Noun Morphology

Ekaitz Santazilia

Please, cite this work as follows:

Santazilia, E. (2013). "Noun Morphology" In M. Martínez-Areta (Ed.), *Basque and Proto-Basque: Language-Internal and Typological Approaches to Linguistic Reconstruction* [= Mikroglottika 5] (pp. 223-281). Frankfurt am Main: Peter Lang.

6. Noun Morphology

Ekaitz Santazilia
(Public University of Navarre, UPNA/NUP
/ University of the Basque Country, UPV/EHU)

6.1. Introduction[*]

This chapter deals with the functioning and grammaticalization of the Basque case system, including the marking of determination and number in the noun phrase. First, I present a synchronic description of that system, and later I analyze the data from a diachronic perspective, summarizing the different hypotheses proposed by Bascologists, adding to these the latest contributions, and giving an account of the most recent avenues for research.

6.2. Description of today's Basque noun declension system

The Basque noun phrase presents a canonical structure which can be linearly represented in the way shown in Figure 6.1.[1]

Det.1   Noun   Adjective   Det.2   Number   Case[2]

Figure 6.1.   Canonical structure of the Basque NP

As can be observed in (1a), it is the whole noun phrase that is marked for number and case in Basque, the morpheme being attached to the end of the

[1]   In this section, I am following Lafon (1999 [1970]), Trask (*HB*) and Martínez-Areta (2009a).

[2]   From the point of view of most morphologists and syntactic theoreticians, this formula would not really be linear, nor would all of its elements be interconnected in the same manner. For the sake of clarity, however, let me be allowed to resort to this simplification.

whole NP, unlike in languages like Latin (1b), which mark each element of the phrase:

(1) a. *Bi*   *gizon*   *handi*   *hori-ei*

         two   man   big   that-DAT.PL

         'To those two big men'

     b. *Ist-is*   *duo-bus*   *magn-is*   *homin-ibus*

         that-DAT.PL   two-DAT.PL   big-DAT.PL   man-DAT.PL

         'To those two big men'

Basque has a single declensional paradigm, which is applied to all nouns (and adjectives) alike, even if the union between noun and declensional morpheme triggers some phonological changes. Thus, a vowel-ending noun presents the paradigm shown in Table 6.1 in Standard Basque.[3]

Table 6.1.   Noun declension of vowel-ending nouns

etxe 'house'	Definite		Indefinite
	Sg.	Pl.	--
Absolutive	*etxe-a-Ø*	*etxe-ak-Ø*	*etxe-Ø*
Ergative	*etxe-a-k*	*etxe-ek*	*etxe-k*
Genitive	*etxe-a-r-en*	*etxe-en*	*etxe-r-en*
Dative	*etxe-a-r-i*	*etxe-ei*	*etxe-r-i*
Instrumental	*etxe-a-z*	*etxe-ez*	*etxe-z*
Inessive	*etxe-an*	*etxe-eta-n*	*etxe-ta-n*
Allative	*etxe-ra*	*etxe-eta-ra*	*etxe-ta-ra*
Ablative	*etxe-tik*	*etxe-eta-tik*	*etxe-ta-tik*

Although not in an entirely precise manner, by and large definite forms are used whenever we refer to concrete nouns which are known to both the speaker and the addressee, whereas the indefinite paradigm is reserved for elements which are not determined or not concrete:

---

3   There exist dialectal differences in the Basque declension, but as Lafon (1999 [1970]: 167) recalls: "Dès les plus anciens textes, c'est-à-dire dès XVIe siècle, elle est la même, ou à peu près la même, qu'aujourd'hui", with notable exceptions, such as the soc. *-gaz* and the abl. *-rean* in the west, or the dat. pl. *-er* in the east (Michelena 1987 [1981a]: 301).

(2) a. *Bi*   *etxe-Ø*   *eros-i*   *d-it-u-t*

        two   house-ABS.INDEF   buy-PFV   PRS-ABS3PL-ROOT-ERG1SG

        'I have bought two houses'

    b. *Bi*   *etxe-ak*   *eros-i*   *d-it-u-t*

        two   house-ABS.PL.DEF   buy-PFV   PRS-ABS3PL-ROOT-ERG1SG

        'I have bought the two houses'

As can be observed, only the definite paradigm distinguishes number (sg. or pl.). In (2b), it is evident that when we refer to more than one house (two, in this case) and they are definite, the verbal marker must be pl. The same holds true for most dialects and Standard Basque when they are indefinite, as in (2a), but in far eastern varieties and some old western texts, the verbal marker would be sg. in such a case.

In the paradigms of this section, I have provisionally included a total of eight cases. As a matter of fact, we will see that Bascologists disagree as to the exact number of cases in Basque, either because some of them consist of several suffixes (e.g. the destinative *etxe-arentzat* 'for the house' can be formally split into the genitive *-aren* plus the suffix *-tzat*, which is essive if it is directly attached to the root), or because some cases do not seem to behave like the others.[4] The cases not included in these paradigms have been discussed in §6.5 and §6.6. Be it as it may, it has been stated that Basque has a relatively rich case-system in comparison with many other languages of the world (cf. Iggesen 2011), but it is also true that, as e.g. in the Uralic family, a considerable number of those cases are derived by attaching postpositions to other (primary) cases (cf. §6.5).

As I have said, Basque has only one declensional paradigm, unlike other languages like Latin. This has sparked discussions among Bascologists about the existence of a declension proper (Rotaetxe 2008: 591). If we consider that we have one single paradigm, and that this is formed, at least in origin, by means of postponed agglutinated morphemes, it might be more suitable to speak about postpositions, and reserve the term *declension* for inflecting languages. Here, however, we cannot get tangled up in theoretical issues, and for the sake of convenience I will employ the term declension in the rest of the paper.

Consonant-final nouns are declined in the way shown in Table 6.2. It should be pointed out that, in that table, the pl. and indef. paradigms are similar at the phonological level (except for the abs. and dat.), due to the epenthesis of the eu-

---

4   For a description of the meaning of these and other cases, cf. Trask (*HB*: 92 ff.).

phonic *-e-* which is inserted in the indefinite paradigm in order to prevent the clash of two consonants.

Table 6.2.   Noun declension of consonant-ending nouns

jolas 'game'	Def.		Indef.
	Sg.	Pl.	--
Abs.	*jolas-a-Ø*	*jolas-ak-Ø*	*jolas-Ø*
Erg.	*jolas-a-k*	*jolas-ek*	*jolas-e-k*
Gen.	*jolas-a-r-en*	*jolas-en*	*jolas-en*
Dat.	*jolas-a-r-i*	*jolas-ei*	*jolas-i*
Instr.	*jolas-a-z*	*jolas-ez*	*jolas-e-z*
Ines.	*jolas-e-an*	*jolas-eta-n*	*jolas-e-ta-n*
All.	*jolas-e-ra*	*jolas-eta-ra*	*jolas-e-ta-ra*
Abl.	*jolas-e-tik*	*jolas-eta-tik*	*jolas-e-ta-tik*

The cases from the abs. to the instr. present *-a(r)-* in their singular paradigm, which is simply the article, and the rest of the cases have *-(e)ta* in the pl. and indef. The latter are the so-called 'local' cases, which —as will be shown in §6.3.2— display special features.

Finally, in Table 6.3 I present the paradigm of proper nouns, which are always definite.

Table 6.3.   Noun declension of proper nouns

	Def. Sg.	
	Vowel-ending	Consonant-ending
Abs.	*Iruri-Ø*	*Gasteiz-Ø*
Erg.	*Iruri-k*	*Gasteiz-e-k*
Gen.	*Iruri-r-en*	*Gasteiz-en*
Dat.	*Iruri-r-i*	*Gasteiz-i*
Instr.	*Iruri-z*	*Gasteiz-e-z*
Ines.	*Iruri-n*	*Gasteiz-e-n*
All.	*Iruri-ra*	*Gasteiz-(e-r)a*
Abl.	*Iruri-tik*	*Gasteiz-(e)-tik*

Proper nouns lack articles, which on the contrary are neatly present in the singular paradigm of common nouns: cf. abs. *lagun-a-Ø* vs. *Iruri-Ø*. On the other hand, the suffixes beginning with a vowel are directly attached to consonant-ending roots (*Gasteiz-i*), but if the root ends with a vowel, an epenthetic *-r-* shows up which we also saw in the paradigm of common nouns (cf. *Iruri-r-i*, and also *etxe-r-i*, *etxe-a-r-en*, etc.).

