

Book Review

KAMUSI YA KISWAHILI SANIFU. TOLEO LA PILI. (A STANDARD SWAHILI-SWAHILI DICTIONARY. SECOND EDITION).

Nairobi: Oxford University Press. 2004. 477 pp. ISBN 0195732227.

Reviewed by Arvi Hurskainen
University of Helsinki, Finland

The introduction to the second edition of *Kamusi ya Kiswahili Sanifu* (hence referred to as *Kamusi II*) states that there are about 2,000 such new words which were not included in the first edition of 1981 (hence referred to as *Kamusi I*). These new words have been found in Swahili newspapers and books of various genres from Kenya and Tanzania. Also domain-specific vocabularies standardised by *Baraza la Kiswahili la Taifa (BAKITA)* as well as other Swahili dictionaries are said to have contributed to the new edition.

A large number of people, about fifty, have participated in the compilation of *Kamusi II* in various roles, excluding the technical editing process. However, the team, which did the actual job, was much smaller. The dictionary was compiled at the Institute of Kiswahili Research, University of Dar-es-Salaam.

Kamusi II contains, no doubt, a number of improvements to the first edition of 1981. These include a spacious layout, which is very easy to read. The solution to print the head-words with a different colour adds to clarity. There are also illustrations as well as charts for describing inflectional paradigms and other grammatical data. An excellent addition is the comprehensive list of names of countries and nationalities at the end of the book.

There are also grammatical charts, but they are oddly located in various places, apparently according to the first occurrence of a word with such information. It would have been much better to place them all in a grammatical appendix and make references to it in each relevant head-word.

My main motivation to write this review was to test how well the compilers have managed in their main target, i.e. in including the frequently occurring words missing in *Kamusi I*. For this purpose I used the computational testing method applied in testing *Kamusi I* (Hurskainen 1994) and four other Swahili dictionaries (Hurskainen 2002, 2004a). Because an electronic version of *Kamusi II* was not available, it was not possible to construct an evaluation program particularly for this edition. Instead I used the program designed for *Kamusi I*.

The testing program is a morphological analyser, which imitates the dictionary as closely as possible. It recognises in text only those words, which are in the dictionary as dictionary entries. All other words in text are marked as unknown. Also such word-forms that do not have in the dictionary sufficient morphological information are marked as unknown.

The dictionary was tested with Helsinki Corpus of Swahili (hence HCS), which contains more than 12 million words of Standard Swahili, mostly news texts and some fiction from Tanzania. The corpus represents broadly the currently used Swahili and suits well for the current testing purposes.

Because the testing program was designed on the basis of Kamusi I, all such words that were not found as key-words in Kamusi I were retrieved from the corpus. It was expected that these words (hence Missing Words), missing in Kamusi I, were added to Kamusi II and would appear among the 2,000 new words.

The full list of new words (hence New Words) in Kamusi II was compiled by manually comparing the two editions. It was found that Kamusi II contains 1,270 such new words that are not in Kamusi I. The number is much lower than expected. Part of the explanation is that many such verbs, which in Kamusi I were only in the base form as dictionary entries, while the extended forms were given as possible endings without any definition or explanation, now in Kamusi II were as separate key-words with corresponding definitions. If all these extended verbs are also counted among the new words, then the total number of new words will be close to 2,000.

How well did the dictionary compilers succeed in bringing the dictionary up to date? In a dictionary with limited space a selection strategy is important. The use value of a dictionary is measured by the dictionary user. Therefore, the user expects to find in the dictionary those words that (s)he needs. From this viewpoint the frequency of the word is an important selection principle. In the ideal case, the 1,270 new words would be among the most frequent words missing in Kamusi I. When the two lists, i.e. the Missing Words and New Words, were compared, it was found that only 778 (Appendix 1) of the added words were found in the corpus. A total of 462 (Appendix 2) added words did not occur even once in the corpus.

At the same time HCS contains a large number of words, even frequent ones, which are not found in Kamusi II. This was tested by first retrieving from Missing Words all such words that were recognised by SALAMA, Swahili Language Manager (Hurskainen 1999, 2004b). By this method all misspelled Swahili words as well as non-Swahili words were eliminated, and the remaining list contained only Standard Swahili words. Then all such words were removed that are in Kamusi II. The resulting list (hence Missing Swahili Words) contains such Swahili words that are found in HCS but neither in Kamusi I nor Kamusi II. In all, this list contains over 2,500 words with more than 5,000 word-forms. Part of these words are listed in Appendices 3 and 4.

