A Glossary of UD POS tags and edge types

A glossary of UD POS labels and edge types mentioned in the paper is provided in Table 1.

B Distinguishing Content and Function Words

We adopt the following distinction between function and content words. Function words include (1) grammatical-relation markers (prepositions marking direct and indirect objects, possession, and other types of relations inside NPs); (2) tense-aspect-mood markers including inflected auxiliaries; (3) markers of (in)definiteness; (4) coordinating conjunctions; (5) complementizers; (6) classifiers; (7) copulas and existential predicates; (8) dummy subjects and expletives. Other word types are considered content words, including (1) all other types of predicates, participants, obliques, adverbial and adjectival modifiers; (2) negation markers; (3) discourse markers; (4) quantifiers; (5) spatial/temporal-relations markers.

C Confusion matrices for translation equivalents of POS tags and UD edge labels: raw counts

C.1 POS tags

Raw-count and percentage confusion matrices for translation equivalents of UD POS tags in three parallel corpora are presetned in Tables 2 and 3.

C.2 UD edge types

Raw-count confusion matrices for translation equivalents of major UD edge types in three parallel corpora are presetned in Table 4.

D Translation entropies of UD relations

Translation entropies of UD relations are shown in Table 5.

E Zero-shot Parsing: Implementation Details

We used the AllenNLP (Gardner et al., 2018) implementation of the deep biaffine attention model by Dozat and Manning (2017). The only modification is that we replaced the trainable Glove embeddings with multilingual Bert (Devlin et al., 2018) untrainable embeddings (i.e., we didn't perform any fine-tuning on Bert), using the built-in embeddings in AllenNLP with the default settings. We ignore UD sub-categorization of the edge labels (as the sub-types are often language-specific). The full list of hyper-parameters is given in Table 6.

We trained three models for each language (with the same hyperparameters), using the UD v2.5 English-EWT dataset for the English model, and the GSD datasets (also v2.5) for French, Russian, Chinese, Korean and Japanese. Standard splits were used for all corpora.[1] The models were evaluated on the GSD and PUD datasets.

The per-label F-scores used for linear modelling in the paper are averages of the F-scores of the three models. The following relations were considered: `acl`, `advcl`, `advmod`, `amod`, `appos`, `ccomp`, `compound`, `conj`, `fixed`, `flat`, `nmod`, `nsubj`, `nummod`, `obj`, `obl`, `parataxis`, `xcomp`.[2] Not all of them were present in parser outputs for all languages due to training-set peculiarities, and missing relations were automatically omitted from the model. The R code used for fitting the models is available in the Supplementary Material.

References

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Thimothy Dozat and Christopher D. Manning. 2017. Deep biaffine attention for dependency parsing. In *ICLR 2017*.

Matt Gardner, Joel Grus, Mark Neumann, Oyvind Tafjord, Pradeep Dasigi, Nelson F. Liu, Matthew Peters, Michael Schmitz, and Luke Zettlemoyer. 2018. AllenNLP: A deep semantic natural language processing platform. In *Proceedings of Workshop for NLP Open Source Software (NLP-OSS)*, pages 1–6, Melbourne, Australia. Association for Computational Linguistics.

Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. In *Proc. of ICLR*.

[1] All UD corpora can be found in `https://github.com/UniversalDependencies/`

[2] The list is slightly longer than that used in the automation experiments since thanks to the precision of manual alignment we were able to target relatively rare edge labels in a small corpus.

