Pre-publication version of chapter published in Wallis Reid, Ricardo Otheguy, and Nancy Stern (eds.), *Signal, Meaning, and Message: Perspectives on sign-based linguistics*. Amsterdam: John Benjamins 2002, pp. 3-64

# (What) do noun class markers mean?

Ellen Contini-Morava
*University of Virginia*

## 1. Introduction

The phenomenon of noun classes, including what is usually referred to as grammatical gender, is widespread in some language families, absent in others. Since noun classification is not universal, it has received little attention from those who regard language as an innate, formal system, except insofar as the related phenomenon of what is called grammatical agreement is cited as evidence for the autonomy of syntax. However, linguists with an interest in the linguistic sign and in functional explanations for grammatical phenomena cannot avoid confronting the challenges posed by both noun classification and grammatical agreement: are noun class markers and/or agreement markers signs, and if so what do they mean? I have addressed the second question, about the meanings of so-called agreement markers, elsewhere (Contini-Morava 1996), as have others (e.g., García 1975; Reid 1991); in this chapter I will focus on the first question, in relation to the noun class markers of Swahili. The chapter is organized as follows. In the next section I outline traditional definitions of the notion of noun class and discuss some ways in which the problem of the meaningfulness of noun class markers could be handled from a sign-oriented perspective. In Section 3 I describe the Swahili morphology that is the focus of the present analysis and discuss its role in the categorization of the Swahili lexicon. Section 4 proposes a semantic analysis of the singular noun classes of Swahili that is inspired in part by work in Cognitive Semantics; differences between the present approach and that of Cognitive Semantics are

also discussed. In the concluding section I discuss two alternative approaches to the semantics of the noun class prefixes: an analysis in which the prefixes serve as indices to semantically motivated groupings of nouns, and a cognitive-semantic network analysis of the prefixes themselves. I also discuss the relationship of noun class markers to the traditional dichotomy between lexicon and grammar, and the functional motivation for noun classification.

## 2. Analyses of "noun class"

Most linguists who deal with noun classes implicitly or explicitly subscribe to a definition like the following, from Dixon (1986: 106):[1]

> (a) SIZE. Noun classes involve a grouping of *all* nouns of a language into a smallish number of classes (usually, from 2 to around 20). In many languages each noun belongs to just one class. Where this is not the case there are typically just a small number of nouns that may select more than one class.…
>
> (b) REALISATION. Noun classes always constitute a closed grammatical system, on a par with number and case and tense (where any member can be specified as the complement of the other members of the system e.g., 'not masculine or neuter' must be 'feminine' in Latin). Information about noun class may be fused in a single morpheme with definiteness (as in French), number (as in Bantu languages), or case (as in Latin). Noun classes may be coded as affixes, or as separate grammatical words or clitics such as articles, or 'noun markers' in Dyirbal….
>
> (c) SCOPE. Marking of noun class is never entirely within the noun word. If a noun indicates class by an affix on itself, then this affix will also apply concordially to some other words in the sentence e.g., in Swahili it goes onto all other words in the noun phrase (demonstratives, numerals, adjectives) and is also coded, for certain syntactic functions, onto the verb.

One significant feature of the above definition is that it includes no mention of meaning. In conformity with traditional, syntax-based approaches to grammatical analysis, the distribution of noun-class markers is seen as a purely

formal phenomenon: noun stems are grouped into classes (the basis for this grouping remaining unstated), which are said to select one or another noun-class marker, and so-called concordial markers on demonstratives, adjectives etc. are said to be selected according to the class of the noun. In fact most linguists assume that noun class systems, including systems of grammatical gender, are largely arbitrary from a semantic point of view.[2]

For those who prefer not to think of language as "form for form's sake," but rather as a system of meaningful signs, noun class markers pose a challenge. On one hand, they behave like grammatical signs: as Dixon points out, they constitute a closed system that exhaustively divides the noun universe up in some way. In Boasian terms (1966 [1911]: 39), they constitute an "obligatory category": no noun stem in discourse remains unclassified.[3] On the other hand, they are very resistant to semantic definition: analyses that posit a single, invariant meaning for each noun class marker are either so abstract as to be virtually untestable (e.g., van Schooneveld 1977 for gender in Russian; Pozdnyakov 1993, Chapter III, for noun class in Temne) or are contradicted by counterexamples (e.g., Zawawi 1979, for noun class in Swahili). How, then, are noun class markers to be analyzed from a sign-based perspective? Here are four possible answers to this question.

(a) Allomorphs

Noun-class markers in the aggregate have a meaning such as ENTITY, i.e., they indicate that the associated lexical item is to be interpreted as a noun. Concordial markers are said to attach to elements such as demonstratives, adjectives or the like, have a meaning like IDENTIFICATION OF ENTITY, i.e., they indicate that this element is to be grouped with a noun-referent in the linguistic or extra-linguistic context. (Markers of either type may also convey additional semantic information such as number and case, but this is beside the present point.) However, markers of different noun classes/genders are not distinct from one another in meaning. In effect this amounts to what a Bloomfieldian would call "grammatically conditioned alternants" (cf. Bloomfield's treatment of German gender, 1933 [1984]: 211).[4] This is the most cautious position, since it does not require further investigation of the question why the nouns are grouped the way they are rather than in some other, equally arbitrary way.

But tempting as it may be to wash one's theoretical hands of the issue, there are two problems with this position. First, no such system is totally arbitrary from a semantic point of view. Even the most apparently chaotic systems show some semantic motivation in the assignment of nouns to classes (see, for example, the work of Zubin and Köpcke on gender in German: Zubin and Köpcke 1986a; 1986b). If formally distinct signals are associated with semantically motivated groups of nouns, it is hard to maintain that all the different noun class markers are merely allomorphs of a single morpheme. Second, although there is a statistical correlation between co-occurrence of particular noun stems with particular noun class prefixes, some stems may co-occur with different prefixes, and change of prefix brings change in the meaning of the resulting noun (see below, Section 3.3.1 and 3.3.2). If the prefixes were allomorphs of a single morpheme, replacement of one allomorph by another should not occur at all, and should certainly not have any semantic effects. Third, the association of noun class markers found on noun stems with concordial markers found on demonstratives, etc. is rarely 100% consistent, as it should be if these were indeed all allomorphs. For example, in cases of conjoined noun phrases containing nouns of different classes, or situations in which no specific noun is being referred to, a choice must still be made among the various options provided by the concordial system, and the basis for such choices must be specified. If in fact different alternants are conditioned in different contexts rather than following directly from the (arbitrary) groupings of nouns, this weakens the claim that all concordial markers are representatives of the same morpheme or sign.

(b) Distinct signs with distinct, invariant meanings

Each noun class marker (NCM) is a distinct linguistic sign with a single abstract, invariant meaning that contrasts with the meanings of all the other NCMs. This would overcome the objections to (a) and is desirable in principle, but in practice no such analysis has been convincingly proposed, as mentioned earlier. For example, Zawawi (1979) proposes the meaning "substance of life, singular" as the invariant meaning of the Swahili NCM *m-* (traditionally regarded as two homonymous prefixes, for Classes 1 and 3). Although this fits all the nouns in traditional Class 1, almost all of which designate humans, and many nouns in Class 3 (those designating plants), there are also many nouns with the prefix *m-*

that do not fit this meaning, if it is interpreted literally. Examples include *m-kuki* 'spear', *m-kufu* 'metal chain', *m-ji* 'town', etc.

(c) Distinct signs with polysemous meanings

Each NCM is a distinct linguistic sign, but rather than having a single, invariant meaning, its meaning consists of a network of senses connected to one another both by relations of taxonomic inclusion and by relations of semantic extension such as metaphor and metonymy. In effect, the interconnected senses are abstractions over clusters of meanings of lexical stems that all share a common NCM. Such networks are a staple of cognitive grammar (see e.g., Lakoff 1987; Langacker 1987, 1988, 1990), and noun classification systems have sometimes been analyzed along these lines (e.g., Zubin and Köpcke 1986b; Spitulnik 1987; Selvik 1996; Moxley 1998), though more often semantic networks are used to represent the polysemous structure of a single morpheme or word. The advantage of this type of analysis is its recognition of metaphor and metonymy as regular principles of semantic extension: cases that would have to be regarded as counterexamples to a traditional abstract-meaning analysis as described in (b) above can be explained by these principles. For example, a network analysis can incorporate concepts like "substance of life" (a meaning category suggested by Zawawi 1979, mentioned earlier) into the network associated with a given NCM and show how they motivate the inclusion of the apparent counterexamples: a spear resembles a plant in that it is long and thin (metaphor), and it is also made of wood (metonymy); a metal chain is likewise long and thin (metaphor); a town is an assemblage of people that can figuratively grow and reproduce (metonymy and metaphor). In earlier work (Contini-Morava 1994, 1997) I have proposed network analyses of the Swahili NCMs myself.

From a sign-oriented perspective however, network analyses are subject to two types of criticism. First, the problem of analytical indeterminacy. As has been pointed out by some proponents of cognitive grammar themselves (e.g., Sandra and Rice 1995), a set of related senses can be represented in many different ways, and it may be hard to decide when two senses are different enough to add a new subcategory to the network. Furthermore, it becomes difficult to draw a line between aspects of the network that are stored in memory and productive extensions to new communicative contexts. Analyses that posit a single, invariant meaning are not subject to this criticism, because all senses,

whether actual or potential, are explained by the invariant meaning, and there is no theoretical need for subclassification of senses. A second criticism, in relation to a network analysis of a noun class marker, has to do with redundancy (I thank the editors of this volume for discussion of this issue). If the network consists of a group of related semantic categories that are abstractions over clusters of lexical meanings (those of the noun stems in a given class), attributing the network to the NCM entails stating the lexical structure twice: once as part of the analysis of the lexicon and again as the analysis of the NCM. An alternative to a network analysis that avoids both these criticisms is:

(d) Noun class indices

Each noun class marker is a distinct linguistic sign (leaving aside obvious morphophonemic variation), and each signals that its associated nouns are members of a specific group, i.e., a noun class or gender. This analysis differs from the one outlined in (a) in that the fact of classification itself is considered meaningful in the sense that it provides a basis for identifying referents in discourse. It differs from (b) in that the meaning of each NCM has little in the way of semantic substance as that is ordinarily conceived. Specifically, this analysis builds into the notion of meaning a version of the functional explanation that is often proposed for noun classification: subdivision of nouns into classes allows a parallel subclassification of elements that refer to a noun (demonstratives, numerals, adjectives, etc.). This helps narrow the range of possible reference of such elements, by restricting this range to a noun of such-and-such a class (see e.g., Fodor 1959; Diver 1972; Otheguy 1977; Greenberg 1978; Zubin and Köpcke 1986a; Corbett 1991). In favor of this argument is the fact, noted above in the Dixon quote, that marking of noun class is never limited to the noun-word alone but is always also (and sometimes exclusively) indexed on concordial elements that are not themselves nouns. Of course, the sense of "meaning" used here would not be countenanced in a truth-conditional semantic theory, since a change of noun-class markers would not make a difference in the truth conditions of sentences. However, this analysis would be consistent with a sign-oriented approach in which meanings serve as hints that help the hearer infer an intended message, rather than constituting compositional fractions of a notional whole (cf. Diver 1975: 9-10; Diver 1995: 73-79; Tobin 1990, Ch. 3; Reid 1991: 95; Huffman 1995: 189).

Meanings that index groupings of nouns would aid in connecting modifiers with modifieds and identifying intended referents when no noun is explicitly mentioned; that is they would be processing aids rather than notional fractions of the message.[5] Note that for the purpose of referent-identification even a semantically arbitrary subdivision of nouns into classes A, B, C and the like would serve: the range of reference of particular concordial elements would still be narrowed to a noun falling within one of these groups rather than another. However, as noted earlier, such groupings are never totally arbitrary. The "indices analysis" is an improvement over the "allomorphs analysis" (a above) in that it treats formally distinct signals as separate units; however it must be shown how an indices analysis can deal with the objections that were raised against (a): the fact that noun class groupings are semantically motivated, and the fact that change of noun class prefix can produce a change in meaning.[6]

In what follows I will present an analysis that incorporates aspects of the "polysemy analysis" (c above) as well as the "indices analysis" (d above), but I will argue that these analyses operate at different theoretical levels. Specifically, a polysemy analysis is a plausible representation of what individual speakers construct for themselves in order to make sense out of the allocation of noun stems to the various classes. It is a mnemonic strategy that helps the individual speaker memorize the class affiliation of thousands of nouns. Since different speakers are exposed to different vocabulary items in different contexts of use, one need not assume that every speaker's semantic network exactly matches those of all other speakers of the language. On the other hand, the *fact* of classification indexed by morphemes that are affixed to noun stems is shared by speakers of Swahili (in non-pidginized dialects). The indices analysis transcends the individual speaker and has the communicative function of identifying discourse referents. As such, it is part of *langue*.

Because the analysis to be presented makes reference to semantic networks like those associated with cognitive grammar, it may be helpful to provide a brief comparison between that theory and the Columbia School framework that underlies most of the chapters in this volume. The two frameworks share a conception of language as an inventory of meaningful units, but there are also significant differences. Specific points of contrast

between cognitive grammar and Columbia School theory that are relevant to the present paper include the following:

(a) Cognitive grammar makes no theoretical distinction between linguistic and extra-linguistic knowledge, or between semantics and pragmatics (see for example Langacker 1987: 154), whereas Columbia School theory makes a clear distinction between semantic content that is associated with a specific linguistic form and information that is inferred from context, knowledge of the world, and the like. Only the former is regarded as "linguistic" knowledge from this theoretical perspective. A corollary of this difference in approach is that cognitive grammar analyses tend to build large numbers of context-specific senses into their semantic networks, whereas Columbia School analyses are more parsimonious, building into the meaning only what is invariant, and tracing context-specific senses to inferences from accompanying signs or pragmatic information. In Section 4 below I offer analyses of the Swahili noun classes that resemble the semantic networks of cognitive grammar: sets of nouns sharing a common NCM are arranged into semantic categories, and relationships among the categories are represented by lines indicating taxonomic inclusion, metaphoric or metonymic extension, and so on. These are meant to demonstrate that the various noun classes show internal semantic coherence, without necessarily making a claim that a particular network is mapped into the minds of Swahili speakers. The discourse function of NCMs is served even if the details of networks vary from one individual to another.

(b) A second major difference between cognitive grammar (CG) and Columbia School (CS) theory is that CG makes no theoretical distinction between lexical and grammatical signs; in fact "grammar" in this approach is identified with traditional syntax – constructions, sentence types, and the like (e.g., Langacker 1987: 449). For Langacker, the debate with generative models of language revolves around the question whether a clear boundary can be drawn between irregularities (the province of lexicon, in traditional structuralist and generative theory) and regularity/productivity, the domain of syntax (ibid. 25-26). The question whether individual linguistic units might be analyzed differently based on the type of meaning they convey, or on their participation in closed vs. open systems of opposition, does not arise. Consequently linguistic units are analyzed individually in CG, with little or no attention to whether they

enter into a system of oppositions – i.e., to their "value," in Saussure's sense. In Columbia School theory, on the other hand, the distinction between grammar and lexicon is based in part on the fact that grammatical elements – whether morphological units or word-order patterns – tend to be closed-class items for which opposition with other elements in the class plays a major role in their use and interpretation. Although I will argue in the conclusion that noun class markers have a special link with the lexicon that makes them different from other grammatical signs, this is not the same thing as identifying grammar with syntax, or ignoring the usefulness of the lexicon/grammar distinction.

What I have found helpful in CG analyses is their close attention to the roles of metaphor, metonymy, and prototypicality in semantic networks, and the recognition that semantic relationships based on "family resemblances" may be motivated without necessarily being predictable by means of formal rules. Such relationships are also recognized in CS theory, but the status of polysemy within the two approaches is different. For example, Reid (1991: 108) discusses what he calls "principled polysemy" in the exploitation of a lexical item to convey multiple senses:

> [T]he innovative use of a word for a new kind of message does not necessarily call for its analytical partitioning into two separate signs with different meanings. For so long as there exists *some* connection between the original meaning and the novel message – no matter how tenuous – it is presumably strong enough to have inspired the initial innovation. The principles underlying such innovation – metaphor, metonymy, and similarity of form or function – are an integral part of the creative aspect of language use and are at work to some degree in every act of speech.

Where CS theory differs from CG is in its greater concern for the question where to draw the line between polysemy (of a single linguistic sign) and homonymy (more than one sign). A theory that puts little stock in the distinction between semantics and pragmatics is necessarily less concerned with this issue, because it provides no principled basis for deciding how to draw this line.

In the next two sections of this chapter I will demonstrate the nature of the semantic regularities within the Swahili noun classes, including the productive and semi-productive uses of the NCMs. I return to the question of the semantic contribution of the NCMs in the concluding section.

