Improving the performance of UDify with Linguistic Typology Knowledge

Chinmay Choudhary
National University of Ireland, Galway
`c.choudhary1@nuigalway.ie`

Colm O'Riordan
National University of Ireland, Galway
`colm.oriordan@nuigalway.ie`

Abstract

UDify is the state-of-the-art language-agnostic dependency parser which is trained on a polyglot corpus of 75 languages. This multilingual modeling enables the model to generalize over unknown/lesser-known languages, thus leading to improved performance on low-resource languages. In this work we used linguistic typology knowledge available in URIEL database, to improve the cross-lingual transferring ability of UDify even further.

1 Introduction

State-of-the-art approaches to dependency parsing are supervised approaches that require large manually annotated dataset to be trained on, thus limiting their utility to only few high-resource languages. Multilingual modeling which involves training a model on a mixed polyglot corpus of high-resource source-languages and applying it on a low resource target-language, is an effective way to circumvent this issue of data-sparsity.

In a similar way, as the proficiency of a speaker's previous languages can enhance his/her ability to learn a new language (Abu-Rabia and Sanitsky, 2010), a model which is trained on multilingual dataset can learn to generalize over unknown or lesser-known languages.

UDify (Kondratyuk and Straka, 2019) is the state-of-the-art mBERT based language-agnostic dependency parser, which takes the advantage of multilingual modeling to improve its performance on low-resource languages. The authors of UDify (Kondratyuk and Straka, 2019) trained it on a joint polyglot corpus created by concatenating all training treebanks available in UDv2.3, and evaluated it on all test treebanks in UDv2.3 individually. Results outlined in (Kondratyuk and Straka, 2019) show that for dependency parsing task, the UDify outperforms its baseline monolingual UDPipe Future (Straka, 2018) model by a large margin especially for low-resource languages, as the model

benefit significantly from the cross-lingual transfer learning which occurs due to joint polyglot training.

However, the performance of UDify model on the low-resource languages (less represented in the polyglot training corpus) is still much lower than the performance of it on the high-resource languages which are well represented within the training corpus.

In this work, we use linguistic typology knowledge to improve the cross-lingual transferring ability of UDify model even further, thereby significantly reducing this gap between model's performance on high-resource and low-resource languages.

We induce the linguistic typology knowledge available in URIEL (Littell et al., 2017) database into the UDify model by adding an auxiliary task of linguistic typology feature prediction to it, within the multitasking framework. Sections 3 and 4 will describe the model in more details.

2 Related Work

Cross-lingual *Model-transfer* approaches to Dependency Parsing such as (McDonald et al., 2011; Cohen et al., 2011; Duong et al., 2015; Guo et al., 2016; Vilares et al., 2015; Falenska and Çetinoğlu, 2017; Mulcaire et al., 2019; Vania et al., 2019; Shareghi et al., 2019) involve training a model on high-resource languages and subsequently adapting it to low-resource languages. Participants of CoNLL 2017 shared-task (Daniel et al., 2017) and CoNLL 2018 shared task (Zeman et al., 2018) also provide numerous approaches to dependency parsing of low-resource languages.

Some approaches such as (Naseem et al., 2012; Täckström et al., 2013; Barzilay and Zhang, 2015; Wang and Eisner, 2016a; Rasooli and Collins, 2017; Ammar, 2016; Wang and Eisner, 2016b) indeed used linguistic typology to facilitate the cross-lingual transfer between source and target languages. However, all these approaches directly

38

feed the linguistic typology features into the respective model, whereas we induce the linguistic typology knowledge into UDify model through Multitask learning.

Inducing typology knowledge through Multitask learning rather than directly feeding it along with word-embeddings have following advantages.

1. The model can also be applied to low-resource languages for which many typology feature values are unknown/missing.

2. The auxiliary task should help to improve the performance on the main dependency parsing task as well, since it would make the model give special emphasis on the syntactic typology (specially word-order typology) of language being parsed while predicting the dependency relations.

3 UDify

UDify is a multitasking multilingual BERT based model which performs four key language-processing tasks simultaneously namely *UPOS-tagging*, *UFeat-tagging*, *Lemmatization* and *Dependency Parsing*, in a multitasking framework. The model utilizes a single shared mBERT based encoder, and four individual task-specific decoders, for each of the four tasks respectively.

The *mBERT Encoder* takes in the entire sentence as input, tokenizes it using pre-trained WordPiece Tokenizer (Wu et al., 2016) and subsequently outputs mBERT (Wu and Dredze, 2019) based contextualized-embeddings for each word within the input-sentence. We refer to original UDify (Kondratyuk and Straka, 2019) paper for detailed description of mechanism of computing/fine-tuning such contextualized embeddings.

The decoders for both *UPOS-tagging* and *UFeat-tagging* tasks adopt standard sequence-tagging architecture with softmax layer on the top. These decoders accept the contextual embeddings generated from the mBERT Encoder for each word in the input sentence, and predicts its UPOS/Ufeats tag.

For *Lemmatization* task as well, the model uses a standard sequence-tagger which predicts a class-tag representing a unique edit script, for each word. An edit-script is simply the sequence of character operations to transform a word form to its lemma-form.

For dependency-parsing, the model adopts the pop-

ular deep biaffine architecture (Dozat and Manning, 2016) for graph-based parsing, with LSTM-encoder been replaced by the shared *mBERT Encoder*.

4 Linguistic Typology prediction

To improve the cross-lingual transferring ability of UDify model, we added a fifth auxiliary task of Linguistic Typology prediction to it.

Our *Typology-predictor* is a simple deep feed-forward neural network with *sigmoid* activation function, which predicts the values of all typology features provided by the URIEL database (Littell et al., 2017).

URIEL database is a collection of binary features extracted from multiple typological, phylogenetic, and geographical databases such as WALS (Haspelmath, 2009), PHOIBLE (Moran and Richard Wright, 2014), Ethnologue (M. Paul Lewis and Fennig, 2015) and Glottolog (Harald Hammarstrom and Bank, 2015). URIEL database can be accessed through Pyton PyPi library called *lang2vec*[1].

Let $\hat{N}$ be the number of typology features provided by URIEL database. Our *Typology predictor* would then output the probability vector $Pr_{ty} \in R^{\hat{N}}$ by applying equation 1.

$$Pr_{ty} = sigmoid(e_{</s>} * U + c) \qquad (1)$$

Here $e_{</s>} \in R^d$ is the contextual embedding from the shared *mBERT Encoder* for end-token $</s>$ of the input-sentence. $U \in R^{d*\hat{N}}$ and $c \in R^{\hat{N}}$ are weights and biases respectively. Pr_{Ty} comprises of the probability of value of each URIEL binary feature being as 1, for the specific language being parsed.

The total-loss is computed by simply adding the *Typology Predictor* loss to *UDify* model's (as computed in (Kondratyuk and Straka, 2019))

5 Experiments

This section describes the details of experiments conducted to evaluate our proposed model.

5.1 Experimental Setup

Both baseline *UDify* and the proposed *UDify+Typology-predictor* models are trained on a single large joint-polyglot corpus, created by concatenating all training datasets available in

---

[1]https://pypi.org/project/lang2vec/

UDv2.5[2] together.

Before each training-epoch, we randomly shuffled all sentences in our polyglot training corpus, and subsequently fed mixed batches of sentences from this shuffled corpus into the model being trained, where each batch may contain sentences from any language or treebank (as done by authors of UDify ([Kondratyuk and Straka, 2019])).

