# Models of speech production and perception

Eva Maria Luef

SS 2026

# Word associations


- Perceptual similarities
- <u>Semantic</u>
  - Psychological associations (--> *cat* and *dog*)
    - both are common pets, words remind of each another
  - Shared features
    - share at least one semantic feature (*four-legged, furry, animal*)
  - Taxonomic relationship
    - classification, e.g., subordination: *animal – dog*
  - Co-occurrences
    - *bread + butter, parking + lot*

- <u>Phonological</u>
  - Maximal phonological similarity, not more than one segment difference
    - phone – tone → foʊn - toʊn
    - cat – mat → kæt - mæt

# Cognitive net of words

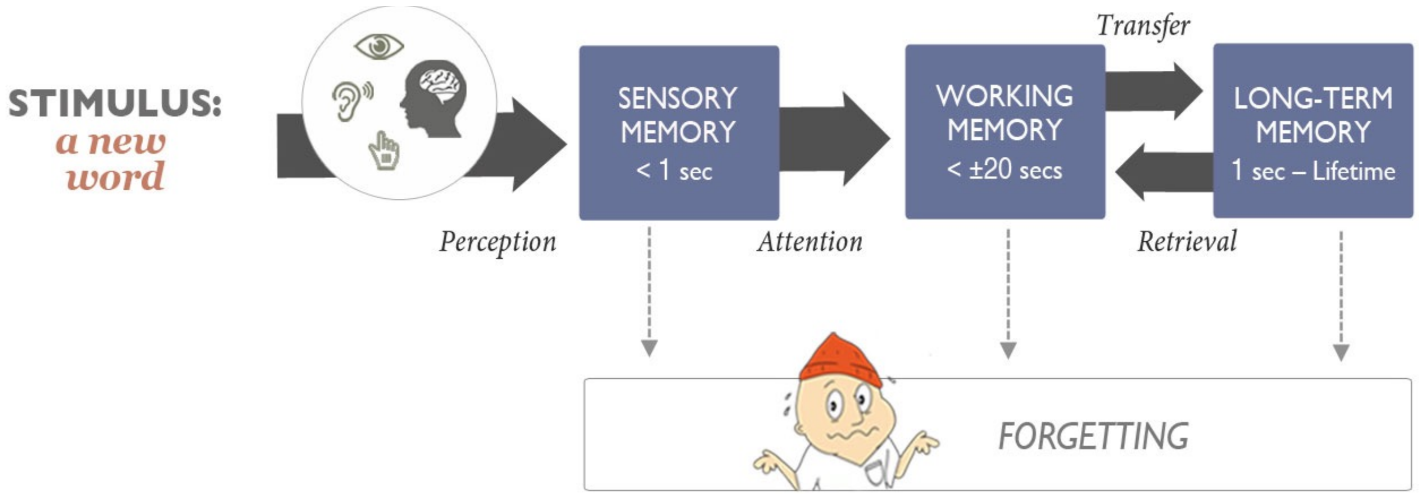
If concepts are interlinked, a network results

Aitchison (2012):
a large structure of word relations exist in the mental lexicon
the mental lexicon can be seen as a "gigantic multidimensional cobweb"

Organized according to perceptual similarity

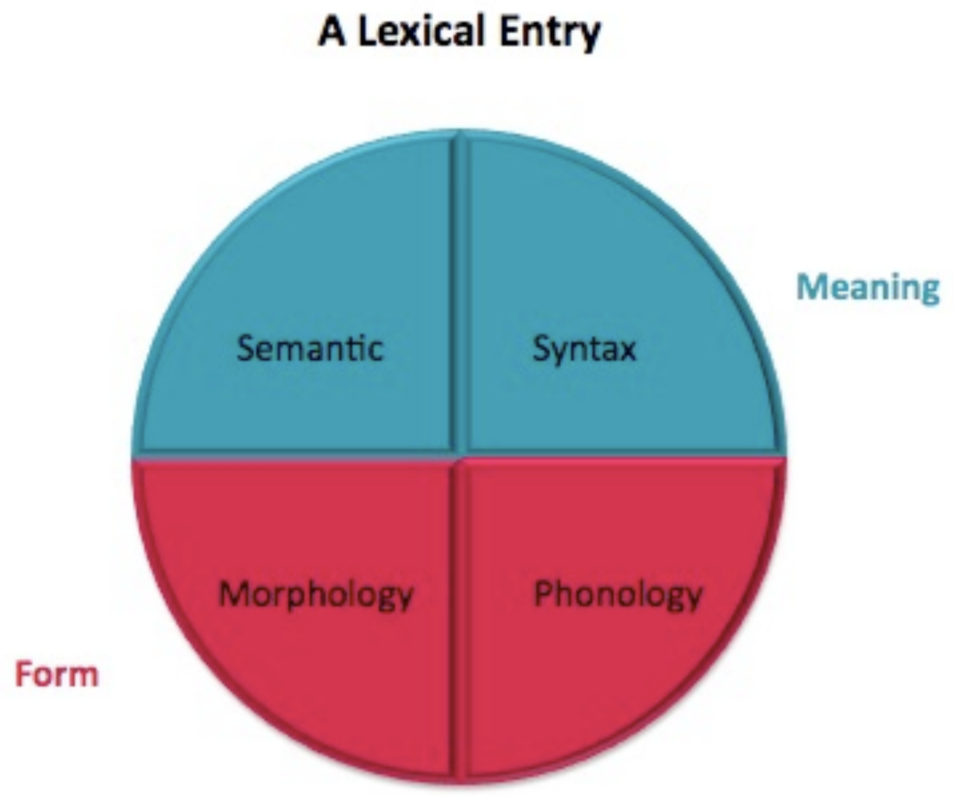
# Long-term memory

- Mental lexicon is part of the long term memory
- Mental lexicon is the part that stores lexical items
  - including information that goes with them (morphology, meaning)
- From functional perspective, long-term memory component of the mental lexicon is comprised of three levels:
  - concept, lemma, phonology



# Lexical access

- Lexical access occurs in speaking when the pronounceable word forms are found which encode the concepts the speaker wishes to communicate
- and it occurs in listening when the meaning is found which is expressed in the word forms the listener has heard
- Dimensions of words