## 3. Noun classification in Swahili

### 3.1 *Definition of "noun class" in Swahili*

The Swahili noun class system is typical of Bantu languages. According to traditional analyses, noun stems are subdivided into eleven classes, each marked by a characteristic prefix (see Table 1 below). Most of the prefixes are paired for singular/plural: noun stems carrying a Class 1 prefix in the singular replace it with a Class 2 prefix in the plural, the prefix of Class 3 (sing.) is replaced by that of Class 4 (plural) etc.[7] Class 11/14 is not paired with a plural prefix of its own; some Class 11/14 nouns are associated with plurals in Class 6 (e.g., *ugonjwa/magonjwa* 'illness/es') and others have plurals in Class 10 (e.g., *ufunguo/Øfunguo* 'key/s').[8] The traditional account is that to each nominal prefix correspond two types of concordial prefix: (a) a prefix attached to adjective and numeral stems that agree with a noun of a given class (called the "primary concord" or "adjectival concord"); (b) a prefix (called the "secondary concord" or "pronominal concord") that is attached to verbs to indicate that the subject or object of the verb is a noun of a given class, and to a miscellaneous collection of other elements including possessives, demonstratives, the relative pronoun, etc. that are said to agree with a noun of a given class. The following table lists the relevant prefixes for each class:

Table 1: *Swahili nominal and concordial prefixes (some morphophonemic alternations ignored)*

| Class (traditional Bantu numbering) | Nominal (NCM) (affixed to noun stems) | Adjectival (AP)[9] (affixed to adjective and numeral stems) | Pronominal (PP) (affixed to V stems as subj. or obj. and to "pronominal" stems: demonst., poss., etc.) |
|---|---|---|---|
| 1 | *m-* | *m-* | *yu-; ye-; w-; a-; m-* (depends on stem[10]) |
| 2 | *wa-* | *wa-* | *wa-* |
| 3 | *m-* | *m-* | *u-* |
| 4 | *mi-* | *mi-* | *i-* |
| 5 | Ø or *ji-* | Ø or *ji-* | *li-* |
| 6 | *ma-* | *ma-* | *ya-* |
| 7 | *ki-* | *ki-* | *ki-* |
| 8 | *vi-* | *vi-* | *vi-* |
| 9 | *-*[11] | - | *i-* |
| 10 | - | - | *zi-* |
| 11/14 | *u-* | *u-* | *u-* |

Various points can be made about the morphological analysis reflected in Table 1. First of all, inspection of the first column of prefixes in the table shows that given the form of a singular noun alone (the odd-numbered prefixes in the table), one would not be able to predict either its corresponding plural (the immediately following even-numbered prefix) nor its associated pronominal concord: for example, some *m*-initial nouns have *wa-* in the plural and *yu/ye/w/a/m* as secondary concords (those traditionally assigned to Classes 1-2, such as *mtoto/watoto* 'child/children'), whereas other *m*-initial nouns have *mi-* in the plural and *i-* as secondary concord (those traditionally assigned to Classes 3-4, such as *mti/miti* 'tree/trees'). Similarly, nouns with no overt prefix in the singular may differ in their plural formation and secondary concords (those assigned to Classes 5 and 9-10 respectively). Therefore the definition of noun class in Bantu languages has traditionally involved reference to either (a) singular/plural prefix pairings, or (b) the concordial affixes associated with nouns of a given class.

In fact, the first of the above criteria is insufficient for defining noun class, because of the existence of nouns that are *singularia* or *pluralia tantum*. Worse, in Swahili the two criteria lead to conflicting results: nouns denoting animate beings may exhibit the singular-plural prefix pairings of various classes

(compare *m-toto/wa-toto* 'child/children', *Ø-baharia/ma-baharia* 'sailor/s', *ki-nyozi/vi-nyozi* 'barber/s', *-sonara*/*-sonara* 'goldsmith/s'), yet all animate nouns are said to "govern" the adjectival, pronominal and verbal concords of Class 1 in the singular, 2 in the plural. Thus according to the morphological criterion of singular/plural prefix pairing, animate nouns would be assigned to different classes, but according to the syntactic criterion of concord" they would all be assigned to the same class.

Swahili grammars and dictionaries usually use the morphological criterion for assigning nouns to classes, and include a statement to the effect that animate nouns are an exception to the normal class-based concord rule.[12] I have argued elsewhere (Contini-Morava 1996) that the "pronominal prefixes" of Swahili should be treated as a separate system whose members make an independent contribution to the message, rather than as a set of meaningless forms automatically governed by a rule of concord. In that discussion I argued that the pronominal prefixes help identify the referent with which the attached lexical item (demonstrative, possessive, verb etc.) is linked, by supplying information about animacy, number, and noun class of the intended referent. However, I remained agnostic on the question what, if any, meaning is signalled by the prefixes that are attached to noun stems (NCMs), i.e. on the question whether the pairing of lexical stems with the various NCMs is semantically motivated. The purpose of this paper is to venture into these murky waters.

The traditional description of noun classes in Swahili reflected in Table 1 and summarized above seems straightforward enough at first sight, but in fact it is subject to criticism on three interrelated points: (a) the notion of concord as automatic marking of targets based on syntactic properties of a controller, which treats agreeing elements as if they make no semantic contribution; (b) the categorization of the lexicon; (c) the analysis of the noun class markers. I have criticized the notion of concord in Contini-Morava (1996). In what follows I will take up the second and third points in turn.

### 3.2 *Categorization of the lexicon*

Swahili is an agglutinative language in which monomorphemic words are relatively rare.[13] That is, most lexical stems co-occur with one or more grammatical affixes indicating such information as noun class/number, animacy/number, tense-aspect-modality, etc. To address the question what the

NCMs may mean entails investigating their distribution in relation to lexical stems. Two issues arise here: (a) the overall co-occurrence of lexical stems with NCMs as opposed to other grammatical signs; (b) the co-occurrence of lexical stems with one or another of the various NCMs. I discuss these in the next two subsections.

### *3.2.1 Parts of speech*

In most languages there is at least a statistical relationship between the distribution of certain grammatical signs (such as those indicating number, comparison, or tense) and classes of lexical signs, a relationship recognized by the traditional notion of parts of speech. Parts of speech have been part of the Western linguistic tradition since at least the time of Plato (see e.g., Dinneen 1967: 78). Although the old terminology has been retained, the original, semantically based definitions of parts of speech are generally replaced by structural definitions these days, and now the view prevails that part-of-speech categorization is an inherent property of lexical items that causes them to select one or another set of grammatical morphemes (e.g., Wurzel 1989: 44; Anderson 1992: 133; Aronoff 1994: 10). Columbia School work has often challenged this view (e.g., Diver 1977, 1995; Otheguy 1977, Chs. 1 and 2; Reid 1991, Chs. 2 and 3). The central point is that "part of speech" should be viewed not as an inherent feature of lexical items, but as a result of the interaction of lexical items with grammatical signs. As Diver (1995: 95) puts it,

> It is not... that there is a part of speech classification existing independently in the lexicon, such that a particular class "takes" certain endings, but that "part of speech" is an effect in the message produced by the meanings of the grammatical systems associated with it in the particular instance.[14]

From this perspective, a noun in Swahili is a combination of a noun class prefix (or its zero-alternant) with a lexical stem, whose effect on the message is that some entity is being referred to. As for the distribution of lexical stems in relation to grammatical affixes, despite the common use of part-of-speech terminology, it is not the case that lexical stems are subdivided into mutually exclusive groups that are in complementary distribution in this regard. For example, a stem like *linda* 'guard' is part of a "verb" when associated with a "subject prefix" and tense-aspect-modality prefix in the word *a-na-linda* [3sg.anim.subj.-pres.tns.-'guard'] 's/he guards' and part of a "noun" when

associated with a NCM in the word *m-linda* [NCM1-'guard'] 'guard'. The stem *refu* 'tall' is part of an "adjective" in the phrase *mlinda m-refu* 'a tall [AP1-'tall'] guard' but part of a "noun" in the word *u-refu* [NCM11/14-'tall'] 'tallness'. In such cases the grammars and dictionaries speak of "nouns derived from verbs," "nouns derived from adjectives" etc. Like Bloomfield's notion of "class-cleavage" (1933: 205), this terminology preserves the appearance of inherent part-of-speech categorization of the lexicon even while acknowledging data that contradict it. A more accurate statement would be something like the following: Certain lexical stems, by virtue of their meanings, are statistically more likely to co-occur with one type of grammatical sign rather than others (e.g., with noun-class morphology rather than tense-aspect-modality morphology). Terms like "noun" and "verb," when applied to lexical stems, indicate the most statistically frequent grammatical-lexical combinations into which these stems enter. The term "derived" indicates a statistically less frequent grammatical-lexical combination, sometimes accompanied by "derivational" morphology.[15]

Having said that much, it is nevertheless true that Swahili lexical classes are more well-defined than English ones. Although every verb stem can also co-occur with noun morphology (i.e. can be nominalized), and every adjective stem can be used to form a derived noun, the reverse is not the case: some lexical stems occur exclusively with noun morphology.[16] Thus the category "noun" as part of speech in Swahili (i.e. a pairing of NCM + a lexical stem) is broader than the category "noun stem" as a lexical class, since the former includes nominalized verbs, nouns derived from adjectives etc., whereas the latter includes only lexical stems that co-occur exclusively with noun morphology.

Having defined "noun" as a grammatical-lexical pairing, one may ask whether these pairings are unique, i.e. whether each lexical stem co-occurs with only one singular/plural NCM pair. As suggested in the introductory section, the answer to this question is no. The next section will outline some of the issues involved.

## 3.3 *Pairing of noun class prefixes with lexical stems*

### *3.3.1 Productive NCM uses*

One of the first questions that may occur to a prospective investigator of noun class semantics is that of productivity: to what extent do lexical stems have

freedom of occurrence with the various NCMs, and what effect on the meaning does such variation have? As it happens, all the NCMs of Swahili can be used productively, in that nouns with predictable semantic properties can be formed by the addition of one or another NCM to a lexical stem, sometimes accompanied by a derivational suffix.[17] Examples:

(a) Class 1-2 prefixes (*m-/wa-*) are used with verb stems to form agentives, e.g., *-cheza* 'play', *m-cheza/wa-cheza* 'player/s'.

(b) Class 3-4 prefixes (*m-*/*mi-*) are used with verb stems to form nominalized verbs, e.g., *-kutana* 'meet', *m-kutano*/*mi-kutano* 'meeting/s';[18] also used with noun stems that do not usually have *m-* prefix to indicate extraordinary size or "monstrosity" (Zawawi 1979: 103; Frankl and Omar 1994), e.g., *ki-kombe* 'cup' (Cl. 7), *m-kombe* 'very large cup' (Cl. 3) ; *fedha* 'money' (Cl. 10), *mi-fedha* 'large amount of money' (Cl. 4).

(c) Class 5-6 prefixes (Ø-/*ma-*) are used with noun stems that are normally in other classes to indicate large size (not as large as Class 3-4), e.g., *m-toto/wa-toto* 'child/ren' (Cl. 1), *Ø-toto/ma-toto* 'large child/ren'; also used with verb stems to derive nominalized verbs, e.g., *-jenga* 'build' (V), *Ø-jengo/ma-jengo* 'building/s' (N, Cl. 5-6); *-fundisha* 'teach', *ma-fundisho* 'teachings' (Cl. 6). Class 6 is also used with noun stems that lack a prefix in the singular (i.e. those of Cl. 9-10) to form collective plurals, e.g., *karatasi* 'paper/s' (Cl. 9-10 nouns are undifferentiated for number), *ma-karatasi* 'collection of papers'.

(d) Class 7-8 prefixes (*ki-/vi-*) are used with noun stems that are normally in other classes to indicate small size, e.g., *m-toto/wa-toto* 'child/ren' (Cl. 1-2), *ki-toto/vi-toto* 'little child/ren'; also used with verb stems to indicate an object associated with verbal process, e.g., *-ziba* 'stop up a hole', *ki-zibo/vi-zibo* 'plug/s, stopper/s'.

(e) Class 11/14 prefix (*u-*) is used with adjective and verb stems, and with noun stems normally in other classes, to form abstract nouns, e.g., *-refu*

'tall', *u-refu* 'height'; *-posa* 'woo', *u-poso* 'marriage proposal'; *m-toto* 'child' (Cl. 1), *u-toto* 'childhood'.

Although these productive uses of the NCMs might seem an appropriate point of departure for a semantic analysis, in fact a problem becomes immediately apparent: the stems that are normally associated with these NCMs (as opposed to those attached to them as productive derivations) do not necessarily fall into the semantic areas defined by the productive uses. Often the reverse is the case. For example, although NCM1-2 can productively derive agentives, not all nouns in Class 1-2 are agentive. Examples: *mtoto* 'child', *mjomba* 'uncle/nephew', *mkongwe* 'old person'. What all Class 1-2 nouns have in common is that they denote animate beings, almost all human. The category "agentive" is easy to see as a more specific extension of "human", because it is humans who tend to be agents, but not the other way around. The same holds for the other classes, although none of them is as semantically cohesive and transparent as Cl. 1-2.[19]

A few more examples should suffice. Although NCMs 3-4 can be used to derive nominalized verbs and "monstrously" large things, most of the membership of this class is comprised of names of plants and trees. NCMs 5-6 can derive augmentatives, but not all nouns in Cl. 5-6 denote large things: Ø-*sikio* 'ear', Ø-*joya* 'spongy substance inside coconut shell', most names of fruits, and so on. NCMs 7-8 can derive diminutives, but not all normal members of class 7-8 are small: *kititi* 'depth of the sea', *kipofu* 'blind person', names of languages, etc. It is also worth noting that the relationships among the various productive uses of a particular NCM are not always easy to connect among themselves: what does 'monstrously large' have to do with 'nominalized verb'?

In short, the productive senses of the NCMs are rather narrowly specialized in relation to the overall membership of their respective classes, and cannot be regarded as diagnostic of the central semantic cores of the classes themselves.[20] This phenomenon is not limited to Swahili but is characteristic of other Bantu languages as well. Accordingly, many scholars divide the noun classes into two subsets: a derived set of classes, whose prefixes are assumed to be meaningful, to which noun stems from any class can be freely assigned with predictable effects on meaning, and an inherent set of classes, whose membership is largely arbitrary (cf. e.g., Givón (1972), for ChiBemba; Heine (1982), who uses the terms "free" vs. "fixed" gender; Reynolds (1989),

Reynolds and Eastman (1989), and Nurse and Hinnebusch (1993) for Swahili). Since there is no formal distinction between the "inherent" and "derived" NCMs or their associated concordial markers, this approach in effect sets up a series of homonymous NCMs.

Although the inherent/derived-class analysis does solve some problems, others still remain. First, the term "derived class" has been used only for the diminutive and augmentative uses of the NCMs (NCM7-8 and 5-6 respectively). However, as pointed out above, the other NCMs can also be used productively, and derived nouns can be formed from adjective and verb stems as well as from noun stems. If one were to create a new, homonymous derived class for every productive derivation listed above, the number of homonymous NCMs would quickly become unwieldy. Furthermore, this tactic does not address the question whether there are any semantic regularities within the inherent classes, and it also ignores semantic relationships, if any, between productive and non-productive uses. Finally, terms like "inherent" or "fixed" gender suggest that each lexical stem is assigned to only one "inherent" class, which is not the case. In fact, there are a number of less productive, but nevertheless systematic patterns in which a given stem can be used with more than one NCM with predictable changes of meaning. I discuss some examples of these below.

### *3.3.2 Less productive NCM alternations*

One very pervasive pattern of interclass relationships is the use of NCMs 3-4 to designate a plant or tree and that of NCMs 5-6 to designate the associated fruit. For example, *m-papai* 'papaya tree' (Cl. 3), *Ø-papai* 'papaya fruit' (Cl. 5). This pattern is productive in the sense that it is extended to loanwords designating plants, whether or not the source word began with *m-* (for example, *mfenesi* 'jackfruit tree', *Ø-fenesi* 'jackfruit', source word begins with *f-* in Hindi). To be sure, not all plants are in Class 3, nor are all fruits in Class 5. But the deviations from this pattern form other patterns that show their own semantic regularities. For example, there is also a regular relationship between NCMs 3-4 to designate a plant, and absence of NCM (i.e. membership in Cl. 9-10) to designate a hard, leathery, or otherwise non-prototypical fruit, seed or other product:

Table 2: *Plant = NCM3-4; non-typical fruit = Class 9-10*

| Class 3 (plant) | | Class 9 (product) | |
|---|---|---|---|
| *mnjugu* | peanut plant | *njugu* | peanut |
| *msandali* | sandalwood plant | *sandali* | sandalwood |
| *msandarusi* | gum copal tree | *sandarusi* | gum copal |
| *msolo* | a thorny shrub | *solo* | seeds of msolo, used as counters in game of *bao* |
| *mtangawizi* | ginger plant | *tangawizi* | ginger root |
| *mtipitipi* | medicinal climbing plant | *tipitipi* | poisonous and ornamental seeds of same |
| *mkarafuu* | clove-tree | *karafuu* | cloves (flower bud of mkarafuu) |
| *mkweme* | a climbing plant | *kweme* | seeds, used for oil |
| *mbangi* | marijuana plant | *bangi* | marijuana |
| *mfiwi* | lima bean plant | *fiwi* | lima bean |
| *mfuu* | kind of tree | *fuu* | small black edible berries of same |
| *mhenzerani* | cane plant | *henzerani* | a cane |
| *mkatani* | sisal hemp plant | *katani* | sisal hemp |
| *mnazi* | coconut tree | *nazi* | coconut (generic term) |
| *mndimu* | lime tree | *ndimu* | lime |
| *mnili* | indigo plant | *nili* | indigo |
| *mnyanya* | tomato plant | *nyanya* | tomato |
| *mpilipili* | red or white pepper plant | *pilipili* | pepper |
| *mpopoo* | areca palm | *popoo* | areca nut |
| *mtumbako* | tobacco plant | *tumbako* | tobacco |
| *mwafu* | wild jasmine plant | *afu* | wild jasmine |
| *mwaridi* | rose busy | *waridi* | rose |
| *mzabibu* | grape vine | *zabibu* | grape, raisin |
| *mzingefuri* | anatta plant | *zingefuri* | cinnabar (orange-red coloring matter) |

Another systematic deviation from the pattern NCM3-4 = plant, NCM5-6 = fruit is one in which NCM3-4 designates the plant, and a long thin or liquid product carries NCM11/14:

Table 3: *Plant = NCM3-4; long thin/liquid product = NCM 11/14*

| Class 3 (plant) | | Class 11 (product) | |
|---|---|---|---|
| *mfagio* | fiber plant used for brooms | *ufagio* | broom |
| *mgomba* | banana plant | *ugomba* | banana plant fiber |
| *mkunde* | runner bean shrub | *ukunde* | runner bean |
| *mlimbolimbo* | thorny hedge plant | *ulimbolimbo* | sap, u/f/birdlime |
| *mtupa* | generic term for plants used for fish poisons | *utupa* | juice of plant, used for fish poison |
| *msufi* | kapok plant | *usufi* | silky cotton from kapok pods (also *sufi*, Cl. 9) |

For plants of which various parts are usable, there is a pattern that combines these:

Table 4: *Plant = NCM3-4; fruit/leaf = NCM5-6; seed/fiber = NCM 11/14*

| Class 3 (plant) | | Class 5 (fruit) | | Class 11 (fiber) | |
|---|---|---|---|---|---|
| *mbuyu* | calabash tree | *buyu* | calabash | *ubuyu* | calabash kernel |
| *mnanasi* | pineapple plant | *nanasi* | pineapple | *unanasi* | inside of plant, used for fiber |
| | | *kuti* | whole coconut leaf | *ukuti* | coconut frond |

From the above examples it should be clear that lexical stems in Swahili are not all bound to a single inherent class. In fact, in the case of stems that can designate either the plant or the fruit depending on whether the prefix from Class 3-4 or 5-6 is used, the decision which class the stem belongs to is essentially arbitrary. Most interclass relationships are not as productive as those described in the last section, and some stems do not alternate between classes as readily as the ones illustrated here, but the existence of patterns like these points to some kind of semantic systematicity both within individual classes and between classes. At the same time, like the productive uses described in the last section, these semi-productive interclass relationships only tell part of the semantic story of each class. To what extent are the classes internally coherent? I turn to this broader question in the next section.