We used batch-size of 32, drop-out probability of 0.01 and the pre-trained mBERT model *cased_L-12_H-768_A-12* downloaded from tensorflow-hub[3]. We fine-tuned these hyper-parameters on *Dev* dataset for *English-EWT* treebank.
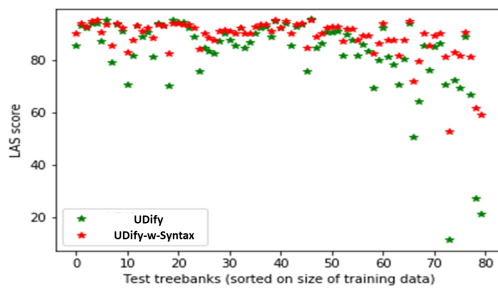

Figure 1: Trends in LAS achieved by *UDify* and *UDify-w-Syntax* models on all 80 test treebanks
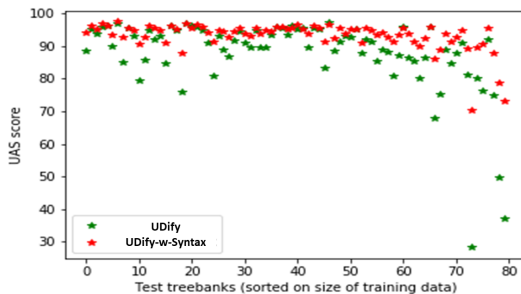

Figure 2: Trends in UAS achieved by *UDify* and *UDify-w-Syntax* models on all 80 test treebanks

6   Results

We evaluated our proposed model on 80 test treebanks available in UDv2.5 datasets individually. Appendix A provides the results achieved on each of these 80 test-treebanks, whereas table 1 outlines the average results on all these 80 treebanks. All scores are evaluated using the official CoNLL 2018 Shared Task evaluation script. We compared

the performance of our model with two baselines namely *UDPipe Fututre*([Straka, 2018]) and *UDify*. URIEL database comprises of three categories of typology features namely *Syntactic*, *Semantic* and *Phonological* features. In this work, we evaluated three variants of our proposed model, based on the categories of features predicted by the typology-predictor within the auxiliary task, namely *UDify-w-Syntax* (predicts only syntactic typology features), *UDify-w-Syntactic+Semantic* (predicts syntactic and semantic typology-features) and *UDify-w-All* (predicts all the URIEL typology-features). Furthermore, we evaluated the performance of *UDify-w-Lang_id* model. The architecture of it is identical to our proposed model but the linguistic-typology predictor is replaced by a simple language-id predictor.

7   Discussion

It is observed that the *UDify-w-Syntax* variant of our proposed model outperforms the other two variants of it, for most of the test-treebanks, despite the fact that the *UDify-w-Syntax+Semantic* and *UDify-w-All* variants utilizes more typology-features than the *UDify-w-Syntax* variant.

The reason being that since all four tasks performed by the UDify model namely *UPOS-tagging*, *UFeats-tagging*, *Lammelization* and *Dependency Parsing* are syntactic tasks, only the syntactic typology-features are relevant to these tasks.

([Henderson, 2004]) proved that, having large number of unrelated features makes it difficult for a neural-network model to effectively learn from provided training-data, and thereby would lead to drop in performance.

Figures 1 and 2 depict the trends in LAS and UAS achieved by the *UDify* and *UDify-w-Syntax* models on all 80 test treebanks. The test-treebanks (as indexes) on the x-axes in these figures are inversely-sorted by the size of their corresponding train treebank, which is part of the training corpus. It is evident in the figures that the *UDify* model shows stronger performance on the high-resource languages which are well represented in the training corpus as compared to the low-resource languages. On the other hand, *UDify+Syntax* shows relatively uniform performance across all languages.

Overall the results in Appendix A, show that *UDify+Syntax* outperforms baselines *UDPipe Future*, *UDify* and *UDify+Lang_id* for almost all 80 test-treebanks. For high-resource languages, the *UD-*

Corpus	Model	UPOS	UFeats	Lemmas	UAS	LAS	Typo F1
Overall (all UDv2.5 test-banks)	UDPipe	94.27	91.37	94.99	86.24	81.78	–
	UDify	94.03	89.33	90.92	87.84	82.83	–
	UDify-w-Lang_id	95.76	90.95	91.52	90.21	85.61	–
	UDify-w-Syntax	**95.89**	**92.05**	**91.87**	**93.18**	**88.4**	**74.6**
	UDify-w-Syntax+Semantic	94.04	88.06	87.09	89.26	83.84	73.33
	UDify-w-All	92.85	85.48	84.33	84.86	79.17	64.88

Table 1: Overall Results achieved by the baseline and all variants of our proposed model. These are average of all results outlayed in Appeandix A.

Corpus	Model	UPOS	UAS	LAS
English-EWT (size: 25377)	UDify	97.73	94.64	90.04
	UDify+	98.32	95.73	91.41
French-GSD (size: 33399)	UDify	98.14	94.74	92.77
	UDify+	99.24	96.19	92.84
Buryat-BDT (size: 19)	UDify	60.23	36.98	21.52
	UDify+	73.73	73.25	59.1
Lithuanian-HSE (size: 2494)	UDify	90.47	80.1	70.38
	UDify+	93.56	90.14	81.6

Table 2: Selected results from Appendix A. **UDify+** refers to *UDify+Syntax* model

Distrb 1	Distrb 2	t-value	p-value
Typo F1	Diff	84.23	3.24e-23
Typo F1	Size	6.98	7.36e-11
Typo F1	UDify	1.42	0.16
Typo F1	UDify+	3.49	6.26e-4

Table 3: Results of t-test for correlation between various performance parmeters. **Typo F1**: Its F1 score achieved by *UDify+Syntax* for auxiliary task.; **UDify, UDify+**:UAS achieved by *UDify* and *UDify+Syntax* models; **Diff**: Improvement in UAS of *UDify+* over *UDify*

Corpus	Model	UPOS	UAS	LAS
Breton-KEB	UDify	63.67	63.97	40.19
	UDify+	62.15	60.65	34.23
Tagalog-TRG	UDify	61.64	64.73	39.38
	UDify+	62.38	63.9	38.31
Faroese-OFT	UDify	77.86	69.28	61.03
	UDify+	77.46	65.57	54.11
Naija-NSC	UDify	56.59	47.13	33.43
	UDify+	55.06	46.61	27.94
Sanskrit-UFAL	UDify	40.21	41.73	19.8
	UDify+	38.08	43.14	15.48

Table 4: Results achieved in zero-shot learning scenario. **UDify+** refers to *UDify+Syntax* model

ify+Syntax model shows only marginal improvement in performance over *UDify* whereas for low-resource languages it shows strong improvement in performance. Such trends can also be observed in Table 2. Table 4 outlines results obtained on selected languages which are not represented in the training data at all (zero-shot learning). For such treebanks, *UDify+Syntax* under-performs *UDify*. Hence it can be inferred that the auxiliary task of linguistic typology prediction, does lead to significant improvement in performance of *UDify* within few-shot learning scenario, but does not lead to any improvement within zero-shot learning scenario.

Furthermore, to ensure that the auxiliary task of linguistic typology-prediction is indeed responsible for the improvement in performance of *UDify*, we conducted numerous statistical t-tests to find the correlation between F1 scores achieved by the *UDify+Syntax* model for the auxiliary-task of typology-prediction, and various other performance parameters including the improvement in performance of *UDify+Syntax* over *UDify*. Table 3 outlays results of these t-tests.

8 Conclusion

In this work we used linguistic typology knowledge available in URIEL database to improve the cross-lingual transferring ability of the state-of-the-art language-agnostic UDify parser. We injected typology knowledge in UDify model through an auxiliary task, in multitasking settings.

References

Salim Abu-Rabia and Ekaterina Sanitsky. 2010. Advantages of bilinguals over monolinguals in learning a third language. *Bilingual Research Journal*, 33(2):173–199.

Waleed Ammar. 2016. *Towards a Universal Analyzer of Natural Languages*. Ph.D. thesis, Ph. D. thesis, Google Research.

Regina Barzilay and Yuan Zhang. 2015. Hierarchical low-rank tensors for multilingual transfer parsing. Association for Computational Linguistics.

Shay B Cohen, Dipanjan Das, and Noah A Smith. 2011. Unsupervised structure prediction with non-parallel multilingual guidance. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 50–61.

Zeman Daniel, Popel Martin, Straka Milan, Hajic Jan, Nivre Joakim, Ginter Filip, Luotolahti Juhani, Pyysalo Sampo, Petrov Slav, Potthast Martin, et al. 2017. Conll 2017 shared task: Multilingual parsing from raw text to universal dependencies. In *CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, volume 1, pages 1–19. Association for Computational Linguistics.