Apart from the suffix *-(e)ta-*, one of the peculiar features of the local cases is the morphological distinction between animates and inanimates. If we bring this distinction into Tables 6.1 and 6.2, a revision of the two would be Table 6.4.

Table 6.4.   Animate and inanimate paradigm of local cases

	Sg.		Pl.		Indef.	
	Inanimate	Animate	Inanimate	Animate	Inanimate	Animate
Ines.	*itsaso-an*	*lagun-a-(ren)-ga(-)n*	*itsaso-eta-n*	*lagun-en-ga(-)n*	*itsaso-ta-n*	*lagun-(en)-ga(-)n*
All.	*itsaso-ra*	*lagun-a-(ren)-gan-a*	*itsaso-eta-ra*	*lagun-en-gan-a*	*itsaso-ta-ra*	*lagun-(en)-gan-a*
Abl.	*itsaso-tik*	*lagun-a-(ren)-gan-dik*	*itsaso-eta-tik*	*lagun-en-gan-dik*	*itsaso-ta-tik*	*lagun-(en)-gan-dik*

The animate paradigm is built upon the genitive, to which a morpheme is attached, which nowadays appears as *-ga(n)*. In modern Standard Basque, this genitive is optional in the sg. and indef., but obligatory in the pl., probably to distinguish it from the indef.[5] In the continental varieties, instead of these forms we find the postposition —of uncertain origin— *bait(h)a*,[6] which takes the genitive (*EMH*: 114), although in old texts it also appears with the abs. Thus, the inessive sg. animate would be *lagun-aren baithan* 'in the friend', the abl. pl. *lagun-en baitha-tik* (or *baitha-rik*, cf. §6.4.2.3) 'from the friends', etc. We also come across this postposition with the meaning of 'house' in the same geographical domain, but Trask (*HB*: 208) considers this to be a specialization of what must have had a generalized local use.

6.3. The article, definiteness and number in noun declension

It is commonly accepted that the article, placed between the root and the case marker, is of recent introduction in the language (Schuchardt 1947 [1923]: 46, Michelena 1987 [1978], Manterola 2009a). It shows differences of use between dialects, the eastern ones being the most reluctant as to its use. I agree with

---

5   Nonetheless, these plural forms of animate nouns formed by *-ga(n)* upon the genitive are relatively recent (cf. Castaños-Garay 1979: 12 ff.). They first appear in Classical Labourdin (17th century), and seem to spread centrifugally. In 16th century texts, we always find animate nouns declined as inanimates, as in RS. *gizon-etan* 'in the men', Laz. *andra-etan* 'in the woman', Dch. *gizon-etarik* 'from the men', etc.

6   It might come from *bait da* 'which is' (cf. Creissels & Mounole 2011: 179).

Lafon (1999 [1954a]: 216) that the heterogeneity of number marking in Basque (depending on the root being vowel- or consonant-ending), and the differences in the formation of the indef., sg. and pl. paradigms, denote a progressive grammaticalization, which makes use of different elements.

In today's Basque, the article not only marks the definite / indefinite division, but also bears the feature of number (Lafon 1999 [1954a]: 211).[7] Let us take as an example the grammatical cases (absolutive, ergative, dative, instrumental, and genitive),[8] which are the ones that most clearly allow us to detect the presence of the article, at least in the sg. We shall then analyze the morphology of local cases, which presents certain peculiar features.

6.3.1. The grammatical cases

Due to its lack of overt case marker, the abs. case (-Ø morph in the indef.) is the clearest for discussing the issue of the article. Hence, I shall take it as a starting point. As can be seen in Table 6.1, we have abs. *etxe* for the indef., *etxe-a* for the def. sg., and *etxe-ak* for the def. pl.

The most widely accepted view, probably since Azkue's *MV*, is that the article —at least the one which grammaticalized from the 3rd grade article— came about in Basque at the same time as in the Romance languages, and in a parallel manner: from demonstratives (Lafon 1999 [1970]: 168, Michelena 1987 [1978], Trask *HB*: 199). According to the data put forward by Manterola (2006, 2008), we must locate that grammaticalization to around the 10th century AD. Thus, the *-a* of the sg. article would simply be the grammaticalized form of the 3rd grade demonstrative **har*.

There are several arguments in favor of this theory. To begin with, in the Aquitanian inscriptions (from approximately the 1st-3rd centuries AD) there is no trace of the article.[9] Secondly, in western dialects the 3rd grade demonstrative is still *a*: *etxe a* 'that house there'. Thirdly, Manterola (2008) has brought into the discussion several late medieval place names from Alava in which the article,

---

7   Lafon himself (ibid.: 217) suggests, not without hesitation, that an indefinite plural may also have existed, but I believe that he bases this on a wish to obtain a symmetric paradigm rather than on linguistic evidence.

8   Although there are several possibilities for grouping and naming the Basque cases, here they will be grouped into grammatical ones and local ones (pseudo-declensional cases are left out of either group, cf. §6.6). Within the grammatical cases, I shall call those which trigger an agreement marker on the verb "core cases" (i.e. abs., erg. and dat.).

9   This argument, however, is not conclusive, as the attested data in those inscriptions are proper nouns, which would not be expected to take the article, even if this had already grammaticalized.

not fully grammaticalized in its modern form, keeps the aspiration, as in *El-horzaha*, *Artazaha* (today's *Elorza*, *Artaza*), etc.[10] Fourthly, there is evidence that in a not very remote period, the 3rd grade demonstrative was not the only one grammaticalized as an article: we have medieval names like *Lope Ederro* (instead of *ederr-a* 'handsome'), or present-day family names like *Ezquerro* which may reasonably be derived from an article of proximity coming from the 1st grade demonstrative **hau(r) > ho(r) > -o* (Irigoyen 1981: 385-9, §7.2.4).[11] In western varieties, there is still an article in *-ok*, side by side with the common *-ak* (*HB*: 94).[12] From the 2nd grade demonstrative *(h)ori*, we would derive forms like *gorpuzori* (*gorputz* 'body'), found at a *fase sparita* in Dechepare's LN. dialect (16th century), and in somewhat later Labourdin authors (Michelena 1987 [1979b]: 450 ff., Castaños-Garay 1979, *HB*: 199).

As for the lexical origin of this **har*, Lakarra (2008a: 484) has proposed an interesting theory. After showing the relevance of the grammaticalization of serial verbs in the formation of numerous affixes in today's Basque, and noticing the origin of the article from a lexeme close to the verb 'take' in a number of languages, he has proposed that lexematic origin for Basque **har* 'that there'. As a matter of fact, this form would not require any asterisk, as the verb *har* 'to take' still presents that form today.

The forms without article have lost ground over history (Lafon 1999 [1954a]: 212), but, as I have said, eastern varieties are more conservative as far as the use of the article is concerned. As we move westwards, the presence of the article is found in increasingly more contexts. On the other hand, the penetration of the determiner has not been homogeneous, and has not run parallel to the Romance languages, for the Basque article is nowadays used in contexts where other Romance languages do not require it (Manterola 2012).

Without going into details (cf. *EMH*), I shall simply mention that the article penetrates more easily into the subject complements —or predicate adjectives—

---

10 It may be mentioned, in passing, that the presence of this aspiration renders any attempt to relate the Basque article to the Berber one *-a* impossible. Cf. Berber *argaz-a* 'the man' (Núñez-Astrain 2003: 288). For more problems concerning this and other genetic relationships, cf. Lakarra (2006b).

