Thanasis Georgakopoulos & Stéphane Polis

# Semantic Change in the Lexicon:

## *Models of Inference Based on Patterns of Coexpression*

ALT-2026: 16th International Conference of the Association for Linguistic Typology

WS6 - Semantic shifts and the dynamics of lexification patterns

1-3 July 2026 Lyon (France)

ARISTOTLE UNIVERSITY OF THESSALONIKI

ΕΛ.ΙΔ.Ε.Κ. Ελληνικό Ίδρυμα Έρευνας & Καινοτομίας H.F.R.I. Hellenic Foundation for Research & Innovation

fnrs LA LIBERTÉ DE CHERCHER

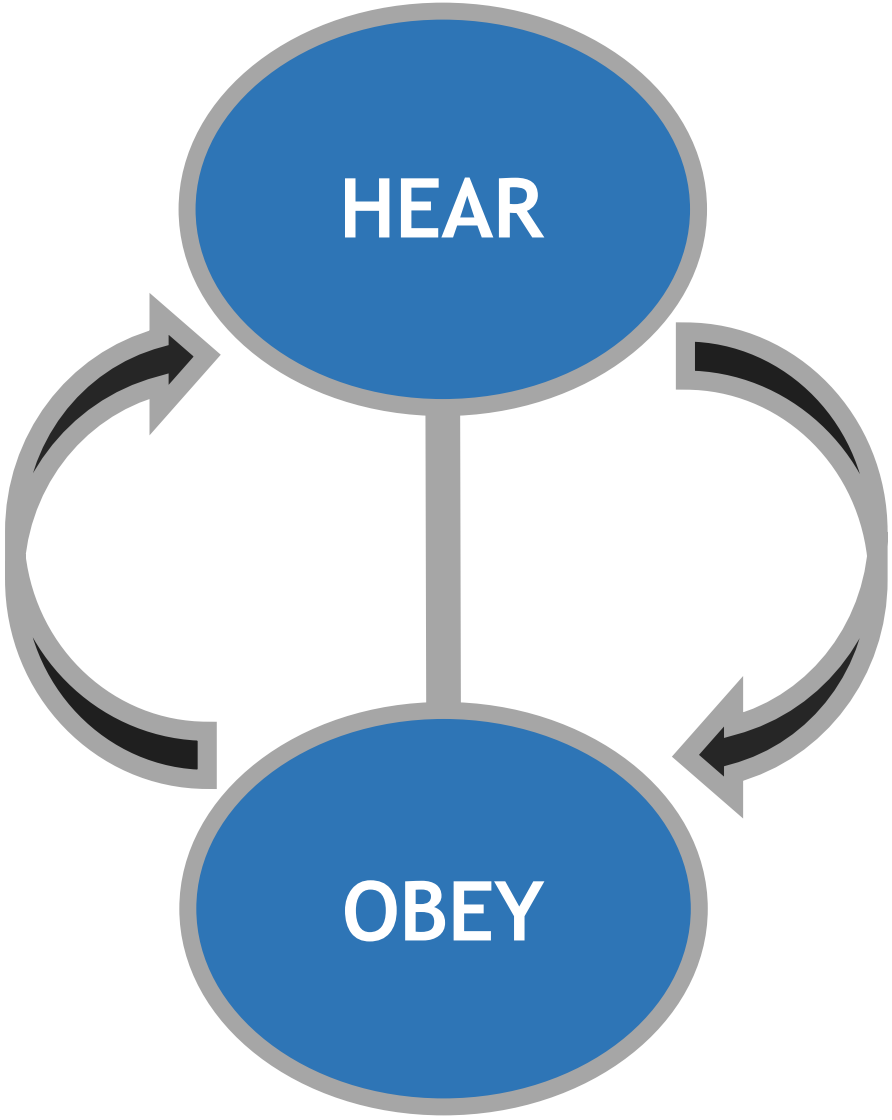
LIÈGE université

# Goal of the talk

**Can semantic change be predicted from synchronic data?**

# Introductory remarks

# Introductory remarks




?

# Introductory remarks

## Why is semantic change difficult to model?

- Meanings change in many possible directions:
  - specific → general
  - general → specific
  - positive → negative
  - negative → positive
  - concrete → abstract
  - concrete → concrete

- Semantic change has often appeared:
  - Irregular, unpredictable, less systematic than phonological or morphological change

"by and large, semantic change operates in a rather random fashion, affecting one word here (in one way) and another form there (in another way)" (Hock, 1991)

# Introductory remarks

## From unpredictability to regularity

- **Research in the last 30 years**
  - Studies have identified regular patterns in semantic evolution (e.g., Sweetser, 1990; Traugott & Dasher, 2002)

- **Semantic change is frequently unidirectional**
  - **Examples:**
    - objective → subjective
    - concrete → abstract
    - lexical → grammatical

- **Mechanisms are increasingly well understood**
    - subjectification
    - metaphor
    - asymmetric priming (Jäger & Rosenbach, 2008)

**Bottom line:** Semantic change is not entirely random

(see Baumann & Hartmann, 2026; Li & Siew, 2022; List, 2023; Urban, 2011; Winter & Srinivasan, 2022; Xu et al., 2017)

# Introductory remarks

- **What we still cannot do well**

  We still lack robust predictive models of semantic change

- **What do we need**

  1. Reliable predictors
  2. Access to large-scale synchronic resources with broad lexical and conceptual coverage for statistical modelling
  3. Extensive diachronic datasets documenting historical semantic shifts across languages

# Today's presentation

1. **Comparing two predictive approaches**
   a) *Approach 1 – Causal Inferences*
      Relies directly on coexpression patterns, applying causal inference algorithms to derive diachronic trajectories from synchronic distributions (Dellert, 2016; 2019; 2024)
   b) *Approach 2* – PredictMe
      Directionality is predicted from psycholinguistic properties attached to the meanings themselves
      (cf. Winter & Srinivasan, 2022; Carling et al., 2023; Baumann & Hartmann, 2026)

2. **How well do the predictions agree with independent diachronic evidence?**
   - Greek, Ancient Egyptian, DatSemShift

**Case study**

PERCEPTION–COGNITION

(cf. Georgakopoulos et al., 2022; Norcliffe & Majid, 2024)

# Perception-Cognition
## Concepts and synchronic data

# Concepts and synchronic data

## Classification of concepts (semantic field)

Concepticon (see List et al., 2016)

Concepticon | Home | Concepts | Concept sets | Concept lists | Languages | Compilers | Sources

### Concept sets

Showing 1 to 21 of 21 entries (filtered from 4,034 total entries)

← Previous 1 Next →

| ID | Name | Definition | Semantic field | Ontological category |
|---|---|---|---|---|
| Search | Search | Search | Sense perc | Action/Proc |
| 21 | TASTE (SOMETHING) | To try the flavor of something. | | |
| 252 | FRIGHTEN | To make someone afraid or anxious. | | |
| 375 | SNIFF | To make a short, audible inhalation, through the nose, as if to smell something. | | |
| 378 | PINCH | To squeeze a small amount of a person's skin and flesh, making it hurt. | | |
| 708 | SHINE | To emit light. | | |
| 1039 | SHOW | To have somebody see something. | | |
| 1079 | FEEL | To be in some emotional state. | | |
| 1408 | HEAR | To perceive sound with the ear, without necessarily paying attention to it. | | |

--any--
Agriculture and vegetation
Animals
Basic actions and technology
Clothing and grooming
Cognition
Emotions and values
Food and drink
Kinship
Law
Miscellaneous function words
Modern world
Motion
Possession
Quantity
Religion and belief
Sense perception

# Concepts and synchronic data

**Classification of concepts (semantic field)**

Concepticon (see List et al., 2016)

Concepticon | Home | Concepts | Concept sets | Concept lists | Languages | Compilers | Sources

## Concept sets

Showing 1 to 21 of 21 entries (filtered from 4,034 total entries)

← Previous 1 Next →

| ID | Name | Definition | Semantic field | Ontological category |
|---|---|---|---|---|
| Search | Search | Search | Sense perc | Action/Proc |
| 21 | TASTE (SOMETHING) | To try the flavor of something. | Sense perception | |
| 252 | FRIGHTEN | To make someone afraid or anxious. | Sense perception | |
| 375 | SNIFF | To make a short, audible inhalation, through the nose, as if to smell something. | Sense perception | |
| 378 | PINCH | To squeeze a small amount of a person's skin and flesh, making it hurt. | Sense perception | Action/Process |
| 708 | SHINE | To emit light. | Sense perception | Action/Process |

--any--
Action/Process
Classifier
Number
Other
Person/Thing
Property

# Concepts and synchronic data

| | A | B | C |
|---|---|---|---|
| 1 | concepticon_id | concept_label | Concepticon_sf_label |
| 2 | 1890 | BELIEVE | Cognition |
| 3 | 1902 | BETRAY | Cognition |
| 4 | 1975 | CHECK | Cognition |
| 5 | 3230 | DECEIVE | Cognition |
| 6 | 1626 | EXPLAIN | Cognition |
| 7 | 1079 | FEEL | Sense perception |
| 8 | 1523 | FORGET | Cognition |
| 9 | 1932 | GUESS | Cognition |
| 10 | 1408 | HEAR | Sense perception |
| 11 | 2411 | HURT (SOMEBODY) | Sense perception |
| 12 | 1410 | KNOW (SOMETHING) | Cognition |
| 13 | 504 | LEARN | Cognition |
| 14 | 1608 | LISTEN | Sense perception |
| 15 | 1819 | LOOK | Sense perception |
| 16 | 378 | PINCH | Sense perception |
| 17 | 1409 | SEE | Sense perception |
| 18 | 2124 | SMELL | Sense perception |
| 19 | 1586 | SMELL (PERCEIVE) | Sense perception |
| 20 | 1587 | SMELL (STINK) | Sense perception |
| 21 | 1929 | STUDY | Cognition |
| 22 | 1841 | SUSPECT | Cognition |
| 23 | 21 | TASTE (SOMETHING) | Sense perception |
| 24 | 709 | TEACH | Cognition |
| 25 | 2271 | THINK | Cognition |
| 26 | 1892 | TOUCH | Sense perception |
| 27 | 1812 | TRY | Cognition |
| 28 | 1536 | UNDERSTAND | Cognition |
| 29 | 1635 | OBEY | Social political relations |
| 30 | 2559 | FEEL (TACTUALLY) | Basic actions technology |
| 31 | 2893 | NOTICE (SOMETHING) | Basic actions technology |