Label	Short Definition
POS tags	
NOUN	Noun.
PROPN	Proper noun.
PRON	Pronoun.
ADJ	Adjective.
ADV	Adverb.
VERB	Verb.
AUX	Auxiliary.
PART	Particle.
NUM	Numeral.
ADP	Adposition.
Clause Elements	
nsubj	Nominal subject.
dobj	Direct object.
ccomp	Clausal complement (finite or infinite), unless its subject is controlled.
xcomp	Open clausal complement, i.e., predicative or clausal complement without its own subject.
advmod	Modifying adverb.
neg	Negation modifier (e.g., "not", "no").
aux	Auxiliary of a verbal predicate, including markers of tense, mood, modality, aspect, voice or evidentiality.
nmod	Oblique: nominal functioning as an adjunct. (*nmods* are also used for nominal modifiers in noun phrases, see below)
Inter-clause Linkage	
conj	Relation between the conjuncts in a coordination to the first conjunct, which is considered the head.
cc	Coordinating conjunction.
advcl	Adverbial clause modifier, including temporal clause, consequence, conditional clause, and purpose clause.
mark	Marker: the word introducing a clause subordinate to another clause, often a subordinating conjunction.
parataxis	Several elements (often clauses or fragments) placed side by side without any explicit coordination, subordination, or argument relation.
Nominal Elements	
det	Determiner.
case	Case marker, including adpositions.
nmod	Nominal modifier of a noun or a noun phrase.
nummod	Numeric modifier.

Table 1: UD POS tags and edge types mentioned in the paper and their definitions.

