

Vector semantics and embeddings unit

- ▶ Lexical semantics, words as vectors (**Today**)
- ▶ Word2Vec (Wednesday)
- ▶ Applications of embeddings (Friday)

Today:

- ▶ The idea of word embeddings
- ▶ Lexical semantics
- ▶ Human-crafted embeddings
- ▶ ML embeddings

Vector semantics is the standard way to represent word meaning in NLP. [Some linguists have proposed] The roots of the model lie in the 1950s when two big ideas converged: [using] a point in three-dimensional space to represent the connotation of a word, and the proposal . . . to define the meaning of a word by its distribution in language use, meaning its neighboring words or grammatical environments.

The idea of vector semantics is to represent a word as a point in a multi-dimensional semantic space that is derived from the distributions of word neighbors. Vectors for representing words are called **embeddings**.

Coming up:

- ▶ Do NBC programming assignment (Mon, Nov 6)
- ▶ Read J&M 6 (Wed, Nov 8)
- ▶ Do Semantle exercise (Fri, Nov 10)

Word2Vec assignment coming. . .