UNIVERSITY OF BERGAMO

School of Doctoral Studies

Doctoral Degree in Linguistics

XXIX Cycle

SSD: L-LIN/01

TITLE

The morphosyntax of number systems: a cross-linguistic study

Advisor

Chiar.ma Prof.ssa Sonia Cristofaro

Doctoral Thesis

Jessica Katiuscia IVANI

Student ID 1031796

Academic year 2015/16

Contents

List of Figures

List of Tables

1. Introduction

1.1. Introduction

Number is one of the most varied domains of grammar. In typological literature on number marking, many studies have focussed on classifying the variety of the number systems in the languages of the world by mapping the presence, in such systems, of the number values and their distribution.

Typological generalizations have been elaborated to describe and, to some extent, explain, this variation. This work focusses on one of the less researched aspect of the nominal number marking domain: how the values are expressed in the number systems, that is, the constructions used to mark number and their interaction with the nominal types from a synchronic and cross-linguistic perspective.

At the state of the art, linguistic studies have explored so far number constructions by investigating the properties of a single phenomenon, like a specific construction type; in other studies, detailed explorations of the number markers are restricted to single languages, or language families.

The aim of this study is to provide for a synchronic and systematic description of the number markers from a cross-linguistic perspective.

This research is built on the following research questions:

Exploration of the construction types used to mark number values

- Which construction types are used to express nominal number in languages?
- What is the distribution of the construction types within the number

values and the lexical categories?

- Are there construction types recurrently associated to a NP type?

Internal distribution of the plural markers within the nominal types

- What is the distribution of the plural markers among the NP types?
- Which nouns tend to signal plurality through the same marker?
- What is the behavior of demonstrative pronouns with respect to plural marking?

Typology of the number systems and the interaction within the number values

- Which number systems can be identified in the languages of the sample?
- What is the distribution of the number systems on each macro-area?
- Are there any relevant features that can be associated to a number system?

1.2. Outline

This dissertation is organized as follows. Chapter 2 provides an overview of previous literature on number marking and broad attested facts related to nominal number expression. The research questions addressed in this study, and partially mentioned here, are fully outlined at the end of the chapter. In Chapter 3 I describe the sample methodology which served as basis for data collection.

In Chapter 4 I introduce the parameters and the methodological approach followed in this work: in the first part of the chapter, I discuss and describe the parameters which serve as backbone of this survey; in the second part of the chapter, I introduce the goals and the recent methodological developments of linguistic typology and I outline the method adopted in this study, the multivariate approach to typology, and illustrate its advantages. The chapter concludes with the description of the formal model of such approach I have developed,

the typological database of nominal number constructions.

Data discussion is presented in Chapter 5 (nominal types and number marker), Chapter 6 (plural marking in nominal types) and Chapter 7 (for a typology of number systems). Final remarks and prospects for future research are illustrated in Chapter 8.

2. Background

2.1. Introduction

In this chapter, I provide an overview of previous literature on nominal number and the most recent developments about the typological study of number and number marking constructions.

The chapter consists of two parts. The first focusses on the broader aspects of number from a typological perspective: this section includes an overview of the typological generalizations proposed on number marking, defined as the Number Hierarchy and the Animacy Hierarchy; a brief description on previous studies on number, from Corbett (2000) until the most recent worldwide surveys developed for databases an atlases of grammatical features, like Haspelmath et al. (2005).

The second part of the chapter illustrates some more narrowed studies about the morphosyntax of number systems, like the origin and the development of number and the number marking systems in a specific language or language family.

(Mithun 1988) illustrates the development of nominal number constructions from markers of verbal plurality in a group of languages from North America. (Iljic 2001), in the case study presented in Section 2.5.2, reconstructs the origin of the plural marker *-men* in Mandarin Chinese; (Dryer 1989b) rather focusses on the synchronic attestation and distribution of plural words in a sample of languages from the world, proposing a classification of the lexical items based partially on their origin. The chapter concludes with an some remarks on the plural constructions attested in pidgin and creoles, whose recent and straight-

forward structure allows, in some cases, the reconstruction of the number marking system.

A summary with the full research questions addressed in this work is provided in Section 2.6.

2.2. Nominal number marking and typology

Typological literature about number marking has focussed on mainly two aspects: the identification and the description of the number systems found in the languages of the world and the distribution of the number values opposition at the noun level within these systems.

The number values identified among the languages of the world are singular, dual, plural, trial, quadral and greater plural [1]. Such values correspond to one, two, more than one, three, four, a few and many real world entities, respectively. Assumed these values, typological observations on available data have argued that their distribution within languages, or 'the prediction of the possible number systems natural languages can have' Corbett (2000: p.38) is constrained by implicational tendencies, on which a Number Hierarchy has been formulated (Comrie 1989, Croft 1990):

singular > plural > dual > trial

Most of this hierarchy can be seen as derived from the assumptions present in Greenberg's universal n. 34:

> No language has a trial number unless it has a dual. No language has a dual unless it has a plural (Greenberg 1966)

The structure of the Number Hierarchy is implicational: this means that the values to the right imply those on the left; paucal and greater plural values

[1] I am explicitly discarding values such as general number, since it refers to languages that do not encode number distinctions obligatorily; in fact, the nominal types involved in such process can be either singular or plural. Such contexts, following Corbett (2000), do posit outside the number systems by definition and therefore they are not included in this overview

lack from the hierarchy scale, and it constitutes also a problem whether they should be incorporated: such issue is discussed in Corbett (2000: p.39-41), who ultimately argues that such values differ in nature from the ones that appear in the Number Hierarchy.

Typological investigations of nominal number have also shown that the distribution of plural marking in a language is conditioned by the animacy of the nominal referents. The version of the Animacy Hierarchy used to account for typical and atypical strategies of number marking on noun and pronouns is the following:

speaker > addressee > 3rd person > kin > human > animate > inanimate

Due to its implicational structure, the Animacy Hierarchy can be read as that the likehood of plural marking increases the more one moves the outmost left of the scale. In many languages, only pronouns and human nouns undergo obligatory plural marking; at the rightmost part of the scale, inanimate nouns may be insensitive to singular/plural distinctions.

The earliest version of the Animacy Hierarchy was presented by Smith-Stark (1974); Smith-Stark was inspired by Forchheimer (1953) and Silverstein (1976); his Hierarchy is clearly akin to what others have defined as the Topicality Hierarchy (Comrie 1989). There are different versions of the Animacy Hierarchy applied to nominal plural marking: one of the latest development is proposed by Haspelmath (2005): his hierarchy does not take into account pronouns (because they often display different encodings of plurality) but, on the other hand, it is more fine-grained within the full nouns scale, adding distinctions between high and lower animates and discrete vs non discrete inanimate nouns.

Smith-Stark paper was a major step forward towards the understanding of number systems: it focusses specifically on the marking of plurality in nouns. Moreover, the examples reported to support his claims and to show the 'splits' which plurality may undergo in the noun domain in languages are taken from a good number of languages and it dealt with aspects of number not covered yet, like the obligatoriness of nominal number and the different morphological means that can be used in a language to express plurality.

Typological hierarchies are explained by invoking abstract principles like economy or frequency: by economy it is implied that a language may show plural distinctions only on certain nouns, like pronouns and human, because they are more relevant and speakers tend to express number distinctions when they need to do so (and inanimates are less relevant); by frequency, it is implied that noun types on the upper part of the hierarchy are more frequent and thus need more accurate number distinctions.

2.3. The expression of number

Number has also been approached through the morphological means it is expressed. Corbett (2000) lists an amount of strategies that can associated with number expression: beside the most common marking expressions that appear in languages (suffixes, suppletive forms and the like), Corbett also reports less common and 'atypical' means of number marking, like subtractive processes, double plurals and cases of inverse number, also defined as polarity (all these phenomena are described in detail in Corbett (2000: p.154-177), who also explores number indexation).

One of the most recent and broad cross-linguistic account on nominal plural marking appears in WALS, and it is the Chapter 33 'Coding Nominal Plurality' by Matthew Dryer (Dryer 2013). In Dryer's survey, conducted on more than 1000 languages worldwide, it is shown the method by which a language indicates plurality in nouns. He does not present the internal distribution of the marker within nouns, nor whether there are different plural marking strategies within the same language: what is captured it is rather the 'primary method' employed by a language, that is, the most common.

Dryer's report shows the geographical distribution of the plural marking strategies among the world, and the number of languages which employ the same construction type. Among the results, the map shows plural suffixes greatly outnumbering all other types shown on the map. They are widely distributed throughout the world; the largest area in which they are not found is

Southeast Asia and most of the area in which Austronesian languages are spoken. The map shows plural prefixes scattered among Austronesian languages in the general vicinity of Indonesia and the Philippines, an area in which plural suffixes are generally lacking, but various other types are also common in this area, including plural by complete reduplication, plural words, and absence of plural altogether. Plural prefixes are completely lacking from two large areas: (i) the entire mainland of Europe and Asia; (ii) an area in the New World stretching from Guatemala south to include all of South America. Plural words and clitics are most common in southeast Asia and among Austronesian languages, in West Africa, and in the Amazon basin in South America. All four languages using tone as the primary means of indicating plural are spoken in Africa.

Dryer (1989b) focusses on the expression of number as well: in this specific case, however, the survey investigates the distribution and the structural properties of a specific construction, plural words in about 50 languages worldwide. In the conclusions he proposes, some facts about the distribution of the plural words are reported: lexical items are more frequent in the languages of the South-East Asia and Australasia. The clearest empirical result of the paper involves word order: plural words exhibit a strong tendency to precede the noun in VO languages, and follow in OV languages.

Daniel (2013) chapter 35 in WALS examines some of the formal correlations between the plural of personal pronouns and nominal plurality. It is a unique study of its kind since the plural of personal pronouns has been considered both as a phenomenon related to nominal number and as something that has nothing to do with it.

Some scholars, like Forchheimer (1953) (and also what one could infer from the Animacy Hierarchy) have been suggesting that, if a language has a number distinction somewhere, it has to be in pronouns and, diachronically speaking, that plurality has spread from pronouns to full nouns. On the other hand, linguists have repeatedly argued that pronominal plurals are different from nominal plurals (Corbett 2000, p.83; Cysouw 2003).

However, Daniel provides an account for how the devices for expressing

pronominal plurality do relate for devices for expressing nominal plurality. Since this aspect is one of the most relevant in the present work, I will report in further detail the methodological approach and results of this survey. The analysis is targeted on full nouns and first and second independent person pronouns; third person pronouns, that could have revealed some interesting features because of their particular referent status halfway between the speaker-addressee dimension and the full nouns, are not involved in the investigation. The results of this analysis show some areal tendencies: suppletion forms in independent personal pronouns are the most common type and they are attested throughout the world. Pronouns which use a number affix are most typical of Asia and northern Australia; they are not infrequent in the Americas and also occur in central Africa. More specifically, pronouns consisting of a person-specific stem plus number marker cluster in eastern Asia and are rare elsewhere, while pronouns consisting of person-number-specific stem plus number marker specific to pronouns are frequent in northern Australia and the eastern Pacific. A number of languages (more than half of the 260 languages investigated) have specific markers to express plurality: 70 languages report to have pronominal-specific number affixes (added to a person stem or to a person/number stem); 40 languages carry the same pronominal affixes on both pronouns and nouns.

Unfortunately, given the values list, the extent of the spread and marking similarities between pronouns and full nouns is not traceable. More specifically, it is not specified whether the plural marker shared with nouns is restricted to a subset of them -and, if that is the case, which noun type- or it applies to all nouns with no regards to animacy. Furthermore, what lacks in this survey is the interaction between the first and second personal pronoun: when the languages show a pronouns-specific number marker, is the same marker for both pronouns or are there two different markers but ultimately pronoun specific? Finally, as aforementioned, independent third person pronouns are excluded from the overview, and this aspect gives a kind of uncompleted frame. These unclear aspects are part of the research questions this study aims to answer.

2.4. The distribution of the number systems

Several studies have ilustrated the distribution and characteristics of the number systems from a cross-linguistic perspective.

Nichols (1992) hypothesizes the existence of a correlation between the morphological characteristics of languages and the types of attested number systems. Grammatical number marking is, for example, very infrequent among the isolating languages of Southeast Asia.

Another comparable result about the distribution of the number systems is the one described in WALS database chapter 33 (Haspelmath 2005), concerning the distribution of obligatory vs non obligatory strategies in nominal number. The study focusses on full nouns, and the different means of plurality expression play no role; only occurrence of the plural marker is investigated, along two dimensions: animacy and obligatoriness. The animacy parameter is set on a binary contrast, human nouns vs inanimate ones. In the obligatoriness dimension,

Haspelmath distinguishes between non occurrence, optional occurrence and obligatory occurrence. The results in the WALS map show evidence for the implicational scale based on animacy seen above, of which Haspelmath gives a richer version: kinship terms > other humans > "higher" animals > "lower" animals > discrete inanimates > nondiscrete inanimates, warning, however, that more fine grained studies on plurals are therefore needed to find convincing evidence. Animate nouns are not contemplated, and this lack of information, among other issues, both structural and methodological makes it impraticable to integrate such data with the language data available from WALS chapter 33 aforementioned. Furthermore, it is often difficult, as Haspelmath also argues, to infer the optionality or the obligatoriness of a marker from a language description and moreover the optionality of plural marking is linked sometimes to other variables, such presence of quantifiers or numerals, since identifying all these micro sub-types and degrees of obligatoriness is indeed complicated. Anyway, data provided by Haspelmath follows Nichols' hypothesis: looking at the map, the distribution of obligatory nominal plural marking appears to be

quite skewed. Africa and the westernmost part of Eurasia are the areas where this type of marking is most commonly found.

Similar results have been attained by Wälchli (2012), on his investigation of frequency and occurrence of nominal plural marking in a sample of 82 languages by means of parallel corpora. In the same study, Wälchli (2012: 255) suggests that attrition in the nominal number domain rarely leads to complete loss. When, as, for instance, in French, number marking on head nouns disappears due to phonological erosion, number-indexation strategies used on demonstratives, articles or even verbs are likely to become more prominent in discourse thus allowing for number distinctions to be still overtly coded through syntax.

2.5. Development of number systems and sources of number

Other studies have rather focussed on the development of the number systems, than their distribution of their features in synchrony. The 'marking reconstruction work' is restricted to small groups of related languages, restricted areas or at a language specific level, since diachronic analysis requires a depth investigation of the possible sources of number and of all the processes that took place and led to the synchronic language frame.

Grammaticalized number markers can be traced to several sources, such as demonstratives, quantifiers, nouns, and markers of verbal plurality. Cristofaro (2012) makes an useful distinction between sources of nominal number markers that are etymologically connected with the encoding of plurality and sources that are instead entirely independent of number. The latter come to be used as number markers as an accidental result of grammaticalization.

Among the first type of sources, she includes (i) collectives and distributives (like 'objects scattered here and there'), where the number markers in result have developed from a context of and already present system of number marking; (ii) expression of multitudes, like quantifiers 'many', 'all', or nouns denoting

group of entities 'men', 'people'[2].

To the aforementioned sources, one could also add markers of verbal plurality: this is the case of a variety of North American languages, in which nominal number marking can be seen as a result of an extension process from markers of verbal number (Mithun 1988) Another source marker that can be added to the list are associative markers: constructions like 'X and others', 'X and associates' can grammaticalize and therefore arise as plural markers: such development can be seen in Mandarin Chinese, where the plural suffix *-men* was in origin an associative marker (Iljic 2001).

The second type of sources for plural markers includes those elements in origin completely independent from the expression of number, like deictic forms and demonstratives: Frajzyngier (1997) analyses markers of nominal (and verbal) plurality in a variety of Chadic languages that developed as a result of grammaticalization from demonstratives. Plural marking on nouns is connected to definiteness; deictic markers that initially had scope over entire NPs, and only marked definiteness, later began to mark plurality on nouns.

2.5.1. Nominal plural from verbal plurality: the case of North American languages

Mithun (1988) shows how verbal markers can give rise to nominal number constructions; empirical data for this assumption comes from North American languages. Number marking in the languages of the North America poses quite far from the traditional ideas of nominal number and agreement one usually encounters in languages, especially Indo-European ones: traditionally, number is considered as an inflectional property of nouns. Number distinctions on other categories, like verbs, are immediately perceived and classified as agreement markers. In the majority of North American languages, noun types are not marked for number; moreover, number can be marked on verbs even when

[2] Cristofaro also reports that generic expressions of multitude like 'all' tend to be grammaticalised and therefore used as plural markers to all nouns, with no semantic restriction; on the other hand, more specific expressions like 'men' show a tendency to develop as plural markers on restricted semantic context, with human and animate entities.

nouns are completely absent. Although some North American languages, like Kiowa or Zuni, show nominal plural inflection on all nouns, in the most common cases only human nouns (or other subsets, like kin terms) happen to have plural forms.

This restriction in distribution of number markers in nouns, Mithun observes, is mostly due to the derivational nature, rather than inflectional, of the markers of verbal plurality they are related to. Another striking difference between 'usual' number markers and the ones found as verbal number constructions in the languages of North America is that the former do enumerate entities, while the latter rather sign aspects of events.

These two facts ((i) the derivational nature of verbal markers and (ii) that they quantify aspects of events rather than plurality) have important consequences on the number systems of these languages: verbal number systems are rich and elaborate and, although multiple events and distributive markers are the most common verbal quantifiers in North American languages, other types of number markers do appear on verbs: they usually specify 'collective agency, collective causation, multiple displacement and iterations of various sorts' Mithun (1988: p.217). This cannot be associated to simple agreement; moreover, within this complexity, the distinction between singular and plural number is not equally pertinent to all verbs; since their function is so specialized, it constitutes as a derivational process which has no reason to spread systematically on nouns.

Nonetheless, verbal number markers may spread to nouns: in such cases, their original function is usually retained and it is mostly connected to distributivity. However, when distributive markers are extended to nouns, their function can shift. Since people can be considered as inherently differentiated, distributives will show a tendency to appear with noun types referring to multiple human beings; in a second phase, markers can be reinterpreted as plurals, but only in that specific domain.

2.5.2. Plural markers from associative markers: the case of *-men* in Mandarin Chinese

Iljic (2001) provides a solid example of a number marker developed from an associative marker construction. Associative plural constructions consist of a noun X (typically of human reference, usually a person's name or a kin term) and some other material, most often an affix, a clitic, or a word. The meaning of the construction is 'X and other people associated with X' (Daniel & Moravcsik 2005). The debate about the development of *-men* relies on the complexity of its etymology, since the mentioned suffix lacks of early written attestation due to its vernacular origin; moreover, the non-phonetic aspect of Chinese writing system, where the meaning of a character chosen for its pronunciation to transcribe a function word is not necessarily related to the grammatical value of the latter, makes the traceability of this construction more complex.

As Iljic reports, the debate revolved around two symmetrical pairs of arguments: the first, related to the nature of this marker (as a result of a borrowing from Mongolian varieties vs native origin), the second pertains to its meaning (plural vs collective value). The difficulty of disambiguating the meaning of this marker applies to its distribution: obligatory with pronouns and human nouns and optional with other nouns, it has been claimed that *-men* could be considered as 'a plural marker with pronouns and a collective marker with nouns' (Iljic 2001: p.75).

This explanation, however, does not account for the optionality of the marker and the non-random property of its assignment to nouns and, moreover, it is not supported by historical evidence. Such ambiguity in synchrony led Iljic to resolve it by invoking diachronic explanations: through a consistent corpus scrutiny, Iljic is finally able to point out that the use of *-men* after pronouns and human nouns cannot be dissociated from each other: that is, '*-men* occurs with nouns when all the conditions for the use of personal pronouns are met' (Iljic 2001: p.90).

This circumstance is satisfied, Iljic observes, when the employ of the suffix after nouns produces a 'group effect', not an 'objective' one, but rather a *personal*

collective, that is, a grouping around a specific subject-locator. Such referentiality is clearly akin to the referential properties of personal pronouns, defined by the speaker point of view and this explains why the marker is mandatory with human nouns. Based on this evidence, Iljic concludes that, as a synchronic result, *-men* can be interpreted as an expansion process of the plural marker from pronouns to nouns, with both the origin and the direction of this 'spread' conditioned by the most durable properties of *-men*, that is its being a 'personal collective', or an associative marker.

2.5.3. Plural markers from demonstratives forms

Frajzyngier (1997) shows that demonstratives[3] may be counted among the historical sources for plural markers. The empirical material he provides comes from African data (Chadic languages). However, this phenomenon is not restricted to Africa and African languages: Dryer (1989b) has found instances of plural words derived from articles from six languages of his sample, Khasi and Hawaiian among others[4]

In the overview presented by Frajzyngier, the plural marker has the same form as a demonstrative or one of its components. Based on these and parallel data from several other Chadic languages, Frajzyngier maintains that the plural markers in Chadic developed from former demonstratives.

Frajzyngier analysis has two specific aims: (i) to demonstrate the similarity between plural markers (both nominal and verbal) in the Chadic language family (ii) to offer an explanation for grammaticalization of demonstratives (also singular forms) into plural markers.

[3] Frajzyngier uses demonstratives as a cover term for anaphoric pronouns and definite articles beside demonstrative pronouns and adjectives. In this study, only demonstrative pronouns are taken into account.

[4] Dryer makes a further crucial distinction between 'real plural words' and plural words which are actually still articles: in the latter case, they are not real plural words in a sense that plurality is marked through the determiner and it codes also definiteness distinction, thus it cannot be referred as nominal number marking strategy anymore: Khasi *kii* plural word is a real example of plural word, according to Dryer's distinction; *nā*, the plural word used in Hawaiian, is instead an article, since it occurs in article position and it carries definiteness value.

The former aspect of his survey appears to be straightforward in the case of plural markers that evolved from plural demonstratives: the latter are readily reinterpreted as plural markers if they lose their deictic function.

For the latter aspect, Frajzyngier discusses several factors that may have contributed to the grammaticalization of singular demonstratives as plural markers. Most importantly, he points out that plural marking in Chadic is often confined to definite nouns marked by an adnominal demonstrative or a related noun modifier. Due to the cooccurrence of definiteness and plural marking, adnominal demonstratives may become associated with the semantic feature of plurality and then they are immediately reanalyzed as plural markers if they lose their deictic function.[5]

2.5.4. Notes on source markers in pidgin and creoles

It has been mentioned how individuation and reconstruction of nominal number source markers is in principle a complex enterprise due to the scarse availability of diachronic data on most languages.

Pidgin and creole varieties, on the other hand, provide less ambiguous and more straightforward information about the development of their structural features, number constructions included.

APiCS database, which covers grammatical and lexical structures of about 80 pidgin and creole languages worldwide, provides a good account of nominal number marking: two chapters cover the occurrence of nominal plurality and the means by nominal plurality is expressed in pidgin and creoles, linking this information to the homonym versions contained in WALS database.

APiCS chapter 25, however, is not linked to any WALS chapter: it describes

[5] Frajzyngier discusses also verbal plurality, and the plural markers which are used by verbs: accordingly to his explanation, the author attributes the morphological resemblance of verbal and nominal plural markers to a common historical origin, claiming that both plural markers developed from demonstratives. Diessel (1999), on his cross-linguistic study on demonstratives, suggests instead that nominal and verbal plural markers originate from demonstratives in two different source constructions: nominal plural markers as derived from adnominal demonstratives that accompany a juxtaposed noun, while verbal plural markers are developed from pronominal demonstratives that cliticize to a verb stem.

the extent of the identity between the nominal plural marker in a language and the third person pronoun. This applies only to languages which have plural words: third person pronouns can function as plural words that follow (or precede) the noun they modify. Looking at the data provided by APiCS, more than half of languages which use a lexical item (as a unique strategy or among other means of plural marking), it happens to be the third person pronoun. This feature is not exclusive of creoles and pidgin languages: it is quite a common construction, present and attested in the languages of the world (as it is shown in this study: see Section 6.8).

However, third person pronoun employed as a plural word is not the exclusive strategy that can be found in pidgin and creoles languages to express plurality: among these markers, one can find the aforementioned expressions of multitude like 'all': 'ol nes' '*nurses*' (Bislama, Crowley 2004); expressions derived from 'group' 'bann vyé fam la' '*these old women*' (Reunion Creole: bann< French 'bande', 'group', employed with human nouns only, Bollée (2013)); expressions meaning 'people': 'nas abúna' '*priests*' (Juda Arabic nas < Arabic 'nas' 'people', used on animate nouns only, Manfredi (2014)).

2.6. Summary

In this chapter, an overview of the main literature on number marking has been illustrated. Typological studies show the presence and the distribution of the number values and the number distinctions, and how they are constrained within the systems.

Some facts involving number marking have not been explored in detail: the cross-linguistic exploration on nominal plurality in WALS excludes the variety of the constructions used in the languages, taking into account only the 'primary strategy' used among the languages; moreover, the distribution of the plural distinction is not fully defined: the related investigations illustrated in WALS do pertain exclusively to human / inanimate nouns, excluding more fine-grained distinctions in nouns and other lexical categories.

Up to date cross-linguistic investigation on pronominal plural marking are restricted to first and second person, excluding the third: the possible 'clustering' of pronouns and full nouns with respect to the plural markers used has not been explored systematically.

The studies pertaining to the history and the reconstruction of the number markers in languages have shown the important role that the origin of the marker, and its development, may have in constraining the distribution of the number distinction within the nominal types. A systematic, when available, report of the source of the number markers may therefore help in tracing these developments and might contribute towards the understanding of the synchronic number marking contexts attested in languages.

Given these facts, the full research questions addressed in this study can be illustrated:

Exploration of the construction types used to mark number values

- Which construction types are used to express the number values?
- What is the distribution of the construction types within the number values?
- Are there any recurrent characteristics in the construction types used to signal a specific number value?
- Which number construction types are used on the nominal elements?
- Are there construction types recurrently associated to a NP type?
- Do the construction types associated with a nominal type show relevant features?

Internal distribution of the plural markers within the nominal types

- What is the distribution of the plural markers among the NP types?
- Which full nouns tend to signal plurality through the same marker?
- What is the distribution of suppletive and morphological strategies in pronouns plural expression?
- Are there languages with pronoun specific plural markers? Which

internal grouping can be identified?

- Are there attested cases of plural markers shared by nouns and pronouns? What is the internal distribution of such constructions?
- What is the behavior of demonstrative pronouns with respect to plural marking?

Typology of the number systems and the interaction within the number values

- Which number systems can be identified in the languages of the sample?
- What is the distribution of the number systems on each macro-area?
- Which groupings, or subtypes within the number systems, can be identified?
- Are there any relevant features that can be associated to a number system?

3. Data collecting and sampling

3.1. Introduction

The present chapter illustrates the language sampling procedure, its methodological approach and how it relates to the research questions outlined in Section 2.6.

This section is structured as follows: first, the state of the art of language sampling and its importance in linguistic typology is briefly introduced. The following subsections include a detailed overview of the method adopted specifically for this research and its theoretical motivations and limits.

The second part of the chapter discusses data collection and organization. A summary of the chapter is presented in Section 3.5.

3.2. Sampling methodology

Based on the configurational structure of its sample, this dissertation may be classified as a typological investigation of macro-areas, following Dryer 1989a terms. Language sampling is complex and debated in typological research; several sampling approaches have been proposed (for an overview of sampling methods in linguistic typology, see Widmann & Bakker (2006); for specific challenges of language sampling, see Bakker (2007) and Dryer (1989a)).

As in any kind of investigation based on empirical data, the construction of a typological sample must reflect the nature of the research questions a study aims to answer (Bakker 2007: p.106). As outlined in Chapter 1, this dissertation explores the number systems in the languages of the world from a construction

approach. I am therefore interested in accounting for the overt constructions employed to express the number values in nouns, their types and distribution, and explore whether there are recurrent associations between constructions and number values and between number markers and nominal types.

From this perspective, the ideal sample shall be able to capture the recurrence of the identified associations within the same language family and between related languages; similarly, I am interested in accounting for the stability and the distribution of the aforementioned associations and patterns both genealogically and areally. Moreover, to ensure typological variety, underinvestigated and isolate languages shall be equally represented in the sample.

Towards this aim, I have built the sample with the purpose of both focussing on linguistic diversity and combining intra-genelogical and inter-genealogical dimensions, well aware that dealing with a world-wide exploratory investigation exposes to multiple sampling biases (genealogical and bibliographic, among others) that is important to limit to the possible extent.

Intra-genealogical comparison leads to the mapping of the synchronic distribution of a specific phenomenon through closely related languages; moreover, it contributes in defining its stability and allows the identification of recurrent associations and patterns. Synchronic patterns may help in the definition of diachronic tendencies of the phenomenon under investigation. Diachronic inference from synchronic patterns has been extensively employed in linguistic typology, and it can be referred to as *dynamic typology*, following definition by Croft (2003: p.247).

The inter-genealogical perspective, more focused on the areal spread and attestation of a specific (or a group of) variables through unrelated languages, may suggest or motivate the limits of cross-linguistic variation of a given phenomenon, or pattern. This areal contribution can provide a more complete and exhaustive overview of a parameter one wants to investigate typologically and may also lead to the discovery of patterns that cut across genealogical affiliation.

To conclude, a data collection with a focus on language isolates could com-

plete the 'variety frame' of the present investigation. The result dataset of this combined sample procedure consists of approximately 250 languages selected from different genealogical groupings from the languages of the world.

3.3. Sampling procedure

The method chosen for this research resembles the one proposed and employed by Dryer 1989a for wide investigations like the cross-linguistic studies on word order universals. Although my sampling method slightly differs from a probabilistic one, I decided to adopt its basic structure especially for the grouping procedure and language classification framework, that has been also employed by important databases like WALS and it is suitable for linguistic investigations of large areas.

The sampling procedure in Dryer's approach is organized as follows (Dryer 1989a: p.267): first, languages are grouped by genealogical groups, roughly comparable, following Dryer's esemplification, to the Indo-European subfamilies, like Romance or Germanic. Each of these groups has been called *genus*, as suggested by Croft and extensively employed in WALS terminology.

By working on *genera* rather than languages, one can control for the most severe genetic bias, since languages within genera are generally fairly similar typologically (Dryer 1989a: p.268).

Subsequently, *genera* are divided into six large continental areas: Africa, Eurasia, North America, South America, Australia and Papunesia. In the present work, such areal sets slightly differ from Dryer's classification, since Australia and Papunesia are grouped together in a single Pacific macro-area.

These geographic boundaries are quite well defined, although there are some instances where the genealogical relations in languages redefine the boundaries: Semitic languages, for example, are treated as part of African languages, because their genealogical relationship bends toward that direction.

When establishing the coding conventions for my database, I used the language coding criteria of Glottolog, also widely used on WALS. Glottolog has

been a reliable source with respect to language classification in a strict sense, since it focusses on language grouping and it provides exhaustive information about languages yet not covered in WALS; moreover, it is constantly updated with the recent developments related to classification and subgrouping.

WALS and Glottolog's labels of classification differ: Glottolog discards the *genera* grouping, since it rather prefers a *isolate, sub-family, top-level-family* labelling. The definition *subfamily* is used to refer to all level of classification below *top-level-family*. This different labelling may rise issues of data comparability, which can be easily solved: "all WALS *genera* exist as subgroups in Glottolog: the genealogical stratifications followed in the two databases are thus largely comparable" (Francesca di Garbo, p.c.).

The sampling procedure just described gives a very useful approach when one needs to balance intra-genealogical and inter-genealogical perspectives; however, a third element crucial to the present analysis is the interest in investigating nominal number at its greatest possible variety.

Thus, this sample technically shows some of the features which pertain to the *variety sample* (Bakker 2007). The primary requisite for a language to be included in a variety sample is to display the variable in object, that, in this case, is nominal number marking. Thus, no language with no nominal number is added to the sample, neither as a control group; in turn, the main condition for its inclusion is the presence of at least one overt number marking construction in at least one nominal type, no matter how restricted is its distribution within the nouns.

Variety samples show two additional characteristics, both fully adopted in this sampling method: first, only a small number of languages from the same language family have been considered; second, language isolates were slightly preferred in the language selection process.

In building my language sample, I have collected at least one language per *genus*; in order to ensure some balance in my sample, I have collected a number of languages per language family in approximate proportion to the size of the family.

I did not follow any mathematical procedure to establish the exact number of languages that should be selected from each subgrouping, *genus* or language family. Thus, the language sample created for this research cannot be used and serve as basis for statistic analysis of the inferential type; the statistical data analysis that can be applied instead is rather descriptive: the obtained frequency distributions of types and patterns identified and identifiable are presented in the results chapters.

All these elements contribute to define this language sample as a convenience one: an additional variable that supports this label is the availability of linguistic data.

It is widely known in typological studies (that rely heavily on reference grammars and language descriptions) that the choice of language in a sample may be strongly influenced by the number and especially and the quality of available linguistic reference.

Grammars may vary to a considerable extent as to their degree of sophistication. This implies that languages with scarce information about the parameter taken into account are excluded and replaced; furthermore, some language families are underrepresented or not represented at all, hence some languages were included in the sample because they were the only ones for which information was available for a reasonable variables of number marking constructions.

All these constraints and the forced necessity to select languages introduce some kind of bias, and the sample selected for this research makes no exception and it also explains why some aspects, like the number of languages collected per language family does reflect an estimate rather than a mathematical approach.

The present research relies on two linguistic samples:

- The main sample, of about 250 languages: it constitutes the backbone of data analysis and it has been used to explore the synchronic distribution of the number systems in the languages of the world; the full survey can be found on Section 7.1;
- The database subsample: it includes about 160 languages, all selected from

the main sample. These languages have been coded in a typological database that it has been used to explore the number constructions within and across the languages.

Both samples and the full list of the languages included are reported in the Appendix A.2

3.4. Data collection

Linguistic data was gathered by consulting three main sources:

- Descriptive and reference grammars and materials, including texts;
- Consultation of native speakers
- Consultation of experts on a particular language or language family.

Not all three sources were available for each of the sampled languages. Most data has been collected during a six months research period at the Max Planck Institute for Evolutionary Anthropology in Leipzig, then at the Max Planck Institute for Science of Human History in Jena. I also profited from a month of research as an invited speaker at the Centre of Excellence for the Dynamics of Language in Canberra. In Leipzig and Jena I had the opportunity to gather data from native speakers and from experts of African language families; in Canberra I completed the sample by adding several languages from the Pacific area, profiting from feedbacks and inputs from fieldwork linguists.

Data were first collected in basic tables, organized by nominal type (nouns, pronouns, demonstratives). The languages selected for the subsample were coded in a nominal number relational database, to ensure better systematicity. An example of database language coding is presented in Section 4.5.3.

3.5. Summary

The sampling method followed in this research aims to combine intra-genealogical and inter-genealogical dimensions, without underestimating language diversity and variation.

The sample is biased both genealogically and bibliographically: from the former perspective, it is biased because it is built on the basis of genealogical relations between subsets groupings; furthermore, it is bibliographically biased since the sample was created on the basis of available sources and not by means of mathematical methods.

These biases imply that the present sample cannot be used in order to make statistical predictions; rather, this investigation aims to account for the identification, the distribution and the stability of recurrent associations between the parameters that play a role in nominal number marking (number values, nominal types, number marking constructions).

4. Parameters and methodological approach

4.1. Introduction

This chapter illustrates the parameters taken into account in this study and the methodological approach used to investigate their interaction.

The chapter consists in three main sections, organized as follows.

Section 4.2 deals with the parameters definition. The main parameters that serve as basis for this study are (i) the nominal types, (ii) the number values and (iii) the number markers. A section is dedicated to each parameter: the types of nouns taken into account are described and their segmentation in more fine-grained elements is described; a list of number values addressed in this research is given; number markers are labelled in types and their formal characteristics are illustrated.

Section 4.3 focusses on the methodological approach followed in this study, that is the multivariate approach. Section 4.3.1 outlines the recent goals and developments of linguistic typology from a methodological approach. Section 4.3.2 introduces the notion of multivariate approach to typology, discusses its major principles and the advantages of its application in the study of nominal number constructions.

Section 4.4 illustrates the application of the methodological framework to the parameters under investigation: the result is a three-level analysis, where the complexity of the number systems is segmented in three dimensions, and each of them explored progressively.

Finally, Section 4.5 illustrates the formal description of this method, whose visible result is found in the development of a typological database of nominal number constructions, built for the purposes of this study.

A summary of the chapter is provided in Section 4.6.

4.2. Parameter units and structural features

This section illustrates the parameters involved in this study, (i) the nominal types (ii) number values (iii) the number markers. These objects can be further divided in simpler elements, in accordance to their features. I what follows, I propose such a decomposition and discuss the criteria that drive it.

4.2.1. The nominal types

The first classification proposed within the noun types set follows the syntactic categories. At this level, it makes sense to define three principal groups:

- **Independent personal pronouns**, distinct in first, second and third person;
- **Nouns**, distinct in kin terms, human nouns, animate and inanimate;
- **Demonstrative pronouns**

Full nouns

The criteria of decomposing full nouns into subtypes are animacy, semantic as well as referential criteria. The motivation for these criteria is as follows. First, the relevance of this partition has been attested in the Animacy Hierarchy (see Section 2.2). Second, the subtypes of this kind are well attested cross-linguistically. The semantic categories of +KIN, +HUMAN, ±ANIMATE are cognitively motivated; what makes a human 'human' is a set of properties that typically does not belong to the language domain. However, morphosyntactic realizations may show interesting perspectives about the means by which language and culture do segment and represent reality: gods and spirits may be

valued as human beings in a language and therefore get the morphosyntactic treatment reserved to human nouns.

This approach in types and parameters selection is top-down, however, it has been combined, in the preliminary phase of research, with a bottom-up parameters identification, conducted on a pilot sample of about twenty languages, following the autotypologising approach.

Nouns referring to spirits, divine entities and high animates, although taken into account and noted separately, do not figure as parameters: the count of languages with specific markers for these noun types was too low to reach 'critical mass' and be promoted to a parameter status[1].

Proper nouns are also excluded from the parameters list; in fact, proper nouns, when modified by a plural marker, show a meaning which is not explicitly related to plurality, but it is rather connected to an associative value, e.g. 'X and associates', 'Y and others'. This trait pertains to the dimension of meaning, rather than of the morphosyntactic expression: including such instances as parameters would have entailed a strike compromise in systematicity.

Nonetheless, any additional or secondary meaning of a number marker (that usually include distributivity, collectiveness and associative value) has been annotated: although the typology outlined in this thesis is synchronic, detailed information about number markers may serve as basis to discuss peculiar synchronic distributions or to link them to diachronic data.

Pronouns

Among all the pronominal paradigms available in a language, this investigation singles out one type of pronouns, which are independent personal pronouns. The choice of selecting a specific pronoun type is due to multiple reasons. First, it ensures cross-linguistic comparability. Many languages lack non-

[1] Adding these variables to the parameters list would have caused redundancy in the database coding, due to the application to all languages of an attribute which would have been left empty in almost all languages. What is really relevant in this combined approach, as fully described in Section 4.3.2, is its ability to include a new parameter when it becomes useful and relevant, without the need to change the parameters that are already defined or altering the already systematized linguistic data

independent pronouns; moreover, independent personal pronouns are more easily identifiable in the grammar of a language. Second, independent person pronouns have already been partially approached and compared to other noun types in number marking literature (Haspelmath et al. 2005, Forchheimer 1953, Corbett 2000: e.g.).

The 'pronominal set' is defined by a number of 'properties', person and clusivity.

The association of a pronoun to a grammatical person is clear and cross-linguistically viable in tendency. Clusivity is an additional dimension that pertains to the interaction between a noun type (first person independent pronoun) and a number value (dual or plural). Clusivity feature shows a high degree of complexity for a comparative study. In the present approach, I have decided to explore the clusivity distinction in pronouns exclusively from a number marking perspective, and only in cases where the identification of clusivity and number markers is quite transparent, to avoid speculation and ambiguity. In this analysis, clusivity does not pose as a parameter *per se*: it is rather an attribute of the first person pronoun.

Demonstratives

The last parameter included in the noun types are demonstrative pronouns (demonstrative adjectives are excluded from this survey). At the state of the art, the literature available on demonstrative pronouns (Diessel 1999, Dixon & Aikhenvald 2003) has been focussing on semantic properties of demonstratives pronouns (like deixis and other qualitative traits), their pragmatic use (e.g., anaphoric) as well as their origin (grammaticalization from third person pronouns, determiners or relative pronouns). Their behavior with respect to number marking has been quite neglected, although studies have shown that demonstrative forms may play an important role for the rise and development of nominal number markers (Frajzyngier 1997).

In most cases, demonstrative pronouns show proximity distinction, usually close/far distance. Many languages have a tripartite system, with close/medi-

um/distal forms; however this variety has been discarded in the present analysis, since only number markers are taken into account; thus, the possible degrees of proximity are ignored, unless each deictic demonstrative form shows a specific number marker linked to the proximal feature: such case, however, has not once occurred in the 300 language main sample of the present study.

4.2.2. Number values

The second main parameter is the number values. In this study, I distinguish the following values:

- **Singular** (including both zero and overt expression)
- **Dual**
- **Plural**
- **Trial**, **Quadral**, **Paucal** (when applicable)

Singular, dual and plural are the 'standard' values on which any typological survey on number expression has extensively worked on, whose results have been summarized in the generalizations presented and described in the Number Hierarchy (see Section 2.2).

In the present work, data has been collected with a particular focus on the expression of the singular value: this can be zero marked or overtly expressed.

Values less common in the languages of the world, like trial and quadral, have been also included. The presence of paucal value is annotated, although results from data show the systematic and widespread distribution of this value in only one language of the sample (Mocovì, Guaicuruan, South America). Paucal is generally defined as 'a reference to a small number of specific world entities, usually between two and seven, sometimes up to fifteen elements' (Corbett 2000: p. 24); the exact quantity is almost never specified in grammars and reference materials; moreover, the concept of paucal is already quite broad and not detailed, since it implies the idea of a quantity which is blurred *per se*.

General number is not treated in the present investigation: Corbett (2000: p. 9) reports cases of languages where the expression of number is less prominent and therefore the meaning of a noun can be expressed without reference to number. This implies that a noun could denote an entity, but the speaker is not committed to specify the quantity: it could be one of more than that. For this reason, general number poses itself 'outside the number system' (Corbett 2000: p. 10) and therefore it is not included in the present survey.

