

# Beyond the Benchmarks: Linguistically-oriented analysis and generalisations in Neural Networks

Paola Merlo

Computational Learning and Computational Linguistics group (CLCL)

University of Geneva and University of Upsala

R&D course, Fall 2021



**UNIVERSITÉ  
DE GENÈVE**

FACULTÉ DES LETTRES

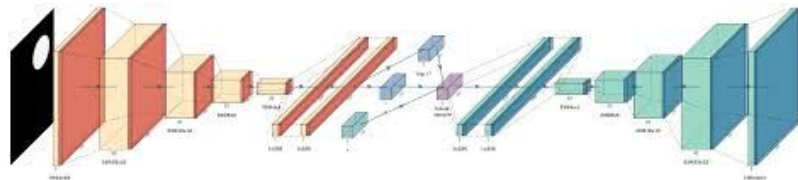


**UNIVERSITÉ  
DE GENÈVE**

CENTRE UNIVERSITAIRE  
D'INFORMATIQUE

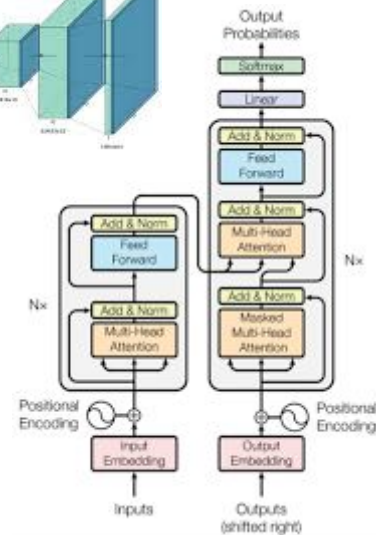
Computational Learning  
and Computational Linguistics  
RESEARCH GROUP

All speakers can understand a sentence never heard before, or derive the meaning of a word from its parts, and children can learn any language.



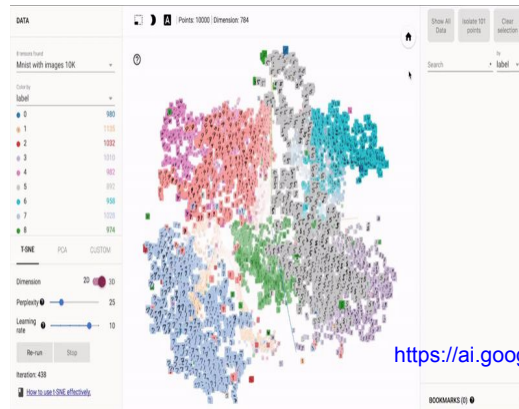
Higgins et al 2017

And yet, these basic linguistic skills have proven very hard to reach by computational models.



Waswani et al 2017

Successes of machine learning architectures are based on computationally expensive algorithms and prohibitively large amounts of data, in fact so prohibitive that they can be deployed in only a few non-representative languages.



<https://ai.googleblog.com/>

# Some properties of word embeddings, neural networks and how we study them

- **Lexical similarity: Monolingual** associations

Rodriguez and Merlo, CoNLL 2020

**Multilingual:** transfer effects in a multilingual lexicon.

Merlo and Rodriguez, CoNLL 2019,

- Linguistically-informed notion of **syntactic similarity**: locality effects, also across languages.

Merlo and Ackermann, CoNLL 2018; Merlo, BBNL 2019; Merlo, Renaud and An, in prep

- Other properties: **structural long distance agreement, verb alternations** (Linzen, Baroni, Gulordava et al, , Thrush et al, see reading list)

# Word associations in monolingual spaces

- Rank
- Asymmetry of similarity
- Violation of the triangle inequality (or lack of transitivity)

Amos Tversky. 1977. Features of similarity. *Psychological review*, 84(4):327–352.  
T.L Griffiths, M. Steyvers, and J.B Tenenbaum. 2007. Topics in semantic representation. *Psychological review*, 114(2):211–244.

