Representations in the Mental Lexicon

Computational Cognitive Science, Lecture 17
Stella Frank, stella.frank@ed.ac.uk
November 12, 2019

Cognitive representations of words

Words are stored in a *mental lexicon:*

- How is this mental lexicon organised?

- How are the words represented?

Cognitive representations of words

Words are stored in a *mental lexicon:*

- How is this mental lexicon organised?

- How are the words represented?

Note: these are questions about representations, not processes or abilities (like learning).

Cognitive representations of words

Words are stored in a *mental lexicon:*

- How is this mental lexicon organised?

- How are the words represented?

Note: these are questions about representations, not processes or abilities (like learning).

Shift in methodology:
analyse evidence from experiments using representations, instead of "argument by model"
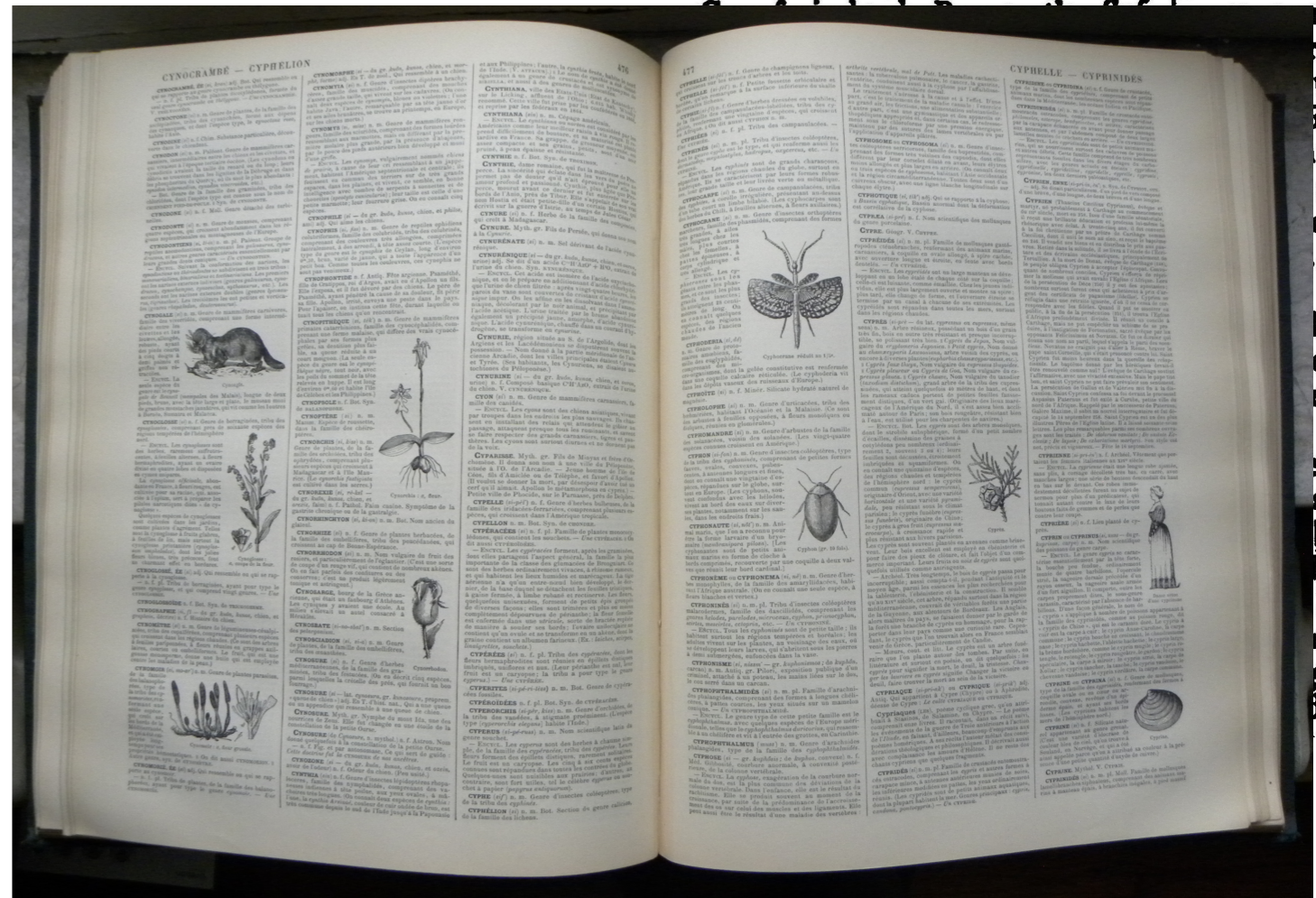
Roadmap

Today:

- Psychological methods for eliciting evidence about the types of word representations used by humans

- Computational methods for representing words and relations between words

Thursday:

- Tying these two together: do the computational methods fit the psychological evidence?

A Possible Lexicon

- Mental lexicon as dictionary?

- List all the words neatly in order

- Define them with a short description

- Find them in O(log *n*)

Cnáwan, he cnæwð; p. cneow; pp. cnáwen. *To* KNOW; noscere, v. oncnáwan.

