
Problems in Swahili Lexicography¹

Albina R. Chuwa, *Institute of Kiswahili Research,
University of Dar es Salaam*

Abstract: Swahili dictionaries have been on the market since 1882. Most of these dictionaries were compiled by foreigners who were either missionaries or political administrators working in East Africa. Although the history of Swahili dictionary compilation seems to have survived a number of years, the field of lexicography is still a very new field in relation to what has been done by Tanzanians in this field. The first Swahili monolingual dictionary by a team of Tanzanian amateurs in dictionary compilation was published in 1981.

Due to inexperience in the field of lexicography the compilers have been facing various kinds of problems arising either from the technical level (whereby solutions for such problems are beyond the dictionary compilers) or from the practical level (whereby solutions depend more on the working team). When the two levels intertwine it makes the work of the dictionary compilers even more difficult, thus requiring support from outside the team. The sample problems discussed in this paper, though not exhaustive, pose some difficulties to Swahili dictionary compilers. Among those problems are:

1. Fiscal constraints which force compilers to embrace a big user target, making the dictionary objectives too wide to address any group satisfactorily.
2. The choice of language (standard versus non-standard varieties) inhibits the expansion of existing vocabulary in Swahili dictionaries.
3. The identification of lemmas and their grammatical categories still requires special attention from Swahili grammarians and structuralists.

Keywords: AMATEUR, CATEGORY, COMPILER, CORPUS, DATA BANK, DIALECT, DICTIONARY ENTRY, FREQUENCY, GRAMMAR, GRAMMATICAL CATEGORY, HEADWORD, LEMMA, LEXICOGRAPHY, MEANING, OBJECTIVE, PROBLEM, SELECTION (OF ENTRY), SOURCE, STANDARD DICTIONARY, SYNONYM, TARGET

Opsomming: Probleme in Swahili leksikografie. Swahili-woordeboeke is op die mark sedert 1882. Die meeste van hierdie woordeboeke is saamgestel deur buitelanders wat óf sendelinge óf politieke amptenare werksaam in Oos-Afrika was. Alhoewel die geskiedenis van die samestelling van Swahili woordeboeke daarop dui dat dit al 'n hele aantal jare bestaan, is die leksikografie steeds 'n baie nuwe veld betreffende werk gedoen deur Tanzaniërs. Die eerste Swahili eentalige woordeboek saamgestel deur 'n span Tanzaniese woordeboekamateurs is in 1981 gepubliseer.

Weens 'n gebrek aan ervaring in die leksikografieveld ervaar die samestellers verskillende soorte probleme wat spruit uit óf die tegniese vlak (waar oplossings buite die vermoë van die

1 This paper was presented at the First International Conference of the African Association for Lexicography, held at the Rand Afrikaans University, Johannesburg, 1st - 2nd July, 1996.

woordeboeksamestellers is) óf die praktiese vlak (waar oplossings meer by die werkspan berus). Wanneer die twee vlakke saamval, maak dit die taak van die woordeboeksamestellers selfs nog moeiliker, wat steun van buite die span vereis. Die voorbeeldprobleme wat in hierdie artikel bespreek word, alhoewel nie uitputtend nie, lewer probleme op vir die Swahili-woordeboeksamestellers. Onder hierdie probleme is:

1. Fiskale beperkings wat samestellers dwing om 'n omvattende doelmark in aanmerking te neem, wat die woordeboekdoelwitte te wyd maak om bevredigend in die behoefte van enige groep te voorsien.
2. Die keuse van taal (standaard- versus niestandaardvariëteite) beperk die uitbreiding van die bestaande woordeskat in Swahili-woordeboeke.
3. Die bepaling van lemmas en hul grammatikale kategorieë vereis steeds spesiale aandag van Swahili-grammatici en -strukuraliste.

Slutelwoorde: AMATEUR, KATEGORIE, SAMESTELLER, KORPUS, DATABANK, DIALEK, WOORDEBOEKINSKRYWING, FREKWENSIE, GRAMMATIKA, GRAMMATIKALE KATEGORIE, TREFWOORD, LEMMA, LEKSIKOGRAFIE, BETEKENIS, DOELSTELLING, PROBLEEM, SELEKSIE (VAN INSKRYWING), BRON, STANDAARDWOORDEBOEK, SINONIEM, DOELWIT

1. A Historical Perspective of Swahili Lexicography

The Swahili language, widely spoken in East Africa (Kenya, Uganda and Tanzania) and parts of the bordering countries (Rwanda, Burundi, Zaire and the Northern areas of Zambia, Malawi and Mozambique), has mainly depended on dictionaries compiled by foreigners who came to East Africa either as missionaries, colonial administrators or tourists.

