

CSCE 638 Natural Language Processing

Foundation and Techniques

Lecture 15: Multilingual NLP

Kuan-Hao Huang

Spring 2025



(Some slides adapted from Graham Neubig)

Assignment 2

- https://khhuang.me/CSCE638-S25/assignments/assignment2_0224.pdf
- Due: 3/17 11:59pm
- Submit a .zip file to Canvas
 - `submission.pdf` for the writing section
 - `submission[x].py` and `submission[x].ipynb` for the coding section
- For questions
 - Discuss on Canvas
 - Send an email to csce638-ta-25s@list.tamu.edu, don't need to CC TA or me

Course Project – Midterm Report

- Due: 4/2
- Page limit: 5 pages
- Format: [ACL style](#)
- The report should include
 - Introduction to the topic you choose
 - Related literature
 - Novelty and challenges
 - Evaluation metrics
 - The dataset, models, and approaches you use
 - Current progress and next steps
- It's a checkpoint to evaluate if you can finish the project!

Multilingual NLP

- Different languages have different linguistic properties
 - Scripts
 - Alphabets, symbols, spaces

I like this restaurant

我喜欢这家餐厅

Me gusta este restaurante

我喜歡這家餐廳

J'aime ce restaurant

나는 이 레스토랑을 좋아한다

मुझे यह रेस्तरां पसंद है

このレストランが好きです

Multilingual NLP

- Different languages have different linguistic properties
 - Scripts
 - Alphabets, symbols, spaces
 - Grammatical rules
 - SVO (Subject-Verb-Object): English, Chinese, French
 - SOV (Subject-Object-Verb): Japanese, Hindi, Korean
 - VSO (Verb-Subject-Object): Arabic, Tagalog, Irish

Value		Representation
	Subject-object-verb (SOV)	564
	Subject-verb-object (SVO)	488
	Verb-subject-object (VSO)	95
	Verb-object-subject (VOS)	25
	Object-verb-subject (OVS)	11
	Object-subject-verb (OSV)	4
	Lacking a dominant word order	189
Total:		1376

Multilingual NLP

- Different languages have different linguistic properties
 - Scripts
 - Alphabets, symbols, spaces
 - Grammatical rules
 - SVO (Subject-Verb-Object): English, Chinese, French
 - SOV (Subject-Object-Verb): Japanese, Hindi, Korean
 - VSO (Verb-Subject-Object): Arabic, Tagalog, Irish
 - Writing systems
 - Left-to-right
 - Right-to-left

I like this restaurant

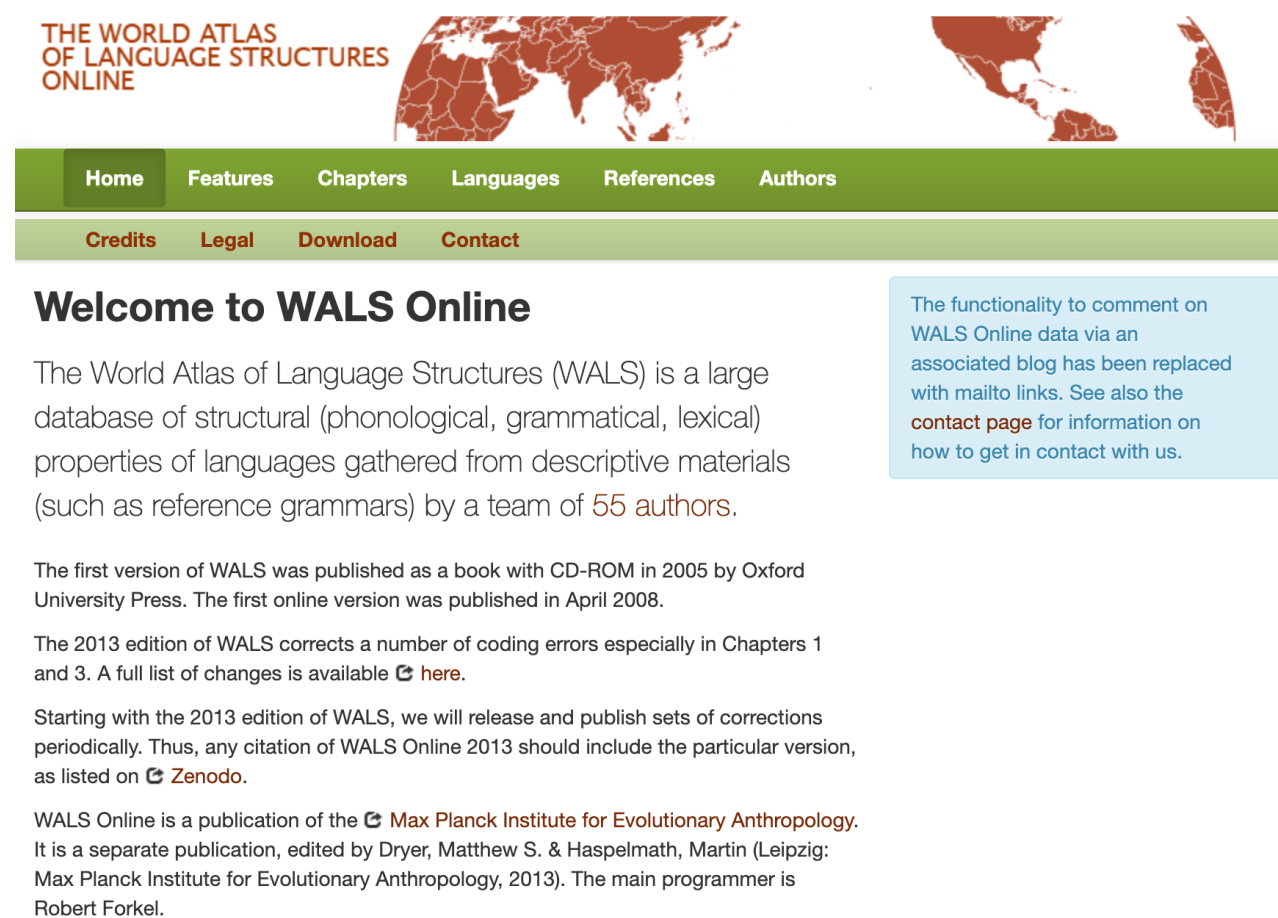
我喜歡這家餐廳

انا احب هذا المطعم

אני אוהב את המסעדה הזו

More About Language Properties

- World Atlas of Language Structures (WALS) database



The screenshot shows the homepage of the World Atlas of Language Structures (WALS) Online. At the top, there is a header with the text "THE WORLD ATLAS OF LANGUAGE STRUCTURES ONLINE" in orange, followed by a world map. Below the map is a green navigation bar with links: Home, Features, Chapters, Languages, References, and Authors. Underneath this is a lighter green bar with links: Credits, Legal, Download, and Contact. The main content area has a heading "Welcome to WALS Online" and a paragraph describing the database. To the right of this paragraph is a light blue box with text about commenting on WALS Online data. Below the main paragraph are three more paragraphs providing historical context and information about the 2013 edition and the publisher.

THE WORLD ATLAS
OF LANGUAGE STRUCTURES
ONLINE

Home Features Chapters Languages References Authors

Credits Legal Download Contact

Welcome to WALS Online

The World Atlas of Language Structures (WALS) is a large database of structural (phonological, grammatical, lexical) properties of languages gathered from descriptive materials (such as reference grammars) by a team of **55 authors**.

The first version of WALS was published as a book with CD-ROM in 2005 by Oxford University Press. The first online version was published in April 2008.

The 2013 edition of WALS corrects a number of coding errors especially in Chapters 1 and 3. A full list of changes is available [here](#).

Starting with the 2013 edition of WALS, we will release and publish sets of corrections periodically. Thus, any citation of WALS Online 2013 should include the particular version, as listed on [Zenodo](#).

WALS Online is a publication of the [Max Planck Institute for Evolutionary Anthropology](#). It is a separate publication, edited by Dryer, Matthew S. & Haspelmath, Martin (Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013). The main programmer is Robert Forkel.

The functionality to comment on WALS Online data via an associated blog has been replaced with mailto links. See also the [contact page](#) for information on how to get in contact with us.

Lang2Vec

Vector type	#Languages	#Features	#Data points	% Coverage
Syntax (from sources)				
syntax_wals	1808	98	78732	44%
syntax_sswl	230	33	6404	84%
syntax_ethnologue	1336	30	18105	45%
Syntax (averaged over sources)				
syntax_avg	2654	103	94227	34%
Syntax (predicted)				
syntax_knn	7970	103	820910	100%
Phonology (from sources)				
phonology_wals	832	27	14358	64%
phonology_ethnologue	543	8	1017	23%
Phonology (averaged over sources)				
phonology_avg	1296	28	15303	42%
Phonology (predicted)				
phonology_knn	7970	28	223160	100%
Inventory (from sources)				
inventory_phoible_aa	202	158	31916	100%
inventory_phoible_gm	428	158	67624	100%
inventory_phoible_ph	404	158	63832	100%
inventory_phoible_ra	100	158	15800	100%
inventory_phoible_saphon	334	158	52772	100%
inventory_phoible_spa	219	158	34602	100%
inventory_phoible_upsid	334	158	75050	100%
Inventory (averaged over sources)				
inventory_avg	1715	158	270970	100%
Inventory (predicted)				
inventory_knn	7970	158	1259260	100%

