Crash course in causal graphs and causal inference for linguists

Seán Roberts, Cardiff University

It's a very exciting time to be a linguist. There's more data than ever and more ways of analysing it. The next generation of Big Data methods even promise to create theories for us. However, with big data comes big trouble. With so many possible connections to test, how do we know which to investigate or control for? How can we integrate theories and evidence from multiple fields and multiple methods?

This workshop will provide a crash-course in causal inference, a powerful way of representing hypotheses about how variables are causally related. It will also present a database of causal hypotheses. CHIELD (https://chield.excd.org/) is a tool for coding existing theories as causal graphs and allowing researchers to locate evidence, discover links between theories and spot weak links. CHIELD focusses on theories of language evolution and language change, but the general approach could be applied to any kind of linguistics.

The workshop will introduce the basics behind causal inference, mainly drawing on Pearl (2018), Daniel (2017), and Rohrer (2018).

Part of the workshop will be delivered asynchronously (links to follow). Before the workshop, you can **view some pre-recorded videos** on:
- The basics of causal inference (20 minutes)
- Confounding and colliders (20 minutes)
- Front door and back door adjustment (20 minutes)
- Model fitting and causal discovery (20 minutes)
- Introduction to the CHIELD database (10 minutes)

The workshop itself will be held live online. It will focus on answering your questions, and a practical task. Before the workshop, **read one of the target papers (links to follow)**, and be prepared to work with others to code it as a causal graph.

References

Roberts, S., Killin, A., Deb, A., Sheard, C., Greenhill, S. J., Sinnemäki, K, Segovia-Martín, J, Nölle, J., Berdicevskis, A., Humphreys-Balkwill, A., Little, H., Opie, C., Jacques, G., Bromham, L., Tinits, P., Ross, R. M., Lee, S., Gasser, E., Calladine, J., Spike, M., Mann, S. F., Shcherbakova, O., Singer, R., Zhang, S. , Benítez-Burraco, A., Kliesch, C., Thomas-Colquhoun, E., Skirgård, H., Tamariz, M., Passmore, S., Pellard, T. & Jordan, F. (2020) CHIELD: The causal hypotheses in Evolutionary Linguistics Database. *Journal of Language Evolution*, lzaa001.

Pearl, J. (2018) The Book of Why: The New Science of Cause and Effect. Penguin.

Daniel, R. (2017) Introduction to causal inference. UK Causal Inference Meeting 2017.

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. Advances in Methods and Practices in Psychological Science, 1(1), 27-42.