Appendix 3 contains words that are not listed in Kamusi II in any form. We see that many of them are quite frequent in the language and deserve to be listed in the dictionary. Part of these missing words are less obligatory in a dictionary with limited space. To this group belong the nouns, which are formed with the prefix *mwana-/wana* (e.g. *mwanasiasa* 'politician', *wanasiasa* 'politicians'). This is a productive method of coining new words and it is not possible to maintain a

full list of all coined nouns. Kamusi II lists 39 *mwana*-initial nouns, Langenscheidts Hanwörterbuch Swahili-Deutsch (Höftmann and Herms 2000) has 63 of them, and SALAMA lists 215 of such nouns. HCS has also 215 *mwana*-initial nouns, here singular and plural forms are counted separately

Appendix 4 contains such missing words that are listed in Kamusi II in some form but not in the form shown in the list. In Kamusi II there are dictionary entries, for which the word is listed but with insufficient morphological information. For example, the noun class affiliation of a noun is missing or defective. On the basis of the base form of many borrowed or coined nouns it is not possible to determine the class affiliation of the noun. As is well known, borrowed nouns may be located in class 5/6 or 9/10, and animate nouns in class 9/6. A number of borrowed nouns have also a double affiliation; they can be located in classes 5/6 or 9/10. Such possibility is usually not indicated. This problem was in Kamusi I and it prevails also in Kamusi II. Also the nouns of classes 11/6 and 11/10 are defectively described, often so that only the singular (i.e. class 11) is given but no plural. On the basis of the singular form it is not possible to determine what the plural should be.

The test shows that dictionary compilation using traditional methods is hardly ever accurate enough to stand critical examination. The dictionary contains words that are rarely used and at the same time frequent words are missing.

It has to be admitted that Kamusi II has vastly improved from Kamusi I, but the basic problems of dictionary compilation are still there. Most of these problems would have been avoided, if the selection of head-words had been done on the basis of a frequency list of Standard Swahili and sufficient morphological information had been provided, especially in noun class affiliation.

REFERENCES

Hurskainen, A. 1994.

Kamusi ya Kiswahili Sanifu in test: A computer system for analyzing dictionaries and for retrieving lexical data. **Afrikanistische Arbeitspapiere** 37: 169–179.

1999 *SALAMA: Swahili language manager.* **Nordic Journal of African Studies** 8(2): 139–157. Available also online at: <http://www.njas.helsinki.fi>.

2002 *Tathmini ya Kamusi Tano ya Kiswahili (Computer Evaluation of Five Swahili Dictionaries).* **Nordic Journal of African Studies** 11(2): 283–300. Available also online at: <http://www.njas.helsinki.fi>.

2004a Computational testing of five Swahili dictionaries. *Proceedings of the 20th Scandinavian Conference of Linguistics, Helsinki, 7–9.1. 2004.* Available online at: <http://www.ling.helsinki.fi/kielitiede/20scl/proceedings.shtml>.

2004b *Swahili Language Manager: A Storehouse for Developing Multiple Computational Applications*. **Nordic Journal of African Studies** 13(3): 363–397. Available online at: <http://www.njas.helsinki.fi>.

Höftmann, H. and Herms, I. 2000.

Langenscheidts Handwörterbuch Swahili-Deutsch. Berlin: Langenscheidt.

APPENDIX 1

New words in Kamusi II that were found in Helsinki Corpus of Swahili. The frequency of each word is shown.

10837	mbalimbali	986	mlalamikaji	526	vilevile
8782	kwani	942	kuwania	505	nyama
6032	uongozi	924	kueleza	500	utumishi
4330	kusini	911	tiketi	494	asasi
4068	nyumbani	894	vinginevyo	486	kikristo
3758	ligi	884	kondomu	483	usafirishaji
3656	mauaji	869	treni	472	raundi
2957	sekta	864	mwakani	464	wodi
2882	kaskazini	859	nccr-mageuzi	458	kukosekana
2586	mafunzo	842	mahojiano	457	mauzo
2389	ugaidi	824	mwakilishi	448	lori
2006	mawasiliano	809	uzinduzi	445	ufunguzi
1924	umuhimu	809	menejimenti	443	shutuma
1830	ajira	794	kulevyo	443	kujihusisha
1823	utekelezaji	788	kitengo	426	nyeti
1749	mtuhumiwa	775	mkakati	422	kuhamasisha
1658	katoliki	773	matatu	420	mafundisho
1650	kuboresha	748	posho	413	upatikanaji
1647	teknolojia	741	mwenendo	408	albamu
1599	matibabu	735	mfanyakazi	407	kujihami
1587	ushauri	733	kusitisha	403	uuzaji
1556	maarufu	728	mwelekeo	403	umwagaji
1460	kupatikana	726	ubinafsishaji	402	uhalali
1448	urais	714	mvutano	401	mwekezaji
1442	mashambulizi	707	mpito	398	utumiaji
1359	marekebicho	701	ushindani	393	udhibiti
1324	mawazo	672	pingamizi	389	lita
1302	pointi	671	kujiuzulu	385	kinyang'anyiro
1296	maandalizi	651	mstaafu	379	kuwekeza
1278	kutangaza	636	tishio	373	malumbano
1267	uzalishaji	628	uendeshaji	360	mantiki
1251	pengine	626	uvamizi	355	isitoshe
1216	mtandao	626	shinikizo	351	kukiri
1207	changamoto	619	dayosisi	351	burudani
1197	mageuzi	616	azma	342	uwakilishi
1191	mfanyabiashara	615	ukarabati	322	video
1188	majeruhi	586	kipaumbele	321	kigogo
1088	operesheni	583	matamshi	316	seminari
1078	awamu	548	maovu	313	uenyekiti
1065	maelfu	544	mshikamano	311	paroko
1047	daladala	537	kuwasiliana	311	lango
1013	majadiliano	529	akaunti	309	upasuaji