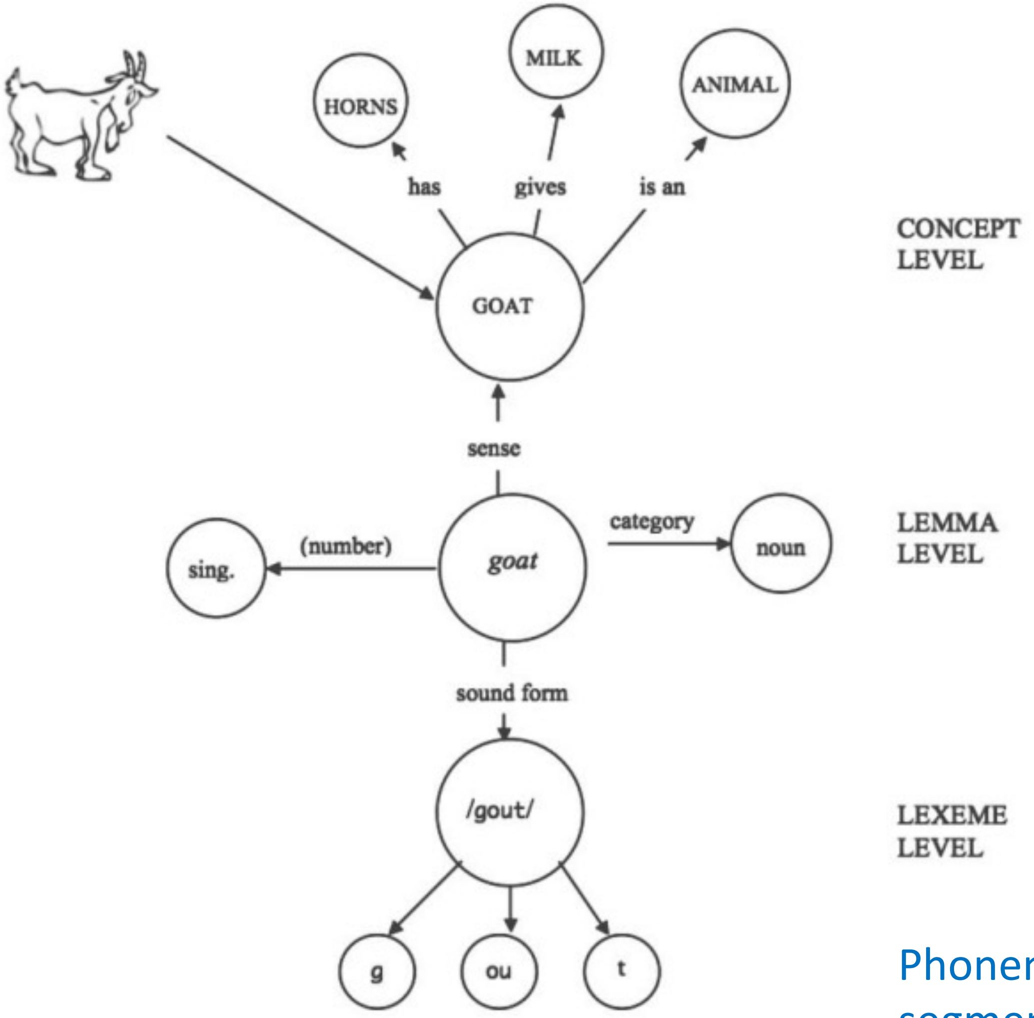



Levelt, 1989

# The existence of lexical entries

- Working definition of word: “mapping sound to meaning”
  - Arbitrary, conventional
  - [noʊz] ‘nose’
    - Presumably non-arbitrary: onomatopoeia “quack”,
      - French “coin”, Danish “rap”
    - Homophony [aɪ] ‘eye’ or ‘I’
- Mental representations of words can be defined and categorized in many ways (Brown et al., 2020)
- One approach to defining words relates to the different in their forms (McLean, 2018)
  - Word **types** consist of each unique word form
    - *record, records, unrecorded*
  - **Lemmas** are made up of a headword and its inflections
    - *record* = head, +infl.
      - Part of speech: noun(*record, records, recording*), verb(*record, recording, recorded*)
  - **Flemmas** are larger classifications of lemmas
    - one headword “record” + all inflections and derivations (*pre-record, recorder*…)

