Zero-shot Cross-Linguistic Learning of Event Semantics

Malihe Alikhani, Thomas Kober, Bashar Alhafni, Yue Chen, Mert Inan, Elizabeth Nielsen, Shahab Raji, Mark Steedman, Matthew Stone

Introduction

- *Tense and aspect rank among the most ubiquitous and problematic features of natural language meaning (Hamm and Bott, 2018).*
- *We leverage image descriptions to offer new insights into*
  - *how can we build models that link speakers' choices of tense and aspect to their communicative goals and the discourse context?*
  - *how can we build models that recognize tense and aspect?*

Lexical Aspect

- *States vs. events:*
  - *Events: run, walk, sleep, dance, sit*
  - *States: know, believe, be, have*
- *Event types*
  - *A man is running in the park. (Atelic)*
  - *A woman arrived at a party. (Telic)*

Dowty 1991

Lexical Aspect Across Languages

Arabic

رجل يمشي بجانب الطريق.
man walking-PRS-MASC-IPFV-3SG nearby street

A man is walking nearby the street.

Chinese

雙層 公共 汽車 正 在 公路 上 行駛
double-decker public bus now IPFV road on drive

Double-decker public buses are driving on the road.

Farsi

اتوبوس‌های دوطبقه در خیابان حرکت می‌کنند.
bus-PL double-decker in street move do

Double-decker buses are moving in the street.

German

Zwei Busse fahren an einer Haltelstelle vorbei.
Two buses drive a bus stop past.

Two buses drive past a bus stop.

Arabic:

فتاة تتحدث على الهاتف

girl talk-PRS-FEM-IPFV-3SG on phone

A girl is talking on the phone.

Chinese:

她 正在 用 手机 通话

She PRS use phone talk

She is using a phone to talk

Farsi:

یک نفر با تلفن صحبت می‌کند.

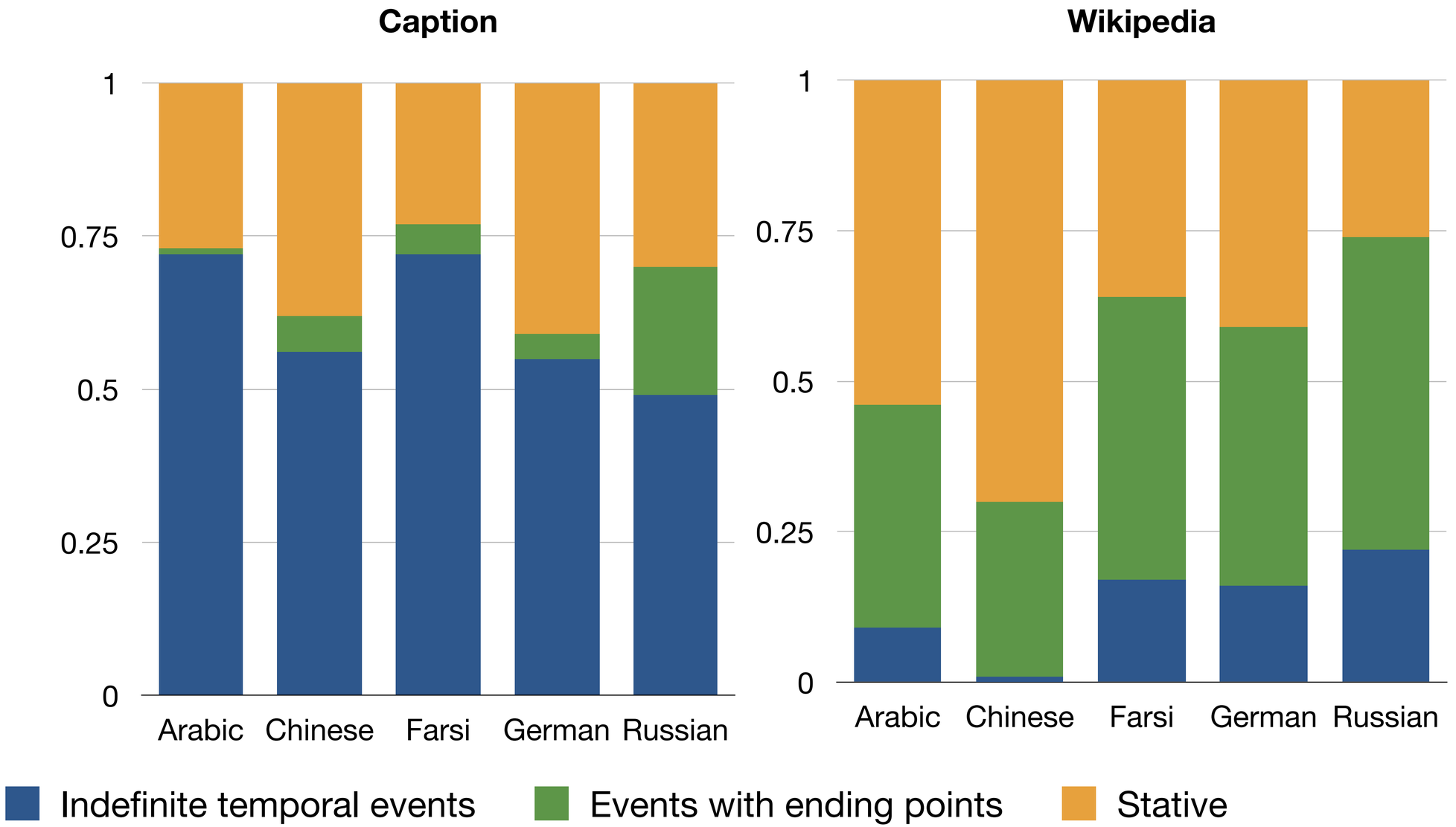
One person with telephone conversation-do-PRS-IPFV-3SG

