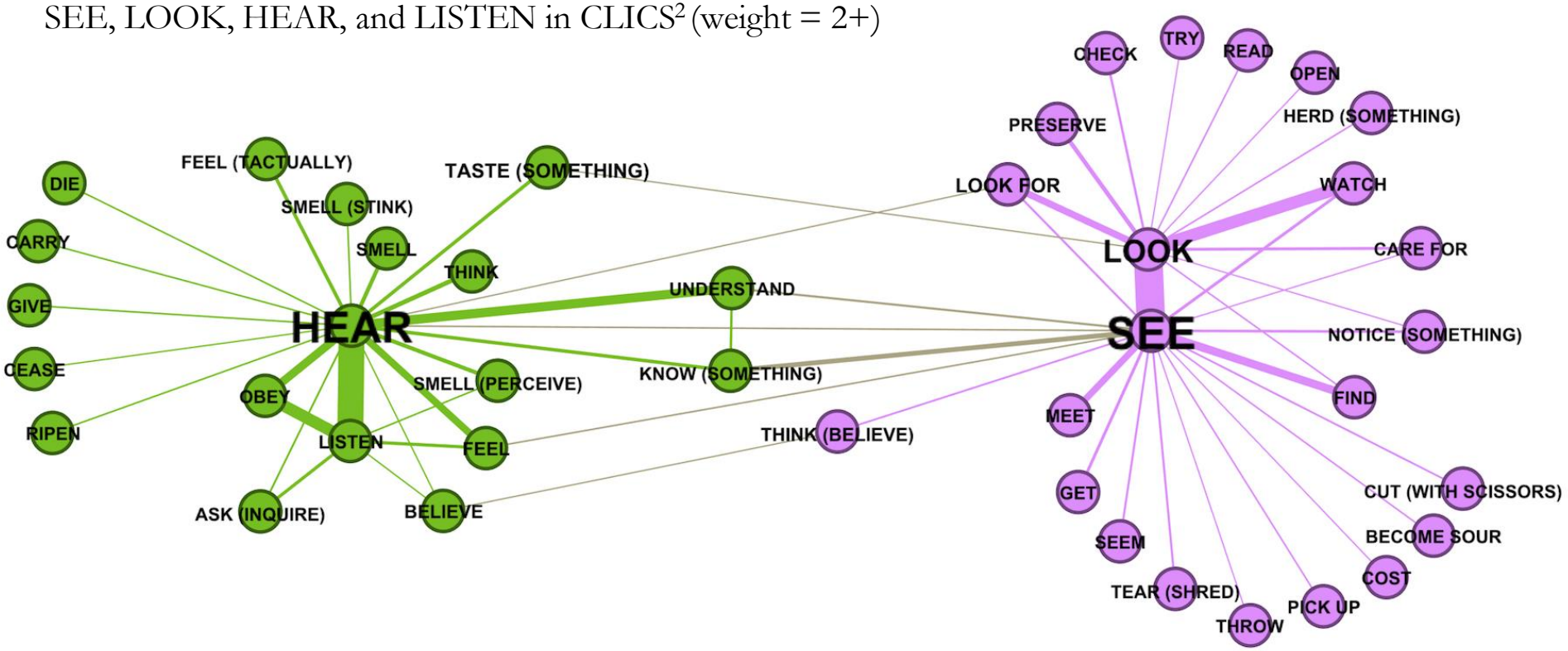
Figure 8. Semantic map for the main colexification patterns of SEE, LOOK, HEAR, and LISTEN in CLICS^2 (weight = 2+)



1. The map respects the economy principle
2. Cognition senses (e.g. KNOW, UNDERSTAND) mediate between the domains of VISION and HEARING
3. Non-controlled experiences (SEE and HEAR) are linked directly to cognition, while controlled activities (LOOK and LISTEN) are not

Associations in the domain of perception

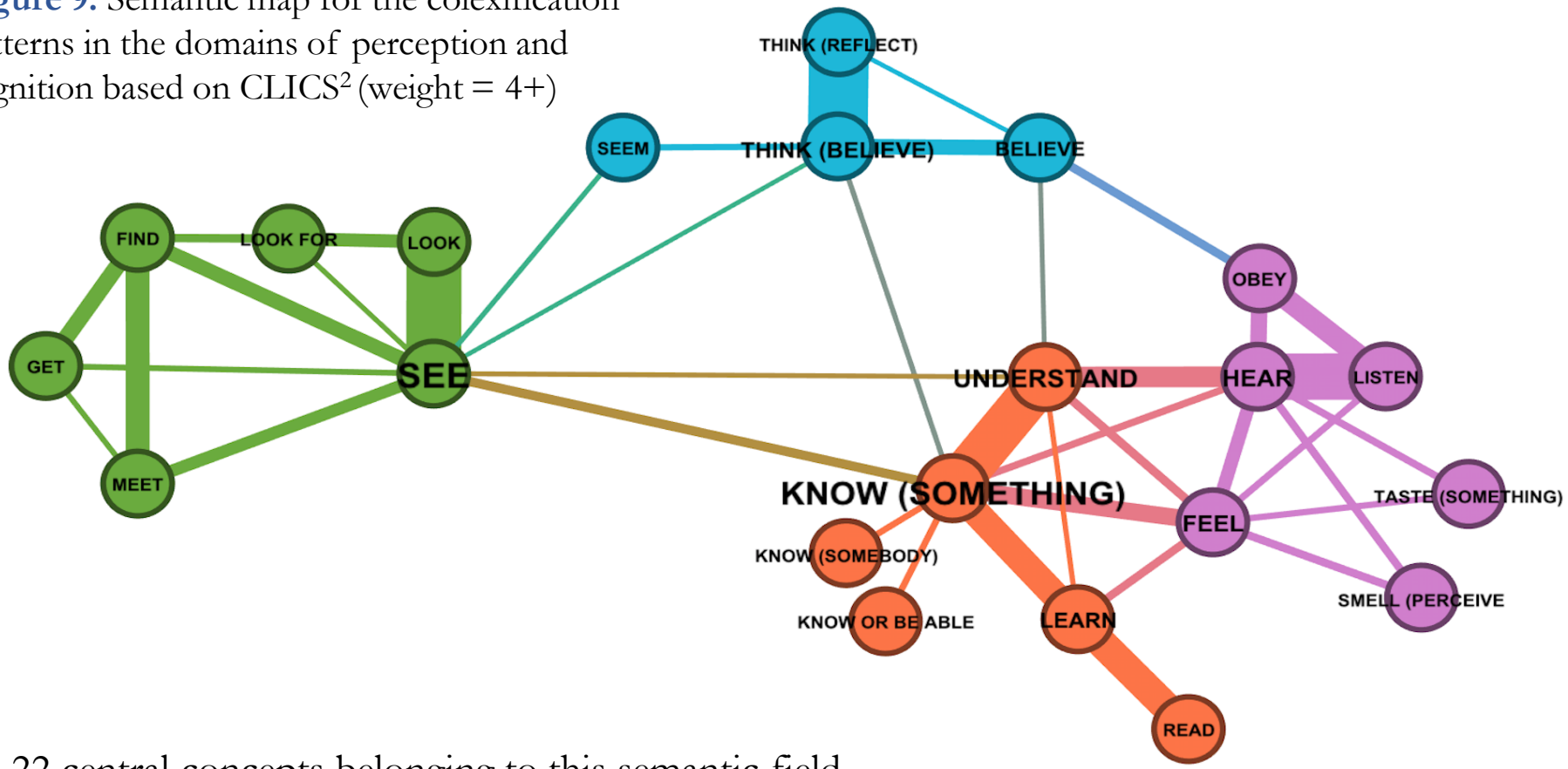
Figure 8. Semantic map for the main colexification patterns of SEE, LOOK, HEAR, and LISTEN in CLICS[2] (weight = 2+)



4. $N_{\text{<SEE, KNOW>}}$: 17 ------- $N_{\text{<HEAR, KNOW>}}$: 11
5. $N_{\text{<HEAR, UNDERSTAND>}}$: 43 ------- $N_{\text{<SEE, UNDERSTAND>}}$: 6).
6. Both cognition meanings are more tightly associated with the HEAR cluster
7. Meanings belonging to the other sensory modalities (such as FEEL (TACTUALLY), SMELL, TASTE) form a group with HEAR

Perception and cognition (a broader view)

