

# The category of Number in Basque: II. Prehistorical and typological aspects

Mikel Martínez

► To cite this version:

Mikel Martínez. The category of Number in Basque: II. Prehistorical and typological aspects. The category of Number in Basque: II. Prehistorical and typological aspects, 2009, The category of Number in Basque: II. Prehistorical and typological aspects. <artxibo-00528573>

**HAL Id: artxibo-00528573**

**<https://artxiker.ccsd.cnrs.fr/artxibo-00528573>**

Submitted on 22 Oct 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The category of Number in Basque: II. Prehistorical and typological aspects

MIKEL MARTÍNEZ ARETA\*

## INTRODUCTION<sup>1</sup>

This is the second part of the article about number in Basque begun in the previous volume of *Fontes Linguae Vasconum* (see Martínez Areta 2009a). Points 1 and 2 (in the previous volume) were a description of the category of number in contemporary standard Basque (Point 1) and of the development of this category in the last centuries as far as it can be tracked by the attested historical data and comparison among dialects (Point 2).

In this second part I aim to step into a more speculative scenario, by discussing what the number category may have been like if we base on the data analysed in Points 1 and 2, but also on internal reconstruction and typological evidence (Point 3). Just as Points 1 and 2, Point 3 of this volume is divided into five subpoints that cover nouns, demonstratives, interrogative/relative and indefinite pronouns, personal pronouns, and finite verbal forms. Point 4 is a concluding remark, in which I compare the conclusions to a hierarchy of splits in number systems that has been proposed by typologists.

\* UPV/EHU.

<sup>1</sup> This paper has been written inside the framework of two Research Projects: 1) one financed by the Spanish Ministry of Education and Science under the heading “Fundamentos para unos Monumenta Linguae Vasconum II: historia, crítica y edición de textos vascos” (HUM2005-08047); and 2) another one financed by the Basque Government under the heading “Lingüística histórica e historia de la lengua vasca”. In both cases, the participants are lecturers or researchers at the Faculty of Arts of the University of the Basque Country, the director being Professor Joseba Lakarra. I would also like to thank Joseba Lakarra, Ricardo Gómez, and Julen Manterola for their helpful comments.

### 3. PREHISTORICAL DEVELOPMENT OF BASQUE NUMBER

I think, as discussed in Point 2, that the evidence in favour of the absence of a grammaticalised number distinction in old Basque is unquestionable for every grammatical category except personal pronouns. Let us gather the facts and try to go back beyond the beginning of the historical attestations in order to envisage what the situation may have been like in prehistorical phases. I shall adduce typological facts as a means to compare possible states of the language with attested empirical data.

#### 3.1. Nouns

In this subpoint, I am going to discuss two things. First, the origin of the plural morphs in non-local (abs. pl. *-ak*, erg. pl. *-ek*, etc.) and in local cases (*-eta-* suffix plus case mark), as well as the rise of the category of number(/definiteness) as it appears in historical phases (see Figure 1). Second, the features of the system which the language may have had before the rise of the historical number(/definiteness) system, as far as the number category is concerned. Supporting typological data will here be displayed.

To explain the origin of the declensional suffixes of the plural forms in non-local cases (I take up the argumentation left in 2.1.4 again), there are various possibilities of analysis, none of them totally conclusive. They can be broadly classified into two approaches: 1) the pl. suffix of nouns *-ak* comes from a bound suffix *-aga*, which can be related to the *-aga* of place and family names (see 2.1.4); 2) the pl. suffix of nouns –or at least some of them like the erg. pl. and the gen. pl.– comes from the 3<sup>rd</sup> grade pl. demonstrative, so that the grammaticalisation –that is, the transformation into article– of the pl. demonstrative is a process which runs parallel to the grammaticalisation of the sg. article, whose demonstrative origin has never been questioned by any Bascologist (see 2.1.1). Some representative works of the first view are Gavel (1920) and Mitxelena (1988 [1964], 1987 [1971]); two of the second Irigoien (1981) and Manterola (2008).

The main advantage of the second approach is that it implies a common development for the sg. and pl. paradigms. On the contrary, its main disadvantage is that, in most dialects, demonstratives end in *-ek* both in the abs. pl. and the erg. pl.: 1<sup>st</sup> grade *hau-ek*, 2<sup>nd</sup> grade *hori-ek*, 3<sup>rd</sup> grade *hai-ek* (see 1.2 and 2.2). Even in eastern dialects where monosyllabic 3<sup>rd</sup> grade demonstratives are present (Leiz. abs./erg. *hek* ‘those there’, etc., see 2.2.), the abs. pl. form is *-ak*. Now, the second approach might explain the evolution of some of the cases, as erg. *gizon hek* ‘those men there (erg.)’ → *gizon-ek* ‘the men (erg.)’, disyllabic forms being built upon the 3<sup>rd</sup> grade demonstrative root (*ha(r) + hek > ha(i)ek*, as suggested by Manterola 2008, see 2.2), but even in this case, the approach does not account for the abs. pl. suffix of nouns in western varieties, *-ak*. This approach also cannot explain the abs. pl. suffix of nouns (*-ak* in all dialects), at least not in an immediate way.

The first approach does not have this problem, for in principle it can account for all the noun suffixes. Thus, abs. pl. *\*-ag-a* would have developed into *-ak* in all dialects, whereas erg. pl. *\*-ag-e-k > \*-aek*, would have developed into *-ak* in the west but into *-ek* in the east. Similarly, the gen. pl. can be explained by as-

suming  $*-ag-e(n) > -aen > -en^2$ . For the dat. pl. we could set out from  $*-ag-i$ , which would develop into  $-ai$  in the west and into  $-ei$  in the east<sup>3</sup>. In all these, the proto-form should be analysed as an agglutination of a plural marker  $-ag-$  plus the corresponding case mark<sup>4</sup>. Nevertheless, although this approach apparently provides a straightforward explanation for the noun suffixes of non-local cases, it is at odds with some phonotactic characteristics that we can reasonably assume for the old common language. In no other part of the autochthonous lexicon can we find any suffix with the structure  $-VC-$ . It would make much more sense to split, say, abs. pl.  $*-aga$  into  $*-a-ga$  than into  $*-ag-a$ . In the first case,  $*-a-$  might be interpreted simply as the article, and  $*-ga$  perhaps as a pl. marker. The phonotactic structure of this  $*-ga$  as a morpheme would be impeccable<sup>5</sup>.

Further details about the possible origins of the plural morphs are given in Martínez Areta (2009b). I am here not going to take any position as to which option should be considered correct, but the rise of the number/definiteness paradigm as we know it in historical periods will be represented by (a variant of) the second approach. That would in fact imply that the pl. paradigm of nouns comes from the grammaticalisation of the 3<sup>rd</sup> grade demonstrative, to which the pl. marker  $-ga-$  followed by the case mark are then added (for instance, dat. pl.  $*gizon-ha-ga-i$  would have the structure stem-demonstr.-pl.mark-case). The rise of the number/definiteness paradigm would be approximately as follows<sup>6</sup>:

<sup>2</sup> The  $-aen$  phase has no asterisk because it is probably attested in Old B and South. (see CASTAÑOS GARAY, 1959). Some examples are: RS. (201) *iusturi-aen aq-ean euri-a* 'after the thunders (= *iusturi*, in gen. pl.) (comes) the rain', Kap. *il-aen* (...) *juzga-eta-ra* 'to judgement of the dead (= *il*)' (but gen. indef. *il-en*), Laz. *beso-aen gain-ean* 'on his/her arms (= *beso*)'. In Kap. and some other Old B. texts we also find forms like *iaquitun-een* 'of the wise (ones)', where gen. pl.  $-een$  seems to be a further step, with assimilation  $-aen > -een (> -en)$ . Finally, Landucci's dictionary of South. words (written in 1562) has *oyal-an puyate-a* 'the washing of cloths (= *oyal*)' and *conquiste-a erri-an-a* 'the conquest of the lands (= *erri*)', where the opposite assimilation, with further abbreviation of the vowel, seems to have occurred ( $-aen > -aan > -an$ ). MITXELENA (1977 [1961]: 117) interpreted these forms as middle points between the  $*-ag-en$  that he reconstructed and  $-en$ .

<sup>3</sup> In fact,  $-ei$  is the dat. pl. suffix of nouns in the standard language (see Tables 1 and 2 in 1.1), as well as in all modern dialects other than, roughly, B and G, where it is  $-ai$ .

<sup>4</sup> Erg. pl.  $*-ag(-e)-k$  might be analysed as pl. marker  $-ag-$  + epenthetic  $-e-$  + erg.  $-k$ . The epenthetic  $-e-$  would avoid the clash between two consonants, as i.e. in erg. indef. *gizon(-e)-k* 'man/men (erg.)'.

<sup>5</sup> In fact, in the Basque lexicon there are at least two different morphemes, in addition to the pl. one, which can be reasonably reconstructed as  $*-ga$ . One is the NORI and NORK 2<sup>nd</sup> p. sg. masc. verbal morph. As this is  $-k$  at the end of the word (*eman di-o-k* 'you (masc.) have given to him/her') but  $-a-$  word-internally (*eman di-a-gu* 'we have given you (masc.)'), it seems sound to postulate  $*-ga$ , which would have apocopated to  $-k$  at the end of the word, and would have lost the  $*-g-$  in intervocalic position. Another would be the ergative suffix  $-k$ . This does not alternate with any  $-ga$  allomorph, but based on alternations like *di-o-k* / *di-a-gu* and some others like the one displayed by the NORI and NORK 1<sup>st</sup> p. sg. morph (*diru-a dauka-t* 'I have money' / *uste du-t diru-a dauka-da-la* 'I think I have money'; compare, without such a contextual conditioning, RS (99) *damin-da sabel-a...* 'let me put (= fill) the belly...'), Lakarra (1995) and some later works have proposed that every  $-T\#$  (where  $T$  = stop) sequence in autochthonous elements must come from  $-DV\#$  (where  $D$  = corresponding voiced stop). As Lakarra (p.c.) points out to me, the erg. and the pl.  $*-ga$  might be originally the same morpheme, for ergativity and plurality can be related ideas (see TRASK, 1979).

<sup>6</sup> The forms in brackets are not grammaticalised at the corresponding period, but syntagmatic combinations of the structure [root + deictic], [root + deictic + pl. marker], etc. There undoubtedly existed more possible syntagmatic combinations (e.g. [*gizon* + *haur*] > *gizon-o*, attested (see 2.1.1)), but I here refer only to combinations that have generated grammaticalised number/definiteness distinctions. I would like to emphasise that I am looking at things from the point of view of the number category alone. Thus, I do not aim at a discussion of when the abs., the erg., the gen. and the dat. began to constitute a grammatical paradigm.

Figure 1. Rise of the number/definiteness axis in non-local cases

Phase I (no nr./def. opposition)	[3 <sup>rd</sup> grade deictic] [gizon-har] [gizon-har-ga] [gizon-har-i] [gizon-har-e(n)]	Neutral paradigm gizon gizon(-e)-ga gizon-i gizon-e(n)	[3 <sup>rd</sup> grade deictic + pl.] [gizon-ha-ga] [gizon-ha-ga-ga] [gizon-ha-ga-i] [gizon-ha-ga-e(n)]
	↓	↓	↓
Phase II (nr./def. opposition developed)	Sg. gizon-a gizon-ak gizon-ari gizon-are(n)	Indef. gizon gizon(-e)-k gizon-i gizon-e(n)	Pl. gizon-ak gizon-ek (W. -ak) gizon-ei (W. -ai) gizon-e(n)

Admittedly, this outline of the development is quite speculative. There is no way of ascertaining that a form like, say, erg. pl. *\*gizon-ha-ga-ga* has really existed at any time. If we dispense with the *-r-* of erg. sg. *\*gizon-ha(r)-ga*, and take into account that the erg. and pl. morphs may have been originally the same morpheme (see Martínez Areta 2009b), this hypothesis might imply that the case axis and the number/definiteness axis developed somehow simultaneously. Thus, at a time when there was no number opposition and the cases were not developed, a (probably periphrastic, not grammaticalised) construction like *\*gizon-ha-ga* might indicate in origin both ergativity and plurality, depending perhaps on the context. Once the two ideas became independent, the need was felt to have, alongside with an erg. (sg.) and an (abs.) pl., also an erg. pl., which was created by attaching another *-ga*, → *\*gizon-ha-ga-ga*.

If we opt for the first approach, then we will have to set out from *\*[gizon-ag-a]* > abs. pl. *gizon-ak*, *\*[gizon-ag(-e)-k]* > erg. pl. *gizon-ek* (W. *gizon-ak*), etc.<sup>7</sup> Here, *-ag-* would be a derivational suffix, perhaps indicating abundance and related ideas, and the element attached to the left would be the case mark. Otherwise, the rise of the number/definiteness category as such would not be very different.

As for the pl. forms of local cases, here the analysis is neater because the historical forms are clearly the agglutination of the pl. mark *-eta* plus the corresponding case mark (ines. pl. *-eta-n*, loc.-gen. *-eta-ko*, etc.). The bulk of the problem has been displayed in 2.1.4. Here again, I shall leave the different options open. To begin with, *-eta(-)* may have or not a Romance origin. If it has, things must have happened approximately as I have described them in 2.1.4. If on the contrary *-eta* is an autochthonous element, it must be related to the other *-(k)eta-s* that appear in the lexicon indicating abundance, *nomina actionis* and gerunds in some regions<sup>8</sup>. Although all these meanings

<sup>7</sup> The Old B. forms like Kap. *il-aen* 'of the dead', which Mitxelena interpreted as support for *\*il-ag-en* (see above) is not conclusive, for it might be also a middle point between *\*il-ha-ga-e(n)* and gen. pl. *il-en*.

<sup>8</sup> Even the conjunction *eta* 'and' could be mentioned here. CREISSELS (2006: 124-125) notes that one of the possible origins of plural marks is the reinterpretation of constructions with comitative sense. He mentions some Basque sentences like [*Patxi eta ...*] *ikus-i dugu* 'we have seen Patxi and the others (lit. Patxi and)', with ellipsis of the second term of the relation. He further suggests that in such

can be ultimately related, the original meaning of *-(k)eta* must have been one of plurality for inanimates<sup>9</sup>, later developing related meanings.

