

Deep Contextualized Word Embeddings in Transition-Based and Graph-Based Dependency Parsing – A Tale of Two Parsers Revisited

Supplementary Material

Artur Kulmizev Miryam de Lhoneux Johannes Gontrum Elena Fano Joakim Nivre

Department of Linguistics and Philology, Uppsala University

{artur.kulmizev,miryam.de.lhoneux,joakim.nivre}@lingfil.uu.se

{johannes.gontrum.4608,elena.fano.3249}@student.uu.se

A Hyperparameters

Component	Specification
w_k dim.	500 (100 * 2 (char) + 300 (word))
ELMo _k dim.	1024
BERT _k dim.	768
BiLSTM output dim.	125
MLP output dim.	100
No. BiLSTM layers	2
BiLSTM dropout rate	0.33
No. training epochs	30
Model selection	best dev. performance

B Error Analysis

B.1 Dependency Length

	GR		GR+B		GR+E		TR		TR+B		TR+E	
	pr	rc	pr	rc	pr	rc	pr	rc	pr	rc	pr	rc
root	0.91	0.91	0.92	0.92	0.92	0.92	0.88	0.88	0.91	0.91	0.90	0.90
1	0.89	0.90	0.90	0.91	0.90	0.91	0.88	0.90	0.90	0.91	0.90	0.92
2	0.83	0.84	0.85	0.86	0.86	0.86	0.82	0.83	0.86	0.86	0.85	0.86
3	0.76	0.77	0.80	0.80	0.80	0.80	0.76	0.75	0.81	0.80	0.80	0.80
4	0.74	0.72	0.78	0.77	0.77	0.76	0.73	0.70	0.78	0.76	0.78	0.76
5	0.70	0.67	0.74	0.72	0.73	0.71	0.67	0.64	0.74	0.71	0.73	0.71
6	0.68	0.66	0.73	0.72	0.72	0.70	0.65	0.63	0.73	0.72	0.70	0.69
7	0.66	0.64	0.72	0.70	0.70	0.69	0.64	0.61	0.72	0.71	0.68	0.67
8	0.65	0.61	0.71	0.67	0.70	0.66	0.63	0.58	0.71	0.67	0.68	0.64
9	0.66	0.64	0.71	0.72	0.69	0.70	0.64	0.63	0.71	0.70	0.68	0.67
>=10	0.69	0.66	0.73	0.71	0.71	0.69	0.65	0.63	0.72	0.71	0.70	0.68

B.2 Distance to Root

	GR		GR+B		GR+E		TR		TR+B		TR+E	
	pr	rc	pr	rc	pr	rc	pr	rc	pr	rc	pr	rc
0	0.91	0.91	0.92	0.92	0.92	0.92	0.88	0.88	0.91	0.91	0.90	0.90
1	0.83	0.81	0.85	0.84	0.85	0.84	0.81	0.80	0.85	0.84	0.84	0.83
2	0.81	0.82	0.84	0.85	0.84	0.84	0.81	0.81	0.84	0.85	0.84	0.84
3	0.81	0.82	0.85	0.85	0.84	0.84	0.81	0.81	0.85	0.85	0.84	0.84
4	0.82	0.82	0.85	0.85	0.84	0.84	0.81	0.82	0.85	0.85	0.85	0.85
5	0.82	0.83	0.85	0.86	0.84	0.85	0.81	0.83	0.85	0.86	0.84	0.85
6	0.82	0.83	0.84	0.86	0.84	0.85	0.82	0.83	0.85	0.86	0.84	0.85
7	0.82	0.83	0.85	0.86	0.85	0.85	0.81	0.83	0.85	0.86	0.85	0.85
8	0.82	0.84	0.83	0.85	0.84	0.84	0.81	0.83	0.84	0.85	0.85	0.85
9	0.83	0.81	0.83	0.82	0.84	0.83	0.79	0.81	0.84	0.83	0.83	0.82
>=10	0.81	0.84	0.84	0.85	0.85	0.87	0.80	0.84	0.83	0.85	0.83	0.87

B.3 Projectivity

	GR		GR+B		GR+E		TR		TR+B		TR+E	
	pr	rc	pr	rc	pr	rc	pr	rc	pr	rc	pr	rc
Non-proj	0.35	0.43	0.36	0.45	0.36	0.45	0.48	0.33	0.49	0.37	0.52	0.40
Proj	0.83	0.83	0.86	0.85	0.85	0.85	0.81	0.82	0.85	0.85	0.85	0.85

B.4 Part of Speech

	GR	GR+B	GR+E	TR	TR+B	TR+E
ADJ	0.82	0.85	0.85	0.82	0.85	0.85
ADP	0.95	0.96	0.96	0.95	0.96	0.96
ADV	0.78	0.80	0.80	0.77	0.80	0.81
AUX	0.93	0.93	0.94	0.92	0.93	0.94
CCONJ	0.85	0.88	0.87	0.84	0.87	0.86
DET	0.94	0.95	0.95	0.94	0.95	0.95
INTJ	0.50	0.44	0.48	0.44	0.45	0.48
NOUN	0.77	0.81	0.80	0.76	0.80	0.80
NUM	0.83	0.84	0.84	0.82	0.84	0.84
PART	0.86	0.89	0.87	0.84	0.90	0.88
PRON	0.83	0.85	0.86	0.82	0.85	0.86
PROPN	0.79	0.84	0.82	0.78	0.84	0.83
PUNCT	0.82	0.85	0.84	0.81	0.86	0.84
SCONJ	0.90	0.93	0.93	0.89	0.93	0.93
SYM	0.85	0.84	0.84	0.82	0.85	0.86
VERB	0.76	0.80	0.79	0.74	0.79	0.78
X	0.64	0.67	0.65	0.64	0.68	0.67

