



# **Pomaci u Neuralnom Mašinskom Prevodjenju**

# **Sta je mašinsko prevodjenje?**

**Idealno celi tekstovi**

## **Zasto nas zanima?**

**50% sadržaja na internetu je na Engleskom  
20% Svetske populacije zapravo govori Engleski**

# **Rule-based sistemi**

# **Statističko mašinsko prevodjenje**

$$P(y/x)$$

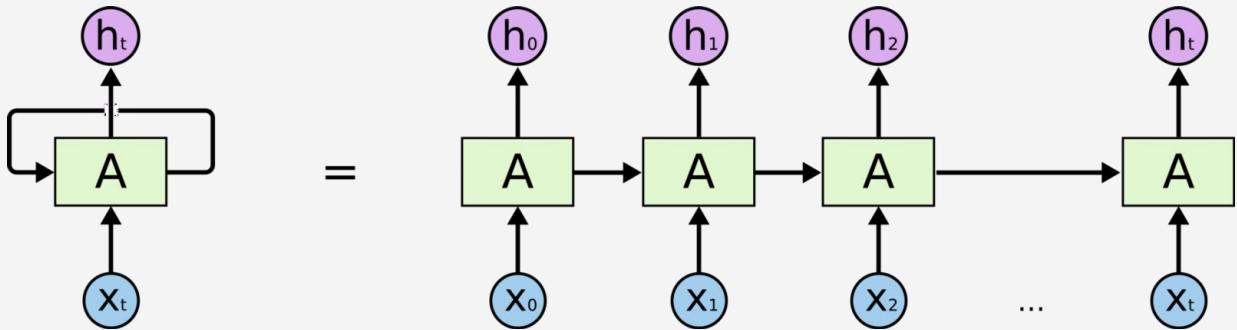
**Modeli najčešće zasnovani na pojedinačnim  
prevodima delova rečenica, fraza**

