

Results of the PolEval 2020 Shared Task 2: Morphosyntactic tagging of Middle, New and Modern Polish

Marcin Woliński, Witold Kieraś, Dorota Komosińska
Institute of Computer Science, Polish Academy of Sciences

Włodzimierz Gruszczyński
Institute of Polish Language, Polish Academy of Sciences

October 26, 2020

The Task

The goal of Task 2 was to disambiguate morphosyntactic interpretations and to guess the interpretation for unknown words — exactly as in Subtask 1A of PoEval 2017.

The data for this year's task covers 400 years of the development of the Polish language and represents three different periods of development of Polish:

- ▶ Middle (KorBa — a corpus of the 17th and 18th century),
- ▶ New (a corpus of the 19th century),
- ▶ Modern (1M subcorpus of the National Corpus of Polish NKJP).

Data split between three periods of time

	train		devel		test	
	# segm.	%	# segm.	%	# segm.	%
Baroque	408,248	28.32%	20,005	49.99%	20,026	50.00%
19th c.	613,914	42.59%	12,001	29.99%	12,001	29.97%
Contemp.	419,346	29.08%	8,010	20.02%	8,018	20.03%
Total:	1,441,508	100%	40,016	100%	40,045	100%

Competition results for Task 2 sorted by overall accuracy

System	Accuracy			
	overall	on known	on unknown	manual not ign
KFTT train+devel	0.9573	0.9607	0.8102	0.6781
KFTT train	0.9564	0.9600	0.7991	0.6661
KFTT train+devel wo_morf	0.9563	0.9595	0.8191	0.6730
Simple Baselines: XLM-R	0.9499	0.9562	0.6770	0.6850
Simple Baselines: COMBO	0.9284	0.9363	0.5838	0.5232
CMC Graph Heuristics	0.9121	0.9214	0.5072	0.1670
Alium-1000	0.8880	0.8985	0.4306	0.2427
Alium-1.25	0.8880	0.8985	0.4295	0.2427

Results by historical period

System	Period	Accuracy			
		overall	known	unknown	manual not ign
KFTT train+devel	Baroque	0.9435	0.9483	0.7943	0.6574
	19th c.	0.9694	0.9715	0.8324	0.6316
	Contemp.	0.9737	0.9748	0.8778	0.8043
KFTT train	Baroque	0.9422	0.947	0.7943	0.6574
	19th c.	0.9692	0.9718	0.7933	0.5789
	Contemp.	0.9727	0.9741	0.8444	0.7609
KFTT train+devel wo_morf	Baroque	0.942	0.9464	0.8054	0.6505
	19th c.	0.969	0.9709	0.8436	0.6316
	Contemp.	0.9733	0.9745	0.8667	0.8043
Simple baselines: XLM-R	Baroque	0.9302	0.9392	0.6535	0.6620
	19th c.	0.9672	0.9715	0.6816	0.6667
	Contemp.	0.9733	0.9749	0.8333	0.8043