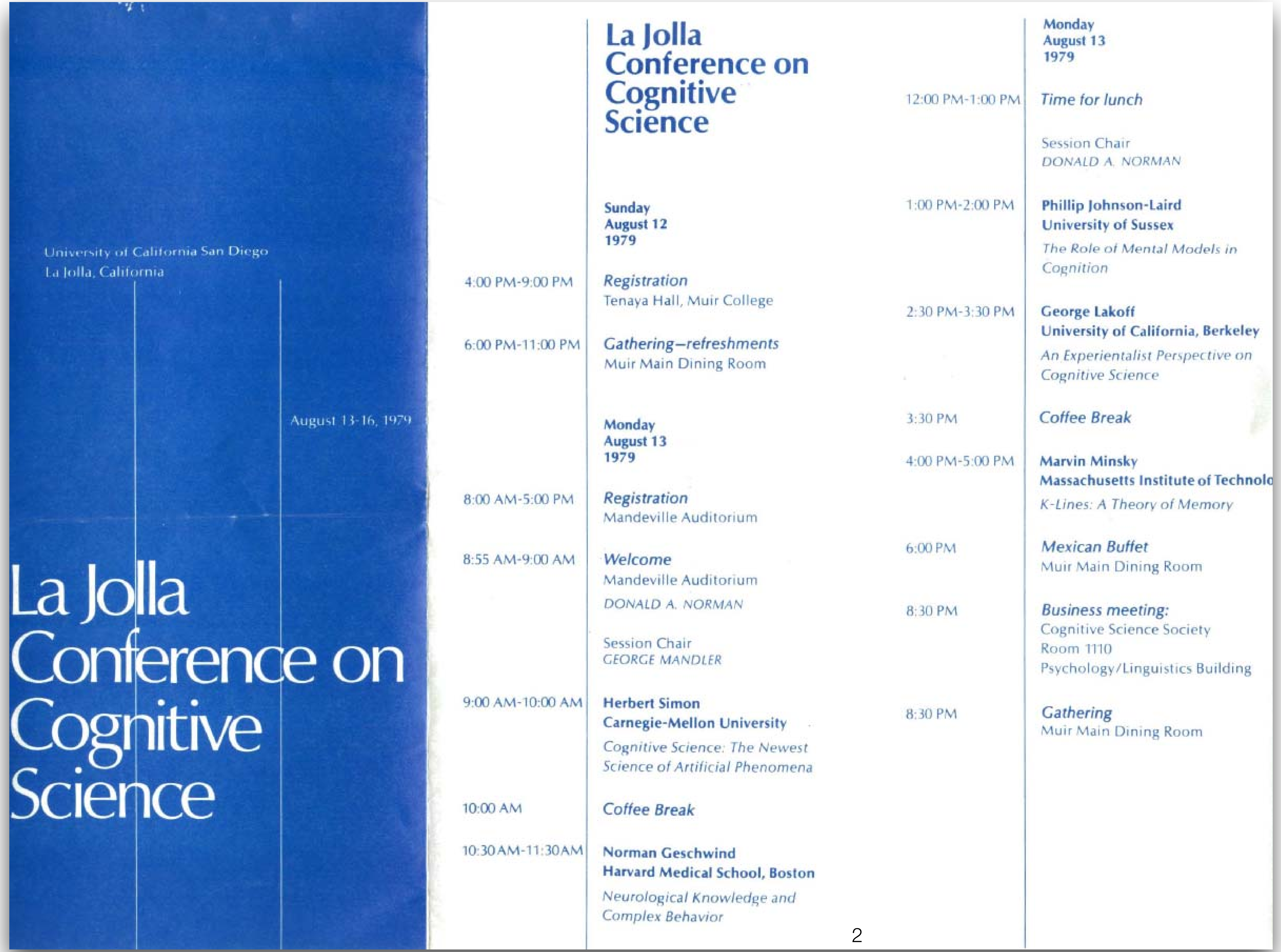
the title of this talk
(yes, it sounds like a journal name)

TOPICS AND TRENDS IN COGNITIVE SCIENCE

equal contribution by all authors

Anselm Rothe *Alexander Rich* *Zhi-Wei Li*

New York University


University of California San Diego
La Jolla, California

August 13-16, 1979

La Jolla Conference on Cognitive Science

La Jolla Conference on Cognitive Science

Sunday August 12 1979

4:00 PM-9:00 PM	*Registration* Tenaya Hall, Muir College
6:00 PM-11:00 PM	*Gathering—refreshments* Muir Main Dining Room

Monday August 13 1979

8:00 AM-5:00 PM	*Registration* Mandeville Auditorium
8:55 AM-9:00 AM	*Welcome* Mandeville Auditorium *DONALD A. NORMAN*
	Session Chair *GEORGE MANDLER*
9:00 AM-10:00 AM	**Herbert Simon** **Carnegie-Mellon University** *Cognitive Science: The Newest Science of Artificial Phenomena*
10:00 AM	*Coffee Break*
10:30 AM-11:30 AM	**Norman Geschwind** **Harvard Medical School, Boston** *Neurological Knowledge and Complex Behavior*

Monday August 13 1979

12:00 PM-1:00 PM	*Time for lunch*
	Session Chair *DONALD A. NORMAN*
1:00 PM-2:00 PM	**Phillip Johnson-Laird** **University of Sussex** *The Role of Mental Models in Cognition*
2:30 PM-3:30 PM	**George Lakoff** **University of California, Berkeley** *An Experientalist Perspective on Cognitive Science*
3:30 PM	*Coffee Break*
4:00 PM-5:00 PM	**Marvin Minsky** **Massachusetts Institute of Technolo** *K-Lines: A Theory of Memory*
6:00 PM	*Mexican Buffet* Muir Main Dining Room
8:30 PM	*Business meeting:* Cognitive Science Society Room 1110 Psychology/Linguistics Building
8:30 PM	*Gathering* Muir Main Dining Room

← Program of the first conference

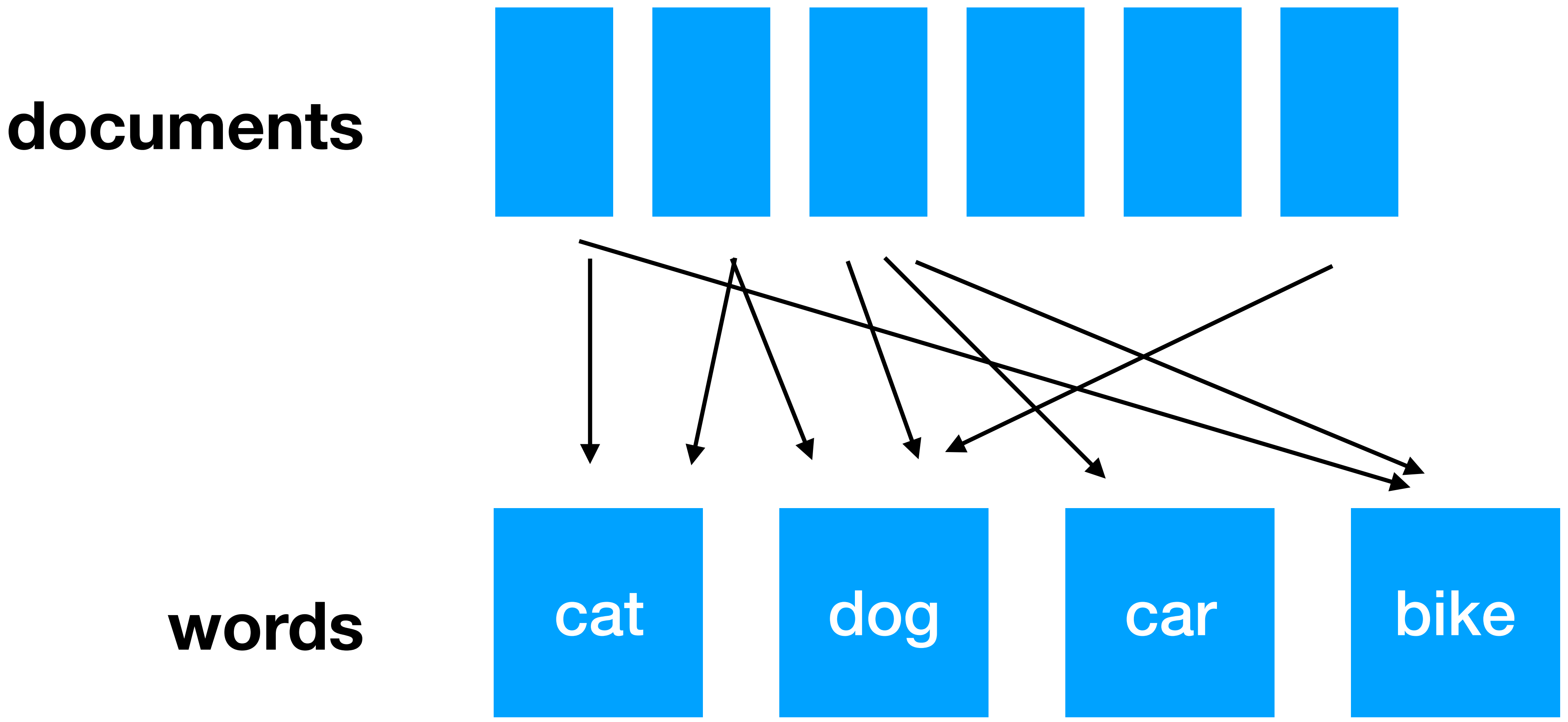
- **CogSci 1979**
  - 1 track
  - 5 categories
  - 42 talks
- **CogSci 2018**
  - 10 parallel tracks
  - 57 categories
  - 213 talks

Overview

- Problem: Deal with complexity of the field, discover topics and trends
- Tool: Dynamic Topic Modeling
- Results: CogSci topics, topic space, global trends, comparison of labs, paper recommendation system

