

## Quantity of linguistic information used during training

BERT, 2019

ELMo, 2018

Sense embeddings, 2015

Retrofitting, 2015

Yu & Dredze, 2014

word2vec 2013, GloVe 2014

fasttext, 2017

Collobert et al., 2011

Collobert & Weston, 2008

Small models

Large models