## 4. Overview of the noun classes

The Swahili noun class system is very complex, as is probably already obvious. The purpose of the present paper is not to go into the details of each class, but rather to make a theoretical point about the kind of meaning signaled by the NCMs, from a functionalist and sign-oriented perspective. However in order to make the theoretical point it is necessary to provide enough data to make a persuasive case, and the reader may want to be given some idea of the system as a whole. I will therefore provide a general overview of an analysis that I have been developing over the past few years, and refer the reader elsewhere for more detail (see Contini-Morava 1994).

### 4.1 *Theoretical orientation*

Attempts to address the question whether noun class markers are meaningful have always begun by looking at the noun-stems that share a common NCM, to see if a common meaning can be abstracted from the set. As mentioned earlier, such attempts have led scholars of Bantu and other noun class languages to either abandon the quest for a common meaning, or else to propose meanings so abstract as to be unfalsifiable, or meanings that are contradicted by the data. The present study also begins by looking at the groupings of nouns into sets sharing a common prefix, but it arrives at a different conclusion. The analysis of noun-classes to be presented below is inspired by recent work on Cognitive Semantics as described by Ronald Langacker and others in numerous publications (e.g., Langacker 1987, 1990; Lakoff 1987). What I have found most useful in this approach is the recognition that linguistic categories may have a complex internal structure, so that membership in a category need not be determined by a set of necessary and sufficient defining properties, but rather by a chain of "family resemblances" (a term borrowed by cognitive grammarians from Wittgenstein). The links among members of a category may involve resemblances based on metaphor and metonymy as well as those based on common properties. Furthermore, categories may exhibit "prototype structure": some members of the category may be viewed as more central, or prototypical, while others are more peripheral.

In representing the semantic structure of the Swahili noun classes, I have found it helpful to borrow some conventions from Cognitive Semantics, especially Langacker (1988). However, as mentioned in Section 2, this does not

entail an endorsement of all the principles of Cognitive Semantics. The important question of the theoretical and psychological status of such semantic network analyses will be taken up in the concluding section of the chapter.

## 4.2 *The data*

The present analysis is based on a computer database of nouns from the *Standard Swahili-English Dictionary* (Johnson 1939), subcategorized according to a number of morphological and semantic features.[21] Data entry is still in progress but the database already contains over 5,000 nouns, so coverage of each class is quite representative. Use of a bilingual dictionary as data source has some drawbacks (see Contini-Morava 1994 for discussion), but hopefully this is compensated for by the large amount of data.

## 4.3 *Notational conventions*

In diagramming the semantic structure of the individual noun classes of Swahili, I have borrowed the conventions of Langacker (1988) for the representation of a linguistic category. Langacker defines two basic types of semantic relationship among the elements in a category:

(a) relations of "schematicity", in which one element is an "elaboration" or "instantiation" of another, more abstract element (represented by solid lines in the diagram);
(b) relations of "extension", in which some feature specifications are suspended or modified, while other features are retained (represented by dotted lines in the diagram). I have also borrowed from Zubin (1995) the use of double lines to indicate productive subclasses. Sample lists of nouns from each subclass may be found in the Appendices.

## 4.4 *Analysis*

The noun classes have differing degrees of internal coherence in their semantic structure. At one extreme are Classes 1-2 (with the prefixes *m-, wa-* respectively, often called the "human" classes), consisting almost solely of noun stems denoting human beings, especially agents of actions.[22] At the other extreme are Classes 9-10, whose noun stems lack prefixes and are undifferentiated for singular/plural. Classes 9-10 house the majority of foreign

loanwords, and consist of a set of noun stems that may already have been fairly heterogeneous semantically even before the major influx of loans from Omani Arabic, which date only from the 17th century (Nurse and Hinnebusch 1993: 320). Classes 9-10 have become a "residual category" (Zawawi 1979: 115 uses the term "catchall"), with neither morphological nor semantic marking, whose existence helps maintain the semantic coherence of the rest of the system. Since Classes 1-2 and 9-10 have a relatively straightforward semantic structure, I will not discuss them further here. Instead I will provide analyses of three classes that fall in the middle, which represent the nucleus of the system.[23]

### *4.4.1 Class 3 (with prefix* m-*)*

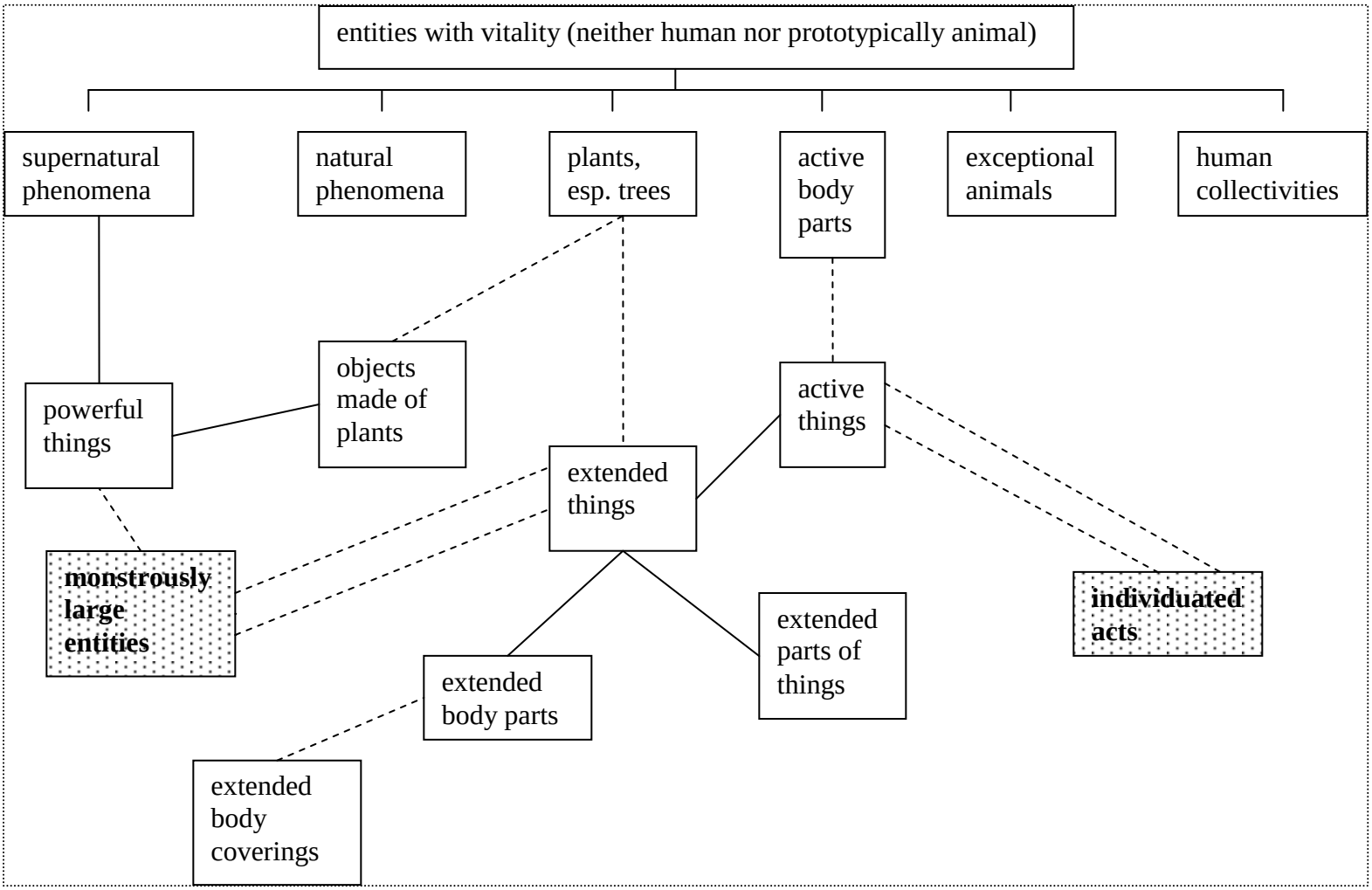
Figure 1 is a schematic representation of the semantic structure of Class 3. The top-to-bottom organization of the diagram moves from the more general to the more specific, but the diagram is not intended to be a taxonomy in the technical sense (cf. Casson 1981: 75-77). Examples of nouns in each category are given in the Appendices.[24]

The topmost category in the chart, "entities with vitality", is what Langacker (1988) calls a "superschema", a maximally abstract category that holds together the various subcategories. "Vitality" is meant to capture various attributes of living beings, including growing and reproducing (true of plants, and metaphorically of human collectivities), but also ability to move (active body parts), to act on or affect other entities in the world or to occur independently of human volition (supernatural and natural phenomena). Most of the categories in the chart are self-explanatory. I will comment only on the ones that are not.

The categories "exceptional animals" and "human collectivities" appear to be in this class not only because of properties that they share with other members of the class, but also because of the opposition between this class and other classes in the system. Classes 9-10 are the "default" for animals, especially mammals, in Swahili and in Bantu languages generally. All animal terms in classes other than 9-10 are therefore exceptional in some way.

The animal terms associated with *m-* (Class 3) denote animals that do not fit easily into established categories, either because of their appearance (swordfish), their behavior (kingfisher, Golden Weaver finch, termite,

Figure 1: A semantic network for Class 3 (with prefix *m*-)



cuttle-fish), or a combination of these (the eel is like a snake, but also like a fish; leeches and intestinal worms have an unusually intimate relationship with the human body – and *mjiko* 'lower bowel' is also in this class, so metonymy could also be operating here).

Human collectivities are entities that include human beings, but are not themselves human, so they fall somewhere between animate and inanimate. Class 3 is a compromise: it has a human-like prefix (*m-*), but a non-animate agreement pattern.

Among all the subcategories of Class 3, that of plants/trees is the most central. It contains the largest number of terms, and loanwords denoting trees and plants are almost always assigned to this class. Trees and plants also form the model for the majority of metaphoric and metonymic extensions in the class.

The subcategories below the second level on the diagram ("extended things", "active things", etc.) are interrelated in several ways: many terms fit with equal ease into more than one of these. My objective here is to suggest plausible avenues for semantic extensions; they need not be mutually exclusive. If a term fits into more than one category, it can also be regarded as well-entrenched within the semantic network of Class 3.

From the category of trees and plants it is easy to derive the property of extendedness in one dimension (a property of 'ramrod', 'metal chain', and extended body parts of humans and animals). The inclusion of objects made of plants ('wooden platter', 'straw mat', etc.), on the other hand, is based on metonymy, and so is that of coverings that are wrapped around the body, metonymically related to long body parts.

The category "powerful things" includes inanimate objects that have effects on human beings, such as substances with curative properties or religious objects (also often derived from plants). "Active things" have movement as a salient characteristic: 'arrow', 'pestle', 'chopper', 'loom pedal', etc. In contrast to the "powerful things", they must be set in motion by a human agent. Like "active body parts", they move but do not have independent volition.

Perhaps the most abstract distillation of the "entities with vitality" category is the use of the Class 3 prefix to derive deverbal nouns referring to the verbal process itself, such as *mparuro* 'a scratching' (from *-parura* 'to scratch'), *mfuo* 'a hammering' (from *-fua* 'to hammer'), *mlio* 'a sound' (from *-lia* 'to make a sound'), etc.[25] Such nouns describe a process as a thing, and so fit well with the other liminal entities in this class, that fall somewhere between animate and

inanimate.

*4.4.2 Class 7 (with prefix* ki-*)*

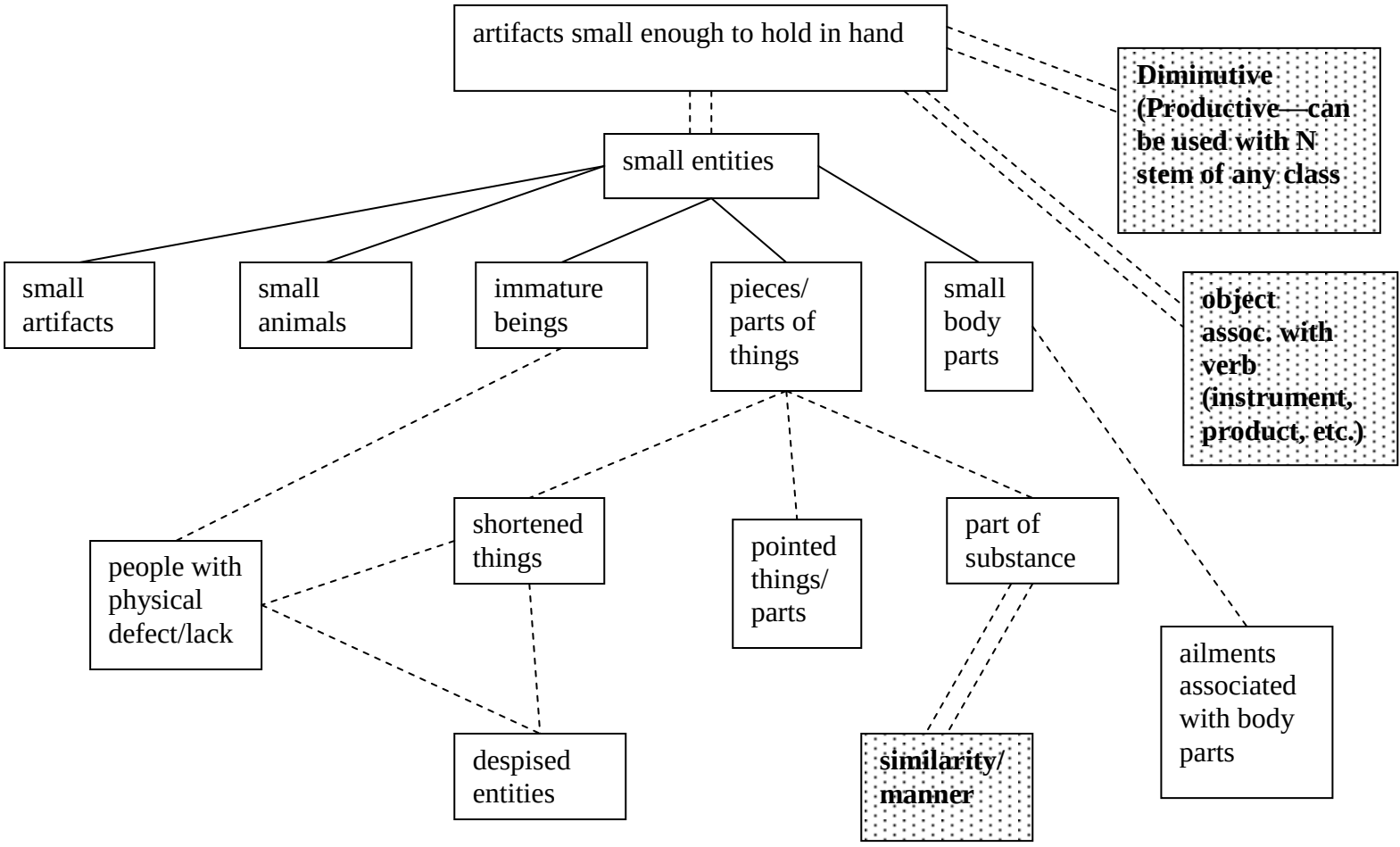
On the following page is a proposed network for Class 7. Denny and Creider (1986 [1976]: 223) state that the "primary meaning [of Proto-Bantu Class 7] is instrumental artifact". If one interprets "primary meaning" as "prototypical meaning", I think this is right, and that it is still true of Swahili. I have added to this the specification 'small enough to hold in the hand', because this applies to the majority of terms for instrumental artifacts in Class 7, and it provides a motivation for the major semantic extension within this class, to 'small entities in general', not all of which are instrumental artifacts.

