[redacted]

between word embeddings

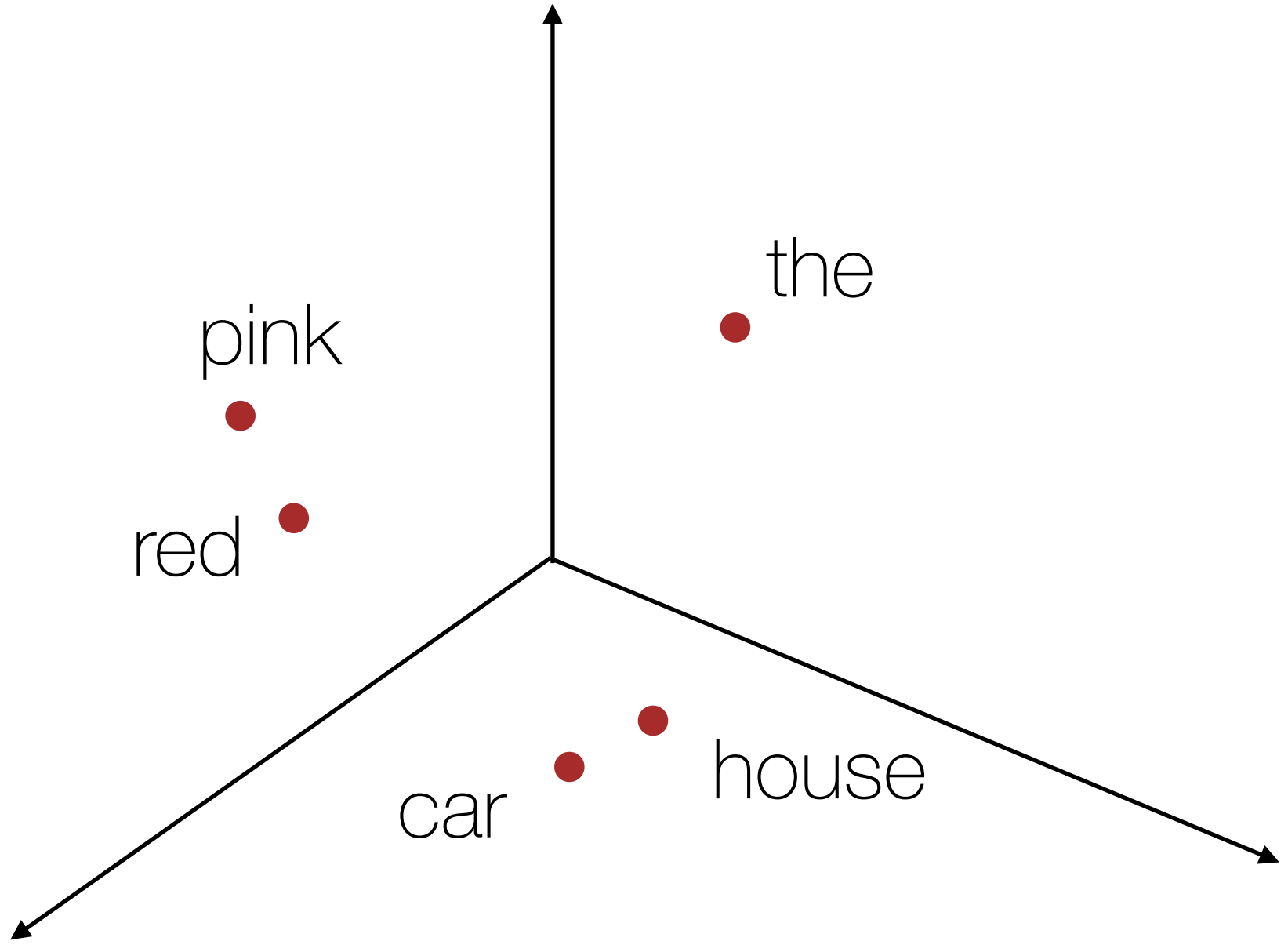
Optimal Transport

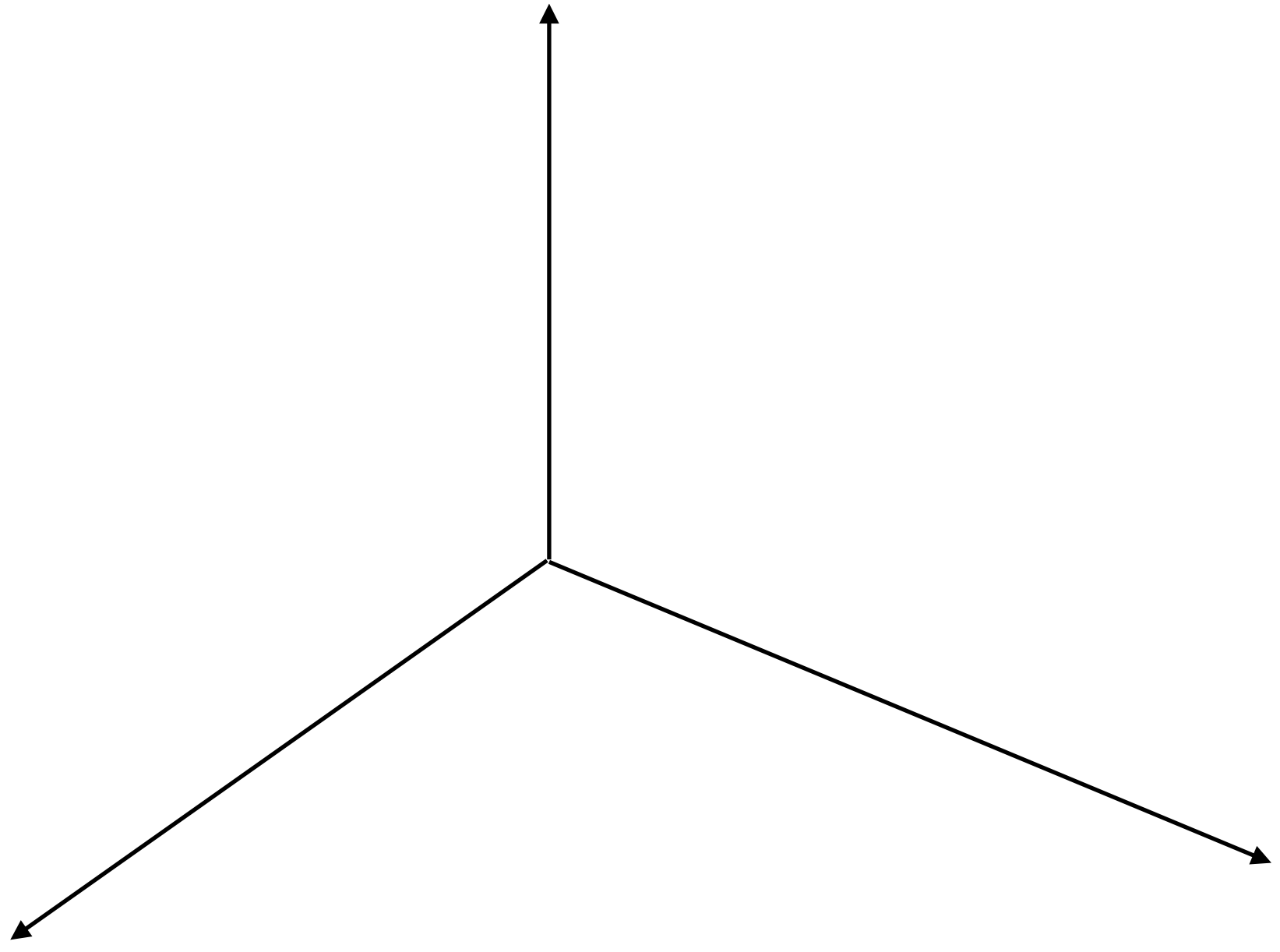
37

- Treat embeddings as support points of discrete distribution

- But this assumes the two spaces are registered (~axes are in correspondence)

- Not true in general for word embeddings in different languages!


