

# Static Word Embeddings

Simon Clematide

Department of Computational Linguistics  
University of Zurich

Machine Learning for NLP 1

# Learning Objectives

- ▶ Understand the basic ideas of distributionalism and its various instantiations
- ▶ Understand the reason behind and advantages of dense representations
- ▶ Grasp the basic ideas behind embeddings in word2vec and glove

# What is Meaning?

What is **bardi**wac?



[YANDEX 2019]

# Short Version of Long History of Distributionalism

- ▶ “Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.” (Ludwig Wittgenstein)
- ▶ “Distributional Structure” [HARRIS 1954]
- ▶ “You shall know a word by the company it keeps!” (J. R. Firth (1957))
- ▶ “words which are similar in meaning occur in similar contexts (Rubenstein & Goodenough, 1965)”
- ▶ “words with similar meanings will occur with similar neighbors if enough text material is available” (Schütze & Pedersen, 1995)
- ▶ “words that occur in the same contexts tend to have similar meanings” (Pantel, 2005)

See [SAHLGREN 2008] for details.

# What is Meaning?

What is **bardiwac**?

- He handed her a glass of **bardiwac**.
- Beef dishes are made to complement the **bardiwac**.
- Nigel staggered to his feet, face flushed from too much **bardiwac**.
- Malbec, one of the lesser-known **bardiwac** grapes, responds well to Australia's sunshine.
- I dined off bread and cheese and this excellent **bardiwac**.
- The drinks were delicious: blood-red **bardiwac** as well as light, sweet Rhenish.



**Bardiwac** is a ...

[YANDEX 2019]