The point of Bantu, Chinese and Romance nominal classification

Paola Crisma, Lutz Marten & Rint Sybesma*

This paper focuses on nominal classification in Bantu, Romance and Chinese.

The relation between numeral classifiers, noun classes and grammatical gender has often been noted in the typological literature (e.g. Craig 1986, Senft 2000), based, on the one hand, on the (perceived) comparable semantic parameters involved in the classification underlying all three types of system (shape, size, animacy, and others), and, on the other hand, on the frequent relation between classification and agreement, both within the nominal phrase and with constituents outside it.

However, despite these apparent similarities, the establishment of a common core of these three systems (if there is one) (represented here by Cantonese classifiers, Italian gender and Swahili noun classes) remains problematic, as well as the question why we have classification at all.

In this paper, we give a fine-grained description of the three systems, showing that Bantu and Romance are very similar, not only for the common pervasive agreement facts, but also because both families allow for derivational uses of noun classes and gender seem to involve similar semantic domains. These two characteristics set Chinese apart.

We propose that a common rationale underlying these classification systems can still be discerned, and is suggested by the fact that the expression of classification and number often, though not always, go hand in hand: in our view, class and number are both part of the individuating process that turns a descriptive "predicate" NP into a referential expression.

1. Introduction

In this paper we discuss nominal classification in languages from three different language families, Bantu, Romance and Sinitic. Our main goal is to find out what is the point of classification. Why do languages go through the trouble of classification? We approach this question by investigating in detail what the elements involved in classification really do and what characteristics they really have. This will lead us to discuss differences and similarities between inflection and derivation and the relation between class and number.

Rivista di Linguistica 23.2 (2011), pp. 251-299 *(received October 2011)*

Section 2 presents a detailed overview of the noun classification system in a Bantu language, namely, Swahili, taking data from other Bantu languages into consideration as well. Section 3 introduces the basics of the system of Italian and several other Romance languages, with the aim of addressing the question as to how different Romance and Bantu/Swahili really are. Section 4 is devoted to Sinitic (discussing Mandarin and Cantonese). The system of Chinese looks quite different, but it drives home the important point that all three systems, despite all their differences, have one thing in common, namely, that the elements that express class are also, directly or indirectly, involved in the expression of number. Section 5 will be concerned with the question why that should be the case.

2. Nominal classification in Bantu / Swahili

2.1. Overview

Bantu noun classes are a well-known example of nominal classification. They are sometimes described as the most grammaticalized classification system, contrasting with classifier systems, such as found in Chinese, at the other end of a lexico-grammatical continuum (Grinevald 2000). In a typical Bantu language, about 15-20 different noun classes are distinguished, and each noun is assigned to a particular class. Noun classes are distinguished by noun class prefixes, a set of class specific agreement markers, as well as, to some extent, the particular semantic content of a given class (Maho 1999). In this section we discuss a number of aspects of Bantu noun classes which will form the background for the comparison with gender in Romance and noun classifiers in Chinese in the subsequent sections. We will illustrate Bantu noun classes with primary reference to Swahili, as Swahili is comparatively well described, and there are several detailed studies of noun classification we can draw on (e.g. Amidu 1997, Carstens 1991, 2008, Contini-Morava 1994, 1996, 1997, 2000, Moxley 1998, Schadeberg 2001). In view of our overall macrocomparative approach, we are particularly interested in three questions: first, the possible semantic base for class membership and whether class membership is arbitrary or semantically motivated; second, the difference between any inherent/static class meaning and semantic regularities in the derivational use of classes, such as in class shift (e.g. diminutives, augmentatives, abstracts, locatives) and nominal derivation (e.g. agent, process, result, state), and, finally, different pos-

sible morphosyntactic analyses of Bantu noun classes, for example, whether Bantu noun classes grammaticalize number and/or gender, and whether noun class prefixes can be analysed as morphological spell-out of (gender) features or as nominalizing heads. Overall, we will show that Bantu noun classes are partly semantically motivated, but not to the extent that predictions about class membership can be made on semantic grounds, and that Bantu noun classes incorporate notions of plurality, but that these have an intermediate status between syntax and semantics.

2.2. Swahili noun classes

Swahili distinguishes 15 noun classes as summarized in Table 1 (see Ashton 1947, Mohamed 2001, Schadeberg 1992). Bantu noun classes are numbered conventionally (Katamba 2003, Maho 1999), and in any given language not all classes may be found – or example, classes 12 to 14 are not found in Swahili, and so in Table 1 class 15 follows class 11. Like in other Bantu languages, most Swahili noun class prefixes are of CV-type, but Swahili has no vocalic pre-prefix or augment as found, for example, in Bemba, Luganda or Zulu (see e.g. Blois 1970, Stump 1993). As mentioned above, classes are distinguished by different noun class prefixes and distinct agreement morphology, and so, in Swahili, classes 1 and 3, as well as 9 and 10, have the same class prefix, but differ in agreement morphology. Agreement morphology in many classes differs from the noun class prefix, although, except for class 1, the different agreement markers of each class can be related to one underlying form. Nouns denoting humans typically show "animate agreement", i.e. concord and (sometimes) referential and possessive concord of class 1/2, irrespective of the class of their noun class prefix. As can be seen from the examples given, classes up to 10 contain many pairs of singular and plural, with odd classes including singulars and even classes including plurals. Plurals of countable class 11 nouns are typically found in class 10, and classes higher than 11 do not take part in the singular-plural pairing. Class 15 contains only infinitives and all verbal roots can be used in this class. The locative classes 16-18 contain only one lexical word, *mahali* 'place'. However any word which can be thought of as a location can be used in class 16-18. The historical class prefixes *pa-*, *ku-* and *mu-* are no longer used with locative nouns in Swahili, where locatives are marked by a suffix *-ni*. However, *pa-*, *ku-* and *mu-* are used as adjective agreement and the various concords. Non-countable nouns are found in all classes (further discussed below), and there are some

exceptions to plural assignment, such that, for example, one singular form may have plural forms in different classes. From a Bantu comparative point of view, singular-plural pairing shows some variation from language to language, especially with higher classes (Katamba 2003, Maho 1999). The typical 'meaning' given in the last column of Table 1 is indicative of some assumed core semantics, which can be related to a number of examples in the relevant class. Classes 1 and 2 are the best examples of a semantic core for noun classes in Swahili, as they include almost exclusively nouns referring to humans,[1] although not all such nouns are found in classes 1 and 2. The following classes are less intuitively coherent, and although many words, for example, for trees and plants are found in classes 3 and 4, the classes also contain a number of other nouns. The same situation is found in classes 5/6, 7/8, 9/10 and 11, while class 15 contains all and only infinitives, and classes 16-18 are exclusively locative. The question to what extent Bantu noun classes are semantically motivated has attracted considerable attention, and we will discuss it in more detail in the following section.

Table 1. Swahili noun classes.[2]

Class	Class prefix	Example word	Concord	Referential Concord	Possessive Concord	'Meaning'
1	m	mtu 'person'	a/yu	ye	wa	People
2	wa	watu 'people'	wa	o	wa	
3	m	mti 'tree'	u	o	wa	Trees, plants
4	mi	miti 'trees'	i	yo	ya	
5	ji/∅	jicho 'eye'	li	lo	la	Round things, liquids, masses, augmentatives
6	ma	macho 'eyes'	ya	yo	ya	
7	ki	kiti 'chair'	ki	cho	cha	Artefacts, tools, manner, diminutives
8	vi	viti 'chairs'	vi	vyo	vya	
9	n/∅	ndege 'bird'	i	yo	ya	Animals, loanwords
10	n/∅	ndege 'birds'	zi	zo	za	
11	u	ubao 'board'	u	o	wa	Long things, abstracts
15	ku	kuimba 'to sing'	ku	ko	kwa	Infinitives
16	(pa)	mahali 'place'	pa	po	pa	Locatives
17	(ku)		ku	ko	kwa	
18	(mu)		mu	mo	mwa	

2.3. Inherent semantics of noun classes

In answer to the question whether Bantu noun classes are semantically based, two main approaches can be distinguished. One approach assumes that noun class assignment is an arbitrary lexical quality, so that it has to be learned during language acquisition and does not reflect any underlying semantic categorisation (e.g., Carstens 2008, Idiata et al. 2000, Idiata 2005, Richardson 1967). The alternative approach proposes that noun classification reflects semantic tendencies, or is built around a semantic core, and that class assignment is semantically motivated (e.g. Contini-Morava 1994, 1996, 1997, 2000, Denny & Creider 1976, Hendrikse 2011, Moxley 1998, Palmer & Woodman 2000, Sagna 2008,[3] Selvik 2001, Spitulnik 1987). The main question with respect to the semantics of noun classes is thus to what extent, if at all, noun classes are semantically based. An early study addressing this question is Denny & Creider's (1976) analysis of reconstructed, Proto-Bantu noun classes, which is based on generic semantic criteria such as animates, kinds, and masses, as well as concepts of shape and spatial configuration, such as "extended solid outline figure". However, the classification has never been applied to an actual, contemporary Bantu language, and more recent approaches have typically employed cognitive linguistic models of explanation. In a study of Shona noun class 3, Palmer & Woodman (2000) use large-scale dictionary comparison and cognitive linguistic methods and show that of the 941 class 3 nouns in their database, 36% refer to concepts such as tree, bush, plant, herb, etc. They propose that the class comprises several related semantic clusters including trees, shrubs, and herbs, and, related to this, the basic scenario of mortar and pestle such as groups and bundles, long, thin things, pounding, grain, noise, repetition, and crushing and witchcraft. Contini-Morava (1994, 1996, 1997, 2000) develops a cognitive linguistic analysis of Swahili noun classes. She proposes that Swahili noun classes are organised as semantic networks, including possibly several prototypical meanings and radial, increasingly peripheral meanings with respect to the relevant prototype, as well as meanings related to the network's core meanings through metaphorical extensions. For example, for class 7, Contini-Morava (1994) proposes a network of meanings built around the core, prototypical meaning of "utilitarian objects small enough to hold in hand". To this core meaning, different meanings are related through metaphorical or metonymical extension, with the most remote meaning being "similarity/manner", a sub-sense of "parts of substance", to explain words for languages and culture-specific

traits found in the class (such as *kizungu* 'in a European style/way').[4] The aim of the work is to explain semantic regularities and tendencies with reference to general cognitive processes, and since different meanings may be related through a variety of meaning relations, the analysis does not aim to predict class membership, but rather shows the relation between noun class membership and the underlying semantic categories which it reflects.

Another approach in the study of the semantics of noun classes is psycholinguistic. Selvik (2001) reports on two psycholinguistic experiments related to noun classes 1, 3, 5 and 7 in Setswana. Speakers were presented with nonsense words such as the class 7 nonsense word *serutsa* and different possible meanings for it, more or less close to the assumed prototypical meaning of class 7. In the second test, conversely, speakers were presented with a meaning such as 'a tool that is used for making soap' and had to choose the best fit from four possible nonsense words from different classes. The majority of responses in both tests showed associations of noun classes with respective prototypical meanings. For example, the majority of speakers identified the relevant class 7 nouns (*serutsa* and *sebôrôlêta*) with a tool for making soap, which was the exemplar for the prototypical meaning for class 7. Selvik (2001) concludes that in Setswana semantic criteria play a role in noun class assignment, and more specifically, that Bantu noun classes are instances of polysemous categories. In contrast, evidence from language acquisition does not seem to support this conclusion. In studies of the acquisition of Sotho (Demuth 2000) and of Isangu (Idiata 2005), no evidence for semantic effects due to noun classification, such as overgeneralization in the use of noun classes, was found, except for instances based on animacy.

