A Simple and Accurate Syntax-Agnostic Neural Model for Dependency-based Semantic Role Labeling

Diego Marcheggiani, Anton Frolov, and Ivan Titov

University of Amsterdam

Yandex

University of Edinburgh

CoNLL 2017, Vancouver

Contributions

- Neural model for dependency-based SRL
- Simple
- Syntax-agnostic
- State of the art on out-of-domain data
- State-of-the-art performance on English, Chinese, Czech, and Spanish

Semantic role labeling

- Predicting the predicate-argument structure of a sentence

Sequa makes and repairs jet engines.

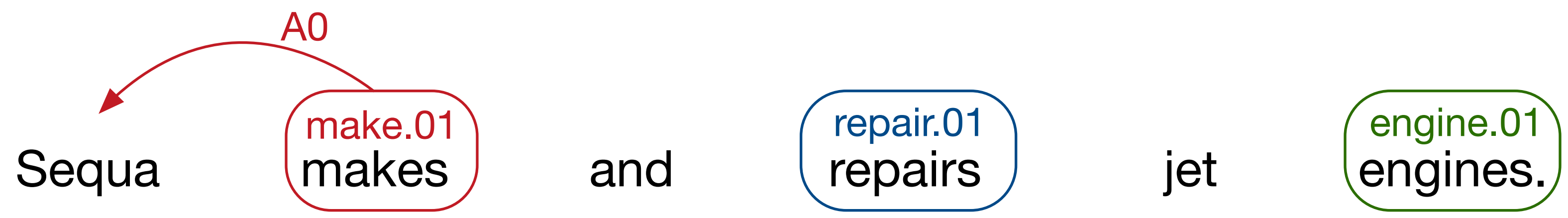
Semantic role labeling

- Predicting the predicate-argument structure of a sentence
  - Discover and disambiguate predicates

Sequa	make.01 makes	and	repair.01 repairs	jet	engine.01 engines.

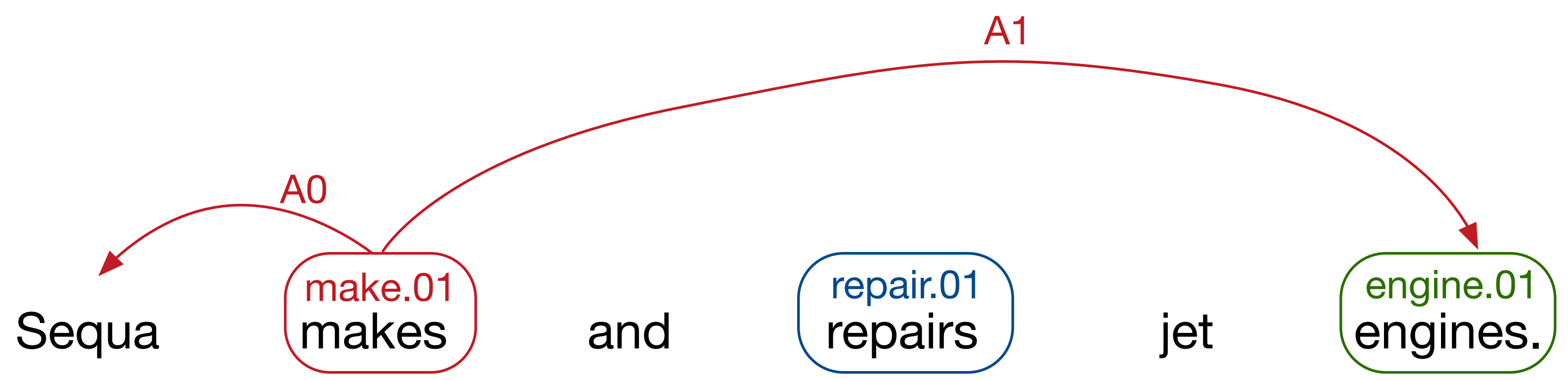
Semantic role labeling

- Predicting the predicate-argument structure of a sentence
  - Discover and disambiguate predicates
  - Identify arguments and label them with their semantic roles



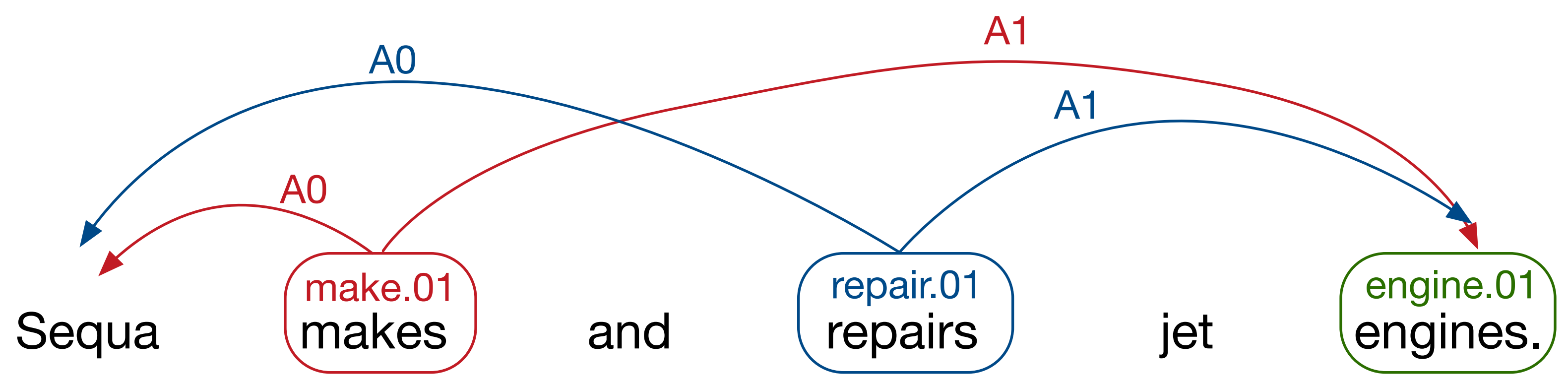
Semantic role labeling

- Predicting the predicate-argument structure of a sentence
  - Discover and disambiguate predicates
  - Identify arguments and label them with their semantic roles



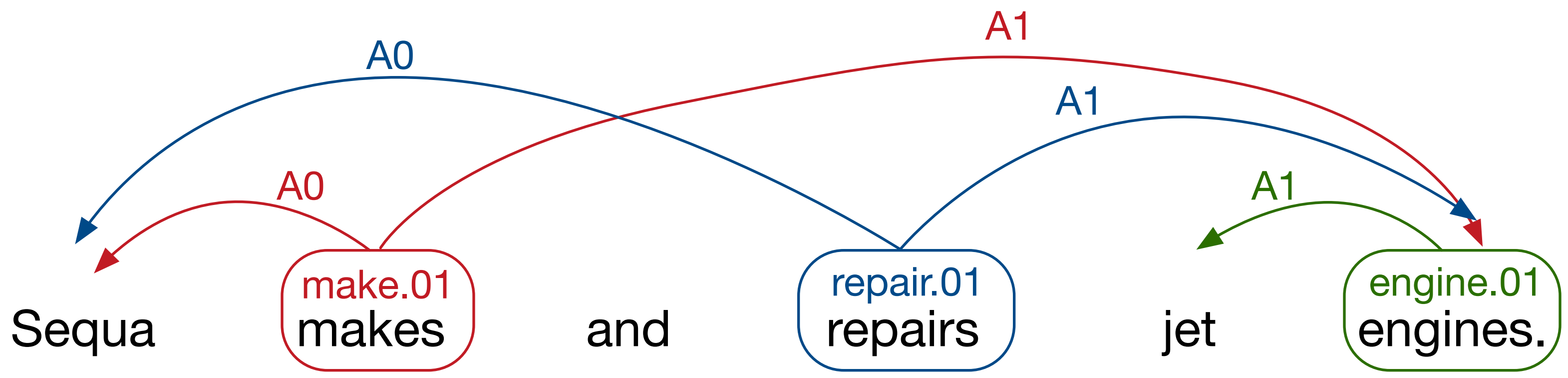
Semantic role labeling

- Predicting the predicate-argument structure of a sentence
  - Discover and disambiguate predicates
  - Identify arguments and label them with their semantic roles



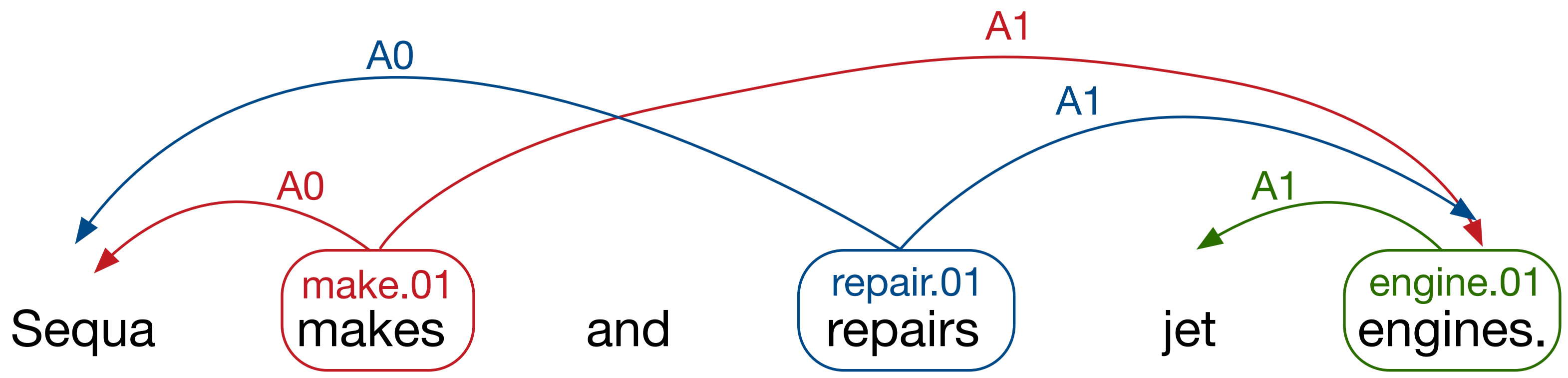
Semantic role labeling

- Predicting the predicate-argument structure of a sentence
  - Discover and disambiguate predicates
  - Identify arguments and label them with their semantic roles



Semantic role labeling

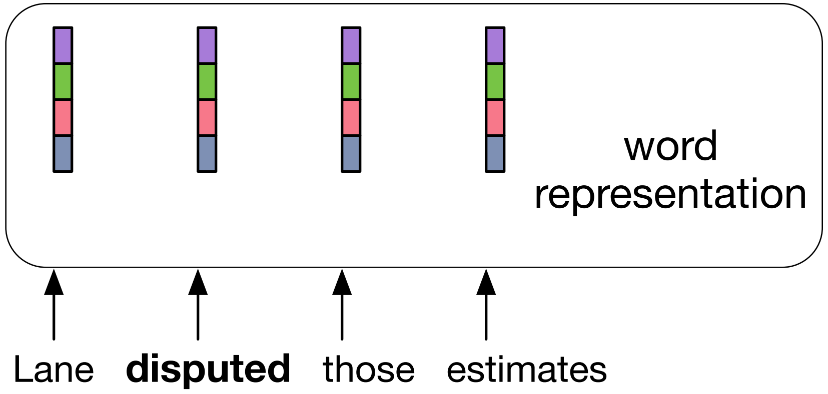
- Only the head of an argument is labeled
- Sequence labeling task for each predicate
- Focus on argument identification and labeling



Model Architecture

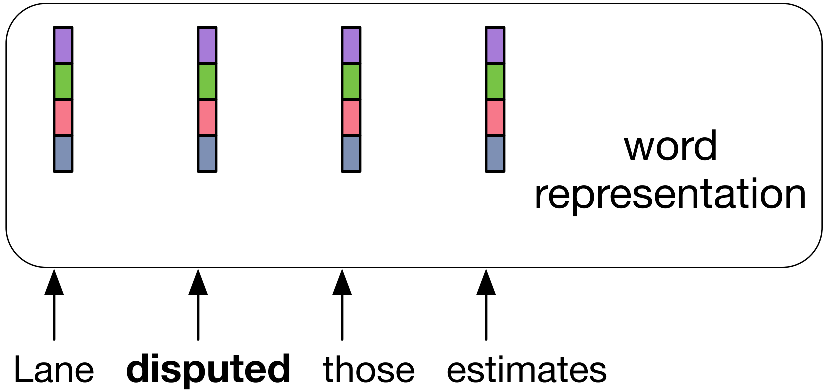
- Word representation
- Sentence encoding (BiLSTM)
- Local classifier