the
pink
red
car
house

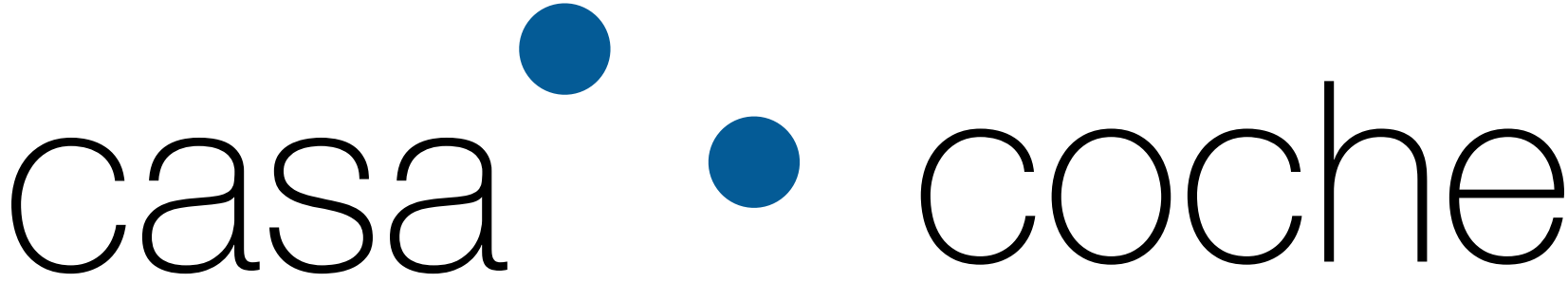


el

la

rosa

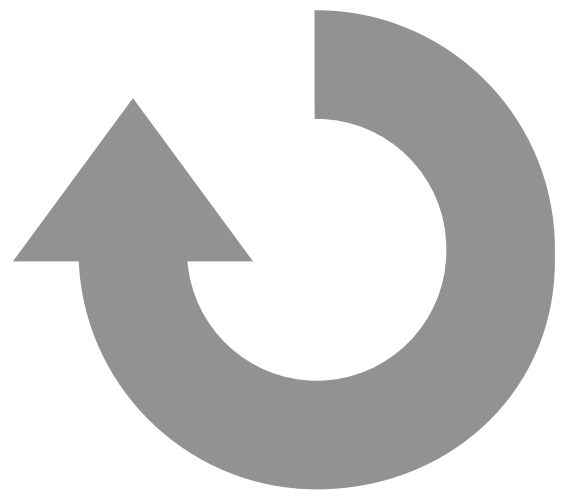
roja


casa
coche

[Kusner et al. 2015]

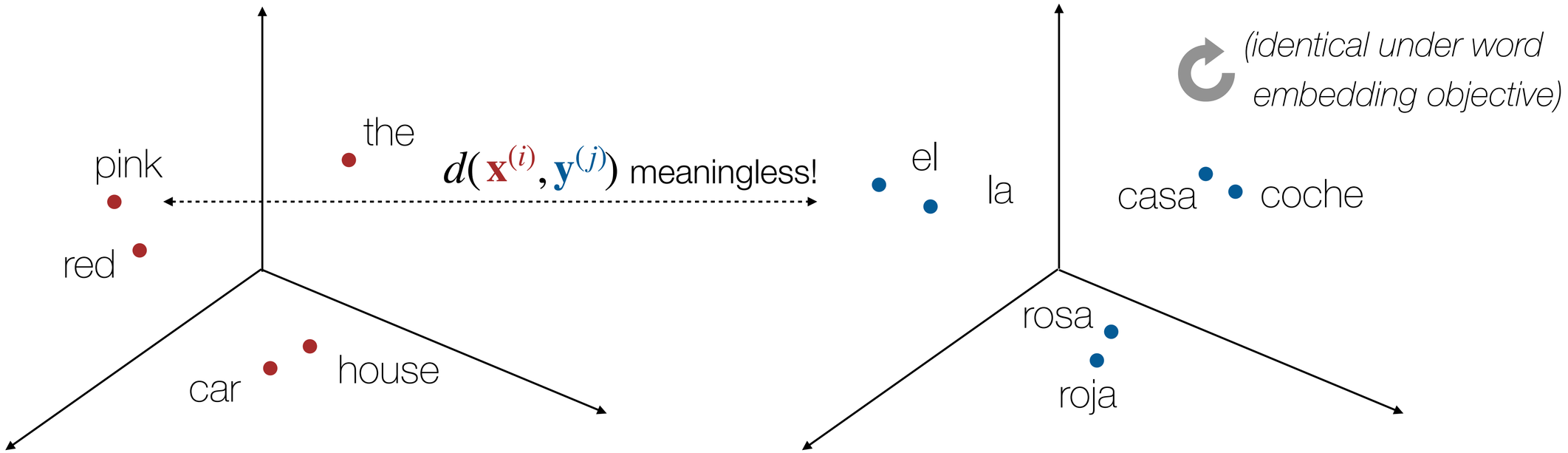
$d(\mathbf{x}^{(i)}, \mathbf{y}^{(j)})$ meaningless!

(identical under word embedding objective)



Optimal Transport
between word embeddings

- Treat embeddings as support points of discrete distribution



- But this assumes the two spaces are registered (~axes are in correspondence)

- Not true in general for word embeddings in different languages!

A general correspondence problem

Data | Two collections of points: $X = \{\mathbf{x}^{(i)}\}_{i=1}^{n}, \mathbf{x}^{(i)} \in \mathcal{X} \subset \mathbb{R}^{d_x}$

$Y = \{\mathbf{y}^{(j)}\}_{j=1}^{m}, \mathbf{y}^{(j)} \in \mathcal{Y} \subset \mathbb{R}^{d_y}$