Saahot Afhi Maskaaka

Saho Corpus

Yatneefecetiyahi santo

User manual

Jama Musse Jama

(With support of Ahmedsaad M Omer)

2022

Canqurko 350 shixtiya takkem yemweesebe ● Dokkumentitko 324 barhsimeh tuumaarime ● canqurko yewce canqurko 30 shixtiya takkem geytime ● 4500 ko bukalle yakke canqur 3 luqhah yunturjume

Over 350 thousand words tagged ● 324 documents parsed and indexed ● Over 4500 head words ● Almost 30 thousand inflected forms generated ● Over 4500 thousand translations in three languages

TABLE OF CONTENTS

1. Introduction

This User Manual aims to introduce different functionalities of the Saho Language and Literature Corpus management platform, developed by Redsea Cultural Foundation, and to allow linguists, researchers, and other potential users of the corpus to access it easily, search and retrieve data from both the corpus and its auxiliary tables (dictionary, synonyms, etc.). It also allows the Committee that oversees the Corpus to follow the correct workflow procedure to make sure that data inserted into the Corpus database is verified and corrected.

The RCF Saho Corpus incorporates a "Word Behaviour" (WB) page, a seminal work of corpus-based analysis of a word in Saho discourse, similar to what has been developed for the Somali language. The WB has been introduced and developed for the Somali Corpus in 2016 and it has been a useful tool for learning the language and at the same time for contributing to the growth of the Somali Corpus itself. New functionality has been introduced in to the WB for the Saho, and now it is based on a powerful querying tool to access the grammatically tagged Saho Corpus to summarize the syntactic context of the searched word; for instance, to show the relationship that the searched word has with other words, which words are mostly associated with the searched word; the frequency of the searched word within specific sub-corpora; the etymology; synonyms and antonyms; spelling variants of the searched word; and finally the definitions from a list of reference dictionaries as well as translations to English, Italian, and Somali languages.

This document covers the following major sections:

1. Access to the system, basic search and data retrieval;

2. Major components of the WB view page: definitions, synonyms, etymology, concordance, translations, grammatical behaviour of the word in pre-defined structures and collocations;

3. *SahoISearch* - an intelligent search engine, incorporated in the WB for the Saho. It can search for words in all their inflected forms. It can also search through translations in English, Italian and Somali, or through synonyms in Saho: for instance, one can search the word *quwa*, which means 'force, energy', and the engine will also search for *xayla*, which means 'energy, power', a near synonym of *quwa.*

4. Exporting data and printing.

## 2. Access to the system, basic search, and data retrieval

Access

RCF Saho Corpus is a web-based online platform, and it needs a valid user id and password to access it and to research data. The system administrator, who also defines a specific role for each approved user, provides the user id and the password upon request. If you do not yet have a valid user id and password, please write to info@somalicorpus.com.

Whether you are connected or not, the system shows you on the top bar of the page some statistics: number of headwords in the dictionary, number of inflected forms automatically generated by the system, number of words tagged in the corpus and number of documents indexed. If you are not connected, at the end of the top bar, you have language options (Saho, Somali, English). When you are connected, at the end of the top bar, you see your name, role and the possibility to logout, as shown below.

Qaamuus: [4,655 Amot canqara | 13,181 Yambedeedele xibar] Maskaaka: [127,810 canqur | 38 Sanad] [Logout: jamamusse/Administrator] Af/waani: EN SO SA

Login

Once you have obtained a user id and password, you can connect to the system as indicated in the first page (see the image below). Put your id in the "*orob*" (login) field and your password in the "*geli-canqara*" (password) field, then click on the "*Maskaaka orbish*" (enter the corpus) button for authentication.
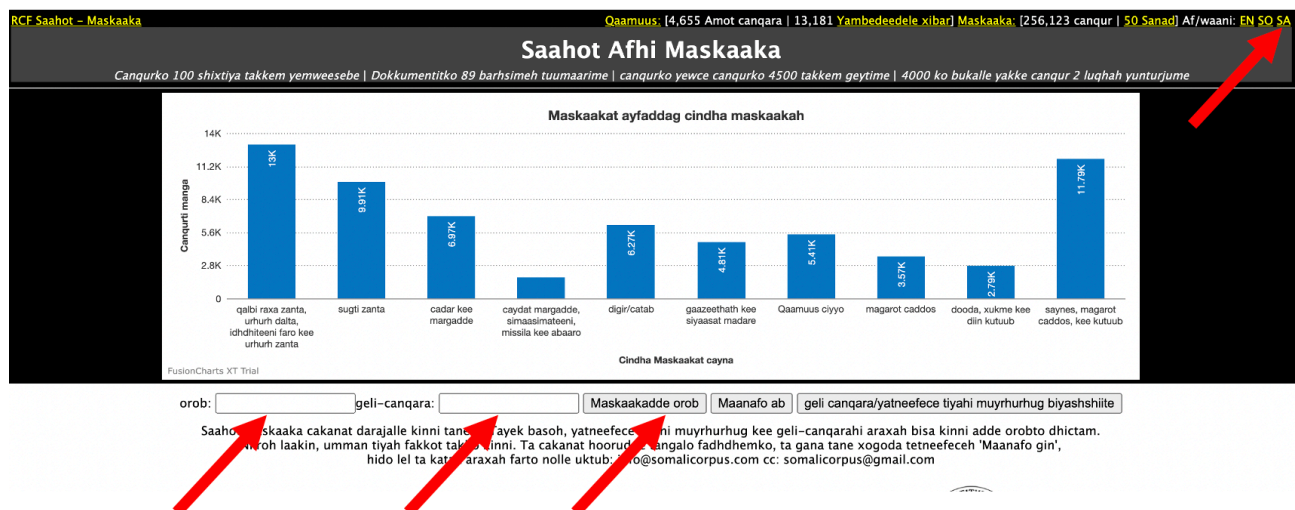
### *Browse*

Statistics given by the system include the distribution of the dictionary in Parts of Speech; distribution of the corpus in eras; and distribution of the corpus in sub corpora. All these distributions are seen in the form of graphs. Links to *Qaamuus* (Dictionary), *Masaaska* (Corpus) and *Sanad* (Documents) are in the top bar. When you click on "*Sanad* (Documents)", in particular,

you can browse the corpus by author: a list of authors is shown ordered by number of words in the contribution to the corpus by that author. When an author's name is clicked, then you see the list of titles written by that author, and each link will take you to the analysis of the single document within the corpus: for instance, frequency of used words, list of successfully tagged words, list of unrecognized words or spelling errors found in the document, etc.