Word encoding



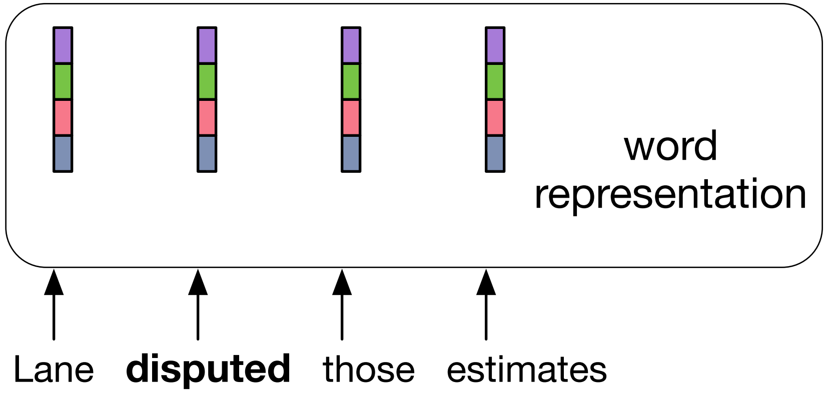
Word encoding

- pretrained word embeddings



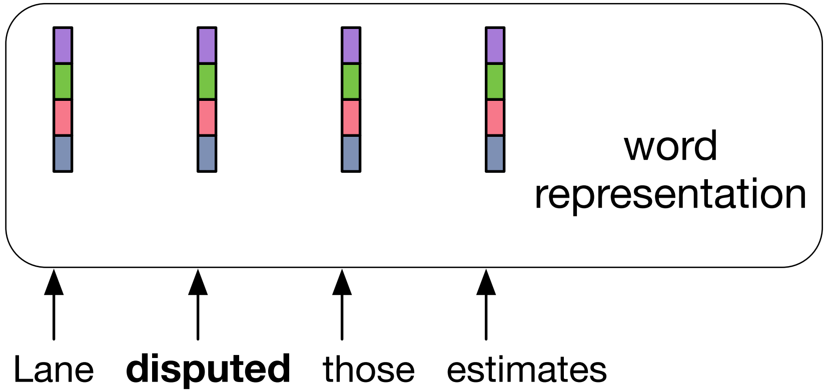
Word encoding

- pretrained word embeddings
- word embeddings



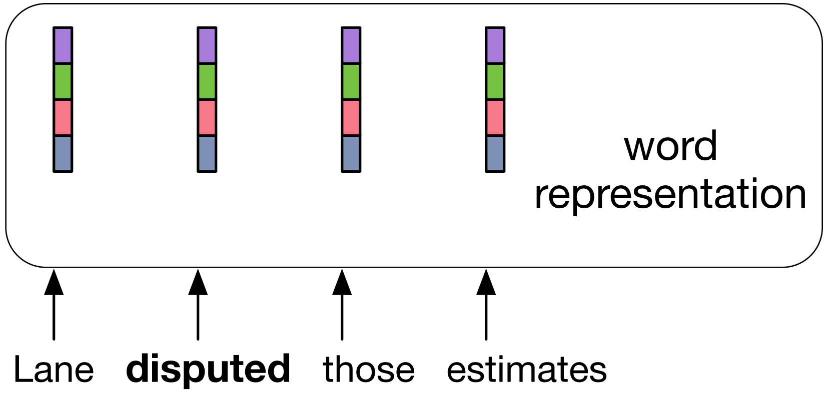
Word encoding

- pretrained word embeddings
- word embeddings
- POS tag embeddings



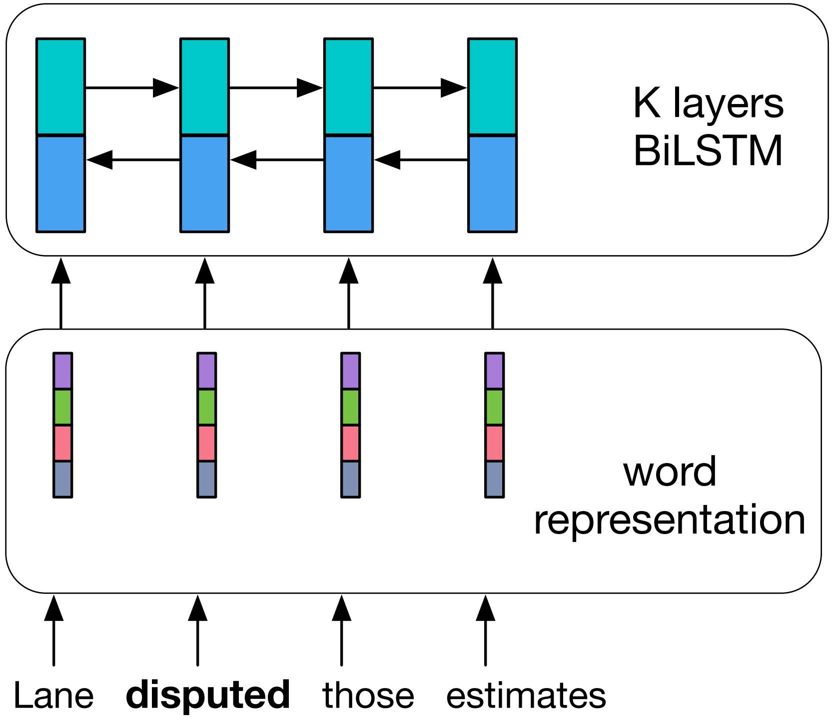
Word encoding

- pretrained word embeddings
- word embeddings
- POS tag embeddings
- predicate lemmas embeddings

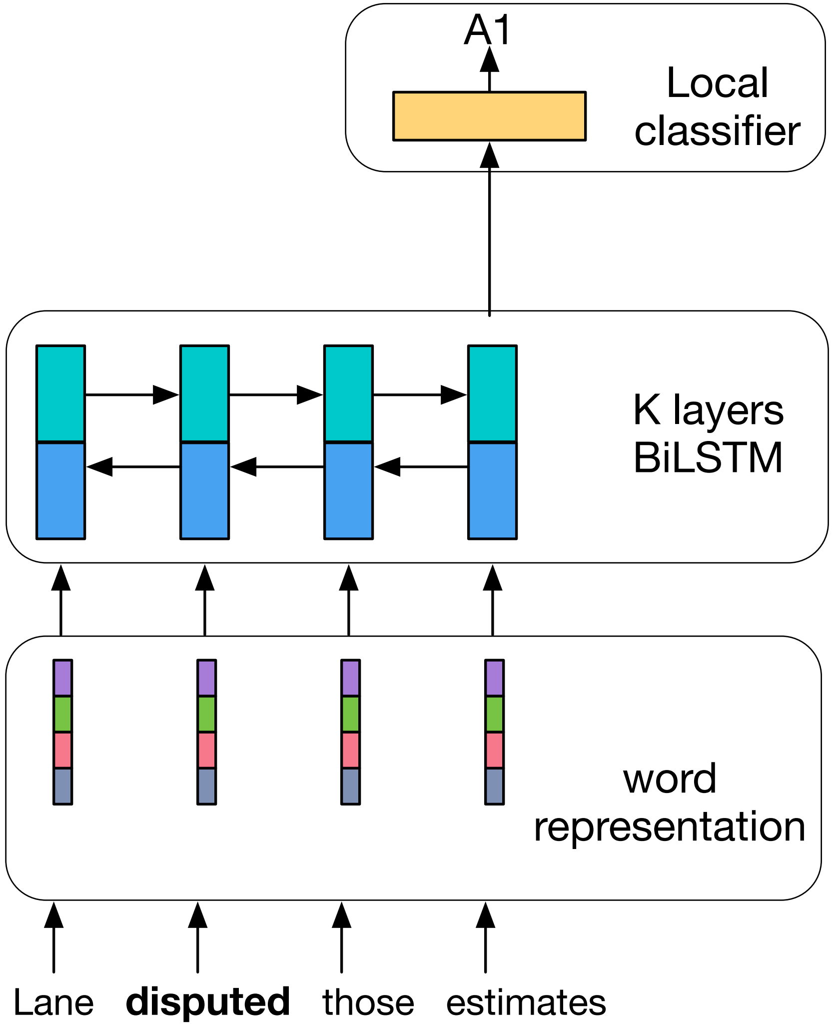


Sentence encoding

- Bidirectional LSTM
  - Forward LSTM encodes left context
  - Backward LSTM encodes right context
  - Forw. and Backw. states are concatenated
  - Stacking of several BiLSTM layers