In sum, while there is agreement that there are no necessary and sufficient semantic criteria for establishing membership in Bantu noun classes, and thus noun class assignment is not predictable on semantic grounds, there are several studies which show that the opposite conclusion, that semantic criteria do not play any role in noun classification in Bantu, is too strong. Based on cognitive linguistic and psycholinguistic evidence, a picture emerges where noun class membership and assignment is at least to some extent related to and motivated by general semantic principles of semantic characterisation. However, another domain of potential relevance for the role of semantics of noun classes is related to derivational processes, which we will discuss in the next section.[5]

2.4. Derivational use of noun classes

In addition to studies of class membership, the derivational use of classes has been studied against the background of investigating semantic regularity, since regular semantic relations between different classes might be taken as evidence for the relevance of semantic criteria for classification (e.g. Givón 1972). In addition, evidence from relations between classes has been used to investigate the importance of grammatical number for noun classes. As noted earlier, some classes appear as singular-plural pairs. This can be explained as a grammatical-inflectional relationship involving the grammatical category of number, or, alternatively, as a lexical-derivational relationship involving semantic notions of individuals and groups, while in terms of grammatical category, class, rather than number, is the relevant feature. This latter position has been proposed for Swahili in Schadeberg (2001), on which a good part of the following discussion is based, and which will be discussed in more detail in Section 3. Similarly, Van der Spuy (2010) argues for Zulu that noun classes are exponents of class, rather than gender, and that plural nouns are derived from singular nouns. A particular aspect of the role of derivation in the noun class system is the status of the locative classes 16-18, as these are almost exclusively used derivationally (but for *mahali*, see above). In many Bantu languages, a locative class prefix is prefixed to an already inflected noun, with its own class prefix, and also in Swahili, the locative suffix *-ni* is suffixed to forms with their own class prefixes. Assuming that inflectional categories are typically only marked once on a given root, locative nouns morphologically reflect their derivational nature:

(1) a. ki-tanda
7-bed
'bed'

b. ki-tanda-ni
7-bed-LOC
'on/at the bed'

For other classes we see that, with respect to count nouns, singular and plural forms are typically distinguished by class membership: *kiti* (class 7) 'chair', *viti* (class 8), 'chairs'. However, noun stems may additionally be found in other classes than 'their' singular and plural classes, and this class shift is usually accompanied by a regular semantic change, as the following examples show (data from Johnson 1939):

(2) a. *mtoto / watoto* 'child, off-spring' (class 1/2)
b. *kitoto / vitoto* 'small child, baby; also childish manner' (diminutive/manner) (class 7/8)
c. *toto / matoto* 'big, fine child; object resembling offspring' (augmentative) (class 5/6)
d. *utoto* 'childhood, dependence' (quality) (class 11)

(3) a. *kivuli / vivuli* 'shadow, shady place; also sometimes ghost, apparition' (class 7/8)
b. *jivuli / mavuli* 'big, large shadow etc.' (augmentative) (class 5/6)
c. *mvuli / mivuli* 'shady place, shade of a tree, etc.' (class 3/4)
d. *uvuli* 'shade, shadiness in general' (quality) (class 11)

The examples show that the noun stems *-toto* 'child' and *-vuli* 'shade' are found in different classes. In both examples, classes 5/6 are augmentative and class 11 denotes a quality related to the meaning of the stem. In (2), *kitoto/vitoto* (classes 7/8) are diminutive, while in (3), *mvuli/mivuli* (classes 3/4) denote the effect or place of shade. While semantic relations in class shift tend to be regular, there are exceptions as well, as this last example shows. The most common regular semantic relations involved in class shift (next to plural, if this is included) are diminutive, augmentative, collective, manner, quality, and fruits (Schadeberg 2001):

(4) a. *mtoto* 'child' (class 1) > *kitoto* 'small child' (class 7) (diminutive)
b. *nyumba* 'house' (class 9) > *jumba* 'big house' (class 5) (augmentative)
c. *rafiki* 'friend' (class 9) > *rafiki* 'friends' (class 10) (plural)
marafiki 'group of friends' (class 6) (collective)
d. *mfalme* 'king' (class 1) > *kifalme* 'royal manner' (class 7) (manner)
e. *mtoto* 'child' (class 1) > *utoto* 'childhood' (class 11) (quality)
f. *mchungwa* orange tree' (class 3) > *chungwa* 'orange' (class 5) (fruit)

In many Bantu languages, in some derived nouns the derivational class prefix is added to a noun which already has a class prefix. Multiple prefixation shows that the lexical use of noun classes is at least in some instances morphologically distinguished from the derivational uses of classes. In Swahili, some instances of multiple prefixes are found, but their use is less regular than in other Bantu languages. In Herero, for example, nouns in some classes, such as class 11, retain their original prefix when the noun is shifted to a different class. However, except for locatives, derived nouns behave syntactically like non-derived nouns in that agreement is typically with the derivational class, irrespective of the number of prefixes (Kavari & Marten 2009):

(5) otji-ru-vyó tj-ándje / *rw-ándje
7-11-knife 7-my 11-my
'my big knife'

With locatives there is some variation. Agreement in Swahili, for example, has to be with the derived, locative class, while in siSwati agreement has to be with the original class (Marten 2010), and in Luganda both options are possible (Marten *forthc*.):

(6) m-oyo-ni mw-angu / *w-angu [Swahili]
3-heart-LOC 18-my / 3-my
'in my heart'

(7) ku-ba-fana ba-mi / *kw-ami [Swati]
LOC-2-boys 2-my / LOC-my
'at my boys'

(8) a. ku-bbalaza kw-ange [Luganda]
17-9.terrace 17-my
'on my terrace'
b. ku-ky-alo ky-ange
17-7-village 7-my
'in my village'

The opposite of class shift, where nouns are found in different classes, are cases of nouns found only in one class. Such 'one-class nouns' do not take part in singular-plural shifts – even though they may be found in classes which have a corresponding singular or plural class. They are found in all classes except the 'human' classes 1/2, and semantically include masses, liquids, abstract nouns, infinitives, and locations.

(9) a. *mchana* 'daytime, daylight' (class 3)
b. *mikambe* 'kicking game played in water' (class 4)
c. *joto* 'heat' (class 5)
d. *mauti* 'death' (class 6)
e. *mafuta* 'fat' (class 6)
f. *kiu* 'thirst' (class 7)
g. *vidondo* 'small chips of wood' (class 8)
h. *virugu* 'anger' (class 8)
i. *njaa* 'hunger (class 9)
j. *ufalme* 'kingdom' (class 11)
k. *udongo* 'clay' (class 11)
l. *kuimba* 'to sing' (class 15)
m.*mahali* 'place' (class 16/17/18)

The existence of these one-class forms may be taken to show that Swahili noun classes are not in a systematic number relation, and that singularity and plurality are semantic concepts associated with the relevant noun classes, although all classes support members without these attributes, as (9) shows. An interesting question in this context is whether one-class nouns are common or rather exceptional. While no systematic study of this question exists, data from Contini-Morava's (1994: Sect. 3.1) lexical study of 4,784 nouns in Johnson's (1939) Swahili dictionary shed some light on it. For paired nouns, the dictionary lists the noun under the odd, singular class, so, for example, 334 class 1/2 nouns are listed as class 1 in the numerical breakdown in (10) below. This would also be true for nouns which are only found in a singular class, so in principle, among the 334 class 1 nouns there could be some which do not have an associated plural in class 2 (although no such example exists). For odd classes of class 3 and higher the data do not show one-class nouns, even though we know that some exist. However, any noun listed as an even (plural) noun is an unpaired, one-class noun, so that, for example for class 4, the data show that there are 21 nouns which do not have a corresponding singular form in another class (as if they did, the noun would be listed as a class 3 noun):

(10) Class 1: 334

Class 3: 924	Class 4: 21
Class 5: 755	Class 6: 94
Class 7: 761	Class 8: 3
Class 9: 1404	Class 10: 34
Class 11/14: 471	

The data show that most one-class nouns are found in class 6 (which includes liquids). Class 4 has 21 and class 10 has 34 one-class nouns, while only three are found in class 8. In class 6 slightly more than 10% of nouns are one-class nouns – in all other classes, one-class nouns constitute only a small sub-group. While thus one-class nouns are found in all classes, overall, they are exceptional: typically Swahili nouns have forms in both the singular and the corresponding plural class.

Another important question is whether the relation between singular and plural forms constitutes in fact a regular singular-plural relation. Schadeberg (2001: 13) notes a number of examples where the relation is "irregular", reminiscent of a lexical-derivational relationship, rather than the encoding of inflectional number:

(11)	a. *uvumbi*	'dust, grain of dust' (cl. 11)	*vumbi*	'dust' (class 10)
	b. *ukuni*	'a piece of firewood' (cl. 11)	*kuni*	'firewood' (class 10)
	c. *usoka*	'brass wire' (general, small piece) (cl. 11)	*masoka*	'thick brass/iron wire' (6)
	d. *pesa*	'pice, money' (class 9)	*mapesa*	'small change' (class 6)
	e. *simba*	'lion' (class 9)	*simba*	'lions' (class 10)
			masimba	'pride of lions' (class 6)
	f. *moshi*	'smoke' (class 3)	*mioshi*	'plumes of smoke' (class 4)

Some of the examples illustrate the semantic tendencies of class shift identified above, e.g. the use of class 6 as augmentative (11c) or collective (11d,e). Examples (11a,b) show that class 11 may function as singularising class for mass nouns in class 10. The data show the complex interaction between mass nouns, class shift and singular-plural pairings in Swahili.

A final set of examples relates to nominal derivation, which involves derivational, noun-deriving suffixes and the assignment of the derived noun to a particular class. Suffixes include agentive *-i* and *-ji*, de-adjectival *-u*, passive *-e*, and *-o* and *-a* which derive states, results, actions or instruments. However, the exact meaning of the derived noun is a function of both the meaning of the suffix and the class to which the noun is assigned. For example, the de-verbal passive suffix *-e* can derive persons ((12a), class 1/2), artefacts ((12b), class 9/10), small things ((12c), class 7/8) and functions and role ((12d), class 3/4), depending on the noun class:

(12)	a. *-shinda* 'conquer, win'	> *mshinde*	'vanquished one, loser' (class 1/2)
	b. *-peta* 'bend'	> *pete*	'ring' (class 9/10)
	c. *-tona* 'fall in drops'	> *kitone*	'small drop of liquid' (class 7/8)
	d. *-tuma* 'send'	> *mtume*	'prophet, apostle' (class 3/4)

Data from class shift, one-class nouns and noun derivation thus provide further evidence for the semantic base of noun classes. Like in the previous section, the data seem to support the conclusion that some semantic regularity can be observed in the derivational use of noun classes, although here, as well, the relevant semantic criteria are not strict enough to permit precise predictions about the function of classes in derivation. With respect to number and class, the data show that plurality does play a role in the system, but it remains unclear whether this reflects inflectional or derivational processes. Further potential evidence for this question comes from two syntactic contexts: the interaction of noun class and conjunction, and pronominalized possessors, discussed in the following section.

2.5. Class and number in syntactic contexts

The interaction of noun class agreement and NP-conjunction has frequently been discussed in the context of analyses of agreement (e.g. Bokamba 1985, Corbett 1983, 1991, Marten 2000, 2005). In Swahili, agreement with conjoined NPs, especially of non-animate nouns, is often avoided, but when it is possible, there is a basic difference between animate and non-animate nouns. A conjunction of animate nouns normally requires class 2 agreement (except in cases of syntactically licensed partial agreement) (13). A conjunction of non-animate nouns can trigger a wider range of agreement, either with the closest conjunct or a default agreement marker, typically of class 8, but also of class 6, 10 or 11, irrespective of whether the conjoined nouns are from different classes (14a) or from the same class (14b):

(13) mw-alimu na mw-anafunzi w-ake wa-li-kuja
1-teacher and 1-student 1-her SM2-PAST-come
'The teacher and her student came' (Marten 2000: 80)

(14) a. ki-su na m-kono w-ake vi-me-loa damu
7-knife and 3-hand 3-his SM8-PERF-be.soaked blood
'The knife and his hand were soaked in blood' (Marten 2000: 81)

b. mi-saada na mi-kopo vi-ta-hatarisha uhuru wetu
4-help and 4-loans SM8-FUT-endanger independence our
'Gifts and loans will endanger our independence' (Schadeberg 1992: 22)

Agreement with conjoined nouns shows that the singular-plural pairing is not fully maintained with conjoined NPs – otherwise, on the evidence of (14b) for example, class 8 would have to be analysed as a plural class of class 4, which is itself a plural class. On the other hand, default agreement is typically with plural classes, rather than singular classes (or in the case of class 11 with an abstract class), so that a notion of plurality as a feature of these classes could be postulated.

A second syntactic context relevant to the question of the role of number in the noun class system is the pronominalization of possessors (cf. Schadeberg 2001: 12). As with conjunction, in Swahili there is a basic split between animate and non-animate nouns with respect to pronominal possessors. With animate nouns, the possessor is expressed according to number (or class 1 or 2): with a singular possessive stem *-ake* 'his/hers' or a plural stem *-ao* 'theirs' (in addition

there are also stems for 1st and 2nd person possessors). The possessive pronoun is built with an agreement prefix, agreeing in class with the possessee, plus the possessive stem.

(15) Mw-alimu na ki-kombe ch-ake
1-teacher and 7-cup 7-POSS1
'the teacher and his/her cup'

(16) Wa-limu na vi-kombe vy-ao
2-teacher and 8-cup 8-POSS2
'the teachers and their cups'

In contrast, with non-animate nouns there is only one, invariant form for all classes (including singular and plural), which is identical to the class 1 form *-ake* (glossed here as PossNonAnim, 'non-animate possessor'):

(17) Ki-kombe na ki-sahani ch-ake
7-cup and 7-saucer 7-POSSNONANIM
'a cup and its saucer'

(18) Vi-kombe na vi-sahani vy-ake
8-cup and 8-saucer 8-POSSNONANIM
'cups and their saucers' (Ashton 1947: 57)

(19) Mi-ti na mi-tawi y-ake
4-trees and 4-leaves 4-POSSNONANIM
'trees and their leaves' (Schadeberg 1992: 20)

The examples show that while number is important for animate nouns, no number (or indeed class) distinctions are expressed with non-animate pronominal possessors. The Swahili situation differs in this respect from many other Bantu languages where both pronominal possessor and pronominal possessee show class agreement, as for example in Herero (Möhlig et al. 2002: 59-60, Möhlig & Kavari 2008: 134-6):

(20) òmù-tí n-òví-yàò vyá-wó
3-tree and-8-leaves 8-Poss3
'the tree and its leaves'

The evidence from pronominal possessors and from agreement with conjoined NPs shows that in Swahili, number distinctions are consistently expressed in class 1/2, that is, with animate nouns.

On the other hand, number plays a far less important role for non-animate nouns, since agreement with conjoined nouns and possessor nouns appears to be computed without reference to number – or indeed to formal class features. In contrast, in Herero, agreement with possessors is strictly defined by class. The cumulative evidence with respect to the semantics of noun class, and the role of grammatical and/or semantic plurality distinctions, thus presents an ambiguous picture. On the one hand, noun class assignment and derivational uses of classes show some sensitivity to semantic distinctions without, however, being fully determined by semantic criteria. On the other hand, number appears to be important for animate nouns, but less so for non-animates. We will address these questions from a comparative perspective further below, but we will first discuss two types of analysis of the morphosyntax of noun classes in the following section.

