Contents of *Linguistic Typology*
Volume 8 (2004)

Noun classes in African and Amazonian languages: Towards a comparison

COLETTE GRINEVALD and FRANK SEIFART

Abstract

Many Amazonian systems of nominal classification have been perceived as constituting a descriptive and typological challenge. The proposal presented here is to consider many of them as emerging noun class systems rather than as a-typical systems that defy integration within an overall typology of nominal classification, at the opposite end from the Niger-Congo systems on a continuum of grammaticalization. First the African noun class systems are reviewed, with an emphasis on the sociolinguistic context of their descriptions and on their common deviations from a prototypical image of them projected in the general linguistic literature. Then a recapitulation of various proposals of a-typicality of the Amazonian systems is given, followed by the presentation of a typology of nominal classification systems that integrates the dynamic dimension of grammaticalization. The application of this typological framework is illustrated with a case study from the Miraña language of Colombia.

Keywords: agreement, Amazonian languages, classifier, concord, gender, grammaticalization, Miraña, Niger-Congo languages, noun class, number, numeral classifier

1. Introduction

The original impetus for a comparison of African and Amazonian linguistics came from the need to respond to the descriptive challenge encountered by field linguists facing the wealth of nominal classification systems of the languages of the Amazon region. It was a decision taken with some sort of *a priori* assumption that there was no particular reason why languages in the Amazon region should be any more "exotic" and indecipherable than languages of other parts of the world. The feeling was that the claimed exoticism of Amazonian systems (as, for instance, in Payne 1987, Derbyshire & Payne 1990, Aikhenvald

1994, and Aikhenvald & Green 1998) was no more than the reflection of their being the ones most recently encountered and the ones having to fit within already established typological patterns. The idea therefore was to give these last languages on board the benefit of the doubt and to search for ways to describe them as if they were "normal" languages, maybe new and surprising to us, but "normal" and calling for a readjustment of our typological approach.[1]

And since one of the striking characteristics of many of those Amazonian systems of nominal classification is their basic concordial nature,[2] a comparison to the well documented and more familiar concordial systems of Niger-Congo languages seemed imperative. In this process of a systematic comparison of the nominal classification systems of the two sides of the Atlantic, the African and the Amazonian ones, the need was felt to reconsider in more detail the relatively stable and overall wholesome picture of the African systems that is projected in the general linguistic literature and to expose their actual irregularities and variations. This more realistic view of the African systems is one of the two ways to draw a closer parallel with the Amazonian systems, the other way being to reconsider their respective degrees of grammaticalization.

It will be argued that, in the end, the comparison of African and Amazonian languages is fruitful on both sides. While it is definitely contributing to the development of a productive descriptive strategy for handling the Amazonian data, it also helps in putting the African systems in a broader typological perspective. In particular it offers a vision of a possible origin for those systems that are now so far advanced in their grammaticalization that the issue of their origin is only a matter of speculation. This comparison of African and Amazonian systems can provide us therefore with an expanded understanding of the life cycles of the omnipresent phenomenon of nominal classification in the languages of the world.

The paper will proceed as follows. Section 2 will offer a review of the state of knowledge of the phenomenon of the noun class systems of the Niger-Congo language family, outlining the type of formal and semantic characteristics that will be drawn on in the comparison with the Amazonian systems. We owe the reconsideration of the systems presented here to Denis Creissels. An independent and fuller account of the notion of prototypical noun class systems and their deviations in Niger-Congo languages can be found in Creissels (2001). Section 3 will articulate the nature of the challenge posed by the Amazonian

---

1. A similar approach of de-exoticization is being proposed for the study of the so-called verbal classifiers of sign languages in Grinevald (2003), which it is argued could benefit from closer comparison with the phenomenon of nominal classification systems of oral languages.

2. The term "concordial" is the one found in most of the traditional literature on such systems; it is used here as an equivalent of the term "agreement" more familiar in contemporary grammatical discussions.

languages, including the relevant sociolinguistic aspects of the issue, some of the claims of deviance of Amazonian nominal classification systems found in the recent literature, and a quick overview of the typological framework proposed to consider many of them as weakly grammaticalized noun class systems (based on Grinevald 2000, 2001, 2002). Section 4 illustrates the descriptive strategy proposed here with a case study by Frank Seifart of a typical Amazonian nominal classification system, that of the Miraña language of Colombia (see Seifart 2002 for a more complete description of the system). The concluding section will review the various points of comparison between African and Amazonian systems.

2.   Niger-Congo classification systems

This reconsideration of the Niger-Congo noun class systems will proceed as follows: it will first seek to attract attention to an aspect rarely considered in the literature, that of the sociolinguistic conditions of their descriptions (Section 2.1). It will then review the essential characteristics of the systems, such as their degree of grammaticalization (2.2) and the notion of "class" as opposed to that of gender (2.3), and the specific ways in which many systems do not conform in fact, in their natural spoken forms, to the prototype image of such systems that is often projected in the general linguistic literature (2.4). The issue of the extent and the nature of the semantic motivation of the classes is considered next, again from the point of view of the languages as spoken today (2.5), and the last section underlines the difficulty encountered in identifying the lexical origin of the proto-system (2.6).

2.1.   *The normalized descriptions of African classification systems*

Our aim is to show that previous studies of Amazonian systems have underestimated the possibilities of drawing parallels between African and Amazonian classification systems, due in particular to differences in the nature of the available documentation. It happens that some African languages have a grammatical tradition that goes back to the 19th century, and that is very often oriented towards standardization in order to facilitate the use of the language in contexts such as school teaching, Bible reading, mass media, etc. The easiest sources of documentation on such languages are therefore pedagogical grammars that by nature tend to minimize variations and irregularities. There are also a number of introductions to African linguistics which, not surprisingly, emphasize the most important regularities in African language systems and address very few of the complications found in the systems of individual languages. The result is that typologists with no first-hand experience of the description of Niger-Congo languages often have a somewhat simplified view of the Niger-Congo classification systems, due to the nature of the most

accessible source of documentation for them, namely such "normalized" presentations.

As will be underlined later, there is nothing similar with Amazonian languages, on three accounts: (i) unlike the situation of Niger-Congo languages, the descriptions of most Amazonian languages remain partial or yet to be done; (ii) very few comparative studies have been possible so far, partly due to lack of data and partly due to the great genetic variety of the languages of that part of the world; (iii) the high level of literacy of some of the languages in the African case, with its corollary standardization process, is in no way comparable yet anywhere in the Amazonian region, where all information on the languages comes from fieldwork on oral traditions that make apparent the variations and irregularities of the systems.

2.2.   *Strongly grammaticalized agreement systems*

The essential features of Niger-Congo classification systems which characterize them as strongly grammaticalized systems are these three: (i) nouns divide into subsets (noun classes) according to their behavior in agreement mechanisms; (ii) the forms involved in these agreement mechanisms (nouns, noun modifiers, pronouns, and verbs) include affixes (class markers) that determine their agreement behavior; (iii) all nouns enter the classification, which is basically a classification of nouns and not of referents.

2.3.   *Gender vs. class vs. number: A descriptive strategy*

Semantically, the most obvious function of noun class markers (but not the only one; see Section 2.5 below) is to encode number (singular vs. plural). An important feature of Niger-Congo classification systems is the impossibility of dissociating noun affixes encoding number from noun affixes expressing gender-like distinctions. Moreover, there is generally no one-to-one correspondence between class distinctions in the singular and in the plural. This is the reason why many descriptions of Niger-Congo class systems do not emphasize the possibility of dividing noun lexemes into genders, but rather start from a division of noun forms into classes in which the singular form and the plural form of a given count noun are treated as two distinct units. In this approach, a gender may be subsequently defined as a pair of classes that correspond with each other in the expression of number.

The regularity of agreement in Niger-Congo languages generally makes it easy to establish into how many classes noun forms divide (i.e., to how many different possible agreement patterns noun forms may belong). However, the idiosyncrasies shown by many nouns in the singular-plural correspondence and the variations observed at this level often make it very difficult to establish the exact number of genders, if one considers that nouns with the same agreement

properties in the singular but different agreement properties in the plural belong to two distinct genders.

The examples in (1) show the division of nouns into 12 classes in Tswana on the basis of the agreement between nouns and adjectives (classes numbered in accordance with their traditional numbering in Bantu studies).[3] The construction pattern of noun phrases with adjectives includes a linking element that varies itself according to class:

(1)			[CLx-N	link.CLx	CLx-Adj]	
	a.	class 1	*mo-sadi*	*yo*	*mo-s'a*	'new woman'
	b.	class 2	*ba-sadi*	*ba*	*ba-s'a*	'new women'
	c.	class 3	*mo-lemo*	*o*	*mo-s'a*	'new medicine'
	d.	class 4	*me-lemo*	*e*	*me-s'a*	'new medicines'
	e.	class 5	*le-saka*	*le*	*le-s'a*	'new cattle kraal'
	f.	class 6	*ma-raka*	*a*	*ma-s'a*[4]	'new cattle kraals'
	g.	class 7	*se-kolo*	*se*	*se-s'a*	'new school'
	h.	class 8–10	*di-kolo*	*tse*	*din-ts'ha*	'new schools'
			di-kgosi	*tse*	*din-ts'ha*	'new chiefs'
	i.	class 9	*Ø-kgosi*	*e*	*n-ts'ha*	'new chief'
	j.	class 11	*lo-kwalo*	*lo*	*lo-s'a*	'new book'
	k.	class 14	*bo-jang*	*jo*	*bo-s'a*	'new grass'
	l.	class 15–17	*go-lema*	*mo*	*go-s'a*	'new way of cultivating'

The agreement system of Tswana involves various types of modifiers such as genitive phrases, demonstratives, etc. The examples in (2) illustrate class agreement with numerals:

(2)		[CLx-N	link.CLx	*le*[5]	CLx-Num]	
	a.	*ba-sadi*	*ba*	*le*	*ba-pedi*	'two women'
	b.	*ma-kau*	*a*	*le*	*ma-pedi*	'two boys'

---

3. The numbering of classes currently used in Bantu studies relies on the correspondence with the reconstructed classes of Proto-Bantu, according to, e.g., Guthrie (1967–1971: Volume 4); "class 8–10" means that Tswana has a class that historically results from the merger of Proto-Bantu classes 8 and 10.

4. Some of the prefixes of Tswana (classes 5, 8, 8–10, and 11) can provoke consonantal mutations that are unpredictable and must be listed as properties of lexical items.

5. *le* is a participial form of the verb 'to be', the class marker and participial form combining in a complex linker between noun and numeral, so that the literal translation of (2a) should be 'women being two'.

c.	*di-fofu*	*di*	*le*	*Ø-pedi*[6]	'two blind
					persons'
d.	*di-ngaka*	*di*	*le*	*Ø-pedi*	'two doctors'

The presence of class markers on every element of complex noun phrases, including relative clauses, is obligatory, as shown in the eight instances of markers for class 1 in example (3) below on noun, linker, adjective, subject marker, and demonstrative:

(3)   *mo-sadi   yo   mo-leele   yo   mo-ntsho*
      CL1-woman   LINK.CL1   CL1-tall   LINK.CL1   CL1-black

      yo   o   opelang   yo-le
      LINK.CL1   SM.CL1   sing.SUB   CL1-DEM
      'this tall woman with dark complexion who is singing'

The examples of (4) further illustrate class markers functioning as subject and object markers (SM, OM) affixed to verbs and as pronouns (PN):[7]

(4)   a.   *Mo-sadi   o   lapile*
           CL1-woman   SM.CL1   be.tired
           'The woman is tired.'
      b.   *Ke   mo   thusitse*
           SM.1SG   OM.CL1   help.TAM
           'I helped her (the woman).'
      c.   *Ke   bua   le   ene*
           SM.1SG   speak   with   PN.CL1
           'I am speaking with her (the woman).'

2.4.   *Formal aspects of prototypical Niger-Congo noun class systems and deviance thereof*

In what precedes, we have summarized characteristics of Niger-Congo languages that are admittedly particularly important in the definition of a prototype of a Niger-Congo class system. We now proceed to emphasize features that are particularly relevant from the point of view of a typological comparison.