Cnáwincg *Knowledge, a knowing*; cognitio, *Som.*

Cnear *A ship, a narrow ship, galley;* navis:—Cnear on flot *the ship on float, Athelst. Vict. Chr.* 938. Nægled on cnear-rum *in nailed ships, Id.*

Cneatung, e; *f. A debate, an inquiry, a search;* disputatio, scrutinium:—*Scint.* 14.

Cnedan, gecnedan; *pp.* cneden [*Dut.* kneden] *To* KNEAD, *ferment;* subigere:—*R. Lk.* 13, 21.

Cneo *a generation,* v. cneores.
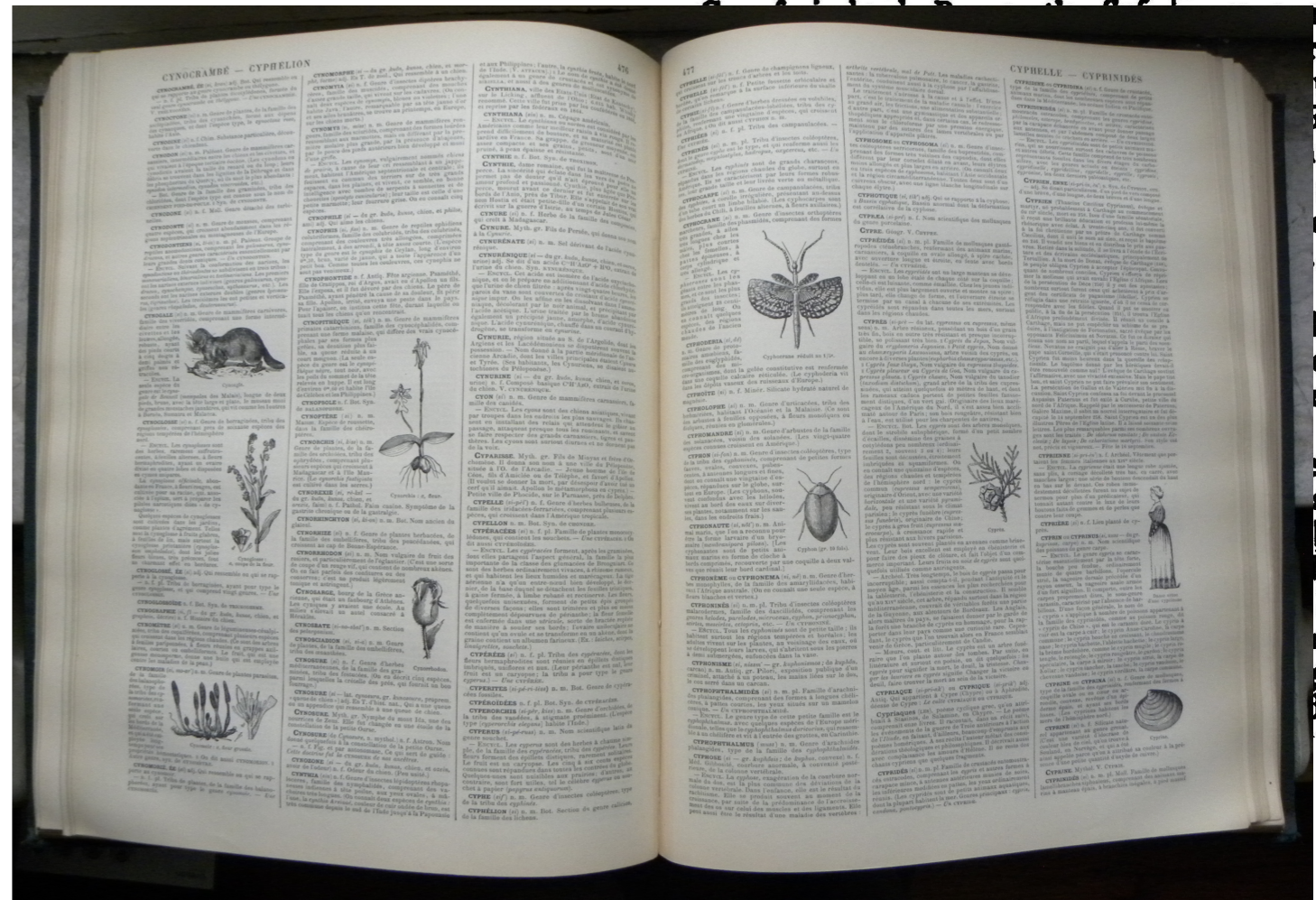
Cneoehte *Knotty;* geniculatus:—*L. Md.* 1, 64.

A Possible Lexicon

- Mental lexicon as dictionary?

- List all the words neatly in order

- Define them with a short description

- Find them in O(log *n*)

What's wrong with this?