2.6. Approaches to the morphosyntactic analysis of Bantu noun classes

Given the ambiguous nature of the evidence from the preceding sections, it is not surprising that different approaches to the analysis of Bantu noun classes have been proposed. One position, developed for example in Carstens (1991, 1993, 2008) assumes that class (or gender) does not play any role in Swahili grammar and is only relevant as a lexical feature of noun stems. On the other hand, Kihm (2005)[6] and Ferrari-Bridgers (2008) assume that noun classes do carry semantic information, and that noun class prefixes are syntactically active (nominal) heads. We will focus on these two approaches in what follows, although other approaches to the morphosyntactic representation of noun classes (e.g. Zamparelli 2008 or Harbour 2008), might well be adopted to address some of the issues discussed here.

Carstens (1991, 1993, 2008) develops an analysis of Swahili noun classes as a system of genders, with each gender – except for one – having singular and plural members corresponding to the traditional noun classes.

(21) Bantu genders (Carstens 2008: 136)

Gender A: stems of Classes 1/2
Gender B: stems of Classes 3/4
Gender C: stems of Classes 5/6
Gender D: stems of Classes 7/8
Gender E: stems of Classes 9/10
Gender F: stems of Classes 11/10
Gender G: stems of Class 14

Gender G reflects the historical distinction between two classes which have merged in Swahili (Proto-Bantu class 11 **lù-* > Swahili *u-*; Proto-Bantu class 14 **βù-* > Swahili *u-*). Class 14 consists of abstract nouns which do not have a plural.

In Carstens' analysis, gender is an arbitrary lexical quality of noun stems, and not related to classes. Noun class prefixes are analysed as gender-specific spell-outs of number features, and not as syntactic heads:

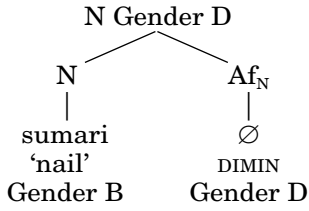
(22) Examples of spell out rules for Swahili class prefixes (Carstens 2008: 136)

[Singular]	*/ki-/*	/_ N	Gender D
[Plural]	*/vi-/*	/_ N	Gender D

The analysis of class/gender as arbitrary explains why class membership cannot be predicted on semantic grounds. However, it does not address cases of semantic regularity in class shift and the derivational use of classes. This, in Carstens' analysis, does not result from any inherent meaning of classes, but is a regular process of word-formation, involving derivational suffixes, which are, however, devoid of phonetic material and thus silent. These phonologically empty derivational suffixes have an inherent gender-specification, and are allocated to different classes to fulfil requirements of identification. For example, the nominal stem *-sumari* 'nail' is of Gender B, but can be combined, in Carstens' analysis, with silent derivational diminutive or augmentative suffixes (23). Since the diminutive suffix in (24), for example, is of Gender D, the noun will be in Gender D, spelled out as class 7/8.

(23) a. *m-sumari* / *mi-sumari* 'nail/s' (class 3/4) (Gender B)
b. *ki-sumari* / *vi-sumari* 'little nail/s' (class 7/8) (Gender D)
c. *sumari* / *ma-sumari* 'big (ugly/nasty) nail/s' (class 5/6) (Gender C)

(24)



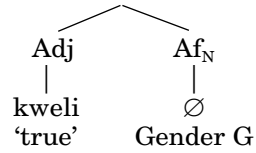
The use of a particular class for diminutives or augmentatives is arbitrary, comparable to, for example, German diminutives (e.g. *-lein*,

-chen) or abstraction suffixes (such as *-heit*, *-schaft*) which derive neuter and feminine nouns respectively, without any particular semantic motivation (cf. Zamparelli 2008: 181). Nouns without singular-plural pairing, such as abstract nouns of class 14 as in (30) and (31) and infinitives of class 15 in (32) and (33) – for which a new Gender H is introduced – can be analysed in a similar manner in Carstens' analysis:

(25) a. kweli
true
'true'

b. u-kweli
14-true
'truth'

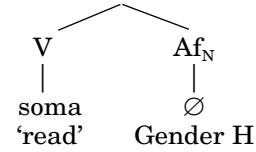
c. N Gender G



(26) a. soma
read
'read (verb)'

b. ku-soma
15-read
'reading/to read (noun)'

c. N Gender H



The comparatively high number of zero suffixes in Swahili (as opposed to, for example, Romance) is related to the gender/noun class system, as the presence of several genders can identify unambiguously a greater number of null elements.

A potential problem for this approach are overt derivational suffixes, as briefly discussed above. For example, from the verb *-angalia* 'look at' the nominal stem *-angalifu* can be derived by the addition of the noun-deriving suffix *-u* (which triggers the phonological insertion of /f/):

(27) a. *-angalia* 'look at, observe'
b. **angalifu* 'carefulness, attention'
c. *uangalifu* 'carefulness, attention' (class 11)

However, since in this case the suffix is overt, it should not require identification through assignment to a noun class. Yet, as (27b) shows, the derived noun – like all nouns – has to be assigned to a noun class and cannot be used without such assignment. Furthermore, when assigned to class 11 (or class 14 in Carstens' approach), the meaning is predictably an abstract quality. But this meaning results in Carstens' analysis from zero-suffixation, and

not from the semantics of the noun class. In the case of (27c), then, a second derivational process of suffixation with a zero suffix has to be postulated, without substantive empirical support. Furthermore, since Carstens' analysis assumes that it is the relevant derivational suffixes which determine gender/class assignment, the possibility to use nouns with an overt derivational suffix in different classes, as for example in (28), is surprising:

(28) a. *-soma* 'read'
b. *somo / masomo* 'lesson' (class 5/6)
c. *msomo / misomo* 'act/means/method of reading' (class 3/4)

A possible solution would be to assume that there are several suffixes *-o*, each with its own gender specification, but the postulation of such an ambiguity for all derivational suffixes would appear to be poorly motivated.

In sum, Carstens (1991, 1993, 2008) proposes a consistent analysis of Swahili noun classes as gender system, with gender features analysed as part of the lexical information of noun stems, without any role in the morphosyntax of nouns. The analysis is consistent with views which regard noun classes as semantically arbitrary, and provides an alternative explanation of regular derivational uses of noun classes in terms of phonologically empty derivational suffixes. However, it is not quite clear how the analysis can be extended to nouns formed with overt derivational suffixes.

An alternative approach to the morphosyntax of noun classes is provided by the analysis of Manjaku (Niger-Congo) in Kihm (2005) and of Luganda in Ferrari-Bridgers (2008). Kihm (2005) proposes that the lexicon consists of un-categorized roots which are combined with a functional (categorical) head for their projection to syntax, and that class/gender is the nominal exponent of a functional head *n*. Furthermore, while gender, for example in Romance, has no lexical-semantic content, class in Niger-Congo does have lexical-semantic content, and the difference in content is related to the position of the noun class/gender marker and the fusion between the root and these markers: noun class markers are typically distinctly identifiable prefixes, while gender markers are fused suffixes. To the extent that class/gender has semantic content, this reflects, according to Kihm (2005: 475) the fact that "it belongs to the inherent properties of nouns that they denote entities that can (perhaps must) be allotted to different classes of *things*, by virtue of innate and culturally informed cognitive processes, diversely expressed in languages".

Kihm's (2005) proposal is based on derivational relations between nouns, such as those involving the noun stem *-konj* 'finger' in (29), which is taken by Kihm as evidence for analysing "singular-plural" relations in Manjaku as derivational, similar to the proposal for Swahili by Schadeberg (2001) discussed above:

(29) a. *pë-konj* 'finger' (class 9)
b. *kë-konj* 'fingers (discrete number of fingers not exceeding ten)' (class 10)
c. *i-konj* 'fingers (unknown/ indefinite, generic)' (class 8)

Similar evidence from class shift in Luganda is presented by Ferrari-Bridgers (2008: 46) to support her analysis of noun classes in Luganda as derivational:

(30) a. *b-beere* 'breast' (class 5) > *ki-beere* 'udder' (class 7)
b. *ki-batu* 'palm of the hand' (class 7) > *lu-batu* 'handful' (class 11)
c. *bu-ligo* 'filth' (class 14) > *mu-ligo* 'dirty person' (class 1)
d. *ndiga* 'sheep' (class 9) > *ka-liga* 'lamb' (class 12)

However, the structure of Manjaku nouns differs from Swahili nouns, since Manjaku has verbal lexical roots from which nouns are formed by using class prefixes:

(31) a. *lik* 'to draw water' > (*pë-*)*lik* 'drawing water/to draw water'
b. > *pë-lik* / *i-lik* 'well(s)' (class 9/8)
c. > *m-lik* 'water' (class 6)
d. > *ka-lik* 'fruit juice' (class 7)

While, as shown above, class shift is common in Swahili, the derivation of nouns from verbal roots typically involves derivational suffixes, and so it has to be kept in mind that Kihm's analysis of Manjaku may not be entirely applicable to Bantu.

As Kihm (2005) points out, his analysis has implications for the comparison between Manjaku noun classes and Spanish gender. According to him, both Manjaku and Romance share class/gender, but class markers in Manjaku have more lexical content than gender markers in Romance, and this difference results in different positions of class and gender markers. In Manjaku, both the stem and the noun class markers are lexical roots. The class root provides a nominal functional feature and functions as head of the nominal construction, taking the stem as the complement:

(32) Manjaku *ka-lik* 'fruit juice' (Kihm 2005: 21)

$[_{nP}$ √ka $[_{√P}$ √lik]]

In contrast, in Spanish, even though class and gender features occupy the same head position, no lexical content is associated with gender, and so the lexical stem moves from complement to head position and is adjoined to the gender marker, resulting in a fused suffix as opposed to an analytic prefix as in Manjaku example in (32):

(33) Spanish *gat-a* 'cat' (Kihm 2005: 20-21)

a. $[_{nP}$ nII $[_{√P}$ √gat]]
b. $[_{nP}$ √gat+nII $[_{√P}$ t]]

While thus class and gender are similar in terms of morpho-syntactic structure, they differ in lexical representation of the case/gender markers, and this difference accounts for the differences in surface manifestation.

A final aspect of Kihm's (2005) analysis addresses the question of plurality and number. In this respect, too, Manjaku and Romance (and English) are analysed differently, as a consequence of the different lexical content of gender and class markers. In Romance plurality is associated with denoting nominal roots, projected on a syntactic Number Phrase. In contrast, in Manjaku, plurality is associated with class roots (or "protonouns", elements with a category feature *n*), and the relation between sets of roots {n} and plural roots {n ∩ Plurality} is an independent, although significantly predictable, relation holding in the lexicon, not derived in the syntax. Singular-plural relations in a class language like Manjaku are thus lexical, derivational relations, but are inflectional in a gender language like Spanish (but see sections 3.3 and 3.4 below).

The analysis of Luganda and Italian nouns by Ferrari-Bridgers (2008) comes to similar conclusions as Kihm's (2005) study with respect to class. Ferrari-Bridgers (2008) proposes that class corresponds to a nominalizing head *n*, combining with a lexical root along similar lines to Kihm (2005). However, Ferrari-Bridgers assumes that Italian gender, too, corresponds to *n*, and that, thus, class and gender are the same. The analysis is thus similar to Carstens's (2008) analysis in that class and gender are said to be identical, but, according to Ferrari-Bridgers, and contra Cartsens, this is not because class/gender are lexical qualities of stems, and number is the only relevant grammatical feature, but rather because class/gender is the relevant grammatical feature in both Luganda and Italian.

2.7. Conclusions

Bantu noun class systems provide a rich empirical base for the study of nominal classification. However, aspects of Bantu noun classes present a challenge for analysis, as we have demonstrated with reference to Swahili. The question of whether Bantu classes are arbitrary or semantically motivated is probably best answered by a compromise: classes are to some extent semantically motivated, but not semantically determined – although animacy plays a more prominent role than other factors.[7] Accordingly, the question of whether class and/or gender play a role in the syntactic representation of Bantu nouns has prompted different analyses, some assuming that class/gender does not play any role beyond the lexicon, others proposing that class prefixes themselves have semantic content and function as syntactic nominal heads.

3. A Romance – Bantu comparison

In this section, we make a systematic comparison between Bantu and Romance. At first sight, Romance looks quite different from Bantu, with just two genders as opposed to the 10 to 20 classes in Bantu. However, especially when we investigate the notions of inflection and derivation, and elements in both languages associated with these notions, Bantu and Romance turn out to be quite alike, for many differences are quantitative rather than qualitative. The conclusion that will be drawn is that 'inflectional' and 'derivational' might be handy descriptive terms, but they cannot be converted into discrete analytical tools.

Before doing so, we give an outline of the main characteristic of gender in Romance, with Italian taken as a representative.

