Network science methods

A potential toolkit for cognitive linguistics?

Cognitive Linguistics Research Group

Edinburgh, 14 May 2021

Tobias Ungerer, University of Edinburgh
t.ungerer@ed.ac.uk

Outline of the talk

1 Basic concepts in network science

2 Applications to linguistic networks

3 Cognitive-linguistic applications: *one* example

4 Open questions

Basic concepts

What is ‘network science’?

Definition

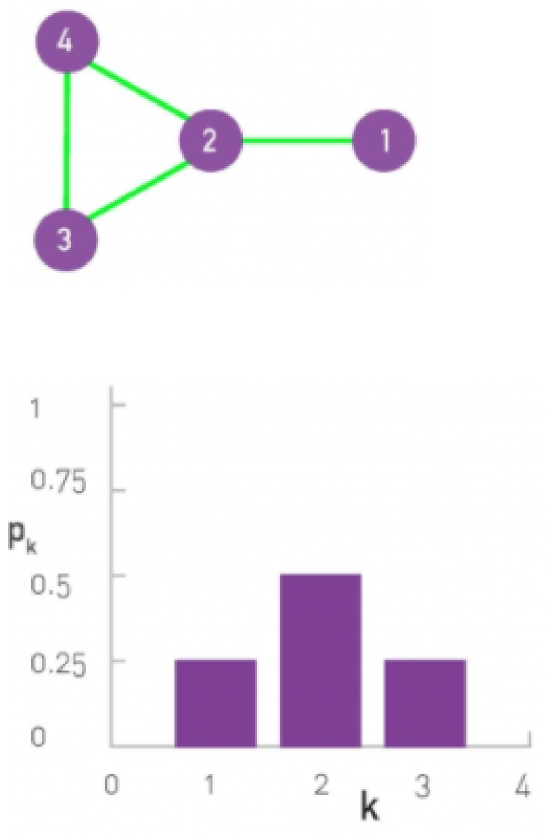
- The formal study of network structures across domains, using a common set of mathematical tools (based on Barabási 2016)
- Foundational papers: Watts & Strogatz (1998) on ‘small world’ networks; Barabási & Albert (1999) on ‘scale-free’ networks

Some major applications

- Sociology (e.g. networks of friendship or kinship ties)
- Technology (e.g. the internet)
- Economics (e.g. trade relations)
- Biology and health (e.g. genetic structures, virus spread)
- Neuroscience and artificial intelligence (neural networks)