11 Although in my opinion it is not as strong as Irigoyen's proposal, the possibility —already hinted at by Michelena (1987 [1969]: 138)— that *-o* has arisen by analogy with the gender distinction in Romance, cannot be ruled out. Thus *ederra* was interpreted as feminine and *ederro* as masculine, especially in already Romanized areas. The feminine nouns that show up in both Irigoyen's and Michelena's works are very few, and among them, those which present that form in *-o* are practically nonexistent.

12 And hence corresponding case forms across the paradigm such as ines. *etxeotan*, soc. *etxeokin*, etc.

at *individual level* (Zabala *GB*: 329)[13] and into grammatical core cases (especially erg. or dat.),[14] and more easily into countable nouns than into uncountable ones in object function (Manterola 2006, Santazilia 2010), or even in word-listing contexts. In this respect, two attestations worth comparing are Aimery Picaud's word list and that of Arnold von Harff (cf. *TAV*: 49-51, 63-65, respectively, and §3.3), as both were gathered in Navarre —the former in the 12th century and the latter in the 15th century— and appear to reflect the spread of the article in progress.[15] As Table 6.5 illustrates, whereas in Picaud's vocabulary countable concepts are given in the def. sg. form and uncountable concepts are given in the indef. form (with the exception of *ur-ic*, in the partitive case), in von Harff's list uncountable concepts also bear the article, with the exception of *aragui*. Note that in some cases the listed words are exactly the same.

Table 6.5.   Aimery Picaud's and Arnold von Harff's word lists

	Picaud (12th c.)	Von Harff (15th c.)	Translation
countable nouns	*Urci-**a***		'sky (/ God)'
	*belaterr-**a***		'presbyter'
	*eregui-**a***		'king'
	*andre-**a***		'Lady'
	*iaon-**a***	*(hytzoko)sanm-**a***	'Sir'
uncountable nouns	*araign-**Ø***		'fish'
	*o(r)gui-**Ø***	*oge-**a***	'bread'
	*aragui-**Ø***	*aragui-**Ø***	'meat'
	*ardum-**Ø***	*arduw-**a***	'wine'
	*ur-**ic***	*oyr-**a***	'water'
	*gari-**Ø***		'wheat'
		*oluw-**a***	'oats'
		*gaz-**a***	'salt'
[*TAV*: 49-51, 63-65; adapted version, not including all words]			

---

13   In fact, at *stage level* the article still does not appear today: *irakasle-Ø dago* 's/he is as a teacher' vs. *irakasle-a da* 'he is a teacher'.

14   However, in the historical period we still have cases like Ax. (L, 17th century) erg. indef. *aingiru-k* (*aingiru* 'angel' + erg.), which would be ungrammatical today, unless we use them with an indefinite determiner (*hainbat aingeru-k* 'some angels').

15   Another word list often mentioned within this set of attestations is that gathered by Lucius Marineus Siculus in the 16th century (*TAV*: 146-147 and §3.3). In it, all items are given with the article. I have left it out of Table 6.5 because it corresponds to the western dialect, where the spread of the article took place faster, and hence it is not exactly comparable to the testimonies of Picaud and von Harff.

This earlier emergence of the article in countable nouns is entirely expected on typological grounds. This is no doubt due to the fact that the notion of singularity or plurality is semantically easier to capture in a countable element than in an uncountable one (Martínez-Areta 2009a: 76). It should not be forgotten that in Basque, number is morphologically expressed only by the article, a feature inherited from the demonstrative from which it comes. Hence, the only way to emphasize plurality is to use it, thus giving rise to an ambiguity: a form like *arraultze-ak*, with the plural article, can mean both 'eggs' and 'the eggs'. There is no overt way of expressing indefinite plurality. Hence, as the marking of number has become more important than that of definiteness, the article has even crept into contexts which are semantically indefinite, in order to mark singularity or plurality (Martínez-Areta 2009b: 254), as this example of *arraultzeak* shows.

Example (3), taken from a corpus from the end of the 19th century (Santazilia 2010: 192 ff.), shows a synchronic alternation. In (3a), uncountable nouns like 'bread' and 'wine' are expressed by the indefinite form, but so are those of plural semantics like 'eggs', undoubtedly due to their indefinite nature. It is in (3b) that we find a stage at which the morphological marking of plural is present in semantically plural phrases like 'eggs'. As already mentioned, definiteness and number go hand in hand, but in (3b) the former is obviated by the growing importance of the latter, thus marking in the same way as a definite NP a noun which is semantically indefinite. At the following stage, uncountable nouns will also end up being definite:[16]

(3) a. *Betea*   *da*   *mutil-ez*   *zeinek*   *baitute*   *ogi-Ø*   *eta*
         full   is   boy-INSTR.PL   who   have   bread-ABS.INDEF   and

         arno-Ø,   *arraultze-Ø*   *eta*   *gasna*
         wine-ABS.INDEF   egg-ABS.INDEF   and   cheese.ABS.INDEF(?)

     b. *Betea*   *da*   *mutil-ez*   *zeinek*   *baitute*   *ogi-Ø*   *eta*
         full   is   boy-INSTR.PL   who   have   bread-ABS.INDEF   and

         arno-Ø,   *arraultze-ak*   *eta*   *gasna*
         wine-ABS.INDEF   egg-ABS.PL   and   cheese.ABS.INDEF

         'It is full of boys who have bread, wine, eggs and cheese'

---

16   For a typological overview of all this, cf. Martínez-Areta (2009b: 254 ff.).

Therefore, it is evident that, before the creation of the article, neither number nor determination was expressed by agglutinated suffixes. In fact, this had already been proposed by Gavel (1929: II, 44).[17]

It is only logical, then, to consider that at first the article arose to create an opposition between the indefinite and the definite, not to mark number (*HB*: 199), although the latter function is inherent in the article, thanks to its former nature as sg. demonstrative. In the noun paradigm, we would initially only have markers like the *-Ø* of *etxe-Ø* in the abs. indef., and the *-a* of *etxe-a* in the def., the latter with a connotation of singularity inherited from the demonstrative. If we look at Tables 6.1 and 6.2 in their sg. paradigm, we can see that it is precisely in these grammatical cases[18] that the article *-a(r)-* is perfectly identifiable.

What about the pl. article *-ak*? It is commonly accepted that this —and actually the whole pl. paradigm—[19] arose later than the sg. article.[20] Michelena (1987 [1971b]: 145) suggested —although not very forcefully— deriving *-ak* from *-aga*,[21] a suffix present in place names (*Iturri-aga*, *Elorri-aga*, etc.). The

---

17 The question mark in (3a) requires an explanation. On phonetic grounds, we cannot know for sure whether in *gasna* the sg. article (*-a*) is present or not, for it is an *a*-ending root. However, its coordination with *arraultze-ak* suggests that it has the article (since in Basque, two coordinated NPs can take the article, and so can only the second, but it would be ungrammatical if the first had it and the second did not). In some varieties, the indef. form and the def. sg. one are homophonous in such roots. In others, by contrast, there are some phonetic differences between the indef. and the def. form. In the west and a part of the center, a dissimilation *-a + -a > -ea* occurs when the article *-a* is attached to the final *-a* (in the whole def. sg. declension, except in the ines.), as in *alaba-a > alabea* 'the daughter' ~ *alaba* 'daughter'. In Souletin, the *-á* of the definite abs. sg. is accented, unlike that of the indef. (*alhabá ~ alhába*, respectively), and in Salazarese, an epenthetic *-r-* is inserted between the final *-a* and the article (*alaba-r-a ~ alaba*, respectively).

18 Perhaps with the exception of the genitive, these cases can be considered the oldest ones (*HB*: 201).

19 In fact, this shows a greater dialectal diversity than the singular paradigm (cf. Castaños-Garay 1979: 65).

20 Michelena (1987 [1981a]: 48-9) puts forward an internal argument in favor of this. In western varieties, *a*-ending roots undergo a dissimilation *-a-a > -ea* when taking the sg. article: *hanka* 'leg' *+ -a > hankea* 'the leg'. This does not occur with the pl. article, where we find *hankak* (or *hankaak* in Old B and Laz.), not **hankeak*, a fact explained by Michelena by suggesting that when the pl. article was grammaticalized the dissimilation was no longer operative.