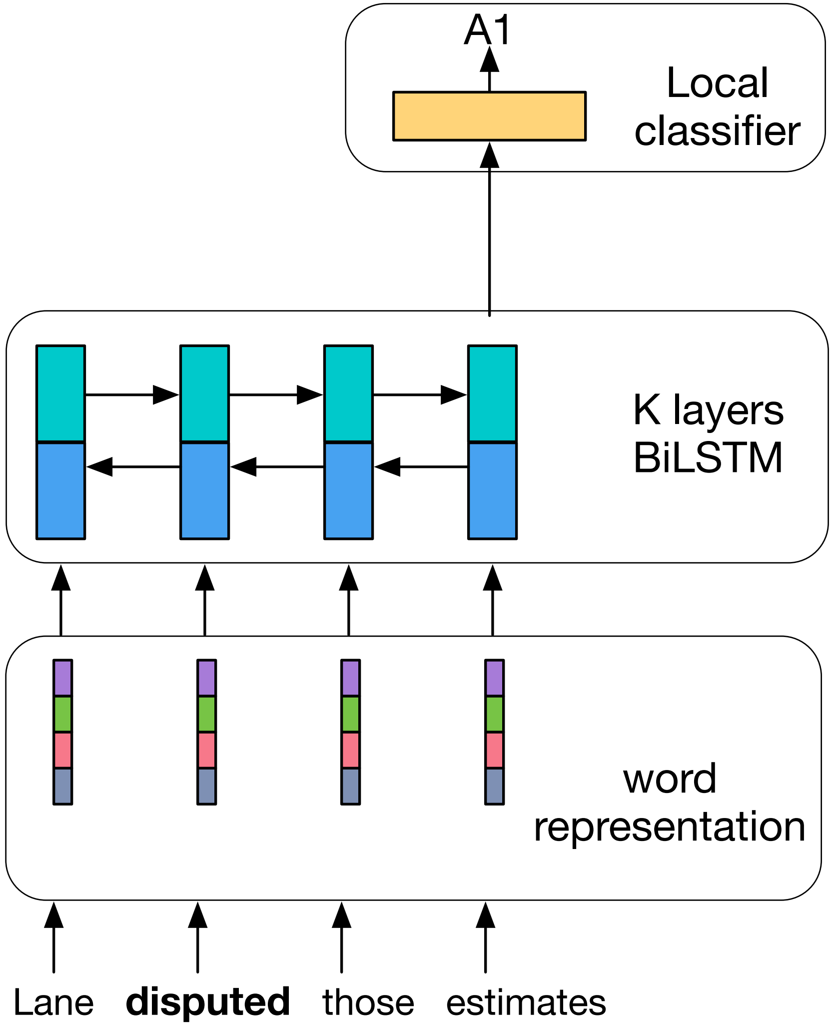
Local classifier



Local classifier

$$p(r|v_i, p) \propto \exp(W_r v_i)$$

role

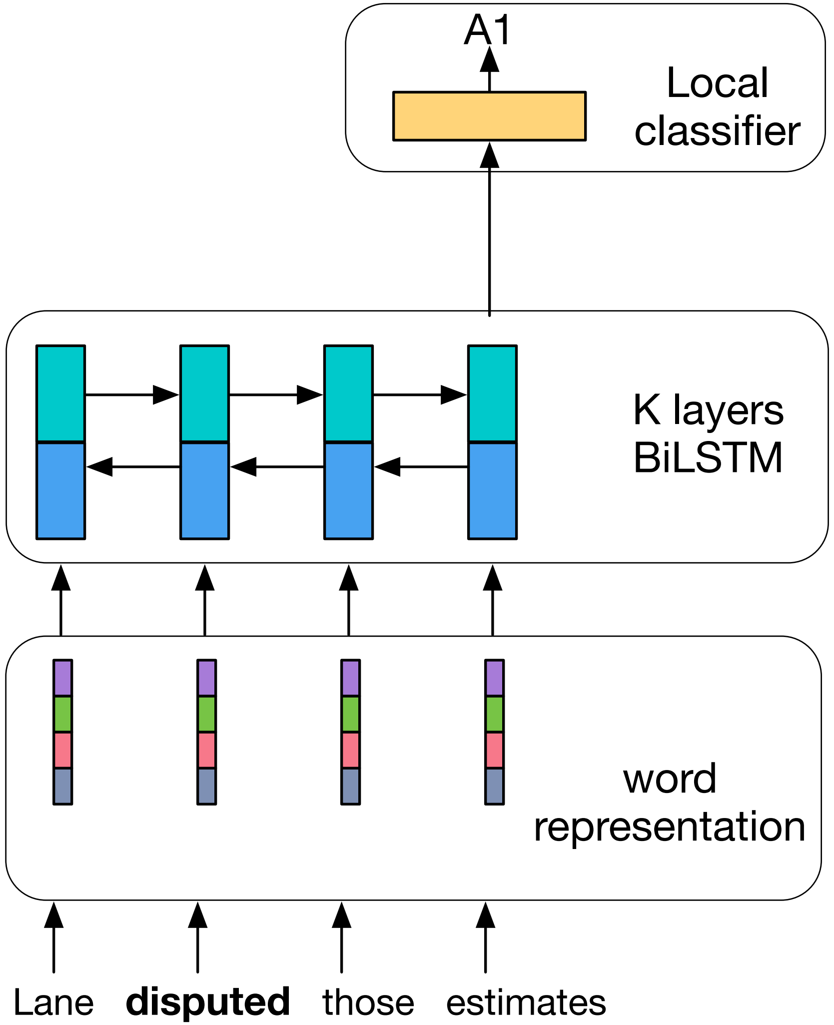


Local classifier

$$p(r|v_i, p) \propto \exp(W_r v_i)$$

role

candidate argument representation



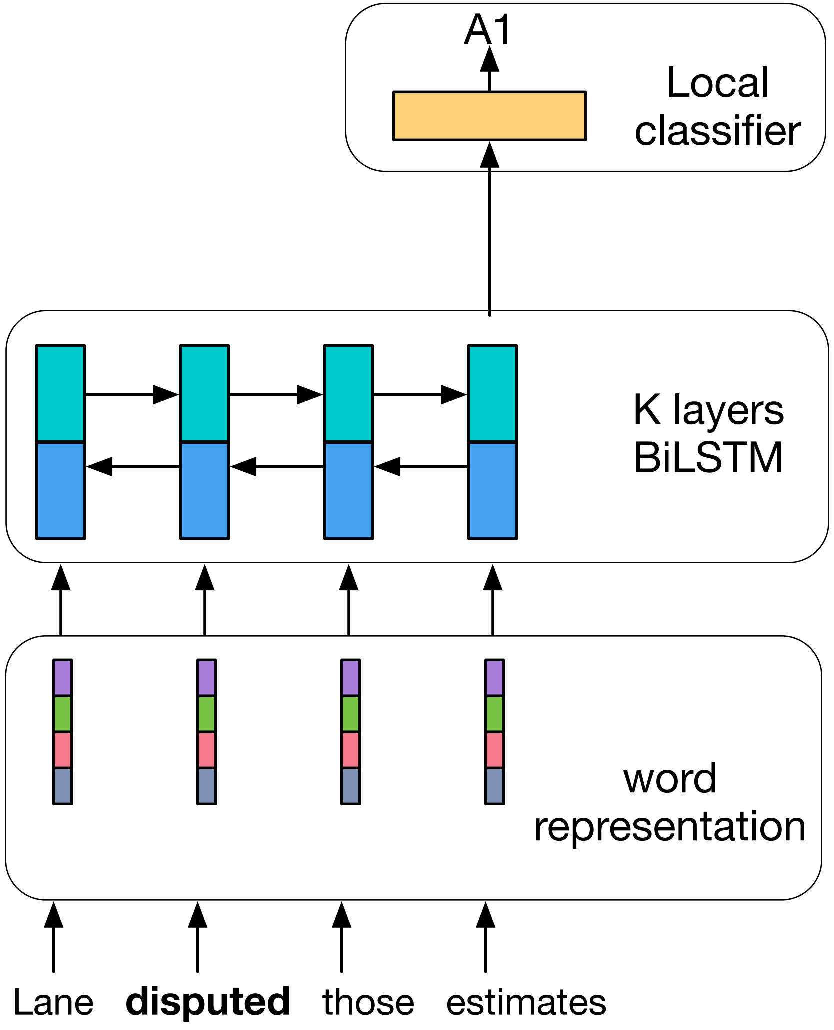
Local classifier

$$p(r|v_i, p) \propto \exp(W_r v_i)$$

role

candidate argument representation

predicate



Re-encoding

$$p(r|v_i, p) \propto \exp(W_r v_i)$$

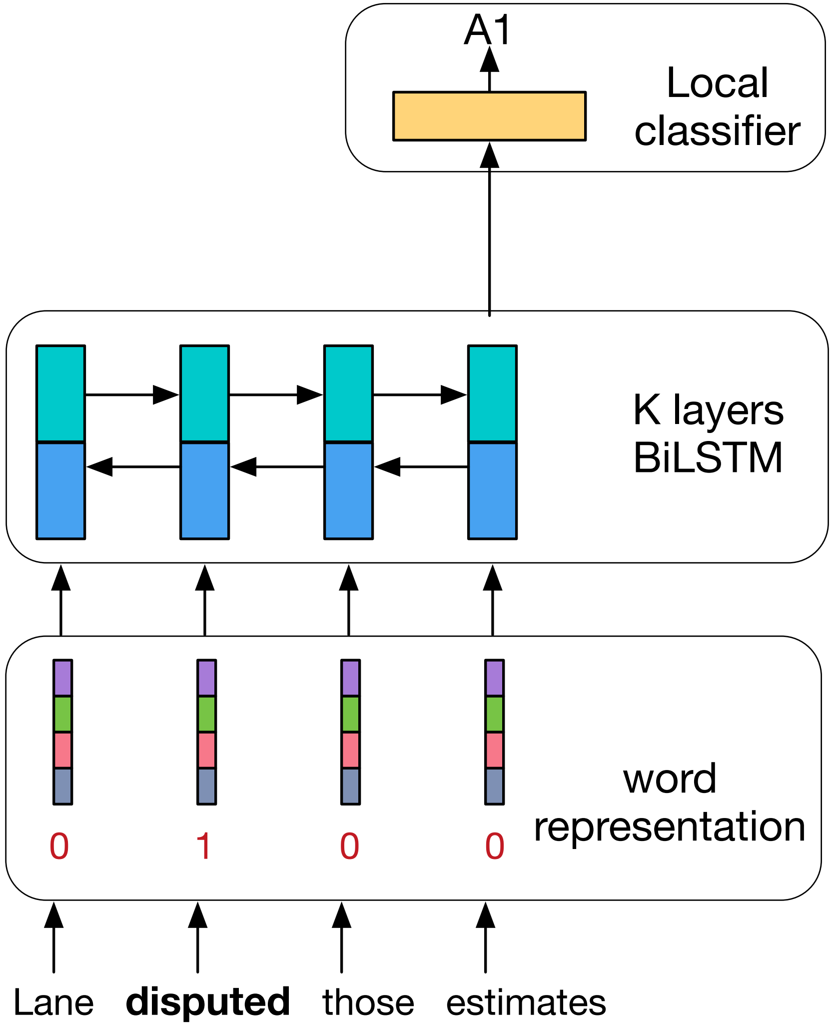
role

predicate

candidate argument representation

Encoding predicates

- Add a predicate flag to word representation
- For each predicate the sentence is re-encoded



Re-encoding

$$p(r|v_i, p) \propto \exp(W_r v_i)$$

role

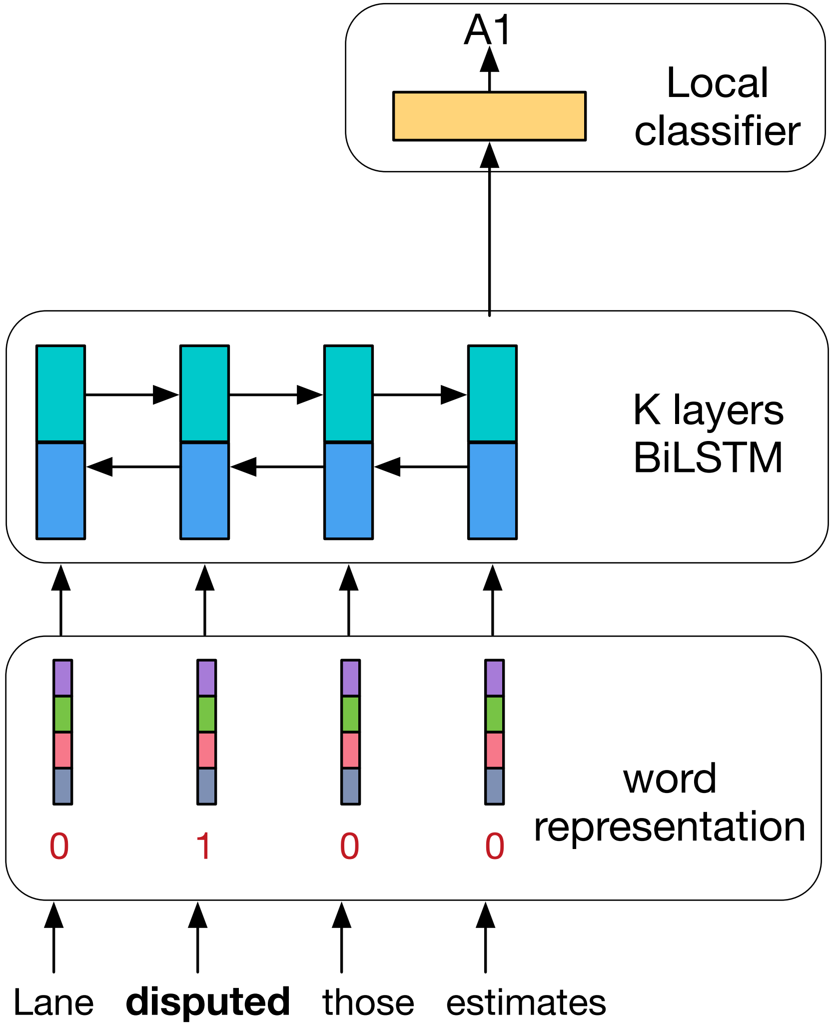
predicate

candidate argument representation

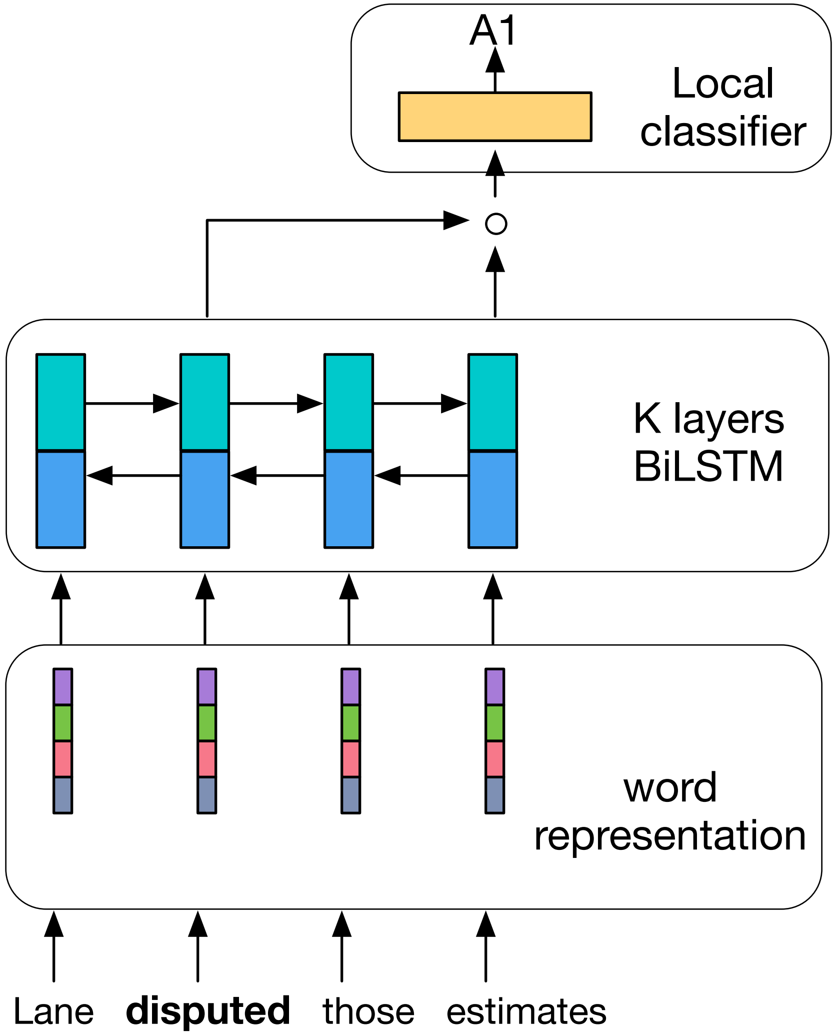
Encoding predicates

- Add a predicate flag to word representation
- For each predicate the sentence is re-encoded

Zhou and Xu, (2015)



