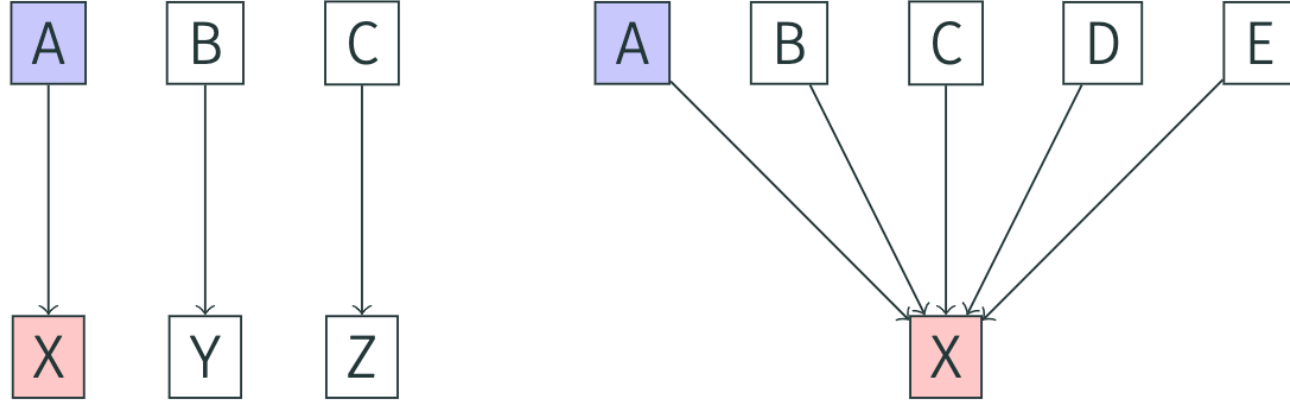


# A network-based approach to lexical dynamics

Maria Copot      Andrea D. Sims

## Background

- Information theory used to describe morphological systems and patterns within
- Common measures don't differentiate between situations that feel different to the linguist



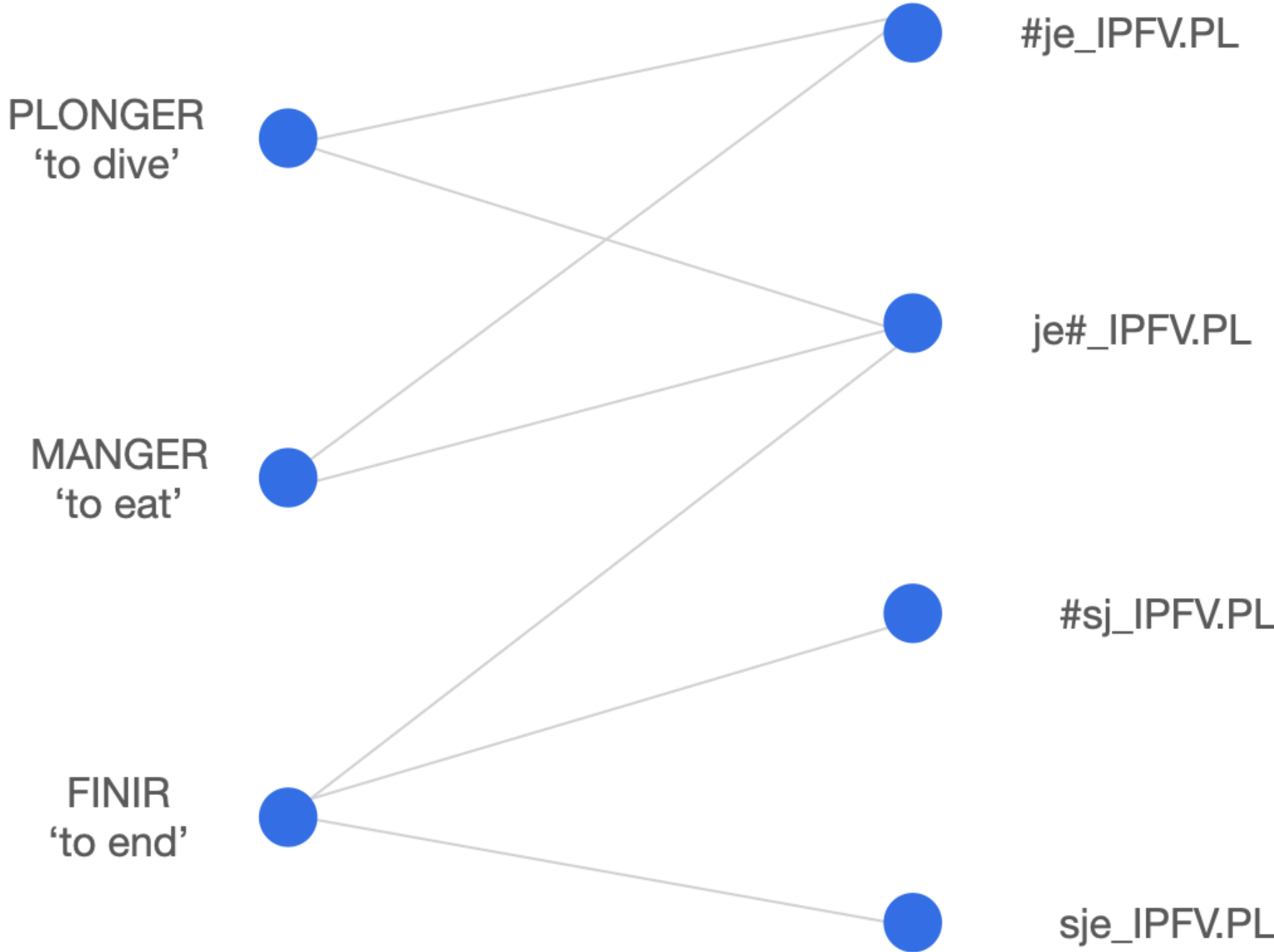
e.g. the entropy of default options and strong paradigmatic implications is the same

