## **FastText**

Represent a word as a bag of character n-grams, e.g. for n = 3:

$$G_{where}$$
: \_wh, whe, her, ere, re\_, \_where\_

Model a word vector as a sum of sub-word vectors:

## **SGNS**:

$$sim(u, v) = \langle \phi_u, \theta_v \rangle$$

## **FastText:**

$$sim(u, v) = \langle \phi_u, \theta_v \rangle$$
  $sim(u, v) = \sum_{g \in G_v} \langle \phi_u, \theta_g \rangle$ 

Code and pre-trained embeddings: <a href="https://fasttext.cc/">https://fasttext.cc/</a>

P. Bojanowsky et al. Enriching Word Vectors with Subword Information, 2016.