Search

When you log into the system, you will see the distribution of the dictionary in Parts of Speech. You can perform three types of simple search and the Advanced Query Language search type: First, to search a root word, choose "*Qaamuus*" (dictionary) at the search drop-down and insert characters in the field "search". The system will perform a word prediction process while you are typing and, if a match is found in headwords, the word prediction system will guide you to the right spelling. You can also put here an inflected form of the word you are looking for, and again, if what you are typing matches to a headword in the dictionary, the system will retrieve it and show the headword definition.
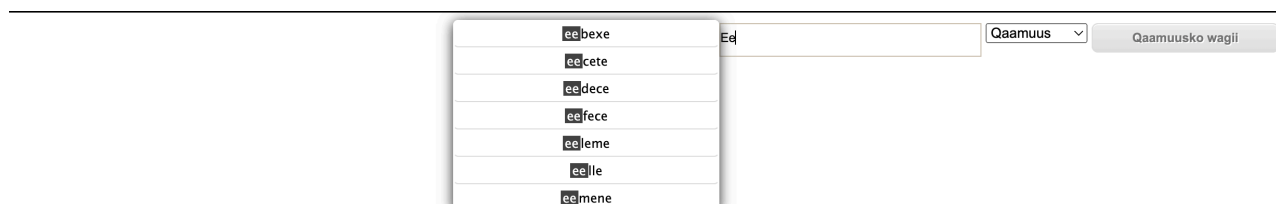
The second type of simple search is on the corpus. Choose "Misaaska" (corpus) in the search drop-down, and insert the word. The system will not perform a prediction and will search the inserted word in the corpus and not in the dictionary of headwords.

The third type of simple search function is through its translations in different languages (at present English, Italian, French and Swedish are available for some words). To perform this search, follow the link *"ku baadh afafkan"* at the left side of the screen. It will open a new popup window.
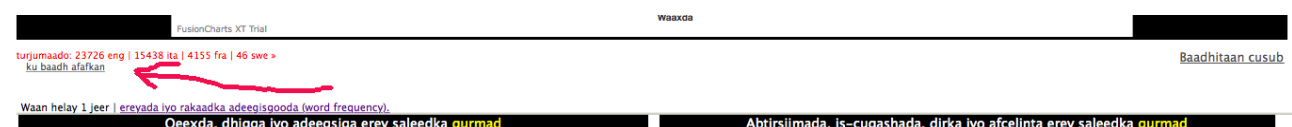
Insert one or all your English, Italian, France and Somali words in the corresponding fields and click *"luqhaat goronish"*. [This function might not work on some browsers].

| English | love | Italiano | odio | Français | | Somali | cilmi |

kaxano – nm
love – love
kixina – vIII
to love, to like, to prefer – to love, to like, to prefer
ikxine – vI
to like, to desire, to fall in love – to like, to desire, to fall in love

nacabo – nm
odio, ostilità, avversione, antagonismo] rel to niciba – odio, ostilità, avversione, antagonismo] rel to niciba

waxaad i waydiisey somali cilmi .... waan raadinayaa ....

luqhaat goronish   finisira alif

You do not need the words to have the same meaning as the system will perform separate and parallel searches. The results will be shown in different colours as shown below (red for English, green for Italian, white for French and yellow for Somali). The corresponding Saho word is shown through an active link which, when clicked, will open a new search in the dictionary.

To close and return back to the main page, click "finisira alif" (close this window) button.

Whatever service you use (browse, search, translation), once the word is identified, the system will bring you to the Word Behaviour view structure. See Chapter 3 for more details.

3. Word Behaviour view structure

As the above screenshot indicates, the Word Behaviour view has 7 major components: Box (1) indicates the number of times the headword is found in the dictionary (meaning also different Parts of Speech); box (2) contains the definitions of the headword as provided in the different dictionaries; box (3) contains the links to the synonyms of the headword; box (4) contains the concordance of the headword (and all its inflected forms); box (5) contains the etymology of the word (if known); box (6) and collocations defined through pre-established configurations (for example if the headword is a Noun then we extract the collocation with the conjunction expression *"kee"*, etc.(see chapter 4 for pre-configured clusters of rules for each part of speech; box (7) contains all the inflected forms of verbs and defined forms of nouns generated by *SaCoFlexor*; and box (8) contains the translations of the word in other languages (English, Italian, and Somali). Below is a capture screen of each section for the word *eebexe* (to sell).