Timothy Dozat and Christopher D Manning. 2016. Deep biaffine attention for neural dependency parsing. *arXiv preprint arXiv:1611.01734*.

Long Duong, Trevor Cohn, Steven Bird, and Paul Cook. 2015. Low resource dependency parsing: Cross-lingual parameter sharing in a neural network parser. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 845–850.

Agnieszka Falenska and Özlem Çetinoğlu. 2017. Lexicalized vs. delexicalized parsing in low-resource scenarios. In *Proceedings of the 15th International Conference on Parsing Technologies*, pages 18–24.

Jiang Guo, Wanxiang Che, David Yarowsky, Haifeng Wang, and Ting Liu. 2016. A representation learning framework for multi-source transfer parsing. In *Thirtieth AAAI Conference on Artificial Intelligence*.

Martin Haspelmath Harald Hammarstrom, Robert Forkel and Sebastian Bank. 2015. Glottolog 2.6.

Martin Haspelmath. 2009. *The typological database of the World Atlas of Language Structures*. Berlin: Walter de Gruyter.

James Henderson. 2004. Discriminative training of a neural network statistical parser. In *ACL'04: Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics*, pages 95–102. Association for Computational Linguistics.

Dan Kondratyuk and Milan Straka. 2019. 75 languages, 1 model: Parsing universal dependencies universally. *arXiv preprint arXiv:1904.02099*.

Patrick Littell, David R Mortensen, Ke Lin, Katherine Kairis, Carlisle Turner, and Lori Levin. 2017. Uriel and lang2vec: Representing languages as typological, geographical, and phylogenetic vectors. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 8–14.

Gary F. Simons M. Paul Lewis and Charles D. Fennig. 2015. Ethnologue: Languages of the World, Eighteenth edition.

Ryan McDonald, Slav Petrov, and Keith B Hall. 2011. Multi-source transfer of delexicalized dependency parsers.

Daniel McCloy Moran, Steven and editors Richard Wright. 2014. PHOBIA Online.

Phoebe Mulcaire, Jungo Kasai, and Noah A Smith. 2019. Low-resource parsing with crosslingual contextualized representations. *arXiv preprint arXiv:1909.08744*.

Tahira Naseem, Regina Barzilay, and Amir Globerson. 2012. Selective sharing for multilingual dependency parsing. The Association for Computational Linguistics.

Mohammad Sadegh Rasooli and Michael Collins. 2017. Cross-lingual syntactic transfer with limited resources. *Transactions of the Association for Computational Linguistics*, 5:279–293.

Ehsan Shareghi, Yingzhen Li, Yi Zhu, Roi Reichart, and Anna Korhonen. 2019. Bayesian learning for neural dependency parsing. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3509–3519.

Milan Straka. 2018. Udpipe 2.0 prototype at conll 2018 ud shared task. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 197–207.

Oscar Täckström, Ryan McDonald, and Joakim Nivre. 2013. Target language adaptation of discriminative transfer parsers.

Clara Vania, Yova Kementchedjhieva, Anders Søgaard, and Adam Lopez. 2019. A systematic comparison of methods for low-resource dependency parsing on genuinely low-resource languages. *arXiv preprint arXiv:1909.02857*.

David Vilares, Carlos Gómez-Rodríguez, and Miguel A Alonso. 2015. One model, two languages: training bilingual parsers with harmonized treebanks. *arXiv preprint arXiv:1507.08449*.

Dingquan Wang and Jason Eisner. 2016a. The galactic dependencies treebanks: Getting more data by synthesizing new languages. *Transactions of the Association for Computational Linguistics*, 4:491–505.

Dingquan Wang and Jason Eisner. 2016b. The galactic dependencies treebanks: Getting more data by synthesizing new languages. *Transactions of the Association for Computational Linguistics*, 4:491–505.

Shijie Wu and Mark Dredze. 2019. Beto, bentz, becas: The surprising cross-lingual effectiveness of bert. *arXiv preprint arXiv:1904.09077.*

Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, et al. 2016. Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144.*

Daniel Zeman, Jan Hajic, Martin Popel, Martin Potthast, Milan Straka, Filip Ginter, Joakim Nivre, and Slav Petrov. 2018. Conll 2018 shared task: Multilingual parsing from raw text to universal dependencies. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual parsing from raw text to universal dependencies*, pages 1–21.

A Results

This section outlines the results obtained by the three variants of our proposed models namely *UDify-w-Syntax* (predicts only syntactic typology features), *UDify-w-Syntactic+Semantic* (predicts syntactic and semantic typology-features) and *UDify-w-All* (predicts all the URIEL typology-features), as well as the baselines.

Table 1: Results achieved on all 80 test tree-banks

Corpus	Model	UPOS	UFeats	Lemma	UAS	LAS	Typo F1
	Begin of Table						
Afrikaans-AfriBooms (size: 1315)	UDPipe	98.25	97.66	97.46	91.26	88.46	–
	UDify	95.31	91.34	94.5	88.79	85.17	–
	Multi-w-Lang_id	96.61	92.64	94.84	90.15	87.87	–
	Multi-w-Syntax	96.73	93.51	95.04	94.36	89.96	82.27
	Multi-w-Syntax+Semantic	94.8	90.51	88.31	83.91	90.63	74.82
	Multi-w-All	93.73	88.8	86.48	81.5	85.96	64.94
Arabic-PADT (size: 21864)	UDPipe	96.83	94.11	95.28	88.29	83.69	–
	UDify	95.35	99.35	99.97	88.6	84.42	–
	Multi-w-Lang_id	96.64	99.33	99.66	89.92	87.13	–
	Multi-w-Syntax	96.76	99.31	99.59	93.78	89.24	81.75
	Multi-w-Syntax+Semantic	96.41	92.77	94.39	90.83	84.12	74.5
	Multi-w-All	96.16	89.76	90.53	86.93	81.51	69.26
Armenian-ArmTDP (size: 1975)	UDPipe	93.49	82.85	92.86	79.65	72.3	–
	UDify	94.42	76.9	85.63	87.01	79.99	–
	Multi-w-Lang_id	96.02	80.58	86.62	89.06	84.3	–
	Multi-w-Syntax	96.15	83.06	87.19	93.43	86.23	83.55
	Multi-w-Syntax+Semantic	92.3	82.87	86.93	85.35	84.2	71.52
	Multi-w-All	91.5	81.24	85.22	79.53	81.21	60.95
Basque-BDT (size: 5396)	UDPipe	96.11	92.48	96.29	86.8	83.55	–
	UDify	95.45	86.8	90.53	85.47	81.5	–
	Multi-w-Lang_id	96.71	88.85	91.16	88.02	85.28	–
	Multi-w-Syntax	95.45	94.95	98.46	92.96	87.4	88.3