3.1. Gender in Romance: an overview

In Romance there are two genders, masculine[8] and feminine, and two numbers. Similar to Bantu noun classes, gender, together with number, participates in various agreement relations inside and outside the DP. Inside the DP, gender agreement is realized with adjectives and most determiners and quantifiers (cf.(34a)). Outside the DP, subjects agree in gender with adjectival predicates (cf. (34b)) and, in unaccusative and passive constructions, with the past participle (when present, cf. (34c)); object agreement is realized only when the object is a clitic pronoun that appears on the left of a past participle (cf. (34e) as opposed to (34d)):

(34) a. Ho rotto molt-e brocch-e ross-e
have-PRES.1SG broken-M.SG many-F.PL jugs-F.PL red-F.PL
'I broke many red jugs'

b. Quest-e brocch-e sono ross-e
these-F.PL jugs-F.PL are-pres.3PL red-F.PL
'These jugs are red'

c. Le brocch-e sono cadut-e
the-F.PL jugs-F.PL are-PRES.3PL fallen-F.PL
'The jugs have fallen down'

d. Ho comprat-o le brocch-e
have-PRES.1SG bought-M.SG the-F.PL jugs-F.PL
'I bought the jugs'

e. (A proposito delle brocche) le ho comprat-e
(as for the jugs) them-F.PL have-PRES.1SG bought-F.PL
'(As for the jugs) I bought them'

Italian nouns are also organized in three major declension classes plus a series of minor ones,[9] a legacy of the richer system of Latin. The three main declension classes are:

- *o*/*i*: first declension class, which comprises mostly masculine nouns – ex. *tett-o* M.SG, *tett-i* M.PL 'roof(s)';
- *a*/*e*: the second declension class comprises mostly feminine nouns – ex. *brocc-a* F.SG, *brocch-e* F.PL 'jug(s)';
- *e*/*i*: the third declension class comprises both masculine and feminine nouns – ex. *pont-e* M.SG, *pont-i* M.PL 'bridge(s)', *bott-e* F.SG, *bott-i* F.PL 'barrel(s)'.

Adjectives are also organized in declension classes, one analogous to the first/second nominal declension class (singular *-o* or *-a*, plural *-i* or *-e* – ex. *ross-o* M.SG, *ross-i* M.PL, *ross-a* F.SG, *ross-e* F.PL 'red'), the other analogous to the third nominal declension class (singular *-e*, plural *-i* – ex. *verd-e* M/F.SG, *verd-i* M/F.PL).[10] Note that Noun-Adjective agreement, inside and outside the DP, is done by gender and not by declension, therefore declension mismatch is very common: *il tett-o*/*pont- e ross-o*/*verd-e* 'the red/green roof/bridge', *la brocc-a*/*bott-e ross-a*/*verd-e* 'the red/green jug/barrel'.

Whenever a target category displays gender agreement with a noun, it also displays number agreement, but the reverse is not true: inflected verbs agree in number but not in gender with their subjects:

(35) a. Il gatt-o / la gatt-a miagol-a
the-M.SG tomcat-M.SG the-F.SG she-cat-F.SG meows-PRES.3SG

b. I	gatt-i	/	le	gatt-e	miagol-ano
the-M.PL	tomcats-M.PL		the-F.PL	she-cats-F.PL	meow-PRES.3PL

This characteristic is very important, because provides strong empirical evidence that in Italian (as is the case in Romance in general) number is recognized as a distinct feature from gender. Similar evidence does not exist for Bantu, where only the cases of coordination and pronominal possessors discussed in section 2.5 may constitute some evidence, albeit ambiguous, in this sense.

Grammatical gender is unpredictable and uninterpretable with inanimate nouns, while it tends to match natural gender with animate[11] nouns:

(36) a. *sedia*-F.SG 'chair' vs. *sedile*-M.SG 'seat'
b. *bambina*-F.SG 'girl' vs. *bambino*-M.SG 'boy'

The present situation in Romance languages results from the simplification of the Latin system, which also had a neuter gender. However, it is likely that the typical IE tripartite distinction is an innovation of late Proto-IE, and that in older stages there were only two classes, animate vs. inanimate: the former class was then split into masculine and feminine, the latter remained as neuter (see Meillet 1931). There are three main arguments supporting this hypothesis: first, the fact that Anatolic languages (e.g. Hittite), which are commonly assumed to be the branch that first separated from the common language, only have the animate-inanimate distinction. Second, in all the IE languages that have a masc./fem. grammatical distinction, this device is normally not used to mark sex distinctions in cases where it really matters, and different roots are used instead: so, though we find Lat. *lupus* 'wolf' vs. *lupa* 'she-wolf', we have series like IE *wak- Lat. *vacca* 'cow', IE *tauro- Lat. *taurus* 'bull', IE *g^{w}ou- Lat. *bos*, *bovis* 'bull or cow indifferently'. Third, the neuter inflection is formally distinct from the masc./fem. inflection in two respects, namely that NOM-ACC-VOC are always identical, both in the singular and in the plural, and that NOM-ACC-VOC plural is always *-a*.

If this hypothesis is correct, the present unpredictability of gender in Romance would be the result of a series of re-organizations of the noun classification system, where the initial situation was a transparent animate/inanimate opposition. It must be said, however, that already in Latin masculine and feminine, rather than neuter, were widely used for inanimate nouns.

3.2. Inflection or derivation?

As discussed in section 2.6, there is no general consensus as to the (non-)affinity between Bantu noun classes and Romance gender/number. This is due to the nature of the observable facts, for there are obvious similarities, but also non-trivial differences. In fact, while both in Romance and in Bantu nouns are always associated with a gender/class marker, Bantu classes are more numerous than Romance genders and in some cases they are clearly derivational, and correspond to derivational affixes in Romance (e.g. augmentative, pejorative, locative etc., see 2.4). Also, while in both groups of languages pervasive agreement involving noun class/gender is observed, both inside and outside the DP, agreement patterns may be different. So, for example, object agreement is limited to constructions like (34e) in Romance while is systematic in Bantu; also, in all Romance languages inflected verbs only agree in number and not in gender with the subject, a separation of the two features that is not attested in Bantu languages, apart from the much less straightforward constructions discussed in section 2.5.

It has been suggested (see especially Schadeberg 2001 and Kihm 2005, with certain differences between them) that the formation of Singular-Plural pairings in noun class languages is a derivational process that cannot be told apart from noun classes themselves; this would make noun class languages different from gender languages, where number is to be analyzed as inflectional morphology. This view will be discussed in detail is sections 3.3-3.5.

Before turning to the empirical data, it should be pointed out that the inflectional/derivational distinction is by no means a clear-cut uncontroversial one. Plank (1994) proposes 28 criteria to distinguish between inflectional and derivational morphology, Beard (1998) proposes 4, Stump (1998) proposes 5: some generalizations can be drawn, in particular that inflection is more regular and productive than derivation, and that it is more relevant to syntax than to the lexicon. However, all the authors stress the fact that it is often the case that a given morpheme cannot be safely classified as inflectional or derivational, because different criteria might put it in different classes; so, for example, English plural is classified as derivational by 6 out of 28 of Plank's criteria. Plank (1998: 1677) also makes the point that the distinction inflectional vs. derivational is not a discrete one: "Insofar as many of these distinctions are gradual rather than categorical, it is an oversimplification to provide only two values, (a [=inflectional]) and (b [=derivational])".

3.3. "Irregularities" in plural formation in Romance

As discussed in section 2.4, Schadeberg (2001) proposes that in Swahili the relation between singular and plural forms is a derivational one, for, according to some of Plank's criteria, it has the properties typical of derivation. In this section, we test how these criteria score in Romance.

First are Plank's (1994) criterion 18, which looks at whether "the semantic contribution of the morphological category is uniform for all bases (=more inflectional) or diverse (=more derivational)", and criterion 19, which looks at whether "the semantic relationship between Complex Non-Compound Words (CNCW) and their bases is transparent for all occurrences of the morphological category (=more inflectional) or at least occasionally opaque (=more derivational)". Schadeberg provides the examples in (11) above to show that "so-called number distinctions [in Swahili] can represent other meanings than singular/plural with inanimate nouns" (Schadeberg 2001: 13), and, according to Plank's criteria 18 and 19, they pattern with derivational morphology. Similar examples can be found in Romance:[12]

(37)	It.	*fum-o* M.SG 'smoke'	*fum-i* M.PL 'intoxication (e.g. from alcohol)'
		acqu-a F.SG 'water'	*acqu-e* F.PL 'amniotic fluid'
		ari-a F.SG 'air'	*ari-e* F.PL 'snobbish attitude' (or musical arias)
	Sp.	*víspera* F.SG 'day before'	*vísperas* F.PL 'evensong'
		agua F.SG 'water'	*aguas* F.PL 'amniotic fluid'
		deber M.SG 'duty'	*deberes* M.PL 'school assignments'
	Cat.	*aire* F.SG 'air'	*aires* F.PL 'snobbish attitude'
		aigua F.SG 'water'	*aigües* F.PL 'amniotic fluid'
		ullera F.SG 'telescope'	*ulleres* F.PL 'glasses'
		fum M.SG 'smoke'	*fums* M.PL 'snobbish attitude'
	Pt.	*água* F.SG 'water'	*águas* F.PL 'amniotic fluid'
		ar F.SG 'air'	*ares* F.PL 'snobbish attitude'
		cobre M.SG 'copper'	*cobres* M.PL 'coins'
		papel F.SG 'paper'	*papéis* F.PL 'documents' (also 'pieces of paper')
	Fr.	*air* F.SG 'air'	*air-s* M.PL 'snobbish attitude'
		devoir M.SG 'duty'	*devoirs* M.PL 'school assignments'
		course F.SG 'race'	*courses* F.PL 'shopping' (also 'races')
	Rom.	*aer* M.SG 'air'	*aere* F.PL 'snobbish attitude'
		apă M.SG 'water'	*ape* F.PL a metaphor for strong emotions
		fumo M.SG 'smoke'	*fumûri* F.PL 'snobbish attitude'

Schadeberg also makes use of Plank's criterion 20 "The applicability of the morphological category to bases of particular word-classes is unlimited (=more inflectional) or limited in one way or another (=more derivational)", and criterion 21 "There are no (=more inflectional) or some (=more derivational) CNCWs expressing the morphological category whose base is only attested in those CNCWs themselves". According to Schadeberg, "infinitives (class 15), locatives (classes 16 to 18) and different kinds of mass nouns (classes 6 and 11) resist plural formation (assuming that these are singular nouns). [...] Unpaired ("one-number") inanimate nouns occur in all classes except classes 1/2" (p. 13-14). The examples provided are in (9) above.

However, nouns without sg/pl alternations also exist in Romance. Mass nouns admit plural forms to the extent that they admit a taxonomic reading, otherwise the result of pluralisation is odd at best; also, various nouns, often the same across Romance, are only found in the plural:

(38)		"Singular-only"[13] nouns	"Plural-only" nouns
	IT.	*fame* F.SG 'hunger'	*averi* M.PL 'possessions'
		sete F.SG 'thirst'	*% spoglie* F.PL 'spoil'[14]
		ossigeno M.SG 'oxygen'	*% compere* F.PL 'shopping'
		salute F.SG 'health'	*interiora* F.PL 'entrails'
			esequie F.PL 'funeral'
			paturnie F.PL 'blues'
	SP.	*sed* F.SG 'thirst'	*anales* M.PL 'annals'
		cariz M.SG 'look'	*tijeras* F.PL 'scissors'
		oxígeno M.SG 'oxygen'	*comestibles* M.PL 'food'
		salud F.SG 'health'	*exequias* F.PL 'funeral'
	CAT.	*set* F.SG 'thirst'	*afores* M/F.PL 'outskirts of a town'
		fam F.SG 'hunger'	*postres* F/M.PL 'dessert'
		oxigen M.SG 'oxygen'	*tovalles, estovalles* F.PL 'tablecloth'
		salut F.SG 'health'	*queviures* M.PL 'food supply'
			golfes F.PL 'attic of a house'
	PT.	*oxigénio* M.SG 'oxygen'	*férias* F.PL 'vacation'
		cais M.SG 'pier'	*óculos* F.PL 'glasses'
		fé F.SG 'faith'	*compras* F.PL 'shopping'
			núpcias F.PL 'nuptials'
			trevas F.PL 'night darkness'
	FR.	*% faim* F.SG 'hunger'	*% alentours* M.PL 'neighborhood'
		% soif F.SG 'thirst'	*% ciseaux* M.PL 'scissors'
		oxygène M.SG 'oxygen'	*ténèbres* F.PL 'night darkness'

	% santé F.SG 'health'	*vivres* M.PL 'provisions'
		obsèques F.PL 'funeral'
ROM.	*foame* F.SG 'hunger'	*zori* M.PL 'dawn'
	sete F.SG 'thirst'	*tărâțe* F.PL 'bran'
	sănătate F.SG 'health'	*plate* F.PL 'hair'
	faimă F.SG 'fame'	*cumpărături* F.PL 'shopping

A further argument is based on Plank's criterion 4 "The specification of the morphological category is grammatically obligatory (=more inflectional) or non obligatory (=more derivational) for the relevant CNCWs of the relevant word-class". In this respect, Schadeberg mentions the fact that Swahili infinitives "are unmarked for number nor can their number be inferred from agreement" (p. 13).

Though Italian infinitives used as nouns are different from Swahili ones, for the latter are part of the noun class system (cf. (26)) while the former do not enter agreement relations, they are similar for they are unmarked for number as well (displaying default 3rd person masculine singular agreement) and cannot be pluralized even when it is obvious that the infinitive denotes a series of repeated occurrences:

(39) a. Il suo continuo chiedere prestiti ai vicini
the-M.SG his-M.SG continuous-M.SG asking loans at-the neighbors
lo ha messo nei guai
him has-PRES.3SG put in-the troubles

b. * I suoi continui chiedere/**i prestiti ai vicini
the-M.PL his-M.PL continuous-M.PL asking loans at-the neighbors
lo hanno messo nei guai
him have-PRES.3PL put in-the troubles
'His continuous asking for loans from the neighbors caused him troubles'

The other Romance languages pattern alike, as shown by the following Catalan and Portuguese examples:

(40) a. Després de tant demanar el que volia,
after of so-much asking-INF the that wanted-3SG
no ho va poguer obtenir.
not it-CL go-PAST.3SG can-INF obtain-INF

b. * Després de tants demanars el que volia, ...
after of so-many asking-INF.PL the that wanted-3SG
'After so much asking for what s/he wanted, s/he couldn't obtain it.'