## Goals

- We propose a network-based approach to lexical structure. Strengths:
  - attention to configurations in both local and global dynamics
- Quantitative, gradient, interpretable
- We apply the approach to the verbal system of French and the nominal system of Bosnian-Croatian-Montenegrin-Serbian (BCMS) as a proof of concept.

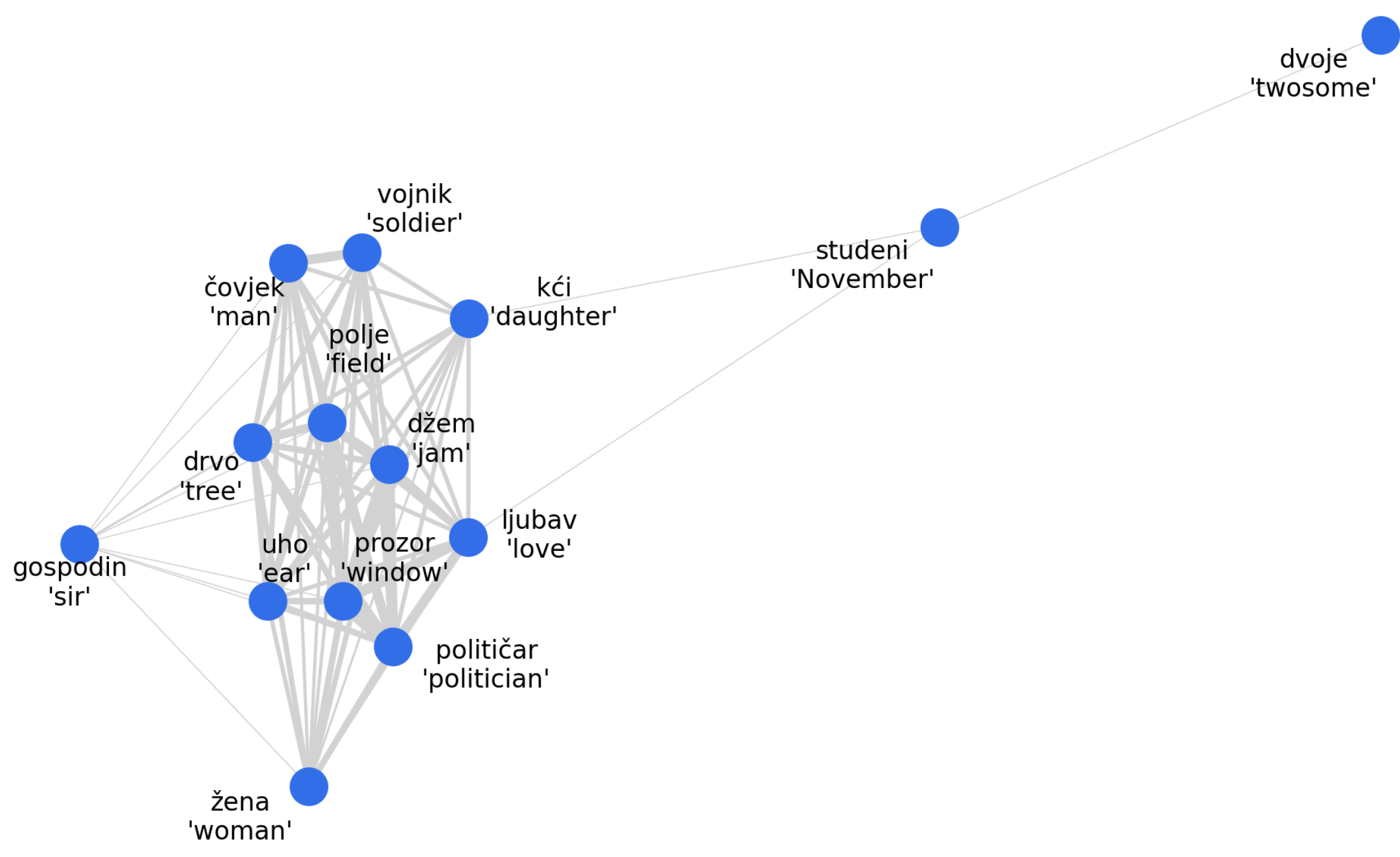
## Creating the network

- Inflectional lexicon → segmentation into exponents (Beniamine & Guzmán-Naranjo, 2020; Beniamine & Carroll, 2023) → removal of nondiscriminative formatives → conversion of exponents to triphones, tagged for cell.
- Data from French verbs (Beniamine et al., 2023) and BCMS nouns (Batsuren et al., 2022).



French bipartite network (subgraph)

