Towards Better UD Parsing: Deep Contextualized Word Embeddings, Ensemble, and Treebank Concatenation
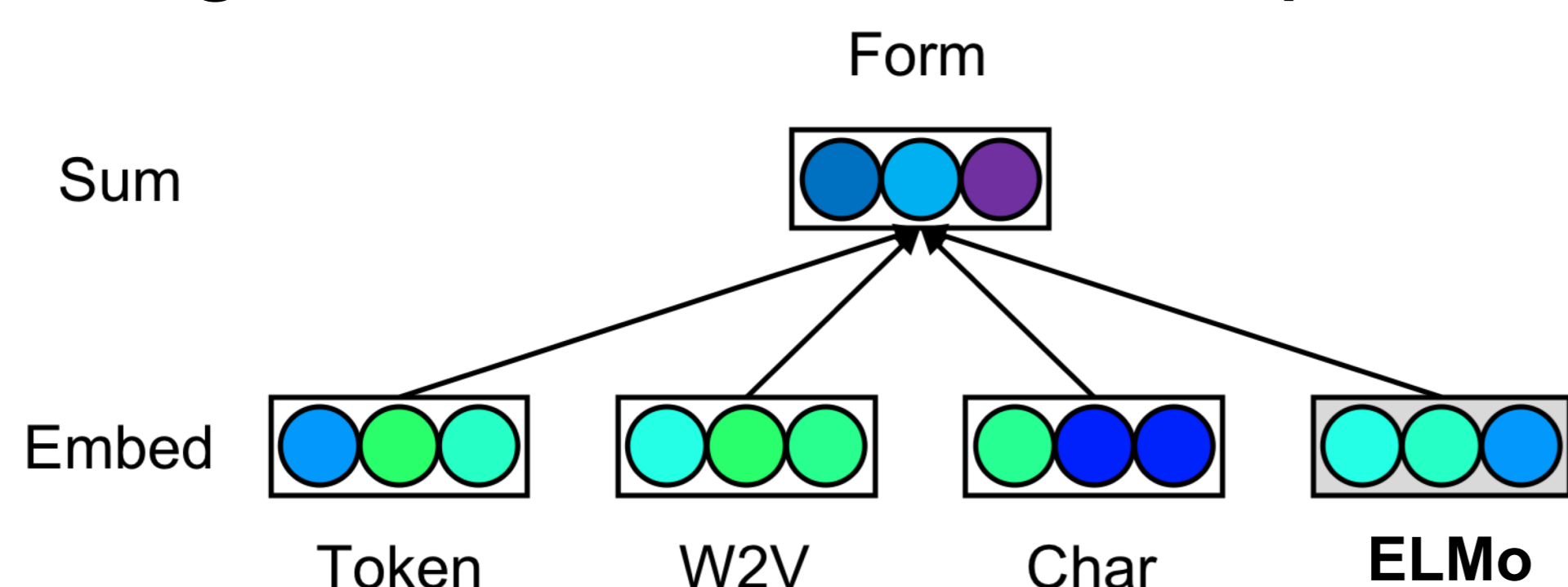
Wanxiang Che · Yijia Liu · Yuxuan Wang · Bo Zheng · Ting Liu
Harbin Institute of Technology

Released ELMo @ https://github.com/HIT-SCIR/ELMoForManyLangs

Extensions to Dozat et al. (2017)

- Using ELMo as additional input

- Using ensemble on both the tagger and parser

Contributions of Each Technique

- ELMo: +0.84
- Ensemble: +0.55
- Treebank Concat.: +0.42 (estimated on Dev set.)

