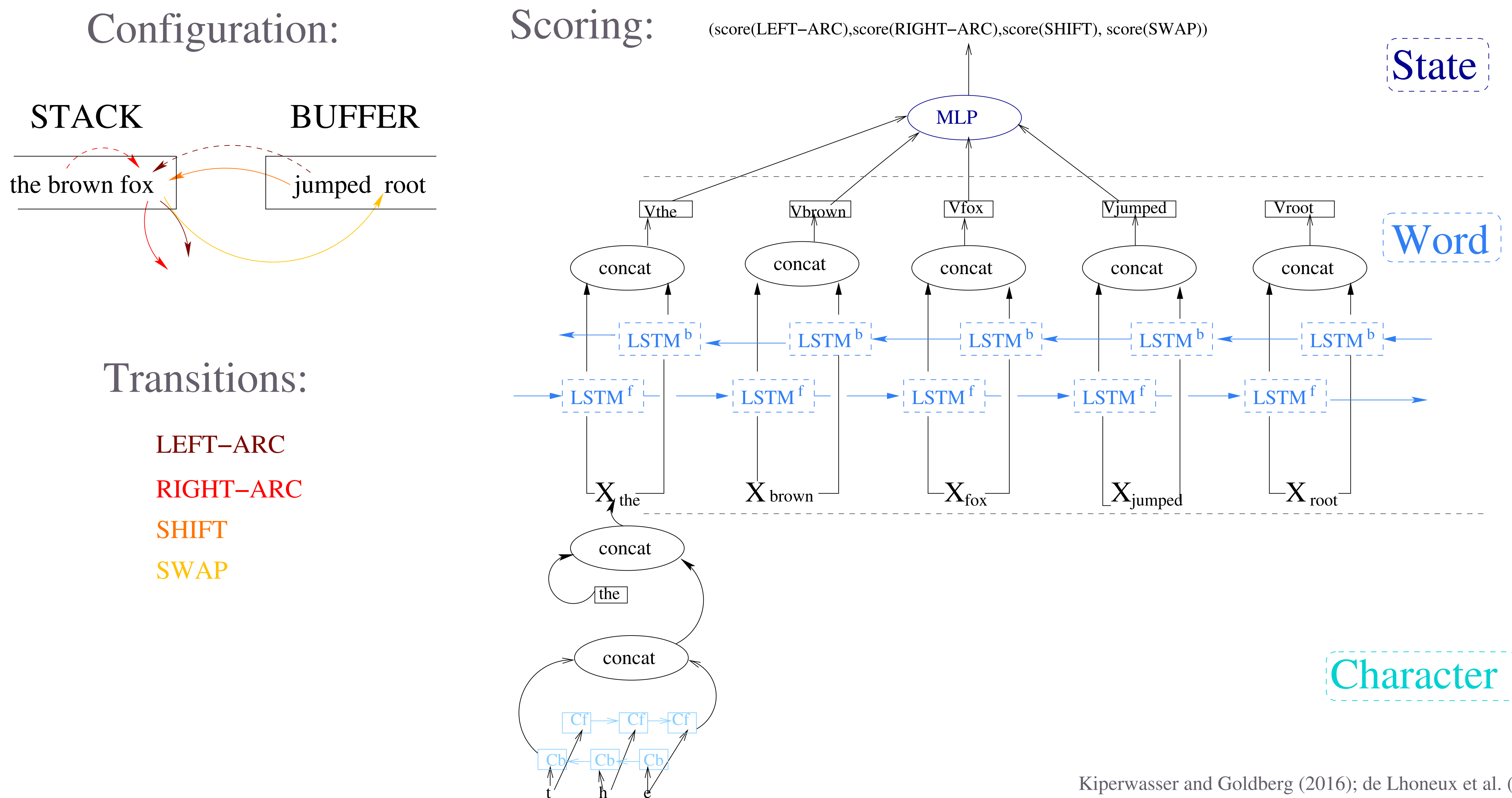


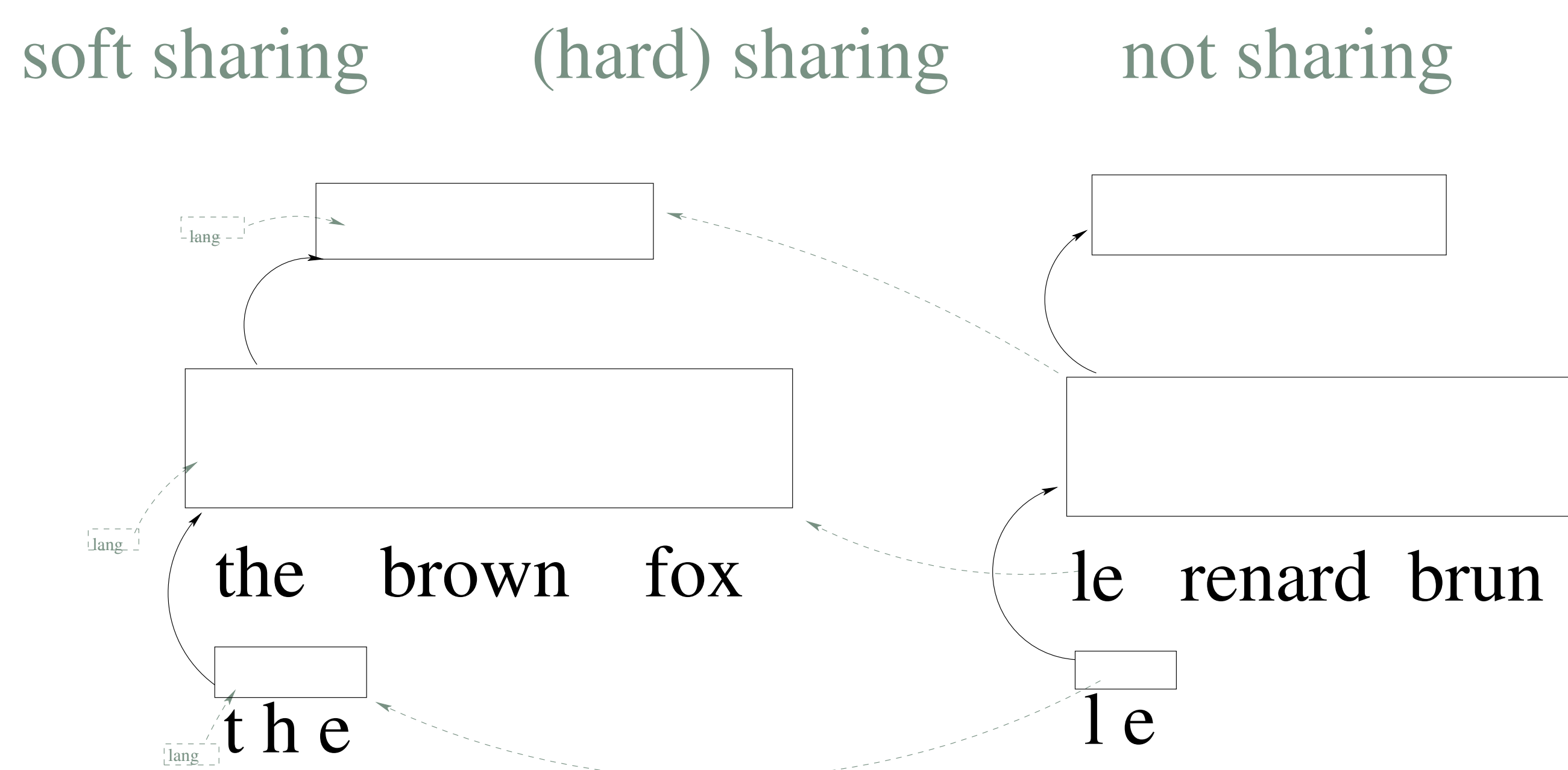


Parsing Architecture



Kiperwasser and Goldberg (2016); de Lhoneux et al. (2017)

Sharing strategies



STRATEGIES:

NOT SHARING: MLP or word/character lookup and BiLSTM parameters not shared
 HARD SHARING: MLP or word/character lookup and BiLSTM parameters shared
 SOFT SHARING: Sharing + concatenating a language embedding to the configuration vector, or word or character vectors at the input of the BiLSTM.

We test the $3^3 = 27$ combinations on 5 language pairs.