**Table 1.** The 30 concepts from the domains of COGNITION and SENSE PERCEPTION included in the analysis

✓ **Semantic field**

- At least one concept of the pair classified as **COGNITION** or **SENSE PERCEPTION** (three exceptions)

✓ **Ontology**

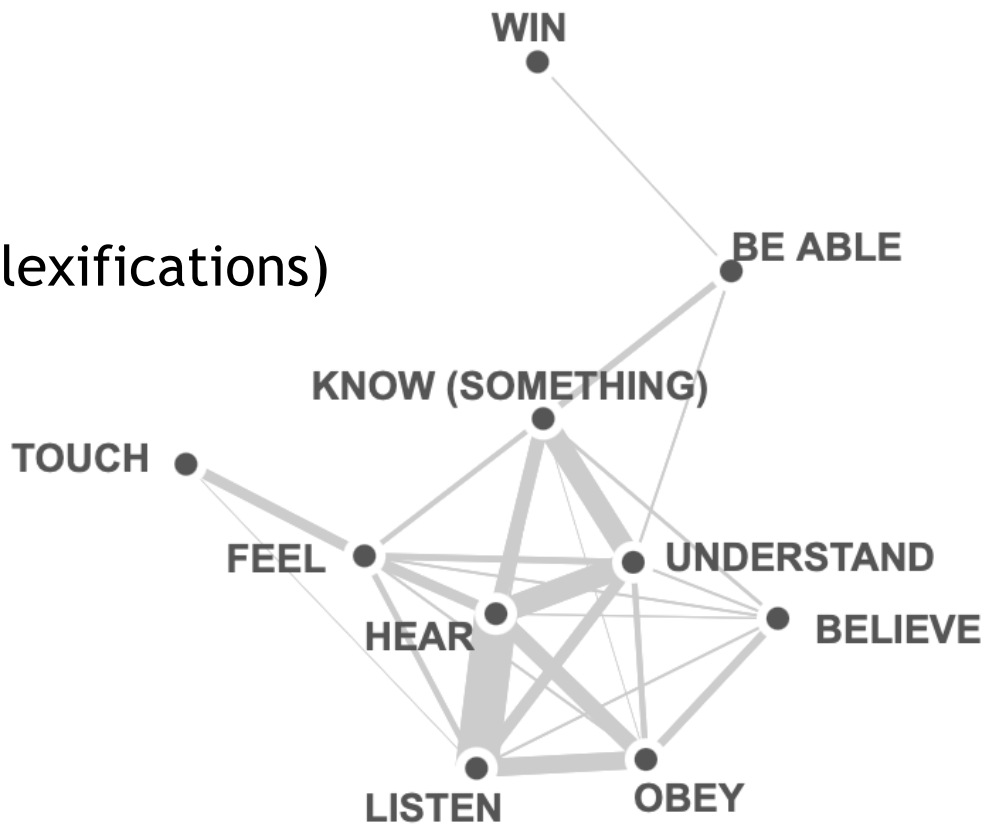
- Both concepts classified as **ACTION/PROCESS**

→ 30 concepts, the shared basis for both methods

# Concepts and synchronic data

**Synchronic data sourced from** CLICS³

- CLICS³ database (see Rzymski et al., 2020)
  - 3156 language varieties
  - Network:
    - 2,919 nodes
    - 4,228 edges (colexifications)



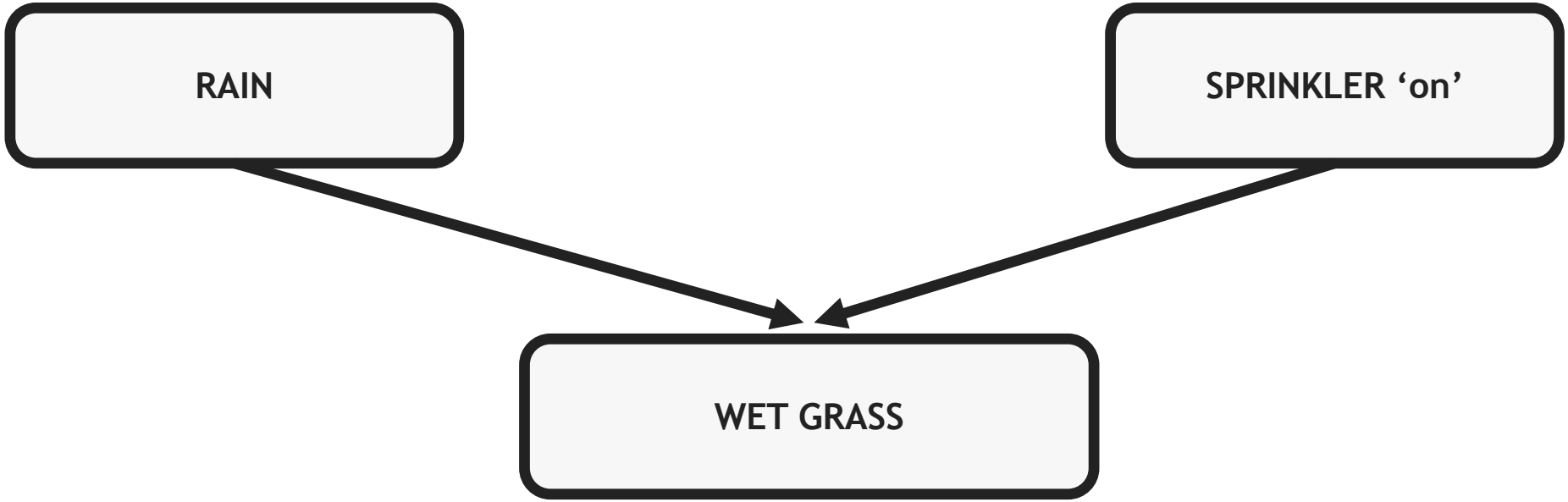
# Approach 1:
# *Causal inference on coexpression patterns*

# What is Causal Inference?

**Causal inference**: recover causal structure from observation alone, using patterns of (conditional) independence.

**The collider signature** ("explaining away"): two independent causes become dependent once their common effect is observed.

*Rain and Sprinkler are independent — until we observe that the grass is wet.*




No collider → different structures can be statistically indistinguishable: direction is not always recoverable.

# From the principle to language

- Meanings = variables; colexification across languages = co-occurrence.
- The PC* algorithm (Spirtes, Glymour & Scheines 2000), applied to colexification by Dellert (2016, 2024).
- It orients an edge only where a collider licenses it; otherwise the edge is left bidirectional.
- Stability via bootstrap resampling (N = 1,000).

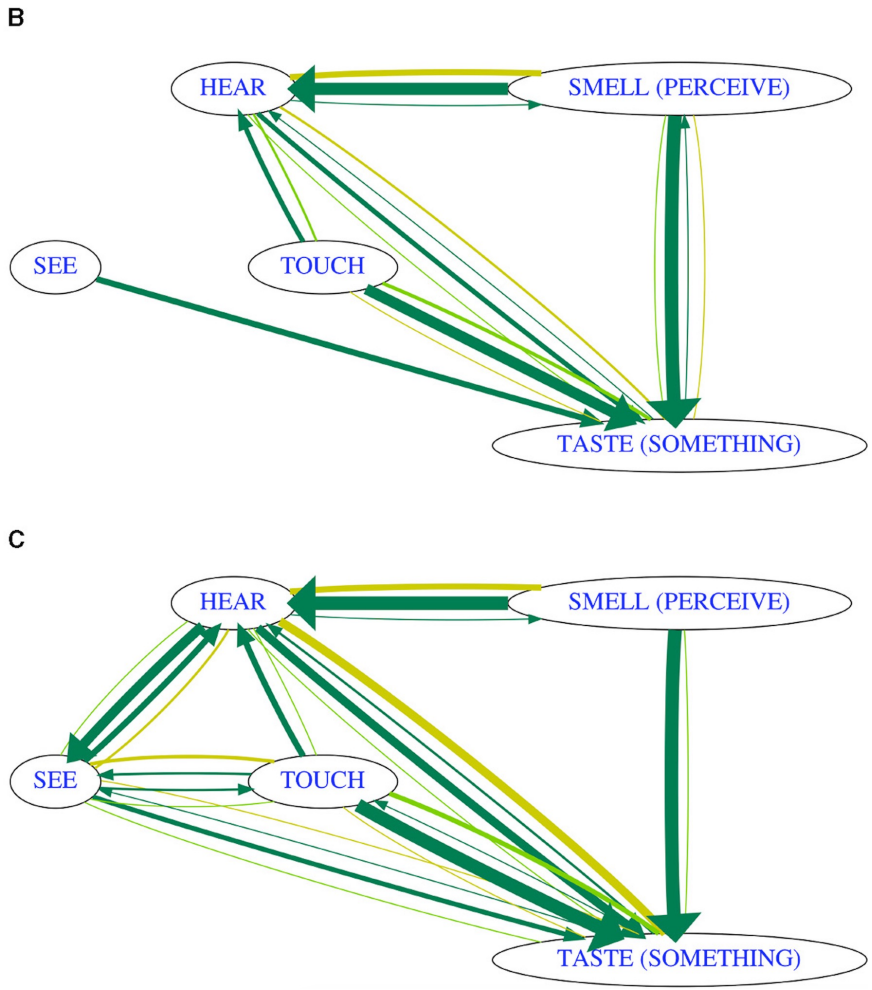



**Figure 1**. Dellert (2024: Fig. 4). (b) 293 languages, all meanings attested; (c) 1,071 languages, at most two meanings missing.