Compositional classifier

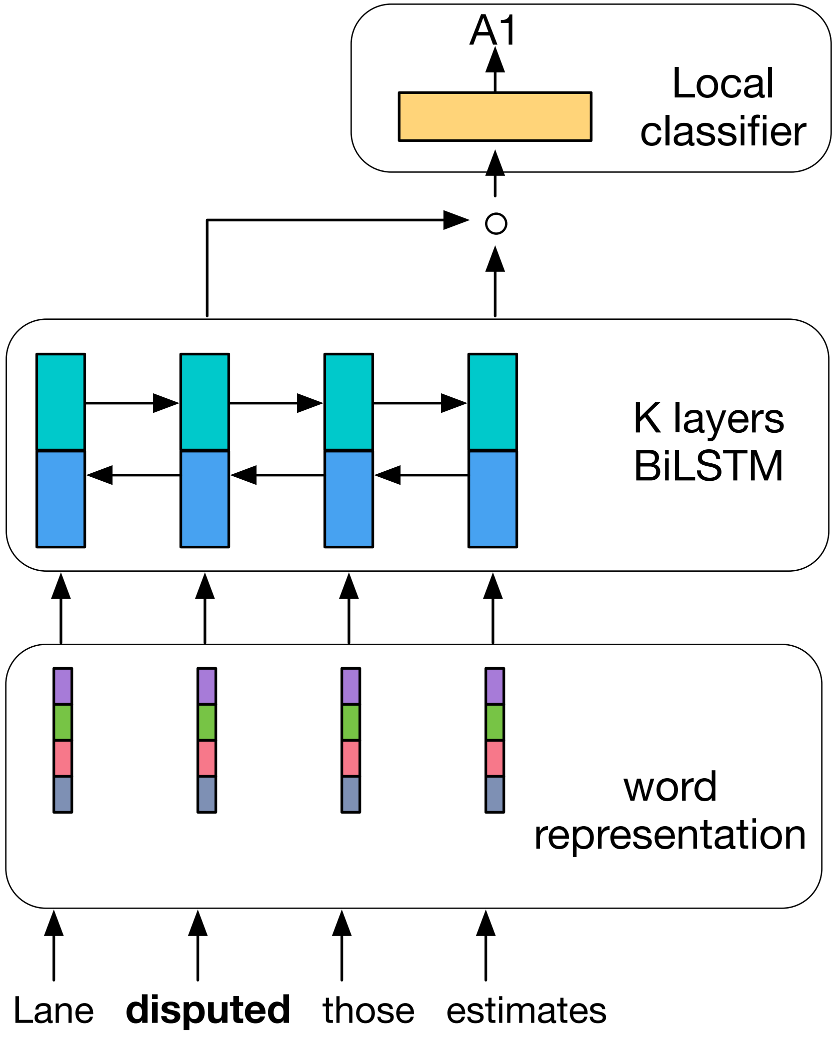


Compositional classifier

$$p(r|t_i, t_p, l) \propto \exp(W_{l,r}(t_i \circ t_p))$$

predicate representation

candidate argument representation



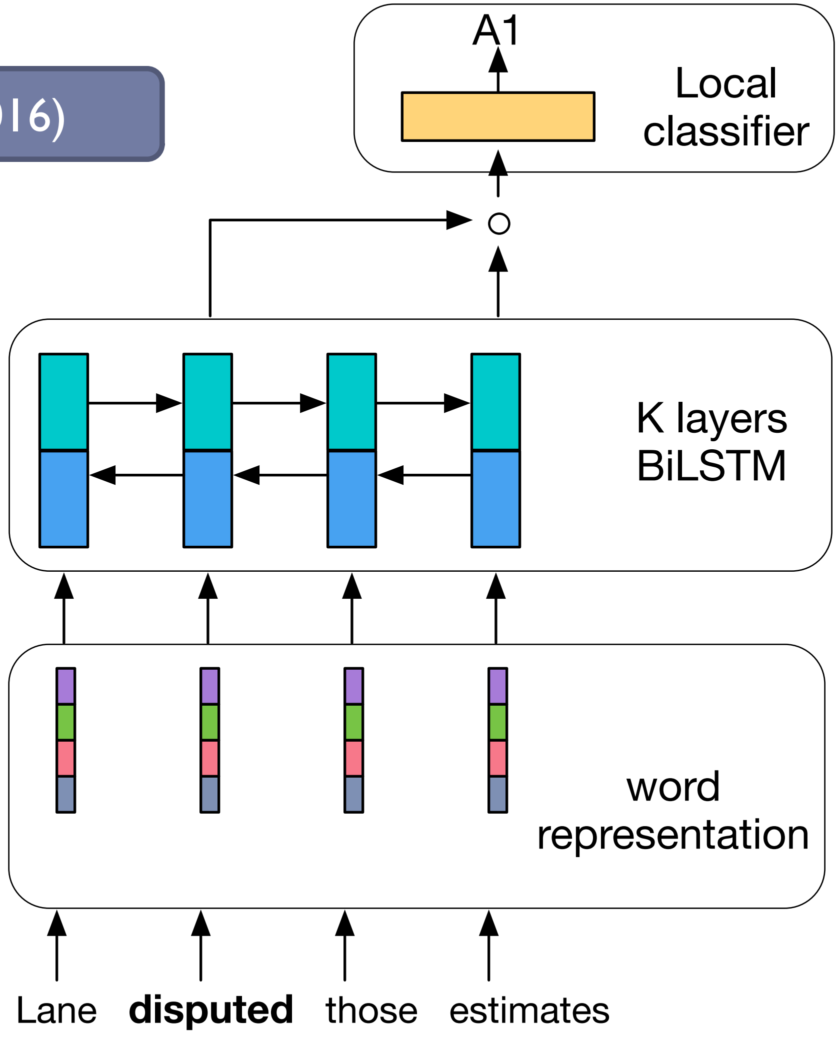
Compositional classifier

Kiperwasser and Goldberg, (2016)

$$p(r|t_i, t_p, l) \propto \exp(W_{l,r}(t_i \circ t_p))$$

predicate representation

candidate argument representation



Compositional classifier

Kiperwasser and Goldberg, (2016)

$$p(r|t_i, t_p, l) \propto \exp(W_{l,r}(t_i \circ t_p))$$

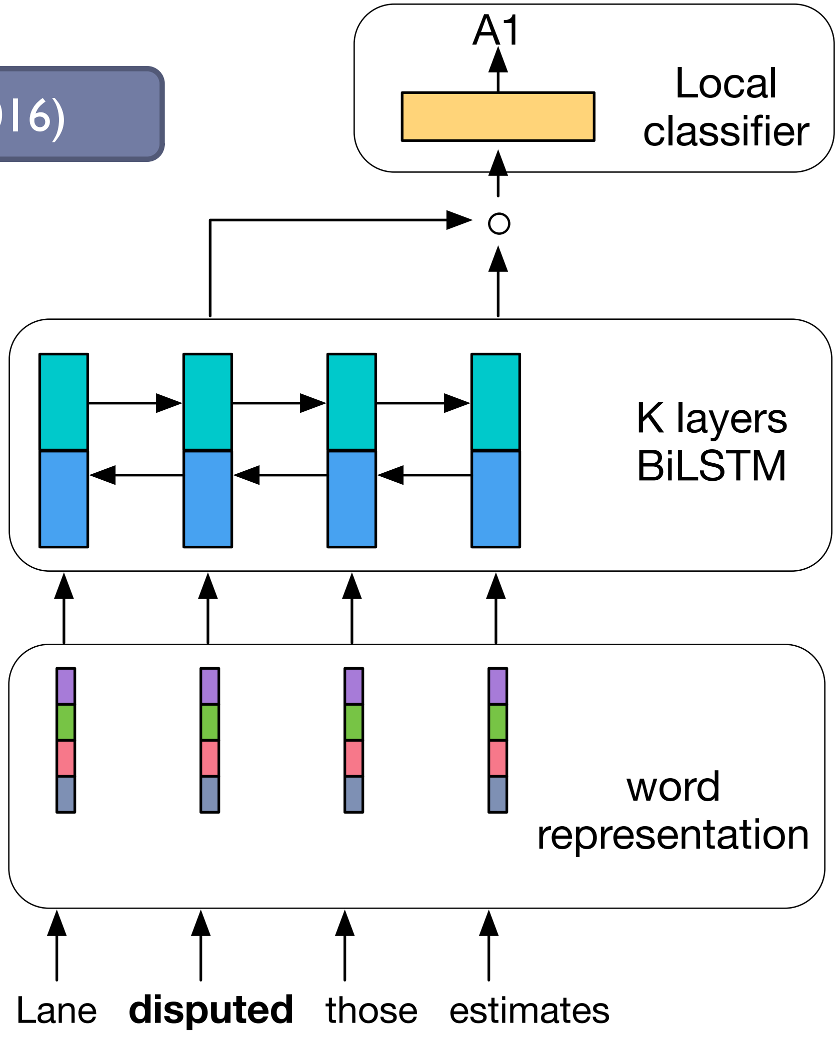
predicate representation

candidate argument representation

$$W_{l,r} = ReLU(U(q_l \circ q_r))$$

predicate lemma embeddings

role embeddings



Compositional classifier

Kiperwasser and Goldberg, (2016)

$$p(r|t_i, t_p, l) \propto \exp(W_{l,r}(t_i \circ t_p))$$

predicate representation

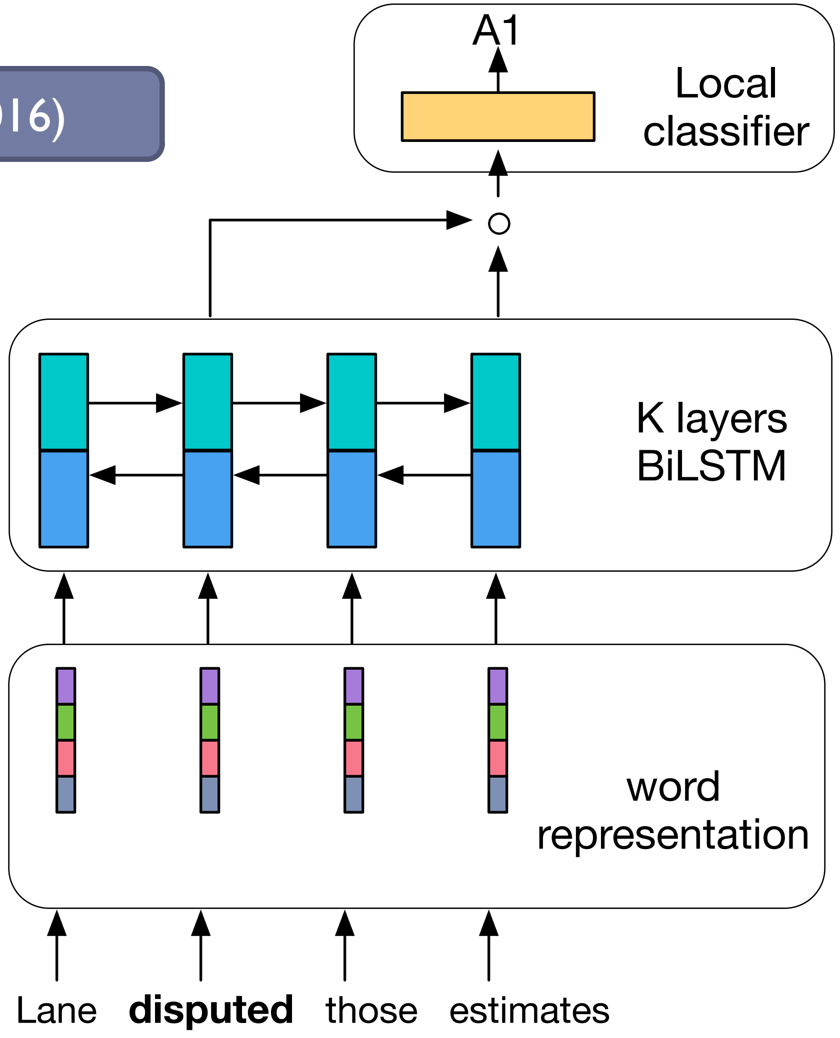
candidate argument representation

$$W_{l,r} = ReLU(U(q_l \circ q_r))$$

predicate lemma embeddings

role embeddings

Fitzgerald et al., (2015)