The second thing that I aimed to consider in this section is what the number category may have been like before the rise of the number/definiteness system present in historical periods. If we look at Figure 1, the first question that comes to mind is what meanings the bare root at Phase I may have contained. The issue is not new in Basque historical research. I mentioned Irigoien's insights on this point in 2.1.1<sup>10</sup>. Mitxelena (1987 [1970]: 292-293) was also aware of the decrease of the functions of the indef. forms during the historical phases: "Ahora bien, a lo largo de toda la corta historia de la lengua vasca, e incluso en lo que se puede reconstruir o adivinar de su prehistoria, se observa un rasgo constante, cada vez más manifiesto: la progresiva sustitución de la declinación indeterminada, que casi no subsiste más que residualmente, por la determinada"<sup>11</sup>. As an example, he mentions Leiz. (Rom. XI, 8):

(25)	<i>Eman</i>	<i>ukan</i>	<i>drau-e</i>	<i>Iainco-ac</i>	<i>spiritu</i>	<i>itho-bat;</i>	<i>eta</i>
	give	have	vb.aux.-3pl	God-ERG.SG.	spirit	drowned-one	and
	<i>begui-Ø,</i>	<i>ikus</i>	<i>ez-teçaten-çát:</i>		<i>eta</i>	<i>beharri-Ø,</i>	
	eye-ABS.INDEF.	see	not-vb.aux.subj.-'so that'		and	ear-ABS.INDEF.	
	<i>ençun</i>	<i>ez-teçaten-çát</i>		<i>egun-go</i>	<i>egun-erano</i>		
	hear	not-vb.aux.subj.-'so that'		day-L.GEN.	day-TEL.SG.		

'God gave them a sorrowful spirit; and eyes, so they could not see; and ears, so they could not hear up to the present day.

constructions the mark of additive or comitative meaning may have been reinterpreted as a mark of plurality of accompaniment and eventually become a normal plural mark. As a matter of fact, the Basque elliptic construction with *eta* is attested only in southern authors from the 19<sup>th</sup> century on (see 'eta' in Euskaltzaindia 1987-2005), but an etymological relation between *eta* 'and' and the pl. morpheme *-eta* (as in *etxe-eta-n* 'in the houses') is otherwise conceivable. Creissels draws the following parallel from Bédiondo (a Sara language, in southern Chad):

- (i) a. *kòdā*      *g'ē*      *nàg'ī*  
          Kody      with      Nadji  
          'Kody and Nadji'
- b. *kòdā*      *g'ē*  
          Kody      PL.  
          'Kody and his comrades'
- c. *bātī*      *g'ē*  
          ram      PL.  
          'Rams'

Although it is irrelevant from a diachronic point of view, it may be interesting to point out that both Old B and Old G have a coordinative structure of the type 'X-with Y', meaning 'X and Y': RS. (246) *ogui-a-gaz hur-a, or-a-gaz hero-en elicatura* 'bread and water (lit. 'with bread, water'), food of the fool and the dog (lit. 'of, with the fool, the dog)'; (G, 1705) *asto-a-rekin idi-a* 'the donkey and the ox (lit. 'with the donkey, the ox)'.<sup>9</sup>

<sup>9</sup> *-eta* on inanimates also appears from the oldest texts, as in Leiz. (Math. XXVI, 45) *bere discipuluetara* 'to(wards) his disciples' and RS. (391) *bildur bat da auere-etan ta asco guizon-etan* 'there is one fear in animals, and many in men'. But the use of *-eta-* on animates is nowhere systematic. Loc.-gen. pl. *-eta-ko*, for instance, never appears on inanimates with purely genitive meaning. There is ground to think, then, that this encroachment of *-eta* on animates is an innovation.

<sup>10</sup> See also LAFON (1999 [1948]).

<sup>11</sup> "Now, during all the brief history of the Basque language, and even in what we can reconstruct or guess for its prehistory, there is a constant, increasingly evident feature: the progressive substitution of the undetermined [= indefinite] declension, which only marginally survives, for the determined one" (translation mine).

The phrase *spiritu itho-bat* is determined by *bat*, but *begui* and *beharri* are not, they appear in the bare form, even if they are semantically plural. In contemporary dialects, for the bare form to appear it has to be determined by an indefinite determiner, such as *zenbait begi* ‘some eyes’ (see 1.3 and 2.3). Nowadays, instead of *begui* and *beharri* in (25) one would say *begi-ak* and *beharri-ak*, with abs. pl. In fact, even if the abs. pl. is in origin definite, it shows indefinite uses from the beginning of the written tradition, as in the following examples:

- (26) a. *Begui-ac dituzue-laric, ez-takusa-zue?* (Leiz., Mark VIII, 18)  
 eye-ABS.PL. vb.aux.-temp.suffix NOT-‘sees’-2pl.  
 ‘Having eyes, do you not see?’
- b. ... *obra on-ac eguin behar ditugu-la* (Ax., 30)  
 deed good-ABS.PL. do must vb.aux.-compl.  
 ‘... that we have to do good deeds’
- c. *Ydi-ac eta vey-ac yl ezau-z* (RS, 429)  
 ox-ABS.PL. and cow-ABS.PL. kill vb.aux.(imperative.)  
 ‘Kill (you, sg.) oxen and cows’
- d. *Es espada zapat-ac ta pantofle-ac bassasa-cayti* (Mik.)  
 no but shoe-ABS.PL. and slipper-ABS.PL. bog-CAUSAL  
 ‘No, but (- I prefer to wear) shoes and slippers, due to the bog’

As we can see by comparing both sentences by Leizarraga, i.e. the *begui-Ø* of (25) to the *begui-ac* of (26a), the substitution of the abs. indef. form –the bare root form– through the abs. pl. on these functions, then, is a process which is apparently in course in the 16<sup>th</sup> century. The logic of this development is clear. As we saw in 2.1.1, although the article *-a* is in origin definite –it is a demonstrative in origin– it has come to be used as a sg. marker. The same seems to be true for the abs. pl. *-ak*. As the whole pl. paradigm, it is morphologically definite, but since it has acquired the secondary meaning of indicating plurality, it has developed a number of indefinite uses in which its function is to indicate plurality.

In 2.1.2, I showed some more examples of indef. uses with some other grammatical cases. In fact, the whole declension has most probably undergone the same process. What seems to have happened is that, before the grammaticalisation of the number category as represented in Figure 1 had taken place, the bare root was the only grammaticalised, i.e. non-periphrastic way, of expressing number/definiteness categories. Consequently, it covered meanings which were later taken over by the sg. and pl. paradigms, once these emerged. More particularly, the bare root form must have contained at some point a general meaning implying the broad idea of the concept concerned, with no specification as for number. This happens in a number of languages, as we are going to see. This would imply that a sentence like, say, *gizon ikus-i dut* could mean “I have seen a man (= *gizon*)” or “I have seen several men”, depending on the context. Iturrioz (1985) called this character of the historical indefinite ‘transnumeral’, and I think it is also an appropriate term for referring to the bare root form in archaic phases of the language.

Let us now take a look at the typological situation of languages without grammaticalised number opposition on nouns. Over the world, there are

languages in which the category of number is not clearly definite, unlike in most European languages, so that the bare root may be different from any form with number specification. Corbett (2000: 10) calls this ‘general number’ (Iturrioz’s (1985) ‘transnumeral’). In the Cushitic language Bayso (southern Ethiopia), for instance, the form of the nouns represents their general meaning, and number specifications are indicated by different suffixes, as in (27):

- (27) a. *lúban*  
lion(GRAL.)  
‘One lion / (any number of) lions ...’
- b. *lubán-titi*  
lion-SG.  
‘One lion’
- c. *luban-jaa*  
lion-PAUCAL  
‘Few lions’
- d. *luban-jool*  
lion-PL.  
‘Lions’

[Corbett (2000: 10-11)]

Thus, in the sentence *lúban foofe* ‘I saw lion’, *lúban* is morphologically non-marked, which implies general meaning; the sentence can translate as ‘I saw a lion’ or as ‘I saw several lions’, depending on the context.

Another parallel is Fula (Fouta Jalon dialect), in Guinea. This language has an unmarked general form, and marked sg. and pl. forms: *toti* ‘toad(s)’, *totii-ru* ‘toad’, *totii-ji* ‘toads’. As Corbett (2000: 12) points out, general forms are used when number is irrelevant, as in:

- (28) *Ko*            *biini*    *tun*            *waawi*    *marde*    *beere*  
PARTICLE    bottle    only            can.PF    preserve    beer  
‘Only a bottle/bottles can preserve beer’

[Corbett (2000: 12)]

Nonetheless, an important difference between Fula and Basque is that in Fula not all nouns are able to show general number. In fact, only 11.5 percent of all nominals are. There is, moreover, a restriction: the form without the suffix must have at least two syllables. Thus, *nyaari peday* means ‘a cat scratches’ or ‘cats scratch’, *nyaari* being non-marked general; but *pucc-u latay* can also mean both ‘a horse kicks’ and ‘horses kick’, even if *pucc-u* is morphologically marked as sg. The reason is that otherwise, the bare root would have only one syllable.

In Igbo, plural is not grammaticalised either. As pointed out by Anagbogu (1995: 50): ‘A word can express either a singular or a plural meaning. Plurality is usually specified in one of the following ways: by numerals, by context, or by reduplication. But, whereas numerals are specific, reduplication is vague and non-specific’. Thus, in some cases the plural sense added by the reduplication has a double meaning; it indicates unspecified plurality as well



as large quantity of certain things to the exclusion of others. The contrast is exemplified in (29a, b):

- (29) a. *o nwèlu ji nà be yā*  
 he has yam(s) in house his  
 ‘He has yams in his house’
- b. *o nwèlu jī ji nà be yā*  
 he has ~ yams and only yams in house his  
 Double meaning:  
 1) ‘He has plenty of yams in his house’  
 2) ‘He has only yams in his house’  
 [Anagbogu (1995: 50)]

The three examples shown hitherto are taken from African languages, but many other regions also show non-distinctiveness as for number. An interesting case in point is Korean (see Song, 1997). This language lacks dual, trial and paucal marking, but uses the suffix *-tul* to signal plurality. This pl. mark, however, is not always obligatory. When a bare root like *salam* ‘person(s)’ appears with no pl. marker and is not determined by the definite article *ku*, it can be interpreted as either sg. or pl. If *-tul* is optional, it tends to appear most frequently with human nouns, less frequently with nonhuman animate nouns, and much less frequently with inanimate nouns (see Point 4).

A case in which *-tul* is optional occurs when the noun is preceded by a pl. quantifier such as *myech* ‘a few, many’, *yeles* ‘several’ or pl. numerals such as *seys* ‘three’. Here, *-tul* tends to be considered as redundant and thus usually not used. But when the noun is ‘definitised’ by *ku* and it refers to a pl. entity, then pl. marking is obligatory:

- (30) a. *salam(-tul)* (optional)  
 person(-PL.)  
 ‘Person(s)’
- b. *yele salam(-tul)* (redundant, usually omitted)  
 several person(-PL.)  
 ‘Several people’
- c. *ku salam-tul* (obligatory, if semantically plural)  
 ART. person-PL.  
 ‘The persons’  
 [Song (1997: 206-207)]

Song (1997: 206) concludes: “Thus, while definitized nouns without plural marking must always be understood to be singular, e.g. *ku salam* ‘the person’, the plural marker must always be overtly used to express the plurality of definitized nouns, e.g. *ku salam-tul* ‘the persons’”. The parallel with Basque is quite remarkable on this point, for there is a direct relationship between definiteness and number.

According to Aikhenvald (1994: 432), most languages of South America have no grammaticalised number opposition either. In Tariana (1994: 432 ff.), plural marking is optional and –nearly as in historical Basque, where

only the last word of any NP bear the ending– pl. affixes tend to appear either on a head noun or on an attribute, but rarely on both:

- |      |              |                    |              |                    |
|------|--------------|--------------------|--------------|--------------------|
| (31) | <i>apiya</i> | <i>taradite</i>    | <i>apiya</i> | <i>tarada-peni</i> |
|      | pig          | alive+CLASS:GEN.AN | pig          | alive-PL           |
|      |              |                    | 'a live pig' | 'live pigs'        |

[Aikhenvald (1994: 432)]

Plenty of parallels could be put forward from many other parts of the world, though no comprehensive list can be included within this analysis. Let us instead succinctly gather some of the features that recur across languages when we read typological literature on languages without grammaticalised number opposition on nouns, and briefly comment on the presence or not of some of these parameters in any phase of Basque.

1) The use of number marks (on nouns) according to pragmatic considerations, which can be contextual, emphatic, etcetera.

2) The optionality of these marks. This is obviously a consequence of the previous feature. As I have shown, it happens in Korean and Tariana, among other languages.

Has this optionality ever featured in Basque? I believe so. In fact, we can find traces of it in alternations like Etx. abs. indef. *punizione-Ø* ~ abs. sg. *goalardon-a* (compare (16a) and (16b)) and Leiz. abs. indef. *begui-Ø* ~ abs. pl. *begui-ac* above (compare (25) and (26a)), in similar contexts. This strongly suggests that during the process of abs. and pl. suffixes becoming number markers, there must have been phases in which they were optional.

3) The variability of pl. marks depending on categories like the animacy or concreteness of the pluralised noun. We have seen this feature in Korean, and it also happens in Tariana (Aikhenvald 1994: 432 ff.), although here the discussion about pl. markers overlaps with that about classifiers in general. Thus, whereas *-peni* is a pl. number for certain animate nouns (as in the example in (31) above), *-peri* can be used to mark plurality on inanimate nouns (32a), as well as to mark agreement on collective nouns (32b) or even on abstract nouns (32c):

- |         |                           |                    |
|---------|---------------------------|--------------------|
| (32) a. | <i>kadusi</i>             | <i>matsia-peri</i> |
|         | fan                       | good-COLLECT.      |
|         | 'Beautiful fans'          |                    |
| b.      | <i>mawina-yape</i>        | <i>matsia-peri</i> |
|         | pineapple-juice           | good-COLLECT.      |
|         | 'Pineapple juice is good' |                    |
| c.      | <i>pi-pitana</i>          | <i>matsia-peri</i> |
|         | 2sg.-name                 | good-COLLECT.      |
|         | 'Your name is beautiful'  |                    |
- [Aikhenvald (1994: 433)]

Similarly, in eastern Huasteca Nahuatl (Kimball 1990), animate nouns generally form the pl. by removing the abs. suffix and adding *-meh* (33a); on inanimate nouns, the abs. suffix is instead replaced by *-tinih* (33b). Abstract nouns, or nouns referring to a mass of objects, form no plural:

- |         |              |                 |
|---------|--------------|-----------------|
| (33) a. | <i>siwaλ</i> | <i>siwa·meh</i> |
|         | 'woman'      | 'women'         |

- b. *a-ma*            *a-matinih*  
       ‘paper’            ‘papers’  
       [Kimball (1990: 200)]

Further detail on this point will be given in Point 4, where it is argued that Basque has preserved no trace of such a phase, for in historical phases, the addition or omission of number marks is not dependant on the animacy / concreteness nature of the concept involved. Whether the language has ever had such a split can only be discussed on typological –not internal– criteria.

