

Morphological Segmentation
Annotation Guidelines
Appendix: Swahili
Linguistic Data Consortium

© [2018] Trustees of the University of Pennsylvania

*****This document is unpublished and intended solely for the use of the individual or entity to whom it was delivered. Redistribution is strictly prohibited without the express authorization of the Linguistic Data Consortium.*****

This document contains examples of common linguistic patterns in Swahili that are relevant to the Morphological Segmentation task. They represent classes of morphological patterns in the language that could be annotated in the task. Note that these examples do not necessarily indicate what the annotation should be.

Table 4.1: Swahili noun classes

Class	Noun prefix	Adjective prefix	Pronominal prefix	Example	
Applies to:	nouns	adjectives, numerals	demonstratives, relatives, possessives, verbs		
1	m-, mw-	m-, mw-	yu-, ye-, w-, a-	mtoto mdogo yule	'that small child'
2	wa-, w-	wa-	wa-	watoto wadogo wale	'those small children'
3	m-, mw-	m-, mw-, (mu-)	u-	mti mdogo ule	'that small tree'
4	mi-	mi-	i-	miti midogo ile	'those small trees'
5	ji-, Ø-	ji-, Ø-	li-	jino dogo lile	'that small tooth'
6	ma-, m-	ma-	ya-	meno madogo yale	'those small teeth'
7	ki-, ch-	ki-, ch-	ki-	kiatu kidogo kile	'that small shoe'
8	vi-, vy-	vi-, vy-	vi-	viatu vidogo vile	'those small shoes'
9	N-, Ø-	N-	i-	ndoo ndogo ile	'that small bucket'
10	N-, Ø-	N-	zi-	ndoo ndogo zile	'those small buckets'
11/14	u-, w-	u-, w-	u-	ufagio mdogo ule	'that small broom'
15	ku-, kw-	ku-, kw-	ku-	kutembea kudogo kule	'that little bit of walking'
16	—	pa-, p-	pa-	mahali pazuri pale	'that nice spot'
17	—	ku-, kw-	ku-	mahali kuzuri kule	'that nice (general) area'
18	—	mu-, mw-	mu-, m-	mahali muzuri mle	'that nice area (inside)'

Underived noun	Diminutive	
mlima 'mountain'	kilima 'hill'	
kombe 'platter'	kikombe 'cup'	
mji 'town'	kijiji 'village'	
mtoto 'child'	kitoto 'infant'	kijitoto 'very small infant'
jiko 'kitchen, stove'	kijiko 'spoon'	
mkono 'person'	kikono 'person with an impaired arm'	
-pofu 'blind'	kipofu 'blind person'	
kwapa 'armpit'	kikwapa 'armpit odor'	
duka 'shop'	kiduka 'small shop'	kijiduka 'small, insignificant shop'

Table 4.4: Diminutives (class 7 or 8)

Underived noun	Augmentative
mto 'river'	jito 'large river'
nyumba 'house'	jumba 'large house'
nyoka 'snake'	joka 'large snake'
mtu 'person'	jitu 'large, strong person'

Table 4.5: Augmentatives (class 5 or 6)

Morphology

Stem	Class 1	Class 11/14
-winda 'hunt'	mwindaji 'hunter'	uwindaji 'hunting'
-omba 'ask, beg, pray for'	mwombaji 'professional beggar'	uombaji 'begging'
-shona 'sew'	mshonaji 'tailor'	ushonaji 'tailoring'
-cheza 'play'	mchezaji 'athlete, actor'	uchezaji 'style of playing'
-pasua 'operate'	mpasuaji 'surgeon'	upasujaji 'surgery'
-pigana 'fight (hit each other)'	mpiganaji 'fighter'	upiganaji 'fighting'

Table 4.7: -aji derivation

Stem	Class 1	Class 3	Class 5/6	Class 7	Class 10	Class 11/14
-sikia 'hear'			(ma)sikio 'ear(s)'			
-unda 'construct'					nyundo 'hammer'	
-andika 'write'		mwandiko 'handwriting'	andiko 'document'			
-pima 'measure'				kipimo 'measure(ment)'		
-penda 'love'		mpendapendapo 'shrub used for making love potions'				upendo 'love'
-umba 'create'			(ma)umbo 'shape(s), structure(s)'			
-lipa 'pay'			malipo 'payment'			
-lipwa 'be paid'	mlipwa 'payee'					

Table 4.8: -o derivation

Stem	Class 1	Class 11/14
-tulia 'be quiet'	mtulivu 'a gentle person'	utulivu 'calmness'
-kosa 'err, offend, sin, miss'	mkosefu 'guilty person'	ukosefu 'lack, deficiency'
-chelewa 'be late'		uchelevu 'lateness'
-potea 'be lost'	mpotevu 'wasteful person, prodigal'	upotevu 'waste, prodigality'
-angalia 'watch, be careful'	mwangalifu 'a careful person'	wizi 'theft'
-amini 'trust, have faith in'	mwaminifu 'trustworthy person'	uaminifu 'honesty, reliability'

