Lecture 2:
Concepts & Categorization

COGS 153

What's in a word?

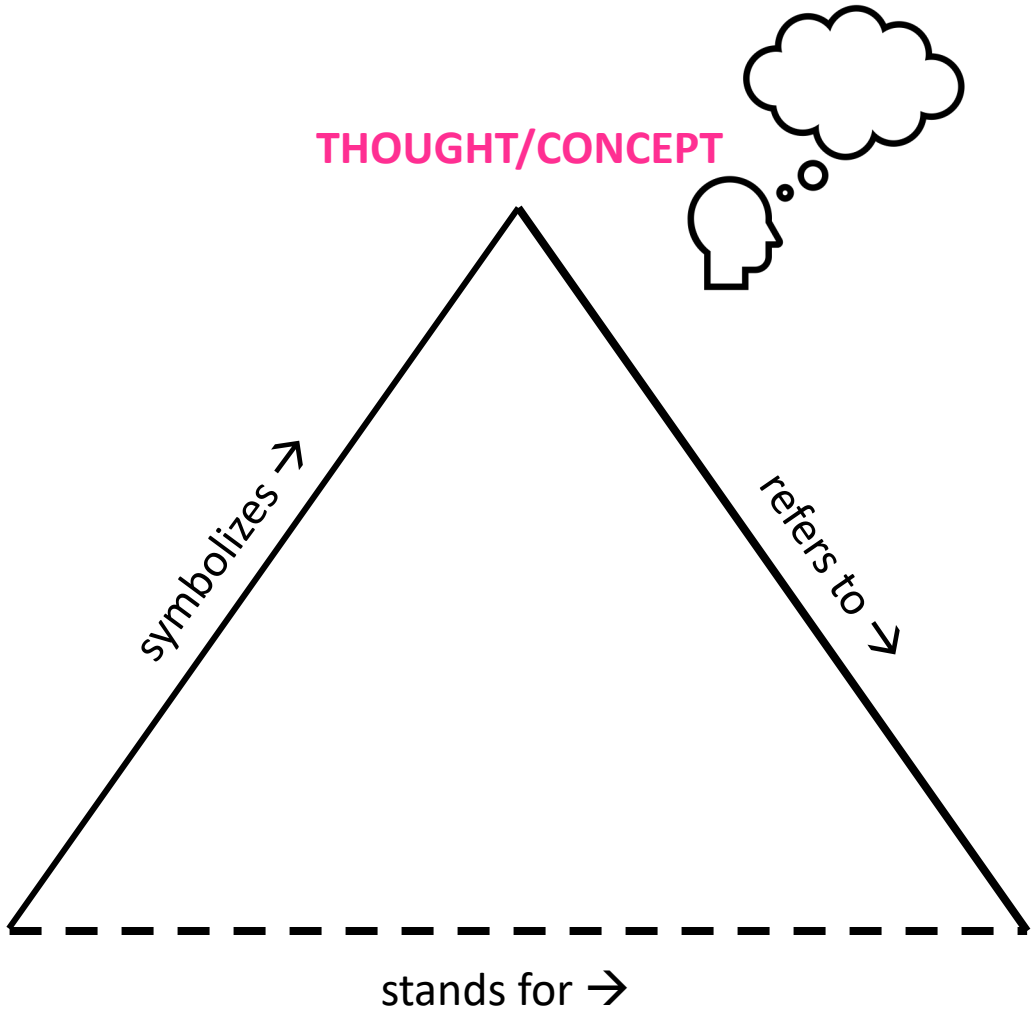
- Words are (usually) arbitrary symbols
  - *arbitrary symbol* – the representational form is arbitrarily related to the meaning of the referent
    - i.e. there is no inherent connection (or perceptual resemblance) between a sign and a referent

- Semanticity: a communication system conveys meaning through a set of fixed relationships among symbols & referents
  - Specific signals can be matched to specific meanings (simplified)
  - Referentiality: meanings are about things in the world

"That which we call a rose by any other name would smell as sweet"

The "Semiotic Triangle"

Ogden & Richards, 1923
Peirce, 1903



REFERENT/OBJECT IN REALITY

LINGUISTIC FORM
"tree"

linguistic forms
are *arbitrary*

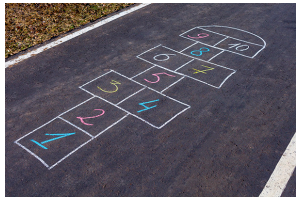
What’s a “chair”?

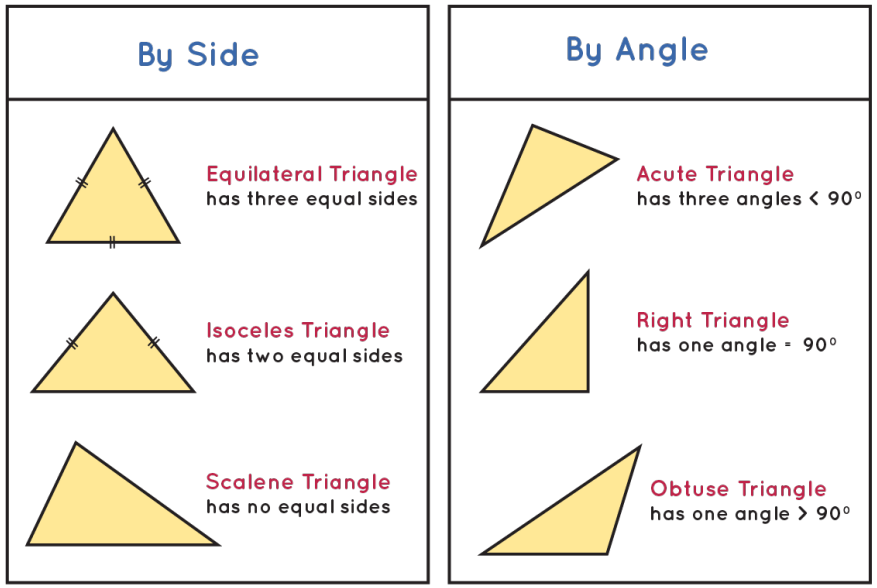
What’s a “chair”?