307	utangazaji	185	dampo	109	krismasi
306	ubinafsi	183	mjamzito	105	oili
294	mwanamuziki	182	hekta	105	kada
282	maliasili	176	utaifa	101	mwendesha_
277	kutokana	175	uandikishaji		mashtaka
275	ukabila	172	kombora	101	kujiona
273	ishirini	169	kusherehekea	100	uvutaji
270	fatuma	168	pembejeo	100	umwagiliaji
266	kujieleza	167	uwajibikaji	98	machungu
262	kujitenga	167	kinyemela	96	miundo
261	mwongozo	165	mwamko	96	maudhui
258	makwao	161	mamluki	96	kabumbu
256	usambazaji	158	malighafi	95	seneti
254	waziwazi	154	majungu	95	kuwakilisha
250	netiboli	153	ujio	94	kwacha
249	masafa	148	motisha	93	trafiki
247	mswada	148	chaguo	93	kukosoa
247	msukumo	145	kujifanya	93	kasheshe
246	mahubiri	145	kujiamini	92	mtumiaji
234	vitimbi	139	kivutio	91	kustawisha
230	kushinikiza	138	tovuti	90	mwendeshaji
228	brigedia	138	studio	89	uwakala
227	uhamisho	137	rais_mstaafu	89	themanini
222	nyanda	136	theluthi	88	ufundishaji
218	mwanajeshi	136	kisanga	86	mchuano
218	mhusika	135	kliniki	84	kutosheleza
215	mabano	134	mzaliwa	84	disko
214	nyenzo	132	pembezeni	83	ulaghai
214	lishe	132	nusura	83	takrima
212	maelewano	131	nyuzi	83	megawati
212	kutayarisha	130	nyayo	83	kuumwa
211	kujulikana	130	mwanadiplomasia	82	usaili
205	makumbusho	130	mhudumu	82	mwinjilisti
205	dhehebu	130	kupumzika	82	kuzidiwa
204	olimpiki	130	kansa	80	uenezi
204	mapokezi	126	kutofautiana	79	mkatoliki
204	chafu	123	kigango	77	polepole
203	ombaomba	122	kidete	77	kontena
199	utaalamu	121	uhandisi	76	miondoko
195	matayarisho	121	stendi	74	usindikaji
194	kurahisisha	118	kujisaidia	74	kushindikana
192	mkandarasi	117	krosi	74	hisabati
190	ugenini	115	kuania	72	diplomasia
188	malazi	114	udiwani	71	uchakavu
188	kurudia	112	nyanya	70	mbingu
187	mpangilio	111	mlemavu	70	kimeta
186	ukeketaji	110	transfoma	69	usanii
186	kikomunisti	109	uingizaji	68	mwajiriwa