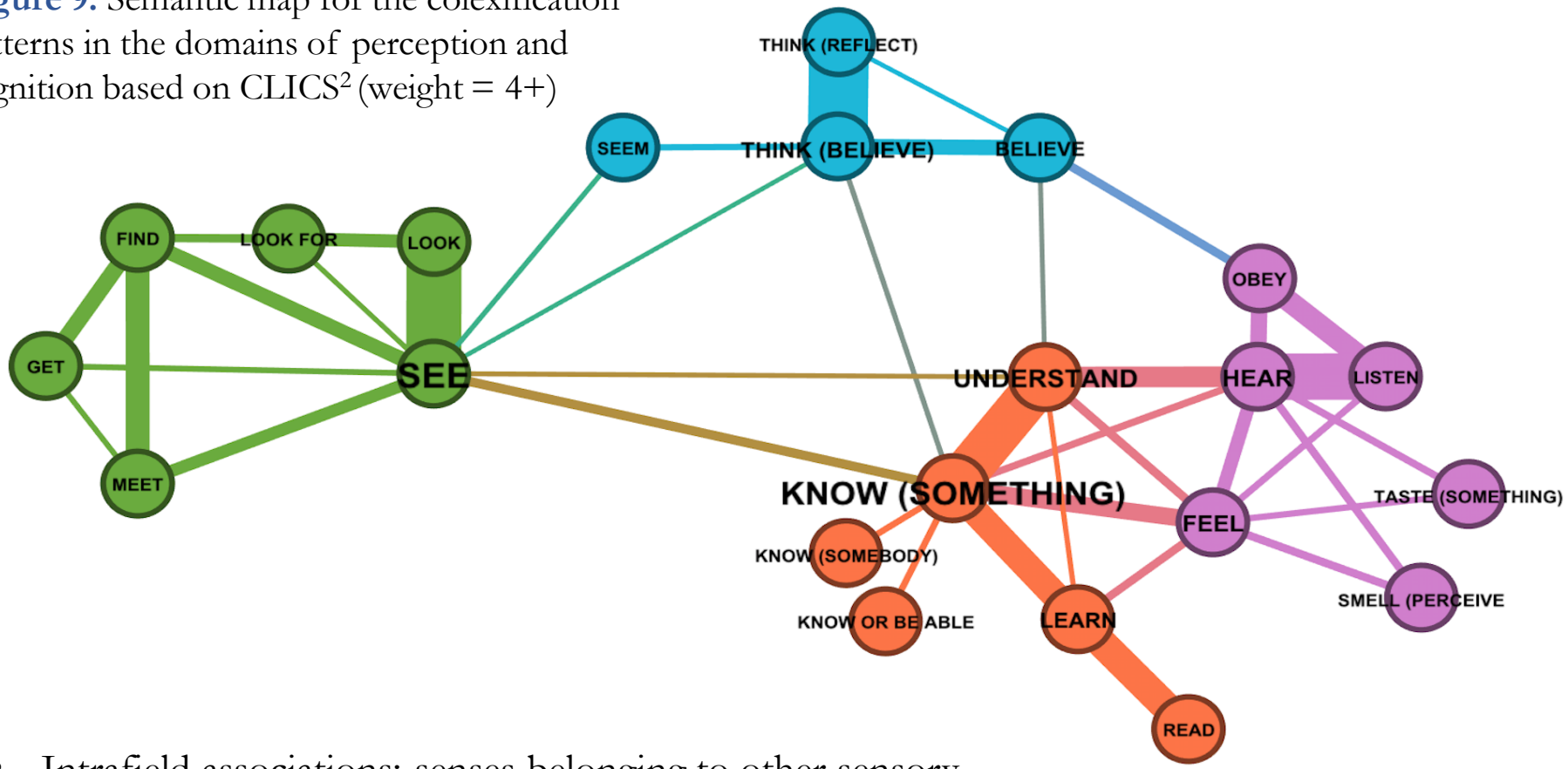
Figure 9. Semantic map for the colexification patterns in the domains of perception and cognition based on CLICS[2] (weight = 4+)



- 22 central concepts belonging to this semantic field
- We extracted all the verbs that colexify at least two meanings from this set of meanings
  - 962 colexification patterns
  - 873 unique forms

Perception and cognition (a broader view)

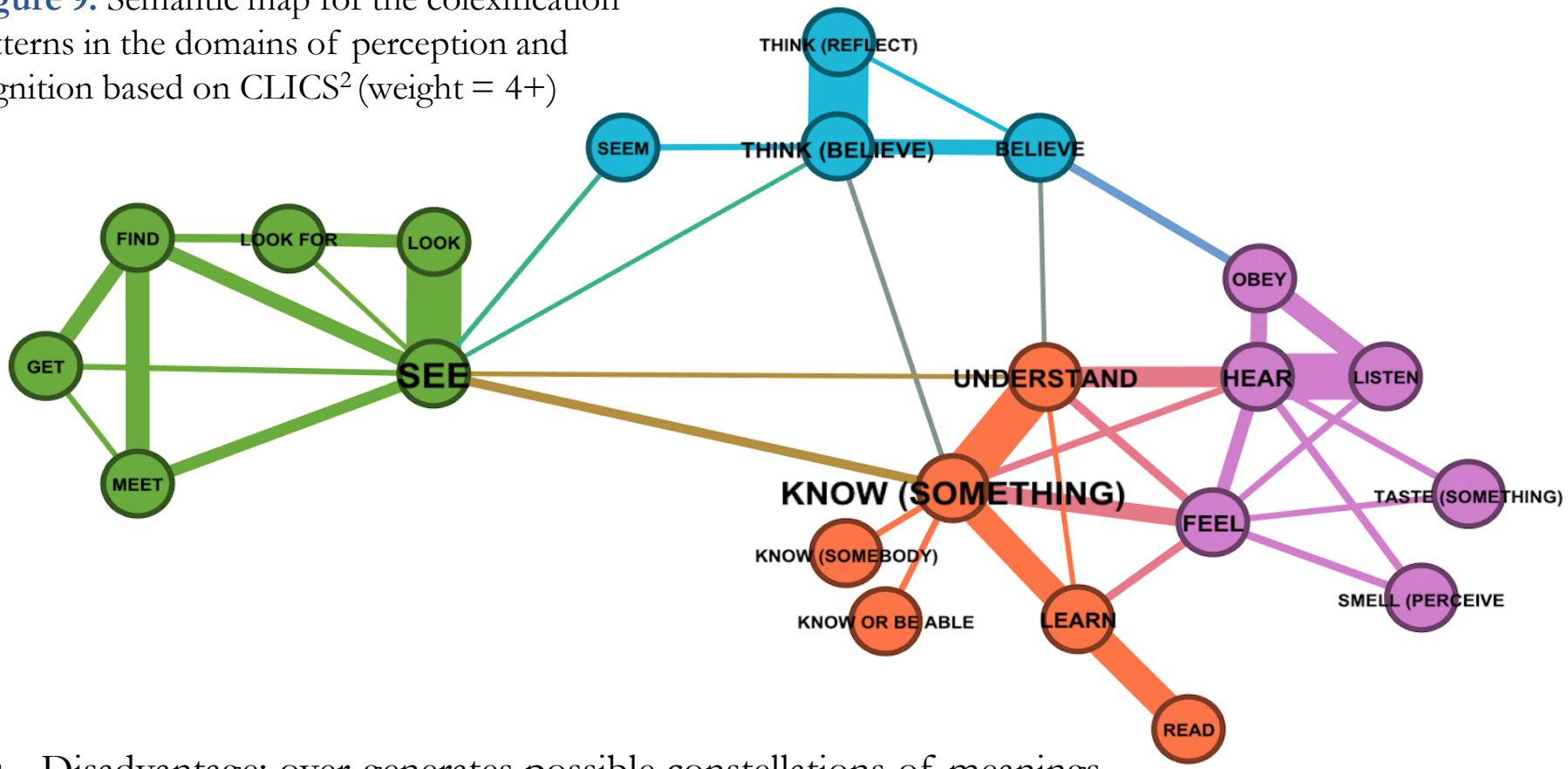
Figure 9. Semantic map for the colexification patterns in the domains of perception and cognition based on CLICS[2] (weight = 4+)



- Intrafield associations: senses belonging to other sensory modalities, i.e., TASTE and SMELL are grouped with HEAR
- FEEL behaves as a kind of hypernym

Perception and cognition (a broader view)

Figure 9. Semantic map for the colexification patterns in the domains of perception and cognition based on CLICS[2] (weight = 4+)



- Disadvantage: over-generates possible constellations of meanings
  - Predicts patterns that are possible but unattested
  - Predicts very unlikely patterns

Perception and cognition (a broader view)