What's a "chair"?

The problem with definitions…

- In a list of features, how many are **necessary and sufficient**?
  - works for some concepts (*triangle*)
  - ...but not for others (*game*)
- Is something like a perfect dictionary definition enough?





Categories, concepts, and meaning

- **Category**: a collection of instances that are treated as having something in common
  - Afford inferences about the members
  - Humans are good at this!
- **Concept**: knowledge about a category
  - Determines how things are related or categorized
- How does the concept of DOG relate to the members of the category DOG?
- How is the meaning of DOG represented in the mind?

Categories, concepts, and meaning

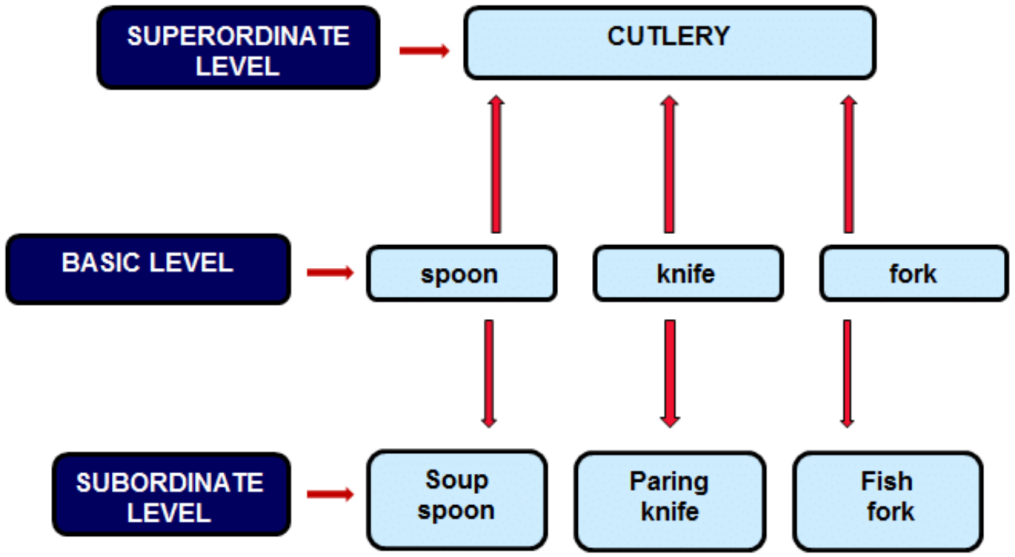
- **Category**: a collection of instances that are treated as having something in common
  - Afford inferences about the members
  - Humans are good at this!
- **Concept**: knowledge about a category
  - Determines how things are related or categorized

- All words have an underlying concept, but not all concepts have words
  - “schadenfreude” in German vs other languages

What is the organization of semantic memory?

- **Categories are not arbitrary** (Rosch, 1978)
  - Determined in part by how we perceive the structure of the world
  - e.g., color names relate to perceptual constraints

- Some categories are hierarchically-organized
  - Superordinate level
    - most abstract and general
  - Basic level
    - most frequently used in language
  - Subordinate level
    - most specific



- A trade-off between information and economy

Amodal (or arbitrary symbol) theories of categorization

- **Amodal**: without modality (modality is referring to senses)
  - e.g. information from visual modality are transduced into an *amodal representation* such as a frame, semantic network, or feature list
- Classical theories
  - necessary and sufficient theories (slide 7)
- Network theories
  - hierarchical semantic network theory
  - spreading activation theory
- Family resemblance theories
  - prototype theory
  - exemplar theory
- Distributional semantics
  - latent semantic analysis
  - BERT
- Frame semantics
  - FrameNet