21 Earlier authors like Albert León and Henri Gavel (1920: 339, 362; 1929 II, 53-55) had already envisaged this. As will be shown below, today we know that the bisyllabic nature of this morpheme renders its existence in Proto-Basque impossible, and requires segmentation.

evolution would be as follows: *-aga > *-ag > -ak*, i.e. first apocope of *-a*, and then devoicing of the stop in final position.[22]

Nevertheless, as de Rijk (1998 [1981]: 218) recalls, Michelena himself poses problems to this explanation. On one hand, this suffix is more locative than abundantial,[23] and hence it presents semantic problems to being linked to the notion of plurality. On the other hand, as already mentioned, in western varieties the dissimilation *-a*-ending root + *-a > -ea* does not occur in the plural, but it does occur if the attached element is the suffix *-aga*, as we have the family and place name *Oleaga*, from *ola* 'factory, forge' + *-aga*.[24]

Let us recall that we start from a linguistic situation in which neither determination (definite / indefinite) nor, consequently, number is marked by means of clitics (Martínez-Areta 2009b: 250). At a later stage, the sg. article was created, with the sole function of opposing definiteness to indefiniteness, but bearing the notion of singularity, inherited from its previous stage as a determiner. Finally, the pl. article was created, which came to oppose to the notion of singularity of the article **(h)a(r)* by adding *-k* to this article.

---

22 We have examples of similar processes (-DV > *-D > -T, where D = voiced stop, T = voiceless stop) in the verbal personal endings, where we find *daki-da-la* 'that I know', but *daki-t* 'I know'. This happens at the synchronic level in many current varieties. But, in addition, in Roncalese the voicedness is kept in spite of the apocope in forms like *ba-daki-d* (*ba-daki-da), vs. the more widespread *ba-daki-t* 'I know', and in Ultzama (North.-HN), forms like *zai-re* (< *zai-da), without apocope but with *d > r* / V_V are still common, instead of the more widespread *zai-t* (bipersonal verbal auxiliary, 's/he (abs.) - to me (dat.)'). Also in RS (B, 16th century), we find fossils like *dike-a-da* 'I will give you', where *-da* is the marker of 1st p. sg. erg., instead of *-t* (cf. *FHV*: 235-6, Domene 2005: 1st Vol. II, 86).

23 A well-known and often cited example is the family name *Arespacochaga* (already attested in Biscay in 1468), formed by *aretx* 'oak' + *bakotx* 'one' + *-aga*. The presence of 'one' is at odds with the idea of plurality of the suffix *-aga*, which, on the contrary, interpreted as a locative would enable a meaning close to 'place of the solitary oak'.

24 Without making reference to this particular case, Artiagoitia (1990: 332, 340) had already shown that, in Proto-Basque, stops could not stand in syllable-final position. On this impossibility and on the syllable structure in Basque, cf. Lakarra (1995), who shows that, in Proto-Basque, the syllable had to be CV(R)(S), rendering the existence of ***-ag* impossible. For a good summary of Lakarra's theory about the monosyllabic root in Proto-Basque, cf. Chapter 5.

Phases	1	2	3[25]
	- determination - number	+ determination - sg. number	+ determination + sg./pl. number

Morphemes	-∅	Indef.: -∅
		Def. sg.: -a[26]
			Indef.: -∅
			Def.:
			    Sg. -a
			    Pl. -ak

Figure 6.2.   Grammaticalization of the article

According to what we have seen so far, case morphemes were directly attached to the root (Lafon 1999 [1965]: 175, *HB*: 202) before the grammaticalization of the article.[27] Later, the article came about from the 3rd grade demonstrative in order to morphologically mark the definite / indefinite distinction, and finally sg. number. Still later, and in an independent manner, it would be the pl. article's turn, from the previously mentioned *-aga*.

Martínez-Areta (2009b), quite reasonably and following Artiagoitia (1990) and Lakarra (1995), considers that the morpheme *-aga*, from which the grammatical pl. cases arose, could not be derived from *-ag-* (cf. for instance, erg. pl. **-ag-e-k > *-aek > -ak / -ek*), for this would be phonotactically very strange. The segmentation *-a-ga* (instead of *-ag-a*) would be more plausible, with the first element corresponding to the article, and the second to a pl. marker (2009b: 251).

In Table 6.6 below, I reproduce an adaptation of Martínez-Areta's table (2009b: 252), which represents his reconstruction of the morphemes prior to the total grammaticalization of the noun paradigm of grammatical cases.[28] Martínez-Areta's view would imply a simultaneous grammaticalization of case and number. Thus, according to him there was a period in which neither was grammaticalized.[29] A bare root like *gizon* had a generic value and, depending on the con-

---

25 As pointed out, in this phase —which is the present one— the penetration of the article as a feature of number has been a gradual process.

26 Or some other morphemes, now disappeared, like those mentioned: *-o*, *-ori*, etc.

27 According to Lafon (1999 [1970]: 171), the fact that in the paradigm of proper place nouns (cf. Table 6.3), which are definite in themselves, we have forms like *Iruri-n* or *Iruri-k* with no trace of the sg. article, bears witness to the previous phase, in which the opposition definite / indefinite did not exist.

28 The author himself indicates that the table shows the elements which take part in this grammaticalization, but that he does not attempt to postulate that this paradigm has ever existed as such in the diachronic axis. It is, then, a theoretical approach.

29 Martínez-Areta (2009a: 84) also believes that the richness of derivational suffixes with abundantial meaning in Basque, like *-(t)z(a)*, *-(t)zu*, *-tsu*, *-d(u)i*, etc. (cf. *dirutza* 'big amount of money', *mukizu* 'squirt (lit. having mucus)', *koipetsu* 'greasy', *pagadi* 'beech

text, it could refer to a sg. or pl. element.[30] The appearance of the 3rd grade demonstrative (*gizon-ha(r)), later grammaticalized as an article —marking both definiteness and number—, would initially have served to differentiate definite and indefinite NPs, but since the presence of that demonstrative also denoted singularity, another pluralizing morpheme became necessary, to cover the semantic slot which had thus emerged. This morpheme was *-ga, after the annexation of which the resulting form was *gizon-ha-ga. But this morpheme *-ga is also the ergative marker (or at least homophonous with it).[31] Thus, it could represent both the plural absolutive and the singular or plural ergative, when the case system was not fully developed.[32] According to Martínez-Areta (2009b: 260), this is the explanation for the lack of an indefinite plural. In any case, he is aware of the speculative nature of his paradigm, and assigns it a character of theoretical construct rather than of reconstruction.

Table 6.6.   Grammaticalization of grammatical cases hinted at in Martínez-Areta (2009b)

gizon 'man'	Def.		Indef.
	Sg.	Pl.	--
Abs.	*gizon-har*	*gizon-ha-ga*	*gizon*
Erg.	*gizon-har-ga*	*gizon-ha-ga-ga*	*gizon(-e)-ga*
Gen.	*gizon-har-e(n)*[33]	*gizon-ha-ga-e(n)*	*gizon-e(n)*
Dat.	*gizon-har-i*	*gizon-ha-ga-i*	*gizon-i*

6.3.2. The local cases

In the local cases, the features of definiteness and sg. number are not represented by means of the inclusion of the article. Instead, the case marker is directly attached to the root in vowel-ending roots (*etxe-ra*), with epenthesis of an *-e-* in consonant-ending roots (*jolas-e-ra*). Martínez-Areta (2009a: 80) sets out the problem in a clear manner, as Michelena (1987 [1971b]) already had. On one

---

wood'), might also be evidence for a period in which there was no declension, but an abundant derivation.

30   Cf. the examples in (3), which show that we can find, still in a recent period, roots without an article with pl. value.