# Causal Inference - mincolex & open network

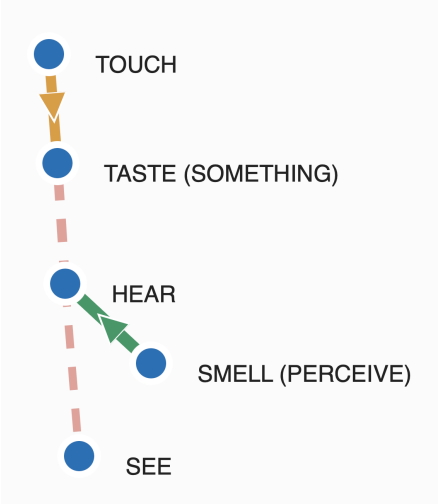



Figure 2a. mincolex 10

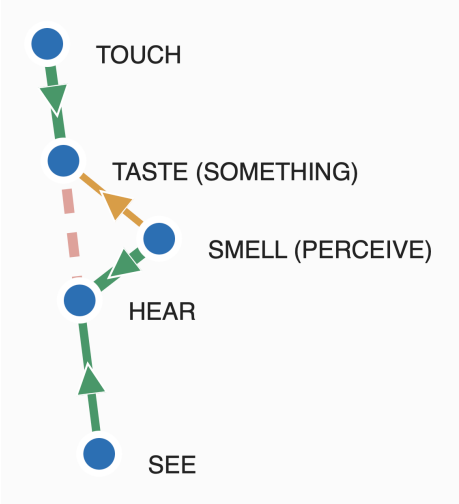



Figure 2b. mincolex 5

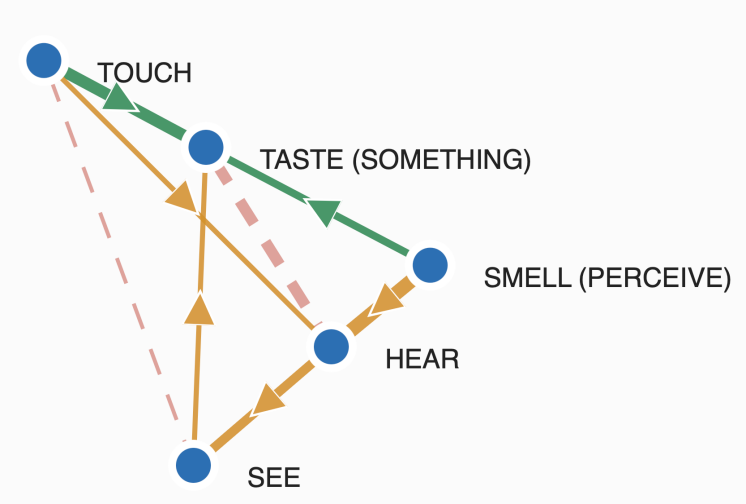



Figure 2c. mincolex 1

# Causal Inference - mincolex & open network

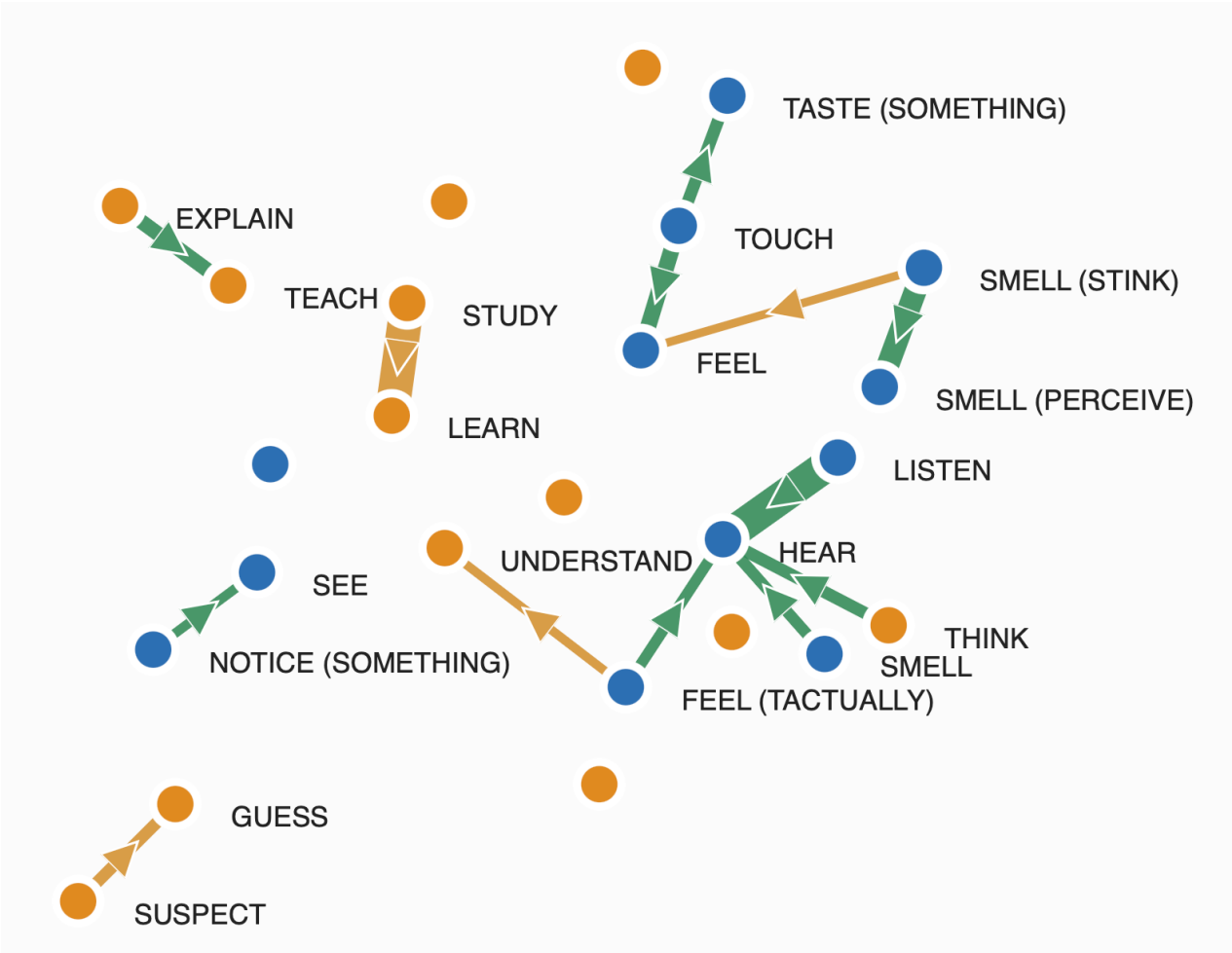



Figure 3. Closed network — directed edges only (mincolex 5)

# Causal Inference - mincolex & open network

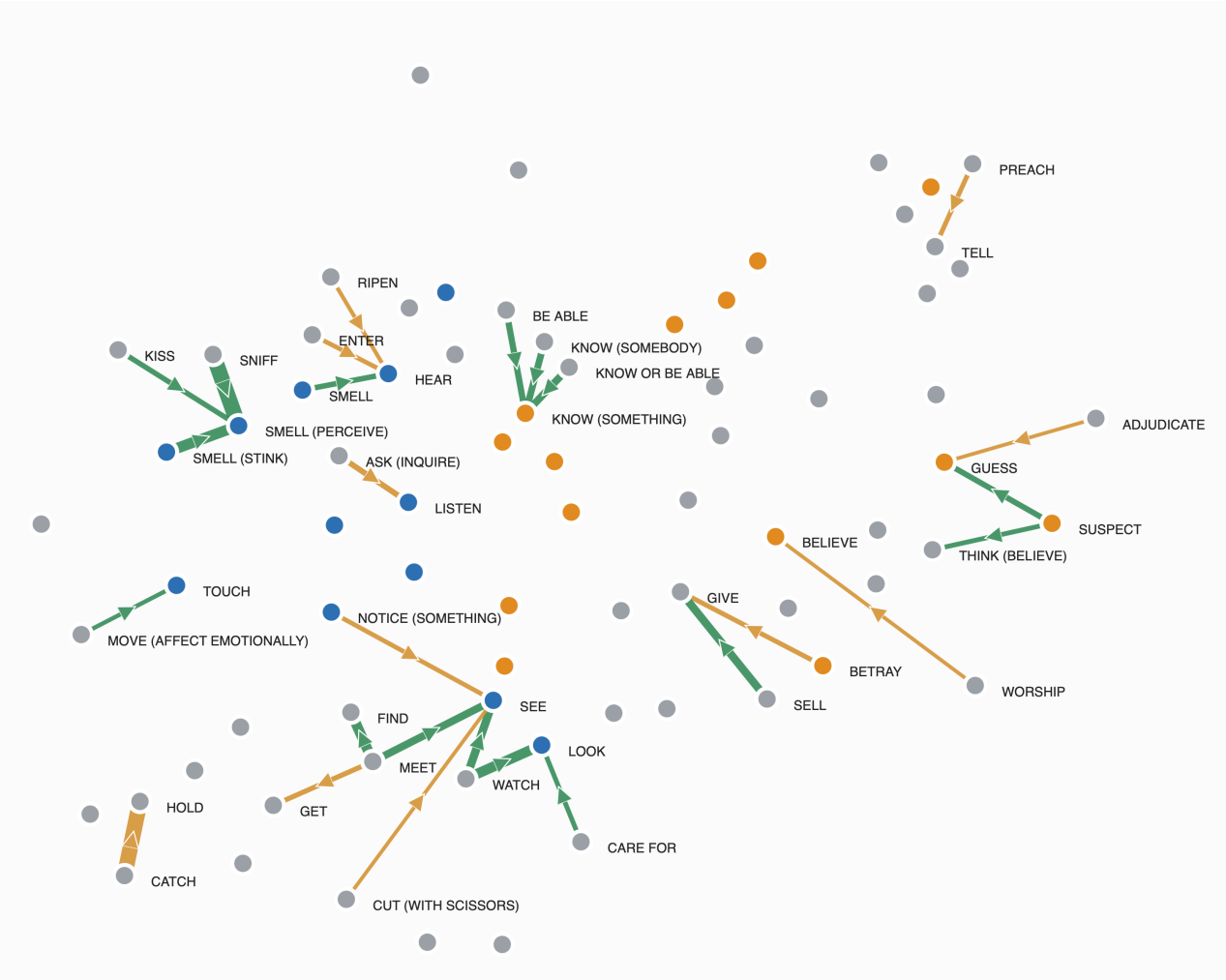



**Figure 4.** Open network — directed edges only (mincolex 5)

**Admitting the periphery is productive.** At Action·Process / mincolex 5 (227 edges): 18 directed, 194 bidirectional. The robust orientations are coherent *specific* → *general* extensions:

- **Perception** (cline parallels a drop in agentivity): WATCH → SEE 0.99, SNIFF → SMELL 1.00, WATCH → LOOK 0.96, KISS → SMELL 0.98, CARE FOR → LOOK 0.89.
- **Cognition**: SUSPECT → GUESS 0.94, SUSPECT → THINK 0.91.
- Signal converges by N ≈ 250 resamples.