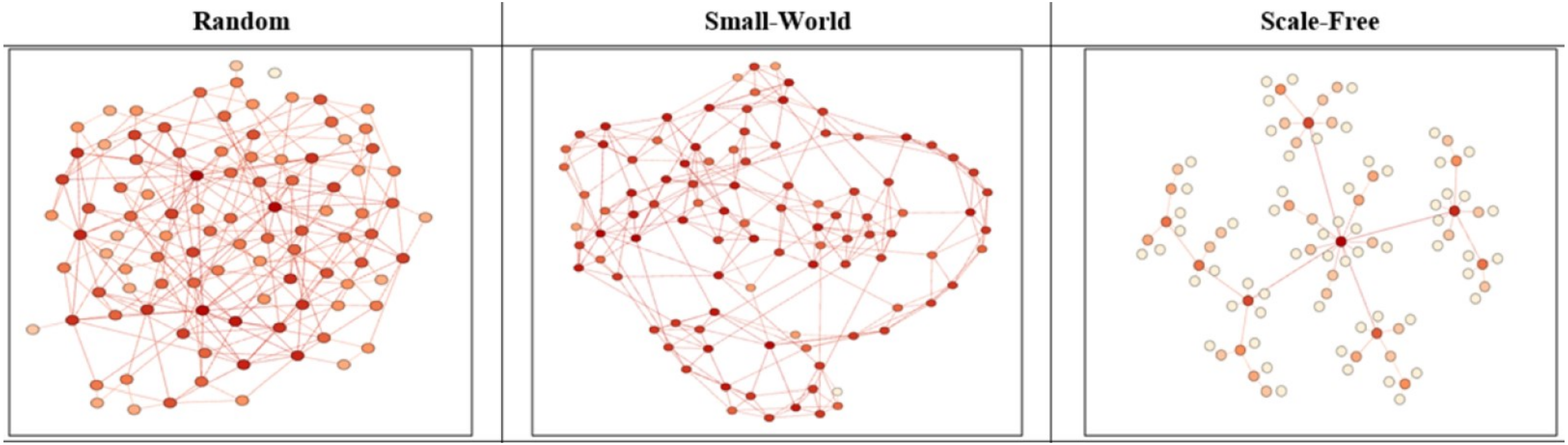
Basic concepts and network measures

- A network (= graph) consists of nodes (= vertices) and links (= edges)
- Types of networks: directed vs. undirected; weighted vs. unweighted; …
- Network measures (local and global):
  - Degree *k* of a node (= how many links a node has)
  - Degree distribution *P(k)* (= proportion of nodes with degree *k*)
  - Average shortest path length *L* (= how many steps from one node to another)
  - Clustering coefficient *C* (= probability that two neighbours of a node are linked to each other)
  - Modularity Q (= how many links are within network communities vs. between communities)



(Barabási 2016: Image 2.3)

Three topological structures



Short average path length

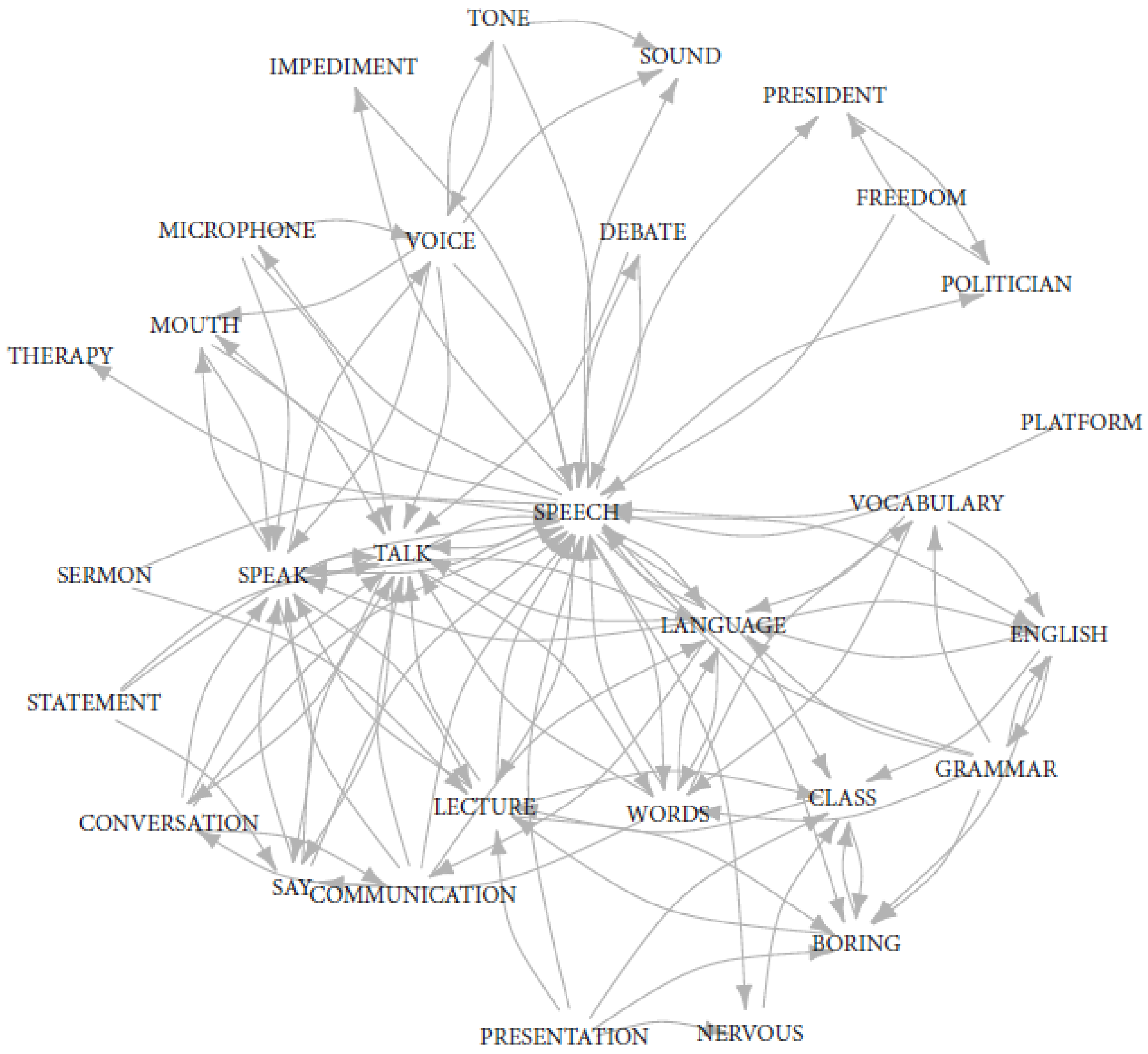
Short average path length, large clustering coefficient