   It happens that all Niger-Congo noun class systems depart more or less from the ideally simple situation in which a form likely to be involved in agreement would always show an overt (2.4.1) and a non-ambiguous class marker(2.4.2),

---

6. There is a consonant alternation *b/p* with this zero class 8–10.
7. Subject and object markers are written as separate words in the examples according to Tswana orthography. They are in fact affixes.

that would be identical for all nouns sharing the same agreement properties (2.4.3), devoid of allomorphic variations (2.4.4), and obligatory (2.4.5), as considered in turn in the following sections.

2.4.1. *Class is not always overtly indicated in noun morphology.*   This situation arises in two cases: cases of zero class markers and cases of non-obligatory use of the marker. In the former case, for instance, a Proto-Bantu prefix **n-* of class 9 can be reconstructed, but synchronically, in many Bantu languages, no prefix can be isolated in noun forms of class 9 (see example (1i) *Ø-kgosi* 'chief'). In the latter case, in some Niger-Congo languages noun classes are only optionally marked with overt affixes under certain conditions. For instance, still in Tswana, the prefixes of classes 5, 7, 8-10, 11, and 14 become optionally marked on the noun in the presence of a modifier itself marked for the class. This is the case for instance in the presence of the genitive marker *-a*, as illustrated in (5).

(5)   a.   *(le-)saka   l-a   di-kgomo*
             (CL5-)kraal   CL5-GEN   CL10-cattle
             'cattle kraal (cow enclosure)'
       b.   *(di-)kgomo   ts-a   Ø-kgosi*
             (CL10-)cattle   CL10-GEN   CL9-chief
             'the cattle of the chief'

2.4.2. *Two nouns with different agreement properties may show phonologically identical class markers.*   In such cases descriptions of Niger-Congo languages generally assume that the two nouns belong to distinct classes with homonymous class markers. This situation may be illustrated by Bantu classes 1 and 3, as shown by examples (1a) and (1c), repeated here:

(1)   a.   class 1   *mo-sadi <u>yo</u> mo-s'a*   'new woman'
         c.   class 3   *mo-lemo <u>o</u> mo-s'a*   'new medicine'

This situation is important for the evolution of noun class systems, since the homophony of noun affixes may lead to confusion in their agreement properties, and eventually to total confusion between two classes originally distinct. This is particularly true in cases when two or more classes include nouns with zero marking.

2.4.3. *Two nouns with different class markers may have exactly the same agreement properties.*   Linguists describing Niger-Congo languages, when faced with such situations, do not always offer the same solution. They sometimes assume that two classes may be distinguished only in noun morphology. But this questions the very notion of noun class, and may lead, practically, to

a needless proliferation of "classes", since it makes it possible to identify as distinct class markers allomorphs not readily reducible to the same abstract phonological form. Another solution is to assume that nouns with the same agreement properties but with different class markers belong to two sub-classes of the same class. For example, descriptions of Bantu languages traditionally recognize a subclass 1a of class 1 in which nouns have exactly the same agreement properties as the other nouns of class 1, but, unlike them, do not combine with a class marker.

2.4.4.   *Morphemes involved in the same agreement pattern are far from being always phonologically uniform.*   Contrary to what handbook examples that heavily emphasize the regularity of Niger-Congo agreement systems may suggest, morphemes involved in the same agreement pattern (i.e., morphemes that indicate the same class but attach to words of different syntactic status – such as nouns, noun modifiers, pronouns, or verbs) are far from being always phonologically uniform. For instance, in Tswana, next to regular patterns one can also find phonologically heterogeneous sets of markers, as in examples of (1) repeated below: uniform ones as in (1b, e, g, j), but non-uniform ones in (1a, c, d, f, h, i, k, l):[8]

(1)

		[CLx-N	link.CLx	CLx-Adj]	
	uniform				
b.	class 2	*ba-sadi*	*ba*	*ba-s'a*	'new women'
e.	class 5	*le-saka*	*le*	*le-s'a*	'new cattle kraal'
g.	class 7	*se-kolo*	*se*	*se-s'a*	'new school'
j.	class 11	*lo-kwalo*	*lo*	*lo-s'a*	'new book'
	non-uniform				
a.	class 1	*mo-sadi*	*yo*	*mo-s'a*	'new woman'
c.	class 3	*mo-lemo*	*o*	*mo-s'a*	'new medicine'
d.	class 4	*me-lemo*	*e*	*me-s'a*	'new medicines'
f.	class 6	*ma-raka*	*a*	*ma-s'a*	'new cattle kraals'
h.	class 8–10	*di-kolo*	*tse*	*din-ts'ha*	'new schools'
		di-kgosi	*tse*	*din-ts'ha*	'new chiefs'
i.	class 9	*Ø-kgosi*	*e*	*n-ts'ha*	'new chief'

8. The cases of non-uniformity of forms are found with the linkers and the demonstrative from which they derive, but never with the adjectives.

k.	class 14	*bo-jang*	*jo*	*bo-s'a*	'new grass'
l.	class 15–17	*go-lema*	*mo*	*go-s'a*	'new way of cultivating'

While this situation may be the result of a confusion of classes that were origi-nally distinct, it is important to emphasize that, synchronically, agreement rules apply with great regularity, irrespective of the degree of phonological irregu-larity of the sets of markers involved.

2.4.5.   *The presence of class markers in certain contexts may be subject to discourse conditions.*   The types of situations in which a noun may not carry an overt mark of class have been considered in Section 2.4.1 above: zero class marker and optional use in the presence of class-marked modifier. In terms of obligatoriness of agreement on modifiers and predicates several situations obtain. There are languages in which agreement is obligatory even in cases when the class to which the noun belongs is not overtly indicated on the noun itself. This situation corresponds with the Tswana examples (1i) above, for instance. There is also the case of languages in which some of the modifiers do not show agreement with the noun they modify while others do. For example, adjectives may divide into a subset of adjectives that show agreement with the noun they modify and a subset of invariable ones. As a rule, however, once a modifier has the ability of showing agreement with the noun, agreement is obligatory. There seems to be no case of optional modifier agreement in those languages.

At the level of argument indexation on verbs, subject indexation is obligatory in most Niger-Congo languages that have noun class systems. On the other hand, object indexation is generally more discourse dependent and applies only if the argument corresponding to the syntactic function object is topicalized and not represented in the same clause by a noun phrase. The contrast in the treatment of subject and object class markers is illustrated with the Tswana examples in (6):

(6)   a.   *Mo-sadi   o   lapile*
              CL1-woman   SM.CL1   be.tired
              'The woman is tired.'
        a'.   *O   lapile*
              SM.CL1   be.tired
              'She is tired.'
        a''.   **Mosadi lapile*
        b.   *Ke   thusitse   mo-sadi*
              SM.1SG   help.TAM   CL1-woman
              'I helped the woman.'

b′. *Ke   mo   thusitse*
    SM.1SG OM.CL1 help.TAM
    'I helped her.'
b″. **Ke mo thusitse mo-sadi*
    'I helped the woman.'[9]

    The main point of this survey was to underline the degree of variability of the systems and the numerous non-prototypical features of many of them in contrast to the impression of a more canonical functioning of those systems created in some of the general linguistic literature and in the African linguistics literature with more pedagogical than descriptive aims. In addition, the formal aspects of Niger-Congo noun class systems considered here were chosen specifically for their relevance in the comparison with Amazonian systems to be proposed later.

2.5. *Issues in the semantics of Niger-Congo classification systems*

The question of the semantic motivation of the division of nouns into genders is rather controversial in Niger-Congo studies. Elementary presentations have popularized the idea that Niger-Congo classes have very transparent meanings, on the basis of the reconstruction of a noun class system in Proto-Bantu (Richardson 1967, Givón 1970, Creider 1975, reconsidered in Denny & Creider 1986, for instance) for which the semantic motivation was claimed to be based mainly on shape and configurational meanings. Things are not so simple, however, and on the whole, the question of the semantic basis of Niger-Congo classes is not very different from that of the semantic basis of Indo-European gender. The fact is that the assignment of individual nouns to genders in today's languages is clearly neither random nor predictable on the basis of fairly obvious categories such as "animals", "trees", "body parts", etc.

    The discussion of the degree of semantic motivation of Niger-Congo classification systems presented here will cover cases of lack of such motivation (Section 2.5.1), partial motivation in loanwords (2.5.2) and in some derivational uses of the class markers (2.5.3), as well as in the special case of an "absolute" anaphoric use of some nouns (2.5.4).

2.5.1. *Lack of semantic transparency.*   In most noun class languages of the Niger-Congo family, the only semantic distinction which is obviously reflected in noun classes is [±human]. The tendency of nouns sharing certain other semantic features to group into certain genders is undeniable, but it is not easy

---

9. This sentence, however, would be perfectly grammatical, with appropriate changes in the tonal structure, if interpreted as a case of right dislocation of the object, as in *I have helped her, the woman*.

exactly to determine the relevant semantic features, since these features often turn out to be fairly abstract, and the relationship between genders and semantic features is rarely absolute, but rather of a statistical nature. It must be kept in mind that Niger-Congo noun class systems have been grammaticalized a very long time ago, and that evolutions of the kind described above in Section 2.4 (especially 2.4.2, viz. nouns with different agreement properties showing phonological identical class markers, and 2.4.3, viz. noun with different class markers having the same agreement properties), may have distorted the original relationship (if any) between classes and meanings. The difficulty in determining to which extent the assignment of nouns to classes is predictable on a semantic basis must not conceal, however, the fact that some aspects of these systems reveal very clear links between certain class markers and certain semantic features, to be considered below.

2.5.2. *Semantic motivation vs. arbitrariness in the treatment of loanwords.* It is common for the phenomenon of lexical borrowing to reveal some aspect of the functioning of the host language. This can be the case with borrowed nouns in Bantu languages with noun class systems that are sometimes assigned to a class on a semantic basis. For example, in Tswana, *mo-fine* 'wine' (borrowed from Dutch) entered class 3; this is clearly neither a phonologically motivated assignment nor a default assignment, and the only possible explanation is that class 3 includes nouns of other psychoactive substances such as *mo-re* 'medicine', *mo-tsoko* 'tobacco', *mo-tokwane* 'marijuana'.

More often class assignment of borrowed nouns proceeds on a morphological basis. In a language with class prefixes, for example, the initial syllable of a borrowed noun that is phonologically similar to some class prefix may be reanalyzed as representing the prefix in question, or, if the language has a noun class in which nouns are devoid of overt class marker, borrowed nouns may be assigned to this class by default.

2.5.3. *Semantic motivation in the "derivational" use of class markers.* A striking feature of the Bantu noun class systems is the blurring of the traditional distinction between derivation and inflection.[10] Of interest here is that the derivational use of class markers, by nature, can reveal semantic motivations of class assignment which are not immediately apparent in the class assignment of non-derived noun stems. As already noted at the beginning of this section, the most obvious function of commutations between class markers is to express number, which is admittedly a typically inflectional distinction.

---

10. For a detailed discussion of the status of Bantu class markers as derivational or inflectional, see Mufwene (1980).

When noun stems have the ability of combining with more than two class markers, the meanings involved are typically derivational and refer to different semantic domains such as 'tree' vs. 'fruit', 'individual' vs. 'collective', 'concrete' vs. 'abstract', 'diminutive', etc. Tswana examples in (7) show the derivational use of class markers in combination with the noun stem *sadi* 'woman' beyond the use of a class marker to mark plurality (CL2):

(7) a. class 1   *mo-sadi*   'woman'
         class 2   *ba-sadi*   'women'
         class 7   *se-sadi*   'feminine behavior'
         class 9   *Ø-tshadi*   'group of women'
         class 11   *lo-sadi*   'group of women'
         class 14   *bo-sadi*   'womanhood'
     b. class 3   *mo-retlwa*   'tree of the species moretlwa'
         class 2   *me-retlwa*   'trees of the species moretlwa'
         class 9   *Ø-thetlwa*   'fruit of the moretlwa tree'
         class 10   *di-thetlwa* 'fruits of the moretlwa tree'
         class 11   *lo-retlwa*   'thicket of moretlwa trees'

Examples in (8) on the other hand show the relative regularity of the semantic motivation of class markers used with derivational function:

(8) a. *mo-sadi*   'woman'   / *bo-sadi*   'womanhood'
                      (class 1)   (class 14)
         mo-nna   'man'   / *bo-nna*   'manhood'
                      (class 1)   (class 14)
         mo-loi   'witch'   / *bo-loi*   'witchcraft'
                      (class 1)   (class 14)
     b. *mo-retlwa* 'moretlwa tree' / *Ø-thetlwa* 'moretlwa fruit'
                      (class 3)   (class 9)
         mo-rula   'morula tree'   / *Ø-thula*   'morula fruit'
                      (class 3)   (class 9)
         m-milo   'mmilo tree'   / *Ø-pilo*   'mmilo fruit'
                      (class 3)   (class 9)

Obvious semantic regularities can also be observed in the assignment of deverbal nouns to classes, or in the substantivization of adjectives, as illustrated in the Tswana examples (9), exhibiting the parallel use of class markers expressing derivational meanings in combination with noun stems (a) and adjectival stems (b).