# Causal Inference - take away

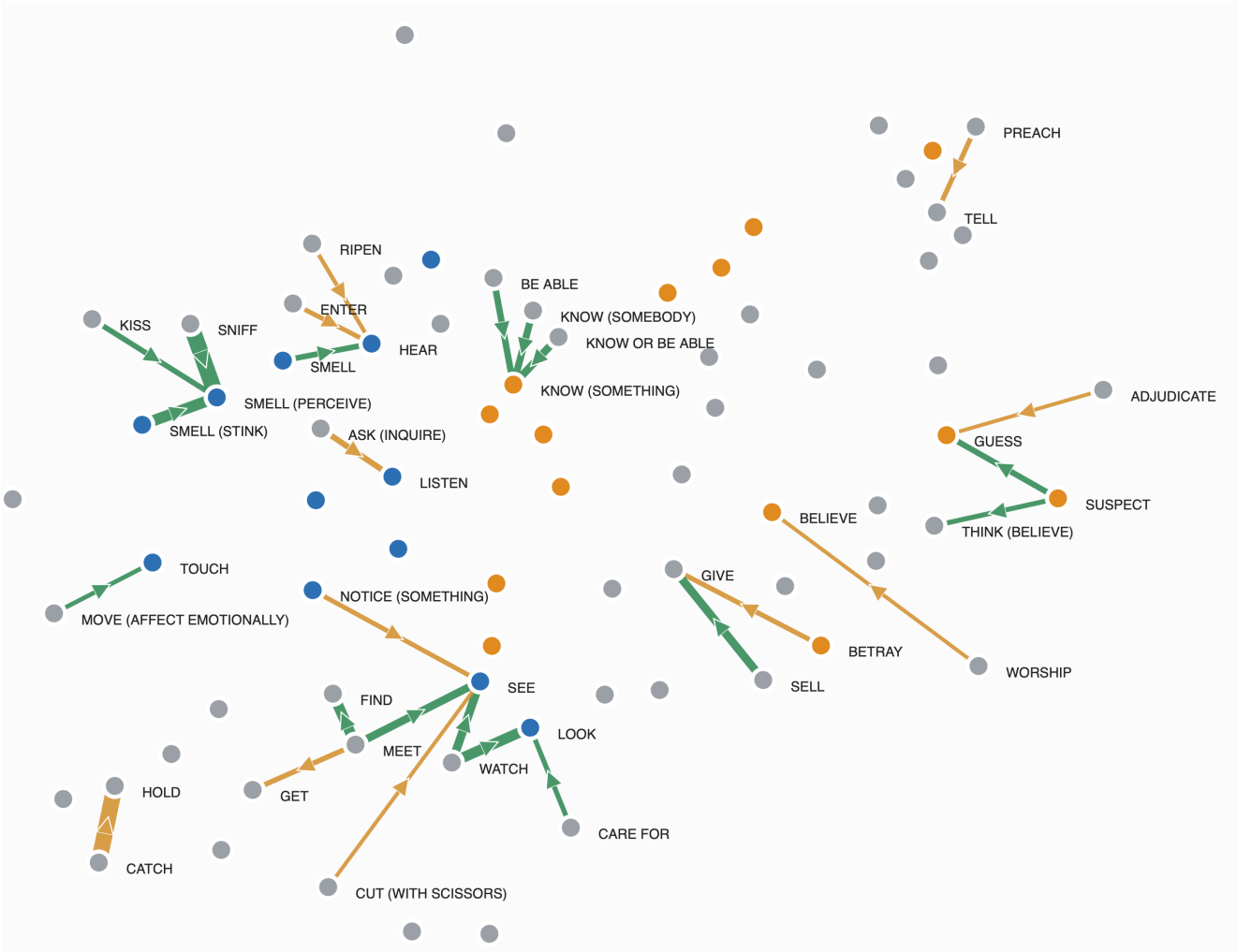


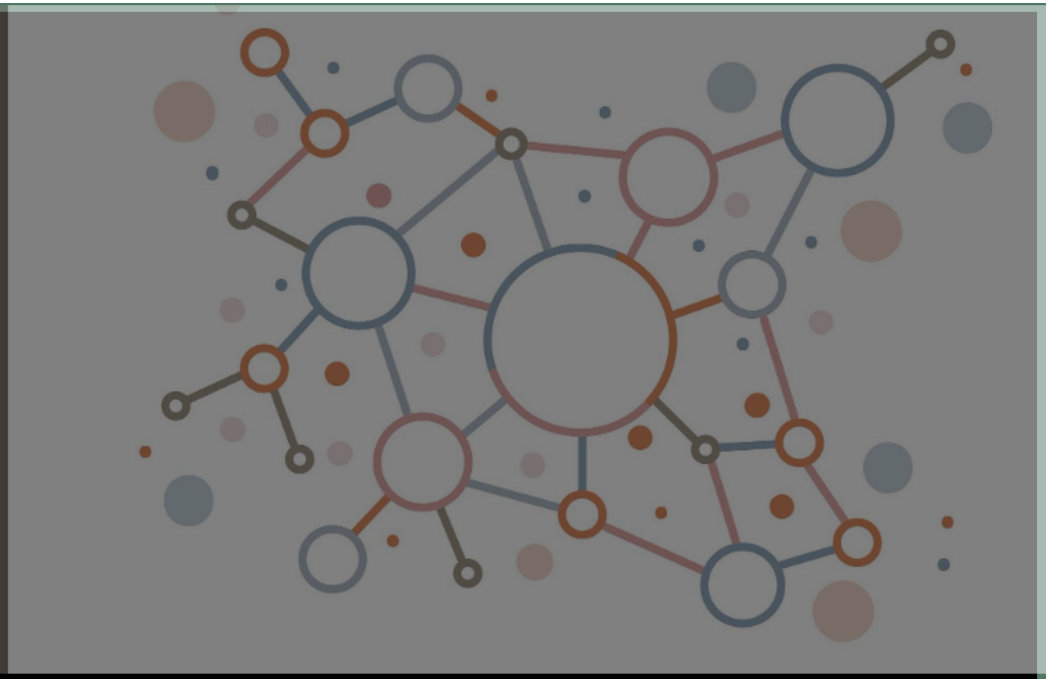

Figure 5. Open network — directed edges only (mincolex 5)

- Recovers a semantic skeleton and *credible source* → *domain* extensions that needs to be evaluated against diachronic data
- But abstains on every perception ↔️ cognition bridge (0 oriented)

**PredictMe addresses these questions**

PredictME

From Networks to Semantic Change:
A Multifactorial Approach to the
Predictability of Meaning Evolution



# Approach 2:
# *Concept properties*

# Predict directionality from psycholinguistic properties

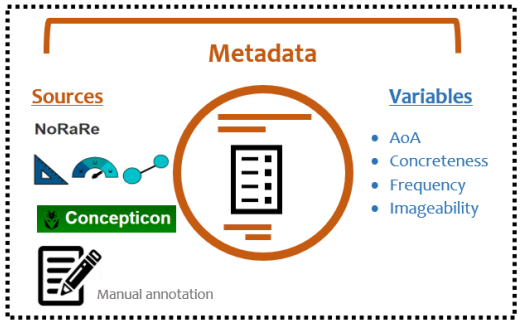
# Predictive model construction: The synchronic data

**NoRaRe** (see Tjuka et al., 2022)

**Focus on four psycholinguistic variables**

a) Concreteness
b) Imageability
c) Frequency
d) Age of Acquisition

- These factors serve as predictors of semantic change direction
- Each concept must have at least one value for each predictor.




Tjuka, Annika, Robert Forkel, and Johann-Mattis List. 2022. Linking Norms, Ratings, and Relations of Words and Concepts Across Multiple Language Varieties. Behavior Research Methods 54. 864–884. DOI: 10.3758/s13428-021-01650-1

# Predictive model construction: The synchronic data

NoRaRe

| Variable | Number of datasets | Number of languages | Languages |
|---|---|---|---|
| Frequency | 11 | 20 | Basque, Chilean Spanish, Chinese, Czech, Dutch, English, Estonian, Finnish, French, German, Greek, Italian, Māori, Polish, Portuguese, Russian, Spanish, Swahili, Tok Pisin, Turkish |
| Concreteness | 9 | 6 | Chinese, English, French, Italian, Polish, Spanish |
| Imageability | 11 | 6 | Chinese, English, German, Italian, Polish, Spanish |
| Age of Acquisition (AoA) | 7 | 26 | Arabic, Catalan, Croatian, Danish, Dutch, English, French, German, Greek, Hebrew, Hungarian, Italian, Kannada, Malay, Mandarin Chinese, Norwegian, Persian, Polish, Portuguese, Serbian, Slovak, Spanish, Swedish, Turkish, Ukrainian, Vietnamese |

**Table 2.** Datasets and language coverage of the psycholinguistic predictors

# Predictive model construction: The synchronic data

- **Summary: final dataset for analysis**

**Inclusion criteria**

✓ **Semantic field**

- At least one concept classified as **COGNITION** or **SENSE PERCEPTION**

✓ **Ontology**

- Both concepts classified as **ACTION/PROCESS**

✓ **Psycholinguistic data**

- Each concept must have at least one value for each predictor

⟶ 1919 colexification pairs

# Predictors of semantic change

# Predictors of semantic change

💡 **Concreteness**

*H1*: In a colexification pair, the concept with the highest concreteness score will most likely be the source meaning in a process of change (see Xu et al., 2017; cf. Winter and Srinivasan, 2022)

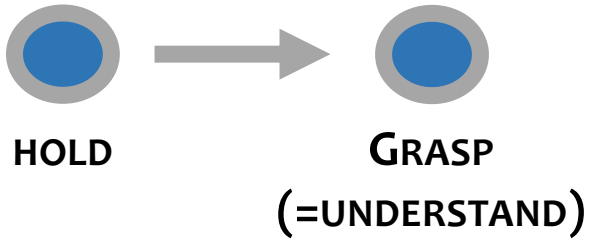
🧠🔁

See the asymmetrical relationship between concrete and abstract entities, as theorized in Conceptual Metaphor Theory (e.g., Lakoff & Johnson, 1980; cf. Dunn 2015; on diachrony, see Heine et al.1991)

# Predictors of semantic change

## 💡 Concreteness

*H1*: In a colexification pair, the concept with the highest concreteness score will most likely be the source meaning in a process of change (see Xu et al., 2017; cf. Winter and Srinivasan, 2022)




| Dataset in NoRaRe | Property | HOLD | GRASP (=UNDERSTAND) | RATING SCALE |
|---|---|---|---|---|
| Scott-2019-Concreteness-ENGLISH | Concreteness | 4.303 | 3.853 | 7-point |

**Variable ENGLISH_CONCRETENESS_MEAN**

Number of Scale Points: 7 Concreteness is a measure of how concrete or abstract something is.