Topic Modeling

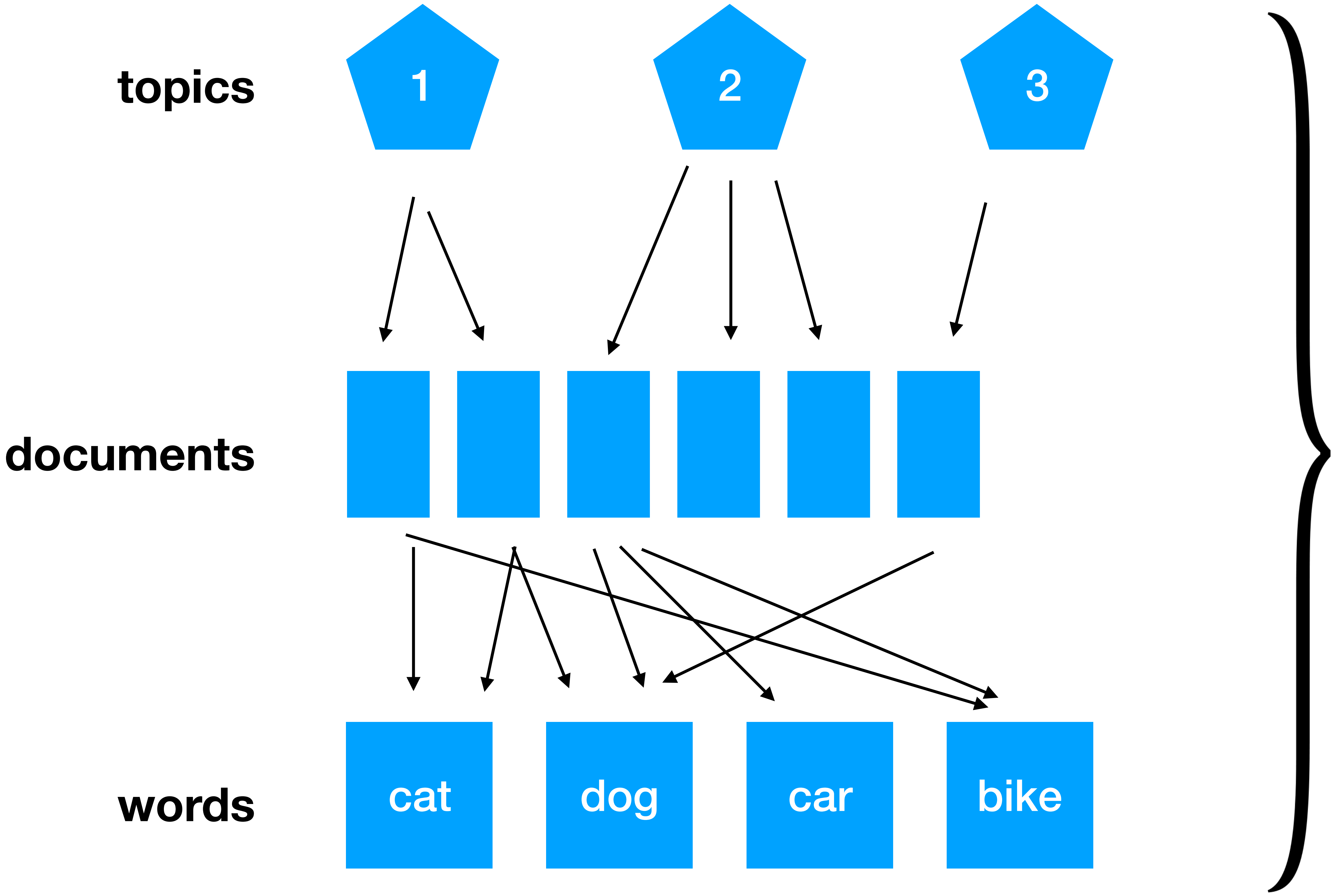
- Capture structure in large text corpora and make them more human-understandable
  (Blei, Ng, & Jordan 2003, Cohen Priva & Austerweil 2015)



(papers, news articles, customer reviews, etc.)

Topic Modeling

- Capture structure in large text corpora and make them more human-understandable
  (Blei, Ng, & Jordan 2003, Cohen Priva & Austerweil 2015)

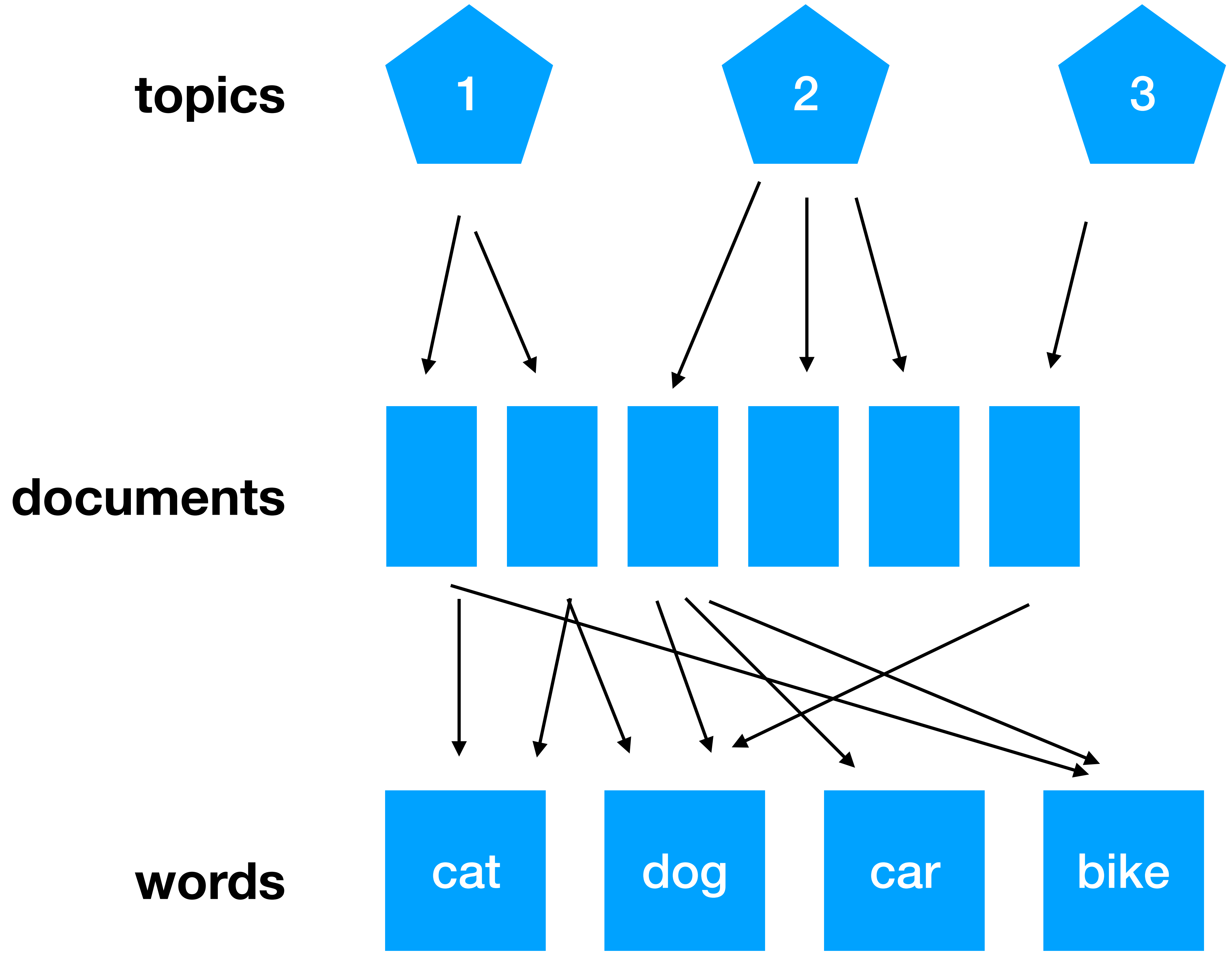


Assumed **generative model**

- a document is sampled from a topic
- a word in a document is generated given this topic
- based on sample probabilities (arrow weights)