# Levels of representation

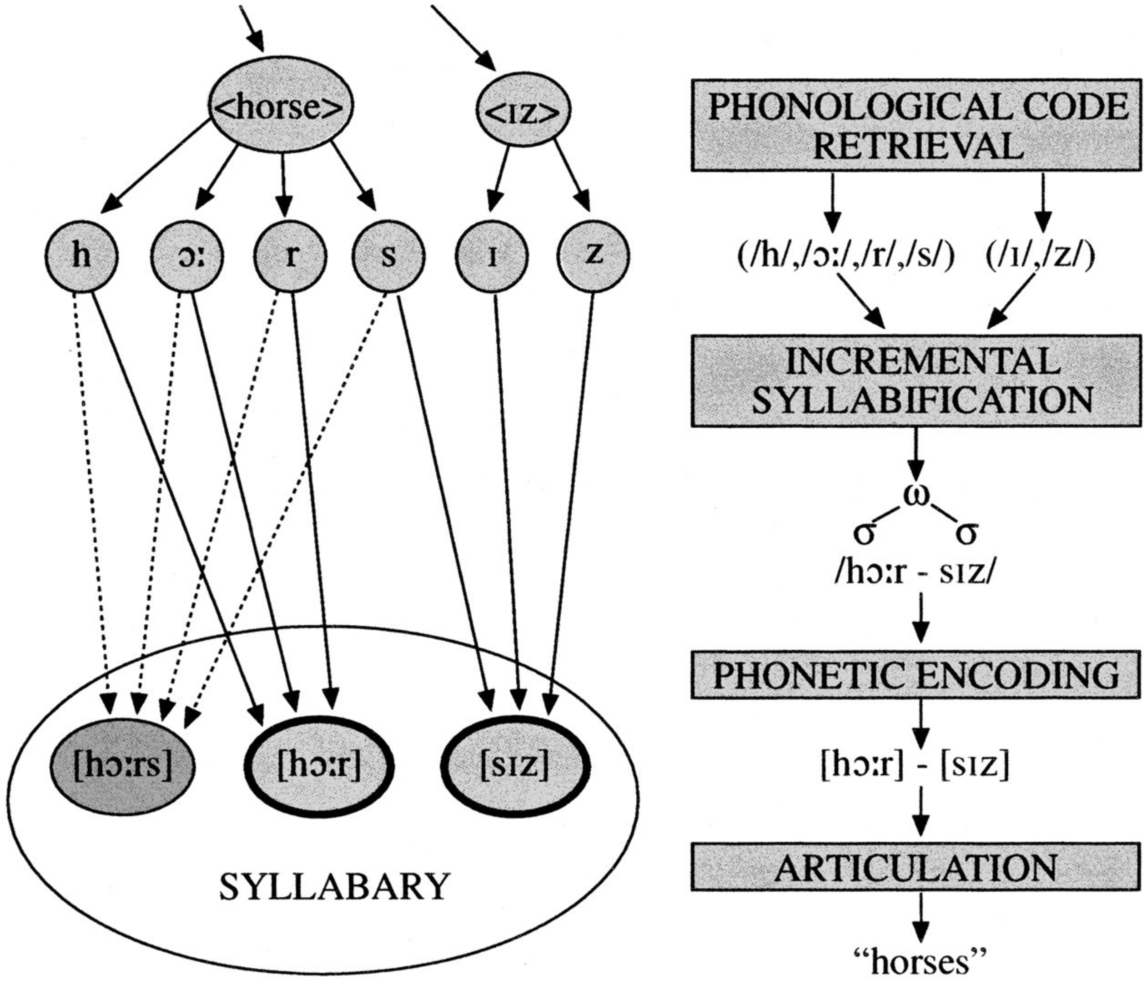



Phonemes/
segments

Bock & Levelt, 1994

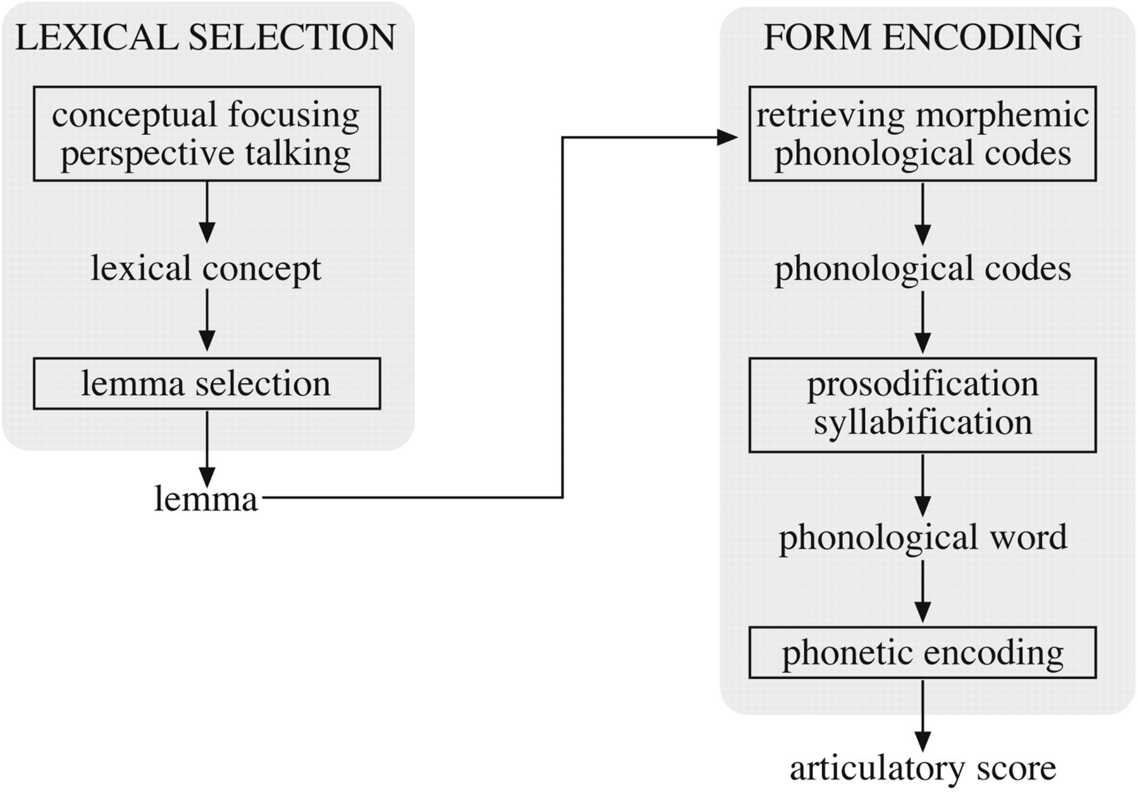
# Segment level

- Phonological encoding
  - assembling words from lexical entries
  - prepare articulation in speech production
  - constructing the phonological form of a target token before articulatory gestures can be prepared in spoken word production (Dell, 1986; Levelt et al., 1999; Caramazza et al., 2001)



# Phonological pre-activation

- phonological pre-activation or phonological preparation (Li et al., 2017) is initiated at the start of the target word
  - Dutch word generation: faster responses in a homogeneous context (e.g. ***h**ut*, ***h**eks*, ***h**iel*—hut, witch, heel) relative to a heterogeneous context
- phonological segments of target words are sequentially activated (Oppermann et al., 2010)
  - segment = level of preparation in Germanic languages
  - onset segment(s) attributed a crucial role in the process (Meyer, 1991; O'Séaghdha et al., 2010)
    - implicit phonological prime
    - loner, local, lotus → preparation faster and more accurate
      - as opposed to "loner, beacon, major"



# ‘Proximate units’

- not all languages reap word-activation benefits from the preparation of the **onset** phoneme
  - in Japanese phonological encoding begins at the level of the mora (Kureta et al., 2006)
  - in Mandarin Chinese at the syllable (Chen et al., 2016)
- concept of ‘proximate unit’ (O’Séaghdha, 2015)
  - = preparation unit
  - different types of phonological units that serve as activating mechanisms for encoding
  - defined as the “first selectable phonological unit below the level of the word or morpheme” (O’Séaghdha et al., 2010: p. 286)
  - in Germanic languages = segment

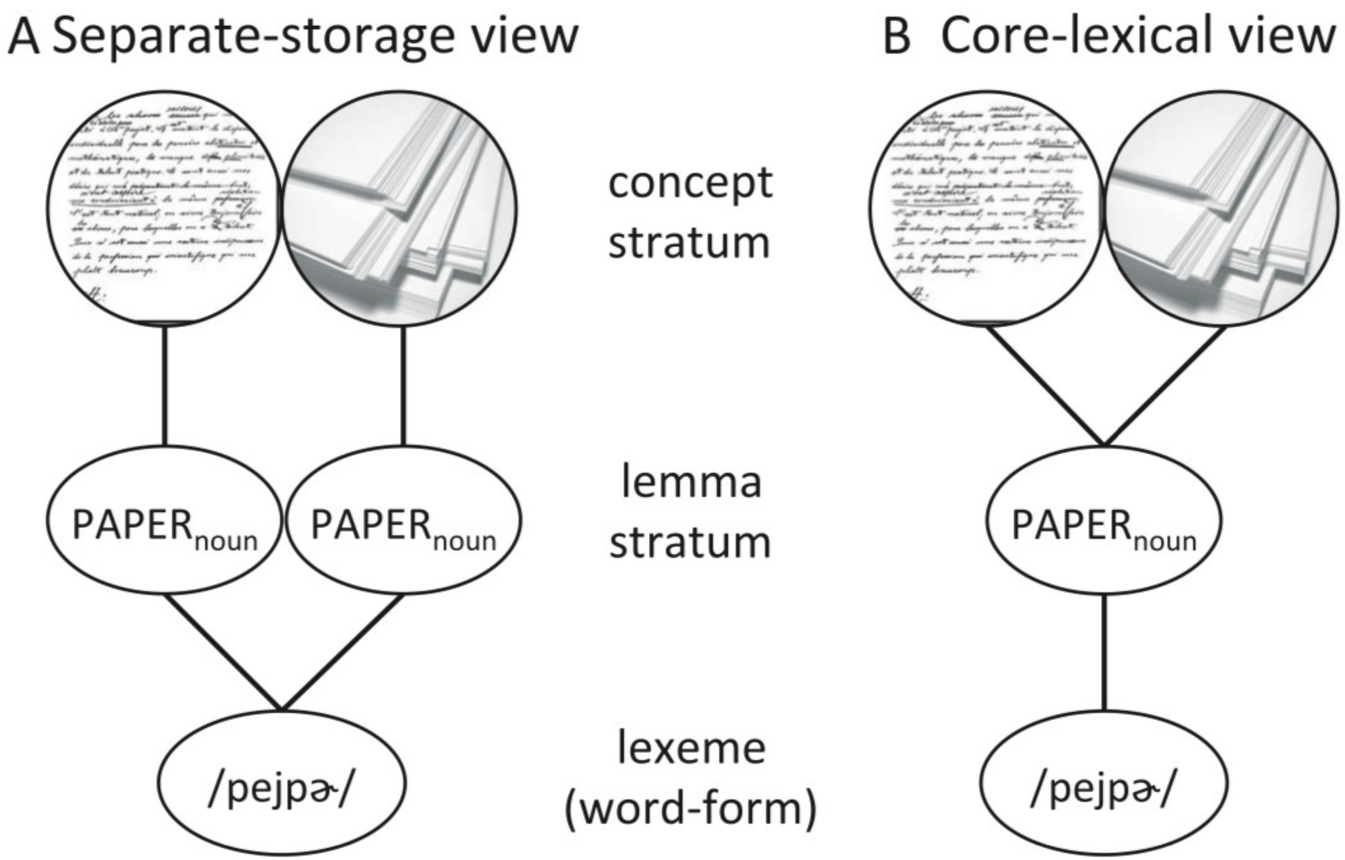
- proximate units are influenced by frequency effects (Stemberger and Macwhitney, 1986; Levelt and Wheeldon, 1994)
  - high-frequency onsets and syllables generally facilitating phonological construction of a target word (Aichert and Ziegler, 2004; Luef and Resnik, 2023; Macizo and Van Petten, 2007)