4.2.3. Constructions

The last parameter are the constructions used to express the number values associated to the noun types.

Under the current approach, construction types are defined as the morphosyntactic strategies that express a specific number value in a defined context (the noun type(s)) they refer to.

Number markers can be described as a combination of a form and a type. A construction form is language specific — it is the explicit morphological (or non-morphological) material used in a language to address a number value, e.g., the plural marker *-s* in English. The construction type instead is a result of a tagging procedure which characterizes the marker in accordance to its morphosyntactic etc. properties. For instance, the English plural marker *-s* is tagged as a suffix.

A list of the number construction types is proposed; the following paragraph focusses on the complexity and the features that the construction forms can bear and show in languages.

Construction types

Literature on number lists the possible means of number expression, by grouping them in two main sets (Corbett 2000, Haspelmath et al. 2005):

- Strategies involving a modification in the morphological form of a noun.
- Morphemes that occur somewhere else in the noun phrase.

Among the first group of strategies one can distinguish:

- **Affixes** distinct in prefixes, infixes and suffixes. They are the most number marking strategies and involve the add of morphological material to the NP type;
- **Stem change** or stem modification, is the morphological process where the stem, or the root, is modified; it may co-occur with other morphological processes, such as affixation;
- **Suppletion** is the replacement of one stem with another, which has no phonological similarity to the previous one;
- **Reduplication** involves the repetition, partial of complete, of the root;
- **Tone** where number information is marked on the noun type through a pitch element.

The second group of strategies includes morphological material that occurs outside the noun phrase (Dryer 2013):

- **Lexical items**, that include specific words separate from the noun type;
- **Clitics**, syntactically free, but phonologically bound to another word of noun or the noun phrase.

These markers and the related labels and properties have been adopted 'in bulk' in the present survey. Although there is a current ongoing debate on the status of forms like clitics or affixes, this matter does not pertain directly to the purpose of this research: these definitions serve as labels for the markers individuated in a given language; moreover, in database design, such 'tags' are kept separately from the construction forms, in order to avoid early data aggregation and ambiguity.

Table 4.1.: Trumai (Trumai, South America), Guirardello 1999

	SG	DU	PL
1	ha	ka a (incl) ha a (excl)	ka wan (incl) ha wan (excl)
2	hi	hi a	hi wan
3	ine (m), inatl (f)	ink a	ink wan
DEM	ni'de (m), ni'datl (f)	ni'dak a	ni'dak wan
N animate	kiki 'man'	kiki a 'two men'	kiki wan 'men'

Structural properties of number construction forms: a general overview

Number markers in languages show a high degree of variety and complexity in the way the express number distinctions on the noun types. In what follows, I report a brief overview of the structural complexity of the number markers, that can be summarized with the following concepts:

Regularity

The most regular number system signals number expression through one marker per number values, spread regularly throughout the NP types. This structure, albeit the 'simplest', is not common among the languages of the world. Trumai language (Table 4.1) provides a good example.

Diversity

Most languages have different markers to express the same number value opposition: Burushaski uses more than ten different constructions to express nominal plural. Consider the following example.

(1) Burushaski Isolate, (Lorimer (1935: p.33))

a. dʌsin, dʌsi-wʌnts
girl, girl-PL
'girl, girls'

b. mi, mi-tsəro
mother, mother-PL
'mother, mothers'

c. ha, ha-kicʌŋ
house, house-PL
'house, houses'

d. hal, hal-jo
fox, fox-PL
'fox, foxes'

Distribution

Number markers display different distributions within the noun types across languages. In Navajo, the same marker is shared by specific NP types (inanimate nouns and pronouns) while other nouns take other markers.

(2) Navajo (Athapaskan), Young and Morgan (1972: 2)

a. kò, da-kò
fire, PL
'girl, many fires'

b. 'ashkii, 'ashkii-ké
boy, boy-PL
'boy, boys'

c. sik'is, sik'is-
sibling, sibling-PL
'sibling, siblings'

In Soteapan Zoque, human nouns and kin terms use the same marker; animates and inanimates signal plural through a specific construction.

(3) Soteapan Zoque (Mixe-Zoque), Elson (1960: 53)

a. yomo, yomo-tam
girl, girl-PL
'girl, girls'

b. itawa, itawa-tam
brother, brother-PL
'brother, brothers'

c. tak, tak-yah
house, house-PL
'house, houses'

d. yoya, yoya-yah
pig, pig-PL
'pig, pigs'

Quantity

Number categories can be expressed through complex and compositional constructions. Pronouns in Tu/Mangghuer language express dual through the simultaneous use of three different morphemes (suppletion, a lexical item and a clitic form). This is illustrated in the following example.

Table 4.2.: Witoto (Huitotoan), de Piñeros & Roselli 2000

	SG	DU	PL
1	kue	koko	kai
2	o	amiko	amoi
3	naiñeño	kip	pik
DEM	bie	/	/
N	rada 'branch'	/	radaei 'branches'

Table 4.3.: Gender/number markers in Tunica (Tunica) Haas (1946: p. 46)

	SG	DU	PL
M	-ku, -ku'hu	-ʔu'nima	-sɛ'ma
F	-hči -hči'hi	-si'nima	

(4) Mangghuer (Mongolic), Slater (2003:98)

beghe ge, beghe-si
tree SG, tree-PL
'tree', 'trees'

Complexity

Number marking in a language may show great inherent complexity, with the combination of multiple markers on the same noun type and strategies with a high fusional level, making the disambiguation of number markers almost impossible. Witoto language is an example of this.

Table 4.4.: Wadjiginy (Wadjiginy) number system Ford (1990b: p.48)

	singular	dual	trial	plural
1	nace	nana (incl) narra-kani (excl)	narra-pana-kani (F) narra-pana-kani (F)	narrara (incl) narra (exclu)
2	kane	nawarra-kani	nawarra-pana-kani (F) nawarra-pena-kani (M)	nawarra
3	camuyic (M) cenmiyic (F)	porra-kani	porra-pana-kani (F) porra-pena-kani (M)	parrmuyic

Cumulation

The last property that can be identified within the marker is cumulation; number constructions may express also other grammatical features, like gender, case and definiteness; number suffixes in Tunica (Table 4.3) express at the same time gender distinction; in Wadjigini (Table 4.4) trial markers in pronouns mark also gender.

All these elements illustrate two major issues in data analysis: first, an accurate and reliable description and coding of these structures; second, the cross-linguistic viability of data, that has to be pursued without leveling language specific traits, procedure that would ensure a better comparability but at the same time it would entail to the discard of linguistic data, compromising internal and language specific variation.

4.3. Methodological approach

4.3.1. Recent methodological developments in linguistic typology

Linguistic typologic has undergone through a systematic change in its goals and aims in the last decades. In the last century until recently, typology was mostly used to explore the limits of possible human languages and thereby contribute

to an universal theory of grammar. Nowadays, typology has fully emancipated from this framework, shifting to new goals that involve the understanding of linguistic diversity and therefore the development of theories able to explore the principles governing the distribution of structural features across languages. In other words, instead of asking what is possible or not, more and more typologists have started to ask ”what’s where why” (Bickel 2007). As Bickel points out, asking “what’s where” targets universal preferences as much as geographical and genealogical skewing; more important, asking ”why?” has its basis on the assumptions that typological distributions are historically grown and interrelated with other distributions as well.

Motivating distributions as historically grown means that the synchronic distributions are seen as a product of transition and diachronic processes. It is a matter of current debate whether universal preferences are guided by (i) general principles or (ii) locally motivated pathways of change. In order to answer such questions and to switch to this new framework, where finding and testing distributions is a priority, typological research focussed on developing variables that measure similarity and differences between languages. Furthermore, these variables and sets of variables, need to be close to observable data more prominently that in past research. The causes behind this decision are mainly two: first, the more abstract the variable is, the more difficult is to test it in large areas and language samples wider than the past; second, more general and abstract the variables, more complicated it becomes capturing linguistic diversity.

A huge step towards the answering of these questions was made easier by major advances in methodology. The past decade has witnessed an increasing number of typological databases, of which The World Atlas of Language Structures (WALS) provides the most prominent example. Simultaneously, statistical methods of exploring cross-linguistic distributions have been redefined and adjusted accordingly to the data at hand (Cysouw 2005: for an overview).

The most important consequence of this joint development of goals and advance in technology relies in the deep transformation that the very methodology in exploring cross-linguistic variation has undergone. Modern methods

cannot longer operate with mass types of traditional typology, but they instead work with much more fine-grained variables; a similar tendency has been identified in individual languages as well, moving away from broad typologizing (and thereby reducing linguistic diversity) to individual structural elements (like constructions or relations) which are comparable cross-linguistically and allow to measure linguistic diversity. How such complex linguistic structures can be compared without reducing diversity is discussed in the next section.

4.3.2. Analysis of individual structures: the multivariate approach

In order to capture and test the distribution of structural features in languages from a cross-linguistic perspective, the typological community uses variables (known also as parameters, properties, features) that allow to measure the degree of similarity and diversity between languages.

But how do we compare languages in order to measure such similarity (or diversity)? Two structures (S1 and S2) are similar if they are identical in some variables, but differ in other aspects.

In order to compare two structures, between languages and within the same language, one therefore needs to decompose such structures into sets of variables able to capture all the aspects which these structures can be identical or different. This procedure is known as multivariate analysis (Bickel 2010), who has extended the use of this term from statistical analysis of sets of variables.

Once this procedure has been explained, the next step is how to define a proper variable. Ideally, a set of variables should satisfy a number of criteria. First, it shall be large enough to capture the full range of linguistic diversity; however, research interests dictate upper limits. Moreover, what is practical important in the choice of sets of variables is the logical independence of one variable from the other, thus avoiding lumping together parameters that might have geographical or genealogical distributions or interesting individual historical profiles and thereby not cross-linguistic viable; on the other hand, the set of variables shall remain close to the observed data.

How are variables estabilished in a research or in a exploratory study? Variables do represent concepts called ”typological types”, ”values of typological features” or ”comparative notions” that are required for the analysis of at least one language. There are two possible strategies to establish such sets of typological types.

On one hand, they can be defined *a priori*: a set of crosslinguistic types is defined in a functional domain and each language of the sample is assigned to a type (such procedure is also referred as *conceptual space* and *etic grid*; the limitations of such approach are discussed extensively in Bickel & Nichols (2002) and Cysouw (2005)).

The approach in large part adopted in the database realization follows the ”autotypologizing method” (Bickel & Nichols 2002).

In autotypologizing databases, no *a priori* crosslinguistics types are assumed: lists of typological types are instead compiled in a bottom-up fashion during data collection as needed for distinguish observed structures. If an observed typological type is different from the ones present in the types list, a new type is established, and the previous types selected can be reconsidered, if necessary. This procedure can be completed without altering data already structured nor discarding it. This method seeks to abstract away from language-specific categories to exactly that degree that is needed to capture all language-specific distinctions of the phenomenon at hand encountered in a sample of languages. The resulting typological sets of variables have the advantages of being not defined a priori, but rather inferred inductively and not based on models of ”possible vs impossible” realizations. What happens in practice, the two strategies (a priori and by inductive derivation) are often used both and combined in compiling typological types: one starts in a top-down fashion with a predefined list of the commonest types based on previous investigations on the phenomenon under exam, but when a number of new unclassifiable types reaches ’critical mass’, it assurges to a full type status.

Table 4.5.: Huallaga Quechua, Weber (1989: p.37)

	SG	PL
1	noqa	noqanchi[a] noqakuna[b]
2	qam	qamkuna
3	pay	paykuna
N	wasi 'house'	wasi-kuna 'houses'

[a] inclusive 1 pronoun
[b] exclusive 1 pronoun

4.4. Three level of analysis

4.4.1. First level of analysis: NP types and number markers

The first level involves the interaction between two variables: the noun types and the construction types, at a language specific level. Each of these units, although defined as basic, are structurally rich: the essentiality relies in the relation between these units.

In the first level of analysis, each NP type is taken into account (e.g., third independent person pronoun) and linked to the construction type (e.g., suffix) used to express a given number value (e.g., plural). In Table 4.5 the nominal number system of Huallaga Quechua is presented as an example.

The first level of data exploration would output the following context for Huallaga:

Huallaga Quechua

Singular

- First person pronoun exclusive, marker: zero marking
- Second person pronoun , marker: zero marking
- Third person pronoun , marker: zero marking
- Kin , marker: zero marking
- Human , marker: zero marking

- Animate , marker: zero marking
- Inanimate , marker: zero marking

Plural

- First person pronoun inclusive, marker: suffix '-anchi'
- First person pronoun exclusive, marker: suffix '-una'
- Second person pronoun , marker: suffix '-una'
- Third person pronoun , marker: suffix '-una'
- Kin , marker: suffix '-una'
- Human , marker: suffix '-una'
- Animate , marker: suffix '-una'
- Inanimate , marker: suffix '-una'

In each language, there will be as many relations as noun types that express number distinctions. It is quite clear that this type of description exhibit at least three limitations:

- It is *redundant*: different noun types may express the through the same construction type;
- It is *sketchy*: many attributes are lumped together in one entity; 'first independent person pronoun inclusive' covers a variety of attributes (syntactic type, person, clusivity) and this merging limits the cross-linguistic viability of the variable;
- It is *not specific*: it does not capture properly the marking dimension. There is no information about the construction form. A 'suffix' can be either suffix in a language: such data is not intra-linguistic comparable.

Given these shortcomings, however, this first level analysis might prove useful in accounting for the following aspects:

- An exhaustive description of the presence of number values and distribution of number distinctions in the sample explored;
- Detailed occurrence of a construction type with a NP type and the infer-

ence of recurrent associations between a marker and a NP type.

4.4.2. Second level of analysis: referential types and construction forms

The second level of analysis aims to explore the plural markers in the number systems of given languages, with a particular attention on the variety of constructions that may be used to express plural and how they distribute within the NP types. This level of analysis introduces two main aspects.

First, the perspective adopted switches from the construction type to the construction form: the label attached to a marker, like 'suffix', is too ambiguous and not exhaustive enough to be taken into account. Languages may show different context in the plural distinction distributions: data exploration needs to be language specific. At the same time, data need to be compared between languages, in order to infer possible tendencies and types.

The necessity of capturing the internal variety of languages to the maximum extent and ensuring cross-linguistic viability at the same time, introduces the second aspect that pertains directly to this level of analysis, the concept of *referential type*.

A referential type may be defined as *a set of nominal expressions that exhibit identical behavior in regards to number marking*.

This concept allows to overcome the shortcomings of the first level of analysis and to undertake a detailed exploration over the targets of the second level of analysis. The example provided in the previous subsection illustrates the nominal number system in Huallaga Quechua. If this language is analyzed in terms of referential types, we can easily see that the only contrast it makes is between the first person pronoun inclusive (which is marked with '-anchi' in plural) and everything else (which is marked with '-una' in plural). Therefore, for Huallaga Quechua we have two referential types and can describe them as follows:

Huallaga Quechua *referential types*

RT1 Includes first person inclusive pronoun. Marked with suffix '-anchi' in plural

RT2 Includes first person exclusive pronoun, second and third person pronoun as well as all nouns. Marked with suffix '-kuna' in plural and zero-marked in singular

This model is able to describe all the possible number marking distributional constraints in a language. All the examples and cases presented in this work have been coded following the referential type model, without ignoring or discarding data. Moreover, this model is able to capture the most intricate structural features that are attested in number construction forms, which have been summarized and exemplified in in Section 4.5.3.

It has been mentioned how the referential types are viable cross-linguistically. This is possible since the NP types are not inherently defined. The features that belong to a NP type (syntactic type, animacy, gender, clusivity) do not constitute the nominal element itself: they are rather linked as additional properties to the noun, and do not interfere with the definition of the referential type. These dimensions, or attributes, when present, are *superimposed* on the referential sets, not inherent. This procedure allows the 'insensitivity' of the referential sets to the nominal types properties, implying that the referential types can be formed by nouns, pronouns, demonstratives, all of them combined or partially. The only restriction that applies is the constructions form used by the members of the set, that has to be shared. The full potential of this model is practically illustrated in Section 4.5.3.

The second level of analysis aims to illustrate the following contexts:

- Types of referential sets attested in the languages of the world;
- The cross-linguistic presence and distribution of the referential types.

Table 4.6.: Onge (Andamanese), Dasgupta & Sharma (1982)

	SG	DU	PL
1	mi	/	eti
2	ɲi	/	ni
3	gi	/	ekwi
DEM	li	/	/
N	inene-da 'foreigner'	inene-dena 'two foreigners'	inene-di 'foreigners'

Table 4.7.: Yupi'k, Miyaoka (2010: p. 350)

	SG	DU	PL
1	wii - wiiŋa	waŋkuk	waŋkuta
2	ɨɬpɨt	ɨɬpɨtɨk	ɨɬpɨci
3 (abs)	ɨɬii	ɨɬkɨk	ɨɬait
3 (rel)	ɨɬiin	ɨɬkɨnka	ɨɬaita
DEM	riya-u-na	u-ku-k	u-ku-t
N		-k	-t

4.4.3. Third level of analysis: a typology of the number systems

The third level of data exploration partially collects the results of the previous two and compares data at the whole number system level. It is possible to explore the number systems from both the number values perspective and the constructions used. Results from this type of survey may also reveal interesting information about the structural diversity of the number systems in the languages of the world. Two languages are provided as examples, Onge (Table 4.6) and Yupi'k (Table 4.7). These languages display the same number system, with an overt singular, dual and plural marking. However, their systems have developed following completely different 'paths'.

Onge number system has a SG (overt) vs DU vs PL distinction. The number

systems structure reveals that overt singular and dual values appear only on nouns and these values are not distinct elsewhere in the nominal domain. The demonstrative pronoun is insensitive to number marking, while pronouns display suppletive form throughout the whole paradigm. The constructions used to express each number value in nouns are suffixes, whose forms seem related: such parallelism in both types and internal distribution seems to suggest a develop of Onge number system that occurred independently on different 'paths' on nouns and pronouns.

Yupi'k number system distinguishes SG (overt) vs DU vs PL as well. Each number value is expressed on all NP types with the same distribution (except nouns, although this loss seems a recent development, confirmed by neighboring and genealogical affiliated varieties like Aleut,Bergsland 1997). Number markers are cumulative with absolutive case. Furthermore, they display the same construction type (suffixes). These similarities in the construction type, distribution and additional features carried by the number markers, suggest that the number values have developed jointly, from an unique process.

The last level of data exploration aims to describe the following contexts:

- Typology of the number systems and their areal distribution;
- Relevant features associated to specific number systems;
- Interactions between the number values, and the number values and the NP types.

4.5. Database realization

4.5.1. Introduction

A systematic collection of data on languages is essential on typological research to gain a better understanding of the principles governing the structural features and their interactions of a given phenomenon. An ideal tool to collect, organize and analyze data is a database. The use of databases in linguistics has grown exponentially in the last years, from small research projects on a

specific language or language feature (Everaert et al. 2009: a good collection is provided by) to large scale atlases and databases like WALS (Haspelmath et al. 2005). Database design has many practical benefits: it helps in organizing and systematizing data in a more efficient way, it improves the quality of the data collected and provides the ability of a analyzing data in a variety of ways. Nonetheless, building a database presupposes the creation and the development of a strong descriptive model that would reflect a specific system, in this case nominal number marking. A database has a high level of explicitness, which forces the researcher to think deeper at some aspects that in many cases could remain unnoticed (Everaert et al. 2009). In what follows, the overall design of the database is presented.

4.5.2. The typological database: a descriptive model

A typological database on nominal marking expression has been developed and populated with a subset of the languages from the main sample. At the state of the art, there are about 170 complete descriptions of languages in the database.

This database can be seen as a formal description of the methodological approach that serve as basis for the three level data analysis aforementioned. The model regards the three main parameters under investigation (nominal elements, marker, number category) as linked and connected through relations; i.e. a certain marker (or marker combination) is used to express a certain number category in respect to a certain reference type.

The reference types are defined as nominal expressions that exhibit identical behavior in regards to number marking. This concept serve as backbone of the database design, since it permits the intra-linguistic and cross-linguistics comparison which constitute the main outcome of such data collection.

This definition is fully operationalizable and therefore can be applied across languages in a transparent manner[2]. The reference types are further specified

[2] the database has been used to create a number of high-level aggregations (which are algorithmically defined within the R suite for statistical computing), with the goal of identifying cross-linguistic patterns in the domain of nominal number marking.

Figure 4.1.: Database representation for Huallaga Huánuco Quechua

```
Language = "Huallaga Huánuco Quechua"
ISOCODE = "qub"

#' ## Reference types
ref.N.X1pro.exc.X2pro.X3pro %=% ReferenceType(N,Pro(Person=1, Clusivity="
    exclusive"),Pro(Person=2),Pro(Person=3))
ref.X1pro.inc %=% ReferenceType(Pro(Person=1, Clusivity="inclusive"))

#' ## Markers
marker.kun %=% Marker(Type="suffix", form="-kun")
marker.anchi %=% Marker(Type="suffix", form="-anchi")

#' ## Distributions

describe("SG", {
  c(ref.N.X1pro.exc.X2pro.X3pro, ref.X1pro.inc) %marked_with% ZERO
})

describe("PL", {
  ref.N.X1pro.exc.X2pro.X3pro %marked_with% marker.kun
  ref.X1pro.inc %marked_with% marker.anchi
})
```

in respect to the syntactic and semantic properties of the items they represent, such as animacy or syntactic type.

This conceptual framework offers a number of advantages: i) linguistic data is captured and described in such a way that no data is discarded or ignored, ii) at the same time, the conceptual core of the model is abstract enough to ensure the validity of cross-linguistic comparisons, iii) as the database aims to represent the empirical data rather than lumped high-level aggregation. This makes my database easily extensible as well as highly reusable and sustainable. Because the raw data is always accessible, the database is compatible with a wide range of data analysis methodologies and can be reused for different purposes.

4.5.3. Example of database coding

Fig. 4.1 shows an example of database coding. The language illustrated is Huallaga Huánuco Quechua, which number system has been reported in Table 4.5 and discussed on page 46. The major source of the linguistic data is Weber (1989: p.37).

The data is inputted into the database in form of structured text. For each language, dedicated markup elements describe how many referential types are distinguished, which nominal types they consist of, which number markers exist and how specific referential types are marked in regards to specific number categories.

In the example Fig. 4.1, lines 5 and 6 describe the two reference types, one of which consists of first person inclusive pronoun and another of all other nominal types that exist within the language. The lines 9 and 10 describe the markers. Finally, line 14 and below describes how number is marked on individual reference types, that is, what is the relation between the number markers, number category and the reference type in the language.

4.6. Summary

In this chapter, the main parameters have been outlined and described.

To provide for a robust basis for comparison of number constructions across and within languages, the parameters must be decomposed into sets of variables. The necessity to develop a set of fine-grained and viable variables led to the adoption of the multivariate approach as a theoretical model, which benefits have been illustrated.

The application of the model in this study has been outlined from a theoretical and analytic perspective, whose results are seen in the decomposed three levels of analysis of the number marking domain.

The practical application of the model is shown in the coding structure of the database, which has been built and developed for the purpose of this research.

5. Nominal types and number markers

5.1. Introduction

This chapter deals with the first level of the analysis of the nominal number marking domain.

The parameters for this preliminary survey are the noun types, the construction types and the number values. The noun types addressed in this chapter are:

- Independent first person pronoun (with clusivity distinction, when applies);
- Independent second person pronoun
- Independent third person pronoun
- Kin terms
- Nouns denoting human entities
- Animate nouns
- Inanimate nouns
- Demonstrative pronouns

The number values targeted in this chapter are:

- Singular (both unmarked and overtly expressed)
- Dual
- Plural

- Trial

The construction types used as a result of the interaction between the NP types and the number meaning are illustrated throughout the chapter. It shall be kept in mind that at the present level of analysis only the types of markers are described, not the construction forms (the formal representation of the construction, language-specific and not labelled).

The main goal of this section is to explore the basic relationship between these entities. This relationship can be described by the following statement: in a language, in a given number value context, a specific construction type is used by a noun type to express that context.

The description outlined in this chapter is fully synchronic. Linguistic data used as source for this overview consists in the 160 languages annotated and coded in the database of nominal number marking (the database language sample is provided in Appendix A.2).

This chapter is structured as follows: Section 5.2 explores the expression of singular value in the nominal types. For convenience, each section is split according to the lexical categories (pronouns, nouns, demonstrative) addressed in the analysis. Tables and maps show the occurrence and the distribution of each interaction. Relevant cases from languages are presented as examples. The same procedure is carried up for dual number (Section 5.3) and plural (Section 5.4).

A summary of the chapter in provided in Section 5.5.

Table 5.1.: Singular marker types on pronouns

Marker Type	1 Pro	1 Pro excl	2 Pro	3 Pro
0	14	13	44	66
clitic	1	-	-	-
infix	1	-	-	1
stem modification	13	15	29	19
suffix	-	-	1	5
suppletion + suffix	-	-	-	1
suppletion	64	29	73	52
suppletion + infix	-	-	1	-
suppletion + tone	-	1	-	-
tone	-	-	1	1

5.2. Singular expression

Among the main facts attested in number marking, two pertain to the expression of singular in languages: (i) singular, more than the other number values, tends to be left unmarked, especially in nouns. On the other hand, (ii) pronouns have a tendency to show suppletive person/number stems, implying the presence of suppletive forms in singular as well.

This tendency is confirmed by a preliminary data survey on the languages of the sample. In what follows, I will show how overt singular expression is not common among languages, and how it is even less widespread on specific lexical categories, like pronouns.

5.2.1. Singular marking in independent personal pronouns

Table 5.1 reports the occurrences of the marker types used for singular expression in the 160 languages of the sample:

The well-known tendencies aforementioned are confirmed by the figures outlined in Table 5.1. The most common marking strategy is suppletion and other

stem modifications; this applies to all pronominal persons, from first to third, with no exception. Marking strategies that involve the addition of morphological material to the pronominal root are quite rare: the marker types involved are suffixes, infixes. Tone is attested in one language. Zero marking in pronominal forms is attested on all pronominal forms, and increases moving to the rightmost side of the referential hierarchy.

In what follows, each nominal types is explored in detail.

The most widespread singular marking strategy in first person pronoun singular (with first pronoun exclusive, that includes languages with clusivity distinction in first person pronoun) is suppletion (paired with modification of the stem). Such strategies are attested in approximately 136 languages out of 160. The second most common strategy is the bare root left unmarked: it is attested in 14 languages of the sample, and it corresponds to a marked counter-strategy in plural through affixation. The overt markers found in singular are suffixes, infixes and tone. These markers appear on the whole pronominal paradigm of the respective languages: they are found also in second and third person pronouns.

The behavior of second person pronoun is striking similar to the one assessed for the first person form. The non-suppletive strategies found have the same occurrence and they are attested in the same languages of the first person pronoun, indicating the high specificity of the phenomenon. The main difference in singular marking between the first and second forms relies on the counting occurrences of suppletive VS unmarked forms: zero marked forms in second person pronoun are found in 44 languages, with 27 occurrences only in first person pronouns.

The most common singular marking strategy in third person pronoun is suppletion (with stem alternation altogether): it is attested on 72 languages; zero marking is strongly attested, with 66 occurrences. Overt singular is expressed by suffixes in 2 language. The other strategies (tone, infix) are shared with first and second person pronouns, so they are attested within the same languages.

Among the languages that show zero-marking in pronouns, cases can be

Table 5.2.: Zero marking in singular independent personal pronouns

Language	Person	Form	Type
Sierra Popoluca (Mixe-Zoquean)	1	ÃČ	UNMARKED
Armenian (Indo-European)	2	du	UNMARKED
Babine (Athabascan)	3	ʔen	UNMARKED

mentioned from Sierra Populuca (Mixe-Zoquean, Elson 1960), Armenian (Indo-European, Dum-Tragut 2009) and Babine (Athabascan, Hargus 2007) with unmarked singular forms in first, second and third independent pronouns respectively (Table 5.2).

The examples provided come from three different pronominal persons from three languages: the main purpose is to show how every pronominal form can be found unmarked. However, the presence of an unmarked singular in one specific pronominal person in a given language does not imply the same marking strategy on the whole pronominal paradigm[1].

Table 5.3 illustrates three cases of suffixes, affixes and tone found in the languages of the sample: Cavineña (Pano-Tacanan, Guillaume 2004: p.78) uses suffixes, Nandi (Nilotic, Creider & Creider 1989: p. 57) signal singular marking though infixation, Northern Pumi (Sino-Tibetan, Ding 1998: p. 90) indicates singular by tone.

It is worth noting how, when morphological strategies are attested, they are widespread on the whole pronominal paradigm. The addition of morphological material as a singular marking strategy is restricted to a specific pronominal person only in third person, although quite uncommon. Singular overt marking through suffixes in third person is found in 2 other languages: they are Wanano

[1] Languages presented as examples are representative of these different contexts that can be observed within the pronominal paradigm: while Sierra Populuca shows zero marking on all pronominal singular persons, Armenian and Babine use different number marking strategies within their respective pronouns. In Armenian, second and third person singular pronouns are zero-marked; on the other hand, first person singular is suppletive. In Babine, only third person pronoun singular is unmarked, while independent first and second pronouns select suppletion and stem modification respectively.

Table 5.3.: Overt singular marking in independent personal pronouns

Language	Person	Form	Type
Cavineña	1	ike	SUFFIX
	2	mike	SUFFIX
	3	tuke	SUFFIX
Nandi	1	áne:	INFIX
	2	ínye:	INFIX
	3	íne:	INFIX
Northern Pumi	1	3H	SUPPLETION AND TONE (HIGH)
	2	neR	TONE (RISING)
	3	niF	TONE (FALLING)

(Tucanoan) and Blackfoot (where demonstrative pronoun is used as free third person pronoun).

5.2.2. Singular marking in full nouns

Singular overt marking in nouns reveals more variety when compared to pronouns: both morphological and non morphological means are used to express singular value. Singular markers can be restricted to a specific noun type or they can be spread to all nouns with no respect to animacy. Moreover, specific different singular marking strategies may be identified among the full noun types within a language.

Table 5.4 reports the occurrences of overt singular marking in full nouns.

The most common strategy is zero-marking: nouns are left unmarked, fact that confirms the facts well known and attested on number marking.

Suffixal strategy is the preferred morphological mean of singular expression: suffixes may attach to any nominal type, suggesting that they do not constitute as a 'preferred marker' of a specific noun type. Similarly, tone covers all noun types and so does the only non morphological mean of singular expression, the 'singular word' or lexical item that signals singular number in nouns, e.g. in

Table 5.4.: Singular marker types on nouns

Marker Type	Kin Terms	Human Terms	Animates	Inanimates
0	133	133	133	96
lexical item	2	2	2	2
prefix	1	1	1	-
stem modification	7	7	-	-
suffix	9	9	9	5
suppletion	20	20	-	-
tone	1	1	1	1

Table 5.5.: Overt singular marking in nouns

Language	**Form**	**Type**	**Animacy spectrum**
Wanano (Stenzel 2004: p.161)	-ro	SUFFIX	HUMAN
Blackfoot (Frantz 1991: p. 21)	-ists(i)	SUFFIX	INANIMATES
Duna (San Roque 2008: p.182)	-na	SUFFIX	ANIMATES, HUMAN, KIN
Bagirmi (Stevenson 1969: p.28)		SUPPLETION	HUMAN
Kumiai (Miller 2001: p. 114)		STEM MODIFICATION	KIN TERMS
Sipacapense (Barrett 1999: p.63)		STEM MODIFICATION	HUMAN, KIN
Ngiti (Kutsch Lojenga 1994: p.132)	mu-	PREFIX	HUMAN
Masalit (Edgar 1989: p.46)		TONE	ALL NOUNS
Yapese (Jensen 1977: p.154)	rea	LEXICAL ITEM	ALL NOUNS

Yapese (Austronesian).

The only strategies (pertaining to singular value) which seem to be related to a specific noun type are suppletion and stem alternation, used with to human and kin terms: the phenomenon is frequently attested in languages and areas of the world, like in Bagirmi (Africa), Kumiai (North America) and Sipacapense (South America).

Table 5.5 describes examples from the construction types that are found for

each nominal type in the languages explored. Kin terms and human nouns may indicate singular value with suffixes (Duna, Duna), lexical items (Yapese, Austronesian), stem modification (Jamul Tiipay, Yuman and Sipacapense, Mayan), tone (Masalit, Maban), suppletion (Bagirmi, Bongo Bagirmi) and prefixes (Ngiti, Lendu). Animate nouns are found with suffixes (Duna), tone (Masalit) and lexical items (Yapese). Singular marking in inanimate nouns has been found expressed with suffixes (Blackfoot, Algonquian), lexical items (Yapese) and tone (Masalit).

When approaching this piece of information, an important disclaimer applies: the specific forms in the examples provided do not necessarily imply a restriction in the number marking distribution in the language. For instance, the inanimate suffix in Blackfoot shown in Table 5.5, does not mean that only inanimate nouns do show singular marking in the given language: it rather signals, for the purpose of this subsection, that inanimate nouns may have reserved markers, and the ones found are all suffixes.

Ngiti language (Nilo-Saharan) uses a prefix to express singular number: such marker is restricted to human nouns only. Another language that similarly employs prefixes to signal singular value is Tuscarora language (Iroquoian): in the specific case of the North American language, the prefixes are three: two are animate specific, and further indicate gender distinction; the other marker is used on inanimate nouns.

5.2.3. Singular marking in demonstratives

Demonstratives do not show great variety in the construction types used to express singular value. Table 5.6 outlines the occurrences of singular overt marking in demonstratives attested in the languages of the sample.

The constructions found include stem modification (Present Andamanese, Andaman), suppletion (Sapuan, AustroAsiatic) and suffixes (Wanano, Tucanoan) and are exemplified in Table 5.7.

The relative low counting occurrence is related to the tendency of demonstrative pronouns to be indifferent to number marking distinctions. The most

Table 5.6.: Overt singular marking — types of markers (Demonstratives)

Marker.Type	DEM
0	70
infix	1
stem modification	8
suffix	5
suppletion	17

Table 5.7.: Overt singular marking in demonstratives

Language	**Form**	**Type**
Andamanese (Abbi 2009: p.113)		STEM MODIFICATION
Sapuan (Jacq & Sidwell 1999: p.27)		SUPPLETION
Wanano (Stenzel 2004: p.161)	-ro	SUFFIX

common strategy is zero marking, followed by suppletion and stem modifications. Suffixes are attested in Cavineña (Paco-Tacanan), Wanano, Blackfoot and Kambera (Austronesian): in all cases, these demonstrative forms function as third person pronouns in their respective languages.

5.2.4. Summary

Singular expression is, in most cases, indicated by suppletive forms (and various modification of the stem) in pronouns and by zero marking in nouns.

Non suppletive singular overt strategies that have been found in pronominal forms include mostly suffixes. Alternative marking strategies include infixes (Nandi, Nilotic) or tone (Northern Pumi, Qiangic). No other markers have been attested in the sample. Less variety is shown by demonstrative forms, where the constructions attested are suppletion and suffixes. Nouns indicate singular number through a quantity of markers, from prefixes to lexical items. Human

nouns and kin terms display the richest variety of construction types, and they tend to show, more than other full nouns type, suppletive or modified stems to express singular.

5.3. Dual expression

Dual number, because of its specificity, is always overtly expressed. There are two main aspects connected to a classification of dual expression from a formal perspective; the first relies on the richness of the construction forms that are used to express dual, complexity that is revealed mostly in pronominal forms, where dual may be indicated through combined constructions and particles on the same nominal element.

The second aspect involves specifically pronouns. It has been mentioned how pronouns tends to mark plural through suppletive forms; dual category makes no exception in this regard. Among the most widespread constructions for dual value, suppletive person/number stem are included, to which another marker, usually a suffix, is added[2]. The root stem may follow two different 'suppletion paths': (i) it can be either identical (or strictly related) to the plural person-/number stem, developing a non-singular generic stem, where dual and plural may be disambiguated through additional markers; or (ii) dual person/number stem may derive or be directly connected to the related singular form, and dual disambiguation is ensured by a specific marker.

5.3.1. Dual expression in independent person pronouns

Table 5.8 illustrates the occurrences of dual construction types in pronouns. From the figures provided in table, it is noticeable the variety and complexity of the marking constructions:

The most common strategy is suppletion, with the exception of third person

[2] However, 'plain' dual suppletive forms, without any additional marker, are also common and they can be specific for dual and unrelated to both singular and plural person / number stems: this happens in Nivkh -Isolate, (see Gruzdeva 1998: among others)

Table 5.8.: Dual marker types on pronouns

Marker Type	1 Pro	1 Pro excl	1 Pro incl	2 Pro	3 Pro
clitic	1	-	-	2	1
clitic+tone	-	-	-	1	1
infix	-	-	-	1	-
lex. item	1	1	-	2	2
lex. item + clitic	-	-	-	-	1
stm mod. + lex. item	-	-	-	-	1
reduplication	-	-	-	-	1
stm mod.	1	2	4	4	2
stm mod + infix	-	-	1	-	-
suffix	2	4	3	9	13
suffix + lex. item	-	-	1	-	-
stm mod + lex. item	2	4	4	5	6
splt + suffix	4	4	5	9	7
splt + lex. item + suffix	-	1	-	-	-
suppletion	8	6	8	10	9
splt + clitic	-	1	-	-	1
splt + clitic + tone	-	1	1	-	-
splt + infix	1	1	-	1	1
splt + lex. item	-	-	1	1	-
splt + lex. item + clitic	1	-	-	1	-

Table 5.9.: Dual marking in independent pronouns, combined and regular marking strategies

Language	Person	**Form**	**Type**
Ineseño (Applegate 1966: p. 168)	1	k-iskiˀ	SUPPLETION + INFIX
Tu (Slater 2003: p.83)	1	da ghu=la	SUPPLETION + LEXICAL ITEM + CLITIC
Northern Pumi (Ding 1998: p.90)	1	eL-dzɑ̃H	SUPPLETION + CLITIC
Abun (Berry 1995: p.44)	2	nin-ka-we	SUPPLETION + CLASSIFIER + SUFFIX
Wadjiginy (Ford 1990a: p.97)	3	porra-kani	SUPPLETION + SUFFIX
Trumai (Guirardello 1999: p.27)	3	inak a	STEM ALTERNATION + LEXICAL ITEM
Nung (Sun et al. 2009: p.72)	1	ŋɑ31-iɯŋ55-si31	LEXICAL ITEM + SUFFIX
Atzingo (Muntzel 1986: p.89)	all	-kwe	SUFFIX
Yokuts (Gamble 1978: p.101)	1, 2	-aḵ̓	SUFFIX
Angami (Giridhar 1980: p.32)	all	=niē	CLITIC
Matses (Fleck 2003: p.120)	all	daëdi	LEXICAL ITEM
Maung (Capell & Hinch 1970: p.54)	3	janadinjanad	REDUPLICATION

pronoun, which shows a slight tendency toward suffixes. First person pronoun is the nominal type that shows the most heterogeneous variety: beside suppletion and stem modification, composite strategies that involve lexical items, clitics, suffixes and infixes are attested. In some cases, dual marking includes three different morphemes. Second person pronoun follows such behavior, and third person pronoun slighlt prefers 'regular' constructions, a tendency that increases in plural marking.

Examples from languages are provided in Table 5.9 and describe more efficiently these contexts.

In Ineseño Chumash (Chumashan), the first person dual pronoun is made of a non-singular stem, */-k-/*, a straightforward dual infix */-is-/*, and a formative pronominal suffix, */-ki/*. In Tu (Mongolic), dual in pronouns is formed by a non-singular stem plus a lexical item */gha/* originally meaning 'two' and a

comitative enclitic particle, */=la/*. Northern Pumi (Sino-Tibetan) shows a non-singular base, disambiguated in dual by the mean of a clitic, */=dzɑ̃H/*. This clitic is not related to the numeral form and its origin is uncertain.

Abun (West Papuan) and Wadjiginy (Wagaydyic) display both dual forms (the examples are from first person and third respectively) made of a non-singular root and a suffix (in Abun a classifier meaning 'person', *ka*, is also present). Lexical items can co-occur in dual marking strategy along with a modified stem: this is the case of Trumai language (Isolate, South America). In Nung (Sino-Tibetan), dual is expressed by adding to the person stem the plural suffix (*/ɳɯŋ55/*) and the dual specific marker, */si/*.

Both morphological and non-morphological regular strategies in dual marking are also attested. They are illustrated in the second section of Table 5.9. The non suppletive strategies used to signal dual are suffixes, clitics and lexical items. No other construction types have been attested among the languages of the sample.

There is one case of reduplication, in Maung (Iwaidjan), restricted to the third person pronoun (first and second person are insensitive to dual marking): the form *janadinjanad* is the repetition of the third person singular pronoun. Thus, the dual form is derived from the singular (*janad*, as expected by the iterative nature of the strategy). The plural form is *wenad*.

The other relevant aspect in the expression of dual in pronouns is the selection of the stem root. In most languages, the suppletive root used in dual pronouns is identical to the plural stem, configuring a singular VS non-singular opposition, where dual and plural are further disambiguated by the use of additional markers. In a number of languages, dual stem is identical to the singular root, in opposition to a plural suppletive stem, building a singular and dual VS plural opposition. Table 5.10 reports a number of cases where dual stems are linked or derived from the singular roots.

For instance, Hainan Cham (Austroasiatic), shows a first person dual pronoun formed by the singular root stem, the lexical item *thua*, 'two' and the suffix *za:ŋʔ*, which means 'people'. In Nahali (Isolate), the second person form shows this

Table 5.10.: Dual marking in independent personal pronouns, dual marker relates to singular marker

Language	**Person**	**Singular**	**Dual**	**Plural**
Hainan Cham (Thurgood 2010: p. 146)	1	kaw	kawthuaza:ŋʔ	**taza:ŋʔ** (incl) **mi** (excl)
Nahali (Kuiper 1962: p.27)	2	nē	nēko	**lā**
Angami (Giridhar 1980: p.32)	3	puô	puôniē	**ūkô**

Table 5.11.: Dual marker types on nouns

Marker Type	Kin Terms	Human Terms	Animates	Inanimates
clitic	4	4	4	4
lexical item	3	3	3	1
suffix	13	13	13	5
suppletion	2	2	2	-

form, with a dual suffix *-ko* added to the singular form, and a suppletive and unrelated plural construction.

