

# Distributional semantics and linguistic research

Louise McNally

Universitat Pompeu Fabra

This talk, which reports on a position paper in progress with Gemma Boleda, Josep M. Fontana and Alessandro Lenci, begins with a brief introduction to distributional semantics, an approach to meaning representation which has come to dominate computational semantics in the last 10 years. Distributional semantics traces its modern origins to Zelig Harris and J.R. Firth, as well as to Latent Semantic Analysis (Landauer and Dumais 1997) and "vector space" methods in information retrieval. Distributional representations are usage-based, non-discrete, not limited to truth-conditionally relevant components of meaning, integratable with multi-modal sources of information (e.g. image, sound), and amenable not only to compositional but also to "decompositional" views of meaning. Despite the existence for some time of notable (and very varied) efforts within computational semantics to connect distributional semantics to the logical semantics tradition (a representative sample includes Coecke, et al. 2010, Garrette et al. 2011, Copestake and Herbelot 2012, Lewis and Steedman 2013, Baroni et al. 2014 and the papers in Boleda and Herbelot 2016), the method has had virtually no impact on formally-oriented theoretical linguistic research. We want to change this by highlighting the potential of distributional methods, when properly combined with formal methods, not only to afford insight into the semantic problems for which they are already well known, such as the analysis of polysemy, but also to illuminate certain claims in the syntax literature that have lacked any meaningful connection to semantic theory (arguably to the detriment of both subfields), as well as patterns of semantic change.