Table 4.9: -fu/-vu derivation

Noun 1	Noun 2	Compound
mwana 'child'	sheria 'law'	mwanasheria 'lawyer'
mwana 'child'	isimu 'linguistics'	mwanaisimu 'linguist'
mwana 'child'	nchi 'country'	mwananchi 'citizen'
hali 'state'	joto 'heat'	halijoto 'temperature'
punda 'donkey'	milia 'stripes'	punda milia 'zebra'

Table 4.11: Swahili noun-noun compounds

Noun	Modifier	Compound
pembe 'corner'	-tatu 'three'	pembetatu 'triangle'
mja 'human being'	-zito 'heavy'	mjamzito 'pregnant woman'
asili 'source'	mia 'hundred'	asilimia 'percent; percentage'

Table 4.12: Swahili noun-modifier compounds

Noun 1	Noun 2	Compound
mwana 'child'	sheria 'law'	mwanasheria 'lawyer'
mwana 'child'	isimu 'linguistics'	mwanaisimu 'linguist'
mwana 'child'	nchi 'country'	mwananchi 'citizen'
hali 'state'	joto 'heat'	halijoto 'temperature'
punda 'donkey'	milia 'stripes'	punda milia 'zebra'

Table 4.11: Swahili noun-noun compounds

Noun	Modifier	Compound
pembe 'corner'	-tatu 'three'	pembetatu 'triangle'
mja 'human being'	-zito 'heavy'	mjamzito 'pregnant woman'
asili 'source'	mia 'hundred'	asilimia 'percent; percentage'

Table 4.12: Swahili noun-modifier compounds

Verb	Noun	Compound
-fanya 'do'	biashara 'business'	mfanyabiashara 'businessperson'
-sema 'speak'	kweli 'true'	msemakweli 'honest person'
-shona 'sew'	viatu 'shoes'	mshonaviatu 'cobbler'
-piga 'hit'	picha 'pictures'	mpigapicha 'photographer'
-fungua 'open'	mimba 'pregnancy'	kifunguamimba 'first-born child'
-paza 'increase'	sauti 'sound'	kipazasauti 'loudspeaker'
-fyonza 'suck up'	vumbi 'dust'	kifyonzavumbi 'vacuum cleaner'
-amsha 'wake up (s.o.)'	kinywa 'mouth'	kifunguakinywa 'breakfast'

Table 4.13: Swahili verb-noun compounds

	Subject markers	Object markers	Free-standing pronouns
1SG	ni-	ni-	mimi
2SG	u-	ku-	wewe
3SG	a-	m-	yeye
1PL	tu-	tu-	sisi
2PL	m-	wa-	ninyi/nyinyi
3PL	wa-	wa-	wao
Reflexive		ji-	

Table 4.14: Personal pronouns

Class	verbal prefix	amba-	verbal suffix
1	-ye-	ambaye	-ye
2	-o-	ambao	-o
3	-o-	ambao	-o
4	-yo-	ambayo	-yo
5	-lo-	ambalo	-lo
6	-yo-	ambayo	-yo
7	-cho-	ambacho	-cho
8	-vyo-	ambavyo	-vyo
9	-yo-	ambayo	-yo
10	-zo-	ambazo	-zo
11/14	-o-	ambao	-o
15	-ko-	ambako	-ko
16	-po-	ambapo	-po
17	-ko-	ambako	-ko
18	-mo-	ambamo	-mo

27

Table 4.28: Swahili verb template²

Neg.	Subj.	TAM	Rel.	Obj.	Root	Ext.	FV	Rel.
ha-	tu-	ta-		ku-	chok	-esh	-a	'we will not tire you out'
	ni-				chek		-a	-ye 'I who laugh'
	a-	li-	ye-	wa-	fund	-esh	-a	'(one) who taught them'
	u-			mw-	on		-e	'you may see him/her'

Table 4.30: Object markers

Person/Class	Object marker	Example
1SG	ni-	aliniona 's/he saw me'
2SG	ku-	alikuona 's/he saw you (sg.)'
3SG/Class 1	m(w)-	alimwona 's/he saw him/her'
1PL	tu-	alituona 's/he saw us'
2PL	wa-	aliwaona 's/he saw you (pl.)'
3PL/Class 2	wa-	aliwaona 's/he saw them'
3	u-	aliuona 's/he saw it'
4	i-	aliiiona 's/he saw them'
5	li-	aliliona 's/he saw it'
6	ya-	aliyaona 's/he saw them'
7	ki-	alikiona 's/he saw it'
8	vi-	aliviona 's/he saw them'
9	i-	aliiiona 's/he saw it'
10	zi-	aliziona 's/he saw them'
11/14	u-	aliuona 's/he saw it'
15	ku-	alikuona 's/he saw it'
16	pa-	alipaona 's/he saw it'
17	ku-	alikuona 's/he saw it'
18	m(w)-	alimwona 's/he saw it'