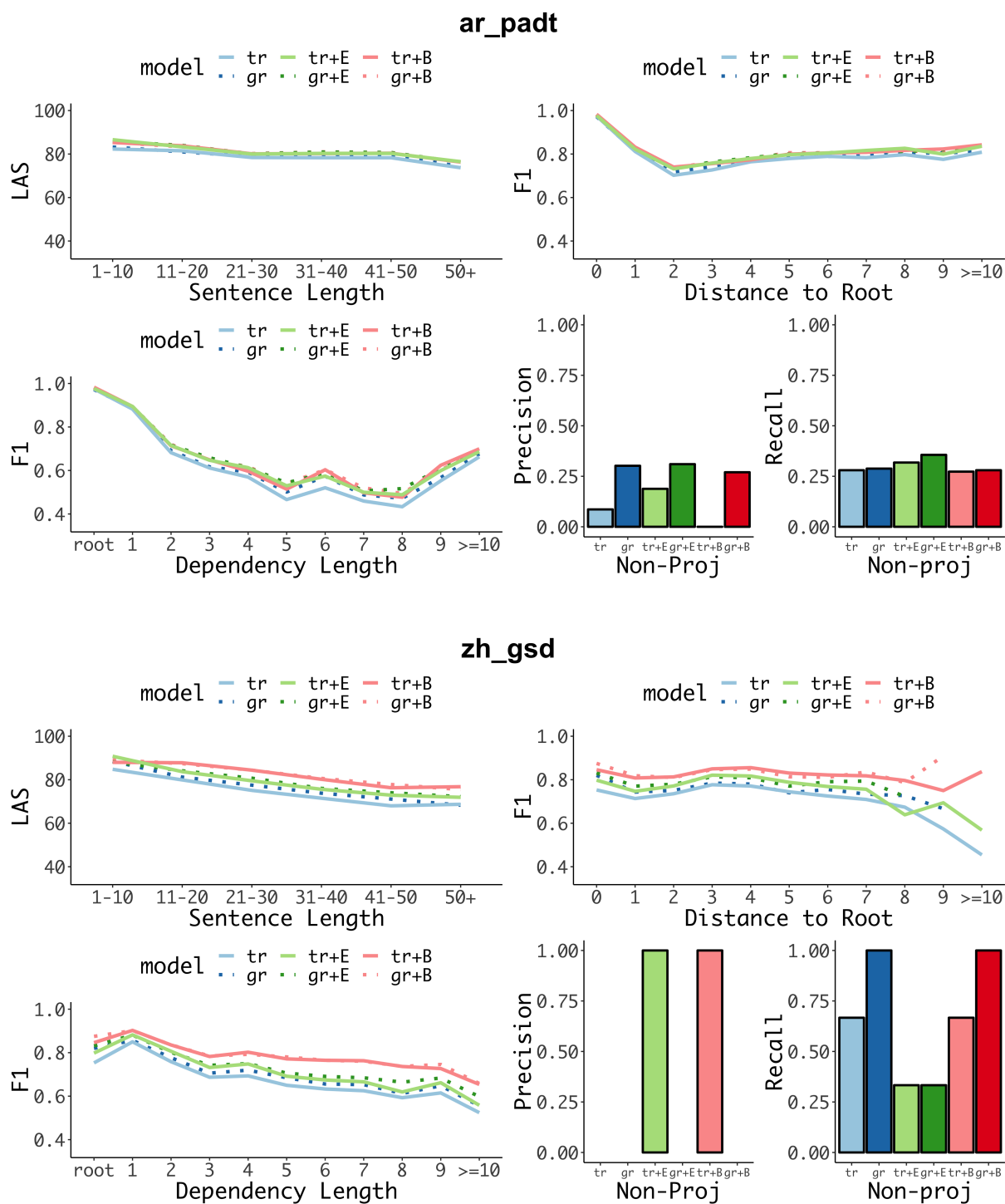
B.5 Dependency Relation

	GR		GR+B		GR+E		TR		TR+B		TR+E	
	pr	rc	pr	rc	pr	rc	pr	rc	pr	rc	pr	rc
acl	0.69	0.69	0.75	0.76	0.73	0.73	0.66	0.66	0.75	0.75	0.73	0.73
advcl	0.68	0.69	0.68	0.71	0.70	0.72	0.66	0.67	0.69	0.71	0.70	0.71
advmod	0.77	0.77	0.79	0.78	0.80	0.79	0.76	0.76	0.79	0.80	0.80	0.80
amod	0.86	0.87	0.88	0.89	0.89	0.90	0.85	0.87	0.88	0.89	0.88	0.90
appos	0.55	0.54	0.61	0.63	0.59	0.60	0.54	0.57	0.62	0.64	0.59	0.61
aux	0.95	0.95	0.95	0.95	0.96	0.96	0.95	0.95	0.95	0.95	0.96	0.96
case	0.95	0.96	0.96	0.97	0.96	0.96	0.95	0.95	0.96	0.97	0.96	0.96
cc	0.84	0.85	0.87	0.88	0.87	0.87	0.83	0.83	0.87	0.88	0.86	0.86
ccomp	0.67	0.64	0.71	0.67	0.71	0.66	0.65	0.60	0.72	0.68	0.71	0.68
clf	0.87	0.90	0.94	0.95	0.92	0.92	0.89	0.87	0.95	0.94	0.92	0.90
compound	0.85	0.84	0.86	0.87	0.87	0.86	0.84	0.84	0.87	0.87	0.86	0.87
conj	0.68	0.69	0.76	0.75	0.72	0.72	0.64	0.65	0.74	0.74	0.71	0.70
cop	0.84	0.85	0.86	0.86	0.87	0.87	0.83	0.84	0.85	0.86	0.86	0.87
csubj	0.59	0.51	0.62	0.54	0.63	0.56	0.55	0.49	0.64	0.53	0.63	0.55
dep	0.68	0.56	0.74	0.60	0.72	0.58	0.65	0.56	0.73	0.60	0.69	0.57
det	0.93	0.94	0.94	0.95	0.94	0.95	0.92	0.93	0.95	0.95	0.94	0.95
discourse	0.74	0.61	0.78	0.62	0.75	0.65	0.72	0.60	0.79	0.65	0.76	0.56