Lang2Vec

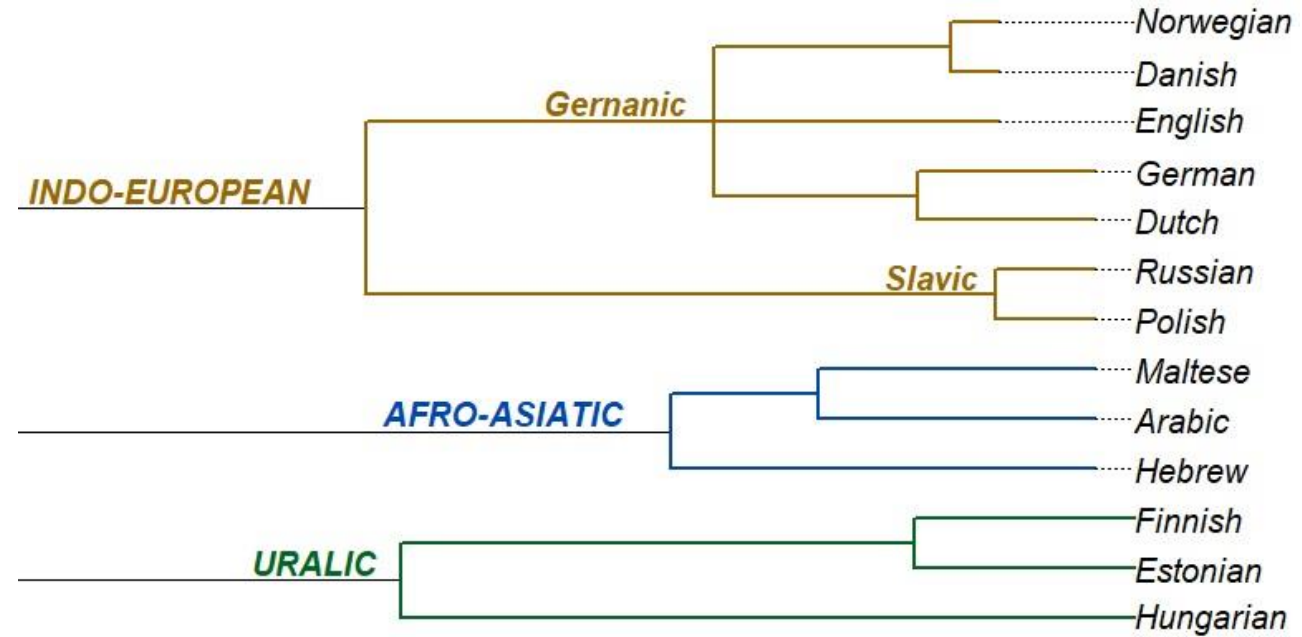
- <https://github.com/antonisa/lang2vec>

```
>>> features = l2v.get_features(["eng", "fra"], "geo")
>>> features["fra"]
[0.7378000020980835, 0.7682999968528748, 0.7982000112533569, 0.6941999793052673, ...]

>>> features = l2v.get_features("eng fr", "syntax_wals")
>>> features["eng"]
[1.0, 0.0, 0.0, 0.0, 0.0, 0.0, ...]
>>> features["fr"]
[1.0, 0.0, 0.0, 0.0, 0.0, 0.0, ...]
```

```
>>> import lang2vec.lang2vec as l2v
>>> features = l2v.get_features("eng", "geo")
>>> features["eng"]
[0.7664999961853027, 0.7924000024795532, 0.8277999758720398, 0.7214000225067139, ...]
>>>
>>> l2v.distance("syntactic", "eng", "fra")
0.4569
```

Language Hierarchy



Multilingual Tokenization is Challenging

I like this restaurant

我喜歡這家餐廳



我 喜歡 這家 餐廳

一個半小時



一個 半小時

(A half hour)

一個半小時



一個半 小時

(One and a half hour)

這幾天天天天氣不好



這幾天 天天 天氣 不好

Translation

- A sequence-to-sequence task
- Need **parallel** data

I like this restaurant

我喜欢这家餐厅

Me gusta este restaurante

我喜歡這家餐廳

J'aime ce restaurant

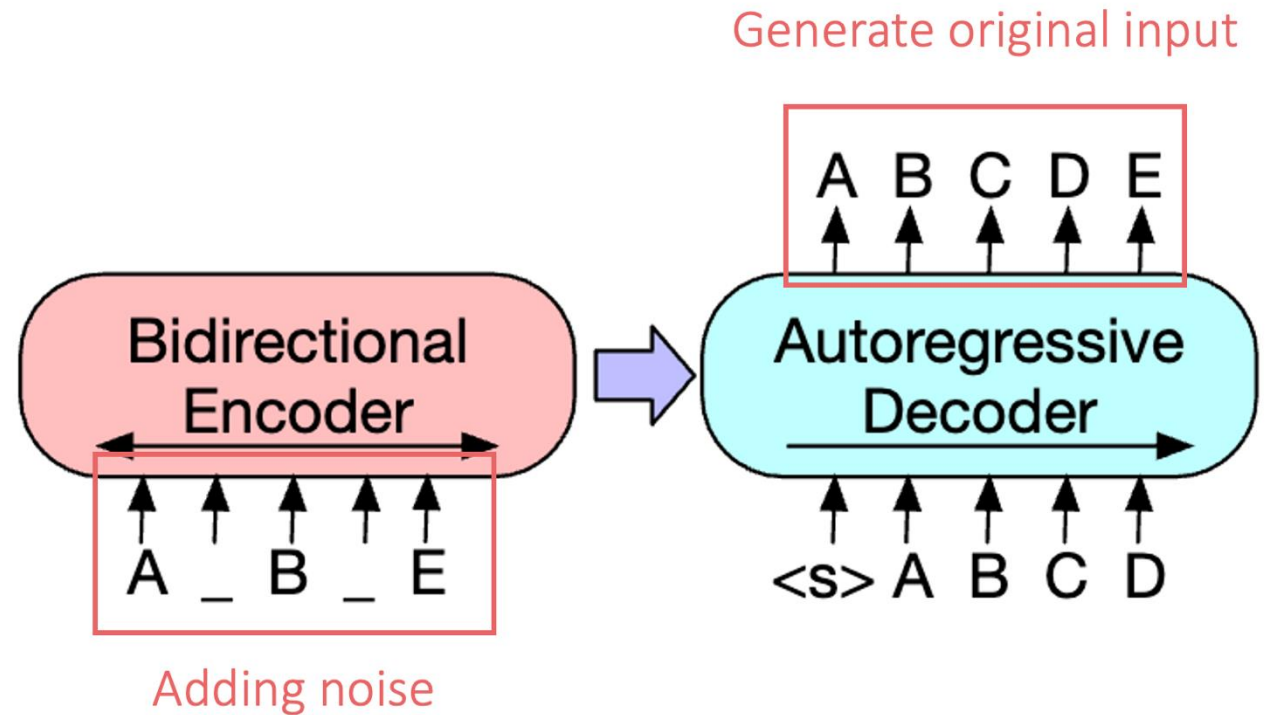
나는 이 레스토랑을 좋아한다

मुझे यह रेस्तरां पसंद है

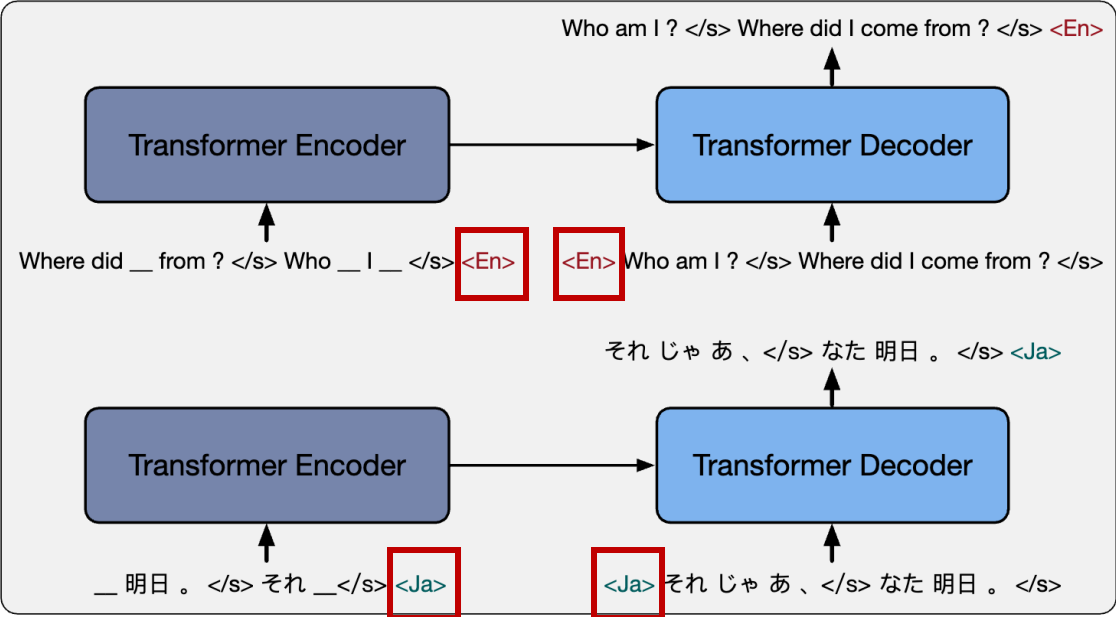
このレストランが大好きです

Recap: BART – Denoising Objective

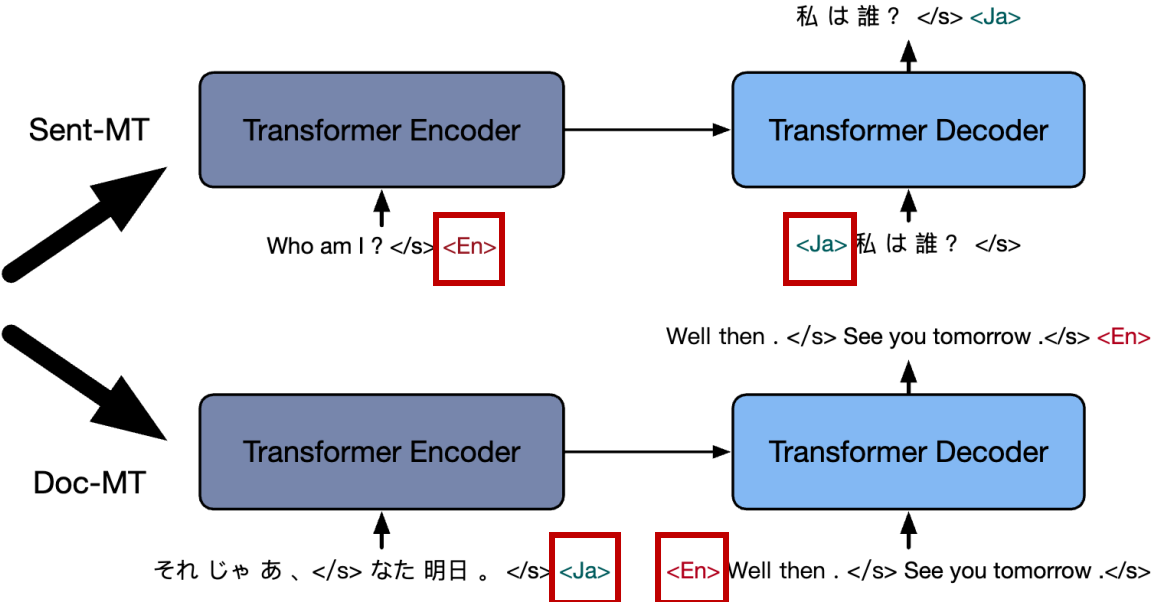
- Token Masking
 - A<mask>CD<mask>F. → ABCDEF
- Token Deletion
 - ACDF. → ABCDEF.
- Text Infilling
 - A<mask>D<mask>F. → ABCDEF.
- Sentence Permutation
 - FG. ABC. DE. → ABC. DE. FG.
- Document Rotation
 - E. FG. ABC. D → ABC. DE. FG.



