# Another look at Lexicographic Parameters

by
Vincent M. Chanda
University of Zambia
Email: mchandav@gmail.com

==========

*Abstract*

*All lexicographers and, presumably, most dictionary users know that there are several types of dictionary, which include, among others, general dictionaries, specialized dictionaries (also called special purpose dictionaries), concise dictionaries and pocket dictionaries. They also know that not all dictionaries provide the same amount of information about headwords and other types of dictionary entry heads. What this means is that all dictionaries share a set of principles, which is why they are all dictionaries, provided by the definition of a dictionary. This paper uses insights from a linguistic theory known as Principles and Parameters (P&P) to give names to the various factors or options explaining the existence of various types of dictionaries as well as differences among dictionaries belonging to the same type. The paper proposes six parameters, namely a) the One-Language Parameter, underlying the monolingual/ bilingual / multilingual trichotomy, b) the User Parameter, underlying the differences among dictionaries meant for specific groups of users, c) the History Parameter, underlying the fact whether a dictionary is etymological or not, d) the Scope Parameter, underlying the difference*

*in the amount of content among dictionaries, (e) the Head Parameter, underlying the nature of dictionary entry heads such as the so-called headwords and f) the Language Parameter, underlying the fact that differences in contents among dictionaries may be due to differences in grammar among the languages. Bearing in mind the concepts of Principles and Parameters, the paper also examines, before concluding, the structure of dictionary entry heads and the information about them.*

*Keywords:* Dictionary, lexicography, dictionary entry, dictionary entry head, parameter, One-Language Parameter, User Parameter, History Parameter, Scope Parameter, Head Parameter, Language Parameter.

*Introduction*

Lexicography, whether theoretical or practical, is the science of dictionary-making and is part of applied linguistics. A survey of the definitions of the dictionary shows that most of them state that a dictionary is a book made of items, usually words, arranged in alphabetical order, each followed by information about them. The most important and frequent piece of information provided about an 'item' is either (a) meaning in the same language, for monolingual dictionaries, or (b) the equivalent in one or more other languages, for bilingual and multilingual dictionaries. However, as will be seen below, such definitions do not capture the nature of all dictionaries. As a matter of fact, while all dictionaries share a number of principles, there are parameters, that is, options or factors that explain not only the existence of

several types of dictionaries, apart from the distinction between monolingual dictionaries, on the one hand, and bilingual and multilingual dictionaries, on the other. This being the case, the number of languages involved is just one of the parameters, or factors. As will be seen, there are many other parameters, or factors, determining other distinctions.

The set of the item and the information provided about it is generally known as a dictionary entry or lexical entry while the item about which some information is given is called headword, entry head, main entry head or citation form. In this paper, the term 'entry head' will be used, to avoid the terms 'headword' 'entry word' and 'main entry word' because not all items about which some information is given are words: as will be seen, some of such items are non-words such as morphemes or phrases. It is worth noting that, in a few types of dictionaries, some or all such items are sentences: this is the case of dictionaries of proverbs.

## The Structure of a Dictionary

Most definitions of the dictionary state that a dictionary is a list of items, in the form of words and phrases, followed by some information, especially and, sometimes only meaning. However, a dictionary, as a book, also comprises other parts, namely a cover page and, often, other parts such as a preface, or foreword, a list of abbreviations, a grammatical digest, etc. Hence, the following tripartite dictionary is sometimes made regarding the structure of a dictionary:

(i) Megastructure: the structure of the dictionary as a book;
(ii) Macrostructure: ordering of dictionary entries, the primary components of the body of a dictionary; and
(iii) Microstructure: the internal structure of any dictionary

entry.

The body of a dictionary is made of entries, often referred to as dictionary entries or lexical entries, each entry having two parts, namely a headword, or entry head, followed by some information, especially and, in most cases, including meaning. Let A, B and C stand respectively for dictionary entry, entry head (in the sense of 'dictionary entry head') and information about the entry head. The structure of the dictionary may be represented, in formula form, as in (1):

(1) Structure of the body of a dictionary

a. dictionary = An

b. A = B + C

Although (1) applies to all dictionaries and is, therefore, a universal Dictionary Principle, there are several types of dictionaries, for example determined by factors referred to in this paper as parameters. One such parameter, called Language Parameter, determines the existence of monolingual dictionaries, bilingual dictionaries and multilingual dictionaries. Being a key term in the paper's title, the term 'parameter' is discussed in the following section.