dislocated	0.40	0.32	0.43	0.32	0.44	0.39	0.28	0.26	0.34	0.36	0.35	0.44
expl	0.75	0.81	0.79	0.82	0.79	0.81	0.75	0.79	0.79	0.83	0.81	0.82
fixed	0.78	0.72	0.81	0.76	0.82	0.77	0.78	0.73	0.82	0.76	0.82	0.78
flat	0.79	0.78	0.85	0.83	0.83	0.81	0.79	0.78	0.83	0.82	0.82	0.82
goeswith	0.71	0.44	0.92	0.31	0.68	0.44	0.61	0.28	0.84	0.41	0.75	0.46
iobj	0.83	0.81	0.82	0.80	0.86	0.82	0.83	0.78	0.85	0.79	0.85	0.83
list	0.31	0.28	0.32	0.32	0.38	0.33	0.28	0.23	0.34	0.33	0.41	0.36
mark	0.89	0.87	0.91	0.91	0.91	0.91	0.88	0.87	0.91	0.91	0.91	0.90
nmod	0.78	0.78	0.81	0.81	0.80	0.81	0.77	0.77	0.81	0.81	0.80	0.81
nsubj	0.79	0.80	0.82	0.83	0.82	0.83	0.77	0.78	0.82	0.83	0.81	0.82
nummod	0.87	0.90	0.88	0.91	0.88	0.91	0.87	0.89	0.88	0.92	0.88	0.92
obj	0.79	0.80	0.82	0.83	0.83	0.84	0.77	0.78	0.83	0.83	0.84	0.83
obl	0.74	0.75	0.77	0.77	0.77	0.78	0.73	0.73	0.77	0.77	0.77	0.78
orphan	0.29	0.11	0.31	0.12	0.35	0.11	0.21	0.10	0.35	0.13	0.38	0.15
parataxis	0.77	0.70	0.80	0.72	0.77	0.72	0.75	0.70	0.80	0.74	0.79	0.73
punct	0.82	0.82	0.85	0.85	0.84	0.84	0.81	0.81	0.86	0.86	0.84	0.84
root	0.90	0.90	0.92	0.92	0.92	0.92	0.90	0.88	0.92	0.91	0.92	0.90
vocative	0.43	0.28	0.43	0.32	0.65	0.43	0.53	0.35	0.58	0.38	0.52	0.38
xcomp	0.68	0.62	0.71	0.66	0.70	0.68	0.65	0.62	0.74	0.69	0.73	0.70