Multilingual BART (mBART)



Multilingual Denoising Pre-Training (mBART)



Fine-tuning on Machine Translation

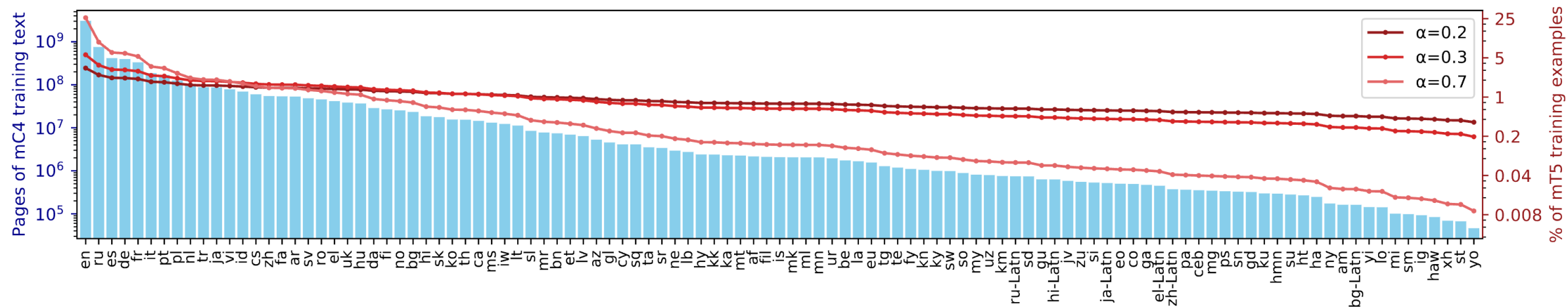
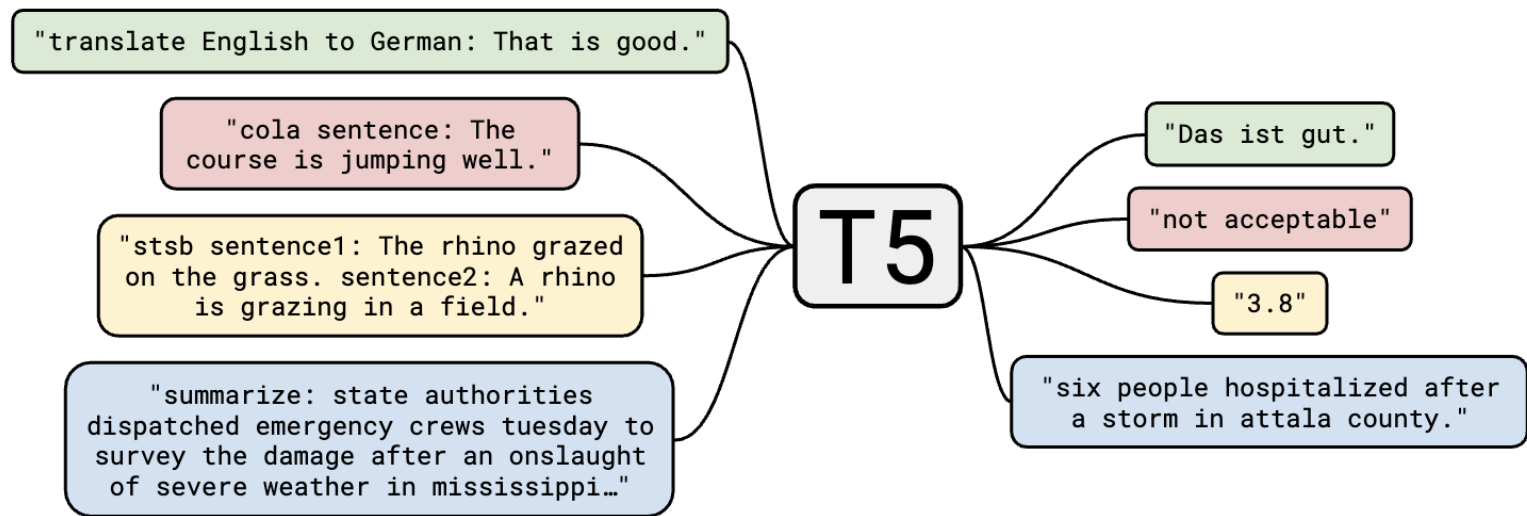
Multilingual BART (mBART)

- mBART25 and mBART50

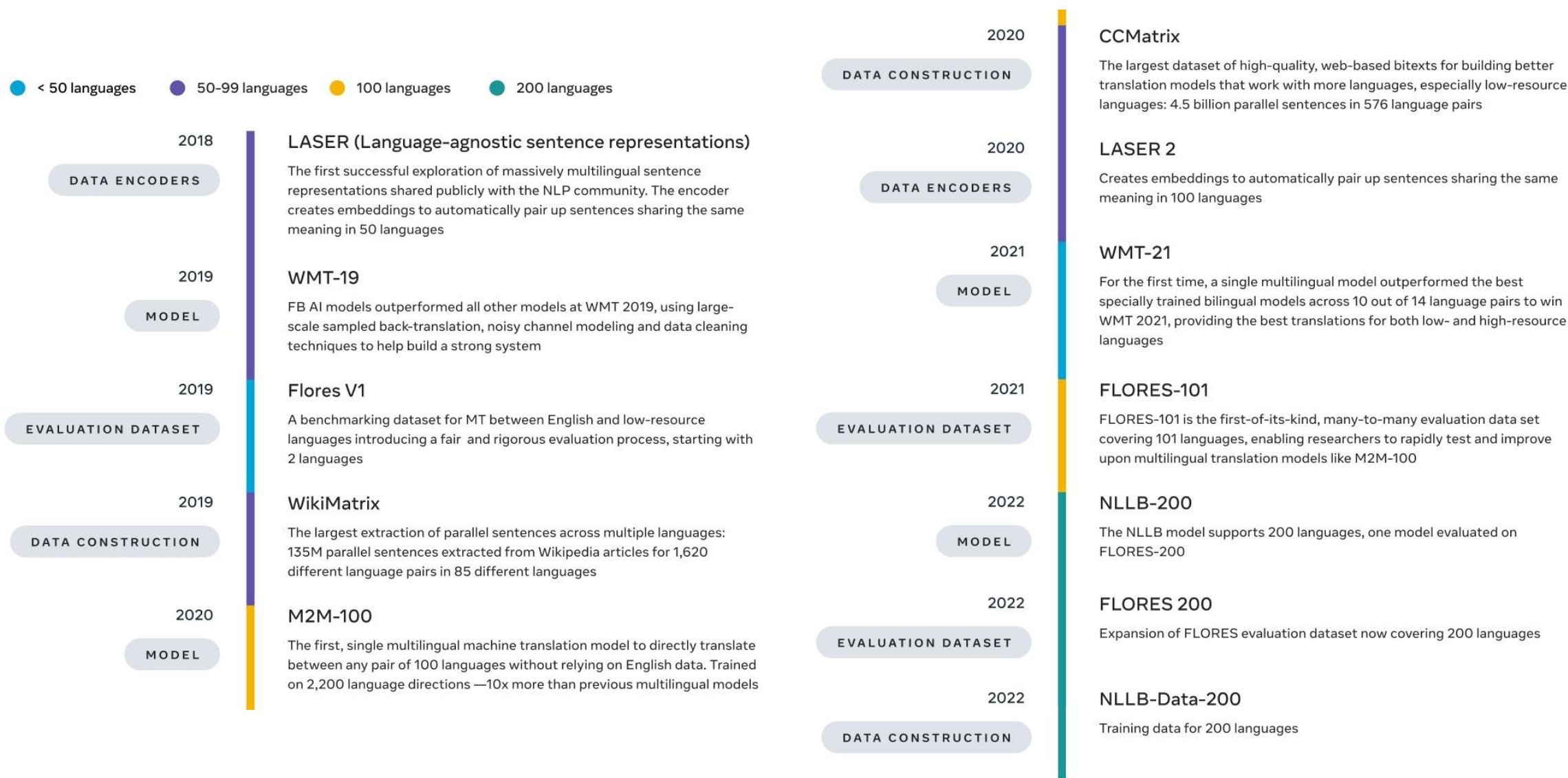
Code	Language	Tokens/M	Size/GB
En	English	55608	300.8
Ru	Russian	23408	278.0
Vi	Vietnamese	24757	137.3
Ja	Japanese	530 (*)	69.3
De	German	10297	66.6
Ro	Romanian	10354	61.4
Fr	French	9780	56.8
Fi	Finnish	6730	54.3
Ko	Korean	5644	54.2
Es	Spanish	9374	53.3
Zh	Chinese (Sim)	259 (*)	46.9
It	Italian	4983	30.2
Nl	Dutch	5025	29.3
Ar	Arabic	2869	28.0
Tr	Turkish	2736	20.9
Hi	Hindi	1715	20.2
Cs	Czech	2498	16.3
Lt	Lithuanian	1835	13.7
Lv	Latvian	1198	8.8
Kk	Kazakh	476	6.4
Et	Estonian	843	6.1
Ne	Nepali	237	3.8
Si	Sinhala	243	3.6
Gu	Gujarati	140	1.9
My	Burmese	56	1.6

Data size	Languages
10M+	German, Czech, French, Japanese, Spanish, Russian, Polish, Chinese
1M - 10M	Finnish, Latvian, Lithuanian, Hindi, Estonian
100k to 1M	Tamil, Romanian, Pashto, Sinhala, Malayalam, Dutch, Nepali, Italian, Arabic, Korean, Hebrew, Turkish, Khmer, Farsi, Vietnamese, Croatian, Ukrainian
10K to 100K	Thai, Indonesian, Swedish, Portuguese, Xhosa, Afrikaans, Kazakh, Urdu, Macedonian, Telugu, Slovenian, Burmese, Georgia
10K-	Marathi, Gujarati, Mongolian, Azerbaijani, Bengali

Multilingual T5 (mT5)



Translation: No Language Left Behind (NLLB)



Translation: No Language Left Behind (NLLB)