Swahili dictionaries can be grouped into three publishing phases. In the first phase we find dictionaries published in the last decade of the 19th century and these include the works of the following personalities: Krapf, *Swahili-English Dictionary* (1882); Sacleux, *Dictionnaire Français-Swahili* (1891); Madan, *English-Swahili Dictionary* (1894) and *Swahili-English Dictionary* (1903). The second phase covers the period from 1939 to 1958. In this second phase there are works of Sacleux, *Dictionnaire Swahili-Français* (1939), Johnson F., *Kamusi ya Kiswahili yaani Kitabu cha maneno ya Kiswahili (A Swahili monolingual dictionary)* (1935), Johnson F., *A Standard Swahili-English Dictionary* (1939a), Johnson F., *A Standard English-Swahili Dictionary* (1939b) and Snoxall, *A Concise English-Swahili Dictionary* (1958). The third phase runs from 1981 to the present day. In this last phase a number of dictionaries have been published, such as the Institute of Kiswahili Research's² (henceforth IKR) *Kamusi ya Kiswahili Sanifu* (1981) (*A Standard Swahili Dictionary* — henceforth KKS), Lenselaer, *Dictionnaire Swahili-français* (1983), Miachina, *Kamusi ya Kiswahili-Kirusi* (1987)

2 In the reference section this Institute appears in its Swahili name 'Taasisi ya Uchunguzi wa Kiswahili'.

(*Swahili-Russian Dictionary*), Legere, *Wörterbuch Deutsch-Swahili* (1990), Baba Malaika, *The Friendly Modern Swahili-Modern English Dictionary* (1991) and Bakhressa, *Kamusi ya Maana na Matumizi* (1992) (*A Dictionary of Meaning and Usage*). Members of the lexicographical section of IKR have been involved in different capacities either as compilers or as reporters and the following has been realized through their joint efforts: In 1991 the IKR published *Kamusi ya Isimu na Lugha* (A linguistic dictionary), plus *Kamusi ya Biolojia, Fizikia na Kemia* (A dictionary of Biology, Physics and Chemistry). In 1992 IKR was co-author of the *French-Swahili Dictionary*, and in 1996 co-author of the *Swahili-French Dictionary*. Other lexicographical projects of IKR include the 2nd edition of the *Kamusi ya Kiswahili Sanifu* (in press); and the *English-Swahili Dictionary* (in press).

As seen from the above summary Swahili lexicography has focussed both on monolingual and bilingual dictionaries. The most used dictionaries, particularly in Tanzania, have been those of Johnson and the IKR.

2. The Art of Dictionary Making

Looking back from the time when the first Swahili dictionary by Tanzanians was published, and in relation to what has been done by Tanzanians in the field of lexicography, one would concur with the idea that lexicography is still a very new field in Tanzania. As mentioned above, Swahili dictionary compilers, especially in the 1st and 2nd phase, had been foreign missionaries or colonial administrators who depended on their linguistic intuition plus information they collected from their informants. Johnson's dictionaries in particular have valuable linguistic information which tell us that Johnson had a good linguistic background which has made his dictionaries authoritative and much depended upon for such a long time. Most dictionary compilers who published after 1939 used his dictionaries as their main sources.

The first Swahili dictionary by Tanzanians, the KKS, was compiled by a team of amateurs. The only expert was the team leader, Prof. R. Ohly, a Polish citizen who was then employed by the University of Dar es Salaam. With the little experience shared by the dictionary team, they very much depended on Johnson's dictionaries plus their personal intuition. The team set out to compile a standard language dictionary which was not well defined. The working team consisted of people not only of different ethnic groups but also people from different Swahili dialects. With such a mixed group therefore we cannot rule out the possibility that every one of them would have tried to include in the so-called standard dictionary some words from his own dialect. All these contributed to the rather poor quality of the product in comparison to Johnson's dictionaries. Too many synonyms and homonyms found their way into the *Standard Swahili Dictionary*.

Experience shows that like any new undertaking, several problems are

encountered by dictionary compilers, particularly when most of them are either amateurs or have little experience.

Dictionaries made by IKR after the KKS of 1981 had more advantages because the team included trained lexicographers. This factor has helped to speed up the process of dictionary making and a more critical approach was taken especially in the selection of entry words. Swahili lexicography has faced several problems and the main ones are discussed below.