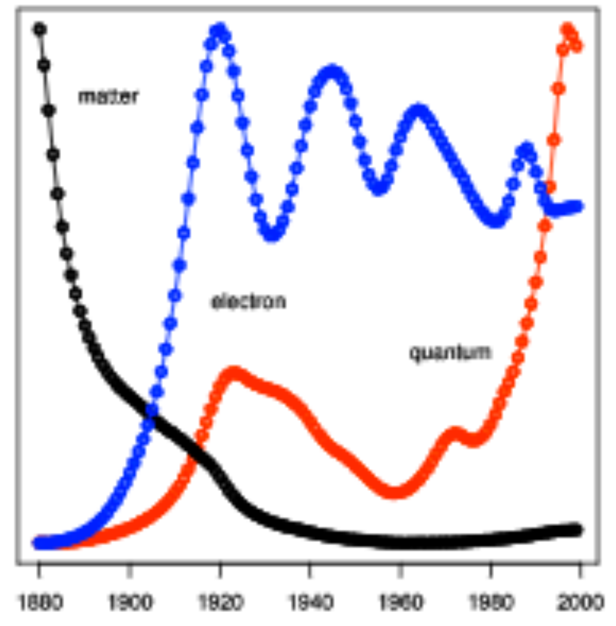
Topic Modeling

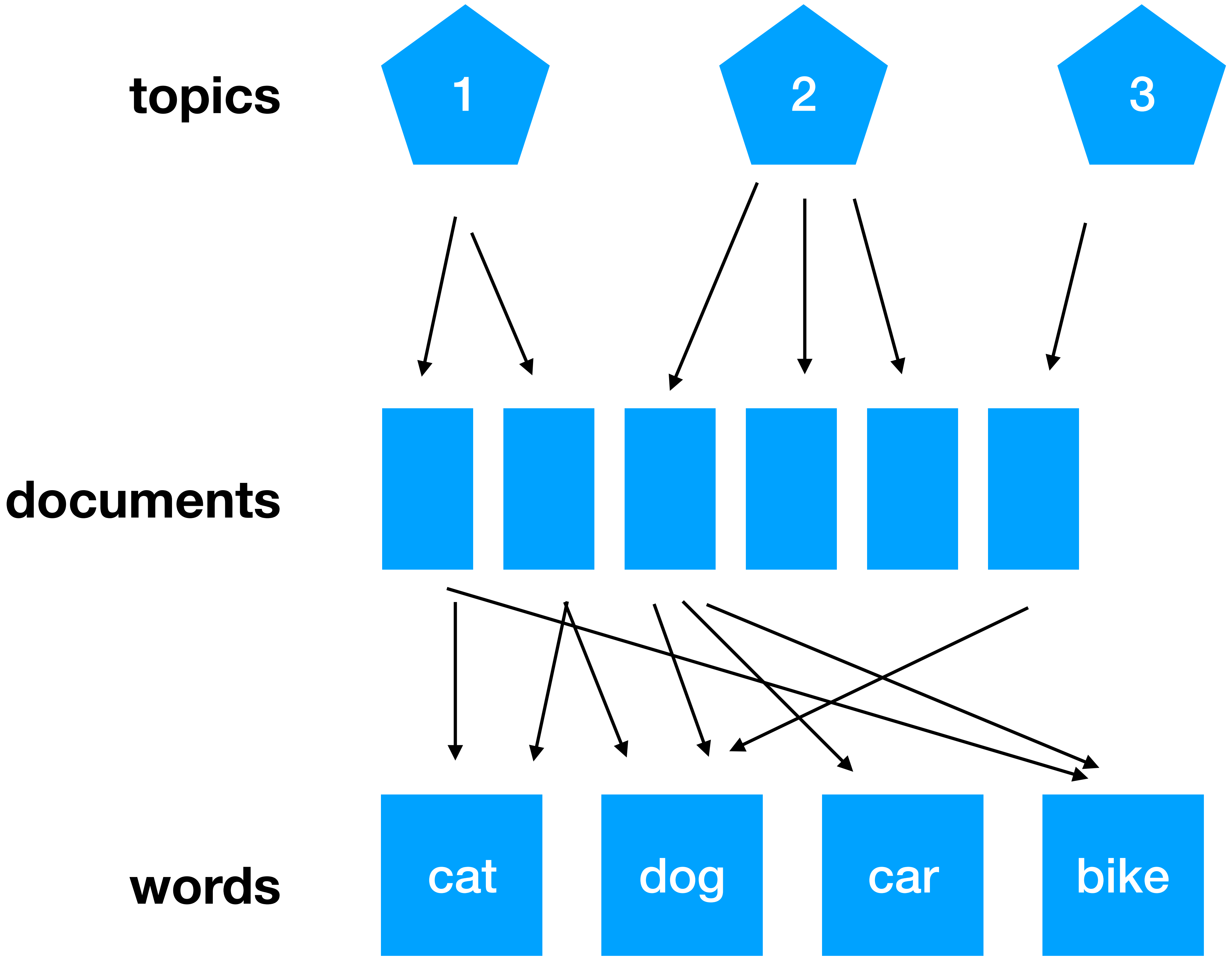
- Capture structure in large text corpora and make them more human-understandable
(Blei, Ng, & Jordan 2003, Cohen Priva & Austerweil 2015)



- **Topic Model** makes reverse inference over the generative model
  - given document x (with the words $\mathbf{w}$), how likely was topic y_1 its generating topic
  - simultaneously: which document-topic assignment works best across all documents
  - indirectly: which words go with which topic

Dynamic Topic Modeling





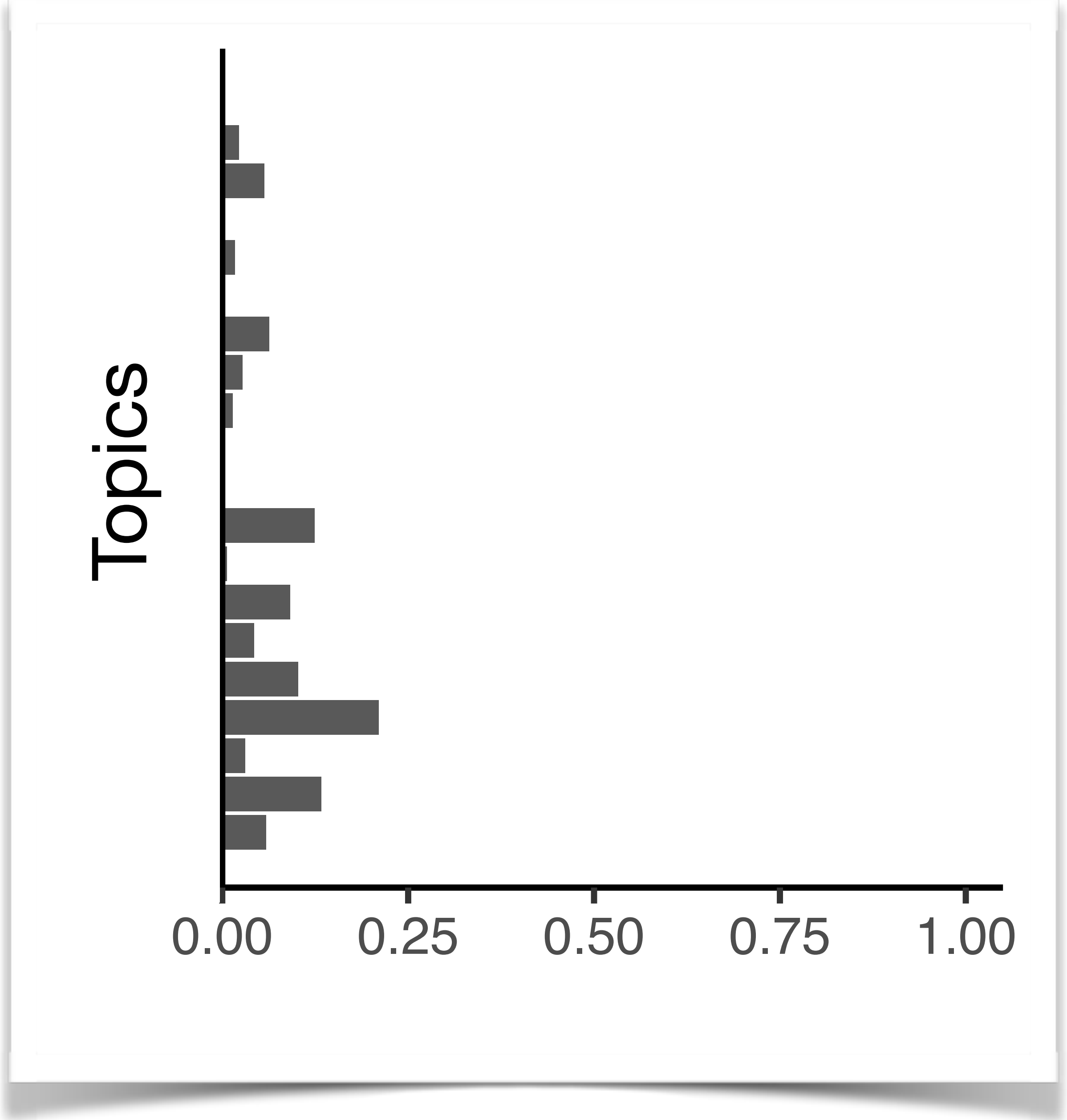
- ***Dynamic* Topic Models** are capturing change over time (Blei & Lafferty 2006)
- **Example**: The topic "Communication" might have a high likelihood for "fax machine" in news articles from 1990 but for today instead a high likelihood for "internet"
- => Which words go with which topic is allowed to change over time

Data set: 18 years of CogSci

- Conference proceedings are particularly useful because there are many fresh ideas rapidly published

- 6920 PDFs from the CogSci archives
- Years 2000-2017 (usable PDFs)
- Preprocessing
  - PDFtoText
  - tokenizing (splitting text into words)
  - lemmatizing (e.g., walked → walk)
  - removing very rare and very frequent words
  - etc.
- Final vocabulary: 9710 words

Output



- For each document (CogSci paper) we have
  - a score of each topic for this document (all the score make a “topic vector”)
  - the words of the document
  - the year of the document

  - from that we can extract a bunch of visualizations and insights

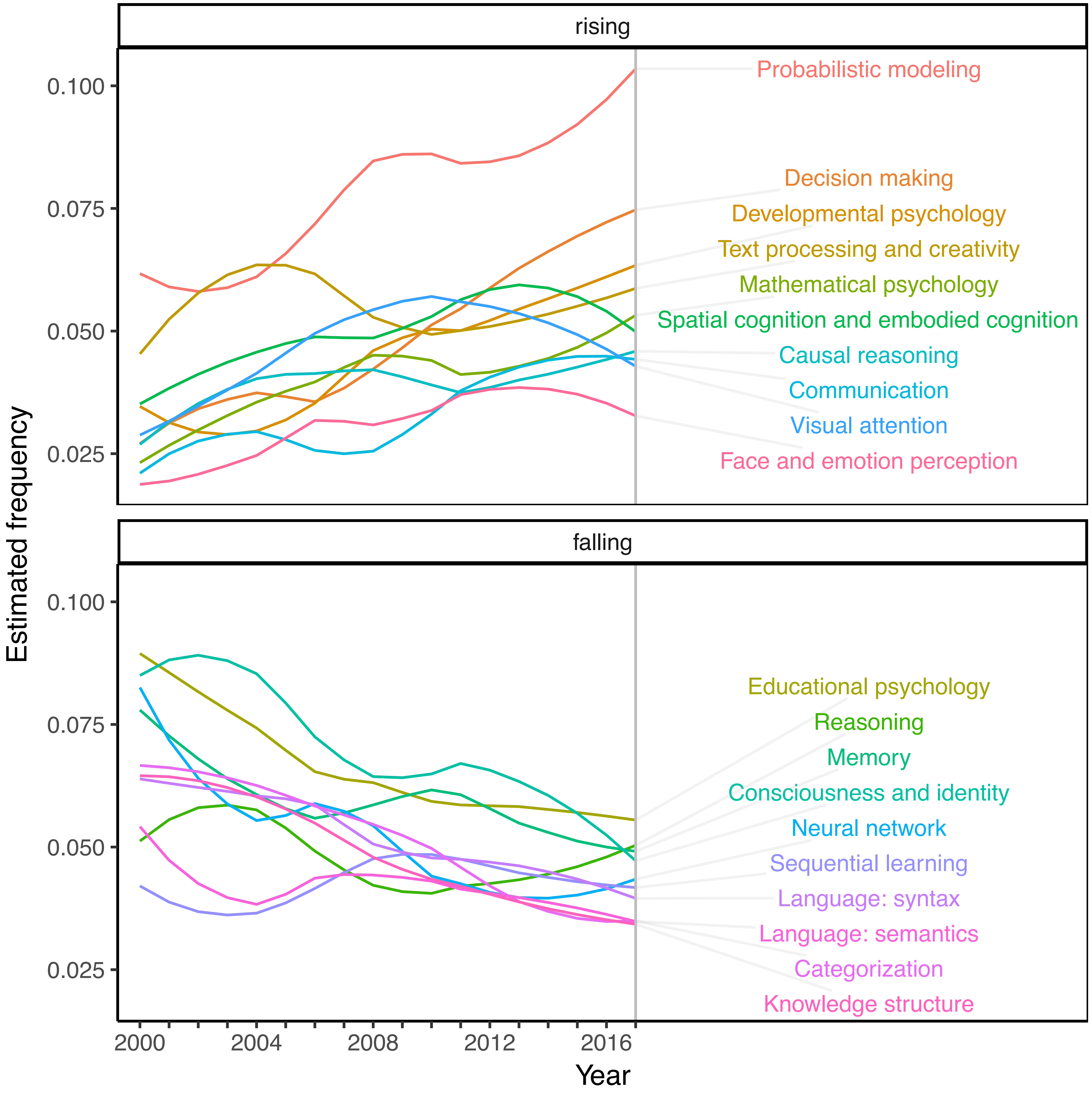
Estimated frequency
0.100
0.075
0.050
0.025
2000
2004
2008
2012
2016
Year
Probabilistic modeling
Decision making
Developmental psychology
Text processing and creativity
Educational psychology
Mathematical psychology
Reasoning
Spatial cognition and embodied cognition
Memory
Consciousness and identity
Causal reasoning
Communication
Neural network
Visual attention
Sequential learning
Language: syntax
Categorization
Language: semantics
Knowledge structure
Face and emotion perception