(9) a. *mo-sadi* / *ba-sadi* (class 1 / class 2) 'woman / women'
         bo-sadi (class 14)   'womanhood'
         se-sadi (class 7)   'feminine behavior'

b.	*-ntle*	'good, beautiful'
		(adjectival stem)
	bo-ntle (class 14)	'goodness, beauty'
	se-ntle (class 7)	'well'

2.5.4. *The "absolute" use of modifiers.*   Generally, it is not uncommon that words typically used to modify nouns or to anaphorically refer to some noun can also have "absolute" uses in which they are neither syntactically nor anaphorically linked to any noun. In this absolute use, the class prefix itself clearly conveys a meaning (example 10), whereas in the modifying or anaphoric use of such words, the choice of the class marker is simply determined by the class to which the head noun (or antecedent, respectively) belongs.

Tswana examples in (10) contrast the absolute use of modifiers of classes 1/2 for persons with those of classes 7/8 for inanimate things.

(10) a. *mo-ngwe*   /   *se-ngwe*
          CL1-one   CL7-one
          'someone'   'something'

      b. *o-pe*   /   *se-pe*
          CL1-not any   CL7-not any
          'nobody'   'nothing'

      c. *b-a*   *motse* /   *ts-a*   *motse*
          CL2-GEN village   CL8-GEN village
          'villagers'   'things concerning the village'

The absolute use of modifiers where the prefixes carry general meanings such as 'person', 'thing', 'place', or 'manner' is an important fact for typological comparison. It is a discourse phenomenon rarely mentioned in descriptive grammars of African languages, in which the class marker of a word typically used as a modifier is interpreted as referring to a particular noun, even if it is impossible to establish an anaphoric relation with a previous occurrence of this noun.[11]

2.6.   *Origins and stage of grammaticalization of Niger-Congo class systems*

In situating the Niger-Congo class systems from the point of view of grammaticalization processes three parameters will be considered: (i) the age of the

---

11. For example, in Tswana, class markers of class 5 can have a default interpretation as referring to *letsatsi* 'day' when it is not possible to establish an anaphoric link with some other noun of class 5 in the context or situation: instead of *ka letsatsi le le latelang* lit. 'on the day that follows', it is possible to say simply *ka le le latelang*, lit. 'on the one (class 5) that follows', and if the context does not suggest an anaphoric link with some other class 5 noun, the hearer will immediately interpret this as 'on the following day'.

system, (ii) the lexical origin of the class markers, and (iii) the degree of grammaticalization of the systems. They are important parameters that will be taken into account in the final comparison between African and Amazonian classification systems.

2.6.1. *Ancient origin of the systems.*   All the available evidence points to an old age of Niger-Congo classification systems. The classification system reconstructed for Proto-Bantu is very similar, in all respects, to the systems of many modern Bantu languages. As for Niger-Congo, there is so far no real reconstruction of a proto-language, but the other branches of Niger-Congo do not seem to provide any evidence supporting the reconstruction of a less grammaticalized noun class system at the level of Niger-Congo languages. Noun class systems that are somehow "incomplete" in comparison with the Bantu prototype are very common in various branches of Niger-Congo. However, it is clear that these systems are not emerging class systems, but rather the result of the disintegration of former systems of the Bantu type.

2.6.2. *No evidence for a lexical origin.*   In particular, Niger-Congo languages do not provide any concrete evidence for the hypothesis of a lexical origin of class markers. This particular feature will be shown to be in absolute contrast with the situation of the Amazonian systems in general.

2.6.3. *Advanced stage of grammaticalization.*   On the other hand, Niger-Congo languages provide abundant illustration of processes that noun classification systems at an advanced stage of grammaticalization typically undergo, in particular (i) the renewal of class morphology by agglutination of former determiners to nouns, (ii) the disintegration of systems of agreement rules,[12] and (iii) the evolution of highly grammaticalized noun class systems with a relatively high number of genders towards two gender systems transparently based on the [±human] or [±animate] distinctions.

This last process has been described for the concord system of Swahili, in which nouns referring to animates still show their etymological prefixes but obey concord rules that in most cases are the same for all animate nouns, irrespective of their class prefixes. Additionally, it seems that even in the most conservative concord systems, the concord between coordinated NPs in subject function and the verb tends to function on a purely semantic basis (Heine 1982; cf. also Wurzel 1986).

---

12. See for instance Osam (1993) on the loss of the noun class system in Akan and its survival only in "irregular" plural agreement markers in certain adjectives.

Noun class systems of Niger-Congo languages do not seem, however, to have preserved any trace of stages of evolution in which they would have been characterized by a lesser degree of grammaticalization than the one at which they have been reconstructed in Proto-Bantu.

2.7.   *Key points for a comparison with Amazonian classification systems*

This review of the characteristics of the Niger-Congo class systems was meant to highlight two major points that will constitute the basis for a comparison with Amazonian systems. The first one was cast in a sociolinguistic perspective and underlined the effect of the projection in much of the literature of an image of a prototype system that is characterized by much regularity. Meanwhile, the reality of the spoken forms of these languages, away from the standardizing goals of much of the written production, produced in situ and widely echoed in the general linguistics literature used by non-experts of those languages, is that there is great variation from this prototypical image across the languages. Irregularities, both in forms and in functioning, are in fact quite common.

The second point was that there is no doubt about the very old age of the African systems, as demonstrated by their widespread lack of semantic motivation (beyond some limited instances of mostly derivational use) and the impossibility of identifying lexical sources for the class markers. Furthermore, the systems are old enough that some exist today in a number of languages in varying degrees of decomposition, in some cases to such an advanced stage that only a few traces are left.[13]

3.   The challenge of Amazonian nominal classification systems

Unlike the relatively well-known and amply described situation with the nominal classification systems of African languages, the description of the profusion of nominal classification systems found in the languages of the Amazon region of South America is still a matter of exploration and typological debate. While the Niger-Congo systems have provided the data for the establishment of prototypical noun class systems, the discussion around the Amazonian systems has evolved around characterizing them in various non-typical ways, as different concordial systems or different classifier systems.

---

13. This issue of language decay, as it were, seems to also have an interesting sociolinguistic overtone in the case of the discussion of African languages. On several occasions the Amerindianist linguist (Grinevald Craig) was told that one does not talk of "decay" of African classification systems. In much the same way one cannot really talk of the iconicity of the American Sign Language, while it is the basis of most of the research on French Sign Language (LSF, Langue des Signes Française). One should never lose track of the sociopolitical contexts of academic research: in the two instances mentioned here, postcolonialism in one case, civil and linguistic rights movements in the other.

3.1.   *The general situation of the Amazonian languages and of Amazonian linguistics*

Amazonian languages are distributed over nine countries where today they are largely found in remote border areas, and often across national borders. They tend to be the languages of small communities, the average community having between 150 and 300 speakers, and several thousand speakers being an unusually large number. Establishing the number of languages is still difficult because of remaining unknowns and of classification issues, but Dixon & Aikhenvald (eds.) (1999) talk of over 300 languages belonging to about 20 families.[14] A characteristic of this region is that a disproportionate number of these languages are isolates or belong to families with only a handful of members.[15] The Amazonian basin remains today one of the last linguistic "black boxes" of the world, with spotty and still limited knowledge, although it has been the object of increasing attention in the last decades, as witnessed by recent collective publications accessible to general linguists (Payne (ed.) 1990; Derbyshire & Pullum (eds.) 1986/1990/1991/1998; Dixon & Aikhenvald (eds.) 1999).[16]

What must be clear is that, in comparison to the situation of African languages, there is by and large no written or writing tradition in those languages, and therefore no general textbooks or reference books, including pedagogical material to which to turn for their study, and no access either to reading material to use as basic text material for building databases. Literacy efforts, in the limited number of communities where they are being started, are all in the making, with limited resources and local scope. Amazonian linguistics is therefore generally still in a pioneer state, and fieldworkers are often dealing with possibly nomadic as well as, too often, not only threatened languages but

---

14. A quick overview of the Amazonian languages as endangered languages is given in Grinevald (1997), and a comprehensive survey organized country by country and with detailed maps can be found in Queixalós & Renault-Lescure (eds.) (2000).

15. The proportion of stocks or families to languages is strikingly different between South America and Africa. Nettle & Romaine (2000: 37) give the following comparative figures: for Africa, a large number of languages (1,995) for a relatively average number of stocks (20) and for South America, an average number of languages (419) but an unusually high number of stocks (93). Amazonia is no doubt the single most genetically diversified region of the world. (For comparison, Europe is considered to have 209 languages from 6 stocks.) Adelaar (2000) draws attention to the fact that numerous distinct language families of the region are represented by isolate languages.

16. A new crop of U.S. trained Brazilian linguists, most of them coming out of programs of the Museu Goeldi of Belém, is also emerging and changing the Brazilian fieldwork and academic landscape, among other things in changing the ratio of foreign missionary to non-missionary linguists. Several projects of documentation of endangered languages of Brazil are presently funded by the new Volkswagenstiftung-DOBES and SOAS-ELDP programs.

threatened populations.[17] Therefore one must not forget the challenges of field-work for the documentation of this great variety of yet largely undescribed languages.

In strictly linguistic terms, the field conditions mean working with oral tradition languages in their basic state of extreme variation, on the background of the extreme diversity of the languages – a situation quite different from the one found in Africa. And since Amazonian systems of nominal classification are found across many genetically unrelated languages, the possibilities of reconstructing proto systems are very limited, unlike in the situation of Bantu languages (cf., e.g., Guthrie 1967–1971: Volume 4). The fact that nominal classification is a widespread feature in Amazonia in fact seems to be largely due to areal diffusion rather than to common origin. Documented cases of nominal classification systems affected by areal diffusion include Cubeo (Gomez-Imbert 1996), Tariana, and Resígaro (Aikhenvald 2001).

3.2. *The typological challenge of Amazonian nominal classification systems*

The study of Amazonian languages, though recent and still limited, has turned out to be very productive from a typological point of view. Various aspects of the languages of that region of the world have provoked renewed discussion in on-going debates, such as in the areas of word order (Payne 1990) and ergativity (Gildea 1997). And clearly one of the interesting typological characteristics of that region of the world is the presence of nominal classification systems – at least in the western part of the Amazon. The descriptive challenge they have presented is reflected in repeated claims of originality and non-conformity of those systems when they are taken in the context of then current typological studies of the phenomenon of nominal classification. There have indeed been two types of approaches for handling the large Amazonian nominal classification systems, that can be interpreted as echoes of two phases in the development of a typology of such systems.

In a first phase the descriptions and typologies of Amazonian nominal classification systems took as reference the typology of Allan (1977). This first

---

17. It is probably worth reminding non-specialists of this region of the world that the threat to the physical survival of many of these populations is very high, particularly those that have only recently been contacted or are still to be contacted. If language shift to dominant languages, regional or national (Brazilian Portuguese), is statistically the major cause of language endangerment, the major cause of the alarming levels of deaths in many of these populations is the same as at the time of the early colonization of the whole continent: that of diseases brought in by outsiders. To this, one must add the deadly effect of development projects such as new roads and the advancement of the agricultural frontier on such populations, often less protected still than endangered biological or botanical species. Not to mention multiple guerrilla warfares and transnational wars taking place in territories traditionally occupied by indigenous people.

typology was all-inclusive, lumping together classifier systems and noun class systems in the discussion of their semantic motivation, while setting apart the then better known numeral classifier systems from a broad category of "concordial" systems when considering their morphosyntax. The pioneering articles of Payne (1987) and Derbyshire & Payne (1990)[18] that offered the first overview of the variety and particularities of Amazonian nominal classification systems relied on this early study by Allan (1977). The typological contribution of these articles was the addition to Allan's inventory of another system of "verb-incorporated classifiers", on the basis of the then recent work by Mithun (1986) on such a subtype of classifiers. Derbyshire & Payne's overall analysis of the situation was that "the chief characteristic of most of the Amazonian classification systems [...] is that they cannot be labeled discretely as any one type, but are a mixture of two or all three types" (1990: 243).