	Multi-w-Syntax+Semantic	95.58	87.18	82.9	88.37	81.11	79.26
	Multi-w-All	93.56	84.17	80.34	84.61	79.0	73.01
Belarusian-HSE (size: 319)	UDPipe	93.63	73.3	87.34	80.44	74.58	–
	UDify	96.12	88.36	93.97	91.08	88.59	–
	Multi-w-Lang_id	97.01	95.77	96.72	93.69	89.9	–
	Multi-w-Syntax	97.13	96.22	96.83	95.59	92.26	83.94
	Multi-w-Syntax+Semantic	96.64	92.73	89.73	91.63	92.88	73.98
	Multi-w-All	95.56	90.76	88.24	86.3	88.83	66.82
Bulgarian-BTB (size: 8907)	UDPipe	98.98	97.82	97.94	95.21	92.18	–
	UDify	96.7	96.57	95.1	95.7	92.58	–
	Multi-w-Lang_id	97.54	97.01	95.4	95.05	92.26	–
	Multi-w-Syntax	97.64	97.3	95.57	96.61	93.46	82.07
	Multi-w-Syntax+Semantic	95.06	93.48	87.06	98.7	94.63	74.77
	Multi-w-All	93.42	92.38	84.14	93.69	92.16	69.3
Buryat-BDT (size: 19)	UDPipe	40.34	32.4	58.17	34.07	20.3	–
	UDify	61.73	47.45	61.03	49.61	27.46	–
	Multi-w-Lang_id	73.25	47.9	61.09	56.98	41.08	–
	Multi-w-Syntax	73.73	54.74	62.8	74.42	58.5	82.69
	Multi-w-Syntax+Semantic	72.56	53.23	59.49	77.17	47.95	71.76
	Multi-w-All	70.76	49.35	56.68	73.23	44.38	63.12
Catalan-AnCora (size: 13123)	UDPipe	98.88	98.37	99.07	95.12	92.96	–
	UDify	98.89	98.34	98.14	95.61	93.69	–
	Multi-w-Lang_id	99.0	98.49	98.22	95.9	93.96	–
	Multi-w-Syntax	99.08	98.58	98.26	96.97	93.55	81.77
	Multi-w-Syntax+Semantic	97.47	95.05	95.29	91.97	90.02	71.06
	Multi-w-All	96.12	93.31	92.87	86.69	87.62	60.62
Chinese-GSD (size: 7994)	UDPipe	94.88	99.22	99.99	85.84	81.7	–
	UDify	93.48	99.31	100.0	92.98	84.66	–
	Multi-w-Lang_id	97.46	93.0	75.43	90.79	86.76	–
	Multi-w-Syntax	97.57	93.83	76.49	94.61	89.19	60.82

2

	Multi-w-Syntax+Semantic	95.17	93.62	70.49	86.34	91.08	49.85
	Multi-w-All	94.17	91.89	67.93	81.95	88.31	38.65
Coptic-Scriptorium (size: 792)	UDPipe	94.7	96.35	95.49	87.4	82.79	–
	UDify	27.17	52.85	55.71	28.29	11.53	–
	Multi-w-Lang_id	51.24	60.49	58.89	51.29	32.13	–
	Multi-w-Syntax	52.06	65.65	60.7	71.54	53.73	84.55
	Multi-w-Syntax+Semantic	50.62	59.7	59.62	71.52	49.73	75.59
	Multi-w-All	48.95	56.29	57.39	68.19	43.67	64.17
Croatian-SET (size: 6914)	UDPipe	98.13	92.25	97.27	92.45	88.13	–
	UDify	97.89	88.97	97.15	92.98	90.5	–
	Multi-w-Lang_id	98.33	90.66	97.3	94.29	92.07	–
	Multi-w-Syntax	98.42	91.8	97.38	95.65	92.08	81.92
	Multi-w-Syntax+Semantic	97.08	86.49	92.89	95.4	81.97	72.68
	Multi-w-All	96.44	83.45	90.91	90.35	75.32	61.19
Czech-CAC (size: 102993)	UDPipe	99.37	96.34	98.57	93.48	91.2	–
	UDify	98.14	96.55	97.18	94.74	92.77	–
	Multi-w-Lang_id	98.5	96.99	97.33	93.9	92.84	–
	Multi-w-Syntax	98.59	97.29	97.41	96.04	93.82	82.61
	Multi-w-Syntax+Semantic	96.63	94.17	94.72	97.75	87.74	76.35
	Multi-w-All	96.29	90.57	92.35	91.46	85.34	65.28
Czech-CLTT (size: 102993)	UDPipe	98.88	91.59	98.25	87.86	84.99	–
	UDify	99.17	93.66	98.86	93.7	91.97	–
	Multi-w-Lang_id	99.18	94.58	98.88	94.14	91.71	–
	Multi-w-Syntax	99.26	95.19	98.9	95.13	93.7	82.49
	Multi-w-Syntax+Semantic	98.64	91.41	91.24	86.51	89.79	74.57
	Multi-w-All	96.66	88.17	87.53	80.78	86.25	66.55
Czech-FicTree (size: 102993)	UDPipe	98.55	95.87	98.63	93.32	90.16	–
	UDify	98.18	96.36	97.33	95.77	93.98	–
	Multi-w-Lang_id	98.52	96.83	97.47	95.9	93.27	–
	Multi-w-Syntax	98.61	97.15	97.54	95.3	93.52	82.45

3

	Multi-w-Syntax+Semantic	97.04	95.16	88.41	89.48	94.11	75.34
	Multi-w-All	95.34	92.47	84.58	83.23	87.94	64.9
Czech-PDT (size: 102993)	UDPipe	99.18	97.23	99.02	94.94	92.92	–
	UDify	98.21	98.38	97.55	96.27	93.99	–
	Multi-w-Lang_id	98.54	98.52	97.67	96.08	93.1	–
	Multi-w-Syntax	98.63	98.61	97.74	95.69	94.8	81.59
	Multi-w-Syntax+Semantic	95.22	96.9	94.12	86.46	96.18	71.77
	Multi-w-All	93.41	93.3	91.73	83.01	92.49	65.57
Danish-DDT (size: 4383)	UDPipe	97.78	97.33	97.52	88.25	85.68	–
	UDify	96.02	89.78	91.0	89.76	85.52	–
	Multi-w-Lang_id	97.09	91.34	91.6	92.53	87.77	–
	Multi-w-Syntax	97.2	92.39	91.94	93.76	89.87	82.18
	Multi-w-Syntax+Semantic	96.56	90.73	85.33	93.11	83.86	73.52
	Multi-w-All	95.36	89.03	83.13	87.91	80.49	64.36
Dutch-Alpino (size: 18051)	UDPipe	96.83	96.33	97.09	93.13	90.14	–
	UDify	97.12	92.59	98.23	95.82	92.15	–
	Multi-w-Lang_id	97.82	93.69	98.3	96.25	92.69	–
	Multi-w-Syntax	97.92	94.42	98.34	96.59	93.22	82.11
	Multi-w-Syntax+Semantic	97.58	92.81	90.05	87.22	91.38	74.53
	Multi-w-All	96.59	91.16	88.05	82.43	87.64	64.24
Dutch-LassySmall (size: 18051)	UDPipe	96.5	96.42	97.41	91.82	88.01	–
	UDify	98.89	96.18	93.49	95.73	92.59	–
	Multi-w-Lang_id	99.0	96.68	93.91	96.14	93.95	–
	Multi-w-Syntax	99.08	97.02	94.14	96.05	94.27	82.29
	Multi-w-Syntax+Semantic	96.1	90.53	86.95	85.32	85.91	71.28
	Multi-w-All	94.1	89.36	84.08	82.57	80.57	60.34
English-EWT (size: 25377)	UDPipe	96.29	97.1	98.25	91.21	88.55	–
	UDify	97.73	96.12	95.84	94.64	90.04	–
	Multi-w-Lang_id	98.22	96.63	96.09	93.9	90.07	–
	Multi-w-Syntax	98.32	96.97	96.22	94.76	91.65	81.84