**Bipartite network** (Gross & Yellen, 2006) linking lexemes to tagged triphones.



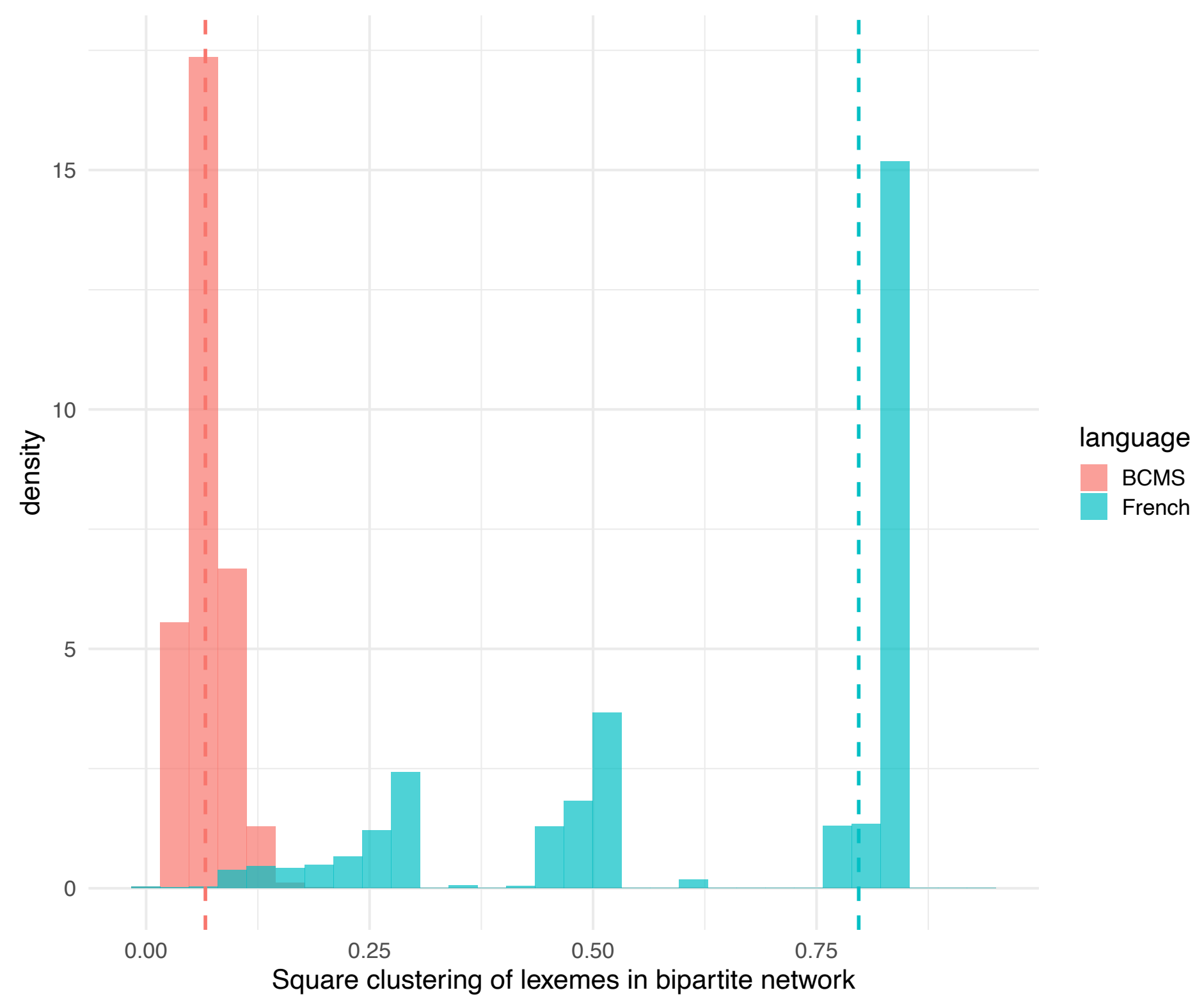
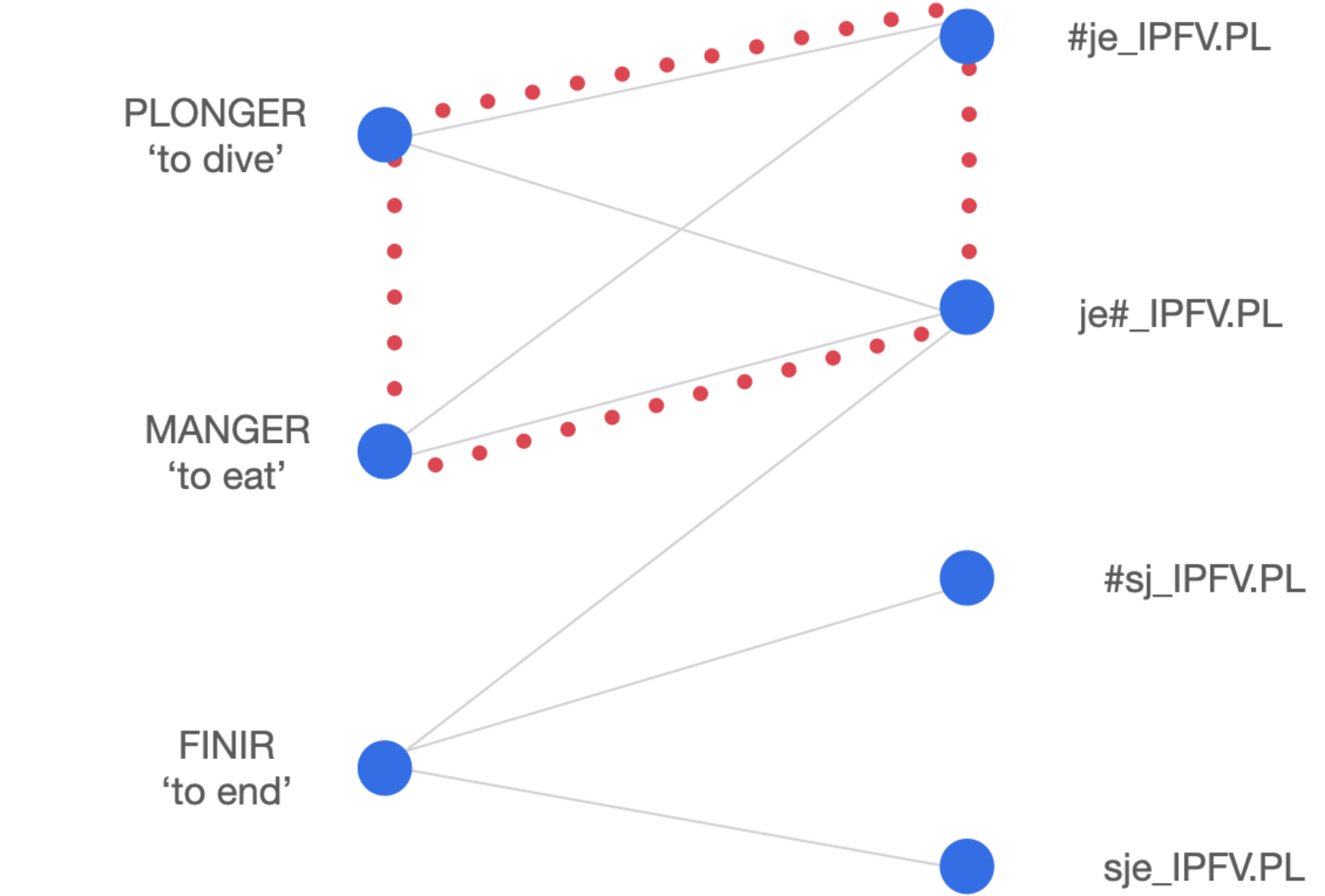
BCMS one-mode network (subgraph)