Cnáwan, he cnæwð; p. cneow;
  pp. cnáwen. To KNOW; nos-
  cere, v. oncnáwan.
Cnáwincg *Knowledge, a knowing*;
  cognitio, *Som.*
Cnear *A ship, a narrow ship,
  galley*; navis:—Cnear on flot
  *the ship on float, Athelst. Vict.
  Chr.* 938. Nægled on cnear-
  rum *in nailed ships, Id.*
Cneatung, e; f. *A debate, an
  inquiry, a search*; disputatio,
  scrutinium:—*Scint.* 14.
Cnedan, gecnedan; pp. cneden
  [*Dut.* kneden] To KNEAD,
  ferment; subigere:—*R. Lk.*
  13, 21.
Cneo *a generation*, v. cneores.
Cneoehte *Knotty*; geniculatus:
  —*L. Md.* 1, 64.

attendant,
modern k
are so ca
serve the s
—Lytel cr
4, 8: *L. l*
Lk. 7, 7:
sende cni
Cnihtcild
  Bd. 4, 8.—
  learning a
Cnihthád *T*
  childhood
  boyhood, K
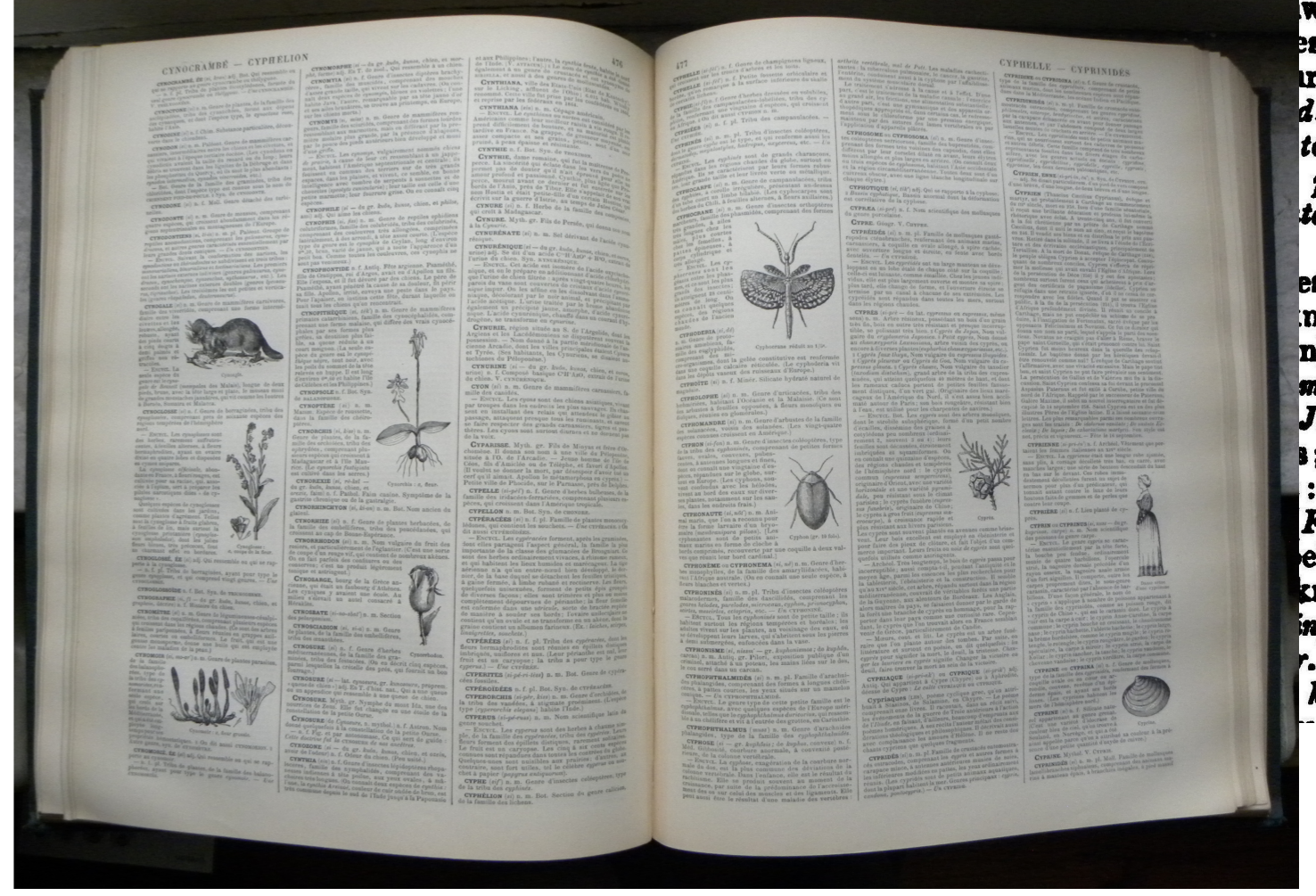 —*Elf. gr.*
Cniht-iuguð
  juventus :
Cnihtlic ; a
  —*L. Guth*
Cnihtwisa
  nwis
  res-l
  urb
  Bd. 8
  to l
  ; *PH*
  ute ;
  es ;
  knol
  en]
  mm
  Jud
  es ;
  s :—
  [*Fr*
  bel
  kno
  ing,
  r. 2
  l kn

A ~~Possible~~ Terrible Mental Lexicon

- Doesn't match human behaviour!

- We don't make alphabetic errors: "garland", "garlic", "garment" are not confused

- Meaning of a word is much broader & flexible than a definition

- People with larger vocabularies don't speak & process language more slowly

A better mental lexicon

Requirement for a good theory of a mental lexicon:

- Words are organised along the same principles as humans use

These principles are only visible via human behaviour, so

- Theory needs to account for human behaviour

A better mental lexicon

Requirement for a good theory of a mental lexicon:

- Accounts for and can predict human lexical processing behaviour

How do we get at human lexical processing behaviour?