# **Neuralno mašinsko prevodjenje**

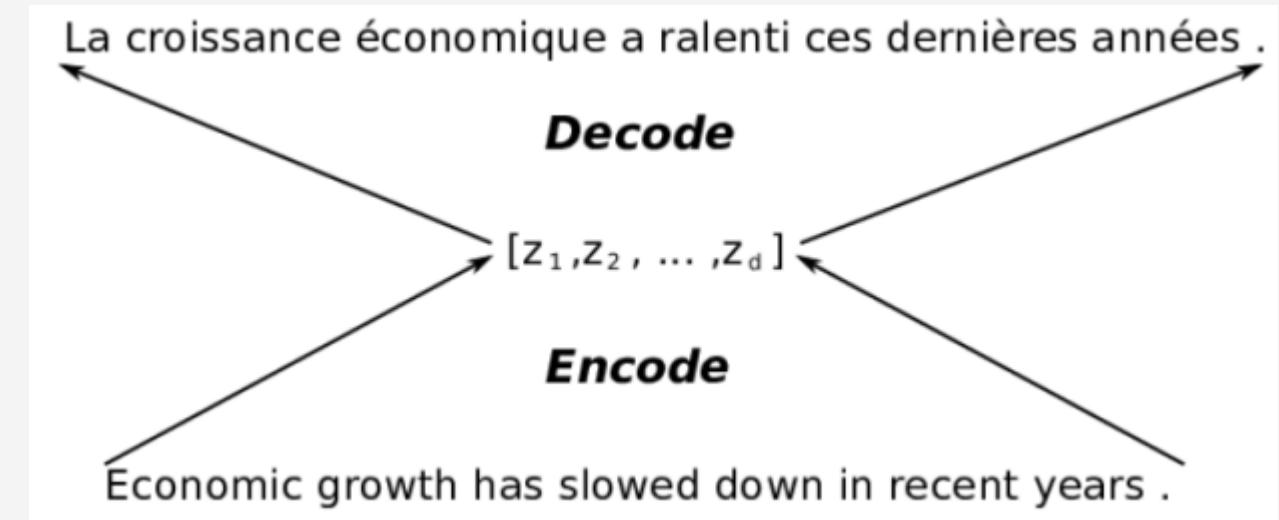
**Sequence-to-sequence mapping**

**Bazirano na rečima!