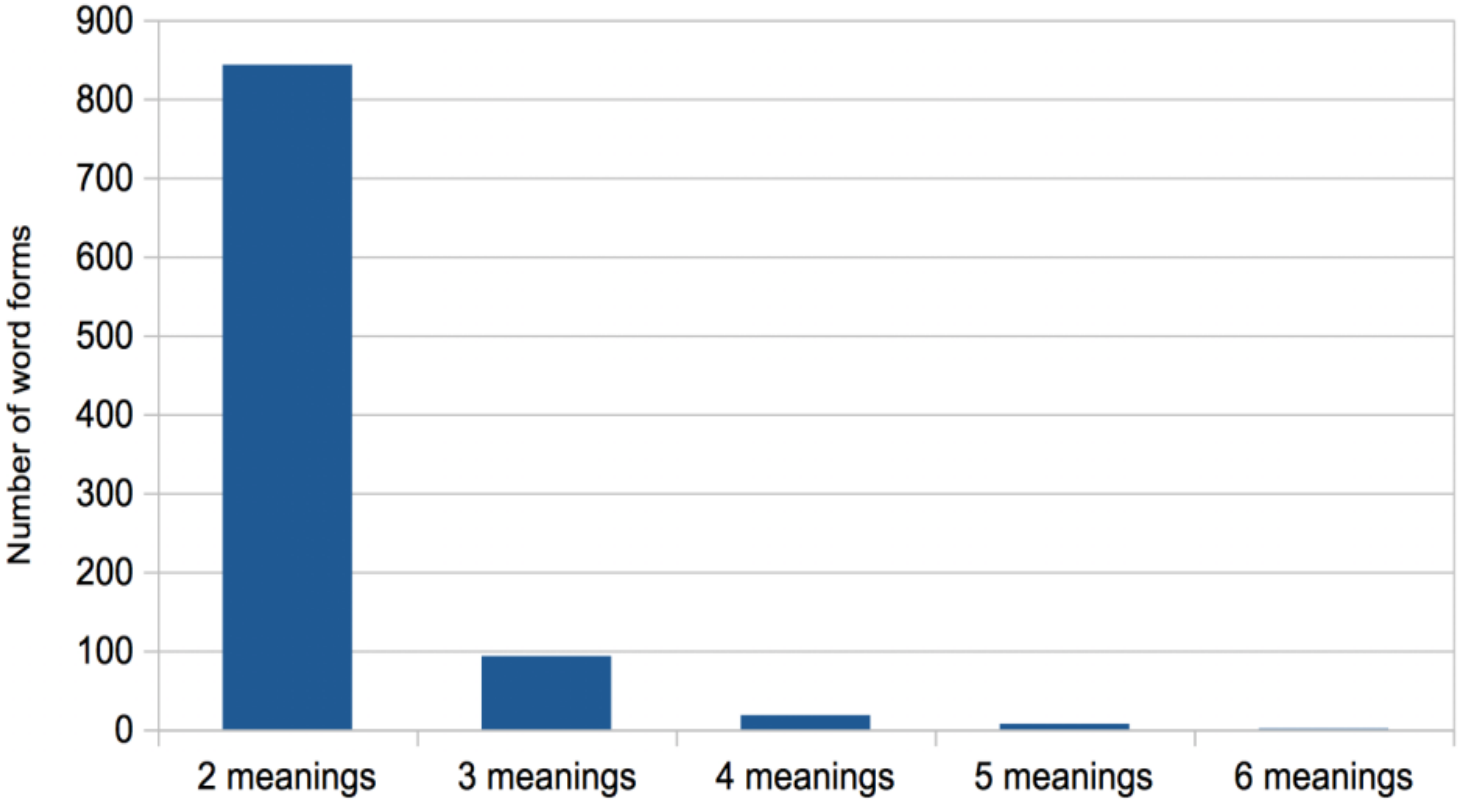
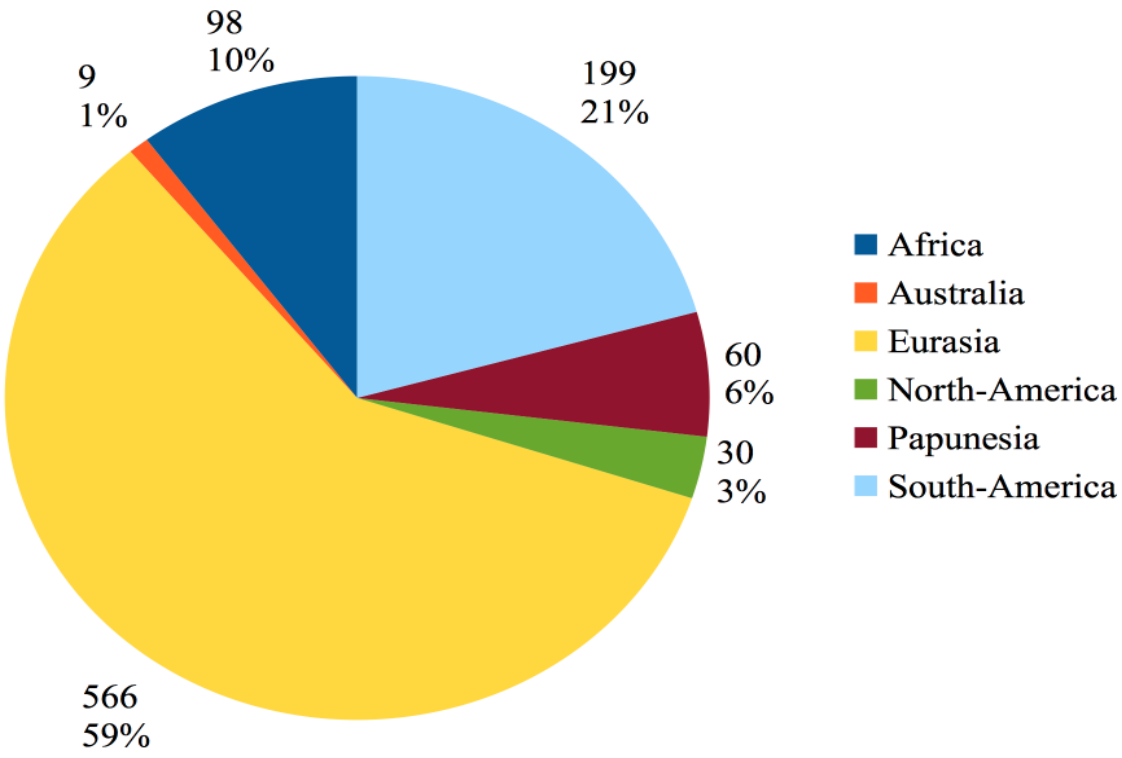


Figure 10. Number of meanings per word-forms in the dataset

Macro-areal patterns in the domains of perception and cognition

Perception and cognition -- Macro-areas



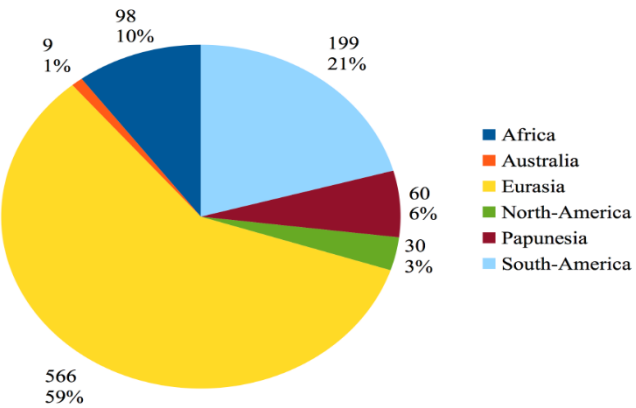
- 22 central concepts
- 962 colexification patterns

Figure 11. Distribution of the data per macro-area in the domains of perception and cognition

Perception and cognition -- Macro-areas

Macro-area	Number of families	Number of language varieties	Number of colexification patterns	Number of meanings colexified	No information about colexification patterns for the meanings
Africa	6	71	98	19	KNOW (SOMEBODY), KNOW OR BE ABLE, SEEM
Eurasia	23	273	566	22	-
Papunesia	3	40	60	17	KNOW (SOMEBODY), KNOW OR BE ABLE, LEARN, READ, SEEM
South America	36	75	199	20	KNOW (SOMEBODY), KNOW OR BE ABLE

Table 4. Distribution of the data per macro-area in the domains of perception and cognition



Macro-areas -- Africa

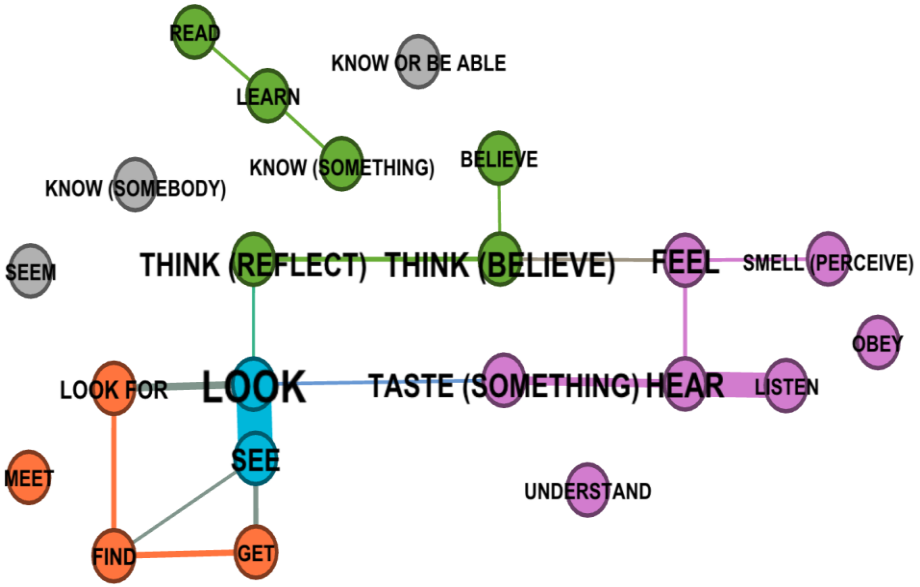


Figure 11a. Semantic map (Africa)