Acehnese (Latin script)	Crimean Tatar	Icelandic	South Azerbaijani	Fijian	Kinyarwanda
Arabic (Iraqi/Mesopotamian)	Welsh	Italian	North Azerbaijani	Finnish	Kyrgyz
Arabic (Yemen)	Danish	Javanese	Bashkir	Fon	Kimbundu
Arabic (Tunisia)	German	Japanese	Bambara	Scottish Gaelic	Konga
Afrikaans	French	Kabyle	Balinese	Irish	Korean
Arabic (Jordan)	Friulian	Kachin Jinghpo	Belarusian	Galician	Kurdish (Kurmanji)
Akan	Fulfulde	Kamba	Bemba	Guarani	Lao
Amharic	Dinka(Rek)	Kannada	Bengali	Gujarati	Latvian (Standard)
Arabic (Lebanon)	Dyula	Kashmiri (Arabic script)	Bhojpuri	Haitian Creole	Ligurian
Arabic (MSA)	Dzongkha	Kashmiri (Devanagari script)	Banjar (Latin script)	Hausa	Limburgish
Arabic (Modern Standard Arabic)	Greek	Georgian	Tibetan	Hebrew	Lingala
Arabic (Saudi Arabia)	English	Kanuri (Arabic script)	Bosnian	Hindi	Lithuanian
Arabic (Morocco)	Esperanto	Kanuri (Latin script)	Buginese	Chhattisgarhi	Lombard
Arabic (Egypt)	Estonian	Kazakh	Bulgarian	Croatian	Latgalian
Assamese	Basque	Kabiye	Catalan	Hugarian	Luxembourgish
Asturian	Ewe	Thai	Cebuano	Armenian	Luba-Kasai
Awadhi	Faroese	Khmer	Czech	Igobo	Ganda
Aymara	Iranian Persian	Kikuyu	Chokwe	Ilocano	Dholuo
			Central Kurdish	Indonesian	Mizo

Translation: No Language Left Behind (NLLB)

 **Hugging Face**



facebook/

nllb-200-3.3B



 like

296

Follow

 AI at Meta

5.08k

 Translation

 Transformers

 PyTorch

 flores-200

 196 languages

m2m_100

text2text-generation

nllb

 License: cc-by-nc-4.0



Train

Deploy

 Use this model

 **Model card**

 Files

 Community

17

NLLB-200

This is the model card of NLLB-200's 3.3B variant.

Here are the metrics for that particular checkpoint.

- Information about training algorithms, parameters, fairness constraints or other applied approaches, and features. The exact training algorithm, data and the strategies to handle data imbalances for high and low

Downloads last month

50,294



 **Inference Providers**

NEW

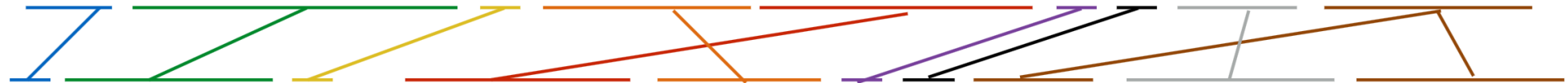
 Translation

This model is not currently available via any of the supported Inference Providers.

The model cannot be deployed to the HF Inference API: The model authors have turned it off explicitly.

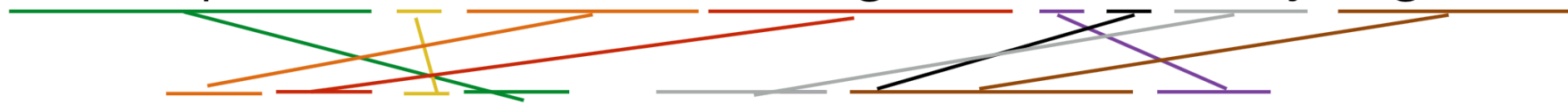
Word Alignment

The development of artificial intelligence is a really big deal.



El desarrollo de la inteligencia artificial es un asunto realmente importante.

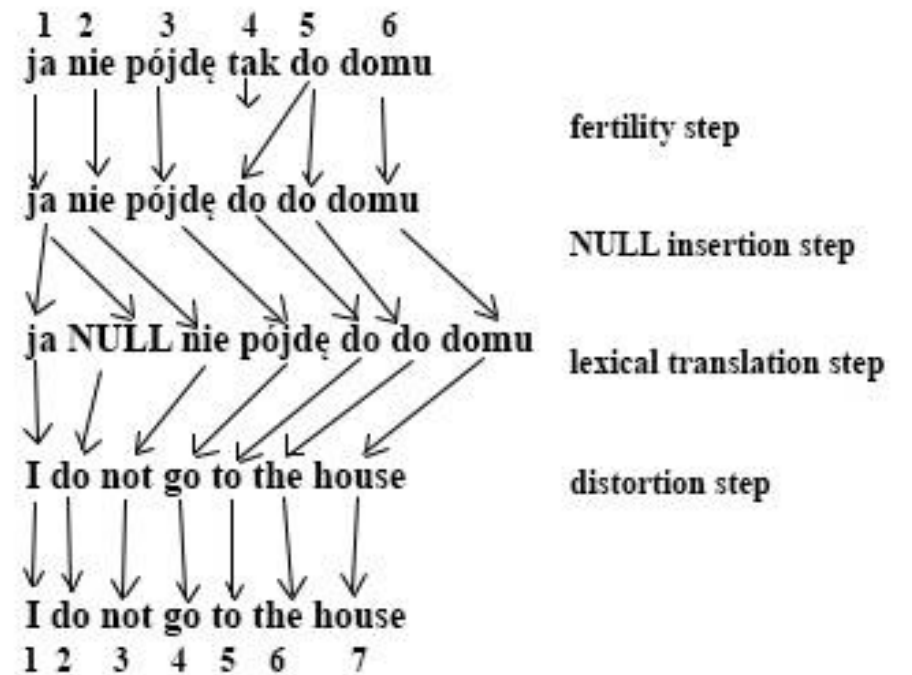
The development of artificial intelligence is a really big deal.



人工知能の発展は本当にすごいことです。

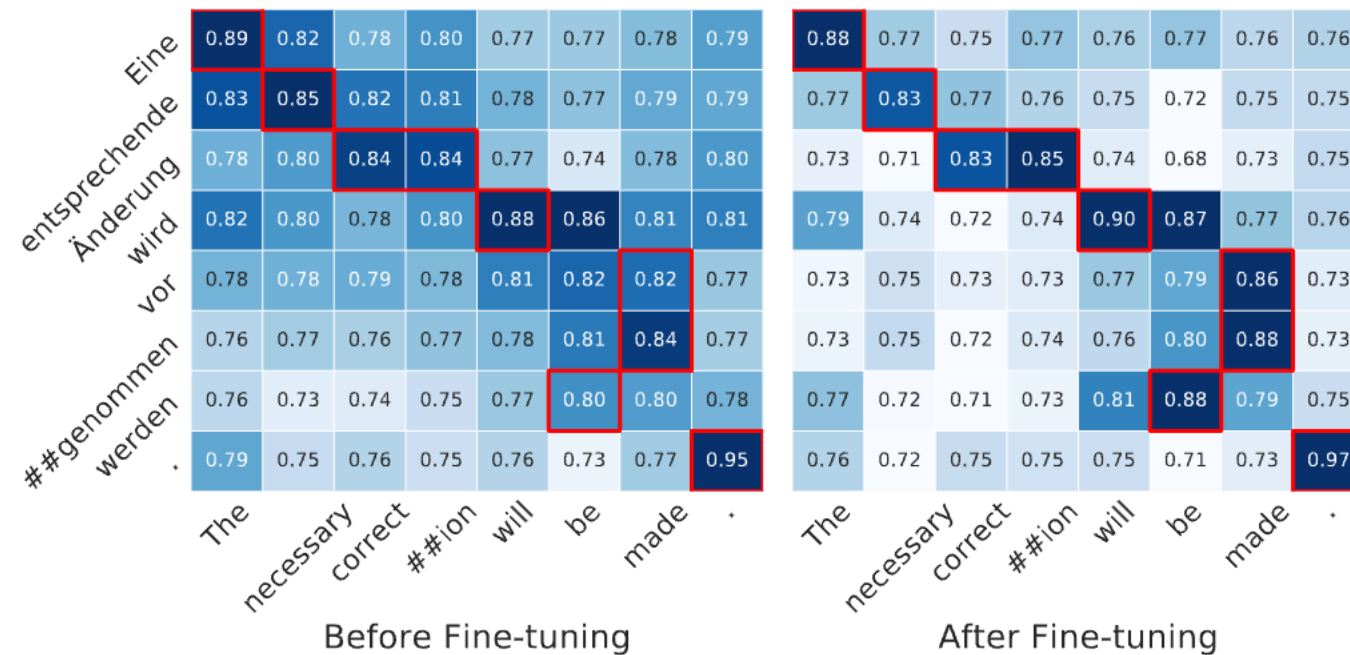
IBM Alignment Models

- Based on statistics
 - Model 1: lexical translation
 - Model 2: additional absolute alignment model
 - Model 3: extra fertility model
 - Model 4: added relative alignment model
 - Model 5: fixed deficiency problem.
 - Model 6: Model 4 combined with a HMM alignment model in a log linear way



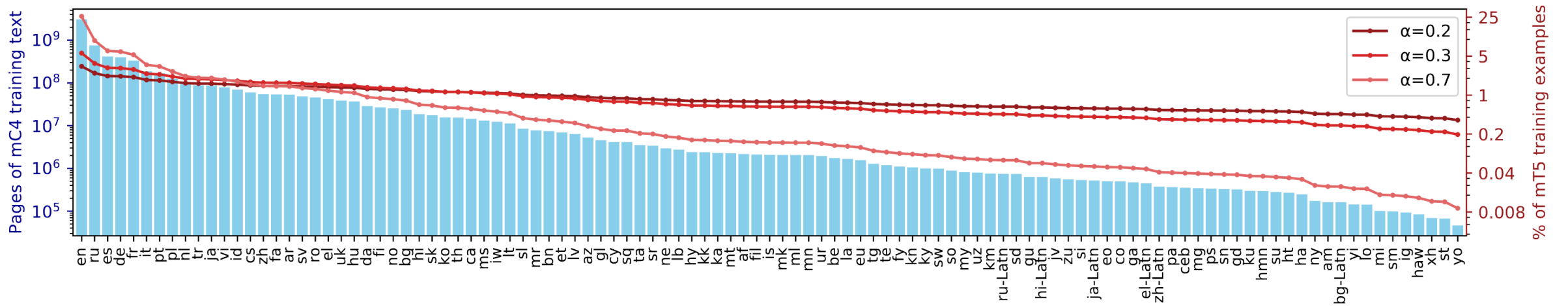
Learning-Based Alignment Models

- Awesome-Align
 - <https://github.com/neulab/awesome-align>



Cross-Lingual Transfer Learning

- Language resources are not evenly distributed



Can we transfer the learned knowledge from languages to languages?