**

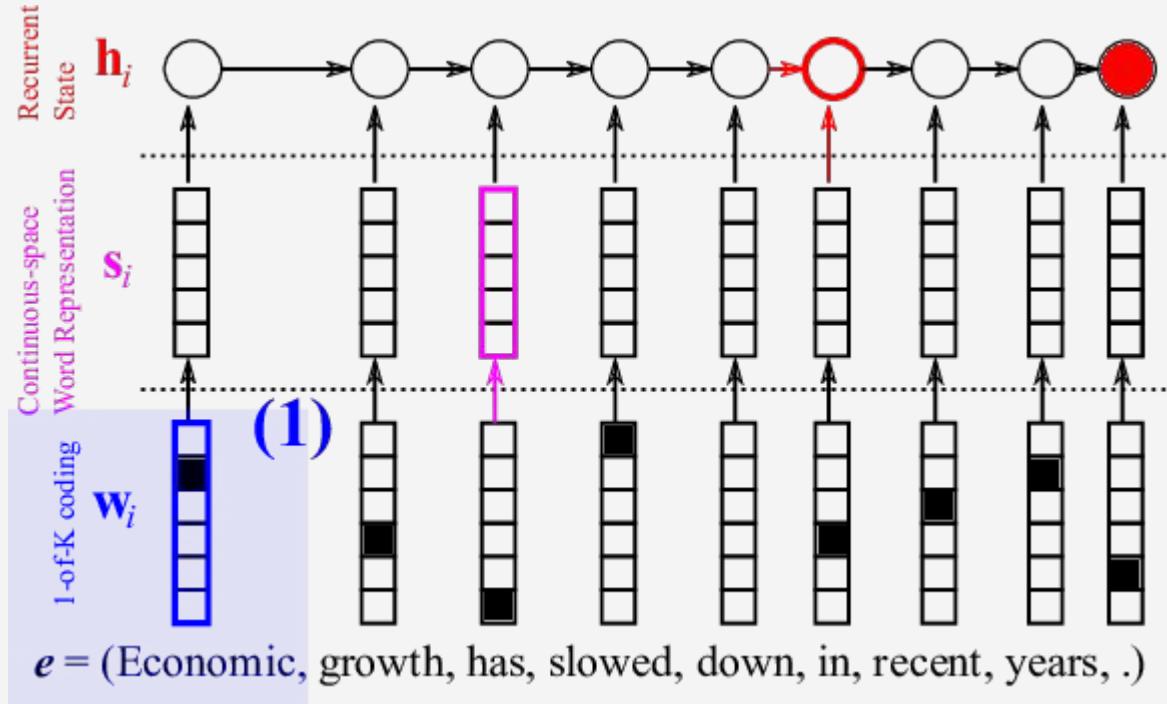
# Recurrent Neural Unit



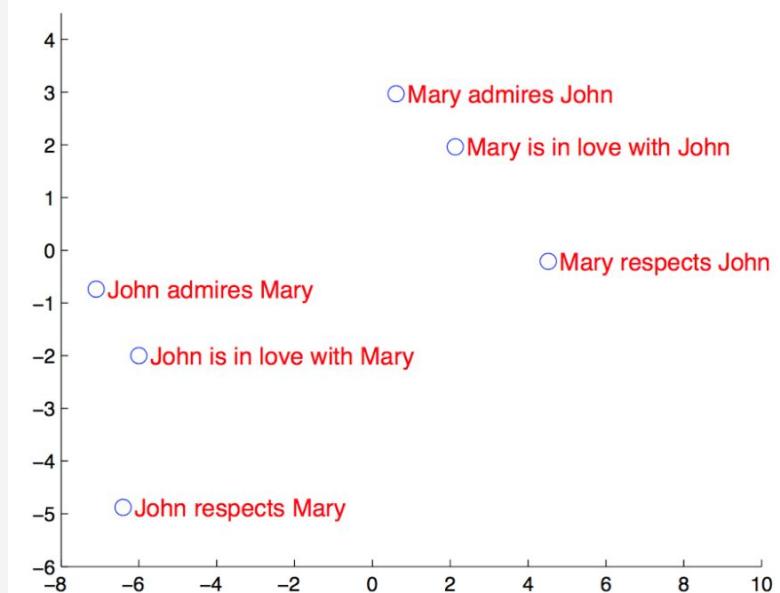