4) The use of reduplication as a pluralising device, where ‘pluralising device’ is understood in its broadest sense. The meaning added by the reduplication can be mathematical plural (more than one), but also distributive, quantifying, collective, etc. (see Moravcsik 1978).

An example of this can again be found in the eastern Huasteca Nahuatl, where (Kimball 1990: 203-205) “pluralization by reduplication has become semantically specialized. It indicates a distribution over space, rather than simply a number bigger than one. All nouns, even nouns that usually do not form any kind of plural, are susceptible to being pluralized in this way”. This is shown in (34):

- (34) *mi-lah*                    *mihmi-lah*  
       ‘planted field’            ‘various planted fields’  
       [Kimball (1990: 203)]

Since Lakarra (1995) and later works, it has been suggested that, in a certain phase of the language –at any rate, a very old one–, Basque noun morphology has had (partial) reduplication as a productive device. Some autochthonous disyllabic words with indisputable reduplicating structure are those in (35a). Those in (35b), instead, have been reconstructed according to a number of phonetic developments that can be postulated for the history of Basque. Their degree of likeness, then, varies.

- (35) a. *go-go* ‘soul; intention’            b. *\*da-dar* > *adar* ‘horn’  
       *go-gor* ‘hard’                            *\*da-dats* > *adats* ‘hair’  
       *ze-zen* ‘bull’                            *\*do-dol* > *odol* ‘blood’  
       *zo-zo* ‘thrush’                            *\*de-der* > *eder* ‘beautiful’  
   *\*di-di* > *idi* ‘ox’  
   *\*za-zal* > *azal* ‘skin’  
   *\*no-nol* > *ohol* ‘plank’

If all these and some others have been reduplications in old Basque, it is clear that the reduplicating device must have been linguistically productive at some point. A more complicated issue is the question of what the original meaning of this device may have been, i.e. the nature of the semantic relationship between a certain monosyllabic root and its reduplicated counterpart. This is not an easy matter, for all these nouns have certainly undergone a great deal of ‘desemantisation’. A possibility, however, is that at least one of the functions of the reduplication was precisely to indicate plurality or some plural idea.

In any case, if reduplication has ever existed as a pluralising device in Basque, it must be inside an extremely old chronology, and it is prudent for us to keep different chronologies distinct. Let it suffice to say, then, that reduplica-

tion may have been a productive pluralising device in a very early period, but that even after the disappearance of it as a morphological device, the language in most sectors of the grammar was still non-sensitive to number.

5) The presence of a singulative suffix to singularise or individuate the quotation form, which, unlike in most European languages, is not necessarily identical with the sg. form from a formal point of view.

We found this feature in the example of Bayso above (see (27b)). In Basque, I have already explained that the article *-a*, which comes from the distal demonstrative *\*har*, has developed a secondary function, which we can call singular or singulative (see 2.1.2). This evolution began within the early history of the language, developed during the centuries previous to the written tradition, and is still happening within early writings. Although we cannot consider today's Basque as a language without grammaticalised number in the standard sense of the term, the language has preserved several structural traits of languages without number, like the fact that the quotation form of words usually has *-a*<sup>12</sup>.

6) The presence of derivational suffixes, or even of periphrases with pluralising classifiers, to pluralise nouns. A remarkable example is Mandarin Chinese, as in the following example:

- (36) *wǒ*      *xǐhuan*      *zhèi*      *xiē*      *háizi*  
 I      like      this      PL.      child  
 'I like these children'  
 [Li & Thompson (1981: 41)]

In (36), *xiē* is a plural classifier which pluralises *zhèi... háizi* 'these children'. In fact, periphrastic plurals are particularly common in isolating languages.

If the Basque plural came from, say, *\*gizon argal ha ga* 'those thin men there → the thin men', as in Figure 1, it might be also interpreted as a periphrastic construction in origin.

7) The use of classifiers when numerals determine a noun. Here are several examples:

- (37) a.      Mandarin Chinese  
           *wǔ*                      *běn*                      *shū*  
           five                      CLASSIF.      book  
           'Five books'
- b.      Korean  
           *mal*                      *du*                      *mali*  
           horse                      two                      CLASSIF.  
           'Two horses'
- d.      Hmong njua  
           *ob*                      *tug*                      *naab*  
           two                      CLASSIF.      serpent  
           'Two serpents'  
           [Croft (2000: 222-223)]

<sup>12</sup> Obviously, this does not happen in dictionary entries, which have to give the radical form to make clear whether a final *-a* is an article or a part of the root (*etxe* 'house', not *etxea* 'the house'). This does happen, in some contexts where we would find the bare form in Spanish, as however, in the answer to the question 'How do you say 'house' in Basque/Spanish?' → Basque *etxe-a*, Spanish *casa*.

In Basque, there are no classifiers in NPs of numeral-noun structure, nor traces that there has ever been. But Manterola (2007) has noticed that, at least in Korean and Turkish –two languages with a (lack of) number category possibly similar to that of Basque–, when numeral determiners determine a noun, this appears in the bare form, as happens in Basque of any historical phase, and probably in old Basque. In fact, in Basque this happens only if the NP is indefinite. If it is definite, the abs. pl. suffix is required: *lau gizon* ‘four men’ – *lau gizon-ak* ‘the four men’ (see (1a) and (1b)). This feature, then, is also present in some languages with no grammaticalised number opposition.

To sum up the discussions in 3.1, it seems to me that the lack of a grammaticalised number category in Basque nouns and NPs has been sufficiently settled on internal criteria. On top of that, typological considerations also concur with these internal facts. It appears that, during the centuries previous to the beginning of the written tradition, Basque developed from a situation in which the bare stem (indef. in historical periods) indicated neither sg. nor pl., but rather a general meaning. If the noun appeared with a sg. demonstrative, then the phrase took on sg. meaning. If the origin of the pl. declension from the pl. demonstratives is admitted, then the same is true for the plural.

The similarities with Korean are particularly striking. In this language, since the definite article *ku* singularises the noun determined by it, *ku salam* is sg. (unlike *salam* ‘person(s)’, which is general), and since *ku salam* is sg. and only sg., if the concept referred to is semantically pl., this plurality is not specified by the context, so that the pluralising *-tul* is obligatory. Note that, just as in Korean, under the analysis according to which *\*gizon-ha-ga* > abs. pl. *gizon-ak* (see Figure 1), the pluraliser *\*-ga* is not suffixed to the bare stem, but to the stem-def.article structure (def.article-stem in Korean). Whether –to bring the analogy to the extreme– a phase /stem + optional pluralising *\*-gal* has ever existed in old Basque, we cannot know. The /stem + *\*-gal* structure that we find in historical Basque is not plural but erg. indef., as in RS. (335) *laru-c* ‘(a) peach-coloured (horse)’, Leiz. *haice handi-c* ‘(a) big wind (erg. indef.)’ (see (18)), and so on. However, as already posited, the erg. *\*-ga* and the pl. *\*-ga* may have been originally the same morpheme, one covering several meanings including transitivity and plurality.

If this were really so, the *\*-ga* which has given rise to the abs. pl. would present the same structural behaviour as the casual marks of the non-local cases in the sg.: it appears after the definite article (which in the sg. seems to have had *-r-*, as in *gizon-har-i*, *\*gizon-har-e(n)*, etc.) and the resulting structure *\*ha(r)-case.mark* (erg. sg. *\*har-ga*, dat. sg. *\*har-i*, abs. pl. *\*ha-ga...*) stands always at the end of the NP. This fact suggests that the marks erg./pl. *\*ga*, gen. *\*e*, dat. *\*i* may have been originally cliticised particles, before they turned into bound morphemes. In such a case, the starting point of the historical plural would be a periphrastic construction, such as *\*gizon ha ga* (> *gizon-ak* ‘the men’) for a noun or *\*gizon argal ha ga* (> *gizon argal-ak* ‘the thin men’) for a longer NP, which remember the above mentioned periphrastic plural of Chinese.

The local cases are somewhat easier to analyse, since the correct segmentation is stem-eta-case.mark in all cases (*etxe-eta-n* ‘in the houses’). More complicated is the question of its historical origin. If the Romance origin is accepted, as Schuchardt proposed (see 2.1.4), then what seems to have

happened is that a derivational suffix has been borrowed and become so productive in the borrowing language that it has been integrated into the declension. If, instead, we prefer to interpret *eta* –perhaps of a more archaic phonetic form in origin, e.g. *\*heta* or even *\*keta* (see 2.1.4)– as an autochthonous element, related not only to the *-(k)eta* of place names but also to some other elements like the *-(k)eta* of *nomina actionis* and to the coordinative conjunction *eta* ‘and’ (see 2.1.4 and 3.1 above), then *etxe-(h)eta-n* ‘in the houses’, *etxe-(h)eta-ra* ‘to the houses’, or even longer NPs like *etxe zaharr-(h)eta-n* ‘in the old houses’, *etxe zaharr-(h)eta-ra* ‘to the old houses’, are perhaps also analysable as periphrastic plurals in origin<sup>13</sup>.

A last point here is how the infix *-eta-* should be analysed if it is not a borrowed but an autochthonous element. In 2.1.6, I said that Trask’s (1997: 204) interpretation of the indef. local forms of vowel-stems (like *mendi-tan*) is simply that *-e-* of the plural *etxe-eta-n* was reinterpreted as a pl. mark, and consequently dropped in the indef. In fact, this explanation is not unanimously accepted, as Lakarra (p.c.) reminds me, for an interpretation of *-e-ta-* as two different morphemes would fit better into the principles of word structure in old Basque, so that in ines. indef. *mendi-ta-n* only *-ta-* appeared, but in ines. pl. *mendi-e-ta-n* both *-e-* and *-ta-*. The former might be in relation to the *-e-* of demonstratives like *hau-e-k*, *hori-e-k*, *hai-e-k*, East. *h-e-k*, etc. (see 3.2), and even to the *-e-* (*-te*, *-de*, see 2.5.2 and 2.5.3) which in the verbal morphology pluralises NORK and NORI 3<sup>rd</sup> (and 2<sup>nd</sup>) person marks<sup>14</sup>.

<sup>13</sup> Under the analysis of Figure 1, neither the *\*ha-ga* of *\*gizon-ha-ga* nor like *\*(h)eta-n* of *\*etxe-(h)eta-n* –which does not appear in the Figure but would be the counterpart in local cases– would be bound suffixes. This means that in principle there is no radical incompatibility with the analysis according to which the noun declensional suffixes come from cliticised demonstratives. In fact, the demonstrative origin is a feature common to the analysis that I have drawn in Figure 1 and Manterola’s view as depicted in (2008), which opposes both to Mitxelena’s bound morpheme analysis (*\*gizon-ag(-a)* > abs. pl. *gizon-ak*, *\*gizon-ag(-e)-k* > *\*gizon-æk* > erg. pl. West. *gizon-ak* / East. *gizon-ek*, etc.). The difference between Figure 1 and Manterola (2008) is that, whereas the former proposes an original agglutinational regularity from which, as a consequence of diverse phonetic developments, all attested variants would have arisen, the latter prefers a more straightforward account of the facts, whereby several attested 3<sup>rd</sup> grade demonstratives would be the direct origin of at least some of the pl. suffixes of the noun declension in the east (from East. abs./erg. pl. *hek* ‘those there (abs./erg.)’, *gizon hek* ‘those men there (abs./erg.)’ > East. erg. pl. *gizon-ek*, from (CSM, 1025) *Arbel-goi-hen* (see 2.1.4), *\*gizon-hen* ‘of those men’ > *gizon-en*, etc.). There are advantages and disadvantages in both views. An advantage of Manterola’s (2008) view is that according to it we can dispense with intricate evolutions like *\*gizon-ha-ga-ga* > erg. pl. West. *gizon-ak* / East. *gizon-ek* or *\*gizon-ha-ga-e(n)* > gen. pl. *gizon-en*. Another advantage is that it provides a better explanation for the accent of the pl. suffixes of local cases in S, as in S. *etxe-éta-ko* ‘of the houses’. The accent of this dialect, which is very conservative and reflects several archaic traits, is paroxytonic in most of the lexicon. Exceptions have diverse explanations. S. *etxe-éta-ko* can be explained if it comes from *etxe hétako* ‘of those houses there’, as the 3<sup>rd</sup> grade demonstrative has an accent on the first syllable (*héta-ko* ‘of those there’, *héta-n* ‘in those there’, etc.). In fact, Manterola’s (2008) account explains the local cases better in several respects, as we have *-eta-kol-nl(...)* in the noun declension of all dialects. 3<sup>rd</sup> grade demonstratives like *ha(i)-eta-n* can be explained by the addition of *héta-kol-nl(...)* to the deictic root *ha-*. The disadvantage of this view is that, in non-local cases, it can not explain: 1) the *-a-* of abs. pl. *gizon-ak* (in both western and eastern dialects; nor 2) the *-a-* of West. erg. pl. *gizon-ak*, dat. pl. *gizon-ai*, and Old B. *il-aen*. It seems that in these cases we have to set out from a form with *ha-* (*-a(g)-* in Mitxelena’s account). This implies that, even if *gizon hek* > erg. pl. (East.) *gizon-ek* and *gizon hen* > gen. pl. *gizon-en*, as well as the corresponding evolutions of local cases, were right, the rest of the cases require different explanations anyway.