Some nodes with very high degree (‘hubs’); many nodes with small degree (the degree distribution follows a ‘power law’)

(Perera et al. 2017: Fig. 1)

Applications to linguistic networks

Lexical/semantic networks



Structure and growth

(Steyvers & Tenenbaum 2005)

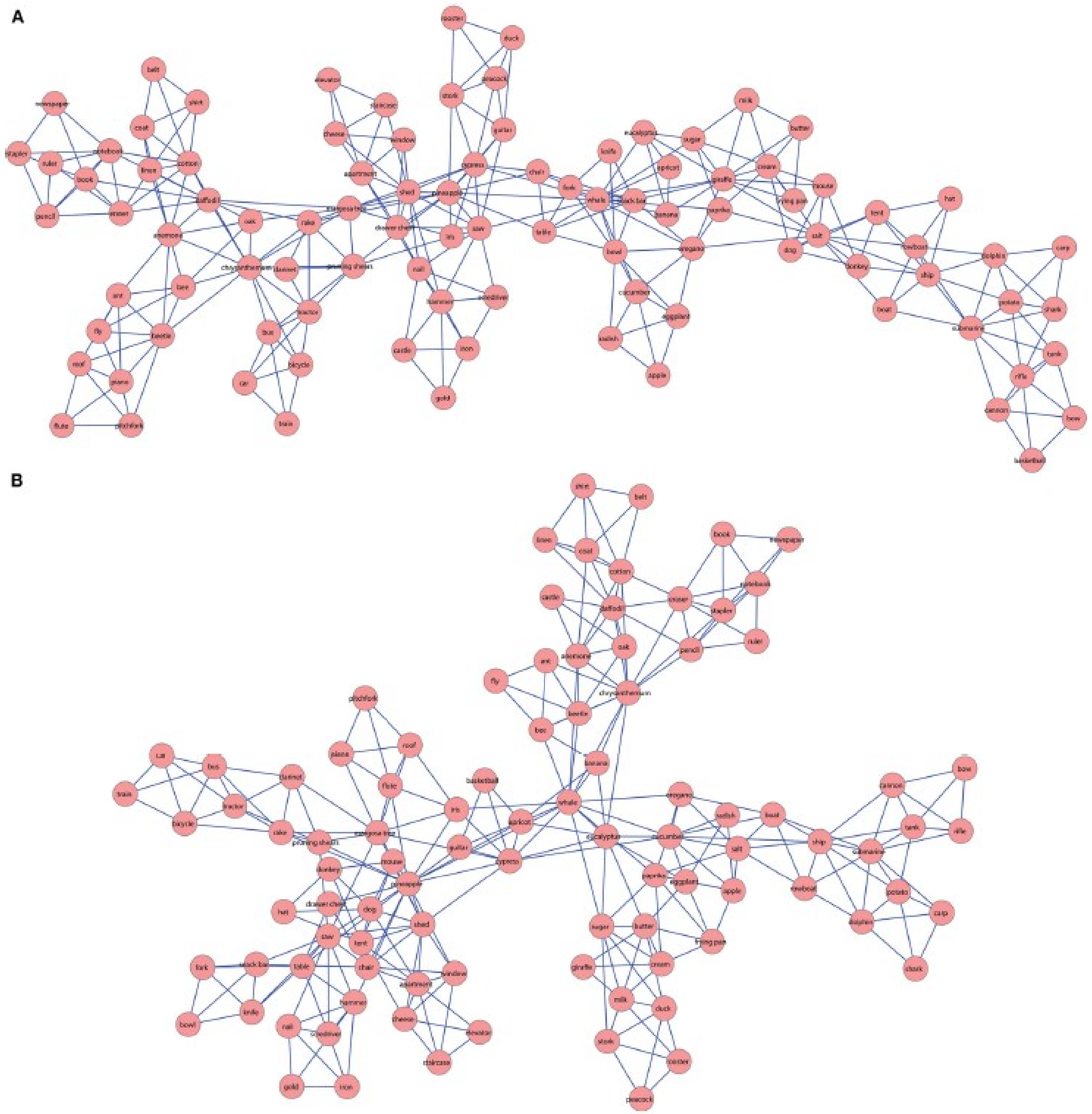
- Semantic networks exhibit a small-world and scale-free structure
- They may grow by preferential attachment ('the rich get richer'): words with more links are acquired first