- Topic labels were manually chosen based on
  - most frequent words
  - most typical papers titles


Estimated frequency
0.100
0.075
0.050
0.025
2000
2004
2008
2012
2016
Year
Probabilistic modeling
Decision making
Developmental psychology
Text processing and creativity
Educational psychology
Mathematical psychology
Reasoning
Spatial cognition and embodied cognition
Memory
Consciousness and identity
decision
choice
strategy
game
making
player
option
reward
behavior
probability

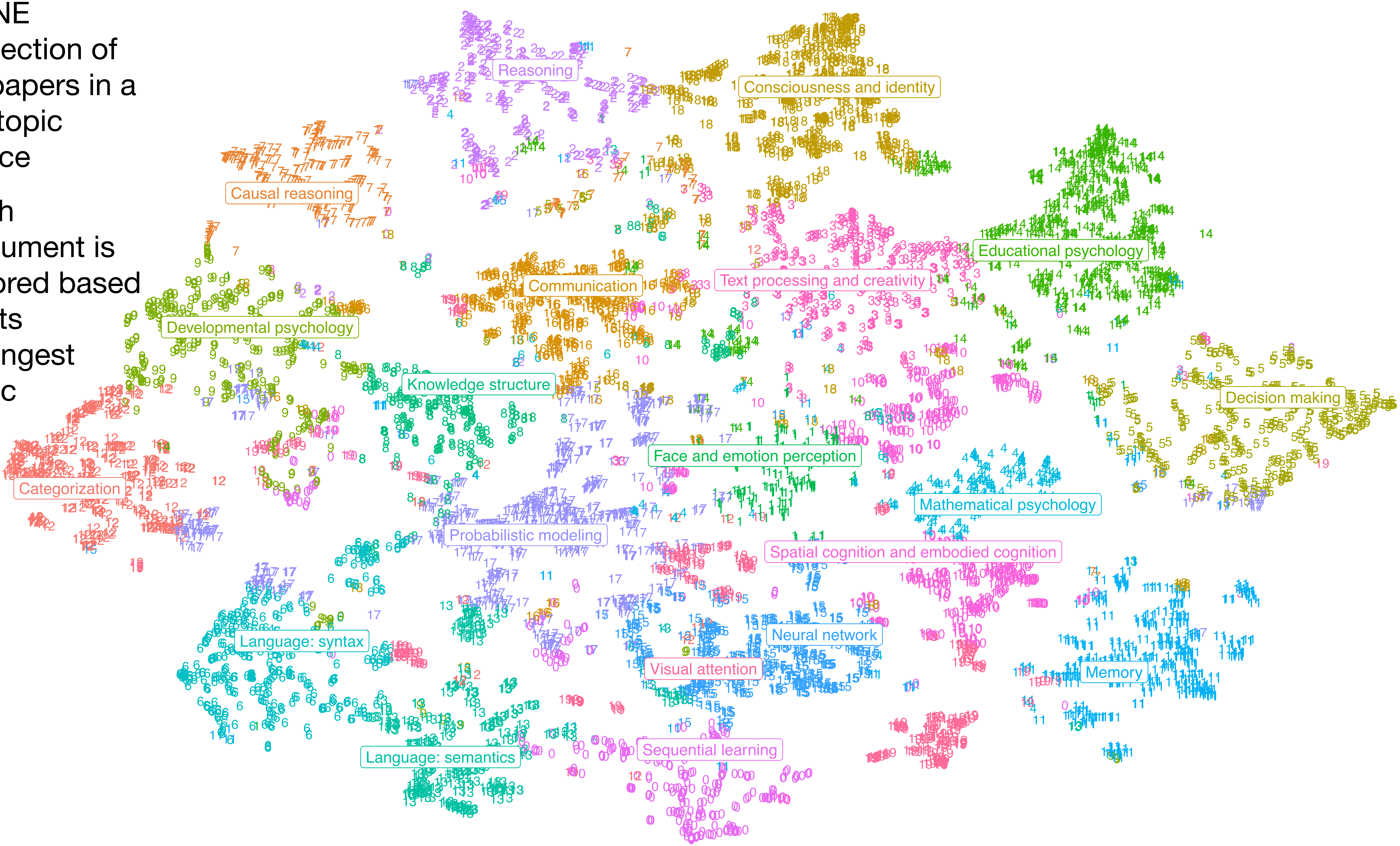
Estimated frequency
0.100
0.075
0.050
0.025
2000
2004
2008
2012
2016
Year
Probabilistic modeling
Decision making
Developmental psychology
Text processing and creativity
Educational psychology
Mathematical psychology
Reasoning
Spatial cognition and embodied cognition
Memory
Consciousness and identity
student
strategy
solving
solution
instruction
transfer
learner
training
skill
goal

0.100
0.075
0.050
0.025
Estimated frequency
2000
2004
2008
2012
2016
Year
Probabilistic modeling
Decision making
Developmental psychology
Text processing and creativity
Educational psychology
Mathematical psychology
Reasoning
Spatial cognition and embodied cognition
Memory
Consciousness and identity
error
trial
recall
retrieval
control
working
list
cue
chunk
accuracy