		ADJ	ADP	ADV	AUX	CCONJ	DET	NOUN	NUM	None	PART	PRON	PROPN	SCONJ	SYM	VERB	X
En-Ru																	
	ADJ	1030	1	52	0	0	28	74	2	29	2	0	22	0	0	57	7
	ADP	4	296	11	0	0	2	12	0	9	1	0	0	1	0	7	3
	ADV	31	10	362	0	25	5	17	2	31	28	1	1	0	0	6	12
	AUX	7	0	2	5	0	0	1	0	3	0	0	0	0	0	80	0
	CCONJ	0	0	0	0	25	0	0	0	2	0	0	0	0	0	0	0
	DET	31	1	3	0	0	29	3	9	4	8	1	0	0	0	1	0
	INTJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	NOUN	204	3	23	0	0	14	3013	4	59	0	8	116	0	5	38	12
	NUM	20	0	1	0	0	1	23	229	2	0	1	0	0	0	0	4
	None	102	44	90	1	10	38	387	6	0	27	61	32	1	0	177	14
	PART	0	1	1	0	0	0	0	0	0	45	0	0	0	0	3	0
	PRON	5	0	2	0	0	127	13	0	22	0	348	1	0	0	1	0
	PROPN	144	0	0	0	0	3	124	0	9	0	0	1247	0	0	2	10
	SCONJ	0	13	3	0	0	0	0	0	0	3	0	0	0	0	0	0
	SYM	0	0	0	0	0	0	2	0	0	0	0	0	0	15	0	0
	VERB	52	7	9	9	0	0	145	0	66	0	1	2	0	0	1444	0
	X	2	0	0	0	0	0	2	0	0	0	0	6	0	0	0	0
En-Fr																	
	ADJ	1006	12	51	0	0	2	126	1	44	0	0	12	0	0	45	2
	ADP	5	388	21	0	1	4	13	0	9	0	0	0	1	0	7	0
	ADV	30	28	416	1	5	2	33	0	43	0	5	0	5	0	7	0
	AUX	0	0	2	4	0	0	0	0	2	0	0	0	0	0	74	0
	CCONJ	0	1	12	0	482	0	0	0	29	0	0	0	0	1	0	0
	DET	25	2	4	0	0	19	4	5	5	0	2	0	0	0	0	0
	INTJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	NOUN	113	9	9	0	0	1	3244	10	46	0	10	13	0	5	38	1
	NUM	8	0	0	0	0	11	18	371	1	0	1	0	0	0	0	0
	None	55	28	70	3	52	5	162	3	0	0	57	3	0	0	170	2
	PART	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0
	PRON	0	1	1	0	0	0	16	0	9	0	409	1	0	0	1	0
	PROPN	70	4	0	0	0	0	380	4	11	0	4	1006	0	0	1	21
	SCONJ	0	62	34	0	6	1	1	0	0	0	0	0	2	0	0	0
	SYM	0	0	0	0	0	0	9	0	0	0	0	0	0	26	0	0
	VERB	63	19	2	36	0	0	100	0	49	0	1	1	0	0	1508	0
	X	1	0	0	0	0	0	3	0	0	0	0	1	0	0	0	6
En-Zh																	
	ADJ	337	6	30	9	0	42	373	66	13	0	6	131	0	0	98	2
	ADP	5	227	3	0	6	2	18	0	34	2	0	0	0	1	117	0
	ADV	46	20	338	9	1	2	79	3	33	0	12	0	0	2	42	1
	AUX	0	0	2	50	0	0	2	0	6	0	0	0	0	0	11	0
	CCONJ	0	4	135	2	243	0	1	0	51	0	0	0	0	2	1	1
	DET	7	0	7	4	0	36	4	6	1	0	3	0	0	0	9	0
	INTJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	NOUN	54	7	4	3	0	7	2671	16	38	4	4	5	0	0	356	10