A second phase corresponds to more recent typologically oriented studies of Amazonian nominal classification systems that have aimed at showing how they are "unusual and complicated". The point of reference taken at that time was a typology meant to establish the existence of various subtypes of classifier systems, using in particular as one of its arguments the co-existence in some languages of two distinct systems (Craig 1987, further developed in Grinevald 2000). The particularity of the Amazonian systems being described on the basis of this typology was the unusual and yet undocumented claim of the co-existence of up to five distinct systems in one and the same language (Aikhenvald 1994, Aikhenvald & Green 1998).

The major difficulty encountered in the typological characterization of many Amazonian systems of nominal classification indeed lies in the apparent profusion of classifying morphemes in a variety of different morphosyntactic loci where they fulfill a variety of different functions. The picture is complicated by a certain amount of overlap of the inventories of such morphemes in different morphosyntactic contexts. When the overlap between the different sets of markers is large enough, one could probably speak of a single system with exceptions, raising the question of whether or not the markers constitute an agreement system of the "concordial" or noun class type. It is only when the various co-existing sets are too distinct that it could be said that different systems co-occur in the language, corresponding to a situation of "multiple classifier language". The crucial question of analysis is therefore whether certain complex nominal classification systems of the Amazon region should be char-

---

18. Made in part possible by their access to unpublished files of SIL colleagues working in the region and their coordinating initiatives to bring out information on Amazonian languages. See the collection of articles in Payne (ed.) (1990), results of efforts of a month-long workshop on Amazonian languages at the University of Oregon supported by the National Science Foundation.

acterized as (concordial) noun class systems or as co-occurring distinct classifier systems. The answer of course cannot be a clear-cut one and has to rely on fine-grained and comprehensive descriptions of individual systems, while allowing for any number of language specific arrangements, and possible continua and overlap. Ongoing research by Corbett and colleagues (Corbett 2003a, b, forthcoming), who argue for a scalar notion of agreement, addresses some of the same issues and may offer converging proposals with the one to be offered here to further refine typologies of systems of nominal classification.

3.3. *A typology of nominal classification systems*

Grinevald (2000) offers a morphosyntactic typology[19] of systems of nominal classification systems meant to identify classifier systems as distinct from other types of nominal classification systems, and to establish the need to recognize different subtypes of classifier systems.

In separating classifier systems from other systems, the typology, as shown in (11), takes into account more lexical systems as well as more grammaticalized systems.

(11)   Systems of nominal classification

       ◄ ···· lexical ············ lexico-grammatical ············ grammatical ···►

         class-terms   CLASSIFIERS   noun classes/gender

         measure terms

On the lexical end, two types of systems are identified, "class terms" and "measure terms", that do not constitute morphosyntactic systems although they do produce a classificatory effect. English examples of class terms would be the term *berry* in *blueberry*, *strawberry*, *boysenberry*, etc., and the measure term *pound* in *a pound of butter*, *a pound of sugar*, *a pound of oranges*, etc. At the more grammaticalized end of the spectrum of nominal classification systems, classifiers are distinguished from noun class systems and gender systems characterized by definition as agreement systems. The standard examples of such systems are the noun class systems of the Bantu type (those discussed in Section 2 above) and gender systems like those of European languages.[20]

---

19. The justification for such an approach to the task of setting up a typology resided mostly in the reality of fieldwork and the assurance that morphosyntax is the part of language most accessible to field linguists, and most reliably describable. It was meant to be a tool offering a descriptive strategy to encourage more systematic and comparable descriptions.

20. In Corbett's (1991) treatment of gender, noun class and gender are united in that they are both realized in agreement classes. Data from Amazonian languages shows that various such agreement system can co-occur in one language, with distinct loci and morphological inventories. Among these systems, there is usually one reminiscent of what has traditionally been called gender and one that is reminiscent of noun classes.

The particular type of nominal classification system labeled "classifiers" is taken to occupy the mid point of this lexical-grammatical continuum, and to include, within the specific type of "classifiers", different classifier subtypes. These subtypes of classifiers are identified first on the basis of their different loci, which in fact defines their name, as sketched out in (12), in that "genitive classifiers" are those occurring in possessive constructions, "numeral classifiers" those occurring in numeral phrases,[21] "noun classifiers" those standing next to a noun, and "verbal classifiers" those found within the verb forms. All, except this last type, occur within noun phrases.

(12)   Classifier systems
      [POSS+CL   Numeral+CL   CL+Noun]   //   Verb+CL
      genitive   numeral   noun   verbal
      classifier   classifier   classifier   classifier

Note that noun classifiers are a subtype of classifiers and, as such, constitute a phenomenon distinct from noun classes. Illustrative examples of each type are given in (13)–(16).[22]

(13)   Noun classifiers
      Jakaltek-Popti' (Mayan; Craig 1986: 264)
      xil   naj   xuwan   no7   lab'a
      saw   CL   John   CL   snake
      '(man) John saw the (animal) snake'

(14)   Numeral classifiers
      Ponapean (Micronesian; Rehg 1981: 130)
      a.   *pwihk   riemen*
          pig   two.CL(ANIMATE)
          'two pigs'
      b.   *tuhke   rioapwoat*
          tree   two.CL(LONG)
          'two trees'

(15)   Genitive classifiers
      Ponapean (Micronesian; Rehg 1981: 184)
      a.   *kene-i   mwenge*
          CL-GEN.1   food
          'my(edible) food'

---

21. It has been noted that in some languages the so-called numeral classifiers can also appear on demonstratives and/or adjectives.

22. Partly on the basis of Amazonian data, the existence of other minor types, such as locative classifiers has been claimed by Aikhenvald (2000).

      b. *were-i   pwoht*
          CL-GEN.1   boat
          'my(transport) boat'

(16)   Verbal classifiers
         Cayuga (Iroquoian; Mithun 1986: 386–388)
      a. *ohon'atatke: ak-hon'at-a:k*
          it.potato.rotten PAST.1SG-CL-eat
          'I ate a rotten potato.'
      b. *so:wa:s akh-nahskw-ae'*
          dog   1SG-CL(DOMESTIC.ANIMAL)-have
          'I have a dog.'
         Gunwinggu (Australia; Oates 1964, in Mithun 1986: 389)
      c. *gugu ga-bo:-mangan*
          water it-CL(LIQUID)-fall
          'Water is falling.'

Arguments for establishing different subtypes of classifiers therefore rely on a combination of the morphosyntactic ground of different loci with the existence of distinct inventories of classifiers in those languages in which two clearly distinct systems co-exist, as shown with the examples of numeral (14) and genitival (15) classifiers of Ponapean.[23] In addition, the different subtypes of classifiers exhibit different kinds of semantic distinctions as well, a tendency particularly noticeable in languages with multiple classifier systems. Numeral classifiers for instance systematically include some classification by physical characteristics such as shape (1D long-rigid, 2D flat-flexible, 3D round being the main ones), in contrast with the almost exclusive functional semantics of genitive classifiers (clothing, food, transport, etc.) and the generic semantics of noun classifier systems (men, women, animals, plants, etc.). The typical semantic alignments of those three types of classifier systems are sketched out in (17):[24]

(17)   Classifier types and preferred semantics
      a.   numeral classifiers = physical categories
           two-ROUND oranges; three-LONG RIGID pencils;
           four-FLAT FLEXIBLE blankets

---

23. An example of a language with multiple classifier systems is found in the description of a Q'anjob'alan Mayan language by Zavala (2000).

24. See Grinevald (2000: 71–74) for results of a comparative study of fifteen classifier systems supporting this claim. The claim is not one of rigid alignment of semantics and classifier types, and it is only a statistical tendency in languages with large systems of numeral classifiers for instance, ones that are characterized by varied semantics. But there is a striking division of labor between well-defined systems co-occurring in the same language and a definite specialization in function for the genitive type.

    b.   genitive classifiers = functional categories
        my-EDIBLE food; his-DRINKABLE potion;
        their-TRANSPORT canoe
    c.   noun classifiers = material/essence categories
        an ANIMAL deer; the ROCK cave; MAN musician

More recent versions of this typology (Grinevald 2001, 2002) insist on the importance of taking into account the dynamic dimensions of the systems in the building of a typology of the phenomenon of nominal classification, taking into consideration multiple and independent variables, such as their age, their vitality, and their degree of grammaticalization. The parameter of grammaticalization is particularly central to the discussion at hand because it was presented from the start (Dixon 1982, 1986) as a determining factor for distinguishing noun class systems from classifier systems, classifiers being considered (in their prototypical form at least, which tends to correspond to the form of the longer and better known numeral classifier systems of South East Asia) as less grammaticalized than noun class systems. On the other hand one can find systems at different stages of grammaticalization within one distinct type of system, as discussed in Grinevald (2002) for the case of noun classifier systems, with Australian ones spanning from minimally to strongly grammaticalized systems, and Q'anjob'alan Mayan ones being unusually extremely grammaticalized, beginning to resemble gender systems in parts of their functioning. The same variation in grammaticalization levels within the numeral classifier type can be noted, as between discourse sensitive numeral classifier systems of languages of South East Asia and very grammaticalized numeral classifier systems of Central America Chibchan languages, for instance.

The proposal of this paper therefore is to approach the Amazonian systems of nominal classification within the typological framework just outlined, and admittedly with the *a priori* assumption that they should be considered as more "normal" than aberrant and a legitimate place within a typology of nominal classification systems of the languages of the world should be found for them. This means that rather than settling for considering many Amazonian systems as being beyond the typology outlined above, on the basis of their being "mixed" systems with characteristics of "classifier" systems and "concordial systems", the working hypothesis is that many of them are in fact noun class systems but with a much lesser degree of grammaticalization than the Niger-Congo noun class systems.[25] It will be suggested therefore that Amazo-

---

25. It is not that it had not been noted earlier that some Amazonian systems of nominal classification have something in common with Bantu noun class systems (see Payne 1986 for an early example), but the aim of the earlier discussions was more to emphasize the supposed differences between the systems in order to establish the uniqueness of the Amazonian patterns than to make sense of their similarities and differences within a unified typological model.

nian and Niger-Congo noun class systems may indeed have more in common than is generally expected, once one looks beyond the regularized picture of the Niger-Congo systems projected in the literature, and into the specifics of their multiple irregularities and variations. A case study of a system of nominal classification considered illustrative of the complexity and "mixedness" of such systems as commonly found in Amazonian languages will now demonstrate the usefulness of the typological approach presented here.

4. The case of Miraña

The presentation of the Miraña data will take up the various themes from the discussion of characteristics of the Niger-Congo noun class systems in Section 2 and of the classifier systems in Section 3. It will cover the basic morphosyntactic properties of this system which invite a comparison with noun class systems (Sections 4.1–4.3), and the semantic and discourse properties which more closely resemble traits that have been attributed to classifier systems (4.4–4.6). A final section (4.7) places this description of the Miraña system in the context of other published accounts of Amazonian languages, in particular the neighboring Eastern Tucanoan languages, in order to reveal the areal vs. language specific features of this system.

4.1. *A typical Amazonian language of Colombia with a typical Amazonian "concordial" classification system*

Miraña is an endangered language spoken in the Colombian Amazon region near the Brazilian border; it has about 100 speakers left, none of them children, from a population of about 400 Miraña. Along with its dialectal variant Bora (cf. Thiesen & Weber forthcoming), Miraña belongs to the Witotoan language family (Aschmann 1993).

The system of nominal classification in Miraña consists of over 60 class markers, a rather large number reminiscent of classifier systems, which are used as phonologically bound forms on a variety of targets to mark agreement within noun phrases and also for anaphora and cross-referencing, much like noun class systems. Example (18) illustrates four instances of one class marker (-*ːba*, see Table 3, No. 56) as agreement markers on a verb, a numeral, an adjective, and as class marker of a noun. When in combination with nouns the function of the class marker is generally derivational, as will be explained presently.