<sup>14</sup> Here we could also list the *-e-* of the noun declension in the east (erg. pl. *gizon-e-k*, etc.), but it would be redundant, for if that *-e-* is not a phonetic result of *-ag-e-k* (Mitxelena) or of *-ha-ga-ga* (Figure 1), then it would be originally a morpheme resulting from *gizon-ek* < *gizon he-k* (MANTEROLA, 2008), so again what we should explain is the *(h)e* of the demonstrative *hek*.

### 3.2. Demonstratives

As I said in 2.2, a major problem here is to explain whether East. *hek* ‘those there (abs./erg.)’ is a contraction of *haiek*, as Azkarate & Altuna (2001: 18) take for granted, or *haiek* came about from the juncture of *ha(r)-* + *(h)ek*, as Irigoien (1981) and Manterola (2008) prefer. As stated in 3.1, the second option is better in several respects, especially as far as local cases are concerned.

As for non-local cases, a problem is why we have abs. *-ek* and erg. *-ek* forms in most dialects, while in the noun morphology we find abs. *-ak* and erg. *-ak* (in the west) / *-ek* (in the east). It most probably has to do with the fact that, since demonstratives are definite by their semantic nature, they do not have to set out from, say, 1<sup>st</sup> grade *\*hau-ha-ga* (*ha-* being the definite article, as in the noun morphology), neither in the abs. nor in the erg. Whether this implies that the *-e-* of *hau-e-k*, *hori-e-k*, *hai-e-k*, etc., is a later added morpheme, which could be related to the *-e-* of the plural of the local cases (as in ines. pl. *mendi-e-ta-n* but ines. indef. *mendi-ta-n*), to the *-e-* of the erg./gen./dat. pl. in the eastern noun-declension, and to the *-e-* of the verbal morphology (see 3.1), I dare not judge. It would certainly deserve a more detailed study than can be offered here<sup>15</sup>.

What seems remarkable to me, though, is the fact that the formation of the plural, at least as far as the 1<sup>st</sup> grade and 2<sup>nd</sup> grade forms are concerned, draws a clear-cut isogloss between (by and large) B and the rest of the dialects. The difference is represented in Table 12:

Table 12. Some plural forms of 1<sup>st</sup> and 2<sup>nd</sup> grade demonstratives in B and in the rest of the dialects

	<u>Biscayan</u>	<u>Rest of the dialects</u>
(1 <sup>st</sup> grade)		
Abs. pl.	<i>ón-(e)ek</i>	<i>hau-ek</i>
Erg. pl.	<i>ón-(e)ek</i>	<i>hau-ek</i>
Gen. pl.	<i>ón-(e)en</i>	<i>hau-en</i>
Dat. pl.	<i>ón-eri</i>	<i>hau-ei</i>
(2 <sup>nd</sup> grade)		
Abs. pl.	<i>órr-(e)ek</i>	<i>hori-ek</i>
Erg. pl.	<i>órr-(e)ek</i>	<i>hori-ek</i>
Gen. pl.	<i>órr-(e)en</i>	<i>hori-en</i>
Dat. pl.	<i>órr-eri</i>	<i>hori-ei</i>

<sup>15</sup> That a pluralising *-k* is recharacterised by *-e-* is historically attested on several indef. determiners, as I showed in 2.3. Sometimes, the sg./pl. and abs./erg. oppositions intertwine. Thus, *bat-zu-k* ‘some, several’ is erg. pl. in Leiç. (opposed to abs. pl. *bat-zu*) but abs. pl. in Ax. and some other 17<sup>th</sup> Lapurdian authors (opposed to erg. pl. *batzu-e-k*). In northern texts of the 18<sup>th</sup> century (see ‘batzu’ in Euskaltzaindia, 1987-2005), there is some tendency towards levelling (both abs. pl. and erg. pl. *batzu-ek*), but in the 19<sup>th</sup> and 20<sup>th</sup> centuries the opposition arises again under the form abs. pl. *bat-zu* / erg. pl. *bat-zu-ek*. In the west, there is no form distinction, *bat-zu-k* being typical of B and *bat-zu-e-k* of G. Whether similar processes have taken place in the demonstratives is not always easy to tell, for in some cases it is not absolutely clear whether what we have is the result of a contraction (*hori-e-k* > *hori-k*) or of the lack of addition of *-e-* (*hori-k* from the outset). It is beyond any doubt, however, that the different results have sometimes morphologised. Thus, Souletin authors of the 17<sup>th</sup> and 18<sup>th</sup> centuries make such distinctions (IRIGOIEN, 1981: 386 ff.). We already saw, in 2.2, Oih. abs. pl. *hau-c*, *hori-c*, *he-c* / erg. pl. *au-ec*, *hori-ec*, *hei-ec*. Belapeyre (S, 1699) has *hoy-ec* for abs. and erg. pl. in the 1<sup>st</sup> grade demonstrative, but in the 2<sup>nd</sup> grade abs. pl. *hori-c* and erg. pl. *hori-ec*, and in the 3<sup>rd</sup> grade abs. pl. *hour-ac* and erg. pl. *hay-ec*. The same forms

The forms of the right-hand column are the ones of the standard language. In fact, plenty of variants exist in both groups. But all variants certainly point to something close to these from a diachronic perspective. The important point is that all B. pl. forms are built upon the root which in the sg. (of any dialect) appears in all cases other than the abs. (*hon-*, *horr-*), whereas the pl. forms of the rest of the dialects are built upon the same root as the abs. sg. (*hau-* (sometimes > *ho-*), *hori-*)<sup>16</sup>. The logical implication of this is that, unlike the plural of the noun declension, the formation of the plural of these demonstratives cannot be brought back to the common language. Therefore, the plural must be older in the noun than in the 1<sup>st</sup> and 2<sup>nd</sup> grade demonstratives<sup>17</sup>.

### 3.3. Interrogative/relative and indefinite pronouns

Apart from fixed expressions (*hitz egin* ‘to speak (lit. ‘to make word’)’), some predicative constructions (*interesgarri suerta-tu zai-t* ‘it has proved interesting to me’) and some other functions (see 2.1.2), the indef. number is today only used when it is determined by an interrogative/relative or indefinite pronoun. When determined by one of these, the indef. is the only grammatical form. This must have been different in prehistorical phases, when the bare root form covered a wider semantic scope (see 3.1).

These pronouns are either autochthonous elements built on the roots *no-* and *ze-* (*no-r* ‘who’, *no-n* ‘where’, *no(i)-z* ‘when’, *ze-r* ‘what’, *ze(i)-n* ‘which’) or combinations of these with some other particles. The meaning added by these particles include ‘some, something’ (*nor-bait* ‘someone’, *noiz-bait* ‘sometime’, *zer-bait* ‘something’, *zen-bait* ‘several’), ‘how much/many’ (*zen-bat* ‘how much/many’), ‘-ever’ (*nor-nahi* ‘whoever’, *zer-nahi* ‘whatever’, etc.), ‘any-’ (*edo-nor* ‘anyone’, *edo-non* ‘anywhere’, etc.), ‘each’ (*nor-bera* ‘each one’). Some of these are probably built on the model of Romance languages, as is the case of *-bait*, < *-baita* ‘also’ (< *bai-* ‘yes’ + *-eta* (see 3.1 and 3.2), Leiz. *bayeta* and, with *ere* ‘too’, Etx. *bayetare*), which can be compared to Lat. *quisque* ‘whoever’, etc. Perhaps of an older chronological layer are *inor* (< *\*e-nor*), *inon* (< *\*e-non*), *inoiz* (< *\*e-noiz*), *e-zer*, *e-zein*. As ‘any-’ in English, these mean ‘some’ in negative sentences (*zer-bait ikus-i du-t* ‘I have seen something’, but *ez du-t ezer ikus-i* ‘I have not seen anything’).

Now, before all these combinations had generated so many different meanings, the semantic scope of the basic roots *nor*, *no-n*, *no(i)-z*, *ze-r*, *ze(i)-n* must have been wider, covering some of the meanings –all indefinite– that

are present in Maister (S, 1757), except for the 3<sup>rd</sup> grade erg. pl. form, which is *he-c* in Maister. In some other cases, the *-Ø/-e-* dichotomy expresses not a abs./erg. but a 1<sup>st</sup> grade / 2<sup>nd</sup> grade opposition, although this is but the result of diverse diachronic vicissitudes. This occurs in the Basque of Baztan (IRIGOIEN, 1981: 373), where the 1<sup>st</sup> grade abs./erg. pl. form is *o-k* (< *(h)au-k*, the traditional form of L), but the 2<sup>nd</sup> grade abs./erg. pl. form *oi-ek* (< *(h)ori-ek*, with loss of intervocalic *-r-*). Another case in point is Ax. 1<sup>st</sup> grade abs./erg. pl. *hau-c*, but 2<sup>nd</sup> grade abs./erg. pl. *hori-ec*.

<sup>16</sup> In some cases, the 1<sup>st</sup> grade and the 2<sup>nd</sup> grade plural forms may undergo several syncretisms (see IRIGOIEN, 1981: 370-372).

<sup>17</sup> Obviously, this argument is true only if we admit that variants like (erg. pl.) West. *gizon-ak* / East. *gizon-ek* or (dat. pl.) West. *gizon-ai* / East. *gizon-ei* are but different phonetic developments of the same morphological proto-form. If we instead set out from different morphological proto-forms to explain the different variants, the formation of the plural in the noun declension cannot be brought back to the common language either.



were later taken over by different periphrases, like the ones just mentioned. As these were expressed by these periphrases, *no-r*, *no-n*, *no(i)-z*, *ze-r*, *ze(i)-n* were limited to the interr./relat. function. This might explain why these basic forms have preserved, up to the present day, some uses other than the pure interr./relat. ones, as in expressions like *nor da nor* ‘each one is each one’, *hobe da noiz*, *ezen ez inoiz* ‘it is better sometime (= *noiz*), than never’, *zein-i bere-a bihur-tze-a* ‘to give each one (= *zein-i*) what is his/hers’, etcetera<sup>18</sup>.

### 3.4. Personal pronouns

The number marking that we should postulate for personal pronouns in the common language is fairly clear. We have *ni* ‘I’ and *hi* ‘you (sg.)’ on the one hand, and *gu* ‘we’ and *zu* ‘you (pl.)’ on the other. 3<sup>rd</sup> person pronouns proper do not exist in Basque<sup>19</sup>. It is interesting to note that the sg. forms have an *i*-vocalism and the pl. forms have an *u*-vocalism. In Martínez Areta (2008: 382-383), I suggested that this might be related to the fact that *i* is, in many a language, phono-symbolically associated to diminutive or diminutive-like meanings. One of these plurality-like meanings is that of singularity (see Diefloth 1994). A relationship between *u* and plurality is not so straight-forward, but I think that further typological research could be made on this point.

Number, then, seems to be grammaticalised on 1<sup>st</sup> and 2<sup>nd</sup> p. pronouns by means of suppletion. On the contrary, in the rest of the grammar there is apparently no grammaticalised opposition. As I shall argue in 4, this is not extraordinary. It happens, for instance, in Asmat (New Guinea), Guaraní (Paraguay) (both in Corbett 2000: 63-64) and Coast Tsimshian (a Penutian language of western North America) (Stebbins 1997: 31). In Asmat, the figure is as follows<sup>20</sup>:

Figure 2. Pronouns in Asmat

	sg.	pl.
1 <sup>st</sup> person	no	na
2 <sup>nd</sup> person	o	ca
3 <sup>rd</sup> person	a	

[Corbett (2000: 64), from Voorhoeve (1965: 143)]

<sup>18</sup> If we look at these meanings in the *Orotarikoa* (Euskaltzaindia, 1987-2005), most of them are not given as pandialectal. Examples exist, however, in the oldest authors. Here are some examples: Leiz. (Ins. E 1r) *nor-i bere-a renda daqui-o-n* ‘let each one be given what is his/hers’, Ax. (39) *nor bere ostatu jaquin-ean beçala* ‘as each one in his known guest house’, Landucci (South., 1562) *ezta nox bere* (= Sp. “nunca”) ‘never’, Leiz. (Joh. x, 3) *bere ardi-ac dei-tzen d-it-u cein bere icen-ez* ‘(he) calls his sheep, each one by its own name’, Etxeberri Ziburukoa (Man. I, 27; L. 1627) *zeñ-ec bere molde-ac* ‘each one (- tends to, prefers) his own manners’. In any case, in combinations like *nor-bait* ‘someone’ and *zer-bait* ‘something’ (*-baita* ‘also’), among others, it seems clear that the meaning of the first element must have been something like ‘someone’ and ‘something’ at some point, in addition to having the interrogative sense historically maintained.

<sup>19</sup> Anaphoric reference can be made by means of the 3<sup>rd</sup> grade demonstrative, as is typical in the east, or by the pronoun *ber-a* (pl. *ber-aiek*, *ber-ak* or B. *eur-ak*), which cannot be strictly considered as a 3<sup>rd</sup> person pronoun because it gathers some other meanings, like ‘idem’ or ‘ipse’. These forms decline as demonstratives in most cases (*ber-aiek* (abs./erg. pl.)), for it is /ber- + 3<sup>rd</sup> grade demonstr./, as adjectives in the rest (*ber-ak* ‘they’, *ber-e* ‘his/her(s)’, *ber-en* ‘their(s)’).

<sup>20</sup> It has an inclusive-exclusive distinction in the first person, too.

The parallel with Basque is quite remarkable, because on top of having the number split at the same point (see point 4), the iconicity for marking number is quite similar: 1<sup>st</sup> and 2<sup>nd</sup> p. sg. marked by a vowel, 1<sup>st</sup> and 2<sup>nd</sup> p. pl. marked by another vowel, and no number distinction in the 3<sup>rd</sup> p.<sup>21</sup>

An observation made by Lakarra (p.c.) which deserves further typological research is the fact that the pronoun *zu*, which we can reasonably reconstruct as ‘you (pl.)’ for the common language, is homophonous –if not the same morpheme in origin– with one of the derivational suffixes indicating abundance in Basque (see 2.1.5). The two corresponding verbal marks are *za-* and *-zu* (see 1.5 and 2.5 for uses and distribution), of clear pronominal origin. The latter is clearly the same morph, and the former is another derivational suffix semantically and formally quite similar to *-zu*. I think that another obvious homophony is that between the 1<sup>st</sup> pl. verbal morph *ga-* (see 1.5 and 2.5) and the (probably) pluralising *-ga* of the declension. On the contrary, the morph *-gu* cannot be compared to anything else.