4

	Multi-w-Syntax+Semantic	95.0	96.22	92.23	88.32	81.7	72.3
	Multi-w-All	94.52	95.07	90.36	81.53	75.82	66.63
English-GUM (size: 25377)	UDPipe	96.02	96.82	96.85	88.4	85.25	–
	UDify	95.44	94.12	93.15	91.01	87.6	–
	Multi-w-Lang_id	96.7	94.96	93.59	92.8	89.74	–
	Multi-w-Syntax	96.82	95.53	93.84	93.3	91.27	82.38
	Multi-w-Syntax+Semantic	96.33	88.65	85.01	91.19	88.07	71.87
	Multi-w-All	95.15	87.52	83.36	88.24	85.02	66.05
English-LinES (size: 25377)	UDPipe	96.91	96.31	96.45	84.79	80.35	–
	UDify	94.55	90.43	94.42	89.56	85.34	–
	Multi-w-Lang_id	96.11	91.88	94.77	91.71	88.13	–
	Multi-w-Syntax	96.23	92.86	94.97	93.51	89.89	82.15
	Multi-w-Syntax+Semantic	92.8	88.52	86.98	83.74	87.1	72.22
	Multi-w-All	92.35	86.44	84.13	79.28	82.34	63.23
English-ParTUT (size: 25377)	UDPipe	96.1	95.51	97.74	91.53	88.51	–
	UDify	96.16	92.61	96.45	94.72	92.02	–
	Multi-w-Lang_id	97.18	93.7	96.65	94.19	92.97	–
	Multi-w-Syntax	97.29	94.43	96.76	94.66	93.07	82.21
	Multi-w-Syntax+Semantic	95.24	87.42	89.92	93.54	87.18	74.59
	Multi-w-All	94.78	84.1	86.6	90.09	82.66	65.91
Estonian-EDT (size: 25749)	UDPipe	97.64	96.23	95.3	88.52	85.7	–
	UDify	96.91	87.45	77.73	91.65	86.97	–
	Multi-w-Lang_id	97.68	89.39	79.3	92.51	87.78	–
	Multi-w-Syntax	98.16	97.34	95.68	93.24	89.49	94.0
	Multi-w-Syntax+Semantic	95.95	88.72	74.83	90.1	86.44	84.33
	Multi-w-All	93.85	85.02	72.23	84.23	84.05	76.18
Finnish-FTB (size: 27198)	UDPipe	96.65	96.62	95.49	90.89	88.1	–
	UDify	94.37	82.8	96.68	88.8	83.21	–
	Multi-w-Lang_id	95.99	85.51	96.86	90.97	85.49	–
	Multi-w-Syntax	96.12	87.33	96.97	94.4	89.35	82.17

5

	Multi-w-Syntax+Semantic	94.63	85.13	95.54	83.78	86.72	72.11
	Multi-w-All	94.1	83.61	93.65	79.29	82.58	66.5
Finnish-TDT (size: 27198)	UDPipe	97.45	95.43	91.45	90.67	88.25	–
	UDify	94.43	90.48	82.89	86.8	82.41	–
	Multi-w-Lang_id	96.03	91.92	84.08	89.67	85.5	–
	Multi-w-Syntax	96.16	92.89	84.76	92.58	89.15	82.76
	Multi-w-Syntax+Semantic	94.19	87.24	77.24	90.98	88.75	71.47
	Multi-w-All	93.01	83.9	75.86	87.38	83.01	65.01
French-GSD (size: 33399)	UDPipe	97.63	97.13	98.35	91.77	89.18	–
	UDify	99.14	95.42	98.32	94.77	92.85	–
	Multi-w-Lang_id	99.16	96.05	98.38	94.18	92.07	–
	Multi-w-Syntax	99.24	96.47	98.42	95.57	93.09	82.08
	Multi-w-Syntax+Semantic	98.52	93.7	97.28	91.49	92.14	71.13
	Multi-w-All	97.28	91.12	93.64	84.77	87.26	64.84
French-ParTUT (size: 33399)	UDPipe	96.93	94.43	95.7	93.97	91.43	–
	UDify	95.91	95.08	96.52	92.24	88.65	–
	Multi-w-Lang_id	97.8	91.27	92.16	93.67	90.63	–
	Multi-w-Syntax	97.91	92.33	92.47	94.05	91.75	78.07
	Multi-w-Syntax+Semantic	96.92	90.9	88.81	89.36	88.05	66.92
	Multi-w-All	95.29	89.88	84.99	85.83	86.02	58.57
French-Sequoia (size: 33399)	UDPipe	98.79	98.09	98.57	93.84	92.2	–
	UDify	98.11	95.92	95.5	93.15	90.27	–
	Multi-w-Lang_id	98.48	96.47	95.77	93.74	90.82	–
	Multi-w-Syntax	98.57	96.83	95.92	94.37	91.27	82.08
	Multi-w-Syntax+Semantic	95.2	90.06	86.83	88.32	86.32	75.4
	Multi-w-All	93.9	86.61	84.06	82.46	84.06	68.37
French-Spoken (size: 33399)	UDPipe	95.91	100.0	96.92	83.08	77.71	–
	UDify	96.23	98.67	96.59	86.42	81.19	–
	Multi-w-Lang_id	97.23	98.76	96.78	90.28	84.23	–
	Multi-w-Syntax	97.34	98.82	96.89	93.39	88.49	81.8

6

	Multi-w-Syntax+Semantic	94.08	94.21	87.61	82.96	81.89	75.73
	Multi-w-All	93.18	92.38	83.82	78.53	76.85	69.6
Galician-CTG (size: 2872)	UDPipe	97.84	99.83	98.58	86.66	84.04	–
	UDify	96.51	97.1	97.08	84.88	81.02	–
	Multi-w-Lang_id	97.41	97.45	97.23	88.55	83.49	–
	Multi-w-Syntax	97.52	97.68	97.32	92.77	88.59	82.22
	Multi-w-Syntax+Semantic	96.75	93.95	89.53	91.93	89.52	75.0
	Multi-w-All	95.31	92.51	85.92	87.76	83.59	69.39
Galician-TreeGal (size: 2872)	UDPipe	95.82	93.96	97.06	83.26	78.23	–
	UDify	94.59	80.67	94.93	85.52	78.21	–
	Multi-w-Lang_id	96.13	83.73	95.24	88.31	81.67	–
	Multi-w-Syntax	96.26	85.79	95.42	92.63	85.98	81.8
	Multi-w-Syntax+Semantic	94.16	84.84	88.28	89.43	84.42	69.56
	Multi-w-All	93.12	82.96	85.17	82.64	81.41	58.73
German-GSD (size: 166849)	UDPipe	94.48	90.68	96.8	87.17	82.71	–
	UDify	97.48	96.63	95.23	88.64	85.15	–
	Multi-w-Lang_id	98.06	97.06	95.52	91.49	86.49	–
	Multi-w-Syntax	97.78	90.7	80.19	93.92	89.53	69.14
	Multi-w-Syntax+Semantic	97.85	91.07	94.13	91.44	83.6	59.88
	Multi-w-All	96.76	89.17	90.6	87.22	80.87	54.63
Gothic-PROIEL (size: 3387)	UDPipe	96.61	90.73	94.75	86.61	80.93	–
	UDify	95.55	85.97	80.57	86.37	80.13	–
	Multi-w-Lang_id	96.77	88.16	81.93	88.24	84.09	–
	Multi-w-Syntax	97.7	92.18	92.64	91.62	87.82	90.23
	Multi-w-Syntax+Semantic	95.55	85.73	75.99	90.14	80.92	82.58
	Multi-w-All	93.8	82.53	71.99	84.73	78.59	71.71
Greek-GDT (size: 1662)	UDPipe	97.98	94.96	95.82	92.9	90.59	–
	UDify	97.08	99.97	98.8	95.91	93.62	–
	Multi-w-Lang_id	97.79	99.78	98.83	94.66	93.51	–
	Multi-w-Syntax	97.89	99.87	98.84	96.56	94.05	81.96