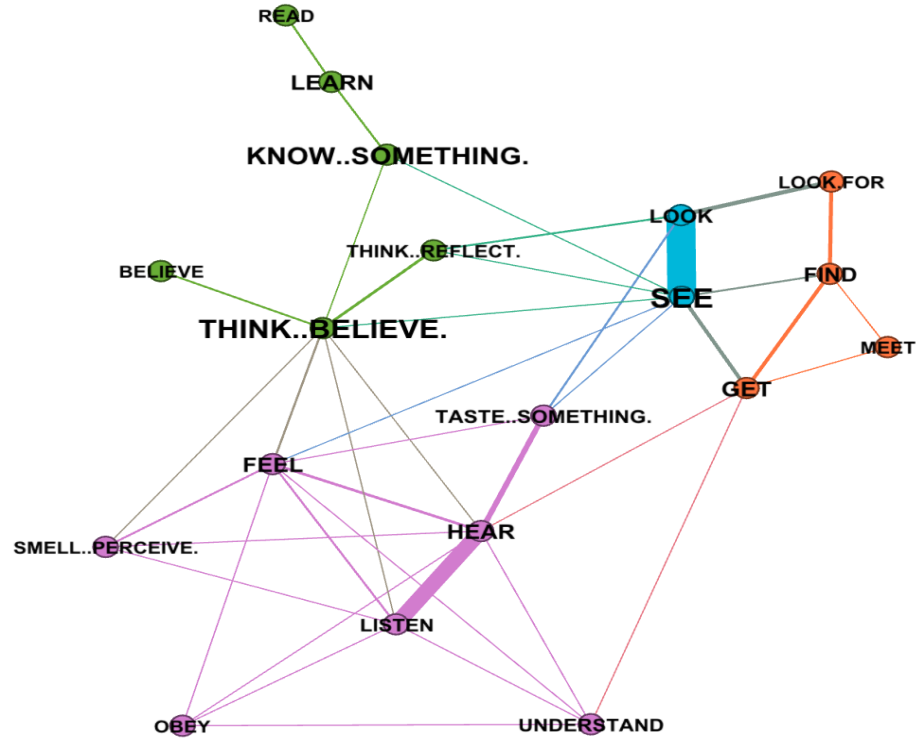
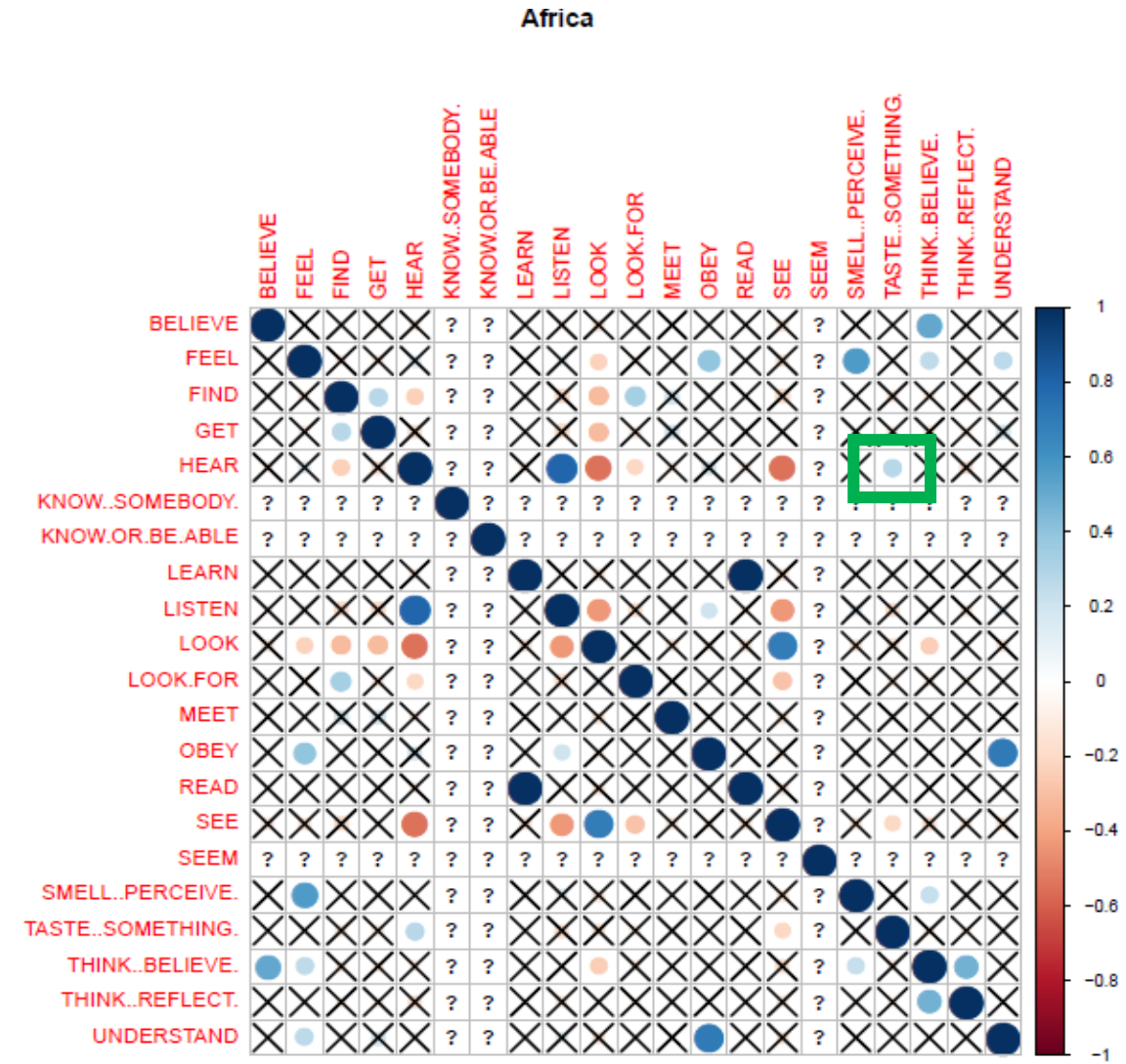


Figure 11b. Colexification network (Africa)

- <HEAR, TASTE> colexification is well represented
- All the sensory modalities are linearly connected in this macro-area
- KNOW and UNDERSTAND do not mediate between (a) VISION (b) other sense modalities

Macro-areas -- Africa



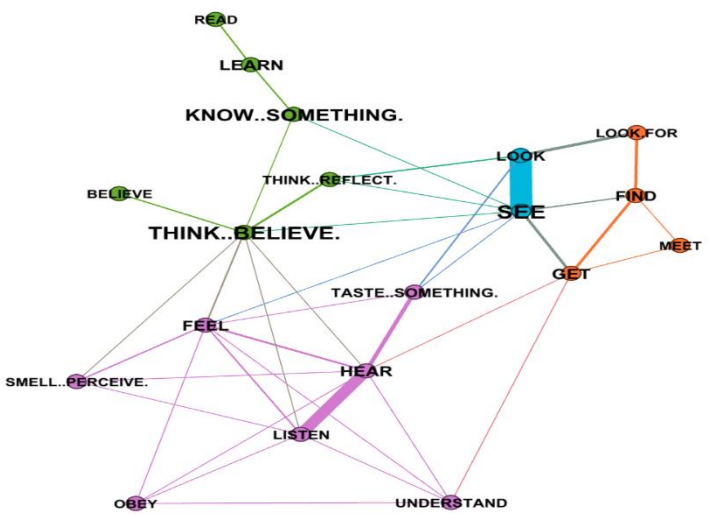


Figure 11c. Correlation plot (Africa)

Macro-areas -- South America

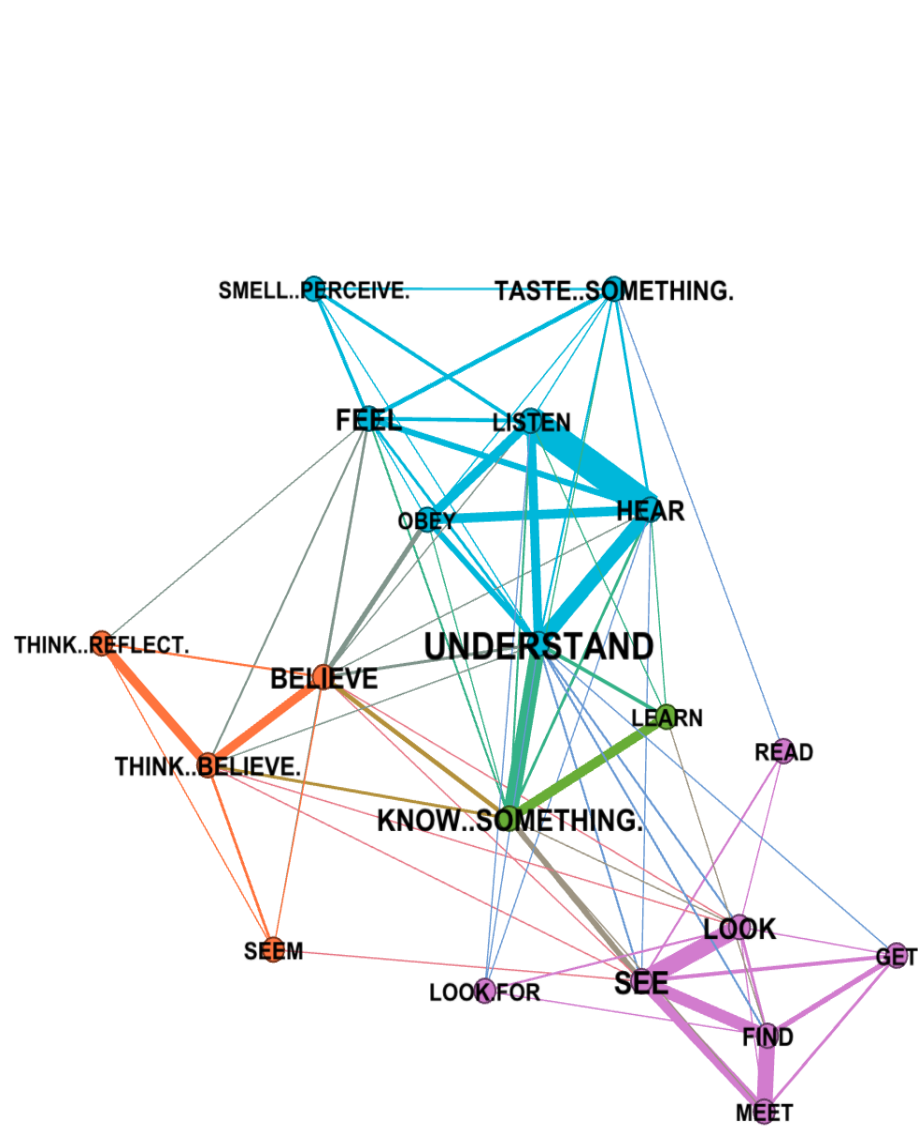


Figure 12b. Colexification network (South America)