A person is talking with the phone.

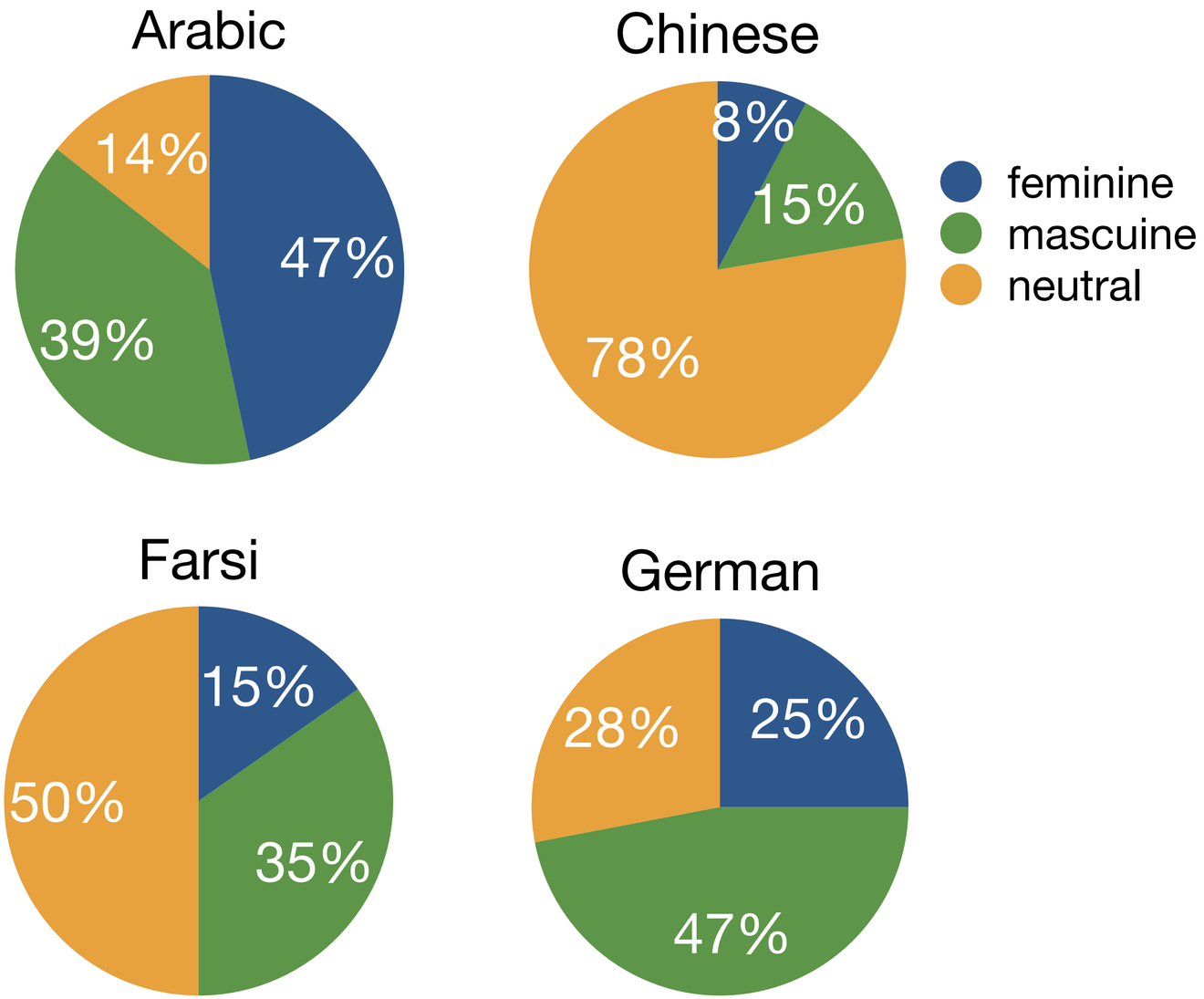
Data Collection

- *We have collected captions in 6 languages for 500 random images from MSCOCO. Arabic, Chinese, Farsi, German, Russian, Turkish*
- *Expert have annotated lexical aspects in those captions.*
- *To compare captions with a unimodal genre, we have annotated 200 Wikipedia sentences across these languages.*
- *Dataset is available on GitHub: https://github.com/malihealikhani/telicity*

Biases in Event Types



Gender across Languages



Computational Experiments

🔬 Research Questions 🔬

- *Mono-lingual* and **cross-lingual** setups:
  - Do distributional semantic models work for the languages in our study?
  - Can we leverage data from other languages?
  - Are there consistent synergies between languages?

Computational Experiments

🎛️ Task Setup 🎛️

- Supervised classification task
  - Our data is annotated with **3 labels**: telic, atelic, state
  - Create **binary classification tasks** for *captions* and *Wikipedia*
    - We drop the telic class from *captions* and the atelic class from *Wikipedia*
  - Evaluation: Accuracy (only mono-lingual), F1-Score per class
- **Mono-lingual**: 10-fold cross-validation
- **Cross-lingual**: Leave-one-language out validation
  - Given *n* languages, train on the data from *n-1* languages, evaluate on the *n*th