# Encoder-Decoder arhitektura



# Encoder

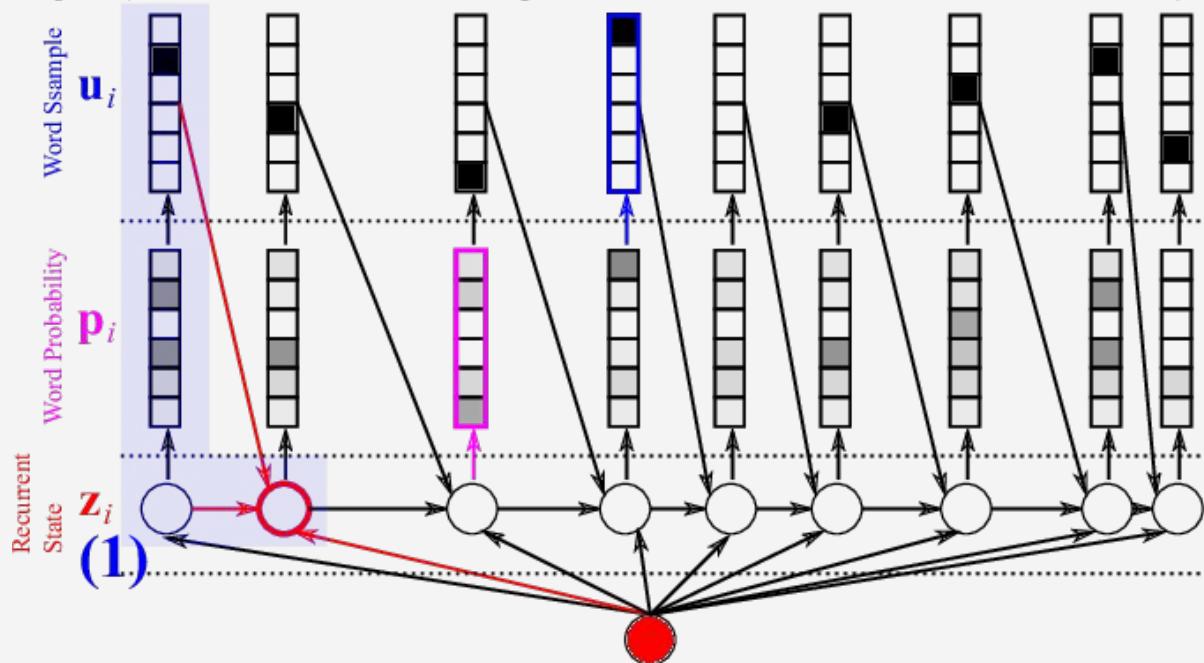


# Enkodirane recenice

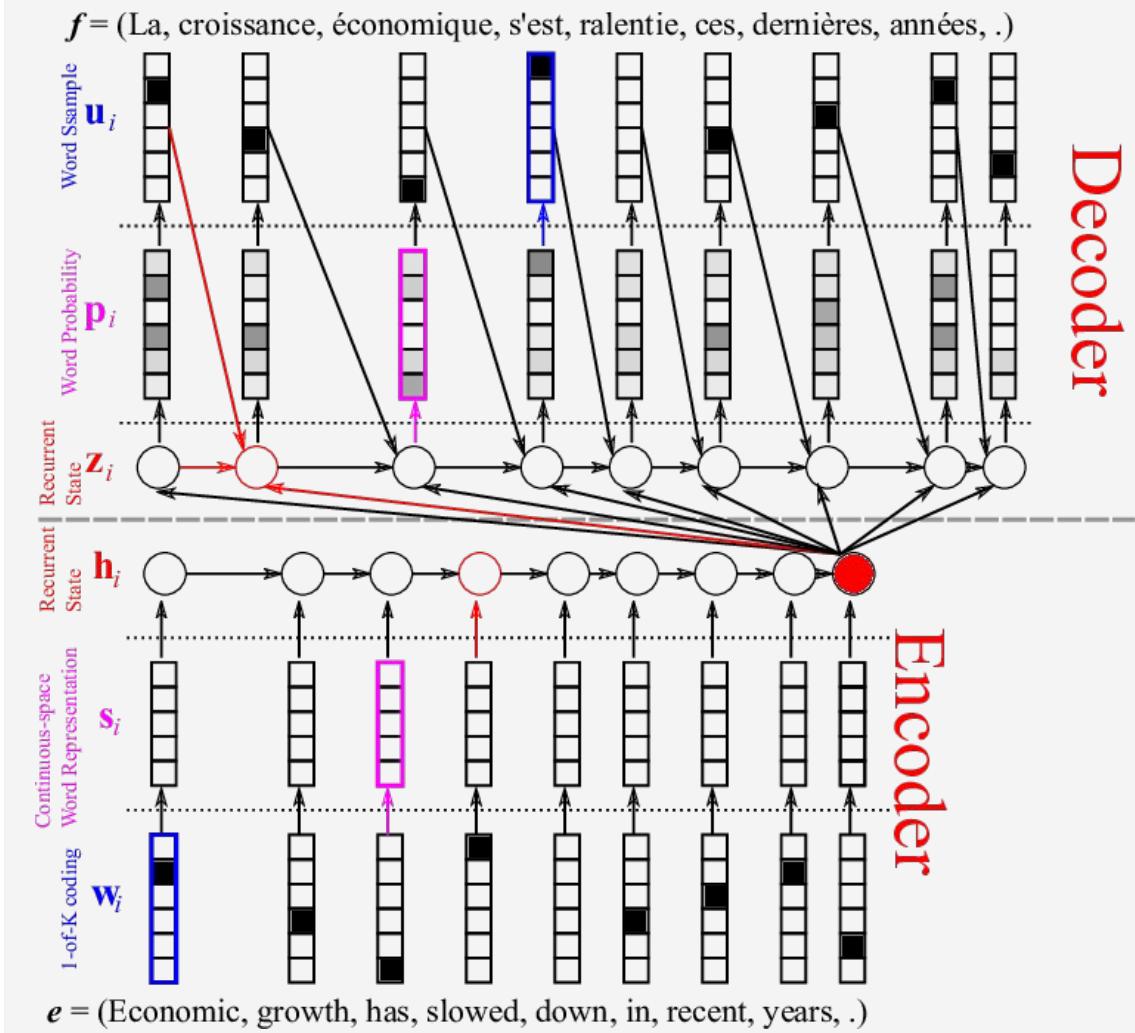


