

# **Unified Parts of Speech (POS) Standard in Indian Languages**

**- Draft Standard –Version 1.0**



**Department of Information Technology  
Ministry of Communications & Information Technology  
Govt. of India**

## CONTENTS

1. INTRODUCTION
2. SCOPE
3. TERMINOLOGY
  - 3.1 POS Tag
  - 3.2 XML Schema
  - 3.3 Metadata
4. WHAT IS A POS TAG
5. REQUIREMENTS OF A POS TAG
  - 5.1 Need of XML Schema in designing common POS format
6. POS TAG SET FOR INDIAN LANGUAGES
7. XML INTERNATIONALIZATION BEST PRACTICES
  - 7.1 What is Internationalization Tag Set (ITS)
8. XML SCHEMA
9. METADATA ON POS
10. ONE TO ONE MAPPING LABELS IN POS SCHEMA
11. POS SCHEMA BLOCK DIAGRAM
12. DRAFT POS SCHEMA FOR INDIAN LANGUAGES USING XML
13. ONE TO ONE MAPPING LABELS FOR INDIAN LANGUAGES
14. ALGORITHM FOR SELECTION OF NODES
15. REFERENCE BASED IMPLEMENTATION
16. REFERENCE

## ANNEXURES

- A. Language Code Table

## 1. INTRODUCTION

Parts of Speech tagging is one the key building blocks (noun, pronoun, verb, demonstrative, etc) for developing Natural Language Processing applications. This POS schema is based on W3C XML Internationalization best practices, ISO 639-3 Language Codes for Language Identification, ISO 12620:1999 as metadata definition and one to one mapping table for all the labels used in POS Schema.

This document sets out the structural part of the XML Schema definition language and also how to make XML POS Schema for tagging. XML Schemas including an introduction to the nature of XML Schemas and an introduction to the XML POS Schema abstract data model, along with other terminology used throughout this document and also specifies the precise semantics of each component of the abstract model, the representation of each component in XML. This document contains block diagram that shows the flow-chart of creating XML scheme for POS tagging. It also includes the algorithm that contains metadata as per ISO 12620:1999.

## 2. SCOPE

The common unified XML based POS Schema for Indian Languages based on W3C Internationalization best practices have been formulated. The schema has been developed to take into account the NLP requirements for Web based services in Indian Languages. This standard specifies XML POS Schema for tagging. This portion of the XML Schema Language discusses labels that can be used in an XML POS Schema.

## 3. TERMINOLOGY

**3.1 POS Tag:** A Part-Of-Speech Tagger (POS Tagger) is a piece of software that reads text in some language and assigns parts of speech to each word.

**3.2 XML Schema:** XML Schemas express shared vocabularies and allow machines to carry out rules made by people and to define a class of XML documents, and so the term "instance document" is often used to describe an XML document that conforms to a particular schema.

**3.3 Metadata:** Metadata describes how and when and by whom a particular set of data was collected, and how the data is formatted.

#### **4. WHAT IS A POS TAG**

A Part-Of-Speech Tagger (POS Tagger) is a piece of software that reads text in some language and assigns parts of speech to each word. Parts of speech include nouns, verbs, adverbs, adjectives, pronouns, conjunction and their sub-categories.

The input to a tagging algorithm is a string of words of a natural language sentence and a specified tag set (a finite list of Part-of-speech tags). The output is a single best POS tag for each word.

#### **5. REQUIREMENT OF A POS TAG**

The POS tagger can be used as a pre-processor. Text indexing and retrieval uses POS information. POS tagger is used for making tagged corpora and Machine Translation System. Speech processing uses POS tags to decide the pronunciation.

POS tagger would be needed to identify the tag for the words that could not be analysed by the morphological analyser. If the Morph gives multiple tags for a word, then the tagger could be used to resolve the ambiguity.

##### **5.1 NEED OF XML SCHEMA IN DESIGNING COMMON POS FORMAT**

The need of XML for creating POS tag-set is to standardize the POS tag framework for all Indian languages.

The main benefits of xml in using POS tag set for IL's are:

- It Supports multilingual documents and Unicode
- XML allows developers to add extra information to a format without breaking applications.
- XML documents can be stored without using database administrator, because they contain meta data in the form of tags and attributes.
- The tree structure of XML documents allows documents to be compared and aggregated efficiently element by element.
- XML documents can consist of nested elements that are distributed over multiple remote servers

It is easier to convert data between different data types.

## 6. POS Tag set for Indian Languages

### POS Categories and Labels

Sl. No	Category			Label	Annotation Convention**	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)			
<b>1</b>	<b>Noun</b>			<b>N</b>	<b>N</b>	
1.1		Common		NN	N_NN	
1.2		Proper		NNP	N_NNP	
1.3		Verbal		NNV	N_NNV	The verbal noun sub type is only for languages such as Tamil and Malayalam)
1.4		Nloc		NST	N_NST	
<b>2</b>	<b>Pronoun</b>			<b>PR</b>	<b>PR</b>	
2.1		Personal		PRP	PR_PRP	
2.2		Reflexive		PRF	PR_PRF	
2.3		Relative		PRL	PR_PRL	
2.4		Reciprocal		PRC	PR_PRC	
2.5		Wh-word		PRQ	PR_PRQ	
2.6		INDEFINITE		PRI	PR_PRI	
<b>3</b>	<b>Demonstrative</b>			<b>DM</b>	<b>DM</b>	
3.1		Deictic		DMD	DM_DMD	
3.2		Relative		DMR	DM_DMR	
3.3		Wh-word		DMQ	DM_DMQ	
3.4		Indefinite		DMI	DM_DMI	
<b>4</b>	<b>Verb</b>			<b>V</b>	<b>V</b>	
4.1		Main		VM	V_VM	
4.1.1			Finite	VF	V_VM_VF	
4.1.2			Non-finite	VNF	V_VM_VNF	
4.1.3			Infinitive	VINF	V_VM_VINF	
4.1.4			Gerund	VNG	V_VM_VNG	
4.2		Verbal		VN	V__VN	<i>paTittam,</i>

						<i>naTattam, naTanam</i>
4.2		Auxiliary		VAUX	V_VAUX	
4.2.1			Finite	VAUX	V_VAUX_VF	
4.2.2			Non-finite	VNF	V_VAUX_VNF	
4.2.3			Infinitive	VINF	V_VAUX_VINF	
4.2.4			Gerund	VNG	V_VAUX_VNG	
4.2.5			PARTICIP LE NOUN	VNP	V_VAUX_VNP	
<b>5</b>	<b>Adjective</b>			<b>JJ</b>		
<b>6</b>	<b>Adverb</b>			<b>RB</b>		Only manner adverbs
<b>7</b>	<b>Postposition</b>			<b>PSP</b>		
<b>8</b>	<b>Conjunction</b>			<b>CC</b>	<b>CC</b>	
8.1		Co-ordinator		CCD	CC_CCD	
8.2		Subordinator		CCS	CC_CCS	
8.2.1			Quotative	UT	CC_CCS_UT	
<b>9</b>	<b>Particles</b>			<b>RP</b>	<b>RP</b>	
9.1		Default		RPD	RP_RPD	
9.2		Classifier		CL	RP_CL	
9.3		Interjection		INJ	RP_INJ	
9.4		Intensifier		INTF	RP_INTF	
9.5		Negation		NEG	RP_NEG	
<b>10</b>	<b>Quantifiers</b>			<b>QT</b>	<b>QT</b>	
10.1		General		QTF	QT_QTF	
10.2		Cardinals		QTC	QT_QTC	
10.3		Ordinals		QTO	QT_QTO	
<b>11</b>	<b>Residuals</b>			<b>RD</b>	<b>RD</b>	
11.1		Foreign word		RDF	RD_RDF	A word written in script other than the script of the original text
11.2		Symbol		SYM	RD_SYM	For symbols such

						as \$, & etc
11.3		Punctuation		PUNC	RD_PUNC	Only for punctuations
11.4		Unknown		UNK	RD_UNK	
11.5		Echowords		ECH	RD_ECH	

## POS for Hindi

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	Noun			N	N	ladakaa, raajaa, kitaaba	
1.1		Common		NN	N_NN	kitaaba, kalama, cashmaa	
1.2		Proper		NNP	N_NNP	Mohan, ravi, rashmi	
1.4		Nloc		NST	N_NST	Uupara, niice, aage, piiche	
2	Pronoun			PR	PR	Yaha, vaha, jo	
2.1		Personal		PRP	PR_PRP	Vaha, main, tuma, ve	
2.2		Reflexive		PRF	PR_PRF	Apanaa, swayam, khuda	
2.3		Relative		PRL	PR_PRL	Jo, jis, jab, jahaaM,	
2.4		Reciprocal		PRC	PR_PRC	Paraspara, aapasa	
2.5		Wh-word		PRQ	PR_PRQ	Kauna, kab, kahaam	
		Indefinite		PRI	PR_PRI	Koii, kis	

3	Demonstrative			DM	DM	Vaha, jo, yaha,	
3.1		Deictic		DMD	DM_DMD	Vaha, yaha	
3.2		Relative		DMR	DM_DMR	jo, jis	
3.3		Wh-word		DMQ	DM_DMQ	kis, kaun	
		Indefinite		DMI	DM_DMI	KoI, kis	
4	Verb			V	V	giraa, gayaa, sonaa, haMstaa, hai, rahaa	
4.1		Main		VM	V_VM	giraa, gayaa, sonaa, haMstaa,	
4.2		Auxiliary		VAUX	V_VAUX	hai, rahaa, huaa,	
5	Adjective			JJ	JJ	sundara, acchaa, baRaa	
6	Adverb			RB	RB	jaldii, teza	
7	Postposition			PSP	PSP	ne, ko, se, mein	
8	Conjunction			CC	CC	aur, agar, tathaa, kyonki	
8.1		Co-ordinator		CCD	CC_CCD	aur, balki, parantu	
8.2		Subordinator		CCS	CC_CCS	Agar, kyonki, to, ki	
9	Particles			RP	RP	to, bhii, hii	
9.1		Default		RPD	RP_RPD	to,bhii, hii	
9.3		Interjection		INJ	RP_INJ	are, he, o	
9.4		Intensifier		INTF	RP_INTF	bahuta, behada	
9.5		Negation		NEG	RP_NEG	nahiin, mata, binaa	
10	Quantifiers			QT	QT	thoRaa, bahuta, kucha, eka, pahala	



10.1		General		QTF	QT_QTF	thoRaa, bahuta, kucha	
10.2		Cardinals		QTC	QT_QTC	eka, do, tiina,	
10.3		Ordinals		QTO	QT_QTO	pahalaa, duusaraa	
11	<b>Residuals</b>			<b>RD</b>	<b>RD</b>		
11.1		Foreign word		RDF	RD_RDF		A word written in script other than the script of the original text
11.2		Symbol		SYM	RD_SYM	\$, &, *, (, )	For symbols such as \$, & etc
11.3		Punctuation		PUNC	RD_PUNC	., : ;	Only for punctuations
11.4		Unknown		UNK	RD_UNK		
11.5		Echowords		ECH	RD_ECH	(Paanii-) vaanii, (khaanaa-) vaanaa	

**\*\* The annotation is to be done using the lowest level tag of the type hierarchy. Once the lower level tag is selected, the higher level tags should be stored automatically.**

### POS for Punjabi

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	Noun			N	N	ਘਰ ਕਿਤਾਬ ਕਹਾਣੀ ਸੜਕ	<b>Gara kiwAba kahANI sadZaka</b>
1.1		Common		NN	N_NN	ਘਰ ਕਿਤਾਬ ਕਹਾਣੀ ਸੜਕ	<b>Gara kiwAba kahANI sadZaka</b>
1.2		Proper		NNP	N_NNP	ਹਰਵਿੰਦਰ	<b>haraviMxara xiYII</b>

						ਦਿੱਲੀ ਤਾਜਮਹਿਲ	wAjamahila
1.4		Nloc		NST	N_NST	ਉੱਤੇ ਥੱਲੇ ਅੱਗੇ ਪਿੱਛੇ	uYwe WaYle aYge piYCe
2	Pronoun			PR	PR	ਮੈਂ ਤੂੰ ਉਹ ਇਹ ਜੇ	mEz wUM uha iha jo
2.1		Personal		PRP	PR_PRP	ਮੈਂ ਤੂੰ ਉਹ	mEz wuM uha
2.2		Reflexive		PRF	PR_PRF	ਆਪਣਾ ਆਪ ਖੁਦ	ApaNA Apa Kuxa
2.3		Relative		PRL	PR_PRL	ਜੋ, ਜਿਸ ਜਿਹੜਾ, ਜਦੋਂ,	jo jisa jihadZA jaxoz
2.4		Reciprocal		PRC	PR_PRC	ਆਪਸ	Apasa
2.5		Wh-word		PRQ	PR_PRQ	ਕੌਣ ਕਦੋਂ ਕਿੱਥੇ	kONa kaxoz kiYWe
2.6		Indefinite		PRI	PR_PRI	ਕੋਈ, ਕਿਸ	koI kisa
3	Demonstrative			DM	DM	ਉਹ ਜੇ ਇਹ	uha jo iha
3.1		Deictic		DMD	DM_DMD	ਇਹ ਉਹ	iha uha
3.2		Relative		DMR	DM_DMR	ਜੋ ਜਿਸ	jo jisa
3.3		Wh-word		DMQ	DM_DMQ	ਕੌਣ	kONa
3.4		indefinite		DMI	DM_DMI	ਕੋਈ ਕਿਸ	koI kisa
4	Verb			V	V	ਆਇਆ ਜਾ ਕਰਦਾ ਮਾਰਾਂਗਾ ਰਹਿੰਦਾ	AiA jA karaxA mArAzgA rahiMxA
4.1		Main		VM	V_VM	ਆਇਆ ਜਾ ਕਰਦਾ ਮਾਰਾਂਗਾ ਰਹਿੰਦਾ	AiA jA karaxA mArAzgA rahiMxA
4.1.2			Non-finite	VNF	V_VM_VNF	ਜਾਂਦਿਆਂ ਆਉਂਦਿਆਂ	jAzxiAz AuzxiAz karaxiAz

						ਕਰਦਿਆਂ ਖਾਕੇ ਜਾਕੇ	KAKE jAKE
4.1.3			Infinitive	VINF	V_VM_VINF	ਗਿਆਂ ਆਇਆਂ ਕਰਿਆਂ	giAz, AiAz, kariAz
4.1.4			Gerund	VNG	V_VM_VNG	ਜਾਣੇ ਖਾਣੇ ਪੀਣੇ ਮਰਨੇ	jANoz KANoz pINoz maranoz
4.2		Auxiliary		VAUX	V_VAUX	ਹੈ ਸੀ ਸਕਿਆ ਹੋਇਆ	hE sI sakiA hoiA
5	Adjective			JJ		ਸੋਹਣਾ ਚੰਗਾ ਮਾਡਾ ਕਾਲਾ	sohaNA caMgA mAdZA kAA
6	Adverb			RB		ਹੌਲੀ ਕਾਹਲੀ	hOI kAhall
7	Postposition			PSP		ਨੇ ਨੂੰ ਤੋਂ ਨਾਲ	ne nUM woz nAla
8	Conjunction			CC	CC	ਅਤੇ ਕਿਉਂਕਿ ਅਗਰ ਕਿ ਸਗੋਂ	awe kiuzki agara ki sagoz
8.1		Co-ordinator		CCD	CC_CCD	ਅਤੇ ਜਾਂ	awe jAz
8.2		Subordinator		CCS	CC_CCS	ਕਿਉਂਕਿ ਕਿ ਜੇ ਤਾਂ	kiuzki ki jo wAz
9	Particles			RP	RP	ਵੀ ਤਾਂ ਹੀ	vI wAz hI
9.1		Default		RPD	RP_RPD	ਵੀ ਤਾਂ ਹੀ	vI wAz hI
9.2		Classifier		CL	RP_CL		Not required
9.3		Interjection		INJ	RP_INJ	ਉਏ ਅਡਿਆ ਨੀ ਜਨਾਬ	ue adZiA nI janAba
9.4		Intensifier		INTF	RP_INTF	ਬਹੁਤ ਬਡਾ	<b>bahuwa</b> <b>badZA</b>
9.5		Negation		NEG	RP_NEG	ਨਹੀਂ ਨਾ ਬਿਨਾਂ ਵਗੈਰ	nahlz nA binAz vagEra
10	Quantifiers			QT	QT	ਥੋੜਾ ਬਹੁਤਾ ਕਾਫੀ ਕੁਝ ਇੱਕ	WodZA bahuwA kAPI kuJa iYka

						ਪਹਿਲਾ	pahilA
10.1		General		QTF	QT_QTF	ਬੋਝਾ ਬਹੁਤਾ ਕਾਫੀ ਕੁਝ	WodZA bahuwA kAPI kuJa
10.2		Cardinals		QTC	QT_QTC	ਇੱਕ ਦੇ ਤਿੰਨ	iYka xo wiMna
10.3		Ordinals		QTO	QT_QTO	ਪਹਿਲਾ ਦੂਜਾ	pahilA xUjA
11	<b>Residuals</b>			<b>RD</b>	<b>RD</b>		
11.1		Foreign word		RDF	RD_RDF		A word written in script other than the script of the original text
11.2		Symbol		SYM	RD_SYM	\$, &, *, (, )	For symbols such as \$, & etc
11.3		Punctuation		PUNC	RD_PUNC	., : ;	Only for punctuations
11.4		Unknown		UNK	RD_UNK		
11.5		Echowords		ECH	RD_ECH	(ਪਾਣੀ-) ਧਾਣੀ (ਚਾਹ-) ਚੁਹ	(pANI-) XANI (cAha-) cUha

\*\* The annotation is to be done using the lowest level tag of the type hierarchy. Once the lower level tag is selected, the higher level tags should be stored automatically.