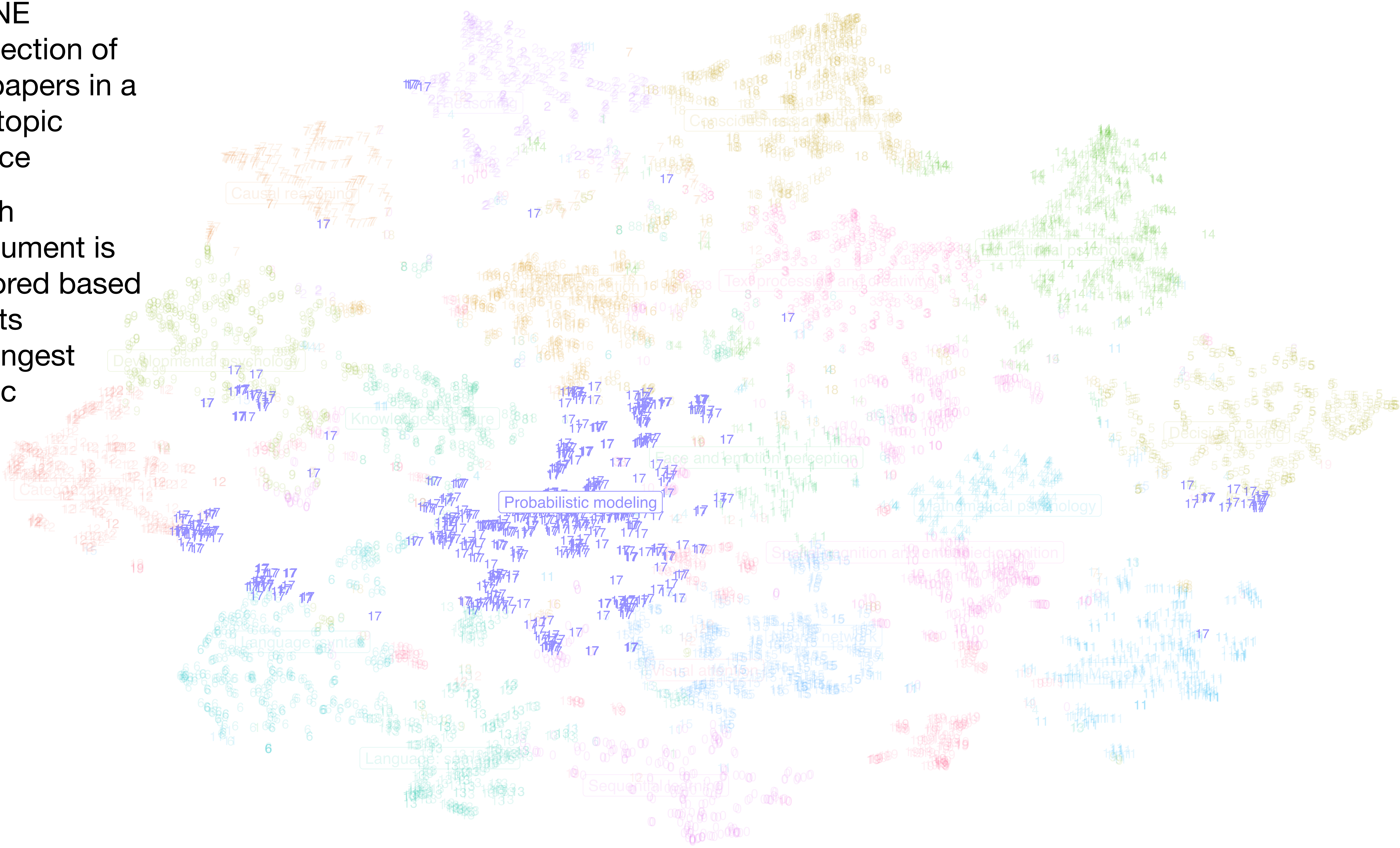


Anselm Rothe, Alexander Rich, Zh

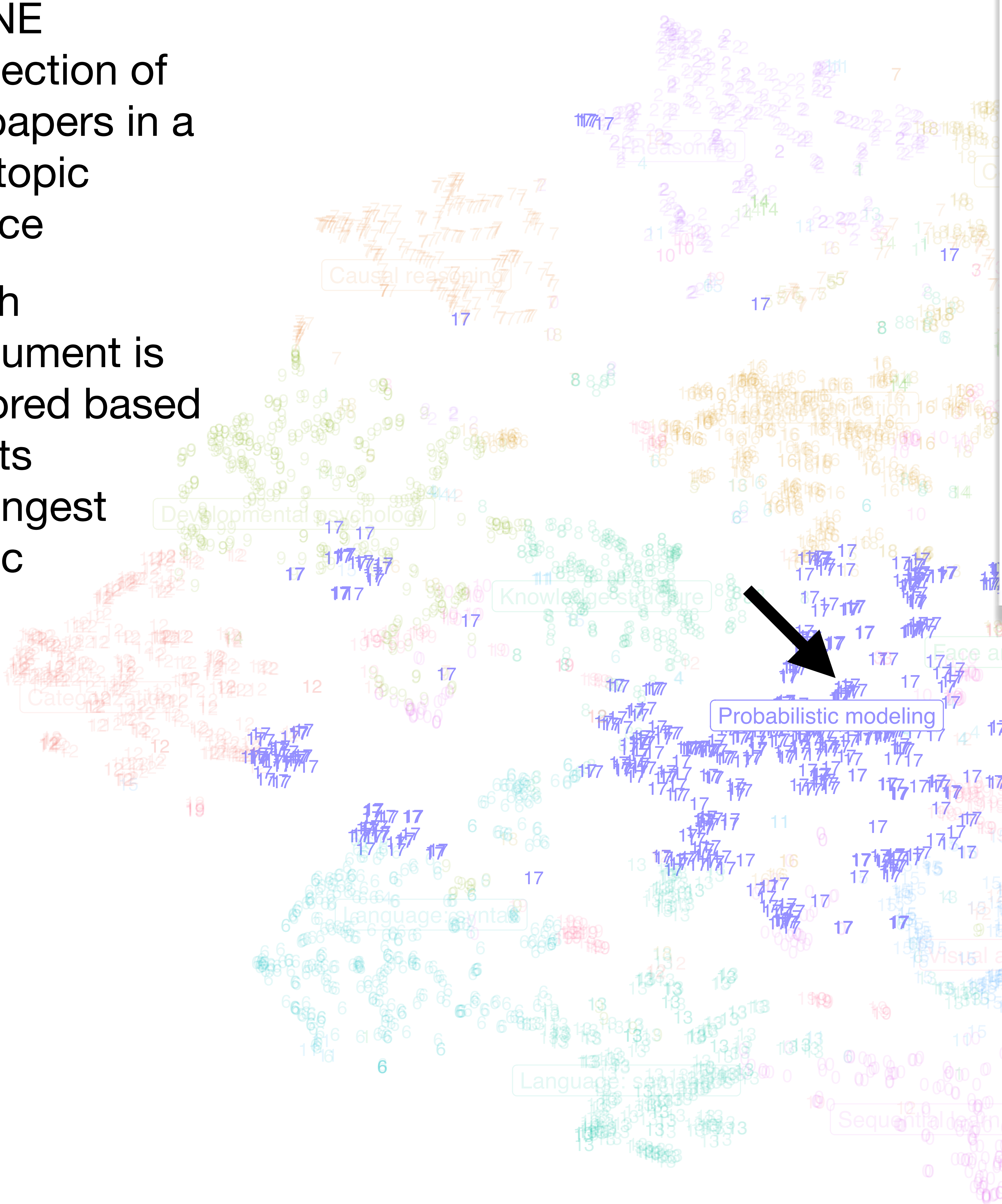
- t-SNE projection of all papers in a 2D topic space
- each document is colored based on its strongest topic

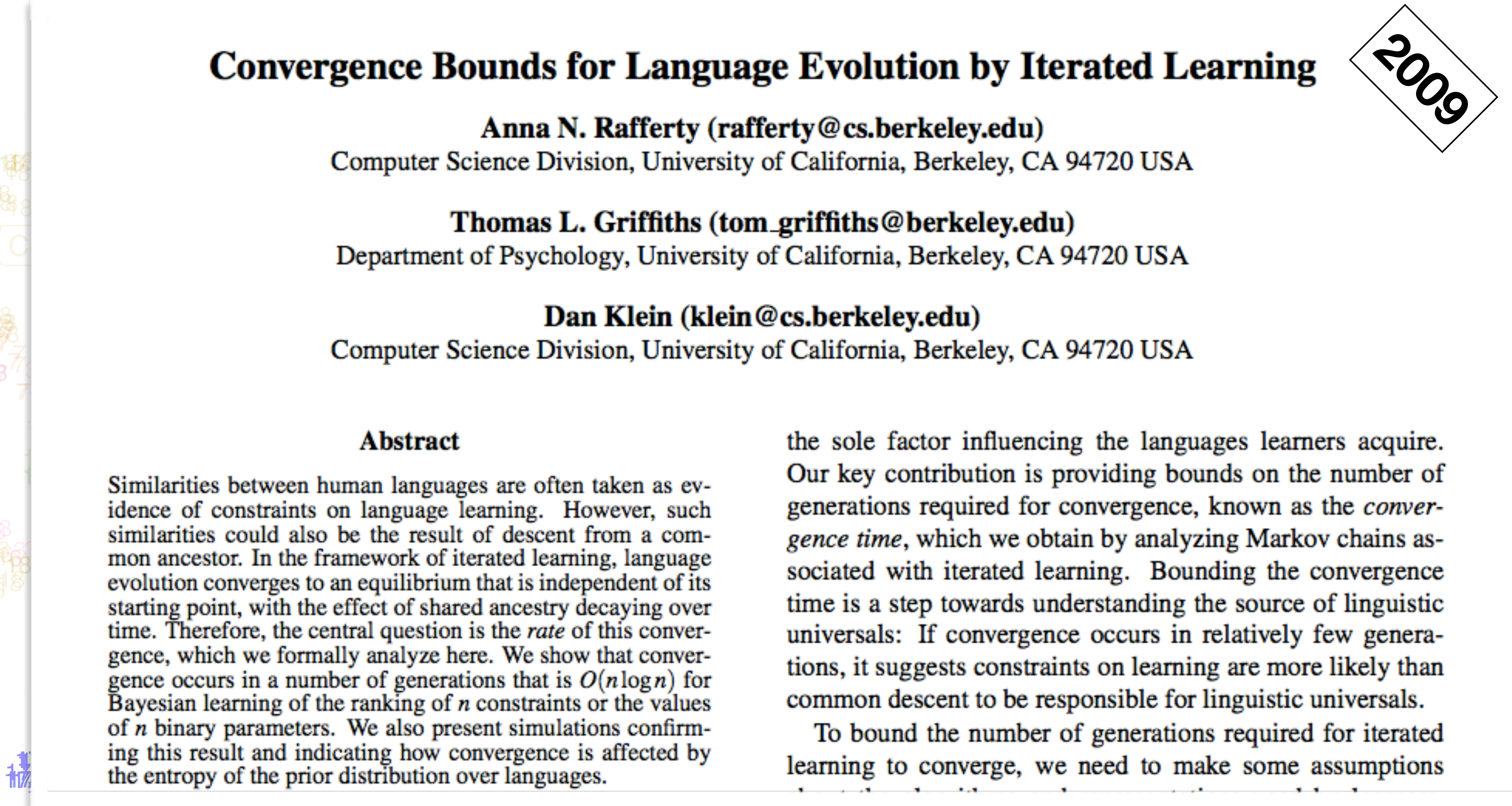


- t-SNE projection of all papers in a 2D topic space
- each document is colored based on its strongest topic



- t-SNE projection of all papers in a 2D topic space
- each document is colored based on its strongest topic





2009

Convergence Bounds for Language Evolution by Iterated Learning

Anna N. Rafferty (rafferty@cs.berkeley.edu)
Computer Science Division, University of California, Berkeley, CA 94720 USA

Thomas L. Griffiths (tom_griffiths@berkeley.edu)
Department of Psychology, University of California, Berkeley, CA 94720 USA

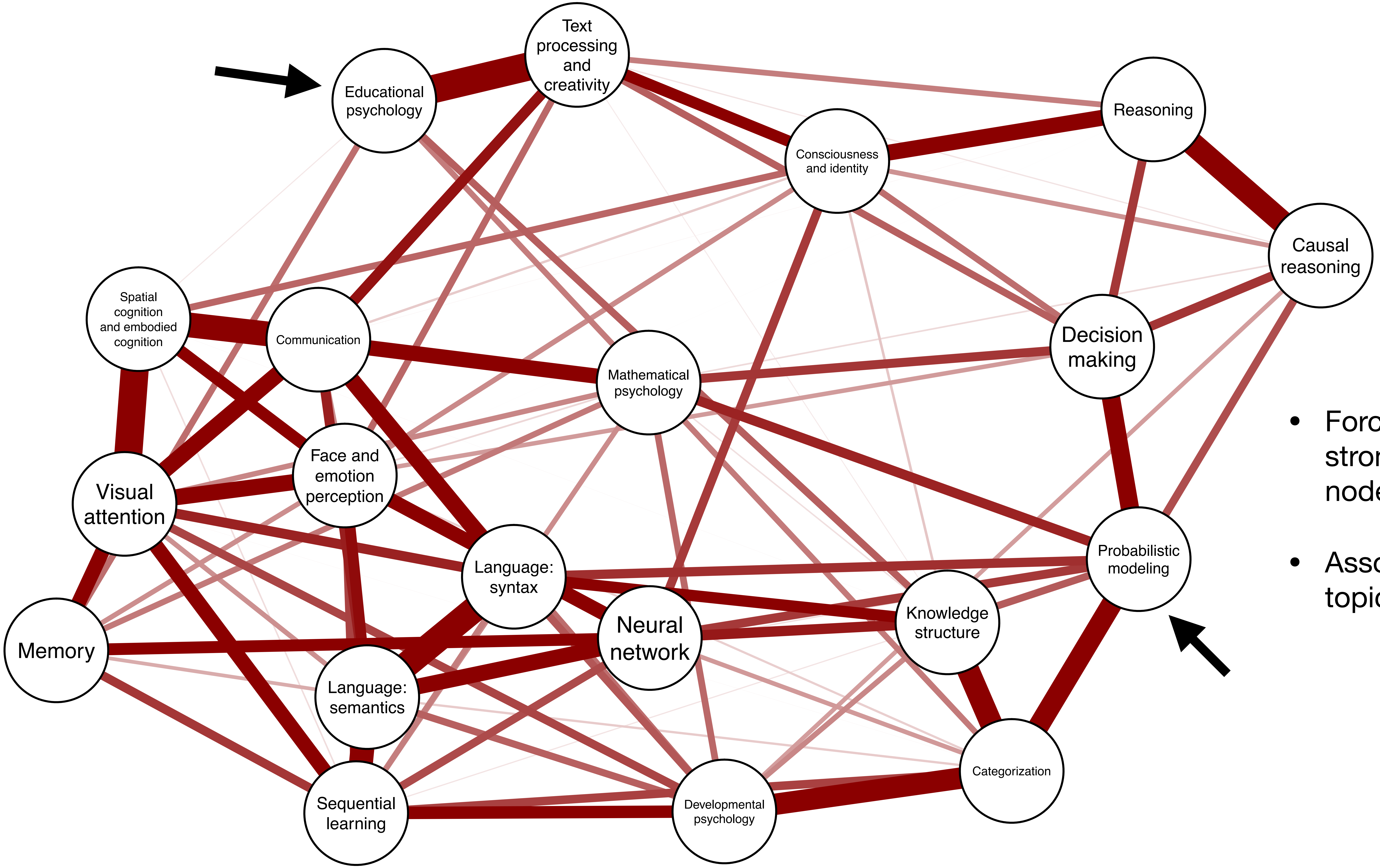
Dan Klein (klein@cs.berkeley.edu)
Computer Science Division, University of California, Berkeley, CA 94720 USA