	NUM	2	0	0	0	0	2	2	194	1	0	0	9	0	0	0	0
	None	49	130	462	48	15	21	559	31	0	11	100	39	1	4	464	23
	PART	0	4	22	9	0	0	0	0	0	0	0	0	0	0	13	0
	PRON	1	0	1	0	0	7	29	0	54	0	526	5	0	0	2	2
	PROPN	6	0	0	0	0	0	192	2	7	2	5	1001	0	0	11	126
	PUNCT	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	SCONJ	1	56	7	2	0	0	3	0	6	0	0	0	0	8	14	1
	SYM	0	0	0	0	1	0	2	0	1	0	0	0	0	0	0	0
	VERB	32	14	25	25	0	2	110	0	54	1	0	0	1	0	1394	2
	X	1	0	0	0	0	0	3	0	0	0	1	0	0	0	0	5
En-Ko																	
	ADJ	152	0	37	0	0	42	554	20	14	2	4	128	0	0	31	0
	ADP	10	0	13	0	0	0	100	0	68	6	1	0	0	0	27	0
	ADV	46	0	202	1	17	6	125	1	28	8	5	2	0	0	21	0
	AUX	7	0	0	23	0	0	6	0	9	1	0	0	0	0	11	0
	CCONJ	2	0	0	0	19	0	2	0	0	0	0	1	0	0	1	0
	DET	5	0	7	0	0	49	13	0	5	2	2	0	0	0	5	0
	INTJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	NOUN	10	0	13	6	0	8	2628	4	39	2	8	29	0	0	46	0
	NUM	0	0	1	0	0	45	21	147	1	0	0	1	0	0	0	0
	None	0	0	0	0	0	0	3	1	0	0	0	0	0	0	0	0
	PART	11	0	2	0	0	0	0	0	2	2	0	0	0	0	26	0
	PRON	3	0	0	1	0	18	17	0	50	0	232	6	0	0	0	0
	PROPN	0	0	0	1	0	0	181	7	10	0	0	1002	0	0	0	0
	SCONJ	0	0	1	0	0	0	14	0	0	1	0	1	0	0	9	0
	SYM	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0
	VERB	32	0	12	5	0	1	804	0	39	30	0	5	0	0	400	0
	X	0	0	0	0	0	0	4	0	0	1	0	2	0	0	0	0
En-Jp																	
	ADJ	101	3	10	2	0	0	185	5	7	0	1	32	0	0	18	0
	ADP	1	425	1	6	0	0	32	0	5	1	0	1	3	0	8	0
	ADV	21	22	43	3	16	0	43	0	17	6	3	0	4	0	7	0
	AUX	1	7	0	99	0	0	5	0	0	1	0	0	1	0	29	0
	CCONJ	0	70	1	18	36	0	0	0	0	0	0	0	17	1	0	0
	DET	4	52	2	2	0	47	12	0	4	0	7	1	0	0	0	0
	INTJ	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
	NOUN	8	0	0	0	0	0	1271	0	24	5	1	18	1	0	27	0
	NUM	0	0	0	0	0	0	5	71	0	0	0	1	0	0	2	0
	PART	1	53	0	15	0	0	0	0	1	0	0	0	1	0	0	0
	PRON	0	5	1	1	1	6	11	0	77	1	108	0	6	0	0	0
	PROPN	0	0	0	0	0	0	139	0	1	0	0	406	0	0	0	0
	SCONJ	0	17	0	1	0	0	6	0	0	3	0	0	14	0	0	0
	SYM	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0
	VERB	7	4	1	18	0	0	53	0	21	0	0	0	0	0	229	0
	X	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0

Table 2: Raw counts of translation mappings of parts of speech.