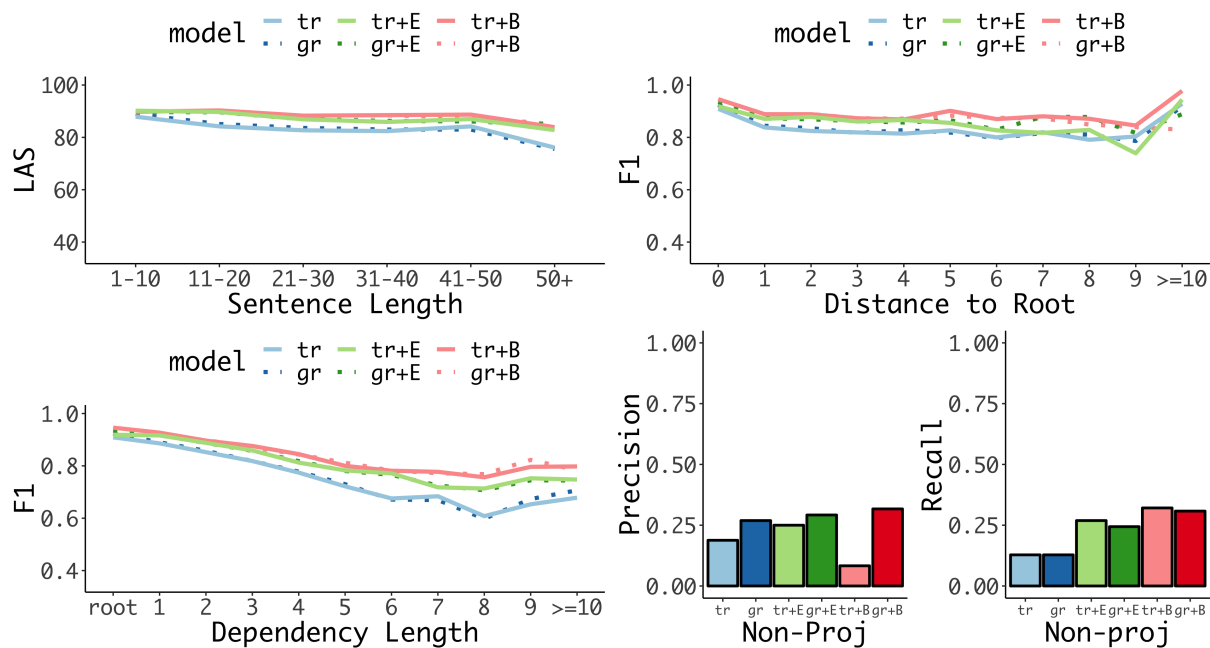
B.6 Sentence Length

	GR	GR+B	GR+E	TR	TR+B	TR+E
1-10	83.50	83.40	84.90	81.80	83.40	84.80
11-20	83.50	85.80	86.20	82.10	85.70	86.00
21-30	83.00	86.20	85.50	82.00	86.10	85.10
31-40	81.90	85.20	84.30	81.00	84.90	84.20
41-50	81.70	84.40	83.80	80.70	84.10	83.60
50+	78.70	81.60	81.50	78.60	82.20	81.70