7

	Multi-w-Syntax+Semantic	97.69	92.93	97.88	96.65	87.99	70.44
	Multi-w-All	95.88	89.61	96.23	91.83	85.26	63.61
Hebrew-HTB (size: 5241)	UDPipe	97.02	95.87	97.12	90.4	87.56	–
	UDify	96.21	96.02	97.28	92.14	89.68	–
	Multi-w-Lang_id	97.21	96.55	97.42	93.99	91.55	–
	Multi-w-Syntax	97.32	96.9	97.5	94.2	92.59	82.29
	Multi-w-Syntax+Semantic	97.32	95.77	94.17	92.57	82.66	72.15
	Multi-w-All	96.81	92.77	91.51	87.58	76.24	66.3
Hindi-HDTB (size: 13304)	UDPipe	97.52	94.15	98.67	94.95	91.93	–
	UDify	98.3	92.22	95.86	95.93	92.2	–
	Multi-w-Lang_id	98.6	93.38	96.1	95.47	93.32	–
	Multi-w-Syntax	98.69	94.15	96.24	95.72	92.11	82.27
	Multi-w-Syntax+Semantic	98.5	91.25	87.06	85.1	94.9	76.16
	Multi-w-All	97.47	89.99	84.47	81.16	89.88	67.11
Hungarian-Szeged (size: 910)	UDPipe	95.76	91.75	95.05	84.17	79.86	–
	UDify	96.36	86.16	90.19	91.01	86.21	–
	Multi-w-Lang_id	97.31	88.31	90.85	91.8	88.59	–
	Multi-w-Syntax	97.42	89.76	91.22	94.65	90.96	82.35
	Multi-w-Syntax+Semantic	94.01	84.9	87.63	94.74	81.06	71.86
	Multi-w-All	93.46	81.98	86.48	91.43	74.44	65.56
Indonesian-GSD (size: 4477)	UDPipe	93.69	95.58	99.64	86.54	80.22	–
	UDify	93.36	93.32	98.37	87.75	81.4	–
	Multi-w-Lang_id	95.31	94.29	98.43	90.96	84.62	–
	Multi-w-Syntax	96.82	90.23	91.52	92.83	87.4	76.11
	Multi-w-Syntax+Semantic	92.65	87.9	93.12	88.35	85.84	64.79
	Multi-w-All	91.91	86.23	89.17	85.59	83.36	55.57
Irish-IDT (size: 858)	UDPipe	92.72	82.43	90.48	81.77	73.72	–
	UDify	90.96	82.09	81.08	79.38	70.65	–
	Multi-w-Lang_id	93.72	84.91	82.4	84.27	76.57	–
	Multi-w-Syntax	93.88	86.82	83.16	90.08	82.83	83.97

8

	Multi-w-Syntax+Semantic	90.39	83.34	81.23	82.63	73.14	76.57
	Multi-w-All	89.53	80.31	78.39	76.84	68.89	71.55
Italian-ISDT (size: 29685)	UDPipe	98.39	98.11	98.66	95.24	93.29	–
	UDify	98.51	98.01	97.72	96.15	94.3	–
	Multi-w-Lang_id	98.74	98.21	97.83	96.57	93.16	–
	Multi-w-Syntax	98.83	98.34	97.89	96.53	94.18	82.27
	Multi-w-Syntax+Semantic	95.14	93.4	94.91	94.87	92.56	71.28
	Multi-w-All	94.13	90.05	90.97	91.84	88.32	62.14
Italian-ParTUT (size: 29685)	UDPipe	98.38	97.77	98.16	93.62	91.45	–
	UDify	99.18	96.69	98.52	95.9	94.05	–
	Multi-w-Lang_id	99.19	97.11	98.57	95.38	94.85	–
	Multi-w-Syntax	99.27	97.39	98.6	96.97	94.76	81.61
	Multi-w-Syntax+Semantic	96.52	90.68	93.6	98.24	89.89	74.21
	Multi-w-All	94.83	88.46	91.05	94.28	87.47	65.06
Japanese-GSD (size: 47926)	UDPipe	98.13	99.98	99.52	95.99	94.66	–
	UDify	98.73	93.44	96.5	95.1	93.43	–
	Multi-w-Lang_id	98.22	94.27	90.14	94.89	93.71	–
	Multi-w-Syntax	98.31	94.93	90.55	95.15	94.23	76.72
	Multi-w-Syntax+Semantic	96.63	91.07	90.11	97.63	91.41	70.03
	Multi-w-All	95.98	90.07	87.13	95.22	88.14	63.07
Kazakh-KTB (size: 31)	UDPipe	55.84	40.4	63.96	55.12	35.2	–
	UDify	91.29	99.58	99.21	74.74	66.63	–
	Multi-w-Lang_id	93.94	99.52	99.21	81.92	72.97	–
	Multi-w-Syntax	94.1	99.48	99.21	88.58	80.55	82.36
	Multi-w-Syntax+Semantic	91.45	96.21	96.02	81.59	80.85	73.43
	Multi-w-All	90.23	93.96	93.1	75.99	74.65	62.45
Korean-GSD (size: 27410)	UDPipe	96.29	99.77	93.4	88.84	85.38	–
	UDify	91.98	99.89	100.0	83.24	75.73	–
	Multi-w-Lang_id	94.4	99.56	99.75	87.17	81.12	–
	Multi-w-Syntax	94.55	99.63	99.59	91.16	84.3	81.2

9

	Multi-w-Syntax+Semantic	92.58	94.34	92.28	90.01	85.69	68.82
	Multi-w-All	92.18	92.19	89.53	87.41	79.76	63.39
Korean-Kaist (size: 27410)	UDPipe	95.59	100.0	94.3	88.62	86.68	–
	UDify	94.67	99.98	85.89	87.9	84.85	–
	Multi-w-Lang_id	96.19	99.64	86.86	90.58	87.0	–
	Multi-w-Syntax	96.31	99.9	87.42	92.87	89.07	83.56
	Multi-w-Syntax+Semantic	94.16	97.98	81.44	94.07	80.1	75.91
	Multi-w-All	93.34	95.61	77.72	88.69	74.55	68.58
Kurmanji-MG (size: 20)	UDPipe	53.36	41.54	69.58	46.16	35.25	–
	UDify	60.23	37.78	58.08	36.98	21.52	–
	Multi-w-Lang_id	74.25	55.98	63.82	64.24	44.36	–
	Multi-w-Syntax	74.72	61.74	65.41	78.91	60.28	85.67
	Multi-w-Syntax+Semantic	72.44	60.99	58.85	72.01	61.32	73.28
	Multi-w-All	71.18	60.04	55.01	66.12	58.02	66.53
Latin-ITTB (size: 34060)	UDPipe	98.34	96.97	98.99	92.35	90.09	–
	UDify	97.71	88.63	94.0	93.22	90.69	–
	Multi-w-Lang_id	98.21	90.38	94.38	94.19	91.25	–
	Multi-w-Syntax	97.8	95.01	94.73	95.24	91.82	82.42
	Multi-w-Syntax+Semantic	97.18	87.84	85.95	88.24	92.54	75.68
	Multi-w-All	96.76	85.14	82.36	84.01	87.23	64.79
Latin-Perseus (size: 34060)	UDPipe	88.4	79.1	81.45	72.86	62.94	–
	UDify	91.5	83.21	80.84	80.24	72.19	–
	Multi-w-Lang_id	94.08	85.85	82.18	84.84	78.38	–
	Multi-w-Syntax	94.24	87.63	82.95	90.66	82.9	83.56
	Multi-w-Syntax+Semantic	90.67	84.65	78.08	91.71	84.56	73.65
	Multi-w-All	90.08	81.38	74.86	86.87	77.64	64.8
Latin-PROIEL (size: 34060)	UDPipe	97.01	91.53	96.32	84.97	80.29	–
	UDify	96.79	89.49	91.79	85.89	81.56	–
	Multi-w-Lang_id	97.6	91.1	92.33	87.94	85.54	–
	Multi-w-Syntax	96.89	89.63	82.71	93.14	88.1	74.65