- Experiments!

  Psycholinguistics: psychology of language

Lexicon is inaccessible to introspection
(like many/most cognitive processes & representations)

As quickly as possible, come up with a word that:

starts with the same sounds as:

" "

As quickly as possible, come up with a word that:

starts with:

"ca-"

As quickly as possible, come up with a word that:

starts with:

"un-"

As quickly as possible, come up with a word that:

starts with:

"gl-"

As quickly as possible, come up with a word that:

ends the same as/rhymes with:

"string"

As quickly as possible, come up with a word that:

ends the same as/rhymes with:

"strange"

As quickly as possible, come up with a word that:

has the same middle syllable as:

"com*pu*ter"

As quickly as possible, come up with a word that:

has the same middle syllable as:

"cog*ni*tive"

What did you notice?

Phonology is part of the lexicon

- *Reaction Time* as a measure of access (look-up) difficulty

- Phonological matching is (sometimes) easy: suggests mental lexicon is organised by sounds

- Certain regions of the word (beginning, end) are more salient/accessible than others (middle)

- Other factors: stress? morphology?

As quickly as possible, come up with a word that:

has a similar meaning to:

" "

As quickly as possible, come up with a word that:

has a similar meaning to:

"plant"

As quickly as possible, come up with a word that:

has a similar meaning to:

"quickly"

As quickly as possible, come up with a word that:

has a similar meaning to:

"unlikeable"

As quickly as possible, come up with a word that:

means the opposite as:

"hot"

As quickly as possible, come up with a word that:

means the opposite as:

"ocean"

As quickly as possible, come up with a word that:

means the opposite as:

"observer"

As quickly as possible, come up with a word that:

is associated with:

"swimming"

As quickly as possible, come up with a word that:

is associated with:

"maths"

As quickly as possible, come up with a word that:

is associated with:

"ties"

As quickly as possible, come up with a word that:

is associated with:

"outside"

As quickly as possible, come up with a word that:

has nothing to do with:

"university"

As quickly as possible, come up with a word that:

has nothing to do with:

"multiply"

As quickly as possible, come up with a word that:

has nothing to do with:

"kitchen"

Semantics is part of the mental lexicon

Words that have related meanings are easier to access together

- Synonyms, antonyms (opposites)

- Hyper/hyponyms ('is a' relation: *cat* is a hypernym of *lion*)

- Topically related words

- ...

Morphology is also part of the lexicon

- Inflectional variants
  chat/chatted, *swim/swam/swum,*
  word/words, lexicon/lexica

- Derivational relations
  swim/swimmer/swimsuit
  plant/plantation
  compress/compressible

- Morphology can be more transparent: *unknowable*
  or less transparent: *inept, disturb*

- Transparency has an effect on lexical access

What are the lexicon entries? Form + Meaning

Candidate form types:

- Morphemes: *lion, swim, lexicon, -sPL, -aPL, -izeV, -ableAdj*

- Lemmas: *lion-N, lionise-VERB, swim-VERB, lexicon-N*

- Wordforms: *swim, swum, swimmer, lionised, lexica*

- Phrases: *strong swimmer, you're welcome*

Lexicon entries: Form + Meaning

Meaning component has to include

- prototype effects (related to categorisation problems)
    "bird" is more than

- flexible and idiosyncratic links to many other concepts

- ability to use words creatively and metaphorically;
  to extend their meaning

Not a dictionary definition: a fixed description of necessary and sufficient conditions for correct word usage.

Lexicon entries: (Form + Meaning) + Links

Along with individual entries, lexicon also has to include

• a way of marking relations between words

and/or

• a way of measuring similarity between words

What kinds of relations exist between words?

What makes two words 'similar'?

As quickly as possible, come up with a word that:

can follow:

"blond"

As quickly as possible, come up with a word that:

can follow:

"raining cats and"

As quickly as possible, come up with a word that:

can follow:

"strong"

As quickly as possible, come up with a word that:

can follow:

"fire"

Contextual Relations

Words that appear together frequently

- idioms: fixed non-compositional phrases
  raining cats and dogs

- collocations: compositional but still somewhat fixed
  strong tea versus *powerful tea*

- syntagmatic relations: words that broadly co-occur
  (*match, game, goal*) or (*match, candle, light*)

Contextual relations link words across syntactic classes

Psycholinguistic methods: Implicit measures

Measuring reaction time:

• Lexical decision task: "Quick, is *glivk* an English word?"

• Priming tasks: After you hear word *cat* (are primed),
  are you faster or slower at reading/recognising *dog*? *hat*?

• Lexical completion task: Fill in a frame, such as *ca_, _rt*

Reaction time results

Similarity effects can manifest in both

- *facilitation*: faster reaction times

- *inhibition*: slower reaction times

due to either

- *spreading activation* from neighbours (facilitates)

- *competition* from many or high-frequency neighbours (inhibits)

Modality (spoken or written) has a strong effect

Reaction time results

Similar

- *facili*

- *inhib*

due to

- *inter*   tes)

- *com*   bits)

Modali

Effect of different kinds of neighbors in different tasks.