68	msamaria	39	tapeli	27	utu_uzima
68	geti	39	patashika	27	ngurumo
67	mgawo	38	mwanakijiji	27	mwanamgambo
66	mwanariadha	38	maonevu	27	makuzi
64	veto	38	kulisha	26	rasta
64	seneta	37	mesenja	26	polio
63	juisi	37	kisirisiri	26	ovyoovyo
62	utapiamlo	37	fizikia	26	mikiki
62	mkereketwa	37	elektroniki	26	kujuana
62	kilima	36	serena	25	supu
62	chembechembe	36	mlundikano	25	oda
60	uimarishaji	36	mlalahoi	25	marathoni
60	maada	36	maombezi	25	kurundikana
60	kurudiana	35	ulamaa	25	dezo
60	kizamani	35	tungo	24	silika
58	kurugenzi	35	mchakato	24	mbia
57	tela	35	maongozi	24	kriketi
57	mdhibiti	35	kusambaratisha	23	vvu
55	saluni	35	kujipendekeza	23	komoni
54	vocha	35	bwalo	23	kimwana
53	zoni	34	mwovu	23	kifurushi
53	zebaki	34	masumbwi	22	maraha
53	biolojia	33	sketi	22	kuzurura
52	uzururaji	33	masalia	21	mtazamaji
52	bantam	33	makaa	21	mfadhaiko
51	kisoshalisti	33	kujidai	21	maudhi
49	mwanafalsafa	32	upokeaji	21	kreti
49	kandambili	32	uchoraji	21	alizeti
47	wahenga	32	mkanganyiko	20	mlaji
47	uelewano	32	kutimka	20	dhifa
47	uduni	32	kujinadi	19	ustawishaji
47	tiifu	31	mukama	19	uoto
47	hakimiliki	31	kutosheka	19	tamathali
46	mwanasoka	31	kitini	19	mwambata
46	kupitiwa	31	gia	19	mawanda
46	kujidanganya	30	wadi	18	uzunguni
45	mnadhimu	30	tochi	18	mchakamchaka
45	mchumi	30	makucha	18	kuchelea
45	chipsi	30	kujitia	18	fasheni
44	maridhio	29	ujangili	17	mtoano
44	komandoo	29	trela	17	lebo
43	uzamili	29	morali	17	kusikilizana
43	shangingi	29	mdau	17	kuelimika
42	tamthiliya	29	kuvusha	17	kamisaa
42	mrajisi	29	boksi	17	hadhira
42	kujizatiti	28	makumi	16	zingatio
40	kaseti	28	kikohozi	16	ukewenza
40	changudoa	28	judo	16	satalaiti

16	nguvukazi	11	njiti	7	mwasho
16	mjenzi	11	mlenda	7	mkutubi
16	mitishamba	11	mchecheto	7	matanda
16	mirungi	11	mazalio	7	masengenyo
16	mchepuo	11	kilowati	7	lira
16	mbinde	11	glovu	7	kustua
16	hafidhi	11	drama	7	kitumbi
16	albino	11	bwanga	7	azima
16	alau	10	urushaji	6	umwinyi
15	usafishaji	10	ulanzi	6	ulokole
15	ugumba	10	tasnia	6	simo
15	mwanazuoni	10	simenti	6	paredi
15	mvamizi	10	sarufi	6	mwanakondoo
15	kuongozana	10	palizi	6	mlei
15	kasma	10	modeli	6	kwato
15	geda	10	marejea	6	kupitiliza
15	boi	10	kusemekana	6	koroboi
14	uhamaji	10	kibantu	6	korido
14	shomvi	10	elementi	6	kabrasha
14	mamantilie	10	chirwa	6	jitu
14	chapati	9	usoshalisti	6	hatimiliki
14	brigedi	9	uegeshaji	6	gruneti
14	anguko	9	tobo	6	divisheni
13	udokozi	9	shushushu	6	chinjachinja
13	runinga	9	neli	6	boneti
13	protokali	9	mtovu	6	boharia
13	msalani	9	mkweche	5	ubuyu
13	kuumbika	9	mkoni	5	sentimeta
13	kutukuza	9	mhafidhina	5	pundamilia
13	kujibanza	9	limbikizo	5	mshawasha
13	juju	9	kidude	5	miraa
13	afueni	9	jiolojia	5	mfamasia
12	rapu	8	ukachero	5	matini
12	pekupeku	8	silabasi	5	kutunuku
12	nishai	8	sementi	5	kusisimka
12	mwanana	8	pesheni	5	kujikaza
12	mviko	8	oyee	5	juakali
12	mini	8	mikogo	4	windo
12	mhitimu	8	mchoyo	4	wanzuki
12	mdundiko	8	kuume	4	upesiupesi
12	matazamio	8	kukokotoa	4	ulozi
12	kusaza	8	jaketi	4	ubishani
12	kujisifu	8	disketi	4	telegramu
12	bukta	8	chambilecho	4	tamimu
11	wasionacho	7	tarazo	4	soketi
11	ulanguzi	7	repa	4	sadakta
11	sosholojia	7	ndwele	4	pumziko
11	nyato	7	mwezini	4	pinku