### Tagset for Dravidian Languages (Telugu, Kannada, Malayalam and Tamil)

Sl. No	Category			Label	Annotation Convention**	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)			
<b>1</b>	<b>Noun</b>			<b>N</b>	<b>N</b>	
1.1		Common		NN	N_NN	
1.2		Proper		NNP	N_NNP	
1.3		Nloc		NST	N_NST	
<b>2</b>	<b>Pronoun</b>			<b>PR</b>	<b>PR</b>	
2.1		Personal		PRP	PR_PRP	
2.2		Reflexive		PRF	PR_PRF	

2.3		Relative		PRL	PR_PRL	
2.4		Reciprocal		PRC	PR_PRC	
2.5		Wh-word		PRQ	PR_PRQ	
<b>3</b>	<b>Demonstrative</b>			<b>DM</b>	<b>DM</b>	
3.1		Deictic		DMD	DM_DMD	
3.2		Relative		DMR	DM_DMR	
3.3		Wh-word		DMQ	DM_DMQ	
<b>4</b>	<b>Verb</b>			<b>V</b>	<b>V</b>	
4.1		Main		VM	V_VM	
4.1.1			Finite	VF	V_VM_VF	
4.1.2			Non-finite	VNF	V_VM_VNF	
4.1.3			Infinitive	VINF	V_VM_VINF	
4.1.4			Gerund	VNG	V_VM_VNG	
4.2		Verbal Noun	Verbal noun	NNV	N_NNV	Verbal Noun
4.3		Auxiliary		VAUX	V_VAUX	
4.3.1		Non-finite		VNF	V_VM_VNF	
4.3.2		Infinitive		VINF	V_VM_VNF	
<b>5</b>	<b>Adjective</b>			<b>JJ</b>		
<b>6</b>	<b>Adverb</b>			<b>RB</b>		Only manner adverbs
<b>7</b>	<b>Postposition</b>			<b>PSP</b>		
<b>8</b>	<b>Conjunction</b>			<b>CC</b>	<b>CC</b>	
8.1		Co-ordinator		CCD	CC_CCD	
8.2		Subordinator		CCS	CC_CCS	
8.2.1			Quotative	UT	CC_CCS_UT	
<b>9</b>	<b>Particles</b>			<b>RP</b>	<b>RP</b>	
9.1		Default		RPD	RP_RPD	
9.2		Classifier		CL	RP_CL	
9.3		Interjection		INJ	RP_INJ	
9.4		Intensifier		INTF	RP_INTF	

9.5		Negation		NEG	RP_NEG	
<b>10</b>	<b>Quantifiers</b>			<b>QT</b>	<b>QT</b>	
10.1		General		QTF	QT_QTF	
10.2		Cardinals		QTC	QT_QTC	
10.3		Ordinals		QTO	QT_QTO	
<b>11</b>	<b>Residuals</b>			<b>RD</b>	<b>RD</b>	
11.1		Foreign word		RDF	RD_RDF	A word written in script other than the script of the original text
11.2		Symbol		SYM	RD_SYM	For symbols such as \$, & etc
11.3		Punctuation		PUNC	RD_PUNC	Only for punctuations
11.4		Unknown		UNK	RD_UNK	
11.5		Echowords		ECH	RD_ECH	

\*\* The annotation is to be done using the lowest level tag of the type hierarchy. Once the lower level tag is selected, the higher level tags should be stored automatically.

## POS for Tamil

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
<b>1</b>	<b>Noun</b>			<b>N</b>	<b>N</b>	<i>paiyan,</i> <i>raajaa,</i> <i>puttakam</i>	
1.1		Common		NN	N__NN	<i>puttakam,</i> <i>kaNNaaTi,</i> <i>paTam</i>	
1.2		Proper		NNP	N__NNP	<i>moohan,</i> <i>ravi,</i> <i>maalati</i>	
1.3		Nloc		NST	N__NST	<i>meel,</i> <i>kiiz,</i> <i>mun,</i> <i>pin</i>	

<b>2</b>	<b>Pronoun</b>			<b>PR</b>	<b>PR</b>	<i>itu,atu,avan</i>	
2.1		Personal		PRP	PR__PRP	<i>naan, nii, avaL, avarkaL</i>	
2.2		Reflexive		PRF	PR__PRF	<i>taan,</i>	
2.3		Relative		PRL	PR__PRL	<i>yaar, etu, eppootu, enkee</i>	
2.4		Reciprocal		PRC	PR__PRC	<i>oruvarukoruv ar, avanavan, parasparam</i>	
2.5		Wh-word		PRQ	PR__PRQ	<i>yaarum, yaaraavatu, yaaroo, etuvum</i>	
<b>3</b>	<b>Demonstrative</b>			<b>DM</b>	<b>DM</b>	<i>a-, i-, e-</i>	
3.1		Deictic		DMD	DM__DMD	<i>anta, inta, enta</i>	
3.2		Relative		DMR	DM__DMR	<i>enta</i>	
3.3		Wh-word		DMQ	DM__DMQ	<i>enta, yaar eetaavatu, yaaraavatu</i>	
<b>4</b>	<b>Verb</b>			<b>V</b>	<b>V</b>	<i>vizu, poo, tuunku, aaku</i>	
4.1		Main		VM	V__VM	<i>vizu, poo, tuunku, ciri</i>	
4.1.1			Finite	VF	V__VM__VF	<i>vizuntaan, pooneen, cirittaaL</i>	
4.1.2			Non-finite	VNF	V__VM__VNF	<i>vizunta, poonaal</i>	
4.1.3			Infinitive	VINF	V__VM__VINF	<i>viza, pooka, cirikka</i>	
4.1.4			Gerund	VNG	V__VM__VNG	<i>vizutal, cirittal, tuunkutal</i>	
4.2		Verbal		VN	V_VN	<i>paTippu, naTai, naTattai, ceykai</i>	
4.3		Auxiliary		VAUX	V__VAUX	<i>aakum, veeNTum, muTiyum</i>	
<b>5</b>	<b>Adjective</b>			<b>JJ</b>		<i>iniya, periya, azakaana</i>	
<b>6</b>	<b>Adverb</b>			<b>RB</b>		<i>veekamaaka, viraivaaka</i>	

7	<b>Postposition</b>			<b>PSP</b>		<i>paRRi, kuRittu, viTa</i>	
8	<b>Conjunction</b>			<b>CC</b>	<b>CC</b>	<i>maRRum, eenenRaal, aanaal</i>	
8.1		Co-ordinator		CCD	CC__CCD	<i>-um(raamanu m) maRRum, aanaal, allatu</i>	-um is a co-ordinator which can be added to noun and verb.
8.2		Subordinator		CCS	CC__CCS	<i>enRu, ena, enpatu, enRaal</i>	
8.2.1			Quotative	UT	CC__CCS__UT	<i>enRu, ena</i>	
9	<b>Particles</b>			<b>RP</b>	<b>RP</b>	<i>maTTUm, kuuTa</i>	
9.1		Default		RPD	RP__RPD	<i>maTTUm, kuuTa</i>	
9.2		Classifier		<del>CL</del>	<del>RP__CL</del>		Not required
9.3		Interjection		INJ	RP__INJ	<i>ayyoo, teey, aamaam</i>	
9.4		Intensifier		INTF	RP__INTF	<i>ati, veku, mika</i>	
9.5		Negation		NEG	RP__NEG	<i>illai</i>	
10	<b>Quantifiers</b>			<b>QT</b>	<b>QT</b>	<i>koncam, niRaiya, oru, mutal</i>	
10.1		General		QTF	QT__QTF	<i>koncam, niRaiya</i>	
10.2		Cardinals		QTC	QT__QTC	<i>onRu, iraNTu</i>	
10.3		Ordinals		QTO	QT__QTO	<i>mutal, iraNTaam</i>	
11	<b>Residuals</b>			<b>RD</b>	<b>RD</b>		
11.1		Foreign word		RDF	RD__RDF		A word written in script other than the script of the original text
11.2		Symbol		SYM	RD__SYM	<i>\$, &amp;, *, (, ), ruu.</i>	For symbols such as \$, & etc
11.3		Punctuation		PUNC	RD__PUNC	<i>., : ;</i>	Only for punctuations
11.4		Unknown		UNK	RD__UNK		
11.5		Echowords		ECH	RD__ECH	<i>vaNTi kiNTi, paal kiil</i>	



## POS for Malayalam

Sl. No	Category			Label	Annotation Convention**	Examples	Examples in Malayalam
	Top level	Subtype (level 1)	Subtype (level 2)				
<b>1</b>	<b>Noun</b>			<b>N</b>	<b>N</b>	<i>avan</i> <i>mOhan</i> <i>vItu</i>	
1.1		Common		NN	N__NN	<i>vItu</i> , <i>vellam</i> , <i>pattam</i>	
1.2		Proper		NNP	N__NNP	<i>mOhan</i> , <i>ravi</i> , <i>sIta</i>	മോഹൻ രവി സീത
1.3		Nloc		NST	N__NST	<i>mEle</i> , <i>tAze</i> , <i>munpil</i> , <i>pinnil</i>	മേലെ താഴെ മുൻപിൽ പിന്നിൽ
<b>2</b>	<b>Pronoun</b>			<b>PR</b>	<b>PR</b>	<i>avan</i> , <i>aval</i> , <i>at</i> <i>u</i> , <i>itu</i> ,	അവൻ അവളു്, അത്, ഇത്
2.1		Personal		PRP	PR__PRP	<i>naan</i> , <i>nii</i> , <i>avaL</i> , <i>avar</i>	ഞാൻ,നീ, അവളു്, അവർ
2.2		Reflexive		PRF	PR__PRF	<i>tanne-taan</i>	തന്നെ ത്താൻ
2.3		Relative		PRL	PR__PRL	<i>aaro</i> ,	ആരോ
2.4		Reciprocal		PRC	PR__PRC	<i>tammiltammi</i> <i>l</i> , <i>parasparam</i>	തമ്മിൽ തമ്മിൽ

							പരസ്പരം
2.5		Wh-word		PRQ	PR__PRQ	<i>aaru, evan</i>	ആര്, എവൻ,
<b>3</b>	<b>Demonstrative</b>			<b>DM</b>	<b>DM</b>	<i>aa-, ii-,</i>	ആ, ഈ
3.1		Deictic		DMD	DM__DMD	<i>atu, itu</i>	അത്, ഇത്,
3.2		Relative		DMR	DM__DMR	<i>eetu</i>	ഏത്
3.3		Wh-word		DMQ	DM__DMQ	<i>eetu, ennane</i>	ഏത്, എങ്ങനെ
<b>4</b>	<b>Verb</b>			<b>V</b>	<b>V</b>	<i>pO, kazhi, Annu,ciri</i>	പോ, കഴി ആൺ(Copula), ചിരി
4.1		Main		VM	V__VM	<i>pO, kazhi, cirri,Annu(copula)</i>	പോ, കഴി, ആൺ, (copula), ചിരി
4.1.1			Finite	VF	V__VM__VF	<i>pOyi, cirikkum, kazhikkunnu Akunnu(copula)</i>	പോയി, ചിരി ക്കും, കഴിക്കുന്നു, ആകുന്നു (copula)
4.1.2			Non-finite	VNF	V__VM__VNF	<i>pOya, ciricca, kazhicca</i>	പോയ, ചിരിച്ച, കഴിച്ച,
4.1.3			Infinitive	VINF	V__VM__VINF	<i>pOkku, cirikkukayAl kazhikkee, varAn/varuv An</i>	പോക്ക്, ചിരിക്കുകയാൽ, കയാൽ,

							കഴിക്ക്, വരാൻ/ വരുവാ ൻ
4.2		Verbal		VN	V__VN	<i>paTittam, naTattam, naTanam</i>	പഠിത്തം നടത്തം നടനം
4.3		Auxiliary		VAUX	V_VAUX	<i>kolluka, talluka, kAnuka, nOkkuka</i>	കൊല്ലുക , തല്ലുക, കാണുക, നോക്കു ക
5	<b>Adjective</b>			<b>JJ</b>		<i>valiya, ceRiya azakulla</i>	വലിയ, ചെറിയ, അഴകു ള്ള
6	<b>Adverb</b>			<b>RB</b>		<i>veegam, ativeegam, kUtutal.</i>	വേഗം, അതിവേ ഗം, കൂടുതൽ
7	<b>Postposition</b>			<b>PSP</b>		<i>paRRi, kUte,</i>	പറ്റി, കൂടെ
8	<b>Conjunction</b>			<b>CC</b>	<b>CC</b>	<i>pakshe, , enniTTum, ennAl,ennalu m, enkilum</i>	പക്ഷെ, എന്നിട്ടും , എന്നാൽ, എന്നാ

							ലും എങ്കിലും
8.1		Co-ordinator		CCD	CC__CCD	<i>-um (rAmanum) pakshe,</i>	ഉം (രാമനും) പക്ഷെ,
8.2		Subordinator		CCS	CC__CCS	<i>ennu, enna, ennAl</i>	എന്ന്, എന്ന, എന്നാൽ
8.2.1			Quotative	UT	CC__CCS__UT	<i>ennu, enna</i>	എന്ന്, എന്ന,
<b>9</b>	<b>Particles</b>			<b>RP</b>	<b>RP</b>	<i>kute,mAtram</i>	കൂടെ, മാത്രം
9.1		Default		RPD	RP__RPD	<i>mAtram</i>	മാത്രം
9.2		Classifier		C	RP__CL	<i>peer</i>	പേർ
9.3		Interjection		INJ	RP__INJ	<i>ayyoo,</i>	അയ്യോ,
9.4		Intensifier		INTF	RP__INTF	<i>pala, valare,</i>	പല, വളരെ
9.5		Negation		NEG	RP__NEG	<i>illa, alla</i>	ഇല്ല, അല്ല
<b>10</b>	<b>Quantifiers</b>			<b>QT</b>	<b>QT</b>	<i>kuracchu, niraccu, oru, dharalam</i>	കുറച്ച്, നിറച്ച്, ഒരു, യാരാളും
10.1		General		QTF	QT__QTF	<i>kuraccu, niraccu, dharalam</i>	കുറച്ച്, നിറച്ച്, യാരാളും

10.2		Cardinals		QTC	QT__QTC	<i>onnu,rantu</i>	ഒന്ന്, രണ്ട്
10.3		Ordinals		QTO	QT__QTO	<i>onnAm,ranta m</i>	ഒന്നാം, രണ്ടാം
11	<b>Residuals</b>			<b>RD</b>	<b>RD</b>		
11.1		Foreign word		RDF	RD__RDF		
11.2		Symbol		SYM	RD__SYM	\$. &, *, (, ), ruu.	\$. &, *, (, ), രൂ
11.3		Punctuation		PUNC	RD__PUNC	., : ;	., : ;
11.4		Unknown		UNK	RD__UNK		
11.5		Echowords		ECH	RD__ECH		

## POS for Bangla

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	<b>Noun</b>			<b>N</b>	<b>N</b>		
1.1		Common		NN	N_NN	kalama, cashmaa	
1.2		Proper		NNP	N_NNP	Mohan, ravi, rashmi	
1.4		Nloc		NST	N_NST	upare, niche, bhitara	
2	<b>Pronoun</b>			<b>PR</b>	<b>PR</b>		
2.1		Personal		PRP	PR_PRP	se, tumi, AmAra	
2.2		Reflexive		PRF	PR_PRF	nijera,	
2.3		Relative		PRL	PR_PRL	ye, yakhana, yena, yAra	
2.4		Reciprocal		PRC	PR_PRC	paraspara,	
2.5		Wh-word		PRQ	PR_PRQ	ke, kakhana,	

						kena, kAra,	
2.6		Indefinite		PRI	PR_PRI	keu	
<b>3</b>	<b>Demonstrative</b>			<b>DM</b>	<b>DM</b>	Vaha, jo, yaha,	
3.1		Deictic		DMD	DM_DMD	sei, oi, o, se	
3.2		Relative		DMR	DM_DMR	ye, yei	
3.3		Wh-word		DMQ	DM_DMQ	kono,	
3.4		Indefinite		DMI	DM_DMI	keu	
<b>4</b>	<b>Verb</b>			<b>V</b>	<b>V</b>		
4.1		Main		VM	V_VM		
4.1.1			Finite	VF	V_VM_VF	karachhilAm a, yAba, khAYa	
4.1.2			Non-finite	VNF	V_VM_VNF	kare, kheYe, karale, khete,	
4.1.3			Infinitive	VINF	V_VM_VINF	karate, khete, yete	
4.1.4			Gerund	VNG	V_VM_VNG	yAoYa, AsA, khelA, karA	
4.2		Auxiliary		VAUX	V_VAUX	chhila, habe, chAi	
5	Adjective			JJ		sundara, bhAla, lAla,	
6	Adverb			RB		tA.DAtA.Di , Aste, haThAt	
7	Postposition			PSP		theke, abadhI, madhye, diYe	
8	Conjunction			CC	CC		
8.1		Co-ordinator		CCD	CC_CCD	Ara, eba.n, athabA, kimbA	
8.2		Subordinator		CCS	CC_CCS	ye, kintu, noile,	

						tAhale	
8.2.1			Quotative	UT	CC_CCS_UT	----	Not required
<b>9</b>	<b>Particles</b>			<b>RP</b>	<b>RP</b>		
9.1		Default		RPD	RP_RPD	to, ye,	
9.2		Classifier		CL	RP_CL	jana, khAnA	
9.3		Interjection		INJ	RP_INJ	Are, ei, hAya	
9.4		Intensifier		INTF	RP_INTF	bhiShaNa, khuba, sA~NghAtik a	
9.5		Negation		NEG	RP_NEG	nA, naYa, chhA.DA	
<b>10</b>	<b>Quantifiers</b>			<b>QT</b>	<b>QT</b>		
10.1		General		QTF	QT_QTF	kichhu, alpa, aneka	
10.2		Cardinals		QTC	QT_QTC	eka, dui, tina	
10.3		Ordinals		QTO	QT_QTO	prathama, paYalA, dvitIYa	
<b>11</b>	<b>Residuals</b>			<b>RD</b>	<b>RD</b>		
11.1		Foreign word		RDF	RD_RDF		A word written in script other than the script of the original text
11.2		Symbol		SYM	RD_SYM	\$, &, *, (, )	For symbols such as \$, & etc
11.3		Punctuation		PUNC	RD_PUNC	., : ;	Only for punctuations
11.4		Unknown		UNK	RD_UNK		
11.5		Echowords		ECH	RD_ECH	jala Tala, khAbAra dAbAra	

\*\* The annotation is to be done using the lowest level tag of the type hierarchy. Once the lower level tag is selected, the higher level tags should be stored automatically.