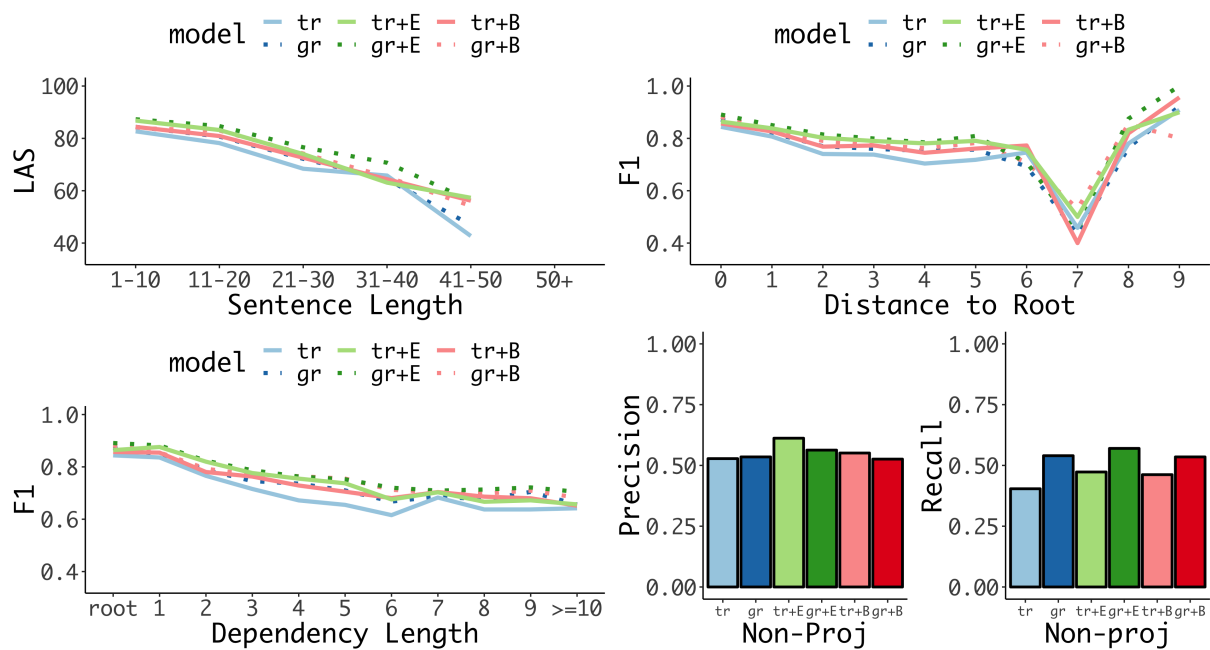
C Per-Language Error Analysis



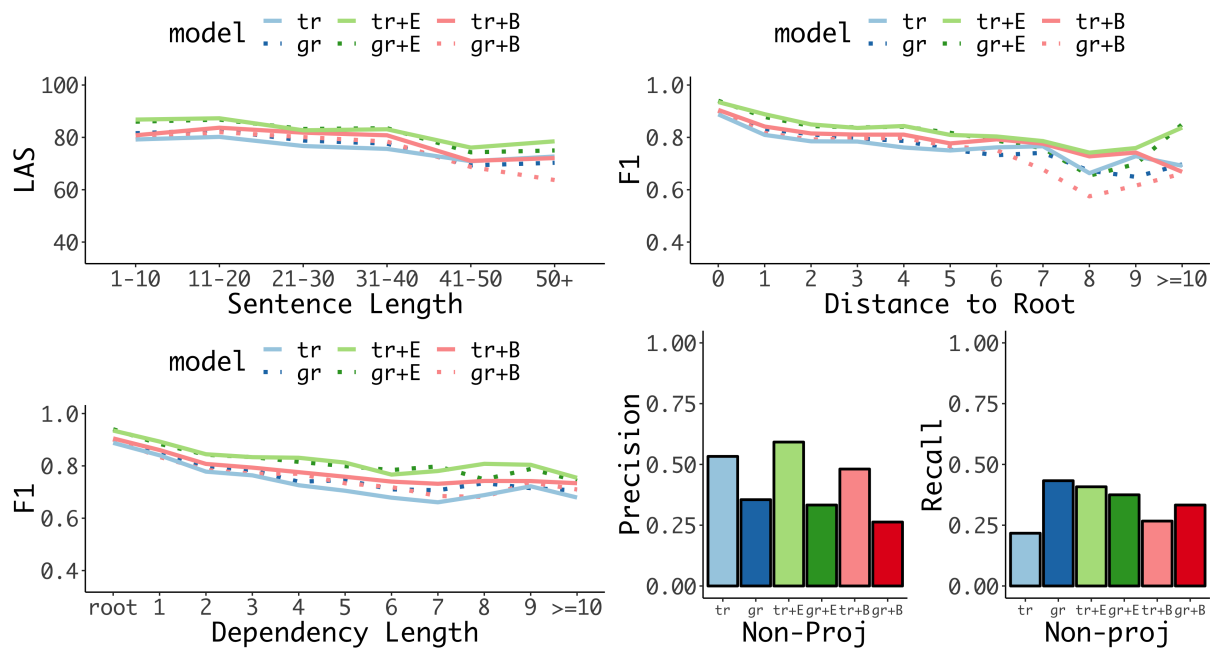
en_ewt