4	pimajoto	3	dingi	1	usumaku
4	penseli	3	chimbo	1	urukaji
4	padiri	3	chalo	1	upunjaji
4	opereta	3	bwanashamba	1	umamluki
4	ngoi	2	upoteaji	1	ukalafati
4	mwungwana	2	uana	1	udurusu
4	mtungaji	2	tyubu	1	toni
4	moramu	2	twaibu	1	tokwa
4	mkandamizaji	2	tumbuizo	1	tasinifu
4	mjeuri	2	teolojia	1	tashihisi
4	metali	2	teleksi	1	tamari
4	makapi	2	steki	1	takriri
4	lamba	2	spinachi	1	tabaini
4	kutuama	2	shifti	1	soseji
4	kushitua	2	samawati	1	sinki
4	kujikusuru	2	pindukia	1	singu
4	kujifaragua	2	pangaboi	1	simsim
4	kontrakta	2	nikotini	1	shikizo
4	kiyoyazi	2	mula	1	shifaa
4	kitegauchumi	2	muadhamu	1	shembea
4	kitafunio	2	mtawanyiko	1	reki
4	kipodozi	2	mpotevu	1	propela
4	kipando	2	mnyofu	1	pitio
4	kioski	2	mchapuko	1	piano
4	kimea	2	mawi	1	parachuti
4	kiambatisho	2	makwenu	1	nitrojeni
4	kafeteria	2	makulaji	1	nimonia
4	giligili	2	maanguko	1	mwekahazina
4	flamingo	2	luteni-jenerali	1	mwarubaini
4	fashisti	2	kutawaza	1	mtori
4	drafti	2	kuripea	1	mtimko
3	tanbihi	2	kukopi	1	mtekanyara
3	sitari	2	kukeneka	1	mparachichi
3	mwega	2	kujikamua	1	mnyanyasaji
3	mlahaka	2	kuchombeza	1	mlegevu
3	mikwala	2	kirai	1	mkukuriko
3	meja-jenerali	2	kijungujiko	1	mkokotoo
3	manifesto	2	kijaluba	1	mkinzano
3	kwashakoo	2	kidokezo	1	mhujuumu
3	kuwiana	2	katalogi	1	mgongaji
3	kumbukizi	2	kareti	1	mgawano
3	kuakisi	2	kanama	1	metriki
3	kisukuku	2	hostia	1	mdhalimu
3	kikaango	2	hidrojeni	1	mchirizo
3	kijito	2	hadubini	1	mbwawa
3	kiimbo	2	elektroni	1	mbonyeo
3	jubilii	2	akiolojia	1	mawio
3	gea	1	welekeo	1	masalo

Nordic Journal of African Studies

1	marudi	1	kuchachafya	1	friza
1	makwetu	1	klachi	1	elezo
1	majununi	1	kitulizo	1	dubwasha
1	maadhura	1	kitaya	1	domino
1	liwati	1	kitasi	1	chuka
1	kwanja	1	kisawe	1	chondechonde
1	kuzengea	1	kipandio	1	chesi
1	kuvisha	1	kifutio	1	chemshabongo
1	kutokosa	1	kichungi	1	burungutu
1	kusogoa	1	kichujio	1	bolti
1	kumotisha	1	kamukunji	1	biogesi
1	kukoki	1	ilhamu	1	binuru
1	kujidukiza	1	harabu	1	bambam
1	kugangamala	1	grisi	1	balagha
1	kudogेशa	1	geji		
1	kudeki	1	gamu		
1	kuchakaza	1	fyongo		

APPENDIX 2

Such new words in Kamusi II that do not occur in Helsinki Corpus of Swahili.

abra	dikoni	hasaisi
adeade	dipfriza	hataza
adharusi	diskaunti	hatimkato
aidini	diski	hipotenyusi
ajua	doezi	holwaho
akisami	doli	homonimu
akselereta	duaradufu	honoraria
alofoni	duhuli	hudihudi
aloi	dumuzi	ifaki
ambatano	ekseli	ikisiri
amonia	eksirei	injekta
anemia	elekezi	iodini
aprikoti	eleveta	istiara
arubastini	eskaleta	istihaki
arudhi	fahiwa	jari
asikirimu	falahi	jazanda
babakrismasi	fanani	jiofizikia
babio	fileti	jiometri
baluni	folio	juludi
batizo	fomeka	kaditama
bepariuchwara	fonolojia	kaimati
biarusi	frikiki	kambakoche
bibliografia	fukusi	kamilisho
cha-cha-cha	fumbu	kategoria
chagizo	fungutenzi	kauliflawwa
chambega	futikamba	kiambajengo
chapasi	fuwele	kiambishi
chelebuni	fyekeo	kiamokibaka
chembembe	gagazi	kibonzo
chembeuzi	galimu	kichapuzi
chipubodi	ghamidha	kichwamaji
chipukio	glukojeni	kidahizo
chochole	gluu	kidhabidhabina
chonda	gowe	kifunguamkoba
choza	gwatemala	kifuto
chudere	haibridi	kigawo
chuguu	hajja	kigutu
chukuchuku	halida	kiheba
chupio	halijoto	kiima
dahalia	halwoho	kiinilishe
dango	hamstasharamia	kikonyo
dawari	hamuma	kikonzo
dhahania	hanchifu	kikukusi
difu	hangale	kilimbo