Dictionary definitions

Definitions are searched in several monolingual dictionaries and for each entry the Part of Speech is shown. There is also an active link for each *PoS* entry found in the major dictionary of reference that shows you the inflected forms of the selected item.

caddos, inki macnale canqur kee amot–canqarahi nafaaca eebexe

amot–caqara: eebexe (vl)
  to sell
  (–)

Synonyms

Synonyms are collected automatically from the main dictionary and subsequently some of the entries were edited manually to link synonymous words. There is an active link that brings you to the WB view for the new word.

inki macnale hido yamseene canqurti amot–canqara candhishe

Sitta–xeela/canqarat aki macna amot–caqara: candhishe (vll)
  candhiirhishe

Concordance

In this box you can see the searched word and all its inflected forms used in the corpus. The first column contains an active link that indicates the sub corpus in which the instance is found. When clicked the link opens a popup notice that tells you the details of the document this sentence belongs to. The rest of each row shows the searched word (or the inflected form of the headword)

	**canqurti hoor: amot–canqarahi nafaaca eebexe**		

canqurti hoor amot–caqara: eebexe (vI)

TIR	*Akatko kursit kee carut akkabaanah*	*yaabixin.*	
TIR	*miyiniinimhi sabbatah amaa baahen badaaca*	*yaabaxoona*	*ayrolle hel arhxi mangum xaftata yinin*
TIR	*Asmaral*	*naabaxo*	*barbare*
TIR	*Ziyaada neskettih*	*aabaxo*	*farhamih sabbatah laakin adde osno kino*
TIR	*lel ta xirkat lifi ica*	*aabaxo*	*kinni yixagara � tih radde akah*
TIR	*Tenadde lel*	*naabaxo*	*kino � akterhxe.*
TIR	*Tee Jumca ilaw ayidhdhah*	*teebexe*	*akyen carelle Ishi lel sik erhex*
TIR	*Asmaralle*	*teebexe.*	
TIR	*Cirheh yaanah lel mango limoh*	*yaabixin.*	
TIR	*ta medderooko walaah ruga akko*	*yaabaxo*	*dhiicak miyine.*
TIR	*Calik beceera ebex akterhxih beceera*	*yeebexe*	*yokyen � akte.*
TIR	*Ziyaada neskettih*	*aabaxo*	*farhamih sabbatah laakin adde osno kino*
TIR	*lel ta xirkat lifi ica*	*aabaxo*	*kinni yixagara tih radde akah tedebbe.*
TIR	*Tenadde lel*	*naabaxo*	*kino akterhxe.*
TIR	*4 Tee Jumca ilaw ayidhdhah*	*teebexe*	*akyen carelle*

Etymology

Etymology is an area that has received very little attention in the academic research on Saho language. In this box of WB, you will find a small number of Saho words of which we do know the etymology, and all loanwords from other languages (Arabic, Italian, English, French, etc.). An active link indicates the language of origin and the word, and if you click on the language, you will find a list of Saho words loaned from that specific language. For Arabic there are some font issues to sort out.

amot–canqarahi maabako, xuggaane, rimid kee tarjama abbamale

Canqarat maabako amot–caqara: abbamale (nm)

From abba [father] and male [without] = abbamale [fatherless]..

Inkiddehi xibar Migac: abbamale

Word behaviour in a pre-defined grammatical structures

Immediately after the etymology box, you have several pre-defined grammatical structures. For each type of Part of Speech, the system checks if there are collocations of linguistic interest. For instance, if it is a noun, as said earlier, the system checks the collocation of the noun with the conjunction expression "*kee*" (and). Another collocation for nouns of particular interest is with the verb "*eelle*" (to have).

Inflected forms of the word

SaCoFlexor is an application written by the author that develops noun and verb derivatives from head words according to the rules defined in the accessible authorative sources published from the time the Saho language was being institutionalized as the national language. One major reference is Vergari & Vergari (2003) [revised edition of 2007], but also other sources including Giorgio Banti, Ahmedsaad M Omer and others. These rules define different ways of recognizing derivatives of the headword and generate all inflected forms. In this box you see the generated

forms in an active link that, when clicked, will show you the concordance of the specific inflected form in the corpus (instead of the lemma). If there are combined words of which the searched word is part of the combination, again here you see the list.

Masarrat kin canqarahi yenkeleelebe xibar eelle

amot–caqara: eelle (vl)
  rimid canqaraako yewce canqur *eelle (vl)* [malammih ab]: eelle eelle aalle aallo teelle taalle taallo eella yeelle yaalle yaallo teelle taalle taallo neelle naalle naallo teellin taallin taalloona eellaa yeellin yaallin yaalloona [remove them all]

amot–canqarahi agleeko yimfillice xibar eelle
  amot canqarahi aglel migeytiminna (eelle)

Translations

Finally, the translation box will show you different translations from different bilingual dictionaries (Saho-Italian, Saho-English, Saho-Somali). Dictionaries of reference are being revised constantly. Other languages can be added to the platform. Here is shown the word *"xurriya"* (freedom).

amot–canqarahi tarjama xurriya

Talyaani: xurriya
  nf 1. libertà 2. indipendenza] var xunriya

Ingiriisi: xurriya
  nf 1. freedom 2. indipendence 1. freedom 2. indipendence

4. Automatic identification of linguistic structures in the corpus

There are few configurations defined within the platform for specific types of Parts of Speech. The aim is to show how the selected word behaves within the discourse (for instance we have already seen if the selected word is Noun, we want to check the collocation of the N with the word "*kee*" (and), so we can figure out all other nouns (N1) where the configuration "N *kee* N1" matches. These results are shown in the Word Behaviour view. Different configurations apply to different types of Parts of Speech.

Verb flexor: tense conjugations

SaCoFlexor is a part of the Saho Corpus, and it learns from the corpus in a progressive way to enrich the corpus itself. For instance, five a verb V1, you can with *SaCoFlexor* calculate all the conjugations of V1 in Perfect, Imperfect, Subjunctive and Imperative form for 1st, 2nd, and 3rd person, singular and plural, as well as gender forms. Once produces all forms, it checks back on the corpus if that flexed form has been used at least one time. You can retrieve the sentence used this form. See example 1 below.

Example 1: how to use *SaCoFlexor* for conjugation of verbs in Saho?

Once connected to the system, Click on "Yambedeedele xibar" (inflected forms) on the top menu bar.