## Some Dictionary Parameters

### *Parameters in linguistic theory*

In contemporary linguistics, the term 'parameter' often recalls to mind Chomsky's Principles and Parameters Theory (P&P), a syntactic theory with two versions known as Government-Binding (GB) theory and Minimalist Program (MP), respectively. The following is a brief account of P&P, motivated by the fact that the meaning of the term 'parameter' as used in both P&P and this paper is the same.

The basic idea of P&P is that all natural languages have the

same fundamentals, which are the subject matter of Universal Grammar (UG), comprising (a) principles, applying in the same way in all human natural languages, and (b) parameters, in the sense of alternatives or options, that explain syntactic differences among languages. One of the parameters is the Null-Subject Parameter, stating that, with regard to the presence or absence of subjects in general, there are languages where in principle every sentence must have an overt subject and languages may have grammatical sentences with no overt subject. Thus, while in English all sentences, apart from imperatives must have an overt subject, in Latin, Italian and some other languages, including all Bantu languages it is normal to have a sentence without overt subjects, as exemplified by (2) and (3), showing that 'Banda', the subject, cannot be dropped in English but can be dropped in Bemba, a Bantu language:

(2) ENGLISH

a. Banda is sleeping

b. *is sleeping

(3) BEMBA (Guthrie's M42)

a. Banda naalaala 'Banda is sleeping'

b. naalaala 'he/she is sleeping'

In Principles and Parameters (P&P), parameters posited by Chomsky and others include. Inter alia:

- Null-Subject Parameter, exemplified above;
- Pro-Drop Parameter, distinguishing languages where some pronouns can be or are dropped in some contexts from those language where this is not the case; and
- Polysynthesis Paramater, distinguishing synthetic, or agglutinative, languages, for example Bantu languages, from isolating languages, such as Chinese.

This paper is an application of Principles and Parameters (P&P) as it takes the view that the presentation and structure of dictionary entries are determined both (a) a set of universal principles, that is, principles followed by all dictionaries, and (b) parameters, that is, variables or factors, explaining the differences among dictionaries, including the typology of dictionaries.

## Background to the proposed dictionary parameters

The parameters proposed below are based on the formula in (1), above, reproduced for convenience sake as (4) below:

(4) Structure of the dictionary

a. dictionary = An

b. A = B + C

As posited for (1), "A, B and C stand respectively for dictionary entry, entry head (in the sense of 'dictionary entry head') and information about the entry head".As will be seen below, there are parameters for A, B and C.

*The One-Language Parameter*

Any dictionary entry may be represented by the following formula, where A, B and C stand respectively for a dictionary entry, the item about which some information is provided and the information provided about the item:

(5) A = B + C

If B and C are given in the same language, the dictionary is monolingual; otherwise, the dictionary is either bilingual or monolingual. In monolingual dictionaries, both entry heads and the information about them are in the same language. In bilingual dictionaries, entry heads are in one language and the information

about them in another language. In multilingual dictionaries, entry heads are in one language and the information about them in more than two other languages.

*The User Parameter*

The term User Parameter is proposed in this paper to explain mainly the following two typologies of dictionaries:

- General dictionaries versus specialized dictionaries; and
- Learner's dictionaries, intermediate learner's dictionaries and advanced learner's dictionaries.

*The History Parameter*

Some dictionaries are etymological dictionaries. These are dictionaries dealing with the origin and historical development of words. Examples of such dictionaries are:

- Etymological Dictionary of Basque (Trask, 2008); and
- Etymological Dictionary of the English Language (Oxford University Press, 1963)

The following two points are worth noting. First, for non-etymological dictionaries, that is, dictionaries dealing with contemporary languages, it is not indicated, usually, that they are not etymological, If this indication is provided, the term 'contemporary' is used but as an adjective modifying the name of language, for example Dictionary of Contemporary English rather than Contemporary Dictionary of English. And second, some non-etymological dictionaries do provide some etymological information at least for a few items.

*The Scope Parameter*

A lexicographer may determine the scope of the dictionary, to produce, among others:

- Pocket dictionaries;
- Concise dictionaries; and
- Encyclopedic dictionaries, also spelt 'encyclopaedic dictionaries.