**One-mode network** links between lexemes weighted by how many triphones are shared.

## Connectivity patterns

Square clustering captures the strength of implicative relationships.

Implicative relationships between exponents appear much stronger in French than BCMS - French inflectional classes have more distinct inflectional behaviour compared to each other than BCMS classes. In other words, joint exponent probability is much higher on average in French compared to BCMS.



## References

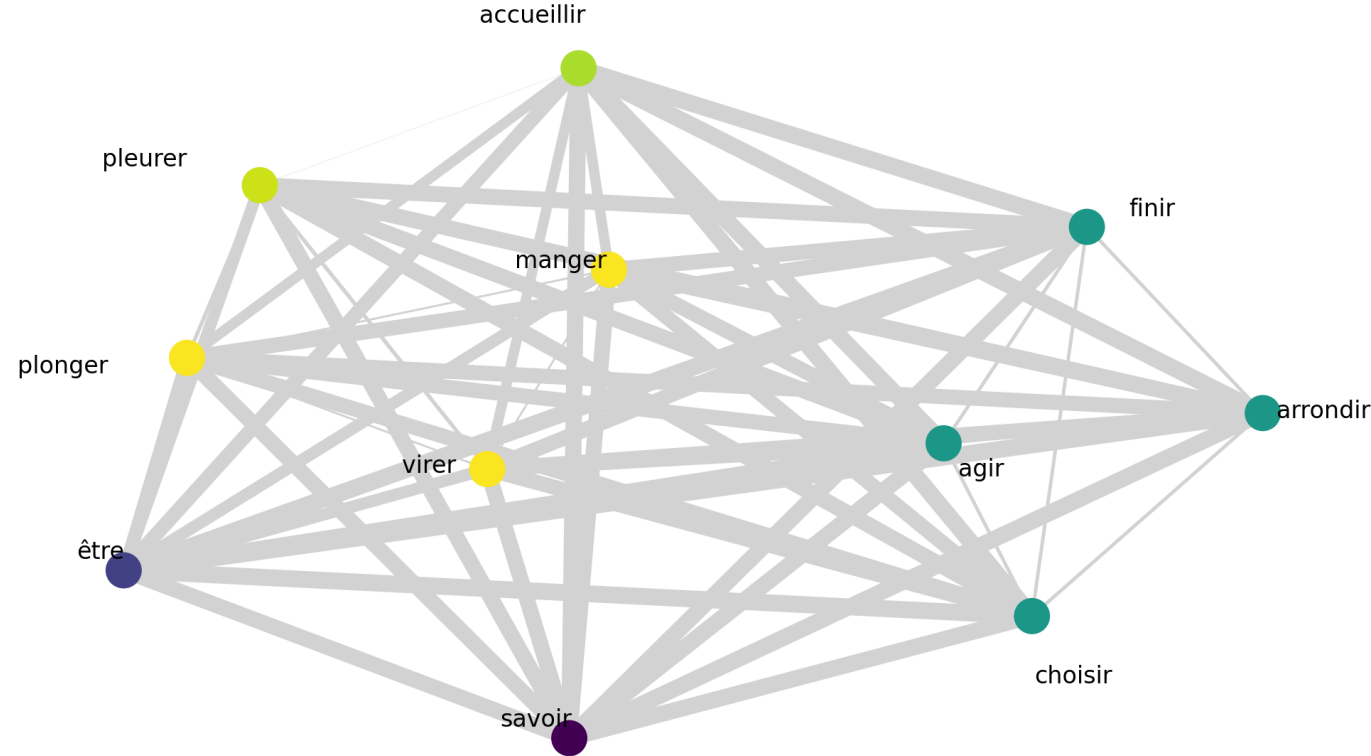
Batsuren, K., Goldman, O., Khalifa, S., Habash, N., Kieraś, W., Bella, G., Leonard, B., Nicolai, G., Gorman, K., Ate, Y. G., Ryskina, M., Mielke, S. J., Budianskaya, E., El-Khaissi, C., Pimentel, T., Gasser, M., Lane, W., Raj, M., Coler, M., Montoya Samame, J. R., ... Vylomova, E. (2022). UniMorph 4.0: Universal morphology. *arXiv*.  
Bavelas, A. (1950). Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America*, 22(6), 725–730.  
Beniamine, S., & Carroll, M. (2023, September). The other perspective on exponence. Talk presented at the International Symposium of Morphology 2023, Nancy, France.  
Beniamine, S., Coavoux, M., & Bonami, O. (2024). Vlexique 2.0 (2.0) [Data set]. Zenodo.  
Beniamine, S., & Naranjo, M. G. (2021). Multiple alignments of inflectional paradigms. In A. Ettinger, E. Pavlick, & B. Prickett (Eds.), *Proceedings of the Society for Computation in Linguistics 2021* (pp. 216-227). Association for Computational Linguistics.  