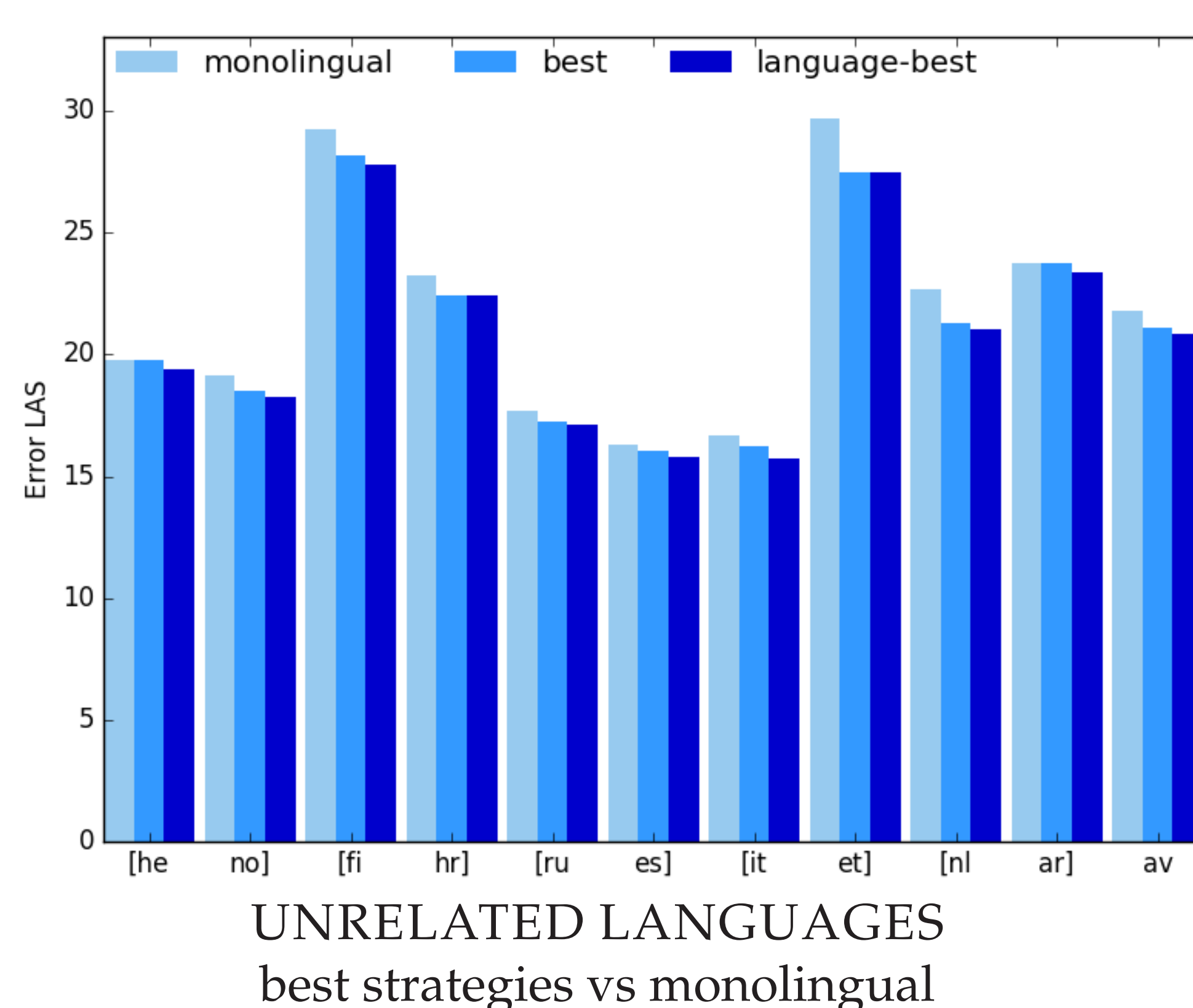
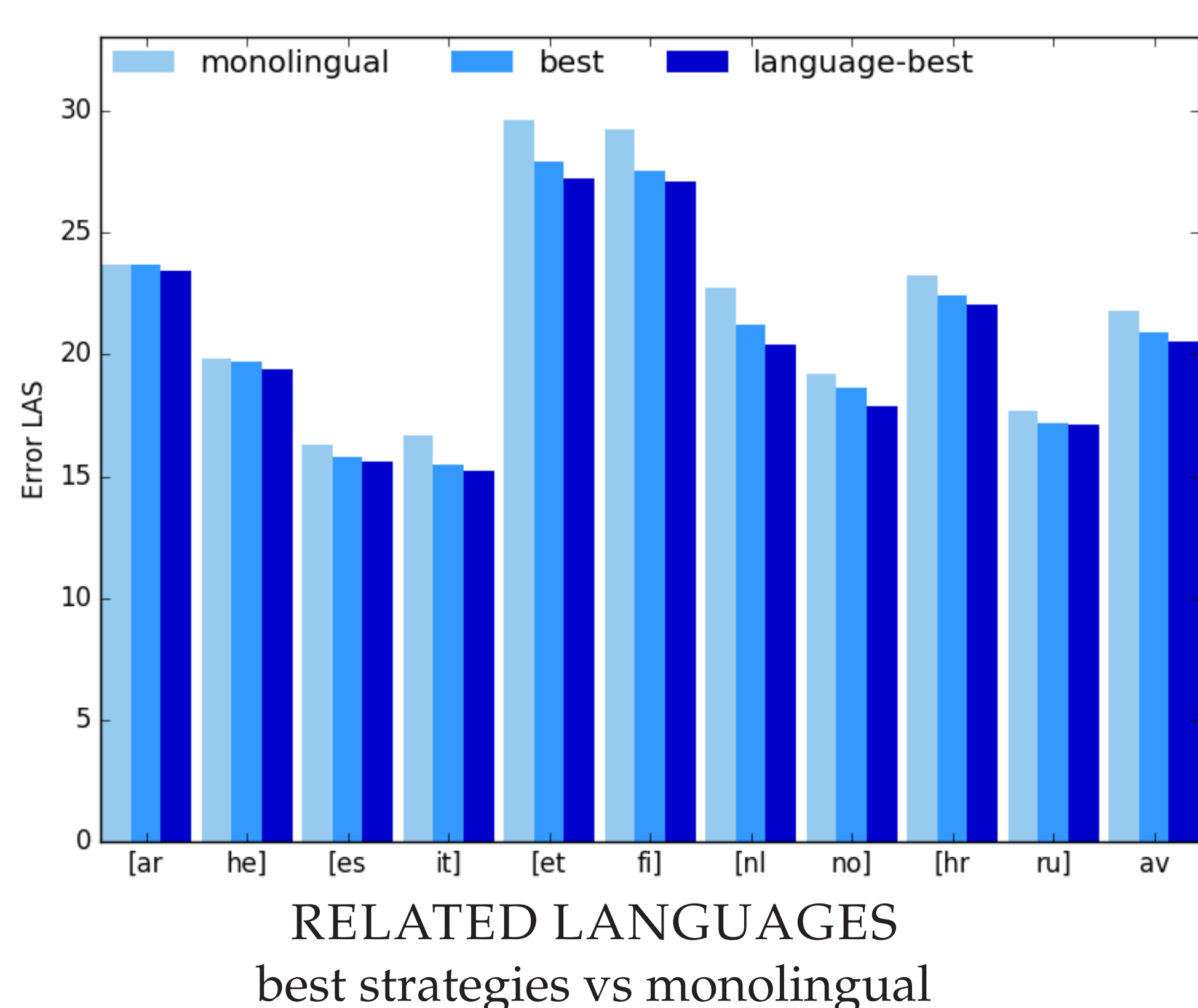
Results