Tense/Aspect	Form	Description	Example
Past	li-	an action or state which took place in the past	nilitembea 'I walked' ulitembea 'you walked'
Perfect	me-	action/state that took place in recent past and still has some effects on present; also used to describe a current state	nimetembea 'I have walked' umetembea 'you have walked'
Already	mesha- mekwisha-	emphasizes the recency of an event; often translated with "already" or "just"	nimeshatembea 'I have already/just walked' umekwishatembea 'you have already/just walked'
Present (Definite)	na-	action or state of being that is ongoing or takes place at a point in time; usually refers to present	ninatembea 'I am walking' unatembea 'you are walking'
Indefinite time	a-	statement of fact and question asked without relation to any particular point in time; often used as simple present	natembea 'I walk' watembea 'you walk'
Future	ta-	an action or state which will take place in the future	nitatembea 'I will walk' utatembea 'you will walk'
Habitual	hu-	an action or state which will take place habitually or recurrently; uses no subject marker	(mimi) hutembea 'I regularly walk' (wewe) hutembea 'you regularly walk'
Consecutive	ka-	an action or state which follows another action; or the first event which took place in a series of multiple events	nikatembea 'and then I walked' ukatembea 'and then you walked'

Table 4.32: Swahili conditionals

Conditional	Form	Description	Example
Factual conditional	kama	an event which is highly likely to take place (or “virtually realized”)	kama ukimwona ‘if you see him/her’
	ikiwa		ikiwa amefika ‘if s/he has arrived’
Probable conditional	ki-	a condition which may be fulfilled in the future; can also be used to indicate simultaneity of two actions; in the future, may refer to an event which will regularly take place	ukimwona ‘if you see him/her’ akifika ‘if s/he arrives’
Improbable conditional	nge-	a condition which hypothetically could be fulfilled in the present; also used to express polite requests	ungemwona ‘if you saw him/her’ angefika ‘if he arrived’
Unreal conditional	ngali-	an unrealistic condition that would have been fulfilled in the past (= past conditional)	ungalimwona ‘if you had seen him’ angalifika ‘if s/he had arrived’

Table 4.34: Negative subject markers

Person/ Class	Subject marker	Neg. subj.	Example (with present definite)
1SG	ni-	si-	sibaki 'I do not remain'
2SG	u-	hu-	hubaki 'you (sg.) do not remain'
3SG/Class 1	a-	ha-	habaki 's/he does not remain'
1PL	tu-	hatu-	hatubaki 'we do not remain'
2PL	m(w)-	ham-	hambaki 'you (pl.) do not remain'
3PL/Class 2	wa-	hawa-	hawabaki 'they do not remain'
3	u-	hau-	mmea hubaki 'the plant does not remain'
4	i-	hai-	mimea haibaki 'the plants do not remain'
5	li-	hali-	jengo halibaki 'the building does not remain'
6	ya-	haya-	majengo hayabaki 'the buildings do not remain'
7	ki-	haki-	kiti hakibaki 'the chair does not remain'
8	vi-	havi-	viti havibaki 'the chairs do not remain'
9	i-	hai-	nyumba haibaki 'the house does not remain'
10	zi-	hazi-	nyumba hazibaki 'the houses do not remain'
11/14	u-	hau-	ukuta haubaki 'the wall does not remain'
15	ku-	haku-	kula hakubaki 'eating does not remain'
	ku-	kuto-	kutobaki 'to not remain'
16	pa-	hapa-	hapabaki ('a place) does not remain'
17	ku-	haku-	hakubaki '(a general place) does not remain'
18	m(w)-	ham-	hambaki '(a place inside) does not remain'

Extension	Form	Description	Example	Example Translation
Passive	-w	passivizes the verb; the object becomes the subject, and the subject (agent) is not expressed	kuandika	'to write'
			kuandikwa	'to be written'
			kulipa	'to pay'
			kulipwa	'to be paid'
Stative	-ik/-ek	describes a state; usually does not say who caused the state to come about	kuvunja	'to break'
			kuvunjika	'to be broken'
			kuchoma	'to burn'
			kuchomeka	'to be burnt'
Causative	-ish/ -esh/ -iz/ -ez	changes the meaning of a verb to indicate that someone caused an action to take place	kuingia	'to enter'
			kuingiza	'to insert'
			kula	'to eat'
			kulisha	'to feed'
Applicative	-i/-e	adds a participant, and suggests that the action is done to, for, or about that participant	kusoma	'to read'
			kusomea	'to read for/to'
			kukata	'to cut'
			kukatia	'to cut with/for'
Reciprocal	-an	describes a reciprocated action, where the action is done to one another	kupiga	'to hit'
			kupigana	'to hit each other'
			kusaidia	'to help'
			kusaidiana	'to help each other'
Reversive	-u	changes the meaning of a verb so that its action is undone	kufunga	'to close'
			kufungua	'to open'
			kuzika	'to bury'