

7 Addendum A: Description of Adjusted HCS Tagset

This addendum describes the POS tags featured in the training set used for the experiments described in this paper. Description is mostly reproduced from [9], but some parts of speech have been adjusted from the original version of HCS.

Tag	Description	Example
Basic Tags		
A-INFL	inflecting adjective	ndogondogo
A-UINFL	uninflecting adjective	tofautitofauti
ABBR	abbreviated word form	Bw.
AD-ADJ	adjective qualifier	kabisa
ADJ	adjective	kamili
ADV	adverb	leo
AG-PART	agentive particle	na
AR	Arabic Origin	mashtakiwa
CAP	Capitalized word form	Akina
CC	co-ordinating conjunction	na
CONJ	conjunction	kwamba
DEF-V:ni	uninflected verb	ni
DEM:hV	demonstrative pronoun of type hV(C)V	hichi
EMPH	emphatic form	ndiye
ENG	English Origin	constitution
EXCLAM	exclamation	eti
GEN-CON	genitive connector	la
IDIOM	idiom	kura
IMP	imperative	twende
INTERROG	interrogative	kwa_nini
N	noun	bomu
NA-POSS	possessive particle	na
NEG	negative	hakuna
NUM	numeral	26,1999
PREP	preposition	katikati_ya
PRON	pronoun	ile
PROPNAME	proper name	Kamau
REL	relative marker	ambaye
RHET	rhetorical	sio
SELFSTANDING	selfstanding subject particle	M
V	verb	laitikisa

Tag	Description	Example
Noun/verb class concords		
1/2-SG-SP	singular noun (class 1,2) subject prefix	ana
1/2-PL-SP	plural noun (class 1,2) subject prefix	walio
3/4-PL	plural noun (class 3,4) prefix	mioto
3/4-SG-SP	singular noun (class 3,4) subject prefix	uliopo
3/4-PL-SP	plural noun (class 3,4) subject prefix	iliyoko
5/6-SG-SP	singular noun (class 5,6) subject prefix	lipo
5/6-PL-SP	plural noun (class 5,6) subject prefix	ya
7/8-SG-SP	singular noun (class 7,8) subject prefix	kilio
7/8-PL-SP	plural noun (class 7,8) subject prefix	vilivyo
9/10-SG-SP	singular noun (class 9,10) subject prefix	iko
9/10-PL-SP	plural noun (class 9,10) subject prefix	zilizomo
11-SG	singular noun (class 11) prefix	uwakilishi
11-SG-SP	singular noun (class 11) subject prefix	ulipo
15-SG-SP	singular noun (class 15) subject prefix	kusiko
16-SG-SP	singular noun (class 16) subject prefix	pasipo
17-SG-SP	singular noun (class 17) subject prefix	kuna
18-SG-SP	singular noun (class 18) subject prefix	mna
SG1-SP	1st person singular subject prefix	nina
SG2-SP	2nd person singular subject prefix	ulipo
SG3-SP	3rd person singular subject prefix	yuna
PL1-SP	1st person plural subject prefix	tunalo
PL2-SP	2nd person plural subject prefix	mna
PL3-SP	3rd person plural subject prefix	wanasoma

Tag	Description
Punctuation	
EQUAL-MARK	=
EXCLAMATION	!
FULL-STOP	.
HYPHEN	-
LEFT-PARENTHESIS	(
RIGHT-PARENTHESIS)
LEFT-SQUAREBRACKET	[
RIGHT-SQUAREBRACKET]
SINGLE-QUOTE	'
DOUBLE-QUOTE	”
COMMA	,
DOLLAR-SIGN	\$
PERCENT-MARK	%
SEMICOLON	;
COLON	:
QUESTION-MARK	?