# Predictors of semantic change

## Imageability

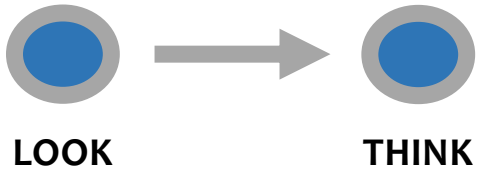
*H2*: In a colexification pair, the concept with the highest imageability score will most likely be the source meaning in a process of change

Domains that are less imageable (e.g., TIME) generally depend on more imageable domains (e.g., SPACE). In line with many Conceptual Metaphor Theorists (e.g., Lakoff & Johnson, 1980)

# Predictors of semantic change

## 💡 Imageability

*H2*: In a colexification pair, the concept with the highest imageability score will most likely be the source meaning in a process of change




| Dataset in NoRaRe | Property | LOOK | THINK | RATING SCALE |
|---|---|---|---|---|
| Clark-2004-ImageryFamiliarity-ENGLISH_IMAGEABILITY_MEAN | Imageability | 3 | 2.59 | 7-point |

**Variable ENGLISH_IMAGEABILITY_MEAN**

Mean value of imageability ratings based on a 7-point scale.

# Predictors of semantic change

## 💡 Frequency

*H3*: In a colexification pair, the concept with the highest frequency score* is more likely to be the source meaning

* As indexed by word frequency metrics from various languages

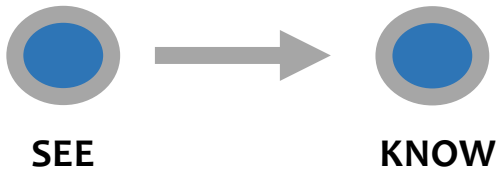
- Important role in the studies of the lexicon (Zipf, 1949)
- Significant factor in understanding various aspects of language behavior (see Bybee, 2007)
- May function as an important driver of language change (Pagel et al. 2007; cf. Mooers & Pappas, 2011 for a critique).
- Higher-frequency words tend to be more polysemous (Zipf, 1945; Casas et al., 2019)

→ Frequent meanings are more likely to serve as source meanings for semantic extensions (see Winter & Srinivasan, 2022)

# Predictors of semantic change

## Frequency

*H3*: In a colexification pair, the concept with the highest frequency score is more likely to be the source meaning (see Winter & Srinivasan, 2022)




| Dataset in NoRaRe | Property | SEE [SP._VER] | KNOW [SP_SABER] |
|---|---|---|---|
| Cuetos-2011-Frequency-SPANISH | frequency | 4.659535907154217 | 4.319189369204891 |

**Variable SPANISH_FREQUENCY_LOG**

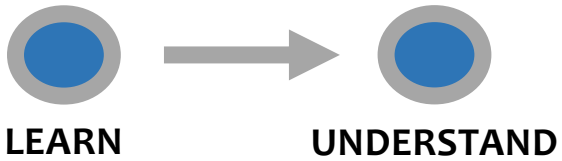
Log10 of subtitle word frequency.

# Predictors of semantic change

## 💡 Age of Acquisition

*H4*: In a colexification pair, the concept that is acquired earlier will most likely be the source meaning.

(see also Baumann & Hartmann, 2026)




In a process of change that involves the pair <LEARN, UNDERSTAND>, we assume that LEARN will most likely be the source meaning, since this is acquired earlier than UNDERSTAND

| Dataset in NoRaRe | Property | LEARN [SP._APRENDER] | UNDERSTAND [SP._ENTENDER] |
|---|---|---|---|
| Alonso-2015-AoA-SPANISH | AoA | 4.54 | 5.44 |

Variable SPANISH_AOA_MEAN

Average score of AoA ratings on a scale from 1 to 11.

# Predictor preprocessing and source score computation

# Predictor preprocessing & source score computation

## Data preprocessing

1. Data harmonization

   **a. Scale alignment**
   - Reverse-code variables whose direction differed
     - Imbir (2021) concreteness:
       - original: higher = **less** concrete

   **b. Standardization**
   - z-transformation
   - performed **within each dataset**

# Predictor preprocessing & source score computation

## Data preprocessing

2. Aggregation

**Computing concept-level predictor scores**

**For each predictor**

- retain all available standardized scores
- average across datasets

**Result**

- Frequency_mean
- Concreteness_mean
- Imageability_mean
- AoA_mean

**Finally**

- reverse-code AoA_mean
- higher values = earlier acquisition

# Predictor preprocessing & source score computation

## Source score

- Predicting the originating sense
  (each predictor contributes equally)

$$S_i = \frac{C_i + F_i + A_i + I_i}{4}$$

**where**

- C = concreteness
- F = frequency
- A = reverse-coded AoA
- I = imageability

**For each colexification pair**

- compare the two source scores
- **higher score → predicted source**

# Predictor preprocessing & source score computation

## Confidence threshold

- Confidence-based prediction

For each pair,

$$\Delta = |S_1 - S_2|$$

**Large Δ**
→ clear prediction
**Small Δ**
→ concepts receive very similar scores
(excluded from directional interpretation)

# Predictor preprocessing & source score computation

## Confidence threshold

- Confidence-based prediction

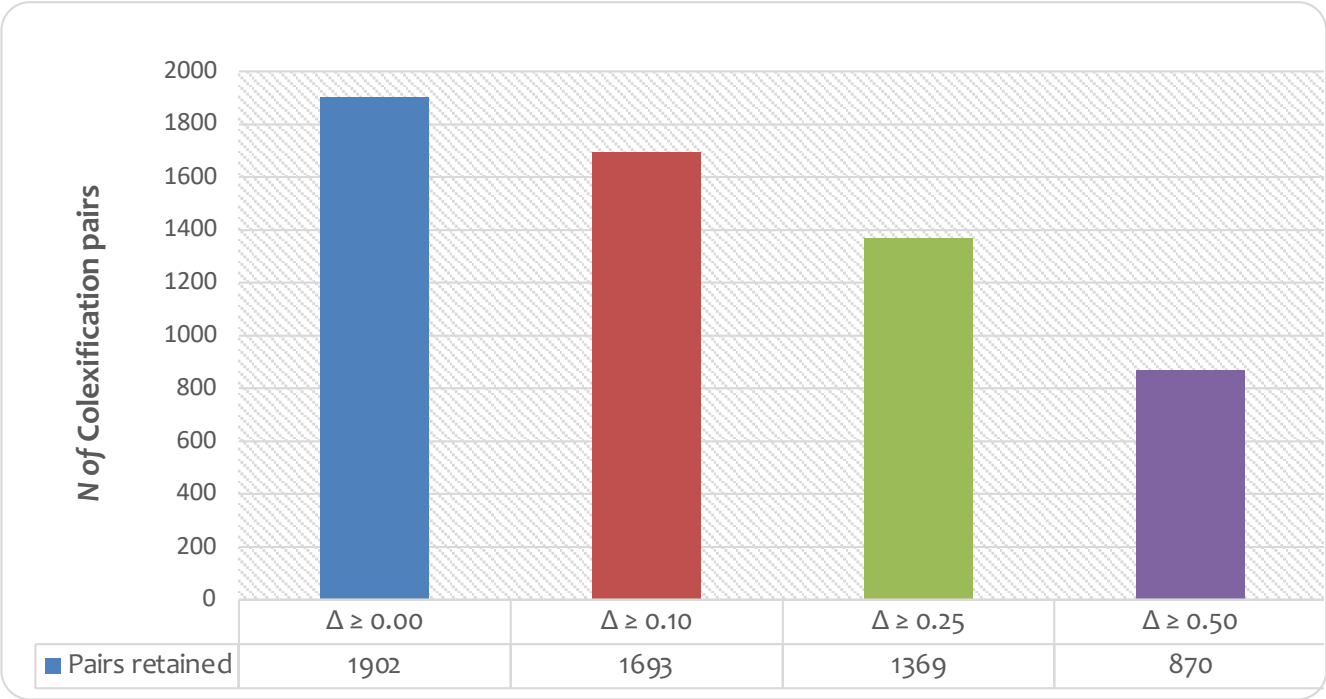



**Table 3**. Number and proportion of colexification pairs retained under alternative confidence thresholds (Δ)

- **Final choice:**

$$\Delta \geq 0.25$$

- Removes weak predictions
- Preserves most of the dataset

# Predictor preprocessing & source score computation

## How the predictive model works: two examples

Correct prediction | Arrow = predicted direction (from left to right concept) | = Concept

**Source score (S) = (C + F + A + I) / 4**

C = Concreteness, F = Frequency, A = reverse-coded AoA, I = Imageability

### 1 KNOW (SOMETHING) → ADMIT

| KNOW (SOMETHING) | ADMIT |
|---|---|
| Source score (S) | Source score (S) |
| -0.50 | -1.02 |

-0.50 > -1.02 → Predict KNOW (SOMETHING) as source (matches attested direction)

### 2 SEE → FIND

| SEE | FIND |
|---|---|
| Source score (S) | Source score (S) |
| 0.52 | -0.23 |

0.52 > -0.23 → Predict SEE as source (matches attested direction)

### How the source score is computed

1. For each concept, obtain standardized (z-scored) values for: Concreteness (C), Frequency (F), Reverse-coded AoA (A), Imageability (I)
2. Average the four aggregated z-scores:
   S = (C + F + A + I) / 4
3. For each pair, compare the two source scores: Higher score = predicted source of semantic change.