Neighbor/Task type	Behavior
Form Neighbors	
Visual word recognition	Facilitation
Visual word recognition with high frequency neighbor	Inhibition
Spoken word recognition	Inhibition
Spoken word production	Facilitation
Semantic neighbors	
Near neighbors	Inhibition
Distant neighbors	Facilitation

Chen & Mirman (2012)

Psycholinguistic methods: Explicit measures

- Similarity ratings on a Likert scale: How similar is *cat* to *dog*, on a scale from 1-5? *given* to *absolutely*?

- Word association: What words come to mind when you see *likeable? prophet? impending?*

https://smallworldofwords.org/

Findings/results so far

Psycholinguistic experiments show words are organised in the mental lexicon along multiple dimensions:

- phonological

- semantic

- morphological

- collocational

- orthography, (others)

But we don't yet have a way of modelling or replicating these representations
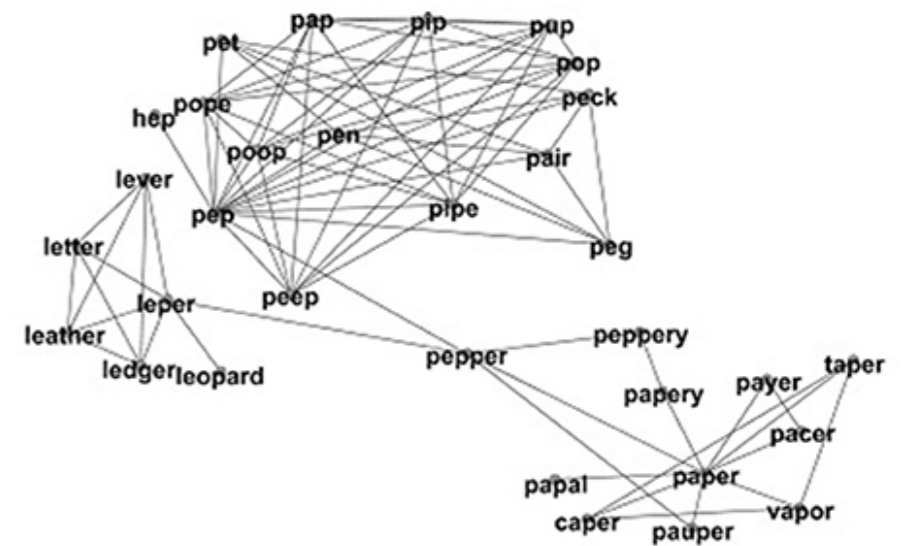
The Mental Lexicon as a network

- Connect related words, e.g phonological neighbours

- Distance: count hops between words

- Use techniques from Network Science
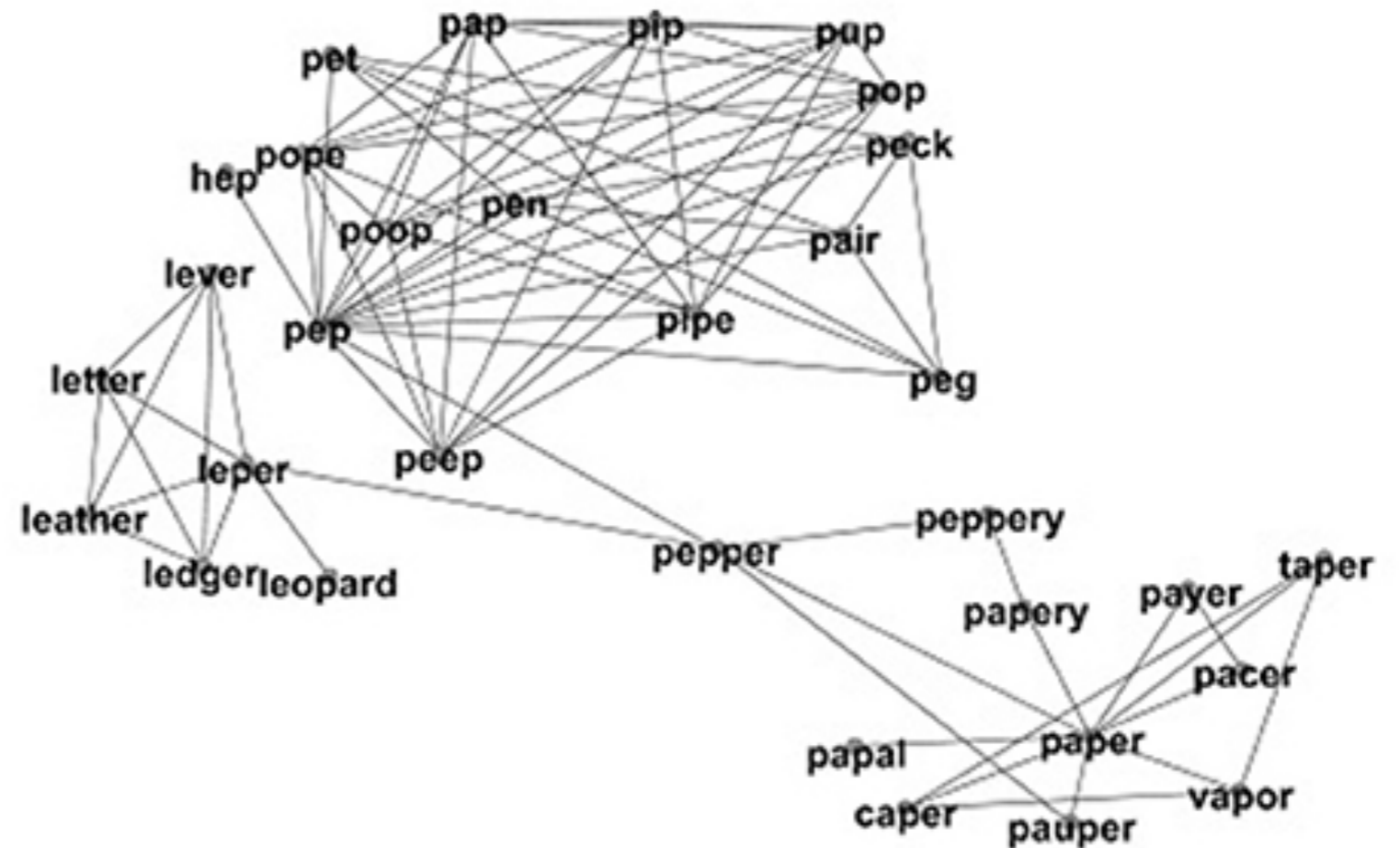
The Mental Lexicon as a graph/network