RCF Saahot – Maskaaka | Qaamuus: [4,655 Amot canqara | 13,181 Yambedeedele xibar] Maskaaka: [127,810 canqur | 1,119 Sanad] Af/waani: EN SO SA

Saahot Afhi Maskaaka

Canqurko 100 shixtiya takkem yemweesebe | Dokkumentitko 89 barhsimeh tuumaarime | canqurko yewce canqurko 4500 takkem geytime | 4000 ko bukalle yakke canqur 2 luqhah yunturjume

You will see the list of lemmas for which so far, the system recognized and implemented its conjugations in case of verbs, or expanded all inflectional forms, categorizing them in 4 types of lemmas: a) not yet implemented, b) fully expanded, c) partially expanded and for those irregular cases, 4) manually expanded). You can search a root word and see what inflected forms generated from it. For instance, search the word *"eebexe"* (to sell) in this box and click on *"Qaamuuska Ogii"* (ask the dictionary)

Canqarat maabako: | Canqarat gabbabbac || dhibrhi warartalle gax | DB–idde cusub dokkumeti os | geyshisheenitti Arux || Yatneefece mara wararah haa | Documents, xato kee ware | Yi Maskaaka

☞Total lemmas: 4655: not expanded 4261 » fully expanded 321 » partially expanded 72 » manually expanded 1 »
total inflected forms: 13181

eebexe | Qaamuusko wagii
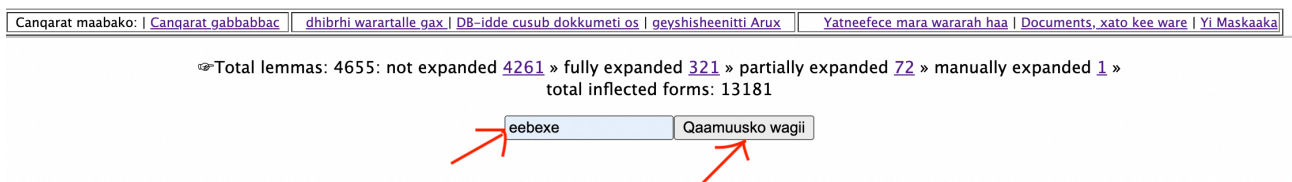
The result will be showing you on the left side a tree-like structure that gives you the word *'eebexe'* (to sell) and its inflected forms generated, and on the right side you will see if these generated words has been already used in the Corpus at least one time.

| canqara orbish | | Qaamuus ⌄ | Qaamuusko wagii |

Canqarat maabako: eebexe | Canqarat gabbabbac | dhibrhi warartalle gax | DB–idde cusub dokkumeti os | geyshisheenitti Arux | Yatneefece mara wararah haa | Documents, xato kee ware | Yi Maskaaka

lemma: eebexe [vI]: *to sell*

eebexe: *[vI_R000]*
teebexe: *[vI_R001]*
yeebexe: *[vI_R002]*
neebexe: *[vI_R003]*
teebexin: *[vI_R004]*
yeebexin: *[vI_R005]*
aabix: *[vI_R006]*
taabixe: *[vI_R007]*
yaabixe: *[vI_R008]*
naabixe: *[vI_R009]*
taabixin: *[vI_R010]*
aabaxo: *[vI_R011]*
yaabaxoona: *[vI_R012]*
naabaxo: *[vI_R013]*
eebexe: *[vI_PERF_1s]*
aabixe: *[vI_IMPE_1s]*
aabaxo: *[vI_SUBJ_1s]*
teebexe: *[vI_PERF_2s]*
taabixe: *[vI_IMPE_2s]*
taabaxo: *[vI_SUBJ_2s]*
ibix: *[vI_IMPV_2s]*
yeebexe: *[vI_PERF_3sM]*
yaabixe: *[vI_IMPE_3sM]*
yaabaxo: *[vI_SUBJ_3sM]*
teebexe: *[vI_PERF_3sF]*
taabixe: *[vI_IMPE_3sF]*
taabaxo: *[vI_SUBJ_3sF]*
neebexe: *[vI_PERF_1p]*
naabixe: *[vI_IMPE_1p]*
naabaxo: *[vI_SUBJ_1p]*
teebexin: *[vI_PERF_2p]*
taabixin: *[vI_IMPE_2p]*
taabaxoona: *[vI_SUBJ_2p]*
ibixa: *[vI_IMPV_2p]*
yeebexin: *[vI_PERF_3p]*

rimid canqaraako yewce canqur: eebexe

	eebexe: eebexe * aabixe * aabaxo * ibix			
	eebex	aabix	aabax	ibix
1s	eebexe ☒	aabixe ☒	aabaxo ✔	
2s	teebexe ✔	taabixe ☒	taabaxo ☒	ibix ☒
3sM	yeebexe ✔	yaabixe ☒	yaabaxo ✔	
3sF	teebexe ✔	taabixe ☒	taabaxo ☒	
1p	neebexe ☒	naabixe ☒	naabaxo ✔	
2p	teebexin ☒	taabixin ☒	taabaxoona ☒	ibixa ☒
3p	yeebexin ☒	yaabixin ✔	yaabaxoona ✔	
geye 408 gedda				

Noun behaviour

Nouns have different predefined configurations: we can observe for now two of them 1) the collocation with the conjunction "*kee*" (and) is an interesting collocation that highlights the two nouns that go together more frequently (see example 2 below on *xan* (milk) and *xazo* (meat)); and 2) the collocation with verb *eelle* (to have) also rises an interesting observation.

Example 2: Collocation of the noun *xan* (milk) with "*kee*" (and).

Although the corpus is currently very limited in terms of number words to be indicative, however, it is very interesting to observe that the noun *"xan"* (milk) is always (almost 70% of the current corpus presence) collocated with *"xazo"* (meat) in relation to *"kee"* (and) conjunction.