Cross-Lingual Transfer Learning

- Training examples in a **source language**
- Testing examples in a **target language**

I like this restaurant because its food is good.



我喜欢这家餐厅，因为它的食物很好。

?

J'aime ce restaurant car sa cuisine est bonne.

?

Ich mag dieses Restaurant, weil das Essen gut ist.

?

मुझे यह रेस्टोरेंट पसंद है क्योंकि इसका खाना अच्छा है।

?

A Simple Baseline: Translation-Train

- Training examples in a **source language**
 - Translate training examples to the **target language**
 - Train the model f with translated training examples
- Testing examples in a **target language**
 - Test the model f

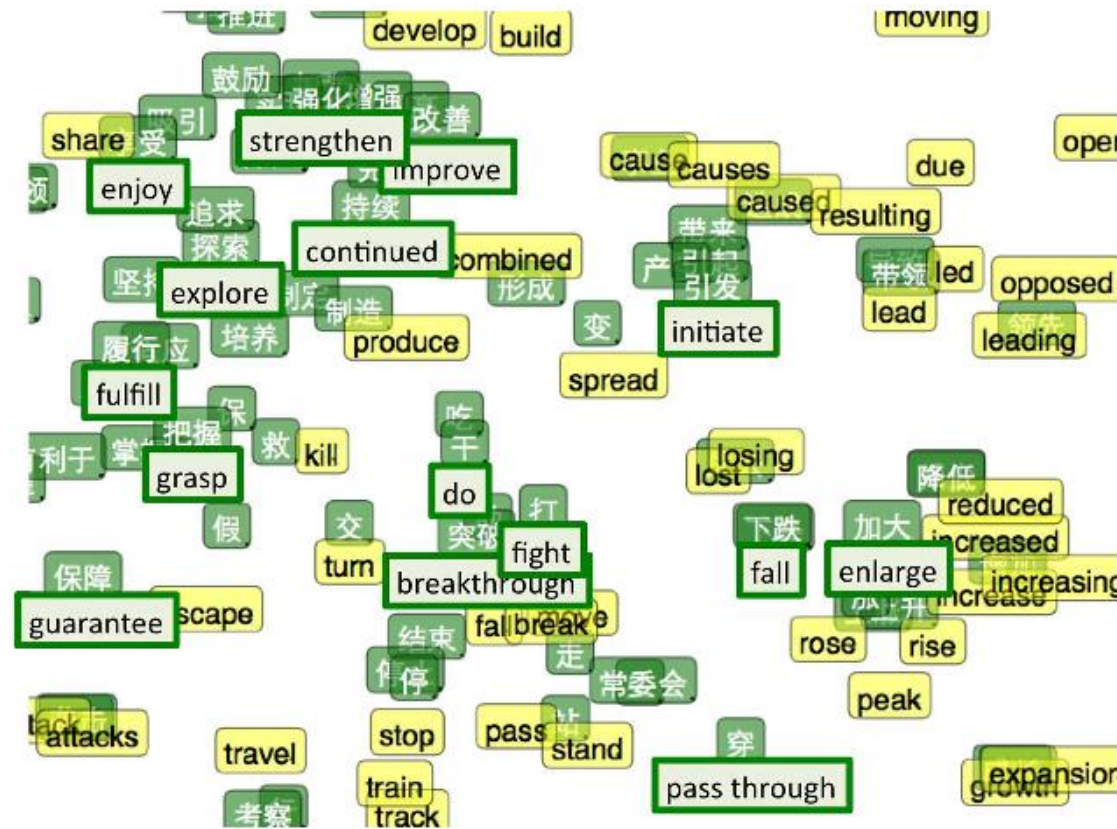
A Simple Baseline: Translation-Test

- Training examples in a **source language**
 - Train the model f with training examples
- Testing examples in a **target language**
 - Translate testing examples to the **source language**
 - Test the model f on translated examples

What if we don't have translator?

Multilingual Embedding

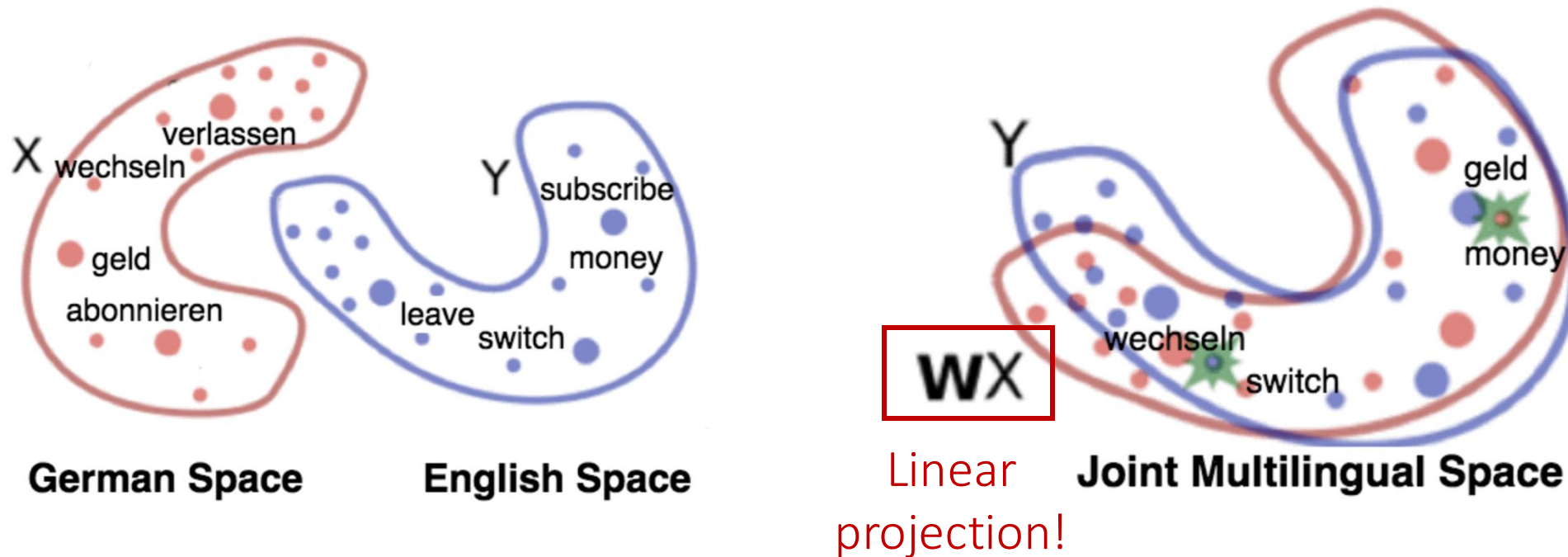
- Train the model f in multilingual embedding



How to learn multilingual embedding?

Multilingual Embedding Alignment

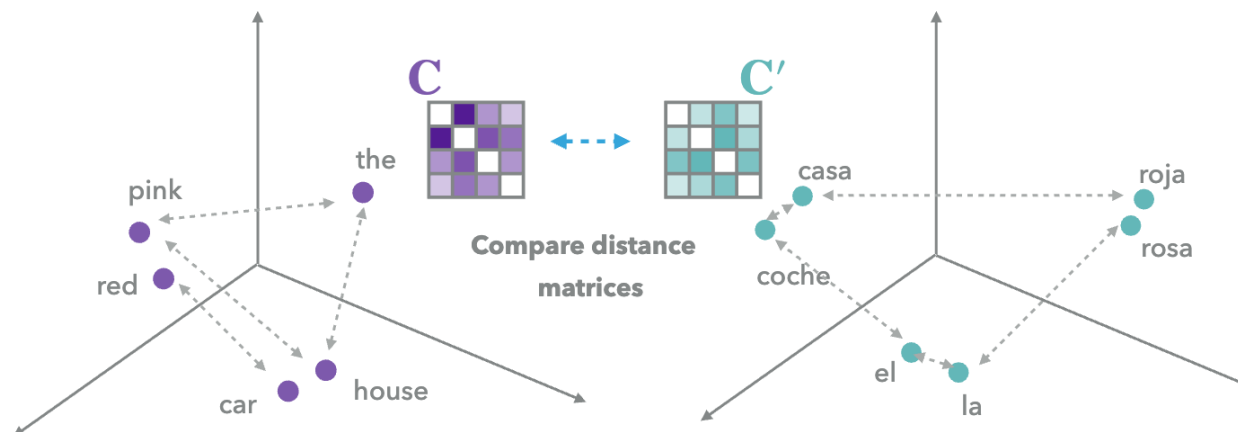
- Train word embeddings for languages separately
- Alignment different embedding spaces to one universal space



Multilingual Embedding Alignment

The Gromov-Wasserstein Distance

- Generalizes OT to the non-registered case
- Main idea: compare **distances** instead of absolute **positions**