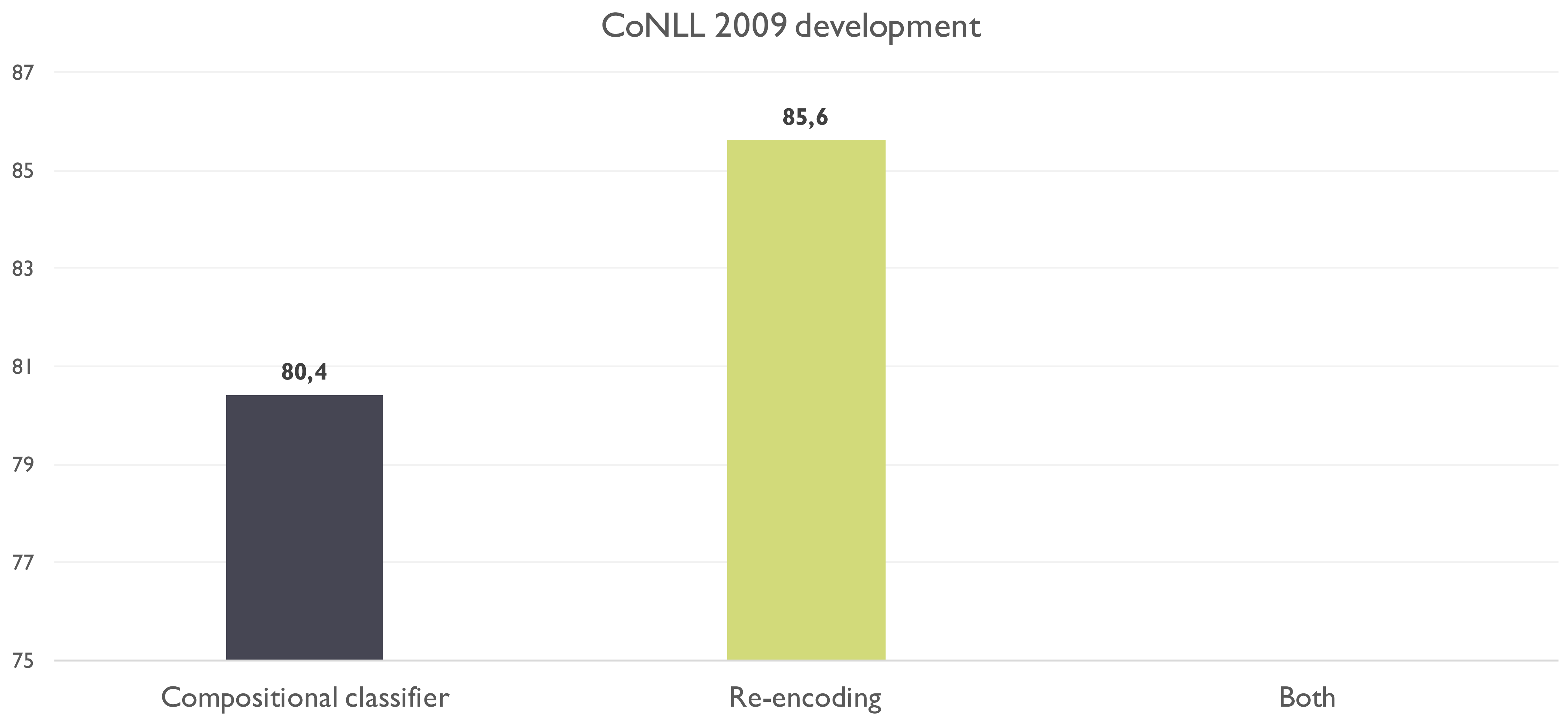


Experimental setting

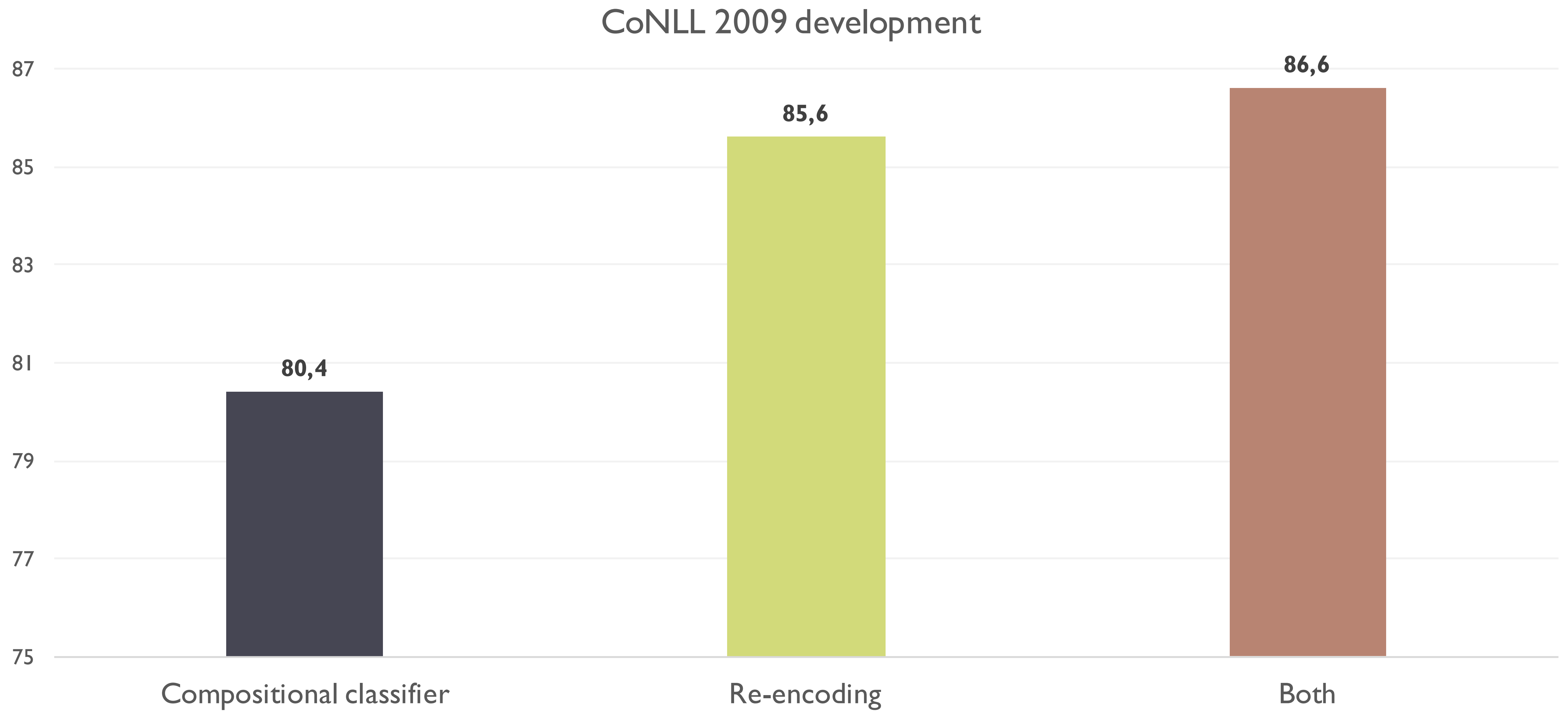
- CoNLL-2009 dependency-based SRL dataset (standard split)
  - English, Chinese, Czech, Spanish
  - F1 as evaluation measure
- State-of-the-art predicate disambiguation models
- Hyperparameters tuned on English dev set
- Adam optimizer