Phonological network:

- *Small word* structure: many small highly-connected clusters with sparse connections between clusters

- A word's community structure influences its processing: *neighbourhood density effects*
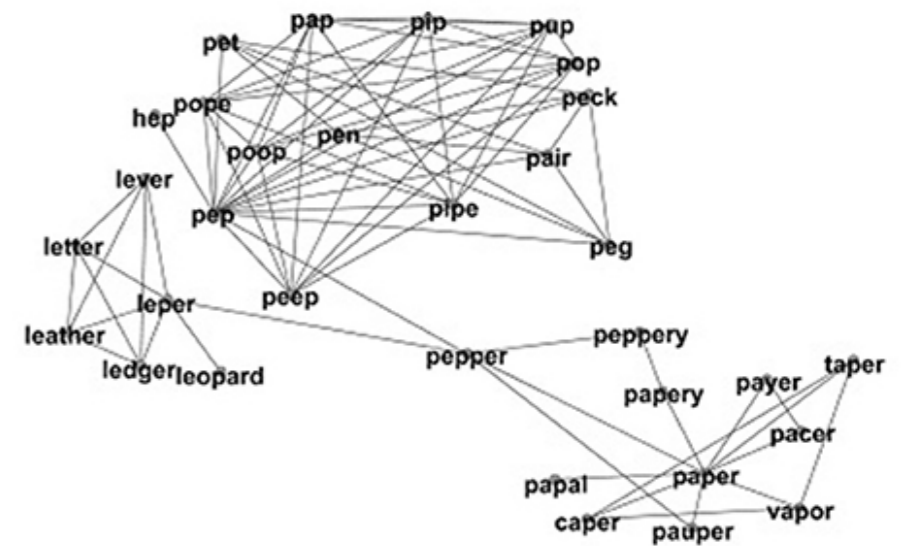