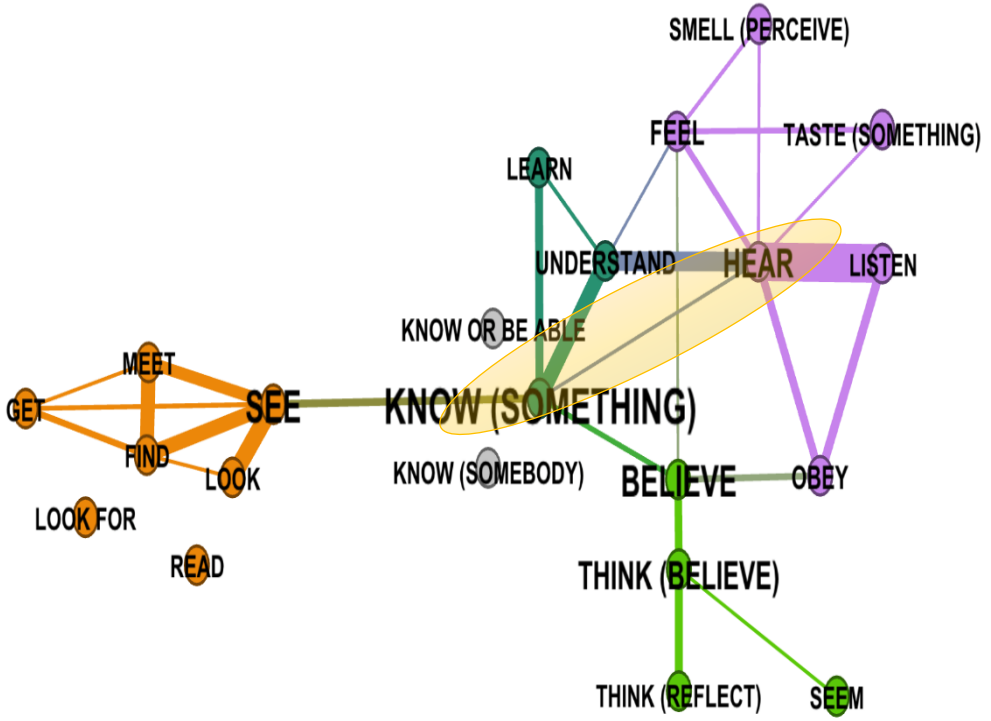


Figure 12a. Semantic map (South America)

Macro-areas -- South America

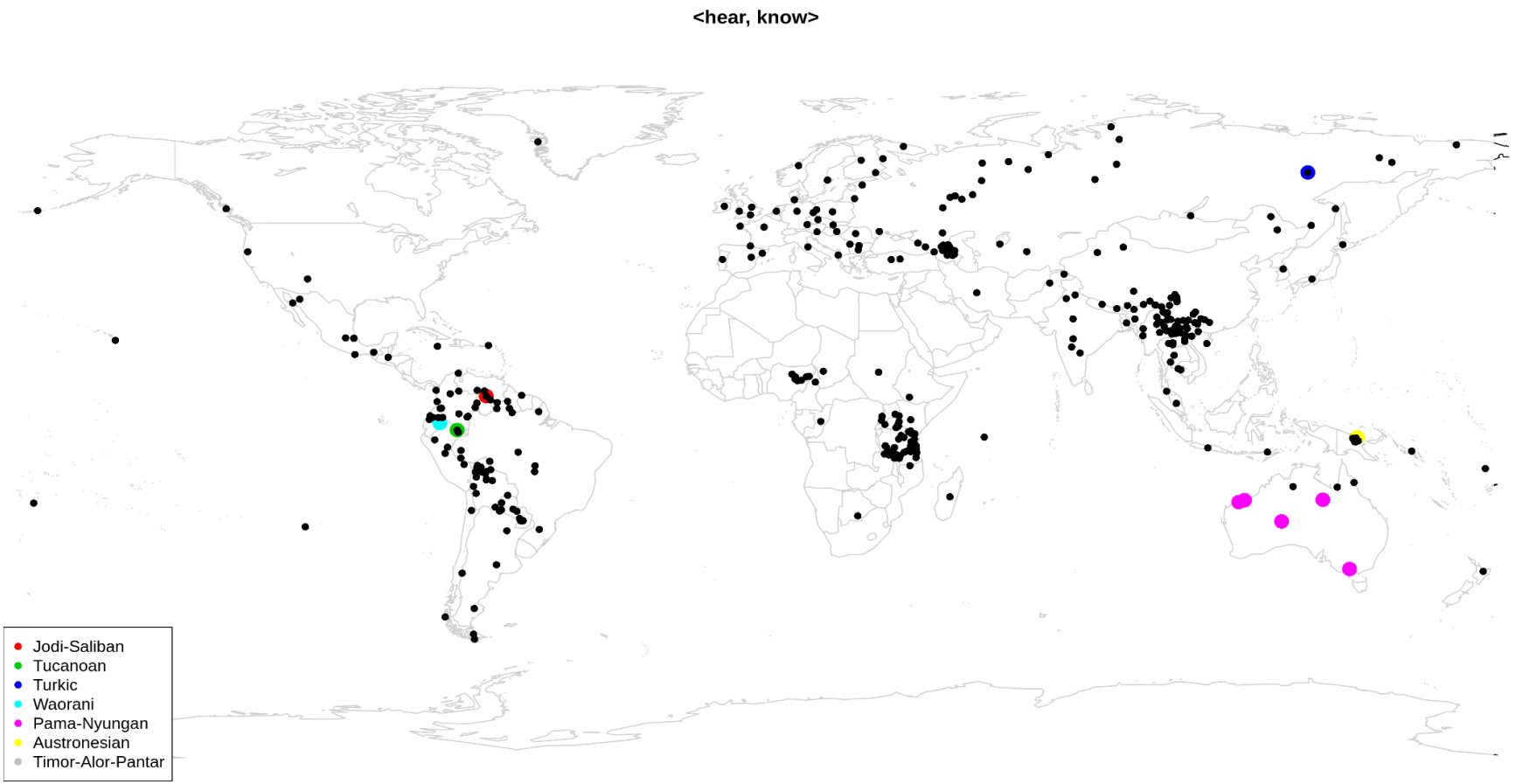


Figure 12c. The <HEAR, KNOW> colexification worldwide (Black-colored dots represent no attestation of the colexification pattern and other-than-black-colored dots signal that the language varieties show this pattern.)