Gross, J. L., & Yellen, J. (2006). *Graph theory and its applications*. Boca Raton, FL: Chapman & Hall.

## Acknowledgements

The research was funded by a Presidential Postdoctoral Scholarship at OSU, awarded to Maria Copot.

## Uniqueness of exponence

**Closeness centrality** (Bavelas, 1950) captures the extent to which a lexeme's exponents take part in the system



The more uncommon the exponents, and the greater the number of uncommon exponents, the lower the lexeme's closeness. Lexemes with low closeness are peripheral to the system.

### French:

- Defective lexemes (e.g., *accroire* 'make sb believe', *quérir* 'seek')
- Suppletive verbs (e.g., *avoir* 'to have', *être* 'to be')
- Verbs with very rare or unique exponents (e.g., *faire* 'to do')

### BCMS:

- Pluralia tanta neuters (e.g., *pluća* 'lungs')
- Nouns with unusual exponent(s) (e.g., *oba* 'both', *gen.pl obaju*)
- Collectives (e.g., *dugmad* 'buttons')

## Main Takeaways

- Network science is a promising framework for investigating lexical dynamics
  - Centers a relational and configurational approach
  - Lossless description of a system, promising for typology
  - Well-developed toolkit for quantitative methods and gradient approaches
- Proof of concept: we discuss two dimensions that characterise the role of a lexeme and its exponents in a network. Many more network-based measures relevant to the organisation of lexical systems exist, and will form the object of further work.

## OSF Repository