Among "small entities" there are several subgroups, most of which are self-explanatory. I will comment on the ones that are less obvious. First, the category "pieces/parts of things". Parts or subdivisions of things are smaller than the whole, so this category includes both reference to size and an implicit comparison between the part and the whole (Zawawi 1979: 115 defines this class as "comparison of size or manner"). This part-whole comparison is carried over into a further extension, to "shortened things", that is things that have been truncated through being worn down or cut, and to terms for people with physical defects, conceived as not-whole. These latter terms generally have derogatory connotations in Swahili which, not coincidentally, has a single adjective for 'whole', 'healthy', and 'adult/mature', i.e. *-zima*. As pointed out by Denny and Creider (1986 [1976]), "it is a fairly natural extension from 'used object' to 'despised object'". I am arguing here that the metaphor of size plays a role in this extension in Swahili.[26]

The salient characteristic of terms in the subcategory "pointed things" is that a point or angle occupies a small amount of space. Even if the whole is large, such as a mountain, the pointed part ('peak') is relatively small. In the case of pointed parts, too, there is an implicit comparison between part and whole.

The category "part of substance" is more abstract than the ones just discussed, but its connection to part-whole relations is nonetheless apparent. This category includes terms denoting subdivisions of time and space. Height, depth, and units of measurement divide and delimit potentially extended spaces

Figure 2: A semantic network for Class 7 (with prefix *ki*-).



or spans of time into measurable parts. Here, too, there is an implicit comparison between the measured entity and the undelimited remainder.[27]

Swahili grammars often point out that the prefix *ki-* is used to derive "adverbs of manner" (cf. Ashton 1944: 165; Polomé 1967: 100), but they make no connection between this function and the noun class meaning of *ki-*. Zawawi's (1979: 115) suggestive definition for the Class 7 prefix *ki-*, "comparison of size or manner", is the first to connect these ideas explicitly, but Zawawi does not explain *how* they are connected. I believe the link lies in a metaphorical extension of the part-whole relation to qualities or attributes, in a way reminiscent of the English expression "a chip off the old block". The relationship of similarity can be thought of as a partial overlap in substance between the entities that are regarded as similar. Thus a very sweet banana overlaps in part with the sweetness of sugar; a butterfly that flies slowly partakes, so to speak, in the quality of slowness. The same rationale can be extended to deverbal nouns denoting human beings who habitually perform the action expressed by the verb.

Cross-linguistic studies of morphology associated with the message "diminutive" have shown a frequent connection between the use of such morphology and connotations of contempt (e.g., Haas 1972; Jurafsky 1993, 1996). Given the connection between small size, immaturity, and incompleteness/ lack or truncation, this association seems to be operating in Swahili Class 7 as well.[28] A number of noun stems regularly used with the *ki-* prefix (which have no diminutive connotations) designate entities that are viewed negatively by Swahili speakers. These include humans with undesirable character traits (*kibahaluli* 'fool'), odd physical features (*kigundu* 'person with unusually protruding buttocks'), or traits of ill omen in Swahili culture (*kigego* 'child whose upper teeth grow in before lower teeth'). They also include despised animals (*kiweto* 'hen that doesn't lay'), despised objects (*kibapara* 'old garment'), names of various evil spirits, and various kinds of afflictions (*kitembe* 'lisp, speech defect').

Given the salience of "despised entity" as a subcategory of Class 7, one could argue that it is unnecessary to delineate the separate subcategory "ailments associated with body parts". I have left this category separate, though connected to "despised entities," because ailments associated with body parts could be derived from the body part names with equal plausibility via the part-whole relation or the "despised" connotation.

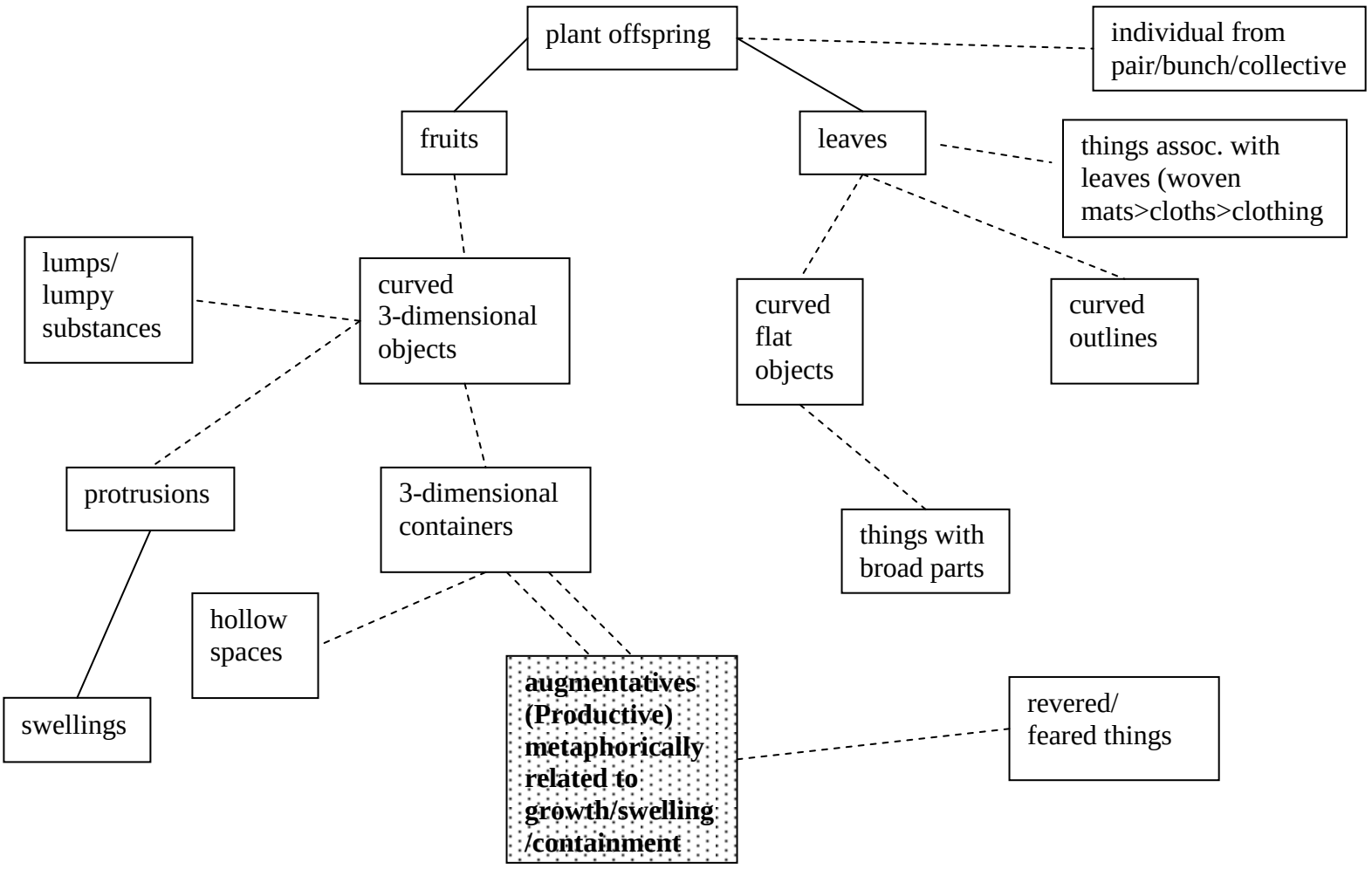
The above categories account for over 90% of the Class 7 nouns in the database. Of the rest, in some cases not enough information is provided in the dictionary gloss to decide whether they fit or not (for example, *kilendo* 'kind of fish'; *kilua* 'kind of sweet smelling flower'– are they small?). Some may be deverbal nouns or "similarity" nouns whose source is now obsolete (e.g., *kisutuo* 'food received after a task has been completed'; *kifabakazi* 'Nandi flame tree'). Some are loanwords that had *ki-* as initial syllable, and so may have been placed in this class for phonological reasons.

There is one set of apparent anomalies that deserves additional comment, however. This is a group of terms referring to large, dangerous animals or birds: *kifaru* 'rhinoceros', *kiboko* 'hippopotamus', *kingugwa* 'large spotted hyena', *kipungu* 'eagle', and *kipanga* 'Dickinson's falcon'. These terms are strikingly anomalous: why should large, predatory animals be placed in a class whose most prominent characteristic is small size, often with a connotation of insignificance? Interestingly, three of these terms ('hippopotamus', 'rhinoceros', and 'eagle') are replacements for terms that originally were in Class 9 (the "animal" class) in Proto-Sabaki (Nurse and Hinnebusch 1993: 296). In other words, these three animals were moved from the "animal" class to Class 7 (*kipanga* 'Dickinson's falcon' was already in Class 7 in PSA, and Nurse and Hinnebusch do not give a reconstructed form for 'spotted hyena'). One possible explanation is that these terms started out as euphemisms. Putting names of large, dangerous animals in the class of small, manipulable things could be a way of figuratively neutralizing or diminishing their power.[29] An alternative explanation, suggested by David Zubin (p.c.), is that these terms fit into the "despised animals" category, which is only indirectly related to size (via the association of small size with contempt).

### *4.4.3 Class 5 (with prefix* ji- ~ ∅*)*[30]

Denny and Creider (1976 [1986]) include Proto-Bantu Class 5 in the set of "configurational" classes, i.e. "prefixes which classify according to the spatial configuration of the objects classified" (p. 3). Within this group they oppose 5 to 9, the former defined as "solid shape," the latter as "outline shape," but both sharing the feature "non-extended" (i.e. "rounded, protruded, bunched, humped, etc.," p. 5). In Swahili it appears that 3-dimensionality is indeed a salient aspect of Class 5, and a large number of the terms denoting solid 3-

Figure 3: A semantic network for Class 5 (with prefix *ji-* ~ Ø)



dimensional objects, protrusions, swellings, and lumpy substances are reconstructed for Class 5 all the way back to Common Bantu. What seems to be an innovation is the inclusion of terms for containers and hollow spaces in Class 5 as well. Although some of the terms for 3-dimensional containers are reconstructed for Proto-Sabaki by Nurse and Hinnebusch (1993) and are apparently inherited from Common Bantu (*ganda* 'husk, rind, shell'; *bia* 'earthenware vessel'), a larger number are reconstructed for Proto-Sabaki only, have reconstructions that are dubious, are derived from verbs, or have changed their class affiliation from 9 to 5 (*koo* 'throat', *ziga* 'vessel for burning embers', *kaka* 'empty shell', *zizi* 'cattle enclosure', *gamba* 'outer shell'). The same is true of terms for hollow spaces, only one of which (*kwapa* 'armpit') is reconstructed back to Common Bantu.

A second major difference between the semantic structure of Class 5 in Swahili and that described by Denny and Creider (1976) for Proto-Bantu is the inclusion of terms for broad, flat surfaces and things with broad parts. In this case there is more support for the pre-Swahili existence of the subdomain within Class 5 (terms for 'leaf', 'lake', 'hoe', and 'axe' are reconstructed to Common Bantu), but here, too, the domain seems to have been extended in more recent times. Several of the terms have moved to Class 5 from other classes (*para* 'bald patch' and *panga* 'machete', both from 11; *kafi* 'paddle', from 9), some fluctuate between Class 5 and another class (*konde* 'cultivated field', also 9; *kosi* 'nape of neck', also 7); others are reconstructed to Proto-Sabaki only (*tanga* 'sail'; *kuti* 'coconut leaf'; *paa* 'roof').

The most productive semantic category within Class 5 is the category of terms for fruits. As pointed out by most Swahili grammars (see also Section 3.3.2 above), there is a regular relationship between Class 5 and Class 3: a noun stem with the prefix of Class 3 designates a plant, and the same stem in Class 5 designates the associated fruit.

The category "fruit" is productive also in the second sense: it is the basis for several kinds of semantic extension, most notably with respect to shape and size. Fruits are typically 3-dimensional, round, large (in their mature, desirable state), and can be viewed either as solid objects or as containers (skin contains fruit meat as well as seeds). Hence the extension to protrusions, swellings, lumpy substances, 3-dimensional containers, and hollow spaces (by association with containers). These associations – growth/swelling and containment especially – in turn motivate a more abstract extension, to large things in

general. And in fact a noun stem normally associated with any of the other NCMs, if given the NCM of Class 5, acquires the connotation of large size, sometimes negatively evaluated as clumsy or ungainly.

The use of the Class 5 prefix to derive augmentatives apparently can be traced at least as far back as Proto-Northeast Coast Bantu (Nurse and Hinnebusch 1993: 342); elsewhere in the Bantu family there are specialized class prefixes for this purpose (those of Classes 20, 21 and 22), for which there is no evidence in PNEC (ibid. 346). As in the case of the use of *ki-* (Class 7) to derive diminutives, this function may be seen as a natural extension from the size associated with prototypical members of the class.

It is also worth making a further point in this regard. Some might find it counterintuitive that manufactured objects – baskets, cooking vessels, and the like, which form the nucleus of Class 7 – would have the connotation of small size whereas fruits and vegetables – which are, after all, objectively smaller than many manufactured objects – would have the connotation of large size. However, this situation simply reinforces one of the central findings of recent research on linguistic categorization: that human beings classify things in the world linguistically according to their human, culturally mediated perspective, not according to objective characteristics of the things themselves. Manufactured objects are small in relation to the human body: they can be easily picked up and manipulated. Fruits are large in relation to their earlier stages of growth, and it is when they become large that they are of most value to humans. Each of these size connotations is logical in its own way, even if the resulting categorization conflicts with size as determined by objective measuring principles.

In summary, it seems that Class 5 in Swahili has retained its semantic integrity despite having lost its prefix in most contexts, making identification of a bare noun stem as Class 5 or 9-10 dependent on either associated "agreement" morphology or on method of plural formation (see Note 31). The semantic structure of 5 has probably been influenced by the development of 9-10 as residual member of the class system, however. If Denny and Creider (1976) are correct in claiming that "non-extended, outline figure" was a feature of Class 9-10 in Proto-Bantu, then the expansion of 5 into this domain could be motivated not only by the internal structure of 5 (fruits as containers), but also by the loosening of internal coherence in 9-10.

## 5. Discussion

In Section 2 of this chapter I suggested four ways in which NCMs could be analyzed from a sign-oriented perspective. The first, called the allomorphs analysis, treats all NCMs as grammatically conditioned variants of a single morpheme or sign. This analysis was rejected on the grounds that it fails to explain the fact that the groupings of noun stems with a common prefix are semantically motivated, and also on the grounds that it fails to explain how a change of prefix could lead to a change of meaning.

The second analysis discussed in Section 2, in which each NCM is analyzed as a separate linguistic sign with an abstract, invariant meaning that contrasts with the meanings of all the other NCMs, was rejected on the grounds that the analyses along these lines that have been proposed to date are either so abstract as to be untestable, or if testable, are falsified by counterexamples. The third analysis, in which the semantic networks described in Section 4 are attributed to the NCM as well as to the lexical items to which the NCM is affixed, was criticized on the basis of analytical indeterminacy and redundancy.

The fourth analysis, called the indices analysis, treats each NCM as a distinct linguistic sign whose meaning consists merely of pointing to a particular set of noun stems – the ones associated with this NCM. That is, a given NCM would mean 'the associated noun stem belongs to Class X' (as opposed to Class Y, Z etc.).[31] Since the meanings of the NCMs themselves would provide no further information as to the semantic properties of Classes X, Y, Z, etc., the differences among the NCMs would be differences of *value* rather than of *substance*, to use Saussure's terminology. The question that remains to be addressed is how (or whether) an indices analysis can avoid the objections raised against the other analyses.

First, the problem of testability. Since the meanings of the NCMs in the indices analysis consist only of identifying a noun stem as a member of one rather than another class, the analysis is essentially true by definition: a noun stem with a particular NCM belongs to a given class by virtue of the fact that it is associated with the NCM of that class rather than with a different NCM. However, exactly because the indices analysis does not specify any further semantic properties that distinguish one class from another, in principle it would be consistent with a situation in which the membership of each class was completely arbitrary from a semantic point of view. Yet as we have seen in the

discussion of Swahili Classes 3, 7, and 5, this is not at all the case in Swahili. On the contrary, each class can be shown to be organized according to readily identifiable semantic principles, even though the resulting structure may defy paraphrase in terms of a single, abstract meaning.

Furthermore, as also demonstrated in this chapter, there exist both productive and semi-productive alternations among NCMs such that substituting one NCM for another leads to a predictable change of meaning of the noun (e.g., *m-papai* 'papaya tree'/*Ø-papai* 'papaya fruit', cf. section 3.3.2). How can an indices analysis explain these semantic regularities? By itself it obviously cannot, if an explanation would consist of showing that the semantic regularities derive directly from the meanings hypothesized for the NCMs.[32] If the meaning of, say, Swahili *m-* is 'the associated noun stem belongs to Class 3', whereas that of *ji-* ~ Ø is 'the associated noun stem belongs to Class 5', these meanings in themselves cannot account for the difference in message between *m-papai* 'papaya tree' and *Ø-papai* 'papaya fruit', nor can they explain why the noun stem *-fupa* 'bone' is associated with the prefix *m-*, whereas *-cho* 'eye' is associated with the prefix *ji-*. Must the indices analysis be abandoned then? Not necessarily. What is needed is an alternative avenue of explanation for (a) the semantic principles discernible in the assignment of nouns to classes; (b) the semantic effects of substituting one NCM for another on a given noun stem. Such an explanation will be suggested below.

