

The Distributional Hypothesis

*You shall know a word by the company it keeps.
(J. R. Firth, 1957)*

Words with **high similarity** occur in the **same contexts** as one another.

A bit of foreshadowing:

- A **word** ought to be able to **predict its context** (word2vec Skip-Gram)
- A **context** ought to be able to **predict its missing word** (word2vec CBOW)