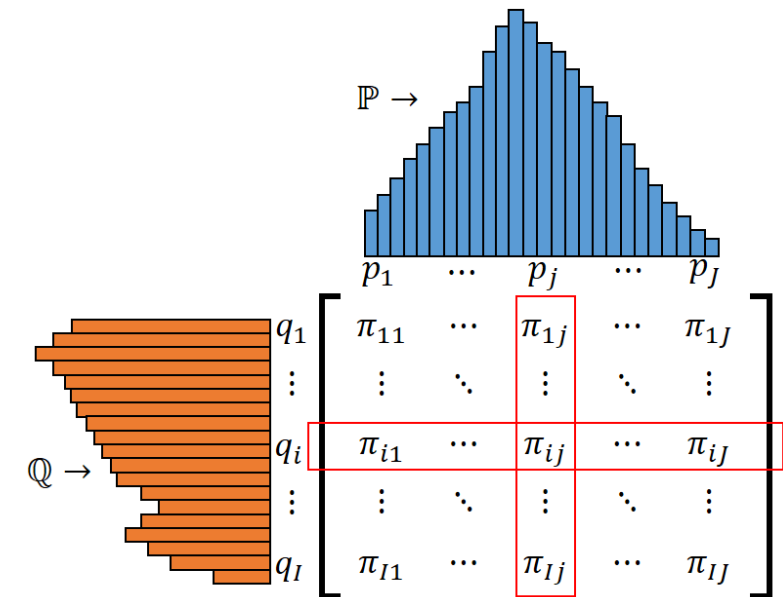
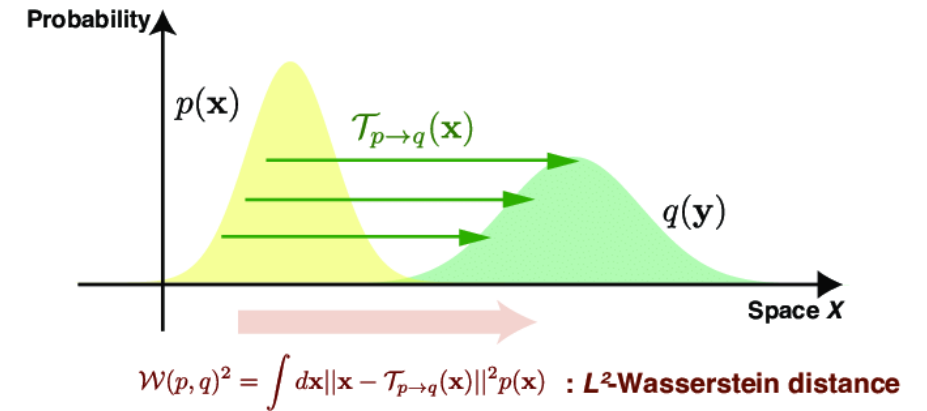


cost of matching $\mathbf{x}^{(i)}$ to $\mathbf{y}^{(j)}$ and $\mathbf{x}^{(k)}$ to $\mathbf{y}^{(l)}$ =

$$\mathcal{L}(d(\mathbf{x}^{(i)}, \mathbf{x}^{(k)}), d(\mathbf{y}^{(j)}, \mathbf{y}^{(l)}))$$

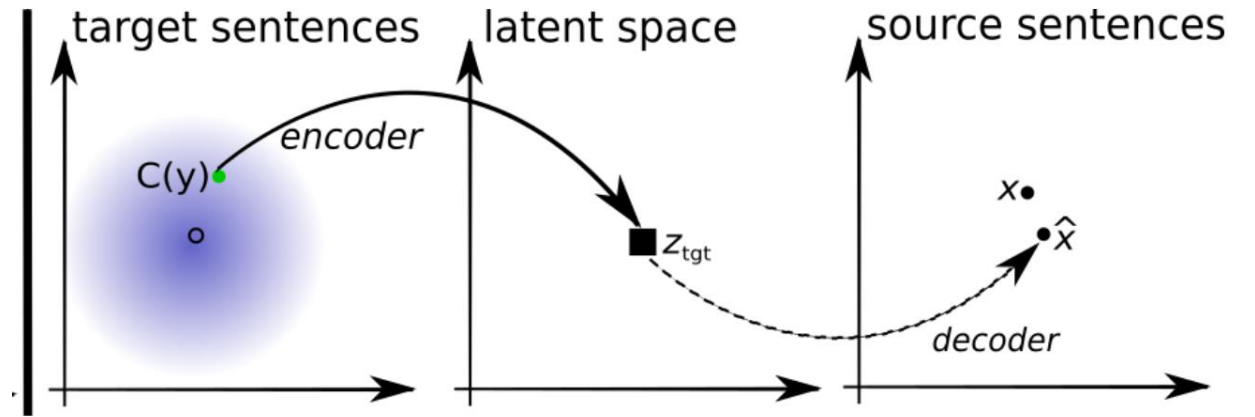
- The objective:

$$\text{GW}(\mathbf{C}, \mathbf{C}', \mathbf{p}, \mathbf{q}) = \min_{\Gamma \in \Pi(\mathbf{p}, \mathbf{q})} \sum_{i,j,k,l} \mathcal{L}(\mathbf{C}_{ik}, \mathbf{C}'_{jl}) \Gamma_{ij} \Gamma_{kl}$$



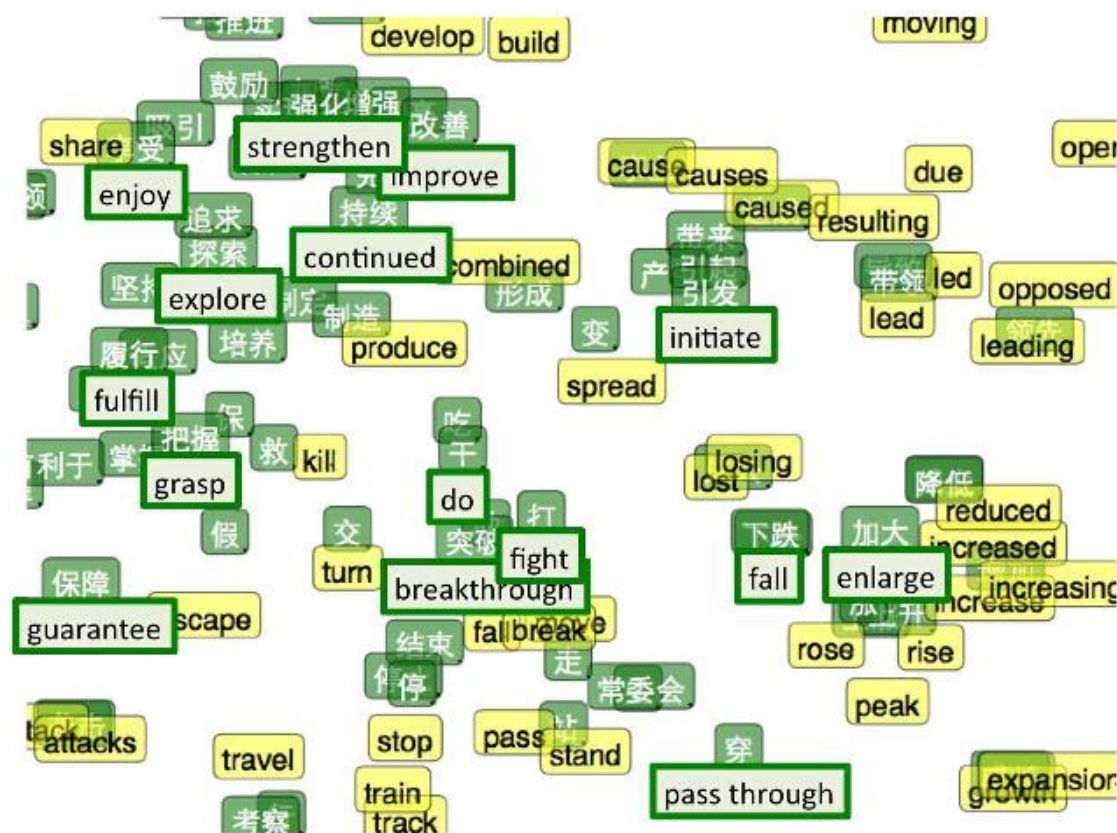
Unsupervised Machine Translation

- Unparallel corpus
 - Corpus in language 1
 - Corpus in language 2
- High-level idea
 - Unsupervised multilingual word alignment
 - Word-level translation
 - Learn syntax from corpus



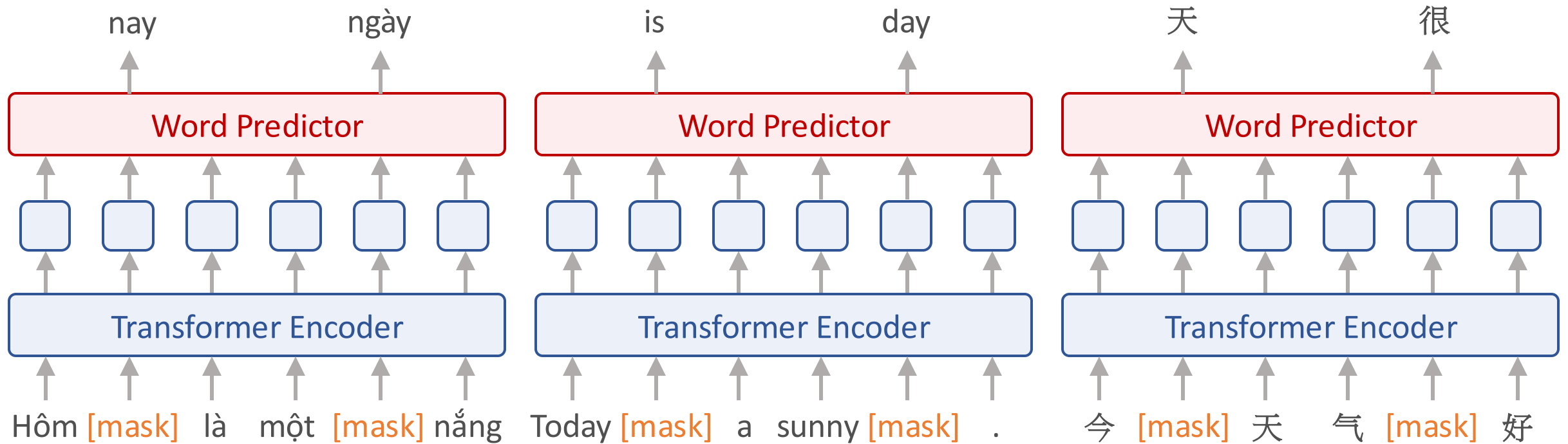
Multilingual Embedding

- Train the model f in multilingual embedding



Joint Learning for Multilingual Embedding

- BERT/RoBERTa
 - Masked language modeling



Multilingual version: mBERT / XLM-R

Impressive Cross-Lingual Transfer Performance

Model	D	#M	#lg	en	fr	es	de	el	bg	ru	tr	ar	vi	th	zh	hi	sw	ur	Avg
<i>Fine-tune multilingual model on English training set (Cross-lingual Transfer)</i>																			
Lample and Conneau (2019)	Wiki+MT	N	15	85.0	78.7	78.9	77.8	76.6	77.4	75.3	72.5	73.1	76.1	73.2	76.5	69.6	68.4	67.3	75.1
Huang et al. (2019)	Wiki+MT	N	15	85.1	79.0	79.4	77.8	77.2	77.2	76.3	72.8	73.5	76.4	73.6	76.2	69.4	69.7	66.7	75.4
Devlin et al. (2018)	Wiki	N	102	82.1	73.8	74.3	71.1	66.4	68.9	69.0	61.6	64.9	69.5	55.8	69.3	60.0	50.4	58.0	66.3
Lample and Conneau (2019)	Wiki	N	100	83.7	76.2	76.6	73.7	72.4	73.0	72.1	68.1	68.4	72.0	68.2	71.5	64.5	58.0	62.4	71.3
Lample and Conneau (2019)	Wiki	1	100	83.2	76.7	77.7	74.0	72.7	74.1	72.7	68.7	68.6	72.9	68.9	72.5	65.6	58.2	62.4	70.7
XLM-R_{Base}	CC	1	100	85.8	79.7	80.7	78.7	77.5	79.6	78.1	74.2	73.8	76.5	74.6	76.7	72.4	66.5	68.3	76.2
XLM-R	CC	1	100	89.1	84.1	85.1	83.9	82.9	84.0	81.2	79.6	79.8	80.8	78.1	80.2	76.9	73.9	73.8	80.9
<i>Translate everything to English and use English-only model (TRANSLATE-TEST)</i>																			
BERT-en	Wiki	1	1	88.8	81.4	82.3	80.1	80.3	80.9	76.2	76.0	75.4	72.0	71.9	75.6	70.0	65.8	65.8	76.2
RoBERTa	Wiki+CC	1	1	91.3	82.9	84.3	81.2	81.7	83.1	78.3	76.8	76.6	74.2	74.1	77.5	70.9	66.7	66.8	77.8
<i>Fine-tune multilingual model on each training set (TRANSLATE-TRAIN)</i>																			
Lample and Conneau (2019)	Wiki	N	100	82.9	77.6	77.9	77.9	77.1	75.7	75.5	72.6	71.2	75.8	73.1	76.2	70.4	66.5	62.4	74.2