# Special cases of words: Homophony, polysemy

- ubiquitous across languages (Antilla 1989)
- how they are created is an important question in psycholinguistics (Wedel, Kaplan, and Jackson 2013)
  - Can arise from sound change (/k/ dropped in „knight“ – homophonous to „night“; see Lutz, 1988)
  - Polysemy (*mouse*: rodent, computer device)
    - Continuum of semantic relationship
    - Zero conversion: to water sth – water, sheep - sheep
- two lemmata linked to one phonological and phonetic form
  - bank ($), bank (river)
  - cut (V), cut (N)

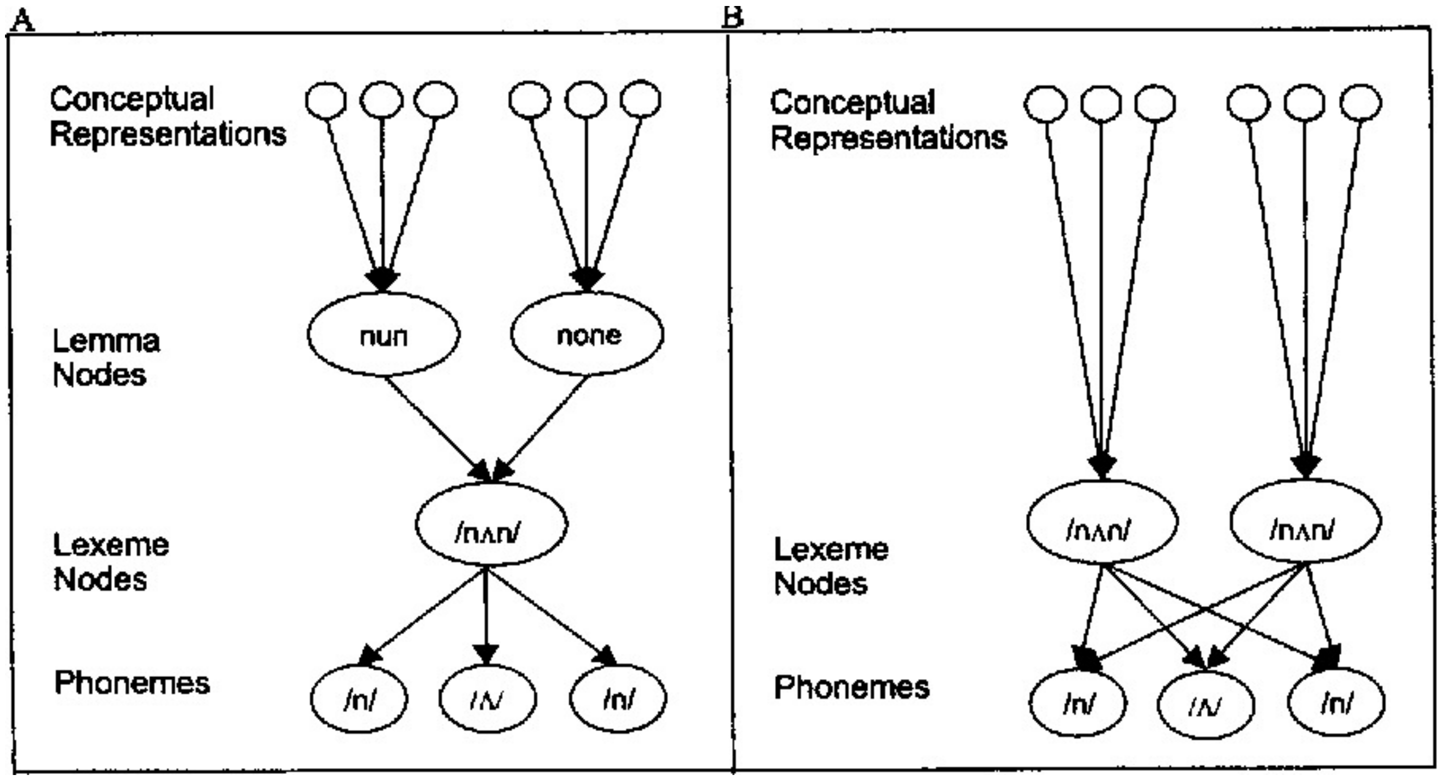
# Polysemy/ homophones

What if the two lemmas showed different word forms?

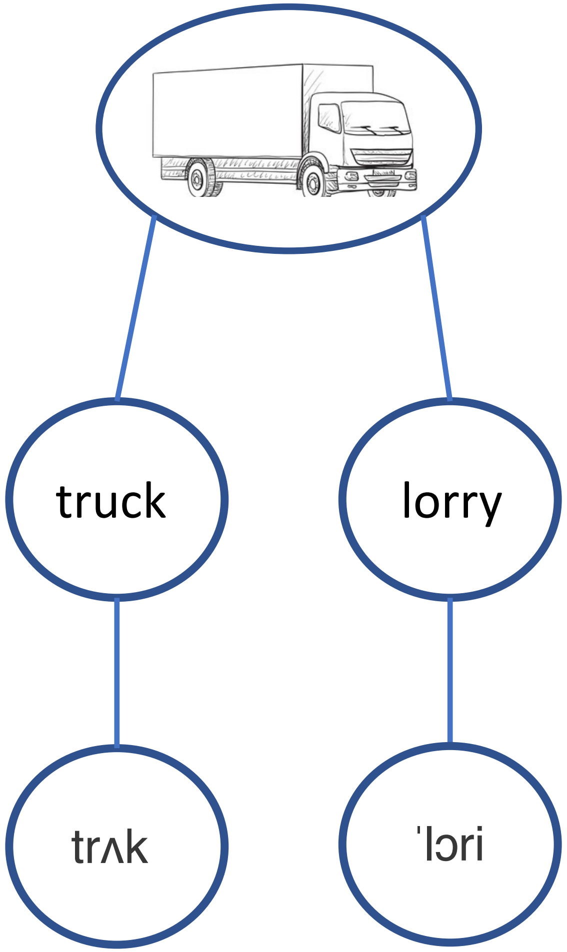


# Frequency effects

- Frequency effects reduce a word phonetically
  - e.g., *time – thyme* (Gahl, 2008; Lohmann, 2017)
    - HF “time” is shorter
  - HF “none”
  - LF “nun”
    - Phonetic realizations are not identical
    - Lemma “frequency inheritance” → phonetic realizations should be identical