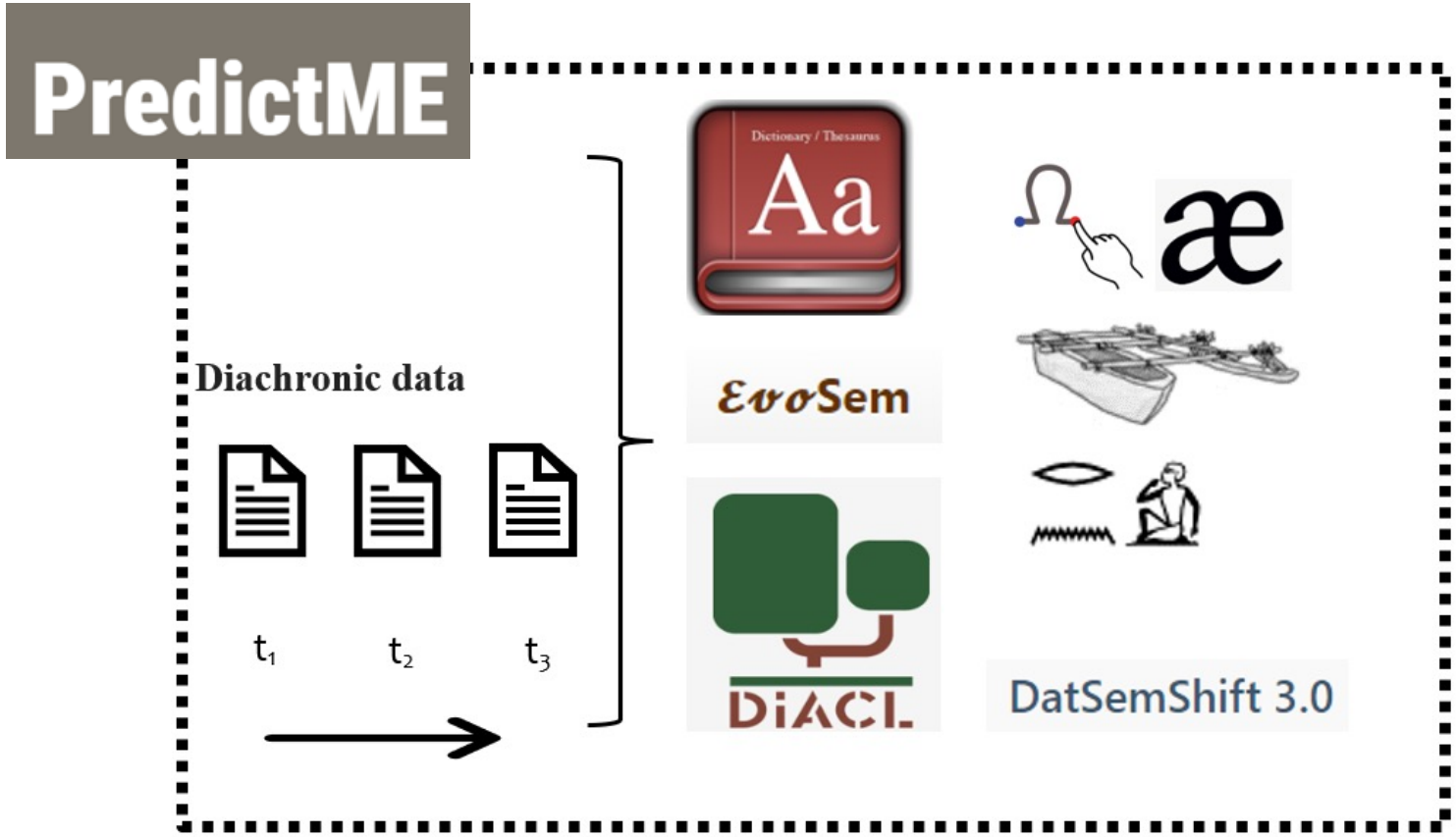
Example: KNOW (SOMETHING)

| Predictor | Aggregated z-score |
|---|---|
| Frequency (F) | 1.08 |
| Concreteness (C) | -1.81 |
| Imageability (I) | -1.19 |
| Reverse-coded AoA (A) | -0.08 |

Source score (S) = (C + F + A + I) / 4 = (-1.81 + 1.08 - 0.08 - 1.19) / 4
= (-2.00) / 4
= -0.50

# Evaluation of the Models

# Model evaluation:
# Diachronic validation data



# Model evaluation: Diachronic validation data

## TODAY: data from DatSemShift

- Database of Semantic Shifts in languages of the world (Zalizniak et al., 2012; Zalizniak ,2018; DatSemShift 3.0, https://datsemshift.ru/)

| Meaning pair | Form(s) | Language(s) | Realization type |
|---|---|---|---|
| **hear → obey** | šemû | Akkadian | Synchronic polysemy |
| **hear → understand** | entendre → entender | French → Spanish | Cognate |
| **arc → rainbow** | Bogen → Regenbogen | German | Morphological derivation |
| **moon → Monday** | mānadagher → maanantai | Old Swedish → Finnish | Borrowing |
| **perceive smell → know** | sapio → savoir | Latin → French | Diachronic sem. evolution |

Table 4. The different realization types of semantic shifts in the DatSemShist

Database of Semantic Shifts in languages of the world

DatSemShift 3.0

# Model evaluation: Diachronic validation data

## TODAY: data from Greek

- Proceed **onomasiologically**: identify the lexemes in Ancient Greek that express meanings in SENSORY PERCEPTION and COGNITION
- **Time span:** (8th c. BC – 17th c. AD)

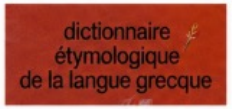
### Lexical Resources

- **Dictionaries**

- Brill Dictionary (Montanari, 2015)

- LSJ (Liddell & Scott, 1996)

- Cambridge Greek Lexicon (Faculty of Classics, University of Cambridge & Diggle, 2021)




- Chantraine's (1968) Etymological dictionary

- Kriaras (1968-) Dictionary of Medieval Greek Vernacular Literature (1100–1669)

- Relevant studies on (ancient) Greek PERCEPTION and COGNITION verbs (e.g. Luraghi, 2020; Luraghi & Sausa, 2019)

# Model evaluation: Diachronic validation data

| Domain | Concepts | Ancient Greek Words |
|---|---|---|
| PERCEPTION | SEE | *horáō* |
| PERCEPTION | SEE | *dérkomai* |
| PERCEPTION | SEE | *leússō* |
| PERCEPTION | SEE | *sképtomai* |
| PERCEPTION | HEAR | *klúō* |
| PERCEPTION | HEAR | *aíō* |
| PERCEPTION | HEAR | *akoúō* |
| COGNITION | THINK (REFLECT) | *phronéō* |
| COGNITION | UNDERSTAND | *ennoéō* |
| COGNITION | KNOW | *oîda, gignṓskō* |
| COGNITION | LEARN | *manthánō* |

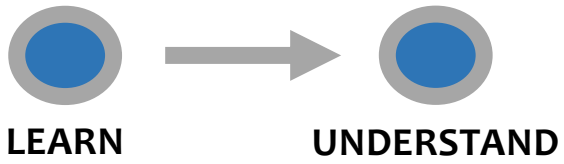
**Table 5.** A subset of Ancient Greek lexemes expressing the concepts in the domains of SENSE PERCEPTION and COGNITION

# Model evaluation: Diachronic validation data

- **Proceed semasiologically**
  - List the different meanings of the lexemes identified (in different stages)

**Diachronic data: Greek**

| Lexeme | Diachronic stage | Meaning 1 'understand' | Meaning 2 'learn' |
|---|---|---|---|
| *manthánō* | 1_Homeric Greek | 0 | 1 |
| | 2_Classical Greek | 1 | 1 |



# Model evaluation: Diachronic validation data

- **Proceed semasiologically**
  - List the different meanings of the lexemes identified (in different stages)

(1)

| | | | | | | |
|---|---|---|---|---|---|---|
| ἀλλ’ | ἐπεὶ | οὖν | δὴ | ἔργα | κάκ᾽ | ἔμμαθεν |
| *all’* | *epeì* | *oûn* | *dḗ* | *érga* | *kák’* | *émathen* |
| but | when | PTCL | PTCL | deed(N):ACC.PL | evil:ACC.PL.N | learn:AOR.3SG |

‘But since he has learned only deeds of evil’ (Hom., *Od.* 22.226)

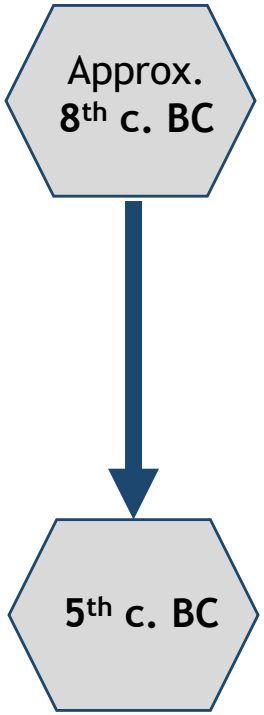
(2)

| | | |
|---|---|---|
| **Dionysus** | μανθάνεις; | |
| | *manthánеis* | |
| | learn:PRS.2SG | |
| **Xanthias** | πάνυ | μανθάνω |
| | pánu | *manthánō* |
| | ADV | learn:PRS.1SG |

‘**Dionysus**: do you understand?