(Siew et al. 2019: Fig. 1)

Lexical/semantic networks

Creativity (Kenett et al. 2014)

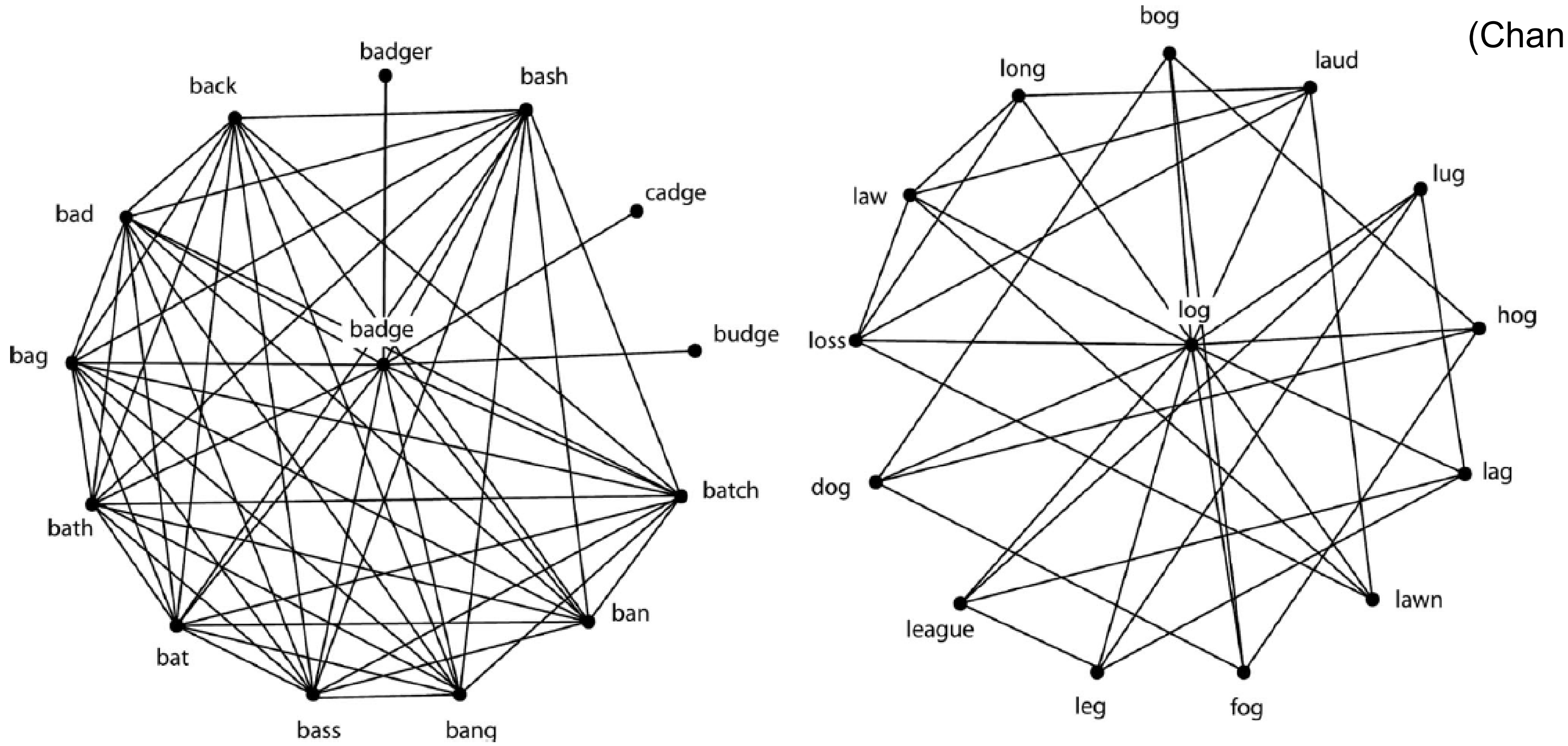
- More creative persons have networks with higher clustering coefficient and smaller average path length (i.e. more ‘small-worldness’), as well as lower modularity in community structure than less creative persons
- Creativity = flexible structures?