Macro-areas -- South America

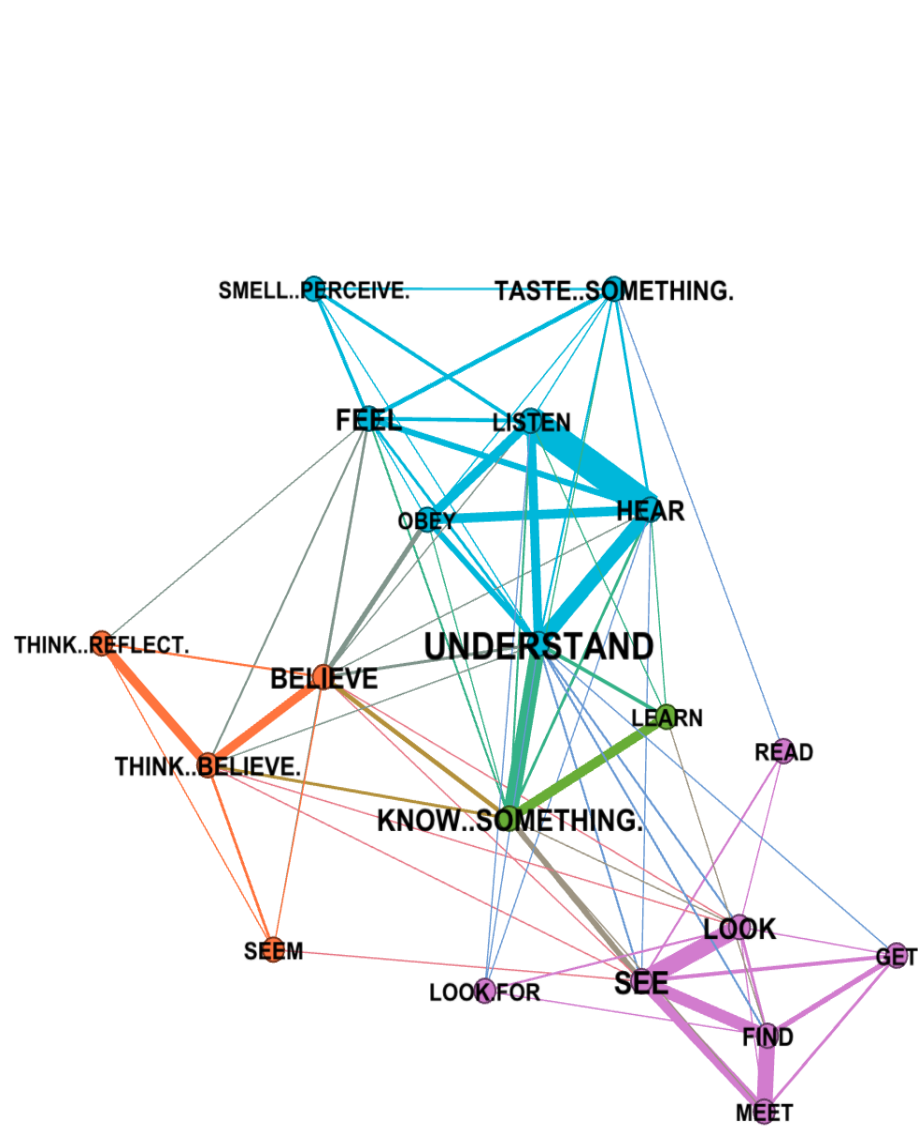


Figure 12b. Colexification network (South America)

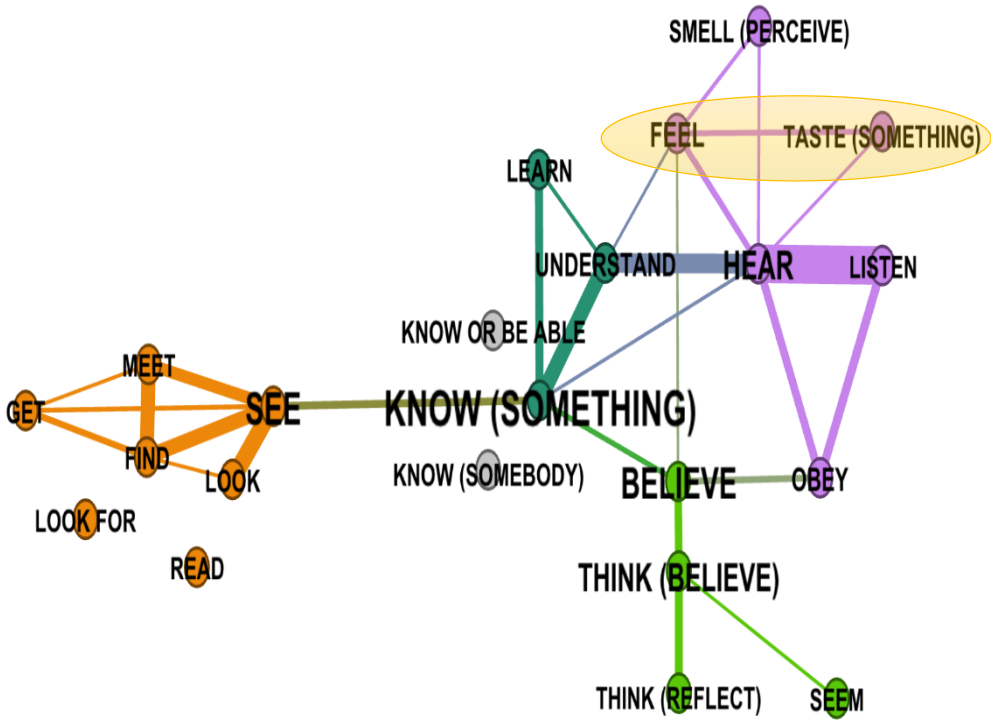


Figure 12a. Semantic map (South America)

Macro-areas -- South America

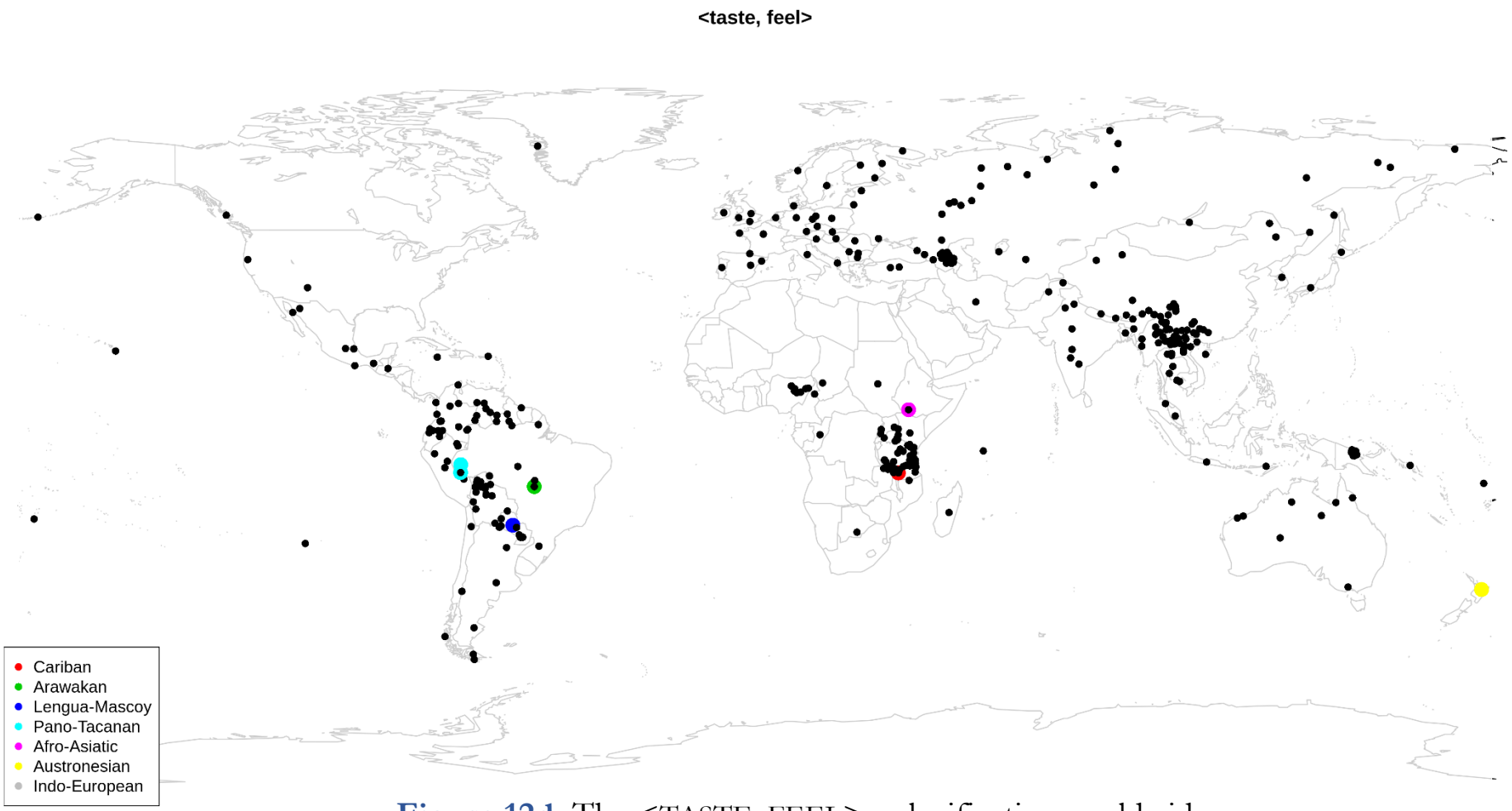


Figure 12d. The <TASTE, FEEL> colexification worldwide (Black-colored dots represent no attestation of the colexification pattern and other-than-black-colored dots signal that the language varieties show this pattern.)