	Noun behavior for *xan* with the collocation *KEE.*		
poe	Lay	kee	xan nakkeyya saaku
TIR	Atin yetta balih shaahi	kee	xan kixinitiini
TIR	udduuniyalle nanem fan sittiyadde xan	kee	lay nekkih marno.
TIR	Amayimko lel xan	kee	xazo akko beetan.
TIR	Gaalatto yi gaalatto Mece xan	kee	xazo lito Meerrah malamma malto.
TIR	lel macaashah amaa sacahi xan	kee	xazo yamneefecin.
TIR	amaa sinni medderooko geyan xan	kee	xazo atteeley maran.
TIR	Amayimko lel xan	kee	xazo akko beetan.
TIR	Gaalatto yi gaalatto Mece xan	kee	xazo lito Meerrah malamma malto.
TIR	lel macaashah amaa sacahi xan	kee	xazo yamneefecin.

Still more interesting that if we focus on *xazo* (meat), and we observe the match goes up almost

88% of the presence of the corpus.

Noun behavior for xazo with the collocation KEE.			
TIR	Amaa lellec	*kee*	Xazo faylishen lellec aarhigi ukfure istiqhfire
TIR	Amayimko lel xan	*kee*	xazo akko beetan.
TIR	Gaalatto yi gaalatto Mece xan	*kee*	xazo lito Meerrah malamma malto.
TIR	lel macaashah amaa sacahi xan	*kee*	xazo yamneefecin.
TIR	amaa sinni medderooko geyan xan	*kee*	xazo atteeley maran.
TIR	Amayimko lel xan	*kee*	xazo akko beetan.
TIR	Gaalatto yi gaalatto Mece xan	*kee*	xazo lito Meerrah malamma malto.
TIR	lel macaashah amaa sacahi xan	*kee*	xazo yamneefecin.

4. Translation of user interface (Saho, English)

Compiled by Ahmed Saad and Dr Jama Musse Jama.

English to Saho

English	Saho
100 most frequently used	mangih gabbabbaacitta 100 canqaratiya
Add new corpus content	Maskaakat cusub dhibrho os
Add new document to the DB	DB-idde cusub dokkumeti os
Add or remove these documents to the personal sub corpus of XXXXXX. There are not at the moment documents in your personal sub corpus, select from the dropdown list and add.	Tamaa cindha ishi maskaakaako XXXXXX ta dokkuments asac-ishe kee adde ose tik. Kadit waqhtedde ku cindha maskaakadde walaa inki dokkumenti mitaneeki, guballe oobeh yane hoorko doorteh adde os.
All	inkoh
Analyse	Mokkaaxish
Analyse documents by this author	tati yuktubeh yane dokkumentit imirmir
Analysing	amarmar
Analysing selected corpus: [generate it again - it may take long time]	Doorimteh tane Maskaaka mokkaaxishsheh: [malammih bahto icikkin – tayim dherh waqhte beyto dhicta
Analyze corpus and list frequency	Maskaaka mokkaaxishsheh gabbabbaacittam wararah haa
and much more	amaye kee osa
Approved	kabbaatime
Author	yuktube tiya
Author name	yuktube tiyahi migac
Back to the list of contents	dhibrhi warartalle gax
Building now	kaado ginnimtam
cancel	xab
categories of the corpus	Maskaakat karim
close window	finisira alif
Combinations formed from the head-word	amot-canqaraako ginniime aglel
Commands	amre
Comment	mablo
Common behavior of the	Inkiddehi xibar
concordance	canqurti hoor
Concordance: use of head-word	canqurti hoor: amot-canqarahi nafaaca

Confusing: provide Corpus ID or Document ID - Not both of them	Amlolokka: Maskaakat muyrhurhug hido dokkumentiiti muyrhurhug kab-ish, kibah lammik maki
Contact the Corpus Manager	Maskaaka Baclayshatiyaalih geytim
Content subtype	Cayna
Content type	Cayna
Contents	dhibrho
Corpus	Maskaaka
Corpus types	Maskaakat cayna
could not find any document in which the word is used	ta canqara adde tane walaa inki dokkumenti migeytiminna
could not find this	Teyim geytimo mirhiicinna
Create an account	Maanafo ab
Creator	abe tiya
Definition, synonym and use of the head-word	caddos, inki macnale canqur kee amot-canqarahi nafaaca
Dictionary	Qaamuus
Dictionary entry for the word	Qaamus orbeena canqara tiya
Distribution of the corpus by era	Maskaakat ayfaddag gedih
Distribution of the corpus by Parts of Speech	Maskaakat ayfaddag waanit karimih
Distribution of the corpus by sub corpora: 10 types	Maskaakat ayfaddag cindha maskaakah
do it again	malammih ab
document	document
document	Sanad
Document ID	dokkumentiiti muyrhurhug (ID)
document	Sanad
documents	Sanad
Documents, help and news	Documents, xato kee ware
Edit	eymemec
Edit user profile	Yatneefece tiyahi Addoodim eymemec
edited	yemmemece
Edited	adde temmemece waqhte
Enter the corpus	Maskaakadde orob
Era	Ged/waqhte
Error	akke-wayne
Error in request of corpus analysis	Maskaakat marmarahi wagiinan qalath kinni

error: check connection to database	error: check connection to database
error: could not find that book	qalath: amaa kitaab migeytimo
essays	magarot caddos
Etymology of	Canqarat maabako
Etymology of the word	Canqarat maabako
Etymology, collocations, root and translations of the headword	amot-canqarahi maabako, xuggaane, rimid kee tarjama
Expanding the word	Expanding the word
Facebook	Facebook
Fifty mostly used words	Mangih nafacadde-asse kontom canqaratiya
First Degree	Afti digri
Flexed forms of the headword	Masarrat kin canqarahi yenkeleelebe xibar
Flexed from a combination of the headword	amot-canqarahi agleeko yimfillice xibar
Focusing on	adde hammitaanam
head words	Amot canqara
Home	meela
Hundred mostly used words	Mangih nafacadde-asse bool canqaratiya
I do not know	maarhige
I forgot userid/passowrd	geli canqara/yatneefece tiyahi muyrhurhug biyashshiite
I found	geye
in	addadde
inflected forms	Yambedeedele xibar
Information	masxassa
Insert a word	canqara orbish
insert one or more words	Canqara hido canqur sayish
inserted	sayeh yane
Inserted	adde orobte waqhte
Inserted by	orbishshem
is synonym of	Xeelta
is the plural form	mangat xibar kinni
Language	Af/waani
Last search	abayshimat goron
Last users seen	sarroorimalle yetneece mari yumbulluyin
Last words searched	sarroorimalle goronsime canqur
lemma	lemma
lexicographical works	Qaamuus ciyyo
list	Hoor/warar