Other Helpful Techniques

- Improve POS tagger with ELMo
- Improve tokenization with ELMo and semi-supervised features (zh_gsd: +6.6, ja_gsd: +4.1, vi_vtb: +7.2)

Training Details for ELMo

- Supporting Unicode range
- Training with *sample softmax*
  - use 8,192 surrounding words as negative samples
  - more stable training and better performance
- Training one language takes 3 days on Nvidia P100

Results

Systems	Word Seg.	UPOS	LAS	MLAS	BLEX
ours	98.12	90.19	**75.84**	59.78	65.33
Uppsala	**98.18**	**90.91**	72.37	59.20	32.09
UDPipe F.	97.04	89.37	73.11	**61.25**	64.49
TurkuNLP	97.42	89.81	73.28	60.99	**66.09**

Effect of Treebank Concatenation

|	Dutch		Swedish	
	apino	lassysmall	lines	talbanken
Single	91.87	86.82	84.64	86.39
Concate.	**92.08**	**89.34**	**85.76**	**86.77**
	Korean		Italian	
	gsd	kaist	isdt	postwita
Single	82.05	**87.83**	**92.01**	80.79
Concate.	**83.73**	87.61	91.80	**82.54**

Final Evaluation of Shared Task

- Rank 1st according to LAS (+2.6 than the 2nd)
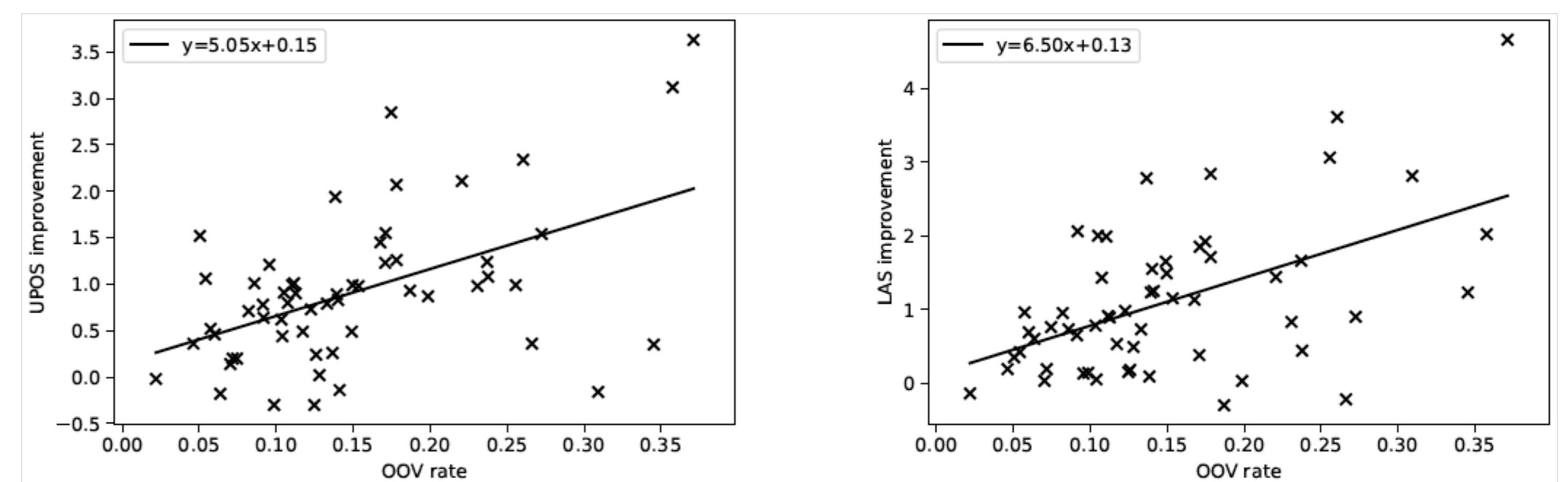
```
1.   HIT-SCIR (Harbin)        75.84 ± 0.14 [OK]  (p<0.001)
2.   TurkuNLP (Turku)         73.28 ± 0.14 [OK]  (p=0.039)
3-5. UDPipe Future (Praha)    73.11 ± 0.13 [OK]  (p=0.221)
3-5. LATTICE (Paris)          73.02 ± 0.14 [OK]  (p=0.461)
3-5. ICS PAS (Warszawa)       73.02 ± 0.14 [OK]  (p<0.001)
6.   CEA LIST (Paris)         72.56 ± 0.14 [OK]  (p=0.036)
```
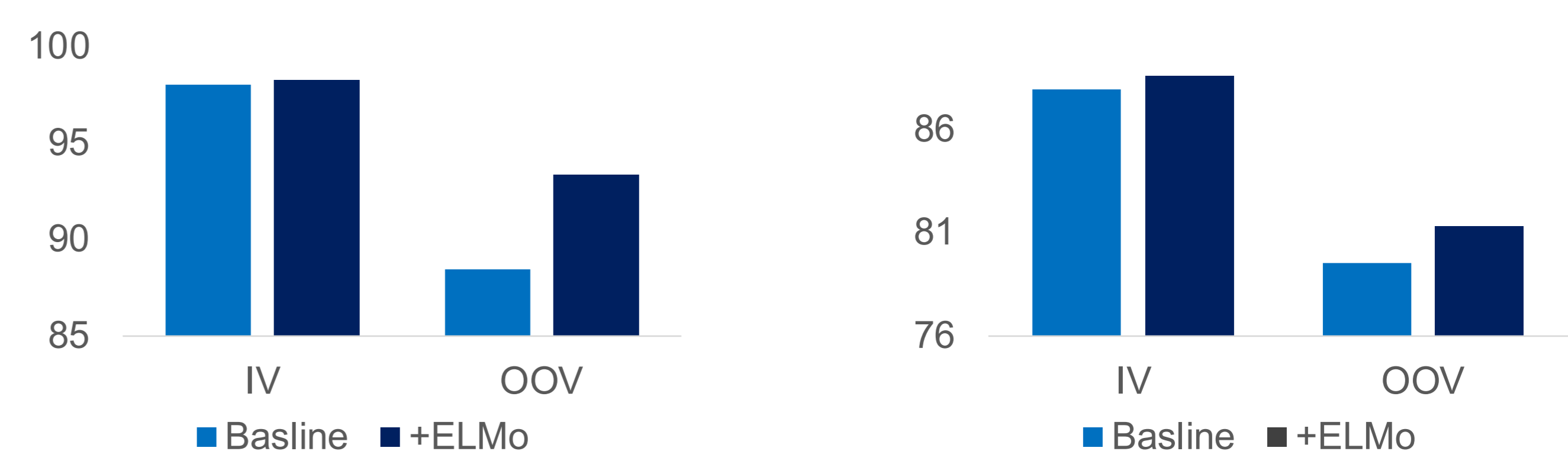
Post-contest Evaluation on ELMo