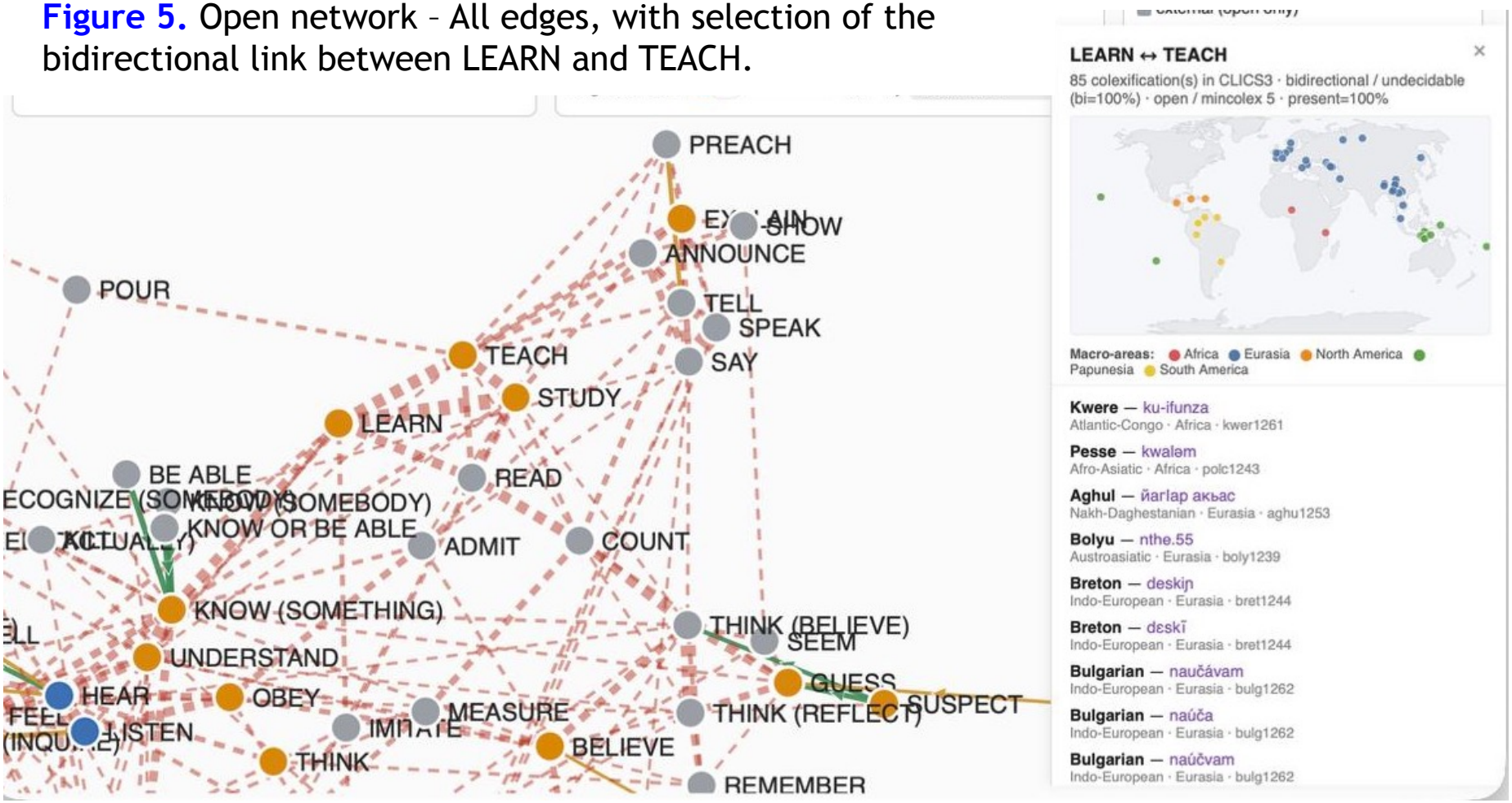
**Xanthias:** I understand perfectly’ (Aristoph. *Frogs* 195)

Context: Dionysus is checking whether Xanthias has caught his point/joke



# Model evaluation:
# Greek & DatSemShift record vs. Causal inference

**Figure 5.** Open network – All edges, with selection of the bidirectional link between LEARN and TEACH.




- **Of the 32 attested directionalities in Greek & DatSemShift: 0 correspond to oriented edges.**

- Where Greek & DatSemShift attest both directions (KNOW ↔ LEARN, LEARN ↔ TEACH), the bidirectional verdict is faithful to the record.

# Model evaluation:
# Greek & DatSemShift record vs. PredictMe

## External Validation of PredictME

### PREDICTION DATASET
(constructed using colexification pairs with Δ ≥ 0.25)

- Psycholinguistic predictors
- ↓
- Source score computation
- ↓
- Confidence threshold (Δ ≥ 0.25)
- **1,902 colexification pairs** with predicted semantic change directions

### EXTERNAL VALIDATION DATASET

**32** attested semantic change directionalities

- **30** from Ancient Greek
- **2** from DatSemShift

Each directionality corresponds to one Concepticon pair with historical evidence for semantic change.

### INTERSECTION
Directionalities present in both datasets (available for evaluation)

### 13 NOT REPRESENTED
Directionalities not present in the prediction dataset because their corresponding colexification pairs did not meet Δ ≥ 0.25.

(40.6%)

### 19 EVALUATED DIRECTIONALITIES

**12** Correct predictions (63.2%)

**7** Incorrect predictions (36.8%)

### PREDICTION ACCURACY

All evaluated pairs (n = 19)
**12** correct **63.2%**

Excluding bidirectional cases (n = 17)
**12** correct **70.6%**

### BIDIRECTIONAL SEMANTIC CHANGE

- 2 evaluated pairs exhibit historical evidence for semantic change in both directions.
- Because the model predicts a single direction for each pair, these cases were conservatively treated as incorrect. → Excluding these cases increases accuracy to 70.6% (12/17).

# Model evaluation:
# Greek & DatSemShift record vs. PredictMe

## External validation using semantic change directionalities

*(matches Ancient Greek and DatSemShift evidence)*

19 pairs found | 12 correct | 7 incorrect | Accuracy: 63.2%

Correct prediction
Incorrect prediction
Bidirectional (both directions attested)
Arrow = predicted direction (from left to right concept)

### CORRECT PREDICTIONS (12 pairs)

1. KNOW (SOMETHING) → ADMIT
2. CATCH → UNDERSTAND
5. TEACH → EXPLAIN
6. SEE → FEAR (BE AFRAID)
7. UNDERSTAND → FEEL
8. SEE → FIND
10. HEAR → UNDERSTAND
12. LISTEN → LEARN
14. LOOK → THINK
15. LISTEN → UNDERSTAND
18. SEE → UNDERSTAND
19. TOUCH → GET

### INCORRECT PREDICTIONS (7 pairs)

3. HIDE (CONCEAL) → DECEIVE
4. HIDE (ONESELF) → DECEIVE
9. LEARN ↔ HEAR — BIDIRECTIONAL EVIDENCE
11. LEARN ↔ KNOW (SOMETHING) — BIDIRECTIONAL EVIDENCE
13. STUDY → LEARN
16. LOOK → SEE
17. UNDERSTAND → LOOK

Arrow indicates predicted direction of semantic change (from left concept → right concept).

Green = prediction matches the semantic change direction in Ancient Greek and DatSemShift.
Red = prediction does not match the semantic change direction in Ancient Greek and DatSemShift.
Orange = semantic change attested in both directions in Ancient Greek and DatSemShift.

# Model evaluation:
# Greek & DatSemShift record vs. PredictMe

## External validation using semantic change directionalities

*(matches Ancient Greek and DatSemShift evidence)*

*— excludes bidirectionalities —*

17 pairs found | 12 correct | 5 incorrect | Accuracy: 70.6%

Correct prediction

Incorrect prediction

Arrow = predicted direction (from left to right concept)

### CORRECT PREDICTIONS (12 pairs)

1. KNOW (SOMETHING) → ADMIT
2. CATCH → UNDERSTAND
5. TEACH → EXPLAIN
6. SEE → FEAR (BE AFRAID)
7. UNDERSTAND → FEEL
8. SEE → FIND
10. HEAR → UNDERSTAND
12. LISTEN → LEARN
14. LOOK → THINK
15. LISTEN → UNDERSTAND
18. SEE → UNDERSTAND
19. TOUCH → GET

### INCORRECT PREDICTIONS (5 pairs)

3. HIDE (CONCEAL) → DECEIVE
4. HIDE (ONESELF) → DECEIVE
13. STUDY → LEARN
16. LOOK → SEE
17. UNDERSTAND → LOOK

Arrow indicates predicted direction of semantic change (from left concept → right concept).

Green = prediction matches the semantic change direction in Ancient Greek and DatSemShift.

Red = prediction does not match the semantic change direction in Ancient Greek and DatSemShift.

# Model evaluation – Take away (Greek & DatSemShift)

## 0 % vs ~70 %

**Causal inference:** 0 / 32 attested directions are oriented

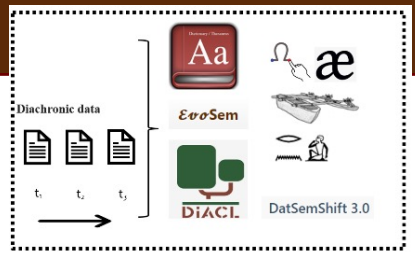
**PredictME:** ~70% correct

- Because PredictME ignores colexification frequency, it can be scored on far more Greek pairs than causal inference

| mincolex | unidirectional correct | accuracy |
|---|---|---|
| ≥ 1 | 12 / 17 (+2 bi-attested) | 70.6 % |
| ≥ 3 | 7 / 10 (+1) | 70.0 % |
| ≥ 5 | 7 / 9 (+1) | 77.8 % |
| ≥ 10 | 4 / 6 (+1) | 66.7 % |

# Limitations and future directions

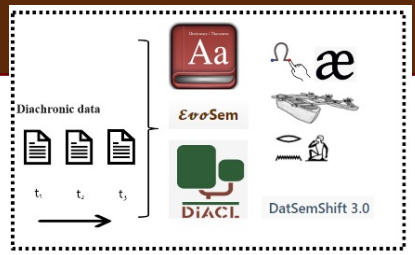
# Limitations and future directions




1. **Optimizing predictor contributions**
   - Estimate the relative contribution of each predictor included in the model
   - Determine which psycholinguistic properties have the greatest predictive validity for the directionality of semantic change
   - Replace equal weighting with empirically derived weights
   - Examine issues of multicollinearity

2. **Refining source score estimation**
   - Examine cases with small score differences
   - Assess whether predictions are driven by a single predictor
   - Introduce robustness and sensitivity analyses

# Limitations and future directions




## 3. Incorporating network structure

- Integrate network-based predictors (e.g. degree, betweenness, closeness, eigenvector centrality)
- Exploit the topology of colexification networks

## 4. Broadening semantic coverage

- Expand beyond the domains of COGNITION and SENSE PERCEPTION
- Extend the model to all semantic domains represented in CLICS and Concepticon

## 5. Expanding diachronic validation

- Extend diachronic validation with additional languages (including data from Ancient Egyptian, English, Greek)
- Incorporate evidence from additional resources (e.g. DatSemShift, Zalizniak, 2018; DiACL, Carling et al., 2018 EvoSem, Dehouck et al., 2023)