# Data

- **Word embeddings:**

- BERT (Devlin et al., 2018):  
in context and not in context

- **Word associations:** The University of South Florida free association, rhyme, and word fragment norms (Nelson et al., 2004)

Cue	Targets
abdomen	stomach, belly, organ, body, muscle
yell	shout, scream, whisper, loud, cheer
saw	see, hammer, look, cut, tool
rise	lift, fall, stand, wake, shine
necessary	important, need, must, money, object

# Association rank

## Data

**Word embeddings:** BERT (Devlin et al., 2018): in context and not in context

**Word associations:** The University of South Florida free association, rhyme, and word fragment norms (Nelson et al., 2004)

Rank	Unlemmatized		Lemmatized	
	Median Rank	P@K (%)	Median Rank	P@K (%)
1	4	13.02	3	24.18
2	12	28.09	10	43.26
3	35	43.64	27	55.86
4	94	53.59	69	64.53
5	230	61.62	157.5	69.94

	YELL	ABDOMEN	SAXOPHONE	SAW	RISE	NECESSARY
Human associations	<b>shout (1)</b> <b>scream (5)</b> whisper (27) loud (189) cheer (194)	<b>stomach (1)</b> <b>belly (4)</b> organ (3399) body (4418) muscle (8368)	<b>trombone (2)</b> <b>clarinet (3)</b> trumpet (6) flute (8) guitar (10)	<b>see (2)</b> hammer (207) look (239) cut (294) tool (350)	lift (32) fall (37) stand (38) wake (72) shine (73)	important (22) need (27) must (263) money (11869) object (13096)
BERT predictions	<b>shout</b> yells yelled yelling <b>scream</b>	<b>stomach</b> abdominal torso <b>belly</b> groin	saxophonist <b>trombone</b> <b>clarinet</b> banjo harmonica	sees <b>see</b> seen seeing Saw	rises risen rising Rise rose	required needed essential unnecessary appropriate

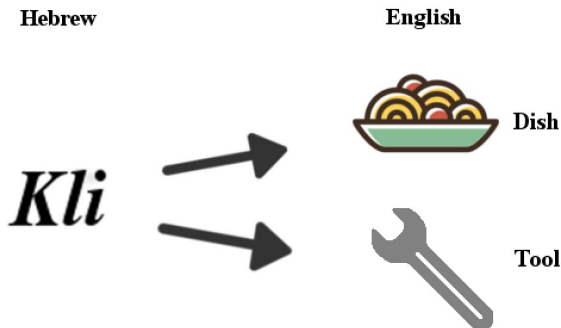
**Do cross-lingual word embeddings  
have the same structure as the  
bilingual lexicon?**

# Shared translation effects

translation pairs      shared translation pairs

wood-legno  
wood-bosco      legno bosco

block-blocco      blocco ceppo  
block-ceppo      blocco bloccare  
block-bloccare      blocco ostacolare  
block-ostacolare      ceppo bloccare



Cross-lingual models show higher mean similarity scores for L2-words that share a common L1 source than the monolingual model ( $p < 0.021$ ).



## Beyond the lexicon: similarity and locality in syntax

Neural networks work in practice, but do they work in theory?  
(Steedman, LTA 2018)

# Similarity and locality: Object relatives intervention and number

## Object relatives, singular (MATCH)

Jules sourit à l' **étudiant** que l'**orateur** endort <étudiant> sérieusement depuis le début.

*Jules smiles to **the student** that **the speaker** is putting seriously to sleep from the beginning.*

## Object relatives, plural (MISMATCH)

Jules sourit **aux étudiants** que l'**orateur** endort <étudiants> sérieusement depuis le début.

*Jules smiles to **the students** that **the speaker** is putting seriously to sleep from the beginning.*

## Completive, singular

Jules signale à l' **étudiant** que l'**orateur** baille sérieusement depuis le début.

*Jules points out to **the student** that **the speaker** is yawning seriously from the beginning.*

