





Motivation: Unsupervised Word Translation

OT with Global Invariances

35

5

• Assume access to monolingual corpora only

- First step: find corresponding between words

• dead: word co-occurrence patterns should be similar for a word and its translation





• Word Embeddings: compact representation of current information

EN



ES



*"The neighbor's garden is twice the size of my whole **house**"*      *"... La **casa** no tiene jardín..."*

*"**Home** price gains keep shrinking, but housing is still overpriced..."*      *"El precio de esta **casa** es ridículo ..."*

EN



ES



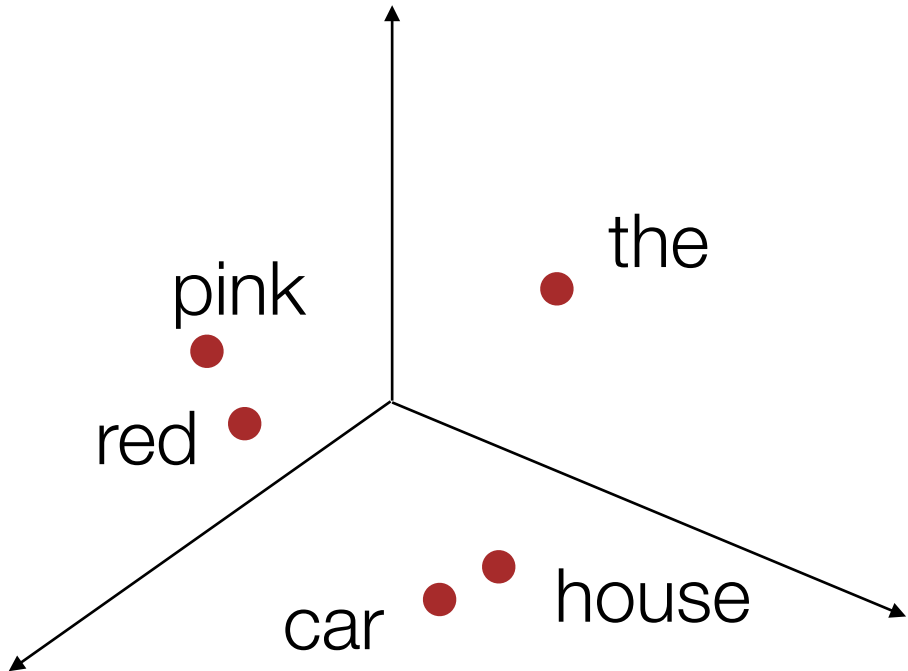
E.g. *word2vec*,

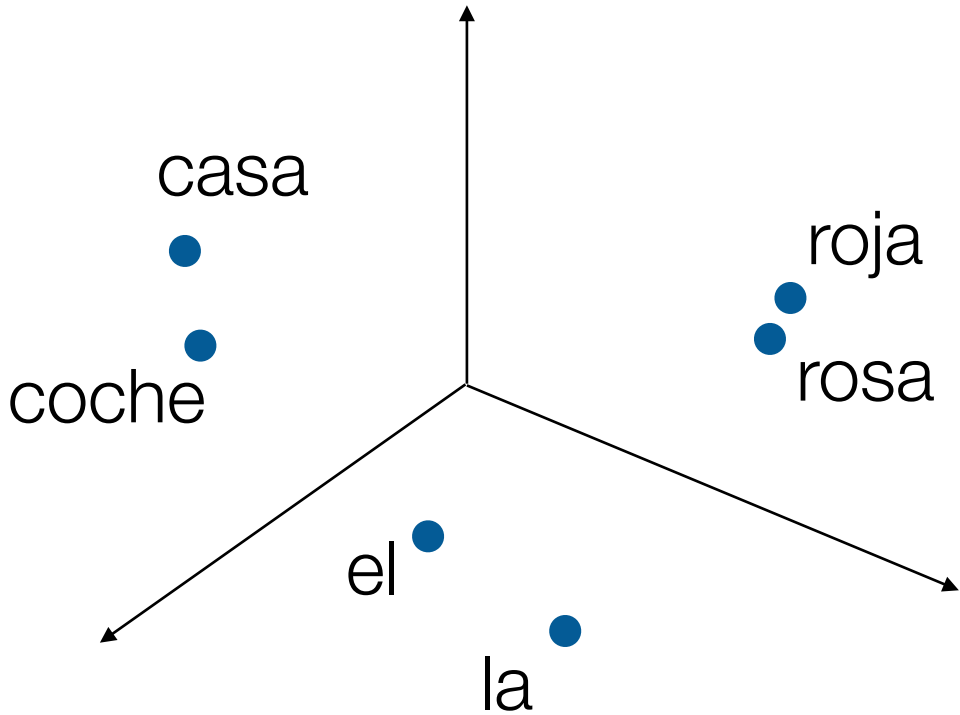
*Glove, etc*











# OT with Global Invariances

## Motivation: Unsupervised Word Translation

- Assume access to monolingual corpora only
- First step: find correspondences between words
- Idea: word co-occurrence patterns should be similar for a word and its translation

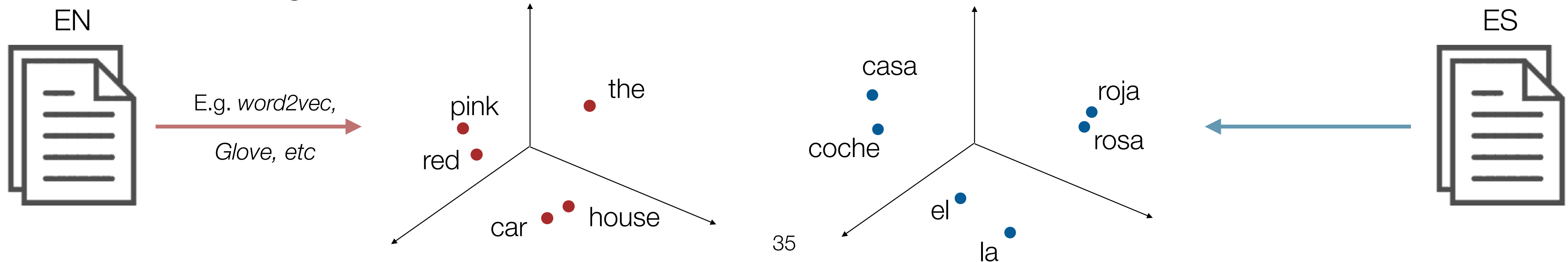
*"The neighbor's garden is twice the size of my whole **house**"*

*"... La **casa** no tiene jardín..."*

*"**Home** price gains keep shrinking, but housing is still overpriced..."*

*"El precio de esta **casa** es ridículo ..."*

- Word Embeddings: compact representation of co-occurrence information



# Optimal Transport between word embeddings

- Treat embeddings as support points of discrete distribution