**POS for Marathi**

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	Noun			N	N	मुलगा (mulagaa-boy), राजा (raajaa-king), पुस्तक (pustaka-book)	
1.1		Common		NN	N_NN	पुस्तक (pustaka-book), लेखणी (lekhaNi-pen), चष्मा (chashmaa-goggles )	
1.2		Proper		NNP	N_NNP	मोहन (Mohan), रवी (Ravi), रश्मी (Rashmi)	
1.3		Verbal		NNV	N_NNV	NA	Not Required
1.4		Nloc		NST	N_NST	वर(var- up), खाली(khaalee-down), पुढे(pudhe-ahead), मागे(maage-back)	Where it is separate it is NST
2	Pronoun			PR	PR	येथे(yethe-here), तेथे (tethe-there),	



						जो(jo-who), तो(to-he)	
2.1		Personal		PRP	PR_PRP	तो(to-he), मी(mee-I), तू(tu-you), ते(te-they), तुम्ही(tumhi-you)	
2.2		Reflexive		PRF	PR_PRF	स्वतः(swatha-myself), आपण(aapana-ourselves)	
2.3		Relative		PRL	PR_PRL	जो(jo-who), ज्याने(jyaane-who), जेव्हा(jevhaa-while), जिथे(jeethe-where)	
2.4		Reciprocal		PRC	PR_PRC	परस्पर(Paraspara-reciprocally ), एकमेक(ekmek - mutually)	
2.5		Wh-word		PRQ	PR_PRQ	कोण(kona-who), केव्हा(kevha-when), कुठे(kuthe-where)	
2.6		Indefinite				कोणी(kona	
3	Demonstrative			DM	DM	तो(to-he), हा(haa-this), जो(jo-who)	

3.1		Deictic		DMD	DM_DMD	इथे(ithe-here), तिथे(tithe- there)	
3.2		Relative		DMR	DM_DMR	जो(jo-who) ज्याने(jyane- who)	
3.3		Wh-word		DMQ	DM_DMQ	कोणता(konta- which), कोणी(kona- who),	
4	Verb			V	V	(padalaa-fell down), गेलाला(gelaa- went), झोपला(jhopala a-slept), आहे(aahe-is),	
4.1		Main		VM	V_VM	पडला (padalaa-fell down), गेलाला(gelaa- went), झोपला(jhopala a-slept), आहे(aahe-is),	
4.1. 1			Finite	VF	V_VM_VF	-	This subtype WILL NOT be used for Hindi as Hindi does not have enough information at the word level.
4.1. 2			Non-finite	VNF	V_VM_VNF	-	--do--
4.1. 3			Infinitive	VINF	V_VM_VINF	-	--do--
4.1.			Gerund	VNG	V_VM_VNG		--do--

4							
4.2		Auxiliary		VAUX	V_VAUX	आहे (is), लागला (started),	
5	Adjective			JJ		सुंदर(sundara- beautiful), चांगला(chaang alaa-good), मोठा(moThaa- big)	
6	Adverb			RB		लवकर(lavakar - fast ), हळूहळू(haLuu aLuu-slowly)	
7	Postposition			PSP		Not in Marathi	
8	Conjunction			CC	CC	आणि(aaNi- and), कारण(kaaraN- because)	
8.1		Co-ordinator		CCD	CC_CCD	आणि(aaNi- and), पण(paNa- but), परंतु (parantu-but)	
8.2		Subordinator		CCS	CC_CCS	कारण की (kaaraN- because of), का की(kaaraN kii-because of), जर- तर(jara-tara- if-then)	
8.2.1			Quotative	UT	CC_CCS_UT	असा, म्हणून	
9	Particles			RP	RP	तर(tara),	
9.1		Default		RPD	RP_RPD	तर(tara) (then)	
9.2		Classifier		CL	RP_CL		Not required
9.3		Interjection		INJ	RP_INJ	अरेरे!(arere),	

						ओहो!(oho-oh!)	
9.4		Intensifier		INTF	RP_INTF	खूप(khoop-lot, very ), बराच(baraach-too much), अतिशय(atisha ya- too much, very)	
9.5		Negation		NEG	RP_NEG	नको(nako-not), न(na-Na)	
10	Quantifiers			QT	QT	थोडे(thode-few), जास्त(jaasta-lot), काही(kaahi-few), एक(eka-one), पहिला(pahilaa-first),	
10.1		General		QTF	QT_QTF	थोडे thoDe-few), जास्त(jaasta-lot), काही(kaahi-few)	
10.2		Cardinals		QTC	QT_QTC	एक(eka-one), दोन(dona-two)	
10.3		Ordinals		QTO	QT_QTO	पहिला(pahilaa-first), दुसरा(dusaraa-second)	
11	Residuals			RD	RD		
11.1		Foreign word		RDF	RD_RDF		A word written in script other than the script of the original text

11.2		Symbol		SYM	RD_SYM	\$, &, *, (, )	For symbols such as \$, & etc
11.3		Punctuation		PUNC	RD_PUNC	., : ;	Only for punctuations
11.4		Unknown		UNK	RD_UNK		
11.5		Echowords		ECH	RD_ECH	જેવણબિવણ(jev anbivaNa- meal/dinner), ડોકેબિકે(Doke bike- head) (Paanii-) vaanii, (khaanaa-) vaanaa	

\*\* The annotation is to be done using the lowest level tag of the type hierarchy. Once the lower level tag is selected, the higher level tags should be stored automatically.

### POS for Gujarati

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	Noun			N	N		
1.1		Common		NN	N_NN	kalam, chashmA 'pen', 'spectacles'	
1.2		Proper		NNP	N_NNP	mohan, ravi 'Mohan', 'Ravi'	
1.3		Nloc		NST	N_NST	upar, nilche, ahIM 'up', 'down', 'in front'	
2	Pronoun			PR	PR		
2.1		Personal		PRP	PR_PRP	huM, tuM, te 'me', 'you',	

						'he/she'	
2.2		Reflexive		PRF	PR_PRF	pote, jAte,svayam  'herself/him self'	
2.3		Relative		PRL	PR_PRL	je, te, jyAM  'who', 'where'	
2.4		Reciprocal		PRC	PR_PRC	aras-paras, paraspar  'mutually', 'e ach other'	
2.5		Wh-word		PRQ	PR_PRQ	koN, kyAre, kyAM  'who', 'when', 'where'	
2.6		Indefinite				kol, kaIMK, kashuM  'someone', 'something'	
<b>3</b>	<b>Demonstrative</b>			<b>DM</b>	<b>DM</b>		
3.1		Deictic		DMD	DM_DMD	A  'this'	
3.2		Relative		DMR	DM_DMR	je, jeNe  'which/who', 'whom'	
3.3		Wh-word		DMQ	DM_DMQ	koN,shuM,ke m  'who', 'what', 'why'	
3.4		Indefinite				kol, kaIMK, kashuM  'someone', 'something'	
<b>4</b>	<b>Verb</b>			<b>V</b>	<b>V</b>		
4.1		Main		VM	V_VM	khAshe,khAd hu  'will eat',	

						'ate'	
4.2		Auxiliary		VAUX	V_VAUX	chhe,hatuM,k aryuM  'is', 'was', 'did'	
5	Adjective			JJ			
6	Adverb			RB			
7	Postposition			PSP			
8	Conjunction			CC	CC		
8.1		Co-ordinator		CCD	CC_CCD	ane,ke  'and', 'or'	
8.2		Subordinator		CCS	CC_CCS	tethl, evuM, kAraNke  'so', 'like that', 'because'	
9	Particles			RP	RP		
9.1		Default		RPD	RP_RPD	paNa,ja,tO  'but', emph, topic	
9.2		Interjection		INJ	RP_INJ	hE !!, arrrE !!,O !!	
9.3		Intensifier		INTF	RP_INTF	bahu,ghaNu M  'very', 'much'	
9.4		Negation		NEG	RP_NEG	nahi,na  'no'	
10	Quantifiers			QT	QT		
10.1		General		QTF	QT_QTF	thoduM,ghaN uM  'little', 'much'	
10.2		Cardinals		QTC	QT_QTC	eka,be traN  'one,two,thr ee'	
10.3		Ordinals		QTO	QT_QTO	paheluM,bljl  'first'(neu),	

						'second' (fem)	
11	Residuals			RD	RD		
11.1		Foreign word		RDF	RD_RDF	tv, perasitemol	
11.2		Symbol		SYM	RD_SYM	\$, *,&	
11.3		Punctuation		PUNC	RD_PUNC	, : ; { } ( )	
11.4		Unknown		UNK	RD_UNK		
11.5		Echowords		ECH	RD_ECH	kAm- bAm,pANi- bANi  'work and the like', water and the like'	

### POS for Konakani

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	Noun			N	N		
1.1		Common		NN	N_NN	पुस्तक ,रुख ,आंबो , माड	
1.2		Proper		NNP	N_NNP	रामायण, बायबल, कुराण, गोंय, कोंकणी, कपिला	
1.3		Nloc		NST	N_NST	भायर, भीतर, वयर, सकयल	
2	Pronoun			PR	PR		
2.1		Personal		PRP	PR_PRP	हांव, तूं, तो, तें, ते, त्यो, तुमचें, आमचें, तांचे	
2.2		Reflexive		PRF	PR_PRF	आपूण, स्वता	



2.3		Relative		PRL	PR_PRL	जातूंत, जो	
2.4		Reciprocal		PRC	PR_PRC	एकामेकाक, आपसांत	
2.5		Wh-word		PRQ	PR_PRQ	कोण, किते, खंयचो	
2.6		Indefinite				कोणूय, कितेंय, खयचेंय	
3	Demonstrative			DM	DM		
3.1		Deictic		DMD	DM_DMD	तो, हो	
3.2		Relative		DMR	DM_DMR	जो	
3.3		Wh-word		DMQ	DM_DMQ	कोणें, कसलें	
3.4		Indefinite				कोणाचेंय, कसलेंय	
4	Verb			V	V		
4.1		Main		VM	V_VM	येवप	
4.1.1			Finite	VF	V_VM_VF	आयलो, आयला, आयिल्लो	
4.1.2			Non-Finite	VNF	V_VM_VNF	येतकच, येवन, आयिल्लयान, येवंक, येवपाक, येवपाचें, येवचें	
4.1.3			Infinitive	VINF	V_VM_VINF	आसूं, व्हरूं, केल्यार	
4.1.4			Gerund	VNG	V_VM_VNG	खावप, वचप, खावपी, जेवपी, समजुपी	
4.2		Auxiliary		VAUX	V_VAUX	NA	
4.2.1			Finite		V_VAUX_VF	केल्लें आसत, आयला आसत	
4.2.2			Non-Finite		V_VAUX_VNF	करता जायत, करता आसतलो, येतीत	
5	Adjective			JJ		सोबीत, सुंदर	
6	Adverb			RB		फाल्यां, सवकास,	

						अशें	
7	Postposition			PSP		खातीर, पासत, बगर, कडेन, लागीं	
8	Conjunction			CC	CC		
8.1		Co-ordinator		CCD	CC_CCD	आनी, वा	
8.2		Subordinator		CCS	CC_CCS	जाल्यार, जर-तर, देखून, म्हणल्यार, पुणून	
8.2.1			Quotative	UT	CC_CCS_UT	अशें, की	
9	Particles			RP	RP		
9.1		Default		RPD	RP_RPD	बी, आदी, इत्यादि	
9.2		Classifier		CL	RP_CL	(पांच) जाण	
9.3		Interjection		INJ	RP_INJ	आरे, चूप	
9.4		Intensifier		INTF	RP_INTF	उपाट, भरपूर	
9.5		Negation		NEG	RP_NEG	ना, न्यह	
10	Quantifiers			QT	QT		
10.1		General		QTF	QT_QTF	थोडे, चड, कांय, खूब	
10.2		Cardinals		QTC	QT_QTC	एक, दोन	
10.3		Ordinals		QTO	QT_QTO	पयलें, दुसरें	
11	Residuals			RD	RD		
11.1		Foreign word		RDF	RD_RDF		
11.2		Symbol		SYM	RD_SYM	&, \$	
11.3		Punctuation		PUNC	RD_PUNC	.,?-/	
11.4		Unknown		UNK	RD_UNK		
11.5		Echowords		ECH	RD_ECH	जोवण-बिचण	

**POS for Maithili**

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	Noun			N	N		
1.1		Common		NN	N_NN	पोथी, कलम, पंडित, खवास	
1.2		Proper		NNP	N_NNP	अरुण, दिनेश, अतुल	
1.3		Nloc		NST	N_NST	आगू, पीछू, ऊपर, नीचा, एखन, आब, बीच, कतहु	
2	Pronoun			PR	PR		
2.1		Personal		PRP	PR_PRP	तौं, हम, ई, ओ, अहाँ	
2.2		Reflexive		PRF	PR_PRF	अपना, अपने, स्वयं, स्वयंमेव	
2.3		Relative		PRL	PR_PRL	जे, जिनका, जिनकर, जकरा	
2.4		Reciprocal		PRC	PR_PRC	एक-दोसरकें, आपस, परस्पर	
2.5		Wh-word		PRQ	PR_PRQ	के, की, कथी ककर	
		Indefinite				केओ, किछु/ किउछ, कोनो	
3	Demonstrative			DM	DM		
3.1		Deictic		DMD	DM_DMD	ओ, ई, ऊ	
3.2		Relative		DMR	DM_DMR	जे, जाहि	
3.3		Wh-word		DMQ	DM_DMQ	के, की, कोन	
		Indefinite				केओ, किछु/	

						किउछ, कोनो	
4	Verb			V	V		
4.1		Main		VM	V_VM	चलबैत, रौपेत, पढइत, खाइत, सुतैत, हँसैत	
4.2		Auxiliary		VAUX	V_VAUX	अछि, छल, होएब, थिक	
5	Adjective			JJ		नीक, मोटका, ललकी,	
6	Adverb			RB		भने, अनायास, क्रमशः, एकाएक, अवश्य, पुनः फेर	
7	Postposition			PSP		सँ, कें, लेल	
8	Conjunction			CC	CC		
8.1		Co-ordinator		CCD	CC_CCD	आओर, परंच, मुदा, वा	
8.2		Subordinator		CCS	CC_CCS	जँ, तँ, कि, यदि	
9	Particles			RP	RP		
9.1		Default		RPD	RP_RPD	भरि, यौ, हौ, रौ	
		Classifier		CL	RP_CL	टा, गोटा, गो	
9.3		Interjection		INJ	RP_INJ	ओह-ओ, अहा, वाह, हा	
9.4		Intensifier		INTF	RP_INTF	बहुत, बेसी, खूब, नितान्त	
9.5		Negation		NEG	RP_NEG	न, नहि, जुनि	
10	Quantifiers			QT	QT		
10.1		General		QTF	QT_QTF	कनेक, बहुत, किछु	
10.2		Cardinals		QTC	QT_QTC	एक, एकटा, दुई, बीसगोट,	

						तीन, चारि	
10.3		Ordinals		QTO	QT_QTO	पहिल, दोसर, तेसर, चारिम	
11	Residuals			RD	RD		
11.1		Foreign word		RDF	RD_RDF		A word written in script other than the script of the original text
11.2		Symbol		SYM	RD_SYM	\$, , *, (, )	For symbols such as \$, & etc
11.3		Punctuation		PUNC	RD_PUNC	., : ;	Only for punctuations
11.4		Unknown		UNK	RD_UNK		
11.5		Echowords		ECH	RD_ECH	जलखे (तलखे), मोंट (सोंट)	

\*\* The annotation is to be done using the lowest level tag of the type hierarchy. Once the lower level tag is selected, the higher level tags should be stored automatically.

