Words in the Mind

Priming and Semantic Memory
The Lexical Decision Task

- Originally used by Meyer and Schvaneveldt (1971).
- Subjects judge whether a string of letters is a word.
- Reaction time is the primary index of performance.
- Key findings:
  - Words recognized faster than non-words.
  - Frequent words recognized faster than infrequent words.
  - Presence of neighborhood competitors slows responses.
  - Priming speeds responses.
Varieties of Priming

- **Repetition priming**: the prime and the target are equivalent
- **Form priming**: the prime and the target have similar orthography (e.g. SHIP and SHOP)
- **Morphological priming**: the prime and the target share parts of their morphological structure:
  - *Semantically transparent*: TALKING and TALK
  - *Semantically opaque*: CORNER and CORN
- **Semantic priming**: the prime and the target have similar meaning (e.g. SHIP and BOAT)
Why is there Priming?

Spreading Activation

- A form or meaning is activated
- Activation spreads to related forms or meanings
- By measuring relative amounts of priming, it may be possible to infer how particular forms or meanings are related.
Sample Priming Tasks

- **Primming Across Trials**
  - Lag = 1
  - Sentence verification

  **Trial n**
  - "True" response
  - Prime: Truck has wheels

  **Trial n + 1**
  - "False" response
  - Target: Apple has legs

  **Trial n + 2**
  - "True" response
  - Target: Car has brakes

- **Primming Within Trials**
  - SOA = 500 ms
  - Lexical decision

  **Trial n**
  - Prime: Fruit
    - 500-ms SOA
    - Target: Fruit, Apple
      - "Yes" response

  **Trial n + 1**
  - Prime: Mammal
    - 500-ms SOA
    - Target: Mammal, Pune
      - "No" response
Studying Priming with the Lexical Decision Task

<table>
<thead>
<tr>
<th>Top String</th>
<th>Bottom String</th>
<th>Correct Response</th>
<th>Sample Stimuli</th>
<th>Mean RT (ms)</th>
<th>Mean Percentage Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>Associated word</td>
<td>Yes</td>
<td>Nurse–doctor</td>
<td>855</td>
<td>6.3</td>
</tr>
<tr>
<td>Word</td>
<td>Unassociated word</td>
<td>Yes</td>
<td>Bread–doctor</td>
<td>940</td>
<td>8.7</td>
</tr>
<tr>
<td>Word</td>
<td>Nonword</td>
<td>No</td>
<td>Book–marb</td>
<td>1,087</td>
<td>27.6</td>
</tr>
<tr>
<td>Nonword</td>
<td>Word</td>
<td>No</td>
<td>Valt–butter</td>
<td>904</td>
<td>7.8</td>
</tr>
<tr>
<td>Nonword</td>
<td>Nonword</td>
<td>No</td>
<td>Cabe–manty</td>
<td>884</td>
<td>2.6</td>
</tr>
</tbody>
</table>
How are Word Meanings Related?

As a first stab at how word meanings are related in semantic memory, we might suppose that categories of objects form the backbone of structure.

• Categories are hierarchical
  – ‘dog’ vs. ‘mammal’
  – ‘dog’ vs. ‘collie’

• Activation spreads through memory at a fixed rate

• If memory is limited, cognitive economy is critical
  – Properties are linked only to the most general category to which they apply.
Collins and Quillian's (1969) Hierarchical model of semantic memory

Animal

Bird

Canary

Concepts are organized in a hierarchy.
Collins and Quillian's (1969) Hierarchical model of semantic memory

Properties are associated with the most general concepts for which they are true.
Collins and Quillian's (1969) Hierarchical model of semantic memory

- Animal
  - Has skin
  - Eats
  - Breathes

- Bird
  - Has wings
  - Can Fly

- Canary
  - Can sing
  - Is yellow

Retrieving information from the network takes time.
Do canaries have skin?

Yes. (1433 ms)

Retrieving information from the network takes time.
Two computational predictions

Bird
  - Has wings
  - Can Fly

Animal
  - Has skin
  - Eats
  - Breathes

Canary
  - Can sing
  - Is yellow

Time to move up in the hierarchy and time to retrieve property information is additive.
Two computational predictions

- Has wings
- Can Fly
- Has skin
- Eats
- Breathes

Time to retrieve a property from a node is independent of the level of the node.

- Canary
  - Can sing
  - Is yellow
Collins and Quillian's (1969) Experiment

- The participants were asked to evaluate the truth value of sentences thought to reflect the organization of semantic memory.

- Is a canary a bird? Yes (+ RT)
- Do canaries swim? No (+ RT)
## Materials

<table>
<thead>
<tr>
<th>Superset relations</th>
<th>Property relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>A canary is a canary (S0).</td>
</tr>
<tr>
<td>1</td>
<td>A canary is a bird (S1).</td>
</tr>
<tr>
<td>2</td>
<td>A canary is an animal (S2).</td>
</tr>
</tbody>
</table>
Results

Superset relations

Property verification
Collins and Quillian (1969)

- Conclusions:
  - The assumptions of the hierarchical model are supported.
  - Moving up one level in the hierarchy takes 75 ms and retrieving a property takes 225 ms.
Problems with the hierarchical model

- Although Collins and Quillian's model is parsimonious and initially received empirical support, it has important problems.
- The most central of these is that the model assumes that categories in semantic memory are Aristotelian (a/k/a classical or rule-governed categories).
- This leads to a variety of problems resulting from the facts that:
  - not all category members are equally good (viz. typical) category exemplars.
  - Boundaries of many categories are fuzzy.
Aristotelian Categories