Perhaps further advance in the historical grammar of Basque can be done by pursuing this direction, that is, by comparing pronominal morphemes –whether pronouns proper or verbal morphemes of pronominal origin– with other elements which have played some role in the grammar of Basque at some point. Among the 1<sup>st</sup> and 2<sup>nd</sup> sg. pronouns and pronominal morphs, it might be promising to investigate in what sense the 1<sup>st</sup> sg. active *\*-da* (as in *dakar-t* ‘(I) bring’, < *\*-da*) can be related to the *\*da* which is possibly an imperfective particle on finite verbs (*da-kar* ‘(s/he) brings’, etc.) and which might be related to some other morphemes, like the complementiser *-la* and the causative infix *-ra-* (*e-karr-i* ‘to bring’ → *e-ra-karr-i* ‘to attract’), as suggested in Lakarra (2005). Another interesting comparison is the one between the 2<sup>nd</sup> sg. (masc.) active *\*-ga* and the *\*ga* that I have shown in 3.1. Obviously, if there is such a relation the latter cannot have pl. sense, but perhaps erg. sense.

In any case, for any hypothesis of this type to be solid we should require strong evidence. That is why typological parallels could prove particularly helpful in explaining how common morphemes may have bifurcated semantically.

### 3.5. Verbal forms

The issue of grammatical number gets increasingly complicated as we turn to consider finite verbs. Several explanations have been put forward in order to explain them. According to Gómez & Sainz (1995: 265-268), who draw the syntax of the Old Irish verb as a parallel, the order of the morphological elements on finite verbs reflects a previous syntactic order in which the first position was occupied sometimes by the subject, *n(V)(C)-*, *h(V)(C)-*, *g(V)(C)-*, *z(V)(C)-* (see 1.5), and sometimes by a sentence connective, *da-*, *ze-*,

<sup>21</sup> The state of affairs that I have conjectured for the old language up to here, then, is: no number distinction except on existing pronouns. Nonetheless, this number distinction on pronouns is suppletive, not morphological, as is the case in Mandarin Chinese, for instance. In this language, there is no grammaticalised number except for pronouns either. Here (NORMAN, 1988: 158-159), the pronouns for the three persons, *wǒ* ‘I’, *nǐ* ‘you’, *tā* ‘s/he, it’, are pluralised by the addition of the suffix *-men*, which in any case is not obligatory in any context: *wǒ(-men)* ‘we’, *nǐ(-men)* ‘you (pl.)’, *tā(-men)* ‘they’.

*le-*, *be-* (see 1.5), so that the verb appeared always in the second position. Even if, according to Guivon's principles, it can be correct that today's morphological order often represents yesterday's syntactic order, I consider this to not necessarily be the neutral, non-marked order. Indeed, the fact that every finite verb has an agreement mark for every 1<sup>st</sup> and 2<sup>nd</sup> p. argument of the sentence (and also for 3<sup>rd</sup> p. sg. datives, and in a later phase 3<sup>rd</sup> p. pl. absolutes, ergatives and datives, see 2.5.4) makes me think that the structure of finite verbs came about in topicalised constructions, in which clitics representing all the elements of the action were attached to the verbal root.

As for number, it was assumed in 2.5 that the 1<sup>st</sup> and 2<sup>nd</sup> person marks were attached first. This implies a suppletive number marking for these marks, and also indicates that the marks for 3<sup>rd</sup> person pl. came about much later. This approach was already envisaged by Trask (1977: 206), who proposed the following reconstruction and development for the synthetic verb *egon* 'to be' (NOR conjugation) in the present:

Figure 3. Evolution of *egon* 'to be' according to Trask (1977)

1 <sup>st</sup> sg.	* <i>na-da-go</i>	>	<i>nago</i> 'I am'
2 <sup>nd</sup> sg.	* <i>ha-da-go</i>	>	<i>hago</i> 'you (sg.) are'
3 <sup>rd</sup> sg.	* <i>da-go</i>	>	<i>dago</i> 's/he, it is'
1 <sup>st</sup> pl.	* <i>ga-da-go</i>	>	* <i>gago</i> 'we are'
2 <sup>nd</sup> pl.	* <i>za-da-go</i>	>	* <i>zago</i> 'you (pl.)'
3 <sup>rd</sup> pl.	* <i>da-go</i>	>	* <i>da-go</i> 'they are'

This development is previous to the extension of the abs. pl. marker of the 3<sup>rd</sup> person<sup>22</sup>. After this, a pl. mark began to distinguish 3<sup>rd</sup> person pl. from 3<sup>rd</sup> person sg. forms. This process –I stick for the moment with the abs. pl. marks in the NOR conjugation– affected all dialects, and took place with some similarities but also with important differences. The verb put forth by Trask, for instance, has *dago-z* as the pl. form in B, but *dau-de* (< \**dago-de*) further to the east. This is apparently the only intransitive verb which takes *-de* as pl. marker (leaving aside forms like *dira-de* 'they are', *zira-de-n* 'they were'..., of the verb *izan* 'to be' (see 2.5.4), where it is actually a redundant morpheme). Nonetheless, Lakarra (2006: 605) suggests that *dago* hides also in the general *dau-tza* 'they lie', < \**dago-tza*. Oscillation between *-tza* and *-tz* is also found in the verb *ibili* 'walk' (B. *dabil-tz* 'they walk' / *dabil-tza* further to the east), and between *-tza*, *-(t)zi* and *-(t)z* in the verb *joan* 'go' (S. *doa-tza* 'they go', BN. *d(o)a-zi* / *doa-z* further to the west).

As pointed out in 2.5.1 following Lakarra (2006), this morpheme seems to be the same as the derivational suffix *-tz(a)* of place names and the derivational suffix of the lexicon (*diru-tza* 'lot of money'). If this is correct, we can further wonder how this morpheme may have arisen. Has it generated in the noun morphology, and passed from there to the verb morphology, or has it generated in the verb itself? This poses a set of very interesting questions, for both processes are abundantly attested across the languages of the world. In fact, some areas of the world show some tendency to one or the

<sup>22</sup> An alternative analysis is proposed by De Rijk (1992), according to which *n(i)-da-go* > *nago*, etc., but this does not affect the question concerning number.

other possibility. Many Indo-European languages, for instance, have number morphemes on verbs, but in many of these cases those are verbal marks that have come about as a result of some kind of ‘copying rule’ or agreement with a plural subject –agreement with a direct object, as in Basque, is much rarer– of the corresponding sentence. On the contrary (Mithun 1988: 218): “Verbal number marking is pervasive in North America, but it does not represent simple agreement with nouns. It operates independently, modifying the verbs themselves”. On the other hand (Mithun 1988: 218): “The number marking typical of Indo-European nouns enumerates entities, while that of North American verbs usually quantifies aspects of events”.

In other words, number marking in Indo-European languages usually marks plurality of some argument –typically the subject in nominative/accusative languages– and most often starts off on nouns and passes from there to verbs by means of some ‘copying rule’ or some agreement process, whereas number marking in North-American languages usually appears on verbs and indicates not plurality of arguments or entities, but iterativity of the action and related meanings<sup>23</sup>.

If we leave aside the theoretical component of the issue, what we should establish is whether the pluraliser *-tz(a)* has arisen on nouns and jumped from there to verbs –and if so, how–, or whether it has arisen directly inside the verb. I see three possible interpretations of the process: 1) as a process of agreement between the noun and the verb (*dakar belar-tza* → *dakar-tza belar-tza* “s/he brings plenty of grass”)<sup>24</sup>; 2) as a “migration” of the morpheme from the noun to the verb (*dakar belar-tza* → *dakar-tza belar*); 3) as a creation of the pl. mark on the verb itself (*dakar belar* → *dakar-tza belar*).

The first possibility is abundantly attested, although, as I said, in languages of nominative/accusative morphology the most typical ‘plural copying’ agreement is that with the subject. A ‘plural copying’ of the direct subject is attested in the Austronesian language Wasika:

- (38) *nu buruk nunga kida-nd-am*  
 3sg. pig PL. cook-pl.-3sg.PAST  
 ‘He cooked the pigs’  
 [Stebbins (1997: 6-7)]

This is actually a periphrastic pl. construction, *nunga* being an element implying obj. pl., but not bound to the object itself (*buruk*). However, this is not the most common way of expressing ‘He cooked pigs’, for *nu buruk nunga kidi-am* (with no mark on the verb) and *nu buruk kida-nd-am* (the

<sup>23</sup> DURIE (1986: 365) states: “A wide variety of evidence shows that Number in natural languages is not inevitably a nominal category. By stem suppletion, and by more productive morphological devices, languages express Number categories which are inherent to the verb, linked to verbal semantic structure. This provides the potential for concord between verbal Number and NP Number, but this concord shows the expected properties of semantic selection rather than agreement”. LANGDON (1992: 406-407) is of the same opinion, proposing a ‘noun agreement’ in the Yuman languages, whereby it is not the verb that attracts the pl. morpheme from the noun, but the other way round: the pl. mark has arisen on the verb and the noun has developed an agreement with it. As LANGDON (1992: 417) points out, this is a language-specific feature of many American Indian languages, which tend to be verb oriented, and sometimes even contain sentences consisting of just a single verb.

<sup>24</sup> The example is fictitious, its only aim being representing the process. I have included in it the features that I assume for common Basque, like  $\emptyset$  for indefinite direct object and VO order.

mark on the verb being the only pl. marker) are likewise correct, and the latter is actually the option preferred by speakers. In this and a number of other languages (see Stebbins' whole paper), double-coding will be used or avoided depending on the economy of the language<sup>25</sup>.

The third possibility –the rise of plural marking in the verb itself– is typical of North-American languages. Langdon (1992) reports several parallels from the Yuman languages, and Mithun (1988) from all over North America<sup>26</sup>. Let us look at a couple of examples from Inseño Chumash displayed by Mithun:

- (39) a. *s-iy-axi-* *kum*  
 3-pl.-iterative dance  
 'They are dancing'
- b. *s-iy-axi-* *kum* *ha-ku*  
 3-pl.-iterative dance the person  
 'The people are dancing'  
 [Mithun (1988: 212)]

Mithun (1988: 212) further notes: "There is no nominal source for the number specification, nothing for a copying rule to copy". And two paragraphs below: "In many of these situations, plural forms of the nouns simply do not exist". In fact, the conception of number marking in North American languages is different from the European conception. (39a, b) should not be interpreted as a plural subject whose overt marking has arisen in the verb. What *-iy-* and other verbal marks of other languages indicate has more to do with verbal aspect than with the arguments of the verbal action. What *-iy-* –along with *-axi-*– actually indicates is distributive iterativity, that is, that the action was done performed by each of the people involved in it. Similarly, in Central Pomo (Mithun 1988: 224-225), a single suffix *-ʔa-* can indicate multiplicity of subjects, of patients, and of indirect objects, depending on the context.

I think that the verbal number marking of Basque should be interpreted in terms of agreement rather than of aspect. Most 1<sup>st</sup> and 2<sup>nd</sup> person marks on

<sup>25</sup> Somewhat more complicated is the 'copying rule' that occurs in Korean (SONG, 1997). In this language, the pl. marker of the subject nominal is 'copied' and suffixed to non-subject elements including adverbs, complementisers, illocutionary markers, etc. For instance:

(ii) *ai-tul-i* *mwul(-tul)-ul* *masi-ess-ta*  
 child-PL-NOM water(-PL)-ACC drink-PST-IND  
 'The children drank water'  
 [Song (1997: 204)]

In this sentence, *mwul(-tul)-ul* is the direct object, but it does not refer to any plural or abundant notion of the water, that is, it is not an Intrinsic Plural Mark (IPM) (in fact, in a sentence like *han ai-ka mwul(\*-tul)-ul masi-ess-ta* 'one child drank water', the pl. marker *-tul* would be ungrammatical); it rather refers to the subject, so that it is an Extrinsic Plural Marker (EPM). The standard interpretation is that it is a redundant marker, used optionally and at random. But according to SONG (1997: 205), EPM is neither optional nor random, but has its own important semantic and pragmatic functions to perform. These functions are usually distributional. The underlying meaning in the sentence under discussion would be, according to Song, 'For each of the children, it was water that s/he drank'. On this issue, see also LEE (1991). Another case of 'direct object copying' is found in Burushaski (northwestern Pakistan), described in TIFFOU & PATRY (1995), who actually mention Basque as a parallel.

<sup>26</sup> DURIE (1986) lists examples from other parts of the world, and claims that cross-linguistically plural marking is not a noun category any more than verbal.

verbs are clearly of pronominal origin. As for *-tza*, the internal evidence suggests that it is more reasonable to assume the direction noun  $\rightarrow$  verb. Assuming the opposite direction would imply a ‘noun agreement’ (see Langdon 1992: 406-407) from the verb to the noun unparalleled in Basque for any other morph. Moreover, it is quite conceivable that *-(t)za* bears some kind of semantic, functional, and even phonotactic relation to the also abundant *-(t)zu*, and this is only present in the noun morphology. At any rate, it seems certain that the verbal *-(t)z(a)* indicated plurality of the abs. argument, not iterativity of the verb or any related idea. It marked plurality of entities, not of events.

It is important to note, however, that the nominal derivational *-tza* of historical periods is neither semantically nor grammatically plural; it simply implies just an uncountable abundance which is in itself morphologically sg. and requires sg. agreement on the verb (*Peru-k diru-tza dauka* ‘Peru has a lot of money’). The meaning of the verbal *-tza* of historical periods is somewhat different. In old eastern texts (see 2.3.), a finite verb requires *-(t)z(a)*—or any of the other NOR pl. marking morphs— if the corresponding abs. argument in the sentence is abs. pl. —hence morphologically definite— or alternatively if an abs. indef. argument is arithmetically plural and this arithmetical plurality is explicitly expressed<sup>27</sup>. Consequently, if the nominal *-(t)z(a)* and the verbal *-(t)z(a)* are the same morpheme in origin, some semantic shift must have taken place in one of them during prehistorical phases of the language.