Computational Experiments

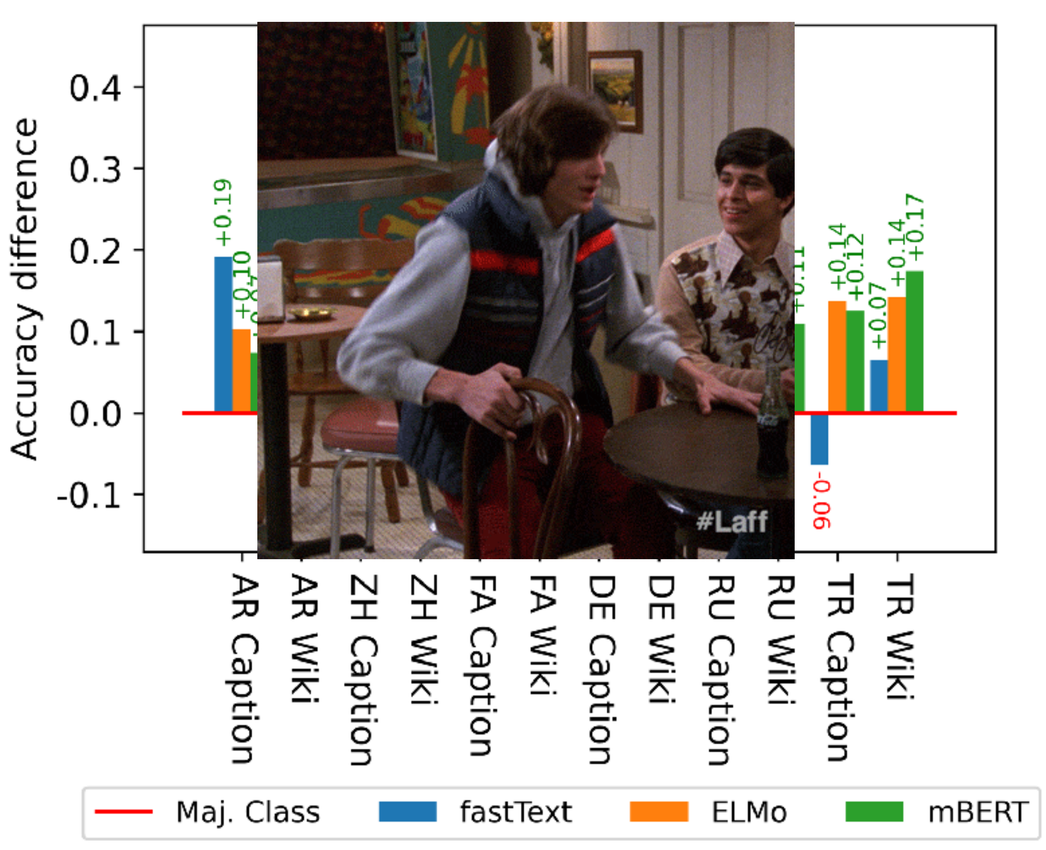
☃️ Model Setup ☃️

- ***Off-the-shelf*** logistic regression classifier
- 3 distributional semantic models:
    - fastText - averaged embeddings
    - BERT
    - ELMo
- *Encode the whole sentence* and use its vector representation *as the input* to the logistic regression classifier.