The last example is taken from a Sino-Tibetan language, Angami: the third person pronoun dual involves the presence of a clitic added to the singular base, */=nīe/*, probably related to the numeral form 'two', *kenie* to which is in complementary distribution. However, what makes the case of Angami particularly interesting is the presence of an alternative third person dual form, *uniē*, likely derived from the plural base. No distributional constrains of these two forms are mentioned in the grammar.

5.3.2. Dual expression in full nouns

Table 5.11 reports the occurrence of dual construction types on full nouns.

The most common dual markers on nouns are suffixes. Less common strategies include non-morphological constructions like lexical items and clitic forms. The presence of dual marking tends to be regular on animate nouns. Examples

Table 5.12.: Dual marking in nouns

Language	Form	Type	Animacy Spectrum
Wagaya (Pama-Nyungan) (Breen 1974: p.55)	-wiy	SUFFIX	kin
Kuot (Kuot) (Lindström 2002: p.2)	-ip-ien	SUFFIX	animates, human, kin
Lavukaleve (Central Solomon) (Terrill 2003: p. 95)	-l	SUFFIX	all nouns
Angami (Sino-Tibetan) (Giridhar 1980: p.28)	=nīe	CLITIC	all nouns
Bilua (Central Solomon) (Obata 2003: p. 54)	kidi	LEXICAL ITEM	human
Yapese (Austronesian) (Jensen 1977: p.154)	mu-	LEXICAL ITEM	all nouns

Table 5.13.: Dual marker types on demonstratives

Marker Type	DEM
clitic	2
stem modification	1
suffix	11
suppletion	3

of dual construction types found in nouns are summarize in Table 5.12.

The examples presented show that dual markers may be added to the bare singular base, like in Lavukaleve (Central Solomon); on the other hand, Kuot (Kuot) dual nouns show both the dual (*-ien*) and the plural suffixes (*-ip*).

5.3.3. Dual expression in demonstratives

Table 5.13 illustrates the occurrence of dual markers types in demonstrative pronouns.

Dual expression in demonstratives is not common. The strategies found among the languages of the sample include suffixes (Duna, Isolate, Papua New Guinea), stem modification (Bilua, Isolate, Papua New Guinea) and lexical items (Trumai, Isolate, Brazil). The clitic form is found in Cavineña (Paco-Tacanan),

Table 5.14.: Dual marking in demonstratives

Language	Form	Type
Bilua (Obata 2003: p. 54)		STEM MODIFICATION
Trumai (Guirardello 1999: p.27)	a	LEXICAL ITEM
Duna (San Roque 2008: p.150)	-yane	SUFFIX

and it corresponds, also in this case, to the third person form. Examples are illustrated in Table 5.14.

5.3.4. Summary

Dual constructions in pronouns are particularly composite. This richness relies on a series of different aspects: (i) dual constructions in pronouns are in a quantity of cases modular, with a number of markers carried cumulatively by the pronominal form; (ii) suppletive forms are widely attested: dual stems may be formally independent from both singular and plural roots or they may either follow the plural (non-singular) root (in most cases) or adopt the singular stem. Moreover, also 'regular' plain dual markers are found, like suffixes or lexical items.

Dual marking is attested on all NP types; the constructions attested involve suffixes, lexical items and clitics. Reduplication as a dual signal has not been found on the languages of the sample, although it is attested in pronouns (one case, restricted to third person pronoun, in Maung).

Dual distinctions are less common in demonstratives. The construction types found are mostly suffixes; lexical items and stem modification are attested.

5.4. Plural expression

Plural is the number category that offers more variety in number marking constructions.

Table 5.15.: Plural marker types on pronouns

Marker Type	1 Pro	1 Pro excl	1 Pro incl	2 Pro	3 Pro
0	-	-	-	-	1
clitic	1	-	-	2	1
infix	1	-	-	2	2
lexical item	3	2	1	6	10
prefix	-	-	-	1	1
stem modification	9	13	10	20	14
suffix	11	10	5	31	58
stem modification + suffix	5	4	5	10	6
splt + suffix	8	9	10	13	5
suppletion	49	16	22	57	43
splt + clitic	1	1	2	2	2
splt + clitic + tone	-	1	1	-	-
splt + lex. item	1	-	1	2	-
splt + prefix	2	-	-	1	-
tone	-	1	-	1	2

5.4.1. Plural expression in independent personal pronouns

This subsection explores the constructions used to mark plural number in pronouns. As seen in the subsections pertaining to singular and dual, number in pronouns is likely to be expressed by suppletive forms, often combined with other strategies, like affixes. Plural number offers the richest complexity in construction types: all pronominal forms may express plural through a person-/number stem and morphological and non-morphological constructions.

Table 5.15 reports the occurrence of plural marker types in pronouns.

In Table 5.15, the suppletive forms and the affixes that cluster together to express plural, are collapsed together. For convenience, data discussion in pronouns is split in three subsections, treating each pronominal form separately.

Table 5.16.: Plural expression in first person pronoun

Language	singular	dual	plural
Kokborok (Sino-Tibetan, Eurasia) Karapurkar (1976: p. 31)	aŋ	-	čɯŋ
Nahuatl (Uto-Aztecan, Central America) Launey & Kraft (1992: p. 35)	neʔhuatl	-	teʔhuantin
Sapuan (Austroasiatic, Eurasia) Jacq & Sidwell (1999: p. 26)	ʔăj	-	mu ɲa
Ineseño Chumash (Chumashan, North America) Applegate (1966: p. 168)	noʾ	kiskiʾ	kiykiʾ
Navajo (Na-Dene, North America) Creider & Creider (1989: p. 8)	shí	nihí	danihí
Xerente (Macro-Jé, South America) de Sousa Filho (2007: p.118)	wa	-	wanõri
Didinga (Nilo-Saharan, Africa) Rosato 1980: p. 12	nana	-	naga
Iu Mien (Hmong Mien, Eurasia) Court (1985: p. 113)	yia	-	yia bua

Plural expression in first person pronoun

As shown in Table 5.15, the most common plural marking strategy in first person pronoun is, among the languages of the sample, bare suppletion. This strategy, grouped together with stem modification and including languages with an inclusive and exclusive form, is attested in more than 100 languages of the sample. Plain suffixal strategy is also attested. Combined strategies are also present, and the most common plural marking of this type involves the presence of a suppletive form and a suffix. Examples from languages are provided in Table 5.16.

Kokborok provides an example of suppletion in first person pronoun. Complex and composite constructions that comprise a suppletive person/number stem and the presence of an additional marker are common. Many combinations are attested among the languages of the sample. Nahuatl signals plurality in first person pronoun with a suppletive form and an additional suffix, *-huantin*. In Sapuan, the plural suppletive first person pronoun coexists with a lexical item which precedes the root form, *mu*. Chumash distinguishes a non-singular stem, used on both dual and plural; the specific plural marker is an infix, *-iy-*. Similarly, Navajo language shows a non-singular person stem; plural is disam-

Table 5.17.: Plural expression in second person pronoun

Language	singular	plural
Dime (Afro-Asiatic, Africa) Seyoum (2008: p. 65)	yáay	yesé
Midob (Nilo-Saharan, Africa) Werner (1993: p. 36)	íin	ùnngú
Pacoh (Austroasiatic, Austronesia) Alves (2000: p. 64)	maj	ʔipɛː
Kokborok (Sino-Tibetan, Eurasia) Karapurkar (1976: p. 31)	nɯŋ	nɔ- rɔk
Karo (Tupian, South America) Gabas (1999: p. 49)	ẽn	*kaʔto*
Choctaw (Muskogean, North America) Broadwell (2006: p. 93)	chishno’	hachisno'
Huastec (Mayan, Central America) Edmonson (1988: p. 133)	tataːʔ -	tataːʔčik
Angami (Sino-Tibetan, Eurasia) Giridhar (1980: p. 32)	nō	niēkô
Miskitu (Misumalpan, Central America) Salamanca (1988: p. 250)	man	man nani

biguated by the prefix *daa*. ‘Regular’ plural markers on first person pronoun include suffixes, like in Xerénte. Didinga language indicates plural through the infix form /-*g*-/.

Plural value may be expressed in first person pronoun by non morphological strategies. A good example is found in Iu Mien language, where the first person pronoun carries the plural word *bua*.

Plural expression in second person pronoun

Table 5.15 reports data also for second person pronoun. The behavior that can be observed is similar to what found in first person pronoun. Suppletion and stem modification are the most common strategies. Second person pronoun follows the trend, already attested in first person and increasing in third, of bare suffixes as plural markers in pronouns.

Table 5.17 illustrates some typical cases of plural expression in the second person pronoun. Suppletive plural forms are also attested, as shown in Dime. Plurality in second person pronouns can be expressed by complex constructions: modification of the stem can combine with a suffix, as in Midob. Sup-

Table 5.18.: Plural marking in independent personal pronouns

Language	Person	Form	Type
Zoococho Zapotec (Sonnenschein 2005: p.31)	3	lhegake'	INFIX
Ngiti (Kutsch Lojenga 1994: p.192)	3	abádhí	TONE
Mumuye (Shimizu 1983: p.57)	3	yi	SUPPLETION/LEXICAL ITEM
Navajo (Muntzel 1986: p.8)	3	daa-ho	PREFIX

pletive forms may cooccur with a suffix, a lexical item or clitic: such cases are found in Pacoh, Kokborok and Karo, respectively. In Karo, plurality in second person pronoun is indicated by a suppletive form and a clitic.

Among the morphological means of plural marking, both prefixes and suffixes are included: Choctaw indicates plurality with the prefix particle *ha*; in Huastec, a suffix *-čik* is used.

Clitics and lexical items (plural words) may express plurality in second person pronouns: in Angami, the clitic particle *-kô*, derived from the plural article, expresses plurality in some restricted nouns and second person pronoun. Miskitu language marks plurality in pronouns by the means of a plural word, *nani*.

Plural expression in third person pronoun

Independent third person forms show less suppletion in plural than the respective first and second persons. Table 5.18 summarizes the 'regular' construction types attested in the language sample. Tone, prefixes, suffixes and affixes are attested.

The third person plural form in Mumuye (Niger-Congo) is defined as a lexical item, although it is a suppletive forms with respect to the singular counterpart. Such 'labelling' choice is justified by the function of the third person pronoun, that is used on noun as a plural word.

Table 5.19.: Plural marker types on nouns

Marker Type	Kin Terms	Human Terms	Animates	Inanimates
circumfix	1	1	-	-
clitic	10	10	10	10
clitic+prefix	1	1	1	-
infix	1	1	1	-
lexical item	18	17	18	10
prefix	6	6	6	3
reduplication	13	15	14	9
stem modification	6	5	-	-
suffix	108	107	111	74
suffix + rdp	1	1	1	-
suppletion	17	21	18	2
tone	3	3	3	2

5.4.2. Plural expression in full nouns

Table Table 5.19 illustrates the counting occurrences of marker types in full nouns. The most widespread strategy are suffixes, widely attested on all NP types. Lexical items are also quite homogeneous: their regular occurrence on all the noun types indicates that plural words, when used, do appear on all the nominal types with no restrictions. Clitics show an in interesting feature on this regard: when the are attested, their presence is spread on all nouns. More specific strategies include suppletion, that, like the singular counterpart, is restricted to human nouns, as well as other strategies like circumfixes. Reduplication is used more frequently on animate nouns.

Table 5.20 provides a quantity of representative cases of plural construction types on nouns.

The most common plural marking strategy are suffixes. Infixes, prefixes and circumfixes are also attested. Suppletion and other modifications of the stem, like tone, are among the means plurality is addressed in nouns, as well as redu-

Table 5.20.: Plural marking in nouns

Language	Form	Type	Animacy Spectrum
Apuriña (Facundes 2000: p.264)	-wako-ro, -wako-ru	SUFFIX	human
Bargam (Hepner 2006: p.30)	-an	SUFFIX	kin
Burushaski (Lorimer 1935: p. 26)	-do	SUFFIX	animates only
Abun (Berry & Berry 1999: p.21)		TONE	all nouns
Sakao (\cite)1564	r-	PREFIX	kin
Bilua (Obata 2003: p. 54)		REDUPLICATION	human
Northern Paiute (Thornes 2003: p. 100)		REDUPLICATION	kin + human
Svan (Schmidt 1991: p. 18)	la- -r	CIRCUMFIX	kin
Didinga (Rosato 1980: p. 7)	-g-	INFIX	kin
Biak (van den Heuvel 2006: p. 101)	=na	CLITIC	inanimates
Ambulas (Wilson 1980: p. 36)	béré	LEXICAL ITEM	all nouns
Anufo (Smye 2004: p.12)	n'-=m	PREFIX + CLITIC	human

plication. Plural words and clitics are also found. The following table reports some examples: Plural infixes have been found only on kin terms (Didinga, Nilo-Saharan) and other particular constructions like circumfixes appear on kin (Svan, Kartvelian) and human terms (Atayal, Austronesian): in both cases, these markers are connected to verbal forms. In Atayal, for example, the circumfix */kin- -an/* is used on verbs which have the meaning of 'have power over' to form nouns that denote 'the ruled area' (Rau 1992: p. 116). This marker is also used on human nouns to express plurality.

5.4.3. Plural expression in demonstratives

Table 5.21 illustrates the occurrence of the marker types used to express plurality in demonstrative pronouns. The most common strategy are suffixes. Plural words are attested, as well as complex structures like stem modification com-

Table 5.21.: Plural construction types in demonstratives

Marker Type	DEM
0	3
clitic	5
infix	2
lexical item	3
stem modification + lex. item	1
prefix	2
stem alternation	5
suffix	63
suffix+ stem modification	1
suppletion	15
suppletion + stem modification	1

Table 5.22.: Plural marking in demonstratives

Language	**Form**	**Type**
Badaga (Hockings & Pilot-Raichoor 1992: p.100)	ive-go	STEM MODIFICATION + SUFFIX
Choctaw (Broadwell 2006: p.239)	oklah	LEXICAL ITEM
Andamanese (Yadav 1985: p.170)	-ni	SUFFIX

bined with other markers.

Table 5.22 reports some relevant examples of construction types used to mark plural value on demonstratives.

5.4.4. Summary

Plural value is expressed by a number of strategies: pronouns do prefer suppletive forms, or suppletive forms in combination with further affixal strategies. A tendency of 'regular' addition of morphological material, like suffixes, is attested in the first person and it progressively increases in third person pronoun

plural, where suffixes 'outnumber' suppletive strategies.

Full nouns are in tendency marked by suffixes; clitics, when attested, tend to cover all the NP types in the given language. Suppletion and stem modification are widespread on human nouns and kin terms.

Demonstratives are, in most cases, indifferent to number marking; however, the most widespread marking type found across this lexical category are suffixes and suppletion.

5.5. Summary of the chapter

In this chapter, the marker types used on the nominal types to express number meaning have been discussed.

Section 5.2 explored the singular dimension in nouns, pronouns and demonstratives. Singular number tends to be expressed by suppletion in pronouns and zero marking in nouns. Overt marking strategies in nouns tend to be restricted to human and kin terms. Suppletion in full nouns has been attested on human nouns. Inanimate nouns tend to select the overt marker used by the other full nouns: this means that they do not show specific constructions.

Dual is expressed by a quantity of constructions types, in most cases quite complex and variegate. Pronouns indicate dual through combined constructions involving, in some cases, three different morphemes or particles. The pronominal stem found in dual is usually related to the plural root. However, pronominal systems with a singular and dual VS plural stem have been found, although quite rare. Dual in nouns, as in demonstratives, is expressed preferentially though suffixes; they can be added to the unmarked noun or, on the other hand, two markers are found: one that expresses plurality and the other dual-specific.

6. Plural marking distribution within languages

6.1. Introduction

This chapter explores the referential types connected to the expression of plural marking in the languages of the sample. A referential type has been defined as a set of referential expressions which are marked in the same way (see Section 4.5). The main goal of this chapter is the identification of the referential expressions (the nominal types) that share the same plural construction form within and among the languages.

The survey is conducted by addressing the following research questions:

- What is the distribution of the plural constructions on the nominal elements of the languages of the world? Do languages use a single marker for all nominal types or do they select different markers within the same language?
- What is the areal distribution of the two contexts aforementioned?
- In case of multiple plural markers among the NP types, which groupings can be identified? Which referential types are more common?
- Are there pronominal-specific plural markers? What is their internal distribution? Which construction types are used? Where are they attested?
- How do nouns and pronouns interact? Which referential sets of both nouns and pronouns can be identified?
- How do demonstratives behave with respect to plural marking?

This survey has been conducted on about 160 languages, selected in the subsample and coded in the nominal number marking database. Due to the complexity of the phenomenon in object, I have decided to proceed step by step by splitting the exploration of the referential types in five main blocks. The first, addressed in Sections 6.2 and 6.3 describes the internal groupings and the related plural construction forms found in full nouns only. The aim of this part is the identification of the most common referential types between kin terms, human noun, animate and inanimate nouns and describe any relevant feature in these sets and in the construction forms selected. The Section 6.4 investigates plural marking and the pronominal domain: this part describes phenomena like pronominal-specific plural markers, the types of constructions used, the internal referential sets that can be identified and the distribution of these contexts among the languages. The third block in Section 6.5 explores the interaction between nouns and pronouns: the languages with shared plural markers between nouns and pronouns are illustrated, by targeting the referential sets and their distribution. Then, Section 6.8 deals with demonstrative pronouns and plural marking; specifically, it focusses on the construction used to mark plurality in demonstratives, and whether these markers are shared with nouns, pronouns or both and the distribution of these contexts. Finally, Section 6.2.2 discusses briefly the languages of the sample with no plural distinctions.

Table 6.1.: Plurality sub-groupings, full nouns

Type	Count
Single marker on all nouns	52
Single marker on animate nouns	14
Single marker on human nouns	11
Single marker on kin terms	8
No plural marking on nouns	25
Plural marking on nouns only	6
Split plurality	27

6.2. Referential types and construction forms in full nouns

This section focusses specifically on full nouns and their behavior with respect to plural marking constructions. Data exploration identifies the following contexts on distributional constrains in plural marking:

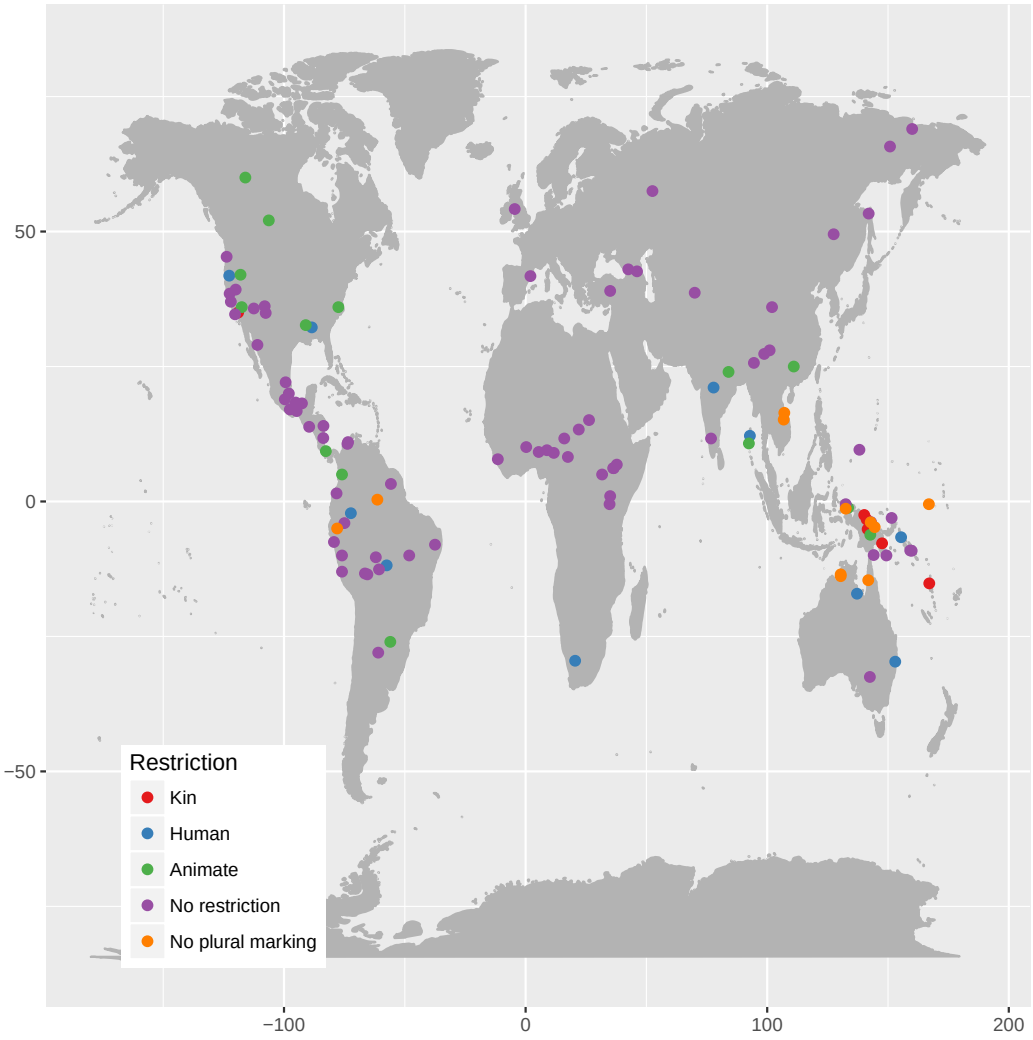
- A unique plural construction shared by all full nouns.
- A unique marker on a restricted plural marking context (e.g., a language distinguishes plural opposition on human nouns only and one marker only is used).
- Plurality split constructions in full nouns: different markers are used to express plural in NP types.

Two additional contexts can also be identified:

- Languages with no plural marking on nouns.
- Languages with plural marking restricted to full nouns (pronouns and demonstratives excluded).

Table 6.1 show the number of occurrences of each context. The geographical distribution is shown in Fig. 6.1.

Figure 6.1.: Plural marking on full nouns: types and distribution



6.2.1. Languages with one plural marker on all full nouns

This subsection describes the first type, which gathers all the languages that show a unique marker for all nouns (it has to be pointed out that only full nouns are investigated in this section). Many marker types have been attested, with different distributions: suffix type is the most widespread among the plural markers with no internal restrictions. Other strategies, like tone, prefix and reduplication, are less common. Clitic markers spread on all the full nouns are also attested. In what follows, the languages where these markers are found are described. For convenience, languages that have a single marker for all NP types are grouped by plural marker and then described.

Tone as a unique plural marking strategy in full nouns

Tone is attested, as a unique plural marking strategy widespread on all full nouns, on one language only, Abun, a Papuan language spoken in the Bird's Head Peninsula (Berry 1995: p.21). Tone is rarely used as the only plural strategy on all nouns, and even in Abun, based on what reported in the language corpus available in the grammar, full nouns have started to be indifferent to plural distinctions.

Prefixes as a unique plural marking strategy on full nouns

Plural prefixes as the only marking strategy are also rare. The only language within the subsample it occurs is Atzingo Matlazinca, a Otomanguean language (Mexico). The plural prefix /*nee*/ (Muntzel 1986: p.79) appears regularly on all full nouns. No information about the origin of this marker is provided.

Clitics as a unique plural marking strategy on full nouns

A preliminary disclaimer is necessary: since the definition of clitic is highly debated, I make no claims about the clitic label of a particular marker in a language: I just take for granted the information provided in the reference language documentation.

Table 6.2.: Clitics as a unique plural marking strategy on full nouns

Area	Family	Language	Marker
Africa	Igboid	Igbo	*=ga*
Eurasia	Sino-Tibetan	Angami	*=ko*
Eurasia	Sino-Tibetan	Northern Pumi	*=ɹəH*
South America	Isolate	Cayubaba	*=mɐ*
South America	Tupian	Karo	*=to?*
North America	Chibchan	Rama	*=ga*
Pacific	Papuan	Savosavo	*=gha*

Clitics are attested as 'full' plural markers in six languages of the subsample. Table 6.2 provides a report of the languages and the related constructions.

Reduplication as a unique plural marking strategy in full nouns

Reduplication as a widespread and unique plural marking strategy is quite uncommon; the only case (within the language sample) where plurality is attested exclusively by reduplication is found in Daga (Transguinean, Dagan): its areal location confirms the distribution of the reduplication processes identified in Dryer 2013, where eight languages (out of 1000+, which confirms also the rarity of the phenomenon) display reduplication and they are all located in the area that includes part of the South East Asia, Northern Australia and Papua New Guinea.

Reduplication is a quite 'anarchical' productive plural marking strategy in languages: instances of reduplicative processes (both full and partial) are found everywhere, with no clear distribution nor semantic criterion. Some facts can be pointed out about the distribution and other features of the reduplication processes:

- Reduplication processes seem to be more extensive and attested in the North American languages: full reduplication is widespread, to various ex-

Table 6.3.: Plural words as unique marking strategy, list of languages

Area	Family	Language	Marker
Africa	Atlantic Congo	Mumuye	*yi*
Africa	East Chadic	Lele	*kinye*
Africa	Defoid	Yoruba	*awon*
Africa	Afroasiatic	Mwaghawul	*mo*
Eurasia	Sino-Tibetan	Chang	*shoung*
South America	Otomanguean	Chalcatongo Mixtec	*xina´ʔa*
North America	Misupalman	Miskitu	*nani*
Pacific	Papuan	Ambulas	*béné*

tents, in Pima Bajo, Pipil, Kawaiisu and Classical Nahuatl, Northern Paiute (all of them Uto-Aztecan), Washo, Tillamook and Bella Coola (Salishan), Kwakiutl (Wakashan), Ineseño Chumash (Chumashan) and Chochtaw (Muskogean), among others.

- In many cases, reduplication serves as a base where further plural marking processes are added, like suffixes. This happens in some African languages, like Dahalo (Cushitic) or Midob (Nilo-Saharan);
- Reduplication has a strong bond with distributive meaning. In some cases, reduplicated forms are used to express distribution over space rather than plurality; such fact is particularly straightforward in North American languages, following Mithun 1988, and it is attested in African, Austronesian and Asiatic languages as well. Distribution over space implies a plurality of referents, which can be objects but also different 'human types' in a group, with specific characteristics. It is not rare to find reduplication used with all types of nouns, human and not human.

Plural words as a unique plural marking strategy in full nouns

Nine languages of the sample use exclusively a lexical item to express plurality in all nouns. Table 6.3 provides the language list and related figures.

Table 6.4.: Mumuye (Niger-Congo, Africa) Shimizu (1983: p.80)

	singular	plural	translation
Nouns	zaa	zaa-yi	'dogs'
	shon	shoo-yi	'people'
Pro 3	wu	yi	'they'

The areal distribution of plural words as general and unique pluralization strategy is heterogeneous attested in all macro-areas. As provided in table, one is located in Eurasia (Chang, Sino-Tibetan), another is North American (Miskitu, Misupalman). South America and Australia/Papua New Guinea areas count two languages each.

Linguistic data provide information about the origin, or the meaning for some of the plural words found in the sample. Some cases are straightforward: both the African languages of the type, Mumuye, Yoruba and Mwaghawul (aka Mupun) (Atlantic Congo and Afroasiatic respectively), employ the third person plural form as a plural word. Mupun language uses the particle /*mo*/, identical to the third person plural form (see Frajzyngier 1993: p.46). The Mumuye example is provided in Table 6.4.

The plural word used in Miskitu (Misupalman, Central America), /*nani*/, means 'people' (Adam 1891: p.80). The Chalcatongo Mixtec (Otomanguean) lexical item /(*xina´ʔa*/ simply means 'plural' (Macaulay 1996: p. 97). Papuan language Ambulas (Ndu) uses a plural word, /*béré*/, whose meaning is 'and associates': it is fully a plural marker. The form is similar to the second person dual pronoun, *béné* (Wilson 1980: p. 97).

Suffixes as a unique plural marking strategy in full nouns

Suffixes are the most common strategy used as a unique marker on all full nouns. More than thirty languages coded in the database show a unique suffix which marks plurality on full nouns.

The areal distribution of the types is quite asymmetrical. Most of the languages are located in Eurasia (13). Cases from South America and Africa are attested with eight languages each. Languages from Papua New Guinea and Australia are less represented: only Duna (Isolate) and Meriam (Eastern-Trans Fly) have been reported. Finally, Ohlone (Miwok-Costanoan) is the only North American language belonging to the type.

In most cases, language descriptions lack detailed information about plural markers; number marking, especially in the nominal domain, tends to be treated superficially: often basic information like obligatoriness of the markers is not mentioned. However, some languages have better coverage in this regard. In what follows, I report more specific data that I have found in some of the languages of this subtype.

Mochica (5) and Urarina (6), both spoken in Peru, South America, use plural suffixes similar to the ones employed on verbs.

(5) Mochica (Chimuan, Peru) Hovdhaugen (2004: p.21)

a. cqolu cqolu-aen
girl, girls-PL
'girl, girls'

b. taec-aen
go-PL
'they go'

It has to be pointed out that verbal plurality in Mochica is indeed rare, if not exceptional: the only case attested with regular plural markers is the verb 'to go', reported in the example. On the other hand, /*-kuru*/, the plural marker in Urarina (6) is so extensively and heterogeneously used in verbal plurality that, 'making any kind of generalization about the use of *-kuru* in verbal plurality is impossible' (Olawsky 2006: p.96). It is however worth to note that such marker has also an associative value with proper names.

(6) Urarina (Isolate, Peru) Olawsky (2006: p. 94)

a. sa, sa-uru
rat, rat-PL
'rat, rats'

b. ku-uru-a
go-PL-3psA
'they go'

Two further South American languages, Jaqaru (Aymaran) and Huallaga Quechua (Quechua) show a common plural suffix, /*-kuna*/. This marker derives from the sequential form meaning 'also' (Johnson-Weiner 2001: p. 13) and, as it has been reported for Jaqaru, this form has been extended to nouns by younger speakers due to the influence of Spanish. Jaqaru pronouns do not distinguish plural (see the following example).

(7) Jaqaru (Aymaran, South America) Johnson-Weiner (2001: p.18)

a. wawa, wawa-kuna
child, child-PL
'child, children'

On the other hand, in Quechua (8), /*-kuna*/ is also found on all pronominal forms[1].

(8) Huallaga Quechua (Quechuan, South America) Weber (1989: p.70)

a. warmi-kuna
woman-PL
'women'

Huastec (Mayan) nouns take a plural suffix, /*-čik*/, which is also present on verbs, but with a slight difference in meaning; specifically it indicates distribu-

[1] In Adelaar (2004), another source reports the plural suffix /*-kuna*/ has also spread to Jaqaru pronominal forms.

tion on space (Edmonson 1988: p. 177). The example below illustrates this.

(9) Huastec (Mayan, Central America) Edmonson (1988: p. 178)

a. warmi-kuna
woman-PL
'women'

b. yab k-a ka-a7 čik-i-čik
not IMP-A2 go-out-CAUS burn-TS-PL
'(Watch the tortillas); don't remove them with little spots all over'

Nivkh (Isolate, Manciuria) — see example (10) — distinguishes plurality in nouns by the means of a suffix, /*-ku*/ (with allomorphs), which resembles, in its formal structure, the plural comitative /*-ko*/[2].

(10) Nivkh (Isolate, Eurasia) Gruzdeva (1998: p. 17)

a. N'i nakr-ux k'ek-ko hyjk-ko zif-ku n'ry-d'
I snow-LOC.ABL fox-COM.PL hare-COM.PL track-PL see-FIN
'I saw foxes' and hares' tracks on the snow'

Chang (Austroasiatic) plural suffix /*-oung*/ derives from the noun *shoung*, meaning 'group'. Such marker may also be suffixed to pronouns.

Similarly, in Paakantyi, a Pama-Nyungan language spoken in Australia, expresses noun plurality with the marker */-lugu/*, that involves a slight collective meaning, like 'group of' (Evans 2007: p. 181).

6.2.2. Languages with restricted plural distinction

This section explores the constructions used on the languages where plural marking is distinguished on restricted full noun types only.

Data exploration provides the following subtypes:

[2] Nivkh language has two specific comitative markers, one for dual and one for plural: dual number is however not longer expressed in nouns.

- Languages with plural distinction restricted to kin terms (4 languages);
- Languages with plural distinction restricted to kin and human nouns (13 languages);
- Languages with plural distinction restricted to animate nouns (kin, human and generic animate, 14 languages).

Plural marking restricted to kin terms only

All the languages included in this subtype belong to the Pacific area: Nimboran, Mian, Suena, Oksapmin. All of them use suffixes to mark plurality on kin, except for Suena, where the markers reported have been labelled as enclitics.

Nimboran (Nimboran) selects two suffixes to mark plurality on kin terms: /*-naŋ*/ and /*-yap*/; both of them means 'all'. Their distribution within the kin terms is complementary: however, the range of application of /*-yap*/ as compared to /*-naŋ*/ has not been determined (May 1997: p. 155). The main difference in usage between these markers is that /*-naŋ*/ is also used by pronouns with the same function.

(11) Nimboran (Nimboran, Papua New Guinea) May (1997: p.155)

a. babu-naŋ
grandfather-PL
'grandfathers, all the grandfathers'

b. belendyo-yap
aunt-PL
'aunts, all the aunts'

Another Papuan language that selects a suffix to express plurality on kin terms only is Mian (12). Nominal morphology is rare in Mian, and one of the few nominal markers is the suffix /*-wal*/ which is used on kin terms with pure plural meaning and on proper names with an associative value, 'X and others'. This marker is not found elsewhere. Other nouns, like human or dyadic terms,

tend to be vague with respect to number opposition. When plurality needs to be expressed, such feature is carried by 'collective articles' (Fedden 2011: p.111), which never occur on kin terms.

(12) Mian (Aymaran, South America) Fedden (2011: p.110)

a. Kasening-wal
Kasening-PL
'Kasening's family'

b. biěm-wal
mother-PL
'mothers'

Suena language expresses plurality by the means of several markers, all defined as enclitics; partial reduplication is also present, although it is restricted to a few kin terms *apie*, 'grandfather', *apipie* 'grandfathers' (Wilson 1974: p.111). Enclitic forms are more common: their distribution is not clear, and similarly their meaning: however, one of these markers, /=*mai*/ is formally identical to the noun that means generically 'person', *mai*, but I do not claim any hypothesis on eventual correlations. The full list of markers is shown in the example below.

(13) Suena (Trans New Guinea, Papua New Guinea) Wilson (1974: p.110)

a. meta, meta=mai
nephew, nephew-PL
'nephew, nephews'

b. nami, nami=sini
uncle, uncle-PL
'uncle, uncles' (father's brother)

c. mama, mama=na
father, father-PL
'father, fathers'

Kin terms in Oksapmin (Trans-New Guinean, Ok), are inflected for number through the addition of a suffix, /*-xel*/. Oksapmin language has two different sets for kin terms: (i) kin terms that inflect for number only (example 14a) and (ii) kin terms inherently possessed which inflect for both number of the referent and person of the possessor. Kin terms of the latter set do inflect for number through the addition of a suffix, usually *-il* (example 14b).

(14) Oksapmin (Nuclear Trans New Guinea), Loughnane (2009: p.110)

a. ənan, ənan-xel
aunt, aunt-PL
'aunt, aunts'

b. inəp, inəp, inəp-il
wife, wife, wife-PL.3POSS
'wife, his wife, his wives'

Plural marking restricted to human nouns only

This section deals with the languages that show a plural distinction restricted to kin and human terms only. The count of languages, their distribution and the characteristics is presented in Table. (15 lngs).

The only language from Eurasia that provides further information about its markers is Darai (Indo-European). The pluralizer marker in Darai, *-səb*, derives directly from the the noun *səbəi*, meaning 'all'. It is reported (Dhakal 2012: p.30) that *səbəi* is still used as an independent lexical item in some neighboring varieties, but not in Darai anymore, where the marker has undergone through grammaticalization process. This suffix works as number agreement marker in

verbs (example 15a) and it is also used with big, 'collective' groups (example 15b).

(15) Darai (IndoEuropean, Nepal) Dhakal (2012: p.30)

a. beta-səb əi-tahə-səb
boy-PL, come-NPSt-PL
'The boys come'

b. əni bəllə gorkhali radza-kə phəudz-səb phərki-kun dzəi-lə
then only Gorkhali king-GEN army-PL return-SEQ go-PST
'Then, having returned, the army of Gorkhali king went'

Two languages from the 'Pacific sample' have a pluralization strategy restricted to human nouns, Maori (Oceanic) and Meyah. Maori language is a textbook example of stem modification restricted to a small group of human nouns and kin terms as a plural marking signal. The phonological processes involved in the stem modification are vowel lengthening, like *wahine, wāhine* 'woman, women' or less predictable modifications like *tamaiti, tamariki* 'child, children' (Harlow 2007: p.115).

Meyah human nouns, a Papuan language spoken in the East Bird's Head peninsula, select the clitic *-ir* as a pluralizer (example 16a). Such form derives from the third person plural object clitic, *-ir* (example 16b). The plural marker can be exceptionally used with high animates, dogs and pigs only, possibly because of their cultural importance (example 16b):

(16) Meyah (East Bird's Head, Meax, Papua New Guinea) Gravelle (2004: p.79)

a. Me-ona-ir me-ohca mar erek ke-uma
1PL.EXC-male-PL 1PL.EXC-search thing like NOM-that
'We men look for things like that.'

b. Beda mes-ir-uma ri-osok
then dog-PL-that 3PL-climb
'Then those dogs went up.'

The next cases will explore some languages from North America (Slave, Babine, Assiniboine, Northern Paiute) and from the South American macro-area (Embera, Awa Pit).

Two Athabaskan languages, Slave and Babine-Wit'suwiten, belong to this subtype. Northern Slavey variety employs a pluralizer clitic on human nouns and high animates (dogs only). This marker, defined as a 'group' suffix, is identical to the human plural prefix *ke-*, to which is probably connected. Moreover, their distribution is complementary: when the subject of a sentence is marked for plurality by the suffix *-ke*, it is not necessary to use the human plural prefix *ke-* in the verb. The following example illustrates this.

(17) Northern Slave (Athapaskan, North America) Rice (1989: p.248)

a. dezoa-ke kare nágoyeh
child-PL outside play
'The children are playing outside'

Witsuwiten (Table 6.5) plural suffix is *-ni* (other forms are reported, but *-ni* is the most productive, and it is found also on pronouns). The plural marker in this language seems to be related as well to verbal plurality, more specifically to noun derivation from verbal forms. Most nouns in Witsuwiten are of verbal origin, and these deverbal nouns carry the nominalizer/human plural suffix *-ni*, as can be seen in the examples; this fact may suggest that the plural marker used for human nouns and on human nouns only is strictly connected to the deverbal nominalization process that is necessary to derive such nouns.

Another human plural marker probably related to verbal forms is used on Assiniboine (Siouan). Human nouns in this language select a pluralizer, /*-pi*/, that also occurs as an obligatory enclitic marker on verbs, used to express animate plural. (Cumberland 2005: p. 315).

Northern Paiute (Uto-Atzecan, Table 6.6) marks plurality on human nouns through the suffix *-mi*, such suffix, strictly related to the marker employed in

Table 6.5.: Babine-Witsuwiten (Na-Dene, North America) Hargus (2007: p.268)

form	translation	note
nənihni	'food servers'	lit. 'those who distribute'
tenedədlini	'religious people'	lit. 'those who are religious'
ləqətni	'Pentecostal people'	lit. 'those who clap'

Table 6.6.: Northern Paiute (Uto-Atzecan, North America) Thornes (2003: p.103)

singular	plural	meaning	construction type
moko?ni	mommoko?ni	'women'	reduplication
woho	wohomɨ	'enemies'	suffix
nana	naana	'men'	suppletion

pronouns as well, is not, however, the only strategy employed in this language to express plurality in human nouns. Cases of partial reduplication and suppletion are also reported and they all seem linked to the means verbs use to express plural distinction. In fact, reduplication processes in Paiute are fairly productive, especially on verbs, where the initial CV reduplication indicates regularly distributive action. Suppletive forms are also common in verbs, always related to plural expression, like in *yadua*) 'talk.SG' and *apiča* 'talk.PL' (Thornes 2003: p.320):

Jamul Tiipay makes plural distinctions on kin terms and human nouns only: Langdon (1970: p.194) successfully proves how these distinctions do not as a result of inflectional processes, but rather because such nouns derive from verbs: as the verb stems make plural distinctions, so do behave the derived nouns (all kin and occasionally human).

The last two examples provided for this subtype come from languages of South America. Embera (Chocoan) human nouns show the suffix, not obligatory, *-rã* to express plurality. This form is not related to any verbal plural

marker (verbs take suffix *-da* to signal plural number, (see Mortensen 1999: p. 41)). Nominals (human) in Awa Pit (Barbacoan) do not have number as a grammatical category, with the except of personal pronouns. However, there is a marker, defined as the 'collective action suffix' (see Curnow 1997: p.131) which can be attached to nouns to express a meaning similar to number. Rather than simply marking multiple referents, the suffix *tuzpa* marks the referents form a coherent group; additionally, the group has to have been acting together to perform some process. This is illustrated in the following example.

(18) Awa Pit (Barbacoan, South America) Curnow (1997: p.131)

a. ampu-tuzpa kal ki-ni-ma-ti
man-COLL work work-PROSP-COMP-TERM
'Together the men went off to work'

Plural marking restricted to animate nouns only

The last subtype includes language with plural distinctions restricted to animate nouns (kin, human, generic animate).

Languages belonging to this subtype lack detailed information about their plural markers. The examples provided are from four languages only: Garawa, an Australian Aboriginal language from the Northern Territory; Choctaw (Muskogean) is from North America macro-area as Shasta language (Shastan, extinct, Northern California); Andamanese (Great Andamanese) is one of the languages of the Andaman Islands, located in South-East Asia.