# Decoder

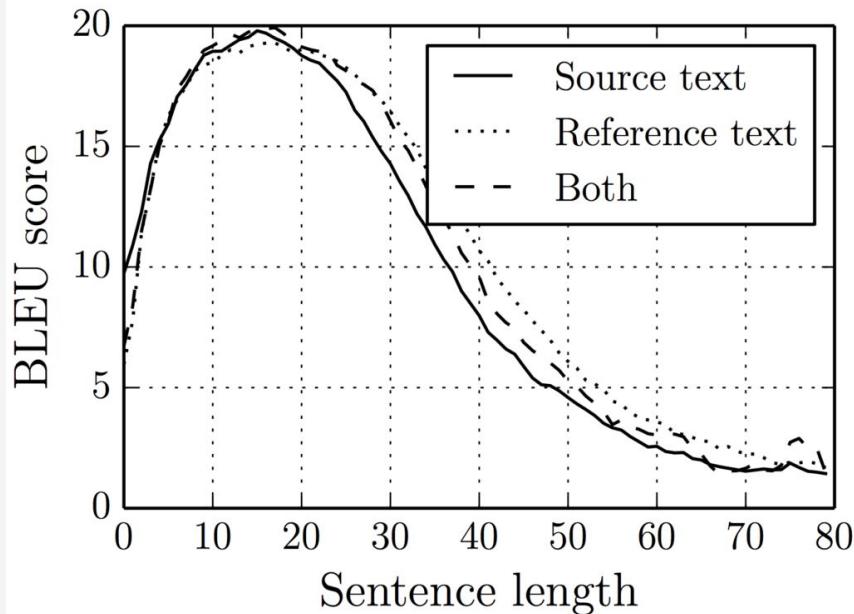
$f = (\text{La, croissance, économique, s'est, ralenti, ces, dernières, années, .})$