## Completive, plural

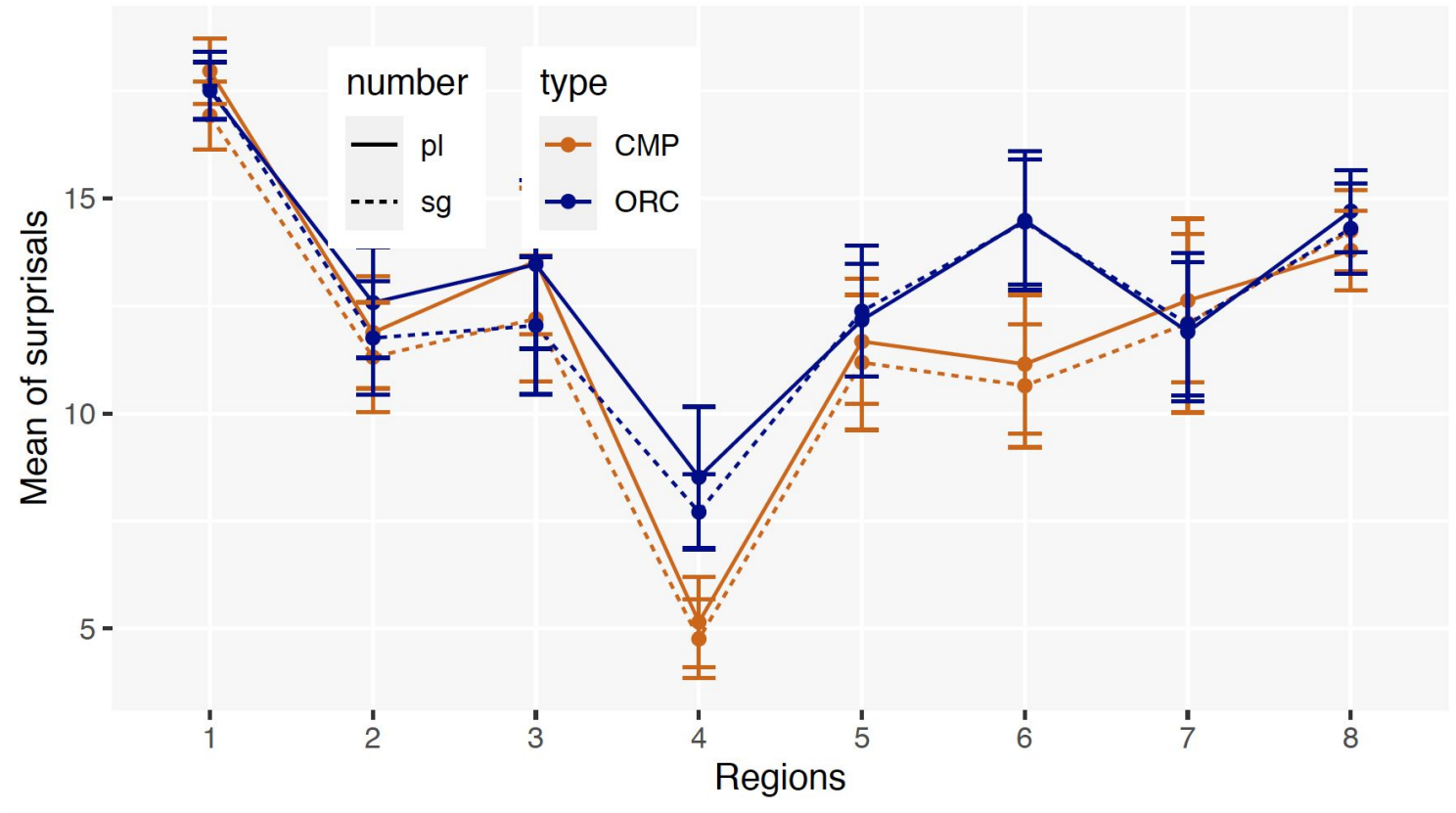
Jules signale **aux étudiants** que l'**orateur** baille sérieusement depuis le début.

*Jules smiles to **the students** that **the speaker** is yawning seriously from the beginning.*

Data kindly provided to us by Sandra Villata and Julie Franck in French. Our translations in Italian and English.

Julie Franck, S. Colonna S., and Luigi Rizzi. 2015. Task-dependency and structure dependency in number interference effects in sentence comprehension. *Frontiers in Psychology*, 6.

# FRENCH



$S(\text{ORC}) > S(\text{CMP})$  ✓  
 $S(\text{sg}) == S(\text{pl})$  ✓ ✗  
 $S(\text{ORCsg}) > S(\text{ORCpl})$  ✗

Jules    sourit    à l'étudiant    que    l'orateur    endort    sérieusement    depuis le début.  
 signale    aux étudiants                baille

## Conclusion and further research questions

Current word embeddings have the same structure as **free association and the bilingual lexicon**.

Word embeddings and the similarity spaces they define **do not encode all properties in long-distance dependencies**: they can encode agreement but not intervention.

**What other linguistic phenomena can we also study?** Verb alternations?

**What other languages can we use to replicate the results?**

How about Swedish?

**Does the architecture we use make a difference? Does the use of attention make a difference?**

Thanks!

Any questions?