Predicate encoding

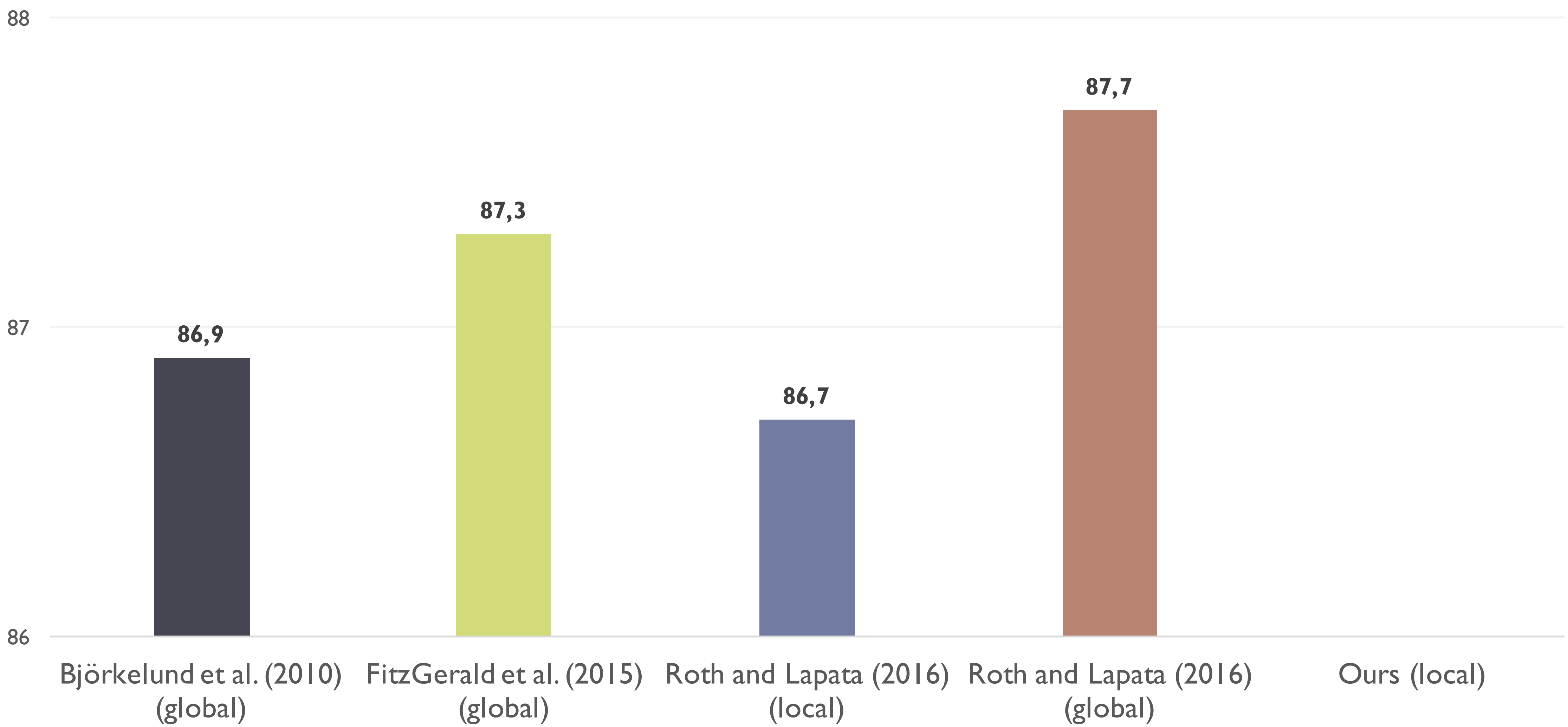
Predicate encoding



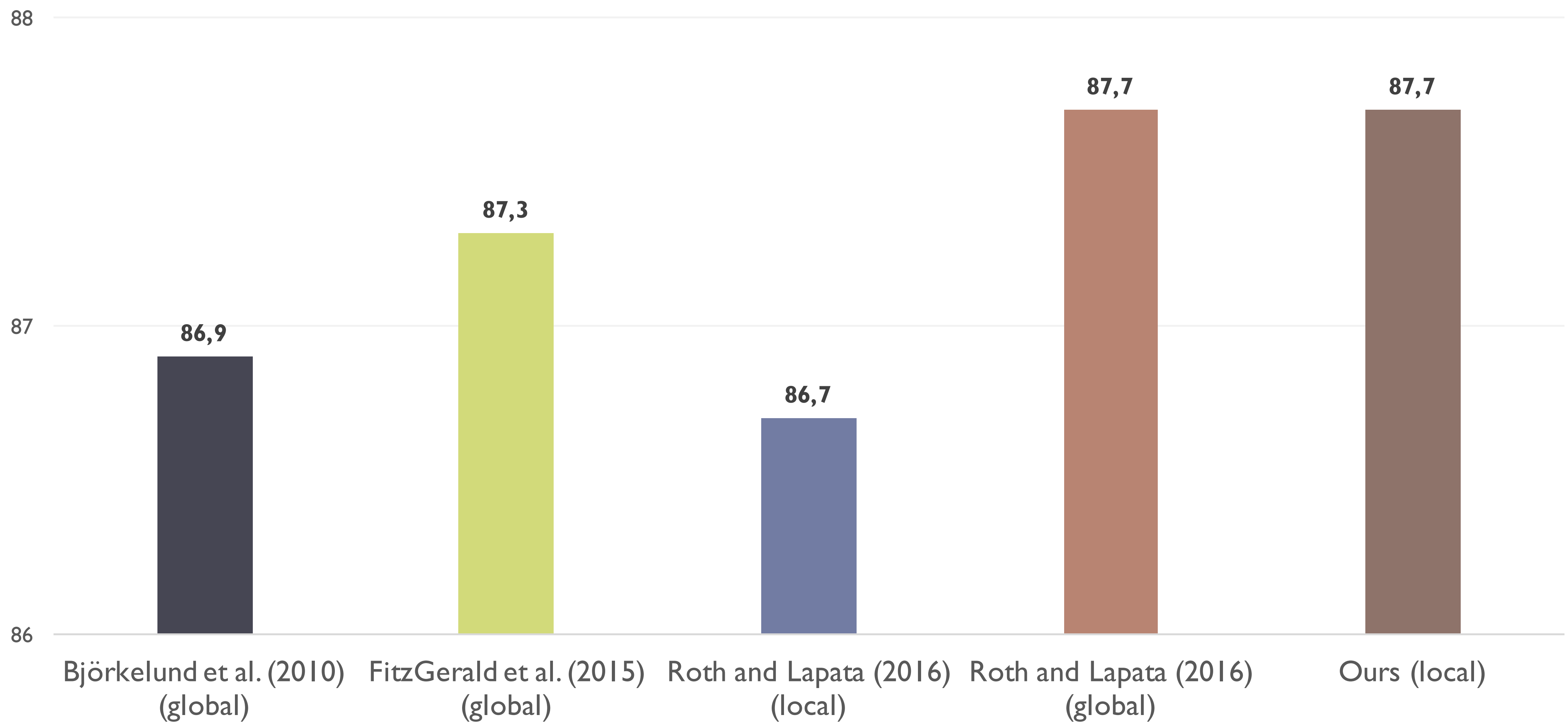
Predicate encoding



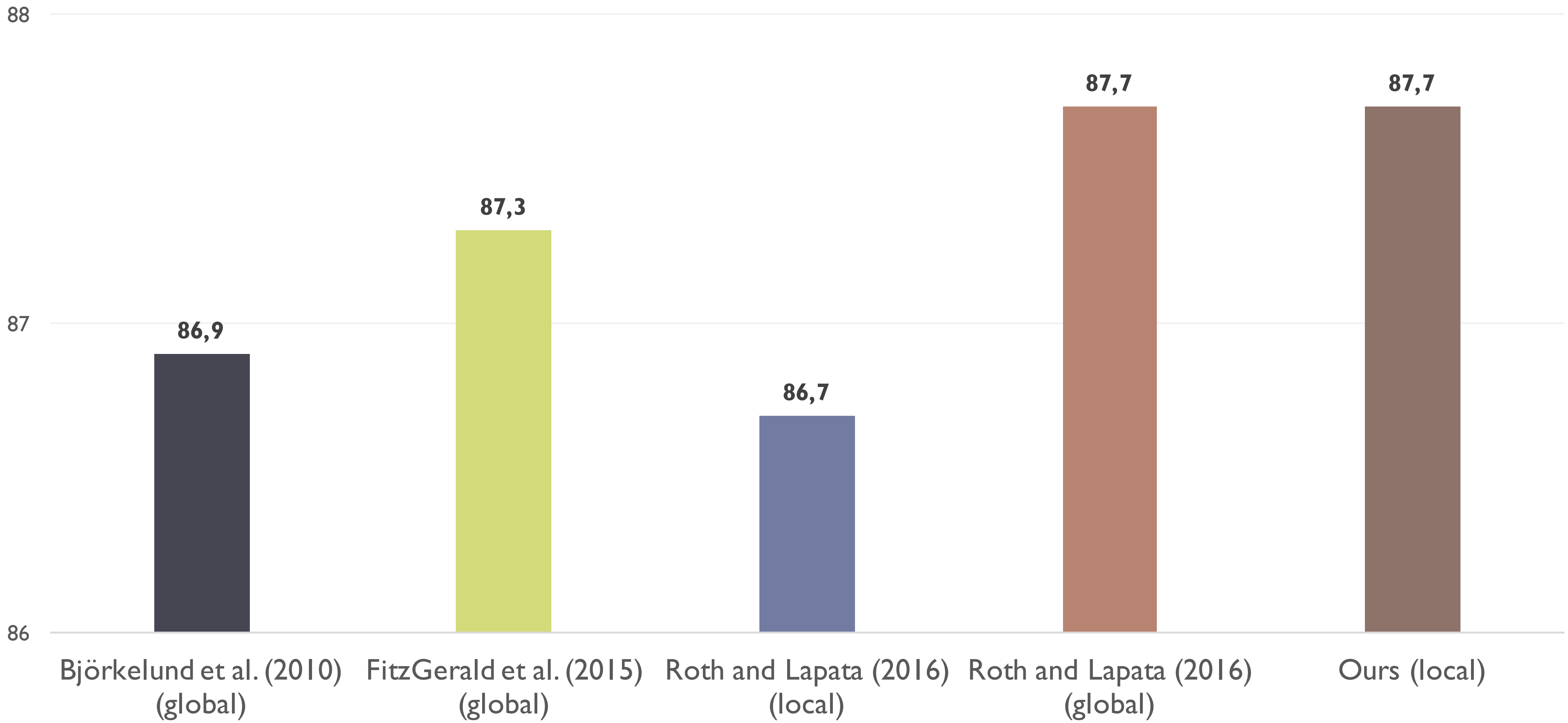
English Test set



English Test set

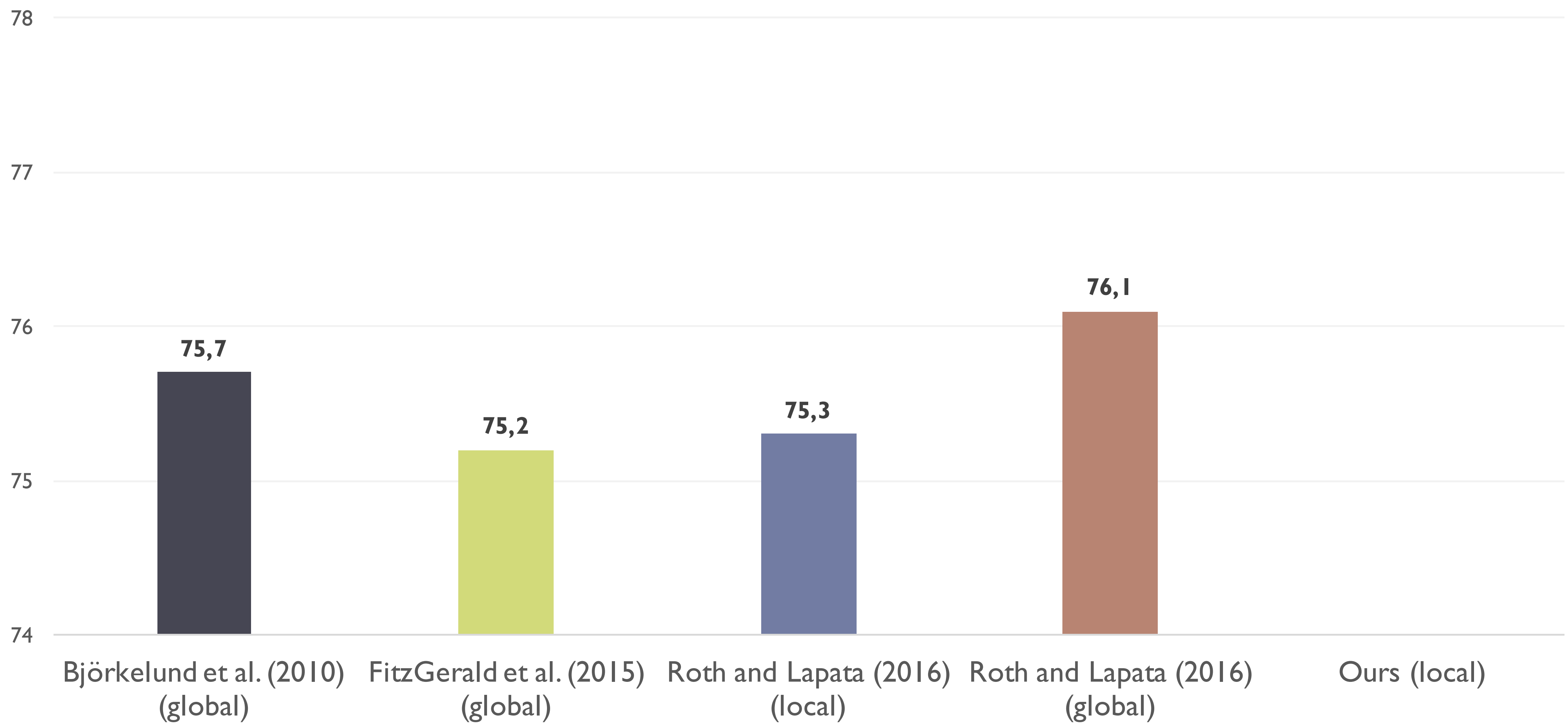


English Test set

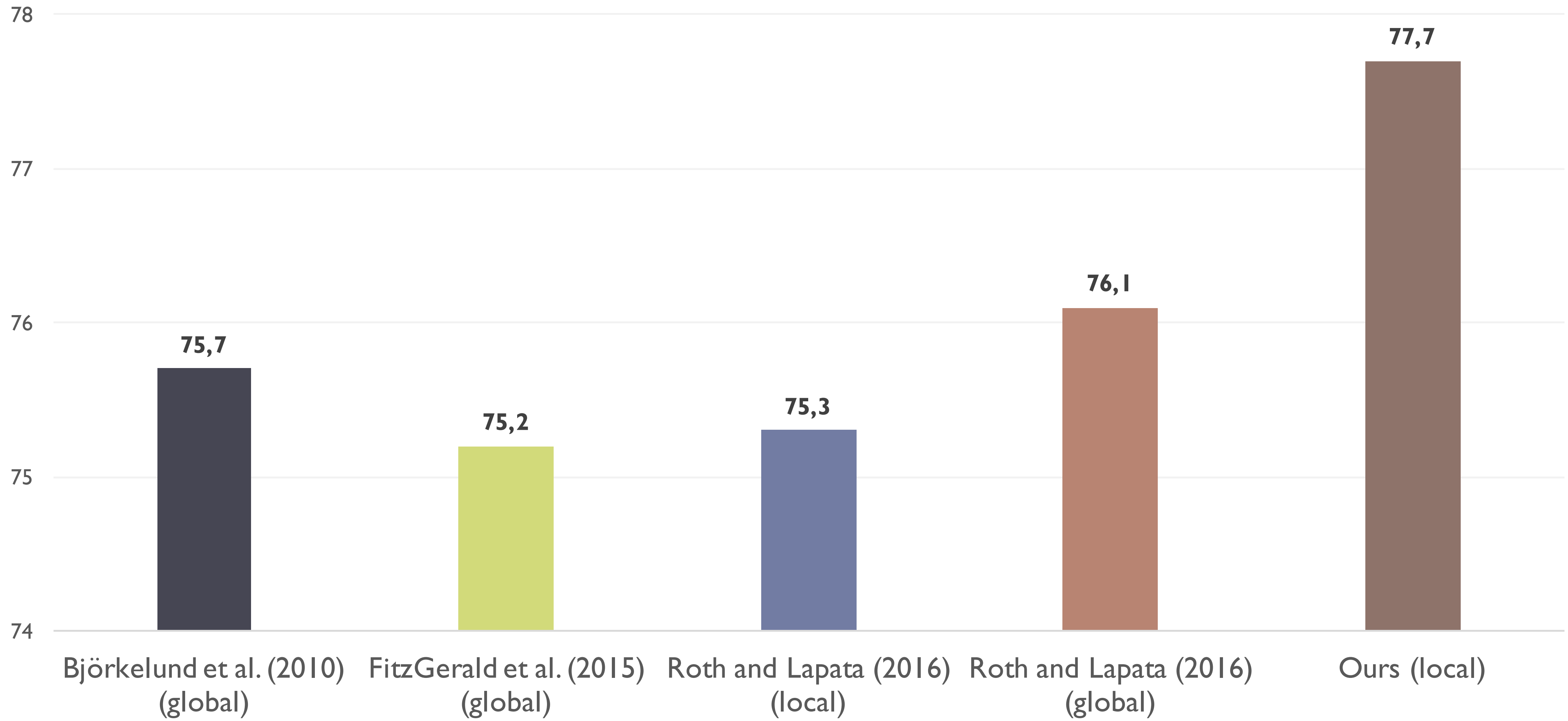


On par with global syntax-rich models

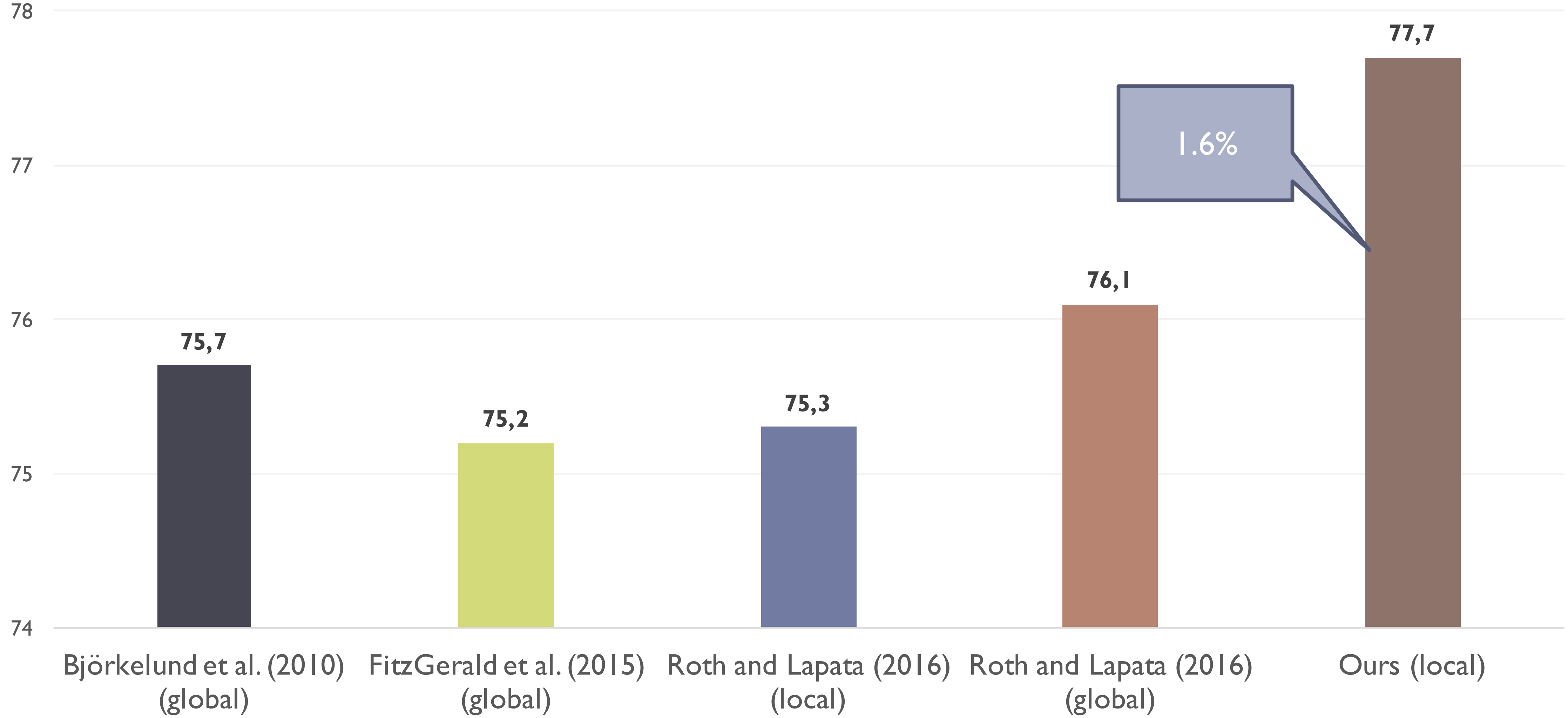
English out-of-domain



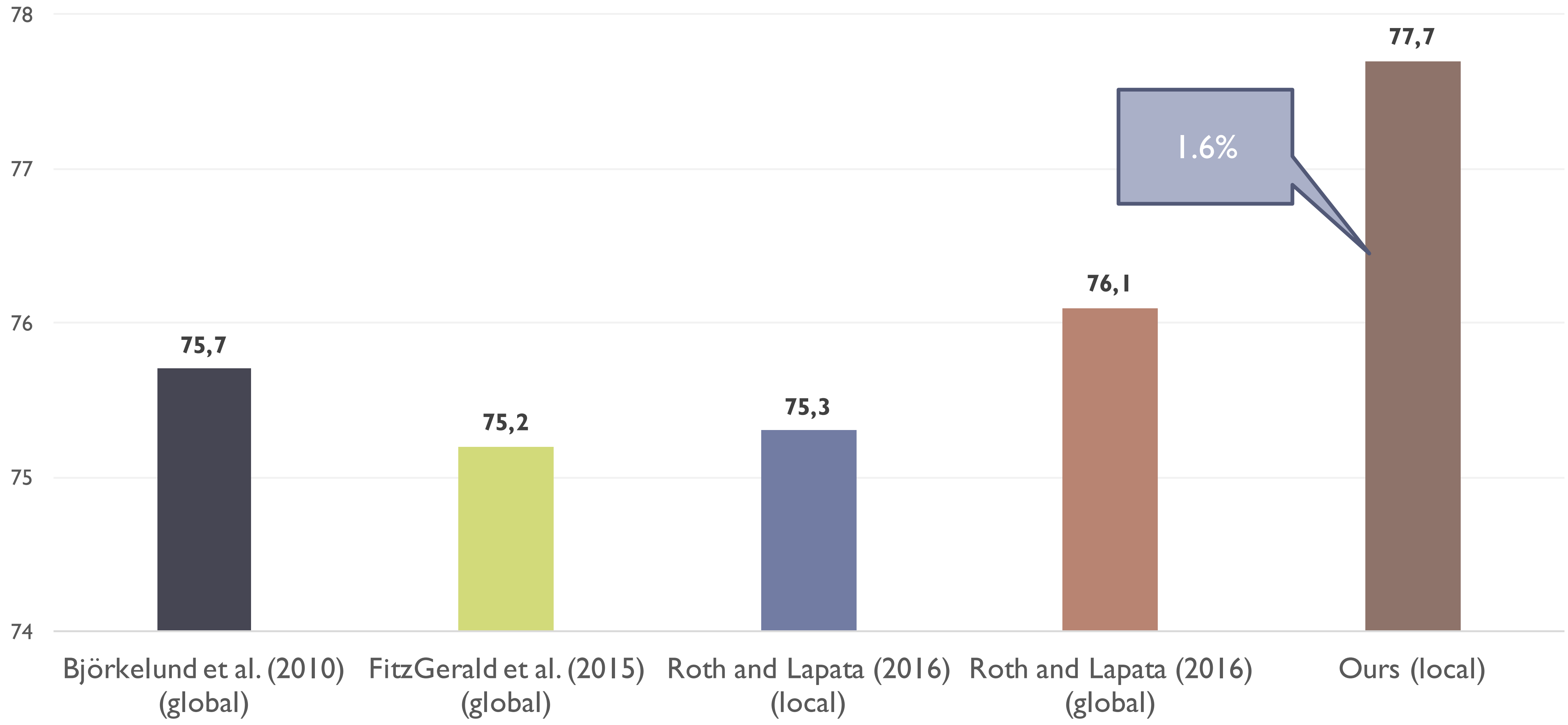
English out-of-domain



English out-of-domain

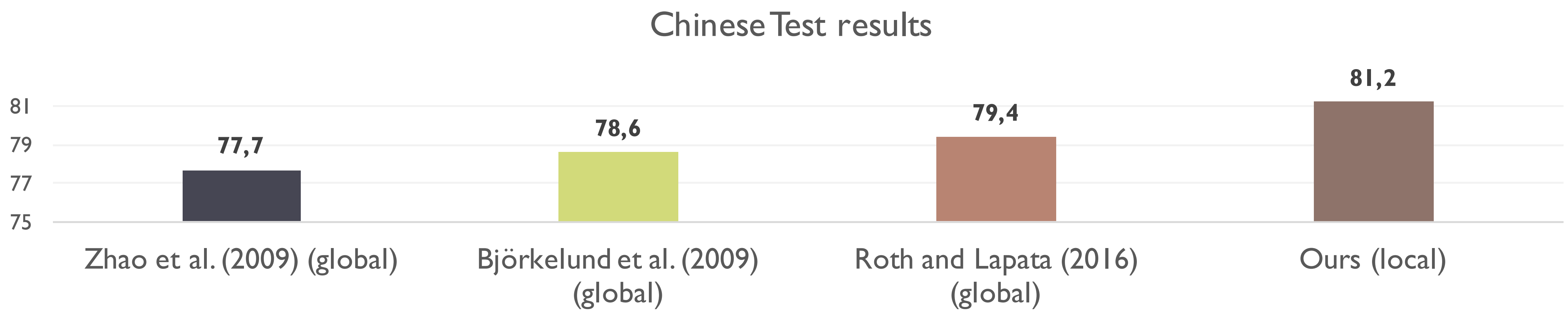


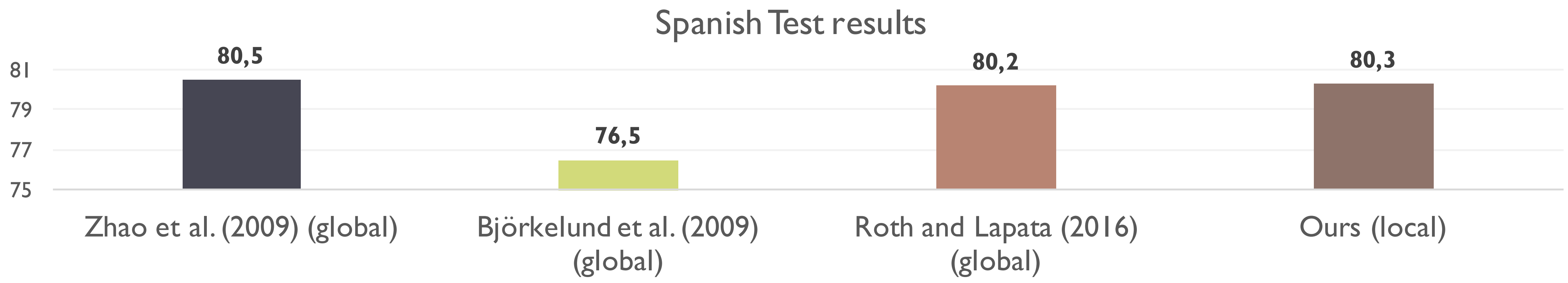