		ADJ	ADP	ADV	AUX	CCONJ	DET	NOUN	NUM	None	PART	PRON	PROPN	SCONJ	SYM	VERB	X
En-Ru																	
	ADJ	79	0	4	0	0	2	6	0	2	0	0	2	0	0	4	1
	ADP	1	86	3	0	0	1	3	0	3	0	0	0	0	0	2	1
	ADV	6	2	68	0	5	1	3	0	6	5	0	0	0	0	1	2
	AUX	7	0	2	5	0	0	1	0	3	0	0	0	0	0	82	0
	CCONJ	0	0	0	0	93	0	0	0	7	0	0	0	0	0	0	0
	DET	34	1	3	0	0	32	3	10	4	9	1	0	0	0	1	0
	INTJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
	NOUN	6	0	1	0	0	0	86	0	2	0	0	3	0	0	1	0
	NUM	7	0	0	0	0	0	8	81	1	0	0	0	0	0	0	1
	None	10	4	9	0	1	4	39	1	0	3	6	3	0	0	18	1
	PART	0	2	2	0	0	0	0	0	0	90	0	0	0	0	6	0
	PRON	1	0	0	0	0	24	3	0	4	0	67	0	0	0	0	0
	PROPN	9	0	0	0	0	0	8	0	1	0	0	81	0	0	0	1
	SCONJ	0	68	16	0	0	0	0	0	0	16	0	0	0	0	0	0
	SYM	0	0	0	0	0	0	12	0	0	0	0	0	0	88	0	0
	VERB	3	0	1	1	0	0	8	0	4	0	0	0	0	0	83	0
	X	20	0	0	0	0	0	20	0	0	0	0	60	0	0	0	0
En-Fr																	
	ADJ	77	1	4	0	0	0	10	0	3	0	0	1	0	0	3	0
	ADP	1	86	5	0	0	1	3	0	2	0	0	0	0	0	2	0
	ADV	5	5	72	0	1	0	6	0	7	0	1	0	1	0	1	0
	AUX	0	0	2	5	0	0	0	0	2	0	0	0	0	0	90	0
	CCONJ	0	0	2	0	92	0	0	0	6	0	0	0	0	0	0	0
	DET	38	3	6	0	0	29	6	8	8	0	3	0	0	0	0	0
	INTJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
	NOUN	3	0	0	0	0	0	93	0	1	0	0	0	0	0	1	0
	NUM	2	0	0	0	0	3	4	90	0	0	0	0	0	0	0	0
	None	9	5	11	0	9	1	27	0	0	0	9	0	0	0	28	0
	PART	0	33	67	0	0	0	0	0	0	0	0	0	0	0	0	0
	PRON	0	0	0	0	0	0	4	0	2	0	93	0	0	0	0	0
	PROPN	5	0	0	0	0	0	25	0	1	0	0	67	0	0	0	1
	SCONJ	0	58	32	0	6	1	1	0	0	0	0	0	2	0	0	0
	SYM	0	0	0	0	0	0	26	0	0	0	0	0	0	74	0	0
	VERB	4	1	0	2	0	0	6	0	3	0	0	0	0	0	85	0
	X	9	0	0	0	0	0	27	0	0	0	0	9	0	0	0	55
En-Zh																	
	ADJ	30	1	3	1	0	4	34	6	1	0	1	12	0	0	9	0
	ADP	1	55	1	0	1	0	4	0	8	0	0	0	0	0	28	0
	ADV	8	3	57	2	0	0	13	1	6	0	2	0	0	0	7	0
	AUX	0	0	3	70	0	0	3	0	8	0	0	0	0	0	15	0
	CCONJ	0	1	31	0	55	0	0	0	12	0	0	0	0	0	0	0
	DET	9	0	9	5	0	47	5	8	1	0	4	0	0	0	12	0
	INTJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
	NOUN	2	0	0	0	0	0	84	1	1	0	0	0	0	0	11	0
	NUM	1	0	0	0	0	1	1	92	0	0	0	4	0	0	0	0
	None	3	7	24	2	1	1	29	2	0	1	5	2	0	0	24	1
	PART	0	8	46	19	0	0	0	0	0	0	0	0	0	0	27	0
	PRON	0	0	0	0	0	1	5	0	9	0	84	1	0	0	0	0
	PROPN	0	0	0	0	0	0	14	0	1	0	0	74	0	0	1	9
	PUNCT	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
	SCONJ	1	57	7	2	0	0	3	0	6	0	0	0	0	8	14	1
	SYM	0	0	0	0	25	0	50	0	25	0	0	0	0	0	0	0
	VERB	2	1	2	2	0	0	7	0	3	0	0	0	0	0	84	0
	X	10	0	0	0	0	0	30	0	0	0	10	0	0	0	0	50
En-Ko																	
	ADJ	15	0	4	0	0	4	56	2	1	0	0	13	0	0	3	0
	ADP	4	0	6	0	0	0	44	0	30	3	0	0	0	0	12	0
	ADV	10	0	44	0	4	1	27	0	6	2	1	0	0	0	5	0
	AUX	12	0	0	40	0	0	11	0	16	2	0	0	0	0	19	0
	CCONJ	8	0	0	0	76	0	8	0	0	0	0	4	0	0	4	0
	DET	6	0	8	0	0	56	15	0	6	2	2	0	0	0	6	0
	INTJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
	NOUN	0	0	0	0	0	0	94	0	1	0	0	1	0	0	2	0
	NUM	0	0	0	0	0	21	10	68	0	0	0	0	0	0	0	0
	None	0	0	0	0	0	0	75	25	0	0	0	0	0	0	0	0
	PART	26	0	5	0	0	0	0	0	5	5	0	0	0	0	60	0
	PRON	1	0	0	0	0	6	5	0	15	0	71	2	0	0	0	0
	PROPN	0	0	0	0	0	0	15	1	1	0	0	83	0	0	0	0
	SCONJ	0	0	4	0	0	0	54	0	0	4	0	4	0	0	35	0
	SYM	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
	VERB	2	0	1	0	0	0	61	0	3	2	0	0	0	0	30	0
	X	0	0	0	0	0	0	57	0	0	14	0	29	0	0	0	0
En-Jp																	
	ADJ	28	1	3	1	0	0	51	1	2	0	0	9	0	0	5	0
	ADP	0	88	0	1	0	0	7	0	1	0	0	0	1	0	2	0
	ADV	11	12	23	2	9	0	23	0	9	3	2	0	2	0	4	0
	AUX	1	5	0	69	0	0	3	0	0	1	0	0	1	0	20	0
	CCONJ	0	49	1	13	25	0	0	0	0	0	0	0	12	1	0	0
	DET	3	40	2	2	0	36	9	0	3	0	5	1	0	0	0	0
	INTJ	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
	NOUN	1	0	0	0	0	0	94	0	2	0	0	1	0	0	2	0
	NUM	0	0	0	0	0	0	6	90	0	0	0	1	0	0	3	0
	PART	1	75	0	21	0	0	0	0	1	0	0	0	1	0	0	0
	PRON	0	2	0	0	0	3	5	0	35	0	50	0	3	0	0	0
	PROPN	0	0	0	0	0	0	25	0	0	0	0	74	0	0	0	0
	SCONJ	0	41	0	2	0	0	15	0	0	7	0	0	34	0	0	0
	SYM	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
	VERB	2	1	0	5	0	0	16	0	6	0	0	0	0	0	69	0
	X	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0