Computational Experiments

📈 Experiments & Results 📈

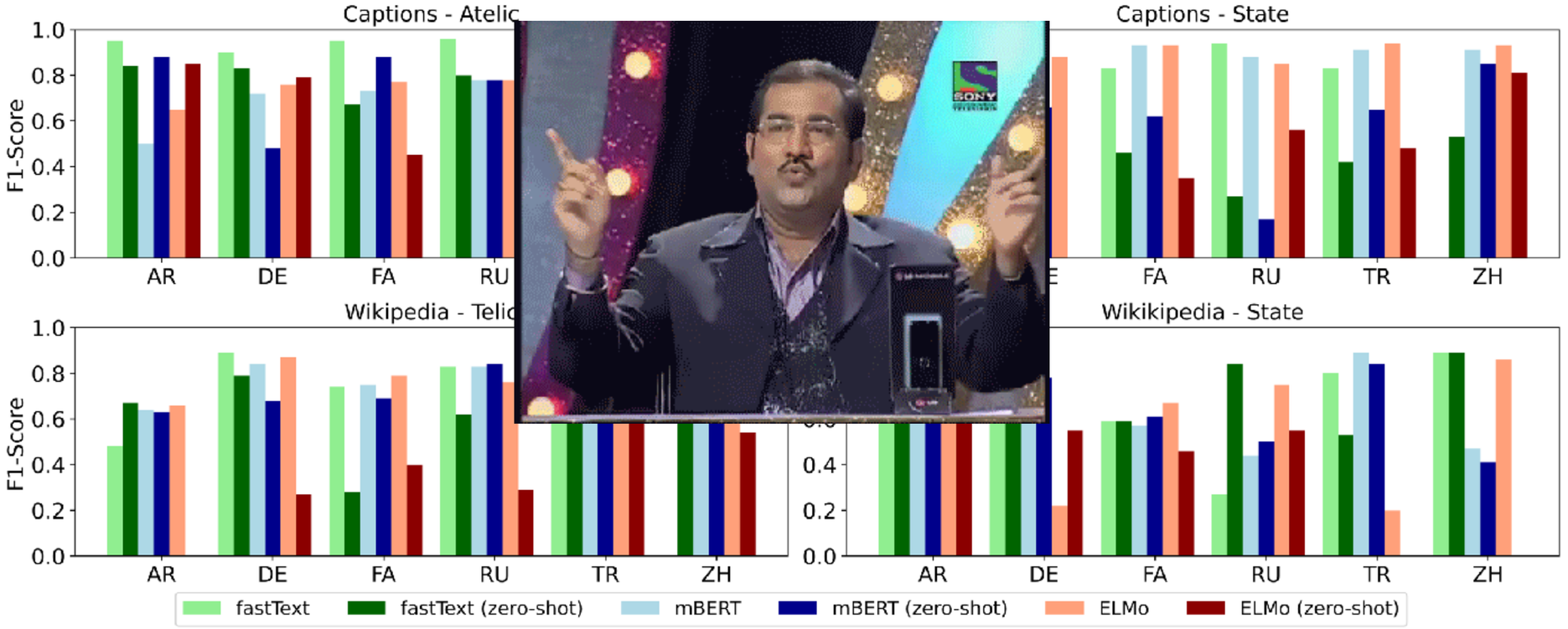
- ***Mono-lingual:*** Compare model performance to majority class baseline (*Accuracy*)
  - *Do distributional semantic models work for the languages in our study?* **(Yes)**