List corpus contents from DB	DB ko maskaakat dhibrho waraarish
list of books: all items	kitobti wararta: umbuka cale
List users	Yatneefece mara wararah haa
loanword from	Abud canqara __ ko
Login	orob
Make the writing less than 7450	Kutuub 7450 xarfiyaako mifeerishin
most frequently used	Mangih gabbabbaacittem
Most searched words	ummanimko ziyaada goronsime canqur
mostly used words	Mangih nafacadde-ase canqur
My corpus	Yi Maskaaka
New	cusub
New search	cusub goron
newspapers and political speeches	gaazeethath kee siyaasat madare
no answer from the server	Server ko modobba mitane
No base word found for this word:	ta canqarah siddimat canqara akah migeytiminna
No combinations found for the headword	amot canqarahi aglel migeytiminna
No concordance found	canqurti hoor migeytiminna
No less than 5,000 and not more than 10000 words	5,000 ko lee mangum 10,000 ko lee dago canqur yemmemeceh yane
No permission to use.	Tatnaafaco izne malito.
no translations found	tarjama migeytiminna
No valid url or no connection	akkinnixe URL hido shabaka mitane
No word to search on dictionary is provided. Select from the frequency list below.	Qaamusudde wagiyan canqara mamxayinna. Ta guba yane gabbabbaacita hoorko door
non synonym words found	Sitta-xeela akkinnixe canqur geytime
notes	xunsus
noun	Migac
novels, short stories, travelogues, and flash fiction	qalbi raxa zanta, urhurh dalta, idhdhiteeni faro kee urhurh zanta
Number of characters	Xuruuf manga
Number of documents	Dokkumentiiti manga
Number of letters	Ta dokkumentidde yane xuruufhi manga.
Number of times	gabbabbacnan manga
Number of words	canqurti manga
Number of words	Canqurti manga
only one time	Inki gedda
paragraph	faqhra

Parts of Speech	Waanit karim
Password	geli-canqara
PhD	Doktoora (PhD)
plays	digir/catab
Plural	Manga
poetry and songs	cadar kee margadde
Post	Post
Provide at least one of the corpus ID or single document ID	Maskaakat muyrhurhugko daggoyemko inketta bah hido inki dokkumentih muyrhurhug kab-ish
RCF Saho Corpus	RCF Saahot Maskaaka
Redo	gabbac/malammih ab
Reduce the number of concordances requested	fafhdhishime canqurti-hoor uynuqhus
Reduce the number of frequency words requested	fafhdhishime canqurti-gabbabbac uynuqhus
Reference of the target corpus or single document	fanti Maskaakah tuumar hido inki dokkumenti
Relation	zamoodinno
remove them all	remove them all
Restricted area	deesimte cara
result	maararo
Save the content	dhibrho ixfiz
science and mathematics textbooks, essays, and articles	saynes, magarot caddos, kee kutuub
search	Goron/wagiyo
Search AQL	AQL goronish
search from dictionary	Qaamuusko wagii
Search on corpus	Maskaaka goronish
search on languages	luqhaat goronish
search on translations	tarjamadde goronish
Search Open AQL	fakkite AQL goronish
Second Degree (Master)	Malammit digri (Majisteer)
Secondary/intermediate school	Malammit daraja/fanti madrasa
see	Ubul
See also	gantah ubul
see the document	Sanad ubul
Social Media	geyshisheenitti Arux
social, juridical and religious texts	dooda, xukme kee diin kutuub

Status	gondo
Subcategory	cindha karrimta
Sub corpora of the corpus	Maskaakat cindha maskaaka
sub corpus type	Cindha Maskaakat cayna
synonym of the word	Sitta-xeela/canqarat aki macna
Synonym or related words of the headword	inki macnale hido yamseene canqurti amot-canqara
The 300 most frequently used words are	Mangih xiyaw tetneefece 300 canqaratya lel xxxx(XXXX) kinni.
the headword	amot-canqara
The word	Canqara
times	gedda
Title	migac
Top five words used	nafaacadde asse agani koona canqara
Total letters should not be less than 5,000 and more than 10,0000	Bukaako xuruuf manga 5000ko guballe takko, hido lel 10,000ko aganalle takke alle mitane.
Total lines	firsim sittiyadde
Total number of letters	Bukaako xuruuf manga
Total unique root words is: XXXX. I did not recognise (XXXXX); but I did recognize (xxxx) which is equivalent of yy.yy%.	Masarrat/rimid canqurti manga XXXX kinni. Tayko (XXXXX) aysharhaarhago dhiiceh aniyo. (xxxx) laakin aysharhaarhago mirhiicinniyo, tayim lel boolko xx.xx% yaabbirhe.
Total unique words	cuzub canqur sittiyadde
Total words	canqur sittiyadde
Total words of the current corpus is: XXXX. The number of identified words is: XXX which is equivalent of xx.xx%.	Ta kaadit Maskaakadde yane canqur hi manga sittiyadde: XXXX kinni. Tayek basoh yimidhdhige canqqur hi manga lel: XXX yekkeh, boolko xx.xx% yabbirhe.
traditional songs, riddles, proverbs, and curses	caydat margadde, simaasimateeni, missila kee abaaro
traditional stories	sugti zanta
translations	tarjama
Translations of the headword	amot-canqarahi tarjama
type of verb that I do not how to generate flexed forms	type of verb that I do not how to generate flexed forms
Unknown	Amidhdhiginnixem
Unknown command	amidhdhiginnixe amre
Upload content to the server corpus	Maskaakat Dhibrho serverille korish