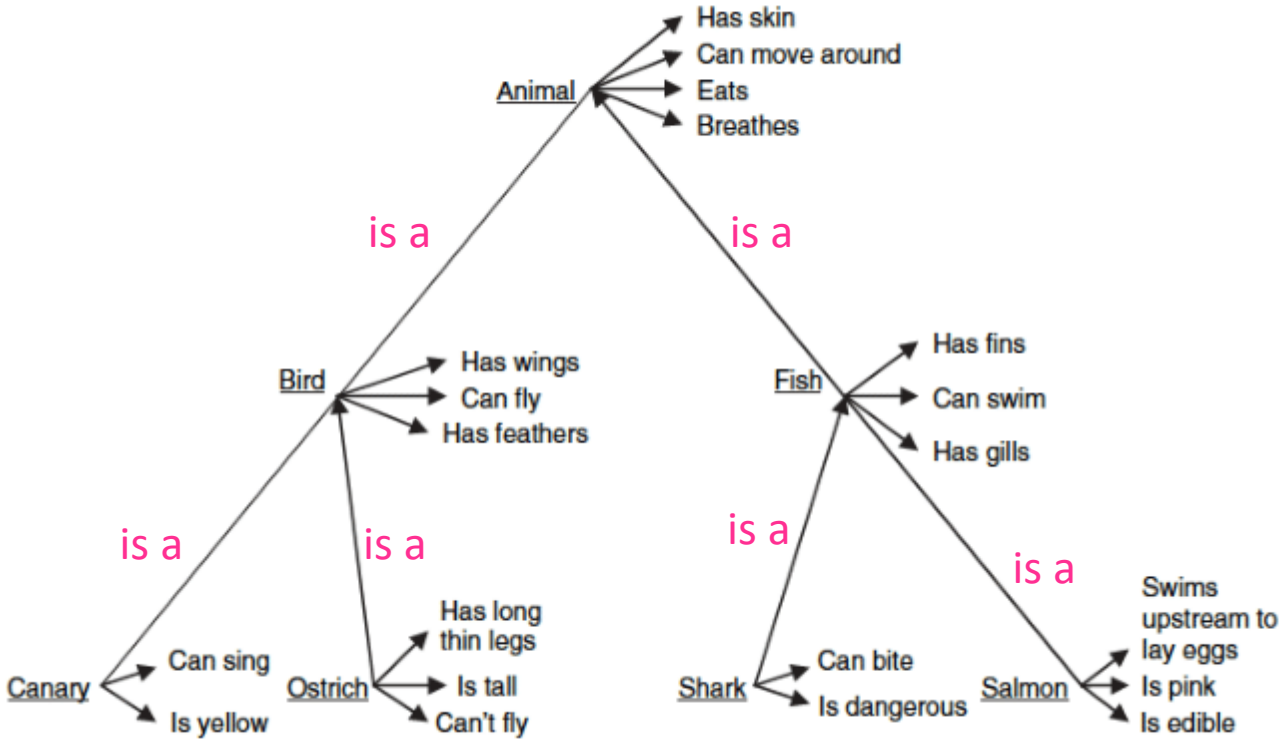
perception of a chair

Chair
4 legs
1 seat
1 back

amodal representation of a chair

Hierarchical semantic network models

- A concept's meaning depends on its relationship to other concepts in a hierarchically organized network
  - Has "is a" links (e.g. a canary IS A bird)
  - Has attribute/feature links (e.g. *can sing*)
  - Attributes of higher nodes are true for all lower nodes attributes (e.g. birds eat)



Collins & Quillian (1972)

Is hierarchical organization psychologically real?

Procedure:

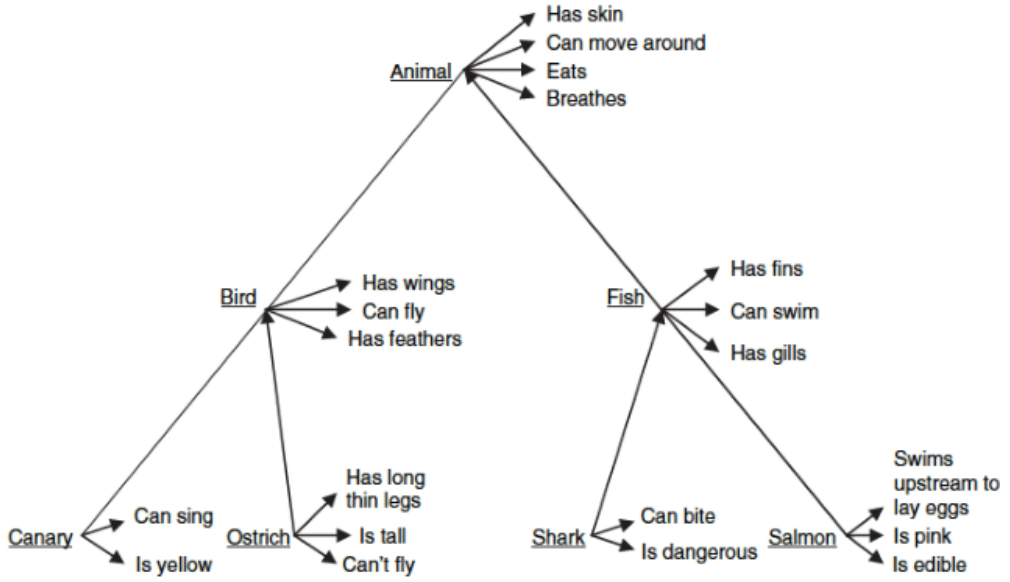
Sentence verification task with a two-alternative-forced-choice task (2AFC: true/false)

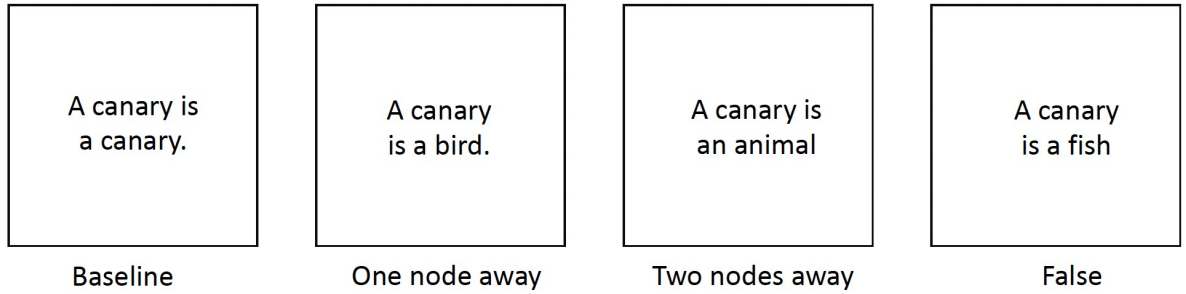
Dependent variable: reaction time (RT)

Results:

baseline RT < 1 node away RT < 2 nodes away RT

→ Faster RT's for shorter semantic distances





Collins & Quillian, 1969

Is hierarchical organization psychologically real?