At any rate, it seems to me that the NOR pl. marking on verbs (of any conjugation) must have started in the 3<sup>rd</sup> person, where it created a distinctive opposition, and later extended to 1<sup>st</sup> and 2<sup>nd</sup> person plurals, where it was redundant. This is what seems to be the development from the situation just sketched (by Trask) for the verb *egon*. To take another example with a more pan-dialectal development, the evolution of the verb *ibili* ‘to walk’ may have been something like the following:

Figure 4. Proposed evolution of number marking of the NOR morphemes of an intransitive verb (*ibili* ‘to walk’)

	Phase I		Phase II		Phase III
1 <sup>st</sup> sg.	na-bil	>	na-bil	>	na-bil
2 <sup>nd</sup> sg.	ha-bil	>	ha-bil	>	ha-bil
3 <sup>rd</sup> sg.	da-bil	>	da-bil	>	da-bil
1 <sup>st</sup> pl.	ga-bil	>	ga-bil	>	ga-bil-tza
2 <sup>nd</sup> pl.	za-bil	>	za-bil	>	za-bil-tza
3 <sup>rd</sup> pl.	da-bil	>	da-bil-tza	>	da-bil-tza

<sup>27</sup> This is the situation in Leiz., for instance, where we find (Phil. III, 17) *eta considera-it-za-çue ha-la dabil-tza-n-ac* ‘and beware (you pl.) those who walk like that’. With plural numerals in indefinite abs. arguments, the verbal agreement is always plural, as in (Acts XXI, 23) *laur guizon citiagu* ‘we have (– there are) four men’. Semantically plural indef. determiners, on the contrary, show sg. verbal agreement ((Cor. 2 IX, 2) *çue-taric heldu den zelo-ac anbitz persona incita-tu ukan du* ‘the zeal that comes from you (pl.) has incited many people’). But this is different if the abs. argument is the subject of an intransitive sentence, instead of the direct object of a transitive one, as in (Math. XXII, 14) *anbitz dirade deithu-ac* ‘many are the called ones’.

A similar development of the NOR pl. marks has probably also taken place in the rest of the conjugations. Thus, in the NOR-NORK verb *ekarri* 'to bring', there was in origin no distinction between *dakar* 's/he brings (one thing)' and *\*dakar* 's/he brings (several things)', but at a certain point the second one took on a suffix *-tza* (as already suggested, perhaps from the noun morphology), > *dakar-tza*, and this extended later to the 1<sup>st</sup> and 2<sup>nd</sup> pl. forms, so that *\*ga-kar* > *ga-kar-tza* 's/he brings us' and *\*za-kar* > *za-kar-tza* 's/he brings you (pl.)'<sup>28</sup>. The even later development whereby the 2<sup>nd</sup> pl. forms becomes sg. (> *za-kar-tza* 's/he brings you (unm.)') and a new 2<sup>nd</sup> pl. form is created by a recharacterising pl. morph (*za-kar-tza-te* 's/he brings you (pl.)') is not yet fully accomplished in some old B. texts.

A more complicated issue is the one of the forms in which the sg./pl. opposition is marked by some other processes, as in the case of the auxiliary verb of the NOR conjugation (*da* 's/he, it is' / *d-ir-a* 'they are'; see Table 6 in 1.4) or in the auxiliary verb of the NOR-NORK conjugation: *du* 's/he, it has' / *d-it-u* 'they have'). The origin of these oppositions is certainly mysterious<sup>29</sup>. As Trask (1997: 222-223; see also 2.5.1) points out, they might be remains of an older situation in which, since the sg./pl. opposition was not grammaticalised, it was expressed by various means. Especially astonishing is the opposition *du* 's/he has (NOR sg.)' / *d-it-u* 's/he has (NOR pl.)' and the rest of this type (with a pl. marking infix synchronically analysable as *-it-*). It is clear that it is old, for it appears in the oldest texts of all dialects. In fact, the infix *-it-* to mark NOR pl. appears also on auxiliary verbs whose root is *\*zan*, as in *ekar deza-gu-n* 'that we bring (subj.) (abs. sg.)' / *ekar d-it-za-gu-n* 'that we bring (subj.) (abs. pl.)'. About its origin, we can only speculate<sup>30</sup>.

<sup>28</sup> These developments have most probably taken place with intermediate stages in which the verbal plurality was a lexical, derivational process, thence an optional device. Even if, as already shown, the direction of the agreement is perhaps the reverse one, it is possible that among the Yuman languages a similar development has possibly occurred. Thus (LANGDON, 1992: 406), in the Delta-California subgroup verbal plurality is a derivational category, whereas in the Pai languages it is an inflectional one. According to Langdon, in Proto-Yuman verbal plurality was derivational, not obligatory, and it is the Pai group that has innovated.

<sup>29</sup> The segmentation *d-ir-a* and *d-it-u* is based on a purely synchronic analysis.

<sup>30</sup> CASTAÑOS GARAY (1979: 95-96) proposes that *ditu* comes from the prefixation of *e-*, the same pluraliser that according to him (see 2.1.6) was abstracted from *-e-ta*, to *du*, so that > *\*e-tu*. Then, this was closed into *\*i-tu* and finally to this form the present marker *d-* was again added. The morphological procedure that he suggests has no parallel in the Basque morphology. If we apply the regularity of transitive synthetic verbs to *du* / *d-it-u*, and consider the fact that the verbal root hiding in it is *\*dun*, *du* must come from *\*da-du* (B. *dau*, and even *nau* 's/he has me' in all dialects), and *d-it-u* should be analysed as *\*da-it-du* (*za-it-u* 's/he has you (sg.)', = East. *z-it-u*). The problem is that an infix *-it-* has a phonotactic structure that is at odds with the morpheme structure that we can guess for the old language. It seems as if *d-it-u* < *\*da-it-du* were the result of something longer which has suffered drastic erosion due to a very frequent use, as is to be expected in a verb which is not only 'to have' but the auxiliary verb of any indicative periphrastic verb of the NOR-NORK conjugation. Could that be *\*da-eta-du* or *\*da-beta-du*, where *-eta-* would be a pluraliser and the whole structure a periphrastic verbal plural, parallel to the *\*etxe-(h)eta-n* in the noun morphology? From a phonetic point of view, it would make sense. From a morphological point of view, however, it presents several problems. Among them: 1) the fact that the old verb structure in Basque should be *\*da-CVC*, with the lexematic root always in second position (although this would not be a problem for *\*da-eta-du* any more than for any theory, inasmuch as we have to set out from a structure *\*da-X-du* anyway); and 2) the fact that in all cases of *-(k)eta* (see 2.1.4) that we have seen, this element is a suffix located after the lexical root, not before it.

The NORK pl. marking morphemes are also developed at the beginning of the written tradition, both in the NOR-NORK and in the NOR-NORI-NORK conjugations, as shown in 2.5.2. As already explained, the ones of the 1<sup>st</sup> and 2<sup>nd</sup> persons are pronominal, either at the beginning or at the end of the word, except for *-da* as 1<sup>st</sup> sg. mark and *\*-ga/-na* (masc./fem.) as 2<sup>nd</sup> mark in present conjugations. These marks, which must in origin have been cliticised to the verb, do not bear any erg. suffix, unlike their corresponding erg. arguments in the main clause. This is most probably a remnant of a phase in which the erg. mark was not grammaticalised.

As for the 3<sup>rd</sup> person number marks, the 3<sup>rd</sup> sg. is always  $\emptyset$ , and the 3<sup>rd</sup> pl. *-de*, *-te* or *-e*, *\*-de* probably being the oldest form. Sadly, any discussion about the origin of this suffix immediately grows very hazy. There is another suffix, *-te*, which in modern dialects has tended to be reinforced by *-ke*, > *-teke* (*izan daiteke* ‘s/he, it can be’) and which in origin probably implied some kind of future, intemporal present or what may happen at any moment, used with intransitive verbs (Azkarate & Altuna 2001: 164 ff. and 217-219), as in Leiz. *haren beldur da-te-la badakusa* ‘if s/he sees that s/he is afraid of him/her’. Semantically, one can relate it to the modal prefix *le-* by means of a “migration” from the beginning to the end of the word; formally, one would need to assume that *\*de*<sup>31</sup> > *le-* (modal prefix), and the migrating form, > *\*-de*, took on an allomorph in devoicing contexts, > *-te*, which was eventually the only surviving allomorph. But any semantic link between this modal *-te* and the erg. pl. *-del-tel-e* would certainly be meagre.

Lakarra (2006: 604) suggests a connection of erg. pl. *-del-tel-e* with some cases of *le-* on nouns (see 2.5.2), but he has put forward a more elaborate analysis in Lakarra (2008: 475-479, 483-484). Here, he investigates the origin of the preterite ending *-en*, as in *nen-torr-en* ‘I came’, *ze-karr-en* ‘s/he brought’, etc., and based on compounds like *ats-eden* ‘rest, break’ (lit., ‘rest (*eden*) for breath’) indicates that *eden* ‘to interrupt, to finish’ (with perfective *\*ede-ten*, from which > *ete-ten* with assimilation, and from here > *eten*, another existing variant with the same meaning) might have been the outset of the preterital *-en*. Thus, clusters like *\*thor-* + *-(d)en*, *\*khus-* + *-(d)en*, *\*khar-* + *-(d)en* (that is, the verbal roots of ‘to come’, ‘to see’, ‘to bring’, respectively, + ‘to finish’) would be, in origin, perfective periphrases, which over the course of time –along with the transformation of Basque from an aspect-sensitive language to a tense-sensitive language– became preterital forms<sup>32</sup>. If this were so, this *-den* would have dropped the *-d-* as the perfective (> pret.) marker, but Lakarra suggests that an allomorph of it, *\*-de(n)*, may have specialised in the 3<sup>rd</sup> p. erg. pl. marking meaning. There are parallels of lexemes implying ‘to finish’ which become 3<sup>rd</sup> pl. markers, as Zavala (2006) indicates

<sup>31</sup> According to LAKARRA (2006: 604), from the agglutination of *\*da-* + *-e*, that is, just like *\*e-* + *-da-* (which has given rise to causative verbs like *jausi* ‘to fall’, < *\*e-da-dus-i*) but the other way round.

<sup>32</sup> These clusters would be, in origin, serial verb constructions. Similarly, present forms like *da-tor* ‘s/he comes’, *da-kar* ‘s/he brings’, etc., would come from *\*dar-tor*, *\*dar-kar*, etc., and would be serial verbs with imperfective meaning, *\*dar-* being the radical form of the verb *jarri* ‘to put’. That verbs of position, usually close to ‘to put’, give rise to imperfective markers and verbs implying ‘to finish’ give rise to perfective markers, is abundantly attested, and particularly in languages with serial verbs (see LAKARRA, 2008, for plenty of references).



in reference to the Mixe-Zoquean family, where it is apparently an areal feature.

Be it as it may, whatever origin we propose for the erg. pl. mark *-del-tel-e*, there are two key findings that indicate a relatively recent formation of the opposition sg./pl. in 3<sup>rd</sup> person erg. verbal marks. One of them is the fact that the opposition sg./pl. in the 3<sup>rd</sup> person erg. marks is expressed by *-Ø* vs. an overt mark. The other one is the presented analogy with other sectors of the grammar, in which the secondary origin of the number opposition is not hypothetical but consistent. Together, these make the assumption of a relatively recent formation –at any rate more recent than the agglutination to the verb of 1<sup>st</sup> and 2<sup>nd</sup> erg. marks– quite possible. If we take as an example the present tense of a normal synthetic NOR-NORK verb (*ekarri* ‘bring’), the evolution might be something like the one in the following figure:

Figure 5. Proposed evolution of number marking of the NORK morphemes of a transitive verb (“*ekarri*” ‘to bring’)

	<i>Phase I</i>		<i>Phase II</i>
1 <sup>st</sup> sg.	*da-kar-da	>	*da-kar-da
2 <sup>nd</sup> sg.	*da-kar-ga/-na	>	*da-kar-ga/na
3 <sup>rd</sup> sg.	da-kar	>	da-kar
1 <sup>st</sup> pl.	da-kar-gu	>	da-kar-gu
2 <sup>nd</sup> pl.	da-kar-zu	>	da-kar-zu
3 <sup>rd</sup> pl.	da-kar	>	da-kar-te (< *-de)

Unlike in the NOR morphemes, in this case the pl. mark has not extended to the 1<sup>st</sup> and 2<sup>nd</sup> pl. forms with a pl. recharacterising function. It has extended, however, to the 2<sup>nd</sup> pl. form when *dakar-zu* has become 2<sup>nd</sup> sg. of politeness, whereby > *dakar-zue*. But this development must have been relatively recent.

In the 1<sup>st</sup> and 2<sup>nd</sup> persons, the NORI marking morphemes are the same as the NORK ones of present tense conjugations, these being after the verbal root but before the NORK morpheme, if there is one. But in the 3<sup>rd</sup> person, the situation is different from the NOR and the NORK forms. Unlike these, the 3<sup>rd</sup> person sg. NORI mark is overt, most commonly *-o*<sup>33</sup>. Let us suppose that this comes from (*h*)*aur* ‘this’ (see 2.5.3)<sup>34</sup>. This would imply that

<sup>33</sup> I shall not further discuss the controversy of whether the variant *-a* comes from the demonstrative *\*har* or is a morpho-phonological variant of *-o*, as it does not crucially affect the issue of number (see 2.5.3).