(18)   *ó-di*   *íhka-ːba*   *tsa-ːba*
       1SG-POSS COP-CL(CONTAINER) one-CL(CONTAINER)
       mʉ́hʉː-ːba   *ʔʉ́βiː-ːba*
       big-CL(CONTAINER) basket-CL(CONTAINER)
       'I have one big basket.'

4.2.   *The derivational use of class markers*

The use of class markers on nouns for derivational purposes was mentioned for Bantu noun class systems (Section 2.5.3 above) in anticipation of a comparison with a similar use in Miraña. While this derivational use is relatively limited in the African systems and generally receives little attention, it is one of the most striking and productive features of the nominal classification system of Miraña. Example (19) illustrates the common process of derivation of denominations of different parts of plants, in this case the ubiquitous banana:[26]

(19)   a.   *tɯ́hɨ*
        banana
        'banana(s)'

      b.   *tɯ́hɨ-ʔo*
        banana-CL(OBLONG)
        'a banana (fruit)'

      c.   *tɯ́hɨ-ko*
        banana-CL(SHAFT)
        'a banana plant'

      d.   *tɯ́hɨ-hɨ*
        banana-CL(DISC)
        'a banana seed'

      e.   *tɯ́hɨ-ʔi*
        banana-CL(BUNCH)
        'a bunch of bananas'

      f.   *tɯ́hɨ-gwa*
        banana-CL(PLANK)
        'one half of a banana'

      g.   *tɯ́hɨ-ʔbábaj*
        banana-CL(BAG)
        'a bag of bananas'

The derivational nature of the class markers is further confirmed by the possibility of successive affixation of different class markers on the same noun, as shown in (20):

(20)   a.   *tɯ́hɨ-kó-ʔáːmɨ*
        banana-CL(SHAFT)-CL(LEAF)
        'a leaf of a banana plant'

      b.   *tɯ́hɨ-ʔó-βíːɯ*
        banana-CL(OBLONG)-CL(CHUNK)
        'a chunk of a banana'

      c.   *tɯ́hɨ-dʒíːhɯ-ro*
        banana-CL(POWDER)-CL(BOTTLE)
        'a bottle of pulverized dried banana'

The following examples illustrate the derivational nature of the class markers in a clear process of lexicogenesis, with nouns derived from other nouns (21a, b) and nouns derived from nominalized verbs (21c, d).

---

26. It is worth noting that the semantics of these class markers is reminiscent of the semantics of numeral classifiers of both the sortal (a–d) and the mensurative kind (e–f).

(21) a. *ájβɛ-hpájko*   b. *kɨ́ːhɨ́ɡwa-hpájko*
          pain-CL(LIQUID)   fire-CL(LIQUID)
          'liquor'   'gasoline'

      c. *kaːnɨ́ɨ-ko*   d. *kaːnɨ́ɨ-i*
          to.pound.NMZ-CL(SHAFT)   to.pound.NMZ-CL(STICK)
          'mortar'   'pestle'

Not all nouns carry a class marker at all times. When bare, they denote non-individuated concepts (22) or generic terms for biological species (19a).

(22) a. *ʔɨ́ɨβi*   b. *ʔɨ́ɨβi-ːba*
          basket   basket-CL(CONTAINER)
          'basket(s)'   'a basket'

Faunal nouns that can be used as a bare, non-individuated form have a collective interpretation, as denoting an unspecified number of individuals (23).

(23) a. *míta-nɛ́*   *nɛː*   *bahɨ́ɨ pajnɛ́-ɾi*
          much/many-CL.INAN mosquito forest inside-LOC
          'There are many mosquitoes in the forest (lit., there is much mosquito in the forest).'
      b. *nɛ-ʔba*   *dí-ʔɨ́mɨ-ɾi*
          mosquito-CL(3D)   2SG.POSS-face-LOC
          'There is a mosquito on your face.'

The individualizing function of the class marker is evidenced by the fact that number marking presupposes the presence of a class marker, which is then followed by a dual or plural marker, as shown in (24).

(24) a. *koːmɨ*
          milpeso.palm
          'milpeso(s)'
      b. *kóːmɨ-ɯ*
          milpeso.palm-CL(ROUND.3D)
          '(a) milpeso palm fruit'
      c. *kóːmɨ-ɨ́-kɯ*
          milpeso.palm-CL(ROUND.3D)-DU
          'two milpeso palm fruits'
      d. *kóːmɨ-ɨ́-ːnɛ*
          milpeso.palm-CL(ROUND.3D)-PL
          'milpeso palm fruits (>2)'

4.3.   *Class markers as agreement markers*

Besides their very productive derivational use, class markers are also used in widespread agreement patterns both within the noun phrase and beyond on the verb.

4.3.1.   *Class markers within the noun phrase.*   Class markers combine with all the modifiers of a noun within a noun phrase. It is noteworthy that this agreement pattern is obligatory, making the system very similar in this aspect to the noun class system of the Bantu language Tswana described in Section 2. Example (18) above has already illustrated agreement on numerals. The following examples illustrate agreement on adjectives (25) and on relative clauses (26):

(25)   a.   *múíhɯ-hɨ*   *kúːmɯ-hɨ*
              big-CL(DISC)   turtle-CL(DISC)
              'a big turtle'

       a′. *múíhɯ-ko*   *pihúí-ko*
              big-CL(SHAFT)   to.fish.NMZ-CL(SHAFT)
              'a big fishing rod'

       b.   *kúːβɛ-gwa*   *boʔdó-gwa*
              dark-CL(PLANK)   paddle-CL(PLANK)
              'a dark paddle'

       b′. *kúːβɛ-hɯ*   *ajnúí-hɯ*
              dark-CL(TUBE)   shooting-CL(TUBE)
              'a dark rifle'

(26)   a.   *boːtá-hɨ*   [*o*   *ɨːtɛ*]-*hɨ*   *káːtɯβɛ́-ʔi*
              button-CL(DISC) [1S.SUB see]-CL(DISC)   fall.down-PRED
              'the button that I saw fell down'

       b.   *ó*   *iːté-ʔí*   *gwatsíʔhɯ-gwa*   [*oː-kɛ*   *ɯ*
              1S   see-PRED   machete-CL(PLANK)   [1S-ACC   2S.SUB
              áhkɯ]-*gwa*
              give]-CL(PLANK)
              'I saw the machete that you gave to me'

Class markers are also found on other modifiers of the noun, such as demonstratives, quantifiers, and possessive. These can be used as actual modifiers next to a noun (27), although their major use is in an anaphoric function where they stand alone as the head of the noun phrase (28):

(27)   a.   *í-gwa*   *gwatsíʔhɯ-gwa*
              DEM.PROX-CL(PLANK)   machete-CL(PLANK)
              'this machete'

   b.   *tsi-gwa*   *gwatsíʔhɯ-gwa*
        another-CL(PLANK)   machete-CL(PLANK)
        'another machete'

(28)   a.   *í-gwa*
            DEM.PROX-CL(PLANK)
            'this one (plank, bench, etc.)'
       a′. *í-hɨ*
            DEM.PROX-CL(DISC)
            'this one (coin, button, etc.)'
       b.   *tsi-gwa*
            another-CL(PLANK)
            'another one (plank, bench, etc.)'
       b′. *tsi-hɨ*
            another-CL(DISC)
            'another one (coin, button, etc.)'
       c.   *tájʔnɛ́-gwa*
            my-CL(PLANK)
            'mine (plank, bench, etc.)'
       c′. *tájʔnɛ́-hɨ*
            my-CL(DISC)
            'mine (coin, button, etc.)'

4.3.2.   *Class markers on verbs.*   Finally, the class markers are used for the
indexing of the subject on predicates in the case of intransitive verbs as well
as transitive verbs, as illustrated in (29).[27] Unlike in Niger-Congo languages,
objects are never cross-referenced by class markers on the verb in Miraña.

(29)   a.   *gwáʔdáɨnɯ-ːbɛ́*   *ahɨ*
            cut-CL.MASC.SG   palm.species
            'He cuts the palm tree.'
       b.   *káːtúβε-gwa*
            fall.down-CL(PLANK)
            'The (plank, etc.) fell down.'

The agreement patterns of the class markers within the noun phrase and as
indexing of subject on the verb are therefore overall rather similar to patterns
described in Section 2 for the noun class system of Niger-Congo languages
such as Tswana.

---

27. When the subject noun phrase immediately precedes the predicate, the class marker may be
    replaced by *-ʔi* (PRED), as in (26).

4.4.   *Discourse sensitivity of class markers*

While the basic agreement function of Miraña class markers outlined above is reminiscent of African systems, a pervasive feature typical of Amazonian systems is their discourse sensitivity. Although some variability in the use of noun class markers was reported in Tswana, as for instance in the case of the discourse sensitive use of object agreement markers (illustrated in (6) of Section 2.4.5), the phenomenon was presented as being rather exceptional in African systems. It is however much more widespread in Miraña, where it takes several forms. The possible contrast between absence vs. presence of class markers in order to indicate a difference between concept/generic nouns vs. individualized participants in discourse chains was one such circumstance already mentioned in Section 4.2 (see examples 19, 22, 23, 24). Two other aspects of the use of class markers in Miraña that characterize the system as discourse sensitive are the possibility of choice between more general and more specific markers (4.4.1) and their widespread anaphoric use (4.4.2), particularly in the phenomenon of the "absolute" use of class markers (4.4.3).

4.4.1.   *Choice of general vs. specific class markers.*   Miraña has in fact two semantic types of class markers, where some more general class markers can stand in for some more specific ones. While the semantics of these two types of class markers will be described in Section 4.5 below, the emphasis here is on the choice of the speaker at the moment of speech, which is reminiscent of numeral classifiers in languages of South East Asia. It needs to be said that the choice between general and specific class markers in Miraña applies only to inanimates and non-human animates. Humans are always class-marked animates (with a distinction of natural gender and number), while animals and objects vary between specific classes of shape or function and general classes of animate or inanimate, as illustrated in (30). Note that the re-classification of nouns according to [±human] or [±animate] distinctions has also been described for Niger-Congo languages (Section 2.6.3).

(30)   a.   *ó-di*   *íhka-ko*   *tsa-ko*
              1SG-POSS   COP-CL(SHAFT)   one-CL(SHAFT)
              pihúú-ko
              to.fish.NMZ-CL(SHAFT)
              'I have one fishing rod.'
        a′.   *ó-di*   *íhka-nɛ*   *tsa-nɛ*
              1SG-POSS   COP-CL.INAN   one-CL.INAN
              pihúú-ko
              to.fish.NMZ-CL(SHAFT)
              'I have one fishing rod.'

b. *kúːmɯ-hɨ*   *áːkítɛ-hɨ*
turtle-CL(DISC)   fall.down-CL(DISC)
'The turtle fell down.'

b'. *kúːmɯ-hɨ*   *áːkítɛ-ːbɛ*
turtle-CL(DISC)   fall.down-CL.MASC.SG
'The turtle fell down.'

This choice between general and specific class markers depends on various factors. It sometimes appears to correlate with the level of formality of speech registers as well as the age of the speaker, both features signaled in the contemporary use of numeral classifiers in languages such as Chinese or Japanese. It is also clearly a matter of discourse factors such as topicality, focus, and saliency, particularly in the widespread anaphoric use of such markers.

4.4.2. *Anaphoric use.* Most of the modifiers in Miraña can be used as the sole anaphoric element of a noun phrase with a class marker affixed to it that establishes the anaphoric link. Both types of class marker can be used in this function, depending on the degree to which the speaker wishes to specify the participant at a given point in discourse. This anaphoric use of class markers can be observed in the following examples which are taken from a text, where a Miraña speaker explains the making of a blowgun, *tódʒiː-hɯ* (blowgun-CL(TUBE)). While the speaker uses the general inanimate class marker *-ne* several times in different kinds of expressions that refer to the blowgun (as in 31), when he wishes to specify the blowgun to a higher degree, he switches to the specific class marker *-hɯ* (CL(TUBE)), in this particular case in combination with the pronominal stem *tɛ-* (as in 31).

(31) a. *tɛ-nɛ*   *ími-nɛ*
         PN-CL.INAN good-CL.INAN
         'It (the blowgun) was good.'

      b. *áːbáhá-hpi-kɛ*   *ó*   *áhkɯ-ːʔi*
         owner-CL.MASC.SG-ACC 1SG give-PRED.FUT

         tɛ́-hɯ-βɯ
         PN-CL(TUBE)-ADL
         'I will give it (the blowgun) to the owner.'