ISO	Lang	Tokens	Family	Word order
ar	Arabic	208,932	Semitic	VSO
he	Hebrew	161,685	Semitic	SVO
et	Estonian	60,393	Finnic	SVO
fi	Finnish	67,258	Finnic	SVO
hr	Croatian	109,965	Slavic	SVO
ru	Russian	90,170	Slavic	SVO
it	Italian	113,825	Romance	SVO
es	Spanish	154,844	Romance	SVO
nl	Dutch	75,796	Germanic	No dom. order
no	Norwegian	76,622	Germanic	SVO

Dataset

	W	C	OURS	MONO	δ
ar	x	x	77.2	77.1	0.1
he	✓	x	80.0	79.8	0.3
et	x	ID	71.4	70.5	0.8
fi	x	x	71.6	71.6	0.1
hr	✓	x	77.9	78.0	-0.1
ru	✓	x	83.5	82.7	0.8
it	ID	✓	85.0	84.0	1.0
es	ID	✓	84.3	83.8	0.5
nl	ID	✓	75.5	74.1	1.4
no	x	ID	81.1	80.1	1.0
av.			78.8	78.2	0.6

LAS on the test sets of the best of 9 sharing strategies and the monolingual baseline. δ is the difference between OURS AND MONO.



Conclusions

- Generally, multi-task learning helps.
- Sharing the MLP parameters always helps. It helps to share MLP parameters when training a parser on a pair of related languages, and it also helps if the languages are unrelated.
- Sharing word and character parameters is differently helpful depending on the language.
- Sharing too many parameters does not help, when the languages are unrelated.

References

Eliyahu Kiperwasser and Yoav Goldberg. 2016. Simple and accurate dependency parsing using bidirectional LSTM feature representations. *Transactions of the Association for Computational Linguistics*, 4:313–327.

Miryam de Lhoneux, Sara Stymne, and Joakim Nivre. 2017. Arc-Hybrid Non-Projective Dependency Parsing with a Static-Dynamic Oracle. In *Proceedings of the 15th International Conference on Parsing Technologies*, pages 99–104, Pisa, Italy. Association for Computational Linguistics.