31   The change *-ga > -k* of the erg. will be commented on in §6.4.1.1. Castaños-Garay (1979) and Lakarra (2005) believe that the *-k* of the abs. pl. and that of the erg. are originally the same morpheme.

32   In central and western varieties today, *-ak* is still the morpheme for the erg. sg. and pl., as well as for the abs. pl. However, in many of those varieties the plural paradigm has a different accentuation.

33   The *-n* in parenthesis will be explained in §6.4.1.2.

hand, the local cases seem to be more recent than the grammatical ones, due to their more lexical nature and their greater phonological content, but on the other hand they seem to preserve traces of a more archaic state, as they do not contain the article. This is precisely part of the evidence behind the assertion that at a certain point in time case morphemes were attached directly to the root, prior to the grammaticalization of the article (Lafon 1999 [1965]: 175, *HB*: 202).

Two further peculiarities of local cases, already touched upon but still worth recalling, are, firstly, the impossibility of their being attached to an animate root, unless we do so by mediation of the morpheme *-(ren)gan(-)* (cf. Table 6.4), and secondly, the appearance of the morpheme *-(e)ta-* in plural and indefinite forms (Martínez-Areta 2009a: 80).

Much has been written about the morpheme *-eta-*. Authors like Trask (1995b) and Urgell (2006) opt for an equivalence with the derivational suffix *-(k)eta*, present in words like *lapurr-eta* 'theft', used to denote an action, and also as a gerund in western varieties: *pents-etan* 'thinking'. As Humboldt and Vinson had already noticed (apud Castaños-Garay 1979: 66), it also appears in place names like *Arrieta*, from *harri* 'stone'. In the light of these place names, Schuchardt attributed a collective meaning to it, and considered it to be of Latin origin, i.e. as derived from the collective *-ēta* (pl. of *-ētum*) (1947 [1923]: 46). Bascologists such as Lafon (1999 [1970]: 168) have shown agreement with Schuchardt on this etymology, but Manterola (2009b: 97 ff.) has recently defended the patrimonial nature of this morpheme, arguing that morphological elements are hardly ever borrowed. Some other problems which he poses to Schuchardt's view are the following: why the pl. should be borrowed but the sg. should not; why the presence of this morpheme conveys not only plural meaning, but also definite value;[34] and why some of the earliest attestations of the morpheme in medieval place names, like *Sansoheta* (Manterola 2008), have aspiration. In fact, the form with aspiration would better explain the variant *-keta* of the derivational suffix mentioned above, which appears in words like *hizketa* 'speech', from *hitz* 'word' (cf. more examples of this alternation in §6.5.2.1).

To sum up, there are two views. Either the suffix *-eta* was borrowed and grammaticalized by entering the declensional paradigm, or, as Manterola defends, the plural paradigm was formed, just like the sg., by starting from the grammaticalization of the demonstrative of the 3rd grade, which employs the patrimonial suffix **-(h)eta* or **-keta* in the local cases. Let us explain the latter view in more detail.

---

34 In grammatical cases, the article had to specify both definiteness and number, but why should a borrowed morpheme like this do so?

According to Manterola, the whole declensional paradigm, in both local and grammatical cases and in both pl. and sg., would come from the grammaticalization of the 3rd grade demonstrative (not only the sg. of grammatical cases). But, what is more, that demonstrative would have grammaticalized not in its bare form, but already as a declined element. Thus, for instance, if the abs. sg. of the root *mutil* 'boy' is *mutil-a*, which as stated above comes from **mutil ha(r)*, the erg. sg. form *mutil-ak* should not be reconstructed as **mutil(h)a(r) + -k*, but as *mutil ha(r)k*, where the demonstrative itself already has the case marker attached.[35] That said, Manterola suggests the paradigm illustrated in Table 6.7.

Table 6.7.   Adaptation of the paradigm according to Manterola (2012)

mendi 'mountain'	Sg.		Pl.	
	Form	Origin	Form	Origin
Abs.	*mendi-a*	*mendi ha*	*mendi-ak*	*mendi hak*
Erg.	*mendi-ak*	*mendi hak*	*mendi-ek*	*mendi hek*
Gen.	*mendi-aren*	*mendi haren*	*mendi-en*	*mendi hen*
Dat.	*mendi-ari*	*mendi hari*	*mendi-ei*	*mendi hei*
Instr.	*mendi-az*	*mendi haz*	*mendi-ez*	*mendi hez*
Ines.	*mendi-an*	*mendi han*[36]	*mendi-etan*	*mendi hetan*
All.	*mendi-ra*	*mendi hara*	*mendi-etara*	*mendi hetara*
Abl.	*mendi-tik*	*mendi hatik*	*mendi-etatik*	*mendi hetatik*

Manterola's theory is advantageous from the point of view of economy, since it does not require the postulation of any reconstructed form: as can be observed, nearly all forms with an ungrammaticalized demonstrative are perfectly attested and employed today. It also introduces evident changes in the relative chronology of the grammaticalization of the paradigm. I shall describe this in more detail below, where case forms will be individually handled. In spite of this, this theory also has some drawbacks. Although it is true that the forms above are all attested, they never appear in the same variety. Leaving some secondary developments aside (cf. §7.2.3), Basque dialects do not distinguish between the abs. pl and the erg. pl. in 3rd grade demonstratives. Thus, the abs. and erg. pl. *hak* and *hek* are two dialectal variants of a demonstrative which represents both the abs. and the erg.

---

35 Without going so far, Castaños-Garay (1979) had already proposed that the plural, at least as far as the absolutive is concerned, grammaticalized first in the paradigm of demonstratives, and only later passed to the noun declension.

36 Although the 3rd grade singular demonstrative is *hartan* today, Manterola —just like Lafon (1999 [1970]: 168)— considers it an innovation, and believes that the old form is *han*, still preserved as an adverb meaning 'here'.

How can we explain the variant of *-eta-* which appears as *-ta-* in the indefinite local paradigm (cf. e.g. abl. *mendi-ta-tik*)? Trask (*HB*: 204), following Schuchardt (1947 [1923]: 47), thinks that it is the result of a reanalysis. If we take a look at Table 6.1, all grammatical forms present an *-e-* in the plural, with the exception of the abs. This fact would have triggered a reanalysis of the *-e-* of *-eta-* as a marker of plurality, thus giving rise to the form *-ta-* for indefinite forms,[37] since, in the grammatical paradigm, indefiniteness is marked by the absence of an article (Castaños-Garay 1979: 58-9).[38] This is reinforced by the paradigm of consonant-ending roots (cf. Table 6.2), where the *-e-* of *-eta-* continued to be used in indefinite forms, as an epenthesis between two consonants. The disagreement between Trask and Schuchardt is that the latter believed that the number distinction had passed from the local cases to the grammatical ones, not the other way round. Schuchardt considered *-eta* to have had a pluralizing (not locative) semantics, and that once the *-e-ta* segmentation had taken place, *-e-*, interpreted as a plural marker, spread to the grammatical cases —which had already grammaticalized the sg. article— in order to mark the pl. Cf., for instance, erg. sg. *lagun-ak* 'the friend' vs. erg. pl. *lagun-ek* 'the friends'.

Like Manterola, Lakarra (p.c.) argues for an internal explanation for *-eta-*, and also segments *-e-ta-*, but, unlike Schuchardt and others, does not consider it a reanalysis. He believes that *-ta* is in origin an undetermined locative morpheme which is related to the 3rd p. sg. prs. copula *da* 'is', later grammaticalized as an indefinite pl. marker. The first element *-e-*, in turn, would be the 3rd p. pl. morpheme which is found in the verbal system, and which has the variants *-te* and *-de* (cf. B. *dab-e* < **dau-e* '(they) have', *dauka-te* '(they) have', and *dau-de* < **dago-de* '(they) are').[39] If, following Trask (*HB*), the variant with voiced stop is the oldest one, according to Lakarra that pl. marker would be the result of the grammaticalization of the Proto-Basque verbal root **den* 'to finish', found in forms such as *e-ten* 'to interrupt, to section', *lehen* 'in the past' (< **den-en*), and the verbal marker of the past (cf. *zetorr-en* '(s)he came') (Lakarra 2008a: 475-9). This terminative nature would endow the following morpheme *-ta* —in prin-

---

37 And it would later have analogically spread to demonstratives and pronouns, which today require it when taking a case marker: cf. instr. *zer-ta-z* 'of what', ines. *hone-ta-n* 'in this', etc. (*HB*: 206).