Table 3: Percentages of translation mappings of parts of speech.

		acl	advcl	advmod	amod	appos	ccomp	compound	conj	fixed	flat	nmod	nsubj	nummod	obj	obl	parataxis	xcomp	Collapsed	Other	MCOP (%)	MCOP
En-Ru																						
	acl	151	2	1	5	1	2	0	1	0	0	27	7	0	5	3	3	13	2	89	15	nmod+acl
	advcl	11	68	2	0	0	9	0	11	0	0	2	0	0	1	22	4	2	1	79	5	advcl+xcomp
	advmod	0	0	332	11	0	0	0	1	0	1	4	0	1	1	16	2	2	27	129	11	advmod+nummod
	amod	29	0	8	893	4	4	0	1	0	5	60	4	7	1	3	0	0	57	77	18	det
	appos	0	0	0	0	43	0	0	1	0	15	21	1	1	0	0	1	0	1	35	7	nmod+flat
	ccomp	1	2	1	0	0	48	0	0	0	1	0	0	0	1	0	6	6	0	32	11	ccomp+xcomp
	compound	0	0	2	244	7	0	4	3	0	48	173	0	10	2	4	0	0	153	97	22	nmod+nmod
	conj	0	2	0	1	0	0	0	387	1	1	0	0	1	0	1	1	0	5	108	10	conj+conj
	fixed	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	9	1	1	nmod+flat
	flat	0	0	0	6	3	0	2	0	0	154	13	0	0	0	1	0	0	8	19	7	flat+flat
	nmod	4	1	11	40	6	1	0	5	0	10	632	5	7	6	31	0	0	5	309	67	det
	nsubj	1	1	1	1	0	0	0	1	0	0	14	829	0	22	26	0	0	3	261	48	nsubj+xcomp
	nummod	0	0	1	10	0	0	0	0	0	1	3	0	153	0	1	0	0	14	4	1	compound+nummod
	obj	1	0	4	4	0	1	0	0	0	0	58	27	0	393	62	0	3	25	106	26	iobj
	obl	0	5	44	2	0	0	0	3	0	0	38	17	0	27	463	3	2	10	311	67	iobj
	parataxis	0	1	0	0	0	1	0	5	0	0	0	0	0	0	0	41	0	0	22	2	acl+obl
	xcomp	0	5	9	0	0	7	0	2	0	0	1	3	0	3	15	0	91	4	40	15	iobj
En-Fr																						
	acl	104	4	1	13	3	52	0	3	0	0	12	1	0	4	3	0	24	2	80	8	acl+xcomp
	advcl	2	96	2	0	0	19	0	4	0	0	0	0	0	0	16	4	5	0	70	9	xcomp+advcl
	advmod	1	0	321	7	2	0	0	1	2	0	8	1	0	6	17	0	6	31	108	18	advmod+xcomp
	amod	1	0	9	854	3	14	5	2	0	0	97	3	3	1	5	0	0	61	80	16	det
	appos	0	0	0	2	88	1	0	1	0	0	7	0	0	0	1	0	0	0	27	4	appos+nmod
	ccomp	0	1	0	0	0	51	0	0	0	0	0	0	0	2	0	5	11	0	28	10	ccomp+xcomp
	compound	0	0	1	143	32	0	47	5	0	21	300	1	12	4	5	0	2	102	46	10	nmod+nmod
	conj	1	2	0	0	2	2	0	405	0	0	3	0	0	1	0	6	1	7	96	9	nmod+conj
	fixed	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1	2	1	appos+nummod
	flat	0	0	0	5	32	0	5	0	0	140	4	0	8	0	0	0	0	8	15	7	appos+flat
	nmod	0	0	6	17	3	0	1	2	0	1	702	4	12	13	34	0	1	15	253	37	nmod+nmod
	nsubj	4	0	1	0	1	2	0	0	0	0	6	870	0	12	8	0	0	0	197	59	nsubj+xcomp
	nummod	0	0	0	3	3	0	0	1	0	0	41	0	169	0	1	0	0	7	12	7	det
	obj	2	1	2	0	1	1	1	1	0	0	37	8	0	520	55	0	7	21	63	13	obj+nmod
	obl	2	4	22	3	0	1	0	1	0	0	23	9	0	57	640	0	7	7	195	27	obl+nmod
	parataxis	1	0	1	0	0	4	0	9	0	0	0	0	0	0	0	37	0	0	23	2	obl+parataxis
	xcomp	0	6	6	0	1	3	0	0	0	0	1	1	0	6	8	0	141	7	18	3	obj+amod
En-Zh																						
	acl	96	2	1	5	1	0	6	0	0	0	3	25	0	6	3	0	3	2	143	8	obj+advcl
	advcl	1	44	1	0	0	4	0	0	0	1	0	0	0	1	3	0	34	0	116	14	dep
	advmod	4	8	227	7	0	2	7	0	0	0	2	5	1	15	26	0	13	52	227	19	advmod+obj
	amod	46	1	20	194	1	0	351	0	0	0	56	28	38	5	8	0	1	160	206	35	compound+compound
	appos	1	0	0	0	53	0	9	2	0	2	5	2	0	0	0	0	1	3	29	2	compound+compound
	ccomp	0	1	2	0	0	37	0	0	0	0	0	0	0	0	0	0	5	0	54	4	ccomp+advcl
	compound	2	0	0	15	3	0	317	4	0	14	26	9	1	2	3	0	0	250	109	64	compound+compound
	conj	2	51	0	0	0	1	7	225	0	1	1	0	0	0	0	0	0	10	200	23	dep