English out-of-domain



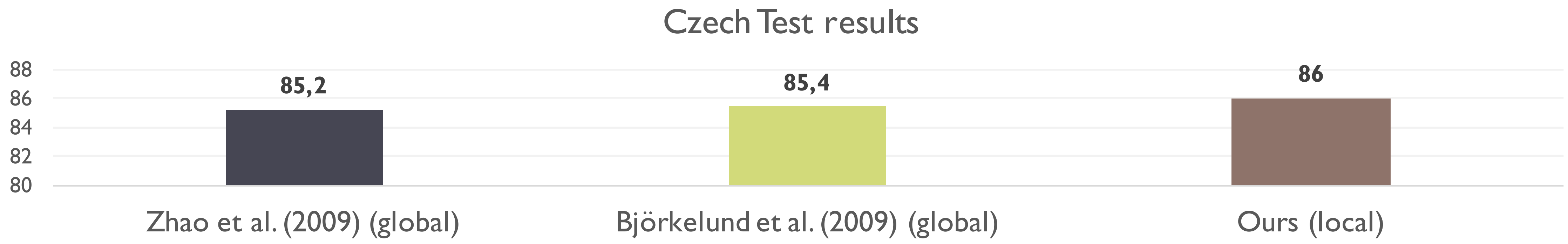
Robust on out-of-domain data

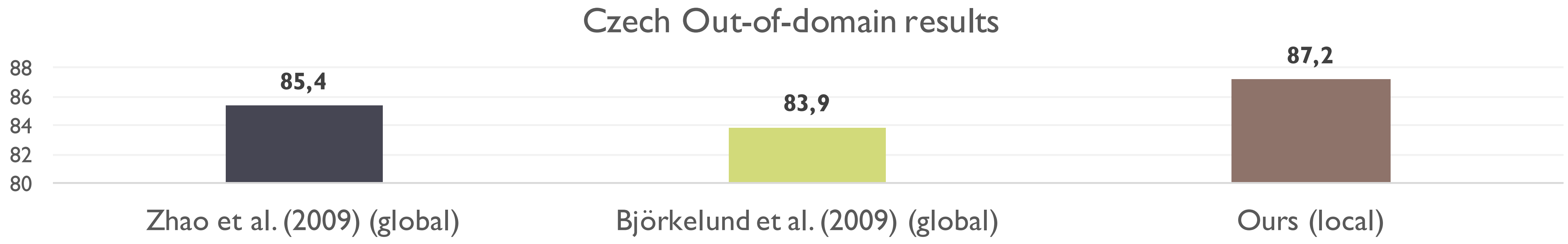
Chinese and Spanish Test set



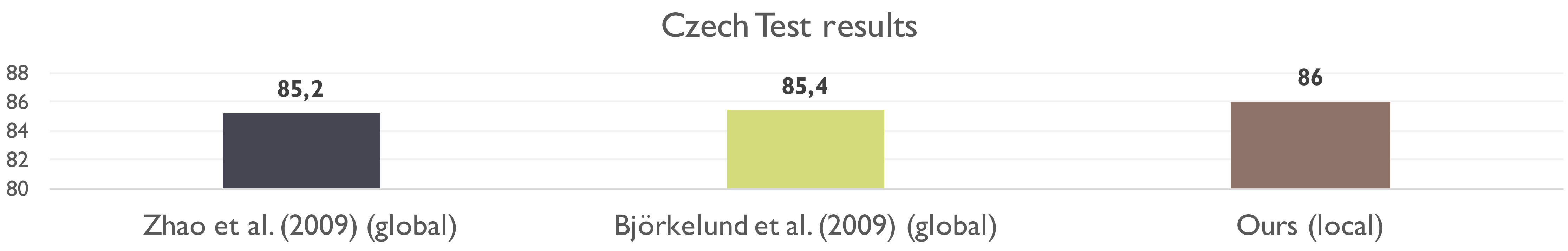


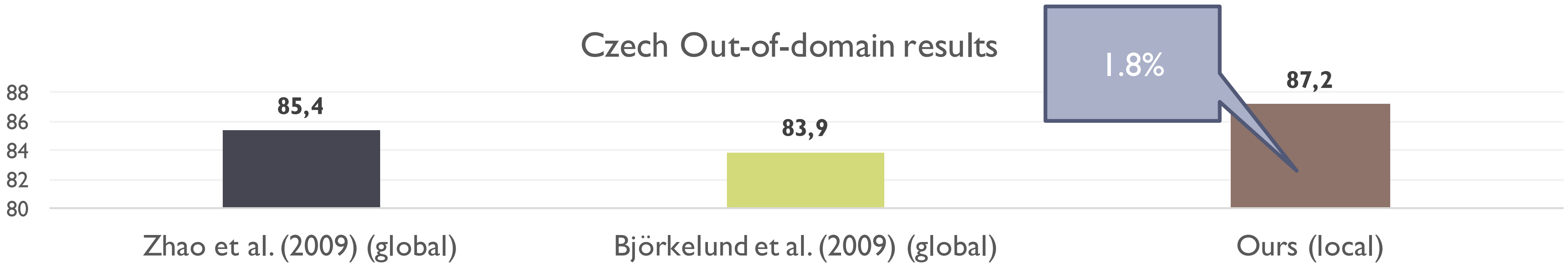
Czech test set and out-of-domain



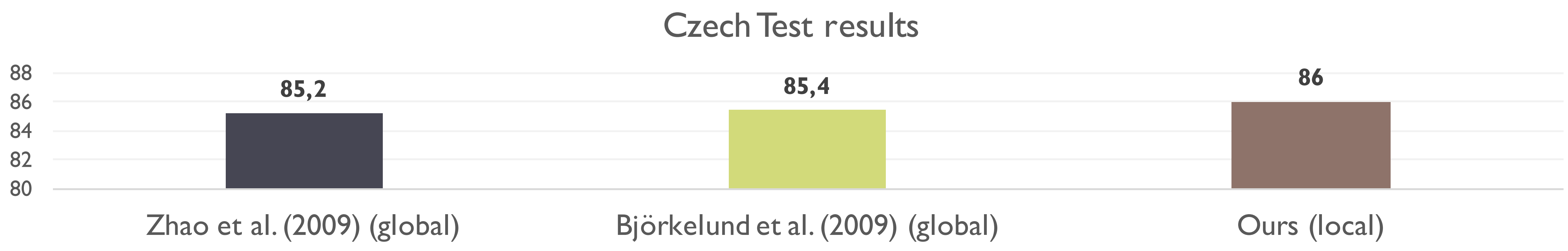


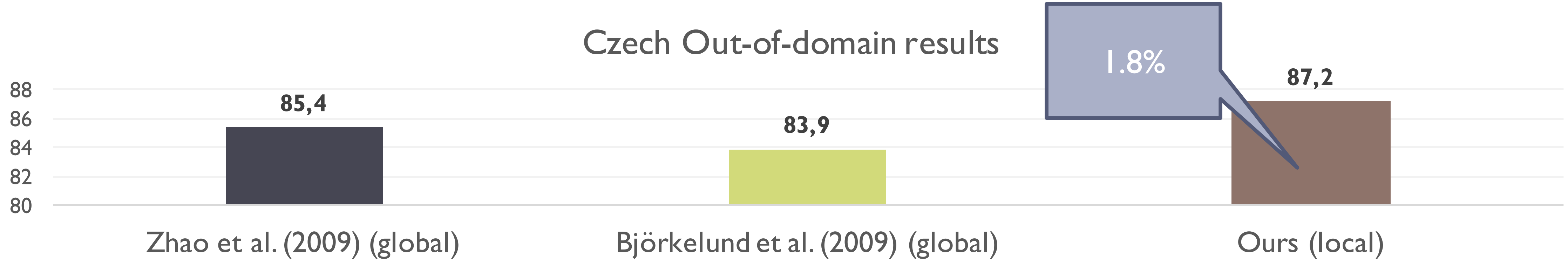
Czech test set and out-of-domain





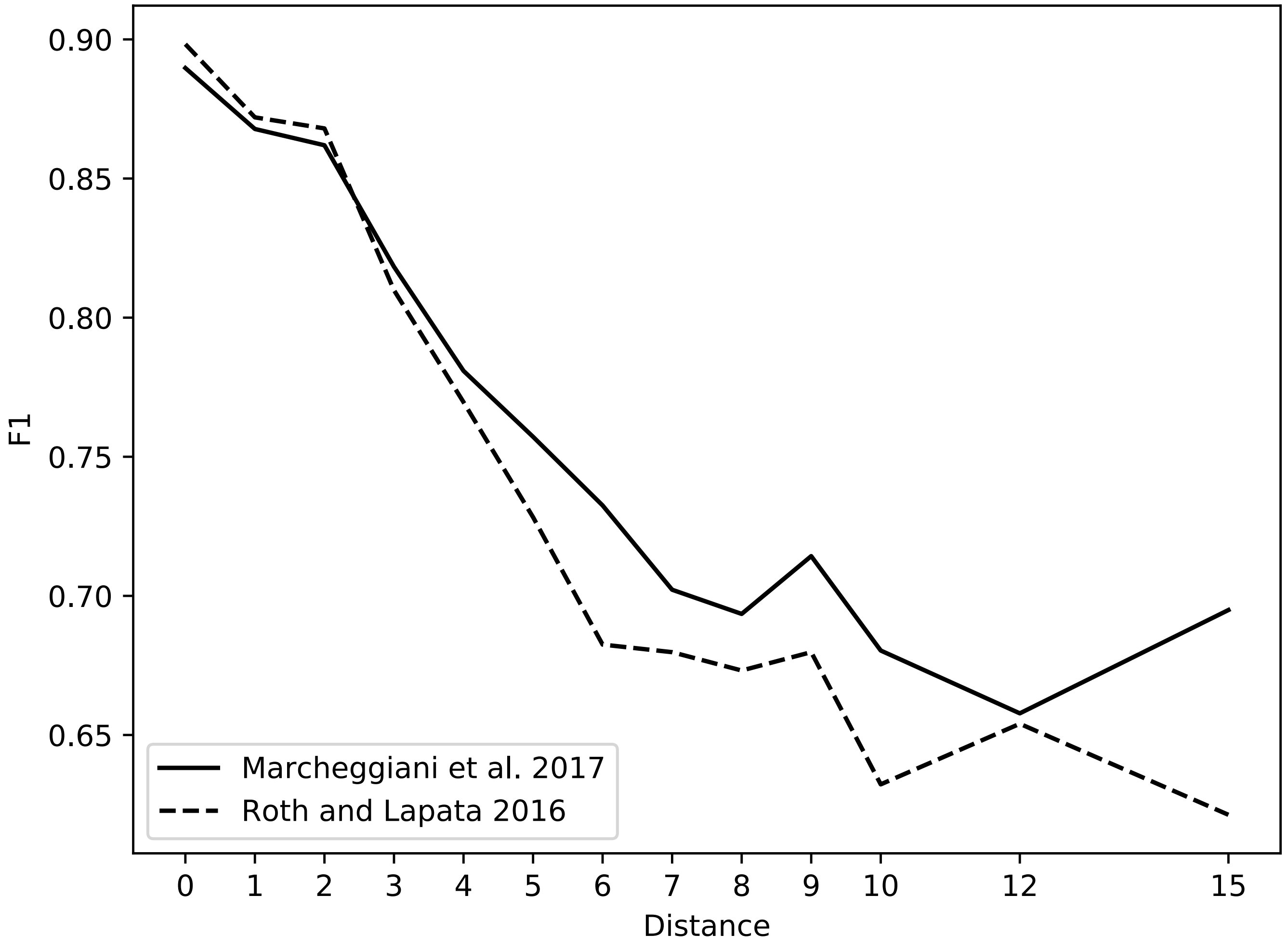
Czech test set and out-of-domain



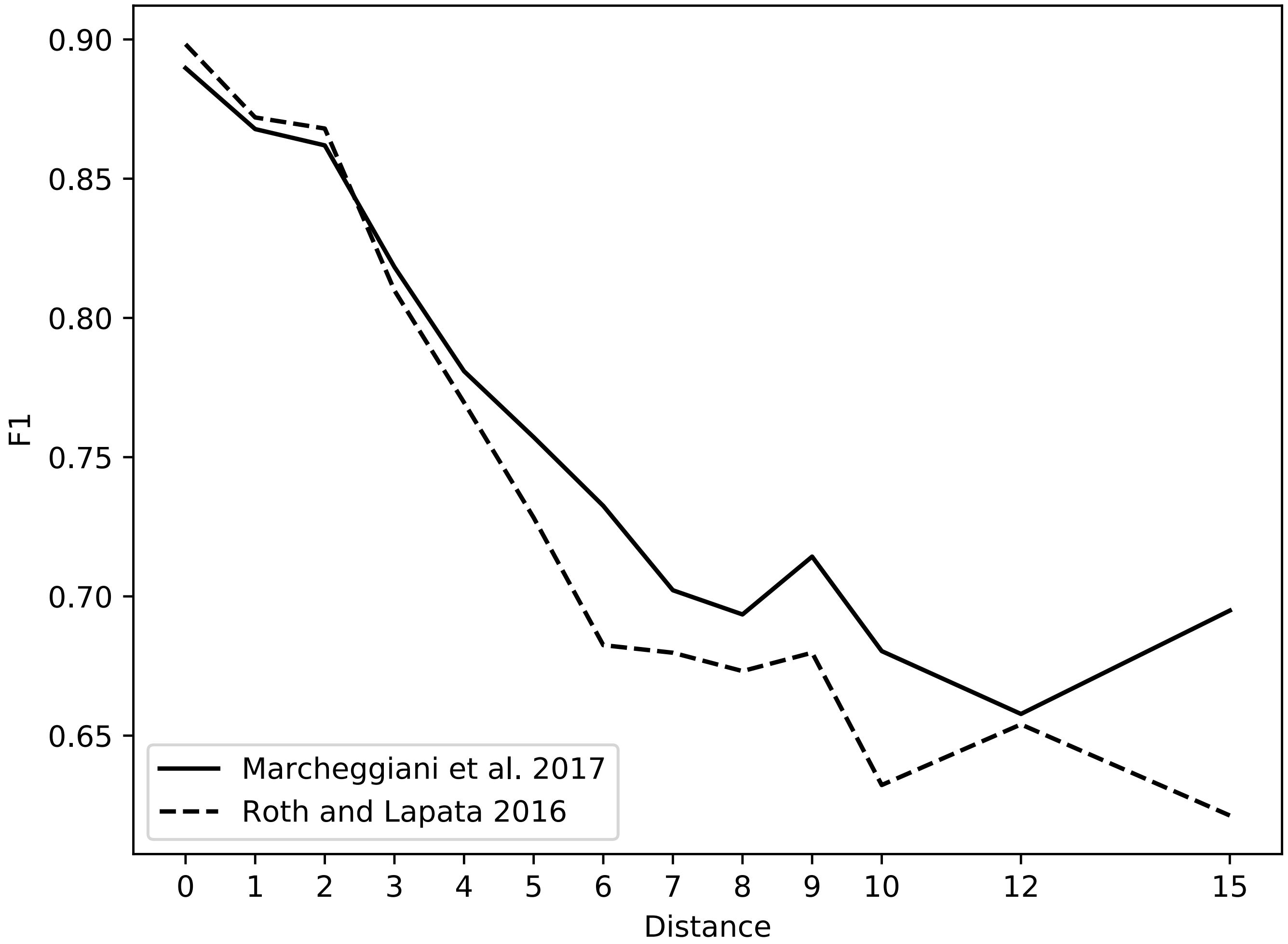


Language independent

Distance analysis



Distance analysis



Long-range dependencies are better captured

Conclusion

- Simple syntax-agnostic dependency-based SRL model
- Very robust on out-of-domain data
- Building block for syntax-aware models (Graph convolutional nets)
    - Marcheggiani and Titov, EMNLP 2017

Conclusion

- Simple syntax-agnostic dependency-based SRL model
- Very robust on out-of-domain data
- Building block for syntax-aware models (Graph convolutional nets)
  - Marcheggiani and Titov, EMNLP 2017

github.com/diegma/neural-dep-srl

- Funding:
  - ERC StG BroadSem 678254
  - NWO VIDI 639.022.518
  - Amazon Web Services (AWS) grant