Shasta nouns, a language from North America, takes a plural animate marker, *-yáwar*: the meaning of this lexeme is a kind of collective, since it means 'a group of': *úpitaktakyáwar*, 'a bunch of hounds' (Silver 1966: p.190). Choctaw main pluralization strategy is a particle, defined as a preverb, meaning 'people' that appears between animate nouns and verbs. In the example below, the particle *oklah* can be seen as a plural marker (19a) and in its meaning (19b).

(19) Choctaw (Barbacoan, South America) Broadwell (2006: p.206)

a. Hattak-at oklah tachi' at apa-tok
man-NM PL corn come:and eat-PT
'The men came and eat corn'

b. Ai okla ha mominchit vbit
Ai people-AC all-PART kill-PART N-complete:L-NEG-TNS-until
ik tahlo hokvto, ik isso tok oka.
3N-quit-NEG-PT-for
'For he did not quit until he had killed all the people of Ai.'

Andamanese animate plural marker enclitic, */=n/* has the same meaning, 'people': it derives from the plural word *ne/nu/ni* (Abbi 2013: p. 113).

(20) Great Andamanese (Andamanese, South East Asia) Abbi (2013: p.113)

a. thire=ni kona-bi rališu-k-o
child-PL kona-ABS finish-FA-DST.PST
'Children finished the tendu fruit'

Finally, Garawa nouns show a plural suffix /*-muku*/, which is obligatory with human nouns to signal plurality, while it seems to be optional with animate non human (example 21a); when *-muku* attaches to a possessive pronoun, the meaning is 'X's mob' (i.e., all the people associated/related to X) (example 21b), suggesting a link with an original associative meaning.

(21) Garawa (Garrwan, Australia) Mushin (2012: p.81)

a. kuyu yalu-nya langki baya-muku waluwijba
bring 3pl-ACC north child-PL lead
'(He) takes them north, the kids, leading them'

b. ngaki-nmuku=yili yalu, jila nayi
1sgdatPL=hab 3plnom walk this
'All my mob used to come here'

6.3. Split plurality: different constructions on full nouns

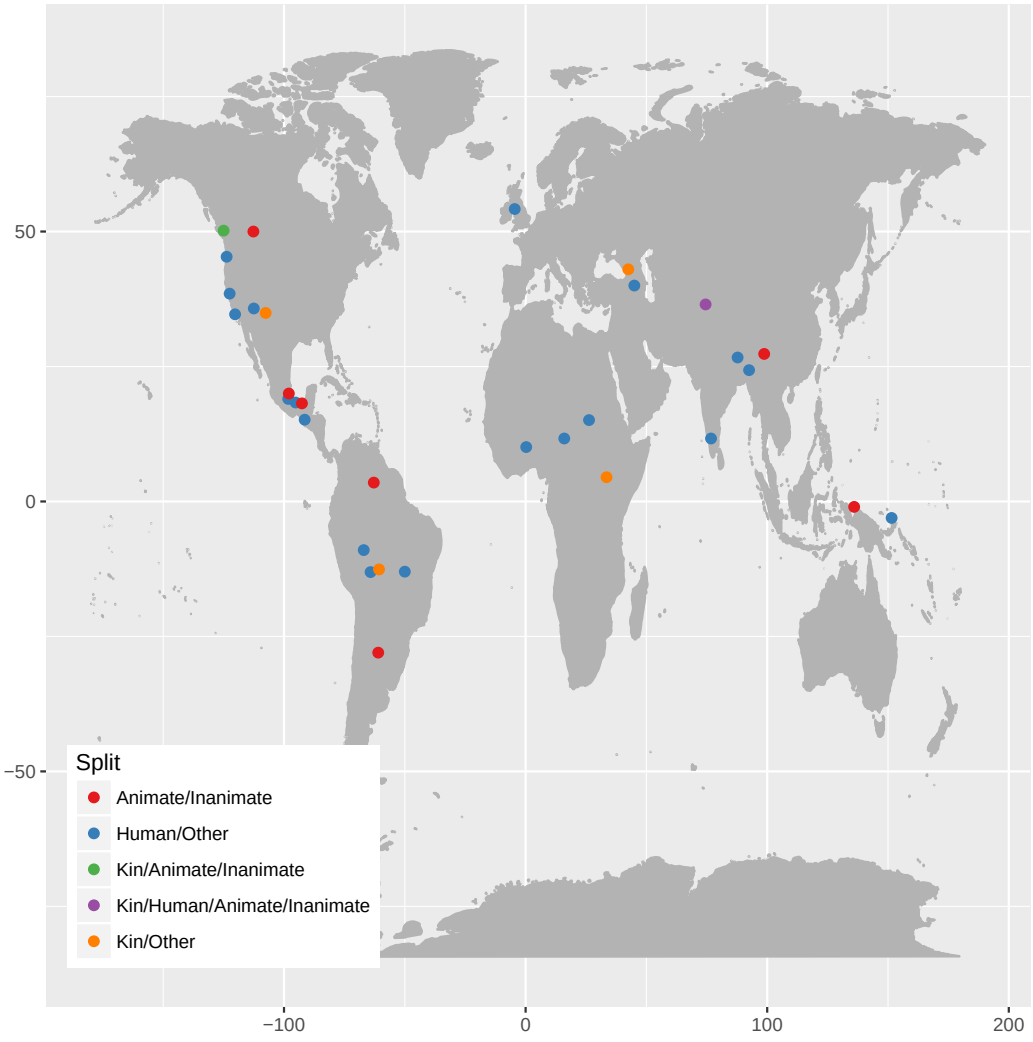
This section describes the languages that express plurality in full nouns through different constructions. The main goal is to identify the groups of noun types that share the same construction as opposed to others, and which are the noun types that tend to express plurality through the same markers.

For all cases that will be presented in this section, however, an important disclaimer is necessary: the restrictions in the use of one marker rather than another are rarely *clear-cut* on the noun types. In fact, most languages that present more than one marker to express plural distinctions for each (or some) nominal type(s), show a *preference* or a *tendency* for a specific marker to appear on a noun type, rather than a 'fixed' and mandatory strategy. The boundaries are therefore 'blended': this means that the groupings identified shall not be considered in absolute terms and in strict opposition, since exceptionsare always present.

The survey of the data shows four distinct systems of split plurality on full nouns. These systems are: (i) Human vs. other nouns (ii) Animate vs. inanimate nouns (iii) Kin vs. human vs. animate vs. inanimate nouns (iv) Kin vs. animate vs. inanimate nouns (v) Kin vs. other nouns

The geographic distribution is shown in Fig. 6.2. The remaining part of this section discusses the individual systems in more detail.

Figure 6.2.: Split plurality on full nouns: groupings and distribution



6.3.1. Human/Other

The first grouping, which is also the most populated subtype, involves the presence of a specific marker for human nouns[3], and another marker restricted to the other full nouns types, from animates to inanimates.

In some cases, the different markers adopted by human nouns and all other nominal types respectively do reflect a formal similarity that is complex to disambiguate.

Anufo language (Atlantic Congo), for example, signals plurality in nouns by the means of a clitic form *-m*; human nouns take the same suffix, but, on the other hand, add also a prefix, *n'-*. Further observations suggests that *n'-* is used mainly as a topicalizer with also a definiteness feature, but it seems to have lost this latter function (see the example below).

(22) Anufo (Atlantic Congo) Smye (2004: p.45)

a. mbàam
PL.child.PL
'children'

Another African language spoken in Sudan, Midob (Nubian) marks plurality on human nouns through a suffixal strategy, the marker *-edi* (and allomorphs), while animate and inanimate nouns do employ the marker *-ti* (Werner 1993: p. 27).

Similarly, in Baure (Arawakan, Bolivia) which generic nouns pluralize through the suffix *-nev*; human nouns, employ instead the suffix *-anev*: morphological */a/* has been identified as a 'linker morpheme' (Danielsen 2007: p. 102).

In Dhimal (Sino-Tibetan) human nouns (also pronouns and nouns indicating ethnicity, and maybe this semantic link may indicate an associative origin of he marker) do pluralize through the suffix *-lai*; regular nouns take the suffix *-gelai*. The main difference between these two markers, actually to be labelled as two

[3] Human do involve also kin terms; language descriptions always 'take for granted' that whether a marker is restricted to humans, it will also include kin terms; I have nonetheless always looked for counterexamples in texts, when available. I have reported any single case which goes against this assumption.

allomorphs, is that *-lai* is no longer productive in Dhimal (T. 2009: p.57).

Another Sino-Tibetan language with distinct markers for human nouns vs both animate and inanimate is Kokborok (23). In Kokborok language, plural in human terms is expressed by the suffix /*-sɔŋ*/, while animate and inanimate nouns select the markers *-rɔk* and *-kəbaŋ* (not respectively, the distribution is free).

(23) KokBorok (Sino-Tibetan, Eurasia) Karapurkar (1976: p.44)

a. takhuk, takhuk-sɔŋ
brother, brother-PL
'brother, brothers'

b. amiŋ, amiŋ-rɔk
cat, cat-PL
'cat, cats'

c. manuy, manuy-kəbaŋ
thing thing-PL
'thing, many things'

Such markers, defined as 'collective adjectives' (Karapurkar 1976: p.45) with the meaning of 'many', 'much', may be used with human nouns to denote associative meaning.

Armenian (Indo-European) main plural markers are usually constrained by phonological rules: the suffixes *-er*, *-ner* do appear on monosyllabic and polysyllabic nouns respectively.

However, some vestigial markers inherited from Classical Armenian reveal a semantic restriction: the most common is the marker *-k'*, used on a small set of nouns denoting human and on names denoting origin. It may be hypothesized that there is a relation between this marker and the plural marker *-nk'*, used with pronouns and on personal names with an associative value (Dum-Tragut 2009: p.65).

Ineseño (Chumashan) employs two main strategies to express plural marking on nouns: the first, and the most common, is reduplication: partial reduplication of the first syllable is widespread on all types of nouns, from kin terms (*mak?aniš, makankaniš*, 'my paternal uncle, my paternal uncles', inherently possessed) to the term that indicates 'people' (*ku, kuhku*) to objects; the other marker, not as used as reduplication, is attested on some nouns, all human, and it is the suffix *-wun*: /*išonuš-wun*/, 'twins', /*alapšawa-wun*/ 'people from Sawa' (note how this marker can reveal an associative force). Although this marker is less widespread on nouns, it is the most typical pluralization strategy with pronouns and also demonstratives, since reduplication is not used. Moreover, CVC reduplication in Ineseño Chumash seems to carry a collective value (Applegate 1966: p.562), rather than a plural one.

Another North American language, Navajo (Athapaskan) has two sets of markers: one specific for animate and inanimate nouns, the other restricted to human: the latter set include the suffixes *-ké* and *-yóó*. There is no information available about the possible origin of these markers; there is more data about the marker used on animate and inanimate nouns, which is a prefix (*daa-*) and it has a distributive origin, and it is used regularly on verbs: *kó*, *daakó* 'fire, many fires' (Young & Morgan 1980: p.2).

Wappo (Yukian) has a number of markers, some of them vestigial remnants; they can be synthetically split in two groups: (i) the marker specific for human nouns, *-te*, is connected to the verbal iterative plural *-te* (Radin 1929: p.119); (ii) *-le*, defined as a collective marker, used on animate and inanimate nouns. Radin observes how 'In present-day Wappo all the old methods of forming the plural are disappearing and the pronoun *oni*, 'they', seems to be carrying the whole burden. There is also a tendency to use *-ti*, the nominal plural suffix, and *mul*, 'all' (Radin 1929: p.121).

Washo (Isolate, California, North America) — example (24) — is a particular case, which is however worth discussing here. This language employs two pluralization strategies: the most important is final reduplication of the syllable of the noun. This marking strategy is used on nouns, mostly humans and animates,

while the latter, that is the suffix *-kic* is used on animates and inanimates. For the groupings I have set, this language does not seem to belong to the human vs all opposition. I decided to include Washo language in this specific subsection because there is however a strict restriction that involves human terms: human nouns do not take the suffix *-kic*: *-kic* occurs also on verbs with 'an unquestionable collective or distributive meaning' (Kroeber 1907: p. 273). This distributive meaning can be inferred also from the only noun attested with both forms, *moko*, 'knee', where the reduplicated form */mokoko/* means 'knees' and */moko-kic/* 'pile of knees, knees scattered about'.

(24) Washo (Isolate, North America) Kroeber (1907: p.272)

a. tamomoo-mo
woman-PL
'women'

b. aiyas-kic
wolf-PL
'wolves'

c. dik-milu-lu pakarec l-ecli-kic-i
each-of-my-friends a-head-of-beef I-give.
'I gave each of my friends a piece of beef'

d. dik-milu-lu pakarec 1-ecl-i
my-friends a-head-of-beef I-give.
'I gave my friends a piece of meat'

The last two relevant cases of presence of distinct markers for human and non human nouns are found in Sierra Populuca (Central America) and Apuriña (South America). In Sierra Populuca language (Mixe-Zoquean) (Elson 1960: p.73) two suffixes /-tam/ and /-yah/ are predominantly used. The former occurs as a human nouns pluralizer, the latter is taken by animate and inanimate nouns; the suffix *-tam* is used to express plural on first and second person pro-

noun; *-yah* is restricted to the third pronominal form (see example 25). Moreover, Populuca has a collective suffix, used on all types of nouns to express great quantities and large groups.

(25) Sierra Populuca (Mixe-Zoquean, Central America) Elson (1960: p.54)

a. yomo-tam
woman-PL
'women'

b. kúy-yah
tree-PL
'trees'

c. ca-ánhõh
rock-PL
'a pile of rocks, many rocks together'

d. pášiñ-ánhõh
person-PL
'a great crowd of people'

Finally, Apuriña language (26) has two specific plural suffixes, one reserved for human */-wako/* and the other used on both animate and inanimate */-nu/*. When the suffix restricted to human is used on animate or inanimate nouns, it 'humanizes' them. Gender is also expressed in Apuriña with suffixes (one for masculine, the other for feminine gender) not cumulative with number, although they are in a kind of dependent relation the one with the other: number markers requires the presence of gender formatives. The same plural markers are found on verbs.

(26) Apuriña (Arawakan, South America) Facundes (2000: p.260)

a. nurumane-wako-ro
relative-PL-F
'female relatives'

b. nurumane-wako-ro
relative-PL-M
'male relatives'

c. kema-nu-ru
tapir-PL-M
'male tapirs'

d. kema-nu-ro
tapir-PL-F
'female tapirs'

6.3.2. Animate/Inanimate

This section will explore and describe the languages of the sample that show a generic animate vs inanimate distinction in the plural markers. Baré (Arawakan) has two plural markers, not obligatory and they present the following tendencies: *-nu* tends to be used with human and animate nouns; *-be* tends to be used with inanimate nouns. There are also some nouns which can employ both suffixes, and in these cases the difference in meaning between these two markers is therefore revealed; the 'animate' plural suffix *-nu* has a collective force, while *-be* is rather distributive: */ʧabati-nu/*, 'birds' (a group of), */ʧabati-be/* 'birds' (one by one), as described in Aikhenvald (1995: p. 19).

The two animate and inanimate plural clitics used in Biak language (Austronesian) *=si* and *=na* respectively, correspond ultimately to the forms of the third person plural pronoun, which is distinct for animacy and do correspond to the markers on the verb (Table 6.7).

Table 6.7.: Biak (Austronesian, Australia) van den Heuvel (2006: p.101)

	Free pronoun	Marker of plurality in the noun phrase	Inflection on the verb
ANIMATE	si	=si	s- / si-
INANIMATE	na	=na	n- / na-

Huehuetla Tepehua, a Totonacan language spoken in Mexico, employs a wide range of markers that can be used to express plurality, although nominal number marking does not seem to be obligatorily expressed (27). Albeit the marker selection by the nominal types seems to revolve around both semantic and phonological processes whose predictability is not inferable (Kung 2007: p.344), some tendencies can be nonetheless extracted: there are two prefixes, used to mark nominal plurality; the plural prefix *lak-* is the default marker for inanimate nouns: this prefix is quite likely related to the verbal prefix *lak-* that serves to code distributive action on the verb. The other marker, again a prefix, has the form *7a-*: in the examples provided by the language description, the marker is used on animate nouns, especially human and kin terms, and it is cognate with the verbal prefix *7a-*, that co-indexes a plural indefinite on the verb[4]

(27) Huehuetla Tepehua (Totonacan, Central America) Kung (2007: p. 347)

a. lak-chaqa7
PL-house
'houses'

b. 7a-maaxkawanini7-n
PL-hunter-PL
'hunters'

[4] Kung also reports a number of suffixes used to express nominal plurality, where the common denominator is the phoneme */n/*. This suffix can be seen in (27b). The origin is unknown. *-Vn* and allomorphs are used to express plurality on all nominal types with no semantic distinction. However, it seems that these suffixal forms are slowly falling out of use, leaving the prefixes to carry completely the expression of nominal plurality

Another language with differentiated markers for animate and inanimate, is Shiriana (or Xiriâna), a now extinct language spoken in Northern Brazil. In Shiriana, the marker used to indicate plurality in animate nouns is *pik*, which is no other than the form of the third person plural pronoun. The strategy used to express plural number in inanimate nouns like plants and objects is the partititive-collective marker *-k*, see the example below.

(28) Shiriana (Arawakan, South America) Gomez (1990: p. 80)

a. pore pik
evil spirit
'evil spirits'

b. hokomo-k
sweet
'sweet potatoes'

A similar association between marker of third person pronoun for animate nouns and collective marker used as a plural suffix on inanimate nouns is provided by Chontal Maya language. In San Carlos variety of Chontal (29), the suffix *-op* is identical to the third person plural enclitic form attached to the verb (Knowles 1984: p. 80). On the other hand, the inanimate nouns pluralizer is the marker *-e(l)* used also as a collective form in the related Chol language (Schumann 1973: p. 26)[5].

(29) Chontal Maya (Mayan, Central America) Knowles (1984: p. 202)

a. čitam-lop'
pig-PL
'pigs'

[5] In the Chontal Maya grammar it is also reported another marker, */-il-ba/* used with both animate and inanimate, very productive and meaning abundance or extension */ʔotot-il-ba/* 'an abundance of houses', */čitam-il-ba/*, 'an abundance of pigs'.

b. ʔotot-e
house-PL
'houses'

The last two languages of the sample which provide a differentiation between animate and inanimate nouns in plural marking and also reveal interesting information are Blackfoot[6] and a language from South East Asia, Nung.

Blackfoot (Algonquinan) employs two different suffixes, for animate plural and inanimate plural (it also distinguishes overtly singular animate and inanimate: a more exhaustive description of Blackfoot markers is provided in the chapter about the number values interaction), both present in verbal forms; the same markers are also carried by demonstrative pronouns (Taylor 1969).

To conclude this subsection, the last language with distinct markers of which the grammar provides some information is Anong, a Sino-Tibetan endangered language spoken in China and Burma. Anong has two plural suffixes: *-zɿ31ɳɯ̀31*, used on animate nouns, and *-mɯ* restricted on inanimate but also used on some animate nouns. Both markers are classifiers which have fully developed as plural markers. These markers can be labelled as associative markers, since linguistic data reports that 'these markers are equivalent in meaning to the Chinese plural marker 'men" (Sun & Liu 2009: p.56), where such grammaticalization path from associative marking to generic plural formative has been successfully proven by Iljic 2001 and it is also typical of other Sino-Tibetan languages, like Iu Mien (Court 1985). I have however noticed that the form of these markers resembles two collective classifiers: the form *-zɿ31ɳɯ̀31*, used as a plural marker on animate nouns only is employed as a classifier meaning 'flock'; *-mɯ* recalls the classifier *ba31mɯ̃53*, meaning 'some'.

[6] Another Algonquinan language in the sample which presents two different markers for animate and inanimate nouns is Ojibwe, Valentine (see 2001)

6.3.3. Kin/Human/Animate/Inanimate

Burushaski language (Isolate, Eurasia) is the only one among the languages of my sample that presents a distinctive marker for each noun type, as illustrated in the example below.

(30) Burushaski Isolate, (Lorimer (1935: p.33))

a. dʌsin, dʌsi-wʌnts
girl, girl-PL
'girl, girls'

b. mi, mi-tsəro
mother, mother-PL
'mother, mothers'

c. ha, ha-kicʌŋ
house, house-PL
'house, houses'

d. hal, hal-jo
fox, fox-PL
'fox, foxes'

Unfortunately there is no information about the origin of such markers: however, this high degree of internal differentiation suggests an independent development of these markers.

6.3.4. Kin/Animate/Inanimate

Comox language has three different strategies to express plurality: a suffix, *-tan* is used on kin terms only; its origin is unknown. Inanimate nouns, especially plants and small insects, express plurality through a suffix (defined as a lexical suffix), which is rather a collective form. Other nouns do express plural number through reduplication processes. Refer to the next example.

(31) Comox (Salishan, North America) Harris (1981: p. 96)

a. kupa, kupa-tan
grandfather, grandfather-PL
'grandfather, grandfathers'

b. sa?an, sa?sa?an
salmon, salmon-RDP
'salmon, salmons'

c. t'am?xw, t'am?xw-cəlli
gooseberry, gooseberry-COLL
'gooseberry, gooseberries'

6.3.5. Kin/Other

It has been shown how usually kin terms express plural number through directly deverbal strategies, while other nouns, within the same language (for instance Comox) use suffixes or inflectional markers. One interesting exception to this tendency is revealed by Acoma (Keresan, North America), where regular nouns, both animate and inanimate, do pluralize through a marker of verbal origin *-dyaiM*, 'to have several', and a nominalizer, *-iši*. Kinship terms take the suffix *-ši* (32) Such plural strategy is also widespread in nearby varieties like Laguna Keres (Lachler 2006: p.66).

(32) Acoma Keresan (Keres, North America) Maring (1967: p. 148)

a. ṁáḱa, ṁáḱatyaimiši
dipper, dipper-PL-NOM
'dipper, dippers'

b. nánaš, ṅåisdi´aši
father, father-PL
'father, fathers'

Kwaza language (33) has two markers: one for kinship terms, inherently possessed, *-mɛ*, and another suffix, *-nahere*, which has an associative value with personal names and tribes and a simple plural value with common nouns.

(33) Kwaza (Kwaza, South America) Van Der Voort (2004: p. 235)

a. tala-ʻmɛ, ta'la-nahere
uncle-PL, uncle-COLL
'my uncles, the uncles'

b. Teteru-nahere
Teteru-PL
'family of Teteru'

Didinga language (Surmic, Africa) expresses plurality through a number of strategies: the basic generalisation that can be made is that kinship terms employ the suffix *-gi*, used also by pronouns and demonstratives, while animate and inanimate nouns take other markers like *-nya, -wa, -yok* (see Rosato 1980: p. 8).

Svan language (34) employs different pluralizers: the most frequent is *-ær*, with both animate and inanimate. Kin terms are pluralised through a circumfix of verbal origin. Other plural forms include collective *-ra* for plants and inanimates and the plural/associative marker which is derived from the genitive from 'of the X clan, of the X family'.

(34) Svan (Kartvelian, Eurasia) Schmidt (1991: p. 18)

a. di, læ-dj-a
mother, CFX-mother-CFX
'mother, mothers'

b. icx, icx-ra
pear, pear-COLL
'pear, pears'

c. zural, zural-ær
woman, woman-PL
'woman, women'

d. set'el-š-e:r, set'el-š-a
member of
'member of the Setl clan, members of the Setl clan'

6.4. Referential types and construction forms: pronominal domain

This section deals with the pronominal domain, and it describes the complexity and the distribution of the pluralization strategies within the pronominal forms.

This survey has been conducted, as for the full nouns data exploration, on the subsample of about 160 languages coded in the nominal number marking database.

As a preliminary phase, I have decided to control for the 'full suppletive variable' in my sample: it is widely known that pronouns tend to mark plurality through person/number suppletive stems. This tendency is confirmed by the results on suppletive pronominal plurals in my data; of about 160 pronominal paradigms coded in the database, 50 are fully suppletive, on all three persons.

For a better data analysis, I have therefore split the data in two big groups:

- Languages with a full suppletive plural paradigm;
- Languages with a non suppletive form in at least one pronominal person.

Another set can be added to these two groups, and it includes the languages with no pronominal forms.

- Languages without number distinctions in pronouns.

By excluding the suppletive form and the languages without number distinc-

tions on pronouns, one can narrow the scope of the analysis to the about 100 languages left that belong to the second main context, which includes all the languages with a non suppletive form in at least one pronominal person.

This set can be further split in two more fine-grained sub groupings:

- Languages with pronoun-specific plural markers;
- Languages with plural markers shared with nouns.

The structure of this section is structured along the groupings that have been identified: first, a list and a map plot of the languages that present full suppletive pronominal paradigms are provided and discussed. Subsequently, the languages with pronominal specific markers are described and further split in subtypes to ensure a systematic classification and a better data exploration. The languages with pronominal specific plural markers are about 30 out of 104; the remaining 70 languages have one marker shared by at least one pronominal person and one full noun.

6.4.1. Languages with suppletive forms in pronominal plurals

Data coded in the database shows that out of the 160 languages coded, about 50 take full suppletive in plural in the whole pronominal domain (first, second and third person). The geographical distribution of these languages is shown in Fig. 6.3. The detailed breakdown by area and family is in Table 6.8

6.4.2. Languages with pronominal specific plural markers

Out of the languages in the sample, 31 have plural markers specific to pronouns. A further survey of these markers may uncover more specific aspects on these constructions:

- Formal characteristics of the pronominal specific plural markers: for instance, which construction types (i.e., suffixes, plural words) are taken exclusively by pronouns to signal plurality.

Figure 6.3.: Plural in pronouns: full suppletive systems

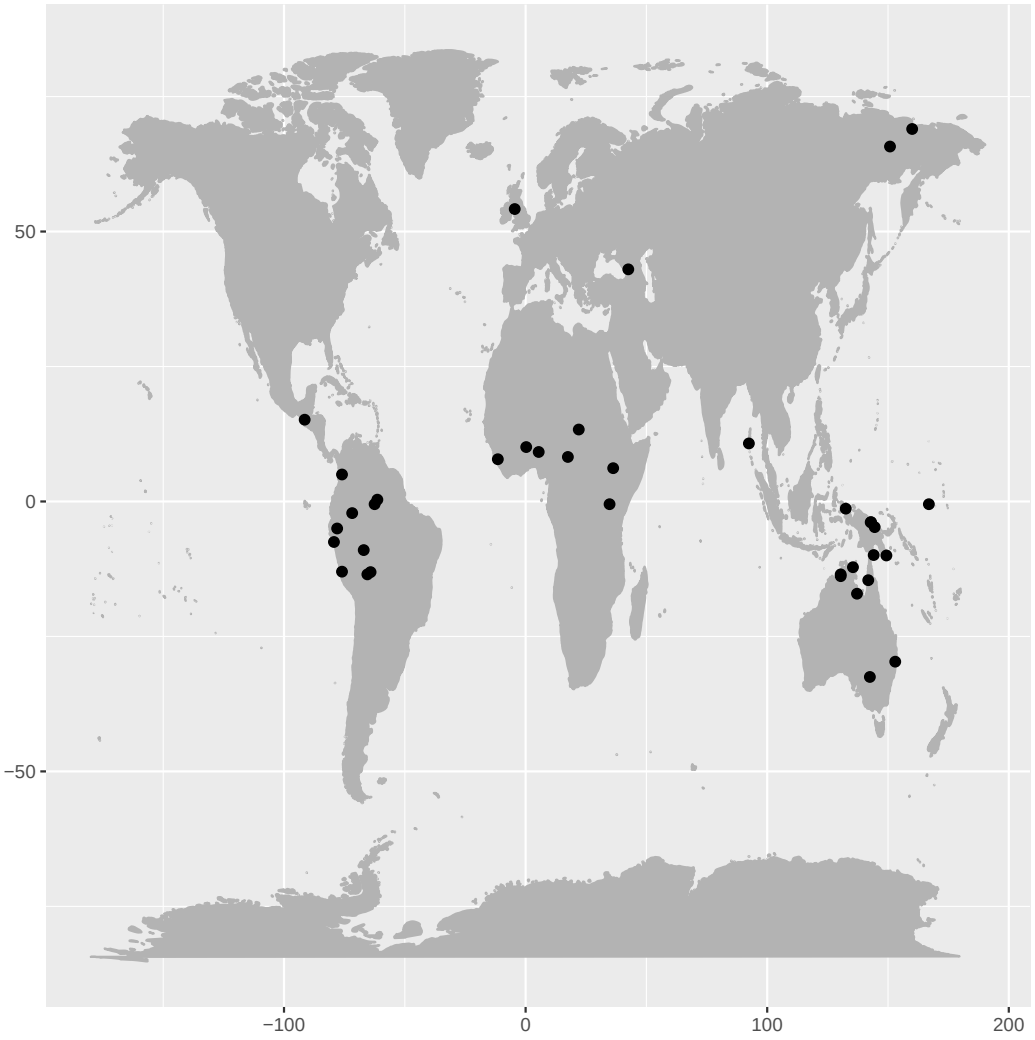
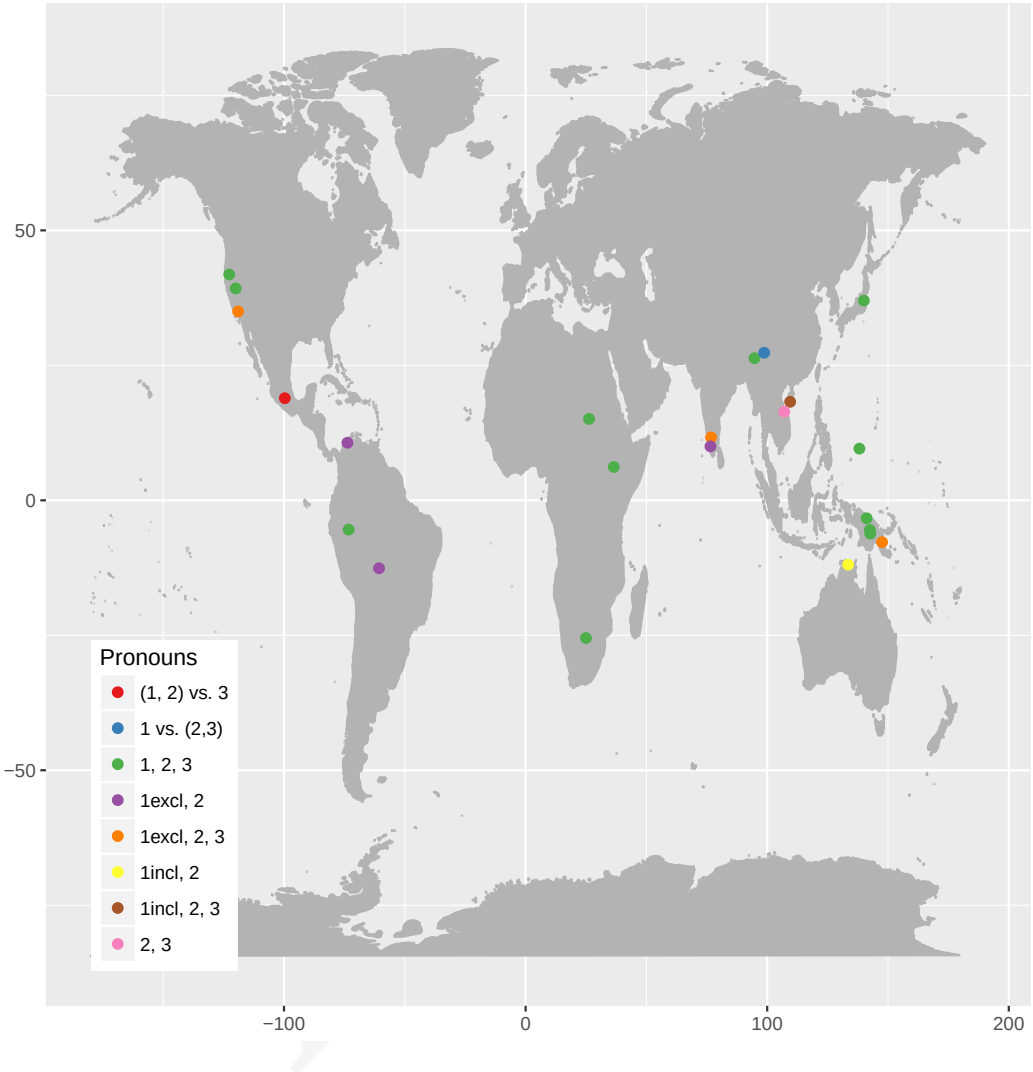


Table 6.8.: Plural in pronouns: full suppletive systems

Area	Family	Languages
Africa	Afro-Asiatic	Dime
	Niger-Congo	Anufo, Mende (Sierra Leone), Nupe-Nupe-Tako
	Nilo-Saharan	Luo (Kenya and Tanzania), Masalit, Mbay
Eurasia	Andamanese	Önge
	Indo-European	Manx
	Kartvelian	Svan
	Yukaghir	Southern Yukaghir
Pacific	Australian	Darling, Djambarrpuyngu, Madngele, Thayore, Wadjiginy
	Austronesian	Nauru
	Dagan	Daga
	Lower Sepik-Ramu	Rao
	Sepik	Ambulas
	West Papuan	Mai Brat
	Western Fly	Meriam
South America	Arawakan	Apurinã, Baré, Baure
	Aymaran	Jaqaru
	Cayuvava	Cayubaba
	Chim˙an	Mochica
	Choco	Emberá-Chamí
	Huitotoan	Murui Huitoto
	Jivaroan	Aguaruna

Figure 6.4.: Pronominal specific plural markers: shared markers



- Internal distributional constraints of such markers.

The geographical distribution is shown in Fig. 6.4. The detailed breakdown by area and family is found in Table 6.9.

Pronominal specific construction types

The most common plural construction type within this type are suffixes. Suffixes outnumber any other pluralization strategy; out of the 31 languages with specific plural marking constructions, only six take a different construction

Table 6.9.: Pronominal specific plural markers: shared markers

Area	Family	Languages
Africa	Afro-Asiatic	Dizi
	Nilo-Saharan	Midob
Eurasia	Austro-Asiatic	Pacoh
	Austronesian	Tsat
	Dravidian	Badaga, Malayalam
	Japanese	Japanese
	Sino-Tibetan	Chang Naga, Nung (Myanmar)
North America	Hokan	Shasta
	Oto-Manguean	Atzingo Matlatzinca
	Penutian	Yokuts
	Washo	Washo
Pacific	Australian	Maung
	Austronesian	Yapese
	Border	Imonda
	Bosavi	Edolo
	Trans-New Guinea	Duna, Suena
South America	Kwaza	Kwaza
	Panoan	Matsés

Table 6.10.: Imonda (Border, Papua New Guinea) Seiler (1985: p. 44)

	singular	plural
1	ka	ka id
2	ne	ne id
3	ehe	ehe id

type. The other type of markers include plural words (four languages), prefixes (one language, Choctaw) and infixes (one language only, Zoogocho Zapotec).

In what follows, a short description of the less widespread marker types within this subgroup is provided.

Four languages of this subgroup employ lexical items as plural markers: Imonda, Matsés, Nung and Sapuan. These languages do not have any genetic nor areal affiliation: Imonda is a Papuan language belonging to the Border family; Matsés,Panoan, is spoken in the Amazonian basin. Nung and Sapuan are both Asiatic, with the former linked to the Sino-Tibetan family and spoken mainly in China and the latter, Sapuan, Austroasiatic, whose few hundreds speakers are located in Laos.

The plural word employed by all Imonda pronouns (Table 6.10) is *id*, meaning 'men' (Seiler 1985: p.44). Pronominal forms are used for human referents only[7].

In Anong (Table 6.11), the plural words *iɯŋ55* and *ŋɯŋ55* are pronouns specific: the former is attached to the first person plural only (both inclusive and exclusive), the latter groups together second and third person plural pronoun. There is no specific information about the possible sources of such constructions (a classifier origin of the noun markers is instead reported Sun et al. 2009: p. 72): although they have different plural markers, it is claimed that first and

[7] Kinship terms in Imonda do express plural, through a suffixal strategy. A small group of human nouns is inherently plural: they select non-plural (that is, singular and dual number) by the means of a suffix, *-ianèi*, whose source is clearly partitive, with the meaning of 'one among the crowd'. Such strategy can be employed with personal pronouns only, where nonetheless the partitive meaning has been retained: *ka-ianèi-m ainam iaha-t* 'One of us would die quickly' (Seiler 1985: p. 44)

Table 6.11.: Anong (Sino-Tibetan, Eurasia) Sun et al. (2009: p.72)

	singular	plural
1	ɑ31	ɑ31 iɯŋ55
2	ŋɑ31	ŋɛ31 ŋɯŋ55
3	ŋ31	ŋ31 ŋɯŋ55

second person pronoun in Anong are both Tibeto-Burman and, ultimately, Sino-Tibetan in their origin (see Sun et al. 2009: p. 73).

Sapuan language selects for first and second person pronoun the plural word *mu*, which precedes the pronominal forms. There is no information about the marker: the only instance found is its similarity with the lexical noun *mu*, 'friend', borrowed from Lao neighboring language (Jacq & Sidwell 1999: p. 47), but there is no room for claims. Sapuan lacks third person pronoun, where demonstrative forms are instead used. Sapuan speakers also uses kinship terms to refer to themselves or others[8], but unfortunately it is not specified in the grammar nor by texts whether the same or another specific plural marker applies in such cases (nouns in Sapuan are unmarked for number: classifiers and numerals are used).

Zoogocho Zapotec (Otomanguean), third person pronoun takes the infix *-gak-* as a plural marker. First and second person are suppletive in plural, nouns[9] and demonstratives are left unmarked for number. Third person pronoun distinguishes animate and inanimate forms, and both of them take the *-gak-* pluralizer.

Choctaw (Barbacoan) — described in Table 6.12 — takes a pronominal specific plural prefix on first and second person plural pronoun; third person, as in Sapuan, is not present and demonstratives or lexical items meaning 'man'

[8] This is typical of some languages from South East Asia, like Vietnamese (Bernard Comrie, personal communication).

[9] Sonnenschein (2005: p.97) reports only two occurrences of plural marking in nouns out of 2000+ text examples: the particle *ka*, of unspecified origin: *yego ka* 'rivers' and *benhe bila ka* 'sister' , where *benhe* is the generic classifier for 'person'.

Table 6.12.: Choctaw (Barbacoan, South America) Broadwell (2006: p.206)

	singular	plural
1	ano'	pishno' 'we few', hapishno' 'we many',
2	chisno'	hachisno' 'you all'

'woman' 'child' according to the speaker or the referent, are employed. This prefix, with the form *ha-* is reported to be used when large number or large groups are intended. This is signaled by the meaning of the two forms of the first person plural pronoun, where there seems to be a distinction of 'paucal' vs 'multiple', as also suggested by others (e.g. Nicklas 1974: p. 30)[10].

Matsès language uses a pronominal specific enclitic to mark plural on all three pronominal forms: the marker *tedi*, which means 'all of', is also strictly connected to the dual pronominal enclitic *daëdi*, 'both of' (Fleck 2003: p. 240). Matsès human nouns take the pluralizer *-bo*, strictly and consistently forbidden on pronouns, with the only exception being the rare and archaic form *mistsbo*, with a slight difference in meaning, implying 'you all'. The marker *chedo*, 'too', 'et cetera' can be added to pronouns (it is also used on nouns, but never with *-bo*) adding an associative meaning: *ubi chedo* 'me and others'.

Pronominal specific subgroupings

The most common context involves the presence of a specific plural marker equally distributed on all three pronominal persons: almost half of the languages (14 out of 31) reflects this structure.

The second main subgrouping comprehends languages with a specific construction restricted to first and second person only, with four different languages showing this trait. Also in this case two subgroupings can be identified, where the shared construction involves first pronoun exclusive and second form (three languages) and first person inclusive and independent second

[10] *hapi-* is not a clusivity marker, as proven by Davies (1986: p. 31).

person pronoun (one language).

There are no languages, among the ones coded in the database, that occur with a specific plural pronominal construction on first person only or on second person only (with an exception that will be now described in further detail); there are, however, languages with specific plural pronominal markers restricted to third person (three languages). This is the third grouping.

The last group includes, conversely, languages where the same pronominal-specific plural marker is assigned to second and third person: three languages follow this subtype, Pacoh, Ottawa and Anong.

The use of more than one pronoun-specific plural marker is not common, in the 31 languages explored. The overall preference of these languages goes significantly towards either the employ of a single marker restricted to pronouns or, when more than one marker is attested, it is likely that at least one it is shared with nouns. Nonetheless, three languages out of more than 160 attest the presence of more than one pronominal specific plural marking strategies: Ottawa, Anong and Aztingo, which are grouped together in the fifth subtype.

In what follows, these types are described in detail.

Subtype 1: pronominal specific plural markers on all persons

This subtype gathers together all the languages with a pronominal specific marker on all three persons: the total number of languages grouped by this trait is 17. Two of them have already been described, Imonda and Matsès, the only varieties within the group carrying a plural word.

Since detailed information about the markers in the remaining languages is not available, only their areal distribution is briefly described. Language pertaining to this subtype are attested in all the macro-areas of the world, with the exception of South America; the presence of varieties from South America is also rather scarce among the other subtypes.

The area with a higher percentage of languages that belong to such group is the Pacific, with languages from Papua New Guinea and Australia: Edolo, Duna, Imonda and Yapese have this configuration (Imonda uses plural words,

Table 6.13.: Edolo (Trans-New-Guinea, Papua New Guinea) Gossner (1994: p. 27)

	singular	plural
1	ne	nilï
2	di	dili
3	e	ili

Table 6.14.: Tsat (Austronesian, Austronesia) Thurgood (2010: p. 146)

	singular	plural
1	kaw	taza:ŋʔ^{incl}, mi^{excl}
1	ha	haza:ŋʔ
1	naw	nawza:ŋʔ

^{incl} inclusive 1 pronoun
^{excl} exclusive 1 pronoun

the others take suffixes); no information is unfortunately available on the markers belonging to these varieties; it is reported, as an example, the pronominal paradigm of Edolo (Table 6.13), whose pronouns express plural through the marker *-li*.