	fixed	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	60	1	1	appos+flat
	flat	0	0	0	0	18	0	20	2	0	99	0	0	0	0	0	0	0	19	53	39	appos+flat
	nmod	18	0	4	13	6	1	200	4	0	3	259	38	13	44	37	0	0	31	427	29	compound+compound
	nsubj	2	2	5	0	0	1	5	0	0	0	5	674	0	31	14	0	1	1	391	92	nsubj+advcl
	nummod	0	0	0	1	0	0	0	0	0	1	0	0	74	0	0	0	0	39	57	41	nummod+clf
	obj	4	4	1	0	1	7	6	0	0	0	3	6	2	396	29	0	7	26	180	50	ccomp+nsubj
	obl	4	6	21	0	0	13	8	0	0	0	8	34	0	138	159	0	18	14	425	49	obj+advcl
	parataxis	0	2	1	0	0	6	0	0	0	0	0	0	0	0	0	0	1	0	71	18	dep
	xcomp	1	5	11	0	1	24	0	0	0	0	0	1	0	9	2	0	42	5	62	22	aux
En-Ko																						
	acl	73	4	3	0	0	0	4	0	0	0	2	2	0	4	0	0	0	0	106	7	acl+nsubj+acl
	advcl	2	24	9	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	95	8	advmod+acl
	advmod	5	9	169	4	0	0	13	0	1	0	1	5	1	4	0	0	0	14	168	30	aux
	amod	68	1	20	106	0	0	339	0	0	5	69	16	16	3	0	0	0	87	168	41	det
	appos	25	1	0	0	7	0	16	0	0	3	2	2	0	0	0	0	0	0	37	4	acl+compound
	ccomp	0	18	2	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	39	5	advcl+advcl
	compound	2	1	4	0	0	0	354	7	0	21	14	4	2	4	0	0	0	143	67	13	compound+compound
	conj	2	47	2	0	0	0	2	201	0	3	2	0	0	1	0	0	0	2	118	7	conj+compound
	fixed	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	3	2	1	compound+compound
	flat	6	0	0	0	0	0	28	0	0	101	1	1	0	0	0	0	0	9	22	5	compound+flat
	nmod	10	1	26	3	5	0	139	4	0	3	207	21	6	12	4	0	0	8	366	51	acl+advmod
	nsubj	5	1	18	1	0	0	7	0	0	0	2	374	0	23	0	0	0	1	358	55	nsubj+advcl
	nummod	0	0	2	0	0	0	3	0	0	0	0	0	100	0	0	0	0	1	40	14	nummod+nmod
	obj	5	5	30	0	0	0	16	0	0	0	2	21	0	249	0	0	0	11	143	27	aux+obj
	obl	5	7	210	0	0	0	12	0	0	0	2	24	0	35	2	0	0	4	350	25	advmod+compound
	parataxis	1	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	5	advcl+advcl
	xcomp	3	12	17	0	0	9	9	0	0	0	0	1	0	7	0	0	0	2	61	11	aux
En-Jp																						
	acl	23	4	0	0	0	0	1	0	0	0	5	1	0	0	0	0	0	0	28	2	acl+advcl
	advcl	0	3	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	5	1	nsubj+advcl
	advmod	1	3	20	2	0	1	3	0	0	0	3	1	3	2	3	0	0	9	67	10	aux
	amod	34	2	3	26	0	0	75	0	0	0	84	2	4	4	1	0	0	102	57	9	compound+compound
	appos	1	0	0	0	9	0	11	0	0	0	10	0	0	0	0	0	0	1	16	4	nmod+nmod
	ccomp	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	8	1	aux
	compound	1	0	0	1	0	0	116	0	0	0	54	0	0	1	0	0	0	124	30	8	compound+compound
	conj	2	11	0	0	0	0	0	0	0	0	41	1	0	0	1	0	0	2	74	18	nmod+nmod
	fixed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	0	0	acl+acl
	flat	0	0	0	0	0	0	19	0	0	0	34	0	0	0	0	0	0	13	20	9	nmod+nmod
	nmod	0	0	0	0	2	0	31	0	0	0	188	1	0	5	3	0	0	23	133	29	nmod+nmod
	nsubj	0	0	1	0	0	0	0	0	0	0	6	105	0	4	7	0	0	1	103	11	nsubj+advcl
	nummod	1	0	0	0	0	0	5	0	0	0	0	0	37	0	0	0	0	0	17	9	nummod+compound
	obj	0	1	0	0	0	0	7	0	0	0	5	12	0	52	2	0	0	3	26	7	iobj
	obl	0	1	4	0	0	0	3	0	0	0	6	3	0	4	42	0	0	2	86	22	iobj
	parataxis	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	8	1	acl+acl
	xcomp	2	2	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	8	3	aux