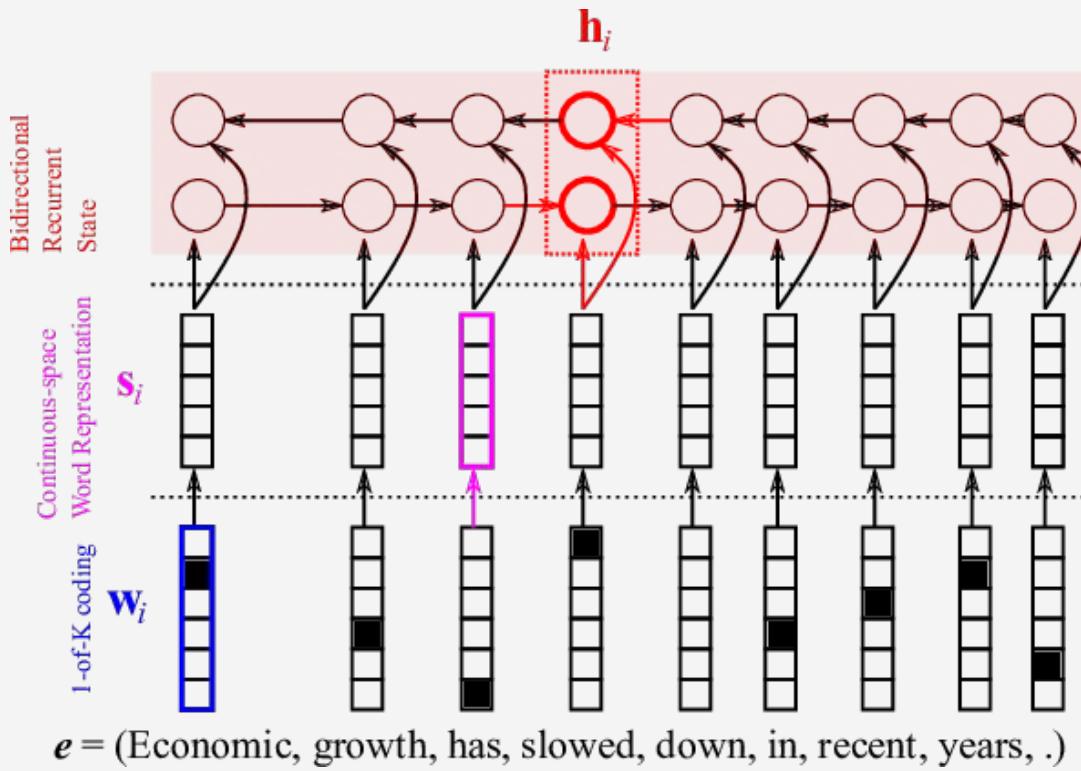
$e = (\text{Economic, growth, has, slowed, down, in, recent, years, .})$