## POS for Urdu

Sl. No		Category		Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	<b>Noun</b> (اسم-ism)			<b>N</b>	<b>N</b>	لڑکا (laRkaa), راجا ((raajaa), کتاب (kitaab)	
1.1		Common (نکرہ- (nakeraa		NN	N__NN	کتاب (kitaab), قلم (qalam), چشمہ (cashma)	
1.2		Proper (معرفہ-)		NNP	N__NNP	موہن ((Mohan), رشمی	

		((m'aarefa				،(Rashmi) روی (Ravi)	
1.3		Verbal (حاصل مصدر- haasil-e- (masdar		NNV	N__NNV	،(jalan) جلن ،(calan) چلن ،(bahaa) بہاؤ بناوٹ (banaavat)	May be considered for Urdu-Hindi, too.
1.4		Nloc (ظرف-zarf)		NST	N__NST	،(upar) اوپر ،(niice) نیچے ،(aage) آگے (piiche) پیچھے	
2	<b>Pronoun</b> (ضمیر-zamiir)			<b>PR</b>	<b>PR</b>	،(yih) یہ ،(voh) وہ (jo) جو	
2.1		Personal (ضمیر شخصی- zamiir-e- (shakhsii		PRP	PR__PRP	،(voh) وہ ،(tum) تم (maim) میں	In Urdu, unlike Hindi, voh is used both for singular and plural.
2.2		Reflexive (ضمیر معکوسی- zamiir-e- m'aakoosii )		PRF	PR__PRF	،(apnaa) اپنا ،(khud) خود اپنے آپ (apne aap)	
2.3		Relative (ضمیر موصولہ- zamiir-e- mausoolaa(		PRL	PR__PRL	،(jo) جو ،(jab) جب ،(jis) جس جہاں (jahaM)	
2.4		Reciprocal (ضمیر راجع- zamiir-e- raaje')		PRC	PR__PRC	،(baaham) باہم درمیان (darmiyaan) آپس (aapas)	

2.5		Wh-word (ضمير استفهامیہ- zamiir-e- istafhaamiy aa)		PRQ	PR__PRQ	کون (kaun)، کب (kab)، کہاں (kahaaM)	
3	<b>Demonstrative</b> (ضمیر اشارہ- zamiir-e- ishaaraa)			<b>DM</b>	<b>DM</b>	یہ (yih)، وہ (voh)، ان (inn)، ان (unn)	
3.1		Deictic (اشارے- ishaare)		DMD	DM__DMD	یہ (yih)، وہ (voh)	
3.2		Relative (ضمیر اشارہ- موصولہ- zamiir-e- ishaaraa mausoolaa)		DMR	DM__DMR	جو (jo)، جس (jis)	
3.3		Wh-word (ضمیر اشارہ استفہامیہ- zamiir-e- ishaaraa istafhaamiy aa)		DMQ	DM__DMQ	کون (kaun)، کس (kis)، کتنا (kitnaa)	According to Urdu grammar words like <i>koi</i> , <i>kisi</i> , <i>kuch</i> do not come under Wh-word; they are used for indefinite person. For them, another category (subtype), i.e., tankiir (indefinitive) is used. Under this category

							following words are also placed: <i>chand</i> ,  <i>b‘aaz</i> , <i>fulaan</i> , <i>sab</i> , <i>bahut</i> . Can we have a category/  subtype like indefinitive demonstrative (DMI)?
4	Verb (فعل- <i>f‘el</i> )			V	V	گرا (gira) ، گیا (gayaa) ، سونا (sonaa) ، ہنسنا (haMstaa)	
4.1		Main		VM	V__VM	گرا (gira) ، گیا (gayaa) ، سونا (sonaa) ، ہنسنا (haMstaa)	
4.1.1			Finite (محدود- mahdoo (d	VF	V__VM__VF		This subtype  WILL NOT  be used for  Hindi as  Hindi does  not have  enough  information at  the word  level.



4.1.2			Nonfinite (غیر محدود) ghair -d mahdoo (d	VNF	V__VM__VNF		-- do--
4.1.3			Infinitive (مصدر - (masdar	VINF	V__VM__VINF		-- do--
4.1.4			Gerund (حاصل مصدر - haasil-e- (masdar	VNG	V__VM__VNG		-- do--
4.2		Auxiliary (فعل امدادی - f'el-e- (imdaadi		VAUX	V__VAUX	ہے (hai) رہا (rahaa) ہوا (huaa)	
5	Adjective (صفت-sifat)			JJ		دلکش (dilkash)، سفید (safed)، سیاہ (siyaah)، چوڑا (cauRaa)، اونچا (uuMcaa)	
6	Adverb (متعلق فعل - mut'alliq-e- (f'el			RB		تیز (tez)، جلد (jald)	
7	Postposition (جار موخر - jaar-- (e-moakkhar			PSP		سے (se)، نے (ne)، کو (ko)، میں (meiM)	
8	Conjunction (عطف - (atf			CC	CC	اور (aur)، اگر (agar)، کیوں کہ (kyoMki)	

8.1		Co-ordinator (حرف وصل- (harf-e-vasl		CCD	CC__CCD	اور (aur)، وہ (voh)، یا (yaa)، کہ (ki)، بلکہ (balki)	
8.2		Subordinator (تابع کنندہ- taab'e (kunindaa		CCS	CC__CCS	اگر (agar)، کیوں کہ (kyoMki)، تو (to)	
8.2.1			Quotative (اقتباسی- iqtabaas (ii	UT	CC__CCS__UT		Not required
9	<b>Particles</b> (حالیہ- (haaliyaa			RP	RP	تو (to)، ہی (hii)، بھی (bhii)	
9.1		Default (ڈیفالٹ- Default)		RPD	RP__RPD	تو (to)، ہی (hii)، بھی (bhii)	
9.2		Classifier (درجہ بند- (darja band		CL	RP__CL		Not required
9.3		Interjection (فجائیہ- (fajaa'iyaa		INJ	RP__INJ	اے (e)، او (o)، ارے (are)، جی (jii)، ابا (ahaa)، واہ (vaah)	
9.4		Intensifier		INTF	RP__INTF	بہت (bahut)،	

		(حرف تاکید- harf-e- (taakiid				بے حد (behad)، البتہ (albattoo)، ضرور (zaroor)، خبردار (khabardaar)	
9.5		Negation  (حرف نہی- harf-e- (nahii		NEG	RP__NEG	نہ (na)، نہیں (nahiiM)	
10	<b>Quantifiers</b>  (کمیت نما- <b>kamiiyat</b> (numaa			<b>QT</b>	<b>QT</b>	چند (cand)، متعدد (muta'addad) قلیل (qaliil)، کثیر (kasiir)	
10.1		General  (عام- 'aam)		QTF	QT__QTF	تھوڑا (thoRaa)، بہت (bahut)، کچھ (kuch)	
10.2		Cardinals  (اعداد مطلق- -a'adaad (e-mutlaq		QTC	QT__QTC	ایک (Ek)، دو (do)، تین (tiin)	
10.3		Ordinals  (ترتیبی اعداد- tartiibii (a'adaad		QTO	QT__QTO	اول (avval)، دوم (doam)، پہلا (pahalaa)، دوسرا (duusaraa)	
11	<b>Residuals</b>  (باقی ماندہ- <b>baaqi</b> (maandaa			<b>RD</b>	<b>RD</b>		
11.1		Foreign		RDF	RD__RDF		A word

		word (بدیسی لفظ- bidesii (lafz					written in script other than the script of the original text.
11.2		Symbol (علامت- (alaamat‘		SYM	RD__SYM	\$, &, *, (, ) & , \$	Such symbols are not used in Urdu. They are written ڈالر (dollar), پاونڈ (pound) etc.
11.3		Punctuatio n (اوقاف- (auqaaf		PUNC	RD__PUNC	- ‘, ‘, ‘;	Only for Punctuation s
11.4		Unknown نامعلوم- (m‘aaloom		UNK	RD__UNK		
11.5		Echowords (گونج دار الفاظ- goonjdar (lafz		ECH	RD__ECH	(دل-) ول (dil-) vil) (پیار-) وبار (pyaar-) vyaar) (چائے-) وائے (caa‘e-) vaa‘e)	

\*\* The annotation is to be done using the lowest level tag of the type hierarchy. Once the lower level tag is selected, the higher level tags should be stored automatically.

## 7. XML INTERNATIONALIZATION BEST PRACTICES

To make the common POS Schema for Indian Languages completely interoperable extensible and web enabled, W3C XML Internationalization best practices guidelines and ISO Metadata standard are adopted in the above framework.

### 7.1 WHAT IS INTERNATIONALIZATION TAG SET (ITS)

ITS is a technology to easily create XML which is internationalized and can be localized effectively.

ITS for Schema developers:

User will find proposals for attribute and element names to be included in their new schema (also called "host vocabulary"). It leads to easier recognition of the concepts represented by both schema users and processors. [For more details <http://www.w3.org/TR/2007/REC-its-20070403/>]

Main Attributes:

Defining mark-up for natural language labelling (*xml:lang*- defined for the root element of your document, and for any element where a change of language may occur), Defining mark-up to specify text direction (*its:dir* - defined for the root element of your document, and for any element that has text content), Indicating which elements and attributes should be translated (*its:translateRule*- elements to indicate which elements have non-translatable content), Providing information related to text segmentation (*its:withinTextRule*- elements to indicate which elements should be treated as either part of their parents, or as a nested but independent run of text), Defining mark-up for unique identifiers (*xml:id*- elements with translatable content can be associated with a unique identifier), Defining mark-up for notes to localizers (*its:locNote*- allows content authors to provide localization-related notes as attribute values, or to point to the location of the relevant note text using). [For more details <http://www.w3.org/TR/xml-i18n-bp/>]

## 8. XML SCHEMA

XML Schemas express shared vocabularies and allow machines to carry out rules made by people and to define a class of XML documents, and so the term "instance document" is often used to describe an XML document that conforms to a particular schema. It provides a means for defining the structure, content and semantics of XML documents. [For more details <http://www.w3.org/TR/1999/NOTE-xml-schema-req-19990215/>]

## 9. METADATA ON POS

Metadata:

Metadata describes how and when and by whom a particular set of data was collected, and how the data is formatted. It is essential for understanding information stored in data warehouses and has become increasingly important in XML-based Web applications.

XML Metadata:

Metadata built into the document. Every element has a tag to tell you where the data is stored in the document. Descriptive tags give structure to the document and tell you what the data means (sort of).

“Sort of” because it only tells the tag name, so this only has meaning to someone who already understands what the element or attribute means.

### METADATA AS PER ISO 12620:1999

Metadata ()

{

<?xml version="1.0"?>

<datasm-categorySelection xmlns="http://www.isocat.org/ns/dcif" dcif-version="1.0">

<globalInformation>.....</globalInformation>

<languageSection>

    <language>en</language>

    <identifier>..... </identifier>

    <version>1.0.0</version>

    <registrationStatus>standard</registrationStatus> // registered as a standard //

    <origin>ISO 12620:1999

    <author>.....</author>

    <domain>.....</domain>

    </origin>

    <creation>

        <creationDate>1999-01-01</creationDate>

    </creation>

<descriptionSection>

```

<definitionClass>
  <definition xml:lang="en">.....</definition>
  <source>ISO 12620:1999</source>
</definitionClass>

```

```

</descriptionSection>

```

```

</languageSection>

```

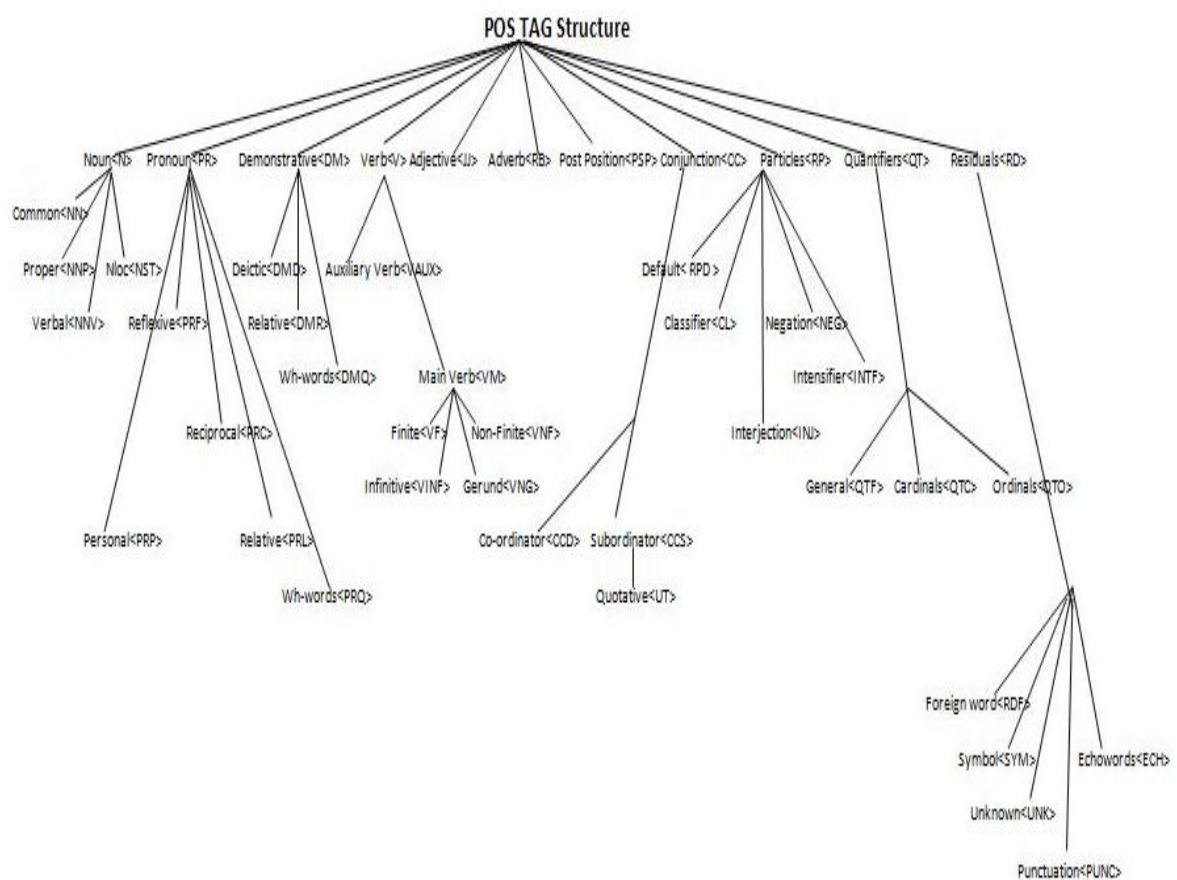
```

}

```

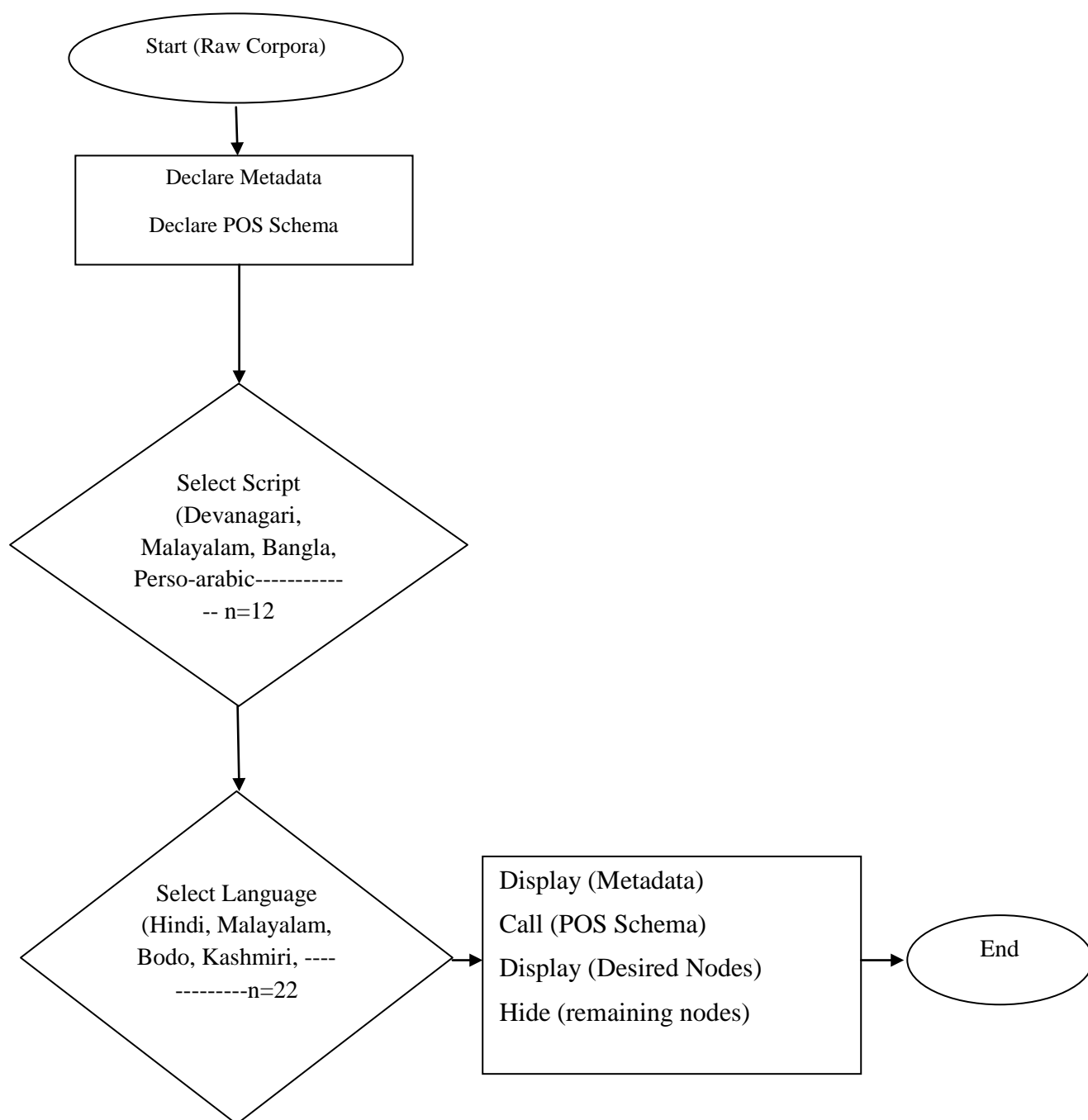
## 10. ONE TO ONE MAPPING LABELS IN POS SCHEMA

In order to develop common framework of XML based POS schema in all 22 Indian Languages, it is necessary that labels defined in POS Schema for English to have one to one mapping for Indian Languages. The XML schema needs to have a complete tree structure as depicted in fig. below:



The common XML Schema would select a particular Indian Language by and the Schema then needs to be transformed into POS Schema for that particular language. The language specific POS Schema could be enabled by making a particular branch of the tree structure 'off'. It is schematically represented in the next heading. i.e. POS schema block diagram

## 11. POS SCHEMA BLOCK DIAGRAM





## 12. DRAFT POS SCHEMA FOR INDIAN LANGUAGES USING XML

Pos schema ()

{

<?xml version="1.0" encoding="UTF-8"?>

<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

<file Desc>

<titleStmnt>

<title>POS tag in multilingual language</title>

<script>..... </script>

<language>multilingual</language>

<label language>.....</label language>

<type>multimodal</type>

[Languages taken: Hindi, Bodo, Malayalam, Kashmiri, Assamese, Konkani, Gujarati]

### Noun Block

<xs:element name="cat" POS cat="noun" hin-cat="संज्ञा" brx-cat="मुंमा" mal-cat="നാമം" kas-cat="नाउत" asm-cat="विशेष" kok-cat="नाम" guj-cat="संज्ञ" tag="N">

<xs:attribute name="type" subcat="common" hin-cat="जातिवाचक" brx-cat="फोलेर दिन्थिग्रा" mal-cat="സാമാന്യ നാമം" kas-cat="عام" asm-cat="जातिवाचक" kok-cat="जातवाचक नाम" guj-cat="जतिवाचक" tag="NN">

<xs:attribute name="type" subcat="Proper" hin-cat="व्यक्तिवाचक" brx-cat="मुं दिन्थिग्रा" mal-cat="സംജ്ഞാ നാമം" kas-cat="خاص" asm-cat="व्यक्तिवाचक" kok-cat="व्यक्तीवाचक नाम" guj-cat="व्यक्तिवाचक" tag="NNP">

<xs:attribute name="type" subcat="Verbal" hin-cat="क्रियामूलक" brx-cat="हाबा दिन्थिग्रा" kas-cat="क्राؤताؤی" asm-cat="क्रियावाचक" kok-cat="क्रियामूलक नाम" guj-cat="क्रियावाचक" tag="NNV">

<xs:attribute name="type" subcat="Nloc" hin-cat="देश-काल सापेक्ष" brx-cat="थावनि  
दिन्थिग्रा मुंमा" mal-cat="ആയാതീക നാമം" kas-cat="ناوتہ جاہیہ ہاو " asm-cat="स्थानवाचक"  
kok-cat="थळसापेक्ष-काळ- नाम" guj-cat="स्थानवलय" tag="NST">

### -----Pronoun Block-----

<xs:element name="cat" POS cat="Pronoun" hin-cat="सर्वनाम" brx-cat="मुंराइ" mal-  
cat="സർവ്വനാമം" kas-cat="پَرِناوُت " asm-cat="सर्वनाम" kok-cat="सर्वनाम" guj-  
cat="सर्वनाम" tag="PR">

<xs:attribute name="type" subcat="Personal" hin-cat="व्यक्तिवाचक" brx-cat="संबु  
दिन्थिग्रा" mal-cat="പുരുഷ സർവ്വനാമം" kas-cat="شخصیاتی " asm-cat="व्यक्तिवाचक"  
kok-cat="पुरुश सर्वनाम" guj-cat="पुरुषवलय" tag="PRP">

<xs:attribute name="type" subcat="Reflexive" hin-cat="निजवाचक" brx-cat="गाव  
दिन्थिग्रा" mal-cat="നീചവാചി സർവ്വനാമം" kas-cat="ماکوسی " asm-cat="आत्मवाचक"  
kok-cat="आत्मवाचक सर्वनाम" guj-cat="प्रतिबिंबित" tag="PRF">

<xs:attribute name="type" subcat="Reciprocal" hin-cat="पारस्परिक" brx-  
cat="गावजों गाव सोमोन्दो" mal-cat="സംബന്ധവാചി സർവ്വനാമം" kas-cat="باہمی " asm-  
cat="पारस्परिक" kok-cat="संबंदी सर्वनाम" guj-cat="परस्परवली" tag="PRC">

<xs:attribute name="type" subcat="Relative" hin-cat="सम्बन्ध वाचक" brx-  
cat="सोमोन्दो दिन्थिग्रा" mal-cat="പാരസ്പരീക സർവ്വനാമം" kas-cat="رابطاوی " asm-  
cat="सम्बन्धवाचक" kok-cat="एकमेकी सर्वनाम" guj-cat="सापेक्ष" tag="PRL">

<xs:attribute name="type" subcat="Wh-words" hin-cat="प्रश्नवाचक" brx-cat="सोंथि  
दिन्थिग्रा" mal-cat="ചോദ്യവാചി സർവ്വനാമം" kas-cat="ک لفظ " asm-cat="प्रश्नबोधक  
सर्वनाम" kok-cat="प्रस्नार्थी सर्वनाम" guj-cat="प्रश्नार्थवलय" tag="PRQ">

## -----Demonstrative Block-----

<xs:element name="cat" POS cat="Demonstrative" hin-cat="निश्चयवाचक" brx-cat="थावनि दिन्थिग्रा" mal-cat="നീർദേശകം" kas-cat="بَاوَن پَرَناوَتِي" asm-cat="নির্দেশবাধক" kok-cat="दर्शक" guj-cat="દર્શક" tag="DM">

<xs:attribute name="type" subcat="Deictic" hin-cat="" brx-cat="थि दिन्थिग्रा" mal-cat="പ്രത്യക്ഷ സൂചകം" kas-cat="وَأَنبَأَوِي" asm-cat="প্রত্যক্ষ নির্দেশক" kok-cat="" guj-cat="উল্লেখ্যদর্শক" tag="DMD">

<xs:attribute name="type" subcat="Relative" hin-cat="सम्बन्ध वाचक" brx-cat="सोमोन्दो दिन्थिग्रा" mal-cat="സംബന്ധവാചി നീർദേശകം" kas-cat="رَأْبَأَوِي" asm-cat="সম্বন্ধবাধক" kok-cat="संबंदी दर्शक" guj-cat="સાપેક્ષ" tag="DMR">

<xs:attribute name="type" subcat="Wh-words" hin-cat="प्रश्नवाचक" brx-cat="म सोंथि दिन्थिग्रा" mal-cat="ചോദ്യവാചി നീർദേശകം" kas-cat="ك لَفْظ" asm-cat="প্রশ্নবাধক অব্যয়" kok-cat="प्रस्नार्थी दर्शक" guj-cat="પ્રશ્નવાચી" tag="DMQ">

## -----Verb Block-----

<xs:element name="cat" POS cat="Verb" hin-cat="क्रिया" brx-cat="थाइजा" mal-cat="ക്രിയ" kas-cat="کَرَاوَت" asm-cat="ক্রিয়া" kok-cat="क्रियापद" guj-cat="અભ્યત" tag="V">

<xs:attribute name="type" subcat="Auxiliary Verb" hin-cat="सहायक क्रिया" brx-cat="लेडाइ थाइजा" mal-cat="സഹായക ക്രിയ" kas-cat="كُرَاهِي كَرَاوَت" asm-cat="সহায়কাৰী ক্রিয়া" kok-cat="पालवी क्रियापद" guj-cat="" tag="VAUX">

<xs:attribute name="type" subcat="Main Verb" hin-cat="मुख्य क्रिया" brx-cat="गुबै थाइजा" mal-cat="പ്രധാന ക്രിയ" kas-cat="راے کراوَت" asm-cat="মূখ্য ক্রিয়া" kok-cat="मुखेल क्रियापद" guj-cat="મુખ્ય" tag="VM">

<xs:attribute name="subtype" subcat="Finite" hin-cat="परिमित" brx-cat="जाफुंजा थाइजा" mal-cat="പൂർണ്ണ ക്രിയ" kas-cat="بَشْرَ ہاو" asm-cat="সমাপিকা" kok-cat="निश्चित क्रियापद" guj-cat="પૂર્ણ" tag="VF">

<xs:attribute name="subtype" subcat="Infinitive" hin-cat="अनंत" brx-cat="जाफुडि थाइजा" mal-cat="ക്രിയാരൂപം" kas-cat="بَشْرَ كَھاو" asm-cat="অসমাপিকা" kok-cat="सादरण रूप" guj-cat="હેત્વર્થ" tag="VINF">

<xs:attribute name="subtype" subcat="Gerund" hin-cat="क्रियावाचक" brx-cat="जाफुबाय थानाय दिन्थिया" kas-cat="کَرَاوَتِه ناؤت" asm-cat="নিমিত্তার্থক সংজ্ঞা" kok-cat="क्रियावाचक नाम" guj-cat="વર્તમાનકૃદન્ત" tag="VNG">

<xs:attribute name="subtype" subcat="Non-Finite" hin-cat="गैर परिमित" brx-cat="जाफुडि थाइजा" mal-cat="അപൂർണ്ണ ക്രിയ" kas-cat="نا بَشْرَ ھاو" asm-cat="অসমাপিকা" kok-cat="अनिश्चित क्रियापद" guj-cat="અપૂર્ણ" tag="VNF">

### -----Adjective Block-----

<xs:element name="cat" POS cat="Adjective" hin-cat="विशेषण" brx-cat="थाइलालि" mal-cat="നാമ വിശേഷണം" kas-cat="بَاؤت" asm-cat="বিশেষণ" kok-cat="विशेषण" guj-cat="વિશેષણ" tag="JJ">

### -----Adverb Block-----

<xs:element name="cat" POS cat="Adverb" hin-cat="क्रिया विशेषण" brx-cat="थाइजानि थाइलालि" mal-cat="ക്രിയാ വിശേഷണം" kas-cat="لَگَہ بَاش" asm-cat="ক্রিয়া বিশেষণ" kok-cat="क्रियाविशेषण" guj-cat="ક્રિયાવિશેષણ" tag="RB">

### -----Post Position Block-----

<xs:element name="cat" POS cat="Post Position" hin-cat="परसर्ग" brx-cat="सोदोब उन महरथि" mal-cat="അനുമുഖപ്രയോഗം" kas-cat="پُوت جائے" asm-cat="অনুসর্গ" kok-cat="संबन्दी अव्यय" guj-cat="અનુગો" tag="PSP">

### -----Conjunction Block-----

<xs:element name="cat" POS cat="Conjunction" hin-cat="योजक" brx-cat="दाजाब महरथि" mal-cat="സമുച്ചയം" kas-cat="واٹون" asm-cat="সংযোজক" kok-cat="जोड अव्यय" guj-cat="સંયોજકો" tag="CC">

<xs:attribute name="type" subcat="Co-ordinator" hin-cat="समन्वयक" brx-cat="लोगो महर" mal-cat="ഏകോപിത സമുച്ചയം" kas-cat="وائت " asm-cat="समन्वयक" kok-cat="समानाधीकरण जोड अव्यय" guj-cat="ಸಹಕರಿಸಿದ ಶಬ್ದ" tag="CCD">

<xs:attribute name="type" subcat="Subordinator" hin-cat="" brx-cat="लेडाइ लोगो महर" mal-cat="ആശ്രിതസമുച്ചയം" kas-cat="تحتون " asm-cat="" kok-cat="आश्रीत जोड अव्यय" guj-cat="ಗೌಢೀಕೃತ ಶಬ್ದ" tag="CCS">

<xs:attribute name="subtype" subcat="Quotative" hin-cat="उक्ति-वाचक" mal-cat="ഉദ്ധാരണവാചി സമുച്ചയം" brx-cat="मुख'थि" kas-cat="دَپَن نِشانِه " asm-cat="" kok-cat="अवतरणअर्थी- उतर" guj-cat="" tag="UT">

### -----Particles Block-----

<xs:element name="cat" POS cat="Particles" hin-cat="अव्यय" brx-cat="महरथि" mal-cat="നീപാദം" kas-cat="تَوْتِه وَتِي " asm-cat="आनुषंगिक अवयव" kok-cat="अव्यय" guj-cat="ನಿಪಲಿತ" tag="RP">

<xs:attribute name="type" subcat="Default" hin-cat="व्यतिक्रम" brx-cat="गोरोन्धि" mal-cat="സാമാന്യം" kas-cat="دُفَالْت " asm-cat="" kok-cat="सरभरस अव्यय" guj-cat="स्वयंभू" tag="RPD">

<xs:attribute name="type" subcat="Classifier" hin-cat="वर्गीकारक" brx-cat="थि दिन्थिग्रा दाजाबदा" mal-cat="പരഗ്ഗുഹം" kas-cat="وَرگِه " asm-cat="निर्दिष्टतावाचक शर्ग" kok-cat="वर्गक अव्यय" guj-cat="" tag="CL">

<xs:attribute name="type" subcat="Interjection" hin-cat="विस्मयादिबोधक" brx-cat="सोमोनांनाय दिन्थिग्रा" mal-cat="പ്രാക്ഷേപഹം" kas-cat="زَهْت " asm-cat="विस्मयबोधक" kok-cat="उमाळी अव्यय" guj-cat="" tag="INJ">

<xs:attribute name="type" subcat="Negation" hin-cat="नकारात्मक" brx-cat="नडि दिन्थिग्रा" mal-cat="നീഷേദം" kas-cat="نَه كَأرى " asm-cat="नकारक" kok-cat="न्हयकारी अव्यय" guj-cat="ನಕಲರಿದ ಶಬ್ದ" tag="NEG">

<xs:attribute name="type" subcat="Intensifier" hin-cat="तीव्रक" brx-cat="गुन  
दिन्थिग्रा" mal-cat="തീവ്ര മനോഭാവം" kas-cat="شدت بار" asm-cat="" kok-cat="तीव्रकारी  
अव्यय" guj-cat="મત્રસૂચક" tag="INTF">

### -----Quantifiers Block-----

<xs:element name="cat" POS cat="Quantifiers" hin-cat="संख्यावाची" brx-cat="बिबां  
दिन्थिग्रा" mal-cat="സംഖ്യാവാചി" kas-cat="گرینڈ" asm-cat="अभिमाणवाचक" kok-  
cat="संख्यादर्शक" guj-cat="પરિમાણસૂચકો" tag="QT">

<xs:attribute name="type" subcat="General" hin-cat="सामान्य" brx-cat="सरासनस्रा"  
mal-cat="പൊതുസംഖ്യാവാചി" kas-cat="عمومی" asm-cat="आधाबण" kok-  
cat="सामान्य" guj-cat="સામાન્ય" tag="QTF">

<xs:attribute name="type" subcat="Cardinals" hin-cat="गणनासूचक" brx-cat="गुबै  
बिसान" mal-cat="അടിസ്ഥാന സംഖ്യാവാചി" kas-cat="آنکونہ گرینڈ" asm-  
cat="अंशवाचक" kok-cat="संख्यावाचक" guj-cat="સંખ્યાવલયક" tag="QTC">

<xs:attribute name="type" subcat="Ordinals" hin-cat="क्रमसूचक" brx-cat="फारि  
बिसान" mal-cat="കരമ്മവാചി" kas-cat="وئی گرینڈ" asm-cat="क्रमवाचक अंशवाचक  
शब्द" kok-cat="क्रमवाचक" guj-cat="ક્રમવલયક" tag="QTO">

### -----Residuals Block-----

<xs:element name="cat" POS cat="Residuals" hin-cat="अवशेष" brx-cat="आद्रा" mal-  
cat="അവശിഷ്ടപദം" kas-cat="باقیاتی" asm-cat="" kok-cat="हेर" guj-cat="શેષ" tag="RD">

<xs:attribute name="type" subcat="Foreign word" hin-cat="विदेशी शब्द" brx-  
cat="गुबुन हादरारि सोदोब" mal-cat="അന്യഭാഷാപദം" kas-cat="غار ملکی لفظ" asm-  
cat="विदेशी शब्द" kok-cat="विदेशी" guj-cat="પરદેશી શબ્દો" tag="RDF">

<xs:attribute name="type" subcat="Symbol" hin-cat="प्रतीक" brx-cat="नेर्सीन" mal-  
cat="ചിഹ്നം" kas-cat="علامت" asm-cat="प्रतीक" ki="कुरु" guj-cat="સંકેત" tag="SYM">

```
<xs:attribute name="type" subcat="Unknown" hin-cat="अज्ञात" brx-cat="मिथियि"
mal-cat="ഇതരപദം" kas-cat="أزون" asm-cat="अज्ञात" kok-cat="अनवळखी" guj-
cat="अज्ञात शब्द" tag="UNK">
```

```
<xs:attribute name="type" subcat="Punctuation" hin-cat="विरामादि-चिह्न" brx-
cat="थाद' सिन खान्थि" mal-cat="വിരാമ ചിഹ്നം" kas-cat="أهجون" asm-cat="যতি
চিন" kok-cat="विरामकूरु" guj-cat="વિરામચિહ્ને" tag="PUNC">
```

```
<xs:attribute name="type" subcat="Echowords" hin-cat="प्रतिध्वनि-शब्द" brx-
cat="रिंखां सोदोब" mal-cat="മാററ്റാലിവാക്ക" kas-cat="پوت دنى لفظ" asm-
cat="ধ্বন্যাত্মক শব্দ" kok-cat="पडसादी उतरां" guj-cat="અનુરણભાષ" tag="ECH">
```

```
</xs:attribute>
```

```
</xs:element></xs:schema>
```

```
}
```

### 13. ONE TO ONE MAPPING LABELS FOR INDIAN LANGUAGES

To incorporate such facility in the xml Schema the common one to one mapping table for the labels has been developed as presented in the Table 1, Table 2 and Table 3