# Limitations and future directions

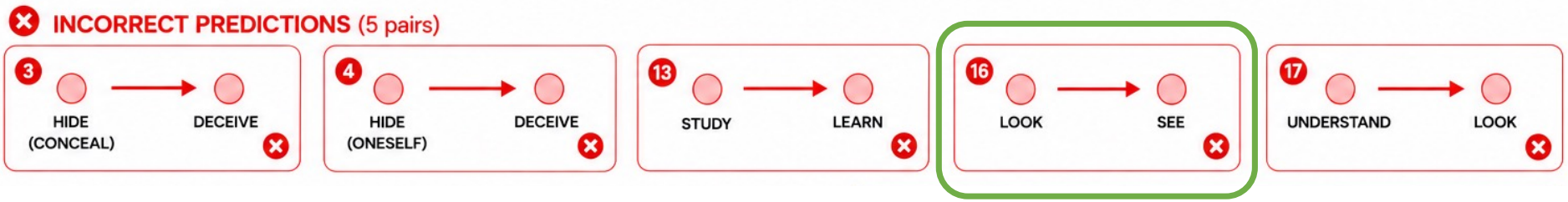
- PredictME orients HEAR → OBEY and LISTEN → OBEY

→ Egyptian corroborates where Greek is silent with
*sḏm* “to hear” > “to obey”

# Limitations and future directions

- PredictME orients HEAR → OBEY and LISTEN → OBEY

→ Egyptian corroborates where Greek is silent with
*sḏm* “to hear” > “to obey”




- Ancient Egyptian also validate predictions that are currently invalidated by the Greek data

*ptr* “to look” → “to see”

*nw* “to watch, to look” → “to see”

# References


Baumann, A., & Hartmann, S. (2026). The chicken and the egg: Unraveling aspects of semantic change and how they relate to lexical acquisition. *Cognition, 258*, Article 106301. https://doi.org/10.1016/j.cognition.2025.10630

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford University Press.

Carling, G., Larsson, F., Cathcart, C. A., Johansson, N., Holmer, A., Round, E., & Verhoeven, R. (2018). Diachronic Atlas of Comparative Linguistics (DiACL)—A database for ancient language typology. *PLOS ONE, 13*(10), e0205313. https://doi.org/10.1371/journal.pone.0205313

Carling, G., Cronhamn, S., Lundgren, O., Bogren Svensson, V., & Frid, J. (2023). The evolution of lexical semantics: dynamics, directionality, and drift. Frontiers in Communication 8. https://doi.org/10.3389/fcomm.2023.1126249

Casas, B., Hernández-Fernández, A., Català, N., Ferrer-i-Cancho, R., & Baixeries, J. (2019). Polysemy and brevity versus frequency in language. Computer Speech & Language, 58, 19–50. https://doi.org/10.1016/j.csl.2019.03.007

Chantraine, P. (1968). *Dictionnaire étymologique de la langue grecque* (Tome 1). Klincksieck.

Dehouck, M., François, A., Kalyan, S., Pastor, M., & Kletz, D. (2023). EvoSem: A database of polysemous cognate sets. In *Proceedings of the 4th Workshop on Computational Approaches to Historical Language Change* (pp. 66–75). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.lchange-1.7

Dellert, J. (2016). Using causal inference to detect directional tendencies in semantic evolution. In The Evolution of Language: *Proceedings of the 11th International Conference (EVOLANGX11)*. Bristol: Evolang Scientific Committee.

Dellert, J. (2019). *Information-Theoretic Causal Inference of Lexical Flow*. Language Science Press.

Dellert, J. (2024). Causal inference of diachronic semantic maps from cross-linguistic synchronic polysemy data. *Frontiers in Communication, 8*, 1288196. https://doi.org/10.3389/fcomm.2023.1288196

# References


Dunn, J. 2015). *Modeling abstractness and metaphoricity*. *Metaphor and Symbol*, *30*(4), 259–289. https://doi.org/10.1080/10926488.2015.1074801

Faculty of Classics, University of Cambridge & James Diggle. 2021. *The Cambridge Greek lexicon*. Cambridge University Press.

Georgakopoulos, T., Grossman, E., Nikolaev, D. & Polis, S. (2022). Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain. *Linguistic Typology, 26*(2), 439-487. https://doi.org/10.1515/lingty-2021-2088

Haspelmath, M. (2023). Coexpression and synexpression patterns across languages: Comparative concepts and possible explanations. *Frontiers in Psychology, 14*, 1236853. https://doi.org/10.3389/fpsyg.2023.1236853

Hock, H. H. (1991). *Principles of historical linguistics* (second). Mouton de Gruyter.

Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. University of Chicago Press.

Hollmann, B. W. (2009). Semantic change. In Jonathan Culpeper, F. Katamba, P. Kerswill, & T. McEnery (Eds.), *English language: Description, variation and context* (pp. 301–313). Palgrave Macmillan.

Jäger, G. & Rosenbach, A. (2008). Priming and unidirectional language change. *Theoretical Linguistics, 34*(2), 85-113. https://doi.org/10.1515/THLI.2008.008

Imbir, K. K. (2021). Corrigendum: Affective norms for 4,900 Polish words reload (ANPW_R): Assessments for valence, arousal, dominance, origin, significance, concreteness, imageability, and age of acquisition. *Frontiers in Psychology, 12*.

Kriaras, E. (ed.). 1968- . *Λεξικό της Μεσαιωνικής Ελληνικής Δημώδους Γραμματείας, 1100- 1669* [Dictionary of Medieval Greek Vernacular Literature (1100–1669)]. Kentro Ellinikis Glossas.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

# References


Li, Y., Siew, C.S.Q. (2022). Diachronic semantic change in language is constrained by how people use and learn language. *Mem Cogn 50*, 1284–1298. https://doi.org/10.3758/s13421-022-01331-0

Liddell, H., & Scott, R. (1996). *A Greek-English lexicon* [Revised and complemented throughout by Henry Stuart Jones]. Clarendon Press.

List, J.-M. (2023). Inference of partial colexifications from multilingual wordlists. *Frontiers in Psychology* 14, 1156540. https://doi.org/10.3389/fpsyg.2023.1156540

List, J.-M., Tjuka, A., Blum, F., Kučerová, A., Ugarte, C. B., Rzymski, C., Greenhill, S., & Forkel, R. (Eds.). (2025). *CLLD Concepticon 3.4.0* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14923561

Luraghi, S. (2020). *Experiential verbs in Homeric Greek: A constructional approach* (Vol. 27). Brill.

Luraghi, S., & Sausa, E. (2019). Aspects of aural perception in Homeric Greek. In E. Mocciaro & W. Short (Eds.), *The embodied basis of constructions in Greek and Latin* (pp. 149–175). De Gruyter Mouton.

Montanari, F. (2015). *The Brill dictionary of ancient Greek*. Brill.

Mooers, A. O., & Pappas, P. A. (2011). Phylogenetic methods in historical linguistics: Greek as a case study. *Journal of Greek Linguistics, 11*(2), 198–220.

Pagel, M., Atkinson, Q. D., & Meade, A. (2007). Frequency of word use predicts rates of lexical evolution throughout Indo-European history. *Nature, 449*(7163), 717–720. https://doi.org/10.1038/nature06176

Rzymski, C., Tresoldi, T., Greenhill, S. J., Wu, M.-S., Schweikhard, N. E., Koptjevskaja-Tamm, M., Gast, V., , … List, J.-M. (2020). The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. *Scientific Data*, 7(*1*), Article 1.

Sweetser, E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. CUP.

# References


Tjuka, A., Forkel, R., & List, J.-M. (2022). Linking norms, ratings, and relations of words and concepts across multiple language varieties. *Behavior Research Methods, 54*(2), 864–884.

Traugott, E. C., & Dasher, R. (2002). *Regularity in semantic change*. CUP.

Urban, M. (2011). Asymmetries in overt marking and directionality in semantic change. *Journal of Historical Linguistics, 1*(1), 3–47. https://doi.org/10.1075/jhl.1.1.02urb.

Winter, B., & Srinivasan, M. (2022). Why is Semantic Change Asymmetric? The Role of Concreteness and Word Frequency and Metaphor and Metonymy. *Metaphor and Symbol, 37*(1), 39–54.

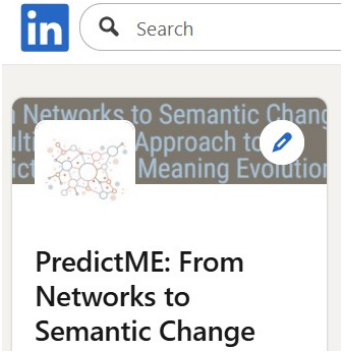
Xu, Y., Malt, B. C., & Srinivasan, M. (2017). Evolution of word meanings through metaphorical mapping: Systematicity over the past millennium. *Cognitive Psychology, 96*, 41–53.

Zalizniak, A. A., Bulakh, M., Ganenkov, D., Gruntov, I., Maisak, T., & Russo, M. (2012). The Catalogue of Semantic Shifts as a database for lexical semantic typology. *Linguistics, 50*(3), 633–670. https://doi.org/10.1515/ling-2012-0020

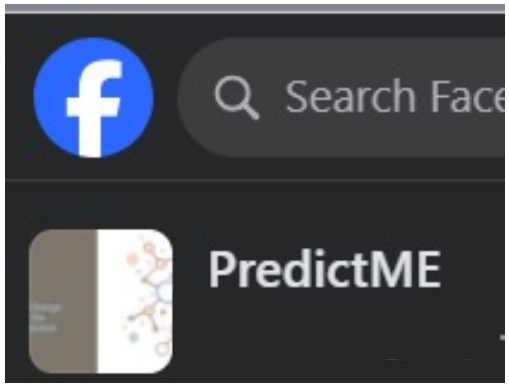
Zalizniak, A. (2018). The Catalogue of Semantic Shifts: 20 Years Later. Russian *Journal of Linguistics, 22*(4), 770-787. https://doi.org/10.22363/2312-9182-2018-22-4-770-787

Zipf, G. K. (1945). The meaning-frequency relationship of words. *Journal of General Psychology, 33*(2), 251–256. https://doi.org/10.1080/00221309.1945.10544509.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley Press.




https://www.linkedin.com/company/110316536/admin/dashboard/




https://www.facebook.com/groups/1506748170428080

# Thank you!

ΕΛ.ΙΔ.Ε.Κ.
Ελληνικό Ίδρυμα Έρευνας & Καινοτομίας
H.F.R.I.
Hellenic Foundation for Research & Innovation