Procedure:

Sentence verification task with a two-alternative-forced-choice task (2AFC: true/false)

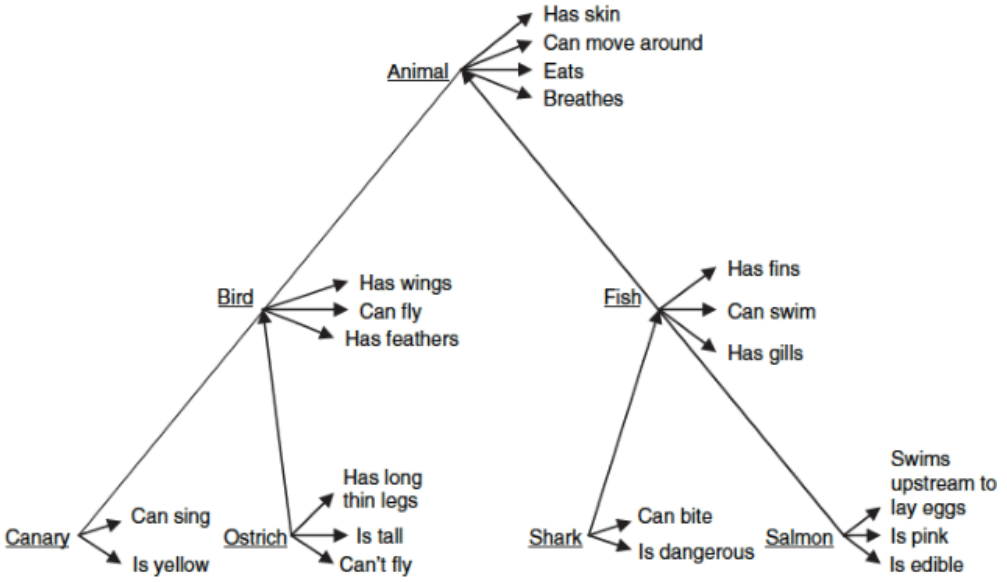
Dependent variable: reaction time (RT)

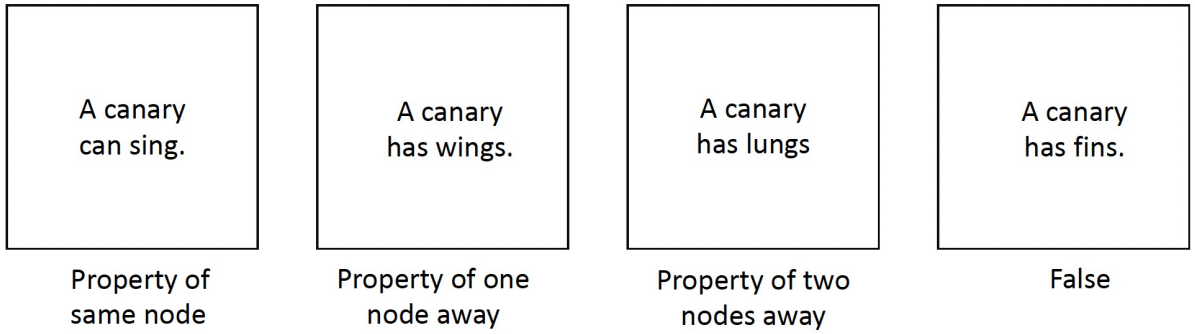
Results:

property of same node RT < property of 1 node away RT

< property of 2 nodes away RT

→ Faster RT's for shorter semantic distances





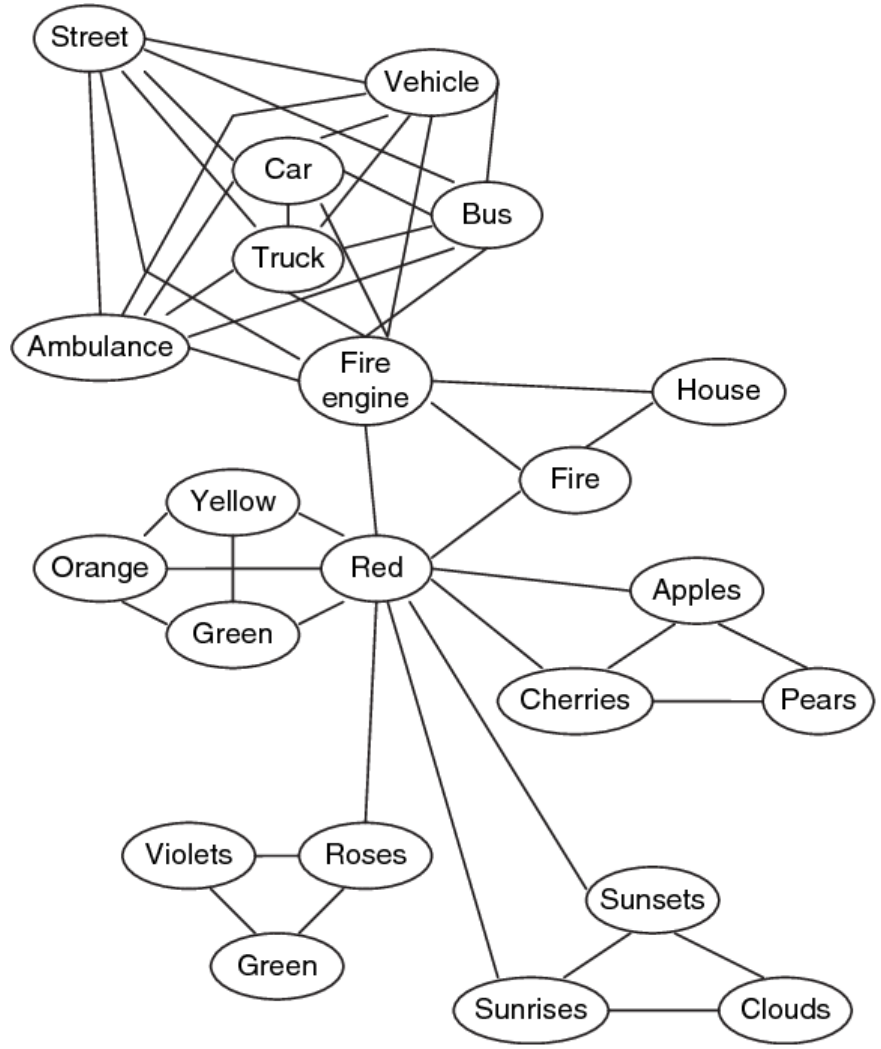
Collins & Quillian, 1969

Critique of Collins & Quillian's hierarchical model

- Not all concepts are hierarchically organized
  - e.g., relationship between truth, justice, law
- Methodological confound: semantic distance and conjoint frequency
  - e.g., robin and bird co-occur in sentences more frequently than robin and animal
- Controlling for conjoint frequency gets rid of the hierarchical relationship effects
- Makes some incorrect predictions
  - RTs for "Cow is a mammal." > "Cow is an animal."
  - RTs for "A robin is a bird." < "A penguin is a bird."
  - Relatedness effects: Rejecting "A pine is a flower." > "A pine is a church."