Two further subtypes may be identified within this set, where the distinctive variable is clusivity of the first person pronoun: (i) languages where the pronominal specific plural marker is restricted to first person inclusive only, second and third and (ii) first person exclusive only, second person and third pronoun. The latter grouping is slightly more common than the former among the languages investigated. The first micro-type is found on one language only, Tsat, an Austronesian variety spoken in China (Table 6.14).

In the example provided, it can be seen how Tsat pronouns use the pluraliser *za:ŋʔ*: this construction means 'people' and it is also present in dual forms, in combination with numeral *thua*, 'two'. The only exception is first pronoun ex-

Table 6.15.: Suena (Trans-New-Guinea, Papua New Guinea) Wilson (1974: p.15)

	singular	plural
1	na	nakai^{incl}, nakare^{excl}
2	ni	nikare
3	nu	nukare

^{incl} inclusive 1 pronoun
^{excl} exclusive 1 pronoun

clusive, that is fully suppletive in the paradigm. The first person plural inclusive, on the other hand, signals plurality through a suppletive stem[11] combined with the pluralizer 'people'.

The other subtype includes three languages, Badaga (Dravidian), Suena (Trans New Guinea) and Yokuts (Yokutian). Suena pronouns form plural *-kare* (see Table 6.15). The same marker appears on Suena verbs to express plurality, as illustrated in the example below.

(35) Suena (Trans New Guinea, Papua New Guinea) Wilson (1974: p.95)

a. nakare ego noisonakare nowenakare
we.exclPL here lived-PL-remote live-PL-always
'We have always lived here'

Subtype 2: pronominal specific plural markers on first and second person only

The second main subset includes the languages where first and second pronouns form plurality through the same construction. Among these languages it can be mentioned Sapuan, where the third person is a demonstrative with unmarked

[11] Tsat is one of the few languages in my sample that takes a shared stem in singular and dual in opposition to a specific plural person/number root, counter trending the general tendency of a singular VS non-singular pronominal roots distinction

Table 6.16.: Pilagá (Guaiucuran, South America) Vidal (2001: p.127)

	singular	plural
1	hayem	qom'i
2	am	am'i

plural, while first and second pronoun use a plural word presumably meaning 'friend'. The other language is Chochtaw.

The third language is Pilagá (Guaicuran) spoken in Argentina (see Table 6.16). Pilagá lacks a specific third person pronoun: its role is covered again by a demonstrative pronoun, which has its own specific plural marker, unrelated to the many pluralizers used by nouns. The plural marker *'i* is attached to the first and second person-number stems: the first person-number stem */Qom/* can be traced to a lexical source *qom*, the name that Pilagá, Toba and Mocoví people use to refer to themselves. This root is combined with the number suffix *-i* which also appears on possessive pronouns (see Vidal 2001: p. 128). The root */am/* participates in both first and second, singular and plural pronominal forms and represents an archaic person root.

There are three languages that belong to the first subset of this grouping, that comprises languages with the same marker for first person exclusive and second person: two are from South America, Kwaza and Arhuaco, the third is Malayalam, a Dravidian language spoken in South-East Asia.

The formal correspondence between the first person exclusive and second person in Kwaza (Isolate, Brazil) is the marker *-'tsɛ*. Third person pronoun *i* is unmarked for number, and it is a distinct form from the demonstrative pronouns, also left unmarked. In these contexts, where there is a strong formal connection between first plural exclusive and second person pronoun, one may hypothesize that *-'tsɛ* does more likely signal the association with a third person (since in this specific case third person pronoun does not take such pluralizer) rather than a plural meaning. A brief comparison with neighboring languages

Table 6.17.: Kwaza, Kanoê, Aikanã (Van Der Voort 2004: p. 239)

		singular	plural
Kwaza	1	si	tsi'tsɛ^{excl}
	2	xyi	'tsɛ
Kanoê	1	aj	aj'tɛ
	2	mī	mī'tɛ
Aikanã	1	(hi')sa	sa'tɛ
	2	hi'ða	hiða'za

^{excl} exclusive 1 pronoun

seems to suggest that *-'tsɛ* is rather a plural marker, or, at least, that Kwaza pronominal system was originally based on number distinction rather than on person association. Two facts seems to support this claim.

First, the neighboring unclassified languages Kanoê and Aikanã show some apparent relics of plural number inflection in the deictic pronominal system. Although their pronominal elements do not resemble Kwaza and their pronominal systems do not involve an inclusive / exclusive distinction, the plural morpheme *-tɛ* strongly resembles the Kwaza element *-tsɛ*; in Aikanã, such marker can be traced on first person only (Table 6.17).

Moreover, the Kwaza form for the first person inclusive, *txa'na*, may originate in a loan from Tupi-Guaraní languages, where the first person inclusive has been reconstructed as **jané* by Jensen (1998: p. 498).

The only case where a pronominal-specific plural marker is restricted to first person inclusive and second person pronoun is attested in Maung (Iwaidjan) where first person inclusive and second person pronoun form plural with *-wuri*) (Maung pronominal system show other interesting typological traits).

Subtype 3: pronominal specific plural markers on third person only

The third main subtype comprises languages where the third person pronoun has a specific number marker shared neither with nouns nor other pronominal persons. Three languages follow this structure: in most cases, third person pronouns with a non suppletive plural rather indicate plural by a marker shared with nouns.

Specific pronominal plural markers restricted on third person are found in Zoogocho Zapotec, Pima Bajo (Uto-Aztecan) and Kambaira (Austronesian).

Pima Bajo demonstratives serve also as third person pronouns: they are marked with the suffix *-ma*: *ida, idama* 'this, these', *hugai, hugama* 'that, those' (Shaul 1982: p. 45). The same marker appears on the oblique forms (with no case cumulation).

Kambera language has two separate forms for third person pronoun and demonstratives: they share the same marker, the affix */d/*. Third person pronoun and demonstrative forms in Kambera show high formal correlation: *nyuna / nyuda* 'he, she, it/ they', *nina, nida* 'this, these', *nana, nada* 'that, those'(Klamer 1998: p. 108). First and second person pronoun are similar in form and they change stem to express plurality.

Subtype 4: pronominal specific plural markers on second and third person

The fourth type contains only one language, Pacoh (Austroasiatic): Pacoh's relevant feature in pronominal number marking is the use of a specific suffix restricted to second and third person, *ʔapɛː*; first person pronoun is suppletive (Table 6.18).

The form *ʔapɛː* is used as a quantifier on human nouns only, meaning 'a number of': *ʔapɛː naʔ*, 'a few people'. Moreover, *ʔapɛː* shares phonological material with numeral *pɛː*, 'three'.

Table 6.18.: Pacoh (Austroasiatic, Vietnam) Alves (2000: p. 73)

	singular	dual	plural
1	kɨ:	ɲa.	hɛ:
2	maj	ʔiɲa:	ʔipɛ:
3	dɔ:	ʔaɲa:	ʔapɛ:

Table 6.19.: Atzingo Matlatzinca (Otomanguean, Mexico Muntzel 1986)

	singular	dual	plural
1	kaakhi	ka-kwe^{incl}, ka-kwe-ɓi ^{excl}	ka-kho^{incl}, ka-khA-ɓi^{excl}
2	kaačʔi	kačʔi-kʔwe	kačʔi-kʔo
3	lihtʔA	lihtʔA-nkwe	lihtʔA-hna

^{incl} inclusive 1 pronoun
^{excl} exclusive 1 pronoun

Subtype 5: multiple pronominal specific plural markers

Only three languages indicate plurality in pronouns through different but specific constructions: one is Anong (sino-Tibetan), with two distinct markers distributed on first person in opposition to both second and third.

Another language, with the same grouping pattern (1 VS 2,3) is Ottawa (Algic), (Valentine 2001), where the suffix *-win* is shared by first person inclusive and exclusive; on the other hand, second and third person use the suffix *-waa*.

Atzingo Matlatzinca (Otomanguean) shows a different grouping (1, 2 VS 3) first and second person share the same marker, *-kho*, while third person takes a specific suffix (*-hna*). This is described in Table 6.19.

6.5. Plural markers shared between nouns and pronouns: types and internal distribution

This section explores the degree of interactions between nouns and pronouns in the use of plural marking constructions. Data analysis is performed on about 70 languages which pronouns do not show a full suppletive behavior (50 languages) nor any pronominal specific plural markers (31 languages).

Linguistic data exploration has been conducted by addressing the following research questions:

- Which pronominal person is more likely to share the plural marker with nouns (and which types of noun)?
- Are there any 'groupings' of pronouns and nouns that tend to select the same marker and what is their distribution?

Preliminary data analysis lead to the definition of two generic contexts:

- The same marker is shared by all noun types, nouns and pronouns indifferently.
- Different constructions are found among the noun types, grouping some NP types in contrast to others.

The structure of this section follows the groupings identified: the languages with a shared marker on both nouns and pronouns are treated first; table with figures and relevant information, like areal distribution and genealogical affiliation are reported.

The second half of the section deals with languages with a multiple markers, shared within nouns and pronouns: internal groupings imply the presence of further subtypes. To ensure a systematic approach, the results are shown by taking into account each pronominal person by comparing it to the related noun types in a given language, with the purpose to identify plural marking symmetries.

6.5.1. Languages with plural marking shared between nouns and pronouns

This section illustrates the subset of languages that share a plural marker on all noun types.

Preliminary results show that about ten languages of the sample belong to this set and they do concentrate mainly in two macroareas: Eurasia and Central-South America. For the 'Eurasiatic block' the languages are Tu (Mongolic), Nivkh (Nivkh), Godoberi (Nakh-Daghestanian), Prinmi and Angami (Sino-Tibetan); Meso-Southern American languages include Bororo (Bororoan), Nahuatl and Huastec (Uto-Atzecan), Rama (Chibchan) and Xerente (Je).

The plural markers used within these languages include clitic forms (Rama, Prinmi), suffixes (Xerente, Tu, Nivkh, Huastec, Nahuatl) and plural words (Miskitu).

Rama plural clitic *-lut* (with the related allomorph *-dut*), is also the 'main grammatical signal that a constituent is functioning as a nominal' (Grinevald 1988: p.90). This may suggest a relation between the plural marker and processes of nouns derivation from verbal forms, as illustrated in (36), where deverbal nominals are formed by a participial form *-ima*, a derived noun *-tahma*, an adjective and a determiner functioning anaphorically as the heads of the noun phrase.

(36) Rama (XXX) Grinevald (1988: p.91)

a. almaling-ima-lut, king-tahma-lut
die-PART-PL, head-NEG-PL
'all the dead ones, the ones without a head'

The plural suffix in Tu (Mongolic) is the marker *-si*, regularly distributed on nouns, pronouns and also demonstratives as shown in Table 6.20.

The etymological history of this marker is reconstructed: it descends from a generic Mongolic suffix, where it is attested as a single consonant suffix *-s*; the epenthetic vowel *-i*, found in Tu seems to be related to a recent development and reinterpretation, by Mangghuer speakers, of *-si* as a postnominal phonological

Table 6.20.: Mangghuer (Mongolic, Eurasia) Slater (2003)

	singular	plural	translation
Noun	bulai ge	bulai-si	'child'
	mori ge	mori-si	'horse'
Pro 1	bi	dasi	'we'
2	qi	tasi	'you'
3	gan	gansi	'they'
Dem	ni	nisi	'these'

separate word, probably in analogy with the singular marker *ge*, derived from numeral particle 'one' (Slater 2003: p.146).

The lexical item *nani*, meaning 'people', that serves as a plural marker in Miskitu language (Misupalman), has already been mentioned (see Section 5.4.1); it signals plurality on nouns and all pronominal persons, with the only exception of first person plural inclusive: *yang*, 'I' *yawan* "we" (incl), *yang nani* 'we' (excl) (Salamanca 1988: p. 250).

Huallaga Quechua follows the same structure, with the marker *-kuna* widespread on nouns and pronouns, the first person plural inclusive uses a full suppletive form *noqanchi* and the first person exclusive indicates plurality by the form *-kuna*, *noqakuna* (Weber 1989: p. 37)

Pronouns and nouns (restricted) shared constructions

This subsection described separately languages with shared marker between pronouns and nouns, where the number opposition in the latter category are restricted to some nouns only.

Trumai (Trumai) distinguishes plural opposition in nouns, pronouns and demonstratives with a plural word. Inanimate nouns are indifferent to number, and therefore left unmarked.

In Gelao (Tai-Kadai) and Iu Mien (Hmong-Mien), the same marker is shared

by pronoun and nouns, but in the latter category it is restricted human referents: it has been illustrated in Section 6.3.2 how in both languages the respective markers have undergone through processes which are similar to the suffix *men* in Mandarin Chinese, that corresponds to an associative marker progressively reanalyzed as a plural marker (Iljic 2001).

Three languages share plural marking strategy between the whole pronominal paradigm and kin terms only: one is a Surmic (African) language, Didinga. The others, Nimboran (Nimboran) and Oksapmin (Nuclear-Trans Guinean), are both Papuan. The plural marker in Nimboran, suffix (*nan*), means 'all' (Anceaux 1965).

6.5.2. Different constructions: internal groupings

The following section counts and describes the interactions between nouns and pronouns. Each pronominal person is taken into account and compared to the other noun types to attest types of occurrences between the pronominal person (or group of pronominal persons) and noun types in plural marking.

6.5.3. First person pronoun and interaction with nouns

Data exploration on the languages of the sample reports no cases where first person pronoun only takes the same marker as nouns (or any noun); the necessary condition for this interaction to apply is for the first person pronoun to 'join' another pronominal person in plural marking. This implies that the set *markerX(1Pro, N)* is not present, leaving up to evaluation two other possible groupings: (i) *markerX(1, 2, N)*; (ii) *markerX(1, 3, N)*. Type (iii), *markerX(1, 2, 3, N)*, has already been described.

The first subgroup is found in three languages: Cavineña (Pano-Tacanan), Ohlone (Yokutian) and Sierra Populuca (Mixe-Zoquean)

In Cavineña, the clitic marker *=kwana* attaches to first and second pronouns and to all nouns, from kin terms to inanimate nouns. Third person pronoun pluralizes with *=na*. The plural marker *=kwana* is homophonous to the parti-

Table 6.21.: Mutsun (Yokutian, North America) Roberts (1987)

	singular	plural	translation
Nouns	wimmah	wimmahmak	'wings'
	innis	innismak	'sons'
Pro 1	ka·n	makke^{incl} makse^{excl}	
2	me·n	makam	
3	wa·k	haysa	

^{incl} inclusive 1 pronoun
^{excl} exclusive 1 pronoun

cle *-kwana*, which means 'uncertain', giving a 'vagueness' connotation to the noun it modifies. According to Guillaume (2004: p.716), 'Although the two forms are possibly historically related, they are synchronically clearly different morphemes. They have quite different semantics. They also have a different distribution (the plural marker only occurs within an NP whereas =kwana 'UNCERT' marks all sorts of constituents)'.

Ohlone language (Yokutian) signals plurality on nouns through two suffixes, *-mak* and *-kma*, with the former occurring after noun stems ending in consonants, and the latter with ones ending in vowel. As pointed out by Okrand Okrand (1977: p. 136), albeit first and second personal pronouns are defined as suppletive with respect to plural marking, they seem to be built on a root *mak-*, identical of the post-consonantial form of the noun plural suffix (Table 6.21).

Sierra Populuca (Table 6.22) indicates plurality through two suffixes: *-tam* in first and second person, while *-yah* is used on third person and demonstratives (which both cover animate and inanimate referents). These two markers are attested on nouns as well, and their distribution is constrained by animacy: *-tam* occurs on kin terms and human nouns only, while *-yah* attaches to animate and inanimate nouns with no restrictions.

The only language attested that follows the type *markerX(1, 3, N)* is Chontal Maya (Mayan), described in Table 6.23. Chontal Maya forms plurality in three

Table 6.22.: Sierra Populuca (Mixe-Zoquean, Central America) Elson (1960: p.54)

	singular	plural	translation
Nouns	yomo	yomo-tam	'women'
	kúy	kúy-yah	'trees'
Pro 1	Ʌ̃č	Ʌ̃č–tam	
2	mĩč	mĩč–tam	
3	hĕ	hĕ-yah	

incl inclusive 1 pronoun
excl exclusive 1 pronoun

Table 6.23.: Chontal Maya (Mayan, Central America) Knowles (1984: p.202)

	singular	plural	translation
Nouns	winik	winik-lop'	'men'
	išiktak	išiktak-lop'	'women'
Pro 1	kandeʔ	kande-la*incl* kande-t'okop'*excl*	
2	ʔande	ʔande-la	
3	ʔunde	ʔunde-lop	

incl inclusive 1 pronoun
excl exclusive 1 pronoun

ways: a collective plural marker, *-il* used on nouns with no animacy restrictions; an 'abundance' marker, highly productive, used on nouns as well. The third plural marker is a suffix, *-'op* and related allomorphs (identical to the third plural dependent pronoun verbal enclitic), found on human and important animates: the same marker is shared by third person pronoun and first person plural, but only in the exclusive form. First person inclusive and second person plural use a different pronominal-specific plural marker, *-la*.

6.5.4. Second person pronoun and its interaction with nouns

The present subsection illustrates the interaction between second person pronoun and the other noun types in plural marking. Second person pronoun, unlike first person, may be grouped with a full noun type without being necessarily associated with a pronominal form.

Languages that follow this context are, for example, Armenian, Haida and Karo: in Armenian (Indo-European) and Haida (Isolate, Alaska), the same plural marker is used by second person pronoun, kin and human terms; in Karo (Tupian), the plural marker used on second person pronoun is also taken by all nouns, from kin to inanimate ones.

Armenian vestigial plural suffix *-k'* is found on second person plural form, *duk'*, in opposition to the singular *du*. First and third person plural (as well as demonstrative pronouns) do share the same marker, *-Vnk'*, that can be used on personal names with an associative meaning (Dum-Tragut 2009).

Haida, an endangered isolate language spoken in Alaska, expresses optionally plurality in human nouns by the means of a suffix, *-lʌŋ*: *łtaxu, łtaxulʌŋ*, 'friend, friends'. The same marker can be found on second person pronoun, *dʌŋ, dalʌŋ* 'you, you all' (Levine 1977: p. 171).

In Karo language (Tupian), plurality in animate and inanimate nouns is expressed by the means of a clitic, *-ʔto*. Such marker is present on second person plural form as well, as seen in Table 6.24. The other pronominal forms do mark plural number by suppletion, with the third person plural form, *tap*, completely unrelated with the singular form and which is also used on nouns, where it expresses associative meaning; first person plural inclusive form *iʔtə* resembles closely the plural clitic on second person pronoun, confirming the relation between these two pronominal persons (as also seen in Chontal Maya aforementioned).

Second person pronoun may 'group together' with third person pronoun and share the same plural marker with full nouns. Examples are provided by two Indo-European languages, Darai and KokBorok. Both languages use a plural

Table 6.24.: Karo (Tupian, South America) Gabas (1999)

	singular	plural	translation
	cibekonnoʔ	cibekonn=ʔto	'vultures'
	inãwroʔ	inãw=ʔto	'mortars'
1	õn	iʔtə^{incl}, té^{excl}	
2	ẽn	kaʔto	
3	at	tap	

^{incl} inclusive 1 pronoun
^{excl} exclusive 1 pronoun

marker denoting quantity in origin: in Darai, the suffix *-səb* means 'all' and it used on second and third person plural forms and all nouns, both animate and inanimate (Dhakal 2012). Similarly, the lexical item *rɔk*, meaning 'many', pluralizes second and third person pronouns (third person pronoun has three different forms, for human animate and inanimate referents respectively and all take the same marker) and nouns 'down' to inanimates (Karapurkar 1976).

6.5.5. Third person and nouns interaction

Third person pronoun is the pronominal form that, more than others, has a intertwined relationship with full nouns from the plural marking perspective. Third person pronouns may share the same plural marker with a restricted set of nouns only, like kin and human terms, or all nouns; furthermore, it has been seen how (mostly) second and third person pronouns may 'group together' and share the same marker with full nouns types.

In addition, third person pronoun can serve as plural markers itself: this context is strongly attested and widespread in the languages explored. In what follows, these dimensions are described.

Table 6.25.: Ineseño Chumash (Chumashan, North America) Applegate (1966: p. 168)

	singular	plural	translation
	coyini	coyini-wun	'people'
	iśon	iśon-wun	'twins'
Pro 3	kay	kay-wun	

Third person and human nouns

In two languages of the sample, Chumash and Slave, the third person plural marker is present in kin and human terms.

Ineseño Chumash (Chumashan) third person pronoun is pluralized by the means of a suffix, *-wun*, which is the main pluralization strategy in human nouns (and also used in demonstrative pronouns)[12]. Suffix *-wun* is used on verbs as a plural object marker: */k-yik-us-wun/*, 'I gave it to them', (Applegate 1966: p.168). The suffix *-wun* is used on verbs as a plural object marker: */k-yik-us-wun/*, 'I gave it to them' (Id.).

The context just presented is similar to the one identified in Northern Slavey (Athapaskan), where human nouns (as well as kin terms) and third person pronoun share the same plural clitic (see Table 6.26), */ke/*, that indicates also plurality in verbs.

[12] As mentioned in section X, first and second person pronouns in Ineseño do instead pluralize through a pronominal-specific plural affix, *-ki-*. This is shown in Table 6.25. Another fact that suggests two different development in Chumash pronominal system relies on the fact that third person pronoun in Chumash lacks a dual form, which is present on first and second person pronoun. Applegate (1966: p.169) suggests that this may be due to the demonstrative origin of the third person pronoun, replacement that had the purpose to fill the gap in the pronominal system.

Table 6.26.: Northern Slave (Athapaskan, North America) Rice (1989: p.248)

	singular	plural	translation
	t'eere,	t'eereke	'girls'
	dene,	deneke	'people'
PRO 3	ʔedeni̧,	ʔekedeni̧	

Third person and animate nouns

The plural marker may be shared by third person pronouns and animate nouns, not just human and kin, as it happens in Andamanese language (Great Andamanese). The plural marker in nouns is the clitic *-nu* and allomorphs, which original meaning is 'people'; this marker also appears on third person pronoun. The third person pronoun in Andamanese is 'covered' by the demonstrative form: *ḍu ḍune* 'that, those' (Abbi 2013: p.113).

Languages may show an animacy distinction on third pronoun plural, with the presence of two different forms: such structure reflects on nouns, that use both markers, according to their animacy. The examples that follow are from two languages of the sample: Biak and Tiriyó.

The case of Biak has already been described Section 6.3.2: there it has been shown how number markers in Biak (Austronesian) are the third person plural forms, which distinguish two separate forms (one for animate, the other for inanimate nouns) that ultimately correspond to the number markers on verbs. In Tiriyó (Cariban) the distribution of the animate/inanimate plural markers in third person pronoun is different: the third person animate plural marker, *-jamo* is shared with the other pronominal forms; the inanimate third person plural marker (and collective) *-tomo* is found on nouns, both animate and inanimate. Tiriyó human nouns, however, seems to retain, in some isolate cases, the plural marker *-jamo*: *notipe noti-jan*, 'grandmother, grandmothers' (Meira 1999: p. 220).

6.5.6. Third person pronoun as a plural marker

It has been shown how the third person pronoun forms are used to express plurality on nouns: in some languages, the third person plural form is itself used as the plural marker: the examples that follow come from two languages from South America, Jarawara and Shiriana, and two African languages, Mumuye and Mupun.

In Jarawara language (Arawan), the third person plural pronoun *mee* is used as pluralizer particle on animate nouns. It is likely that the 3nsg pronoun is a development from the noun *madi* 'people', by changes *madi* > *mai* > *mee* (Dixon 2004: p. 34).

Shiriana variety (Yanomam), another Southern American language spoken in Brazil, uses the third person plural form *pik* to mark plurality on animate nouns; inanimate nouns select the collective-partitive marker *-k*, that is found also on third person plural inanimate. It has been suggested (Gomez 1990: p. 49) a further segmentation of the third person animate plural form *pik* in order to explain its origin: *pi* may be related to the singular third person form *pa* and third person possessive pronoun *pe*, with the meaning of 'animacy', while *-k* corresponds to the partitive-collective suffix.

The last two cases presented are from Mumuye (Table 6.27) and Mupun (Table 6.28), both spoken in Nigeria, Africa, although they belong to two different linguistic families: Mumuye is affiliated to the Niger-Congo languages, Mupun belongs genetically to the Afro-Asiatic family.

In both languages, the third person plural is used as the plural marker for both animate and inanimate nouns. Frajzyngier (1993) reports that such process is typical in other Chadic-Afro-Asiatic languages; moreover, it has to be seen as a sign of recent innovation: this fact is proven by the existence, in Mupun, of suppletive human plural forms which can also take the pluralizer *mō*, as shown in the example. Moreover, such presence indicates that the inherently and suppletive plural nouns may be on the way to lose their plural meaning. The suppletive human plural forms do not show any common characteristics that would have made a possible reconstruction of a plural marker no longer in

Table 6.27.: Mumuye (Niger-Congo, Africa) Shimizu (1983: p.80)

	singular	plural	translation
Nouns	zaa	zaa-yi	'dogs'
	shon	shoo-yi	'people'
Pro 3	wu	yi	'they'

Table 6.28.: Mupun (Afro-Asiatic, Africa) Frajzyngier (1993: p.46)

	singular	plural	translation
Nouns	rēep	jìráp / jìráp mō	'girls'
	krem	krem mō	'mats'
Pro 3	wùr	mō	'they'

use. It is therefore likely that Mupun had any other plural marker other than *mō*. Among the other functions reported for this marker, there is the associative meaning with a personal name: *james mō*, "James and others", (Frajzyngier 1993: p.162).

6.5.7. Third person and all other nouns

The most frequent interaction with respect to plural marking between third person pronoun and nouns is the one that involves the presence of a shared marker between third person pronoun and all nouns, from animate to inanimate. Some of these cases have already been described above; linguistic data show the presence of more than fifteen languages where the plural marker selected by third person pronoun is shared by all nouns; this counting excludes all the contexts of 'partial sharing' with only restricted noun types or the cases of animacy distinctions already discusses throughout this section.

Table 6.29.: Third person pronoun and nouns shared markers

		singular	plural	info
Abun (West Papuan)	N	ndam	ndám	'birds'
	Pro	an	án	Berry 1995
Catalan (Indo-European)	N	perro	perros	'dogs'
	Pro	ell	ells	Wheeler et al. 1999
Kabardian (Abkhaz- Adyge)	N	wona	wonahar	'houses'
	Pro	er	eher	Colarusso 1992
Lavukaleve (Papuan)	N	mikat	mikatev	'centipedes'
	Pro	fo	fova	Terrill 2003
Udmurt (Uralic)	N	uj	ujos	'nights'
	Pro	so	soos	Winkler 2001
Urarina (Isolate)	N	kirimata	kirimataʉru	'fishes'
	Pro	aka	akaʉru	Olawsky 2006
Wolaytta (Ta-Ne Omotic)	N	mitta	mittata	'trees'
	Pro	ii	eeta	Lamberti & Sottile 1997
Warao (Isolate)	N	sa	satuma	'bats'
	Pro	tai	tatuma	Romero-Figueroa 1997

Such context is attested and equally distributed in the macro-areas, except for North American languages: these varieties seems to be less likely to share plural constructions in third person and all nouns; when it happens, the marker tends to cover only a specific set of nouns, especially kin terms and human.

An overview of the languages following this type is provided in Table 6.29.

Table 6.30.: Languages with plural distinction on nouns only

Area	Family	Language
North America	Igboid	Blackfoot
North America	Siouan	Assiniboine
North America	Iroquoian	Tuscarora
North America	Yuman	Hualapai
South America	Aymaran	Jaqaru
Africa	Bongo-Bagirmi	Mbay

6.6. Languages with plural forms restricted to nouns

Data exploration has revealed the presence of a number of languages with plural distinctions restricted to nouns only, excluding pronouns and demonstratives. They are now reported in Table 6.30, since this context is quite uncommon:

It is important to point out that in the list are found languages that do not have free pronouns (and therefore lack any marking) and languages that actually have independent personal pronouns, indifferent to number. This happens in Tuscarora, Blackfoot and Mbay. Hualapai has independent pronominal forms (for first and second person, a demonstrative pronoun functions as third person pronoun) and plurality is distinguished only through verbal suffixes.

Jaqaru (Aymaran) provides a good example, although its noun plural marker has been signaled as a recent development. However, it is worth mentioning that this recent 'adoption' of the plural form seems to be developed in nouns prior to pronominal forms, fact that suggest a non traditional number marking assignment, according to the implication described by the Animacy Hierarchy.

Table 6.31.: Languages without plural marking in nouns

Area	Family	Language
Pacific	Oceanic	Nauru
	Pama-Nyungan	Thayorre
	Anson Bay	Wadjiginy
	North-Central Bird's Head	Maybrat
	Eastern Daly	Matngele
	Annaberg	Rao
	Kamasau	Kamasau
South America	Waimiri-Atroari	Waimiri Atroari
	Jivaroan	Aguaruna
	Kwaza	Kwaza
	Arhuaco	Ika
	Salishan	Bella Coola
North America	Mixtecan	Chalcatongo Mixtec
Eurasia	Bahnaric	Sapuan
	Kam-Tai	Mulao
	Basque	Basque
	Khasian	Khassi
	Katuic	Pacoh
Africa	Aymaran	Wolof
	Gbeya	Gbeya
	Aymaran	Gbini

6.7. Languages with no plural forms in nouns

For convenience, in this section are reported all the languages of the sample that do not distinguish plural (and number in general) in full nouns: however, it has to be specified that the list can be biased; the sample, the purpose of this research, has been structured in order to gather languages with number distinctions on at least one nominal type. This means that languages with no number distinctions in full nouns are here underrepresented. Nonetheless, a number of languages with no plural distinction in nouns have been integrated as 'control languages'. A list of these languages is provided in Table 6.31 and the geographical distribution is shown as part of Fig. 6.1.

6.8. Notes on demonstratives forms

This chapter concludes with some remarks on demonstrative pronouns and their interaction with nouns and pronouns with respect to number marking.

The main goal of this section is the investigation of demonstrative pronouns from the plural marking perspective, with a focus on specific questions:

- The construction types (suffixes, affixes, plural words) used to express plurality;
- The cross-linguistic distribution of demonstrative pronouns with distinct plural markers in opposition to languages with demonstrative forms insensitive to plural distinction;
- Intra-linguistic distribution of the plural markers used in demonstrative pronouns: plural markers shared with nouns, pronouns and markers restricted to demonstrative pronouns;
- Presence and cross-linguistic distribution of demonstrative pronouns used as third person pronouns.

The data exploration that follows is based on the subsample of the 160 languages coded in the database. 59 languages out of 160 have demonstrative pro-

nouns indifferent to plural marking and 55 with plural distinctions on demonstrative pronouns.

A preliminary classification of the 55 languages with plural demonstrative forms show the following distribution:

- 22 languages show demonstrative-specific plural markers;
- 25 languages have demonstrative forms that form plural as nouns; in such case, demonstrative pronouns may select a strategy used by all noun types or they may attach a plural marker employed by a specific noun type, e.g., human;
- 8 languages have demonstrative forms that share plural marking strategy with the pronominal forms. The grouping can relate to the full pronominal paradigm or it can be restricted to a specific pronominal form.
- 13 languages show a plural marker by all NP types: noun, pronouns and demonstratives altogether.

The geographical distribution of these types is displayed in Fig. 6.5. The detailed breakdown by area and family is found in Table 6.32.

Demonstratives and full nouns

Almost half of the languages with a plural overt marker on demonstratives, show the presence of the same marker in at least one full noun (25 out of 55). Of these 25 languages, 15 have a shared marker between all full nouns and demonstratives, with no restriction.

All macroareas are represented, with a more significative presence of languages from Africa and from the Pacific area; Northern American languages are also attested, with three languages: Siksika (Algic), Chimalapa Zoque and Tillamook. Other languages from the Eurasiatic plateau are attested: Yukaghir, Svan and Korku. The only language from South America following this type is Cayubaba (Isolate).

Demonstrative pronouns may share the plural marker with kin terms only; this happens with languages that show plural distinctions constrained on kin

Figure 6.5.: Plural marking in demonstratives

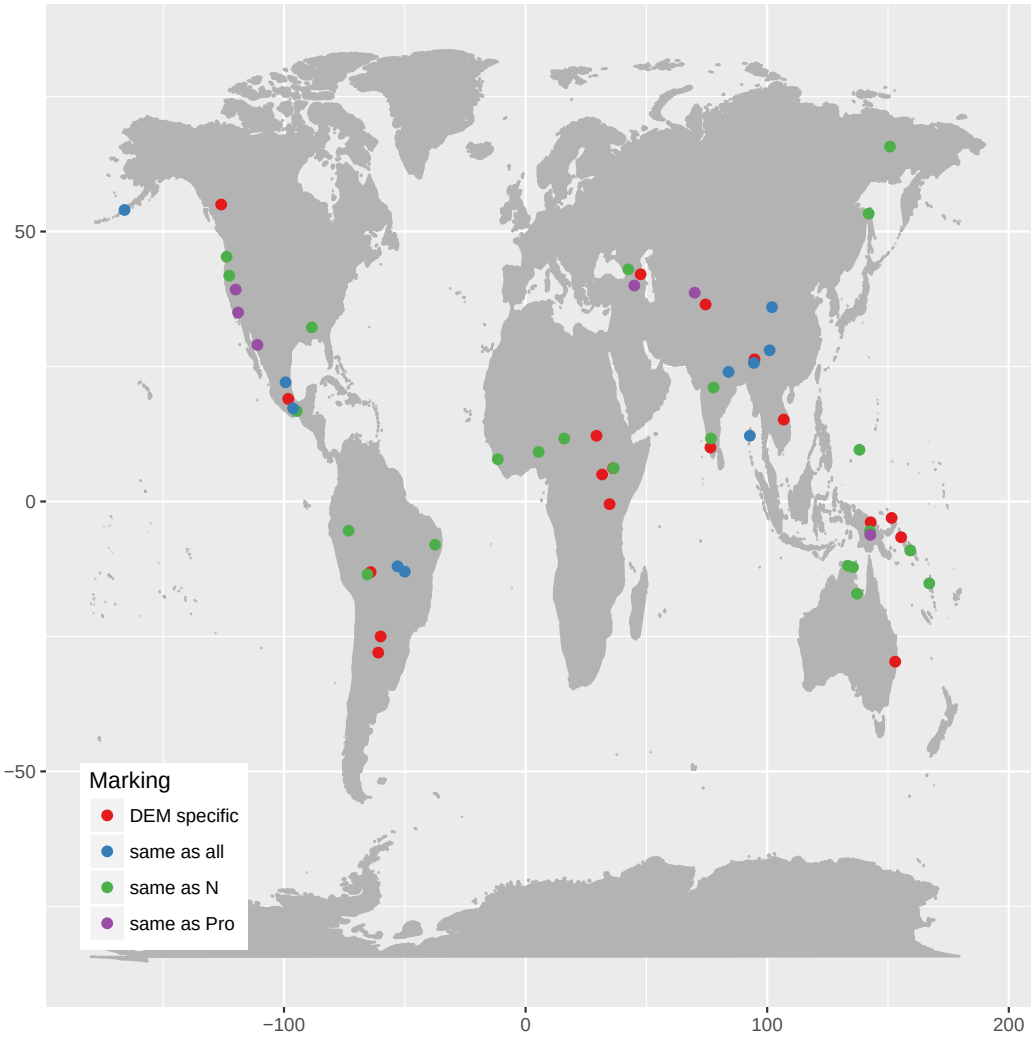


Table 6.32.: Plural marking in demonstratives

Area	Family	Languages
Africa	Afro-Asiatic	Dime, Dizi
	Niger-Congo	Mende (Sierra Leone), Nupe-Nupe-Tako
	Nilo-Saharan	Ama (Sudan), Bagirmi, Bari, Luo (Kenya and Tanzania)
Eurasia	Altaic	Tu
	Andamanese	A-Pucikwar
	Austro-Asiatic	Korku, Sapuan
	Burushaski	Burushaski
	Dravidian	Badaga, Malayalam
	Eskimo-Aleut	Aleut
	Indo-European	Armenian, Darai, Tajik
	Kartvelian	Svan
	Nakh-Daghestanian	Dargwa
	Nivkh	Gilyak
	Sino-Tibetan	Angami Naga, Chang Naga, Northern Pumi
	Yukaghir	Southern Yukaghir
North America	Hokan	Shasta
	Mixe-Zoque	Chimalapa Zoque
	Muskogean	Choctaw
	Oto-Manguean	Zoogocho Zapotec
	Penutian	Yokuts
	Salishan	Tillamook
	Uto-Aztecan	Central Nahuatl, Pima Bajo
	Washo	Washo
Pacific	Australian	Djambarrpuyngu, Maung
	Austronesian	Sakao, Yapese
	Bosavi	Edolo
	East Bougainville	Siwai
	Kuot	Kuot
	Solomons East Papuan	Lavukaleve
	Trans-New Guinea	Duna
South America	Arawakan	Baure
	Cayuvava	Cayubaba
	Guaicuruan	Mocoví, Pilagá
	Macro-Ge	Fulniô
	Panoan	Matsés
	Trumai	Trumaí

terms, like Sakao (Melanesian) and Djambarrpuyngu (Australian), but also in languages with different markers for the various NP types, like Maung (Iwaidjan): in Maung, nouns select different plural marking strategies, like reduplication (on human). Among these less common plural marking strategies, a suffix, *-awg*, which is shared by a small subset of kin terms and it is used by demonstrative forms (Capell & Hinch 1970).

Wakaya language (Australian) indicates plural marking in animate nouns by two suffixes: *-wul*, used optionally on animates only, and *-manha*, specific for human terms, as in *ngarremanha* 'men'; this marker is also found on demonstrative pronouns: *imu imumanha* 'this, these' (Breen 1974: p. 155). Demonstrative pronouns and nouns also share the same dual marker, suffix *-wiy*.

The same context can be identified in languages from other areas: Shasta (Hokan, North America, Silver (1966)) demonstratives take the same collective suffix (*-yă·war*) restricted to animate nouns.

Plural marking strategy shared by all noun types

12 languages explored in the subsample show a regular distribution in the plural strategies: the same plural marker is present on nouns, pronouns and demonstratives with no relevant exceptions. Most of the languages that follow this type are from Eurasia area, for instance in Angami, Darai, Tu and Andamanese. Such regular marking is attested in other areas and language families, like in Huastec (Mayan) for the North-Central America, Xerente and San Miguel Zapotec for Southern America area and Trumai language located in the Pacific. Aleut language provides an example of languages from North America.

In Table 6.33 it is reported an example from Ninam.

Demonstrative-specific plural markers

In about 22 languages demonstrative pronouns show specific plural markers that are not used on any other noun type. In the table that follows all the languages coded which show this feature, grouped by language family, are reported, along with information about the construction type used. In Table 6.34

Table 6.33.: Ninam

	SG	DU	PL
1	ca	cehek	camak
2	wa	wehek	wamak
3 anim	pa	kip	pik
3 inan	thə		thə-k
DEM	ihi	ihi kip	ihi pik
N anim	irit 'child'		irit pik 'children'
N inan	cahi 'meat'		cahi-k 'lot of meat'

Table 6.34.: Lakhota, Ingham 2003: p. 39

	SG	DU	PL
NOUNS			halhate ‘magpies’
1	miś		uŋkiś
2	niś		/
3	iś		/
DEM PROX	le	lenaos	lena
DEM NEU	he	henaos	hena

it is a reported the nominal number system of Lakhota language. In Lakhota nouns, plurality is expressed generically by reduplication, although it is rare. Second and third person pronouns do not distinguish plural; demonstrative pronouns have specific number markers in dual and plural as well.

Demonstrative forms as third person pronouns

The map in Fig. 6.6 explores the distribution of the languages in the sample where the demonstrative pronoun function as a third person pronoun. The list includes approximately 27 varieties out of 200 present in the sample. The detailed breakdown by area and family is found in Table 6.35.