38 Castaños-Garay (1979: 68) shares this view, but does not believe that the grammatical cases have played any role in that alleged "false cut" *-e-ta-*, since neither the erg. nor the dat. present that *-e-* in the plural in all dialects.

39 Gavel (1929) had already established this parallelism, and Castaños-Garay (1979: 95) also discussed it, although the latter did not share it. However, both refer to the *-e-* of the pl. nominal paradigm of grammatical cases, not to that of *-eta-*, as for both authors this morpheme is a suffix borrowed from Latin.

ciple indefinite— with a plural definite semantics. It is possible that this -*(d)e*- is also present in the pl. forms of grammatical cases (Lakarra 2011a: 51).[40] If this were so, the pluralizing morpheme -*e*- might have grammaticalized simultaneously in the local and grammatical cases, against Castaños-Garay's (1979: 77) view that the forms with -*eta*- would be older.

6.4. The origin of the primary cases

As has already been suggested, Basque cases can be classified into different groups. First, we have the core cases (abs., erg., dat.), which —unlike the rest of the cases— are indexed on the corresponding finite verb, and are those which can function as subject, direct object, or indirect object. These three core cases plus the gen. and instr. make up the group of the grammatical cases, which formally present an -*a*- in the sg. and no -*(e)ta*- in the pl. and indef., and semantically are attached to both animates and inanimates. These are opposed to the local cases (ines., all. and abl.), in which the case marker is directly attached to the root in the sg., and -*(e)ta*- appears in the pl. and indef. Besides, these case markers are never —or only exceptionally— attached to animates. Grammatical and local cases together make up the group of primary cases, opposed to the secondary cases (§6.5), which are built upon the primary ones.

The case system cannot be very old, according to Michelena (1987 [1971b]), Trask (*HB*: 246) and Lakarra (2005: 423; 2006b), among others, and it is evident that it has grown in complexity over history. Most case markers are too transparent to be considered old, but even those with scant phonological content, and apparently reduced, have a problem. Although it cannot be discussed here (cf. Trask 1977, Gómez 1994, Gómez & Sainz 1995, Lakarra 2006a, b), there are reasons to believe that Basque may have had a SVO word order, or even VSO or VOS, at some point (see §10.4). This being so, according to Greenberg's universals we should expect the presence of prepositions and a relative scarcity of postpositions in such a period (Lakarra 2006b: 267, 2008a: 455-6). By contrast, the abundance of postpositions is typologically more plausible in SOV languages.[41] Hence, while some cases may have grammaticalized at a stage prior to the SOV order, the latest cases —perhaps the local ones— most probably grammaticalized at a V-final stage.

---

40 For the variant -*heta* (Manterola 2008), we might conceive **ha-(d)e-ta*.

41 Lakarra (2006a, 2006b) defends the view that Basque started from a situation in which the V was not final, and the language was predominantly prepositional. Thus, the language would be in a state of typological drift towards an SOV order and proliferation of suffixes.

6.4.1. The grammatical cases

6.4.1.1. The ergative[42]

While the absolute, the covertly marked case, is used for direct objects, object complements (although this function is usually performed by the essive), subject complements, vocatives, and subjects of intransitive verbs, the ergative is employed as the subject of transitive verbs. It shows dialectal variation in its plural form: while continental varieties and some Navarrese ones have *-ek* for the ergative plural, in the rest we find *-ak* (*Lagunek / lagunak ekarri dute* 'The friends have brought it'). Let us recall that Michelena derived the plural article from *-aga* (§6.3.1). The ergative plural marker would be, for this author and for some others (de Rijk 1998 [1981]: 212-3), simply the agglutination of the pl. article and the ergative marker *-k*, which after the apocope of *-a* (*-aga > *-ag*) would require an epenthesis of *e* to avoid the clash of two consonants (**-ag-e-k*), the intervocalic velar fricative later dropping out, as is usual in Basque.[43] The result is **-aek*, which later phonetically evolved into *-ek* in the east, and into *-ak* in the center and the west.

Trask (*HB*: 200) opts for a second theory, according to which the erg. pl. form would come, exactly like the sg. article, from the 3rd grade demonstrative, but this time from the pl. one. This would allow us to conceive a simultaneous development of the sg. and pl. paradigms, insofar as we would derive the former from the sg. demonstrative and the latter from the pl. demonstrative.[44] Thus, as can be seen in Figure 6.3, in the pl. we would have, just as in the sg., simply the grammaticalization of the demonstrative, with no trace of the morpheme *-aga*.

---

42 Rather than as a purely ergative language, some have defined Basque as an active language, since ergative marking is also used for the subjects of some active, but semantically intransitive, verbs, such as *Patxi-k eskiatu du* 'Patxi has skied', or *Patxi-k saltatu du* 'Patxi has jumped' (Dixon 1994, Etxepare *GB*: 364).

43 Uhlenbeck (1909/10 [1903]: 105) derived the erg. pl. from the agglutination of the pl. article *-ak* and the erg. *-ek*, whence **-akek > *-aek > -ak / -ek*. But this would imply that an intervocalic voiceless stop should drop out, which seems very unlikely in Basque. On the other hand, if Lakarra's assumption (§6.3.2) is right, that *-e-* would not be a mere epenthesis. It could be the plural morpheme, grammaticalized from the Proto-Basque root **den*.

44 Plural demonstratives are homophonous in absolutive and ergative forms (cf. §7.2.1), which makes this explanation valid not only for the erg., but also for the abs., and even for the rest of grammatical cases.

gizon ha(r)-k Erg. > *gizona-k* Erg.Sg.

gizon haik Erg. > *gizon (h)aek* > *gizon-ak / -ek* Erg.Pl.

Figure 6.3.   Grammaticalization of the erg. sg. and erg. pl. according to Trask

Irigoyen (1981) also favors this view, but believes that the erg. pl. *-ek* would come from a demonstrative **heg-*. However, this reconstruction cannot explain the ergative pl. in *-ak* that we have in the center and west of the entire Basque-speaking territory.

In fact, both approaches have problems explaining the absolutive pl. *-ak*, which is the abs. pl. form in all dialects. The former analysis has to face the previously mentioned semantic problems of the suffix *-ag(a)*, whereas the latter considers the absolutive sg. form to be a reduction of *haik* > *-aik* > *-ak*, disregarding the fact that *haik* is a dialectal and not very widespread form in comparison with the more widespread *hek* and *haiek*, starting from which it would be easy to get to erg. pl. *-ek*, but not so easy to get to the central-western variant erg. pl. *-ak*, nor the general abs. pl. *-ak*.

Lakarra (2005: 442-4) suggests that if, according to the phonotactics of Basque, we propose **-ag* > *-ak* for the pl. article, the erg. marker should be handled in the same way. It does not seem reasonable, then, to stop the reconstruction of the erg. pl. in **-ag-e-k*, since the *-k* of the erg. cannot be anything but a voiced stop unvoiced in final position. Consequently, in Lakarra's analysis, the erg. marker also comes from the already known morpheme **-ga*, the semantics of which is exclusively locative. That morpheme would be the same as the one we find in the local paradigm of animates (cf. Table 6.4), in medial position, and hence with no phonetic modification.

But, how can we link a locative morpheme to an erg. marker? As I have said, Lakarra (2005: 442-444) has suggested that **-ga* is the same morpheme which is used in animate NPs when a local case marker is used (cf. ines. *lagun-a-ga-n* 'in the friend'), and the reason appears to be their high position in the agentivity scale. Animate NPs tend to be agents and local functions are only rarely assigned to them. That is why, according to Lakarra, they would require that morpheme **-ga* to incorporate a local case marker. Since this **-ga*, then, would occur only with animate nouns, it took on an agentive connotation, and, perhaps after turning into *-k*, came to mark the agent in ambiguous sentences where it was not clear which was the agent and which the object (as in 'John kissed Mary'). Finally, *-k* spread by analogy to inanimates, with the same function of marking the agent in opposition to the object.