10

	Multi-w-Syntax+Semantic	94.25	87.77	89.66	87.03	80.45	65.26
	Multi-w-All	93.56	84.81	86.89	83.26	77.58	56.93
Latvian-LVTB (size: 10156)	UDPipe	96.11	93.01	95.46	87.6	83.75	–
	UDify	97.5	95.41	94.6	88.94	85.68	–
	Multi-w-Lang_id	98.07	96.04	94.94	91.26	87.31	–
	Multi-w-Syntax	97.74	90.55	93.44	93.79	90.15	81.29
	Multi-w-Syntax+Semantic	97.47	91.74	86.86	87.64	79.88	71.0
	Multi-w-All	95.4	88.6	84.17	84.92	74.62	65.49
Lithuanian-HSE (size: 2494)	UDPipe	81.7	60.47	76.89	53.52	43.71	–
	UDify	90.49	71.84	81.27	81.15	70.38	–
	Multi-w-Lang_id	93.39	74.81	69.51	85.29	76.17	–
	Multi-w-Syntax	93.56	78.07	70.84	90.47	81.32	74.84
	Multi-w-Syntax+Semantic	91.27	77.81	65.57	83.26	80.25	64.85
	Multi-w-All	90.6	76.85	62.99	77.68	76.15	54.86
Maltese-MUDT (size: 1123)	UDPipe	95.99	100.0	100.0	86.18	81.24	–
	UDify	90.56	99.63	82.84	84.65	76.17	–
	Multi-w-Lang_id	93.45	99.48	84.04	88.21	80.22	–
	Multi-w-Syntax	93.62	99.4	84.72	92.02	85.5	82.93
	Multi-w-Syntax+Semantic	92.66	93.39	80.36	83.99	81.66	75.95
	Multi-w-All	92.39	90.17	77.56	77.68	78.32	65.6
Marathi-UFAL (size: 373)	UDPipe	80.1	67.23	81.31	71.59	62.37	–
	UDify	94.29	84.49	87.71	76.46	69.34	–
	Multi-w-Lang_id	95.93	86.92	88.55	82.65	76.35	–
	Multi-w-Syntax	96.06	88.56	89.03	88.99	82.35	82.7
	Multi-w-Syntax+Semantic	94.22	86.29	82.01	83.57	76.35	73.14
	Multi-w-All	93.6	82.85	80.63	77.5	73.46	66.84
Norwegian-Bokmaal (size: 33282)	UDPipe	98.31	97.14	98.64	93.07	91.17	–
	UDify	98.34	91.82	98.13	96.37	93.95	–
	Multi-w-Lang_id	98.63	93.04	98.21	95.16	93.86	–
	Multi-w-Syntax	98.72	93.86	98.25	95.93	92.85	82.54

	Multi-w-Syntax+Semantic	98.67	88.01	92.42	92.62	88.51	76.51
	Multi-w-All	97.17	86.85	89.65	87.88	83.47	66.37
Norwegian-Nynorsk (size: 33282)	UDPipe	98.14	97.02	98.18	93.71	91.63	–
	UDify	97.83	96.17	97.34	95.08	92.93	–
	Multi-w-Lang_id	98.29	96.68	97.48	94.66	92.82	–
	Multi-w-Syntax	98.38	97.01	97.55	96.47	93.01	82.47
	Multi-w-Syntax+Semantic	98.11	96.32	90.59	87.94	85.1	71.03
	Multi-w-All	96.65	95.16	88.84	85.59	79.94	64.62
Norwegian-NynorskLIA (size: 33282)	UDPipe	89.59	86.13	93.93	69.27	61.26	–
	UDify	95.01	93.36	96.13	75.8	70.0	–
	Multi-w-Lang_id	96.41	94.33	96.35	82.11	76.21	–
	Multi-w-Syntax	96.54	94.98	96.48	89.43	82.32	82.45
	Multi-w-Syntax+Semantic	96.47	91.97	93.68	84.95	71.57	75.39
	Multi-w-All	94.4	88.52	92.04	81.15	65.31	66.05
Persian-Seraji (size: 4798)	UDPipe	97.75	97.78	97.44	91.68	88.29	–
	UDify	96.22	94.73	92.55	91.21	87.46	–
	Multi-w-Lang_id	97.22	95.47	93.04	93.1	89.74	–
	Multi-w-Syntax	98.19	92.08	87.04	95.54	91.46	76.91
	Multi-w-Syntax+Semantic	95.33	89.81	89.63	85.58	88.76	65.78
	Multi-w-All	95.08	88.41	86.93	79.6	84.56	56.29
Polish-LFG (size: 31496)	UDPipe	98.8	95.49	97.54	96.77	94.95	–
	UDify	98.97	96.29	94.47	96.82	95.12	–
	Multi-w-Lang_id	99.05	96.78	94.82	96.24	94.76	–
	Multi-w-Syntax	99.13	97.1	95.01	96.58	94.42	82.34
	Multi-w-Syntax+Semantic	95.77	95.71	90.4	97.59	94.27	75.21
	Multi-w-All	95.55	92.96	88.51	92.51	90.71	64.28
Portuguese-Bosque (size: 17992)	UDPipe	97.07	96.4	98.46	91.48	89.16	–
	UDify	97.54	89.36	85.46	93.38	88.75	–
	Multi-w-Lang_id	98.1	90.99	86.46	92.93	90.17	–
	Multi-w-Syntax	97.33	95.97	93.31	93.61	91.43	88.42

	Multi-w-Syntax+Semantic	97.25	86.0	86.6	85.0	87.72	78.71
	Multi-w-All	96.86	83.66	85.4	82.64	85.66	72.05
Portuguese-GSD (size: 17992)	UDPipe	98.31	99.92	99.3	94.28	92.9	–
	UDify	98.04	95.75	98.95	96.21	94.53	–
	Multi-w-Lang_id	98.43	96.32	98.97	95.64	95.03	–
	Multi-w-Syntax	98.52	96.71	98.98	96.08	94.67	82.17
	Multi-w-Syntax+Semantic	94.99	90.1	91.69	89.47	84.48	74.26
	Multi-w-All	93.54	88.1	88.37	85.86	79.15	65.18
Romanian-Nonstandard (size: 21782)	UDPipe	96.68	90.88	94.78	90.07	85.15	–
	UDify	96.85	87.24	92.7	89.73	86.45	–
	Multi-w-Lang_id	97.64	89.22	93.17	92.32	89.02	–
	Multi-w-Syntax	98.17	96.46	95.13	94.02	90.33	84.2
	Multi-w-Syntax+Semantic	95.03	86.69	87.79	94.45	92.69	76.75
	Multi-w-All	94.77	84.06	84.24	89.51	86.1	67.56
Romanian-RRT (size: 21782)	UDPipe	97.96	97.53	98.41	92.72	88.15	–
	UDify	96.94	93.41	94.15	93.43	89.91	–
	Multi-w-Lang_id	97.7	94.37	94.52	93.95	91.22	–
	Multi-w-Syntax	98.31	91.55	94.59	94.39	91.86	82.46
	Multi-w-Syntax+Semantic	97.79	88.58	92.66	96.92	89.67	74.87
	Multi-w-All	96.92	85.79	91.08	91.81	85.74	63.73
Russian-GSD (size: 54099)	UDPipe	97.1	92.66	97.37	89.47	85.69	–
	UDify	97.44	95.13	86.56	89.8	86.94	–
	Multi-w-Lang_id	98.03	95.81	87.48	91.44	88.86	–
	Multi-w-Syntax	98.13	96.26	88.01	92.79	90.68	83.44
	Multi-w-Syntax+Semantic	97.55	90.07	87.88	94.22	88.19	73.55
	Multi-w-All	96.62	86.43	84.78	91.21	83.72	64.12
Russian-SynTagRus (size: 54099)	UDPipe	99.12	97.57	98.53	95.22	93.74	–
	UDify	97.46	89.3	93.8	97.35	95.3	–
	Multi-w-Lang_id	98.04	90.94	94.19	96.42	94.06	–
	Multi-w-Syntax	98.14	92.04	94.42	96.6	95.49	82.8