Table 4: Raw counts of corresponding syntactic relations of frequent UD edge types.

	Ko	Jp	Zh	Ru	Fr
acl	4.769242	3.830784	5.060668	3.652101	4.050027
advcl	5.310528	2.846439	4.835997	4.504237	3.745798
advmod	4.089253	5.092975	4.555394	2.848869	2.767038
amod	3.551876	3.268781	3.642666	1.634234	1.709718
appos	3.880474	3.21408	3.409938	3.297443	2.135566
aux	3.087929	2.565961	2.013037	1.157876	0.508315
case	3.50321	1.988427	3.81068	2.219851	1.333314
cc	3.062907	4.349407	3.109744	1.773011	1.736929
ccomp	4.411585	3.277613	4.158406	2.994698	2.65771
compound	2.309827	2.145895	2.503707	2.988854	2.821717
conj	3.531517	4.126376	4.16811	2.144241	2.039554
cop	2.272808	3.21288	2.217022	3.000973	1.584963
csubj	3.251629	0.811278	3.386637	3.221252	2.879249
det	3.107192	2.769522	2.989865	2.689149	2.197421
fixed	1.842371	0	0.46229	1.35203	2.584963
flat	2.128367	2.381739	2.39878	1.634478	1.932195
iobj	2.321928	NA	2.5	1.5	1.351644
mark	2.96381	3.027169	4.024999	2.7912	2.142446
nmod	4.971409	3.542603	5.006165	3.155123	2.795594
nsubj	4.11521	4.069544	3.196435	2.319142	1.721087
nummod	1.967499	1.946163	2.15335	1.120919	1.478528
obj	3.488357	3.077944	3.095776	2.679303	1.934877
obl	5.265461	4.585266	5.601495	3.664774	2.746977
parataxis	4.446289	3.321928	4.897466	2.711151	3.091764
xcomp	4.609696	3.039149	3.912925	3.054541	1.997426
Average	3.53	3.02	3.48	2.56	2.24

Table 5: Translation entropies of UD relations.

Input	Input dropout rate: 0.3 POS tag embedding dimension: 100
Word-level BiLSTM	LSTM size: 400 # LSTM layers: 3 Recurrent dropout rate: 0.3 Use Highway Connection: Yes Output dropout rate: 0.3
MLP and Attention	Arc MLP size: 500 Label MLP size: 100 # MLP layers: 1 Activation: ReLU Dropout: 0.3
MLP and Attention	Batch size: 128 # Epochs: 100 Early stopping: 50 Adam (Kingma and Ba, 2014) lrate: 0.001 Adam $\beta 1$: 0.9 Adam $\beta 2$: 0.9

Table 6: Hyper-parameters used in our zero-shot experiments.