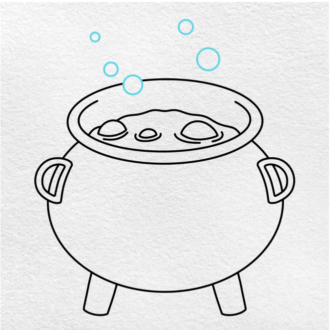


# Special cases of words: (Full) synonyms

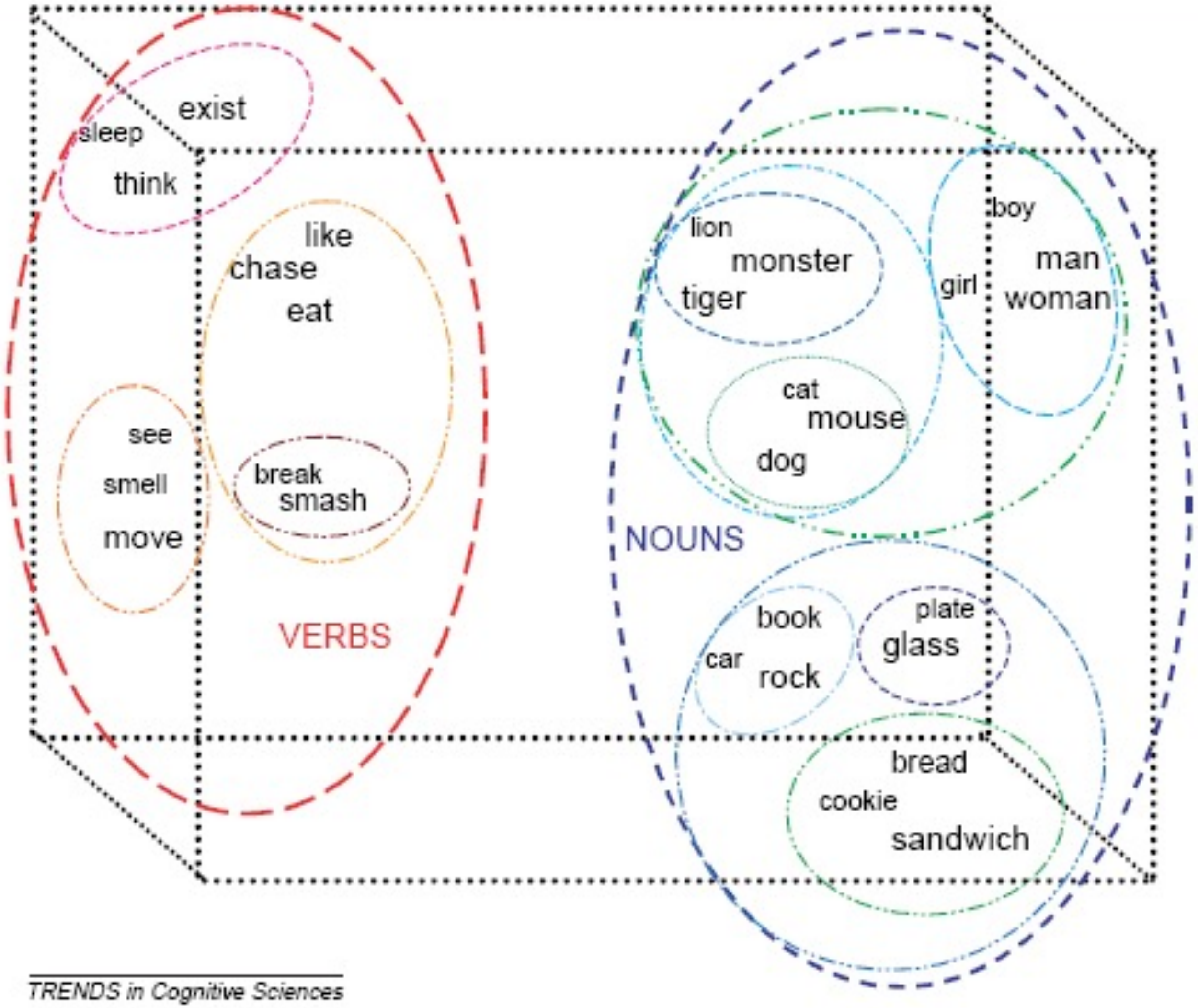



One concept
Two lemmata
Two lexemes

# No “cauldron of lexical soup”

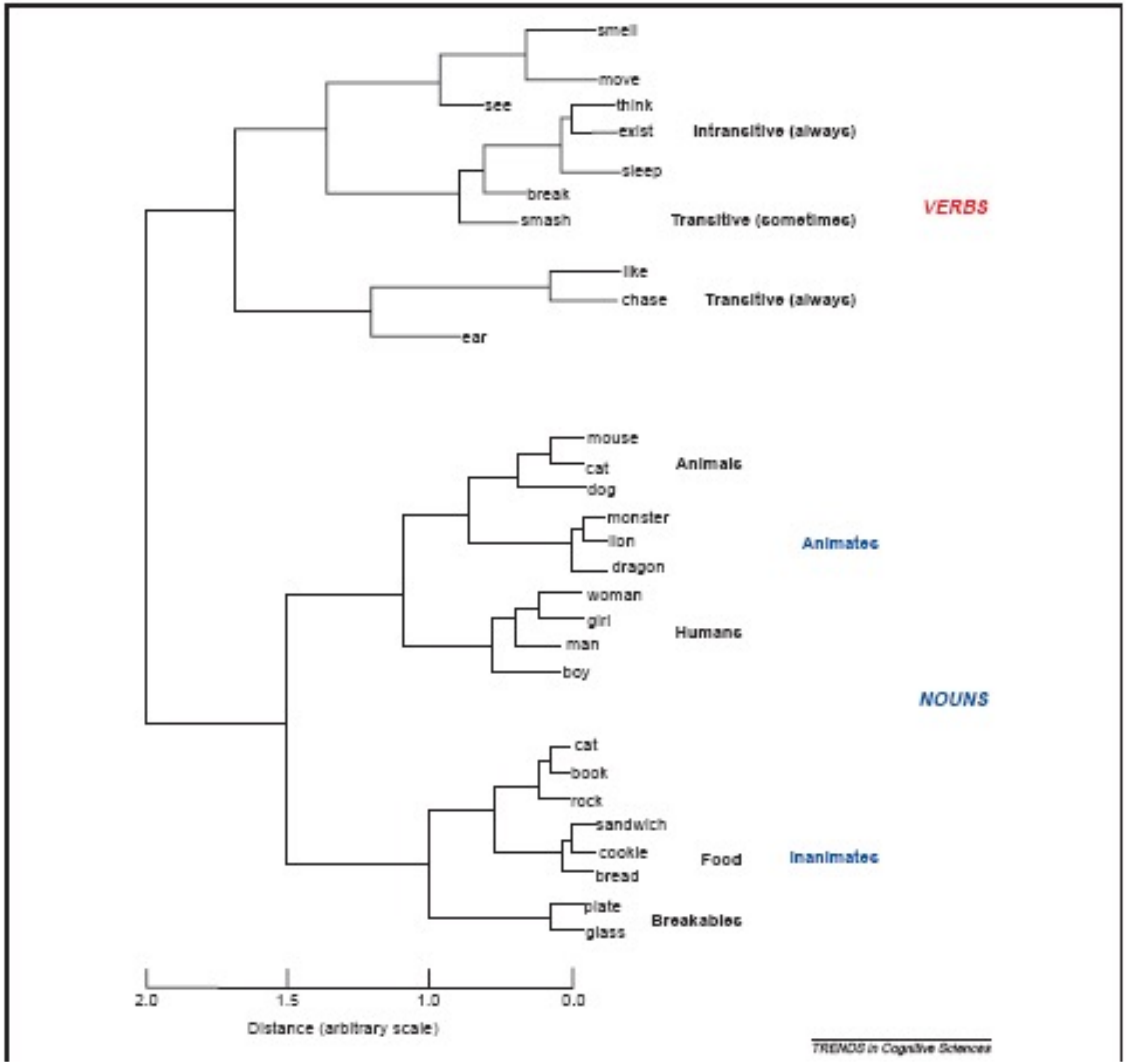



- Word memory is organized
- Words are linked to one another according to certain principles
  - Phonological similarity (*nun - none*)
  - Semantic similarity (*sheep – goat*)
- Interconnections exist
  - Influence priming