(41) a. O cair facilmente impede-o de sair sozinho.
the fall-down-INFIN easily prevents-him from going-out alone
b. * Os caíres contínuos impedem-no de sair sozinho.
the fall-down-INFIN easily prevents-him from going-out alone
'His falling easily prevents him from going out alone'

Another argument used by Schadeberg to show the "number insensitivity" of nominal classification in Swahili is based on quantifiers. He notes that the word *-moja* 'one' can be combined with any noun class, also 'plural' ones. However, the same can be said for Spanish *un* 'one' (e.g. *un caballo* 'a horse'), which is found also with plural nouns as an indefinite existential quantifier: *unos libros* 'some books'. Analogously, he observes that "*-ingi* 'many, much' and *-ote* 'all, whole' are not restricted to a subset of classes with either singular or plural meaning" (p. 12). Again, the argument is very naturally extended to Romance, for the quantifiers corresponding to *-ingi* and *-ote* are equally unselective with respect to number:

(42)

	SINGULAR COUNT	MASS	PLURAL
IT.	tutto il mondo	tutto il vino	tutti gli studenti
		molto vino	molti studenti
SP.	todo el mundo	todo el vino	todos los estudiantes
		mucho vino	muchos estudiantes
CAT.	tot el món	tot el vi	tots els estudiants
		molt vi	molts estudiants
PT.	todo o mundo	todo o vinho	todos os estudantes
		muito vinho	muitos estudantes
	'all the world'	'all the wine'	'all the students'
		'much wine'	'many students'

In sum, the empirical arguments used by Schadeberg to argue that number is derivational in Swahili can be replicated to some extent in Romance. In what follows, some derivational properties of gender in Romance will be discussed.

3.4. "Lexical plurals" (Acquaviva 2008)

The first argument presented by Schadeberg (2001) (his "major objection against the gender-plus-number analysis", p. 10) is the fact that in some cases two prefix series that are distinct in the singular (*JI*, *ji*, *li* 5 and *U*, *mu*, *u* 11) are paired with one prefix series in the plural (*MA*, *ma*, *ya* 6), or the other way round (sg. *U*, *mu*, *u* 11 paired

with pl. *MA*, *ma*, *ya* 6 and *N*, *N*, *zi* 10). In his view, having the same series of singular or plural prefixes for two distinct genders is "inexplicable", and leads one to reject a gender-plus-number analysis for Swahili, favoring the "class-plus-pairing" approach.

However, note that situations analogous to the one described by Schadeberg for Swahili are not uncommon in languages with undisputed genders, though of course the number of morphemes involved in these "inexplicable" pairings is much lower, a trivial consequence of the fact that genders are normally two or at most three. So, for example, the Italian morpheme *-i* can be used to form the plural of three different singular forms, masculine nouns that have *-o* as the singular ending, but also masculine and feminine nouns that have *-e* as the singular ending:

(43) *cont-o* M.SG 'addition' *cont-i* M.PL 'additions'
cont-e M.SG 'count' *cont-i* M.PL 'counts'
font-e F.SG 'source' *font-i* F.PL 'sources'

Similar patterns may be a problem in a class-plus-pairing system, but they are easily captured in a system that treats gender and number as separate features, gender being also signaled by agreement patterns; thus, whatever the endings on the noun, masculine nouns are preceded by the definite article *il* in the singular and have *i* in the plural, while feminine nouns are preceded by *la* in the singular and *le* in the plural.

A notable exception to this generalization is represented by "lexical plurals" (Acquaviva 2008), namely masculine nouns that have a feminine plural form. Apart from Italian, this pattern is observed only in Romanian,[15] with the very robustly attested group of so-called 'neuter' nouns: *ou* M.SG 'egg' ~ *ouă* F.PL 'eggs', *strigăt* M.SG 'scream' ~ *strigăte* F.PL 'screams', *gând* M.SG 'thought' ~ *gânduri* F.PL 'thoughts', *ram* M.SG 'branch' ~ *ramuri* F.PL 'branches', etc. As convincingly argued by Acquaviva (2008: 135-140), notwithstanding the surface similarities, Italian lexical plurals and Romanian neuters should not be analyzed as the same phenomenon. For the purposes of the present discussion, only Italian lexical plurals are relevant:

(44) It. *uov-o* M.SG 'egg' *uov-a* F.PL 'eggs'
dit-o M.SG 'finger' *dit-a* F.PL 'fingers'
ris-o M.SG 'laughter' *ris-a* F.PL 'laughters'
grid-o M.SG 'scream' *grid-a* F.PL 'screams'
? *grid-i* M.PL

url-o M.SG 'scream'	*url-a* F.PL 'screams'
	url-i M.PL
bracci-o M.SG 'arm'	*bracci-a* F.PL 'arms (people)'
	bracc-i M.PL 'arms (of see, of cranes)'
gest-o M.SG 'gesture'	*gest-a* F.PL 'glorious deeds'
	gest-i M.PL 'gestures'
mur-o M.SG 'wall'	*mur-a* F.PL 'fortification walls' (of a city)
	mur-i M.PL 'walls'

Note that with some nouns the feminine plural is the only possible form, while other nouns admit both the masculine and the feminine plural. The latter usually means something more than just 'more than one', and some cases are strikingly similar to some of the examples discussed in section 2.4 as evidence for the derivational nature of noun classes. Recall in particular ex. (11e), the case of *simba* 'lion' (class 9), which has two plurals *simba* 'lions' (class 10) and *masimba* 'pride of lions' (class 6), the latter receiving a collective interpretation. This is exactly the case of *mur-o* M.SG 'wall', whose 'regular' plural *mur-i* M.PL simply means 'walls' while the gender-shifted *mur-a* F.PL is interpreted as the set of walls forming a fortification.

A similar effect is observed in the following example:

(45) Per tutto il pomeriggio ho sentito degliurli /delle urla
for all the afternoon have heard of-the screams-M.PL of-the screams-F.PL
a intervalli di mezz'ora
at intervals of half hour
'During the whole afternoon, I heard screams every half hour'

Unlike *urli* M.PL, *urla* F.PL necessarily has a group interpretation, thus only *urli* M.PL is compatible with a scenario of an isolated scream heard every half hour; *urla* F.PL forces a reading where several screams were heard every half hour.

Consider also the case of *dit-o* M.SG 'finger'. The immediate reaction to a possible masculine plural form *dit-i* M.PL 'fingers' is assured rejection. Note however that it is possible to construct examples in which *dit-i* is the only admissible form:

(46) Gli investigatori hanno trovato a casa dell'indiziato una scatola
the inquirers have found at home of-the suspect a box
contentente 12 ?diti/*dita anulari
containing 12 fingers-M.PL/fingers-F.PL ring
'At the suspect's home, the inquirers found a box containing 12 ring fingers'

This is because *dit-a* F.PL denotes the set of fingers of a hand (or subsets/supersets thereof), while the 12 ring fingers in example (46) are a sum of fingers which necessarily come from different individuals. This is strongly reminiscent of one of Kihm's (2005) arguments for the derivational nature of noun classes in Manjaku, presented in (29): *pë-konj* 'finger' has two plural forms, *kë-konj* or *i-konj*, depending on whether it's the "the normal number [of fingers] a human being is endowed with".

In short, if nouns with more than one singular/plural pairing and their interpretation are taken as evidence for the derivational nature of noun classes in Swahili and Manjaku, one must conclude that lexical plurals constitute comparable evidence for Italian, hence also gender/number in Italian should be analyzed as derivational, much in the spirit of Ferrari-Bridgers (2008).

3.5. Other derivational uses of gender in Romance

As discussed in section 2.4, apart from the singular/plural pairings, in noun class languages there are classes that are clearly derivational, such as infinitives, locatives but also augmentative, diminutives and the like. Notwithstanding the first impression that nothing of this kind exists in Romance, there are cases of clearly derivational uses of gender in all Romance languages. Before turning to the relevant data, a short description of noun-from-noun derivation in Romance will be presented, with Italian taken as the representative language.

In Italian, nouns are derived from other nouns by means of the addition of an affix, most commonly a suffix that is attached to the root before the gender/number ending. Nominal derivational suffixes can be stacked:

(47)	*libr-o* 'book'	*libr-ett-o* 'small gracious book'	*libr-ett-in-o* same
		libr-in-o same	
	scop-a 'broom'	*scop-ett-a* 'short broom'	*scop-ett-on-a* 'big short broom'
		scop-on-a 'big broom'	

These suffixes are specified for being inflected in one or the other declension class, and the resulting noun is inflected according to the declension imposed by the most external suffix. So, for example, an *e*/*i* masculine noun such as *cane* 'dog' moves to the *o*/*i* declension class when the morpheme *-acc-* 'ugly' is added. Conversely, an *o*/*i*

noun such as *gatto* 'cat' moves to the *e*/*i* class when the morpheme *-on-* 'big' is added:

(48) *gatt-o/i* 'cat(s)' *gatt-acci-o* 'ugly/nasty cat' *gatt-on-e* 'big cat'
can-e/i 'dog(s)' *cagn-acci-o* 'ugly/nasty dog(s)' *cagn-on-e* 'big dog'

Derivational morphology does not trigger agreement, which is still in gender/number, not in derivational class:

(49) *un gatt-o ner-o* *un gatt-accio ner-o* **un gatt-accio ner-accio*

In this, Italian is clearly different from Bantu, where augmentative, diminutive, pejorative prefixes enter agreement relations exactly like any other noun class prefix (see section 2.4, ex. (5)). We will see in section 3.5.2, however, that derivational morphology may interact with gender agreement also in Romance.

3.5.1. Derivation via gender shift in Romance

Though unproductive, derivation by means of gender shift is attested in all Romance languages:[16]

(50)	IT.	*mel-o* M.SG 'apple-tree'	*mel-a* F.SG 'apple'
		noc-e M.SG 'nut-tree'	*noc-e* F.SG 'nut'
		legn-o M.SG 'wood'	*legn-a* F.SG 'burning wood'
		buc-o M.SG 'hole'	*buc-a* F.SG 'big/man-made hole'
		frutt-o M.SG 'fruit (count)'	*frutt-a* F.SG 'fruit (mass)'
		cos-a F.SG 'thing'	*cos-o* M.SG 'mysterious, incongruous thing'
	SP.	*manzan-o* M.SG 'apple-tree'	*manzan-a* F.SG 'apple'
		naranj-o M.SG 'orange-tree'	*naranj-a* F.SG 'orange'
		mader-o M.SG 'piece of wood'	*mader-a* F.SG 'wood'
		barc-o M.SG 'ship'	*barc-a* F.SG 'boat'
	CAT.	*full* M.SG 'leaf (of paper)'	*full-a* F.SG 'leaf (of a tree)'
		fi M.SG 'goal'	*fi* F.SG 'end'
		roc M.SG 'stone, pebble'	*roc-a* F.SG 'rock'
		sac M.SG 'sack'	*sac-a* F.SG 'big sack'
		ram M.SG 'bunch of flowers, branch'	*ram-a* F.SG 'ensamble of branches of a tree'
		crit M.SG 'scream' (M.PL *crits*)	*crid-a* F.SG 'call, proclamation' (F.PL *crides*)
	PT.	*ram-o* M.SG 'branch (count)	*ram-a* F.SG 'branch (mass)'
		lenh-o M.SG 'wood, branch (count)'	*lenh-a* F.SG 'burning wood'

	frut-o M.SG 'fruit (count)'	*frut-a* F.SG 'fruit (mass)'
	barc-o M.SG 'boat'	*barc-a* F.SG 'a particular type of boat'
	janel-o M.SG 'small window'	*janel-a* F.SG 'window'
	ov-o M.SG 'egg'	*ov-a* F.SG 'roe'
	gest-o M.SG 'gesture'	*gest-a* F.SG 'heroic historical events'
FR.	*fossé* M.SG 'ditch'	*fosse* F.SG 'pit'/'grave'
	bras M.SG 'arm'	*% brasse* F.SG 'fathom' (measure of distance)
ROM.	*păr* M.SG 'pear tree'	*pară* F.SG 'pear'
	nuc M.SG 'nut tree'	*nuca* F.SG 'nut'
	fruct M.SG 'fruit'	*fructă* F.SG 'fruit' (colloquial)
	fapt M.SG 'fact'	*faptă* F.SG 'completed action'

Though in many cases it is difficult to define the semantic relations between the two members of the masculine/feminine pair, there are some regularities (on this, see in particular Ferrari 2005: 38-44). It is intriguing that very similar regularities are attested also in Swahili, so some of the pairs in (50) parallel those found in (4). Plant~fruit, IT *melo~mela*, *albicocco~albicocca;* SP: *manzano~manzana*, *naranjo~naranja;* ROM: *păr~pară*, *nuc~nuc*a; SW: *mchungwa* 3*~chungwa* 5 (see section 2.4, ex. (4f)). Small~big, CAT: *roc~roca*, *sac~saca*; PT: *saco~saca*, *janelo~janela*; SW: *kitoto* 7 ~ *mtoto* 1 (see section 2.4, ex. (4a)). Unit-collective: IT: *frutto~frutta*, *legno~legna*; SP: *madero~madera*; CAT: *ram~rama*; PT: *ramo~rama*, *lenho~lenha*, *fruto~fruta*, *ovo~ova*; SW: *rafiki* 9 ~ *marafiki* 6 (see section 2.4, ex. (4c)).

