

languages		low-resource ($ \mathcal{D}_\tau = 100$)		high-resource ($ \mathcal{D}_\tau = 10000$)	
τ	σ_i	log-linear	neural	Δ	
gl	—	57.64	49.19	-8.45	89.42 +2.19
gl	es	71.46	76.40	+4.94	89.46 +1.96
gl	ca	67.32	75.40	+8.08	89.32 +1.92
gl	it	63.81	70.93	+7.12	89.50 +2.16
gl	fr	58.22	68.02	+9.80	89.38 +1.46
gl	ro	59.23	67.76	+8.44	89.19 +1.95
fy	—	62.71	58.43	-4.28	91.03 +0.61
fy	nl	68.15	72.12	+3.97	91.01 +0.07
tl	—	58.15	56.98	-1.17	79.03 +4.79
tl	ceb	75.29	81.79	+6.50	79.51 +5.48
uk	—	61.40	60.65	-0.75	87.39 +1.75
uk	ru	70.94	76.74	+5.80	87.42 +1.41
mr	—	42.76	39.02	-3.73	74.95 +4.86
mr	hi	54.25	60.92	+6.67	74.49 +4.04
mr	ur	49.32	58.92	+9.60	74.81 +4.07

Table 2: Results comparing the log-linear and neural CRFs in various settings. We compare the log-linear linear and the neural CRF in the low-resource transfer setting. The difference (Δ) is blue when positive and red when negative.