

Tagset for Malayalam

Sl. No	Category			Label	Annotation Convention**	Example	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	Noun			N	N	സീത, വീട്, മോഹൻ	
1.1		Common		NN	N_NN	വെള്ളം, വീട്, പട്ടം	
1.2		Proper		NNP	N_NNP	രവി, സീത, മോഹൻ	
1.3		Nloc		NST	N_NST	മുകളിൽ, മുന്നിൽ, പിന്നിൽ, താഴെ	
2	Pronoun			PR	PR	അവൻ, അവളു്, അത്, ഇത്	
2.1		Personal		PRP	PR_PRP	ഞാൻ, നീ, ഞങ്ങളു്, നിങ്ങളു്, അത്, ഇത്	
2.2		Reflexive		PRF	PR_PRF	താൻ, തന്നെത്താൻ, സ്വയം,	
2.3		Relative		PRL	PR_PRL	ആരോ	
2.4		Reciprocal		PRC	PR_PRC	തമ്മിൽ, തമ്മിൽതമ്മിൽ, പരസ്പരം	
2.5		Wh-word		PRQ	PR_PRQ	ആര്, എന്ത്	

3	Demonstrative				DM	DM	ആ, ഈ	
3.1		Deictic			DMD	DM_DMD	എന്ത്	
3.2		Relative			DMR	DM_DMR	എന്ത്	
3.3		Wh-word			DMQ	DM_DMQ	ആര്	
4	Verb				V	V	പോ-, കഴി-, ചിരി-, ആണ്	
4.1		Main			VM	V_VM	പോ-, കഴി-, ചിരി-, ആണ്	
4.1.1			Finite		VF	V_VM_VF	കഴിക്കുന്നു, ചിരിക്കും, ആണ് (copula), പോയി,	
4.1.2			Non-finite		VNF	V_VM_VNF	പോയ, കഴിച്ച, ചിരിച്ച, പറഞ്ഞത്, വന്നത്, ആയി	
4.1.2.1				Verbal participle		V_VM_VNF_V BN	വീണു, ഓടി	
4.1.2.2				Relative participle verb		V_VM_VNF_R P	വീണ, ഓടിയ	
4.1.2.3				Conditional		V_VM_VNF_C		

				verb		OND	വീണാ ല്, ഓടീയാല്	
4.1. 2.4				Infinitive verb		V_VM_VNF_I NF	വരാന് (വരുവാന്) (Purposive Infinitive), പറ യുക (Infinitive)	
4.1.3.			Gerund		VNG	V_VM_VNG	വീഴല്, ഓടല്	
4.2		Verbal Noun	Verbal noun		NNV	N_NNV	പഠിത്തം , നടത്തം , ഓട്ടം	Verbal Noun
4.3		Auxiliary			VAUX	V_VAUX	തീരും, കഴിയും	
5	Adjective				JJ		വലിയ, ചെറിയ	
6	Adverb				RB		വേഗം, പെട്ടെന്നു, പതുക്കെ	Only manner adverbs
7	Postposition				PSP		പറ്റി, കൂടെ, കുറിച്ച്	
8	Conjunction				CC	CC	പക്ഷെ, എന്നിട്ടും, എങ്കിലും, എന്നാല് , എന്നാലും	
8.1		Co-ordinator			CCD	CC_CCD	പക്ഷെ, അഥവാ	
8.2		Subordinator			CCS	CC_CCS	എന്നാല് , എന്നിട്ടും , എന്നാലും , എങ്കിലും	

		Quotative			UT	CC_CCS_UT	എന്ന്,എന്ന	
9	Particles				RP	RP	മാത്രം,കൂടി, പോലും	
9.1		Default			RPD	RP_RPD	മാത്രം,കൂടി, പോലും	
9.2		Classifier			CL	RP_CL	തുടങ്ങിയ,മു തലായ, എന്നീ	
9.3		Interjection			INJ	RP_INJ	ഹാവു, അയ്യോ	
9.4		Intensifier			INTF	RP_INTF	വളരെ, ഏറെ	
9.5		Negation			NEG	RP_NEG	ഇല്ല, അല്ല	
10	Quantifiers				QT	QT	കുറച്ച്, ധാരാളം	
10.1		General			QTF	QT_QTF	കുറച്ച്, ധാരാളം	
10.2		Cardinals			QTC	QT_QTC	ഒന്ന്, രണ്ട്	
10.3		Ordinals			QTO	QT_QTO	ഒന്നാം, രണ്ടാം, രണ്ടാമത്തെ	
11	Residuals				RD	RD		
11.1		Foreign			RDF	RD_RDF		
11.2		Symbol			SYM	RD_SYM	\$, &, *, (,), ruu.	A word written in word script other than the script of the original text
11.3		Punctuation			PUNC	RD_PUNC	., : ;	For symbols such as \$, & etc

11.4		Unknown			UNK	RD_UNK		Only for punctuations
11.5		Echowords			ECH	RD_ECH	രണ്ട്, മുണ്ട് പല, തല	