3.5.2. Gender-shifting derivational morphology in Romance

The augmentative suffix *-on-* may change the declension of the noun it attaches to (see (48)). Another property of *-on-* is that, when attached to a feminine noun, it may turn it into a masculine one. This gender shift is not obligatory: attaching the suffix *-on-* to a feminine noun usually results into a doublet, *NOUN-on-e* M.SG and *NOUN-on-a* F.SG. The meaning of gender-shifted *NOUN-on-e* is usually something more than just 'big N', and can be more or less lexicalized. These properties make gender-shifting *-on-* different from the diminutive suffix *-chen* in German or *-aki* in Greek, where the shift to neuter is obligatory and is only rarely accompanied by idiosyncratic semantic nuances.

The cognate of Italian *-on-* is found in the other Romance languages (with the exception of French), with comparable interpretive and gender-shifting properties:

(51)	It.	*bors-a* F.SG 'bag'	*bors-one* M.SG 'big bag'
		febbr-e F.SG 'fever'	*febbr-one* M.SG 'high fever'
		magli-a F.SG 'jersey'	*magli-one* M.SG 'sweater'
		seggiol-a F.SG 'chair'	*seggiol-one* M.SG 'high chair'
	Sp.	*camis-a* F.SG 'shirt'	*camis-ón* M.SG 'night gown'
		mant-a F.SG 'blanket'	*mant-ón* M.SG 'traditional shawl'
		cuchar-a F.SG 'spoon'	*cuchar-ón* M.SG 'ladle'
	Cat.	*finestra* F.SG 'window'	*finestró* M 'big window'
	Pt.	*febre* F.SG 'fever'	*febrão* M.SG 'very high fever'
		festa F.SG 'party'	*festão* M.SG 'very big party'
		maçaneta F.SG 'handle'	*maçanetão* M.SG 'big handle'
		porta F.SG 'door'	*portão* M.SG 'gate'
	Rom.	*fată* F.SG 'girl'	*fătoi* M.SG 'big, healthy girl' can (need not) be pejorative
		uşă F.SG 'door'	*uşoi* M.SG 'entrance door'
		geantă F.SG 'bag'	*gentoi* M.SG 'big bag'
		babă F.SG 'old woman'	*băboi* M.SG 'old hag'

In some cases, the gender-shifted form is lexicalized, and the resulting *NOUN-on-e* M.SG denotes a special type of *NOUN-a* F.SG: this is the case of *maglia~maglione*, *seggiola~seggiolone*. However, gender-shifting *-on-* is in fact very productive and can be attached to virtually any feminine noun; the common interpretation of non-lexicalized gender-shifted *NOUN-on-e* M.SG is that of an exaggeratedly big and therefore slightly ridiculous *NOUN-a* F.SG. This phenomenon is particularly interesting, because it combines an uncontroversial derivational morpheme with what is normally considered inflectional morphology, thus showing how difficult it can be to tell derivation and inflection apart, even in a domain such as Romance, where the distinction is generally felt as reasonably clear-cut. In particular, the fact that the agreement class of a noun may change as a consequence of augmentative derivation makes these cases not entirely different from the Bantu ones, in particular the examples where derivational suffixes and noun classes interact, such as in (12).

3.6. Conclusions

In this section, much of the empirical evidence used in Schadeberg (2001) and Kihm (2005) to argue for the derivational nature of noun classes as opposed to gender/number was replicated in Romance, in some cases with examples that are virtually identical to

the Swahili or Manjaku ones: the unexpected conclusion may be that the gender/number is derivational in Romance as well. In other words, there is no principled difference between the two systems, as they can be described and analysed in similar terms. It is possible that there is a difference in the magnitude of the evidence discussed: where Romance has a handful of examples, Bantu may have hundreds. Quantitative data in this sense are not available for either of the two groups, so for the moment this argument must be set aside. Note however that even if it turned out that a difference in magnitude really exists, the difference between the two groups would be a quantitative rather than a qualitative one. Given that one group has two genders/classes and the other group has at least six, quantitative differences are entirely expected.

A point that has clearly emerged from the data presented is that the inflectional vs. derivational dichotomy is by no means a discrete one: 'inflection' and 'derivation' might be used as descriptive tools, but cannot be treated as grammatical categories. So, a morphosyntactic account of the difference between gender languages and noun-class languages such as the one proposed by Kihm (2005, see (32) and (33)) cannot be motivated on the basis of the derivational function of noun class prefixes, as opposed to the true inflectional nature of gender markers.

4. Chinese

4.1. Introduction

Sinitic languages present a picture which is quite different from the Bantu and Romance languages presented above. One eye-catching difference is that classification is a wholly DP-internal affair: there is no relation (of agreement or otherwise) between (elements involved in) classification and any element elsewhere in the sentence. In fact, in Mandarin the elements we call classifiers play a limited role even in the nominal domain: we only see them with numerals and demonstratives.

In this section, we briefly introduce a number of properties of the classification system in Chinese languages, especially Mandarin and Cantonese, paying attention both to the functional aspects as well as the more contentful, lexical aspects, from which the 'classifiers' draw their name, taking into account aspects of classification, and the elements involved in it, that we have discussed in the previous sections: principles of classification, number and derivation.

4.2. Sortals and measures

In what follows, we use the term 'classifier' to mean 'sortal classifier', to distinguish it from 'measure expressions'. The difference between sortals and measures is that while the latter creates a unit by which we can count or measure, the former simply names the unit that is already present in the semantic denotation of the noun (Croft 1994). This is illustrated in (52):

(52) a. yī běn shū
one CL$^{\text{VOLUME}}$ book
'one/a book'

b. yī duī shū
one pile book
'one/a pile of books'

Measure expressions include real measures (kilo, mile, spoonful) as well containers (cup, bucket, box) and collectors (group, heap, mass); an example is provided in (52b). They bring in their own semantics to impose on the noun a unit for counting and measuring, which has no necessary relation with the noun they combine with (they can be used with count nouns and mass nouns alike). A sortal like *běn* in (52a) is different. It does not impose a unit for counting onto the noun *shū* 'book'; rather, it just names the unit that is part of the semantic denotation of *shū* 'book'. The classifier *běn* is glossed as 'volume' in (52a). However, not all classifiers are glossed as easily, since in English we do not have separate lexical elements for the inherent unit by which things are counted, with the exceptions of a few, such as books.

The words 'sortal' and 'classifier' suggest that some sorting and classifying is being done. This is true to the extent that classifiers combine with nouns that can be seen to belong to one and the same class from one perspective or another: long thin objects (ropes, snakes, roads, rivers, slogans) are counted with the classifier *tiáo*, objects with a flat surface (stamps, maps, tables, beds) with *zhāng*, etc. We will look at the criteria for classification below.

Chinese languages have up to about 100 sortal classifiers. Some of these are very specific, such as *pǐ*, used only for horses and donkeys, and *dǔ* exclusively for walls and fences. Others are more general, such as *tiáo* and *zhāng* we just mentioned, as well as *lì*, which is used for all things that are small and (more or less) round: marbles, beans, grains, pearls, beads, bullets, tears, etc., and *zhī* for (small) animals.

Then there is *ge*, which, according to Loke (1994: 40-41) is the classifier for (a) all humans (regardless of sex, age, social status, occupation, etc.); (b) all abstract entities (such as: family, holiday,

opinion, victory); (c) all kinds of containers (like bucket, refrigerator); (d) rings and frames (e.g. circle, bracelet, doorway); (e) objects with a large enclosed hollow interior (e.g. drum, crystal ball); and (f) other solid saliently 3-dimensional objects (such as: cabbage, cake, alarm clock, car battery, buttocks). One can conclude from this, as Myers (2000) does, that *ge* does not classify at all, and that it is a general classifier in the sense that it is the default or the "elsewhere" option. Alternatively, one can say that *ge* is the sortal for humans, and the "elsewhere" for nouns that do not have a sortal of their own; we return to this point briefly below. The idea that *ge* is a default (not just an "elsewhere" element) can also be supported. Studies such as those by Erbaugh (1986) as well as Liu (2003) show that *ge* is used by speakers with nouns that combine with another sortal according to prescriptive grammars. Liu's data show cases in which at first mention of a noun the "right" classifier is used, while subsequently it is combined with *ge*. Erbaugh (1986, 2006) and Ahrens (1994) show that (respectively) for young children and certain aphasics *ge* also functions as a default in the sense that when they know when a classifier needs to be used, but they are not sure about which one, they fall back on *ge*.

Even though there may be as many as 100 classifiers, speakers tend to make use of a small subset of those in daily speech. Erbaugh (1986) found that adult Mandarin speakers on average use 22 different classifiers. Counting individual classifiers in a spoken corpus, Liu (2003) found 437 tokens and 17 types. Significantly, of the 437 tokens, 342 were *ge*, and 68 were instances of *zhǒng* meaning 'sort'. The remaining 27 tokens instantiated 15 different types. Erbaugh (2002) found that speakers of Cantonese use more different types. There are reasons to believe that *ge* is expanding its territory, at the expense of other classifiers, leading to an impoverished situation in that fewer classifiers are used.

4.3. Function of the sortal classifier

We can conclude from the previous section that, in any case for Mandarin speakers, classification is not the most prominent feature of the sortal classifiers. The reason why they are used seems to be essentially grammatical. The grammar tells us when we have to use one; which one we use is a matter of the lexicon (and language use). In this section we discuss the grammatical function, in the next one we return to the lexical content.

When it comes to the grammatical function of classifiers, the Chinese languages do not behave uniformly (Cheng and Sybesma 2005); we will illustrate this with Mandarin and Cantonese.

For both languages we observe that there is no counting without a classifier; this is why they are often referred to as "numeral classifiers". For Mandarin, this is basically it: we see the classifier only with numerals and with demonstratives (in which case they are often dropped in informal speech).

In Cantonese, however, we see the classifier in other contexts as well (Sybesma 2007 and references given there). The generalization is that in Cantonese, the classifier is used in all contexts in which individuality plays a role, that is, with specific indefinite and definite noun phrases. The following examples illustrate this.

(53) a. [*(bun^{2}) syu^{1}] hai^{2} li^{1}-dou^{6}
CL^{VOLUME} book be here
'the book is here' (definite)
b. ngo^{5} bong^{1} go^{3} hok^{6}-saang^{1} hai^{2} Gwong^{2}-zau^{1}wan^{2}-dou^{2} *(fan^{6}) gung^{1}
1s for CL student at Guangzhou find CL work
'I found a job for a student in Guangzhou' (specific indefinite)
c. gaan^{1} uk^{1} ge^{3} uk^{1}-deng^{2} cap^{3}-zyu^{6} *(zi^{1}) kei^{4}
CL house SUB roof stick-CONT CL flag
'there was a flag on the roof of the house' (specific indefinite)

A probably related property of classifiers is their relation with number (Cheng and Sybesma 1999). In Chinese languages, bare nouns are unspecified for number (see (54a)). [Cl-N] phrases, however, are always singular. The distribution of such phrases is limited in Mandarin (basically, to object position; an example is given in (54b)), but not so in Cantonese, where they can also occupy the subject position, as shown in (53).

(54) a. shū zài zhuōzi-shang
book at table-top
'the book/the books is/are on the table'
b. wǒ xiǎng mǎi běn shū
1s want buy CL book
'I want to buy a book' (sg only)

In (54a) we see that the bare noun can be interpreted as singular or plural; which interpretation it will have, depends on the context. In (54b) we see an example of a [Cl-N] phrase in Mandarin, showing that it, like its counterparts in Cantonese in (53), can only be interpreted as singular.

Both Mandarin and Cantonese also have what may be a plural classifier, *xiē* and *di^{1}* respectively; compare the following examples with (54b) and (53a):

(55) a. wǒ xiǎng mǎi xiē shū
1s want buy CL^{PL} book
'I want to buy some books' (pl only)
b. [di^{1} syu^{1}] hai^{2} li^{1}-dou^{6}
CL^{PL} book be here
'the books are here' (pl only; definite)

Xiē and *di^{1}* are never used in counting. The only numeral they combine with is 'one', leading to the meaning 'some', 'a few': (56) is the same as (55a):

(56) wǒ xiǎng mǎi yi-xiē shū
1s want buy one-CL^{PL} book
'I want to buy some books'

Note that there is just one plural classifier: unlike what we saw in Bantu and Romance, class distinctions disappear in the plural, although from the Bantu perspective, we may say that *xiē* and *di^{1}* are just class markers like all the other classifiers, in their case for (all) plural things.

What we have seen in this section is that the classifier is used to mark individuality, in counting and otherwise (at least in Cantonese definites and specific indefinites), and that it is also related to number.

4.4. Semantics of the classification

Closing off the section on Chinese, we briefly look at the principles behind the classification (Loke 1994, T'sou 1976, Rovira-Esteva 2008). It is important to note that the classification is wholly semantic: the phonological form of the word, for instance, plays no role whatsoever.

An important factor in classification is *geometrical shape/form*, putting together objects on the basis of whether they are spherical/round, long, slender, cylindrical, have a flat surface, a horizontal or vertical orientation, etc. We have seen examples above (*tiáo*, *zhāng*, *lì*).

Another factor is constituted by *natural attributes*, according to which different types of fauna (wild, domestic), flora, arboreals, etc. are categorized. We have seen that horses and donkeys go with *pǐ*; here are three more examples:

(57) birds, insects, land animals: *yī **zhī** máquè* /one CL sparrow/ 'one sparrow'
*yī **zhī** lǎohǔ* /one CL tiger/ 'one tiger'
cattle: *yī **tóu** niú* /one CL cow/ 'one cow'
flora: *yī **kē** mǔdan* /one CL peony/ 'one peony'

Next are the *cultural attributes*, with different classifiers for buildings, clothing, written records, vehicles, (small) instruments and (big) machines.