Columbia School theory recognizes that an understanding of the distribution of linguistic forms arises not only from consideration of the communicative function of language (i.e., from a speaker's desire to convey a particular message by means of linguistic signs), but also from the fact that language is designed for use by human beings, with our characteristic interests, abilities, and limitations:

> The general motivations (here termed "orientations") discovered by analysis are derived from the principles of acoustics, from the characteristics of human physiology in respect to a certain sound-producing mechanism and to the sound-perceiving mechanism..., from normal traits of human behavior that can be grouped under the labels of economy of effort and intelligence, and from principles of communication. (Diver 1995: 110)

As mentioned in Section 2, noun classification itself may be seen as a solution to

a particular communicative problem: that of helping to identify the intended referent of so-called agreeing elements in discourse by restricting their range of possible reference to a noun of a particular class. The overt marking of coreference is consistent with Slobin's (1973: 202) "Operating Principle E [of child language acquisition]: underlying semantic relations should be marked overtly and clearly". From the point of view of connecting modifiers with modifieds, this problem is addressed as long as the assignment of nouns to a given class remains relatively consistent across contexts, and as long as speakers know which nouns belong to which class.[33] Thus meanings that signal mere indexing of noun class can be communicatively useful in that they serve as a mnemonic device for allocating nouns to the appropriate classes.

Although the communicative purpose of indexing nouns as members of a class could be performed just as readily, in principle, regardless whether the resulting classes of nouns formed semantically coherent sets or not, in practice a further pragmatic factor comes into play here besides pure communicative usefulness, and that is ease of learning. If all pairings of NCM + noun stem were semantically arbitrary, the subdivision of thousands of nouns into six or more classes would have to be memorized entirely by rote, which would impose a burden on language learners. Although it is difficult to assess the limits of the human capacity for memorization of linguistic forms (see Sandra 1994: 243), it has long been known, for example, that people find it easier to recall lists of words that have some semantic connection with one another than lists of randomly assembled words (see e.g., Deese 1959). It therefore makes sense for the pairing of a given NCM with a particular set of noun stems to have some semantic rationale, and it also makes sense for noun stems to be assigned to different classes according to different semantic principles, so as to minimize overlap among classes. Thus the coherence within each class and the contrasts among the various classes need not be attributed directly to the meanings of the NCMs, but rather constitute a "human factor" solution to the learning task imposed by the fact of noun classification itself. That is, the meaning of the NCM motivates the semantic coherence within each class not by its own semantic substance, but rather by imposing a subdivision on nouns that speakers must learn. Language learners, faced with the task of memorizing the class affiliation of thousands of nouns, put them into categories that have some internal logic based on perceptual, functional, or cultural similarities – categories that contrast with one another along one or more of these semantic

dimensions.

This brings up the important question of psychological reality: to what extent can we assume that the semantic networks outlined in Section 4 of this chapter are shared by native speakers of Swahili? The data on which those networks were based came mostly from a Swahili-English dictionary that was published in 1939. But the vocabulary of a language is not fixed in time, nor is it a well-defined set, and different speakers have exposure to different words and learn them in different contexts. Furthermore, as mentioned in Section 4, some words could fit equally well into more than one subcategory within a network, and in some cases it is not clear which category has conceptual priority and which is an extension. Moreover, it seems quite possible that different analysts would arrive at different categorizations of the same data – why would native speakers not do likewise? The problem of analytical indeterminacy has already been mentioned as a criticism of a cognitive grammar-style polysemy-network analysis of the NCMs. The reason this is a serious criticism is that such an analysis must attribute the polysemy network to the minds of native speakers, and therefore is responsible for finding criteria for choice among different, equally plausible analyses. This problem is related to the fact that cognitive grammar makes no principled distinction between knowledge of language and pragmatic knowledge (see Langacker 1990: 4) However, the indices analysis does make this distinction, and it claims only that Swahili speakers share – as part of *langue* – a set of NCMs with their associated lexical stems, and the knowledge that formally distinct NCMs index different sets of lexical stems. As long as speakers are consistent with one another in their usage of NCM+stem combinations and the related agreement markers, the NCMs have done their semiotic work of aiding in grouping and coreference. In this analysis, the creation of semantic networks is not part of *langue*, but rather a plausible learning strategy that eases the task of memorizing the class affiliations of nouns. Exactly because the indices analysis claims no more than this about the semantic content of the NCMs themselves, it is not necessary to believe that all Swahili speakers construct identical networks in order to remember which noun stem belongs in which class.

The reader should not conclude from the above that the construction of the networks illustrated in Section 4 was merely an empty exercise, however. Given the high degree of semantic coherence discernible in the various classes, it seems likely that the most robust of these lexical connections are also perceived

by many Swahili speakers. Although the details of the networks of individual speakers may vary in relation to their own experience of Swahili vocabulary, some aspects of the analyses are likely to be shared by large numbers of speakers: specifically, subcategories that are well entrenched in the sense of (a) comprising a large set of nouns and (b) forming the basis for several types of semantic extension. The same is true of subcategories that are productive, i.e. ones that serve as the basis for the classification of neologisms, loanwords, and other creative exploitations of the system. It is also likely that words with high token frequency reinforce connections among semantically related words in the same class (see Bybee 1995 for discussion of the relation between token frequency, "lexical strength," and reinforcement of morphological relations).34 Thus the semantic networks may have psychological reality without being identical for all speakers and without being attributed to *langue*.

But if the NCMs do not directly impose a semantic structure on the nouns of their respective classes, as has been argued here, how then can change of NCM produce a change of meaning of a noun, as in the case of *m-papai* 'papaya tree'/*Ø-papai* 'papaya fruit'? This, too, can be reconciled with an indices analysis as an indirect rather than a direct consequence of the meanings of the NCMs. A NCM just allocates the associated noun stem to a particular class of nouns, to which the speaker has assigned a semantic structure like the ones discussed in Section 4, to maximize ease of learning. In effect, a particular NCM serves as an instruction to the hearer to interpret the attached lexical stem in such a way that it fits into the conceptual structure created by the set of noun stems in that class. Given that the semantic structures of the various classes are internally coherent and contrast with one another, to signal the assignment of a noun stem to a given class by means of a NCM has the semantic effect of attributing to that noun stem the semantic characteristics of the nouns of that class, in contrast with the characteristics of the nouns of some different class to which the stem might have been assigned by a different NCM. Thus an indices analysis is not inconsistent with the fact that the Swahili noun classes are semantically coherent; indeed it motivates that coherence, albeit indirectly, and the semantic structures that speakers construct for the various classes in turn motivate the semantic effects of substituting one prefix for another.

Finally, a point about the relationship of NCMs to other grammatical signs. If one accepts the argument that the function of NCMs is to partition noun stems into classes for the purpose of discourse coreference, NCMs differ

fundamentally from grammatical signs that give instructions about how to integrate lexical items into an ongoing communication, i.e., ones that have "transmissional deixis," in van Schooneveld's terms (e.g., van Schooneveld 1987). A grammatical sign such as a marker of tense or case, for example, must be relatively free in its co-occurrence potential with various lexical signs, since a particular noun stem, say, could play any number of roles in a sentence. (This is why inflectional morphemes are traditionally described as belonging to paradigms, cf. Bloomfield 1933: 223; Anderson 1992: 79.) On the other hand, for the purpose of classifying nouns so as to facilitate grouping with modifiers, a NCM should *not* be able to co-occur freely with any noun stem, but should co-occur fairly consistently with a particular set of stems. That way, modifiers could be grouped with the appropriate noun even in cases traditionally described as ellipsis or, what is more common in Swahili, in cases where several potential referents have been mentioned in the context.[35]

Unlike grammatical signs such as tenses or case markers, the association of a NCM with a given stem is not (only) motivated by the context-specific message that the speaker wants to convey in a particular utterance. As argued by Otheguy (1977), it is the stability of association between a particular NCM and a particular set of stems that makes grouping possible: if one had total freedom of choice among NCMs whenever a particular noun stem were to be mentioned, this would undermine the grouping strategy. Every noun in discourse is assigned to one or another class, so learners of Swahili are always exposed to a noun stem paired with a particular NCM. Furthermore, the NCM-stem pairings are consistent across contexts. The NCM is learned together with the lexical item itself, and it differs from other kinds of grammatical signs in the closeness of this tie. The NCMs create an intermediate category between the meanings of individual lexical items and the semantic substance shared by all Swahili nouns, i.e. ENTITY. They partition entities into various subtypes and assign the associated lexical stem to one of these subtypes.[36] The NCMs provide an organizing principle for lexical meanings (that is useful for cross-reference purposes), and thus straddle the borderline between lexicon and grammar.[37]

**APPENDIX A.** Examples of words in the Class 3 network. [Note: nouns reconstructed to Proto-Sabaki (PSA) or to Common Bantu (CB) by Nurse and Hinnebusch (1993) are so indicated. Loanwords also marked as such.]

Supernatural phenomena

| | |
|---|---|
| *mzimu* | spirit of dead person or ancestor |
| *mtume* | prophet |
| *mwiko* | taboo; something abstained from |
| *mkodi* | kind of spirit |
| *mungu* | God / PSA **mulungu* (currently being reanalyzed as Class 1-2), N&H |
| *mkosi* | bad omen, bad luck |

Natural phenomena

| | |
|---|---|
| *mto* | river / PSA **muto* |
| *moto* | fire / PSA **moto* |
| *mlima* | mountain / PSA **mulima* |
| *moshi* | smoke / PSA **mo*š*i* |
| *mvuke* | vapor produced by heat |
| *mchana* | daytime / PSA **mucana* |
| *mwonzi* | sunbeam; ray of light |
| *mwezi* | moon; month / PSA **mwezi* |
| *mrao* | cool season (<Arabic) |
| *msimu* | northeast monsoon or its season (<Arabic) |
| *mchoo* | rains between July and September |
| *mbwoji* | spring of water |
| *mfo* | flood |
| *mkaragazo* | very heavy shower of rain |
| *mkuranga* | waterless tract of country |
| *mwangwi* | echo |
| *msitu* | land covered w/thick bushes, undergrowth / PSA **musitu* forest |
| *mwamba* | rock; reef / PSA **mwamba* |

Active body parts

| | |
|---|---|
| *moyo* | heart / PSA **moyo* |
| *mdomo* | mouth / PSA **mulomo* |
| *mguu* | foot/leg / PSA **kugulu*, shift in Swahili from Class 15 |
| *mkono* | hand/arm / PSA **mukono* |
| *mkonga* | elephant trunk / PSA **mwilo*, also in Class 3 |
| *mnyiri* | cuttlefish tentacle |
| *mkia* | tail / PSA **mukila* |

Exceptional animals

| | |
|---|---|
| *mchumbururu* | swordfish |
| *mnimbi* | kind of very large fish |
| *mkunga* | eel / PSA **mukunga* |
| *mkizi* | cuttlefish |
| *mdiria* | kingfisher |
| *mnana* | Golden Weaver Finch |
| *mruba* | leech |
| *mchango* | intestinal worms |
| *mnyoo* | intestinal worms |
| *mchwa* | termite / PSA **mucwa* |

Human collectivities

| | |
|---|---|
| *mji* | town, city / PSA **muji* |
| *mtaa* | quarter, division of town / PSA **mutala* |
| *msoa* | large company of people travelling together |
| *mwima* | party of women arranging a funeral |
| *mlango* | door [used figuratively to mean one's relationship to family, clan] / PSA **mulyango* |

Powerful things

| | |
|---|---|
| *mkuyati* | aphrodisiac; tonic (<Arabic) |
| *mdundugo* | charm said to make one invulnerable |
| *mrututu* | copper sulphate, used as medicine for sores, yaws (<Arabic) |
| *mtindi* | buttermilk; any kind of intoxicating liquor |
| *mhirabu* | apse of mosque showing direction of Mecca (<Arabic) |
| *msahafu* | the Koran (<Arabic) |
| *mzumai* | bead of Muslim rosary (<Arabic) |
| *mrungura* | large drum used for summoning people |

Objects made of plants

| | |
|---|---|
| *mkeka* | straw mat / PSA **mukeka* |
| *mfure* | wooden platter |
| *mwiko* | wooden spoon; ladle / PSA **mwiko* |
| *mpira* | rubber; rubber ball; tire [also rubber tree] |

Active things

| | |
|---|---|
| *mshale* | arrow |
| *mkuki* | spear |
| *mundu* | sickle |
| *mshipi* | fishing line |
| *mtepe* | sailboat |

*mtumbwi* canoe
*mlango* door / PSA **mulyango*

Extended things

*mpiko* pole for carrying loads
*munda* harpoon; plank
*mstari* line, row (<Arabic)
*mlia* stripe
*msumari* nail, large pin (<Arabic)
*mkufu* light chain worn as ornament
*mkuruzo* drawstring
*mjeledi* leather whip; thong (<Arabic)
*mfumbi* trench dug to carry away rain water
*mfereji* ditch; irrigation trench (<Arabic)
*mfuo* groove; furrow
*mwanya* gap, narrow crevice
*mrera* lines of ornamental stitching on collar
*mpoporo* line of tattoo marks down nose; Indian file
*mfuradi* poetry verse (<Arabic)
*mwaka* year
*mwezi* month (<moon)
*muda* period of time

Extended body parts

*mfupa* bone / PSA **mufupa*
*`mgongo* back; backbone; ridge / PSA **mugongo*
*muundi* shin / PSA **mulundi*
*mtulinga* collar-bone
*mwili* whole body, inClass head and limbs / PSA **muWili*
*mkano* sinew
*mjiko* lower bowel; rectum
*mshipa* tendons/blood vessels / PSA**mušipa*
*mleli* long tail feathers of cock or bird
*mzizi* root / PSA **muzi*
*mwiba* thorn, spine / PSA **mwiWa*
*mcheche* porcupine spine
*mzoga* corpse

Extended parts of things

*mwamba* ridge pole on roof
*mtenge* canoe outrigger
*mfumbati* bedstead side piece
*mlingoti* mast
*mwimo* upright part of door or windown frame [cf. *kizingiti* top/bottom piece of door/window (Class 7)
*mdakale* opium pipe stem (<Arabic)

| | |
|---|---|
| *mkombo* | rudder handle |
| *msitamu* | kelson, to which foot and mast of vessel are secured (<Arabic) |
| *msukwano* | shaft and barrel of drill (<Arabic) |

Extended body coverings

| | |
|---|---|
| *mfuria* | loose coat (<Arabic) |
| *mdongea* | cloth to throw over shoulders |
| *mtandio* | cloth worn over shoulders [from *-tanda* 'to spread out'] |
| *mkaja* | cloth worn around body after childbirth |
| *mkumbuu* | cincture, sash, shoulder strap |
| *msuani* | woman's shroud or gravecloth |
| *mtafura* | crupper, saddle strap (<Arabic) |

**APPENDIX B**. Examples of nouns in the Class 7 network.

Small artifacts

| | |
|---|---|
| *kikapu* | basket / PSA **kikapu* |
| *kikombe* | cup / PSA **kikombe* |
| *chombo* | utensil, tool, vessel / PSA **kyombo* |
| *chungu* | cooking pot / PSA **kyungu* |
| *kiko* | pipe / PSA **kiko* |
| *kiti* | stool, seat, chair / PSA **kiti* |
| *kinu* | mortar (for grinding) / PSA **kinu* |
| *kisu* | knife / PSA **kifyu* |
| *kioo* | mirror / PSA **kilolo* |
| *kikero* | nose ornament |
| *kitana* | comb |
| *kibinjo* | whistle |

Pieces/parts of things

| | |
|---|---|
| *kipande* | piece / PSA **kipande* |
| *kichane* | something split off |
| *kigaga* | scab, crust / PSA **kigaga* |
| *kigeregenja* | splinter, fragment |
| *kinyunya* | bit of cake |
| *kipaku* | small speck, patch of color |
| *kitambaa* | scrap of cloth |
| *kibanzi* | splinter, chip |
| *kinofu* | scrap of meat |

Small animals

| | |
|---|---|
| *kidiri* | squirrel |
| *kipwe* | Senegal bush shrike |
| *kiduku* | duiker |
| *kijusi* | small lizard |
| *kimburu* | semi-wild cat |
| *chura* | frog / PSA **kyula* |
| *kijino* | whitebait-like fish |

Immature beings

| | |
|---|---|
| *kidaka* | coconut in 1st stage of growth on stem |
| *kitale* | coconut in 2nd stage of growth |
| *kidanga* | fruit in very early stage of maturation |
| *kimatu* | young locust |
| *kijusu* | 4-month-old foetus |
| *kijana* | youth |
| *kigoli* | girl of almost marriageable age |

Small body parts

| | |
|---|---|
| *kidole* | finger/toe |
| *kikoromeo* | Adam's apple |
| *kisigino* | elbow/heel / PSA **kisi(n)gino* |
| *kitovu* | navel / PSA **kitovu* |
| *kitako* | space between buttocks |
| *kizinda* | hymen |
| *kikwaru* | cock's spur |
| *kinena* | groin, pubic area / PSA **kinena* |
| *kiwele* | nipple, udder / PSA **kiWeele* |