#### Languages: Hindi, Punjabi, Urdu, Gujarati, Oriya, Bengali

S.No	English	Hindi	Punjabi	Urdu	Gujarati	Oriya	Bengali
1	Noun	संज्ञा	ਨਾਂਵ	اسم	સંજ્ઞા	ସଂଜ୍ଞା	বিশেষ্য
	common	जातिवाचक	ਆਮ	نكره	જાતિવાચક	ଜାତିବାଚକ	জাতিবাচক
	Proper	व्यक्तिवाचक	ਖਾਸ	معرفه	વ્યક્તિવાચક	ବ୍ୟକ୍ତିବାଚକ	ব্যক্তিবাচক
	Verbal	क्रियामूलक / कृतंत	ਕਿਰਿਆਮੂਲਕ	حاصل مصدر	ક્રિયાવાચક	କ୍ରିୟାବାଚକ	ক্রিয়ামূলক
	Nloc	देश-काल सापेक्ष	ਸਥਿਤੀ ਸੂਚਕ	ظرف	સ્થાનવાચક	ଦଶ-କାଳ ସାପੇક્ଷ	স্থানবাচক
2	Pronoun	सर्वनाम	ਪੜਨਾਂਵ	ضمير	સર્વનામ	ସର୍ବନାମ	সর্বনাম
	Personal	व्यक्तिवाचक	ਪੁਰਖਵਾਚੀ	ضمير شخصي	પુરુષવાચક	ବ୍ୟକ୍ତିବାଚક	ব্যক্তিবাচক
	Reflexive	निजवाचक	ਨਿਜਵਾਚੀ	ضمير معكوسي	પ્રતિબિંબિત	ଆତ્ମବାଚક	আস্ববাচক
	Reciprocal	पारस्परिक	ਪਰਸਪਰੀ	ضمير راجع	પરસ્પરવાચી	ପାରસ્પારિક	ব্যতিহার
	Relative	संबंध- वाचक	ਸੰਬੰਧਵਾਚੀ	ضمير موصول	સાપેક્ષ	ସଂବନ୍ଧବାଚକ	সম্বন্ধবাচক
	Wh-words	प्रश्नवाचक	ਪ੍ਰਸ਼ਨਵਾਚੀ	ضمير استفهاميه	પ્રશ્નાર્થવાચક	ପ୍ରଶ୍ନବାଚକ	প্রশ্নবাচক
	Indefinite	अनिश्चयवाचक	NA	NA	અનિશ્ચિત સર્વનામ	NA	অনির্দেশ্য
3	Demonstrative	निश्चयवाचक/ संकेतवाचक	ਸੰਕੇਤਵਾਚੀ	اشارے	દર્શક	ନିଶ્ଚયବାଚક/ସଂକେତବାଚକ	নির্দেশক
	Deictic	निर्देशी	ਪ੍ਰਤੱਖ ਪ੍ਰਮਾਣਵਾਚੀ	اشاره	ઉલ્લેખદર્શક		প্রত্যক্ষ নির্দেশক
	Relative	संबंधवाचक	ਸੰਬੰਧਵਾਚੀ	اشاره موصول	ਸਾਪੇક્ષ	ସଂବନ୍ଧବାਚକ	সম্বন্ধবাচক
	Wh-words	प्रश्नवाचक	ਪ੍ਰਸ਼ਨਵਾਚੀ	اشاره استفهاميه	પ્રશ્નવાચી	ପ୍ରଶ્ନବାଚକ	প্রশ্নবাচক
	Indefinite	अनिश्चयवाचक	NA	NA	અનિશ્ચિત સર્વનામ	NA	অনির্দেশ্য
4	Verb	क्रिया	ਕਿਰਿਆ	فعل	આખ્યાત	କ୍ରିୟା	ক্রিয়া
	Auxiliary Verb	सहायक क्रिया	ਸਹਾਇਕ ਕਿਰਿਆ	فعل امدادی	સહાયકારી ક્રિયા	ସହાયક ક્રિયા	গৌণ ক্রিয়া
	Main Verb	मुख्य क्रिया	ਮੁੱਖ ਕਿਰਿਆ	فعل لازم	મુખ્ય	ମୁଖ્ય ક્રિયા	মূখ্য ক্রিয়াপদ
	Finite	परिमित	ਕਾਲਕੀ	فعل محدود	પૂર્ણ	ପରିମિત	সমাপিকা
	Infinitive	क्रियार्थक संज्ञा	ਅਮਿਤ	مصدر	હેત્વર્થ	ଅନନ୍ତ	অপূর্ণ ক্রিয়া



	Gerund	ક્રિયાવાચક	કિરિઆવાંચી	حاصل مصدر	વર્તમાનકૃદન્ત	ક્રિયાવારક	પ્રયોજક ક્રિયા
	Non-Finite	ઘૈર-પરિમિત	અકાલકી	فعل غير محدود	અપૂર્ણ	અપરિમિત	અસમાપિકા
	Participle Noun	કૃદન્ત પરક નામ	NA	NA	NA	NA	ક્રિયાજાત વિશેષ્ય
5	Adjective	વિશેષણ	વિસેસ્ટ	صفت	વિશેષણ	વિશેષણ	વિશેષણ
6	Adverb	ક્રિયા-વિશેષણ	કિરિઆ વિસેસ્ટ	متعلق فعل	ક્રિયાવિશેષણ	ક્રિયા-વિશેષણ	ક્રિયા-વિશેષણ
7	Post Position	પરસર્ગ	સઘંપક	جار مؤخر	અનુગો	પરઘર્ગ	પરસર્ગ
8	Conjunction	યોજક	યેનક	حرف عطف	સંયોજકી	સંયોજકી	સંયોગમૂલક
	Co-ordinator	સમન્વયક	સમાન યેનક	حرف وصل	સહક્રિયાદર્શક	સમન્વયક	સમન્વયક
	Subordinator	અધીનસ્થ	અધીન યેનક	حرف تابع کننده	ગૌણક્રિયાદર્શક		શર્ત સંયોજક
	Quotative	ઉક્તિ-વાચક	કથનવાંચી	حرف اقتباسی	NA	ઉક્તિવારક	ઉક્તિવાચક
9	Particles	અવ્યય	નિપાત	حرف حالیه/پابند	નિપાત	અવ્યય / નિપાત	અવ્યય
	Default	વ્યતિક્રમ	ઉત્તરીવાંચક	حرف ڈیفالٹ	સ્વયંભૂ	વ્યતિક્રમ	સાધારણ અવ્યય
	Classifier	વર્ગીકારક	વરગીવિંત	حرف درجه بند	NA	વર્ગીકારક	વર્ગવાચક
	Interjection	વિસ્મયાદિબોધક	વિસમક	حرف فجائیه	વિસ્મયઆદિ બોધક	વિસ્મય ૧૬-૧૮	વિસ્મયાદિબોધક
	Negation	નકારાત્મક	નાંહવાંચી	حرف نهی	નકારદર્શક	નિષેધોત્પન્ન	નકાર્થક
	Intensifier	તીવ્રક	તીવ્રતાવાંચી	حرف تاکید	માત્રાસૂચક	તીવ્રતાવારક	તીવ્રતાબોધક
10	Quantifiers	સંખ્યાવાંચી	સંખ્યાવાંચી	کمیت نما	પરિમાણસૂચકી	સંખ્યાવારકી	પરિમાણવાચક
	General	સામાન્ય	સપારન	عمومی/ عام	સામાન્ય	સામાન્ય	સાધારણ
	Cardinals	ગણનાસૂચક	ગિણતીસૂચક	اعداد مطلق	સંખ્યાવાચક	ગણનાવૃત્ત	સંખ્યાવાચક
	Ordinals	ક્રમસૂચક	ક્રમસૂચક	ترتیبی اعداد	ક્રમવાચક	ક્રમવૃત્ત	ક્રમવાચક
11	Residuals	અવશેષ	ઘાકી	باقی مانده	શેષ	અવશેષ	અવશેષ પદ
	Foreign word	વિદેશી શબ્દ	વિદેશી શબ્દ	بیرونی لفظ	પરદેશી શબ્દ	વિદેશી શબ્દ	વિદેશી શબ્દ
	Symbol	પ્રતીક	સંકેત	علامت	સંકેત	પ્રતીક	પ્રતીક
	Unknown	અજ્ઞાત	અગિઆત	نامعلوم	અજ્ઞાત શબ્દ	અજ્ઞાત	અજ્ઞાત
	Punctuation	વિરામાદિ-ચિહ્ન	વિરામાદિ ચિહ્ન	اوقاف	વિરામચિહ્ન	વિરામ ચિહ્ન	ચિહ્નચિહ્ન
	Echowords	પ્રતિધ્વનિ-શબ્દ	પ્રતિધ્વનિ શબ્દ	گونج دار الفاظ	અનુરણનાત્મક	પ્રતિધ્વનિ	અનુકાર શબ્દ

**Languages: Assamese, Bodo, Kashmiri (Urdu Script), Kashmiri (Hindi Script), Marathi**

S.No	English	Hindi	Assamese	Bodo	Kashmiri	Kashmiri (Hindi)	Marathi
1	Noun	संज्ञा	विशेष्य	मुंमा	नाؤत	नावुत	नाम
	common	जातिवाचक	जातिवाचक	फोलेर दिन्थिग्रा	عام	आम	सामान्य नाम
	Proper	व्यक्तिवाचक	व्यक्तिवाचक	मुं दिन्थिग्रा	خاص	खास	वशिष नाम
	Verbal	क्रियामूलक / कृदंत	क्रियावाचक	हाबा दिन्थिग्रा	کڑاوتأوی	क्रावतांव्य	धातुसाधति नाम
	Nloc	देश-काल सापेक्ष	ज्ञानवाचक	थावनि दिन्थिग्रा मुंमा	ناوتہ جايہ ہاو	नाव त जायि हाव	देश कालवाचक नाम
2	Pronoun	सर्वनाम	प्रवर्णनाम	मुंराइ	پَرناؤت	पर नावुत	सर्वनाम
	Personal	व्यक्तिवाचक	व्यक्तिवाचक	संबुं दिन्थिग्रा	شخصیاتی	शख्सियांती	पुरुषवाचक
	Reflexive	निजवाचक	आत्मवाचक	गाव दिन्थिग्रा	ماکوسی	माकूसी	आत्मवाचक
	Reciprocal	पारस्परिक	पारस्परिक	गावजों गाव सोमोन्दो	بابمی	बाहिमी/ बोहिमी	पारस्परिक
	Relative	संबंध- वाचक	सम्बन्धवाचक	सोमोन्दो दिन्थिग्रा	رابطأوی	रोबितांव्य	संबंधवाची
	Wh-words	प्रश्नवाचक	प्रश्नावोधक सर्वनाम	सोंथि दिन्थिग्रा	ک لفظ	क-लफ़ज़	प्रश्नार्थक
	Indefinite	अनिश्चयवाचक					
3	Demonstrative	निश्चयवाच/ संकेतवाचक	निर्देशवाधक	थावनि दिन्थिग्रा	ہاؤن پَرناؤتی	हावन परनावुत्य	दर्शक
	Deictic	निर्देशी	प्रत्यक्ष निर्देशक	थि दिन्थिग्रा	وانیأوی	वोनयोव्य	
	Relative	सम्बन्ध वाचक	सम्बन्धवाचक	सोमोन्दो दिन्थिग्रा	رابطأوی	रोबतांत्य	संबंधवाच/ संबंधदर्शक
	Wh-words	प्रश्नवाचक	प्रश्नावोधक अव्यय	म सोंथि दिन्थिग्रा	ک لفظ	क-लफ़ज़	प्रश्नार्थक
	Indefinite	अनिश्चयवाचक	NA	NA	NA	NA	NA
4	Verb	क्रिया	क्रिया	थाइजा	کڑاؤت	क्रावुत	क्रियापद
	Auxiliary Verb	सहायक क्रिया	सहायकाबी क्रिया	लेडाइ थाइजा	ٹکھہ کڑاؤت	डख क्रावुत	सहायकारी क्रियापद
	Main Verb	मुख्य क्रिया	मूथ्य क्रिया	गुबै थाइजा	راے کڑاؤت	राय क्रावुत	मुख्य क्रियापद
	Finite	परिमित	प्रभाषिका	जाफुंजा थाइजा	پَشَر ہاو	हिशर हाव	आख्यात क्रियारूप

	Infinitive	अनंत	अप्रभाषिका	जाफुडि थाइजा	بَشْرَ كَهاو	हिशर खाव	भाववाचक कृदंत
	Gerund	क्रियावाचक	बिभित्थार्थक संज्ञा	जाफुबाय थानाय दिन्थिगा	كَراوْتِهْ नाउत	क्राव त नावुत	वभित्तिक्पम कृदंतरूप
	Non-Finite	गैर-परिमित	अप्रभाषिका	जाफुडि थाइजा	نا بَشْرَ हाव	ना हिशर हाव	आख्यातेतर क्रियारूप
	Participle Noun	कृदंत परक नाम	NA	NA	NA	NA	NA
5	Adjective	विशेषण	विशेषण	थाइलालि	बाउत	बावुत	वशिषण
6	Adverb	क्रिया-विशेषण	क्रिया विशेषण	थाइजानि थाइलालि	لَگِهْ बाश	लग बांश	क्रियावशिषण
7	Post Position	परसर्ग	अनूसर्ग	सोदोब उन महरथि	پوْت جاے	पोत जाय	अंत्यस्थान
8	Conjunction	योजक	संयोजक	दाजाब महरथि	واثون	राटवन	उभयान्वयी अव्यय
	Co-ordinator	समन्वयक	सम्बन्धक	लोगो महर	واثت	वाटत/ वाटथ	NA
	Subordinator	अधीनस्थ	NA	लेडाइ लोगो महर	تحتون	तहतून	NA
	Quotative	उक्ति-वाचक	NA	मुंख'थि	دَپِن نِشانِه	दपन निशान	उद्गारवाचक
9	Particles	अव्यय	आनुसंगिक अव्यय	महरथि	ثوْتِهْ وَبَنِي	टोट वन्त्य	अव्यय/ नपात
	Default	व्यतिक्रम		गोरोन्थि	ثِفَالْت	डिफाल्ट	सामान्य
	Classifier	वर्गीकारक	निर्दिष्टतावाचक सर्ग	थि दिन्थिगा दाजाबदा	وَرگِهْ	वरगहा	NA
	Interjection	विस्मयादि बोधक	विस्मयबोधक	सोमोनांनाय दिन्थिगा	زَهْتت	छटत/ छटथ	वस्मयवाचक
	Negation	नकारात्मक	नकारार्थक	नडि दिन्थिगा	نَه كَارِي	नकारय	नपिधात्मक
	Intensifier	तीव्रक		गुन दिन्थिगा	شَدَت هار	शदत हाव	तीव्रतावाचक
10	Quantifiers	संख्यावाची	प्रमाणवाचक	बिबां दिन्थिगा	گَرِيْنَد	ग्रेन्द	संख्यावाचक
	General	सामान्य	प्राधान्य	सरासनसा	عمومي	अमूमि	सामान्य
	Cardinals	गणनासूचक	संख्यावाचक	गुबै बिसान	انكونِهْ गरीन्द	ओकवन ग्रेन्द	गणनावाचक
	Ordinals	क्रमसूचक	क्रमवाचक संख्यावाचक शब्द	फारि बिसान	وَنِي گَرِيْنَد	वेन्य ग्रेन्द	क्रमवाचक
11	Residuals	अवशेष	NA	आद्रा	باقياْتِي	बाक्यांती	शेष

	Foreign word	विदेशी शब्द	विदेशी शब्द	गुबुन हादरारि सोदोब	غار مُلڪى ألفظ	गोर मुल्की लफुज़	वदिशी शब्द
	Symbol	प्रतीक	प्रतीक	नेर्सेन	علامت	अलामत	चन्ह
	Unknown	अज्ञात	अज्ञात	मिथियि	آزون	अज़ोन	अज्ञात
	Punctuation	विरामादि-चिह्न	यति चिह्न	थाद' सिन खान्थि	لہجُون	लहजिवन	वरिमचन्हि
	Echowords	प्रतिध्वनि-शब्द	ध्वन्यात्मक शब्द	रिंखां सोदोब	پوٽ دنى لفظ	पोत देन्य लफ़ज़	नादानुकारी/ अभ्यस्त