Computational Experiments

📈 Experiments & Results 📈

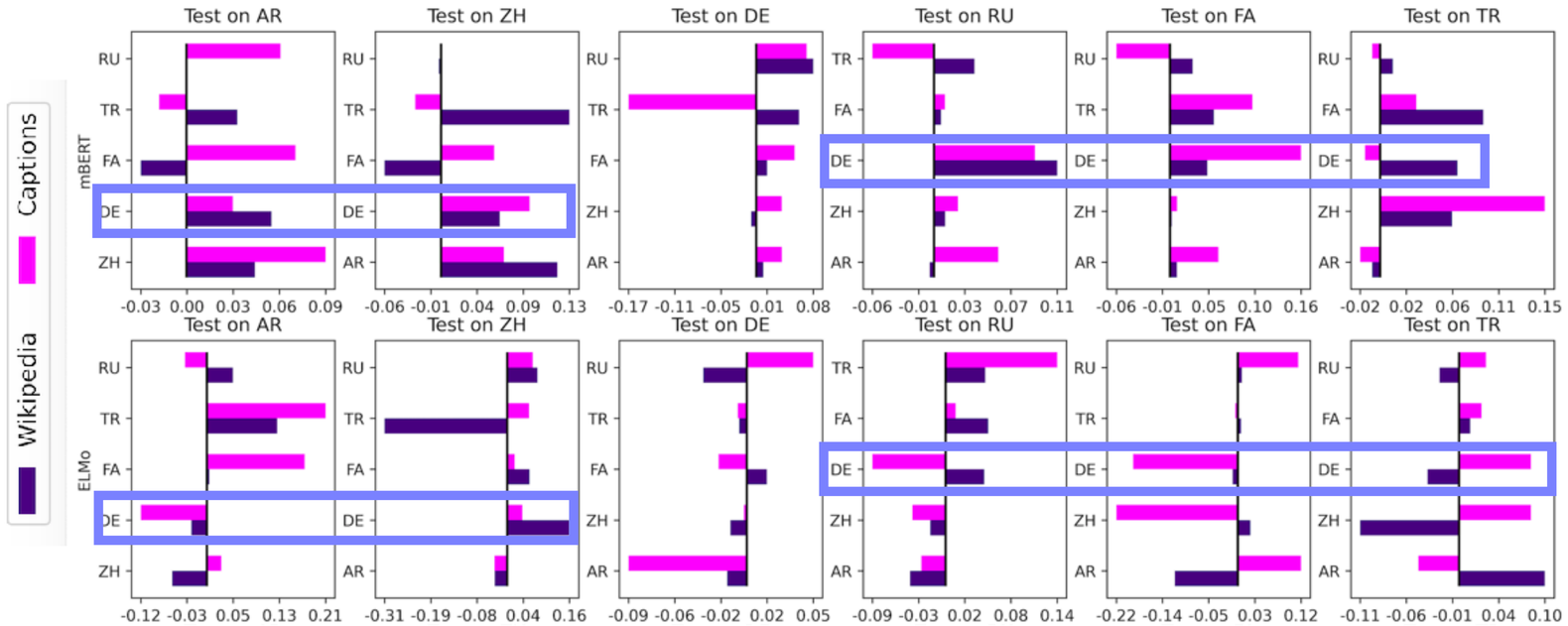
- ***Cross-lingual***: Compare *mono-lingual* and *zero-shot cross-lingual* model performance (class-based **F1-score**)
  - *Can we leverage data from other languages?* **(Also yes)**



Computational Experiments

📈 Experiments & Results 📈

- ***Cross-lingual***: Estimate the *impact of each language* in the zero-shot cross-lingual setup
  - *Are there consistent synergies between languages?* **(Doesn't look like it)**



- Impact of language seems to be **governed** more **by** the **model** than the language. More analyses in future work.

Conclusion

- We show that distributional semantic can reliably predict aspectual classes across languages, and achieves remarkable performance even in zeroshot cross-lingual experiments.
- We have furthermore provided first evidence that aspect can be predicted in a zero-shot cross-lingual manner where a model has not been exposed to any training data in the target language at all.

Thank you!