Spreading activation model of conceptual organization

- Activating one node spreads activation to connected nodes
- Conceptual nodes are linked **via association**
  - Semantic distance based on association (e.g., conjoint frequency)
  - Semantic distance: a metric based on the number of edges separating two nodes in a semantic network
- Activation above threshold brings concept to conscious awareness
- Consistent with data from many paradigms
  - e.g. in lexical decision task, faster RT to *nurse* following *doctor* compared to *pen*
  - faster to bring *nurse* to activation threshold because *doctor* was already activated



Collins & Loftus, 1975

Family resemblance models of categorization

Family members share features, but not all features are present in each member

- **Prototype & exemplar theories**
  - Inspired by theory of family resemblance (Wittgenstein, 1953)

Prototype theory

- Radial categories
  - At the center is **an abstract, schematic average** (a prototype)
  - There is a graded degree of belonging to a conceptual category, i.e. some members are more central than others (robin vs ostrich)
- Prototype: a schema filled with average values in its slots – **the most average member of the category**
  - Schema: an abstract frame for organizing knowledge
    - e.g., some slots in a schema for BIRD
      - Can fly?
      - Beak length:
      - Leg length:
- Categorization involves comparing a member to a prototype

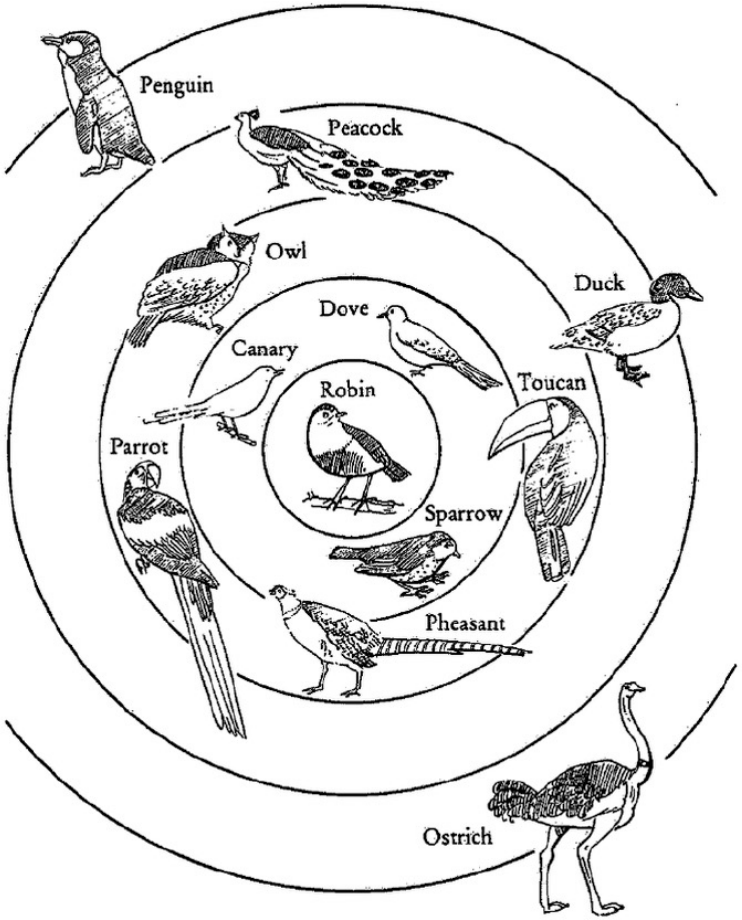


Figure 1 Birdiness rankings

Rosch, 1978

Is prototype theory psychologically real?

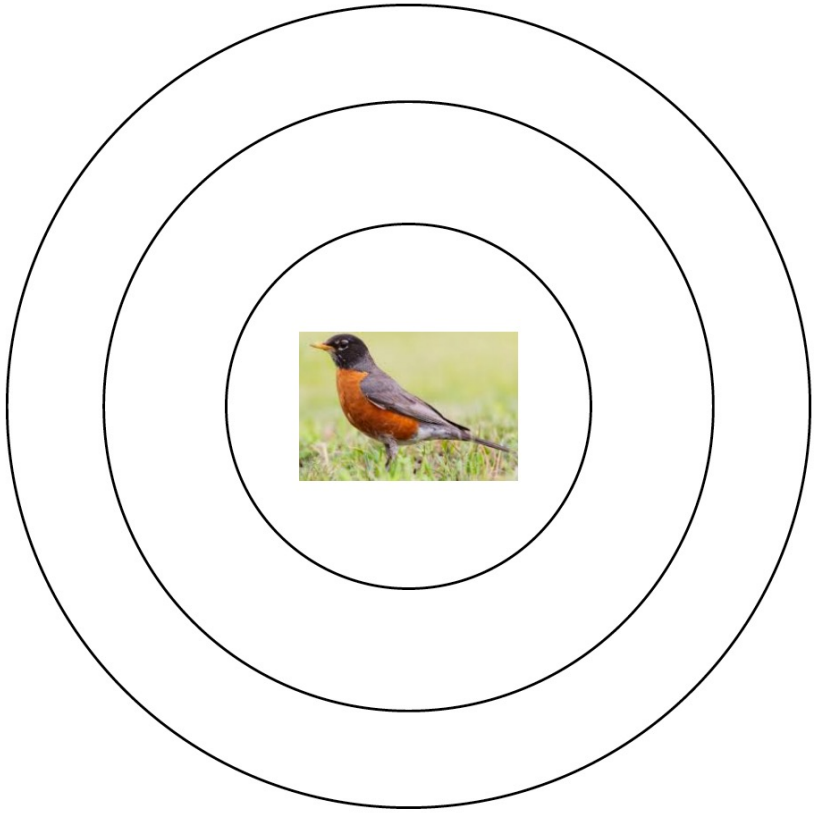
- Sentence verification tasks
  - RTs: “A robin is a bird.” < “An ostrich is a bird.”
- Feature listing & group membership listing tasks (e.g., Kail & Nippold, 1984)
  - Ppts tend to write robin before ostrich for birds
  - Ppts tend to write green before orange when listing features of vegetables
  - Words for prototypical objects tend to be learned earliest
- Prototypical members share more features with other members of the category but minimize overlap with related categories (Rosch & Mervis, 1975)
  - E.g., Apple is close to the prototypical fruit but tomato is not