Use of the word	Canqarat nafaaca
used words	Nafacadde-ase canqur
Users' management	yatneefece marahi asxadaadar
verb	Abitto
Warning	Asuuso
We are working on etymology of the base word	Masarrat kin canqarahi maabako akah geynaakah abak nane
word frequency	Canqarat gabbabbac
Word list management	canqurti warar-hi asxadaadar
Word mostly known in this request	ta esserodde mangih timidhdhige canqara
words	canqur
words collected	Yekette canqur
words derived from the root word	rimid canqaraako yewce canqur
Year	iggida
You are connected as	balih sayteh tanito

Saahot to English

Saho	English
5,000 ko lee mangum 10,000 ko lee dago canqur yemmemeceh yane	No less than 5,000 and not more than 10000 words
abayshimat goron	Last search
abe tiya	Creator
Abitto	verb
Abud canqara __ ko	loanword from
addadde	in
adde hammitaanam	Focusing on
adde orobte waqhte	Inserted
adde temmemece waqhte	Edited
Af/waani	Language
Afti digri	First Degree
akke-wayne	Error
akkinnixe URL hido shabaka mitane	No valid URL or no connection
amarmar	Analysing
amaye kee osa	and much more
amidhdhiginnixe amre	Unknown command

Amidhdhiginnixem	Unknown
Amlolokka: Maskaakat muyrhurhug hido dokkumentiiti muyrhurhug kab-ish, kibah lammik maki	Confusing: provide Corpus ID or Document ID - Not both of them
Amot canqara	headwords
amot canqarahi aglel migeytiminna	No combinations found for the headword
amot-canqaraako ginniime aglel	Combinations formed from the headword
amot-canqarahi agleeko yimfillice xibar	Flexed from a combination of the headword
amot-canqarahi maabako, xuggaane, rimid kee tarjama	Etymology, collocations, root and translations of the headword
amot-canqarahi tarjama	Translations of the headword
amot-canqara	the headword
amre	Commands
AQL goronish	Search AQL
Asuuso	Warning
balih sayteh tanito	You are connected as
Bukaako xuruuf manga	Total number of letters
Bukaako xuruuf manga 5000ko guballe takko, hido lel 10,000ko aganalle takke alle mitane.	Total letters should not be less than 5,000 and more than 10,0000
cadar kee margadde	poetry and songs
caddos, inki macnale canqur kee amot-canqarahi nafaaca	Definition, synonym, and use of the headword
Canqara	The word
Canqara hido canqur sayish	insert one or more words
canqara orbish	Insert a word
Canqarat gabbabbac	word frequency
Canqarat maabako	Etymology of the word
Canqarat maabako	Etymology of
Canqarat nafaaca	Use of the word
canqur	words
canqur sittiyadde	Total words
canqurti hoor	concordance
canqurti hoor migeytiminna	No concordance found
canqurti hoor: amot-canqarahi nafaaca	Concordance: use of headword
canqurti manga	Number of words
Canqurti manga	Number of words
canqurti warar-hi asxadaadar	Word list management
caydat margadde, simaasimateeni, missila kee	traditional songs, riddles, proverbs and

abaaro	curses
Cayna	Content type
Cayna	Content subtype
cindha karrimta	Subcategory
Cindha Maskaakat cayna	sub corpus type
cusub	New
cusub goron	New search
cuzub canqur sittiyadde	Total unique words
DB ko maskaakat dhibrho waraarish	List corpus contents from DB
DB-idde cusub dokkumeti os	Add new document to the DB
deesimte cara	Restricted area
dhibrhi warartalle gax	Back to the list of contents
dhibrho	Contents
dhibrho ixfiz	Save the content
digir/catab	plays
document	document
documents	documents
Documents, xato kee ware	Documents, help and news
Dokkumentiiti manga	Number of documents
dokkumentiiti muyrhurhug (ID)	Document ID
Doktoora (PhD)	PhD
dooda, xukme kee diin kutuub	social, juridical, and religious texts
Doorimteh tane Maskaaka mokkaaxishsheh: [malammih bahto icikkin – tayim dherh waqhte beyto dhicta	Analysing selected corpus: [generate it again - it may take long time]
error: check connection to database	error: check connection to database
Expanding the word	Expanding the word
eymemec	Edit
Facebook	Facebook
fafhdhishime canqurti-gabbabbac uynuqhus	Reduce the number of frequency words requested
fafhdhishime canqurti-hoor uynuqhus	Reduce the number of concordances requested
fakkite AQL goronish	Search Open AQL
fanti Maskaakah tuumar hido inki dokkumenti	Reference of the target corpus or single document
faqhra	paragraph
finisira alif	close window

firsim sittiyadde	Total lines
gaazeethath kee siyaasat madare	newspapers and political speeches
gabbabbacnan manga	Number of times
gabbac/malammih ab	Redo
gantah ubul	See also
Ged/waqhte	Era
gedda	times
geli canqara/yatneefece tiyahi muyrhurhug biyashshiite	I forgot userid/passowrd
geli-canqara	Password
geye	I found
geyshisheenitti Arux	Social Media
gondo	Status
Goron/wagiyo	search
Hoor/warar	list
iggida	Year
Inki gedda	only one time
inki macnale hido yamseene canqurti amot-canqara	Synonym or related words of the headword
Inkiddehi xibar	Common behaviour of the
inkoh	All
kaado ginnimtam	Building now
kabbaatime	Approved
kitobti wararta: umbuka cale	list of books: all items
Kutuub 7450 xarfiyaako mifeerishin	Make the writing less than 7450
lemma	lemma
luqhaat goronish	search on languages
Maanafo ab	Create an account
maararo	result
maarhige	I do not know
mablo	Comment
magarot caddos	essays
malammih ab	do it again
Malammit daraja/fanti madrasa	Secondary/intermediate school
Malammit digri (Majisteer)	Second Degree (Master)
Manga	Plural
mangat xibar kinni	is the plural form
mangih gabbabbaacitta 100 canqaratiya	100 most frequently used