Figure 6.6.: Languages with DEM=3Pro

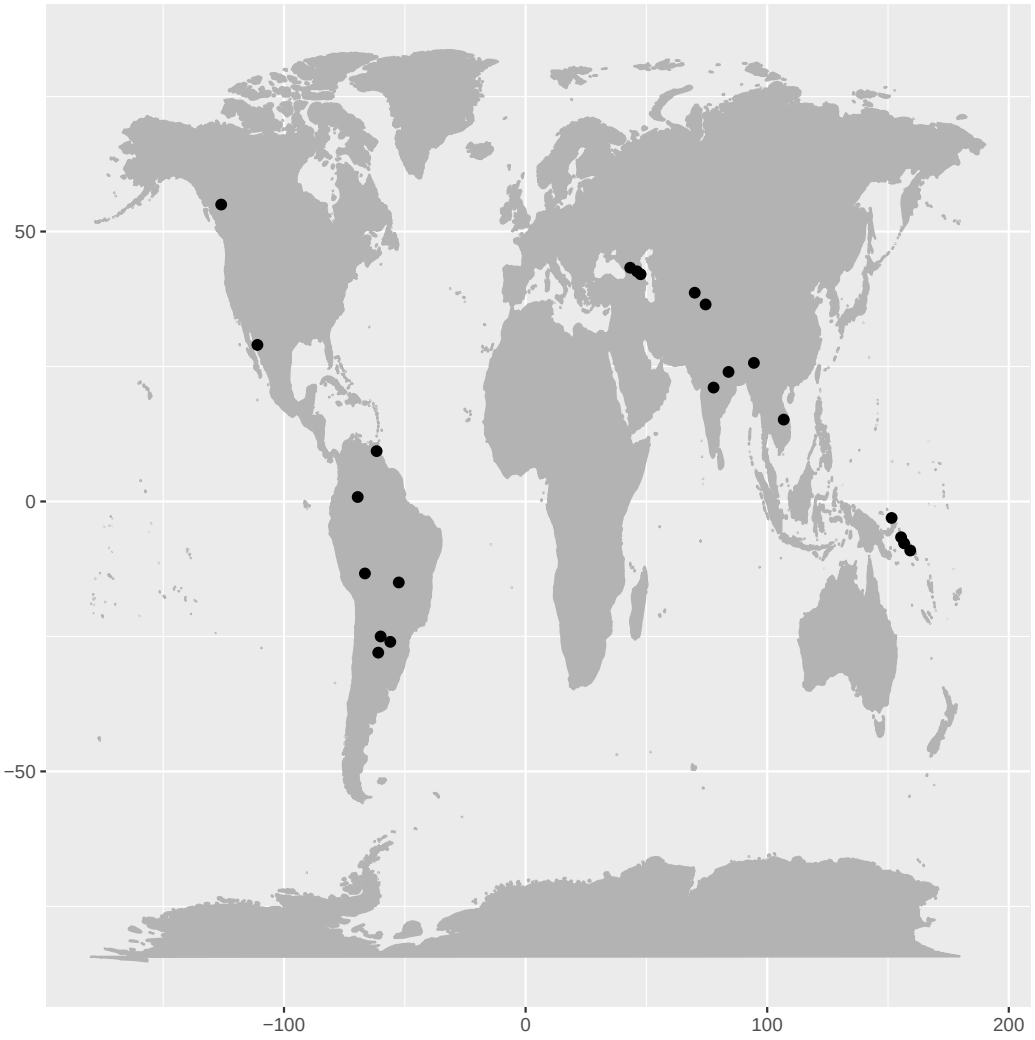


Table 6.35.: Languages with DEM=3Pro

Area	Family	Languages
Eurasia	Austro-Asiatic	Korku, Sapuan
	Burushaski	Burushaski
	Indo-European	Darai, Tajik
	Nakh-Daghestanian	Dargwa, Ghodoberi
	Northwest Caucasian	Kabardian
	Sino-Tibetan	Angami Naga
North America	Uto-Aztecan	Pima Bajo
Pacific	East Bougainville	Siwai
	Kuot	Kuot
	Solomons East Papuan	Bilua, Lavukaleve
South America	Guaicuruan	Mocoví, Pilagá
	Macro-Ge	Xavánte
	Tacanan	Cavineña
	Tucanoan	Guanano
	Tupian	Paraguayan Guaraní
	Warao	Warao

The count of languages show a number of about 27 varieties out of 200, with the following areal distribution.

The areal classification reveals that none of the African languages of the sample show this feature. African languages tend to have rich and detailed pronominal and deictic paradigms, so there are less gaps to be filled. Cases from languages include Kawaiisu (Table 6.36) and Cavineña (c6.37) Uto-Atzecan and Tacanan languages respectively.

Table 6.36.: Kawaiisu, Zigmond et al. (1990: p. 46)

	SG	DU	PL
NOUNS	momoʔo 'woman'		momoʔo-mɨ 'women'
1	nɨʔɨ	tami	tawa incl
2	ʔimi		mumi
DEM PROX	siʔina		siʔimɨ
DEM NEUTRAL	samana		samamɨ
DEM DISTAL	suʔuna		suʔumɨ

Table 6.37.: Cavinena, Guillaume (2004: p.78)

	SG	DU	PL
1	i-ke	yatse	ekwana
2	mi-ke	metse	mikwana
3 neutral	tu-ke	tatse	tuna
3 proximate	riya-ke/ry-ke	reste	rena

Demonstratives and pronouns

Cases from languages where demonstratives do follow personal pronoun only are less common. The most typical context where a demonstrative form has a plural interaction with a pronominal form is the one that involves usually third person pronoun and other noun types. A construction shared exclusively by demonstrative forms and pronoun is more limited, and in some case it is attested in languages where nouns do not distinguish plural forms.

Edolo (Papuan, Gossner 1994) pronouns indicate plurality with the marker *-li*. The same marker is used on demonstratives; nouns do not distinguish plural, but human nouns do sometimes attach a pluralizer suffix. A striking similar structure is given by Washo (Isolate) language: the pronominal-specific number markers (both dual and plural) are taken by demonstrative pronouns, with nouns that select other strategies, like reduplication. Example shown in Ta-

Table 6.38.: Burushaski pronominal system and demonstratives Lorimer (1935: p. 123)

	SG	PL
1	jɛ	mi
2	uːn	ma
3 hum	iːnɛ	uːɛ
3 anim	iːsɛ	itsɛ
3 inan	iːte	iːkɛ
DEM hum	kiːne	kuːɛ
DEM anim	gʊsɛ	gʊtsɛ
DEM inan	gʊtɛ	gʊkɛ

ble 6.38 is from Burushaski, a language isolate spoken in Pakistan.

7. A typology of number systems

7.1. Introduction

This chapter explores the number systems of the languages in the sample. The main purpose of the survey performed in this chapter is to illustrate the distribution of the number systems in the macro-areas, identify relevant subtypes and describe notable interactions at the number system level.

The chapter is organized as follows: for convenience, data exploration is carried on by grouping the languages by the macro-areas they belong to, along with the following partition.

- North America
- South America
- Europe and Asia
- Australia and Papua New Guinea.

For each area, the nominal number systems are attested, and the languages assigned to the respective type. Internal sub-groupings, when attested, are presented and discussed.

The number system of a language are analyzed from both the noun types and the number values perspective and the interaction between these variables is evaluated. In cases linguistic data is available, by exploring this interaction it is possible to infer some tendencies in the development of these systems.

One of the most interesting 'marks' that allows to track the spreading path of a value, or a marker, throughout the number system, is the presence of relevant features associated with a marker, that may recur at a number value level or

restricted to a lexical category: one of this traits is the cumulation of the number markers with other grammatical features.

A whole subsection is thus dedicated to the number-gender interaction. This dimension is examined from three different angles: (i) the nominal types that tend to carry gender/number markers, (ii) the number value(s) that show(s) recurrently gender/number markers, (iii) the areal constraints that are eventually identifiable.

7.2. Gender interaction in nominal number marking

Introduction

This chapter aims to provide an overview of the interaction between number and gender attested in the languages of the sample. The main goal is the description of this interaction by addressing the following core concepts:

- Identification of languages with cumulative gender/number markers;
- Distribution of the gender/number markers within the noun types;
- Distribution of the gender/number markers within the number values;
- Relevant structural features at the number system level.

In order to ensure a systematic approach, 'plain' gender markers not cumulative with number are not taken into account in this data exploration; similarly, gender distinction in pronominal forms is not considered, unless the cumulative gender/number markers are clearly identifiable: such procedure is justified by the consistent number of suppletive forms in pronouns, where the not-segmentable person/number roots do not allow a straightforward recognition of the eventual gender/number markers.

The rich noun classes found in a consistent number of African languages, most of them belonging to the Bantu language family, are also not taken into account: such classes are highly specific of the area, making it difficult to claim

Table 7.1.: Gender/number markers on nouns

Language	Number	Marker type	Nominal type
Tuscarora	SG	prefix	human
Zuni	SG	prefix	all nouns
Tunica	SG	suffix	all nouns
	DU	suffix	all nouns
Cubeo	SG	suffix	human nouns
Desano	SG	suffix	human nouns
Wanano	PL	suffix	kin terms, third person pronoun
Bora	DU	suffix	N animates, pronouns
Rikbaksta	PL	suffix	human nouns, 2 pro
Lavukaleve	SG	suffix	human nouns
Turkana	SG	suffix	all nouns
	PL	suffix	all nouns
Karimojong	SG	suffix	all nouns
	PL	suffix	all nouns
Korana	SG	suffix	human, pro, dem
	DU	suffix	human, pro, dem
	PL	suffix	human, pro, dem

any kind of generalizations; at any rate, they should be treated separately.

Table 7.1 show the languages in the sample with cumulative gender/number markers.

The macro areas with the broader presence of cumulative gender/number markers are Africa and South America. No Eurasiatic language of the sample is present. The only language of the Pacific included in this group is located in Papua New Guinea. Three South American languages, Cubeo, Desano and Wanano, are genetically related, with markers that show strong affiliation and formal similarities. The same applies to the two Nilotic languages of the subgroup, Karamojong and Turkana: their markers seem to be derived from a common source.

7.2.1. Gender/number markers and number values

This subsection is devoted to the description of the gender/number markers as they are linked to a number value: the aim is to capture the detailed presence of these markers and the relevant traits they are associated with (i.e., number value and distribution).

The first step towards a classification of these markers as they appear in languages consists in verifying their amount of spread at the number category level: cumulative markers may appear, on a given language, on all the number categories expressed or either restricted to a number value.

Cumulative gender/number markers on singular only

As aforementioned, the presence of gender/number cumulative markers may be limited to a number value only: in five languages of this subgroup, number markers cumulative with gender are found on singular: this happens in Tuscarora (Iroquonian), Zuni (Isolate, North America), Cubeo and Desano (Tucanoan) and Lavukaleve.

Number markers in Tuscarora are cumulative with gender on singular value: dual and plural markers, which are two prefixes, *neye-* and *kaye-* respectively, do not show any gender distinction and they are restricted to human nouns only. Three different prefixes are used to express singular in Tuscarora nouns, and they do distinguish masculin (*ra-*) and feminine *(e-)* human; an additional marker is used on generic inanimate singular (Williams 1976).

Cubeo, a Southern American language spoken in Venezuela, shows the same structure: singular overt suffixes express also gender, feminine *-ki*☒and masculine *-ko*. Nominal plural marking is assigned to the suffix *-wa*, indifferent to gender distinctions (see Morse & Maxwell 1999). Desano is a Tucanoan language strictly related to Cubeo and such structure is preserved, with some further peculiarities. Desano human nouns consist of a root plus a suffix which indicates gender and number: *-gi* (singular masculine), *-go* (singular feminine). For example, *bahi*, 'child', when inflected becomes *bahigi* 'small boy' or *bahigo*

'small girl'; when inflected with *-ra* (plural) it gives the form *bahira* 'children' . Nominalized verbs which are animate also fall into this category, e.g., *biisigi*: 'the one (MSG) who knows' (see Miller 1999: p. 54).

A similar tendency is reported in Wanano (Tucanoan). Wanano language has a interesting number system which is worth illustrating in further detail. Wanano nominal number marking system involves the presence of gender/number cumulative markers: *-dubia* or *-sadubia* (*/sa/* is the possessive marker) plural feminine, suffix derived from the lexical noun meaning 'woman' (*dobia*), and masculine plural suffix *-su*[illegible]*a*, meaning 'man' (Stenzel 2004: p. 130). The markers are found on kin terms and partially on third person pronoun plural. This 'partial assignment' to the third person plural pronoun is justified by the fact that the pronominal form, in order to express masculine gender and plural number, selects the suffix *-da* and not *-su*[illegible]*a* as expected. Moreover, the suffix *-da* is the gender-indifferent marker used in Wanano to express plurality in human and high animate nouns. Third person pronoun and human nouns do share the same marking strategy also on singular, since they both employ the generic suffix *-ro*. Table 7.2 clarifies the asymmetric distribution of the gender/number markers in this language.

An additional example comes from Zuni language (Table 7.3), where the three noun classes have specific markers which are carried by the singular value; dual and plural do not distinguish gender.

Lavukaleve language reports two vestigial remnants of singular overt markers, *-a* and *-m*: they also mark feminine and masculine gender (see Table 7.4). Non-singular markers do not carry further grammatical information.

Cumulative gender/number markers on plural only

Rikbaksta (Nuclear Macro-Gê, de Jesus Silva 2011) does present cumulative markers only on plural forms; singular is not overt marked. What is striking about such gender/number markers is their distribution: they are found on nouns and also on pronouns, but on second person plural pronoun only.

Table 7.2.: Wanano (Tucanoan, South America) Stenzel (2004: p. 68)

	singular	plural	translation
Nouns	wabi-ro	wabi-sʉba	'older brother, older brothers'
	wabio-ro	wabio-sadubia	'older sister, older sisters'
	die-ro	die-ya	'male dog, dogs'
	die-ro-koro	die-ya	'female dog, dogs'
Pro 1	yʉ'ʉ	badi^{incl}, sa^{excl}	
2	bʉ'ʉ	bu'sa	
3	ti-ko-ro^{f}, ti-ro^{m}	ti-dubia^{f}, ti-da^{m}	

^{incl} inclusive 1 pronoun
^{excl} exclusive 1 pronoun
^{f} feminine
^{m} masculine

Table 7.3.: Number marking in Zuni Newman (1996: p.56)

Noun class	Noun root	SG	PL
1	lu 'ash'	lu-ʔleʔ	lu-weʔ
2	ɬe 'board'	ɬe-mmeʔ	ɬe-weʔ
3	lupa 'box of ashes'	lupa-ʔe	lupa-weʔ

Table 7.4.: Gender/number suffixes in Lavukaleve (Terrill 2003: p. 95)

	SG	DU	PL
M	-m	-l	-v
F	-a		

Table 7.5.: Rikbaksta (Nuclear Macro-Gê, South America) de Jesus Silva (2011: p. 39)

	singular	plural
N	iʃtʃekba	iʃtʃekba-ka^{f}, iʃtʃekba-tʃa^{nf} 'old men, old women'
1	ikɽa^{f}, uta^{nf}	katʃa
2	ikia	ikiaha-ka^{f}, ikiaha-tʃa^{nf}
3	a-tatʃa^{f}, a-ta^{nf}	a-ʃiɽa^{f}, a-ʃa^{nf}

^{f} feminine
^{nf} non feminine

Cumulative gender/number markers on dual only

Bora language, conversely, takes cumulative number markers on dual number only: masculine and feminine are distinguished through the suffixes *-tsi* and *-pi* respectively. Such gender/number suffixes are extensively found on animate nouns, the whole pronominal paradigm (all three persons) and also on demonstratives. Plural meaning is carried by the marker *-mu* on nouns and other specific plural suffixes on pronouns and demonstratives, which, however, do not show any cumulation with gender value Thiesen (1996: p.27).

Cumulative gender/number markers on more than one number category

Finally, some languages may show cumulative markers on more than one number value, or on all number values.

The former case is illustrated by Tunica language (Table 4.3). Tunica has cumulative markers on both singular and dual, as shown in the table below; however, the cumulation between gender and dual value is only partial: while masculine gender does have specific markers for singular, dual and plural, in feminine gender the gender/number opposition is restricted to singular vs non-singular marking.

The latter case, where the gender/number markers are present and widespread

Table 7.6.: Korana (Khoe, Africa) Maingard (1962: p. 17)

	singular	dual	plural
N	-s^f, -p^m	-sara^f, -kara^m	-di^f, -ku^m
1	-te^f, -tir^m	-sam^f, -ka^m	-si^f, -kie^m
2	-sa^f, -tsa^m	-saro^f, -karo^m	-sau^f, -kao^m
3	-s^f, -p^m	-sara^f, -kara^m	-di^f, -ku^m

^{f} feminine
^{m} masculine

on all the number values is attested in the remaining three languages: they are located in Africa, and two of them (Karamojong and Turkana) are also genetically related (they do both belong to the Nilo-Saharan family). The Table 7.6 shows number markers in Korana (Khoe), which spread on all noun types (in Korana they are restricted to human nouns) and on pronouns regularly.

7.2.2. Summary

Based on the linguistic data explored and described in Table 7.1 and along the section, some tendencies about the interaction between gender and number in the number systems may be suggested: the number value that is more inclined to take gender/number cumulative markers is singular. This occurrence is particularly interesting when compared to the total number of singular overt markers attested in the languages of the sample.

The noun types that are more likely to select gender/number markers are mostly nouns that denote living beings, specifically human: this may be due to the fact that human nouns, more than any other noun type, may be in need to specify their gender: this applies in particular in un-specified noun root stems, as shown by the example from Desano language, where a lexical item like *bahi*, meaning generically 'child', is disambiguated through the use of a gender affix (which carries also the number value).

It is also worth to mention that when cumulative gender/number markers

are found on more than one number category, as dual and plural, or even on the whole nominal number system, their distribution on nominal types will cover completely the nominal types as well. The result is a rich, yet regular, nominal number system where every 'slot' is filled (the example from Korana is emblematic, with gender/number distinctions on each pronominal person for all number values).

Such regularity is also confirmed by the formal characteristics of the construction types: their structure is similar for each number value; by exploring the values of the nominal number marking system of Korana (and same conclusions that can be claimed for Turkana and Karamojong), one may notice that singular, dual and plural are all expressed by suffixes, which are clearly related to one another.

An additional consequence of this joint, systematic and regular number marking assignment is the less common and pervasive presence of alternative number marking in these languages, when compared to all the other cases presented in the previous chapters, where it has constantly repeated how number markers tend to 'fluctuate' from one nominal type to the other with no restricted boundaries and all cases of NPtype-specific number markers come with exceptions.

Regular number markers of this kind suggest regular development of the nominal number system, that is, from a common source for all number values: within this context, it is unlikely to hypothesize different times and origin of number markers for each number value; the latter is rather the case of number systems where the construction types used to express singular, dual and plural are formally different, fact that leads, consequently, to different sources markers. The presence and distribution of the different types of number systems are investigated throughout the next subsections.

7.3. Number systems interaction

7.3.1. Introduction

In order to be the most systematic as possible, the number systems of the languages of the sample are arranged and investigated by macro-area. For each macro-area, the following information is reported:

- Different number systems and number of languages for each type;
- Tendencies in distribution of the number values within the languages;
- Other relevant characteristics.

7.4. Number systems in North American languages: types and distribution

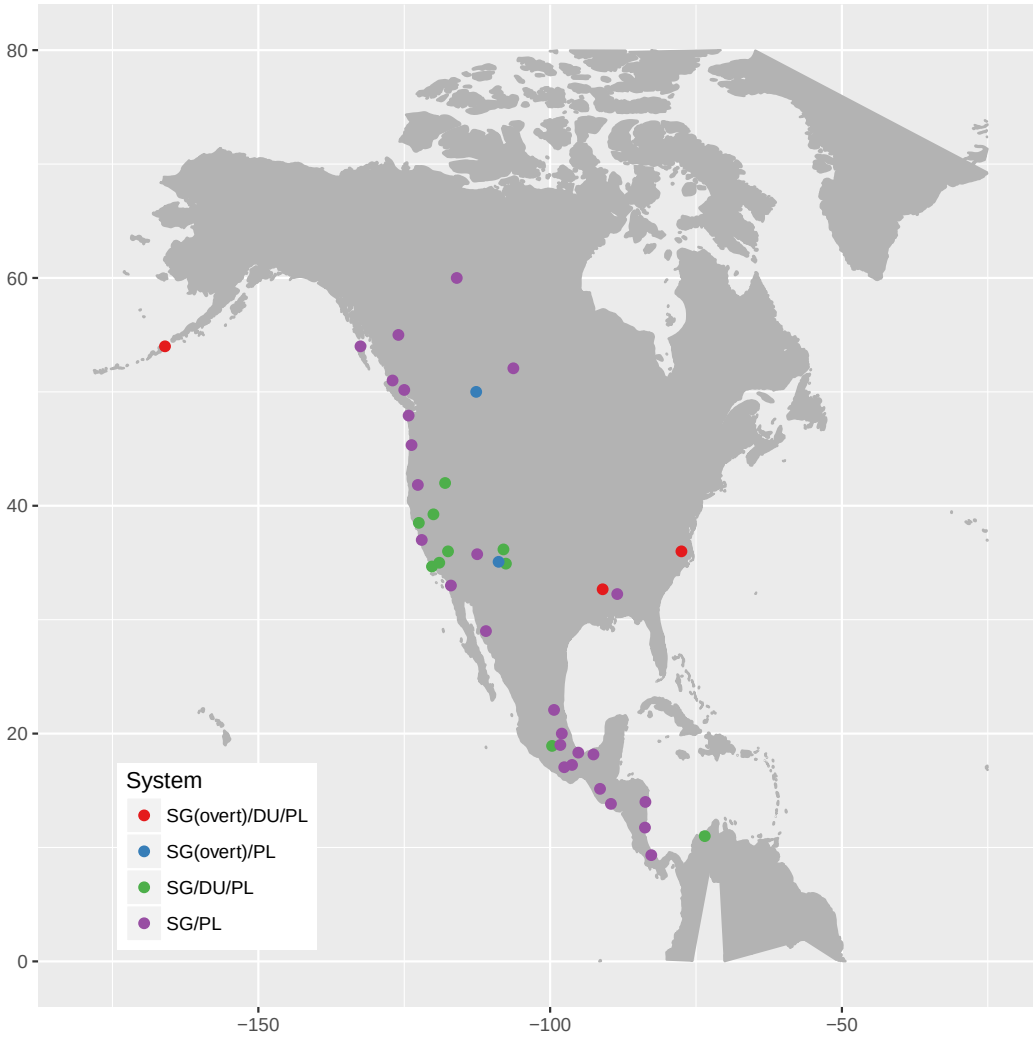
This subsection explores the main number systems identified in all the languages of the North America macro-area included in the sample. The first step of this analysis consists in a general identification of the number systems and in a preliminary classification of the types identified.

Among the North American languages of the sample, four main nominal systems have been identified:

A. Singular/Plural (34 languages)
B. Singular (overt)/Plural (2 languages)
C. Singular/Dual/Plural (14 languages)
D. Singular (overt)/Dual/Plural (6 languages)

The geographical distribution of languages grouped by system types is shown in Fig. 7.1. Within the sample, the most common number system found in the North American languages is type A, where singular form is unmarked and plural is marked. The second most widespread type, C, adds dual expression. The types B and D include overtly expressed singular values and they are less

Figure 7.1.: Number systems in North American languages



attested.

7.4.1. Type A: Singular/Plural

The Singular/Plural system, however, does not present itself with the same distribution in all languages that reflect this type: the internal distribution of these values may vary. Based on this internal variation, further sub-types are determined; since there is only one value that is expressed overtly within this general types, any sub-categorization reflects the distribution of plural number only (seeTable 7.7).

Subtype A1

The first sub-type is defined by the presence of the overt plural marker on all noun types: no main lexical category is left unmarked. The North American languages from the sample that belong to this subset are 15 out of 28, with therefore a tendency of distinguish plurality in an homogeneous way across the nominal categories. It has to be pointed out that the pluralization strategies may differ, as shown extensively in the previous chapters, across the lexical categories and within the same category.

Linguistic data presented in table also provides more detailed information about the distribution of each value: in five of the Northern American languages of subset A1, plural marking is present on all three of the lexical categories, but with a different internal presence and distribution.

In Nahuatl (Uto-Aztecan), Shasta (Shastan) and Witsuwiten (aka Babine, Athabaskan), plural marking opposition is present on pronouns, demonstratives and nouns; within the latter, the number opposition is restricted to animate nouns; Assiniboine (Siouan) and Northern Slavey (Athabaskan) show a further restriction, with plural marking present on all lexical categories but restricted to human nouns.

Table 7.7.: Number systems in North American languages

	Family	Language	Internal Distribution
Type A1	Algonquian	Ojibwe (Nishnaabemwin)	N, P, D
	Athapaskan	Northern Slavey	N(H), P, D
	Athapaskan	Witsuwiten	N(A), P, D
	Aztecan	Nahuatl	N(A), P, D
	Chimakuan	Quileute	N, P, D
	Costanoan	Mutsun	N, P, D
	Mayan	Chontal	N, P, D
	Mayan	Huastec	N, P, D
	Mixe-Zoque	Sierra Populuca	N, P, D
	Salishan	Comox	N, P, D
	Shastan	Shasta	N(A), P, D
	Siouan	Assiniboine	N, P, D
	Tepiman	Nevome	N, P, D
	Tillamook	Tillamook	N, P, D
	Zapotecan	San Miguel	N, P, D
Type A2	Aztecan	Pipil	N(A), P
	Haida	Haida	N(H), P
	Mayan	Sipacapense	N(A), P
	Misupalman	Miskitu	N, P
	Muskogean	Chochtaw	N, P(1,2)
	Rama	Rama	N, P
	Talamanca	Teribe	N(H), P
	Totonacan	Tepehua	N, P
	Zapotecan	Zoogocho Zapotec	N, P
Type A3	Mixtecan	Chalcatongo Mixtec	P(1)
	Wakashan	Kwakiutl	N(H)
	Yuman	Hualapai	N

Subtype A2

The second subset (A2) all the languages that exclude one lexical category from the plural marking distinction.

The results shown in Table 7.7) reveal that the only lexical category left out from the plural marking distinction are demonstrative pronouns: there are no cases, among these languages taken into account, of plural distinction spread on nouns and demonstratives, leaving the pronouns out of the number marking domain, or plural opposition among pronouns and demonstratives excluding nouns.

Nonetheless, internal further distinction in number opposition are present in this subgroup as well; as seen for subgroup A1, the restrictions appear mostly at the full nouns level. Of the ten languages included in subgroup A2, five appear to have a further restriction in plural marking.

In Teribe (Chibchan) and Haida (Haida), plural opposition lacks in demonstratives and non-human nouns; Sipacapense (Mayan) and Pipil (Uto-Atzecan) exclude inanimate nouns (beside demonstratives) from any pluralization strategy.

The remaining language displays a gap in plural opposition as well, but in pronouns: Jamul Tiipay (Hokan) and Choctaw (Muskogean) both exhibit plural marking constructions on nouns and pronoun, but restricted to first and second person: this can be explained by observing the third person pronominal form in Jamul Tiipay, *peya* (Miller 2001: p. 80), that corresponds to the demonstrative pronoun: demonstratives in Jamul Tiipay are insensitive to number distinction and so the derived pronominal form.

Subtype A3

The third subgroup, A3, is the less common and includes languages where the plural opposition is even more restricted. Such cases are displayed by Kwakiutl language (Wakashan), that has a plural marking distinction on a small set of human nouns only, by the means of suppletive forms; Hualapai language, where

Table 7.8.: Number system Type B, North American languages

Family	Language	Singular	Plural
Zuni	Zuni	N	N, P
Algonquian	Blackfoot	N, D	N, D

pronouns do not distinguish plural forms and plural meaning is disambiguates by verbal affixes. Nouns in Hualapai (Cochimi, Yuman), from kin to inanimate, indicate singular/plural distinction through various marking strategies.

7.4.2. Type B: Singular (overt)/Plural

The second type, B, includes the languages that have overt expression in both singular and plural. Two languages do exhibit this structure, Blackfoot and Zuni. In Blackfoot (Algonquinan), singular and plural marking are symmetric: since there are no free pronouns available, singular and plural markers are carried by nouns and demonstratives only. The following table reports the demonstrative pronoun in Blackfoot, with both singular and plural marking. The markers are the same as nouns and they distinguish also animacy, see Table 7.8

The markers are *-a* for the singular animate, *-yi* for singular inanimate. The animacy distinction is also reflected in plural forms, with the markers *-ksi* and *-isti* respectively. Such affinity and parallelism in both construction types (suffixes) and about the structural information they carry (both are distinct for animacy), as well as their distribution seems to suggest a common development of the singular and plural markers.

Zuni language (Isolate) number distinctions are spread through different strategies and distributions across the nominal domain. Demonstratives are excluded, since no data is available. Nouns and pronouns show different behavior with respect to number marking. Plural number is expressed on both nouns and pro-

Table 7.9.: Number system Type C, North American languages

Family	Language	Dual	Plural
Algonquian	Chumash	P	N(H), P, D
Aruak	Damana	P	N, P, D
Athapaskan	Navajo	P(1)	N, P
Keresan	Acoma	P	N, P, D
Numic	Paiute	P(1)	N(H), P
Numic	Kawaiisu	P(1)	N(H), P, D
Otomian	Ocuilteco	P	N, P, D
Wappo	Wappo	N(H), P(3)	N, P
Washo	Washo	P, D	N, P, D
Yokuts	Wikchamni	P, D	N, P, D

nouns: pronominal plural forms modify their stems to express plurality; nouns do add a marker, a suffix, which is not cumulative with any other grammatical feature. Singular number is overtly marked in Zuni nouns: singular suffixes are cumulative with the noun classes existing in Zuni. Such asymmetry does not support a joint 'vertical' development of the singular vs plural distinction in nouns and pronouns in Zuni: it rather indicates that nouns and pronouns have developed the respective number systems separately.

7.4.3. Type C: Singular/Dual/Plural

Ten of the North American languages included in the sample belong to type C, which groups together the languages with a singular (unmarked), dual and plural number system. Table 7.9 reports languages and distributions:

Plural marking can be either present on all lexical categories, or it may exclude one lexical category: this happens in Navajo, Wappo and Northern Paiute, and the lexical category always excluded from plural marking are again demonstrative forms. Northern Paiute, Kawaiisu and Ineseño Chumash show plural distinction restricted to human nouns.

All languages belonging to this type have dual opposition on pronouns only: dual distinction is attested and widespread on all pronominal forms, although four languages reveal additional restriction in dual marking: in Northern Paiute, Kawaiisu and Navajo, dual marking is constrained to first person pronoun; Wappo language distinguishes dual on third person pronoun only. The marking strategy employed by Wappo language to express dual number is connected to the plural form, since it consists in the third person plural marker attached to a different root stem: *tse'pi o'koti tse'koti*, 'He/She They two They' (Thompson et al. 2006: p. 135)

Wikchamni (Yokutian) and Washo (Isolate) are the only languages of this group to show dual marking on demonstrative forms as well, beside pronouns. The two languages are structurally similar on this regard: both have distinct forms for demonstratives and both languages show symmetric marking in pronouns and demonstratives, in dual and plural. Moreover, dual and plural markers are the same for both pronouns and demonstratives. Nouns do behave differently and independently in both languages, since they select different markers unknown to pronouns and demonstrative pronouns. Such structure suggests a common development in the number marking system for pronouns and demonstratives in these two languages, with nouns following other 'marking paths'.

7.4.4. Type D: Singular (overt)/Dual/Plural

The last type groups together all the languages that express overtly singular dual and plural. Four languages of the North American sample follow this structure: Aleut and Yupik, Zuni, Tuscarora and Tunica. The distribution of each number opposition within the lexical categories is illustrated in Table 7.10

Aleut and Yupik, related genetically (Eskimo-Aleut), have an similar nominal number system structure (and the markers across the two languages are clearly related). Singular, dual, plural markers are cumulative with case and they cover all lexical categories and all noun types. For these similarities in construction forms (all cumulative suffixes) and their regularity across the noun types, it is unlikely to argue a independent development of the number distinctions in

Table 7.10.: Number system Type D, North American languages

Family	Language	Singular	Dual	Plural
Aleut	Aleut	N, P, D	N, P, D	N, P, D
Eskimo	Yupik	N, P, D	N, P, D	N, P, D
Iroquoian	Tuscarora	N	N	N
Tunica	Tunica	N	N, P, D	N, P, D

these varieties.

The other three languages with a overt singular vs dual vs plural number system have been mentioned in the subsection related to the interaction between gender and number: they are Zuni, Tuscarora (Iroquoian) and Tunica (see Section 7.2)

Tuscarora lacks free pronominal and demonstrative forms: this implies that the nominal number markers identified are reported on full nouns level exclusively.

Singular number markers in Tuscarora are cumulative with gender; such feature is not present on dual and plural noun markers, both used on human nouns only. The affinity between dual and plural markers is also reflected by the construction forms: dual number is expressed by the prefix *neye-*, while plural employs the prefix *kaye-* (Williams 1976: p. 136). It is therefore more likely that the singular number in Tuscarora has developed independently, and probably with different times from the dual and plural markers, strongly connected and possible output of a common development process.

Tunica (Isolate) number system shows the presence of overt cumulative markers on singular, dual and plural (Haas 1946: p. 46). Such prefixes do correspond to the pronominal forms in Tunica, and for this reason, their distribution is regular on all noun types for each number value. Their construction forms are also related and a common development may be suggested.

7.5. Number systems in South American languages: types and distribution

The South American language sample consists of 38 languages. Within this area, four number systems have been identified: three of them coincide with the types described for the languages of North America in Section 7.4. These types are:

A. Singular/Plural (27 languages)
B. Singular (overt)/Plural (3 languages)
C. Singular/Dual/Plural (7 languages)

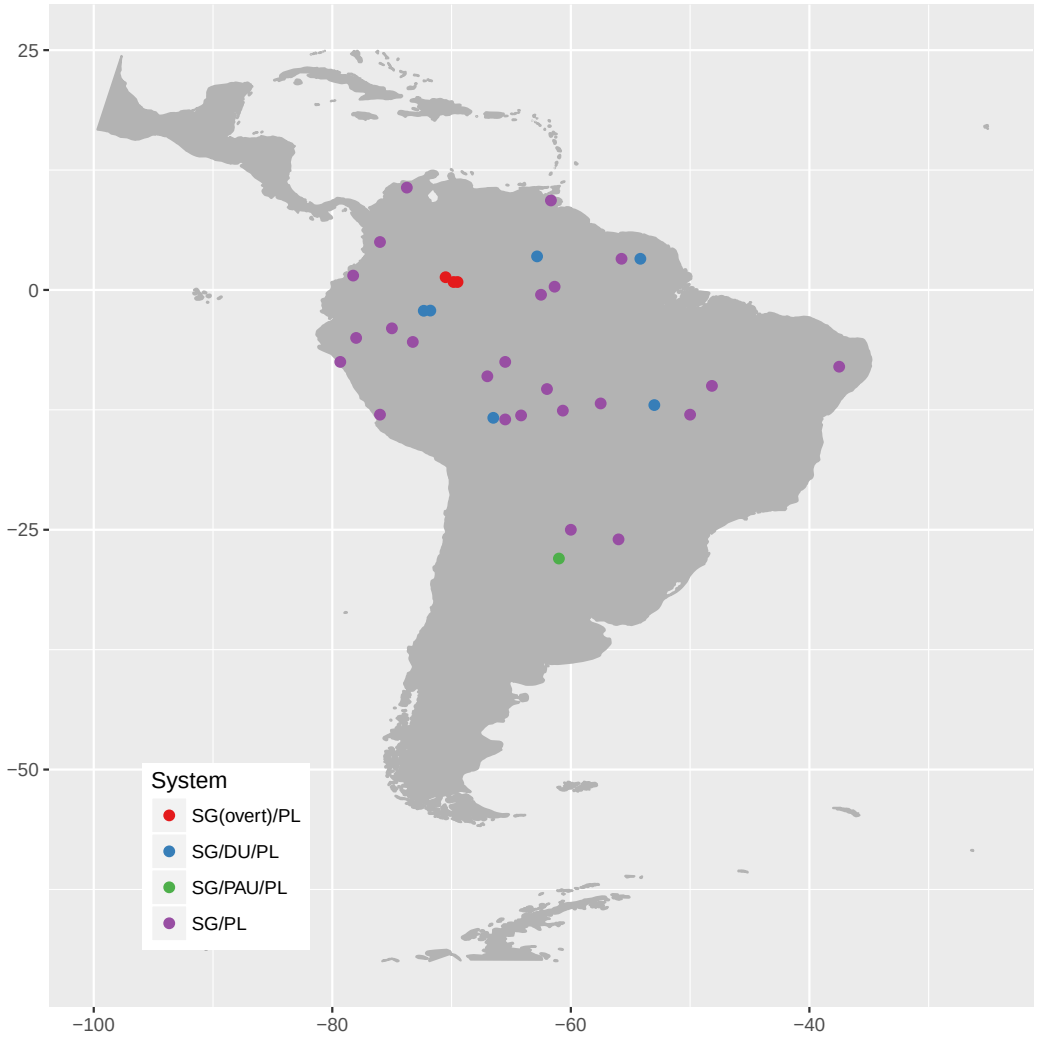
The nominal number system D, with a tripartite structure singular/ dual / plural and the singular value expressed overtly is not present in the South American languages included in the sample.

Furthermore, a single South American language in the sample (Mocovì) exhibits an additional number marking system, which involves a paucal value. This establishes a new marginal type E:

E. Singular/Paucal/Plural (1 language)

The geographical distribution of the South Americal languages in the sample and their respective number system is shown in Fig. 7.2. The distribution of the languages within each nominal number system type is quite similar to the one observed in the North American area: most languages do follow type A, and the presence of overt singular markers is attested on a relatively small amount of languages. A single South American language exhibits a paucal number value. Dual marking in the nominal domain is less widespread in the South American area compared to the North Americas.

Figure 7.2.: Number systems in South American languages



7.5.1. Type A: Singular/Plural

The 'default' number system, that is unmarked Singular VS Plural, is also the most common in both North American and South American languages. 26 languages do belong to this type and, as already described in the previous subsection, the languages within this grouping can be aggregated in further sets based on the different presence and distribution of the number values distinctions, as shown in Table 7.11

Subtype A1

Fifteen languages out of 26 exhibit a singular vs plural opposition on all lexical categories and do therefore belong to subtype A1; such percentage is very close to the one defined among the languages from North America.

Nonetheless, as illustrated in table above, South American languages display less internal variation in the distribution of the plural distinction across the noun types: of these fifteen languages, only one, Guaranì (Tupian) shows an internal restriction: in Guaranì, plural markers are present on all lexical categories and types with the exception of inanimate nouns, which are insensitive to plural marking.

Subtype A2

The second subset, A2, collects all the languages that exclude a whole lexical category from plural marking opposition: nine South American languages follow such sub-type:

The lexical category excluded does again correspond, in all cases, to the demonstrative pronouns. Kwaza (Isolate) and Jaqaru (Aymaran) languages also lack plural distinctions in pronouns: Kwaza third person pronoun does not distinguish plural marking; in Jaqaru, plural is restricted to nouns and first person pronoun (for a more detailed overview about the noun markers in Kwaza see Section 6.3.5).

Table 7.11.: Number system Type A, South American languages

	Family	Language	Plural
Type A1	Arauan	Bare	N, P, D
	Cariban	Tiriyo	N, P, D
	Purus	Apurina	N, P, D
	Mayan	Fulnio	N, P, D
	Mixe-Zoque	Bororo	N, P, D
	Warao	Warao	N, P, D
	Urarina	Urarina	N, P, D
	Panoan	Matsés	N, P, D
	Panana	Baure	N, P, D
	South Guaicuruan	Pilaga	N, P, D
	Choco	Embera	N, P, D
	Chimuan	Mochica	N, P, D
	Tupian	Guarani	N(A), P, D
	Rikbaksta	Rikbaksta	N, P, D
Type A2	Kwaza	Kwaza	N, P(1, 2)
	Quechuan	Quechua	N, P
	Northern Arawakan	Maipure	N, P
	Tupi	Karo	N, P
	Arauan	Jarawara	N, P
	Barbacoan	Awa Pit	N, P
	Aruak	Ika	N, P
	Gê	Xerente	N, P
	Aymaran	Jaqaru	N, P(1)
Type A3	Cariban	Atroari	N(A), P
	Jivaroan	Aguaruna	N

Table 7.12.: Number system Type B, South American languages

Family	Language	Singular	Plural
Tucanoan	Wanano	N(A), P(3)	N, P
Tucanoan	Desano	N(H)	N, P, D
Tucanoan	Cubeo	N(H)	N, P, D

Subtype A3

Two languages belong to the subgroup A3, with even further restriction in plural marking. Aguaruna (Jivaroan) shows plural distinction on pronouns only; Waimiri Atroari (Cariban) distinguishes plural on first person pronoun only.

7.5.2. Type B: Singular (overt)/Plural

Three Tucanoan languages from the South American sample do exhibit this number system (see Table 7.12): Wanano, Desano and Cubeo. All these three languages and the related number markers have already been described in Section 7.2 because of their interesting interaction with the gender feature.

Cubeo and Desano are genetically strongly affiliated and this relation can be noticed also in both the construction forms employed to express number and the distribution of the number values in their respective number systems. Both languages have plural distinctions on nouns, pronouns and demonstratives, while the singular overt marker is restricted to human nouns only, suffixes cumulative with gender expression. My knowledge of the Tucanoan language family from the historical perspective is too little to make any claims or even hypothesis about the possible development of their number markers.

Table 7.13.: Number system Type C, South American languages

Family	Language	Dual	Plural
Trumai	Trumai	N(A), P, D	N(A), P, D
Yanoman	Shiriana	P	N, P, D
Tacanan	Cavinena	N, P	N, P
Boran	Bora	N(A), P, D	N, P, D
Huitoto	Witoto	P	N, P
Cariban	Wayana	P	N, P, D

7.5.3. Type C: Singular/Dual/Plural

The nominal number system C includes six of the Southern American languages of the sample, which represent the only languages that show a dual marking opposition in at least one nominal type, seeTable 7.13

Plural marking distribution within this type confirms the tendencies: it is usually widespread on all lexical categories (3 out of 6 languages), with two languages excluding demonstratives (2 out of 6, Cavineña and Witoto) and one language (Trumai) only with plural marking on all lexical categories with a further restriction on nouns (inanimate are excluded).

Dual marking offers more variation: it can be found on all lexical categories like in Bora, but restricted to human nouns and referents only, since the dual marker is cumulative with gender expression; or in Trumai, where dual marking, as plural, is restricted to animate nouns.

In Cavineña (Paco-Tacanan) dual marking is found on nouns and pronouns; dual marking becomes more specific in Huitoto (Huitotoan) and Shiriana (Yanomam), where it covers pronoun only, and especially in Wayana (Cariban), where it is restricted to first person pronoun.

The interactions and the relation between dual and plural number are clearer and more visible when the dual marking is less restricted: languages that show a richer presence of dual marking are Trumai, Cavineña and Bora.

Trumai language (Isolate, Brazil) conveys a very regular number system: two lexical items, one with dual meaning (*/a/*), the other with plural meaning (*/wan/*) cross cut the whole nominal domain, with the same identical distribution (restriction to animate referent applies to both number values). The regularity in both construction types and distribution may suggest a common and joint origin of these values and therefore of the whole number system.

In Bora (Boran), the context is similar: dual markers are the same on all lexical categories, as plural suffixes: they are both restricted to animate and their distribution is fully symmetric.