exist
sleep
think
like
chase
eat
see
smell
move
break
smash
VERBS
lion
monster
tiger
boy
girl
man
woman
cat
mouse
dog
NOUNS
book
car
rock
plate
glass
bread
cookie
sandwich
TRENDS in Cognitive Sciences

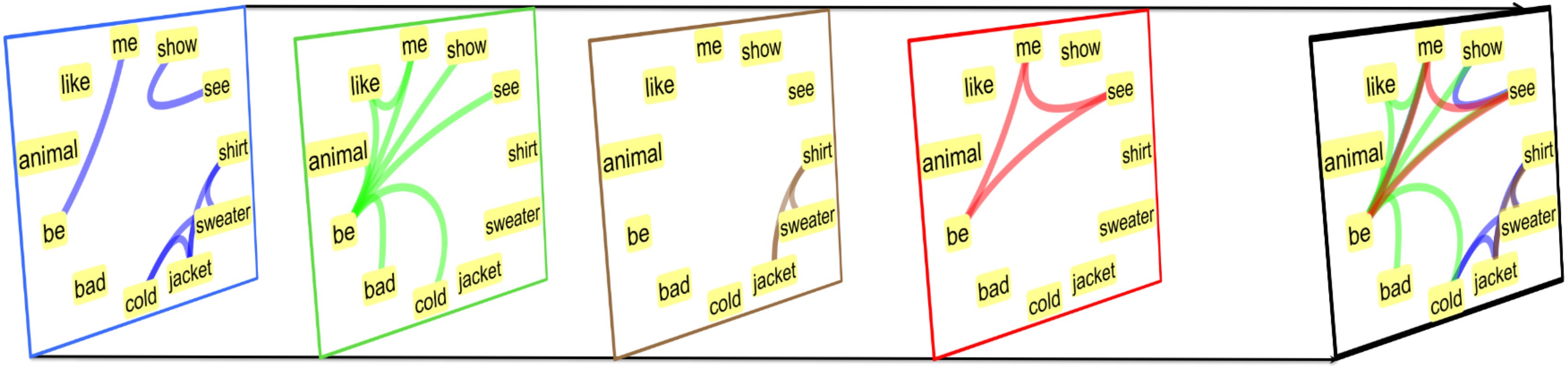
Elman 2004

# Hierarchical trees




Elman 2004

# Organizational principles




**Free Associations**

*Words reminds of each other*

**Co-occurrences**

*Words co-occurr in child directed speech*

**Feature Sharing**

*Words share at least one semantic feature*

**SEMANTIC LAYERS**

*(Semantic memory)*

**Phonological Similarities**

*Words differ in one phoneme*

**PHONOLOGICAL LAYER**

*(Phonology)*

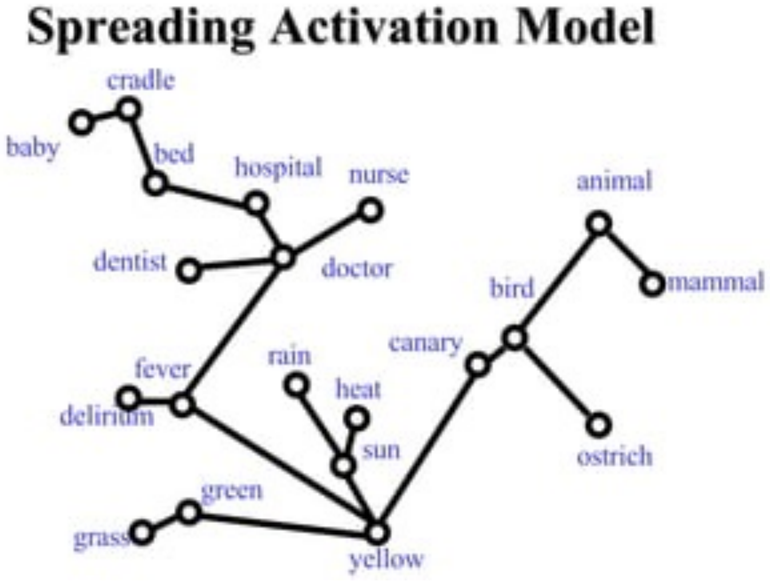
**MULTIPLEX LEXICAL NETWORK**

*(Semantics and phonology of the mental lexicon)*

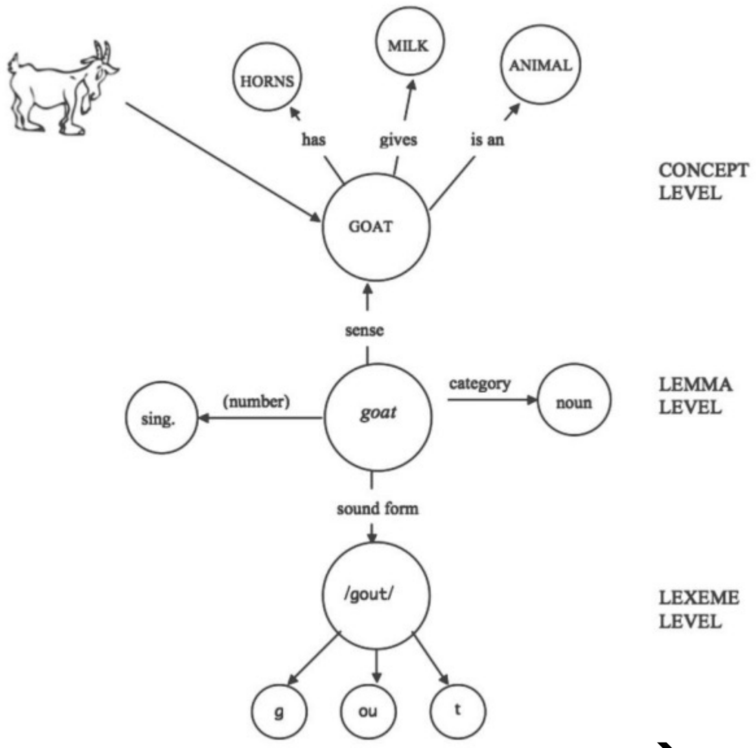
Stella et al., 2018

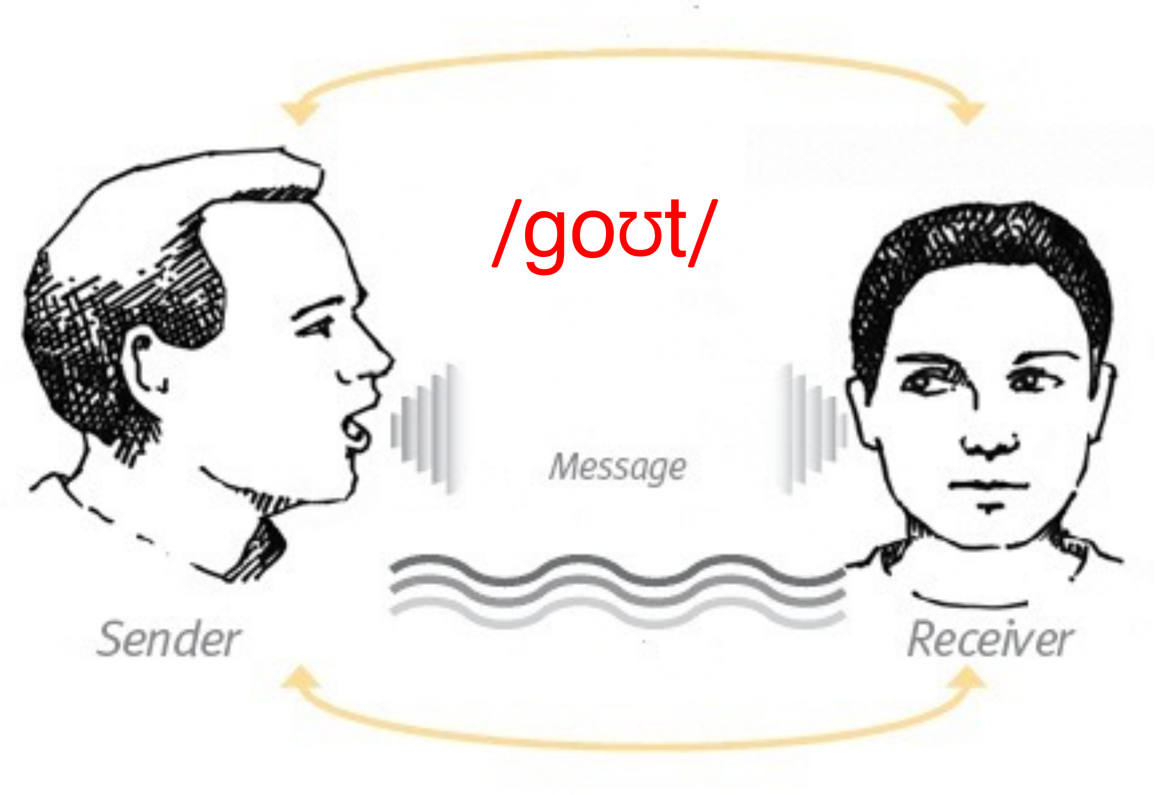
# Activate word representations

- Not in isolation
- With ‘neighbors’
  - On various relationship dimensions (e.g., semantic…)