kimbugimbugi	kontua	kuukeni
kimelea	koro	kuumeni
kimeng'anya	kostebo	kwache
kimenomeno	kotiledoni	kweche
kimono	kreni	kwina
kimori	krioli	labibu
kinegwa	kuatili	lalasalama
kingamwili	kubonga	leksikografia
kingwagu	kubumuta	leksimu
kingwangwa	kubunda	liamba
kinyambuo	kubura	lokesheni
kiowevu	kuchabanga	lugojo
kipashio	kuchokonoa	luteni-usu
kipueo	kudawiri	luva
kipimahewa	kuetuka	luvea
kipimamvua	kufahiri	maarusi
kipimapembe	kufumka	mafu
kipuri	kugaza	mafurungu
kiradi	kugwajuka	maghani
kiraikielezi	kuhemka	magowe
kiraikitenzi	kuhiuta	mahuruji
kiraikiunganishi	kuiliya	majazi
kiraikivumishi	kujibaidisha	malago
kirainomino	kujikunyata	maluuni
kirambamchuzi	kujitukusa	mangazimbwe
kirejeshi	kujuvya	manyenzi
kirekodiasauti	kukihili	mapatilizo
kirutubisho	kukosha	mapisi
kiruu	kukunyaa	marara
kishida	kulau	mararuraru
kisuguru	kuleuka	maruerue
kitapa	kuliliwala	maruwerue
kitendewa	kulinganua	mashutumu
kitendwa	kumangamanga	masimbulio
kithembe	kumwayamwaya	masmuma
kitibegi	kunasibu	masurupwete
kitindio	kung'fu	matukulele
kitomeo	kunururisha	matungamavi
kitone	kurapu	mavani
kivukomilia	kuriladi	mazingazi
kivyere	kushakili	mbatilisho
kiyeyusho	kusimbua	mbeuzi
kizabizabina	kutabahari	mbilikizi
kodiani	kutasiliti	mbinuko
kokoneo	kutinda	mbiombio
komafi	kutukutu	mbwamwitu
kongole	kututu	mchakuro
kongoo	kutwisha	mchanguzo

mcheduara
mchemraba
mchochole
mdadi
mdhalilishaji
mesi
mfuachuma
mfuafedha
mfuanazi
mgoto
mhangi
mhanuni
mharara
midabwada
mikrofoni
mikroskopu
minibasi
mitara
mjao
mjasiriamali
mkikimkiki
mkinzani
mkologwe
mkrismasi
mkugo
mkuruba
mkurupuko
mkurupusho
mkwinini
mlakungu
mlandano
mlariba
mlau
mlimaumwitu
mmeng'enyoy
mmenyuko
mnaa
mngogwe
mnyambuliko
mnyambuoy
mnyimbi
mnyumbuko
mochari
modereta
mofolojia
mpitisho
mpotovu
mraibu

msagunda
msanamaki
msanifuujenzi
mshonaji
msingizaji
mstahimilivu
msu
mswano
mtala
mtongozaji
muholelo
muhu
muungama
mviga
mvimbaji
mvukizo
mwairo
mwamana
mwanaisimu
mwanambuzi
mwangani
mwangusho
mwembemwitu
mwondoko
mwonevu
nadhariatete
nazali
ndifu
nevi
ng'ang'ani
ngisha
nibu
niru
njuzi
nong'ono
ntwe
nuktambili
nuktamkato
nyanyachungu
nyanyamshumaa
nyono
nyuzijoto
pajama
pambamoto
pambanuko
papi
parachichi
paruwanya

patapotea
peasi
pedeli
pembekali
pembemraba
pembenne
pembetatu
pembetatu_mraba
pembetatu_pacha
pembetatu_sawa
petali
peteo
pijini
polisemi
progamu
pwete
rediokaseti
rekebisho
riboni
rovu
sabaia
salalaa
saratangi
sayansikimu
segemnege
semansi
semantiki
shifta
shumizi
shundwa
sijiko
silesi
sinonimu
sintaksia
skonzi
slipa
sombombi
sroveya
spanaboi
stokingi
subana
suitafahamu
suke
sukumawiki
sulula
sumsuma
supamaketi
swahibu

tajamala	uambishaji	upadri
tajiriba	uamilifu	usafidi
takariri	uanazuoni	usakubimbi
tarabushi	ubindo	usenge
tarawanda	uoi	ushirishaji
tarbushi	uchakubimbi	utamaradi
tarijama	udakizi	utendi
tashbihi	uele	utesi
tawasifu	uheke	utungisho
tepurekoda	uitaji	uwasio
tepwetepwe	ujalidi	uwekevu
tikitimaji	ujasiriamali	vamio
tobwe	ujumi	vundo
tolatola	ulumbi	wakatabahu
tombola	umajinuni	wiski
topu	umbijani	yambiwa
tosti	unyafuzi	yambwa
uainisho	unyambulishaji	zaana
uamana	upadiri	zebu

APPENDIX 3

Part of words found by SALAMA from Helsinki Corpus of Swahili but missing in Kamusi II.