Finally, Cavineña offers an analog structure, but with some differences: dual and plural are clearly connected by both the construction type angle (two clitic particles) with the same distribution on nouns and pronouns. The only exception might be, but it is yet difficult to make a claim or building hypothesises, third person plural forms, which are similar but not as regular as the marker that does appear on nouns and first and second pronominal pronouns; the formal difference between the plural marker *=ekwana* and third person plural marker *-na* can be either massive or irrelevant. Third person pronouns in Cavineña have deictic distinctions (proximal / distal) making them function as demonstrative pronouns: a further look to this direction might help in a fully understanding of the position and development of these forms within the nominal number system of Cavineña.

7.5.4. Type E: Singular/Paucal/Plural

Mocovì language (Guaicuruan) is the only member of this nominal number system. Plural distinction is present on all nouns, first and second pronouns and on demonstrative pronoun (which may function as third person form, since it lacks in this language). For all lexical categories, plural is expressed through different strategies (specific suffixes for animate nouns and collective markers for inanimates, suppletion and probably affixes on pronouns, a specific suffix for demonstrative/third pronoun plural), fact that lead to independent developments of the plural markers for each lexical category.

On the other hand, paucal number is found on nouns only, both animate and inanimate. Many different markers are used, which follow, among other aspects, phonological constraints. The only connection that might be hypothesised between paucal and plural values is the presence of a specific paucal marker, *-ri*, which might be connected to the plural affixes found in first and second person pronouns (with the latter that shows transparently the presence of such affix, sg. *qami* pl. *qamiri*, leading to a former plural status of this marker that has been later reanalyzed as paucal.

7.6. Number systems in Eurasiatic languages: types and distribution

This section deals with the nominal number systems of the 50 languages from Europe and Asia of the sample.

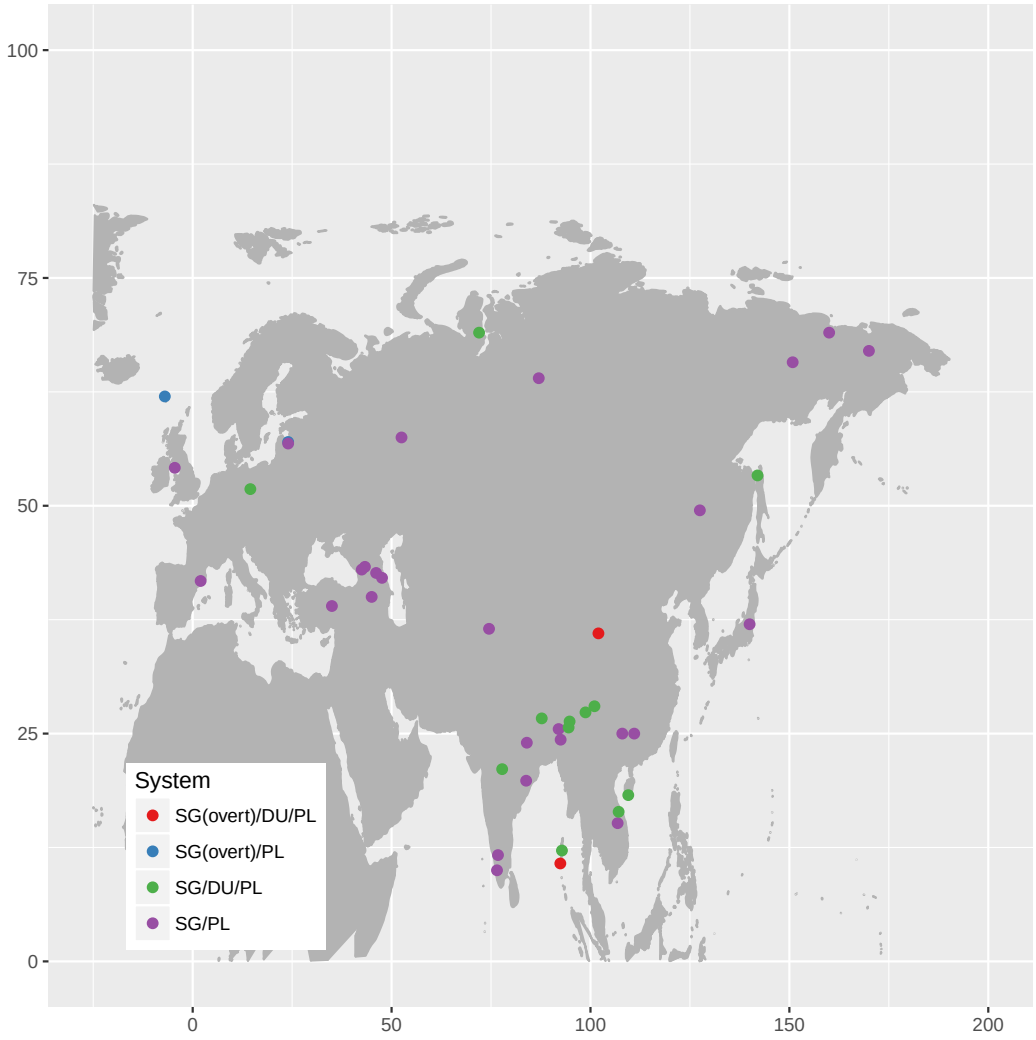
The first step in data analysis is, as seen and done in the languages of other areas, a preliminary grouping of the languages based on the nominal system they show. All 49 Eurasiatic languages of the sample can be aggregated in four main nominal number marking systems:

A. Singular/Plural (35 languages)
B. Singular (overt)/Plural (2 languages)
C. Singular/Dual/Plural (14 language)
D. Singular (overt)/Dual/Plural (2 languages)

The nominal number systems just defined are the same identified for the North American languages; the relative distribution of these languages within these types are also similar between the two areas. The geographical distribution is shown in Fig. 7.3.

Nominal system type A, with the default singular vs plural opposition, is again confirmed as the most predominant number system in the Eurasiatic macro-area as well. The second most common number system is Type C, which differs from Type A by the addition of dual number. Type B and Type D, that

Figure 7.3.: Number systems in Eurasiatic languages



have both the main feature of overt expression of singular value, are the less widespread: only four languages, among the ones of the sample, show the presence of overt singular number (and two also include dual expression).

7.6.1. Type A: Singular/Plural

Table 7.14 reports all the 28 Eurasiatic languages of the sample and their further split in subtypes.

Overt plural behaviour: subtype A1

Of these 28 languages, 21 belong to the sub-type A1, that gathers together the languages with plural distinction on all lexical categories (nouns, pronouns, demonstratives). It can be observed that internal distributional constraints are quite rare within this subgroup: it seems that there in a tendency, among these languages with a singular/plural system, of not leaving 'aside' noun categories and thus in avoiding restrictions in number oppositions on specific lexical categories. Plural number can be expressed through many different strategies in the same language, but in this system type, all nominal elements tend to be covered by a plural marker. Udmurt (Uralic, Winkler 2001) posits itself as an exception, with plural marking present on all lexical categories but with a distributional restriction on nouns (inanimates are excluded from plural marking).

Overt plural behaviour: subtype A2

The aforementioned tendency is also confirmed by the scarse number of languages grouped in the sub-set A2: Khasi (Austroasiatic, Roberts 1891), Tundra Yukaghir (Yukaghir, Maslova 2003), Kokborok (Sino-Tibetan, Karapurkar 1976) and Sapuan (Austroasiatic, Jacq & Sidwell 1999).

The former three languages also reflect the tendency of excluding demonstrative pronouns in case of restrictions of general number marking to lexical categories. Sapuan excludes nouns: it is a rare case when plural marking is

Table 7.14.: Number system Type A, Eurasiatic languages

	Family	Language	Plural
Type A1	Tungusic	Manchu	N, P, D
	Turkic	Turkish	N, P, D
	Finnic	Livonian	N, P, D
	Permic	Udmurt	N, P, D
	Kartvelian	Svan	N, P, D
	Yeniseian	Ket	N, P, D
	Lak-Dargwa	Icari	N, P, D
	Northwest Caucasian	Kabardian	N, P, D
	Dravidian	Badaga	N, P, D
	Dravidian	Malay	N, P, D
	Dravidian	Pengo	N, P, D
	Yukaghir	Kolyma Yukaghir	N, P, D
	Tai-Kadai	Zoulei	N, P, D
	Burushaski	Burushaski	N, P, D
	Japanese	Japanese	N, P, D
	Northern Chukotko-Kamchatkan	Chukchi	N, P, D
	Avar-Andic-Tsezic	Godoberi	N, P, D
	Armenian	Armenian	N, P, D
	Romance	Catalan	N, P, D
	Iranian	Persian	N, P, D
	Indic	Darai	N, P, D
Type A2	Bahnaric	Sapuan	P, D
	Yukaghir	Tundra Yukaghir	N, P
	Celtic	Manx	N, P
	Bodo-Garo	Kokborok	N, P
	Khasian	Khasi	P, D
Type A3	Kam-Tai	Mulao	P
	Hmong-Mien	Iu Mien	P

Table 7.15.: Number system Type B, Eurasiatic languages

Family	Language	Singular	Plural
Germanic	Faroese	N, P(3)	N, P
Balto-Slavic	Latvian	N, P, D	N, P, D

restricted to pronouns and demonstratives. However, it should be pointed out that third person pronoun in Sapuan is the demonstrative form itself.

Overt plural behaviour: subtype A3

Finally, two languages show restricted marking on one category only: in Iu Mien (Hmong-Mien, Court 1985)) and Mulao (Tai-Kadai, Wang & Guoqiao 1993)) only pronouns select plural markers.

7.6.2. Type B: Singular (overt)/Plural

Two languages from the Eurasiatic sample do belong to this type: Faroese and Latvian. Both languages exhibit a rich inflectional system, which does reflect on the nominal number marking system.

In Latvian, singular and plural markers have the same distribution, form and internal gender distinction (masculine / feminine) on the whole lexical category, first and second person pronoun excluded, that have suppletive constructions. Such symmetry, granted by the noun classes paradigms, leads to claim that dual and plural markers in Latvian have originated and spread jointly, see Table 7.15.

7.6.3. Type C: Singular/Dual/Plural

Thirteen languages of the Eurasiatic sample are aggregated in type C number system: they do therefore distinguish dual and plural overtly on at least one noun type. The following table illustrates both the languages and the distribution of the number values within their number system, Table 7.16

Table 7.16.: Number system Type C, Eurasiatic languages

Family	Language	Dual	Plural
Kuki-Chin	Angami	N, P, D	N, P, D
Malayo-Sumbalam	Hainan Cham	P	N, P, D
Munda	Korku	N(A), P, D	N(A), P, D
Samoyedic	Nenets	N, P	N, P
Nivkh	Nivkh	P(1)	N, P, D
Nahali	Nahali	N, P, D	N, P, D
Katuic	Pacoh	P	P
Andamanese	Andamanese	P, D	N(A), P, D
Slavic	Upper Sorbian	N, P, D	N, P, D
Nungish	Anong	P	N, P
Naga	Chang	P	N, P, D
Qiangic	Prinmi	N, P	N, P, D
Dhimalic	Dhimal	P	N, P

The same observations that have been made for plural marking on type A are also valid and confirmed by type B: plural marking is spread on all lexical categories in most languages (eight languages out of thirteen), with some internal restrictions: in Korku (Austroasiatic) and Andamanese (Great Andamanese), plural marking is limited to animate nouns.

Four languages do exclude demonstratives from plural marking: Dhimal, Anong, Hainan Cham (Sino-Tibetan) and Nenets (Uralic). No other restrictions have been identified.

Dual marking behaviour in Eurasiatic languages is more similar to the one observed in South American ones rather than in the languages of North America, since it shows bigger variety. In four languages, Angami (Sino-Tibetan), Korku, Nahali (Indo-European, Indo-Arian) and Upper Sorbian (Indo-European, Sorbian), dual marking is present on all lexical categories (Korku does not distinguish dual in inanimate nouns, as seen in plural marking as well); six languages expresses dual number on pronouns only: Nivkh (Isolate, Kamchatka;

in Nivkh dual marking is restricted to first person pronoun), Pacoh and Chang (Austroasiatic), Anong, Hainan Cham and Dhimal (Sino-Tibetan).

Three languages show dual marking attested on two lexical categories: Prinmi (Sino-Tibetan) and Nenets do exclude demonstrative forms, Andamanese excludes nouns.

Some final remarks can be made on the interactions between dual and plural in these languages: in sampled languages from Europe and Asia, there are no cases where dual is less restricted than plural in its distribution: this implies that when dual marking is equally present to all noun categories, plural will behave equally or it will show broader distribution. For a further discussion, let's look at the languages which number values are in symmetrical distribution on the nominal types.

Six languages do not present any split in number values distribution: these are Angami, Korku, Nenets, Nahali, Pacoh and Upper Sorbian. Angami, Korku, Pacoh and Nenets beside showing the same distributional constraints on their respective number values, have a formal regularity: each language employs two construction forms, one for dual and one for plural, sistematically used on all the lexical categories involved in their respective nominal number system (Angami employs clitics on nouns, pronouns and demonstratives, Korku suffixes on noun animates, pronouns and demonstratives, Pacoh select affixes on pronouns only, Nenets employs suffixes cumulative with case). In all these languages, both the distribution and the characteristics of the markers seems to be connected to a common and joint development of the number values and, ultimately, of the nominal number systems of these varieties.

In Nahali, the number marking paradigm is symmetric, but the construction forms used on each noun category are partially misaligned: dual is regular on nouns, third person pronoun and demonstratives: this can be explained with the fact that third person pronoun and the demonstrative pronoun are formally identical: since first and second person pronoun take a specific and shared dual marker and also their root form shows some resemblance, (while third person pronoun looks completely unrelated), it is likely that the third person pronoun

Table 7.17.: Number system Type D, Eurasiatic languages

Family	Language	Singular	Dual	Plural
Mongolic	Mangghuer	N	P	N, P, D
South Andamanese	Onge	N	N	N, P

is derived from the demonstrative pronoun. The dual marker means literally 'he two' *it-tel* and it is fully used on nouns. Plural marking on nouns is expressed by a suffix *-ta* which is not used on pronouns. The third person/demonstrative form marks plurality though the form *etla*, which seems to be connected to *la* second person plural form, meaning therefore *he.you.ALL*, 'they' (Kuiper 1962).

These facts lead to a strongly intertwined number system, with strong connections between the dual and plural values, as reflected by third person pronoun. However, it seems that the whole synchronic number systems as we see it today is rather the development of many processes than an unique, joint process.

7.6.4. Type D: Singular (overt)/Dual/Plural

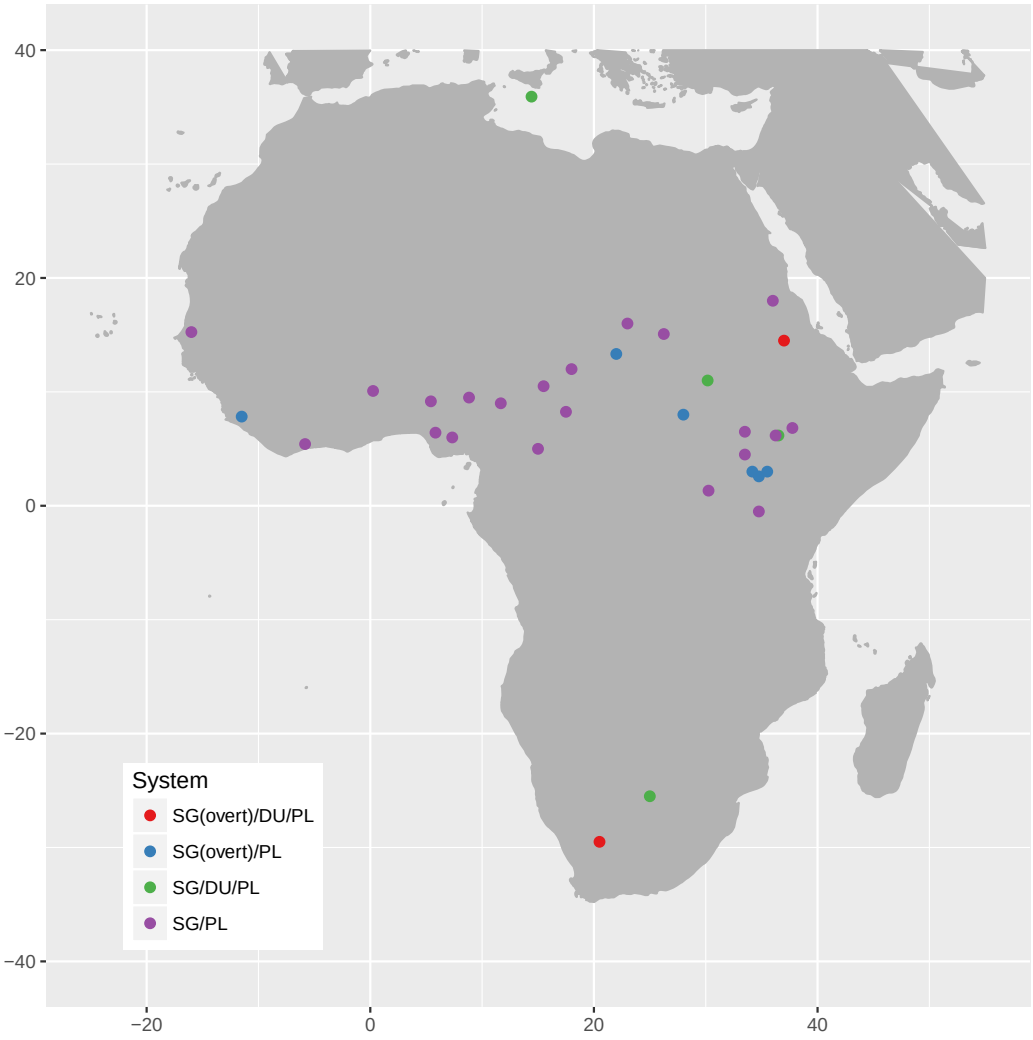
Two languages located in Asia, Mannghuer and Onge, display a nominal number system with overt singular, dual and plural, see Table 7.17. Mangghuer language has an overt singular marker carried by nouns, dual restricted to pronouns and plural marking on all lexical categories. The plural marker is the same suffix on all nominal types. The dual marker is a complex structure used only on pronouns. The singular marker is a lexical item, meaning originally 'one', non cumulative with other features.

This context clearly posits a number system with three separate developments of the number values involved, and most likely taken place at different times.

Onge language has a particularly interesting nominal number marking system. Plural is marked on pronouns and nouns, with singular and dual number

expressed on nouns only. The lack of overt dual markers in pronouns and their presence on nouns is quite rare. The number markers on nouns are quite similar in their structure (no cumulation with other features is involved): the singular suffix correspond to *-da* and allomorphs, the dual is *-dena* and the plural is *-di*. Pronouns are suppletive in plural and do not select these markers. It seems that the development of the nominal number marking system in Onge has been joint and simultaneous in nouns.

Figure 7.4.: Number systems in African languages



7.7. Number systems in African languages: types and distribution

The last two subsections are dedicated to African and languages from the Pacific respectively.

The languages that are part of the backbone of the African sample are around 44; ten additional languages, all from the Bantu / Niger-Congo family, are excluded from the analysis, albeit kept as control languages. The African languages of the sample can be assigned one of the four most common types so far

A. Singular/Plural (22 languages)
B. Singular (overt)/Plural (6 languages)
C. Singular/Dual/Plural (4 language)
D. Singular (overt)/Dual/Plural (2 languages)

The most common number system is, again type A; African languages tend to express overt singular more frequently than the other languages. Dual is less represented, when its presence is taken and compared between the macro-areas. The geographical distribution of number systems in Africa is shown in Fig. 7.4

7.7.1. Type A: Singular/Plural

The table7.18 summarizes the languages that belong to this type and their internal variation (if any).

African languages that follow number system Type A can be further grouped in the sub-types already described for all the other areas.

Subtype A1

20 languages out of 31 exhibit plural distinctions on all lexical categories. Internal restrictions do not seem to apply within this subgroup.

Table 7.18.: Number system Type A, African languages

	Family	Language	Plural
Type A1	Adamawa	Mumuye	N, P, D
	Surmic	Didinga	N, P, D
	Lendu	Ngiti	N(H), P, D
	Chadic	Mupun	N, P, D
	Omotic	Wolaytta	N, P, D
	Omotic	Dime	N, P, D
	Chadic	Hausa	N, P, D
	Nilotic	Luo	N, P, D
	Saharan	Beria	N, P, D
	Nupoid	Nupe	N, P, D
	Kru	Grebo	N, P, D
	Beja	Beja	N, P, D
	Nilotic	Dinka	N, P, D
	Kru	Dajdriwale	N, P, D
	Bongo-Bagirmi	Bagirmi	N, P, D
	Saharan	Kanuri	N, P, D
Type A2	Kwa	Anufo	P, D
	Nubian	Midob	N, P
	Northern Atlantic	Wolof	P, D
	Masa	Masa	N, P
	Blu-Mandara	Wandala	N, P
	Igboid	Igbo	N, P
Type A3	Bongo-Bagirmi	Mbay	N
	Gbeya	Gbeya	P
	Edoid	Bini	P

Table 7.19.: Number system Type B, African languages

Family	Language	Singular	Plural
Mande	Mende	N	N, P, D
Maban	Masalit	N	N, P, D
Kuliak	So	N	N, P, D
Nilotic	Luwo	N	N, P, D
Nilotic	Nandi	N	N, P, D
Nilotic	Turkana	N(H)	N, P, D
Nilotic	Karamojong	N(H)	N, P, D

Subtype A2

Seven languages exclude one lexical category in plural opposition: the lexical category excluded are demonstrative pronouns, as shown in six cases out of seven. The only language which distinguishes plural marking on pronouns and demonstratives excluding nouns is Murle (Surmic).

Subtype A3

The three languages left do belong to subset A3, that involves languages with plural marking on one lexical category only. Gbeya and Bini (both Atlantic-Congo) select plural opposition on pronouns only; in Mbay (Central Sudanic) plural marking is constrained to nouns only, since independent personal pronouns are not present in this language.

7.7.2. Type B: Singular (overt)/Plural

Seven languages express overtly singular meaning in opposition to plural marking. The languages that follow this structure are Ngiti, Mende (Mande), Luwo, Nandi, Turkana and Karimojong (all Nilotic). The internal distribution of the number values within the lexical categories and in the noun types is almost identical among all the languages of the type: plural is present on all noun cat-

egories. Singular value is present, for all of these languages, on nouns only; Turkana and Karamojong operate a further restriction, with singular marking limited to human referents, markers which are also cumulative with gender, as seen in Section 7.2.

Turkana and Karamojong number systems have already been partially described: their regular structure within the lexical categories and the nominal types, the same internal distinctions applied by all the number values that are motivated ultimately by the formal traits of the construction form (that is, the cumulation with gender), lead to a joint development of the number values and ultimately, of the number system, Table 7.19.

The other four languages do not express additional cumulative features in singular; Nandi language provides a good example. In Nandi, plural marking is expressed on all lexical categories: the same marker is shared by nouns and pronouns, demonstratives express plurality through suppletive forms or modification of the stem root. Singular marking is overtly expressed in nouns by the means of a suffix form. Such marker is not found on demonstratives and pronouns. This asymmetry indicates that the number system has developed following independent processes and therefore its origin is not amenable to a single, common phenomenon.

7.7.3. Type C: Singular/Dual/Plural

Type C includes the languages that express overtly dual and plural, with singular value left unmarked. Three languages of the African sample follow this classificationTable 7.20: Ju'Hoan (Kxa), Moro (Heiban), and Dizin (Dizoid). Such languages show a similar distribution of the number values on the noun types as well: plural is expressed on all lexical categories in Ju Hoan and Dizin, while Moro demonstratives are unmarked for number. Dual is restricted to pronouns, with Ju'Hoan and Dizin again grouped together with dual expression restricted to pronominal forms and Moro with a further distributional constraint on first person pronoun.

Table 7.20.: Number system Type C, African languages

Family	Language	Dual	Plural
Khoe	Ju Hoan	P	N, P, D
Heiban	Moro	P(1)	N, P
Dizin	Dizin	P	N, P, D
Semitic	Maltese	N	P, D

Table 7.21.: Number system Type D, African languages

Family	Language	Singular	Dual	Plural
Khoisan	Korana	N(H), P, D(H)	N(H), P, D(H)	N(H), P, D(H)
Kunama	Kunama	N(H), D	P, D	N, P, D

7.7.4. Type D: Singular (overt)/Dual/Plural

The two languages of the African sample that exhibit a tripartite number system with singular dual and plural expressed overtly are Korana (Khoe-Kwadi) and Kunama (Kunama). Korana has already been described in the gender interaction section (see Section 7.2): Korana number markers are cumulative with gender, for all number values, and they occur with the same distribution (restricted to human referents) on all number categories. The development of this system is clearly simultaneous and amenable to a single derivational processTable 7.21.

The nominal number system of Kunama is structurally different. First, the distribution of the number values is not homogeneous. Plural marking appears on all lexical categories, but with different construction forms. The suffix *-e*, that is the plural marker in Kunama, is not used on pronouns and demonstratives, which, in turn, share the same plural marking strategy. Dual number distinctions appear on pronouns and demonstratives, with forms again clearly connected. Singular expression is attested on nouns only, and the suffix *-a* is used. This suffix is not cumulative with other grammatical features, i.e. gender. The partial asymmetry of this number system suggests a joint development

Table 7.22.: Kunama (Nilo-Saharan, Africa) Bender (1996)

	singular	dual	plural
N	-a		-e
1	aba	kiime^{incl} aame^{excl}	kime^{incl} ame^{excl}
2	ena	eeme	eme
3	unu	iime	ime
DEM	‘nnà	'nname	'nnae

^{incl} inclusive 1 pronoun
^{excl} exclusive 1 pronoun

of dual and plural markers on pronouns and demonstratives, with the singular and plural marking in nouns that has conversely originated independently. The number system of Kunama is shown in Table 7.22

7.8. Pacific languages and nominal number systems: types and distribution

This final section deals with the number systems identified among the languages of the Pacific area, both Australian and Papuan. The languages are about 40 and they are aggregated into the following nominal number systems types:

A. Singular/Plural (10 languages)
C. Singular/Dual/Plural (22 language)
D. Singular (overt)/Dual/Plural (3 languages)
F. Singular/Dual//Trial/Plural (4 languages)

The nominal number system type B, that is the overt singular vs overt plural values, has not been found in the Australian - Pacific sample. In addition, the presence of trial number has been attested in four languages, delivering a quadripartite nominal number systems with singular (unmarked), dual, trial, plural overt marking. The geographical distribution is shown in Fig. 7.5.

Figure 7.5.: Number systems in Pacific languages

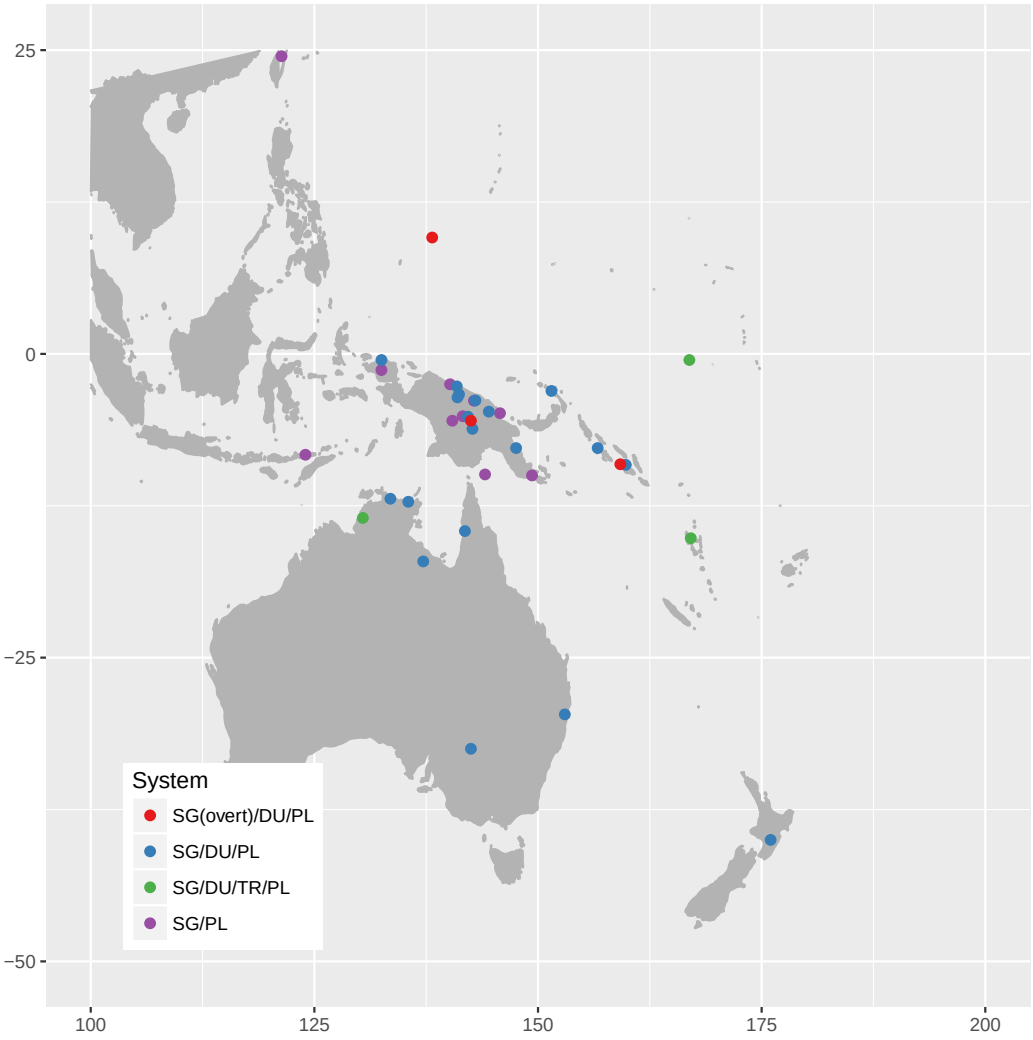


Table 7.23.: Number system Type A, Pacific languages

	Family	Language	Plural
Type A1	Nimboran	Nimboran	N, P, D
	Dagan	Kambera	N, P, D
	Mugil	Bargam	N(H), P, D
	Nuclear-Trans-New Guinea	Mian	N, P, D
Type A2	Austronesian	Atayal	P, D
	Yukaghir	Tundra Daga	N, P
	Torricelli	Kamasau	N, P
	Eastern Trans Fly	Meriam	N, P
	Nuclear-Trans-New Guinea	Wambon	N, P
Type A3	Maybrat	Maybrat	P
	Timor-Alor-Pantar	Adang	P

7.8.1. Type A: Singular/Plural

Languages showing a singular VS plural distinction are relatively underrepresented when compared to other large areas like North American or African languages. Eleven languages do take plural markers leaving the singular unmarked. The languages and their internal distributional constraints are illustrated in Table 7.23:

Overt plural behavior: subtype A1

Five languages do belong to subtype A, expressing plural number opposition on all lexical categories: nouns, pronouns and demonstratives. Mugil (Nuclear Trans New Guinea, Madang), Wambon (Nuclear Trans New Guinea, Awyu-Dumut) and Mian (Awyu-Ok) do limit the respective plural marking strategies to human nouns only.

Overt plural behavior: subtype A2

Daga (Dagan), Kamasau (Nuclear Torricelli), Meryam (Eastern Trans-Fly) and Atayal (Austronesian) can be grouped in the second subtype: plural distinctions are used on nouns and pronouns only by Daga, Meryam and Atayal. Kamasau language excludes nouns rather than demonstratives from plural marking.

Overt plural behavior: subtype A3

Adang (Timor-Alor-Pantar) and Kambera (Austronesian) do distinguish plural number on pronominal forms only.

7.8.2. Type C: Singular/Dual/Plural

This nominal number system type, that involves the presence of overt markers for both dual and plural values, leaving the singular unmarked, is the most frequent among the languages of the Pacific sample. This type includes 22 languages out of 41. Table 7.24 illustrates the languages that belong to type C and the internal distribution of the respective number values.

Plural marking distribution

The distribution of plural marking in this system reflects what observed in both the other nominal systems types and in other macroareas: plural marking tends to cover more lexical categories than singular or dual value. Plural is widespread on all lexical categories in about 11 of the 21 languages grouped within this type; of these 21 languages, five shows restrictions within the related internal distribution. Such restriction always applies to nouns: in Bilua, Maori (Austronesian), Djambarrpuyngu (Pama-Nyungan, Yuulngu) and Maung (Iwaidjan), plural marking is restricted to human nouns; Wakaya (Pama-Nyungan, Ngarna) and Garrwa (Garrwan) exclude inanimate nouns from plural distinctions.

In a similar way, the first lexical category which lacks plural marking corresponds to demonstrative pronouns: seven languages show this feature, and

Table 7.24.: Number system Type C, Pacific languages

Family	Language	Dual	Plural
Border	Imonda	N(K)	N(K), P
Bosavi	Edolo	P, D	N, P, D
Meax	Meyah	N, P	N, P
Kuot	Kuot	N(A), P, D	N, P, D
Ramu	Rao	P	P
Ndu	Ambulas	P, D	N, P, D
Savosavo	Savosavo	N, P	N, P
Sko	Skou	N, P	N, P
Nuclear-Trans-New Guinea	Oksapmin	P	N(K), P
Papuan	Menggwa Dla	P	P
Binanderean	Suena	P	N(K), P
Abun	Abun	P	N, P
Bilua	Bilua	N(H), P, D	N(H), P, D
Malayo-Polynesian	Maori	P	N(H), P, D
Paman	Thayorre	P	P
Yarli-Baagandi	Paakanti	N, P, D	N, P, D
Yuulngu	Djambarrpuyngu	P	N(H), P, D(H)
Pama-Nyungan	Yorta-Yorta	N(A), P	N(A), P, D
Garrwan	Garrwa	N(A), P, D(A)	N(A), P, D(A)
Iwaidjan	Maung	P(3)	N(A), P, D
Pama-Nyungan	Wakaya	N(A), P, D	N(H), P, D

within this subset, three languages reveal a further constraint in plural opposition, with plural markers restricted to pronouns and kin terms in Imonda (Border), Suena (Nuclear Trans New Guinea, Greater Binanderean) and Oksapmin.

Moreover, plural marking may be restricted to one lexical category only: this happens, for instance, in Rao (Lower Sepik-Ramu) and Mengglwa Dla (Senagi), where plural marking is restricted to pronouns only.

Dual marking distribution

Dual marking is expressed on all the lexical categories in five languages: Kuot (Isolate), Wakaya and Garrwa (with restriction to animate nouns), Bilua and Duna (limited on human nouns) and Paakanty. Paakanty (Pama-Nyungan, Yarli-Baagandji) is the only language of the type with dual marking attested on all nominal elements (the other Pacific languages that exhibit such behaviour are Lavukaleve and Yapese, both of them belonging to the nominal system type B).

Five languages exhibit dual markers on a limited set of lexical categories; internal variation in dual marking reveals a trend which is slightly more miscellaneous to the one established by the various types and macro-areas seen so far: Imonda language has dual marking limited to kin terms only; conversely, in Edolo (Bosavi) pronouns and demonstratives select dual markers and in Maung only third person pronoun has a dual form. The general tendency, however, confirms the presence of dual marking opposition on pronouns: examples are provided by Skou (Sko), Rao, Suena and Thayorre (Pama-Nyungan) among others.

Dual and Plural interaction

The interaction between dual and plural marking within the languages follows the traditional tendencies and generalizations about the presence of the number values: dual marking tends to be more limited than plural marking and in general it does not cover nominal types or lexical category which are insensitive to plural markers.

It has been seen how certain shared features between the number values within a language, like the same distribution or similar and someway related construction types may provide further information about the structure and development of the number marking system.

Languages from this area have particularly striking features on this aspect, although not in the direction expected: all the languages grouped in this type show little or no interaction between the number values. lexical categories may share the same constructions, especially nouns and demonstrative pronouns, but this symmetry is not confirmed by a similar distribution in dual marking that would suggest a relation, or a common origin, or 'starting point' between the values: it seems that each lexical category, or even nominal type, has developed independently its own number system, by selecting the most 'appropriate' number markers, rather than adapting to a set of markers, ready to use and with differentiated number meanings incorporated.

A regular and symmetric distribution of the number values can even be misleading. If one examines in detail and compares the constructions used in the five languages with the most regular and harmonic distribution of the number values within the lexical categories, would notice that an homogeneous distribution is not the result of a regular, unique number marking development strategy, but a *collage* of different markers and sharing of strategies that ultimately covers all the nominal elements involved.

Languages like Bilua have 'regular' number systems, from the number values distribution perspective, as shown in Table 7.24. However, this regularity is not counterbalanced construction-wise, as shown in the detailed number system overview in

7.8.3. Type D: Singular (overt)/Dual/Plural

Two Papuan languages show a tripartite nominal number system with singular overtly expressed: Lavukaleve (Isolate) and Yapese (Austronesian, Oceanic). The Table 7.26 reports the detailed internal distribution of each number value.

Although the internal distribution of the number values within the lexical

Table 7.25.: Bilua (Central Solomons, Papua New Guinea) Obata (2003)

	singular	dual	plural
N	-	kidi	RDP
1	anga	eqe^{incl} aniqe^{excl}	aninge^{incl} anime^{excl}
2	ngo	qe	me
3	unu	iime	ime
DEM	nei	nioqi	ni

^{incl} inclusive 1 pronoun
^{excl} exclusive 1 pronoun

Table 7.26.: Languages with gender/number markers

Family	Language	Singular	Dual	Plural
Duna	Duna	N(H)	N(H), P, D	N(A), P, D
Lavukaleve	Lavukaleve	N(H)	N, P, D	N, P, D
Oceanic	Yapese	N	N, P, D	N, P, D

categories and the noun types seems quite regular, the construction types used do confirm the tendency of the Pacific Type C languages of a development of the number systems as a result of different processes internal to the lexical categories if not the single noun types, rather than a unique, joint process of number marking assignment.

The independent processes operated at the lexical category level are clear in both languages: nouns in Yapese mark singular, dual and plural by the means of lexical items, while pronouns have affixes and demonstratives mixing strategies, using lexical items from nouns and specific markers.

Lavukaleve is the only language that uses number markers (specifically, singular) cumulative with gender; such feature is not carried by neither dual or plural noun markers and different strategies can be identified in pronouns and

Table 7.27.: Noun number suffixes in Yapese, Jensen (1977: p. 154)

	SG	DU	PL
Nouns	rea	gəl	pi
1	gaeg	gamow	gamaed
2	guur	gimeew	gimeed
3	qiir	yow	yaed
DEM	chaqneey	gəl chaqneey	tineey

Table 7.28.: Number system Type F, Pacific languages

Family	Language	Dual	Trial	Plural
Micronesian	Nauru	P, D	P, D	P, D
Enindhilyagwa	Enindhilyagwa	P	P	P
Wadjiginy	Wadjiginy	P	P	P
Oceanic	Sakao	P	P	N(K) P, D

demonstrative forms.

7.8.4. Type F: Singular/Dual/Trial/Plural

Four languages have also trial number, beside dual and plural: Nauru and Sakao (Austronesian, Oceanic), Wadjiginy (Wadjiginy) and Enindhilyagwa (Enindhilyagwa). The following table sums up the respective internal distribution of the number values, Table 7.28:

In Enindhilyagwa and Wadjiginy, all number values distinctions are constrained to pronouns only. In Sakao, plural number is marked also on demonstratives and kin terms through a shared marker, while pronouns select a full and rich number marking system. In Nauru, nouns do not select number markers: pronouns and demonstratives are marked for dual, trial and plural, with the

same construction forms. Finally, Wadjiginy trial markers express also gender.

The main difference between the languages within this type and other rich number system encountered among the languages of the Pacific, relies on the fact that in type F number systems it is possible to hypothesize a joint development of the number values oppositions. Number distinction on this type are highly specific, rich and detailed, with the number markers therefore easily identifiable, as illustrated in Table 4.4

8. Concluding remarks

8.1. Introduction

In this chapter I summarize the main results of the present survey and I assess their contribution to a typology of the number systems, viewed in its interactions between nominal elements and number markers. The limitations and weaknesses of this analysis are also pointed out, along with suggestions for future research on this phenomenon.

8.2. Overview of the main results

The approach in data exploration pursued in this work has been built by segmenting the number marking domain on three layers, or levels of progressive interaction, with the aim to investigate the following research foci: (i) the exploration of the construction types used to mark number, (ii) the internal distribution of the number markers within the nominal types and (iii) a typology of the number systems attested in the languages of the world and the interaction between the number values in these systems. These three main research nuclei were approached individually in chapters X, Y and I respectively.

The specific research questions pertaining to each of the chapters are illustrated for convenience, with a summary of the results and the answers achieved.

On the exploration of the construction types used to mark number values

– Which construction types are used to express the distinct number

values?

- What is the distribution of the construction types within the number values?
- Are there any recurrent characteristics in the construction types used to signal a specific number value?
- Which number construction types are used on the nominal elements?
- Are there construction types recurrently associated to a NP type?
- Do the construction types associated with a nominal type show relevant features?

On the internal distribution of the plural markers within the nominal types

- Languages may signal plurality on full nouns through a shared marker or by using different markers assigned to each NP type. What is the distribution of these contexts?
- Which full nouns tend to signal plurality through the same marker?
- What is the distribution of suppletive and morphological strategies in pronouns plural expression?
- Are there languages with pronoun specific plural markers? Which internal grouping can be identified?
- Are there attested cases of plural markers shared by nouns and pronouns? What is the internal distribution of such constructions?

On a typology of the number systems and the interaction within the number values

- Which number systems can be identified in the languages of the sample?
- What is the distribution of the number systems on each macro-area?
- Which groupings, or subtypes within the number systems, can be identified?

- Are there any relevant features that can be associated to a number system?
- Which are the main interactions between the number values in the systems identified?

8.2.1. Construction types and number values

In chapter 3, the interaction between the three basic parameters has been explored. The chapter was divided in three parts, one for each of the number values explored (singular, dual and plural).

For each number value, I examined the construction types used by the different nominal elements.