image from Goldstein and Vitevich, 2017

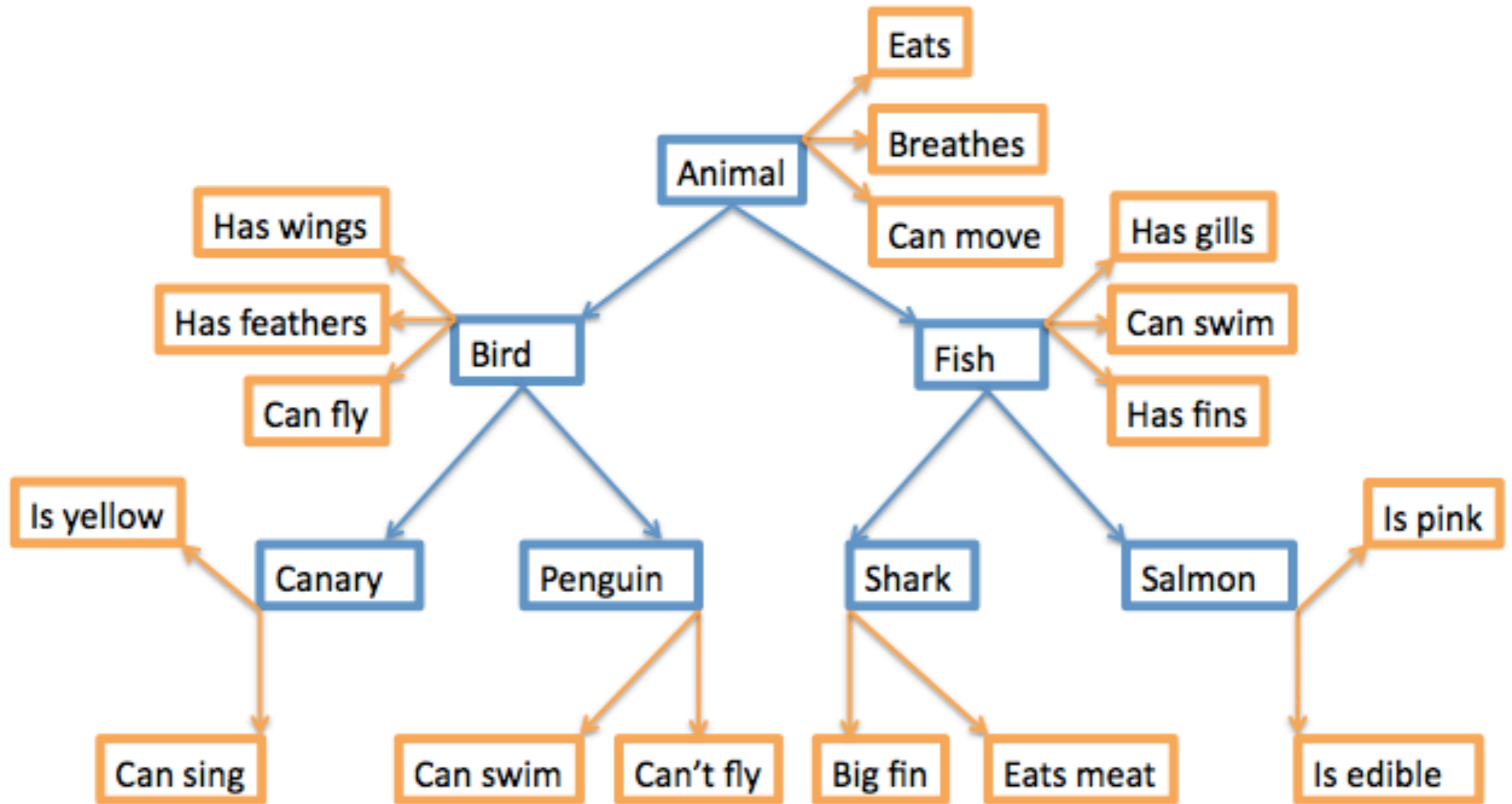# The Mental Lexicon as a graph/network

Advantages:

- Explicit structure makes model more interpretable

Issues:

- Connections are binary (on/off)