1796	uwezekano	119	umakini	61	machinjio
1640	maamuzi	118	intaneti	60	fukwe
1389	mwanasiasa	113	miundombinu	59	mzoefu
890	mwanamaombi	112	utayarifu	59	maripota
787	mafao	107	udikteta	59	manunuzi
687	wanamgambo	107	mkong'oto	59	kiliberali
631	kike	107	milenia	56	waandalizi
538	walemavu	106	ukimya	56	vipeperushi
497	maelekezo	104	mlemavu	56	utakatifu
474	mrengo	103	msisitizo	56	umilikaji
416	maambukizi	102	uelewa	56	mwanafamilia
403	mwanaharakati	101	wosia	56	marathon
344	dhulma	100	gesti	55	paundi
318	muhanga	98	utambulisho	53	umisionari
312	endelevu	98	ukimbizi	53	euro
311	shambulizi	98	makabidhiano	52	ukaribu
266	mwanamichezo	97	mwanataaluma	52	shirikishi
261	mashauriano	92	mrufani	52	majigambo
259	upatanishi	92	mahoteli	51	uchafuzi
234	mwanasayansi	88	wabunifu	50	taftishi
233	vitimbi	84	ulazima	50	mwanamfalme
231	mgao	82	wanyamapori	50	mshambulizi
204	marudiano	82	uniti	49	mwanakwaya
203	mwanandoa	80	mwanahisa	48	nusu-finali
192	ubores haji	78	mlalamikiwa	48	mtengano
187	feki	77	zimamoto	47	ruti
185	ubabe	77	ufinyu	47	mwanamageuzi
182	mbadala	77	al-qaeda	47	mtarajiwa
179	selula	76	punguzo	46	mashambuliano
176	upanuzi	75	uripuaji	46	faksi
167	kansela	73	uokozi	45	mwanabalozi
157	makusanyo	73	mihadarati	45	manyanyaso
149	dala	72	al-qaida	45	mahospitali
143	ubunifu	70	sekretarieti	45	kijenetiki
132	ujoto	70	maingiliano	44	michubuko
131	mbinyo	70	machangudoa	43	ukomavu
127	vijidudu	69	mitaala	42	washambulizi
127	utapeli	69	dili	42	umbele
127	mchujo	67	mwanahabari	42	udereva
123	wachambuzi	66	kihafidhina	42	mtunza
123	maridhiano	64	geti	41	wani
120	mwanamama	63	ofa	41	mwanabiashara

41	mkao	30	kisoshalisti	22	usongo
40	uranium	30	droo	22	rifu
40	mzingiro	30	bodaboda	22	mparaganyiko
40	mwathirika	29	wanasemina	22	mijizi
39	usaidizi	29	ufufuko	22	majibizano
39	uharamia	29	mwongezeko	22	kumtapeli
39	jihadi	29	molekuli	22	jungu
38	mwanakamati	28	uvunjifu	21	zuio
38	mataa	28	mchapakazi	21	utofauti
37	presha	28	kulibinafsisha	21	uingilizi
37	morani	27	theolojia	21	uharifu
36	ukocha	27	mwendelezo	21	muonekano
36	ukereketwa	27	mtindio	21	muathirika
36	mwenda	27	morali	21	matenki
36	mahitajio	27	kuwatapeli	21	mamalishe
35	upembuzi	27	kisekula	21	legelege
35	umakamu	27	chonde	21	antena
35	mwingiliano	26	suali	20	mwanadada
35	mtembea	26	mwanahistoria	20	miniti
35	klorokwini	26	mfumko	20	dodoso
34	washikadau	26	matashi	19	utatanishi
34	wabia	26	masuali	19	ukaazi
34	vipandikizi	26	machapisho	19	rasili
34	manuari	26	kanzela	19	mwonekano
34	maamrisho	25	ukansela	19	mshiko
34	kuzawadiwa	25	ujumla	19	mizozano
33	wanajumuia	25	mrundikano	19	mathara
33	mshukiwa	25	kianzio	19	maadhali
33	mchambuzi	25	kenchi	19	kuswekwa
33	mashirikisho	25	grosari	19	kitimoto
33	madenti	25	bwana_harusi	19	kisovieti
33	fiti	25	ameburuzwa	18	utanda
33	batiki	24	tisaa	18	uaskofu
32	uvunguni	24	swahiba	18	satelaiti
32	promota	24	ripota	18	mcheza
32	mwanamazingira	24	mwanauchumi	18	mapaparazi
32	mapromota	24	mmiminiko	18	mang'amuzi
32	laivu	23	wikiendi	18	kulundikana
32	kutapeli	23	waandaji	14	mwanajamii
31	madodoso	23	vikaragosi	13	wanaushirika
31	lindo	23	utunzi	13	wanakikundi
31	kugushi	23	promosheni	11	wanawarsha
30	wanandugu	23	mwanademokrasia	11	wanausalama
30	usomi	23	mpigania	10	mwana-mfalme
30	uhakikisho	23	mpanda	9	wanaparokia
30	maveterani	23	mashiko	9	wanangangari
30	kuvurunda	23	chupuchupu	7	mwanakisomo
30	kutapeliwa	22	zengwe	6	wanamnyumbuliko