13

	Multi-w-Syntax+Semantic	96.47	91.56	85.24	97.58	95.59	73.68
	Multi-w-All	95.69	88.9	82.75	94.32	92.34	65.16
Russian-Taiga (size: 54099)	UDPipe	93.18	82.87	89.99	76.81	70.47	–
	UDify	95.39	88.47	90.19	85.05	78.83	–
	Multi-w-Lang_id	96.67	90.24	90.85	87.91	83.11	–
	Multi-w-Syntax	96.79	91.44	91.22	92.4	86.68	82.88
	Multi-w-Syntax+Semantic	94.62	84.64	81.78	93.23	76.18	74.81
	Multi-w-All	92.94	82.19	79.46	89.74	71.12	68.8
Serbian-SET (size: 3328)	UDPipe	98.33	94.35	97.36	93.68	90.25	–
	UDify	97.67	97.66	95.44	95.19	92.17	–
	Multi-w-Lang_id	98.18	97.92	95.71	94.34	93.51	–
	Multi-w-Syntax	98.28	98.09	95.87	95.15	92.72	81.84
	Multi-w-Syntax+Semantic	96.82	92.0	94.62	97.62	95.09	69.71
	Multi-w-All	96.12	90.24	91.3	93.18	88.4	62.96
Slovak-SNK (size: 8483)	UDPipe	96.83	90.82	96.4	90.77	87.85	–
	UDify	98.8	87.71	94.04	97.1	95.01	–
	Multi-w-Lang_id	98.94	89.61	94.42	96.52	94.48	–
	Multi-w-Syntax	99.02	90.89	94.63	97.19	94.18	82.57
	Multi-w-Syntax+Semantic	98.24	89.32	86.29	93.11	97.1	72.51
	Multi-w-All	96.77	86.76	82.84	87.17	90.83	61.05
Slovenian-SSJ (size: 8556)	UDPipe	98.61	95.92	98.25	93.75	91.95	–
	UDify	97.72	93.29	89.43	95.75	93.57	–
	Multi-w-Lang_id	98.89	94.4	96.7	94.78	93.36	–
	Multi-w-Syntax	98.97	95.03	96.81	96.14	93.33	88.09
	Multi-w-Syntax+Semantic	98.01	94.46	96.56	99.07	95.06	75.63
	Multi-w-All	96.0	92.62	92.86	93.82	92.85	66.03
Slovenian-SST (size: 8556)	UDPipe	93.79	86.28	95.17	74.89	68.89	–
	UDify	95.4	89.81	95.15	80.89	75.55	–
	Multi-w-Lang_id	96.67	91.36	95.45	86.08	79.37	–
	Multi-w-Syntax	96.79	92.41	95.61	90.85	84.94	82.51

14

	Multi-w-Syntax+Semantic	94.89	91.45	89.54	93.06	77.71	71.56
	Multi-w-All	93.25	89.92	87.1	89.35	72.92	66.05
Spanish-AnCora (size: 28492)	UDPipe	98.91	98.49	99.17	92.85	90.77	–
	UDify	98.53	97.89	98.07	94.72	92.23	–
	Multi-w-Lang_id	98.76	98.11	98.15	94.35	92.69	–
	Multi-w-Syntax	98.84	98.26	98.2	95.05	92.43	82.47
	Multi-w-Syntax+Semantic	98.49	94.01	90.46	86.57	82.26	74.49
	Multi-w-All	97.12	91.6	88.9	82.45	80.2	68.69
Spanish-GSD (size: 28492)	UDPipe	96.85	97.09	98.97	92.14	89.46	–
	UDify	97.1	89.7	91.6	92.22	88.69	–
	Multi-w-Lang_id	97.15	90.15	94.35	92.4	89.44	–
	Multi-w-Syntax	97.26	91.36	94.57	95.03	92.11	84.76
	Multi-w-Syntax+Semantic	96.27	88.24	86.95	88.19	90.98	77.8
	Multi-w-All	95.32	87.05	83.41	84.96	87.81	67.84
Swedish-LinES (size: 7479)	UDPipe	96.78	89.43	97.03	86.97	82.76	–
	UDify	96.83	88.89	89.33	91.31	86.21	–
	Multi-w-Lang_id	97.63	90.59	90.05	93.03	88.63	–
	Multi-w-Syntax	97.73	91.74	90.46	94.14	89.37	82.74
	Multi-w-Syntax+Semantic	96.32	85.52	88.97	95.04	81.05	72.38
	Multi-w-All	94.5	84.43	85.93	88.81	75.54	64.43
Swedish-Talbanken (size: 7479)	UDPipe	97.94	96.86	98.01	90.73	87.71	–
	UDify	98.48	95.81	98.08	92.92	90.61	–
	Multi-w-Lang_id	98.72	96.37	98.16	93.6	91.35	–
	Multi-w-Syntax	98.81	96.75	98.21	93.97	91.65	81.81
	Multi-w-Syntax+Semantic	97.59	91.2	95.15	85.2	88.69	72.94
	Multi-w-All	96.31	88.6	92.99	79.93	86.24	65.96
Tamil-TTB (size: 400)	UDPipe	91.05	87.28	93.92	74.37	66.63	–
	UDify	90.47	70.0	67.17	80.1	70.38	–
	Multi-w-Lang_id	93.4	76.35	82.58	86.07	75.59	–
	Multi-w-Syntax	93.57	79.4	83.33	89.93	83.32	91.56

15

	Multi-w-Syntax+Semantic	89.6	76.93	82.69	82.85	81.06	81.72
	Multi-w-All	89.4	75.18	81.16	78.94	77.48	71.06
Telugu-MTG (size: 1051)	UDPipe	93.07	99.03	100.0	92.74	86.5	–
	UDify	96.58	91.77	73.55	89.46	84.62	–
	Multi-w-Lang_id	95.39	99.3	99.72	94.52	87.82	–
	Multi-w-Syntax	95.53	99.28	99.93	95.94	90.11	98.97
	Multi-w-Syntax+Semantic	92.4	96.82	94.55	85.87	92.79	88.27
	Multi-w-All	91.8	93.97	90.45	80.01	86.46	81.44
Turkish-IMST (size: 3664)	UDPipe	96.01	92.55	96.01	75.11	68.48	–
	UDify	88.59	59.22	72.82	80.85	69.2	–
	Multi-w-Lang_id	92.14	65.81	74.75	84.77	76.11	–
	Multi-w-Syntax	92.33	70.26	75.85	89.8	81.3	84.15
	Multi-w-Syntax+Semantic	89.85	65.69	69.47	91.52	74.33	77.58
	Multi-w-All	88.44	64.48	65.52	89.05	71.95	71.7
Ukrainian-IU (size: 5496)	UDPipe	97.59	92.66	97.23	90.2	87.16	–
	UDify	98.02	89.67	95.34	95.3	91.01	–
	Multi-w-Lang_id	98.42	91.25	95.62	95.87	91.56	–
	Multi-w-Syntax	98.51	92.31	95.78	95.26	92.1	81.89
	Multi-w-Syntax+Semantic	96.75	91.34	94.79	96.5	81.13	75.05
	Multi-w-All	95.12	87.64	93.36	92.39	78.1	63.56
Urdu-UDTB (size: 4043)	UDPipe	93.66	81.92	97.4	89.41	83.53	–
	UDify	93.8	90.38	88.8	88.3	83.33	–
	Multi-w-Lang_id	95.61	91.84	89.56	89.56	86.99	–
	Multi-w-Syntax	95.74	92.82	89.99	93.24	89.31	83.37
	Multi-w-Syntax+Semantic	94.04	92.7	81.71	84.07	86.3	72.71
	Multi-w-All	93.4	90.32	79.35	80.98	83.0	61.73
Uyghur-UDT (size: 1656)	UDPipe	89.87	88.3	95.31	79.97	68.6	–
	UDify	75.88	70.8	79.7	67.78	50.69	–
	Multi-w-Lang_id	83.67	75.48	81.13	76.91	61.43	–
	Multi-w-Syntax	83.99	78.65	81.94	85.84	73.26	83.4

16

Corpus	Model	UPOS	UFeats	Lemmas	UAS	LAS	Typo F1
	Multi-w-Syntax+Semantic	83.15	72.27	79.1	76.14	74.65	77.17
	Multi-w-All	82.6	69.28	76.9	73.88	72.57	67.9
Vietnamese-VTB (size: 1400)	UDPipe	89.68	99.72	99.55	72.2	64.38	–
	UDify	85.59	65.49	77.18	75.29	64.18	–
	Multi-w-Lang_id	90.14	71.05	78.79	81.82	72.35	–
	Multi-w-Syntax	90.36	74.8	79.71	89.12	78.29	83.13
	Multi-w-Syntax+Semantic	88.68	68.43	71.27	91.0	70.24	73.17
	Multi-w-All	86.7	67.48	69.45	87.18	67.23	68.02
Corpus	**Model**	**UPOS**	**UFeats**	**Lemmas**	**UAS**	**LAS**	**Typo F1**

17