<sup>34</sup> This supposition raises several interesting questions. Firstly, the morphology of this mark *haur* is a NOR morphology, as we can see if we compare it to the demonstrative from which it comes (abs. *hau*, *haur* in the east, vs. dat. *honi*). This implies that at the time of this cliticising process, the NORI marking procedures were not grammaticalised, at least on cliticised elements. These conserve, undoubtedly, an archaic morphology previous to the origin of the historical diathetical system, and most probably to the ergative system as this appears in historical periods, for we should notice that cliticised NORK marks bear no erg. marker either (*dakar-gu* ‘we bring’, not *\*\*dakar-guk*). This does not mean that during this particular stage in time Basque was not at all ergative, but that it was perhaps less ergative than in historical phases, when Basque seems to be almost exclusively ergative. As it is widely known, split ergative systems appear in a number of languages, whereby a part of the morphology is abs./erg. and another part is nom./acc. DIXON (1994: 85) presents a hierarchy concerning all this based on logic and typological data. According to Dixon, pronouns

the number marking is of pronominal origin on the 1<sup>st</sup> and 2<sup>nd</sup> p. morphs, and of demonstrative origin on the 3<sup>rd</sup> p. morphs. In any case, if the 3<sup>rd</sup> p. sg. morph was overt, we may assume that in origin –that is, when the number opposition had not extended to 3<sup>rd</sup> p. verbal marks yet– the 3<sup>rd</sup> p. pl. morph was the same. During this phase, then, there would be no number distinction in the 3<sup>rd</sup> p. This would have come about later, when the pl. marker *\*-de* was brought from the NOR/NORK marks. We can exemplify this with the following NOR-NORI conjugation (‘he walks to me, you (sg.), him/her/it...’):

Figure 6. Proposed evolution of number marking of the NORI morphemes of an intransitive verb (*ibili* ‘to walk’)

	Phase I		Phase II
1 <sup>st</sup> sg.	*da-bil-ki-da	>	*da-bil-ki-da
2 <sup>nd</sup> sg.	*da-bil-ki-ga/-na	>	*da-bil-ki-ga/na
3 <sup>rd</sup> sg.	*da-bil-ki-haur (> -o)	>	da-bil-ki-o
1 <sup>st</sup> pl.	da-bil-ki-gu	>	da-bil-ki-gu
2 <sup>nd</sup> pl.	da-bil-ki-zu	>	da-bil-ki-zu
3 <sup>rd</sup> pl.	*da-bil-ki-haur (> -o)	>	*da-bil-ki-o-de (> -e)

To sum up, I contend that the whole process of verbal development happened approximately as follows. At the beginning, the verb was impersonal, that is, it had no attached person mark corresponding to any of the arguments of the clause, and therefore indicated no number opposition. The verb was the sheer verbal root, sometimes with prefixes and/or suffixes of different (probably aspectual) meanings, which according to Lakarra (2006, 2008) must have, in origin, been conglomerations of serial verbs. The lack of person marks on the verb is most typical of isolating languages, or at least of languages with isolating characteristics, like Chinese or English<sup>35</sup>. In a later phase, perhaps in topicalised constructions, pronoun marks corresponding to 1<sup>st</sup> and 2<sup>nd</sup> p. arguments were cliticised to the verbal complex, either as prefixes or as suffixes, in the way explained throughout the paper (see 1.5, 2.5, 3.5). Number marking during this phase, then, was suppletive for 1<sup>st</sup> and 2<sup>nd</sup> p. arguments, but non-existent for 3<sup>rd</sup> p. arguments, which were not marked on the verb, except perhaps if that argument was an indirect object<sup>36</sup>. In this case, *\*(h)au(r)* (> -o-) was suffixed. The morphology of this demonstrative

are the elements least prone to take ergative markers, and this observation is lucid in Basque data too. Another interesting question is why, unlike the corresponding abs. and erg. marks –which are covert–, the dat. verbal mark is overt in the 3<sup>rd</sup> p. The reinforcement of a dat. to distinguish it from the acc. is, with different peculiarities in each case, common to several well known languages (to Spanish and English, for instance: *tengo un perro* ‘I have a dog’, but *pongo a un perro su collar* ‘I put the collar on the dog’; *I gave him the book*, but *I gave it to him*). But what we find in Basque is not an overt abs. versus a reinforced overt dat., but an overt abs. (and erg.) versus a covert dat. I attribute this to the fact that the finite verbal complex probably arose in topicalised constructions. Now, since subjects and direct objects –subjects of passive sentences– are typically more prone to be topicalised than indirect objects –subjects are the topics *par excellence* –, the covert marking (-Ø mark) of NOR and NOR/NORK 3<sup>rd</sup> p. sg. morphemes on the verbs is not unexpected, for they would be redundant. On the other hand, indirect objects are somewhat rarer in the topicalised position, making the presence of a mark that represents it on the verb linguistically relevant.

<sup>35</sup> English has lost every person mark, except for the 3<sup>rd</sup> p. sg. -s in the present.

<sup>36</sup> The structure of the finite verb at this point resembles in some respects that of Seri (Mexico), as described by MARLETT (1990). In this language, subject person agreement on finite verbs is signalled

seems to indicate that at the beginning of the cliticising period, the diathetical system that is present on the verb in historical phases was not yet complete. If these were simply topicalised constructions, it might be that in standard clauses verbs remained impersonal (root plus affixes with no argument mark).

In the course of time, these topicalised constructions became more frequent and less marked, so that the clitics became totally morphologised bound markers and the verbs of topicalised clauses –with argument marks– became the typical in standard clauses. Allocutive morphs were adjoined to the verb, too<sup>37</sup>. At this point, there was no number marking for 3<sup>rd</sup> p. marks (which were overt only for indirect objects), and in fact the language as a whole had no number category. As pl. marks began to grammaticalise –most probably in the noun morphology first–, 3<sup>rd</sup> p. abs. and erg. pl. markers appeared on verbs too, perhaps extended from nouns. The abs. pl. marker was *-za-* –surely with some different variants– and the erg. pl. marker was *-de* (see 2.5.1 and 2.5.2). The former spread finally from 3<sup>rd</sup> p. abs. pl. forms to 1<sup>st</sup> and 2<sup>nd</sup> p. abs. pl. forms, where it was redundant but created a regularity in the system. On the contrary, the latter did not spread to 1<sup>st</sup> and 2<sup>nd</sup> p. erg. pl. forms, but it extended to 3<sup>rd</sup> p. dat. forms to indicate plurality. Leaving aside the transformation of the old 2<sup>nd</sup> pl. form into a 2<sup>nd</sup> sg. form of politeness and the consequences of this shift, this account more or less represents the verbal form situation at the beginning of the Basque written tradition, when the number category seems to be fully developed, but not long in use.

by the set of prefixes on the left-hand column (1990: 514), and direct object person agreement by those on the right-hand column (1990: 521):

(iii)

	<u>Subject person agreement</u>	<u>Direct object person agreement</u>
1 <sup>st</sup> sg.	ʹ, ʹp-	ʹim-, ʹpo- ( <i>ipv. allomorph</i> )
2 <sup>nd</sup> sg.	m-	ma-
1 <sup>st</sup> pl.	ʹa-	ʹiši-
2 <sup>nd</sup> pl.	ma-	maši-

On the contrary, 3<sup>rd</sup> person marks are covert in both cases. The diathetical system of Seri is extremely complex. Some further similarities with Basque can be identified. In spite of its nom./acc. syntax, the language is sometimes sensitive to the final transitivity of the clause. Thus, the allomorphy of the 1<sup>st</sup> p. sg. prefix for subjects is suppletive: ʹ occurs in finally transitive clauses (1<sup>st</sup> sg. erg.), and ʹp- in finally intransitive clauses (1<sup>st</sup> sg. abs.). On the other hand, Seri has one more set of agreement prefixes, referred by MARLETT (1990: 524) as oblique person agreement prefixes. In some cases (1990: 527), they signal agreement with nominals which are more or less agreement indirect objects. Their forms are:

(iv)

	<u>Oblique person agreement</u>
1 <sup>st</sup> sg.	ʹe-
2 <sup>nd</sup> sg.	me-
3 <sup>rd</sup> sg.	ko-

In Basque dat. (NORI) markers are suffixed, but it is remarkable that in this case the 3<sup>rd</sup> p. mark is overt in both cases, unlike in the other categories. In Seri, these oblique marks are identical in shape with the independent pronouns. Stebbins (1997: 32) argues that this lack of number distinctions in oblique verbal agreement marks of Seri has to do with the fact that referents which occur as oblique have the least salient role in the discourse.

<sup>37</sup>The time at which these were incorporated in the verbal complex is a much discussed issue (see TRASK, 1997: 234-236).

This is, of course, a speculative synthesis. I would not like to be too dogmatic about the topicalised origin of the verb<sup>38</sup>. If, as suggested by Gómez & Sainz (1995), the mark *ze-* present in such forms as *ze-go-(e-)n* 's/he was', *ze-karr-(e-)n* 's/he brought', etc., was a sentence connector comparable to the historical conjunction *ezen* 'for, so that', then its corresponding verbal phrase may in fact be in a relationship of sheer coordination with the preceding one. This is an issue which undoubtedly deserves further investigation. Things would not differ much, though, in relation to number. That Basque must have at some point had an impersonal structure is a hypothesis that I personally find to have a very solid basis, independently of this particular issue.

A more relevant thing is the fact that, as mentioned above, contrary to what happens in plenty of languages, in Basque there seems to be no relation between plurality of verb and aspect<sup>39</sup>. A major device in this respect is reduplication (Moravcsik 1978). In Basque, there is apparently no reduplication of the verbal root to mark plurality of action. However, Lakarra (2008: 455 ff.) has suggested that for a series of words a double *-da-/ra-* element can be reconstructed: RS. *jaraunsi* 'to inherit' < \**e-da-ra-dun-ts-i*, *urgatzi* 'to help' < \**e-da-ra-gotz-i*, amongst others. Since *-ra-* is probably the same element as \**-da-*, this might imply a reduplication of the aspect mark. If the order of these elements were old and could be brought back to the serial verbal stage, we could, in origin, have a reduplication of the imperfective aspect-implying verb \**e-darr-i* 'to put', to form verbal constructions of the type \**-dar-dar-CVC* (as suggested by Lakarra 2008: 471-475).

#### 4. SOME FINAL REMARKS

To finish off this outline of the diachronic analysis of number, let us compare the hypothesis reached here to the Animacy Hierarchy in relation to number that was proposed by Smith-Stark (1974), an updated version of which is given in Corbett (2000: 56):

*Figure 7. Animacy Hierarchy in relation to number marking:*

speaker > addressee > 3<sup>rd</sup> person > kin > human > animate > inanimate  
 (1<sup>st</sup> p.            (2<sup>nd</sup> p.  
 pronouns)    pronouns)

This means that, in a given language, if a particular segment of the series shows sg./pl. distinction, the segments that are further up in the Hierarchy

<sup>38</sup> One of the several things that let me think that historical finite verbs are old topicalised clauses is the fact that most person marks on them are bound affixes, formally different from independent pronouns, hence perhaps clitics in origin. This is always so with NOR marks and past tense NORK marks (*ni* 'I' - *ne(n)-*, *gu* 'we' - *gen-*, etc.). Even with NORI marks and present tense NORK marks, the marks of sg. persons are different from pronouns (*ni* - *-da*, *hi* 'you (sg.)' - \**-gal/\*-na*).

<sup>39</sup> In the Seri verbs just mentioned (MARLETT, 1990: 508-509), pl. marks can refer to the subject or to the action. The sg./pl. dichotomy can therefore generate four possibilities. If we take the verb 'to arrive' as an example: *-afp* (sg. subject, sg. action); *-afap-im* (sg. subj., multiple act.); *-aşk-am* (pl. subj., mult. act.); *-aşixk-am* (pl. subj., mult. act.). As in Basque, even if the 3<sup>rd</sup> p. sg. mark is covert, the corresponding pl. is overt.

will also show sg./pl. distinction. Thus, languages with plural opposition on the lowest level of the Hierarchy (inanimate nouns), like most European languages, will be expected to have the same opposition on the other levels as well<sup>40</sup>. Among languages with a split (grammaticalised number on some levels but not on others), the question is where this split is. It can, for example, be between inanimates and animates, as in Marind (in southern Irian Jaya) (Corbett 2000: 59, from Foley 1986: 78, 82-83). A number split located between animates and humans, or even between human non-kins and kins, is present in many North American languages (Corbett 2000: 57-59, Mithun 1988: 212). A split between kins and pronouns occurs in Chinese (Corbett 2000: 61-62, Norman 1988: 159), but there are also languages with grammaticalised number in some pronouns but no such an opposition in others. Among these, in a few languages the split is between the 3<sup>rd</sup> person on one side and the 1<sup>st</sup> and 2<sup>nd</sup> persons on the other. This happens in Asmat and in Mandarin Chinese, as I showed in 3.4., though the latter is on the border of not having grammaticalised number at all, not even on pronouns, for on these the pluralising procedure is performed by means of suffixation, and the suffix is not obligatory (see 3.4). It also happens in Coast Tsimshian (Stebbins 1997: 31, from Mulder 1994: 64)<sup>41</sup>.

It seems to me, then, that old Basque may have at some point been a language of the type with number split between the 3<sup>rd</sup> person and the 2<sup>nd</sup> person, i.e. a language with (suppletive) number marking on 1<sup>st</sup> and 2<sup>nd</sup> pronouns and no distinction in the segments to the right. If this supposition is accurate or very close to the original situation, the expected development from then on would be a later creeping extension of the number grammaticalisations into segments further to the right. A drawback of this line of argumentation, though, is that according to it we should expect the segment next to 2<sup>nd</sup> p. pronouns to acquire number marking; that is 3<sup>rd</sup> p. pronouns (see 3.2). In the case of Basque, this equates to an expectation of number marking on deictics and anaphoric elements. I have already discussed, however, that internal criteria might point to an earlier number distinction on nouns rather than on demonstratives.

Last but not least, historical attestations have not left any trace of a situation of split as suggested by the hierarchy, such as a phase where, say, animates were marked for plural but inanimates were not, or were only optionally. From the beginning of the tradition, pl. NPs are marked as such irrespective of whether they are human, animates or inanimates. Nonetheless, if we consider the possible relationship between the erg. \*-*ga* and the pl.

<sup>40</sup> Some typologists include more parameters in the hierarchy. STEBBINS (1997: 28 ff.), for instance, mentions the case hierarchy subject < object < oblique (see CROFT, 1990), and insists on the idea that in many languages high pragmatic salience increases the possibility of finding number marking. In Basque, however, it seems not to have played any role whether a particular element was subject, object or oblique to number-marking.

<sup>41</sup> The discussion about number splits between 2<sup>nd</sup> person and 1<sup>st</sup> person pronouns is particularly controversial (CORBETT, 2000: 64-66, SMITH-STARK, 1974), but it seems that in languages with no number distinction over the different segments, the first person pronoun tends to be the segment which most easily develops some mechanism towards the grammaticalisation of the sg./pl. opposition (for one case of such a split, see FOLEY, 1986: 70). Finally, some languages seem to have no number opposition at all, not even in any pronouns (CORBETT, 2000: 50-51). Two of these are Pirahã (of the Mura family, in Amazonas) and Kawi (Old Javanese).