Singular number does not show much variety in the construction types used. Pronouns, in most cases, are suppletive. Nouns are generally left unmarked. Controlling for these variables outputs a more heterogeneous picture, albeit limited: pronouns may be found left unmarked. Overt marking in pronouns is usually assigned to a suffix form; tone strategy and infixes have been found, but in a such reduced quantity that suggest rather a language specific number expression than a consolidated morphological mean of singular expression. Clitics and lexical items ('singular word') have not been found in pronouns. Demonstratives pronouns use a more restricted set of overt markers to express singular, that includes stem modification and suppletion and suffixes. Overt marking on nouns is signaled by different strategies, although suffixes are the most common: suppletion and stem modification are attested, as well as prefixes, lexical items and clitic particles. Suppletive form are restricted on human nouns and kin terms, confirming the high specificity of this construction type. Inanimate nouns have rarely been attested with other construction types beside suffixes; when this happens, it is linked to the fact that the 'unusual' construction type is used by all full nouns, and inanimates make no exception.

Dual marking in pronouns is characterized by the presence of composite strategies, where multiple markers are used cumulatively on the noun types. This phenomenon is attested especially in pronouns, where nonetheless also

'plain' marking strategies are found. Data survey confirms the link between non-singular forms: dual and plural show, in most cases, the same person-/number stem. Instances where dual is connected to the singular root are less common and the related marking strategies can be typologically rare (like reduplication in third person dual pronoun). Dual marking in nouns is expressed by more regular constructions, especially suffixes and lexical items, in most cases derived by the respective numeral form in the language. The most variety in dual construction types is found in human and kin terms. The lowest elements on the hierarchy indicate dual number only through a marker that occurs on all noun types. The construction types found on demonstratives include suppletion and suffixes.

Plural is the number category where the greatest varieties of construction types are found. First and second person pronoun signal plurality through composite strategies, usually a non-singular stem with a suffix. Other constructions types, like prefixes, plain suffixes and plural words are attested. Nouns attach different plural markers: stem modifications and suppletion seems to be restricted to kin and human nouns. These and other morphological strategies like circumfixes are found on human nouns only: in most cases, they seem to be connected to deverbal processes.

Among the relevant features that can be identified within the construction type / number value interaction there is cumulation. The grammatical feature that is more likely to be expressed cumulatively with number is gender, and the number value is singular. The presence of cumulative markers has a tendency to be restricted to human referents.

8.2.2. Internal distribution of the plural markers within the nominal types

In chapter 4, the internal distribution of the plural markers within the nominal types have been described. The chapter was split in four main sections. In the first section, the full nouns of each language of the sample have been taken into account and their plural markers explored. Data survey led to a

preliminary classification of the attested distributions of plural markers on full nouns. In about one third of the sampled languages, a single marker is used to express plural on full nouns, with no animacy distinction. Three subgrouping involve the presence of a unique plural marking strategy restricted to specific noun types, while the other show no plural distinction: among these subgroups, the most attested include languages with an unique marker restricted on animates and languages with plural distinctions on humans only. Less languages, concentrated mostly in the Pacific area, show plural distinction on kin only. Languages with different markers on full nouns have been found: they constitute subsets, where some nominal types show a specific plural construction in opposition to others. The groupings identified are: (i) human nouns vs non human; (ii) animate vs inanimate, (iii) kin vs human, and other subsets involving further internal distributions. Pronouns are highly suppletive. In a quantity of cases, they may show pronominal-specific plural markers, that in most languages cover the whole pronominal paradigm. Some are further groupings are nonetheless found, with pronominal specific markers occurring on restricted types: the marker is usually shared by first and second person or restricted to third person only. Further groupings occur, but they are much rarer (one or two cases attested), as well as cases of multiple pronominal specific plural markers. The most common context involves the presence of a shared marker between nouns and pronouns: it is more widespread than suppletion and markers pronominal specific. A marker can cover the whole NP set; in most cases, is third person that share plural markers with nouns. First person pronoun alone never clusters with nouns: it is always grouped with another pronominal person, either second or third. Conversely, second person pronoun has been found individually grouped with nouns. Demonstratives are in a quantity of occurrences, indifferent to plural distinctions. However, when they signal plurality, they tend to use the same marker as nouns. When a construction is shared by both nouns and pronouns, they also tend to cluster as well. Another context, well widespread, sees the presence of demonstrative specific markers, usually suffixes. The less common grouping involves pronouns and demonstra-

tives and they tend to cluster together when plural in nouns is not distinguished or either in contexts where demonstrative pronouns and third person pronouns are formally related.

8.2.3. Typology of number systems

In chapter 5 I described the number systems attested in the languages of the sample and I provided a classification. The chapter was split on five parts, each of them devoted to a specific macro area. The singular vs plural (with unmarked singular) nominal number system is the most attested on all the macro-areas, Pacific excluded, where the unmarked singular vs dual vs plural one is more common. Number systems with unmarked singular, dual and plural immediately follows in all macro-areas, with strong presence in Eurasia and North America; it is also widespread, to a lesser extent, in South America and Africa. Two number systems with overt singular expression have been encountered: the overt singular vs plural, attested, although less common than the systems previously aforementioned, in all areas (especially Africa) except South America. South America area exhibit a singular vs paucal vs plural system, not found elsewhere. Other areal-specific number systems include the presence of trial, attested among the languages of the Pacific.

The noun type set that is more likely to be dropped in number distinctions are demonstrative pronouns and, progressively, nouns. Dual number shows a tendency, for all languages on all number systems and subsystems, to cover the same or the less noun types of plural. This implied that dual number never occurs, as a number opposition, on nominal elements which do not distinguish plural marking. Singular marking, as shown also by results in chapter 3, is used on nouns.

Interaction with gender has been further explored in this chapter, with a focus on how it distributes on number values: although gender/number markers are especially common with to express singular in human nouns, they are may cover the whole number system in a language (like in Korana) and to a lesser extent, in plural only or restricted to dual.

8.3. Prospects for future research

The descriptive model built for this analysis has been a useful tool to perform analysis on such complex data. This model, implemented in a database, has the main benefit to be constantly implemented and expanded, by adding features and performing aggregations not necessarily connected to number marking but that could also unveil further interactions with other grammatical features.

A typological survey reveals tendencies which need specific genealogical and areal studies to be confirmed and motivated.

The most interesting aspect involves the diachronic perspective: the distribution and the information available about the number constructions has revealed a link between the original meaning of the source marker and the distributional constraints attested in synchrony. This applies also to the number systems and the number values distinctions, where the internal processes, rather than other factors, seem to have determined the presence and the distribution of the number distinctions. These tendencies need to be confirmed and further explored.

A. List of languages

A.1. The main language sample

Africa

- Lele (East Chadic)
- Mwaghavul (West Chadic)
- Wolaytta (Omotic)
- Dime (Omotic)
- Arbore (Eastern Cushitic)
- Iraqw (Southern Cushitic)
- Dizi (Omotic)
- Beja (Beja)
- Dahalo (Southern Cushitic)
- Masa (Masa)
- Hausa (West Chadic)
- Malgwa/Wandala (Biu-Mandara)
- Hdi (Biu-Mandara)
- Tamashek (Berber)
- Mina (NA)
- Arabic (Egyptian) (Semitic)
- Hoan (NA)
- Korana (Central Khoisan)
- Sandawe (Sandawe)

- Anufo (Kwa)
- Mumuye (Adamawa-Ubangian)
- Diola-Fogny (Northern Atlantic)
- Mende (Western Mande)
- Wolof (Northern Atlantic)
- Venda (Bantoid)
- Zulu (Bantoid)
- Dagbani (Gur)
- Ogbronuagum (Cross River)
- Gbeya (NA)
- Bini (Edoid)
- Yoruba (Defoid)
- Tima (NA)
- Nupe (Nupoid)
- Fyem (Platoid)
- Orig (NA)
- Orig (Kordofanian)
- Moro (Kordofanian)
- Herero (Bantoid)
- Kinyamwezi (Bantoid)
- Koromfe (Gur)
- Igbo (Igboid)
- Dadjriwale (Kru)
- Jamsay (NA)
- Midob (Nubian)
- Mbay (Bongo-Bagirmi)
- Didinga (Surmic)
- Fur (Fur)

- Ngiti (Lendu)
- Luo (Nilotic)
- Kuku (Nilotic)
- Pokot (Nilotic)
- Masalit (Maban)
- So (Kuliak)
- Murle (Surmic)
- Beria (NA)
- Turkana (Nilotic)
- Karimojong (Nilotic)
- Bagirmi (Bongo-Bagirmi)
- Luwo (Nilotic)
- Kunama (Kunama)
- Koyra Chiini (Songhay)
- Lango (Nilotic)
- Ma'di (Moru-Ma'di)

Eurasia

- Maltese (Semitic)
- Ainu (Ainu)
- Mangghuer (Mongolic)
- Manchu (Tungusic)
- Turkish (Turkic)
- Korku (Munda)
- Sapuan (Bahnaric)
- Pacoh (Katuic)
- Khasi (Khasian)
- Vietnamese (Viet-Muong)
- Tsat (Sundic)

- Basque (Basque)
- Burushaski (Burushaski)
- Chukchi (Northern Chukotko-Kamchatkan)
- Badaga (Southern Dravidian)
- Malayalam (Southern Dravidian)
- Pengo (South-Central Dravidian)
- Tamil (Southern Dravidian)
- Great Andamanese (Great Andamanese)
- Mien (Hmong-Mien)
- Armenian (Eastern) (Armenian)
- Faroese (Germanic)
- Catalan (Romance)
- Catalan (NA)
- Tajik (Iranian)
- Sorbian (Upper) (Slavic)
- Latvian (Baltic)
- Manx (Celtic)
- Darai (Indic)
- Japanese (Japanese)
- Svan (Kartvelian)
- Korean (Korean)
- Dargwa (Lak-Dargwa)
- Godoberi (Avar-Andic-Tsezic)
- Lezgian (Lezgic)
- Nivkh (Nivkh)
- Kabardian (Northwest Caucasian)
- Anong (Nungish)
- Chang (Baric)

- Dhimal (Bodic)
- Kokborok (Baric)
- Lepcha (Lepcha)
- Lahu (Burmese-Lolo)
- Athpare (Bodic)
- Onge (South Andamanese)
- Mulao (Kam-Tai)
- Nenets (Samoyedic)
- Nenets (NA)
- Liv (Finnic)
- Udmurt (Finnic)
- Ket (Yeniseian)
- Yukaghir (Kolyma) (Yukaghir)
- Yukaghir (Tundra) (NA)

North America

- Aleut (Eskimo-Aleut)
- Yupik (Eskimo-Aleut)
- Blackfoot (Algonquian)
- Yurok (Yurok)
- Teribe (Talamanca)
- Rama (Rama)
- Quileute (Chimakuan)
- Shasta (Shasta)
- Hualapai (Yuman)
- Tiipay (Jamul) (Yuman)
- Pomo (Eastern) (Pomoan)
- Tuscarora (Northern Iroquoian)
- Karok (Karok)

- Acoma (Keresan)
- Kiowa (Kiowa-Tanoan)
- Sipakapense (NA)
- Chontal Maya (Mayan)
- Miskitu (Misumalpan)
- Chimalapa Zoque (Mixe-Zoque)
- Popoluca (Mixe-Zoque)
- Choctaw (Muskogean)
- Koasati (Muskogean)
- Kwakiutl (Northern Wakashan)
- Witsuwit'en (NA)
- Zapotec (Zapotecan)
- Chalcatongo Mixtec (Mixtecan)
- Ocuilteco (Otomian)
- Mutsun (Costanoan)
- Wikchamni (Yokuts)
- Wikchamni (NA)
- Klamath (Klamath-Modoc)
- Bella Coola (Bella Coola)
- Tillamook (Tillamook)
- Assiniboine (Nakhota) (NA)
- Lakhota (Siouan)
- Tepehua (Huehuetla) (Tepiman)
- Tunica (Tunica)
- Nevome (Tepiman)
- Kawaiisu (Numic)
- Pipil (Aztecan)
- Nahuatl (Central) (Aztecan)

- Northern Paiute (Numic)
- Chemehuevi (Numic)
- Wappo (Wappo)
- Washo (Washo)
- Zuni (Zuni)

Pacific

- Wambaya (West Barkly)
- Gooniyandi (Bunuban)
- Wardaman (Yangmanic)
- Muna (Sulawesi)
- Tuvaluan (Oceanic)
- Tinrin (Oceanic)
- Rapanui (Oceanic)

South America

- Damana (Aruak)
- Jarawara (Arauan)
- Mapudungun (Araucanian)
- Baré (Arawakan)
- Apurina (Arawakan)
- Baure (Arawakan)
- Jaqaru (Aymaran)
- Aymara (NA)
- Awa Pit (Barbacoan)
- Tiriyo (Cariban)
- Waimiri Atroari (NA)
- Wayana (Cariban)
- Cayuvava (Cayuvava)

- Wari' (Chapacura-Wanhan)
- Ika (NA)
- Mochica (Chim˙an)
- Embera (Choco)
- Mocovi (Guaicuruan)
- Pilaga (Guaicuruan)
- Bora (Huitoto)
- Aguaruna (Jivaroan)
- Kwaza (Kwaza)
- Fulnio (YatÍ)
- Bororo (NA)
- Akwe-Xerentexer (Ge-Kaingang)
- Rikbaksta (Rikbaktsa)
- Mosetén (Mosetenan)
- Hup (VaupÈs-Japur·)
- Matses (Panoan)
- Cavinena (Tacanan)
- Trumai (Trumai)
- Guanano (Tucanoan)
- Desano (Tucanoan)
- Cubeo (Tucanoan)
- Karo (Ramarama)
- Guarani (Tupi-GuaranÌ)
- Urarina (Urarina)
- Warao (Warao)
- Shiriana (Yanomam)
- Ayoreo (Zamucoan)

A.2. The database sample

Africa

- Anufo (Kwa)
- Bagirmi (Bongo-Bagirmi)
- Didinga (Surmic)
- Dime (Omotic)
- Dizin (Omotic)
- Ju-Hoan (Northern Khoisan)
- Kuku (Nilotic)
- Luo (Nilotic)
- Masalit (Maban)
- Mbay (Bongo-Bagirmi)
- Mende (Sierra Leone) (Western Mande)
- Midob (Nubian)
- Mumuye (Adamawa-Ubangian)
- Mwaghavul (West Chadic)
- Nandi (Bantoid)
- Ngiti (Lendu)
- Nupe-Nupe-Tako (Nupoid)
- Wolaytta (Omotic)

Eurasia

- Angami Naga (Kuki-Chin-Naga)
- Badaga (Southern Dravidian)
- Burushaski (Burushaski)
- Catalan (Romance)
- Catalan (NA)
- Chang Naga (Baric)

- Darai (Indic)
- Dhimal (Bodic)
- Godoberi (Avar-Andic-Tsezic)
- Icari Dargwa (Lak-Dargwa)
- Icari Dargwa (NA)
- Iu Mien (Hmong-Mien)
- Japanese (Japanese)
- Ket (Yeniseian)
- Kok Borok (Baric)
- Korku (Munda)
- Malayalam (Southern Dravidian)
- Manchu (Tungusic)
- Manx (Celtic)
- Nivkh (Nivkh)
- Northern Pumi (Qiangic)
- Northern Yukaghir (NA)
- Nung (Myanmar) (Nungish)
- Önge (South Andamanese)
- Pacoh (Katuic)
- Pengo (South-Central Dravidian)
- Persian (Iranian)
- Sapuan (Bahnaric)
- Southern Yukaghir (Yukaghir)
- Svan (Kartvelian)
- Tsat (Sundic)
- Tu (Mongolic)
- Turkish (Turkic)
- Udmurt (Finnic)

North America

- Atzingo Matlatzinca (Otomian)
- Babine (NA)
- Chimalapa Zoque (Mixe-Zoque)
- Choctaw (Muskogean)
- Comox (Central Salish)
- Havasupai-Walapai-Yavapai (Yuman)
- Highland Popoluca (Mixe-Zoque)
- Huehuetla Tepehua (Tepiman)
- Ineseño (Chumash)
- Kumiai (NA)
- Kumiai (Yuman)
- Mískito (Misumalpan)
- Navajo (Athapaskan)
- Northern Paiute (Numic)
- Pima Bajo (Tepiman)
- Pipil (Aztecan)
- Rama (Rama)
- San Miguel El Grande Mixtec (Mixtecan)
- Shasta (Shasta)
- Siksika (Algonquian)
- Sipacapense (NA)
- Southern Haida (Haida)
- Southern Ohlone (Costanoan)
- Tabasco Chontal (Mayan)
- Tillamook (Tillamook)
- Wappo (Wappo)
- Washo (Washo)

- Western Keres (Keresan)
- Yokuts (Yokuts)
- Yokuts (NA)
- Zoogocho Zapotec (Zapotecan)

Pacific

- Abun (North-Central Bird's Head)
- Ambulas (Middle Sepik)
- Bargam (Madang)
- Biak (South Halmahera - West New Guinea)
- Bilua (Solomons East Papuan)
- Daga (Dagan)
- Djambarrpuyngu (Pama-Nyungan)
- Duna (Duna-Bogaya)
- Edolo (Bosavi)
- Imonda (Border)
- Kamasau (NA)
- Kuot (Kuot)
- Lavukaleve (Solomons East Papuan)
- Madngele (Eastern Daly)
- Mai Brat (North-Central Bird's Head)
- Maung (Iwaidjan)
- Meriam (Western Fly)
- Nauru (Oceanic)
- Nimboran (Nimboran)
- Oksapmin (Oksapmin)
- Paakantyi (Pama-Nyungan)
- Rao (Annaberg)
- Sakao (Oceanic)

- Savosavo (Solomons East Papuan)
- Siwai (East Bougainville)
- Suena (Binanderean)
- Thayore (Pama-Nyungan)
- Wadjiginy (Anson Bay)
- Yapese (Yapese)
- Yorta Yorta (Pama-Nyungan)

South America

- Aguaruna (Jivaroan)
- Apurinã (Arawakan)
- Arhuaco (NA)
- Awa-Cuaiquer (Barbacoan)
- Baré (Arawakan)
- Baure (Arawakan)
- Bororo (NA)
- Cavineña (Tacanan)
- Cayubaba (Cayuvava)
- Fulniô (YatÍ)
- Huallaga Huánuco Quechua (NA)
- Jamamadí (Arauan)
- Jaqaru (Aymaran)
- Karo (Brazil) (Ramarama)
- Kotiria (Tucanoan)
- Kwaza (Kwaza)
- Malayo (Aruak)
- Matsés (Panoan)
- Mbyá Guaraní (Tupi-GuaranÌ)
- Mochica (Chim˙an)

- Mocoví (Guaicuruan)
- Murui Huitoto (Huitoto)
- Ninam (Yanomam)
- Northern Emberá (Choco)
- Pilagá (Guaicuruan)
- Trió (Cariban)
- Trumai (Trumai)
- Urarina (Urarina)
- Waimiri-Atroari (NA)
- Warao (Warao)
- Xavánte (Ge-Kaingang)
- Xerénte (Ge-Kaingang)

References

Abbi, Anvita. 2009. Is great andamanese genealogically and typologically distinct from onge and jarawa? *Language Sciences* 31(6).

Abbi, Anvita. 2013. *A grammar of the great andamanese language: an ethnolinguistic study* (Brill's Studies in South and Southwest Asian Languages). Brill.

Adam, Lucien. 1891. *Langue mosquito: grammaire, vocabulaire, textes*. París: Maisonneuve (Kraus Reprint 1968).

Adelaar, Willem F.H. 2004. *The languages of the andes. (with the collaboration of pieter c. muysken)*. Cambridge: CUP.

Aikhenvald, Alexandra Y. 1995. *Bare*. Vol. 100 (Languages of the World / Materials). München: Lincom Europa.

Alves, Mark John. 2000. *A pacoh analytic grammar*. University of Hawai'I at Manoa dissertation.

Anceaux, Johannes C. 1965. *The nimboran language: phonology and morphology*. The Hague: Martinus Nijhoff.

Applegate, Joseph. 1966. *Ineseño Chumash grammar*. University of California, Berkeley dissertation.

Bakker, Dik. 2007. Language sampling. In Jae Jung Song (ed.), *The oxford handbook of language typology*. Oxford: Oxford University Press.

Barrett, E.R. 1999. *A grammar of sipakapense maya*. University of Texas at Austin dissertation.

Bender, M. Lionel. 1996. *Kunama* (Languages of the World: Materials). Munich: Lincom Europa.

Bergsland, Knut. 1997. *Aleut grammar = unangam tunuganaan achixaasix*. Fairbanks: Alaska Native Language Center.

Berry, Keith. 1995. *A description of the abun language*. La Trobe University MA thesis.

Berry, Keith & Christine Berry. 1999. *A description of abun: a west papuan language of irian jaya*. Vol. 115 (Pacific Linguistics, Series B). Canberra: Australian National University.

Bickel, Balthasar. 2007. Typology in the 21st century: major current developments. *Linguistic Typology* 11. 239–251.

Bickel, Balthasar. 2010. Capturing particulars and universals in clause linkage: a multivariate analysis. In Isabelle Bril (ed.), *Clause-Hierarchy and Clause-Linking: The Syntax and Pragmatics Interface*. Amsterdam: Benjamins.

Bickel, Balthasar & Johanna Nichols. 2002. Autotypologizing databases and their use in fieldwork. In Peter Austin, Helen Dry & Peter Wittenburg (eds.), *Proceedings of the International LREC Workshop on Resources and Tools in Field Linguistics, Las Palmas, 26–27 May 2002*. Nijmegen: MPI for Psycholinguistics.

Bollée, Annegret. 2013. Reunion Creole structure dataset. In. *Atlas of Pidgin and Creole Language Structures Online*. Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology.

Breen, Gavan. 1974. *Wakaya Grammar*.

Broadwell, Georg Aaron. 2006. *A Choctaw reference grammar*. Lincoln: University of Nebraska Press.

Capell, Arthur & H. E. Hinch. 1970. *Maung grammar; texts and vocabulary*. The Hague: Mouton.

Colarusso, John. 1992. *A grammar of the kabardian language*. Calgary: University of Calgary Press.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. 2nd edn. Chicago: University of Chicago Press.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Court, Christopher. 1985. *Fundamentals of iu mien (yao) grammar*. University of California at Berkeley.

Creider, Chet A. & Jane Tapsubei Creider. 1989. *A grammar of nandi*. Hamburg: Helmut Buske Verlag.

Cristofaro, Sonia. 2012. Typological universals in synchrony and diachrony: the evolution of number marking. Paper presented at the 45th meeting of the Societas Linguistica Europea.

Croft, William. 1990. *Typology and universals*. Cambridge: Cambridge University Press.

Croft, William. 2003. *Typology and Universals*. Cambridge: Cambridge University Press [2nd edition].

Crowley, Terry. 2004. *Bislama reference grammar*. Vol. 31 (Oceanic linguistics special publication). Honolulu: Univ. of Hawai'i Press.

Cumberland, Linda A. 2005. *A grammar of assiniboine: a siouan language of the northern plains (montana, saskatchewan)*. Indiana University dissertation.

Curnow, Timothy Jowan. 1997. *A grammar of Awa Pit (Cuaiquer): an indigenous language of South-Western Colombia*. Australian National University dissertation.

Cysouw, Michael. 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press.

Cysouw, Michael. 2005. Quantitative methods in typology. In Gabriel Altmann, Reinhard Köhler & Rajmund G. Piotrowski (eds.), *Quantitative Linguistics: An International Handbook*. Berlin: Mouton de Gruyter.

Daniel, Michael. 2013. Plurality in independent personal pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology.

Daniel, Michael & Edith Moravcsik. 2005. The associative plural. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*. Oxford: Oxford University Press.

Danielsen, Swintha. 2007. *Baure: an arawak language of bolivia* (Indigenous Language of Latin America). PhD 2007 Radboud Universiteit Nijmegen. CNWS Publications. 480.

Dasgupta, D. & S.R. Sharma. 1982. *A hand book of onge language*. Calcutta: Anthropological Survey of India.

Davies, William D. 1986. *Choctaw Verb Agreement and Universal Grammar*. Dordrecht: Reidel.

Dhakal, Dubi Nanda. 2012. *Darai grammar* (Languages of the World/Materials). München: Lincom.

Diessel, Holger. 1999. *Demonstratives: form, function, and grammaticalization*. Vol. 42 (Typological Studies in Language). Amsterdam: John Benjamins.

Ding, Sizhi. 1998. *Fundamentals of prinmi (pumi): a tibeto-burman language of northwestern yunnan, china*. Australian National University.

Dixon, R. M. W. 2004. *The Jarawara language of southern Amazonia*. Oxford: Oxford University Press.

Dixon, R.M.W. & A.Y. Aikhenvald. 2003. Demonstratives: a cross-linguistic typology. *Studies in Language* 27(1).

Dryer, Matthew S. 1989a. Large linguistic areas and language sampling. *Studies in Language* 13. 257–292.

Dryer, Matthew S. 1989b. Plural words. *Linguistics* 27.

Dryer, Matthew S. 2013. Coding of nominal plurality. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology.

Dum-Tragut, Jasmine. 2009. *Armenian* (London Oriental and African Language Library 14). Amsterdam/Philadelphia: John Benjamins.

Edgar, John. 1989. *A masalit grammar*. Berlin: Dietrich Reimer.

Edmonson, Barbara Wedemeyer. 1988. *A descriptive grammar of huastec (potosino dialect)*. Tulane University dissertation.

Elson, Benjamín. 1960. *Gramatica Popoluca de la Sierra*. Xalapa, Mexico: Universidad Veracruzana.

Evans, Nicholas. 2007. Reciprocals, grammar-writing and typology. *Revue de linguistique roumaine II(4)*. 479–506.

Everaert, Martin, Simon Musgrave & Alexis Dimitriadis (eds.). 2009. *The Use of Databases in Cross-Linguistic Studies*. Berlin: Mouton de Gruyter.

Facundes, Sidney da Silva. 2000. *The language of the apurinã people of brazil (maipure/arawak)*. State University of New York, Buffalo dissertation.

Fedden, Sebastian. 2011. *A grammar of mian* (Mouton Grammar Library [MGL]). De Gruyter.

Fleck, David W. 2003. *A grammar of matses*. Rice University dissertation.

Forchheimer, Paul. 1953. *The category of person in language*. Berlin: W. de Gruyter.

Ford, Lysbeth. 1990a. *The phonology and morphology of bachamal (wogait)*. Australian National University MA thesis.

Ford, Lysbeth J. 1990b. *The phonology and morphology of bachamal (wogait)*. Canberra: Australian National University dissertation.

Frajzyngier, Zygmunt. 1993. *A grammar of mupun*. Berlin: Dietrich Reimer.

Frajzyngier, Zygmunt. 1997. Grammaticalization of number: from demonstratives to nominal and verbal plural. *Linguistic Typology* 1. 193–324.

Frantz, Donald. 1991. *Blackfoot grammar*. Toronto: University of Toronto Press.

Gabas, Jr. 1999. *A grammar of karo, tupi (brazil)*. University of California at Santa Barbara dissertation.

Gamble, Geoffrey. 1978. *Wikchamni grammar*. Vol. 89 (University of California Publications in Linguistics). Berkeley / Los Angeles: University of California Press.

Giridhar, Puttushetra Puttuswamy. 1980. *Angami grammar*. Manasagangotri, Mysore: Centeral Institute of Indian Languages.

Gomez, G.G. 1990. *The shiriana dialect of yanam (northern brazil)*. Columbia University.

Gossner, Jan David. 1994. *Aspects of edolo grammar*. University of Texas at Arlington MA thesis.

Gravelle, Gilles. 2004. *Meyah: an east bird's head language of papua, indonesia*. Zsfassung in niederländischer Sprache. Ph.

Greenberg, Joseph H. 1966. *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.

Grinevald, Colette G. 1988. *A Grammar of Rama (Report to National Science Foundation BNS 8511156)*. Tech. rep.

Gruzdeva, Ekaterina. 1998. *Nivkh*. Vol. 111 (Languages of the World: Materials). Munich: Lincom Europa.

Guillaume, Antoine. 2004. *A grammar of Cavineña, an Amazonian language of Northern Bolivia*. Victoria: Victoria University of Wellington dissertation.

Guirardello, Raquel. 1999. *A reference grammar of Trumai*. Houston: Rice University dissertation.

Haas, Mary. 1946. A grammatical sketch of Tunica. In C. Osgood (ed.), *Linguistic structures of native america*. New York: Viking Fund.

Hargus, Sharon. 2007. *Witsuwit'en grammar: phonetics, phonology, morphology* (First Nations Languages). UBC Press.

Harlow, Ray. 2007. *Maori: a linguistic introduction*. Cambridge: Cambridge University Press.

Harris, H. 1981. *A grammar of comox*. Ann Arbor: University Microfilms.

Haspelmath, Martin. 2005. Occurrence of nominal plurality. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*. Oxford: Oxford University Press.

Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie (eds.). 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Hepner, Mark. 2006. Bargam grammar sketch. Summer Institute of Linguistics.

van den Heuvel, Wilco. 2006. *Biak, description of an Austronesian language of Papua*. University of Amsterdam dissertation.

Hockings, P. & C. Pilot-Raichoor. 1992. *A badaga-english dictionary*. Berlin: Mouton de Gruyter.

Hovdhaugen, E. 2004. *Mochica* (Languages of the world: Materials). Lincom Europa.

Iljic, Robert. 2001. The origin of the suffix -men 們 in chinese. *Bulletin of the School of Oriental and African Studies, University of London* 64(1). 74–97.

Ingham, B. 2003. *Lakota* (Languages of the world: Materials). Lincom.

Jacq, P. & P. Sidwell. 1999. *Sapuan (sepuar)*. Vol. 302 (Languages of the World/-Materials). München: Lincom Europa.

Jensen, C. 1998. *Comparative tupi-guarani syntax*. Berlin: Mouton de Gruyter.

Jensen, John Thayer. 1977. *Yapese reference grammar*. Honolulu: University of Hawaii Press.

de Jesus Silva, Leia. 2011. *Morphosyntaxe du rikbaktsa*. Université Denis Diderot – Paris 7 dissertation.

Johnson-Weiner, Karen M. 2001. Review of amery, r, 2002 warrabarna kaurna. *OGMIOS: Newsletter of Foundation for Endangered Languages* 2(5). 24.

Karapurkar, Pushpa. 1976. *Kokborok grammar*. Vol. 3 (CIIL Grammar Series). Mysore: Central Institute of Indian Languages.

Klamer, Marian. 1998. *A grammar of kambera*. Vol. 18 (Mouton Grammar Library). Berlin: Mouton de Gruyter.

Knowles, Susan Marie. 1984. *A descriptive grammar of chontal maya (san carlos dialect)*. Tulane University dissertation.

Kroeber, A. L. 1907. *The washo language of east central california and nevada*. Vol. 4.5 (University of California Publications in American Archaeology and Ethnology). Berkeley: University of California Press.

Kuiper, Franciscus B. J. 1962. *Nahali: a comparative study*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.

Kung, Susan Smythe. 2007. *A descriptive grammar of huehuetla tepehua*. University of Texas at Austin dissertation.

Kutsch Lojenga, Constance. 1994. *Ngiti: a central sudanic language of zaire*. Vol. 9 (Nilo-Saharan Linguistic Analyses and Documentation). Köln: Rüdiger Köppe.

Lachler, Jordan. 2006. *A grammar of laguna keres*. University of New Mexico dissertation.

Lamberti, Marcello & Roberto Sottile. 1997. *The wolaytta language*. Vol. 6 (Studia Linguarum Africae Orientalis). Köln: Köppe.

Langdon, Margaret H. 1970. *A grammar of Diegueño: the Mesa Grande dialect* (University of California Publications in Linguistics). Berkeley: University of California Press.

Launey, Michel & Cristina Kraft. 1992. *Introducción a la lengua y a la literatura Náhuatl*. Mexico: Universidad Nacional Autónoma de México.

Levine, Robert. 1977. *The Skidegate Dialect of Haida*. New York: Columbia University dissertation.

Lindström, Eva. 2002. *Topics in the grammar of Kuot, a non-Austronesian language of New Ireland, papua new guinea*. Stockholm: Stockholm University.

Lorimer, D. L. R. 1935. *The burushaski language. volume 1: introduction and grammar*. Vol. 29 (Instituttet for Sammenlignende Kulturforskning, Serie B). Cambridge, Massacgusetts / Oslo: Harvard University Press / Aschehoug.

Loughnane, Robyn. 2009. *A grammar of oksapmin*. The University of Melbourne dissertation.

Macaulay, Monica. 1996. *A grammar of chalcatongo mixtec*. Vol. 127 (University of California Publications in Linguistics). Berkeley: University of California Press.

Maingard, Louis Fernand. 1962. *Korana folktales: grammar and texts*. Johannesburg: Witwatersrand Univ. Press. vii+81.

Manfredi, Stefano. 2014. Qualche osservazione sull'espressione del plurale nominale nel creolo arabo di juba. *RiCognizioni. Rivista di lingue, culture e letterature moderne.* 2(1). 53–61.

Maring, Joel M. 1967. *Grammar of acoma keresan*. Indiana University dissertation.

Maslova, Elena. 2003. *Tundra yukaghir*. Vol. 372. München: Lincom Europa.

May, Kevin. 1997. *A study of the nimboran language. phonology, morphology, and phrase structure*. La Trobe University MA thesis.

Meira, Sérgio. 1999. *A grammar of Tiriyo*. Rice University dissertation.

Miller, Amy W. 2001. *A grammar of jamul tiipay*. Berlin: Mouton de Gruyter.

Miller, Marion. 1999. *Desano grammar*. Vol. 6 (Studies in the languages of Colombia). Dallas: Summer Institute of Linguistics & University of Texas at Arlington.

Mithun, Marianne. 1988. Lexical categories and the evolution of number marking. In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology*. San Diego, California: Academic Press, Inc.

Miyaoka, Osahito. 2010. *A grammar of central alaskan yupik: an eskimo language*. Mouton de Gruyter.

Morse, Nancy L. & Michael B. Maxwell. 1999. *Cubeo grammar*. Texas: University of Texas at Arlington & SIL Publications.

Mortensen, C. 1999. *A reference grammar of northern embera languages*. Vol. 7 (Studies in the Languages of Colombia). Dallas: Summer Institute of Linguistics & The University of Texas at Arlington.

Muntzel, Martha C. 1986. *The structure of ocuilteco*. State University of New York at Albany.

Mushin, Ilana. 2012. *A grammar of (western) garrwa* (Pacific Linguistics). Berlin: De Gruyter Mouton.

Newman, Stanley. 1996. Sketch of the zuni language. In Ives Goddard (ed.), *Handbook of american indians. volume 17: languages*. Washington: Smithsonian Institute.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.

Nicklas, Thurston. 1974. *The Elements of Choctaw*. University of Michigan dissertation.

Obata, Kazuko. 2003. *A grammar of bilua*. Canberra: Pacific Linguistics.

Okrand, Marc. 1977. *Mutsun grammar*. University of California at Berkeley dissertation.

Olawsky, Knut. 2006. *A grammar of urarina*. Berlin, New York: Berlin: Mouton de Gruyter.

de Piñeros, Gabriele Petersen & Carlos Patiño Roselli. 2000. El idioma uitoto. In Maria S. Gonzáles de Pérez & Maria L. Rodríguez de Montes (eds.), *Lenguas indígenas de colombia. una visión descriptiva*. Instituto Caro y Cuervo.

Radin, Paul. 1929. *A grammar of the wappo language*. Vol. 27 (California Univer-

sity Publications in American Archaeology and Ethnology). Berkeley: University of California Press.

Rau, Der-Hwa Victoria. 1992. *A grammar of atayal*. Cornell University dissertation.

Rice, Keren. 1989. *A grammar of slave*. Berlin: Mouton de Gruyter.

Roberts, H. 1891. *A grammar of the khassi language for the use of schools, native students, officers and english residents*. Vol. 21 (Trübner's collection of simplified grammars of the principle Asiatic and European languages). London: Kegan Paul, Trench, Trübner & Co.

Roberts, John R. 1987. *Amele*. London: Croom Helm.

Romero-Figueroa, A. 1997. *A reference grammar of warao*. Vol. 6 (Lincom Studies in Native American Linguistics). München: Lincom Europa.

Rosato, Michele. 1980. *Didinga grammar and dictionary*. Rome.

Salamanca, D. 1988. *Elementos de gramatica del miskito*. Massachusetts Institute of Technology.

San Roque, Lila. 2008. *An introduction to duna grammar*. Australian National University dissertation.

Schmidt, Karl Wilhelm. 1991. Svan. In Alice C. Harris (ed.), *The indigenous languages of the caucasus. volume 1: the kartvelian languages*, 473–556. Delmar, New York: Caravan Books.

Schumann, Otto Gálvez. 1973. *La lengua Chol de Tila (Chiapas)*. México: UNAM.

Seiler, Walter. 1985. *Imonda, a papuan language*. Vol. 93 (Pacific Linguistics, Series B). Canberra: Australian National University.

Seyoum, Mulugeta. 2008. *A grammar of Dime*. Leiden University dissertation.

Shaul, David Leedom. 1982. *A grammar of nevome*. University of California, Berkeley dissertation.

Shimizu, Kiyoshi. 1983. *The zing dialect of mumuye: a descriptive grammar*. Hamburg: Helmut Buske.

Silver, Shirley. 1966. *The shasta language*. University of California, Berkeley dissertation.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Dixon, Robert M. W. (ed.), *Grammatical categories in Australian languages*, 112–171. New Jersey: Humanities Press.

Slater, Keith W. 2003. *A grammar of Mangghuer: a Mongolic language of China's Qinghai-Gansu sprachbund*. London: Routledge Curzon.

Smith-Stark, Thomas Cedric. 1974. The plurality split. *Papers from the 10th Regional Meeting of the Chicago Linguistics Society*.

Smye, Graham. 2004. *A grammar of anufo*. Vol. 7 (Language monographs). Institute of African Studies, University of Ghana.

Sonnenschein, A.H. 2005. *A descriptive grammar of san bartolomé zoogocho zapotec* dissertation.

de Sousa Filho, Sinval Martins. 2007. *Aspectos morfossintáticos da língua akwexerente (jê)* dissertation.

Stenzel, Kristine Sue. 2004. *A reference grammar of wanano*. University of Colorado dissertation.

Stevenson, R. C. 1969. *Bagirmi grammar*. Vol. 3 (Linguistic Monograph Series). Khartoum: Sudan Research Unit, University of Khartoum.

Sun, Hongkai & Guangkun Liu. 2009. *A grammar of anong: language death under intense comtact*. Vol. 5/9 (Brill's Tibetan Studies Library). Languages of the Greater Himalayan Region. E. J. Brill.

Sun, Hongkai, Tsieh Sun & Fengxiang Li. 2009. *A Grammar of Anong*. Brill.

T., King John. 2009. *A grammar of dhimal*. Vol. 8 (Brill's Tibetan studies library : 5, Languages of the greater Himalayan region). Includes bibliographical references (p. [609]-612). Leiden: Brill.

Taylor, Allan Ross. 1969. *A grammar of blackfoot*. University of California at Berkeley dissertation.

Terrill, Angela. 2003. *A grammar of lavukaleve* (Mouton Grammar Library). Berlin: Mouton de Gruyter.

Thiesen, Wesley. 1996. *Gramatica del idioma Bora*. Yarinacocha, Peru: Instituto Linguistico de Verano.

Thompson, S.A., J.S.Y. Park & C.N. Li. 2006. *A reference grammar of wappo* (UC Publications in Linguistics). University of California Press.

Thornes, Timothy Jon. 2003. *A northern paiute grammar with texts*. University of Oregon dissertation.

Thurgood, Graham. 2010. Hainan cham, anong, and eastern cham: three lan-

guages, three social contexts, three patterns of change. *Journal of Language Contact - VARIA* 3.

Valentine, J. Randolph. 2001. *Nishnaabemwin reference grammar*. Toronto: University of Toronto Press.

Van Der Voort, H. 2004. *A grammar of kwaza* (Mouton Grammar Library). De Gruyter.

Vidal, Alejandra. 2001. *Pilagá grammar (guaykuruan family, argentina)*. University of Oregon dissertation.

Wälchli, Bernhard. 2012. Grammaticalization clines in space: zooming in on synchronic traces of diffusion processes. In Bernhard W¨alchli & Bj¨orn Hansen Bj¨orn Wlemer (ed.), *Grammatical replication and borrowability in language contact*, 233–272. Berlin: Mouton de Gruyter.

Wang, Jun & Zheng Guoqiao. 1993. *An outline grammar of mulao*. Canberra: Australian National University.

Weber, David John. 1989. *A grammar of huallaga (huánaco) quechua* (University of California Publications in Linguistics). Berkeley: University of California Press.

Werner, Roland. 1993. *Tìdn-á?al: a study of midob (darfur-nubian)*. Berlin: Dietrich Reimer Verlag.

Wheeler, Max, Alan Yates & Nicolau Dols. 1999. *Catalan: a comprehensive grammar*. London: Routledge.

Widmann, Thomas & Peter Bakker. 2006. Does sampling matter? a test in replicability, concerning numerals. *Linguistic Typology* 10.

Williams, Marianne Mithun. 1976. *A grammar of tuscarora*. New York: Garland.

Wilson, Darryl. 1974. *Suena grammar*. Vol. 8 (Workpapers in Papua New Guinea Languages). Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Wilson, Patrizia R. 1980. *Ambulas grammar*. Vol. 26 (Workpapers in Papua New Guinea Languages). Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Winkler, Eberhard. 2001. *Udmurt*. Munich: Lincom Europa.

Yadav, Yogendra. 1985. Great andamanese: a preliminary survey. In D. Bradley (ed.), *Papers in south-east asian linguistics 9*, vol. 67 (Pacific Linguistics, Series A). Canberra: Australian National University.

Young, Robert W. & William Morgan. 1980. *The navajo language: a grammar and colloquial dictionary*. Albuquerque: University of New Mexico Press.

Zigmond, Maurice L., Curtis G. Booth & Pamela Munro. 1990. *Kawaiisu: a grammar and dictonary with texts*. Vol. 119 (University of California Publications in Linguistics). Berkeley: University of California Press.