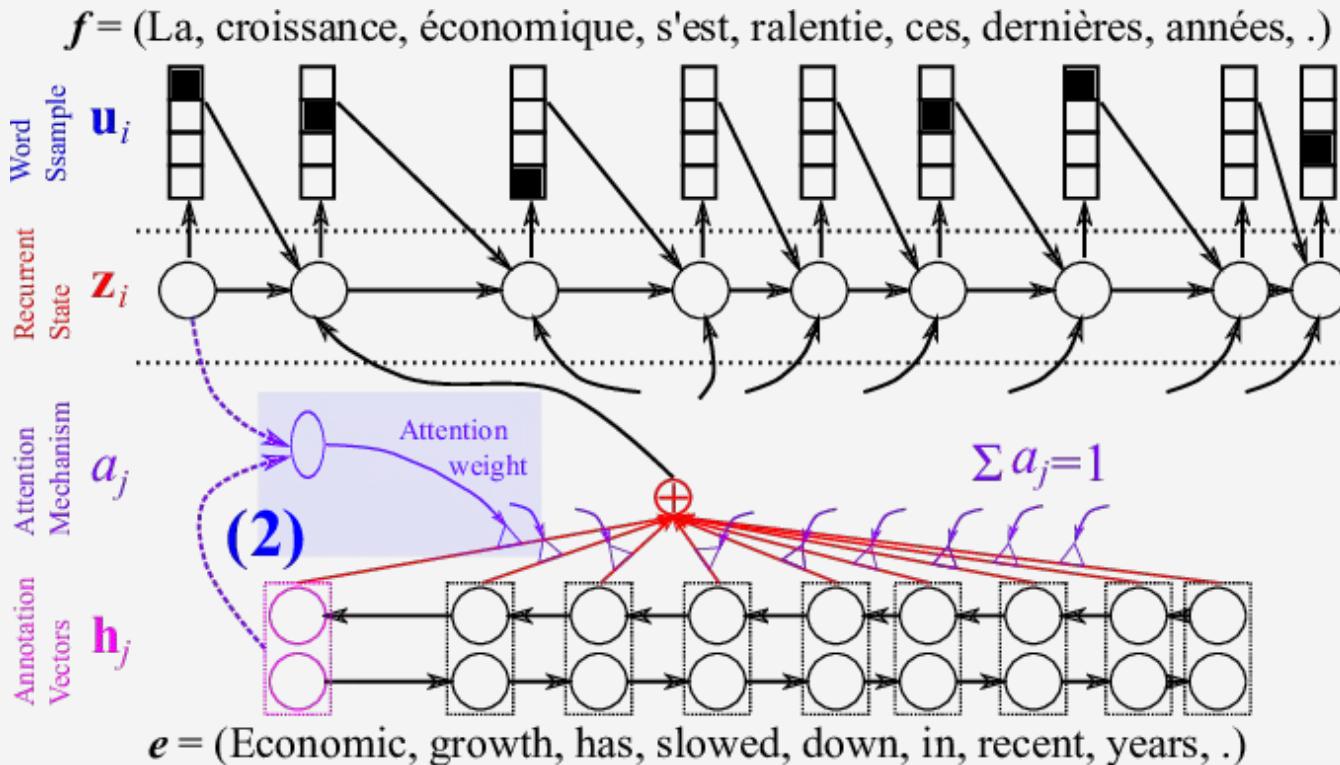
# Problemi: Lokalnost Duzina recenice



# Bidirectional Encoder



# Attention layer



# Attention example

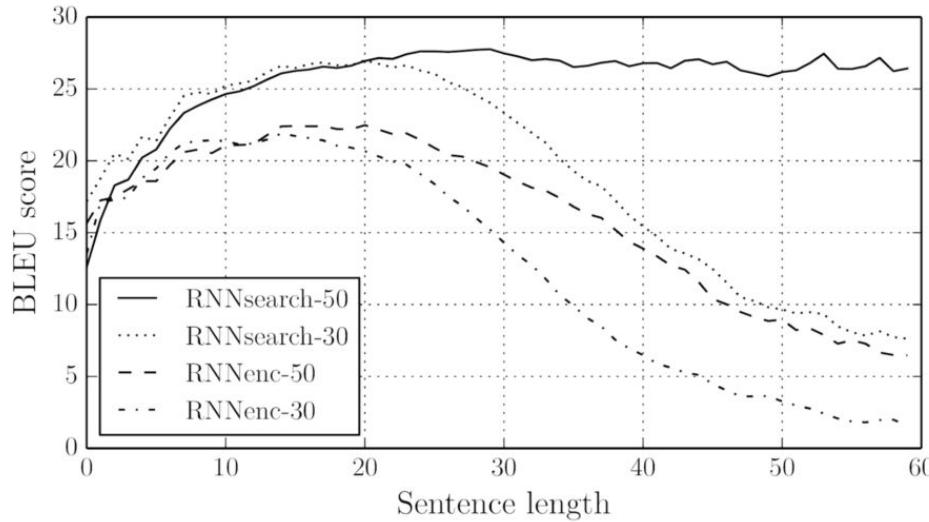
Economic growth has slowed down in recent years .

Das Wirtschaftswachstum hat sich in den letzten Jahren verlangsamt .

Economic growth has slowed down in recent years .

La croissance économique s' est ralentie ces dernières années .

# Rezultati poboljšanja



# Ne preterano praktično

za  $n$  jezika  $\sim n^*n$  različitih modela

This does not scale!