- On gradual scale
  - Related words → strong co-activation
  - Loosely related words → weak co-activation
  - Not related → no co-activation



# Production vs. perception



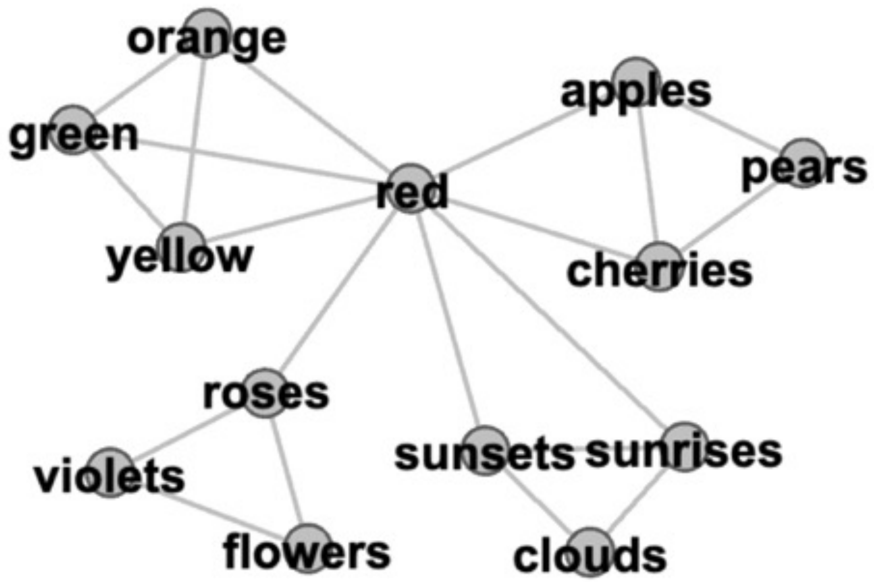




→ Models of spoken word recognition rely on the notion of "phonological neighbor"

# Similarity relationships in lexicon

- Items in the mental lexicon are related in meaning, use, and form
  - What we consider ‘related’ or ‘similar’

- Meaning:
  - Semantic similarity, relatedness
- Use:
  - Collocations
- Form:
  - Phonological word form – sound similarity



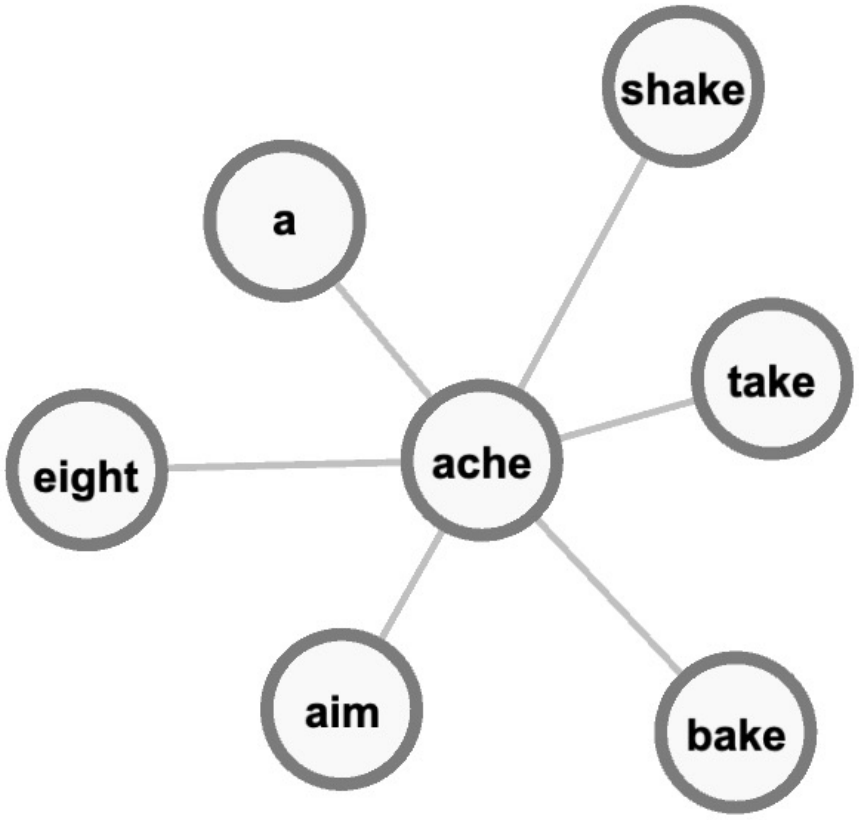
# Phonological neighborhoods

- Similarity bias in the phonological domain is governed by *phonological neighbors*
  - well-studied notion of lexical relationships in psycholinguistics (Goldrick, Folk, & Rapp, 2010; Landauer & Streeter, 1973)
  - string similarity → distance of one piece of information (a phoneme or grapheme) between two words
    - coast – ghost
    - bat – hat

