

Appendix C: POS and Chunk Tag Set for Indian Languages (ILMT Tagset)

S. No.	Category	Tag name	Example
1.1	Noun	NN	
1.2	NLoc	NST	
2.	Proper Noun	NNP	
3.1	Pronoun	PRP	
3.2	Demonstrative	DEM	
4	Verb-finite	VM	
5	Verb Aux	VAUX	
6	Adjective	JJ	
7	Adverb	RB	*Only manner adverb
8	Post position	PSP	
9	Particles	RP	bhI, to, hI, jI, hA.N, na,
10	Conjuncts	CC	bole (Bangla)
11	Question Words	WQ	
12.1	Quantifiers	QF	bahut, tho.DA, kam (Hindi)
12.2	Cardinal	QC	
12.3	Ordinal	QO	
12.4	Classifier	CL	
13	Intensifier	INTF	
14	Interjection	INJ	
15	Negation	NEG	
16	Quotative	UT	ani (Telugu), endru (Tamil), bole/mAne (Bangla), mhaNaje (Marathi), mAne (Hindi)
17	Sym	SYM	
18	Compounds	*C	
19	Reduplicative	RDP	
20	Echo	ECH	
21	Unknown	UNK	

S. No	Chunk Type	Tag Name	Example
1	Noun Chunk	NP	Hindi: ((merA nayA ghara))_NP “my new house”
2.1	Finite Verb Chunk	VGF	Hindi: mEMne ghara para khAnA ((khAyA_VM))_VGF
2.2	Non-finite Verb Chunk	VGNF	Hindi: mEMne ((khAte – khAte_VM))_VGNF ghode ko dekhA
2.3	Infinitival Verb Chunk	VGINF	Bangla: bindu Borabela ((snAna karawe))_VGINF BAlobAse
2.4	Verb Chunk (Gerund)	VGNN	Hindi: mujhe rAta meM ((khAnA_VM))_VGNN acchA lagatA hai
3	Adjectival Chunk	JJP	Hindi: vaha laDaZki hE((suMdara JJ sI RP)) JJP
4	Adverb Chunk	RBP	Hindi : vaha ((dhIre-dhIre_RB))_RBP cala rahA thA
5	Chunk for Negatives	NEGP	Hindi: ((binA))_NEGP ((kucha))_NP ((bole))_VG ((kAma))_NP ((nahIM calatA))_VG
6	Conjuncts	CCP	Hindi: ((rAma))_NP ((Ora))_CCP ((SyAma))_NP
7	Chunk Fragments	FRAGP	Hindi: rAma (jo merA baDZA bhAI hE) ne kahA.
8	Miscellaneous	BLK	