A potential problem with this view is that it is at odds with the agentivity scale. In fact, if there is an ergative / non-ergative split in noun morphology, we

should expect to find the ergative marking with inanimates, and the non-ergative marking with animates, not the other way round. It is precisely inanimate NPs which, due to their low position in the agentivity scale, require an erg. marker to mark them as agents of an action (Dixon 1994: 83 ff.). One way of avoiding this criticism could be to interpret that, once *-*ga* (> -*k*) had become an agent morpheme, it affected all nouns at the same time, not animates before inanimates. Thus, when *-*ga* appeared only with animates, its meaning was still local, without any agentive connotation.

Another explanation is the one proposed by Martínez-Areta (p.c. and 2012), who considers that the rise of the ergative can be better explained if we interpret the whole process syntagmatically. Following Trask (1977: 209, 1979), he believes that the origin of the ergative is to be searched for in an active reinterpretation of passive sentences, and reconstructs a passive sentence in which the verb would be diathetically unspecified. The passive nature of the sentence would be syntagmatically specified by marking the agent complement with the local marker *-*ga*.[45] Thus, a sentence like *Jone dakus Peru-ga* would mean 'Jone is being seen by Peru'. Later, the agent complement was topicalized, thus moving to occupy the initial position: *Peru-ga Jone dakus* 'By Peru Jone is seen'. Finally, this passive structure was actively reinterpreted, and the agent complement became an agent subject of that active sentence, thus giving rise to the historical ergative.

As for the relative chronology, according to typological principles, ergativity most often arises either in VSO languages or in SOV ones, in both of which SO are next to each other (Dixon 1994: 50). Since most Bascologists believe that Proto-Basque was VO (§10.4), the rise of SOV appears to be a *terminus post quem* for the rise of ergative marking, or perhaps both features —erg. *-*ga* (> -*k*) and SOV— arose as part of the same process.

6.4.1.2. The genitive

If, as explained above in §6.3.1, the plural paradigm of grammatical cases developed according to the first theory —the one which defends the position that they derive from the grammaticalization of the suffix -*aga*, not of the pl. demon-

---

45   Typologically, it is common to find the patient subject of a passive marked by a locative morpheme. Cf. the locative preposition used in Spanish, as well as its English translation, as in *ha sido vista **por** mi madre* '(s/he) has been seen **by** my mother'.

strative—, the pl. genitive would have undergone the same development as the other grammatical cases: *-ag(a)-e(n) > *-ag-e(n) > -aen[46] > -en.

It is quite commonly accepted that the old genitive was not -(r)en, but -(r)e (Lafon 1999 [1965]: 173, Michelena 1985 [1972b]: 297), to which the inessive marker -n would have been attached. There are several clear pieces of evidence supporting this view. To begin with, this gen. form is the one that appears in pronouns like *gu-re* 'our', *zu-re* 'your (sg.)', etc. It also appears in cases built upon the genitive (§6.5.2), and in fossilized expressions like *are* 'still more', resulting from the union of the article with the genitive (*HB*: 201). It also seems to be attested in medieval proper nouns such as *Orti-re semea* 'Orti's son' (Iratxe, 1125) (Michelena 1987 [1969]: 96), or even *Orci-re çorita* (1110), which was already *Orçi-ren çorita* in 1284 (Irigoyen *DRFLU II*: 177-178).

Michelena (1987 [1970]: 290, 1985 [1972b]: 297) established a parallel between the genitive morpheme -(r)en (4a) and the one employed to form relative clauses, -(e)n (4b), as well as the past tense marker of the verbal morphology (4c):

(4) a. *Gizon-a-ren*   *liburu-a*

         man-ART-GEN   book-ART.SG

         'The man's book'

     b. *Gizon-ak*   *du-en*   *liburu-a*

         man-ERG.SG   has-RLV   book-ART.SG

         'The book that the man has'

     c. *Zetorr-en*   *gizon-a*

         came-RLV / PST   man-ART.SG

         'The man who came / The man came'

As can be seen in (4), the three morphemes are formally very similar,[47] and the function of -en seems to be the same in both (4a) and (4b, c): to link a dependent to its head. The difference is that while in (4a) the dependent is an NP, in (4b, c) the dependent is a relative clause (and the -(e)n is attached to its finite

---

46 This phase is residually attested in Old Biscayan and Alavese, as in RS. *iusturi-aen atzean* 'after the thunders'.

47 Another morpheme which can be related to these is the superlative, which today appears as -en (e.g. *gizon(-ik) altu-en-a* 'the tallest man'), and which is undoubtedly the same morpheme as the gen. pl. one. This morpheme also shows fossilized traces of an old variant -e, as in *ja(u)be* 'owner' (< *jaun-* 'lord' + -e) and *hobe* 'best' (< *hon-* 'good' + -e) (Lakarra 2006a: 576, 2006c: 581).

verb).[48] Another shared feature is that both the modifying noun in the genitive case and the relative clause precede the phrasal head *liburua*. However, while in old texts, and until recently in far eastern dialects, the order 'phrasal head + relative clause' is productive, the order 'phrasal head + genitive' is rarer.[49] A final difference is that while the gen. *-(r)en* has the variant *-(r)e*, which seems to be older, the *-(e)n* of relative clauses does not have any variant without *-n*. The parallelism between the two linking functions, then, is perhaps to be sought in the simple *-n*, rather than in the whole *-(e)n*.

A final question to clear up in relation to the morphology of the genitive is the origin of the *-r-* which appears in the gen. (and dat.) indef. of vowel-ending roots (Table 6.1), and in the gen. and dat. of proper nouns (Table 6.3). It is usually explained as an epenthetic element, without getting into diachronic considerations. However, in the 19th century, van Eys (apud Castaños-Garay 1979: 27) had already proposed that it came from the grammaticalization of the 3rd grade demonstrative as the article, after an analogical process. If for the abs. sg. we reconstruct **etxe-har*, for the gen. sg. we will reconstruct **etxe-har-e(n)*. From these, after the regular changes we would have abs. sg. **etxe-(h)a(r) > etxe-a* (with the expected loss of *-r* in word-final position), but gen. sg. **etxe-ar-e(n)* (with the expected preservation of *-r-* between vowels). Since in the abs. sg. —probably the most frequently used morphological case in discourse— the article was *-a* after that change, the gen. sg. was reanalyzed as **etxe-ar-e(n) →* *etxe-a-re(n)*, *-r-* now being part of the gen. case marker. After this reinterpretation, the new gen. *-re(n)* was attached to the indef. of vowel-ending roots (→ *etxe-ren*), and to vowel-ending proper nouns (→ *Iruri-ren*).[50]

---

48   As example (4c) shows, in Basque, verbs in the past are marked by an *-n*, but since the relative finite verbs are also formed by attaching *-(e)n*, past forms with and without relative morphemes are formally identical.

49   Recall that we are talking about the possessive genitive. With the local genitive, the order 'phrasal head + loc. gen.' is much more common (cf. §10.3.3).