(Kenett et al. 2014: Fig. 1)

Phonological and orthographic networks

(Chan & Vitevich 2009: Fig. 2)



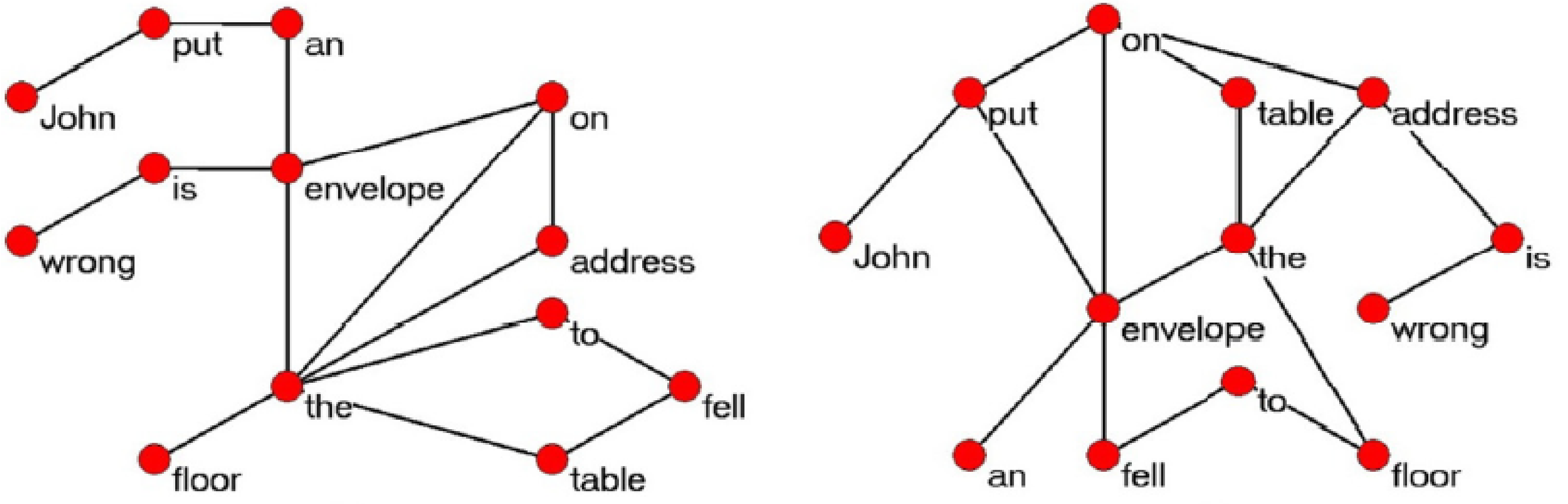
Predicting psycholinguistic effects (Chan & Vitevich 2009)

- Words with low clustering coefficient are recognised faster than words with high clustering coefficient
- The latter receive less activation because they have to 'share' with more neighbours?

Syntactic networks

Small-world and scale-free structure (Ferrer i Cancho & Solé 2001; Ferrer i Cancho et al. 2004)

- … are present in both word co-occurrence networks and syntactic dependency networks
- Network measures provide evidence of the hierarchical organisation inherent in dependency grammars