Abstract

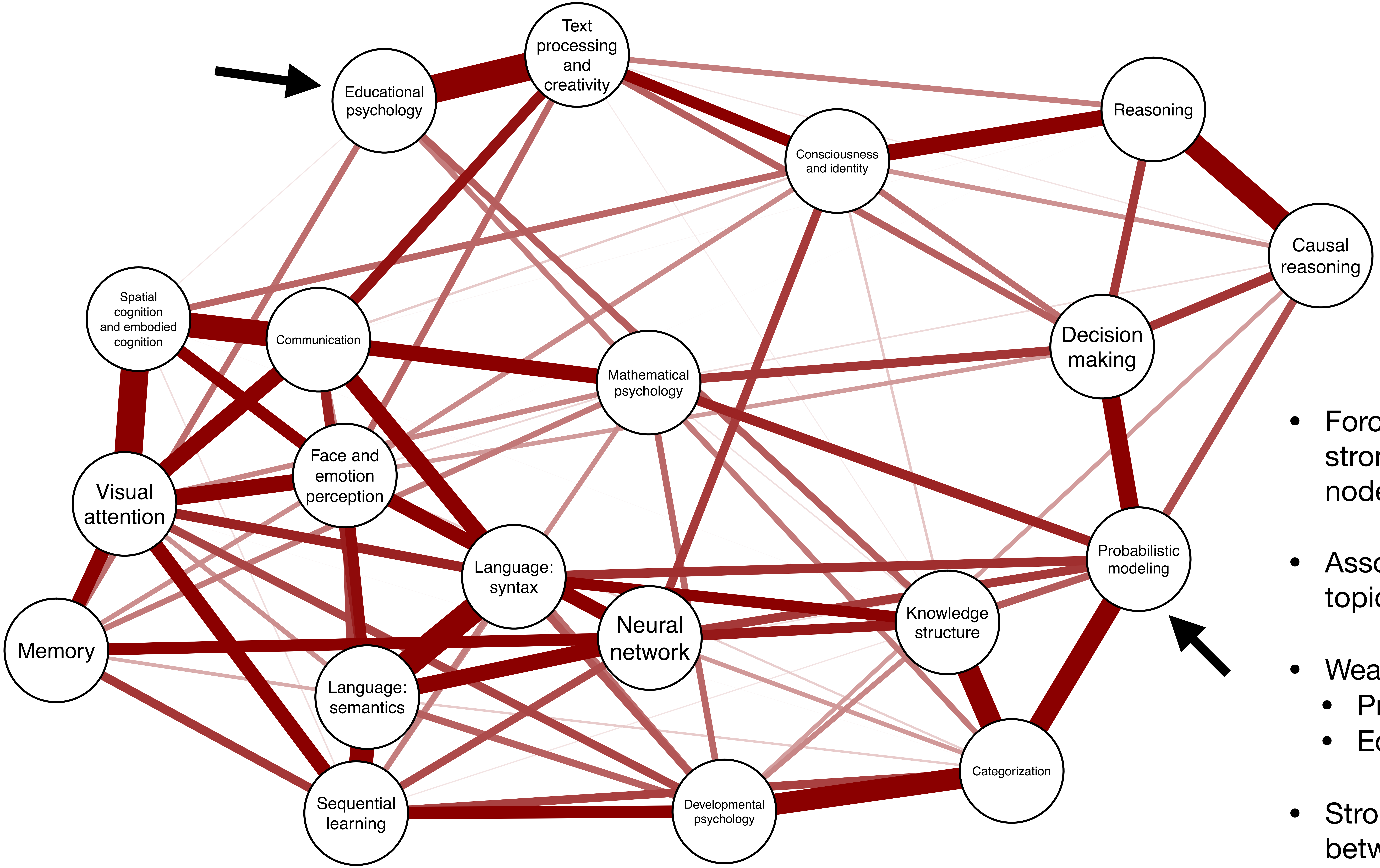
Similarities between human languages are often taken as evidence of constraints on language learning. However, such similarities could also be the result of descent from a common ancestor. In the framework of iterated learning, language evolution converges to an equilibrium that is independent of its starting point, with the effect of shared ancestry decaying over time. Therefore, the central question is the *rate* of this convergence, which we formally analyze here. We show that convergence occurs in a number of generations that is $O(n \log n)$ for Bayesian learning of the ranking of n constraints or the values of n binary parameters. We also present simulations confirming this result and indicating how convergence is affected by the entropy of the prior distribution over languages.

the sole factor influencing the languages learners acquire. Our key contribution is providing bounds on the number of generations required for convergence, known as the *convergence time*, which we obtain by analyzing Markov chains associated with iterated learning. Bounding the convergence time is a step towards understanding the source of linguistic universals: If convergence occurs in relatively few generations, it suggests constraints on learning are more likely than common descent to be responsible for linguistic universals.

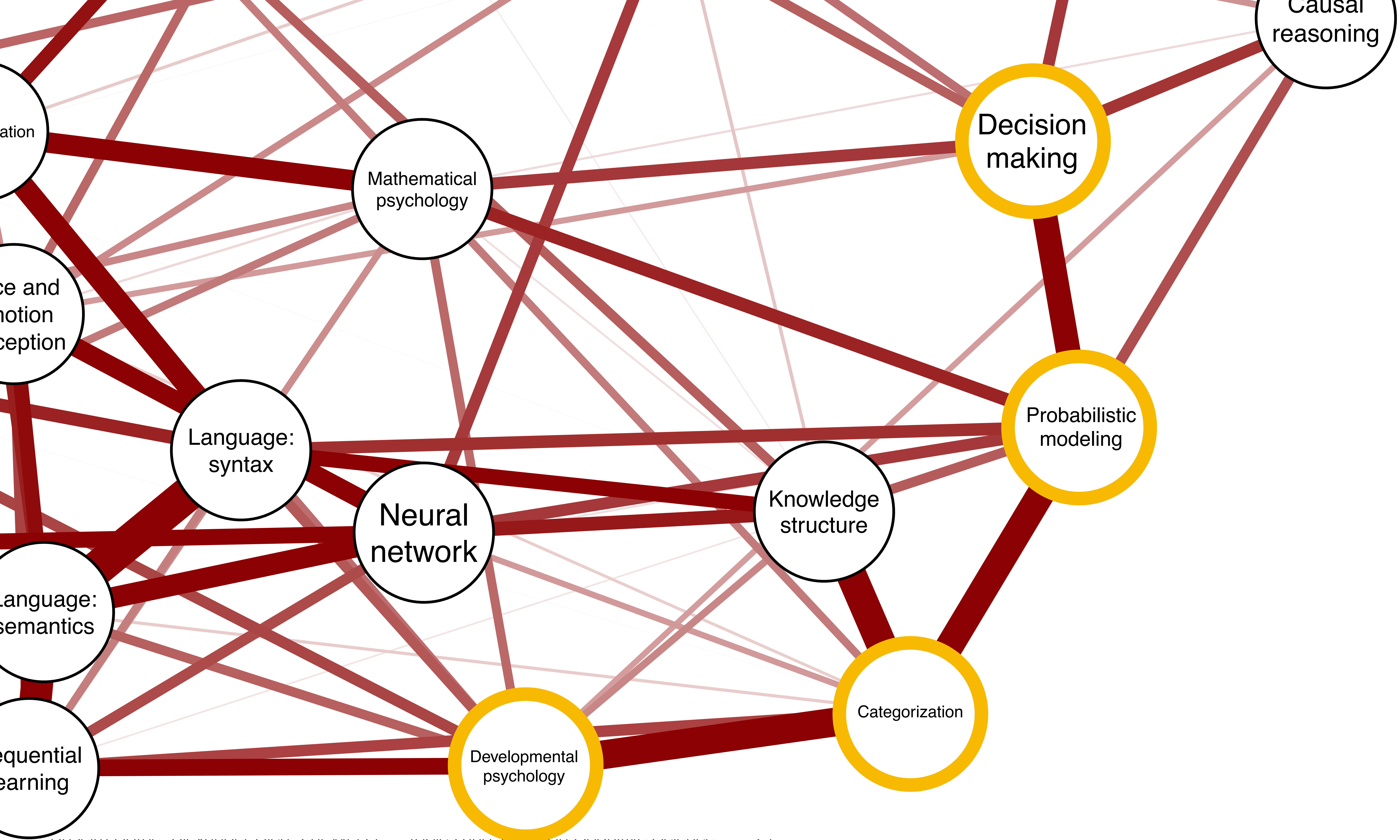
To bound the number of generations required for iterated learning to converge, we need to make some assumptions