Macro-areas -- South America

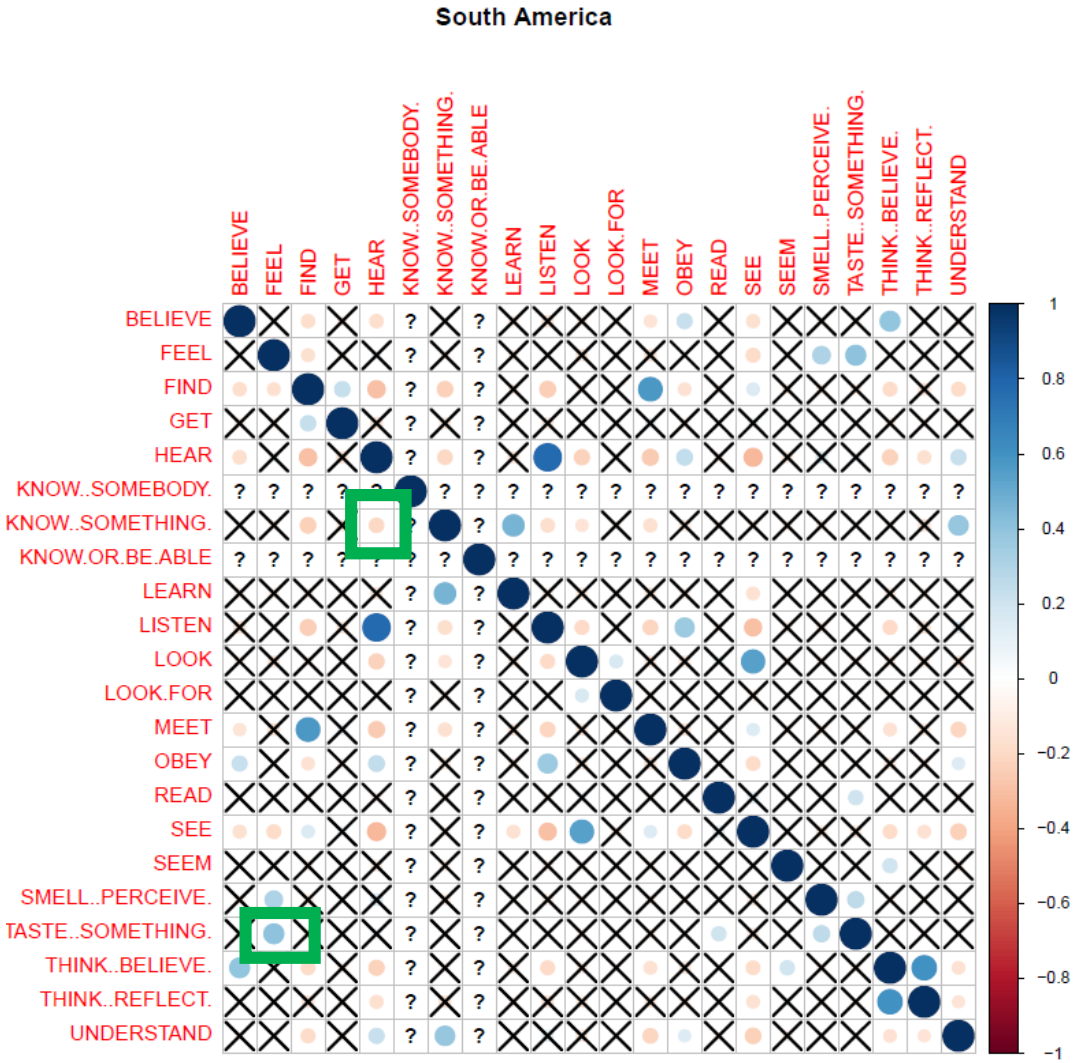


Figure 12e. Correlation plot (South America)

General discussion

General discussion

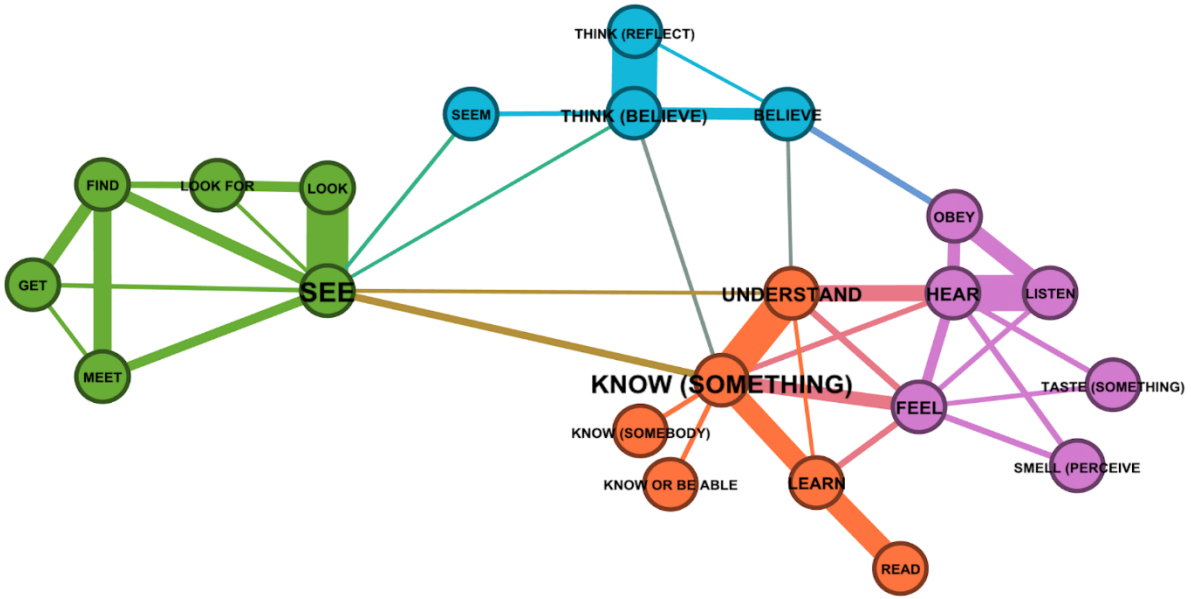
- Generalizations about the cross-linguistic organization of the lexicon are not easily or straightforwardly identified
  - Large datasets make large typological studies on the lexicon possible

- Different methods answer different type of questions
  - **Semantic maps** for unveiling cross-linguistically recurrent semantic structures in the semantic domains of perception and cognition
  - **Correlation plots** and **colexification networks** to uncover patterns that are specific to particular macro-areas (and, potentially) to micro-areas

General discussion

- Our semantic maps show that intrafield connections between verbs of vision and hearing are mediated by interfield connections, i.e., via the cognition domains of knowledge and understanding.

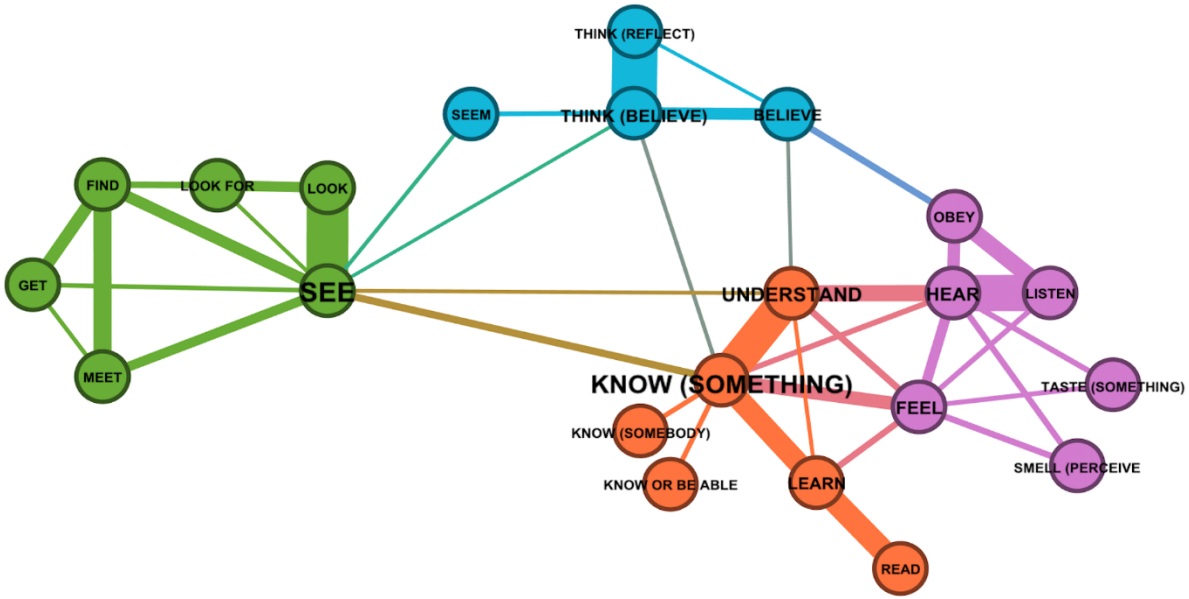
- Controlled activities, as those instantiated by such verbs as *look* or *listen*, are not directly linked to cognition: the verbs expressing uncontrolled experiences (SEE and HEAR) are



General discussion

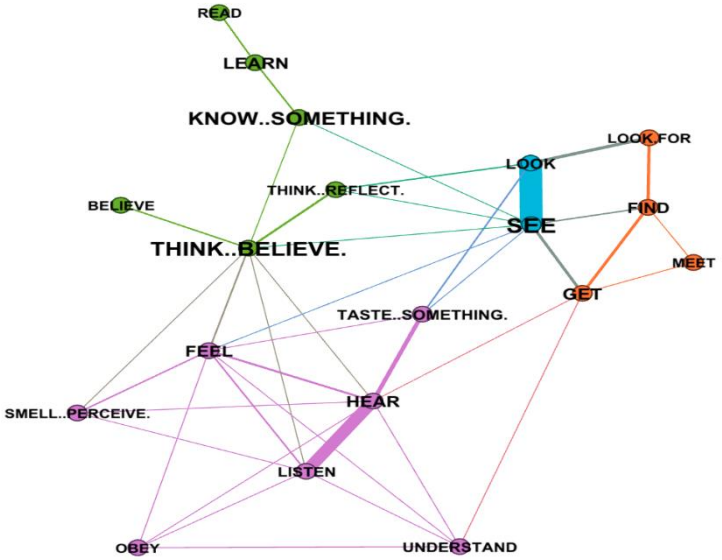
- Knowledge is more closely linked to vision, and mental manipulation (i.e., understanding) to hearing (a correlation which is stronger in Eurasia)