(58) clothing: *yī **jiàn** hànshān* /one CL t-shirt/ 'one t-shirt'
vehicles: *yī **liǎng** zìxíngchē* /one CL bicycle/ 'one bicycle'
machines: *yī **jià** fēijī* /one CL airplane/ 'one airplane'
instruments: *yī **tái** diànnǎo* /one CL computer/ 'one computer'

Finally, there is a mixed bag of objects, for example, the group of things with a handle or that one can grab and hold with one hand, which go with *bǎ*:

(59) a. *yī **bǎ** yǐzi* /one CL chair/ 'one chair'
b. *yī **bǎ** dāo* /one CL knife/ 'one knife'

Earlier on, we mentioned the "elsewhere" function we can ascribe to the classifier *ge*, which we can make a bit more concrete now. From the short discussion about the basics of the classification, we see that it only concerns concrete objects, animate and otherwise, and humans. All "things" which do not have geometrical shape/form, or natural or cultural attributes (i.e., abstract notions and concepts, etc.) cannot be classified, and go with *ge*.

A few more comments about the workings of the system are in order. As it goes, objects and creatures can in principle fall into several categories: something can be a machine and having a flat surface at the same time; a snake is both a long, thin, flexible thing, and an animal; trousers can be seen as long and thin, but they are also clothing. However, generally speaking, nouns combine with one sortal classifier only (although there is some regional and speaker variation). Thus, Mandarin snakes are counted using *tiáo*, the classifier for long, thin, flexible things, rather than *zhī* for animals. Likewise, trousers go with *tiáo* instead of *jiàn* as most clothing does:

(60) a. *yī **tiáo** shé* /one CL snake/ 'one snake'
b. *yī **tiáo** kùzi* /one CL trousers/ 'one pair of trousers'

One does not, then, have a lot of flexibility or freedom to play around with. One cannot use the classifier for animals when counting the boyfriends of one daughter, for instance. It is only in a very limited number of cases that different classifiers can be used and that the choice of one or the other has a semantic effect. Here are two examples:

(61) a. *yī ge yáng* /one CL sheep/ 'one sheep: neutral'
b. *yī zhī yáng* /one CL sheep/ 'one sheep: as a small animal'
c. *yī tóu yáng* /one CL sheep/ 'one sheep: as a domestic, tame animal' (Zhang 2007)

(62) a. *yī duǒ huār* /one CL flower/ 'one flower: emphasizing the flower part'
b. *yī zhī huār* /one CL flower/ 'one flower: flower and stem'

From this we can also deduce that, different from what we saw in Bantu and Romance, Chinese sortal classifiers do not function as derivational elements.

4.5 Summary

In sum, Chinese classifiers have no function outside the nominal domain: they do not play a role in agreement relations with other categories in the sentence. They are used in counting and, in any case in Cantonese, to mark individuation.

5. Discussion: the point of classification

5.1. Introduction

From the discussion above, it appears that a unified treatment of classification for Bantu, Romance and Chinese is probably unattainable. What we do observe, however, is that in Italian, Swahili and Cantonese (in fact, in Romance, Bantu and Chinese more generally), the class/gender morphemes express two things, class and number:

(63) a. *libr-**o*** 'book' (Italian[17])
b. ***bun**²* *syu¹* 'a/the book' (Cantonese)
c. ***ki**-tabu* 'book' (Swahili)

This observation leads us to pose the following questions: 1. To what extent can class and number be distinguished or separated?

2. Why do we have the class/gender markers and classifiers at all? 3. Why is it the case that class and number should go together? We will look at these questions below.

5.2. Class and number

With respect to the first question, the discussion in the previous sections yields a mixed picture. When we look at ITALIAN, for a start, we have evidence that class and number are separate notions: when nouns agree with the verb, only number is at play and class no longer plays a role (cf. (35)). Also, the fact that in Spanish the plural marker is distinct from the class marker and surfaces to its right can be seen as showing that the two notions must be distinguished.

The case of SWAHILI is more complex. Let us summarize what we have said about number in Swahili above. First, in as far as it is possible to distinguish between the derivational and non-derivational classes in Bantu, speakers of Bantu languages have a clear awareness of the semantic number associated with the different non-derivational classes, at least for words which come in pairs. The best example for this are the 'human' classes 1/2, where we could say that class 1 and 2 are both, let's say, [+human], and so the purpose of having two markers is not to distinguish two classes, but to distinguish two numbers.[18]

A second indication can be deduced from the coordination facts we discussed in section 2.5. Looking once more at the Bantu example in (13), we observe that when we coordinate two class 1 Ns, we get class 2 marking on the verb, not class 1. Similarly, in (14), when we coordinate two inanimate nouns, be they singular or plural, we get class 8 on the verb, or 6 or 10 – most readily 8, the class for "plural things". Crucially, for a conjunction of, say, two class 5 nouns, we do not get class 5 (except arguably as cases of closest conjunct agreement). The point here is that if number played no role whatsoever, we would expect to get class 1 on the verb in the first case and 5 in the last. The fact that we do not get these shows that number, or at least plurality, plays a role separate from class.

Turning, finally, to CANTONESE, we recall that there are many different classifiers associated with singular, but only one for plural. This indicates that class and number can be separated. If we do not express the one (in this case, class), we can still express the other (number). A similar conclusion can be drawn considering what happens in what we called the "impoverishment" in the use of classifiers in Mandarin above. What happens here is a decrease in the number of different classifiers in use, not in the use of the category classifier

as such: apparently, classification is not the only reason why we have classifiers.

Kihm (2005) observes the same to have happened in Dakar Wolof. In the varieties of Wolof spoken outside of Dakar, there are eight classes for the singular and two for the plural (incidentally an unbalanced situation very similar to what we have in Cantonese). In Dakar Wolof, presumably an "impoverished" variety, we have one class marker for the singular and one for the plural. In other words, the class marking function of the class markers disappears, but not their number marking function, thus showing that class and number operate separately.

We conclude that in all three language groups, we can find indications that number and class play separate roles, to some extent independent of one another.

5.3. Why number and class?

It is clear that number and class both have roles to play and that they do so in separate ways, but at the same time we acknowledge the fact that many of the elements we have discussed in this paper express both class and number in portmanteau fashion. Why would this be the case?

Ideally, the reason for adding a grammatical marker to a noun should be the same for all cases in all languages, at least at its core. Our hypothesis regarding why we need these markers at all, is close to what Ferrari-Bridgers (2008) and Kihm (2005) propose as summarized above. Van de Velde (2006) proposes that the classification system we see in Bantu languages "may have emerged out of a need to provide referential disambiguation" (p. 219). Kihm and Ferrari-Bridgers assume, as we have seen, that the markers are needed to derive nouns from lexical roots and are in fact little *n*s, and that "the basic function of noun classes [is] noun formation" (Kihm 2005: 471).

We propose that the classificational elements in Italian, Swahili and Cantonese play an important role in the process of singling out individuals. Similar to what Kihm proposes, we think that class markers may be heading the first functional category above the lexical root (whatever the label of this should be). Thus, they are the first step in the path that ultimately leads to referentiality, the latter arguably involving higher functional projection(s) as well, notably, DP (see Szabolcsi 1987, Stowell 1989, Longobardi 1994). Also, we take it that the operation of getting from the uncategorized lexical root to a ref-

erential expression involves marking out an individual unit, or individual units (or a portion of stuff), and this is where number, possibly heading its own functional projection (but see Borer 2005), comes into play. Class and number are therefore involved in the same process and possibly head (near-)contiguous projections, and this may be the reason for their frequent surfacing as a single morpheme.

The next question that comes up is: how does 'class' come in? Why is this operation of making an NP referential by way of marking out individuality related to classification? Why do we get class information when we mark number or individuality?

Our answer to these questions is quite speculative. What we have been calling 'marking number' (and marking individuality) must perhaps be paraphrased as: 'taking one or more individual units belonging to a certain class'. In other words: in marking individuality, we (must) make explicit what *kind* of individual we are dealing with. As we discussed above, in classification, we most rudimentarily distinguish between animates and inanimates, as the Indo-European forefathers did. Marking individual instances of 'people' leads to 'one people-being' or 'one being of people' whereas with stones we get 'one stone-thing', 'one thing of stone'. In more elaborate systems, we distinguish different types of animates (people, sometimes even male vs. female, small animals, big animals, swimming ones, flying ones) and different types of inanimates (long thin flexible things, flat square things, things with a handle): people-person instead of people-being, animal-being, stone-round.thing, etc.

If these speculations make any sense at all, we conclude that marking 'class', and 'number', is part of a process to mark individuality, and that one and the same element may be involved in marking both. It is possible that other factors play a role as well, such as needs for agreement or word formation needs independent of number/individuality. This has to be investigated further.

Addresses of the Authors

Paola Crisma: Dip.di Filosofia, Lingue e Letterature, Androna Campo Marzio 10, 34123 Trieste, Italy <crismap@units.it>
Lutz Marten: School of Oriental and African Studies, Thornhaugh Street, Russell Square, London WC1H 0XG <lm5@soas.ac.uk>
Rint Sybesma: Leiden University, GW-Chinese Studies, P.O. box 9515, 2300 RA Leiden, The Netherlands <R.P.E.Sybesma@hum.leidenuniv.nl>

Notes

* Acknowledgments
Research reported in this paper results from the international network on the structure and interpretation of nouns and noun phrases in Bantu, Chinese and Romance. We are grateful to the Netherlands Organization for Scientific Research (NWO), as well as Leiden University, the University of Trieste and SOAS for financial support for the network. Paola Crisma was also supported by a PRIN grant (PRIN 2008EHLWYE_005). We are also grateful to Lisa Cheng, Romano Lazzeroni, Anna Morpurgo-Davies and Nigel Vincent for helpful suggestions on aspects of the paper. All errors and shortcomings remain, of course, our own.

1 With the exception of the two generic terms *mnyama* 'animal' and *mdudu* 'insect'.

2 Notes: 1) Vowel-initial stems trigger phonological processes not indicated here; 2) Noun class prefixes are also used for adjective 'agreement'; 3) The concord is used in verbal agreement, as subject markers and object markers (except in class 1, where SM = *a-*, OM = *m(w)-*); 4) Class 11 plurals (if possible) are typically in class 10; 5) There is only one word in classes 16-18, these are mainly used in agreement.

3 Sagna's (2008) study is concerned with Gújjolaay Eegimaa, a Niger-Congo language of Senegal and so outside of Bantu, but his approach is very similar to studies of Bantu languages.

4 See Moxley (1998) for a similar proposal.

5 From the foregoing discussion it appears that all studies investigating the semantic base of noun classification are conducted from a cognitive/functional perspective. This is not necessarily the case though. Harbour (2008), for example, develops a formal (Minimalist) analysis of (interpretable) class features in Kiowa (a Kiowa-Tanoan language spoken in Oklahoma, USA), based on the notion of "semantic coherence" which allows cases of a given noun fulfilling semantic criteria for membership in two classes, exceptions to the overall semantic criteria, and subclasses.

6 Kihm's (2005) evidence is based on Manjaku, a Niger-Congo language of Guinea-Bissau, rather than on Bantu, but his conceptual argument is concerned with noun class and gender more generally. We will discuss his proposal in more detail below.

7 Note that animacy has also been argued to underlie the Indo-European gender system (e.g. Meillet 1921), a point to which we return in 3.1 below.

8 Masculine is the default gender.

9 For an analysis of declension classes in Italian, see Lampitelli (2010).

10 We ignore here the small number of invariant adjectives such as *rosa* 'pink'.

11 Or possibly [+ human], see Di Domenico (1997), Zamparelli (2008).

12 Here and below: IT. = Italian, SP. = Spanish (Castillan), CAT. = Catalan, PT. = Portuguese, FR. = French, ROM. = Romanian. We are grateful to our informants Manuel Español Echevarría, Charlotte Galves, Miriam Lemle, Manuel Leonetti, Ana Maria Martins, Carme Picallo, Alla Robu.

13 "Singular-only", "Plural-only": our labels.

14 The symbol % represents the fact that some speaker may accept both the singular and the plural form.

15 And very marginally in French: *amour* M.SG 'love' ~ *amours* F.PL 'love affairs', *orgue* M.SG 'pipe organ' ~ *orgues* F.PL 'pipe organs' (emphatic), *délice* M.SG 'delight' ~ *délices* F.PL 'delices'.

16 Cases of totally unrelated roots that happen to have the same form albeit different genders (e.g. *coll-o* M.SG 'neck' *coll-a* F.SG 'glue') are disregarded. Cases of interpretable gender with [+animate] nouns, as in (36)b, are equally ignored.

17 This property of Italian is not found in all Romance languages, see 5.2 below.

[18] The same can be said about Italian, which has two markers for each gender, one plural and one singular: the purpose of these is not to mark difference in gender, but to mark difference in number.

Bibliographical References

Acquaviva Paolo 2008. *Lexical Plurals: A Morphosemantic Approach.* Oxford: Oxford University Press.

Ahrens Kathleen 1994. Classifier production in normals and aphasics. *Journal of Chinese Linguistics* 22/2. 202-246.

Amidu Assibi A. 1997. *Classes in Kiswahili*. Köln: Köppe.

Ashton Ethel O. 1947. *Swahili Grammar (including intonation),* 2nd edition. London: Longmans.

Beard Robert 1998. Derivation. In Spencer & Zwicky 1998. 44-65.

Blois Kornelis Frans de 1970. The augment in Bantu languages. *Africana linguistica* 4. 85-165. Annales du MRAC (Musée Royal de l'Afrique Centrale), sciences humaines, 68. Tervuren.