- Hard to integrate multiple dimensions: phonology + semantics

- Other spaces are harder to graph: semantically relations

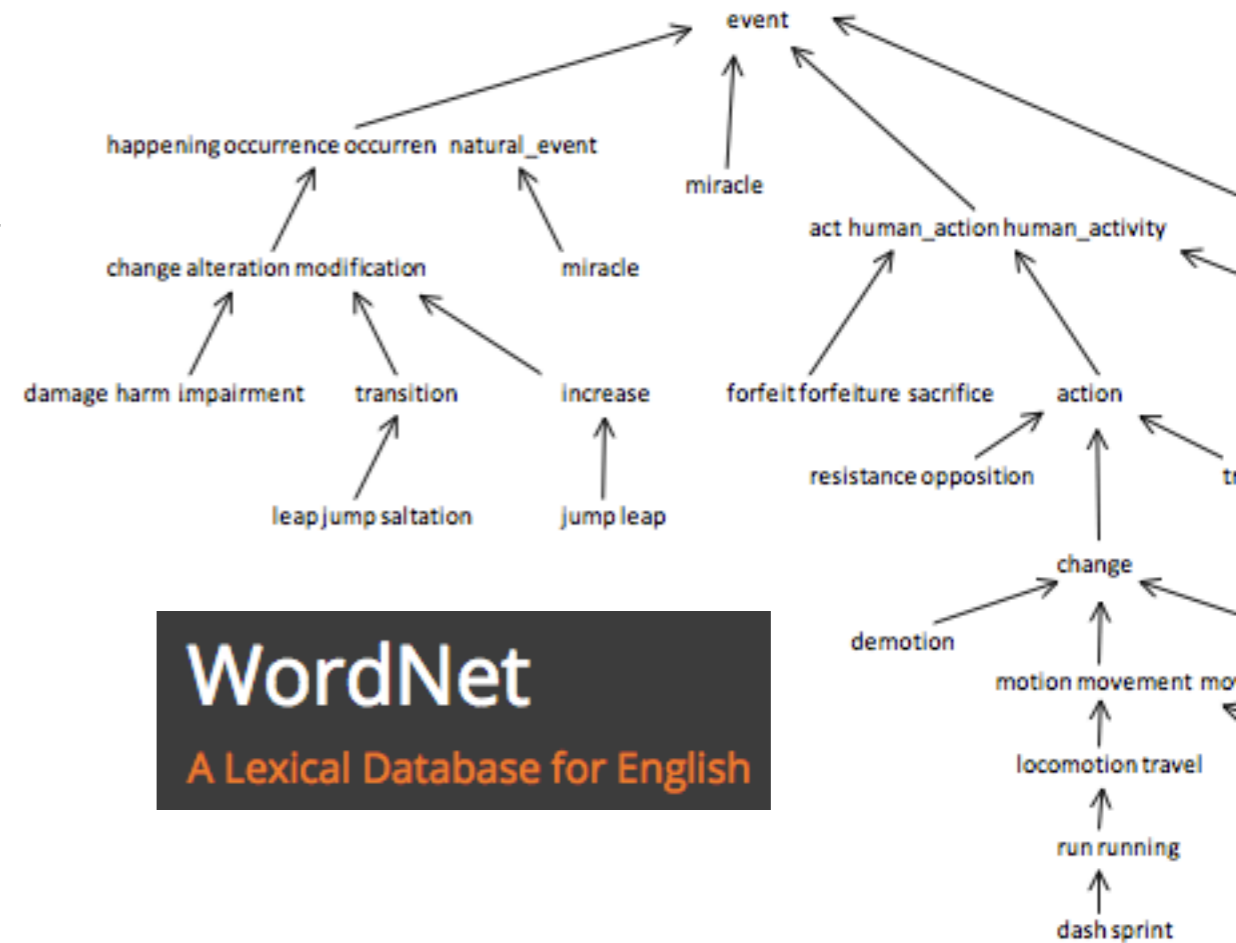image from Goldstein and Vitevich, 2017

Semantic Networks

Adapted from the Hierarchical Model of Collins and Quillian (1969)
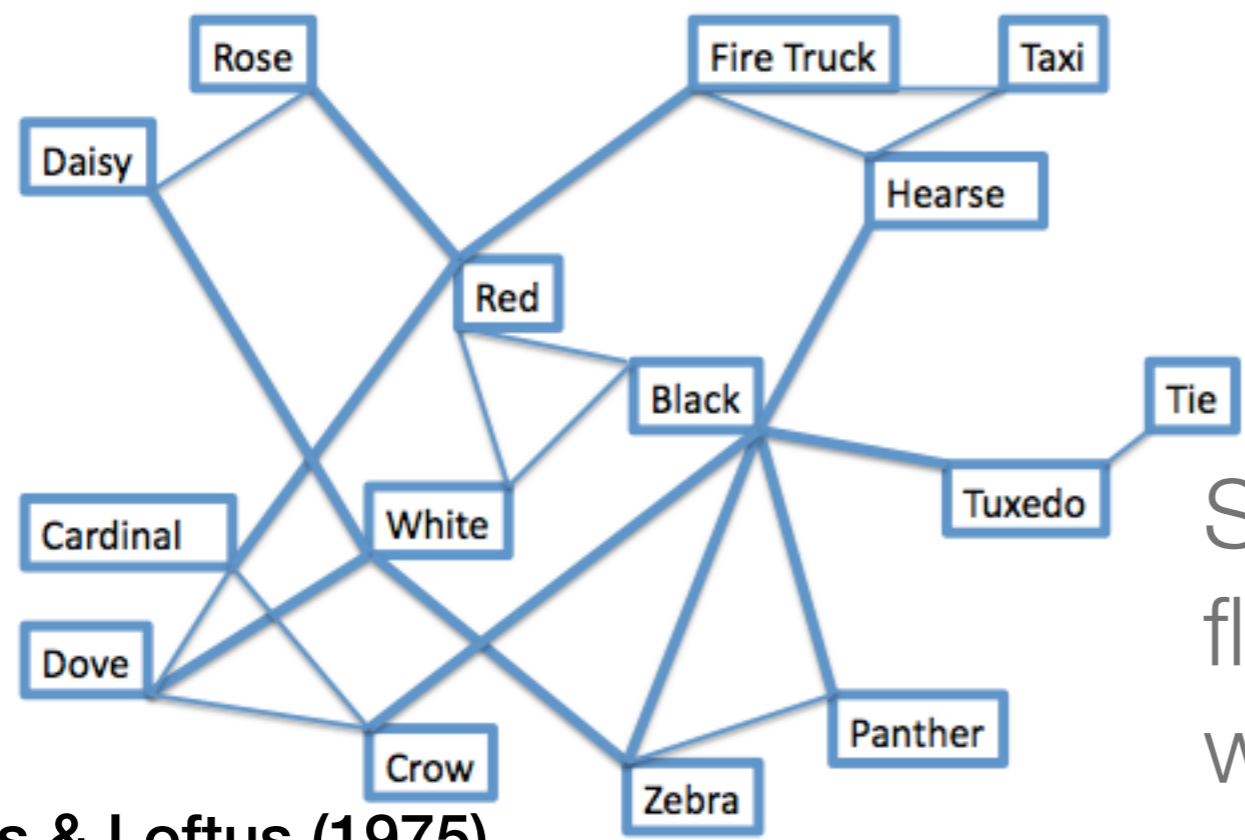By Nathanael Crawford - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=13268578

Semantic Networks

Hierarchical networks don't reflect human behaviour
 - link *dog* to *animal* faster than *dog* to *mammal*

Collins & Loftus (1975)

Spreading activation model: flat organisation with weaker/stronger links

Summary

- Humans have a *mental lexicon* that captures links between words on multiple dimensions: semantic, syntactic, phonological, etc

- Requires a multi-dimensional mental representation of words (and morphemes, lemmas, phrases)

- Psycholinguistics gives us evidence for what these representations must contain

- NLP gives us methods for estimating high-dimensional representations of words

Next:

- Are the NLP word representations (embeddings) useful approximations of cognitive representations?

- For example, can they be used to predict human behaviour on psycholinguistic tasks?
  Despite the fact that they are ungrounded - learned from text only, not interaction with the world?