Ailments associated with body parts

| | |
|---|---|
| *kikope* | inflammation of eyelid [cf. *kope*, Class 5] |
| *kikundu* | hemorrhoids [cf. *mkundu* anus, Class 3] |
| *kimenomeno* | pyorrhoea [cf. *meno* teeth, Class 6] |
| *kimio* | croup, throat abscess [cf. *umio* throat, Class 11] |
| *kitiiti* | breast ailment [cf. *titi* breast, Class 5] |

Pointed things/parts of things

| | |
|---|---|
| *kilembwa* | point, end |
| *kilele* | top/peak [**kilele* listed among "doubtful PSA reconstructions" by N&H] |
| *kigumba* | arrowhead / PSA **kigumba* |
| *kidevu* | chin / PSA **kilevu* |
| *kivi* | elbow |
| *kia* | arm/leg joint |

Shortened things

| | |
|---|---|
| *kibwiko* | club-foot |
| *kiselema* | tool worn down with use |
| *kingune* | tree stunted in growth |
| *kishiku* | tree stump / PSA **kisiki* |

People with physical defects/lack

| | |
|---|---|
| *kipofu* | blind person / PSA **kipofu* |
| *kiziwi* | deaf person |
| *kilema* | lame person / PSA **kilema* |
| *kibahaluli* | fool (<Arabic) |
| *kifefe* | weak, useless person (< Arabic) |
| *kijimo* | dwarf |

Part of substance

| | |
|---|---|
| *kitambo* | short period of time / Proto-Swh. **kitambo* |
| *kipindi* | period of time / PSA **kipindi* |
| *kibunzi* | end of the year |
| *kimo* | height, measure / PSA **kimo* [but SSED lists this as loan from Arabic] |
| *cheo* | measurement, rank, size / PSA **kyelo* |
| *kititi* | depth of the sea |
| *kina* | depth of the sea |
| *kiasi* | measure, quantity, moderate amount (<Arabic) |
| *kibaba* | unit of dry measure (<Arabic) |
| *kiwanja* | piece of ground / PSA **kiWanja* |
| *kivuli* | shadow, shady place [cf. *mvuli* shade, Class 3] / PSA **kivuli* |

Similarity/manner

| | |
|---|---|
| *kisukari* | very sweet banana [*sukari* sugar] |
| *kipolepole* | kind of butterfly [*polepole* slowly] |
| *kihunzi* | secret speech |
| *king'ong'o* | nasal speech [prob. onomatopoetic] |
| *Kiswahili*, etc. | names of languages |
| *kisagaunga* | crab that digs self into sand [-*saga* grind, *unga* flour] |
| *kimulimuli* | firefly [-*mulika* shine] |
| *kichoma mguu* | herb with barbed seeds [-*choma* pierce, *mguu* foot] |
| *kichwa* | head [From Common Bantu *-*cúá* termite; meaning changed to 'termite mound' via metonymy, then to 'head' via metaphor, N&H] |
| *kisamli* | Pemba coconut palm; nuts only used for drinking [*samli* ghee, clarified butter] |
| *kizuu* | evil spirit that enters home in form of rat [-*zua* bore hole, like rat] |
| *kinyozi* | barber [-*nyoa* to shave off hair] |
| *kichinja mimba* | youngest child [-*chinja* cut, *mimba* conception – also fits "immature beings" category] |

Afflictions/character defects

| | |
|---|---|
| *kifafa* | fits, convulsions |
| *kigagazi* | nausea |
| *kiruu* | madness, blind rage (<Arabic) |
| *kiherehere* | heart palpitation |
| *kitembe* | speech defect, lisp |
| *kitwea* | loneliness |
| *kisalisali* | gonorrhoea; bilharzia |
| *kikuli* | horror, fright |
| *kiburi* | arrogance (<Arabic) |
| *kisonono* | gonorrhoea |

Defective/despised animals

| | |
|---|---|
| *kidimu* | fowl w/ruffled neck feathers (not good to eat) |
| *kivunjavunja* | mantis (if killed causes one to break things) |
| *kipasuasanda* | screech owl (bad omen) |
| *kiwavi* | chrysalis;caterpillar with irritating hairs |
| *kiweto* | hen that doesn't lay |

Evil spirits

| | |
|---|---|
| *kibwengo* | evil spirit living in big trees or sea |
| *kitimiri* | evil spirit sometimes used as charm |
| *kizimwi* | ghoul |
| *kizuka* | ghost, spirit, sthg. that appears suddenly |
| *kizuu* | kind of evil spirit in rat form |
| *kinyamkela* | evil spirit of crossroads; whirlwind |
| *chamchela* | whirlwind; spirit that causes them |

Despised humans

| | |
|---|---|
| *kimwondo* | simpleton |
| *kiberenge* | prostitute |
| *kigundu* | someone with protruding buttocks |
| *kifefe* | weak, useless person |
| *kibahaluli* | fool |
| *kikaramba* | old person (contemptuous) |
| *kigego* | child whose upper teeth grow in first (unlucky) |
| *kikausha* | one who brings bad luck [-*kausha* cause to dry up] |

Despised things

| | |
|---|---|
| *kichelema* | over-watery cassava/potatoes |
| *kibapara* | old garment |
| *kibunzi* | anything insignificant |
| *kichepe* | cloth worn threadbare |
| *kichoro* | small illegible writing, scrawl |

| | |
|---|---|
| *kifu* | dead thing |
| *kigosho* | bend, crook, esp. if deformity |
| *kijineno* | silly prattle |
| *kimelea* | parasitic plant |
| *kilema* | deformity, deformed person |
| *kilimi* | bad, abusive speech style |
| *kizimwe* | something dried up, withered |
| *kimba* | corpse, heap of dung |
| *kinyaa* | filth, anything disgusting |
| *kinyesi* | excrement |

**APPENDIX C**. Examples of nouns in the Class 5 network.

Fruits

*tunda* fruit / Proto Northeast Coast Bantu **itunda*
most specific fruit names
[productive: name of associated plant is same stem in Class 3]
Example: *papai* papaya
[cf. *mpapai* papaya plant, Class 3]

Lumps/lumpy substances

| | |
|---|---|
| *kumbwe* | mouthful of food |
| *tonge* | lump of food [also Class 9] / PSA **itonge* |
| *vumbu* | lump in flour |
| *poroja* | porridge or similar mass |
| *kamasi* | nose mucus / PSA *?*ikamasi*, listed as "doubtful" by N&H |
| *povu* | bubble, froth, foam / PSA **ipovu*, CB **podu* |
| *jivu* | ash [also Class 6] / PSA **ivu*, CB **bu* |
| *tope* | mud [also Class 6 and 11] / PSA **itope*, CB **tope* |
| *pumba* | lump, clod / PSA *?*mpumba ~ ipumba* |
| *tumbawe* | coral not yet consolidated |
| *joya* | spongy substance inside coconut shell |
| *shata* | lees from making coconut oil (<Arabic) |
| *shudu* | refuse after seeds crushed for oil (<Arabic) |
| *sira* | dregs, lees (<Arabic) |

Curved 3-dimensional objects

| | |
|---|---|
| *jiwe* | stone / PSA **jjiWe*, CB **bue* |
| *kaa* | lump of charcoal / PSA **ikala*, CB **kada* |
| *yai* | egg / PSA **iyayi (iyagi),*CB **jayi, yagi* |
| *ini* | liver / PSA **ini* |
| *kende* | testicle, scrotum / PSA **ikende*, CB **kende* |
| *tumbo* | belly, womb / PSA **itumbo*, CB *tumbo* |
| *ziwa* | breast / PSA **iziWa* milk, loan from Proto-Cushitic, replaces PSA **iWeele* |
| *konde* | closed fist / PSA **ikonde* |

*pafu* lung / PSA **?ipafu* ("doubtful", N&H)
*jicho* eye / PSA **ijico*, CB **yico*
*sikio* ear / PSA **isikilo* (<-*sikil* hear)
*jino* tooth / PSA **ijino*, CB **yino*
*ua* flower / PSA **iluWa*, CB **duba*
*chozi* *teardrop* / PSA **icozi*, CB **codi*
*jua* the sun / PSA **ijuWa*, CB **yuba, juba*
*wingu* cloud / PSA **iWingu*, CB **bingu*
*gogo* log / PSA **igogo*, CB **gogo*
*shingo* neck / PSA **nshingo* (Class 9), CB **kingo*
*tokoni* coccyx
*zio* half of something round, e.g. coconut (also Class 7)
*tofali* brick, tile (<Arabic)
*takia* large cushion (<Arabic)

3-dimensional containers

*koo* throat (also Class 9) / PSA **nkolo* (Class 9, from *-kol* choke)
*jiko* fireplace, hearth / PSA **ijiko*, CB **yiko*
*gamba* outer covering: fish scale, tree bark, tortoise shell / PSA **igamba* shell
*ganda* husk, rind, shell / PSA **ikanda* skin, bark ? CB *kanda* skin
*kaka* empty shell, e.g. of egg, fruit / PSA **?ika(n)ka* ("doubtful", N&H)
*kununu* empty grain husk, spike
*kumbi* fibrous plant sheath
*bia* earthenware vessel / PSA **iWiga* pot, CB **biga*
*ziga* vessel for burning embers / PSA **iziga*, CB **-dig* become burnt
*zizi* enclosure for keeping animals / PSA **izizi* cattlefold
*kopo* metal can, jug, etc.
*rumbi* large jar
*pakacha* basket for fish, fruit
*toza* tobacco pipe's bowl
*tondoo* thimble
*jamanda* round covered basket (<Portuguese)
*jarife* dragnet, seine (<Arabic)
*kabati* cupboard (<English)
*kandili* lantern (<Arabic)
*sanduku* box, chest, trunk (<Arabic)
*kombora* bomb, shell, mortar (<Arabic)
*tangi* water storage tank (<English)

Protrusions

*tako* buttock / PSA **itako*, CB **tako*
*shavu* cheek, biceps, calf / PSA **icavu*, CB **cagu*
*kano* large sinew or tendon / PSA **ikano*, CB **?kanda* strap
*bega* shoulder / PSA **iWega*, CB **bega*
*fuzi* tip of shoulder / PSA **ifuuzi*, CB **tuudi*
*shungi* plait, hair tress / PSA **isungi* lock of hair, CB **?cuki* hair
*tawi* branch, bough w/growing fruit, grain / PSA **itaWi*, CB **tabi*
*kole* coconut palm branch

*shina* root/stem of tree / PSA **ishina* base of tree trunk, CB **kina*
*tengo* outrigger of canoe / PSA *tengo* (uncertain class), CB **?tengo* stool
*tango* eaves of house
*bawa* wing (also Class 11) / PSA **iWaWa ~ luWaWa ~ ipapa*, CB **baba, papa*
*wimbi* wave / PSA **iWimbi*, CB **bimbi*
*tuta* raised bed for planting
*taya* jaw
*jabali* cliff, jutting rock (<Arabic)
*jukwaa* stage, scaffold (<Arabic?)
*kanadili* projection from stern of vessel (<Arabic)
*omo* forepart of ship (<Arabic)
*tezi* stern, poop of ship (<Portuguese)

Swellings

*jipu* boil, abscess / PSA **ipu*
*lengelenge* blister,pustule / PSA **ilenge* pumpkin, blister, CB **denge* pumpkin
*kaka* whitlow (disease of hand)
*kororo* throat swelling
*tezi* tumor, goiter
*pisho* cautery; mark made by cautery
*jeraha* wound, sore (<Arabic)

Hollow spaces

*pengo* gap, notch, hole / PSA **ipengo*
*shimo* hole, pit, cavity / PSA **isimo*
*tundu* hole (also Class 9) / PSA **ntundu* (Class 9)
*kwapa* armpit / PSA **ikwapa*, CB **yapa, kuapa*
*koongo* hole dibbled for planting seeds
*korongo* ravine, water channel made by rainy-season stream
*komeo* creek, inlet of sea
*kuo* furrow, trench, hollowed out hole
*pelezi* gaps, blunt places in knife or axe
*kara* pause for someone to start reading Koran
*dirisha* window (<Persian)

Large things

A stem from any class can be placed in Class 5, with resulting meaning 'a big one of these', e.g. *mtu* person (Class 1), *jitu* giant (Class 5). Fully productive. Loanwords denoting large things also often in Class 5, e.g. *lori* truck (<English).

Revered/feared things

*koikoi* kind of evil spirit
*zimwi* goblin, ogre
*shetani* demon, devil (<Arabic)

| | |
|---|---|
| *jini* | djinn (genie) (<Arabic) |
| *juju* | Gog (appears on Judgment Day) (<Arabic) |
| *kaburi* | grave, tomb (<Arabic) |
| *kafara* | offering, sacrifice to avert evil (<Arabic) |
| *talasimu* | talisman, charm (<Arabic) |
| *ziara* | tomb, monument, pilgrimage, visit (<Arabic) |
| *tanga* | formal mourning ceremony (also Class 6) / PSA **itanga* |
| *kanisa* | church (<Arabic) |

Leaves

| | |
|---|---|
| *jani* | leaf, grass blade / PSA **ijani*, CB **jani* |
| *kaa* | frond, e.g. of date tree |
| *kuti* | whole coconut leaf [frond is Class 11] / PSA *ikuti* |
| *kanja* | leaf of coconut palm with plaited fronds |
| *tapa* | palm leaf used for umbrella |

Curved flat objects/surfaces

| | |
|---|---|
| *ziwa* | lake, pond, marsh / **iziWa*, CB **diba* |
| *kaa* | the palate / PSA **ikala ~ ilaka*, CB **daka* throat |
| *kosi* | nape of neck (also Class 7) / PSA **ikosi*, CB **koti* |
| *panja* | temples, side of head where hair recedes / PSA **ipanja*, CB **panja* baldness |
| *para* | bald patch on head / PSA **lupala* (Class 11), CB **pada* |
| *kope* | eyelid, eyelashes / PSA **lukope* (Class 11), CB **kope* |
| *tanga* | sail of vessel / PSA **itanga* mat, sail |
| *paa* | roof / PSA **ipala* |
| *konde* | cultivated field (also Class 9) / PSA **nkonde* (Class 9), CB **konde* |
| *toasi* | cymbal, large castanet |
| *jasi* | round earlobe ornament (<Arabic) |

Things with curved outlines

| | |
|---|---|
| *tao* | something curved: arc, arch, river bend, hem |
| *kota* | crook, bend, e.g., of crooked leg condition; cf. CB **-kotam-* become bent |
| *pambizo* | margin, outskirts |
| *timbi* | bracelet, armlet |
| *tinda* | string of beads to go around neck |
| *pochi* | chain wrist bangle (<Persian) |
| *jebu* | braided ornament worn around head (<Hindi) |

Things with broad parts

| | |
|---|---|
| *jembe* | hoe / PSA **igembe ~ ijembe*, CB **jembe ~ gembe* |
| *shoka* | axe / PSA **icoka*, CB **coka* |
| *kafi* | paddle, small steering oar / PSA **nkafi(a)* (Class 9), CB *kapi* |
| *kasia* | oar |
| *panga* | machete / PSA **lupanga* (Class 11) |

| | |
|---|---|
| *kolego* | spade, shovel |
| *parange* | broad bladed knife for clearing forest |
| *tezo* | carpenter's adze |

Individual from pair/bunch

| | |
|---|---|
| *koshi* | slipper (<Persian) |
| *wele* | grain of millet |
| *hindi* | grain of maize |
| *kombamoyo* | roof rafter pole |
| *kara* | splinter, spark, piece |
| *pacha* | twin |

**Notes**

*I would like to thank Erica García, the editors of this volume, and the anonymous referees, for helpful comments on earlier drafts of this chapter. Any errors that have passed through such a discerning net are my own.

1. Dixon was concerned with establishing criteria for differentiating between noun classes and noun classifiers, such as the numeral classifiers of Chinese or Burmese. I have omitted his discussion of noun classifier systems, which are not relevant to the present paper.
2. A particularly forceful, and influential, statement of this position is the classic one from Sapir (1921: 100): "It is almost as though at some point in the past the unconscious mind of the race had made a hasty inventory of experience, committed itself to a premature classification that allowed of no revision, and saddled the inheritors of its language with a science that they no longer quite believed in nor had the strength to overthrow. Dogma, rigidly prescribed by tradition, stiffens into formalism. Linguistic categories make up a system of surviving dogma – dogma of the unconscious. They are often but half real as concepts; their life tends ever to languish away into form for form's sake."
3. This statement conceals some complexities. In Swahili the realization of the noun class marker for some classes or morphophonemic contexts may be zero, in which case identification of noun class must depend on the so-called agreement marker used on accompanying words such as demonstratives or adjectives. (This is the rule in a language like German, where classification occurs exclusively on agreeing elements and is not overtly marked on the noun.) Consequently, for some words that occur only in fixed expressions without a modifier of any kind it may be difficult to determine their noun class affiliation, because markers of different classes may be homonymous. For example, the Swahili word *chafya* 'a sneeze' apparently occurs only in the expression *-piga chafya* 'to sneeze' (literally, 'to strike a sneeze'). The word *chafya* looks like a stem *-afya* (a noun that occurs in Class 9 with the meaning 'health') preceded by the (prevocalic allomorph of the) prefix of Class 7, *ki-*. But the first syllable could also just be part of the stem, in which case

the noun could belong to Class 5 or 9 (both of which often lack an overt prefix). Presumably one could determine the analysis historically by looking at related languages or dialects, but it still would not be clear how Swahili speakers analyze this word synchronically. Grammarians sometimes classify such words as adverbs rather than nouns (and correspondingly have identified what has been called an adverbial use of the noun class prefix *ki*). But this is merely an acknowledgment of the ambiguous status of such words in relation to the noun class system.