\*-ga, and the possibility that erg. \*-ga started off as an animate marker for animates to form local cases (see Lakarra 2005: 442-444), it might not be unconceivable that stem-article-ga constructions appeared on animates earlier than on inanimates. Constructions of the type stem-eta-case.mark, on the contrary, are typical of inanimates, even if they also appear on animates within the oldest texts (see 3.1). Thus, unless more attestations appear in the future, we can only speculate about how number marking has crept to the right from the ‘transnumeral’ phase.

## REFERENCES

- AIKHENVALD, A. Y., 1994, “Classifiers in Tariana”, *Anthropological Linguistics* 36, 407-465.  
 — & DIXON, R. M. W. (ed.), 2006, *Serial verb constructions. A cross-linguistic typology*, Oxford University Press.
- ANAGBOGU, P. N., 1995, “The Semantics of Reduplication in Igbo”, *Journal of West African Languages* 25, 43-52.
- ARTIAGAGOITIA, X. & LAKARRA, J. A. (eds.), 2008, *Gramatika jaietan Patxi Goenagaren omeñez*, San Sebastián/Bilbao, Appendix of ASJU-51.
- AZKARATE, M. & ALTUNA, P., 2001, *Euskal morfologiaren historia*, San Sebastián, Elkarlanean.
- CASTAÑOS GARAY, F., 1979, *La formación del plural en la Lengua Vasca*, Bilbao, Publicaciones de la Diputación Foral del Señorío de Vizcaya.
- CORBETT, G. G., 2000, *Number*, Cambridge University Press.
- CREISSELS, D., 2006, *Syntaxe générale. Une introduction typologique. Volume 1*, Paris, Hermès.
- CROFT, W., 1990, *Typology and universals*, Cambridge University Press.
- , 2000, “Parts of speech as language universals and as language-particular categories”, in Vogel eta Comrie (eds.), *Approaches to the Typology of Word Classes*, Berlin-New York, Mouton de Gruyter, 65-102.
- DIFFLOTH, G., 1994, “*i*: big, *a*: small”, in Hinton-Nichols-Ohala (eds.), *Sound Symbolism*, Cambridge University Press, 107-114.
- DIXON, R. M. W., 1994, *Ergativity*, Cambridge University Press.
- DURIE, M., 1986, “The grammaticalization of number as a verbal category”, *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society. February 15-17, 1986*, Berkeley Linguistics Society, 355-370.
- EUSKALTZAINDIA (ed.), 1981, *Euskalarien nazioarteko jardunaldiak. Iker-1*, Bilbao, Euskaltzaindia.
- , 1987-2005, *Orotariko Euskal Hiztegia (= OEH)*, Bilbao, Euskaltzaindia.
- FOLEY, W. A., 1986, *The Papuan Language of New Guinea*, Cambridge University Press.
- GAVEL, H., 1920, *Éléments de Phonétique Basque Paris (= Revista Internacional de Estudios Vascos 12)*.
- GÓMEZ, R. & SAINZ, K., 1995, “On the origin of the Finite Forms of the Basque Verb”, in Hualde, Lakarra, Trask (eds.), 235-274.
- GREENBERG, J. *et alii* (ed.), 1978, *Universals of human language, III*, Stanford University Press.
- HUALDE, J. I., LAKARRA, J. A., TRASK, R. L. (eds.), 1995, *Towards a History of the Basque Language*, Amsterdam, John Benjamins.
- IRIGOIEN, A., 1981, “Haur ola çirola”, in Euskaltzaindia (ed.), 365-402.
- ITURRIOZ, J. L., 1985, “La función de -a y de -ta a la luz de la dimensión de individuación”, *Euskera* 30, 175-213.
- KIMBALL, G., 1990, “Noun pluralization in Eastern Huasteca Nahuatl”, *IJAL* 56, 196-216.
- LAFON, R., 1999 [1948], “Le nombre dans la déclinaison basque”, in Lafon (1999), 209-217.
- , 1999, *Vasconiana (= Iker-11)*, Bilbao, Euskaltzaindia.
- LAKARRA, J. A., 1995, “Reconstructing the Pre-Proto-Basque Root”, in Hualde, Lakarra, Trask (eds.), 189-206.

- , 2005, “Prolegómenos a la reconstrucción de segundo grado y al análisis del cambio tipológico en (proto)vasco”, *Paleohispanica* 5, 407-470.
- , 2006, “Notas sobre iniciales, cambio tipológico y prehistoria del verbo vasco”, in Lakarra & Hualde (eds.), 561-621.
- & Hualde, J. I. (eds.), 2006, *Studies in Basque and historical linguistics in memory of R.L. Trask*, San Sebastián/Bilbao, Appendix of ASJU-40.
- , 2008, “Aitzineuskararen gramatizarantz (malkar eta osinetan zehar)”, in Artiagaitia & Lakarra (eds.), 451-490.
- LANGDON, M., 1992, “Yuman plurals: from derivation to inflection to noun agreement”, *IJAL* 58, 405-424.
- LEE, H.-G., 1991, “Plural marking copying in Korean”, *Studies in Linguistic Sciences* 21, 81-105.
- LI, Ch. N. & THOMPSON, S. A., 1981, *Mandarin Chinese: a functional reference grammar*, Berkeley, University of California Press.
- MANTEROLA, J., 2007, “-a morfemaren erabilera (eza) ekialdeko euskaretan (The (non) use of the -a morpheme in Eastern Basque dialects)”, *Proceedings to the 2<sup>nd</sup> Conference of the Koldo Mitxelena Chair* (celebrated in October 2007) (draft, in press, available at <http://artxiker.ccsd.cnrs.fr/>).
- , 2008, “Is Basque an agglutinative language? A proposal for the diachrony of nominal morphology”, *Proceedings of the Basque Studies Symposium, May 14<sup>th</sup>, 2008*, to be published in *TINTA, Research Journal of Hispanic and Lusophone Studies, Department of Spanish and Portuguese, UCSB* (draft, in press, available at <http://artxiker.ccsd.cnrs.fr/>).
- MARLETT, S. A., 1990, “Person and number inflection in Seri”, *IJAL* 56, 503-541.
- MARTÍNEZ ARETA, M., 2008, *El consonantismo proto-vasco*, University of the Basque Country.
- , 2009a, “The category of number in Basque: I. Synchronic and historical aspects”, *Fontes Linguae Vasconum* 110, 63-98.
- , 2009b, “The historical reconstruction of proto-Basque: II. Historical morphology of the noun”, *Mikroglottika. 2008-09 Annual Volume* (in press).
- MITHUN, M., 1988, “Lexical categories and the evolution of number marking”, in Hammond, Noonan (eds.), *Theoretical Morphology: Approaches in Modern Linguistics*, San Diego Academic Press, 211-234.
- MITXELENA, K., 1977 [1961], *Fonética histórica vasca*, Appendix of ASJU-4. San Sebastián, Diputación Foral de Guipúzcoa.
- , 1988 [1964], *Sobre el pasado de la lengua vasca*, in Mitxelena (1988), 1-73.
- , 1987 [1970], “Nombre y verbo en la etimología vasca”, in Mitxelena, 1987, 283-309.
- , 1987 [1971], “Toponimia, Léxico y Gramática”, in Mitxelena, 1987, 141-167.
- , 1987, *Palabras y textos* (collection of articles), University of the Basque Country.
- , 1988, *Sobre Historia de la Lengua Vasca, II*, San Sebastián, Appendix of ASJU-10, San Sebastián, University of the Basque Country.
- MORAVCSIK, E. A., 1978, “Reduplicative constructions”, in J. Greenberg *et alii* (ed.), *Universals of human language*, III, 297-334.
- MULDER, J. G., 1994, *Ergativity in Coast Tsimshian (Sm'algyax)*, Berkeley, University of California Press.
- NORMAN, J., 1988, *Chinese*, Cambridge University Press.
- RIJK, R. P. G. de, 1992, “‘Nunc’ Vasconice”, *ASJU* 26, 695-724.
- SARASOLA, I., 1990 [1983], *Contribución al estudio y edición de textos antiguos vascos*. Appendix of ASJU-11, San Sebastián, University of the Basque Country.
- SMITH-STARK, T. C., 1974, “The plurality split”, in La Galy, Fox, Bruck (eds.), *Papers from the Tenth Regional Meeting, Chicago Linguistic Society, April 19-21*, Chicago Linguistic Society, 657-671.
- SONG, J. J., 1997, “The so-called Plural Copy in Korean as a marker of distribution and focus”, *Journal of Pragmatics* 27, 203-224.
- STEBBINS, T., 1997, “Asymmetrical nominal number marking: a functional account”, *Sprachtypologie und Universalienforschung* 50, 5-47.
- TIFFOU, E. & PATRY, R., 1995, “La notion de pluralité verbale: le cas du Bourouchaski du Yasin”, *Journal Asiatique* 283, 407-444.
- TRASK, R. L., 1977, “Historical Syntax and Basque Verbal Morphology: Two Hypotheses”, in Douglass, W. A.; Etulain, R. W.; Jacobsen, W. H. (eds.), *Anglo-American Contribu-*

- tions to Basque Studies: Essays in Honor of Jon Bilbao, Reno, University of Nevada, 203-217.
- , 1979, “On the origins of ergativity”, in F. Plank (ed.), *Ergativity. Towards a theory of grammatical relations*, London Academic Press, 305-404.
- , 1995, “On the History of the Non-Finite Verb Forms in Basque”, in Hualde, Lakarra, Trask (eds.), 207-234.
- , 1997, *The History of Basque*, London, Routledge.
- VOORHOEVE, C. L., 1965, *The Flamingo Bay Dialect of the Asmat Language* (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 46), 's-Gravenhage, M. Nijhoff.
- ZAVALA, R., 2006, “Serial verbs in Olutec (Mixean)”, in Aikhenvald & Dixon (ed.), 273-300.

## LABURPENA

### *Numeroaren kategoria euskaraz: II. Alderdi aurre-historikoak eta tipologikoak*

Euskararen numero gramatikalari buruzko artikulua bigarren eta azken zati honetan, datu historikoetatik abiatuz, garai aurre-historikoetan egoera nolakoa izan zitekeen eztabaidatzen da. 3. atalean, morfologiaren atal guztiak aztertzen dira, hurrenez hurreneko azpi-ataletan: izena, erakusleak, izenordain mugagabeak, pertsona-izenordainak eta aditz-morfemak. Lehenengo bietan, izen-pluralaren jatorriaren arazoa xeheki deskribatzen da, posibilitate bakoitzaren aldeonak eta eragozpenak zerrendatuz. 3.4. azpi-atalean, zenbait pertsona-izenordainen eta zenbait plural-morfemaren arteko harreman formalak adierazten dira, eta 3.5.-an pertsonako aditz-marken izenordainezko jatorria numero gramatikalaren kategoriaren ikuspegitik jorratzen da. 4. atalean, iritsitako ondorioa –1. eta 2. pertsonako izenordainak izan ezik, euskara zaharrak numero gramatikalik ez zuela, alegia– hizkuntz tipologiak numeroaren inguruan aurkeztutako hierarkia inplikazionalarekin aurrez aurre jar-tzen da.

**Gako hitzak:** numeroaren kategoria, numero gramatikala, singularra, plurala, mugagabea, euskara.

## RESUMEN

### *La categoría de número en euskera: II. Aspectos prehistóricos y tipológicos*

En esta segunda y última parte del artículo sobre el número gramatical en euskera, se parte de los datos históricos para debatir cómo podría ser la situación en época prehistórica. En el apartado 3 son analizados, en cinco subapartados sucesivos, nombres, demostrativos, pronombres indefinidos, pronombres personales y morfemas verbales. En los dos primeros, se describe en detalle el problema del origen del plural nominal, enumerándose los pros y contras de cada posibilidad. En el subapartado 3.4 se incide en las relaciones formales entre algunos pronombres personales con ciertos morfemas de plural, y en 3.5 se analiza el origen pronominal de las marcas verbales de persona desde el punto de vista de la categoría del número gramatical. En el apartado 4 se confronta la conclusión alcanzada –la ausencia de número gramatical en euskera antiguo, salvo en los pronombres de 1ª y 2ª persona– con la jerarquía implicacional en torno al número aportada por la tipología lingüística.

**Palabras clave:** categoría de número, número gramatical, singular, plural, indefinido, euskera.



## RÉSUMÉ

*La catégorie du numéro en basque: II. Aspects préhistoriques et typologiques*

Dans cette deuxième et dernière partie de l'article sur le numéro grammatical en basque, on part de données historiques pour analyser sa situation dans la préhistoire. Dans l'aparté 3, des noms, des démonstratifs, des pronoms indéfinis, des pronoms personnels et des morphèmes verbaux, sont analysés en cinq sous-apartés successifs. Dans les deux premiers, on décrit dans le détail le problème de l'origine du pluriel nominal, en énumérant les pour et les contre de chaque possibilité. Dans le sous-aparté 3.4 ont insisté sur les rapports formels entre certains pronoms personnels avec certains morphèmes de pluriel, et dans le 3.5 on analyse l'origine pronominale de marques verbales de personne du point de vue de la catégorie du numéro grammatical. Dans l'aparté 4, on compare la conclusion obtenue –l'absence de numéro grammatical en basque ancien, sauf dans les pronoms de la 1<sup>ère</sup> et 2<sup>ème</sup> personne– avec la hiérarchie implicationnelle autour du numéro apporté par la typologie linguistique.

**Mots clé:** catégorie du numéro, numéro grammatical, singulier, pluriel, indéfini, basque.

## ABSTRACT

*The category of Number in Basque: II. Prehistorical and typological aspects*

In this second and last part of the article about the grammatical number in Basque, it is discussed what the state of affairs in prehistorical periods may have been like, taking historical data as a starting point. In Point 3, the different sectors of the morphology are analysed in successive subpoints: nouns, demonstratives, indefinite pronouns, personal pronouns and verbal morphs. In the first two of them, the problem of the origin of the plural is described in detail, listing pros and cons of each possible interpretation. In 3.4, several formal similarities between some personal pronouns and certain plural morphs are pointed out, and in 3.5 the pronominal origin of the verbal marks of person are analysed from the point of view of the category of grammatical number. In Point 4, the conclusion reached –i.e., the absence of grammatical number in old Basque, except for 1<sup>st</sup> and 2<sup>nd</sup> person pronouns– is confronted with the implicational hierarchy about number put forward by typological research.

**Key words:** category of number, grammatical number, singular, plural, indefinite, Basque.