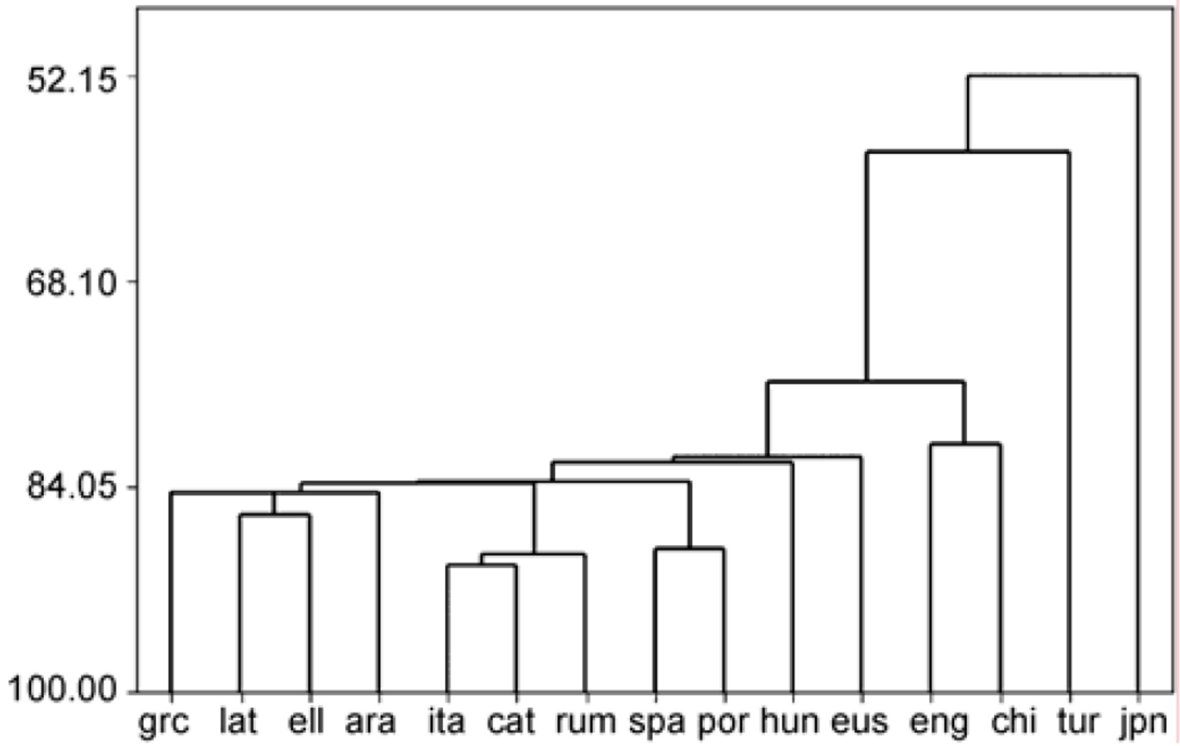
(a) (b)

(Cong & Liu 2014: Fig. 1)

Syntactic networks

Language typology (Liu & Li 2010)

- Parameters of syntactic dependency networks (*L*, *C*, <*k*> etc.) can potentially predict typological similarity, even though the effect might be largely morphologically driven (Liu & Xu 2011)



(Liu & Li 2010: Fig. 6)

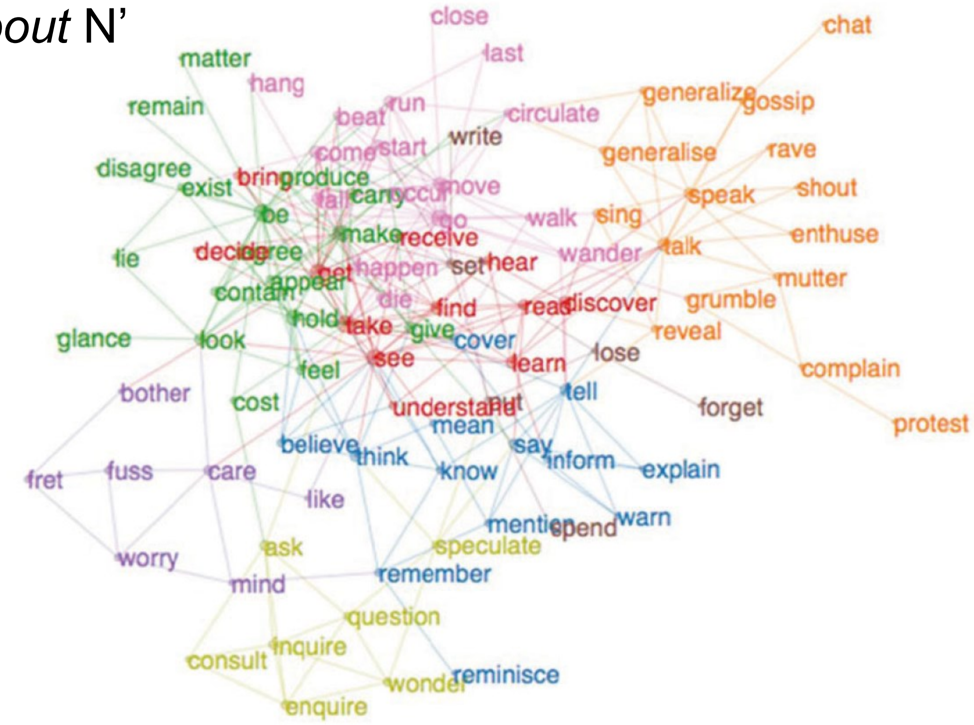
Cognitive-linguistic applications: *one* example