4. A slightly different analysis of Swahili noun class prefixes and associated "concordial" markers is suggested by Harris (1945). Harris treats each noun class prefix and all concordial markers as a discontinuous morpheme, "spread out over a syntactically recognizeable domain" (1945: 125). As for the morphemic status of the noun class prefixes themselves, Harris suggests regarding each prefix as part of the noun stem, analogous to the *sl-* of English *slide, slick*, etc. (ibid., pp. 126-7), i.e. "all nouns have discontinous repeated parts" (126). A more recent analysis, very similar in spirit to that of Harris, has been proposed by Amidu (1997, Ch. 2). The same objection may be made against both types of "allomorph"analysis (see below).
5. Otheguy (1977) analyzes Spanish gender along similar lines; however he does not take the step of claiming that membership in a semantically arbitrary class can be regarded as a kind of "meaning". Carstairs-McCarthy (1994: 741) on the other hand argues that membership in a particular inflection class (including grammatical gender classes) can be regarded as part of the "meaning" of an affix, if one interprets "meaning" as "information content," which subsumes "not just extralinguistic meaning, such as plurality or pastness, but also intralinguistic information such as inflection class or gender or stem shape". Carstairs-McCarthy's general goal is to demonstrate that the psycholinguistic Principle of Contrast (Clark 1987) applies to inflectional morphology, not just to lexical items, rather than to explore why languages would have formally differentiated but semantically arbitrary inflection classes in the first place, but his findings are consistent with a functional motivation for noun classification, as he points out himself (p. 784).
6. The problem of accounting for the distribution of the associated "concordial" markers when there is a lack of concord, cited as a counterargument to the "allomorph" analysis, could be dismissed by a proponent of the "indices" analysis on the grounds that the "concordial" markers would require a separate analysis in order to determine whether or not they would be expected to match up consistently with the noun class prefixes in discourse.
7. Different analyses identify different numbers of classes. The variation depends not on disagreement over the identification of the relevant morphemes, but on whether singular/plural prefix pairs are regarded as a single class or divided into separate classes, each with its own number. The latter expedient is commonly used in Bantu linguistic descriptions in order to facilitate comparative and diachronic studies, and a uniform numbering system is used in which cognate noun class markers are assigned the same number in each language. I use the standard Bantu numbering system here not only because

of its familiarity to Bantuists, but because it makes it easier to address the question whether the "plural classes" are really just plural equivalents of their corresponding singulars. In what follows, I will cite paired noun classes with numbers separated by a hyphen, e.g., Class(es) 1-2. For those familiar with Bantu languages, I should note that I omit Classes 16, 17, and 18 (the so-called locative classes) from discussion here because these prefixes are never attached to noun stems in Swahili. For an analysis of the locatives, see Contini-Morava (1976).

8. Class 11/14 has two numbers because it derives historically from the merger of two formerly distinct classes marked by nominal prefixes **lu* and **Wu*, which lost their initial consonants. The original class distinction is now preserved only in the plural: nouns derived historically from Class 11 have associated plurals in Class 10, whereas those derived from Class 14 either denote abstractions that have no plural (e.g., *uchawi* 'witchcraft') or have associated plurals in Class 6. There is no formal distinction between these classes in the singular.

9. Note that the so-called nominal and adjectival prefixes in Table 1 are identical, with the sole exception of Class 11/14, which has *u-* as nominal prefix and *m-* as adjectival. The question whether to distinguish between nominal and adjectival prefixes, most of which are homonymous, is controversial in Swahili. Some scholars make this distinction (e.g., Ashton 1944: 9); others do not (e.g., Polomé 1967: 94-95; Zawawi 1979: 116; Nurse and Hinnebusch 1993: 344). Space does not allow discussion of this question here; see Contini-Morava (forthcoming) for details.

10. The considerable formal differentiation among the "pronominal" prefixes of Class 1 is doubtless semantically/pragmatically motivated. Class 1 is the most semantically coherent class in Bantu languages, being comprised only of nouns denoting animate beings, almost all human. It is common for languages to make finer distinctions when referring to humans or animates than when referring to inanimate objects, and the asymmetry between singular (Class 1) and plural (Class 2) in this respect is also not unusual. The distribution of the Class 1 "pronominal" prefixes is as follows:

    *yu*: with demonstrative stem, e.g., *yu-le* 'that one (anim. sg.)'

    *ye*: relative pronoun affixed to V, e.g., *a-li-ye-siki-a* '[the one] who heard [3sg subj.-past-rel.pn.-Vstem-indic.]'

    *w*: with possessive stem, e.g., *w-angu* 'my (anim. sg. possessed object)'

    *a*: subject prefix on V, e.g., *a-li-siki-a* 's/he heard'

    *m*: object prefix on V, e.g., *a-li-m-siki-a* 's/he heard him/her'

    For all other classes, identical pronominal prefixes are used in all these contexts.

11. I have analyzed Class 9-10 as having no prefix in modern Standard Swahili. In this respect Class 9-10 differs from Class 5, whose prefix is often realized as zero. The difference is that between a meaningful absence of overt form (in the case of Class 5) and mere absence of

form (in the case of Class 9-10). Class 9-10 nouns remain invariable in singular and plural and are distinguished only by the associated pronominal prefix, whereas in Class 5 absence of a prefix in the singular alternates with presence of an overt prefix in the plural (e.g., ∅ *-shoka/ma-shoka* 'ax/axes'). Space does not allow discussion of details here; see Herbert (1978); Contini-Morava (2000).

12. An exception is Zawawi (1979, Ch. 5), who argues that the notion of noun class as a fixed pairing of a prefix with a set of lexical stems should be abandoned, and proposes instead a classification of prefixes into three systems: a system of nominal indicators (including the prefixes labeled nominal and adjectival in Table 1), a system of pronominal indicators (including the so-called pronominal prefixes of Table 1), and a system of locational indicators (not included in Table 1). She assigns each nominal indicator an invariant meaning, and regards the co-occurrence of prefixes with lexical stems as semantically motivated. Although this approach is obviously very attractive to a sign-oriented linguist, I do not believe it tells the whole story, as I will argue below.

13. In case the reader is curious, a sample paragraph of written Swahili narrative (from Mohamed 1976: 128) analyzed with the morphological parsing program AINI (Schadeberg and Elias 1989) shows that of 129 word tokens only 25, or 19%, are monomorphemic – these being mostly repetitions of proper names. For comparison, in a sample paragraph from an English narrative (Dick 1969 [1991]: 58-59), 180, or 76%, of 237 word tokens are monomorphemic – counting irregular past tenses like *took* as bimorphemic and singular nouns like *face* as monomorphemic.

14. The assumption that part-of-speech categorization is inherent to lexical items has been challenged recently from a generative-minimalist perspective for Straits Salish by Jelinek and Demers (1994).

15 Some of what are said to be derived nouns may become relatively specialized in meaning, when they are the means of labelling a culturally relevant entity, e.g., ∅*-pato* 'profit' (noun, Class 5) < *-pata* 'get' (V). In such cases even though the semantic relationship between the nominal and verbal uses of the lexical stem is transparent, one could argue that the noun has developed a narrower meaning that is not just the sum of its morphological parts. (A similar point is made by Langacker 1988: 16 in reference to the English expression *pencil sharpener*.) Should we then speak of two homonyms, in the present example *-pat*$_1$- 'get' (the verbal use) and *-pat*$_2$- 'profit' (the nominal use)? There can be no general answer to this question, which evidently would depend on the degree of "entrenchment" (to borrow another term from cognitive grammar) that a particular, specialized sense has in general usage. But it seems likely that on the whole, nominal uses of stems would have a tendency toward greater specialization than verbal uses, just because there are many ways that entities can be involved in occurrences: as agents, instruments, outcomes, names for the occurrence itself, etc. Which role will be the most culturally useful to name by a noun will depend on the nature of the occurrence and the messages that speakers find a recurring need to convey by means of a single lexical item.

16. This fact is probably related to the more general observation made by Hopper and Thompson (1984: 745) that "languages often possess rather elaborate morphology whose sole function is to convert verbal roots into N's, but no morphology whose sole function is to convert nominal roots into V's". They suggest a pragmatic explanation for this: human beings find it easier to treat events as entities than vice versa (in examples like English *to table, to dog* etc. an event is named after an associated entity, but this is not the same thing as treating an entity as an event, and is not parallel to reverse cases such as *proposal, excitement*).
17. Class 9-10, which I have analyzed as lacking a prefix (see Note 11), cannot be used productively to form derived nouns. Instead, replacement of a NCM by zero is invariably interpreted as deriving a Class 5 augmentative noun (see c below) – which constitutes further evidence that Class 5 has a zero prefix, whereas Classes 9 and 10 have no prefix at all. For discussion of the derivational productivity of NCMs in Bantu more generally, see Mufwene (1980).
18. The derived noun *mkutano* also includes a suffix *-o*, regularly used to indicate that the attached lexical item is to be regarded as an entity rather than an occurrence (i.e. what is called a deverbal noun). This *-o* may co-occur with all noun class prefixes except those of Classes 1-2 (the so-called animate classes), as it denotes only inanimate entities.
19. The limitation of Class 1-2 to humans is a universal feature of Bantu languages, as is the difference in semantic transparency between this class and the others. The "human factor" explanation for this hardly needs to be spelled out.
20. David Zubin (1995) notes, based on his cross-linguistic study of noun classification, that it is a mistake to assume that productive subclasses are at the center of the semantic structure of a category.
21. The database fields are NOUN, GLOSS, CLASS, HUMAN, ANIMAL, PLANT, INANIMATE, [force of] NATURE, SHAPE, SIZE, NUMBER, AFFECT, DERIVED, LOAN, and SOURCE [in case more dictionaries are added]. The semantic fields are further subcategorized: for example ANIMAL includes invertebrate, snake, reptile [other than snake], mammal, domestic, insect, amphibian, fish, bird, origin [e.g., horn used as musical instrument], made [e.g., honeycomb], part [e.g., fish fin], social organization [e.g., killed as sacrifice]. The database tags are not meant to be mutually exclusive; more than one can be entered for a given noun. They were chosen based in part on my own knowledge of Swahili language and culture, in part on categories found relevant in earlier cross-linguistic work on noun categorization (e.g., Adams and Conklin 1973; Craig 1986), and in part on categories recognized by earlier studies of Swahili and other Bantu languages (e.g Meinhof 1948 [1906]; Ashton 1944; Polomé 1967; Denny and Creider 1976; Hinnebusch 1979; Zawawi 1979; Spitulnik 1987, 1989). The categories are avowedly "etic": they were used in order to make it possible to manipulate large quantities of data with reasonable flexibility, but they do not constitute an analysis. And in fact the analysis does not reflect the database tags in a direct way.

22. The few exceptions are all nouns denoting animate beings: *mdudu*, the generic term for 'insect', *mnyama*, the generic term for 'animal', and a few names of animals that are agentives derived from verbs, e.g., *mpasuasanda* 'nightjar' [literally, 'shroud-tearer', a bird of ill omen].
23. As pointed out earlier, the NCMs signal information about number in addition to indexing a particular noun class. Traditionally, the classes have been paired, with odd numbers assigned to singular prefixes and even numbers to their corresponding plurals. In fact, the relationship between singular and plural in Swahili (and in Bantu more generally) is not as straightforward as this pairing suggests. Instead of a binary opposition between 'one' and 'more than one', it appears that most of the classes are arranged along a scale of three degrees of individuation, with 'singular' and 'plural' as subcategories of the highest degree of individuation:

    figure 4 here – but can't put drawing objects in endnotes

    For details, see Contini-Morava (2000). Since analysis of the System of Number is beyond the scope of the present chapter, I will discuss only the singular classes at the highest level of individuation, namely Classes 3 ($m_2$-), 7 (*ki*-), and 5 (*ji*-), leaving aside Class 1 ($m_1$-) and 9-10 (no prefix), for the reasons mentioned above.
24. The lists of examples in the Appendices include forms reconstructed for Proto-Sabaki (PSA) or Proto-Swahili by Nurse and Hinnebusch (1993), if available. In their notation W means a voiced labial approximant (exact feature specification unclear, cf. N & H: 89ff). Where a word is listed as a loanword in Johnson (1939), I indicate the source language.
25. As mentioned earlier (Note 18), these deverbal nouns also have a suffix *-o*, used with various NCMs to derive nouns from verb stems.
26. Zubin (1995 and p.c.) suggests that the most abstract semantic level for Swahili Class 7 is "minor-counterpart relation," and that the part-whole relation is the metaphorical source for smallness of size rather than vice versa. His argument is based in part on his cross-linguistic research into partitive relations and noun classification. From a logical point of view, it is as easy to derive part-whole from small size as the reverse, so the question is what empirical evidence one can find about the direction of derivation, with regard to Swahili. There are certainly a large number of nouns in Class 7 that denote pieces/parts (e.g., *kibanzi* 'splinter, chip', *kinofu* 'scrap of meat' etc.), but there are also many that denote small entities that are not parts (e.g., *kiko* 'pipe', *kikapu* 'basket', *kisu* 'knife' etc.). Since smallness is common to

both groups, whereas part/whole is not, I prefer to regard the former as more basic in Swahili.

27. This category in Swahili shows interesting parallels to the Ewe morpheme *ví*, discussed in Heine et al. (1991: 84).
28. My thanks to David Zubin (p.c.) for pointing out the significance of this subcategory within Class 7, and for the Jurafsky reference.
29. Khamis (1984) describes this as a "metaphorical" use of *ki-* 'for excess', although 'ironic' might be a more appropriate term. I would like to thank Abdulaziz Lodhi (p.c.) for bringing this paper to my attention. The use of *ki-* in an inversion of its more literal meaning still seems to be productive in Swahili, as evidenced by the following forms from Ohly (1987): *kifurushi* 'large buttocks'(dim. of *furushi* 'bundle', Class 5), *kizigo* 'large buttocks' (dim. of *mzigo* 'load', Class 3), *kibunda* 'a lot of money' (dim. of *bunda* 'bale', Class 5).
30. Historically, nouns of Class 5 used to have a distinctive prefix, reconstructed as phonologically conditioned allomorphs *ji□ and *i□ for Proto-Sabaki by N&H (p. 338). However, in most Swahili dialects the prefix (now *j-* or *ji-*) has been retained only before vowel-initial and monosyllabic noun stems, which are comparatively infrequent, and indeed it is recognizable as a prefix only with the monosyllabic stems. (Of the 656 Class 5 nouns in my database, only 72, or 10.9%, begin with *j-*. However for almost all polysyllabic stems the initial *j-* must be analyzed as part of the noun stem, since it is not replaced by a plural prefix when the noun is pluralized. For example, compare *jani* 'leaf', pl. *majani* with *jicho* 'eye', pl. *macho*.) Thus most nouns of Class 5 have a Ø prefix, and in this respect they are indistinguishable from most nouns of Class 9, the so-called residual class, whose historical prefix **ni-* has been lost. The grammatical distinction between Classes 5 and 9 is maintained not by differences in the form of the noun, but by differences in pluralization and agreement patterns: Class 5 noun stems are associated with the prefix *ma-* in the plural, whereas those of Class 9 do not change in the plural; the pronominal agreement for Class 5 is *li-*, whereas that for Class 9 is *i-*.
31. Note that this analysis need not apply to Swahili Class 1-2, which can be analyzed straightforwardly as meaning 'animate entity' (singular and plural respectively). Also, the analysis being discussed here ignores the fact that the Swahili NCMs signal information about number in addition to noun class. The topic of number is dealt with in a separate paper (Contini-Morava in press); see also Note 23.
32. Such a requirement of meanings is sometimes suggested in Columbia School work: "...a *meaning* in this technical sense is a theoretical construct..., its postulation is justified by its ability to account for the deployment of its associated *signal* by speakers of the language in communicating messages" (Reid 1991: 95). And "In the Columbia School view, the purpose of a grammatical analysis is to explain people's deployment of forms in discourse... The appearance of grammatical forms in discourse is explained by their being signals of instrumental meanings." (Huffman 1995: 210)
33. The same point is made by Heath (1985: 104-5) with reference to Ngandi: "N[oun] C[lass]

cross-reference greatly assists in weeding out incorrect construals." Heath also points out that this function is independent of the degree of semantic coherence within the various classes. I thank the anonymous reviewers of this volume for the Slobin and Heath references.

34. A word frequency count exists only for written Swahili, based on newspapers and literary texts (Bertoncini 1973), but it is suggestive that of the 500 most frequent words in the list, all the Class 3 nouns fit directly into the subcategories of Figure 1 (Section 4): natural phenomena (*mwitu* 'forest', *mchana* 'daylight', *mto* 'river', *moto* 'fire'), active body parts (*moyo* 'heart', *mkono* 'hand/arm', *mguu* 'foot/leg'), active things (*mlango* 'door', *mpira* 'ball'), extended things including periods of time (*mwaka* 'year', *mwezi* 'month', *mwili* 'body', *muda* 'period of time'), human collectivities (*mji* 'town, city'), and deverbal nouns.