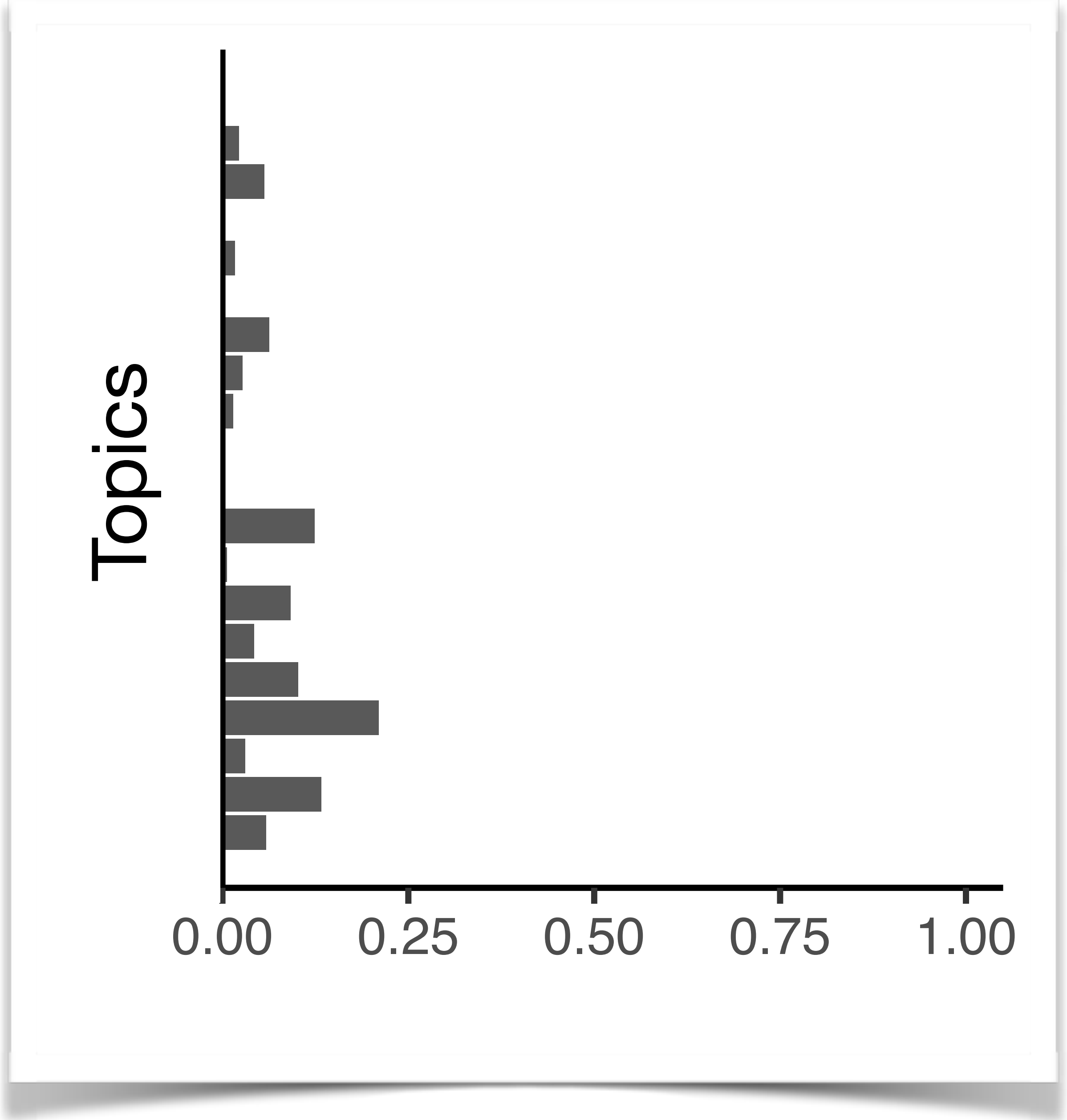
- Force network, where stronger associations pull the nodes towards each other
- Association = cosine between topic vectors

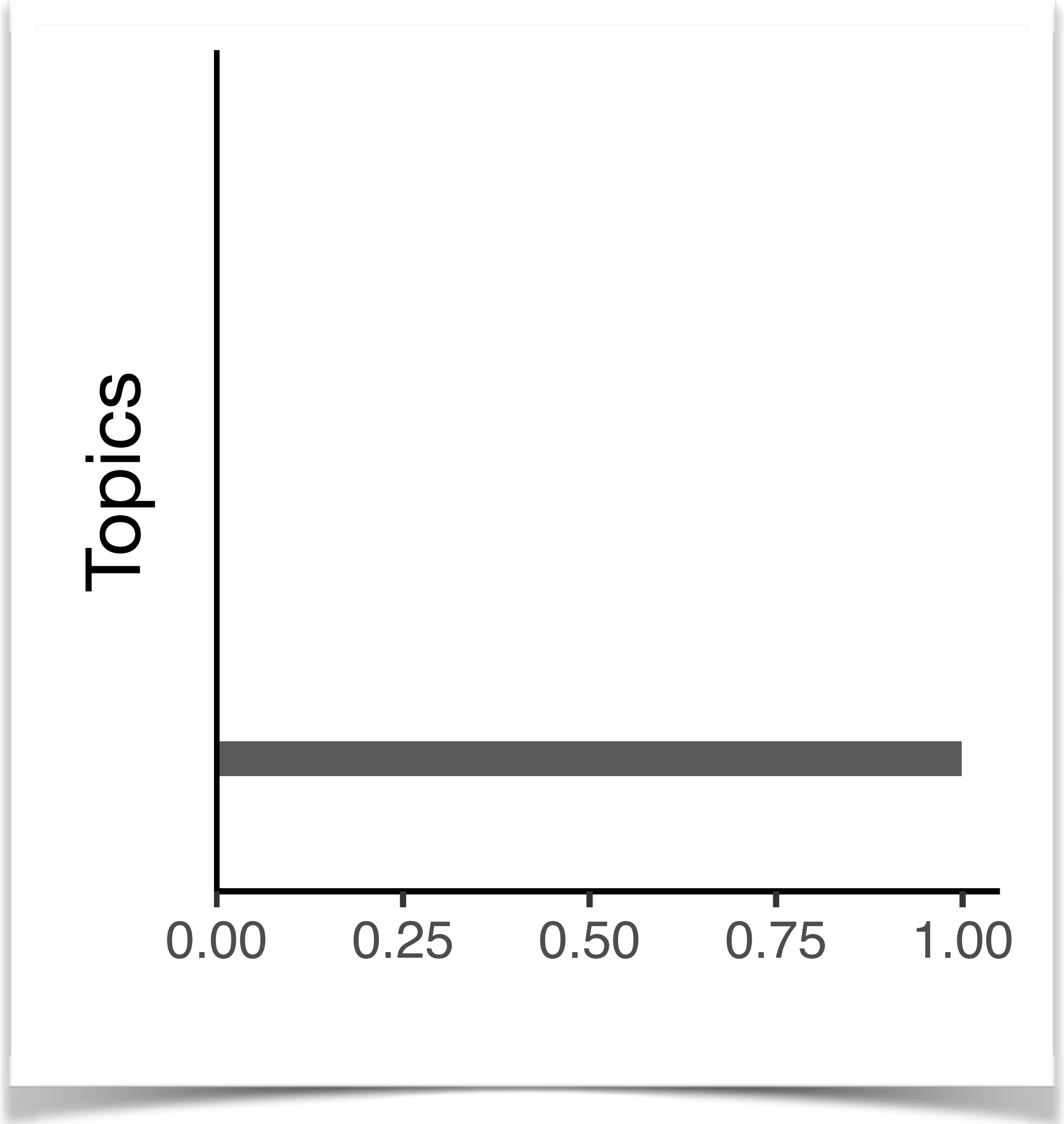


- Force network, where stronger associations pull the nodes towards each other
- Association = cosine between topic vectors
- Weakest association between
  - Probabilistic modeling
  - Educational psychology
- Strongest association between
  - Educational psychology
  - Text processing and creativity

