

SYNTACTIC TAGS IN THE LITHUANIAN TREEBANK ALKSNIS

Vytautas Magnus University, Centre of Computational Linguistics, 2017

Syntactic tags, information about syntactic relations, and dependencies between words are based on the Lithuanian grammar and this book: Jan Hajič, Jarmila Panevová, Eva Buráňová, Zdeňka Urešová, Alla Bémová *ANNOTATIONS AT ANALYTICAL LEVEL. Instructions for annotators* (11.10.1999), UK MFF ÚFAL Praha.

Category	Tag	Notes / Examples
Predicate	Pred	<i>Jonas bėga</i>
Nominal predicate	PredN	<i>Jonas yra gydytojas; Jonas yra gražus</i> [yra is marked as Pred, gydytojas, gražu – PredN]
Verbal predicate	PredV	<i>Noriu eiti; reikia pagalvoti</i> [noriu, reikia is marked as Pred, eiti, pagalvoti – PredV]
Subject	Sub	<i>Jonas bėga; man reikia; atėjo svečių</i>
Object	Obj	Direct and indirect objects are marked identically.
Attribute	Atr	<i>Geras oras; mokyklos direktorė</i>
Adjunct	Adj	<i>Rytoj ateisiu, vaikšto lėtai</i>
Comma that does not mark a syntactic relation	Aux	A comma before coordinating or subordinating conjunctions: <i>ir Jonas</i> , [Aux] <i>ir Petras</i> ; <i>pamatė</i> , [Aux] <i>kad</i> [AuxC]...
Prepositional tag	AuxP	<i>Einu į mišką</i> Pred ->AuxP ->Adj
The last punctuation mark in a sentence	AuxK	
Subordinating conjunctions	AuxC	<i>Žinau, kad...</i> <i>Teisė vairuoti buvo atimta ilgiau negu</i> [AuxC] <i>vieneriems metams...</i>
Coordinating conjunctions and punctuation marks	Coord	<i>Jonas ir Petras šoka ir dainuoja; Jonas, Petras</i>
Parenthesis	Par	<i>Tiesa, pasak, etc.</i>
Auxiliary lexical unit	AuxL	Foreign lexical units, for example: <i>JAV kompanija North American Investment Consulting Inc.</i> [North American Investment Consulting is marked as AuxL; Inc as Atr]
Units connected by coordinating relation (additional tag)	Atr_Co, Adj_Co, Sub_Co, Obj_Co ir	<i>metalų</i> [Atr_Co], <i>statybinių</i> [Atr] <i>medžiagų</i> [Atr_Co], <i>urano</i> [Atr_Co] <i>kainos</i> [Sub] <i>nuolat</i> [Adj] <i>kyla</i> [Pred].

	t. t.	
Ellipsis (additional tag)	ExD	Pvz., Obj_ExD
Types of subordinate clauses	Pred_Atr, Pred_Sub...	A predicate in a composite subordinate sentence is tagged by these or other similar tags designating the type of the subordinate clause, i.e. modifying, subjective, objective, etc. If predicates in different clauses are connected by coordination, then addition Co is added at the end of the tag, e.g. Pred_Obj_Co.
Particle tag	AuxZ	If a particle does not play a syntactic role, then it needs to be marked as AuxZ.

Note: [...] – denotes skipped text (e.g., a phrase in foreign language).