4.4.3. *"Absolute" use.* Like in the case of Niger-Congo languages such as Tswana, Miraña class markers can also have an "absolute" use in which they are neither syntactically nor anaphorically linked to any noun and where the class marker itself clearly conveys a meaning. In the above mentioned story of the blowgun, for instance, when it comes to talking of making a hole in it, this notion is introduced by a construction using the bound pronominal stem *tɛ-* in combination with the class marker *-paːhɨ* (CL(HOLE)), as shown in (32).

Table 1. *Miraña general class markers*

Forms of general class markers	Meaning
-ːbɛ, -di, -hpi, -ːpi	masculine singular
-dʒɛ, -pidʒɛ	feminine singular
-mɯtsi, -tɛtsi	masculine dual
-mɯpɨ, -tɛpɨ	feminine dual
-mɛ, -tɛ, -mɯ	animate plural
-nɛ	inanimate

This construction represents an absolute use, since a participant is introduced by means of a class marker without a full-fledged noun.

(32)   *tɛ́-paːhɨ   mɛ́-kádʒóɨʔkɨ́-ʔi*
      PN-CL(HOLE)   SAP-scrape-PRED
      'One scrapes a hole.'

4.5.   *The inventory of class markers and their semantics*

Compared to African noun class systems, the classes defined by the class markers in Miraña are generally much more semantically motivated, as was already shown in the discussion of the derivational and the absolute uses. There are however varying degrees of semantic transparency among the Miraña class markers. Three subtypes of class markers can be clearly identified, all three also found in the literature on numeral classifiers: general class markers (4.5.1), specific class markers (4.5.2), and repeaters (4.5.3). These three types occur in the same morphosyntactic contexts (on nouns and in agreement positions), but they differ in their agreement function as well as their semantics.

4.5.1.   *General class markers.*   There are six general class markers for animates and inanimates. The animates combine the meanings of animacy, number (singular, dual, and plural), and natural gender (masculine and feminine), as shown in Table 1. They all have allomorphs according to morphosyntactic contexts, including some suppletive forms, and some of them are morphologically complex, characteristics that would point to an older system.

4.5.2.   *Specific class markers.*   The set of specific class markers (over 60) is morphologically and semantically heterogeneous, ranging from monosyllabic forms with relatively broad semantics to polysyllabic ones with very specific semantics. All specific class markers keep the same form in all of the constructions in which they are found, with one exception (No. 5 in Table 2), and they do not express number, which is marked by a separate suffix. These two char-

Table 2. *Miraña specific class markers without a demonstratable nominal origin*

No.	Class markers	Meaning	Examples
1	*-ʔaj*	oval	*tɨ́hkaʔaj* 'beard'
2	*-ʔɛ*	tree, bush, plant	*kóʔhɯʔɛ* 'avocado tree'
3	*-ʔi*	raceme, river	*bɛːʔi* 'bunch of coconuts'
4	*-ʔo*	oblong	*aːʔo* 'maraca fruit'
5	in noun:	a variety of meanings, e.g.,	*nɛ́ːgwajːba* 'stone'
	-ba, -ʔba	thick liquids; fruits;	*maʔníba* 'natural tar'
	in all other	seasons; oblong objects;	
	contexts: -ʔba	musical instruments; etc.	
6	*-dʒi*	a bunch of fibers	*ádʒɨ́ɯɯdʒi* 'eyelashes'
7	*-gwa*	board-like shaped	*boʔdógwa* 'paddle'
8	*-hɨ*	flat, round	*maɾíːmɯhɨ* 'taricaya turtle'
9	*-i*	stick	*koːi* 'wooden stick'
10	*-ko*	shaft	*kaːnɨ́ko* 'mortar'
11	*-hkɛ*	climbing vines	*gwáʔdahkɛ* 'guaya vine'
12	*-hpaj*	liquid, broth	*tɛhpaj* 'broth'
13	*-hto*	pointed	*nɨhɨ́ʔɛhto* 'cumare palm thorn'
14	*-ɾo*	bottle, pipe	*adóɾo* 'drinking bottle'
15	*-ɯ*	round, three dimensional	*íːʔúɯ* 'egg'
16	*-ʔahkɯ*	trunk	*tɨ́maʔáhkɯ* 'tree trunk'
17	*-ʔahtsɨ*	clearing	*dʒáʔahtsɨ* 'patio'
18	*-ʔaːmɨ*	leaf	*gwahákuʔáːmɨ* 'book'
19	*-ʔbabaj*	bag	*gwaráhkoʔbábaj* 'bag to sieve coca'
20	*-ʔɛhu*	hole	*tɯ́ːʔɛhɯ* 'nostril'
21	*-ʔiːgwa*	open space on a riverbank	*pájhtɛʔíːgwa* 'port on a river'
22	*-ʔiːba*	small palm tree	*tóːkɛʔíːba* 'small palm tree, sp.'
23	*-ʔoːba*	short gable	*háʔoːba* 'short gable of a house'
24	*-βiːɯ*	chunk	*kóʔbavíːɯ* 'chunk of wood'
25	*-dʒiːʔo*	bud	*mɛ́mɛ́ʔɛdʒíːʔo* 'bud of chontaduro palm'
26	*-dʒiːhɯ*	powder	*bádʒiːhu* 'ash, powder'
27	*-iːʔo*	little stick	*kaːtɨ́nɯíːʔo* 'pencil'
28	*-iːhɨ*	platform	*gwáiːhɨ* 'raft, floor'
29	*-gwajːʔɛ*	crumbs	*íːbɨ́ʔgwájːʔɛ* 'leftover of coca'
30	*-gwajːʔo*	bundle of sticks	*íːbɨ́ʔgwájːʔo* 'bundle of coca sticks'
31	*-hɯːʔo*	palm leaf	*nihiʔɛhɯ́ːʔo* 'cumare palm leaf'
32	*-kaha*	swamp, creek	*ínɛkáha* 'flooded canangucho plantation'
33	*-mɨha*	wooden stick for illumination	*kógwáʔamɨha* 'copaiba stick for illumination'
34	*-miːʔɛ*	skin, soft shell	*báhimiːʔɛ* 'ray fish skin'
35	*-miːʔo*	hard shell	*kɯ́ːmɨ́himiːʔo* 'charapa turtle shell'
36	*-niːʔo*	female with offspring	*ókáhiníːʔo* 'female tapir with offspring'

Table 2 (cont.). *Miraña specific class markers without a demonstratable nominal origin*

No.	Class markers	Meaning	Examples
37	*-hpajko*	liquid	*ájβɛhpájko* 'liquor'
38	*-htsɯːʔo*	pack	*tɯ́mɛhtsɯ́ːʔo* 'pack of salt'
39	*-padʒi*	planted field	*tɯ́miʔɛ́pádʒi* 'a planted field'
40	*-pahtsɨ*	roll, ring	*móʔopáhtsɨ* 'a roll of liana'
41	*-paːhi*	cave, hole	*tɯ́méʔɛpáːhi* 'a hole in a tree'
42	*-tohko*	corner	*hátohko* 'a corner of a house'
43	*-tsoːʔo*	medium-sized palm tree	*ínéʔɛtsóːʔo* 'medium-sized canangucho palm tree'
44	*-roʔdʒo*	completely twisted	*dɯ́ːrɯ́iɾóʔdʒo* 'a completely twisted candle'
45	*-ɯʔo*	basin	*kaːnɯ́nɯ́ʔo* 'basin for pounding fruit'
46	*-ɯːʔaj*	grains	*nɛ́gwajɯ́ːʔaj* 'sand'
47	*-mɛhkɛi*	thin part	*tɯ́méʔɛ́mɛhkɛ́i* 'thin part of a tree'
48	*-tsaːragwa*	fibers sticking out	*níːgwatɯ́tsáːragwa* 'head with hair sticking out'

Table 3. *Miraña specific class markers with nominal origin*

No.	Class markers	Meaning	Examples	Corresponding nouns
49	*-ʔbaɯ*	round protuberance	*téʔbaɯ* 'hill'	*í-ʔbaɯ* (POSS.3SG-belly) 'his/her belly'
50	*-ha*	cover, shelter	*káːmɛɛha* 'shirt'	*ha* 'house'
51	*-hɯ*	tube, story, song	*bájnɛhu* 'cigarette'	*í-ʔhu* (POSS.3SG-mouth) 'his/her mouth, story, language'
52	*-mɨ*	means of transport	*kɯ́ːhɯ́gwamɨ* 'motorboat'	*miːne* 'canoe'
53	*-mo*	big river	*ókáhimo* 'Caquetá river' (= river of the tapir)	*moːaj* 'big river'
54	*-ta*	metal	*méʔóhtsigwámiːʔóné imítʃota* 'nail clipper'	*ráːta* 'tin can'
55	*-tɯhkɛ*	stem of a fruit or a leaf	*mɯ́tsiːtsibatɯ́hkɛ* 'stem of a pear apple fruit'	*tuhkénɯ* 'beginning'
56	*-ːba*	deep pool, container	*adóːba* 'salty well' *gwájːba* 'hammock'	*báneːba* 'pool'

acteristics distinguish them from the noun class systems of African languages, which can have different agreement markers for the same class and very generally combine class and number. Specific class markers are found on inanimate nouns as well as non-human animates.

The inventory of specific markers of Miraña is rather typical of inventories of nominal classification systems of the region in its size, semantic transparency, and for the fact that for a fair number of them, a nominal origin can be identified. Table 2 contains the class markers for which a nominal origin cannot be demonstrated today, along with an example of a noun that includes this class marker. Table 3 contains the class markers for which a nominal origin can be identified in the language today. These class markers share some phonological material with their source nouns, and are clearly semantically related, but generally have a broader meaning now.

The specific classes are generally highly motivated semantically, that is they contain only nouns denoting a very specific kind of object, e.g., bags (No. 19), bundles of sticks (No. 30), or even only medium-sized palm trees (No. 43). For most classes, at least a prototypical semantic core can be identified, which in some cases extends to other, related meanings. The class determined by *-ʔɛ* (CL(TREE); No. 2), for example, contains mostly names of trees, but also some names of bushes and some names for longish, rigid objects, e.g., *tsɛːɾé-ʔɛ* cricket-CL(TREE) 'a cricket'. Some classes have a core meaning but include exceptions which cannot be related to the semantic core. For example, the class of nouns determined by *-ko* (CL(SHAFT); No. 10) includes mostly nouns denoting objects of the shape of a (thick) stick, but also the name for cahuana, a thick drink made of manioc starch, *kaʔgɯ́nɯ-ko* cahuana-CL(SHAFT) 'cahuana'. Only for a few classes is the semantic motivation not obvious, the most heterogeneous one being the one determined by *-ba*, *-ʔba* (CL(3D); No. 5), which contains names of fruits, musical instruments, drinks, and baskets, among others. No semantic core can be identified in this case.

As is typical of the classification systems of the area, a basic parameter for categorization is physical shape. The three shape categories that are prevalent in systems of numeral classifiers (one-, two-, and three-dimensional) are distributed over various classes in Miraña:

(33)   Shape semantics of Miraña class markers

a. one-dimensional: large: -*ʔɛ* (CL(TREE); No. 2)
    (long and thin)   long: -*ko* (CL(SHAFT); No. 10)
    medium: -*i* (CL(STICK); No. 9)
    small: -*iːʔo* (CL(LITTLE.STICK);
    No. 27)
    pointed: -*hto* (CL(POINTED); No. 13)
    tubular: -*hɯ* (CL(TUBE); No. 51)
    -*ɾo* (CL(BOTTLE); No. 14)

b. two-dimensional: flexible: -*ʔaːmɨ* (CL(LEAF); No. 34)
    (flat)   rigid: -*gwa* (CL(PLANK); No. 7)
    round: -*hɨ* (CL(DISC); No. 8)

c. three-dimensional: general: -*ba*,-*ʔba* (CL(3D); No. 5)
    round: -*ɯ* (CL(ROUND.3D); No. 15)
    oval: -*ʔo* (CL(OVAL); No. 1)
    oblong: *ʔo* (CL(OBLONG); No. 4)

Although categorization by shape is dominant, some categorization appeals more to function, e.g., -*mɨ* (CL(TRANSPORT); No. 52), or to essence/material, e.g., -*ta* (CL(METAL); No. 54) or -*hpajko* (CL(LIQUID); No. 37). Such a combination of shape, function, and material semantics is typical of numeral classifier systems (Grinevald 2000).