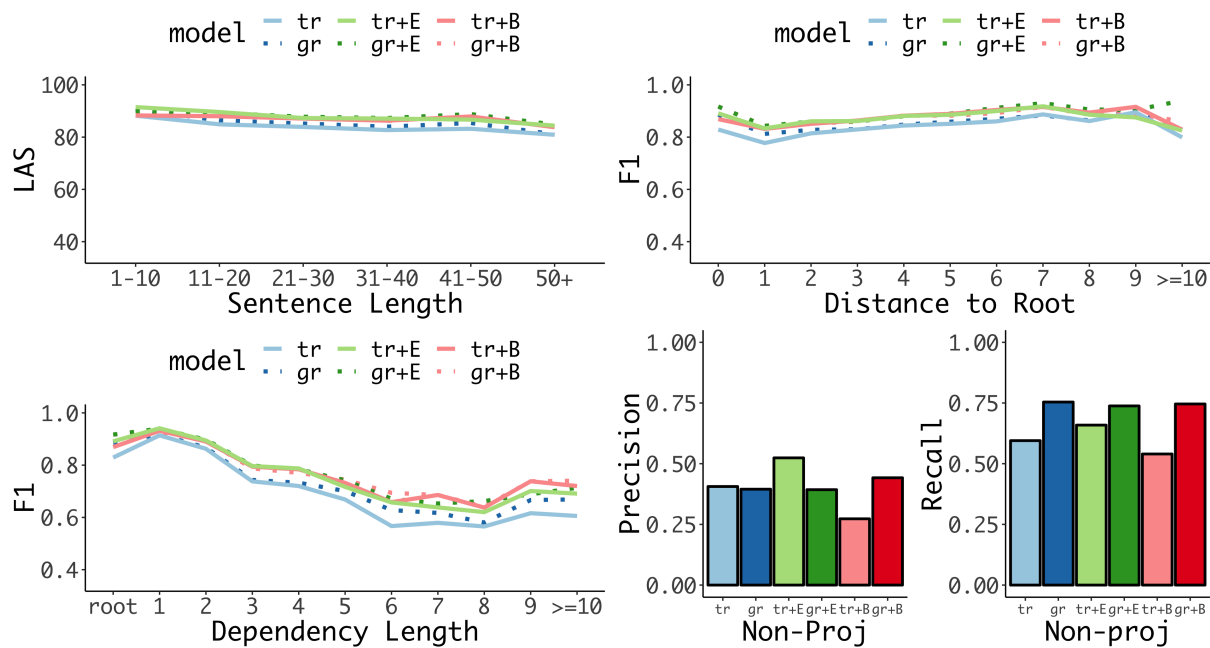
eu_bdt



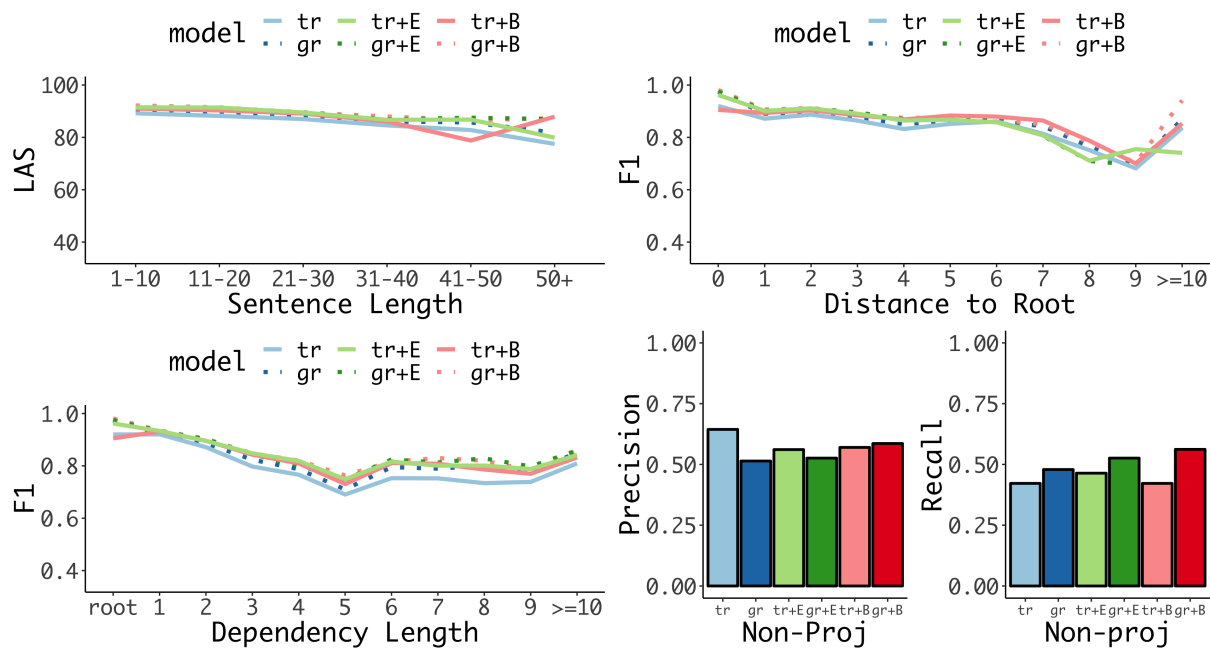
fi_tdt



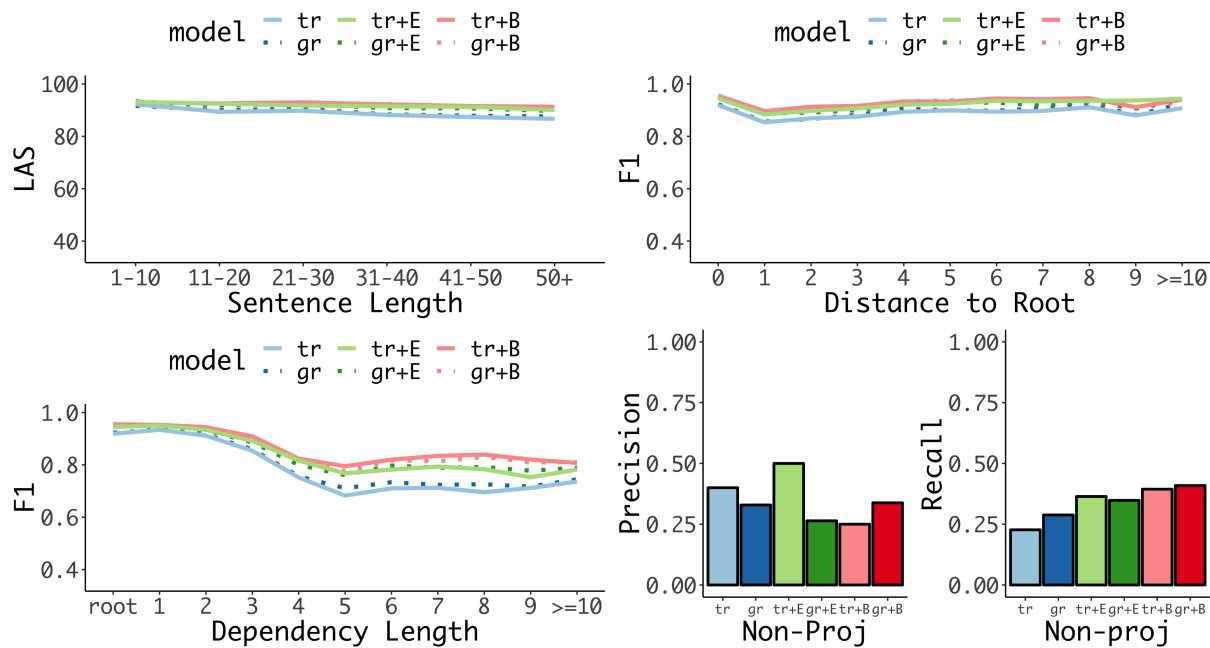