**Languages: Telugu, Malayalam, Tamil, Konkani**

S.No.	English	Hindi	Telugu	Malayalam	Tamil	Konkani
1	Noun	संज्ञा	సంజ్ఞా	നാമം	பெயர்	नाम
	common	जातिवाचक	జాతివచకం	സാമാന്യ നാമം	பொதுப் பெயர்	जातवाचक नाम
	Proper	व्यक्तिवाचक	వ్యక్తివచకం	സംജ്ഞാ നാമം	சிறப்புப் பெயர்	व्यक्तीवाचक नाम
	Verbal	क्रियामूलक / कृदंत	కరయములకం	NA	தொழில் பெயர்	क्रियामूलक नाम
	Nloc	देश-काल सापेक्ष	దేశ-కాల సపేక్షకం	ആയാരിക നാമം	இடப் பெயர்	थळ-काळ-सापेक्ष नाम
2	Pronoun	सर्वनाम	సర్వనామం	സർവ്വനാമം	பதிலீடும் பெயர்	सर्वनाम
	Personal	व्यक्तिवाचक	వ్యక్తివచకం	പുരുഷ സർവ്വനാമം	மனிதர்பெயர்	पुरुश सर्वनाम
	Reflexive	निजवाचक	ఆత్మవచకం	നീചവാചി സർവ്വനാമം	தற்குட்பதிலீடும் பெயர்	आत्मवाचक सर्वनाम
	Reciprocal	पारस्परिक	పరస్పరకం	സംബന്ധവാചി സർവ്വനാമം	பரஸ்பர பதிலீடும் பெயர்	संबंदी सर्वनाम
	Relative	संबंध- वाचक	సంబంధ-వచకం	പാരസ്പിക സർവ്വനാമം	இணைப்பும் பதிலீடும் பெயர்	एकमेकी सर्वनाम
	Wh-words	प्रश्नवाचक	పశ్చర నవచకం	ചോദ്യവാചി സർവ്വനാമം	வினாச் சொல்	प्रस्नार्थी सर्वनाम
	Indefinite	अनिश्चयवाचक		NA	சுட்டும்	अनिश्चित सर्वनाम
3	Demonstrative	निश्चयवाचक/ संकेतवाचक	నరదశకవచకం	നിർദ്ദേശകം	நேர்ச்சுட்டும்	दर्शक
	Deictic	निर्देशी	నరదషట	പ്രത്യക്ഷ സൂചകം	சுட்டும் பதிலீடும் பெயர்	दर्शक उतर
	Relative	संबंधवाचक	సంబంధ-వచకం	സംബന്ധവാചി നിർദ്ദേശകം	வினாச் சொல்	संबंदी दर्शक
	Wh-words	प्रश्नवाचक	పశ్చర నవచకం	ചോദ്യവാചി നിർദ്ദേശകം	வினா	प्रस्नार्थी दर्शक
	Indefinite	अनिश्चयवाचक	NA	NA	தூணை வினா	अनिश्चित सर्वनाम
4	Verb	क्रिया	కరయ	ക്രിയ	மாதன்மை வினை	क्रियापद
	Auxiliary Verb	सहायक क्रिया	సహాయక కరయ	സഹായക ക്രിയ	முற்று வினை	पालवी क्रियापद Auxiliary Finite (पूर्ण पालवी)

						క్రియాపద) Auxiliary Non Finite (अपूर्ण पालवी क्रियापद)
	Main Verb	मुख्य क्रिया	ముఖ్య కరయ	ప్రధాన క్రియ	కూర్తణ ంశ్శమ్	मुखेल क्रियापद
	Finite	परिमित	సమపక	పూర్ణ క్రియ	విణణైప్పయెర్	निश्चित क्रियापद
	Infinitive	क्रियार्थक संज्ञा	తమునసరథకం	క్రియార్థం	విణణ ంశ్శమ్	सादारण रूप
	Gerund	क्रियावाचक	కరయవచకం	NA	పయెరడణ	क्रियावाचक नाम
	Non-Finite	गैर-परिमित	అసమపక	అపూర్ణ క్రియ	విణణయడణ	अनिश्चित क्रियापद
	Participle Noun	कृदंत परक नाम	NA	NA	పిణ్ణురూప	NA
5	Adjective	विशेषण	వశషణం	నామ విశేషణం	ఇణణైప్పొశ్శ శొల	विशेषण
6	Adverb	क्रिया-विशेषण	కరయవశషణం	క్రియా విశేషణం	ఇణణ ఇణణైప్పొశ్శ శొల	क्रियाविशेषण
7	Post Position	परसर्ग	పరసరగ	అనుప్రయో గం	శార్ప ఇణణైప్పొశ్శ శొల	संबंदी अव्यय
8	Conjunction	योजक	సముచయం	సమ్యయం	నిరప్ప ఇడణైప్పొశ్శ శొల	जोड अव्यय
	Co-ordinator	समन्वयक	సమనధకరణం	అంకూపిత సమ్యయం	ఇడణైప్పొశ్శ శొల	समानाधीकरण जोड अव्यय
	Subordinator	अधीनस्थ	వయధకరణం	అంతర్యంత అంక సమ్యయం	మణ్ణిరూప్ప అంక	आश्रीत जोड अव्यय
	Quotative	उक्ति-वाचक	అనుకరకం	అంతర్యంత అంక సమ్యయం	ఇణప్పిరిప్ప అంక	अवतरणअर्थी- उतर
9	Particles	अव्यय	అవయయం	నిరూపం	వియప్పిడణై శొల	अव्यय
	Default	व्यतिक्रम	వయతకరమం	సామాన్యం	ంతిరమణ	सरभरस अव्यय
	Classifier	वर्गीकारक	వరగకరకం	వర్గీకరణం	మికువిప్పాన్	वर्गक अव्यय
	Interjection	विस्मयादिबोधक	వసమయదబోధకం	వ్యక్తిపాదం	అణవణయడణ	उमाळी अव्यय
	Negation	नकारात्मक	నకరతమకం	నిరూపం	పొత	न्हयकारी अव्यय
	Intensifier	तीव्रक	అతీయరథకం	తీవ్ర నిరూపం	ంణ్ణుప్ప పయెర్	तीव्रकारी अव्यय
10	Quantifiers	संख्यावाची	సంఖ్యవచకం	సంఖ్యావాచి	ంణ్ణు మూర్తణైప్ప పయెర్	संख्यादर्शक

	General	सामान्य	సమనయం	പൊതുസം ഖ്യാവാചി	எஞ்சியவனை	सामान्य
	Cardinals	गणनासूचक	గణనసూచకం	അടിസ്ഥാന സംഖ്യാവാചി	அயல் சொல்	संख्यावाचक
	Ordinals	क्रमसूचक	కరమసూచకం	കർമ്മവാചി	கூறியீடு	क्रमवाचक
11	Residuals	अवशेष	అవశేషం	അവശിഷ്ടപദം	தெரியாதது	हेर
	Foreign word	विदेशी शब्द	వదేశ శబ్దం	അന്യഭാഷാപദം	நிறுத்தற்கூறியீடு	विदेशी
	Symbol	प्रतीक	సంకతం	ചിഹ്നം	இரட்டைக்கிளவி	कुरु
	Unknown	अज्ञात	అజ్ఞాత	ഇതരപദം	NA	अनवळखी
	Punctuation	विरामादि-चिह्न	వరమం	വീരാമ ചിഹ്നം	NA	विरामकूरु
	Echo-words	प्रतिध्वनि-शब्द	పరతధవన-శబ్దం	മാറ്റൊലിവാക്ക്	NA	पडसादी उतरां

## 14. ALGORITHM FOR SELECTION OF NODES

If script is Devanagari then

If language is Hindi then

Display (Metadata)

Call (POS Schema)

Display (English and Hindi Nodes)

Hide (remaining nodes)

Eg: {

```
<xs:element name="cat" POS cat="noun" hin-cat="संज्ञा" tag="N">
```

```
  <xs:attribute name="type" subcat="common" hin-cat="जातिवाचक" tag="NN">
```

```
  <xs:attribute name="type" subcat="Proper" hin-cat="व्यक्तिवाचक" tag="NNP">
```

```
<xs:element name="cat" POS cat="Pronoun" hin-cat="सर्वनाम" tag="PR">
```

```
  <xs:attribute name="type" subcat="Personal" hin-cat="व्यक्तिवाचक" tag="PRP">
```

```
  <xs:attribute name="type" subcat="Reflexive" hin-cat="निजवाचक" tag="PRF">
```

```
.....
```

```
}
```

End if

If language is Bodo then

Call (POS Schema)

Display (English, Hindi and Bodo Nodes)

Hide (remaining nodes)

Eg: {

```
<xs:element name="cat" POS cat="noun" hin-cat="संज्ञा" brx-cat="मुंमा" tag="N">
```



```
<xs:attribute name="type" subcat="common" hin-cat="जातिवाचक" brx-cat="फोलेर  
दिन्थिग्रा" tag="NN">
```

```
<xs:attribute name="type" subcat="Proper" hin-cat="व्यक्तिवाचक" brx-cat="मुं  
दिन्थिग्रा" tag="NNP">
```

```
<xs:element name="cat" POS cat="Pronoun" hin-cat="सर्वनाम" brx-cat="मुंराइ" tag="PR">
```

```
<xs:attribute name="type" subcat="Personal" hin-cat="व्यक्तिवाचक" brx-cat="संबुं  
दिन्थिग्रा" tag="PRP">
```

```
<xs:attribute name="type" subcat="Reflexive" hin-cat="निजवाचक" brx-cat="गाव  
दिन्थिग्रा" tag="PRF">
```

```
.....
```

```
}
```

```
End if
```

If language is Konkani then

Call (POS Schema)

Display (English, Hindi and Konkani Nodes)

Hide (remaining nodes)

Eg: {

```
<xs:element name="cat" POS cat="noun" hin-cat="संज्ञा" kok-cat="नाम" tag="N">
```

```
<xs:attribute name="type" subcat="common" hin-cat="जातिवाचक" kok-  
cat="जातवाचक नाम" tag="NN">
```

```
<xs:attribute name="type" subcat="Proper" hin-cat="व्यक्तिवाचक" kok-  
cat="व्यक्तीवाचक नाम" tag="NNP">
```

```
<xs:element name="cat" POS cat="Pronoun" hin-cat="सर्वनाम" kok-cat="सर्वनाम"  
tag="PR">
```

```
<xs:attribute name="type" subcat="Personal" hin-cat="व्यक्तिवाचक" kok-cat="पुरुष
सर्वनाम" tag="PRP">
```

```
<xs:attribute name="type" subcat="Reflexive" hin-cat="निजवाचक" kok-
cat="आत्मवाचक सर्वनाम" tag="PRF">
```

```
.....
```

```
}
```

End if

**Else If script is Malyalam** (Orthographic variation) then

**If language is Malyalam then**

**Call (POS Schema)**

**Display (English, Hindi and Malyalam Nodes)**

**Hide (remaining nodes)**

Eg: {

```
<xs:element name="cat" POS cat="noun" hin-cat="संज्ञा" mal-cat="നാമം" tag="N">
```

```
<xs:attribute name="type" subcat="common" hin-cat="जातिवाचक" mal-
cat="സാമാന്യ നാമം" tag="NN">
```

```
<xs:attribute name="type" subcat="Proper" hin-cat="व्यक्तिवाचक" mal-
cat="സംജ്ഞാ നാമം" tag="NNP">
```

```
<xs:element name="cat" POS cat="Pronoun" hin-cat="सर्वनाम" mal-cat="സർവ്വനാമം"
tag="PR">
```

```
<xs:attribute name="type" subcat="Personal" hin-cat="व्यक्तिवाचक" mal-
cat="പുരുഷ സർവ്വനാമം" tag="PRP">
```

```
<xs:attribute name="type" subcat="Reflexive" hin-cat="निजवाचक" mal-
cat="നിച്ചവാചി സർവ്വനാമം" tag="PRF">
```

```
.....
```

}

End if

Else If script is Perso-Arabic then

If language is Kashmiri then

Call (POS Schema)

Display (English, Hindi and Kashmiri Nodes)

Hide (remaining nodes)

Eg: {

```
<xs:element name="cat" POS cat="noun" hin-cat="संज्ञा" kas-cat="नाؤत" tag="N">
```

```
  <xs:attribute name="type" subcat="common" hin-cat="जातिवाचक" kas-cat="عام"
tag="NN">
```

```
  <xs:attribute name="type" subcat="Proper" hin-cat="व्यक्तिवाचक" kas-cat="خاص"
tag="NNP">
```

```
<xs:element name="cat" POS cat="Pronoun" hin-cat="सर्वनाम" kas-cat="پَرناؤت"
tag="PR">
```

```
  <xs:attribute name="type" subcat="Personal" hin-cat="व्यक्तिवाचक" kas-cat="
شخصیاتی" tag="PRP">
```

```
  <xs:attribute name="type" subcat="Reflexive" hin-cat="निजवाचक" kas-cat="
ماکوسی" tag="PRF">
```

.....

}

End if

Else If script is Bangla then

If language is Assamese then

Call (POS Schema)

Display (English, Hindi and Assamese Nodes)

Hide (remaining nodes)

Eg: {

```
<xs:element name="cat" POS cat="noun" hin-cat="সংজ্ঞা" asm-cat="বিশেষ্য" tag="N">
```

```
  <xs:attribute name="type" subcat="common" hin-cat="জাতিবাচক" asm-
  cat="জাতিবাচক" tag="NN">
```

```
  <xs:attribute name="type" subcat="Proper" hin-cat="ব্যক্তিবাচক" asm-
  cat="ব্যক্তিবাচক" tag="NNP">
```

```
<xs:element name="cat" POS cat="Pronoun" hin-cat="সর্বনাম" asm-cat="সর্বনাম"
tag="PR">
```

```
  <xs:attribute name="type" subcat="Personal" hin-cat="ব্যক্তিবাচক" asm-
  cat="ব্যক্তিবাচক" tag="PRP">
```

```
  <xs:attribute name="type" subcat="Reflexive" hin-cat="নিজবাচক" asm-
  cat="আত্মবাচক" tag="PRF">
```

```
.....
```

```
}
```

End if

Else If script is Gujarati then

If language is Gujarati then

Call (POS Schema)

Display (English, Hindi and Gujarati Nodes)

Hide (remaining nodes)

Eg: {

```
<xs:element name="cat" POS cat="noun" hin-cat="संज्ञा" guj-cat="સંજ્ઞા" tag="N">
```

```
  <xs:attribute name="type" subcat="common" hin-cat="जातिवाचक" guj-
  cat="જાતિવાચક" tag="NN">
```

```
  <xs:attribute name="type" subcat="Proper" hin-cat="व्यक्तिवाचक" guj-
  cat="વ્યક્તિવાચક" tag="NNP">
```

```
<xs:element name="cat" POS cat="Pronoun" hin-cat="सर्वनाम" guj-cat="સર્વનામ"
tag="PR">
```

```
  <xs:attribute name="type" subcat="Personal" hin-cat="व्यक्तिवाचक" guj-
  cat="પુરુષવાચક" tag="PRP">
```

```
  <xs:attribute name="type" subcat="Reflexive" hin-cat="निजवाचक" guj-
  cat="પ્રતિબિંબિત" tag="PRF">
```

```
.....
```

```
}
```

```
End if
```

## 15. REFERENCE BASED IMPLEMENTATION

### Hindi

1. सप्तपुरियों\N\_NNP के\PSP दर्शन\N\_NN से\PSP मिलता\N\_NN है\N\_NN मोक्ष\N\_NN !\RD\_PUNC
2. हिंदू\N\_NN धर्म\N\_NN में\N\_NN तीर्थ\N\_NN का\N\_NN बड़ा\N\_NN महत्व\N\_NN है\N\_NN !\RD\_PUNC
3. यू\RP\_RPD तो\RP\_RPD हर\QT\_QTF तीर्थ\N\_NN बड़ा\N\_NN और\CC\_CCD अहम\N\_NN है\N\_NN ,\RD\_PUNC लेकिन\CC\_CCS सात\QT\_QTC स्थानों\N\_NN की\N\_NN बड़ी\N\_NN महत्ता\N\_NN और\CC\_CCD मान्यता\N\_NN है\N\_NN !\RD\_PUNC
4. ये\DM\_DMD सातों\QT\_QTC धर्मस्थल\N\_NN सात\QT\_QTC नगरों\N\_NN या\RP\_RPD सप्तपुरियों\N\_NNP के\N\_NN रूप\N\_NN में\N\_NN ग्रंथों\N\_NN में\N\_NN वर्णित\N\_NN है\N\_NN !\RD\_PUNC
5. ऐसा\DM\_DMD कहा\N\_NN गया\N\_NN है\N\_NN कि\CC\_CCS चतुर्मास\N\_NNP में\N\_NN इन\DM\_DMD सप्तपुरियों\N\_NNP का\N\_NN दर्शन\N\_NN मोक्ष\N\_NN प्रदान\N\_NN करने\N\_NN वाला\N\_NN होता\N\_NN है\N\_NN !\RD\_PUNC

### Punjabi

1. सप्तपुरीਆਂ\N\_NN ਦੇ\N\_NN ਦਰਸ਼ਨ\N\_NN ਨਾਲ\N\_NN ਮਿਲਦਾ\N\_NN ਹੈ\N\_NN ਮੋਖ\N\_NN !\RD\_PUNC
2. ਹਿੰਦੂ\N\_NN ਧਰਮ\N\_NN ਵਿਚ\N\_NN ਤੀਰਥ\N\_NN ਦਾ\N\_NN ਬਹੁਤ\N\_NN ਮਹੱਤਵ\N\_NN ਹੈ\N\_NN !\RD\_PUNC
3. ਯੂ\RB ਤਾਂ\CC\_CCS ਹਰ\QT\_QTF ਤੀਰਥ\N\_NN ਵੱਡਾ\N\_NN ਤੇ\CC\_CCS ਅਹਿਮ\N\_NN ਹੈ\N\_NN ,\CC\_CCS ਪਰ\CC\_CCS ਸੱਤ\QT\_QTC ਸਥਾਨਾਂ\N\_NN ਦੀ\N\_NN ਬਹੁਤ\N\_NN ਮਹੱਤਤਾ\N\_NN ਅਤੇ\CC\_CCD ਮਾਨਤਾ\N\_NN ਹੈ\N\_NN !\RD\_PUNC
4. ਇਹ\DM\_DMD ਸੱਤ\QT\_QTC ਧਰਮ\N\_NN ਸਥਾਨ\N\_NN ਸੱਤ\QT\_QTC ਨਗਰਾਂ\N\_NN ਜਾਂ\CC\_CCD ਸਪਤਪੁਰੀਆਂ\N\_NN ਦੇ\N\_NN ਰੂਪ\N\_NN ਵਿਚ\N\_NN ਗਰੰਥਾਂ\N\_NN ਵਿਚ\N\_NN ਦਰਜ\N\_NN ਹਨ\N\_NN !\RD\_PUNC
5. ਇੰਝ\N\_NN ਕਿਹਾ\N\_NN ਗਿਆ\N\_NN ਹੈ\N\_NN ਕਿ\CC\_CCS ਚੌਥੇ\QT\_QTO ਮਹੀਨੇ\N\_NN ਵਿਚ\N\_NN ਇਨ੍ਹਾਂ\N\_NN ਸਪਤਪੁਰੀਆਂ\N\_NN ਦਾ\N\_NN ਦਰਸ਼ਨ\N\_NN ਮੋਖ\N\_NN ਪ੍ਰਦਾਨ\N\_NN ਵਾਲਾ\N\_NN ਹੁੰਦਾ\N\_NN ਹੈ\N\_NN !\RD\_PUNC