3. Problems Encountered in Swahili Lexicography

3.1 Dictionary Objectives

The early Swahili dictionaries were compiled by people who wanted to use the product as one of their tools in performing whatever task was ahead of them. As such there was no earmarked target user — it was left open for all individuals who could get access to it. Such a target is too general and as such it is rather difficult to exhaust the language. Because they lacked identified users for their product the objectives were too wide to be embraced to the satisfaction of all users.

The KKS, published forty-two years after the first monolingual dictionary, also faced the problem of generality of objectives. It mentions a very wide range of users which would definitely create a problem in trying to satisfy them all. If we take the objectives mentioned in KKS we read '... it is targeted to students of all levels, teachers, writers, radio listeners, readers of books and newspapers...' (my translation — KKS 1981: xii). To satisfy such a wide range of users in a small-sized dictionary is too big an ambition. A primary school pupil definitely needs a more elementary dictionary whereas a college student requires a more advanced dictionary.

3.2 Choice of Language in Swahili Lexicography

Of the Swahili dictionaries mentioned above only KKS identified the kind of Swahili that was to be featured. IKR set out to compile a *Standard Swahili Dictionary*. The Inter-Territorial Swahili Language Committee under which Johnson published his dictionaries had chosen the dialect of Swahili spoken in Zanzibar at that period as the standard language. But because Swahili was used all over East Africa in the midst of other languages, and the fact that the language was rapidly growing, it lost more and more of its originality by incorporating words from other dialects and ethnic languages found in the mainland.

The problem of compiling a *Standard Swahili Dictionary* as pointed out by IKR arises from the fact that it was very difficult to identify the standard lan-

guage. The basic source of KKS was Johnson's dictionaries which had relied on informants and Johnson's linguistic knowledge. It is difficult to record a standard language without using documented materials which help to give an idea on the frequency use of the chosen words plus their meanings.

There is a time lapse from the time when Johnson's dictionaries were published to the time when the first Swahili monolingual dictionary by Tanzanians was compiled and published in 1981. As such one would expect KKS to be richer and more informative as a result of language growth. But this has not totally been the case because when KKS was declared a standard dictionary it restricted itself too much to the number of entry words purported to be standard, thus leaving dialectal and less standard words out.

Looking at what has so far been published in Swahili dictionaries we can conclude that very little has been done to increase the number of entries to cater for the growing language. This weakness can be attributed to the following problems. The first is the lack of a Swahili corpus — the main question remains — how did KKS collect data for a standard dictionary without a language corpus?

The question posed above still holds because although the Swahili language is rich in printed material covering a wide range of topics, it still lacks modern technology which would allow the team to build up a corpus and a data bank that can be accessed and easily retrieved by dictionary compilers. Lack of modern technology still force dictionary compilers to depend on their intuition plus that of a few informants, thus leaving out a substantial amount of valuable information which is essential not only to Swahili language learners but also to the daily users.

Depending on informants and the intuition of compilers contributed to the poor choice of entry words plus definitions which do not identify the defined. A lexicographical research conducted by Mdee (1996 — in press) on the flora and fauna featuring in KKS revealed that a number of the items listed were not recognized by his informants. For example, he had a total of thirty-five queries on animal names out of which ten were identified by an expert in wild life and only thirteen were identified by an informant who had participated in compiling that same dictionary. Twenty two names could not be identified by both of them. This indicates how much more research is required before we have a standard Swahili dictionary.

Another Swahili monolingual dictionary published in Kenya by Bakhressa based its work on KKS with no new information except for the inclusion of illustrative sentences on all headwords.

3.3 Lemma and Grammar in Swahili Dictionaries

3.3.1 Grammar in Swahili Dictionaries

Compilers of Swahili dictionaries have always used foreign grammar to

describe features found in the Swahili language. When describing grammatical categories the English classification has been used without taking into consideration the differences existing between the two languages. The question of the exact number of Swahili grammatical categories is still a crucial issue to be solved by Swahili grammarians. Identification of Swahili grammatical categories, especially the functional words, is essential if all words existing in the language are to receive fair treatment in dictionaries.

3.3.2 Lemma Form in Swahili Dictionaries

The Swahili language is an agglutinative language which requires some skills in identifying canonical forms that are acceptable as dictionary lemmas. Not all research assistants paid to collect data from printed Swahili materials came out with the required information. A quick review of dictionary cards collected by inexperienced research assistants hired by the Institute of Kiswahili Research (1994) revealed that agglutinated forms were registered as dictionary lemmas. For example:

a-tu-wafikish-e (= subject prefix + object marker + verb + imperative ending = he pacify us)
a-na-kwawa (= subject prefix + present tense marker + verb = he is hesitating)

Obviously one would have expected the two quoted cards above to represent lemmas *wafikisha* and *kwawa* instead of the clauses *atuwafikishe* and *anakuwawa* respectively.