Mangih gabbabbaacittem	most frequently used
Mangih nafacadde-ase canqur	mostly used words
Mangih nafacadde-asse bool canqaratiya	Hundred mostly used words
Mangih nafacadde-asse kontom canqaratiya	Fifty mostly used words
Mangih xiyaw tetneefece 300 canqaratya lel xxxx(XXXX) kinni.	The 300 most frequently used words are
Masarrat kin canqarahi maabako akah geynaakah abak nane	We are working on etymology of the base word
Masarrat kin canqarahi yenkeleelebe xibar	Flexed forms of the headword
Masarrat/rimid canqurti manga XXXX kinni. Tayko (XXXXX) aysharhaarhago dhiiceh aniyo. (xxxx) laakin aysharhaarhago mirhiicinniyo, tayim lel boolko xx.xx% yaabbirhe.	Total unique root words is: XXXX. I did not recognise (XXXXX); but I did recognize (xxxx) which is equivalent of yy.yy%.
Maskaaka	Corpus
Maskaaka Baclayshatiyaalih geytim	Contact the Corpus Manager
Maskaaka goronish	Search on corpus
Maskaaka mokkaaxishsheh gabbabbaacittam wararah haa	Analyse corpus and list frequency
Maskaakadde orob	Enter the corpus
Maskaakat ayfaddag cindha maskaakah	Distribution of the corpus by sub corpora: 10 types
Maskaakat ayfaddag gedih	Distribution of the corpus by era
Maskaakat ayfaddag waanit karimih	Distribution of the corpus by Parts of Speech
Maskaakat cayna	Corpus types
Maskaakat cindha maskaaka	Sub corpora of the corpus
Maskaakat cusub dhibrho os	Add new corpus content
Maskaakat Dhibrho serverille korish	Upload content to the server corpus
Maskaakat karim	categories of the corpus
Maskaakat marmarahi wagiinan qalath kinni	Error in request of corpus analysis
Maskaakat muyrhurhugko daggoyemko inketta bah hido inki dokkumentih muyrhurhug kab-ish	Provide at least one of the corpus ID or single document ID
masxassa	Information
meela	Home
migac	Title
Migac	noun
Mokkaaxish	Analyse
nafaacadde asse agani koona canqara	Top five words used

Nafacadde-ase canqur	used words
orbishshem	Inserted by
orob	Login
Post	Post
Qaamus orbeena canqara tiya	Dictionary entry for the word
Qaamusudde wagiyan canqara mamxayinna. Ta guba yane gabbabbaacita hoorko door	No word to search on dictionary is provided. Select from the frequency list below.
Qaamuus	Dictionary
Qaamuus ciyyo	lexicographical works
Qaamuusko wagii	search from dictionary
qalath: amaa kitaab migeytimo	error: could not find that book
qalbi raxa zanta, urhurh dalta, idhdhiteeni faro kee urhurh zanta	novels, short stories, travelogues, and flash fiction
RCF Saahot Maskaaka	RCF Saho Corpus
Remove them all	remove them all
Rimid canqaraako yewce canqur	words derived from the root word
Sanad	document
Sanad	documents
Sanad ubul	see the document
Sarroorimalle goronsime canqur	Last words searched
Sarroorimalle yetneece mari yumbulluyin	Last users seen
Sayeh yane	inserted
Saynes, magarot caddos, kee kutuub	science and mathematics textbooks, essays, and articles
Server ko modobba mitane	no answer from the server
Sitta-xeela akkinnixe canqur geytime	non synonym words found
Sitta-xeela/canqarat aki macna	synonym of the word
Sugti zanta	traditional stories
Ta canqara adde tane walaa inki dokkumenti migeytiminna	could not find any document in which the word is used
Ta canqarah siddimat canqara akah migeytiminna	No base word found for this word:
Ta dokkumentidde yane xuruufhi manga.	Number of letters
Ta esserodde mangih timidhdhige canqara	Word mostly known in this request
Ta kaadit Maskaakadde yane canqur hi manga sittiyadde: XXXX kinni. Tayek basoh yimidhdhige canqqur hi manga lel: XXX yekkeh, boolko xx.xx% yabbirhe.	Total words of the current corpus is: XXXX. The number of identified words is: XXX which is equivalent of xx.xx%.

Tamaa cindha ishi maskaakaako XXXXXX ta dokkuments asac-ishe kee adde ose tik. Kadit waqhtedde ku cindha maskaakadde walaa inki dokkumenti mitaneeki, guballe oobeh yane hoorko doorteh adde os.	Add or remove these documents to the personal sub corpus of XXXXXX. There are not at the moment documents in your personal sub corpus, select from the dropdown list and add.
Tarjama	translations
Tarjama migeytiminna	no translations found
Tarjamadde goronish	search on translations
Tati yuktubeh yane dokkumentit imirmir	Analyse documents by this author
Tatnaafaco izne malito.	No permission to use.
Teyim geytimo mirhiicinna	could not find this
Type of verb that I do not how to generate flexed forms	type of verb that I do not how to generate flexed forms
Ubul	see
Ummanimko ziyaada goronsime canqur	Most searched words
Waanit karim	Parts of Speech
Xab	cancel
Xeelta	is synonym of
Xunsus	notes
Xuruuf manga	Number of characters
Yambedeedele xibar	inflected forms
Yatneefece mara wararah haa	List users
Yatneefece marahi asxadaadar	Users management
Yatneefece tiyahi Addoodim eymemec	Edit user profile
Yekette canqur	words collected
Yemmemece	edited
Yi Maskaaka	My corpus
Yuktube tiya	Author
Yuktube tiyahi migac	Author name
Zamoodinno	Relation