Problems with prototype theory

- Not all categories have prototypes
  - What is the prototype of truth? (Hampton, 1981)
  - Balan – an aboriginal category for “women, fire, and dangerous things”
  - What is the prototype for Balan? (Lakoff, 1987)
- The graded structure of categories is dynamic, not static
- Prototypicality effects vary by context (Barsalou & Sewell, 1984)
  - “Stacy volunteered to milk the animal.”
  - “Fran pleaded with her father to let her ride the animal.”
- In some situations, central tendencies determine graded structure and in others it is ideals that determine graded structures (Barsalou, 1985)

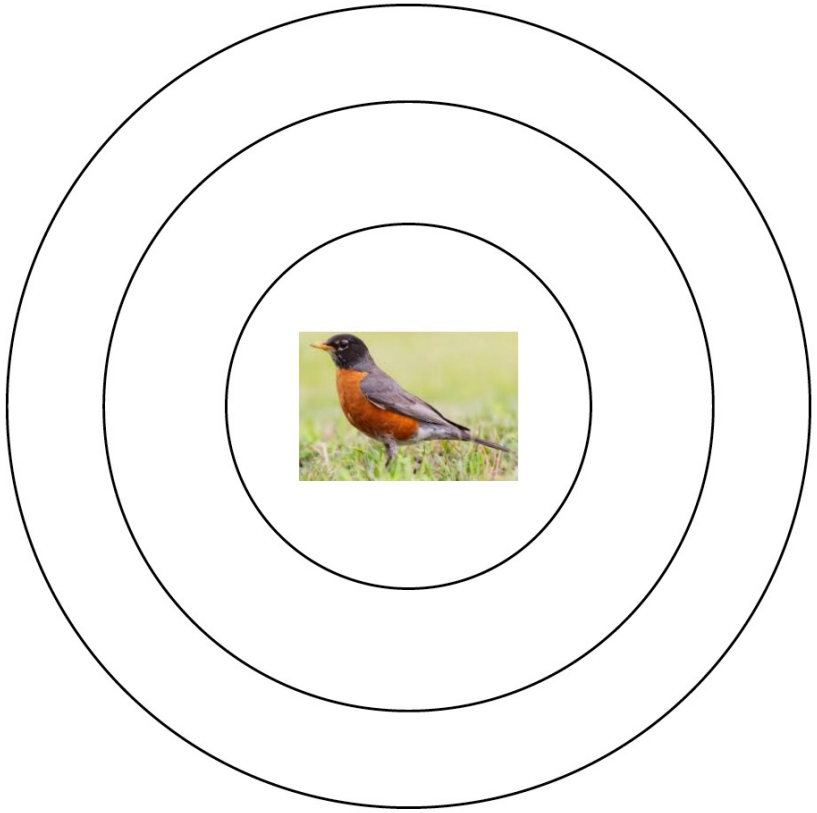
Exemplar theory

- Also has radial categories
  - The center is an EXEMPLAR, the most birdy of birds
  - **Not an average**
- New instances are compared to the specific exemplar instance stored in memory
  - The new instance is assigned to a category based on the greatest number of similarities it holds with exemplar
  - If a new instance is similar enough to the bird examples stored in memory, then it gets categorized in the "bird" category



Exemplar theory

- Also has radial categories
  - The center is an EXAMPLAR, the most birdy of birds
  - Not an average
- Comparisons are made to a specific instance stored in memory
  - the instance is still represented as an abstract feature and relational list, but with values based on a specific example rather than the average
- Makes identical predictions to prototype theory in most cases
- Careful research suggests that people rely on prototypes and exemplars, depending on the context (situation and experience)



Distributional accounts of meaning

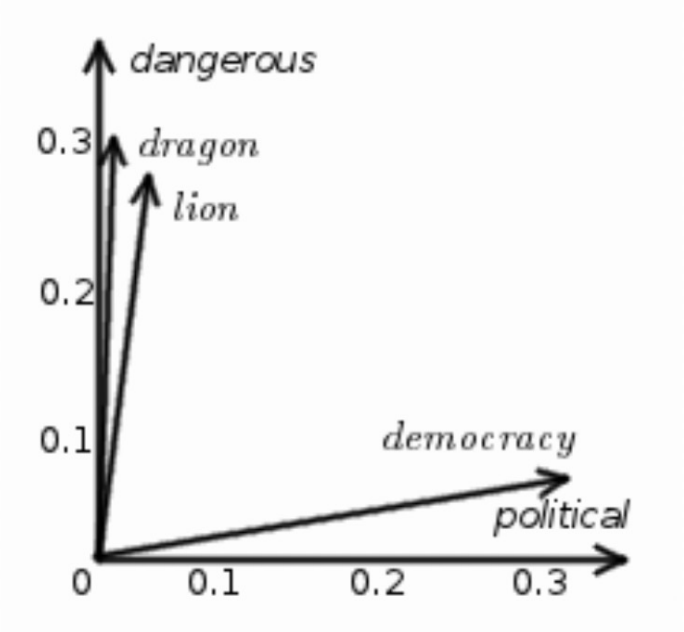
- Idea behind distributional accounts: **The statistical regularities in how words are used can be used to represent their meaning**
- Linguistic items (e.g., words) used in similar contexts have similar meanings.
  - I have a pet cat at home.
  - I have a pet dog at home.
  - I have a refrigerator at home.
  - I have a pet refrigerator at home.*
  - I took my cat to the vet.
  - I took my dog to the vet.
  - I took my refrigerator to the vet.*
  - The electricity went out and my refrigerator won't work.
  - The electricity went out and my lamp won't work.
  - The electricity went out and my cat won't work.*