35. This point is made elegantly by Otheguy (1977: Ch. 2), in relation to the co-occurrence of the Spanish articles *el/la* with lexical items that designate inanimate objects. Otheguy argues that the pairing strategy, whereby a particular word is regularly associated with one or the other *l*-form, makes it possible to identify the intended referent of an *l*-form even if the noun is not explicitly mentioned. In Swahili the "ellipsis" strategy is rather rare: nouns designating inanimate objects (i.e., the great majority of the noun stems subclassified by the NCPs) are not usually sustained as topics of discourse long enough to justify ellipsis. On the other hand, the need for grouping arises when several nouns have been mentioned and need to be grouped with appropriate modifiers. A common construction in Swahili, for example, is Noun + *-a* + Noun, in which the noun following *-a* further specifies the meaning of the noun preceding *-a* (e.g., *ch-uo ch-a Ø-elimu* 'school [Class 7] of [Class. 7] education [Class 9]). The *-a*, called "*-a* of Association" by Ashton (1944), carries the agreement prefix corresponding to the class of the noun that is being further specified. Since it is possible to have chains of nouns linked by *-a*, the agreement helps sort the specified(s) with the intended specifier(s). For example:

    *ch-uo ch-a elimu y-a wa-tu wa-zima ch-a Dar es Salaam*

    school(7) of(7) education(9) of(9) people(2) adult(2) of(7) Dar es Salaam

    'the School of Adult Education of Dar es Salaam'

    Here the phrase "of Dar es Salaam" could potentially specify any of the preceding nouns (school, education, or adults). Its connection to *chuo* 'school' is made clear by the what is called the agreement prefix attached to *-a*.

36. This *formulation* resembles that of van Schooneveld in his analysis of gender in Russian:

    The lexical meaning of the given referent who is identified in the *parole* is assigned to a certain class of lexical meanings. Thus, a tension is created between the identity of the

given referent of a given lexical meaning in the given *parole* and a set of which the given referent is an element. It is to be noted that this set, a given gender category, is neither identical with the set constituted by the given lexical meaning, in other words, the *signatum* of the given lexical morpheme, nor is it identical with the superset constituted by the linguistic category (part of speech) of substantive. A given gender category is a subset of the category of gender. Gender is, in turn, a superset which, so to say, is suspended in the *langue* between lexical meaning and the part of speech of substantive (van Schooneveld 1977: 131).

37. The use of noun classification for cross-reference purposes may also be found in languages with numeral classifiers, which typically have anaphoric functions even though they are not associated with grammatical agreement (see e.g., Lyons 1977: 464; Lucy 1992).

## References


Adams, K. and N. Conklin. 1973. "Toward a theory of natural classification." *Papers from the Ninth Regional Meeting*. Chicago Linguistics Society, 1-10.

Amidu, Assibi. 1997. *Classes in Kiswahili*. Köln: Rüdiger Köppe.

Anderson, Stephen. 1992. *A-morphous Morphology*. Cambridge University Press.

Aronoff, Mark. 1994. *Morphology by Itself.* Cambridge, MA: MIT Press.

Ashton, E.O. 1944. *Swahili Grammar*. London: Longmans.

Bertoncini, Elena. 1973. "A tentative frequency list of Swahili words." *Annali dell'Istituto Orientale di Napoli* 33: 297-363.

Bloomfield, Leonard. 1933. *Language*. University of Chicago Press, reprinted 1984.

Boas, Franz. 1911 [1966]. "Introduction." *Handbook of American Indian Languages*. Reprinted in Preston Holder (ed), *Franz Boas: introduction to handbook of American Indian languages; J.W. Powell: Indian linguistic families of America north of Mexico*. Lincoln: University of Nebraska Press.

Bybee, Joan. 1995. "Regular morphology and the lexicon." *Language and Cognitive Processes* 10(5): 425-455.

Carstairs-McCarthy, Andrew. 1994. "Inflection classes, gender, and the Principle of Contrast." *Language* 70(4): 737-788.

Casson, Ronald. 1981. *Language, Culture, and Cognition: Anthropological perspectives*. New York: Macmillan.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Clark, Eve. 1987. "The principle of contrast: a constraint on language acquisition." In Brian MacWhinney (ed), *Mechanisms of Language Acquisition*. Hillsdale, NJ: Erlbaum, 1-33.

Cole, Pascale and Juan Segui. 1994. "Grammatical incongruency and vocabulary types." *Memory and Cognition* 22(4): 387-394.

Contini-Morava, Ellen. 1976. "Statistical demonstration of a meaning: the Swahili locatives in existential assertions." *Studies in African Linguistics* 7(2).

Contini-Morava, Ellen. 1994. "Noun classification in Swahili." *Institute for Advanced Technology in the Humanities, Research Reports, 2nd Series.* Charlottesville: University of Virginia. http://jefferson.village.virginia.edu/swahili/swahili.html

Contini-Morava, Ellen. 1995. "Introduction: on linguistic sign theory." In E. Contini-Morava and B.S. Goldberg (eds), 1-39.

Contini-Morava, Ellen and Barbara Sussman Goldberg (eds). 1995. *Meaning as Explanation: Advances in linguistic sign theory*. Berlin: Mouton de Gruyter.

Contini-Morava, Ellen. 1996. "'Things' in a noun class language: semantic functions of grammatical agreement in Swahili." In Edna Andrews and Yishai Tobin (eds), *Toward a Calculus of Meaning: Studies in markedness, distinctive features, and deixis.* Amsterdam and Philadelphia: Benjamins.

Contini-Morava, Ellen. 1997. "Noun classification in Swahili: a cognitive semantic analysis using a computer database." In Robert K. Herbert (ed), *African Linguistics at the Crossroads:Papers from Kwaluseni.* Köln: Rüdiger Köppe Verlag, 599-628.

Contini-Morava, Ellen. 2000. "Noun class as number in Swahili." In E. Contini-Morava and Y. Tobin (eds), *Between Grammar and Lexicon*. Amsterdam and Philadelphia: Benjamins.

Contini-Morava, Ellen. Forthcoming. "Swahili morphology." In Alan Kaye (ed), *Morphologies of Asia and Africa*. Winona Lake, IN: Eisenbrauns.

Corbett, Greville. 1991. *Gender*. Cambridge University Press.

Craig, Colette. 1986. "Introduction." In C. Craig (ed), *Noun Classes and Categorization*. Amsterdam and Philadelphia: Benjamins, 1- 10.

Deese, James. 1959. "Influence of inter-item associative strength upon immediate free recall." *Psychol. Rep*. 5: 305-312.

Denny, J. Peter and Chet Creider. 1976. "The semantics of noun classes in Proto-Bantu." *Studies in African Linguistics* 7(1): 1-30. Reprinted in Craig (ed.), *Noun Classes and Categorization*. Amsterdam and Philadelphia: Benjamins, 1986, 217-239.

Dick, Philip K. 1969. *Ubik*. New York: Doubleday. Reprinted by Vintage Books, 1991.

Dinneen, F.P. 1967. *An Introduction to General Linguistics*. New York: Holt, Rinehart and Winston.

Diver, William. 1972. *Avoidance of the Obvious: The pronoun as minimax solution*. Unpublished ms., Columbia University.

Diver, William. 1975. "Introduction." *Columbia University Working Papers in Linguistics* 2: 1-25.

Diver, William. 1977. "A concise grammar of Modern English I." *Columbia University Working Papers in Linguistics* 4: 1-21.

Diver, William. 1995. "Theory." In E. Contini-Morava and B.S. Goldberg (eds), 43-114.

Dixon, R.M.W. 1986. "Noun classes and noun classification in typological perspective." In Colette Craig (ed), *Noun classes and categorization*. Amsterdam: Benjamins, 105-112.

Fodor, Istvan. 1959. "The origin of grammatical gender". *Lingua* 8(1): 1-41; 186-214.

Frankl, P. J. L. in consultation with Yahya A. Omar. 1994. "Diminutives and insignificance, augmentatives and 'monstrosity': examples of class re-assignment in Swahili." *South African Journal of African Languages* 14(3): 113-116.

García, Erica. 1975. The role of theory in linguistic analysis: the Spanish pronoun system. Amsterdam: North-Holland.

Givón, Talmy. 1972. "Studies in ChiBemba and Bantu grammar." *Studies in African Linguistics Supplement* 3.

Greenberg, Joseph. 1978. "How does a language acquire gender markers?" In Greenberg, J. (ed), *Universals of human language* Vol. 3. Stanford: Stanford University Press.

Guillaume, Gustave. 1984. *Foundations for a science of language*. Transl. by Walter Hirtle and John Hewson. Amsterdam and Philadelphia: Benjamins.

Guthrie, Malcolm. 1971. *Comparative Bantu*. Volume 2. Farnborough: Gregg.

Haas, Mary. 1972. "The expression of the diminutive." In M.E. Smith (ed), *Studies in linguistics in honor of G.L. Trager*. The Hague: Mouton, 148-152.

Harris, Zellig. 1945. "Discontinuous morphemes." *Language* 21: 121-127.

Heath, Jeffrey. 1985. "Discourse in the field: clause structure in Ngandi." In J. Nichols and A. Woodbury (eds), *Grammar inside and outside the clause: some approaches to theory from the field*. Cambridge University Press.

Heine, Bernd. 1982. "African noun class systems." In. H. Seiler and C. Lehmann (eds), *Apprehension. Das sprachliche Erfassen von Gegenstanden*, Vol. 1. Tubingen: Narr,190-216.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: a conceptual framework*. University of Chicago Press.

Herbert, Robert. 1978. "Morphological re-analysis in the Bantu nasal class." *African Studies* 37(1): 125-137.

Hinnebusch, Thomas. 1979. "Swahili." In Timothy Shopen (ed), *Languages and their status*. Philadelphia: University of Pennsylvania Press, 209-293.

Hirtle, Walter. 1975. *Time, aspect, and the verb*. Québec: les Presses de l'Université Laval.

Hopper, Paul and Sandra Thompson. 1984. "The discourse basis for lexical categories in universal grammar." *Language* 60(4): 703-752.

Huffman, Alan. 1995. "The purpose of a grammatical analysis." In E. Contini-Morava and B.S. Goldberg (eds), 185-211.

Jelinek, Eloise and Richard Demers. 1994. "Predicates and pronominal arguments in Straits Salish." *Language* 70(4): 697-736.

Johnson, Frederick. 1939. *A Standard Swahili-English dictionary*. Oxford University Press.

Jurafsky, Dan. 1993. "Universals in the semantics of the diminutive." In Joshua Guenter, Barbara Kaiser and Cheryl Zoll (eds), *Proceedings of the 19th Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society.

Jurafsky, Dan. 1996. "Universal tendencies in the semantics of the diminutive." *Language* 72(3): 533-578.

Khamis, Said Ahmed Mohamed. 1984. "The forgotten *ki-* in Kiswahili." *Lugha* 3: 13-19.

Lakoff, George. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. University of Chicago Press.

Langacker, Ronald. 1987. *Foundations of cognitive grammar*. Stanford, CA: Stanford University Press.

Langacker, Ronald. 1988. "An overview of cognitive grammar." In B. Rudzka-Ostyn (ed), *Topics in cognitive linguistics*. Amsterdam and Philadelphia: Benjamins, 3-48.

Langacker, Ronald. 1990. *Concept, image, and symbol: the cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Lucy, John. 1992. *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*. Cambridge University Press.

Lyons, John. 1977. *Semantics*, Vol. 2. Cambridge University Press.

Meinhof, Carl. 1906 [1948]. *Grundzüge einer vergleichenden Grammatik der Bantusprachen*. Berlin. 2nd edition, Hamburg: Dietrich Reimer, 1948.

Mohamed, Mohamed S. 1976. *Nyota ya Rehema*. Nairobi: Oxford University Press.

Moxley, Jeri. 1998. "Semantic structure of Swahili noun classes." In Ian Maddieson and Thomas Hinnebusch (eds), *Language history and linguistic description in Africa*, Asmara: Africa World Press, 229-238.

Mufwene, Salikoko. 1980. "Bantu noun class prefixes: inflectional or derivational?" *Papers from the 17th Meeting of the Chicago Linguistic Society*, 246-258.

Nurse, Derek and Thomas Hinnebusch. 1993. *Swahili and Sabaki: a linguistic history*. University of California Press.

Ohly, Rajmund. 1987. *Swahili-English slang pocket-dictionary*. Vienna: Institut für Afrikanistik.

Otheguy, Ricardo. 1977. *The meanings of Spanish* el, la, *and* lo. Ph.D. dissertation, City University of New York.

Otheguy, Ricardo. 1980. "Meaning in the grammar and in the lexicon." *Columbia University Working Papers in Linguistics* 5: 6-11.

Polomé, Edgar. 1967. *Swahili language handbook.* Washington, DC: Center for Applied Linguistics.

Poplack, Shana, D. Sankoff and C. Miller. 1988. "The social correlates and linguistic processes of lexical borrowing and assimilation." *Linguistics* 26: 47-104.

Pozdnyakov, Konstantin I. 1993. *Sravnit'el'naya grammatika atlanticheskix yazykov: im'ennyye klassy i fono-morfologiya*. Moscow: Nauka.

Reid, Wallis. 1974. "The Saussurean sign as a control in linguistic analysis." *Semiotexte* 1(2): 31-53.

Reid, Wallis. 1991. *Verb and noun number in English: a functional explanation*. London: Longman.

Reynolds, Karl. 1989. *The structure of the Kiswahili nominal*. Ph.D. dissertation, University of Washington.

Reynolds, Karl and Carol Eastman. 1989. "Morphologically based agreement in Swahili." *Studies in African Linguistics* 20(1): 63-77.

Sandra, Dominiek. 1994. "The morphology of the mental lexicon: internal word structure viewed from a psycholinguistic perspective." *Language and Cognitive Processes* 9(3): 227-269.

Sandra, Dominiek and Sally Rice. 1995. "Network analyses of prepositional meaning: mirroring whose mind – the linguist's or the language user's?" *Cognitive Linguistics* 6(1): 89-130.

Sangster, Rodney. 1982. *Roman Jakobson and beyond: language as a system of signs*. Berlin: Mouton de Gruyter.

Sapir, Edward. 1921. *Language: an introduction to the study of speech*. New York: Harcourt, Brace, and World. Reprinted by Harcourt Brace Jovanovich.

Schadeberg, Thilo and P.S.E. Elias. 1989. *AINI: a morphological parser for Kiswahili, version 1.3*. Department of African Linguistics, State University at Leiden, the Netherlands.

Selvik, Kari-Anne. 1996. *Setswana noun classes: conceptual categories marked by grammar?* Cand. Philol. thesis, University of Oslo.

Slobin, Dan. 1973. "Cognitive prerequisites for the development of grammar." In C. Ferguson and D. Slobin (eds), *Studies of child language development*. New York: Holt, Rinehart and Winston.

Spitulnik, Debra. 1987. "Semantic superstructuring and infrastructuring: nominal class struggle in Chibemba." *Studies in African Grammatical Systems*, 4. Bloomington, IN: Indiana University Linguistics Club.

Spitulnik, Debra. 1989. "Levels of semantic structuring in Bantu noun classification." In Paul Newman and Robert Botne (eds), *Current approaches to African linguistics 5*. Dordrecht: Foris, 207-220.

Tobin, Yishai. 1990. *Semiotics and linguistics*. London: Longman.

van Schooneveld, C.H. 1977. "The place of gender in the semantic structure of the Russian language." *Scando-Slavica* 23: 129-138.

van Schooneveld, C.H. 1983. "Programmatic sketch of a theory of lexical meaning." *Quaderni di semantica* 4(1): 158-170.

van Schooneveld, C.H. 1987. "Linguistic structure and autopoeisis." In K. Pomorska and Roman Jakobson Colloquium (eds), *Language, poetry, and poetics: the generation of the 1890's – Jakobson, Trubetzkoy, Majakovskij*. Berlin: Mouton de Gruyter.

Wurzel, Wolfgang. 1989. *Morphology and naturalness*. Berlin: Akademie-Verlag.

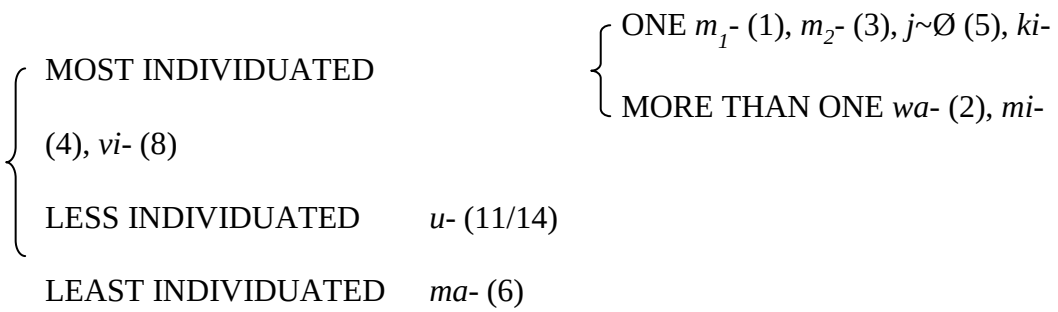
Zawawi, Sharifa M. 1979. *Loan words and their effect on the classification of Swahili nominals*. Leiden: E.J. Brill.

Zubin, David. 1995. *The cognitive typology of nominal classification systems*. Handout of talk given at the Department of Linguistics, University of Colorado, 9/27/95.

Zubin, D. and K-M. Köpcke. 1986a. "Gender and folk taxonomy: the indexical relation between grammatical and lexical categorization." In C. Craig (ed.), *Noun classes and categorization*. Amsterdam and Philadelphia: Benjamins,139-180.

Zubin, D. and K-M. Köpcke. 1986b. "Natural classification in language: a study of the German gender system." *Buffalo Cognitive Science Report* No. 2, SUNY Buffalo.

Figure 4. System of degree of individuation in Swahili.




[Unmarked for individuation: no prefix (Classes 9-10)]