- On gold segmentation
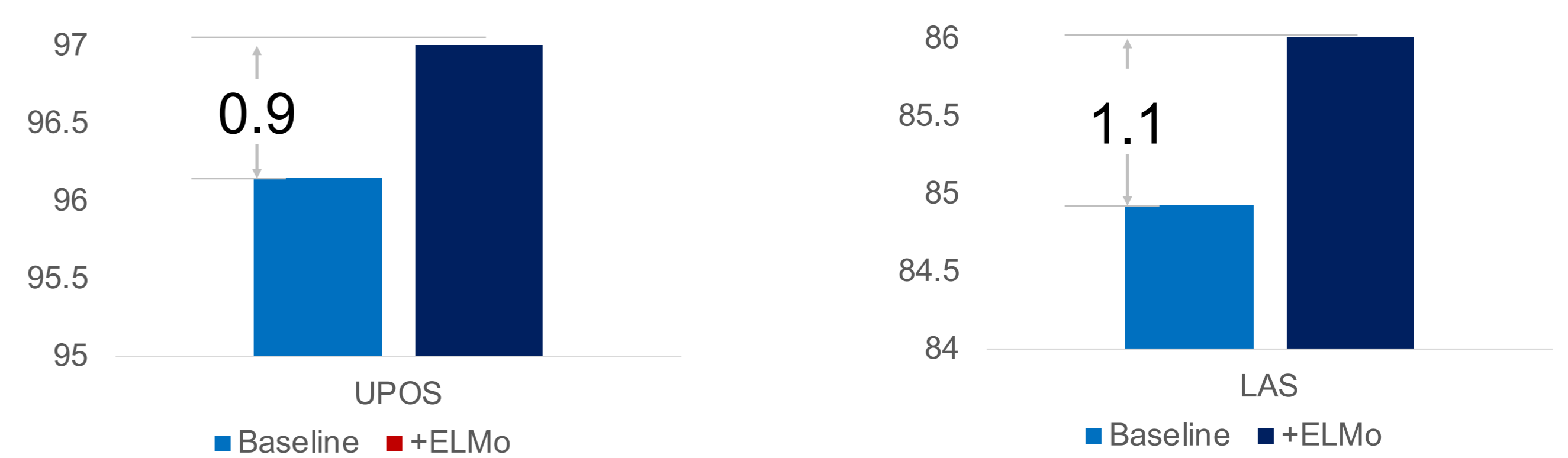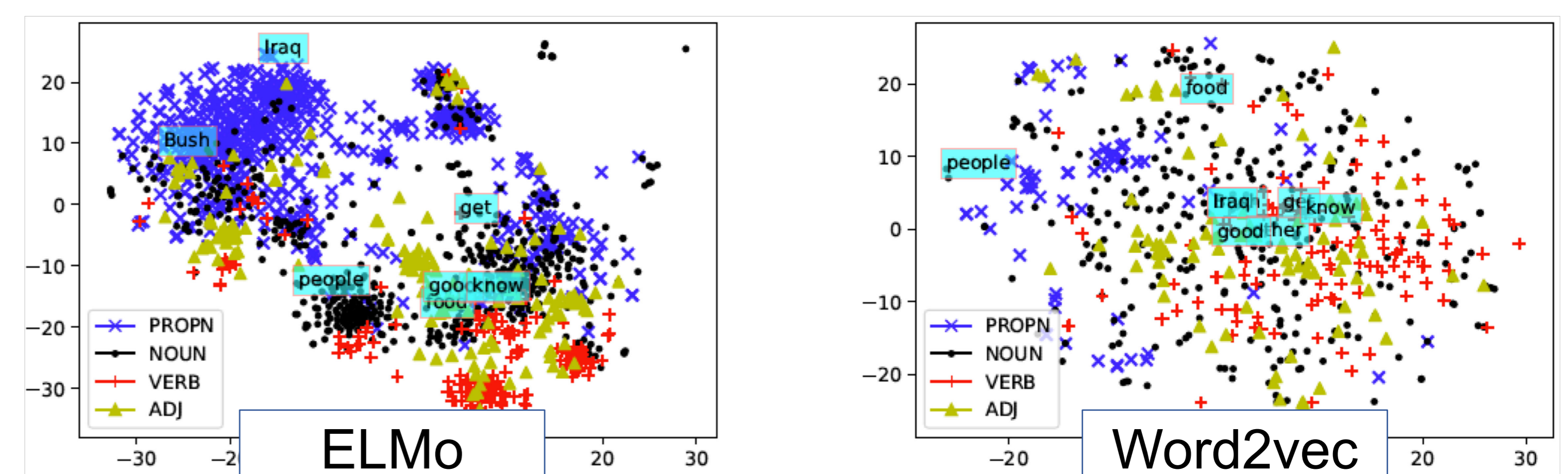

OOV Rate against ELMo Improvements

ELMo on IV and OOV

- ELMo improves more on OOV than IV

OOV Words Visualization

ELMo   Word2vec

Conclusions

- Several enhancement to Dozat et al. (2017) including: ELMo, ensemble, and treebank concatenation.
- ELMo improves the OOV via better abstraction.