"You shall know a word by the company it keeps."

– Firth (1957)

Distributional semantics models

- Evaluating distributional accounts using **computational models**
- Natural Language Processing (NLP)
  - Analyze Corpus data for **word co-occurrences**
    - How often does the word dragon co-occur with the word dangerous?
    - How often does the word dragon co-occur with the word political?
    - How often does the word lion co-occur with the word dangerous?
    - How often does the word lion co-occur with the word political?
    - How often does the word democracy co-occur with the word dangerous?
    - How often does the word democracy co-occur with with the word political?
  - Create a multi-dimensional semantic space model
    - Similar words have similar vectors in the semantic space
      - shorter semantic distance -> closer in meaning
- Very good at modeling word-word similarity judgments made by humans
  - may support the psychological plausibility of this type of representation



Latent Semantic Analysis (LSA)

A distributional semantics model

- Create a matrix of **the relationship between words and their appearance in documents or passages**
- Apply singular vector decomposition (SVD) for maximal dimensionality reduction
- *Semantic similarity* is measured using the cosine between vectors that represent words or documents in LSA space
  - "Human" does not co-occur with "user" but, "Human" co-occurs with "computer" and "User" co-occurs with "computer"

Example of text data: Titles of Some Technical Memos

c1:	*Human* machine *interface* for ABC *computer* applications
c2:	A *survey* of *user* opinion of *computer system response time*
c3:	The *EPS user interface* management *system*
c4:	*System* and *human system* engineering testing of *EPS*
c5:	Relation of *user* perceived *response time* to error measurement
m1:	The generation of random, binary, ordered *trees*
m2:	The intersection *graph* of paths in *trees*
m3:	*Graph minors* IV: Widths of *trees* and well-quasi-ordering
m4:	*Graph minors*: A *survey*

$\{X\} =$

	c1	c2	c3	c4	c5	m1	m2	m3	m4
human	1	0	0	1	0	0	0	0	0
interface	1	0	1	0	0	0	0	0	0
computer	1	1	0	0	0	0	0	0	0
user	0	1	1	0	1	0	0	0	0
system	0	1	1	2	0	0	0	0	0
response	0	1	0	0	1	0	0	0	0
time	0	1	0	0	1	0	0	0	0
EPS	0	0	1	1	0	0	0	0	0
survey	0	1	0	0	0	0	0	0	1
trees	0	0	0	0	0	1	1	1	0
graph	0	0	0	0	0	0	1	1	1
minors	0	0	0	0	0	0	0	1	1

Landauer, Foltz & Laham (1998)

LSA continued

- it's a *"bag of words"* analysis
  - **Word order does not matter in this analysis!**
  - Meaning is based on the contexts that a word does and does not appear in
    - Words with similar meaning tend to share contexts, whereas words that are dissimilar in meaning tend not to share contexts
- The model performs similarly to humans
  - On standard vocabulary & subject matter tests
  - On lexical priming tests
  - Also predicts human judgments of passage coherence, learnability of passages, and the quality and quantity of knowledge contained in essays

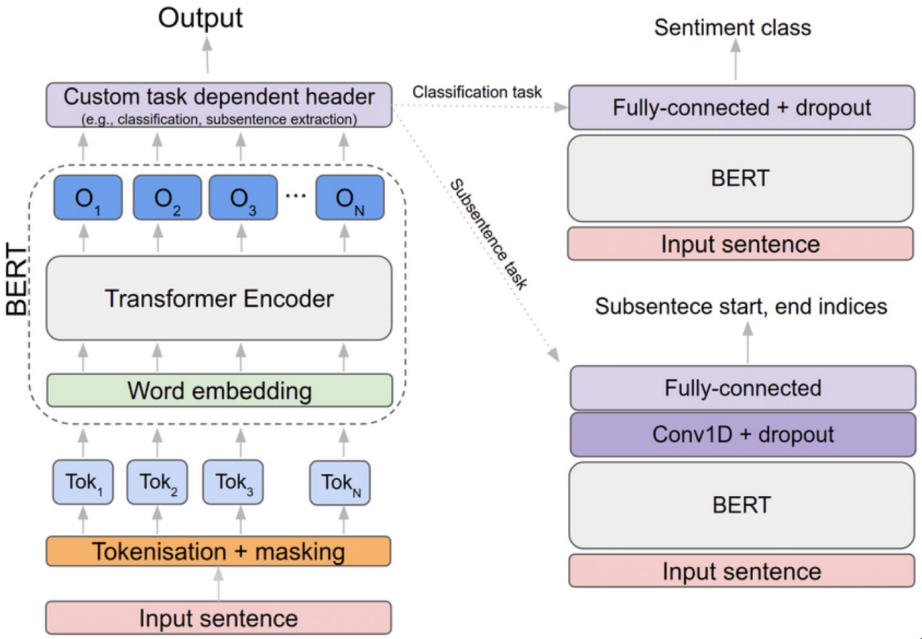
doctor–physician .61
go–went .71
good–bad .65
she–her .98
blackbir–bird .46
blackbir–black .04
telephone– shark .01