4.5.3.   *Unique repeaters.*   The third set of class markers are basically nouns which are "repeated" in a class marker slot, and correspond therefore to the type of so-called repeaters found in many classifier systems (Grinevald 2001). What distinguishes repeaters from specific class markers with a recognizable origin is that they have the same meaning in both their use as noun and as class marker.

Formally, some such repeaters are morphologically identical in their uses as class markers and nouns (and can be labeled as full repeaters; Table 4), while others appear in a truncated form in their use as class marker (those being partial repeaters; Table 5). Example (34) illustrates instances of a full repeater.

(34)   *ó-di   íhka-báhɯ   tsá-bahɯ   báhɯ*
    1SG-POSS   COP-RP(FOREST)   one-RP(FOREST)   forest
    'I have one (stretch of) forest.'

The use of some of the repeaters, especially the partial repeaters which denote culturally important notions, such as the anaconda (No. 26), seems restricted to formal contexts (such as the recitation of mythological texts, often with a magical function). In addition, classes determined by repeaters can be

Table 4. *Miraña full repeaters*

No.	Repeaters	Corresponding nouns	Meaning
1	*-bajhkɛ*	*bájhke*	root
2	*-gwahka*	*gwáhka*	branch
3	*-bahkɯ*	*báhkɯ*	bone
4	*-hkɯba*	*hkɯ́ba*	leg
5	*-htɯʔaj*	*tɯ́ʔaj*	foot
6	*-ʔɛːba*	*í-ʔɛ́ːba* (POS.3S-waist) 'his/her waist'	waist
7	*-ʔohtsɨ*	*ʔóhtsɨ*	hand
8	*-ʔɯɯmɨ*	*ʔɯ́ɯmɨ*	face
9	*-bahɯ*	*báhɯ*	forest, plantation
10	*-dɛhɯʔɾiɯ*	*dɛhɯ́ʔɾiɯ*	tree without roots
11	*-gwaiɾoːhɨ*	*gwaíɾóːhɨ*	very thin
12	*-gwajhko*	*gwajhko*	hook
13	*-iːhɯ*	*iːhɯ*	anteater
14	*-koːhɨ*	*koːhɨ*	day
15	*-koːmi*	*kóːmi*	community
16	*-mihko*	*mɨhko*	corral
17	*-mɯhko*	*mɯ́hko*	place for bathing
18	*-oːʔi*	*oːʔíːbɛ* (jaguar-CL.MASC.SG) 'jaguar, dog'	jaguar or dog
19	*-pɛhko*	*pɛhko*	night
20	*-hkɯːβɛ*	*hkɯːβɛ*	afternoon
21	*-ʔihko*	*ʔíhko*	nest
22	*-ɾoʔpɛgwa*	*ɾoʔpɛ́gwa*	twisted

Table 5. *Miraña partial repeaters*

No.	Repeaters	Corresponding nouns	Meaning
22	*-bo*	*bóːa*	anaconda
23	*-iː*	*íːbíiː*	coca
24	*-iʔdʒo*	*dʒíɾiʔdʒo*	pot
25	*-ma*	*namɛ*	feces
26	*-pe*	*dʒiʔpɛ*	rat, sp.
27	*-to*	*toʔhɯ*	earthworm
28	*-tʃi*	*tʃiʔtʃi*	tucupí sauce
29	*-ra*	*dʒoːɾa*	parrot, sp.

said to be maximally motivated semantically to the extent that they are at the same time "unique" class markers heading classes of only one item.

The noun class system of Miraña therefore exhibits a very high degree of semantic motivation, both in its general and specific classifiers, the semantic motivation being particularly transparent in the case of unique repeaters.

4.6.   *Origins of class markers*

Unlike the situation prevalent with Bantu noun class systems, the origin of the Miraña system is clear, both in terms of the lexical origin of the class markers already considered and of a source for the construction. The syntactic source pattern for the system of nominal classification was most probably a genitive construction which still shares a number of formal and semantic similarities with nominal stem plus class marker constructions.

This genitive construction consists of a noun phrase expressing the determiner juxtaposed to the left of a noun phrase that expresses the determined. A "genitive" low tone docks near the limit of these two according to the number of syllables of both constituents. The fact that certain morphophonological processes occur within these constructions suggests that the two constituents form a unit which might be called a compound.[28] These constructions are very productive in Miraña. The following examples (35a–d) show the formal and semantic similarities that these constructions still share with noun stem plus class marker constructions. Compare the genitive constructions in (35a) and (35c) with the noun stem plus class marker constructions in (35b) and (35d). Forms of lexical nouns are bracketed in the examples.

(35)   a.   *úúhɨ-[ʔɯβíːba]*
             banana-[basket-CL(CONTAINER)]
             'a basket full of bananas'
         b.   *úúhɨ-ʔbábaj*
             banana-CL(BAG)
             'a bag of bananas'
         c.   *úúhɨ-ko-[bájhkɛ]*
             banana-CL(SHAFT)-[root]
             'a root of a banana plant'
         d.   *úúhɨ-ko-ʔáːmɨ*
             banana-CL(SHAFT)-CL(LEAF)
             'a leaf of a banana plant'

---

28. Cf. the development of classifiers from compounding constructions in Sino-Tibetan languages described by DeLancey (1986).

In (35a), *ʔɯβí-ːba* basket-CL(CONTAINER) 'a basket' is a noun that functions as the head in a genitive construction. It cannot occur in any of the environments typical for class markers. In (35b), *-ʔbábaj* (CL(BAG)) and *-ʔáːmɨ* (CL(LEAF)) in (35d) occur in the same position as *ʔɯβí-ːba* basket-CL(CONTAINER) 'a basket'. These forms have been grammaticalized as class markers. They may occur in all the contexts typical for class markers, but never as free forms. In (35c), *bájhke* 'root' is a noun that corresponds to a full repeater (see Table 4).[29] This form can be said to be halfway grammaticalized since it may occur as a repeater in a class marker slot and at the same time as a free form. Given that for a fair number of specific class markers a nominal origin can still be shown, it seems likely that most if not all specific class markers entered the system as repeaters, and then reduced their form, broadened their semantics, and now cannot be used as free forms anymore. Possibly, an agreement pattern involving general class markers already existed before specific class markers emerged, but there is no evidence for that today.

4.7.   *Areal features*

This section points out which of the features of the system of nominal classification in Miraña as described here constitute widespread areal phenomena and what is specific to Miraña.

Large, open systems of nominal classification that have agreement functions to varying degrees, on the one hand, and serve a derivational function for nouns, on the other hand, are an areal feature of Amazonian languages, at least western Amazonian ones. Systems of that type exist throughout the eastern Tucanoan languages (e.g., in Tuyuca (Barnes 1990), Tatuyo (Gómez 1982), Cubeo (Morse & Maxwell 1999)), as well as in Northern Arawak Languages (e.g., Tariana (Aikhenvald 1994), Baniwa (Aikhenvald 1996)), in Yagua (Peba-Yaguan; Payne 1986), and in the unclassified language Kwaza (Voort 2000).

In none of these systems is agreement by noun class as strict as it is in Niger-Congo languages. There seem to be two types of non-strict agreement of such large systems that are realized in a variety of morphosyntactic contexts:

(i)   possible variation between general and specific markers: there is a small subset of class markers with a general meaning that can substitute for more specific class markers in agreement positions (e.g., in Miraña and Yagua); this small subset seems to always distinguish animacy and often also natural gender and number;

(ii)   non-obligatory use: the realization of class markers is not obligatory in positions where agreement is found (e.g., in Tuyuca (Barnes 1990)).

---

29.   The grammatical status of *bájhkɛ* 'root' as a noun in (35c) is confirmed by the tonal patterns that are typical for possessive construction as against noun stem plus class marker constructions.

It is common in Amazonian languages that more than one system of nominal classification, using different markers on different targets, co-exist in the same language. Yagua and Tariana have a small obligatory gender system, which uses a different set of morphemes on different targets in addition to the large system of the Miraña type. Derbyshire & Payne (1990) give examples of other possible combinations in Amazonian languages.

Another possible arrangement, not found in Miraña, is that within one large set of class markers, the subsets of markers that are used on different targets differ to varying degree in their forms and/or semantics. This seems to be the situation in Palikur, an Arawakan language of another subgroup spoken in the North Eastern Amazon (Aikhenvald & Green 1998).

The phenomenon of repeaters which are used as class markers is a common feature of systems of Amazonian languages (cf. Barnes 1990, Payne 1986, Aikhenvald 1994, Aikhenvald & Green 1998, Eraso 2000).

The productive derivational use of class markers, at least with plant nouns, is a widespread feature of Amazonian languages (cf., e.g., Barnes 1990: 283, Aikhenvald 1994: 436–437; Gómez 1982: 116–117; Payne 1986: 122; Eraso 2000; Crofts 1985: 95; Petersen de Piñeros 1994: 39; Voort 2000: 73–75; Montes 2001).

The anaphoric use of class markers to refer to participants once they have been introduced in the discourse world has also been pointed out for various Amazonian languages (cf. Derbyshire & Payne 1990, Barnes 1990, Payne 1986, Aikhenvald 1994: 428, Thiesen 1996; Petersen de Piñeros 1994). In some cases, the use of class markers has also been described as being dependent on factors such as the topicality of a participant (e.g., Aikhenvald & Green 1994: 450–451). The "absolute" use of class markers has also been attested in other Amazonian languages, e.g., Witoto (Petersen de Piñeros 1994: 44–45).

It is worth mentioning that systems that are more like the noun class systems in Niger-Congo languages also exist in the Amazon. Andoke (unclassified; Landaburu 1979), for instance, has a system of six classes that are semantically non-transparent and that constitute an obligatory agreement system.

5.   Conclusions

The Miraña nominal classification system presented here is not the only type found in Amazonian languages, some of which have clearly either gender systems or some type of classifier system, such as numeral or genitival ones, but it is a fairly common one found with minor variations in a variety of languages, and coexisting sometimes alongside other, independent systems. As demonstrated, such a system characteristically combines features of various types of classification systems, such as class terms, numeral classifiers, noun classes, and to some extent gender. It is this richness that makes the description of such

systems challenging. One is confronted with systems that have the semantics and discourse use of a numeral classifier system, the agreement pattern of a noun class system, and the derivational and compounding productivity of class terms. However, the point to be made here is the possibility of describing such a system as essentially a noun class system, but one at a much earlier stage of grammaticalization than the Niger-Congo systems.

Beyond this particular demonstration, much remains to be done to assemble other full descriptions that address all of the issues raised here, of inventory, semantics, and origin of the markers, as well as morphosyntactic distribution and discourse use, descriptions comprehensive enough to permit a more complete comparison. Most descriptions are still partial, focusing on one or another aspect of the system, and are structured to demonstrate the uniqueness of the system, not to fit it within current typologies.

We hope to have demonstrated the interest of a closer comparison of the Niger-Congo noun class systems with the descriptively challenging nominal classification systems of many Amazonian languages, once the impact of some of the differences in the sociolinguistic contexts of their study and descriptions has been taken into account. The similarities and points of divergence that we have emphasized can be summarized as follows.

As to SIMILARITIES, the Amazonian systems and the African systems share an essential characteristic, namely the possibility that markers of noun classification that belong to one and the same system (in opposition to the co-existence of multiple classifier systems, and in spite of some possibility of non-uniformity of the markers themselves) can be simultaneously present on the noun itself, on its modifiers, and on the verb of which it is an argument.

As to DIVERGENCE, the only clear contrasts between Amazonian and African systems concern the facts that in the African systems (i) the presence of the markers of classes in the constructions in which they appear obey a morpho-syntactic constraint and are obligatory, and (ii) there is never any clue to a lexical origin of these class markers.

In response to the descriptive challenge posed by Amazonian nominal classification systems we have proposed, therefore, that many Amazonian systems, including that of Miraña, can be conceived of as emerging noun class systems, when compared to Niger-Congo systems which demonstrate what such systems can become if/when they reach full maturity. Many Amazonian systems are indeed classification systems of a concordial type: they share with African noun class systems a great variety of loci of noun class agreement. While in some languages the agreement is only with certain elements, in others it seems to have generalized to all the co-referring elements of a clause, both within the noun phrase and between noun phrase and verb. In addition, the nominal head can also be marked with a class marker and the agreement morphemes may be more or less regular.

