# Learning to Understand Languages with Neural Networks



#### Jörg Tiedemann

Department of Digital Humanities University of Helsinki jorg.tiedemann@helsinki.fi







#### From explanations

- "It's something to drink"
- "... for adults, you know ..."





#### From explanations

- "It's something to drink"
- "... for adults, you know ..."

#### Grounding

- this is a bottle of gin ...
- this is someone who drank gin ...







#### From explanations

- "It's something to drink"
- "... for adults, you know ..."

#### Grounding

- this is a bottle of gin ...
- this is someone who drank gin ...

### Reinforcement

- "Give me that bottle!"
- "This is alcohol! ... nothing for you, man!







# Learning from multilingual data

#### Translations as explanations

- each translation gives a different view
- translations are like naturally occurring reformulations

#### Grounding through multimodality

- add visual features
- add sound



Methods for Managing Audiovisual Data

#### **Reinforce semantics**

- supervised training objectives that require understanding
- machine translation and other tasks





# Machine translation and meaning

level of understanding



# Semantic representation learning



(From Sutskever et. alet. al: "Sequence to Sequence Learning with Neural Networks")





















## **Multilingual MT & Zero-Shot Translation**





## **Multilingual MT & Zero-Shot Translation**



IELSINGFORS UNIVERSITET



## **Multilingual MT & Zero-Shot Translation**





# **Multilingual MT & Paraphrasing**



#### Fix your language!



Valsch geschreibt is nich gut! Das Pferd hat gelaufen. Ich haben fertig. Wir sein kommen. wat morkelst du denn da rum? Icke geb dir dann och noch wat zu trinken. Dat is nix für meinereiner! Mein Fuß ist brechen! Ich muss nach die dokter.

Falsch geschrieben ist nicht gut! Das Pferd ist gelaufen. Ich bin fertig. Wir kommen. Was hast du denn da zu suchen? Dann gebe ich dir noch etwas zu trinken. Das ist nichts für mich! Mein Fuß ist gebrochen! Ich muss zum Arzt.

#### https://translate.ling.helsinki.fi/fix\_language



# **Multilingual MT & Paraphrasing**



#### Fix your language!

detect language	Afrikaans	Danish I	Outch Cat	alan Esto	onian Finnis	h Faroes	French
Frisian Galicia	n German	Hungaria	Icelandi	c Italian	Norwegian	Occitan	Portuguese
Spanish Swedi	sh						



#### https://translate.ling.helsinki.fi/fix\_language



### What does the system attend to?

we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties .

size of (a) k = 1 bridge

we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties . we cannot afford to lose more of the momentum that existed at the beginning of the N\_ ine\_ ties .

(b) 
$$k = 10$$







![](_page_20_Picture_3.jpeg)

https://blogs.helsinki.fi/fiskmo-project/

### http://blogs.helsinki.fi/language-technology/

audiovisual data & MT in 6 languages

![](_page_20_Picture_7.jpeg)

![](_page_20_Picture_8.jpeg)

Methods for Managing Audiovisual Data https://memad.eu

![](_page_20_Picture_10.jpeg)

semantics & MT > 1,000 languages

![](_page_20_Picture_13.jpeg)