he_htb



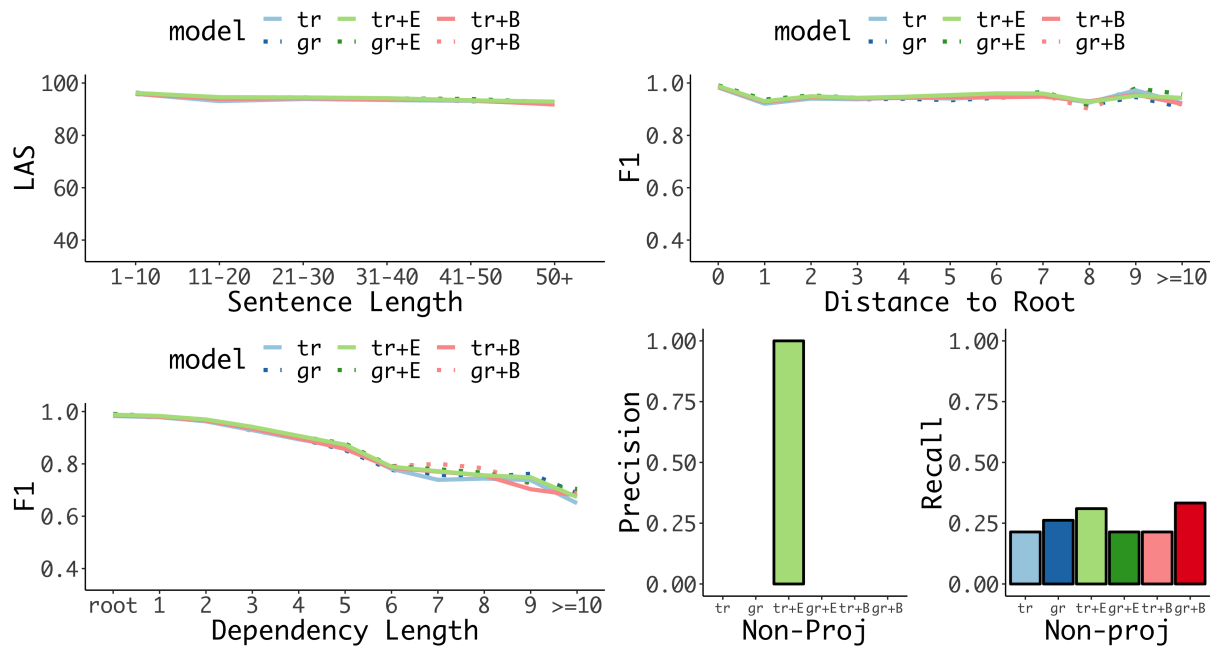
hi_hdtb



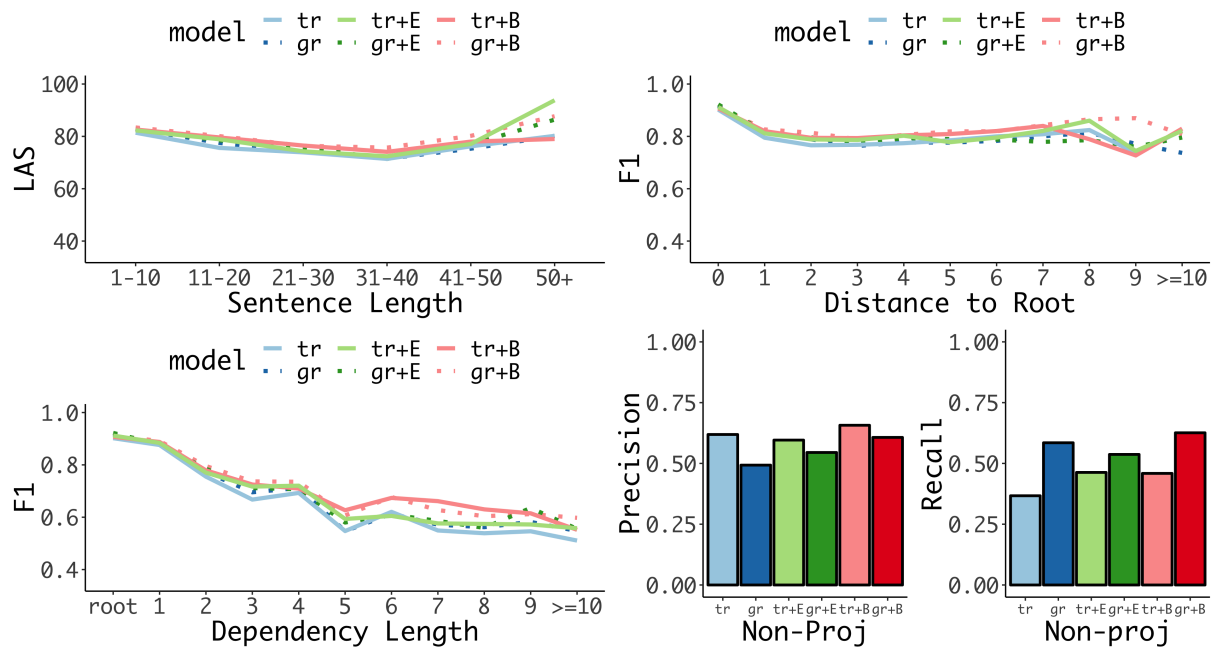
it_isdt