- Meanings belonging to other sensory modalities (TASTE, SMELL (PERCEIVE), FEEL) cluster with HEAR rather than with SEE



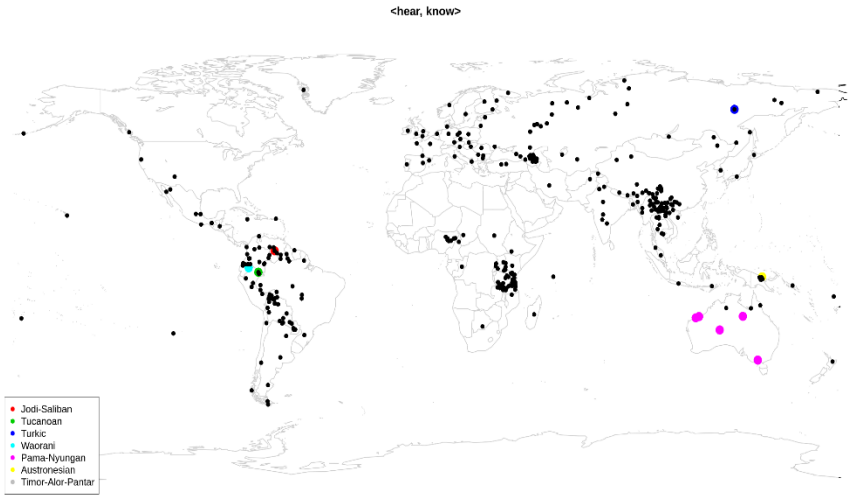
General discussion

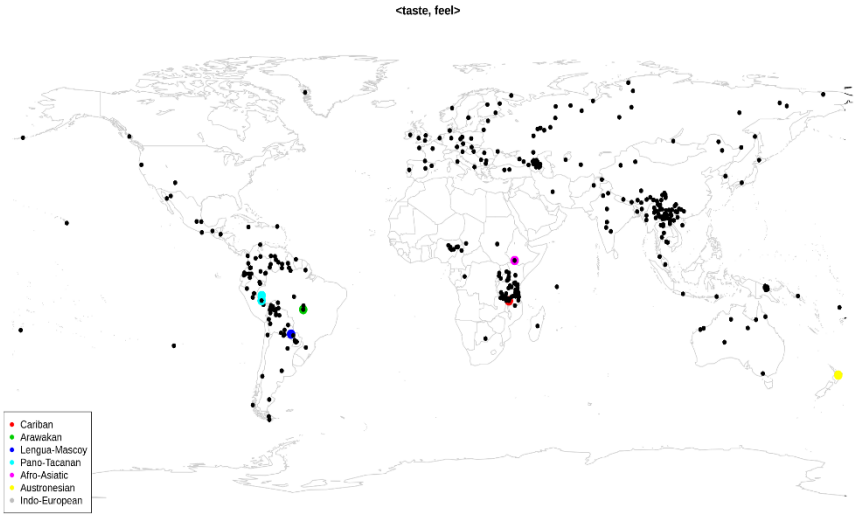
- **In Africa:**
  - The colexification <TASTE, HEAR> is well-attested and the two meanings are positively correlated
  - All the sensory modalities are linearly connected in this macro-area
  - KNOW and UNDERSTAND do not mediate between (a) VISION (b) other sense modalities



General discussion

- **In South America**:
  - Two possible candidates of areally biased colexification patterns have been identified: <TASTE, FEEL> and <HEAR, KNOW>





Some limitations

- The sample is not ideal (skewed towards Eurasia)
- Semantic maps over-generate possible constellations of meanings
  - Gap between predicted and documented colexification patterns
- Even those techniques that allow for significance testing (e.g., correlations) can be fooled by unbalanced samples

Selected references

Aikhenvald Y. Alexandra & Anne Storch (eds.). 2013. *Perception and cognition in language and culture*. Leiden: Brill.

Bastian, Mathieu, Sebastien Heymann & Mathieu Jacomy. 2009. Gephi: an open source software for exploring and manipulating networks. In *Proceedings of the third international AAAI conference on weblogs and social media, Association for the Advancement of Artificial Intelligence*. http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154.

Croft, William. 2001. *Radical Construction Grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William. 2003. *Typology and universals*. 2nd ed.. Cambridge: Cambridge University Press.

Evans, Nicholas & David Wilkins. 2000. In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3). 546–92.

François, Alexandre. 2008. Semantic maps and the typology of colexification: intertwining polysemous networks across languages. In Martine Vanhove (ed.), *From polysemy to semantic change*, 163–215. Amsterdam: Benjamins.

Gast, Volker & Maria Koptjevskaja-Tamm. 2018. The areal factor in lexical typology: Some evidence from lexical databases. In Daniel Van Olmen, Tanja Mortelmans and Frank Brisard (eds.), *Aspects of linguistic variation*, 43-82. Berlin: de Gruyter Mouton.

Georgakopoulos, Thanasis & Stéphane Polis. 2018. The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass* 12 (2).e12270–n/a. doi:. E12270 LNCO-0727.R1.

Georgakopoulos Thanasis, Eitan Grossman, Dmitry Nikolaev, and Stéphane Polis. 2021-forthcoming. Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain. *Linguistic Typology*.

Selected references

Haspelmath, Martin. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In Michael Tomasello (ed.), *The new psychology of language*, 211–242. Mahwah, NJ: Lawrence Erlbaum. Ibarretxe-Antuñano, Iraide. 2013. The relationship between conceptual metaphor and culture. *Intercultural Pragmatics* 10(2). 315–339.

Ibarretxe-Antuñano, Iraide. 2013. The relationship between conceptual metaphor and culture. *Intercultural Pragmatics* 10(2). 315–339.

Koptjevskaja-Tamm, Maria, Ekaterina Rakhilina & Martine Vanhove. 2015. The semantics of lexical typology. In Nick Riemer (ed.), *The routledge handbook of semantics*, 434–454. London & New York: Routledge.

Levinson, C. Stephen & Asifa Majid. 2014. Differential ineffability and the senses. *Mind & Language* 29. 407–427.

List, Johann-Mattis, Simon Greenhill, Cormac Anderson, Thomas Mayer, Tiago Tresoldi & Robert Forkel (eds.). 2018. *Database of Cross-Linguistic Colexifications*. Jena: Max Planck Institute for the Science of Human History. http://clics.clld.org.

Matisoff, James. 1978. *Variational semantics in Tibeto-Burman*. Philadelphia: ISHI.

Östling, Robert. 2016. Studying colexification through massively parallel corpora. In Päivi Juvonen & Maria Koptjevskaja-Tamm (eds.), *The lexical typology of semantic shifts*, 157–176. Berlin and Boston: De Gruyter Mouton.

Regier, Terry, Naveen Khetarpal & Asifa Majid. 2013. Inferring semantic maps. *Linguistic Typology* 17.89–105.

San Roque, Lila, Kobin H. Kendrick, Elisabeth Norcliffe, Penelope Brown, Rebecca Defina, Mark Dingemanse, Tyko Dirksmeyer, NJ Enfield, Simeon Floyd, Jeremy Hammond, Giovanni Rossi, Sylvia Tufvesson, Saskia van Putten & Asifa Majid. 2015. Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics* 26. 31–60.

References

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

Sweetser, Eve. 1990. *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.

Vanhove, Martine. 2008. Semantic associations between sensory modalities, prehension and mental perceptions: A crosslinguistic perspective. In Martine Vanhove (ed.), *From Polysemy to Semantic Change. Towards a Typology of Lexical Semantic Associations*, 341–370. Amsterdam/Philadelphia: John Benjamins.

Viberg, Åke. 1984. The verbs of perception: a typological study. *Linguistics* 2. 123–162.

Viberg, Åke. 2001. The verbs of perception. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals. An International Handbook*, 1294–1309. Berlin & New York: Mouton de Gruyter.

Wälchli, Bernhard & Michael Cysouw. 2012. Lexical typology through similarity semantics: Toward a semantic map of motion verbs. In Maria Koptjevskaja-Tamm & Martine Vanhove (eds.), *New Directions in Lexical Typology*. Special issue in Linguistics 50(3). 671–710. https://doi.org/10.1515/ling-2012-0021

Wälchli, Bernhard. 2016. Non-specific, specific and obscured perception verbs in Baltic languages. *Baltic Linguistics* 7. 53–135.

Winter, Bodo, Marcus Perlman & Asifa Majid. 2018. Vision dominates in perceptual language: English sensory vocabulary is optimized for usage. *Cognition* 179. 213-220.

Wilkins, David. 1981. Natural tendencies of semantic change and the search for cognates. In Mark Durie & Malcolm Ross (eds.), *The comparative method reviewed*, 264–304. New York: Oxford University Press.

Ευχαριστώ πολύ

ageorgakopoulos@enl.auth.gr