3.3.3 Potentiality of Derivational System

The derivational system of Swahili has great potentiality in producing new words. There are a lot of words created through noun and verb derivations. The verb derivational system makes use of suffixes to express such features as applicative (*-ia, -ea*), passive (*-wa*), reciprocal (*-ana*), and stative (*-ka*), plus a prefix for the reflexive form (*ji-*). Although the system is well known to users of the language, it is still not very consistent. Some derivatives carry extra shades of meaning from those realized in the canonical form. KKS (1981) indicated the derivational suffixes at the end of a verb entry with no further comments. Baba Malaika (1991) went a step further by including derived forms of verbs as dictionary lemmas. For a more systematic approach whereby no redundancy is allowed into dictionaries and whereby words with specific meanings are accorded fair treatment in Swahili dictionaries, more research, funds and time should be allocated to the compilers. This would guarantee a better selection of entry words for future Swahili dictionaries.

4. Conclusion

We envisage that due to fiscal problems and time taken in compiling a dictionary, any substantial work on lexicography in Tanzania will still be under institutions for a longer time. Both sides, institutions and compilers, should therefore work more closely together so that the projects put forward are allocated enough time and resources to be able to deliver the goods that will meet the required standard.

References

- Baba Malaika. 1991. *The Friendly Modern Swahili Modern English Dictionary*. MS-tryk Denmark.
- Bakhressa, Salim, K. 1992. *Kamusi ya Maana na Matumizi*. Nairobi: Oxford University Press.
- Johnson, Fredrick. 1935. *Kamusi ya Kiswahili yaani Kitabu cha maneno ya Kiswahili*. Nairobi, Dar es Salaam: Oxford University Press.
- Johnson, Fredrick. 1939a. *A Standard Swahili-English Dictionary*. Nairobi, Dar es Salaam: Oxford University Press.
- Johnson, Fredrick. 1939b. *A Standard English-Swahili Dictionary*. Nairobi, Dar es Salaam: Oxford University Press.
- Krapf, L.A. 1882. *Swahili-English Dictionary*. London: Trubner and Company Ludgate Hill.
- Legere, K. 1990. *Wörterbuch Deutsch-Swahili*. Leipzig: VEB Verlag Enzyklopädie.
- Lenselaer, A. 1983. *Dictionnaire Swahili-français*. Paris: Éditions KARTHALA.
- Madan, A.C. 1894. *English-Swahili Dictionary*. London: Oxford University Press.
- Madan, A.C. 1903. *Swahili-English Dictionary*. London: Oxford University Press.
- Mdee, James S. 'Nafasi ya Msamiati usiotumika katika Kamusi ya Lugha Sanifu. Chuwa and Kihore (Eds.). 1996. *Kiswahili Katika Karne ya 21*. Institute of Kiswahili Research. (in press)
- Miachina, E.N. 1987. *Kamusi ya Kiswahili-Kirusi*. Moscow: Idara Ya Uchapishaji ya Lugha ya Kirusi.
- Sacleux, Charles. 1891. *Dictionnaire Français-Swahili*. Paris: Institut d'Ethnologie.
- Sacleux, Charles. 1939. *Dictionnaire Swahili-Français*. Paris: Institut d'Ethnologie.
- Snoxall, R.A. 1958. *A Concise English-Swahili Dictionary*. Nairobi: Oxford University Press.
- Taasisi ya Uchunguzi wa Kiswahili. 1981. *Kamusi ya Kiswahili Sanifu*. Nairobi/DSM: Oxford University Press.
- Taasisi ya Uchunguzi wa Kiswahili. 1991a. *Kamusi ya Isimu na Lugha*. Dar es Salaam: Dar es Salaam University Press.
- Taasisi ya Uchunguzi wa Kiswahili. 1991b. *Kamusi ya Biolojia, Fizikia na Kemia*. Dar es Salaam: Dar es Salaam University Press.
- Taasisi ya Uchunguzi wa Kiswahili, CREDU and Department of Foreign Languages and Linguistics. 1992. *French-Swahili Dictionary*. Nairobi: CREDU.
- Taasisi ya Uchunguzi wa Kiswahili, CREDU and Department of Foreign Languages and Linguistics. 1996. *Swahili-French Dictionary*. Nairobi: CREDU.
- Taasisi ya Uchunguzi wa Kiswahili. *Kamusi ya Kiswahili Sanifu*. (Second Edition — in press).
- Taasisi ya Uchunguzi wa Kiswahili. *English-Swahili Dictionary*. (In press).