Ellis, Römer & O'Donnell (2016)

Semantic networks of verbs in verb-argument constructions (VACs)

- E.g. patterns like 'V *about* N', 'V *across* N', 'V *as* N'
- VACs are based on COBUILD patterns; similarity metric is based on WordNet database (Fellbaum 1998)
- Findings: semantic networks are well-connected (high *C*), have a few hubs (e.g. *say, see*, *go*) and form communities of related senses (e.g. communication expression, physical movement)

Ex.: 'V *about* N'



Psycholinguistic evidence

- The networks provide a measure of semantic prototypicality ('betweenness centrality'), which predicts how often L1 and L2 speakers generate the verbs in free association tasks
- Other factors are frequency and collostructional attraction

(Ellis et al. 2016: Fig. 3.5)

Open questions

Open questions

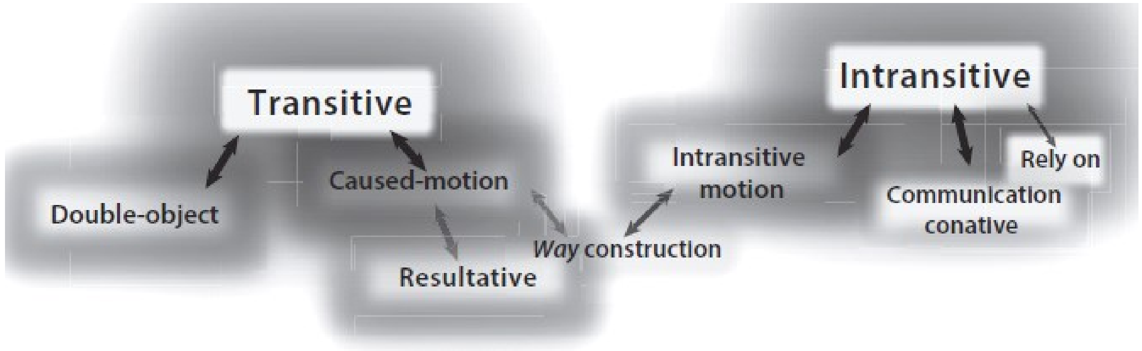
1) What other phenomena and questions in cognitive linguistics could be studied with the help of network science tools?

2) What kind of data would be required to construct the relevant networks?

3) To what extent can these methods be applied to Construction Grammar and Word Grammar networks?

   E.g. a question for CxG: what are the nodes and how many are there (words, low-level schemas, abstract phrase/clause-level cxns)?