Bokamba Eyamba G. 1985. Verbal agreement as a non-cyclic rule in Bantu. In Goyvaerts Didier L. (ed.). *African Linguistics: Essays in memory of M. W. K. Semikenke*. Amsterdam: Benjamins. 9-54.

Borer Hagit 2005. *In Name Only.* [Structuring sense I.] Oxford: Oxford University Press.

Carstens Vicky 1991. *The Morphology and Syntax of Determiner Phrases in Kiswahili*. Los Angeles: University of California at Los Angeles. PhD dissertation.

Carstens Vicky 1993. On nominal morphology and DP structure. In Mchombo Sam A. (ed.). *Theoretical Aspects of Bantu Grammar*. Stanford, Cal.: CSLI. 151-180.

Carstens Vicky 2008. DP in Bantu and Romance. In de Cat & Demuth 2008. 131–165.

Cheng Lisa & Rint Sybesma 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30. 509-542.

Cheng Lisa & Rint Sybesma 2005. Classifiers in four varieties of Chinese. In Cinque Guglielmo & Richard S. Kayne (eds.) *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press. 259-292.

Contini-Morava Ellen 1994. Noun classification in Swahili. Charlottesville: University of Virginia. Ms. http://www3.iath.virginia.edu/swahili/swahili.html (20.09.2010)

Contini-Morava Ellen 1996. 'Things' in a noun-class language. Semantic functions of agreement in Swahili. In Andrews Edna & Yishai Tobin (eds.). *Toward a Calculus of Meaning. Studies in Markedness, Distinctive Features and Deixis*. (Studies in functional and structural linguistics, vol. 43). Amsterdam/Philadelphia: Benjamins. 251-290.

Contini-Morava Ellen 1997. Noun classification in Swahili: a cognitive-semantic analysis using a computer database. In Herbert Robert K. (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 CONTINI-MORAVA Ellen & Yishai TOBIN (eds.). *Between Grammar and Lexicon*, Current Issues in Linguistic Theory 183. Amsterdam/Philadelphia: Benjamins. 3–30.

CORBETT Greville 1983. Resolution rules: agreement on person, number, and gender. In GAZDAR Gerald, Ewan KLEIN & Geoffrey PULLUM (eds.). *Order, Concord, and Constituency*. Foris: Dordrecht. 175–206.

CORBETT Greville 1991. *Gender*. Cambridge: Cambridge University Press.

CROFT William 1994. Semantic universals in classifier systems. *Word* 45. 145–171.

CUYCKENS Hubert & Britta ZAWADA (eds.) 2001. *Polysemy in Cognitive Linguistics: Selected Papers from the Fifth International Cognitive Linguistics Conference*. Amsterdam/Philadelphia: Benjamins.

DE CAT Cécile & Katherine DEMUTH (eds.) 2008. *The Romance-Bantu Connection*. Amsterdam: Benjamins.

DEMUTH Katherine 2000. Bantu noun class systems: loanword and acquisition evidence of semantic productivity. In SENFT Gunter (ed.). *Systems of Nominal Classification*. Cambridge: Cambridge University Press. 270-292.

DENNY J. Peter & Chet A. CREIDER 1976. The semantics of noun classes in Proto-Bantu. *Studies in African Linguistics 7.1*. 1–30.

DI DOMENICO Elisa 1997. *Per una teoria del genere grammaticale*. Padova: Unipress.

ERBAUGH Mary 1986. Taking stock: the development of Chinese noun classifiers historically and in young children. In CRAIG Colette (ed.). *Noun Classes and Categorization*. Amsterdam: Benjamins. 399-436.

ERBAUGH Mary S. 2002. Classifiers are for specification: complementary functions of sortal and general classifiers in Cantonese and Mandarin. *Cahiers de linguistique Asie Orientale* 31/1. 33-69.

ERBAUGH Mary S. 2006. Chinese classifiers: their use and acquisition. In: LI Ping, Li Hai TAN, Elizabeth BATES & Ovid TZENG (eds.). *The Handbook of East Asian Psycholinguistics. Volume 1: Chinese*. Cambridge: Cambridge University Press. 39-51.

FERRARI Franca 2005. *A Syntactic Analysis of the Nominal Systems of Italian and Luganda: How Nouns Can Be Formed in the Syntax*. New York: New York University. PhD dissertation.

FERRARI-BRIDGERS Franca 2008. A unified syntactic analysis of Italian and Luganda nouns. In DE CAT & DEMUTH 2008. 239-269.

GIVÓN Talmy 1972. Studies in ChiBemba and Bantu grammar. *Studies in African Linguistics*, supplement 3.

GREENBERG Joseph H. 1978. How does a language acquire gender markers? In GREENBERG Joseph H., Charles FERGUSON & Edith MORAVCSIK (eds.).*Universals of Human Language: Volume 3: Word Structure*. Stanford, CA: Stanford University Press. 47-82.

GRINEVALD Colette 2000. A morphosyntactic typology of classifiers. In SENFT Gunter (ed.). *Systems of Nominal Classification*. Cambridge: Cambridge University Press. 50-92.

HARBOUR Daniel 2008. *Morphosemantic Number: From Kiowa Noun Classes to UG Number Features*. Dordrecht: Springer.

HARRIS James W. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22(1). 27-62.

HENDRIKSE Andries Petrus 2001. Systemic polysemy in the Southern Bantu noun class system. In CUYCKENS & ZAWADA 2001. 185-212.

IDIATA Daniel Franck, Myles LEITCH & Pierre ONDO-MEBIAME 2000. *Les classes nominales et leur sémantisme dans le langues Bantu de nord-ouest*. Munich: Lincom.

IDIATA Daniel Franck 2005. *What Bantu Child Speech Data Tells Us about the Controversial Semantics of Bantu Noun Class Systems*. Munich: Lincom.

JOHNSON Frederick 1939. *A Standard Swahili-English Dictionary*. London: Oxford University Press.

KATAMBA Francis 2003. Bantu nominal morphology. In NURSE Derek & Gerard PHILIPPSON (eds.). *The Bantu Languages*. London: Routledge. 103-120.

KAVARI Jekura & Lutz MARTEN 2009. Multiple noun class prefixes in Otjiherero. In AUSTIN Peter, Oliver BOND, Monik CHARETTE, David NATHAN & Peter SELLS (eds.). *Proceedings of the 2nd Conference on Language Description and Theory*. London: SOAS. 165-173.

KIHM Alain 2005. Noun class, gender, and the lexicon / syntax / morphology interfaces: a comparative study of Niger-Congo and Romance languages. In CINQUE Guglielmo & Richard KAYNE (eds.). *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press. 459-512. http://www.llf.cnrs.fr/Gens/Kihm/OUPhandbook.pdf

KILLINGLEY Siew-Yue 1983. *Cantonese Classifiers: Syntax and Semantics*. Newcastle upon Tyne: Grevatt & Grevatt.

LAMPITELLI Nicola 2010. Nounness, gender, class and syntactic structures in Italian nouns. In BOK-BENNEMA Reineke, Brigitte KAMPERS-MANHE & Bart HOLLEBRANDSE (eds.). *Romance Languages and Linguistic Theory 2008*. Amsterdam: Benjamins. 195-214.

LIU Hsin-Yun 2003. *A Profile of the Mandarin NP. Possessive Phrases and Classifier Phrases in Spoken Discourse.* München: Lincom.

LOKE Kit-Ken 1994. Is *ge* merely a "general classifier"? *Journal of the Chinese Language Teachers Association* 29/3. 35-50.

LOKE Kit-Ken 1996. Norms and realities of Mandarin shape classifiers. *Journal of the Chinese Language Teachers Association* 31/2. 1-22.

LONGOBARDI Giuseppe 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25.4. 609-655.

MAHO Jouni 1999. *A Comparative Study of Bantu Noun Classes*. Gothenburg: Acta Universitatis Gothoburgensis.

MARTEN Lutz 2000. Agreement with conjoined noun phrases in Swahili. *Afrikanistische Arbeitspapiere 64: Swahili Forum VII*. 75-96.

MARTEN Lutz 2005. The dynamics of agreement and conjunction. *Lingua* 115. 527-547.

MARTEN Lutz 2010. The great siSwati locative shift. In BREITBARTH Anne, Christopher LUCAS, Sheila WATTS & David WILLIS (eds.). *Continuity and Change in Grammar*. Amsterdam: Benjamins. 249-267.

MARTEN Lutz *forthc.*, Agreement in locative phrases in Luganda. *Proceedings of WOCAL*.

MEILLET Antoine 1921. *Linguistique historique et linguistique générale*, vol. 1. Paris: Champion.

MEILLET Antoine 1931. Essai de chronologie des langues indo-européennes. *Bulletin de la Société de Linguistique de Paris* 32. 1-28.

MÖHLIG Wilhelm J. G. & Jekura U. KAVARI 2008. *Reference Grammar of Herero (Otjiherero)*. Köln: Köppe.

MÖHLIG Wilhelm J. G., Lutz MARTEN & Jekura U. KAVARI 2002. *A Grammatical Sketch of Herero (Otjiherero)*. Köln: Köppe.

MOHAMED, Mohamed Abdulla 2001. *Modern Swahili Grammar*. Narirobi: East African Educational Publishers.

MOXLEY Jeri 1998. Semantic structure of Bantu noun classes. In MADDIESON Ian & Thomas J. HINNEBUSCH (eds.). *Language History and Linguistic Description in Africa*. Trends in African Linguistics 2. Trenton, NJ & Asmara: Africa World Press, Inc. 229-238.

MYERS James 2000. Rules vs. analogy in Mandarin classifier selection. *Language and Linguistics* 1/2. 187-209.

PALMER Gary B. & Claudia WOODMAN 2000. Ontological classifiers as polycentric categories, as seen in Shona class 3 nouns. In PÜTZ Martin & Marjolijn H. VERSPOOR (eds.). *Explorations in Linguistics Relativity*. Amsterdam/Philadelphia: Benjamins. 225-249.

PEYRAUBE Alain 1998. On the history of classifiers in archaic and medieval Chinese. In T'SOU Benjamin (ed.). *Studia linguistica serica*. Hong Kong: City University of Hong Kong. 39-68.

PLANK Frans 1994. Inflection and Derivation. In ASHER Ronald E. & J. M. Y. SIMPSON (eds.). *The Encyclopedia of Language and Linguistics*, vol. 3. Oxford: Pergamon Press. 1671-1678.

RICHARDSON Irvine 1967. Linguistic evolution and Bantu noun class systems. In MANESSY Gabriel & André MARTINET (eds.). *La Classification Nominale dans les Langues Négro-Africaines*. Aix-en-Provence: Éditions du Centre National de la Recherche Scientifique. 373-390.

ROVIRA-ESTEVA Sara 2008. Chinese classification categories revisited. *Journal of Chinese Linguistics*, 36(1). 106-120.

SAGNA Serge 2008. *Formal and Semantic Properties of the Gújjolaay Eegimaa (a.k.a. Banjal) Nominal Classification System*. London: SOAS. PhD dissertation.

SCHADEBERG Thilo C. 1992. *A Sketch of Swahili Morphology*, 3rd edition. Köln: Köppe.

SCHADEBERG Thilo C. 2001. Number in Swahili grammar. *Afrikanistische Arbeitspapiere 68: Swahili Forum VIII*. 7-16.

SELVIK Kari-Anne 2001. When a dance resembles a tree: a polysemy analysis of three Setswana noun classes. In CUYCKENS & ZAWADA (2001). 161-184.

SPENCER Andrew & Arnold M. ZWICKY (eds.) 1998. *The Handbook of Morphology*. Oxford: Blackwell.

SPITULNIK Debra A. 1987. *Semantic Superstructuring and Infrastructuring: Nominal Class Struggle in ChiBemba*. Bloomington, Indiana: Indiana University Linguistics Club.

SPITULNIK Debra A. 2001. *Levels of Semantic Structuring in Bemba Noun Classification*. Munich: Lincom.

Stowell Timothy 1989. Subjects, specifiers, and X-bar theory. In Baltin Mark & Antony Kroch (eds.). *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press. 232-262.
Stump Gregory T. 1993. Reconstituting morphology: the case of Bantu preprefixation. *Linguistic Analysis* 23 (3/4). 169-204.
Stump Gregory T. 1998. Inflection. In Spencer & Zwicky 1998. 13-43.
Sybesma Rint 2007. Běifāng fāngyán hé Yuèyǔ zhōng míngcí de kěshǔbiāojì (Markers of countability on the noun in Mandarin and Cantonese). *Yǔyánxué lùncōng* 35. 234-245.
Szabolcsi Anna 1987. Functional categories in the noun phrase. In Kenesei István (ed.). *Approaches to Hungarian, Vol. 2*. Szeged: Jate. 167-189.
T'sou Benjamin 1976. The structure of nominal classifier systems. In Jenner Philip, Laurence Thompson & Stanley Starosta (eds.). *Austroasiatic Studies. Part II*. Honolulu: The University Press of Hawaii. 1215-1247.
Van de Velde Mark 2006. Multifunctional agreement patterns in Bantu and the possibility of genderless nouns. *Linguistic Typology* 10. 183-221.
Van der Spuy Andrew 2010. Generating Zulu noun class morphology. *Language Matters* 41. 294-314.
Zamparelli Roberto 2008. On the interpretability of phi-features. In De Cat & Demuth 2008. 167-199.
Zhang Hong 2007. Numeral classifiers in Mandarin Chinese. *Journal of East Asian Linguistics* 16/1. 43-59.