LSA similarity measures
(from Kintsch & Rawson, 2004)

Landauer, Foltz & Laham (1998)

BERT

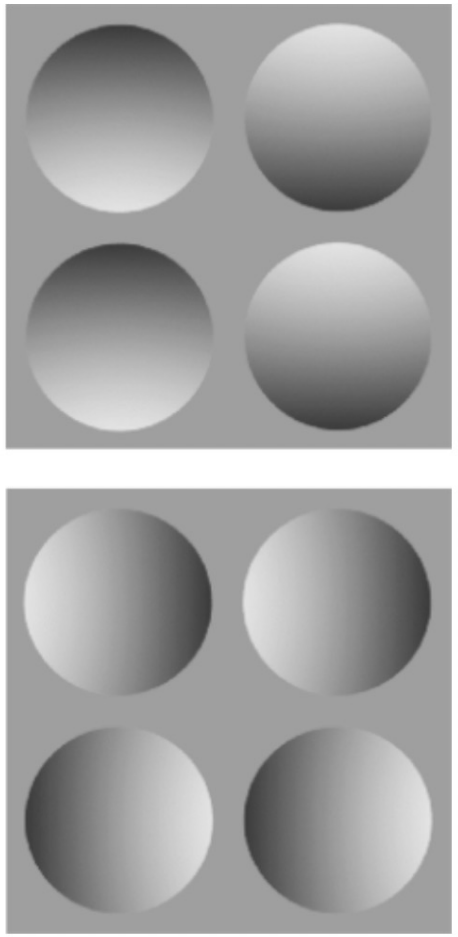
- Distributional model
- BERT: Bidirectional Encoder Representations from Transformers
  - Bidirectional analysis of words in context
  - **Word order matters**
  - **More sophisticated** and powerful than LSA and other bag-of-words models



Devin et al., 2018

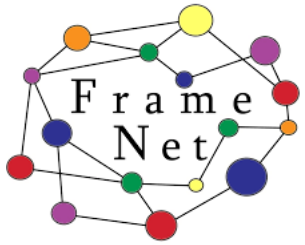
More on distributional accounts

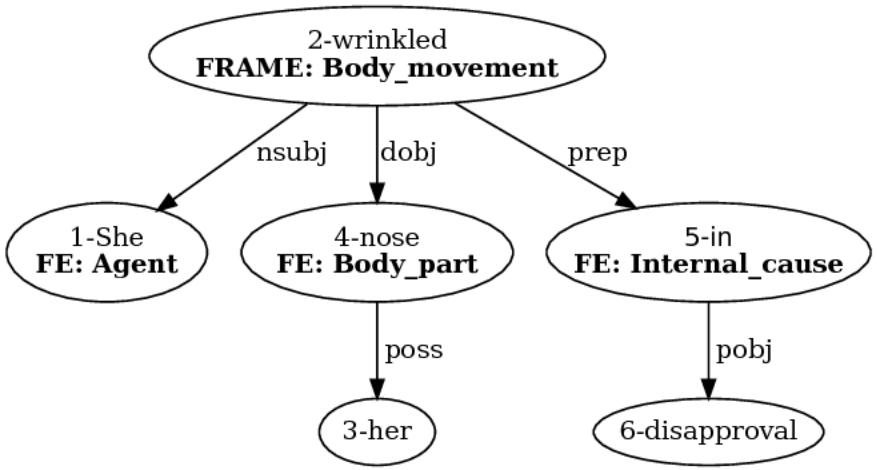
- The brain is very good at picking up on statistical regularities
  - perceptual inference
    - e.g., shape from shading (Freeman, 1994) & motion (Weiss et al 2022)...
  - cognitive judgments
    - e.g., inferences about categorization (Kemp et al, 2004) and the duration of events (Griffiths & Tenenbaum, 2006)
- Statistical models of language comprehension (based on distributional frequencies) can...
  - predict upcoming words in context
    - for humans, prediction can facilitate comprehension
  - predict the gist of meaning and the sentiment of passages
    - critical aspect of comprehension in humans!



Griffiths, Steyvers, & Tenenbaum, 2007

FrameNet





- "The FrameNet corpus is a lexical database of English that is both human- and machine-readable, based on annotating examples of how words are used in actual texts."
  - based on a theory of meaning called Frame Semantics, deriving from the work of Fillmore et al.
  - Frame Semantics: idea that we understand the meaning of a single word by accessing all the essential knowledge that relates to that word (like encyclopedic knowledge)
- "A **frame** is a script-like conceptual structure that describes a particular type of situation, object, or event along with the participants and props that are needed for that Frame"
  - The roles of a Frame "frame elements" (FEs) and the frame-evoking words are called "lexical units" (LUs)

https://framenet.icsi.berkeley.edu/

Animals

Definition:

The LUs in this frame refer to biological entities other than human beings and plants, which are labeled by the FE Animal.

My husband and I adopted a really sweet dog last year.

The cows that live on the farm down the road are always mooing loudly.

FEs:

Core:

Animal [An] — **Semantic Type:** Living_thing
The living organism belonging to the kingdom Animalia.
Did you see that BIRD in the tree?

Non-Core:

Characteristic [Char]
A quality of the Animal.
FOXES *are* sly.

Descriptor []
A characterization of the Animal not covered by another frame element.
The DOGS that live next door are so well-behaved.

Origin []
The location from which the Animal originated.

Frame-frame Relations:

Inherits from: Biological_entity

Summary

- The brain is for moving and acting in the world
  - Llinas, 2001; Wolpert, 2011; Glenberg, 1997; Glenberg, 2020
- Concepts and Categorization
  - There’s fuzzy boundaries!
  - We make inferences about novel situations
  - Generalizing requires some abstraction
- The brain is excellent at detecting statistical regularities from distributional info
  - predictions based on experience