However, some characteristics of the Amazonian systems mark them as less grammaticalized systems, such as their high level of semantic motivation and the clear lexical origin of many of their class markers (which is obvious with the phenomenon of repeaters and other truncated forms of the nouns). Another dimension that marks them as less grammaticalized is their discursive nature and their anaphoric function, features generally considered as more characteristic of large numeral classifier systems. These differences, however, are taken to be a matter of degree rather than essence, and the common patterns of the systems of these two regions of the world have to be seen as being representative of the two ends of a grammaticalization continuum: The Amazonian systems more often are emerging systems, and the African systems cover the entire spectrum from fully grammaticalized systems to the extreme of attrition and reduction of previously fully functioning systems. Our aim was to redress the tendency to consider as prototypical the first linguistic systems encountered and known, in this case the Niger-Congo noun class systems, and as exotic and deviant those discovered later, as has been the case with the perplexing Amazonian ones; we want *a priori* to give those that are more recently described a legitimate and full place in the construction of a typology of nominal classification systems. The approach advocated in order to accomplish this integration of the Amazonian systems was to espouse a dynamic approach to the endeavor that seriously takes into account the essential dimension of the progressive grammaticalization of systems: from their emergence as morphosyntactic systems as illustrated with Amazonian systems to their fully grammaticalized stage in many Niger-Congo languages, often in reality less regular and homogeneous than projected in the general linguistics literature, all the way to cases of the progressive loss of such systems, illustrating from one continent to the other full cycles of grammatical construction.

Correspondence addresses: (Grinevald) Institut des Sciences de l'Homme, 14 avenue Berthelot, 69363 Lyon, Cedex 07, France; e-mail: colette.grinevald@univ-lyon2.fr; (Seifart) Max Planck Institute for Psycholinguistics, PB 310, 6500 AH Nijmegen, The Netherlands; e-mail: frank.seifart@mpi.nl

Abbreviations: 1 1st person / class 1; 2 2nd person / class 2; 3 3rd person / class 3; 4 class 4; 5 class 5 (etc.); 3D three-dimensional; ACC accusative; ADL adlative; CL class marker/classifier; COP copula; DEM demonstrative; DIST distal; DU dual; FEM feminine; FUT future tense; GEN genitive; INAN inanimate; LINK linking particle; LOC locative case; MASC masculine; NMZ nominalization; OM object marker; PL plural; PN pronoun; POSS possessor; PRED predicator; PROX proximal; RP repeater; SG singular; SAP speech act participant; SM subject marker; SUB subordinate clause; TAM tense, aspect, mood.

References

Adelaar, Willem (2000). La diversidad lingüística y la extinción de las lenguas. In Queixalós & Renault-Lescure (eds.), 29–36.

Aikhenvald, Alexandra Y. (1994). Classifiers in Tariana. *Anthropological Linguistics* 36: 407–465.

— (1996). Noun classes, classifiers and agreement in Baniwa (Arawak). *Moscow Journal of Linguistics* 3: 7–33.

— (2000). *Classifiers: A Typology of Noun Categorization Devices.* Oxford: Oxford University Press.

— (2001). Areal diffusion, genetic inheritance, and the problems of subgrouping: A North Arawak case study. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Areal Diffusion and Genetic Inheritance*, 167–194. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. & Diana Green (1998). Palikur and the typology of classifiers. *Anthropological Linguistics* 40: 429–480.

Allan, Keith (1977). Classifiers. *Language* 53: 283–310.

Aschmann, Richard P. (1993). *Proto Witotoan*. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.

Barnes, Janet (1990). Classifiers in Tuyuca. In Payne (ed.), 273–292.

Corbett, Greville G. (1991). *Gender*. Cambridge: Cambridge University Press.

— (2003a). Agreement: The range of the phenomenon and the principles of the Surrey Database of Agreement. *Transactions of the Philological Society* 101-2: 155–202.

— (2003b). Agreement: Canonical instances and the extent of the phenomenon. In Geert Booij, Janet DeCesaris, Angela Ralli, & Sergio Scalise (eds.), *Proceedings of the Third Mediterranean Morphology Meeting (Barcelona, September 20–22, 2001)*, 109–128. Barcelona: Institut Universitari de Lingüistica Aplicada.

— (forthcoming). Agreement: Terms and boundaries. In William Griffin (ed.), *The Role of Agreement in Natural Language: Proceedings of the 2001 Texas Linguistic Society Conference,* Austin, Texas, 2–4 March 2001.

Craig, Colette G. (1986). Jacaltec noun classifiers: A study in language and culture. In Craig (ed.), 263–293.

— (1987). Jacaltec noun classifiers: A study in grammaticalization. *Lingua* 7: 241–284.

Craig, Colette G. (ed.) (1986). *Noun Classes and Categorization*. Amsterdam: Benjamins.

Creider, Chet (1975). The semantic system of noun classes in Proto-Bantu. *Anthropological Linguistics* 17: 127–138.

Creissels, Denis (2001). Les systèmes de classes nominales des langues Niger-Congo: Prototype et variations. In Anna Sörés & C. Marchello-Nizia (eds.), *Invariants et variables dans les langues: Etudes typologiques* (Linx, 45), 157–166. Nanterre: Université Paris X.

Crofts, Marjorie (1985). *Aspectos da língua mundurukú*. Brasília: Summer Institute of Linguistics.

DeLancey, Scott (1986). Toward a history of Tai classifiers. In Craig (ed.), 437–452.

Denny, J. Peter & Chet A. Creider (1986). The semantics of noun classes in Proto Bantu. In Craig (ed.), 217–239.

Derbyshire, Desmond C. & Doris L. Payne (1990). Noun classification systems of Amazonian languages. In Payne (ed.), 243–272.

Derbyshire, Desmond C. & Geoffrey K. Pullum (eds.) (1986/1990/1991/1998). *Handbook of Amazonian Languages*. 4 volumes. Berlin: Mouton de Gruyter.

Dixon, R. M. W. (1982). Noun classifiers and noun classes. Chapter 8 of *Where Have all the Adjectives Gone?* Berlin: Mouton.

— (1986). Noun classes and noun classification in typological perspective. In Craig (ed.), 106–112.

Dixon, R. M. W. & Alexandra Y. Aikhenvald (eds.) (1999). *The Amazonian Languages*. Cambridge: Cambridge University Press.

Eraso, Natalia (2000). El género y los clasificadores nominales en la lengua Tanimuka. Working paper. Bogotá: CCELA/Universidad de los Andes.

Gildea, Spike (1997). Introducing ergative word order via reanalysis: Word order change in the Cariban family. In Joan L. Bybee, John Haiman, & Sandra Thompson (eds.), *Essays on Language Function and Language Type*, 145–162. Amsterdam: Benjamins.

Givón, T. (1971). Some historical changes in the noun class system of Bantu, their possible causes and wider implications. In Chin-Wu Kim & H. Stahlke (eds.), *Papers in African Linguistics*, 33–54. Edmonton: Linguistic Research.

Gómez, Elsa (1982). De la forme et du sens dans la classification nominale en Tatuyo (langue tukano orientale d'Amazonie colombienne). Doctoral dissertation, Ecole Pratique des Hautes Etudes, Paris.

Gomez-Imbert, Elsa (1996). When animals become "rounded" and "feminine": Conceptual categories and linguistic classification in a multilingual setting. In John J. Gumperz & Stephen C. Levinson (eds.), *Rethinking Linguistic Relativity*, 438–469. Cambridge: Cambridge University Press.

Grinevald, Colette (1997). Language endangerment in South America: A programmatic approach. In Lenore A. Grenoble & Lindsay J. Whaley (eds.), *Endangered Languages: Language Loss and Community Response*, 124–159. Cambridge: Cambridge University Press.

— (2000). A morphosyntactic typology of classifiers. In Senft (ed.), 50–92.

— (2001). Classifiers, Linguistics of. In Neil J. Smelser & Paul B. Baltes (eds.), *International Encyclopedia of the Social and Behavioral Sciences*, Volume 3, 1973–1978. Oxford: Elsevier.

— (2002). Making sense of nominal classification systems: Noun classifiers and the grammaticalization variable. In Ilse Wischer & Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 259–275. Amsterdam: Benjamins.

— (2003). Classifier systems in the context of a typology of nominal classification. In Karen Emmorey (ed.), *Perspectives on Classifiers Constructions in Sign Languages*, 91–109. Mahwah, N.J.: Erlbaum.

Guthrie, Malcolm (1967–71). *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of Bantu Languages*. Farnborough: Gregg.

Heine, Bernd (1982). African noun class systems. In Hansjakob Seiler & Christian Lehmann (eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen*, Part I: *Bereich und Ordnung der Phänomene*, 189–216. Tübingen: Narr.

Landaburu, Jon (1979). *La langue des Andoke.* (Langues et Civilisations à Tradition Orale, 36.) Paris: SELAF.

Mithun, Marianne (1986). The convergence of noun classification systems. In Craig (ed.), 379–395.

Montes, R., María Emilia (2001). Los nombres de las plantas, sus partes y sus espacios de crecimiento: Aproximación ethnolingüística a partir de datos de la lengua ticuna, Amazonia

colombiano. In Carlos E. Franky & Carlos G. Zárate (eds.), *Imani mundos: Estudios en la Amazonian colombiana,* 523–558. Bogotá: Editorial Universidad Nacional.

Morse, Nancy L. & Michael R. Maxwell (1999). *Cubeo Grammar*. (Studies in the Languages of Columbia, 5.) Arlington, Texas: Summer Institute of Linguistics.

Mufwene, Salikoko S. (1980). Bantu class prefixes: Inflectional of derivational? *Chicago Linguistic Society* 16: 246–258.

Nettle, David & Suzanne Romaine (2000). *Vanishing Voices: The Extinction of the World's Languages.* Oxford: Oxford University Press.

Oates, Lynette (1964). *A Tentative Description of Gunwinggu Language.* (Oceania Linguistic Monographs, 10.) Sydney: University of Sydney.

Osam, Kweku (1993). The loss of the noun class system of Akan. *Acta Linguistica Hafniensia* 26: 81–106.

Payne, Doris L. (1986). Noun classification in Yagua. In Craig (ed.), 113–132.

— (1987). Noun classification in the Western Amazon. *Language Science* 9-1: 22–44.

— (1990). *The Pragmatics of Word Order*: *Typological Dimensions of Verb Initial Languages*. Berlin: Mouton de Gruyter

Payne, Doris L. (ed.) (1990). *Amazonian Linguistics: Studies in Lowland South American Languages.* Austin: University of Texas Press.

Petersen de Piñeros, Gabriele (1994). *La lengua uitota en la obra de K. Th. Preuss*. Bogotá: Editorial Universidad Nacional.

Queixalós, Francisco & O. Renault-Lescure (organizers) (2000). *As línguas amazônicas hoje.* São Paulo: IRD, ISA, MPEG.

Rehg, Kenneth L. (1981). *Ponapean Reference Grammar*. Honolulu: University Press of Hawaii.

Richardson, Irvine (1967). Linguistic evolution and Bantu noun class systems. In *La classification nominale dans les langues négro-africaines*, 373–390. Paris: Centre National de la Recherche Scientifique.

Seifart, Frank (2002). *El sistema de clasificación nominal del Miraña*. (Lenguas Aborígenes de Colombia: Descripciones, 13.) Bogotá: CCELA/Universidad de los Andes.

Senft, Gunter (ed.) (2000). *Systems of Nominal Classification*. Cambridge: Cambridge University Press.

Thiesen, Wesley (1996). *Gramática del idioma Bora.* (Serie Lingüística Peruana, 38.) Yarinococha, Peru: Instituto Lingüístico de Verano.

Thiesen, Wesley & David Weber (forthcoming). A grammar of Bora. Summer Institute of Linguistics.

Voort, Hein van der (2000). A grammar of Kwaza. Doctoral dissertation, Universiteit Leiden.

Wurzel, Wolfgang Ullrich (1986). Die wiederholte Klassifikation von Substantiven: Zur Entstehung von Deklinationsklassen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39: 76–96.

Zavala, Roberto (2000). Multiple classifier systems in Akatek (Mayan). In Senft (ed.), 114–146.