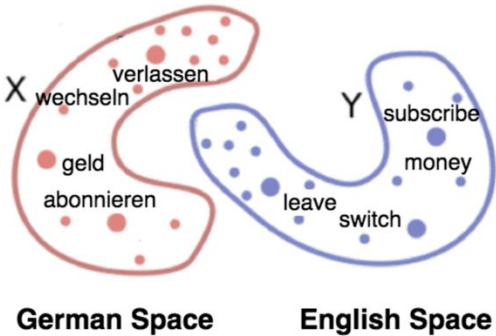
Why Work?

- Anchor word (word-piece overlap)?
 - I like this restaurant
 - J'aime ce restaurant
 - Me gusta este restaurante

Fake-English Test

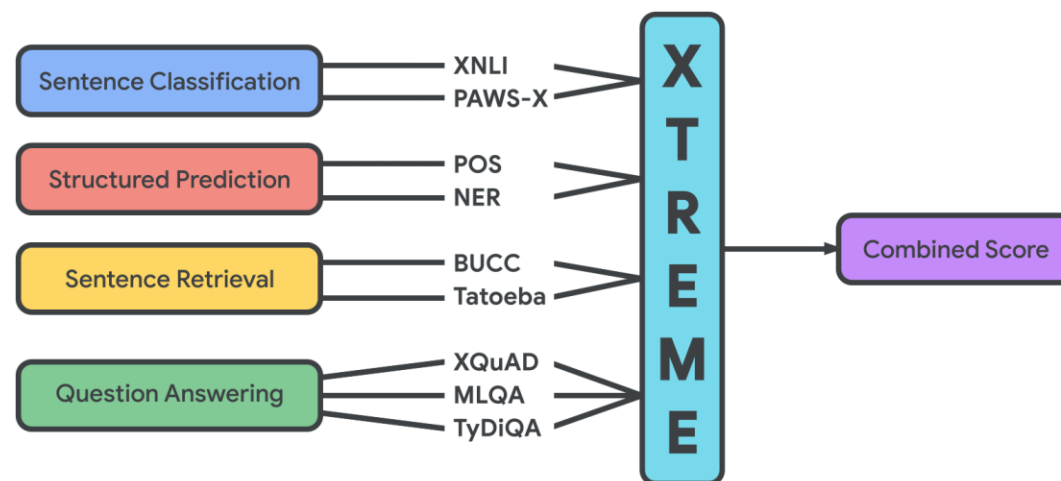
- Fake-English (enfake)
 - A different language than English, but having the exact same properties except word surface forms

B-BERT	Train	Test	XNLI		NER
			Accuracy	Word-piece Contribution	F ₁ -Score
en-es	en	es	72.3	1.4	61.9 (± 0.8)
enfake-es	enfake		70.9		62.6 (± 1.6)
en-hi	en	hi	60.1	0.5	61.6 (± 0.7)
enfake-hi	enfake		59.6		62.9 (± 0.7)
en-ru	en	ru	66.4	0.7	57.1* (± 0.9)
enfake-ru	enfake		65.7		54.2 (± 0.7)
en-enfake	enfake	enfake	78.0	0.5	78.9* (± 0.7)
en-enfake	enfake	en	77.5		76.6 (± 0.8)



The contribution of word-piece overlap is small!

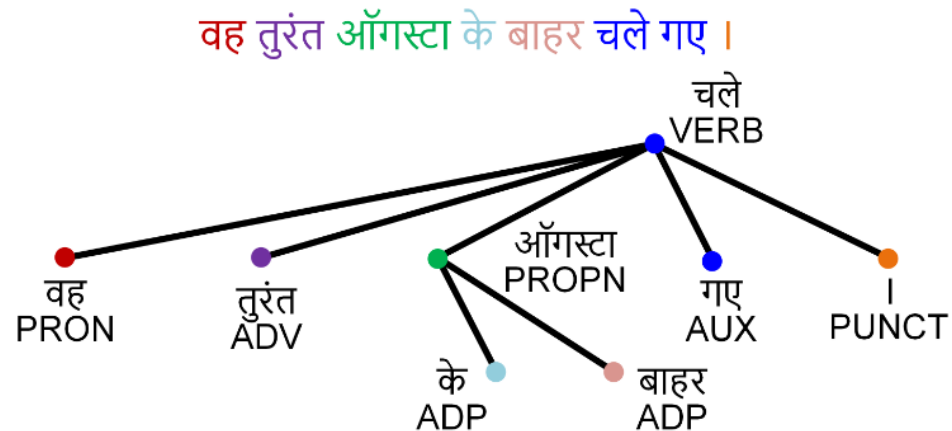
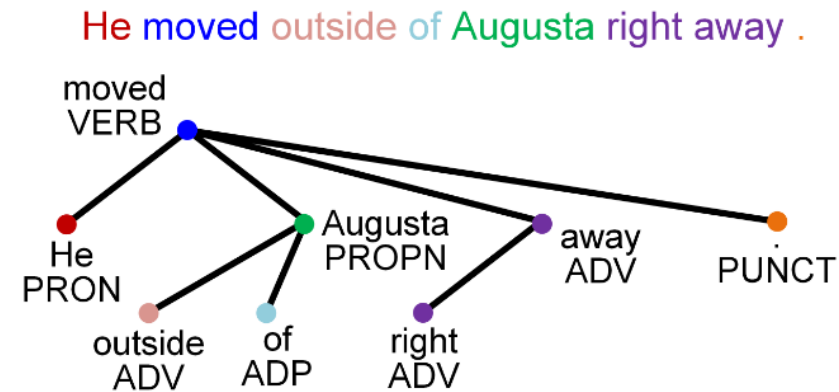
XTREME: Benchmark for Cross-Lingual Transfer



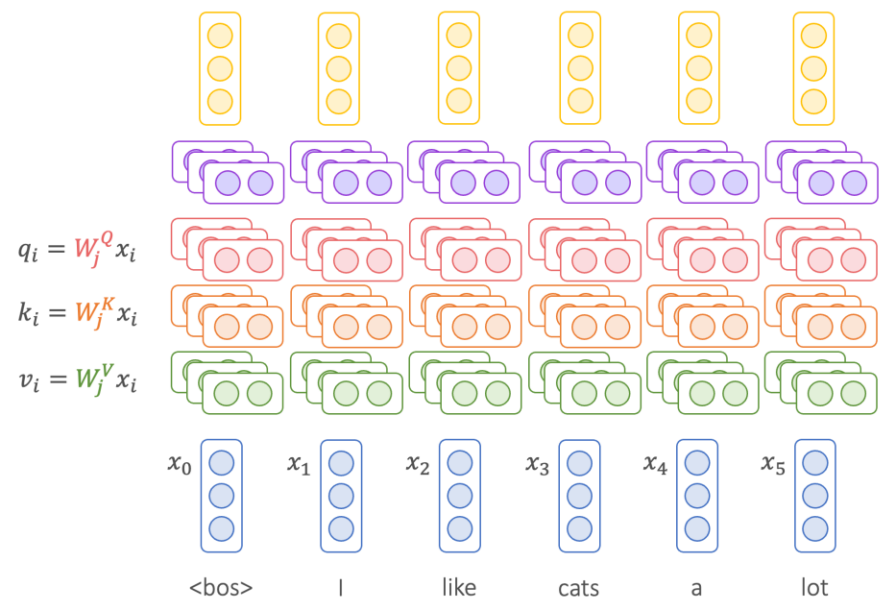
Task	Corpus	Train	Dev	Test	Test sets	Lang.	Task	Metric	Domain
Classification	XNLI	392,702	2,490	5,010	translations	15	NLI	Acc.	Misc.
	PAWS-X	49,401	2,000	2,000	translations	7	Paraphrase	Acc.	Wiki / Quora
Struct. pred.	POS	21,253	3,974	47-20,436	ind. annot.	33 (90)	POS	F1	Misc.
	NER	20,000	10,000	1,000-10,000	ind. annot.	40 (176)	NER	F1	Wikipedia
QA	XQuAD	87,599	34,726	1,190	translations	11	Span extraction	F1 / EM	Wikipedia
	MLQA			4,517-11,590	translations	7	Span extraction	F1 / EM	Wikipedia
	TyDiQA-GoldP	3,696	634	323-2,719	ind. annot.	9	Span extraction	F1 / EM	Wikipedia
Retrieval	BUCC	-	-	1,896-14,330	-	5	Sent. retrieval	F1	Wiki / news
	Tatoeba	-	-	1,000	-	33 (122)	Sent. retrieval	Acc.	misc.