*The Head Parameter*

The parameter proposed here is not to be confused with the Head Parameter in Chomsky's Principles and Parameters (P&P) theory, referring to the place of the head of phrase vis-à-vis its complements, determiners or modifiers, Using the formula A = B + C, above, where A, B and C stand respectively for a dictionary entry, the item about which some information is provided and the information provided about the item, the term 'Head Parameter' as used here refers to the content of B and C. In most dictionaries, what is to be given information about precedes the information: the entry head, B, precedes the information, C. Such dictionaries, which are the majority, are known as onomasiological dictionaries, meaning going from a term to the information, mainly semantic information, about it. There are a few dictionaries in which the opposite happens: the entry head, B in the above formula, is some information and C is the item fitting that information. This is the case of crossword dictionaries or glossaries. The general term for such dictionaries is reverse dictionary, a good example of which is The Oxford Reverse Dictionary (Edmonds, 1999).

*The Language Parameter*

The term 'Language Parameter' is different in meaning from the term 'One-Language Parameter, the name of the first parameter in this paper, is that the One-Language Parameter is about the

number of languages involved while the Language Parameter concerns the nature, in terms of grammar, of a language. The concept of language and its meanings can be the same or related with minor variations (Mkandawire, 2018), but the paramenters in the context stated above is different.

Following the Principles and Parameters (P&P), an instance of Universal Grammar (UG), it may be said that languages are both similar and dissimilar: they are similar in that they have the same set of principles and they differ because of the parameters. It follows that dictionaries of two languages cannot have all their entries identical. However, the number and nature of the differences among languages is smaller in languages with same ancestor language than among languages which are not genealogically related. This has been shown, inter alia, by Kiango's (2005.) Problems of Citation Forms in Dictionaries of Bantu Languages, dealing with what is known in lexicography as lemmatization when the term is used to refer to the choice of the words or/and word forms to be entered as what are called citation forms, another term for headword.

The above-mentioned Kiango's work will be amply used, or referred to in Section 4 dealing with lemmatization, as it has well illustrated what is called Language Parameter in this paper.

## The Structure of Entry Heads

*The syntax of entry heads*

In human natural languages, words are often divided, morphologically, into three categories, namely:

- Simplex word, a word made of one morpheme, for example mouth in English;

- Complex word, a word made of a set of morphemes comprising only one root or stem; for example the English word worker, in which 'work' is a root and 'er' a suffix, and the Swahili word atampiga 'he/she will beat him/her up', in which pig 'beat up' is a root and the other morphemes are affixes; and
- Compound word, made of two or more simplex words or parts of simplex words, for example the English words gunpowder, made of the simplex nouns gun and power, and great fathers. Made of the adjective great and the complex noun fathers, made of the simplex noun father and the plural suffix s.

As a general practice, in dictionaries where entries are not morphemes but words, which is the normal practice, entries are either simplex words or compound words while, for complex words, only grammatically irregular complex words are listed among entries.

Regarding the syntactic nature of entry heads in dictionaries of human natural languages, the following two questions in (6), below, are unavoidable:

(6) a. Are all entry heads words?

b. If the answer to (6a) is positive, can an entry word belong to any word category as described above, that is, a simplex word, a complex word or a compound word?

Let us start with the last question because it is simple. The answer is positive: an entry head can be a simplex word, a complex word or a compound word, as exemplified in English by man, manhood and mankind, respectively.

The answer to (6a) is 'yes' and 'no'. The main cases encountered are as follows:

- Dictionaries in which all entry heads are words;

- Dictionaries in which most entry heads are words and a few are morphemes, especially lexical morphemes; and
- Dictionaries in which all entry heads are sentences, which, to best knowledge of this writer, is only the case of dictionaries of proverbs.

Entry heads that are morphemes do exist but are rare. One example found in Longman Dictionary of Contemporary English is –ee, a suffix in words like employee.

Another type of non-word entry head is found in many Bantu dictionaries, where verb entries are infinitive forms from which the infinitive marker is dropped, as shown in . Perrott's (1965) Concise Swahili and English Dictionary, Swahili-English/ English-Swahili Dictionary:

(7) Swahili (Perrott, 1965)

a. ona to see; feel

b. onana to meet

The forms in (7) are not words but represent infinitives, kuona 'to see/feel' and and kuonana literally meaning 'to see each other'or 'to see one another', from which the infinitive marker, the prefix ku is dropped, a practice found in many Bantu dictionaries. It is to be noted, however, that in some Bantu dictionaries where the practice is used, the verb entry heads are preceded by a hyphen, to mean that something, in this case the infinitive marker, has been dropped. This is found, for example in The White Fathers' Bemba-English Dictionary (The Society of the Missionary for Africa (White Fathers), 1991), as exemplified in (8), below, by -kaka and –kakula, which stand respectively for ukukaka and ukukakula, in which the initial u is an augment preceding the infinitive marker, or prefix, ku, which means that both the augment and the infinitive

prefix are dropped:

(8) BEMBA (White Fathers, 1991; Chanda & Mkandawire, 2013)
  a. -kaka ‘to tie’
  b. kakula ‘to untie/unfasten

At the beginning of Section 4.1, it was stated that there are three major categories of words, in natural languages, namely (a) simplex words, (b) complex words and (c) compound words, the latter being made of “two or more simplex words or parts of simplex words”. However, a survey of dictionaries shows that there are several types of multi-word entry heads, for examples (a) multi-word proper nouns, such as full anthroponyms, such as Noam Avram Chomsky, names of many states, for example United Kingdom and South Africa. In some specialized dictionaries, for example dictionaries of mathematics and dictionaries of law, or legal terms, multi-word terms are, more often than not, entry heads in their own rights Thus, in Clapham and Concise Oxford Dictionary of Mathematics, there is an entry head ‘Gauss, Carl Friedrich’ and separate entry heads for concepts or theories developed or proposed by the German mathematician and astronomer alone or combined with concepts or theories developed by others, listed in (9) below.

(9) The entry head Gauss, Carl Friedrich and related entry heads
  a. Gauss, Carl Friedrich
  b. Gaussian distribution
  c. Gaussian elimination
  d. Gaussian function
  e. Gaussian integer
  f. Gaussian plane
  g. Gauss-Jordan elimination

h. Gauss-Markov Theorem
i. Gauss-Seidel iterative method
j. Gauss's Lemma

Similar examples from Daintith's A Dictionary of Chemistry (Daintith, 2008) are provided in (10 ).

(10) The entry head carbon and related entry heads in the Oxford Dictionary of Chemistry

a. carbon
b. carbon assimilation
c. carbon cycle
d. carbon dating (radiocarbon dating0
e. carbon disulphide carbon bisulphide
f. carbon fibres
g. carbon monoxide
h. carbon suboxide

What is to be gathered from the discussion we have had so far is twofold, namely that (a) in general dictionaries, viz. dictionaries meant for the general public, whether a multi-word term is provided as an autonomous entry head is a matter of an author's or publisher's choice, while (b) in specialized dictionaries, mutli-word entry heads are given as autonomous citation forms.

As stated at the opening of this section, there are dictionaries where some entry heads are morphemes, for example –ness in Longman Dictionary of Contemporary English (Pearson Education Limited, 2003: 1102) and dictionaries where some entry heads are acronyms or other types of abbreviations, for example USA for 'United States of America.

*Lemmatization*

In 4.1, above, it was pointed out that not all entry heads are words and that on the basis of the syntactic nature of entry there are several types of dictionaries, including the following, among others (a) dictionaries in which all entry heads are words, (b) dictionaries in which most entry heads are words and a few are morphemes, especially lexical morphemes and (c) dictionaries in which all entry heads are sentences, which, to best knowledge of this writer, is only the case of dictionaries of proverbs.

The term lemma, of which the plural form is lemmas or lemmata, as used in lexicography, refers to an entry head which is (a) a simplex word, that is, not a complex word or compound word, (a) a word from or *c) a morpheme. Hence the term lemmatization refers to the choice or determination of what are to be picked as lemmas.

As was stated in 2.7, lemmatization is determined by a parameter referred in this paper as the Language Parameter, as it depends on the structure of a language.

One good example of the importance of the Language Parameter, is the way verbs are lemmatized in various languages. Consider the following examples of verb lemmatization:

(11) ENGLISH
a). kick
b). justify

(12) DUTCH
a). lachen ‘to laugh’
b). komen ‘to come’

(13) FRENCH
a). aimer ‘to like/love

b). ‘rire ‘to laugh’

(14) SWAHILI

a. cheza ‘to play/dance’

b. lima ‘to plough/cultivate’

(15) LATIN (Collins, 2011)

a. amo, -are, -avi, -atum ‘to love/like’’

b. lego,-ere -legi. lectum’

A comparison of the data in (11) through (16) shows the following:

- In English (11), Dutch (13) and French (13), the verb form lemmatized is the positive infinitive;
- In Swahili (14), the verb form lemmatized is the positive infinitive form which the infinitive marker, or prefix, ku is dropped, the full positive infinitives being kucheza and kulima;
- In Latin (15a), the verb form lemmatized is the positive indicative present in the 1st person singular, amo, followed by –are, the infinitive ending, the suffix –avi, of the 1st person singular positive perfect tense, followed by –atum, the ending of the singular neuter past participle.
- In Latin (15b), the verb form lemmatized is the positive indicative present in the 1st person singular, lego, followed by –ere, the infinitive ending, the suffix legi, the full form of the 1st person positive perfect tense, followed by lectum, the full form of the singular neuter past participle

Three points are made here. First, we notice that in Latin the lemma is followed by grammatical information and that the two verbs are not always treated in the same way. The information following the lemmas, amo and lego. Is useful because in Latin verbs various ‘conjugations’ on the basis of their verbal morphology, However, for French, which, to a large extent, is like Latin with

regard to verbal morphology, only positive infinitives are picked as lemmas.

Second, verbs in all Bantu language dictionaries are not lemmatized like in Swahili. also a Bantu language, in (14). For instance, in some Bantu language dictionaries, a verb lemma is is hyphenated, as in –lima 'to plough/cultivate' in the White Fathers'Bemba-English Dictionary (White Fathers (1991)). In fact, two dictionaries of the same language may lemmatize differently.

The third and last point is that Kiango (2000) , using the term 'citation form' to refer to 'headword' has well discussed problems of lemmatization in Bantu languages, He has shown that in many cases entry heads in Bantu dictionaries are word-forms or parts of a word. Topics discussed by Kiango's (op. cit.) include, among others, the following:

- Alphabetizing verbs;
- Alphabetizing nouns;
- Alphabetizing adjectives;
- Alphabetizing words with pre-prefixes; and
- Derivational morphology and dictionaries

All in all, Kiango (op. cit.) made good observations and, directly or indirectly, good recommendations concerning lemmatization in Bantu.

**The Content of the Information about Entry Heads**

In the formula A = B + C, which has been used on several occasions, in which A stands for the dictionary entry, B for the entry head and C for the information about the entry head, this section is a brief account of the contents of C. The primary point to

be made is that the content is a function of the various dictionary parameters, including the parameters that have been dealt with in this paper.

It is important to note that one important parameter not mentioned so far in this paper is a parameter for which this writer is hereby proposing the term 'Individual Parameter', which means that a lexicographer may decide what to include and what not to include in the informational about entry heads. As an example, two dictionaries of the same type may differ in that one gives some pieces of grammatical information before the meanings of entry heads while, for no compelling reason, the other only gives the meaning.

**Conclusion**

The paper has presented the typology of dictionaries, showing that there are differences in content of even among dictionaries belonging to the same type, The paper, after acknowledging these, uses insights from the theory known as Principles and Parameters (P&P) to set a number of parameters underlying the typologies of dictionaries as well as the contents of dictionaries. That insights are taken from a linguistic theory, mainly concerning syntax, are used in lexicography is not abnormal as lexicography is part of applied linguistics.

## References


Clapham, C. and J. Cholson .(2009). *Oxford Concise Dictionary of Mathematics.* Oxford: Oxford University Press.

Chanda V. M. & S. B. Mkandawire. (2013). *Speak Zambian Languages.* Lusaka: unza press.

Collins. (2011). *Collins Pocket Latin Dictionary.* Collins

Daintith, J. (2008). *Oxord Dictionary of Chemistry.* Oxford University Press.

Edmonds, D. (1999). *The Oxford Reverse Dictionary*. Oxford University Press.

Kiango, J. G. (2005). Problems of Citation Forms in Dictionaries of Bantu Languages. *Nordic Journal of African Studies* 14(3): 255-273

Longman. (2003). *Longman Dictionary of Contemporary English.* London: Longman

Mkandawire, S. B. (2018). Literacy Versus Language: Exploring their Similarities and Differences. *Journal of Lexicography and Terminology,* 2 (1), 37-55.

Perrott, D.V. (1965). *Concise Swahili and English Dictionary: Swahili-English/ English-Swahili.* London: Hodder and Stoughton (Teach Yourself Books)

Trask, R. L. (2008). *Etymological Dictionary of Basque*. University of Sussex

White Fathers (The Society of the Missionary for Africa) (1991). *The White Fathers' Bemba=English Dictionary*. Ndola, Zambia: Mission Press