- Current psycholinguistics: Levenshtein distance → one-segment distance

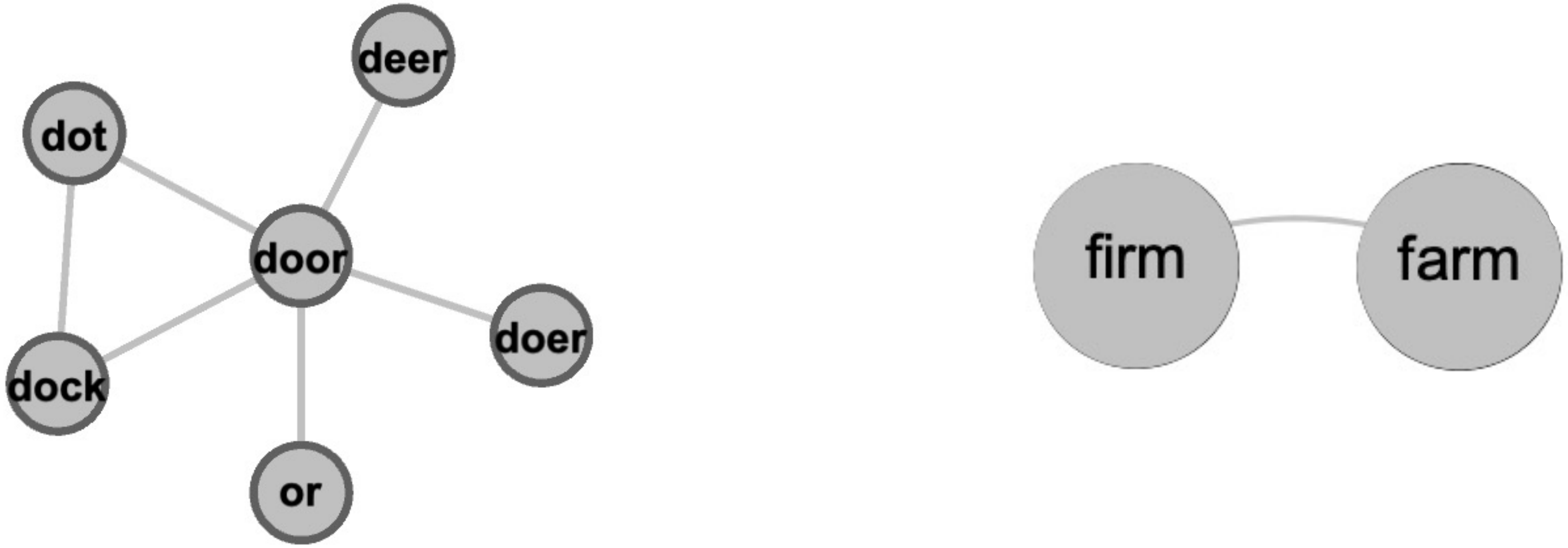
| Substitution | Deletion | Addition |
|---|---|---|
| Cat – hat - sat | Cat - at | Cat - catty |
| *Rhyme neighbors* | | *Onset neighbors* |

# Phonologial neighborhood




k=6

# Sparse and dense neighborhoods




Density is relative

In English, dense neighborhoods >50 members (e.g., cats)

# References


- Arutiunian, V., & Lopukhina, A. (2020). The effects of phonological neighborhood density in childhood word production and recognition in Russian are opposite to English. *Journal of Child Language, 47*(6), 1244-1262.
- Chan, K. Y., & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1934-1949.
- Chen, Q., & Mirman, D. (2014). Interaction between phonological and semantic representations: Time matters. *Cognitive Science, 39*(3), 538-558. doi: https://doi.org/10.1111/cogs.12156
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychologial Review, 93*(3), 283-321.
- Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory & Language, 56*, 490-520.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. G., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychologial Review, 104*(4), 801-838.
- Frisch, S. A. (2011). Frequency effects. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology: Phonological interfaces* (pp. 2137-2163). Chichester, UK: Blackwell Publishing Ltd.
- Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language, 66*(4), 789-806.
- Goh, W. D., Suarez, L., Yap, M. J., & Tan, S. H. (2009). Distributional analyses in auditory lexical decision: Neighborhood density and word frequency effects. *Psychonomic Bulletin & Review, 16*, 882-887.
- Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory & Language, 28*(5), 501-518.
- Goldrick, M., Folk, J. R., & Rapp, B. (2010). Mrs. Malaprop's neighborhood: Using word errors to reveal neighborhood structure. *Journal of Memory and Language, 62*(2), 113-134.
- Gordon, J. K. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language, 82*(2), 113-145.
- Jescheniak, J. D., & Schriefers, H. (1998). Serial discrete versus cascaded processing in lexical access in speech production: Further evidence from the co-activation of near-synonyms. *Journal of Experimental Psychology: Language, Memory, and Cognition, 24*, 1256-1274.
- Kapatsinski, V. (2006). Sound similarity relations in the mental lexicon: Modeling the lexicon as a complex network. *Speech Research Lab Progress Report, 27*, 133-152.

# References


- Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior, 12*, 119-131.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*, 1-36.
- Luce, P. A., Goldinger, S., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics, 62*(3), 615-625.
- Marian, V., & Blumenfeld, H. K. (2006). Phonological neighborhood density guides: Lexical access in native and non-native language production. *Journal of Social and Ecological Boundaries, 2*(1), 3-25.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71-102.
- Marslen-Wilson, W. D. (1990). Activation, competition and frequency in lexical access. In G. T. M. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge, M. A.: MIT Press.
- Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access. *Psychological Review, 101*, 653-675.
- Nooteboom, S. G. (2005). Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech communication, 47*, 43-58.
- Scarborough, R. (2013). Neighborhood-conditioned patterns in phonetic detail: Relating coarticulation and hyperarticulation. *Journal of Phonetics, 41*(6), 491-508.
- Schriefers, H., & Vigliocco, G. (2015). Psychology of speech production. In J. D. Wright (Ed.), *International Encyclopedia of the Social and Behavioral Sciences* (pp. 225-258). Amsterdam et al.: Elsevier.
- Siew, C. S. Q., & Vitevitch, M. S. (2016). Spoken word recognition and serial recall of words from components in the phonological network. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(3), 394-410.
- Vitevitch, M. S. (2002c). Naturalistic and experimental analyses of word frequency and neighborhood density effects in slips of the ear. *Language and Speech, 45*, 407-434.
- Vitevitch, M. S., & Castro, N. (2015). Using network science in the language sciences and clinic. *International Journal of Speech-Language Pathology, 17*, 13-25.
- Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics, 2*, 75-94.
- Vitevitch, M. S., & Rodriguez, E. (2004). Neighborhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders, 3*(1), 64-73.