Parse Trees Help Cross-Lingual Transfer

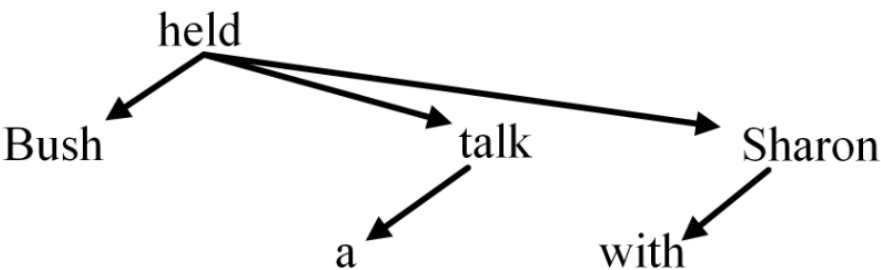
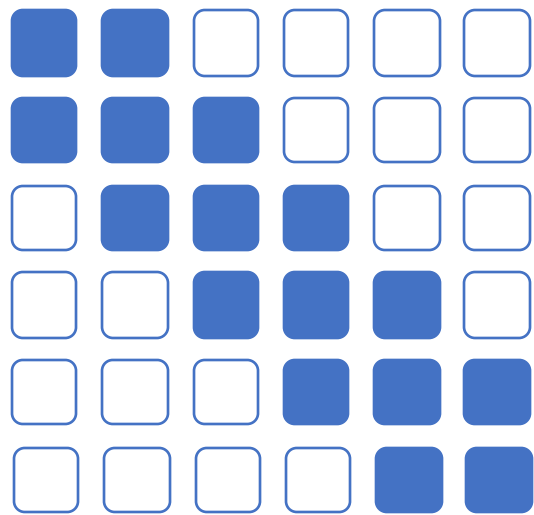
- Translation sentences share similar dependency structures



Parse Trees Help Cross-Lingual Transfer



Attention Mask



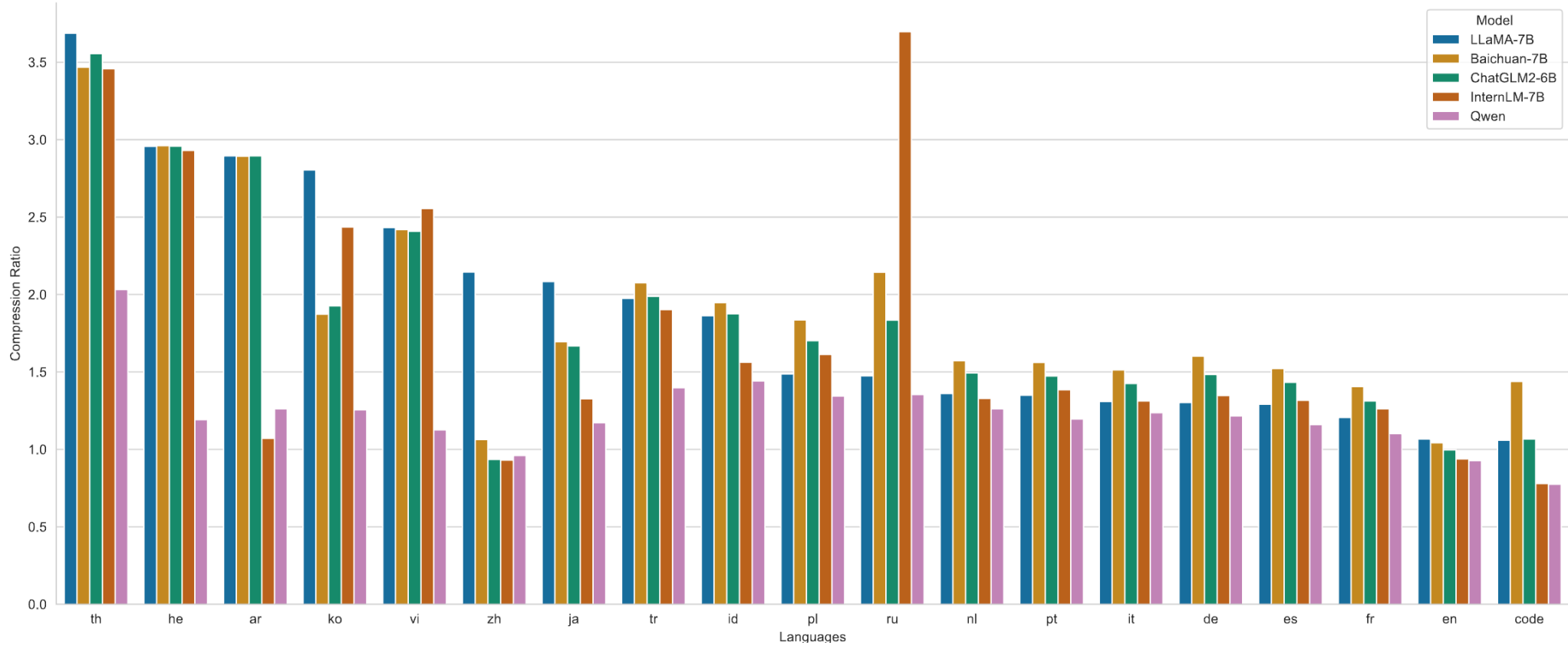
	Bush	held	a	talk	with	Sharon
Bush	1	1	3	2	3	2
held	1	1	2	1	2	1
a	3	2	1	1	4	3
talk	2	1	1	1	3	2
with	3	2	4	3	1	1
Sharon	2	1	3	2	1	1

Parse Trees Help Cross-Lingual Transfer

Model	Event Argument Role Labeling						Relation Extraction					
	En	En	Zh	Zh	Ar	Ar	En	En	Zh	Zh	Ar	Ar
	⇓ Zh	⇓ Ar	⇓ En	⇓ Ar	⇓ En	⇓ Zh	⇓ Zh	⇓ Ar	⇓ En	⇓ Ar	⇓ En	⇓ Zh
CL_Trans_GCN	41.8	55.6	41.2	52.9	39.6	40.8	56.7	65.3	65.9	59.7	59.6	46.3
CL_GCN	51.9	50.4	53.7	51.5	50.3	51.9	49.4	58.3	65.0	55.0	56.7	42.4
CL_RNN	60.4	53.9	55.7	52.5	50.7	50.9	53.7	63.9	70.9	57.6	67.1	55.7
Transformer	61.5	55.0	58.0	57.7	54.3	57.0	57.1	63.4	69.6	60.6	67.0	52.6
Transformer_RPR	62.3	60.8	57.3	66.3	57.5	59.8	58.0	59.9	70.0	55.6	66.5	56.5
GATE (this work)	63.2	68.5	59.3	69.2	53.9	57.8	55.1	66.8	71.5	61.2	69.0	54.3

Multilingual Large Language Models

- Llama series / Qwen series



Cross-Lingual In-Context Learning

English ICL Q

Michael had 58 golf balls. On Tuesday, he lost 23 golf balls. On Wednesday, he lost 2 more. How many golf balls did he have at the end of Wednesday?

ICL A

Michael started with 58 golf balls and lost 23, so he has $58 - 23 = 35$. After he lost 2 more, he has $35 - 2 = 33$ balls.

English ICL Q

Olivia has \$23. ... How much does she have left?

ICL A

5 bagels for \$3 ... The answer is 8.

English ICL Q

... How many lollipops did Jason give to Denny?

ICL A

... The answer is 8.

Thai Test Q

... รายได้ของเธอในสัปดาห์นี้จะเท่ากับกี่ดอลลาร์
(Eliza's rate per hour for the first 40 hours she works each week is \$10. She also receives an overtime pay of 1.2 times her regular hourly rate. If Eliza worked for 45 hours this week, how much are her earnings this week?)

Model Answer

... Elisa's total earnings for the week are \$400 (from the first 40 hours) + \$48 (from that overtime hour) = \$448.



Spanish ICL Q

Michael tenía 58 pelotas de golf. El martes, perdió 23 pelotas de golf. El miércoles, perdió 2 más. ¿Cuántas pelotas de golf tenía al final del miércoles?

ICL A

Michael started with 58 golf balls and lost 23, so he has $58 - 23 = 35$. After he lost 2 more, he has $35 - 2 = 33$ balls.

German ICL Q

Olivia hat \$23. ... Wie viel Geld hat sie übrig?

ICL A

5 bagels for \$3 ... The answer is 8.

Chinese ICL Q

... 杰森给了丹尼多少根棒棒糖 ?

ICL A

... The answer is 8.

Thai Test Q

... รายได้ของเธอในสัปดาห์นี้จะเท่ากับกี่ดอลลาร์

Model Answer

Calculate the earnings for the first 40 hours at the regular rate: $40 \text{ hours} * \$10/\text{hour} = \400 ... earnings for this week will be \$460.



Cross-Lingual In-Context Learning

- Multilingual context-irrelevant sentences also help

 Llama3.1-8B-Instruct			
English	57.10	44.42	55.91
English + CIS-En	55.90	47.88	55.46
English + CIS-Fr	52.10 ^{*3.80↓}	52.82 ^{***4.94↑}	59.40 ^{***3.94↑}
English + CIS-Ja	58.80 ^{2.90↑}	55.96 ^{***8.08↑}	59.86 ^{***4.40↑}
English + CIS-Zh	55.00 ^{0.90↓}	54.68 ^{***6.80↑}	64.66 ^{***9.20↑}
English + CIS-Multi	62.50 ^{***6.60↑}	56.03 ^{***8.15↑}	64.74 ^{***9.28↑}
Multilingual + CIS-Multi	68.60 ^{***6.10↑}	57.24 ^{1.21↑}	66.20 ^{*1.46↑}
 Qwen2-7B-Instruct			
English	43.70	48.46	62.29
English + CIS-En	43.00	50.90	62.31
English + CIS-Fr	43.50 ^{0.50↑}	53.65 ^{***2.75↑}	62.31 ^{0.00–}
English + CIS-Ja	43.80 ^{0.80↑}	56.22 ^{**5.32↑}	63.09 ^{0.78↑}
English + CIS-Zh	42.60 ^{0.40↓}	56.79 ^{***5.89↑}	62.49 ^{0.18↑}
English + CIS-Multi	42.70 ^{0.30↓}	54.94 ^{***4.04↑}	62.51 ^{0.20↑}
Multilingual + CIS-Multi	47.30 ^{***4.70↑}	55.83 ^{0.89↑}	63.51 ^{1.00↑}

How Do LLMs Understand Multilingual Inputs?

- Still an open research problem

Do Multilingual LLMs Think In English?

Lisa Schut¹ Yarin Gal¹ Sebastian Farquhar²

Code-Switching Inputs

- Another open research problem!

这个 weekend 我要去 shopping。

Party बहुत amazing थी, food भी tasty था

Voy a comprar un laptop porque mi old one ya no funciona

cafe مهم بعد ساعة، بعدين نروح meeting أنا عندي