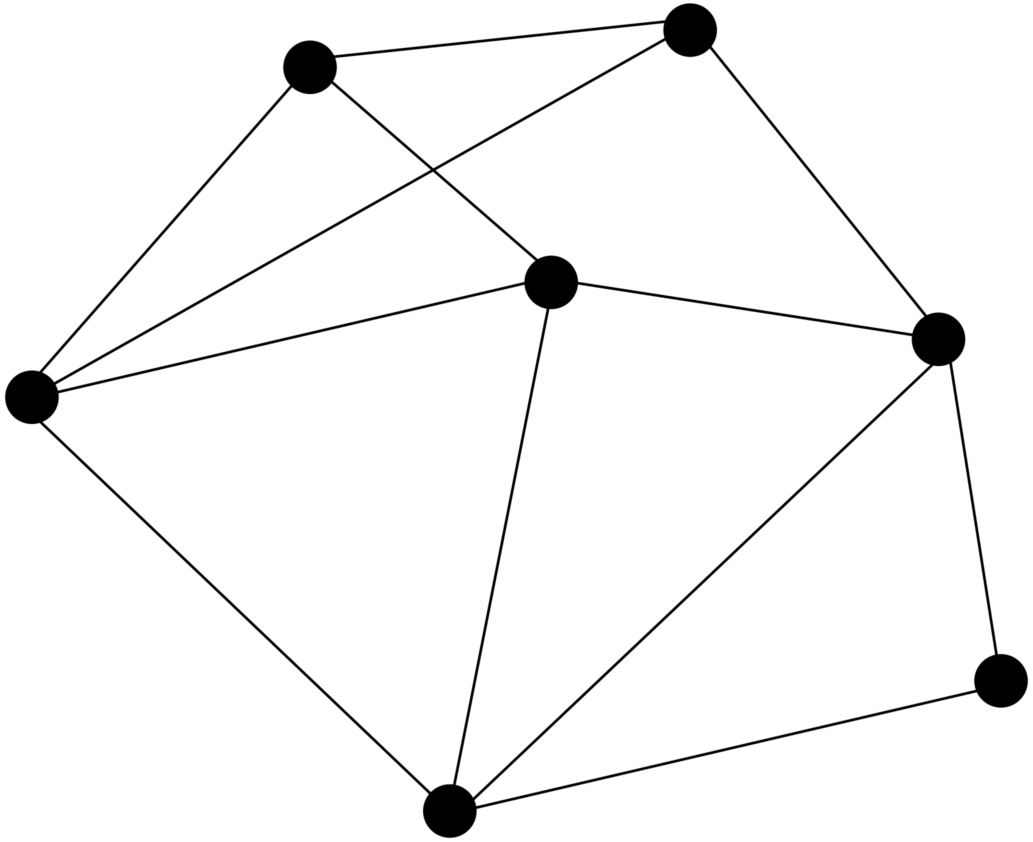
4) Can Ellis et al.'s work be extended to compare verb usage across *multiple* constructions?



(Goldberg 2019: 37)

5) How much can network science measures tell us about psychological processes and representation? How cognitively plausible are the different types of networks (phonological, semantic, syntactic dependency networks)?

Thank



you!

References

Barabási, Albert-László. 2016. *Network science*. Cambridge: Cambridge University Press.

Barabasi, Albert-Laszlo & Reka Albert. 1999. Emergence of scaling in random networks. *Science* 286(5439). 509–512.

Chan, Kit Ying & Michael S. Vitevitch. 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.

Ellis, Nick C., Ute Römer & Matthew Brook O'Donnell. 2016. *Usage-based approaches to language acquisition and processing: Cognitive and corpus investigations of construction grammar* (Language Learning Monograph Series 10). Wiley.

Fellbaum, Christiane (ed.). 1998. *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Ferrer i Cancho, Ramon, Ricard V. Solé & Reinhard Köhler. 2004. Patterns in syntactic dependency networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 69(5 Pt 1). 051915.

Ferrer i Cancho, Ramon & Richard V. Solé. 2001. The small world of human language. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268(1482). 2261–2265.

Goldberg, Adele E. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton: Princeton University Press.

Kenett, Yoed N., David Anaki & Miriam Faust. 2014. Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience* 8.

Liu, HaiTao & WenWen Li. 2010. Language clusters based on linguistic complex networks. *Chinese Science Bulletin* 55(30). 3458–3465.

Liu, Haitao & Chunshan Xu. 2011. Can syntactic networks indicate morphological complexity of a language? *EPL (Europhysics Letters)* 93(2). 28005.

Perera, Supun, Michael G. H. Bell & Michiel C. J. Bliemer. 2017. Network science approach to modelling the topology and robustness of supply chain networks: A review and perspective. *Applied Network Science* 2(1). 33.

Siew, Cynthia S. Q., Dirk U. Wulff, Nicole M. Beckage & Yoed N. Kenett. 2019. Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity* 2019. e2108423.

Steyvers, Mark & Joshua B. Tenenbaum. 2005. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science* 29(1). 41–78.

Watts, Duncan J. 2007. A twenty-first century science. *Nature* 445(7127). 489–489.