6 wanakaya
6 wana-ndoa
5 wanawazuoni

5 wanaikolojia
4 wanakanisa
4 wanafizikia

4 wanabaharia
4 mwanafizikia
3 wanaukoo

APPENDIX 4

Part of words found by SALAMA from Helsinki Corpus of Swahili but missing in Kamusi II. The word has either insufficient morphological information in Kamusi II or it has a spelling not known by Kamusi II.

1773	muafaka	117	padre	44	mapolisi
1568	demokrasia	117	makandarasi	43	uwingi
1260	umasikini	115	rasilmali	43	mahospitalini
1202	uhakika	108	oparesheni	42	kinuklea
1012	magonjwa	107	wanachuo	41	wanazuoni
781	masikini	103	kinyuklea	41	mathalan
716	majibu	102	mahsusu	40	upadre
638	makombora	101	maaskari	40	milki
612	wakaazi	97	repoti	40	makusudio
589	serekali	95	makaratasi	40	maguvu
494	mahusiano	94	kinyukliya	40	lundo
451	masaa	84	luninga	39	watalaam
386	kadri	80	muongozo	39	sentimita
383	hatma	80	madhamana	39	gobore
257	mahesabu	78	opereshini	38	mshituko
256	radio	76	utaalam	38	mine
253	takribani	76	muongo	38	masanamu
253	kinyuklia	75	tathimini	38	makovu
250	makaazi	74	kikoministi	37	muandishi
240	raslimali	73	mwungano	37	mnyumbuliko
238	mauwaji	71	waakilishi	37	mapango
230	kugharamia	70	hiyari	36	paspoti
220	majuzi	68	khofu	36	kujiripua
218	kuchukuwa	64	fitna	35	umbea
216	makambi	63	muuzaji	35	maghorofa
210	mapadre	59	wazoefu	35	kipuuzi
206	malewano	59	taslim	34	kusheherekea
202	kinuklia	58	kuripuka	34	hadith
199	muwafaka	56	mahkamani	33	silka
195	khanga	54	muhusika	33	khitma
179	finali	54	kutathimini	33	hitma
174	nyuklia	52	nuklia	33	barua-pepe
150	nuklea	52	muombaji	32	utayari
150	mahkama	51	maofisi	32	tekinolojia
147	unafuu	49	mashuka	32	penati
141	rukasa	48	biliyoni	32	muakilishi
141	kikomunisti	45	maharusi	32	mpiga_picha
133	kondom	45	jamuhuri	32	matamasha
131	kheri	44	wafanyikazi	32	mashataka
127	makutano	44	mustakbali	32	mahangaiko
120	ankara	44	mme	32	kibayolojia

32	jubilee	28	kinyanganyiro	21	mda
31	ubinaadamu	26	matrafiki	21	dimokrasia
31	hereni	26	makasri	20	watalaamu
31	billioni	26	makambini	20	sekondi
30	ubinafsishwaji	26	kuripua	20	mangumi
30	thelasini	26	hijab	20	maksudi
30	milimita	25	utengamano	20	komputa
30	mazowezi	25	matetemeko	20	chemchem
30	mangapi	25	kuripuliwa	20	baskeli
30	makucha	25	kibaiolojia	19	liliporipuka
30	maboksi	23	mkaazi	19	kaptula
30	helkopta	23	warufani	18	gogoro
30	chemichemi	22	masiku	10	mwanazuoni
29	vizuwizi	22	mashirikiano	5	mwanachuo
29	liliripuka	21	mweleka		