

Multilingual Distributional Semantics

Benno Kruit Sara Veldhoen

January 13, 2015

Outline

Introduction - related work

Our first idea (and why it wouldn't work)

Our new idea

Introduction -
related work

Our first idea (and
why it wouldn't
work)

Our new idea

Our new idea

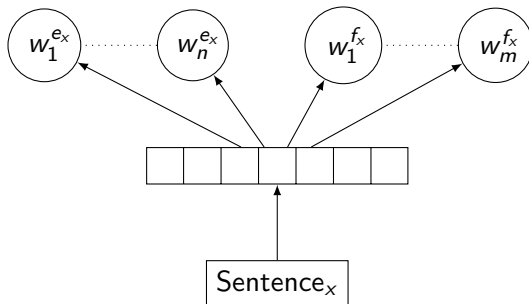


Figure : Bilingual dbow

Our new idea

- ▶ Training a single embedding for parallel sentences
- ▶ Word embeddings are not trained
- ▶ Can be extended to more than two languages

Our new idea

- Use sentence embeddings to obtain word vector:

$$w = \frac{1}{\text{freq}(w, D)} \sum_{s \in D} \sum_{w' \in s} s$$