TRAINING  
TIME

2-3  
weeks per  
model

TRAINING  
DATA

100M+  
per model

MODELS

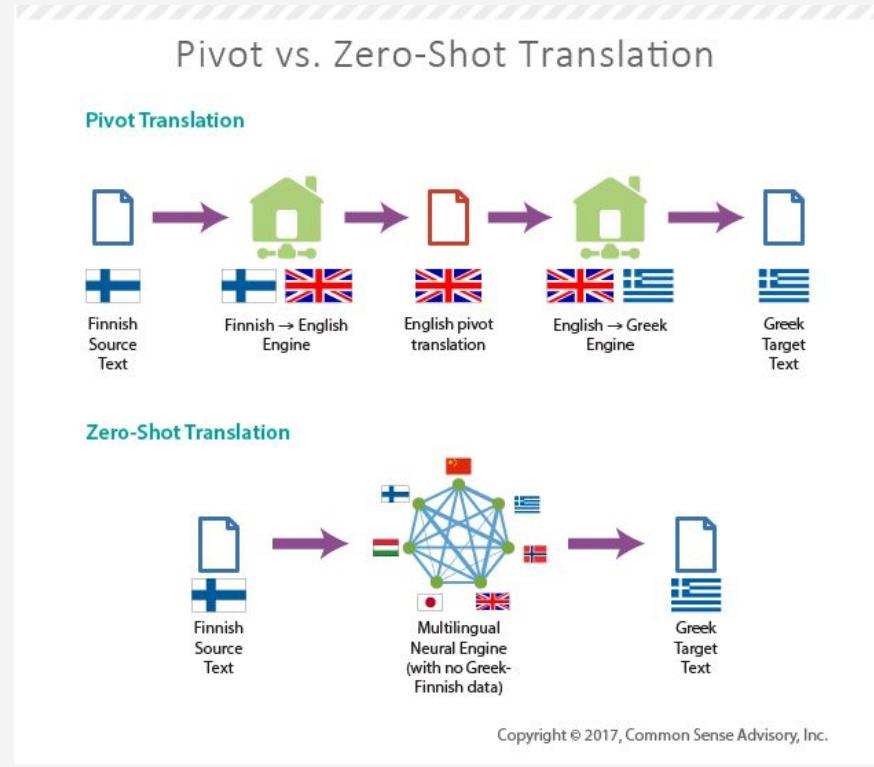
$\sim 10^{3^2}$

# Multilingual model

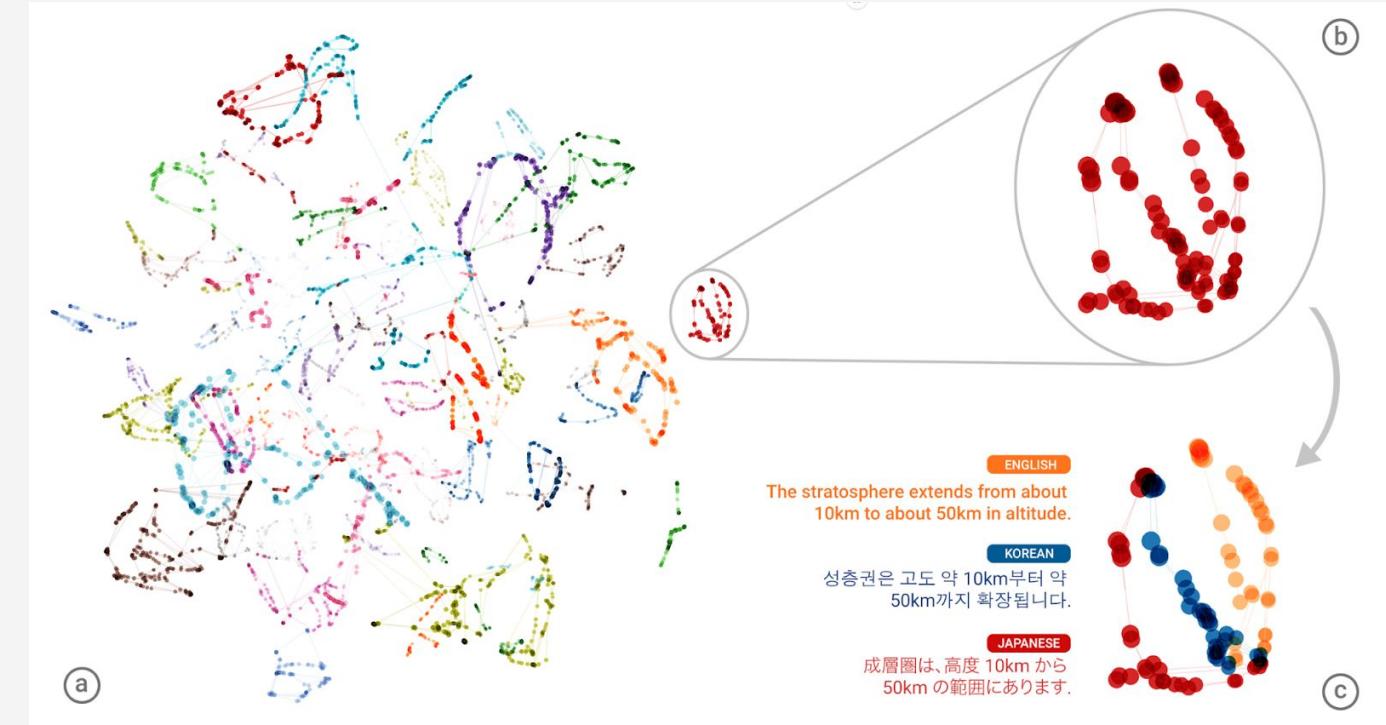
<2es> How are you </s> -> Cómo estás </s>

<2en> Cómo estás </s> -> How are you </s>

# Zero-shot translation



# Interlingua



# Vreme

Encoder-Decoder model	2014
Attention based model	2014
Multilingual solutions	2016 - 2017

**Hvala na paznji!**

**Pitanja?**