## Tamil

1. சப்தபுரிகளை\N\_NN தரிசிப்பதால்\V\_VM\_VNG முக்தி\N\_NN  
கிடைக்கிறது\V\_VM\_VF .\RD\_PUNC
2. இந்து\N\_NNP மதத்தில்\N\_NN புண்ணிய\J இடங்கள்\N\_NN மிகவும்\RP\_INTF  
சிறப்பு\N\_NN வாய்ந்தவை\N\_NN ஆகும்\V\_VAUX .\RD\_PUNC
3. ஒவ்வொரு\QT\_QTC புண்ணியத்தலமும்\N\_NN பெரியது\N\_NN  
மற்றும்\CC\_CCD முக்கியத்துவம்\N\_NN வாய்ந்தது\N\_NN ஆகும்\V\_VAUX  
ஆனால்\CC\_CCS ஏழு\QT\_QTC இடங்கள்\N\_NN மிக\RP\_INTF சிறப்பும்\N\_NN  
மதிப்பும்\N\_NN வாய்ந்ததாகும்\V\_VM\_VF .\RD\_PUNC
4. இந்த\DM\_DMD ஏழு\QT\_QTC புண்ணியத்தலங்கள்\N\_NN ஏழு\QT\_QTC  
நகரங்கள்\N\_NN அல்லது\CC\_CCD சப்தபுரிகள்\N\_NN என்று\CC\_CCS\_UT  
புத்தகங்களில்\N\_NN வர்ணிக்கப்பட்டு\N\_NN இருக்கின்றன\V\_VAUX  
.\RD\_PUNC
5. பெளர்ணமியில்\N\_NN இந்த\DM\_DMD சப்தபுரியின்\N\_NN தரிசனம்\N\_NN  
முக்தியை\N\_NN வழங்குகிறது\V\_VM\_VF என்று\CC\_CCS\_UT  
சொல்லப்பட்டு\N\_NN இருக்கிறது\V\_VAUX .\RD\_PUNC

## Malayalam

1. ഏഴ്\QT\_QTC പുണ്യനഗരികൾ\N\_NN സന്ദർശിക്കുന്നത്\V\_VM\_VNF  
കൊണ്ട്\RP\_RPD മോക്ഷം\N\_NN ലഭിക്കുന്നു\V\_VM\_VF .\RD\_PUNC
2. ഹിന്ദു\N\_NN മതത്തിൽ\N\_NN പുണ്യസ്ഥലങ്ങൾക്ക്\N\_NN വലിയ\J  
മഹത്വം\N\_NN ഉണ്ട്\V\_VAUX .\RD\_PUNC
3. എല്ലാ\QT\_QTF തീർത്ഥാടനസ്ഥലങ്ങളും\N\_NN വലുതും\J പ്രധാനപ്പെട്ടതും\J  
ആണ്\V\_VAUX ,\RD\_PUNC എങ്കിലും\CC\_CCD ഈ\DM\_DMD ഏഴ്\QT\_QTC  
സ്ഥലങ്ങൾക്കും\N\_NN വലിയ\J ശ്രേഷ്ഠതയും\N\_NN ആദരവും\N\_NN  
ഉണ്ട്\V\_VAUX .\RD\_PUNC
4. ഈ\DM\_DMD ഏഴ്\QT\_QTC ധർമ്മസ്ഥലങ്ങളും\N\_NN ഏഴ്\QT\_QTC  
പട്ടണങ്ങൾ\N\_NN അഥവാ\CC\_CCD ഏഴ്\QT\_QTC പുണ്യനഗരികൾ\N\_NN  
എന്ന\CC\_CCD രീതിയിൽ\N\_NN ഗ്രന്ഥങ്ങളിൽ\N\_NN വർണ്ണിച്ചിട്ടുണ്ട്\V\_VM\_VF  
.\RD\_PUNC
5. ചതുർമാസത്തിൽ\N\_NNP ഈ\DM\_DMD പുണ്യസ്ഥലങ്ങളുടെ\N\_NN  
സന്ദർശനം\N\_NN മോക്ഷപ്രദായകമാണെന്ന്\N\_NN പറഞ്ഞിട്ടുണ്ട്\V\_VM\_VF  
.\RD\_PUNC

## Bangla

1. સમ્પૂર્ણિ\N\_NNP દર્શન\N\_NN કરે\V\_VM\_VNF મોક્ષનાભ\N\_NN હય\V\_VAUX I\RD\_PUNC
2. હિન્દુ\N\_NN ધર્મ\N\_NN તીર્થેર\N\_NN ચથેજી મહર્ષ\N\_NN આજે\V\_VAUX I\RD\_PUNC
3. ચદિ\PSP સમ્પૂર્ણિ\QT\_QTF તીર્થ\N\_NN ચથેજી ગુરુશ્વર્ણ\N\_NN તા\PSP સાત\QT\_QTC જાગાર\N\_NN વિશેષ\જી ગુરુશ્વર્ણ\N\_NN ઓ\CC\_CCD મહર્ષ\N\_NN આજે\V\_VAUX I\RD\_PUNC
4. એ\DM\_DMD સાત\QT\_QTC ધર્મશ્લ\N\_NN સાત\QT\_QTC નગર\N\_NN વા\CC\_CCD સમ્પૂર્ણિ\N-NNP નામે\N\_NN પરિચિત\N\_NN I\RD\_PUNC
5. એ\DM\_DMD વળા\V\_VM\_VNG હય\V\_VAUX ચે\RP\_RPD ચર્તુદશીતે\N\_NN એ\DM\_DMD સમ્પૂર્ણિ\N-NNP દર્શન\N\_NN કરલે\V\_VM\_VNF મોક્ષનાભ\N\_NN હય\V\_VAUX I\RD\_PUNC

## Marathi

1. સાતપુરીચ્યા\N\_NNP દર્શનાને\N\_NN મિલતો\VM મોક્ષ\N\_NN .\PUNC
2. હિંદુ\જી ધર્મામધ્યે\N\_NN તીર્થક્ષેત્રાંચે\N\_NN ખૂપ\QT\_QTF મહત્ત્વ\N\_NN આહે\VM .\PUNC
3. તસે\PR તર\RP પ્રત્યેક\QT\_QTF તીર્થ\N\_NN મહત્ત્વાચે\N\_NN આણિ\C\_CCD મુખ્ય\જી આહે\VM પળ\C\_CCD સાત\Q-QTC સ્થાનાંચે\N\_NN મહત્ત્વ\N\_NN આણિ\C\_CCD માન્યતા\N\_NN મોઠી\જી આહે\VM .\PUNC
4. હી\DM સાત\Q-QTC ધર્મસ્થળે\N\_NN સાત\Q-QTC નગરી\N\_NN વા\C\_CCD સપ્તપુરીચ્યા\NNP રૂપાત\N\_NN ગ્રંથાંમધ્યે\N\_NN વર્ણિલેલી\જી આહેત\VM .\PUNC
5. અસે\PR મ્હટલે\VM ગેલે\VAUX આહે\VAUX કી\C\_CCD ચાતુર્માસામધ્યે\N\_NN ચા\C\_CCD સપ્તપુરીચે\NNP દર્શન\N\_NN મોક્ષ\N\_NN દેનારે\V\_VM\_VNF ઠરતે\VM .\PUNC

## Gujarati

1. સપ્તપુરીઓના\N\_NNP દર્શનથી\N\_NN મળે\V\_VM છે\V\_VAUX મોક્ષ\N\_NN
2. હિંદુ ધર્મમાં\N\_NN તીર્થનું\N\_NN ધણું\QT\_QTF મહત્ત્વ\જી છે\V\_VM
3. આમ\RP\_RPD તો\RP\_RPD દરેક\DM\_DMD તીર્થ\N\_NN મહાન\જી અને\CC\_CCD મહત્ત્વપૂર્ણ\જી છે\V\_VM પણ\CC\_CCD સાત\QT\_QTC સ્થાનોની\N\_NN મહત્તા\N\_NN અને\CC\_CCD માન્યતા\N\_NN છે\V\_VM
4. આ\DM\_DMD સાતેય\QT\_QTC ધર્મસ્થળ\N\_NN સાત\QT\_QTC નગર\N\_NN અથવા\CC\_CCD સપ્તપુરીઓના\N-NNP સ્વરૂપે\N\_NN ગ્રંથોમાં\N\_NN વર્ણવાયેલા\V\_VM છે\V\_VAUX
5. એમ\RP\_RPD કહેવાય\V\_VM કે\CC\_CCS ચાતુર્માસમાં\N\_NN આ\DM\_DMD સપ્તપુરીઓના\N-NNP દર્શન\N\_NN મોક્ષ\N\_NN આપનાર\V\_VM હોય\V\_VAUX છે\V\_VAUX



## Konkani

1. ससपुरीचें\N\_NNP दर्शन\N\_NN घेतल्यार\V\_VM\_VNF मोक्ष\N\_NN मेळटा\V\_VM\_VF.\RD\_PUNC
2. हिंदू\N\_NNP धर्मांत\N\_NN तिर्थस्थानाक\N\_NN व्हड\J म्हत्व\N\_NN आसा\V\_VM\_VF.\RD\_PUNC
3. तशें\RB पळोवपाक\V\_VM\_VNF गेल्यार\V\_VM\_VNF सगलींच\QT\_QTF तिर्था\N\_NN व्हड\J  
आनी\CC\_CCD खाशेली\J आसात\V\_VM\_VF पूण\CC\_CCS सात\QT\_QTC स्थळां\N\_NN व्हड\J  
आनी\CC\_CCD म्हत्वाचीं\J अशें\RB मानतात\V\_VM\_VF.\RD\_PUNC
4. ग्रंथांनी\N\_NN ह्या\DM\_DMD सातूय\QT\_QTC धर्मस्थळांचें\N\_NN वर्णन\N\_NN सात\QT\_QTC  
नगरी\N\_NN वा\CC\_CCD ससपुरी\N\_NNP अशें\RB आसा\V\_VM\_VF.\RD\_PUNC
5. चातुर्मासांत\N\_NN हे\PR\_PRP ससपुरीचें\N\_NNP दर्शन\N\_NN मोक्ष\N\_NN मेळोवन\V\_VM\_VNF  
दिवपी\N\_NN थारता\N\_NN अशें\RB मानतात\N\_NN.\RD\_PUNC

## Urdu

1. N\_NNP کی PSP زیارت N\_NN سے\ملتے\V\_VM ہے\N\_NN نجات\N\_NN
2. N\_NNP میں\میں\N\_NN تیرتھ\N\_NN کی\کی\QT\_QTF بڑی\QT\_QTF اہمیت\N\_NN ہے\N\_NN
3. N\_NNP ہیں\ہیں\N\_NN اہم\CC\_CCD اور\اور\N\_NN بڑی\N\_NN تیرتھ\QT\_QTF ہر\ہر\QT\_QTF اتو\N\_NNP  
،\RD\_PUNC کی\کی\QT\_QTF مقامات\N\_NN کی\کی\QT\_QTF بڑی\N\_NN عظمت\N\_NN اور\اور\CC\_CCD  
N\_NN مقبولیت\N\_NN ہے\N\_NN
4. N\_NNP شہروں\شہروں\QT\_QTC سات\N\_NN مقامات\مقامات\QT\_QTC ساتوں\PR\_PRP  
،\RD\_PUNC میں\میں\N\_NN کتابوں\N\_NNP میں\میں\N\_NN شکل\N\_NNP کی\کی\QT\_QTC ساتوں\QT\_QTC  
N\_NN ہیں\N\_NN
5. N\_NNP میں\میں\N\_NN برساتی\برساتی\CC\_CCD کہ\N\_NN ہے\N\_NN گیا\N\_NN کہ\DM\_DMD ایسا  
،\RD\_PUNC کی\کی\QT\_QTC شہروں\N\_NNP کی\کی\QT\_QTC ساتوں\QT\_QTC  
N\_NN ہے\N\_NN بڑی\N\_NN ہے\N\_NN بڑی\N\_NN ہے\N\_NN

## Oriya

1. ସପ୍ତପୁରୀଗୁଡ଼ିକ N\_NN ର PSP ଦର୍ଶନ NN ରୁ PSP ମୋକ୍ଷ NN ମିଳିଥାଏ N\_NNV ।
2. ହିନ୍ଦୁଧର୍ମ N\_NN ରେ PSP ଧର୍ମାନ୍ତ NN ର PSP ବଡ଼ JJ ମହତ୍ତ୍ୱ NN ଅଟେ V\_\_VAUX ।
3. ଏପରିକି RP\_\_RPD ସବୁ PR\_\_PRL ଧର୍ମାନ୍ତ NN ବଡ଼ JJ ଏବଂ CC\_\_CCD ମୁଖ୍ୟ JJ ଅଟନ୍ତି V\_\_VAUX, ଯେଉଁଠି  
CC\_\_CCS ସାତ QT\_\_QTC ସ୍ଥାନଗୁଡ଼ିକର N\_NN ଗୁରୁତ୍ୱେଠିକ୍ JJ ମହନୀୟତା N\_NN ଓ CC\_\_CCD  
ମାନସ୍ୱତା N\_NN ଅଟେ V\_\_VAUX ।
4. ଏହି PR ସାତ QT\_\_QTC ଧର୍ମସ୍ଥଳ NN ସାତ QT\_\_QTC ନଗରଗୁଡ଼ିକ N\_NN ର PSP କିଂବା CC\_\_CCD  
ସପ୍ତପୁରୀଗୁଡ଼ିକ N\_NN ର PSP ରୂପ JJ ରେ PSP ଗର୍ବଧର୍ମଗୁଡ଼ିକ N\_NN ରେ PSP ବର୍ଣ୍ଣିତ N\_NNV ହେଉଅଛି  
V\_\_VAUX ।
5. ଏଭଳି PR କୁହାଯାଇ V\_\_VM ଅଛି V\_\_VAUX କି CC\_\_CCS ଚରୁତ୍ରମାସ N\_NN ରେ PSP ଏହି PR  
ସପ୍ତପୁରୀଗୁଡ଼ିକ N\_NN ର PSP ଦର୍ଶନ NN ମୋକ୍ଷ NN ପ୍ରଦାନ V\_\_VAUX କରିବାବାଲା NN ହେଉଥାଏ  
V\_\_VAUX ।

## 16. REFERENCE

1. ISO 12620:1999, *Terminology and other language and content resources — Specification of data categories and management of a Data Category Registry for language resources*
2. XML Schema Requirements: <http://www.w3.org/TR/1999/NOTE-xml-schema-req-19990215>
3. Best Practices for XML Internationalization: <http://www.w3.org/TR/xml-i18n-bp/>
4. Internationalization Tag Set (ITS) Version 1.0: <http://www.w3.org/TR/2007/REC-its-20070403/>
5. ISO 639-3, Language Codes: <http://www.sil.org/iso639-3/codes.asp>

## ANNEXURE-1

### LANGUAGE TAGS

S.No.	Language Name	Language Tags according to ISO 639-3
1	Hindi	asm
2	Assamese	ben
3	Bangla	brx
4	Bodo	doi
5	Dogri	guj
6	Gujarati	hin
7	Kannada	kan
8	Kashmiri	kas
9	Konkani	kok
10	Maithili	mai
11	Malayalam	mal
12	Manupuri	mni
13	Marathi	mar
14	Nepali	nep
15	Oriya	ori
16	Punjabi	pan
17	Sanskrit	san
18	Santhali	sat
19	Sindhi	snd
20	Tamil	tam
21	Telugu	tel
22	Urdu	urd

## CONTRIBUTERS

- |                                  |   |
|----------------------------------|---|
| 1. Ms. Swaran Lata,              | Department of Information Technology, New Delhi |
| 2. Prof. Girish Nath Jha,        | JNU, New Delhi                                  |
| 3. Dr. Somnath Chandra,          | Department of Information Technology, New Delhi |
| 4. Dipti Misra Sharma,           | LTRC, IIIT-H                                    |
| 5. Somi Ram                      | CDAC, NOIDA                                     |
| 6. Prof. Uma Maheswara Rao G,    | University of Hyderabad                         |
| 7. Dr. Sobha L,                  | AU-KBC, Chennai                                 |
| 8. Menak. S,                     |   |
| 9. Kalika Bali,                  | Microsoft, Bangalore                            |
| 10. Prof. Pushpak Bhattacharyya, | IIT-Bombay                                      |
| 11. Prof. Malhar Kulkarni,       | IIT-Bombay                                      |
| 12. Lata Popale,                 | IIT-Bombay                                      |
| 13. Kirtida Shah,                | Gujarati University, Ahemadabad                 |
| 14. Mona Parakh,                 | LDCIL, Mysore                                   |
| 15. Jyoti Pawar,                 | Goa University                                  |
| 16. Madhavi Sardesai,            | Goa University                                  |
| 17. Ramnath,                     |   |
| 18. Aadil Kak,                   | University of Kashmir                           |
| 19. Nazima,                      | University of Kashmir                           |
| 20. Dr. Richa,                   | LDCIL, Mysore                                   |
| 21. Mazhar Mehdi Hussain,        | JNU, New Delhi                                  |
| 22. Mr. Prashant Verma,          | W3C India, New Delhi                            |
| 23. Swati Arora,                 | W3C India, New Delhi                            |