• Classical, all-or-none, rule-governed
  – Examples:
    \[ \text{even number}=\text{integer divisible by 2} \]
    \[ \text{grandmother}=\text{mother of a parent} \]

• Defining features
  – list of properties common to all the members of a category (necessary conditions = genus proximum)
  – list of properties common only to the members of that category (sufficient conditions = differentia specifica)

• Intension=Extension
Sense and Reference

• Reference is “what is picked out in the world”
• But terms with the same reference don’t always have the same sense.
• Frege: classical example concerns the ‘evening star’ and the ‘morning star’, both of which refer to Venus.
Sense and Reference
Word Sense

The sense of a linguistic expression:
the sum total of all of its *sense-relations*
with other parts of the linguistic system
• synonymy, antonymy, hyponymy…
• paraphrase, contradiction, entailment
Defining Features

People don’t know “defining” features of everyday concepts

Consider for example the proceedings that we call "games". I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all? -- Don’t say: "There must be something common, or they would not be called 'games' " - but look and see whether there is anything common to all. -- For if you look at them you will not see something that is common to all, but similarities, relationships, and a whole series of them at that.

Ludwig Wittgenstein
Philosophical Investigations
Prototype Theory (Eleanor Rosch)

• Concepts are organized around “prototypes” or “central members” of a category
• More & less typical instances
• Fuzzy boundaries
For the following categories, write down the first example that comes to mind.

- Fruit
- Tool
- Vehicle
- Item of clothing
Typicality Effects

Bachelor: Male and not married.
More & Less Typical Instances
(1 = typical, 7 = less typical)

- **Fruit:**
  - Apple = 1.07
  - Orange = 1.08
  - Raspberry = 2.15
  - Raisin = 3.42
  - Tomato = 5.58

- **Even Numbers:**
  - 4 = 1.1
  - 8 = 1.5
  - 10 = 1.7
  - 18 = 2.6
  - 106 = 3.9
Prototypes

• Are supplied as examples of a category
• Are judged more quickly
• Share more common attributes (family resemblance)
  – List features of fruit
  – List features of apple, raspberry, tomato etc.
  – Typical items share more features of category

(Rosch & Mervis)
Problems with the hierarchical model

Response times obtained in semantic categorization tasks do not all support the hierarchical model:

<table>
<thead>
<tr>
<th>Sentences</th>
<th>Response times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canaries (S0) are birds (S1).</td>
<td>1279 ms</td>
</tr>
<tr>
<td>Penguins (S0) are birds (S1).</td>
<td>1466 ms</td>
</tr>
</tbody>
</table>
Both concepts and properties are represented by nodes.
Collins and Loftus' (1975) Spreading activation model

The organization is non-hierarchical. Every node is connected to every node.
Quillian and Loftus' (1975) Spreading activation model

Only one node may be activated at one time.
Quillian and Loftus' (1975) Spreading activation model

However, once activated, nodes send activation to their neighbors in parallel.
Quillian and Loftus' (1975) Spreading activation model

The strength of association between nodes may be represented by distance (or connection strength).
Evidence in support of the Spreading Activation model

It accounts for the data presented in the categorization literature (typicality effects, exemplar effects, semantic categorization response times, …).
Conclusion

• The spreading activation model is very popular because it is intuitive and it provides an explanation for a wide array of phenomena.

• Its major problem is that it is too powerful.
An Alternative Approach: How are Words Learned?

Many avian species possess ctenoid crests.

In my opinion, the fenestration of Sayles Hall is marred by overly heavy muntins.
Word Learning: Adult Resources

- Use other words in the sentence.
- Use morphology
- Use a dictionary
- Ask someone
- Use real-world knowledge
Word Learning: Infant Resources

• Use other words in the sentence.
• Use morphology
• Use a dictionary
• Ask someone
• Use real-world knowledge
Principles and Constraints

- The Whole Object Assumption
- The Taxonomic Assumption
- Mutual Exclusivity
Principles and Constraints

• The Whole Object Assumption

• The Taxonomic Assumption

• Mutual Exclusivity
Whole Object Assumption in Initial Mappings

Evidence

Tendency toward whole object interpretation
- in ambiguous situations
- with inappropriate syntax
- in languages without count/mass distinction

More nouns in early vocabulary

Novel nouns learned faster
Whole Object Assumption

Criticisms

Only roughly 40% of early words are object labels.

Whole Object Assumption

Explanations

Object concepts are richer and more cohesive

Ostensive teaching of nouns more prevalent in middle-class Americans
Whole Object Assumption

Consequences - speculations

Verbs are harder to learn

Adjectives are harder to learn
e.g., color words (Soja, 1994)
Principles and Constraints

• The Whole Object Assumption

• The Taxonomic Assumption

• Mutual Exclusivity
Taxonomic Assumption in Word Extensions

Evidence

Find another dog
Taxonomic Assumption in Word Extensions

Conceptual extensions
Taxonomic Assumption in Word Extensions

Shape Bias

Find another dax
Principles and Constraints

• The Whole Object Assumption

• The Taxonomic Assumption

• Mutual Exclusivity
Mutual Exclusivity

Evidence

Fast mapping of novel label to novel object when familiar object is present

Lack of fast mapping of second labels
Mutual Exclusivity

Value

Fast-mapping

Avoiding redundant hypotheses

Overcoming whole object assumption

Overriding taxonomic assumption
Nature of Principles and Constraints

• Are they language specific?
• How and when do children overcome them?
• Are they applicable to learning of words in other classes?