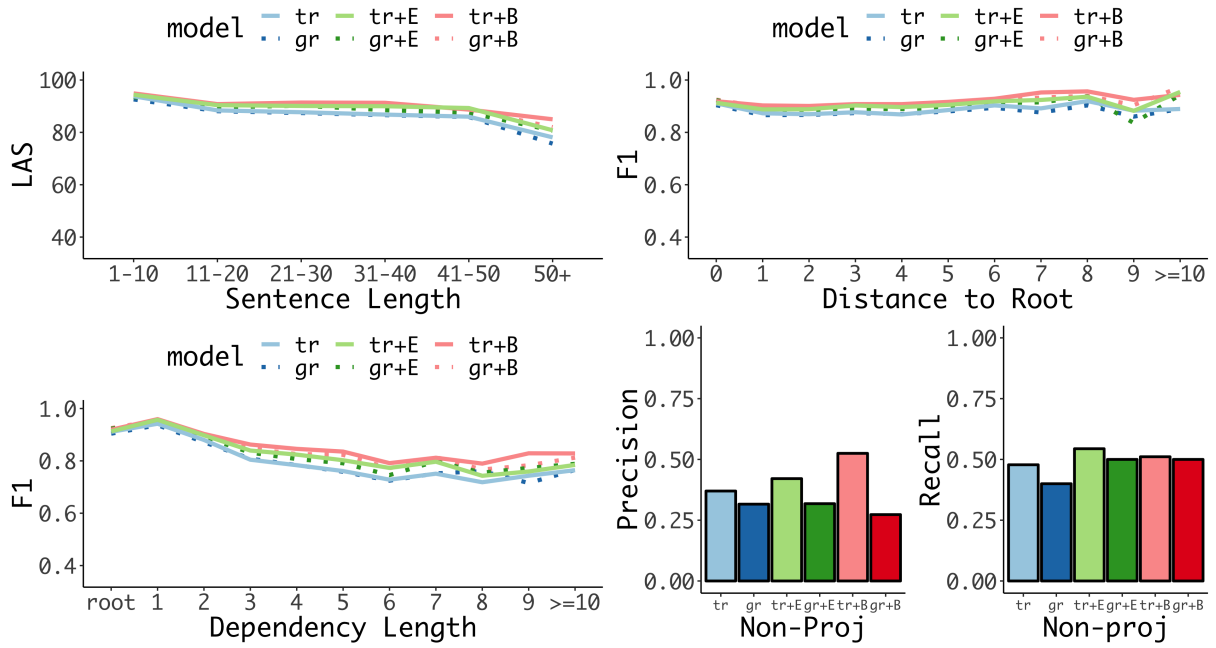
ja_gsd



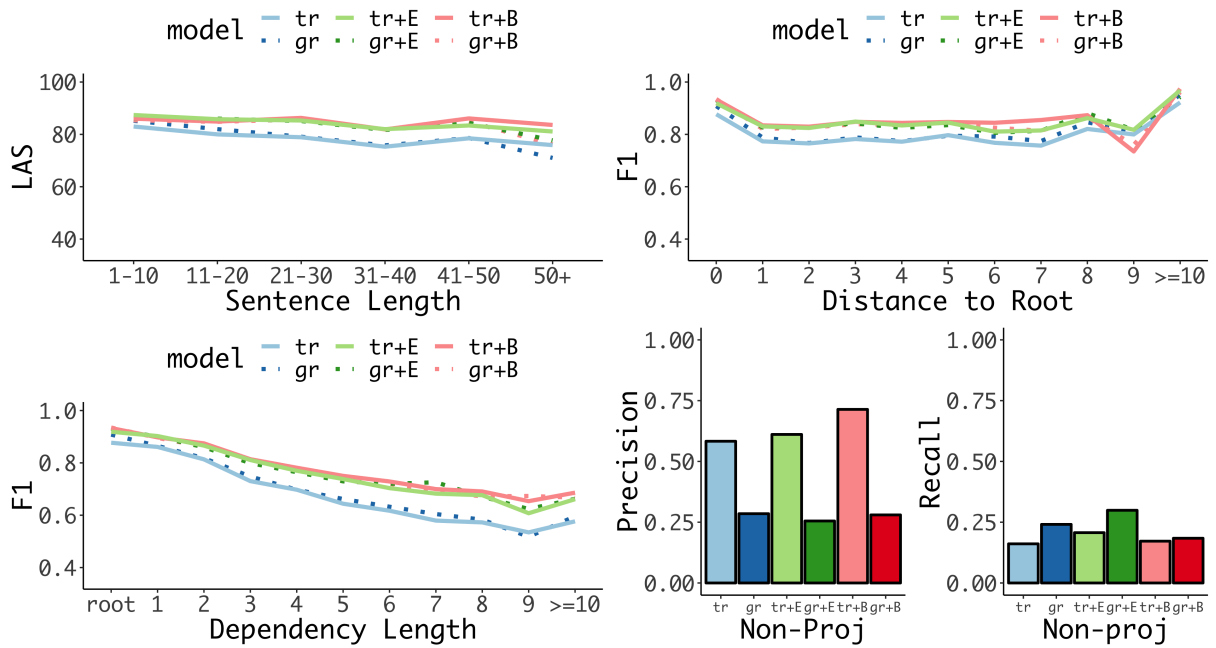
ko_gsd



ru_syntagrus



sv_talbanken



tr_imst