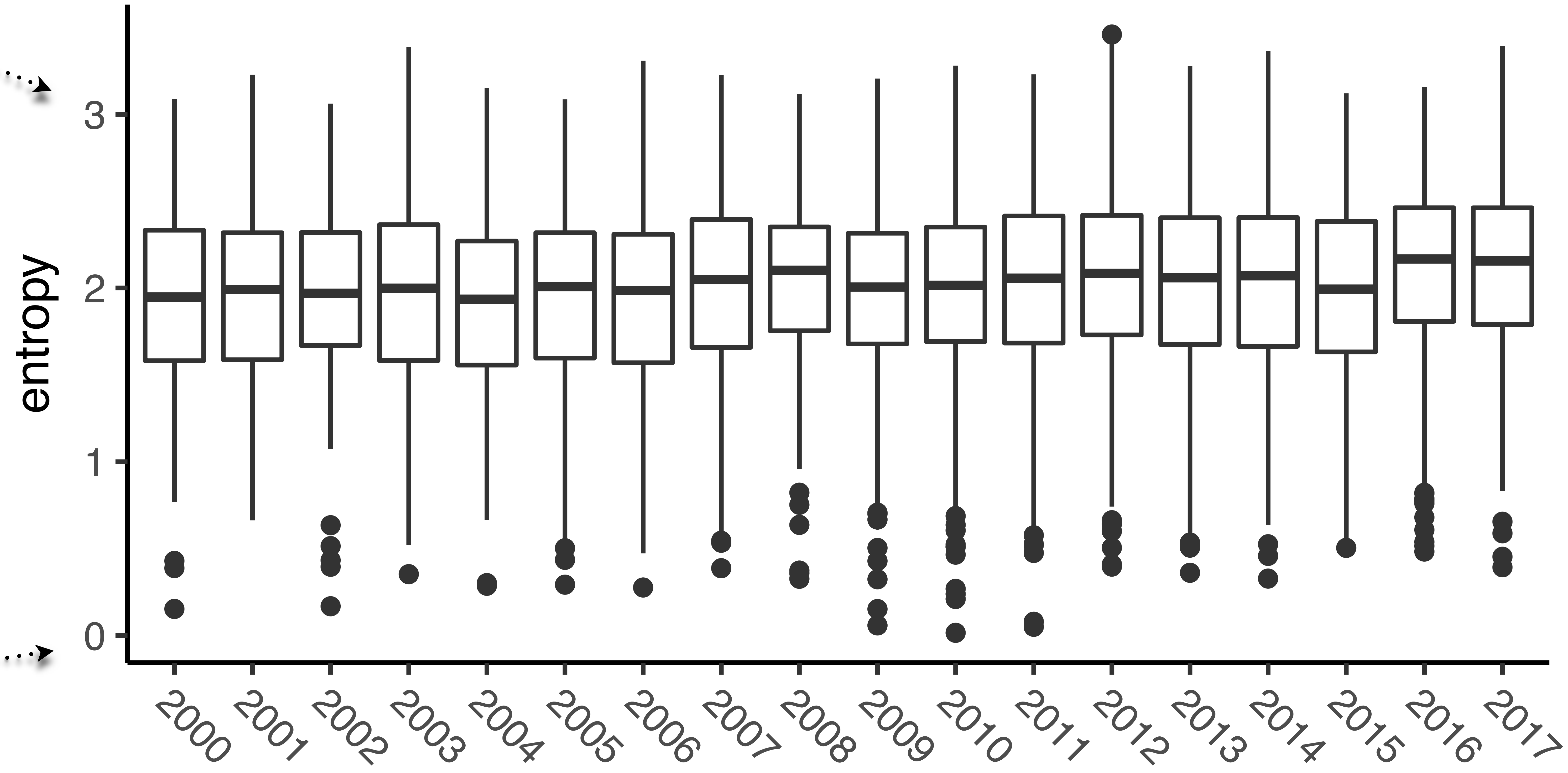

Causal reasoning
Decision making
Mathematical psychology
Probabilistic modeling
Language: syntax
Neural network
Knowledge structure
Language: semantics
Categorization
Developmental psychology


Spatial cognition and embodied cognition
Communication
Decision making
Mathematical psychology
Face and emotion perception
Visual attention
Language: syntax
Probabilistic modeling
Knowledge structure
Neural network
nory
Language: semantics
Categorization
Sequential learning
Developmental psychology



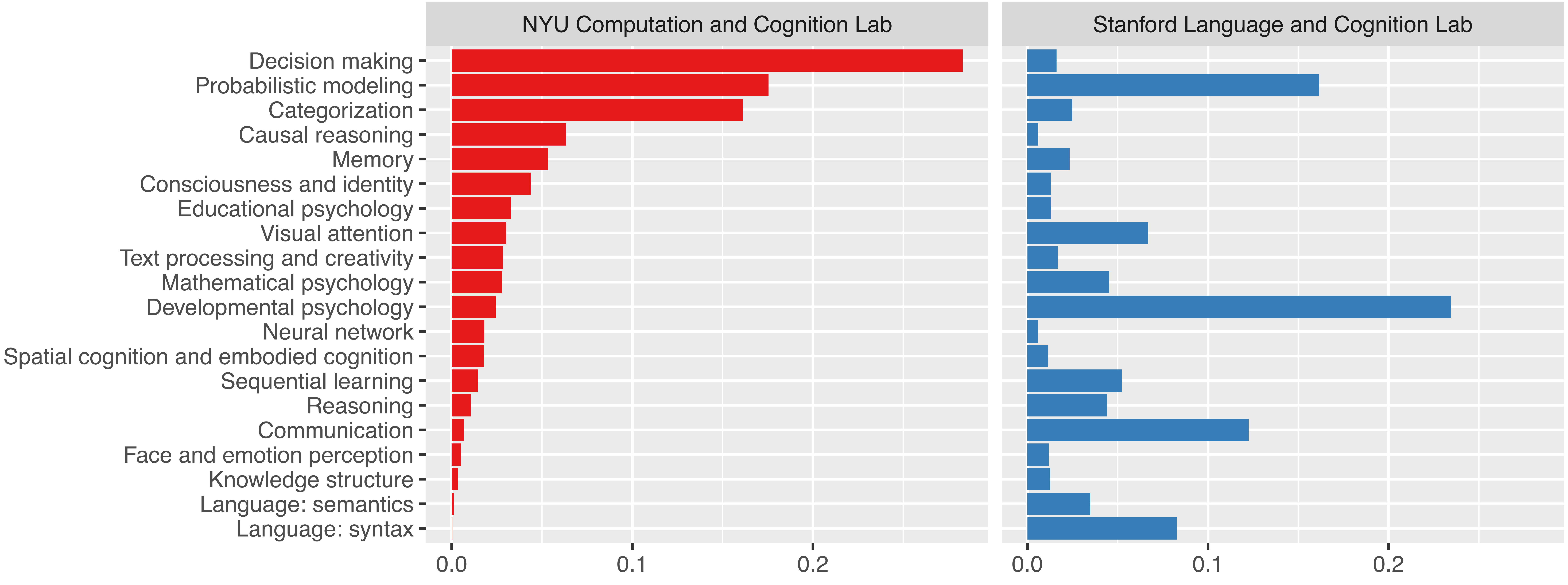


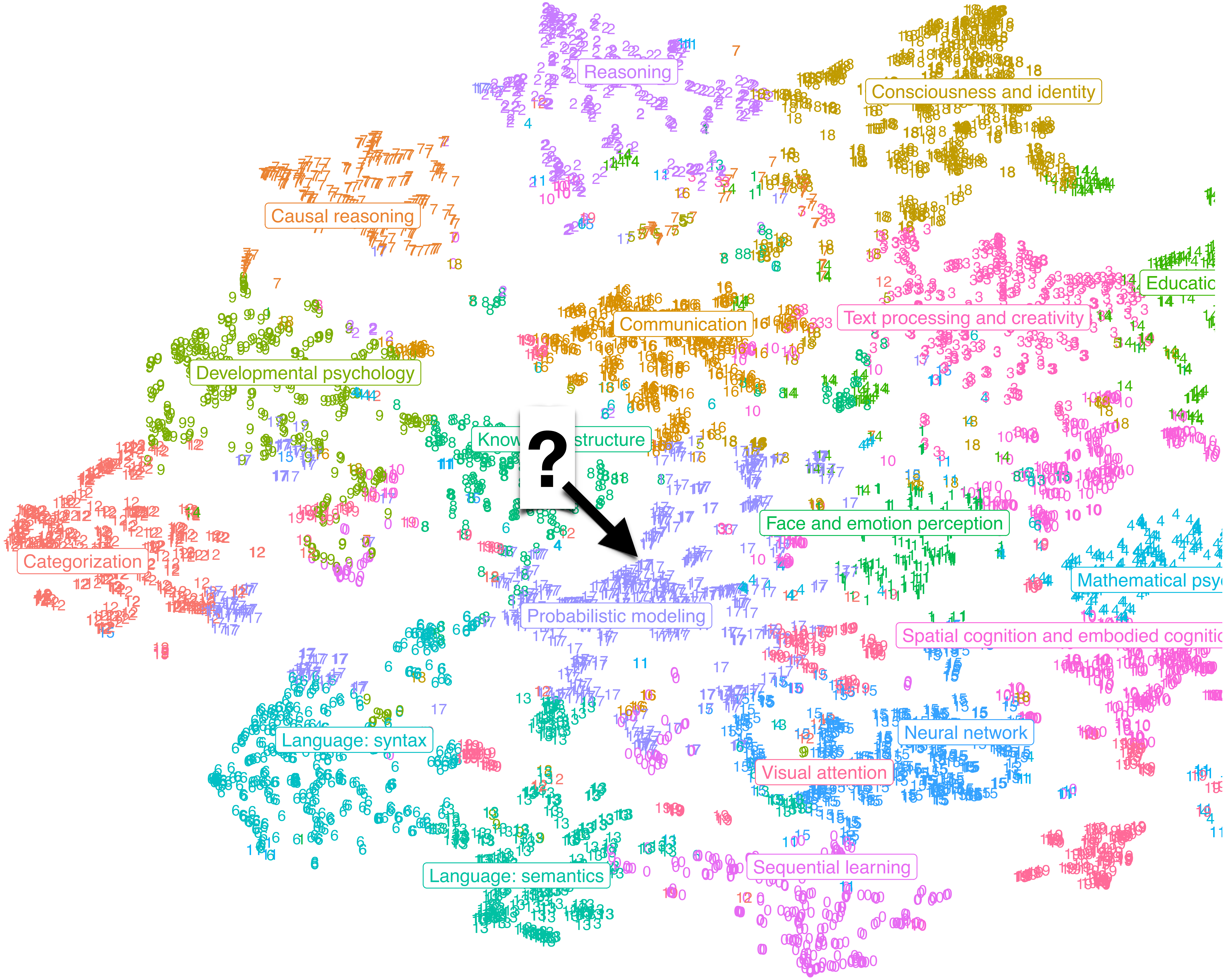
Entropy (within papers) increases by 0.01 per year ($p < 0.001$)



Compare labs

- Average topic proportions across the CogSci papers published by the lab





Paper recommendation system

- Find the 6 most similar papers
- based on the cosine angle between their topic vectors
- anselmrothe.github.io/dtm/

CogSci2018 Archives of the Cognitive Science Society Conference Proceedings

Find similar papers!

Convergence Bounds for Language Evolution by Iterated Learnin

Convergence Bounds for Language Evolution by Iterated Learning

Similarity	Year	Title
0.996	2014	Percentile analysis for goodness-of-fit comparisons of models to data
0.995	2005	A Bayesian View of Language Evolution by Iterated Learning
0.993	2002	A probabilistic approach to semantic representation
0.993	2005	Modeling Individual Differences with Dirichlet Processes
0.993	2013	Inferring Subjective Prior Knowledge An Integrative Bayesian Approach
0.991	2009	Iterated learning in populations of Bayesian agents

Paper recommendation system

- Find the 6 most similar papers
- based on the cosine angle between their topic vectors
- anselmrothe.github.io/dtm/

Summary & Discussion

- We used **Dynamic Topic Modeling** to get an overview over the last 18 years of CogSci publications

- "**Probabilistic modeling**" is on the rise
- The **complexity** (entropy) of the field is slowly increasing

- The **"content" of topics** was allowed to change over time
- The **number of topics** was fixed: 20 as a reasonable middle ground
  - Same as Blei & Lafferty 2006
  - Model log-likelihood improves with more topics (we tested up to 100)

- **Future work**: Enrich data set with google scholar citations + author + institution data

- **Code and results**: github.com/anselmrothe/dtm

THANKS FOR YOUR ATTENTION