50   As is the case for the gen. pl. (cf. §6.4.1.2), western varieties —especially in old texts— also pose particular problems in relation to the gen. sg. Actually, in Old Biscayan we find, in addition to *-aren*, attestations such as RS. *alargun-aen* 'of the widow', Cap. *geure Iaun-een* 'of our Lord', Cap. *Iaungoiko bizi-en-a* 'of the living Lord', Cap. *aiteen* 'of the father' (*aita* 'father'), Bet. *aiten* 'of the father'. According to Castaños-Garay (1957: 65), in all of these, the *-r-* appears to have dropped out. From a common gen. sg. *-aren*, then, we would have *-aren > -aen > -een > -en*, the attestations corresponding to different phases of that development (and the *-a* of *-a*-ending roots being assimilated to the *-e-* of the gen. suffix). A further feature of Old B is the lack of the epenthetic *-r-* that we see in vowel-ending proper nouns, as in Bet. *Eva-en* 'of Eva' or Cap. *no-en* 'of whom?' There are two conceivable explanations: 1) these are also cases of the *-r-* dropping out between vowels, as with *-aren*; and 2) in these cases, the analogical spread of

6.4.1.3. The dative

This case is used mainly as a marker of indirect object. The dat. sg. form comes from the cliticization of the 3rd grade demonstrative (*-har-i* > *-ari*) whereas the *-r-* of the dat. indef. form on vowel-ending roots (*etxe-ri*, cf. Table 6.1) and the one on vowel-ending proper nouns (*Iruri-ri*, cf. Table 6.3) are to be explained as a result of a reanalysis of the type abs. sg. **etxe-har* > *etxe-a*, hence *etxe-ar-i* → *etxe-a-ri* (stem-article-dat.) (cf. §6.4.1.2 for the gen.).

The plural poses special problems. If we start from the same explanation as that of the other grammatical cases, the development would be as follows: **-ag(a)-i* > **-ag-i* > *-ai*. This development would account for the B. G. Al. variant *-ai* (also present in western regions of Navarre), but to the east we still have the variants *-ei* and S. (and adjacent areas) *-er*, for which the origin of the *-e-* would remain unexplained.[51] If we go back to Lakarra's account of the verbal pluralizer *-de* as derived from **den* 'to finish' (cf. §6.3.2), we could interpret that the dative morpheme *-i* has taken this pluralizer *-e-*, thus giving rise to the dat. pl. form *-ei*, in those dialects which have it.

According to Trask (*HB*: 201), *-ei* would be the old form. From this, another form like **lagun-e-r-i* would have been created, *-r-* being an epenthesis analogous to the one that we have already seen in the gen. and dat. indef. of vowel-ending roots. The variant *-er* found from Donibane Garazi to the east would be an innovation with apocope of *-i*. A point in favor of this view is that in the Souletin dialect, where the accent is systematically paroxitone, the dat. pl. has the accent on the last syllable (*lagün-ér* 'to the friends'), which can be explained if we start from **lagün-éri* and assume a later apocope.[52]

The opposite option, according to which the *-r-* may have dropped out between vowels, is not inconceivable, but would be more problematic, since it seems that the loss of intervocalic *-r-* appears to be a recent phenomenon (*FHV*: 311).

To round off the explanation concerning the dative, Lakarra (2008a, 2011a) has recently proposed a lexical source for the pure dative marker —which is *-i* in

---

-ren to vowel-ending proper nouns did not take place. All in all, however, the overall picture of these phenomena appears to be a western innovation.

51 Unless we postulate, quite *ad hoc*, an assimilation *-ai* > *-ei* (which would not explain *-er* anyway).

52 In fact, the dat. pl. *lagun-eri* 'to the friends' is not only attested, but has been spreading in the continent and northeastern region of Navarre from two or three centuries ago onwards (cf. Pikabea 1993: 45-48). However, as Padilla-Moyano (2011: 41) points out, this does not appear in old texts, at least not in the continent. Therefore, it seems dubious that there is a continuity between the **-eri* proposed by Trask as a source of S. *-ér* after apocope and this modern *-eri*.

the historical period—, deriving it from the grammaticalization of the Proto-Basque root *nin* 'to give', which can be found in finite verbal forms of old texts (e.g. Laz. *in-da-zu*, Dch. *i-da-zu* 'give me!') and fossilized expressions, as well as in auxiliary verbs of tripersonal agreement in central dialects (e.g. *errai-ten d-i-da-zu* 'you say to me', *nin* 'to give' being, among others, the original verbal root of *didazu*, according to Lakarra 2008a: 482). In fact, the verb 'to give' is a common lexical source of dative markers (cf. Heine & Kuteva 2002: 153-4).

6.4.1.4. The instrumental

The instrumental has the function, among others, of marking the object by means of which an action is performed: *mailu-az kolpatu du* '(s)he has hit with the hammer'. But it also holds a great number of other miscellaneous values as in many other languages (Narrog & Ito 2007, Narrog 2009, 2010), which can be temporal, modal, material, etc. Sometimes, the instrumental functions simply as the case governed by another element, whether by a verb, a NP, an adjective, or a participle. It can be also used in antipassive constructions, in which the subject moves from being marked as erg. to being marked as abs., and the direct object, previously in abs., takes the instr. (cf. de Rijk 2002). The purely instrumental function of the instrumental case is actually on the wane, and is increasingly being taken over by the sociative, especially in western areas. Formally, it is commonly -*z*, but the variant -*s* is found in part of the High Navarrese dialect. The latter variant appears to be an innovation, not easy to explain. Perhaps the motivation is phonetic, for there is a phenomenon, not extremely widespread but common in several regions, whereby $z > s$ / _ T (*FHV*: 282). Examples of this include the place name *Maeztu / Maestu* and the word *uztaia / ustaia* 'horse collar'. The instrumental appears in exactly this context when it takes the local genitive -*ko* (cf. *altzairu-z-ko* 'of steel'). It could thus turn into -*s*-, and later generalize as dorsal in other phonetic contexts.

   Whether this is the case or not, this development would be phonetically unusual anyway, for there are not many sibilant fricatives in final position outside the verbal morphology (*HB*: 201). As loanwords like *gorputz* 'body' (< Lat. *corpus*) show, in the final position of old lexemes the sibilants were neutralized in favor of the affricate (-*tz* and -*ts*, seldom -*z* and -*s*). To explain the instr. -*z*, then, it would be reasonable to think that an old vowel which came after the sibilant apocopated. In this case, we would originally have had instr. *-zV*, perhaps *-za*. This form would not need any asterisk if the old instrumental of demonstratives and personal pronouns, e.g. *horr(-e)-za-z*, *zu-za-z* (cf. §7.2.3, §7.3.2), were to be analyzed as the old instrumental suffix -*za* plus its apocopated variant -*z* pleonastically attached.

If *-za* were really the old form of the instrumental, its resemblance with the derivational suffix of abundance *-(t)za* (as in *diru-tza* 'plenty of money', *elur-tza* 'snowfall', *Elor-tza* 'place of hawthorns') would be striking. A semantic link between the two, however, is more difficult to figure out. Martínez-Areta (p.c.) speculatively suggests that the instr. may have taken on that semantics from contexts in which it is the case required by verbs and nouns of abundantial meaning, as in *poz-ez bete-rik* 'full of happiness', *jende-z gainezka* 'crammed with people', etc.[53]

The morphology of the instr. resembles that of grammatical cases in most aspects. Thus, the instr. sg. is formed upon the article (*etxe-a-z*), whereas the indef. is formed by attaching the pure case marker directly to the root (*etxe-z*). The plural presents an *-e-* (*etxe-e-z*), which makes it similar to erg. pl. *etxe-e-k* and dat. pl. *etxe-e-i*, but in the instr. this *-e-* appears in all dialects. Hence, unlike for the erg. pl. and the dat. pl., for the instr. pl. a unitary explanation appears to be possible, whether that of Lakarra's spread of the pluralizer *-e-* (cf. §6.4.1.3), or any other.[54]

6.4.2. The local cases

Bascologists usually agree that these cases are more recent than the grammatical ones. They have more phonic substance, more semantic content, and typically present polymorphism across both their chronology and geography (Michelena 1985 [1972b]: 303, Trask *HB*: 202).

On the morphological level, as seen above, local cases make use of *-eta-* in the pl., *-ta-* in the indef., and, unlike in grammatical cases, the article *-a-* is not inserted in the sg. At least on the surface, then, the case marker is directly attached to the root, as with grammatical cases before the grammaticalization of

---

53 An alternative proposal has been made in Igartua & Santazilia (2013). In this proposal, *-(t)za* is regarded as a locative rather than as an abundantial marker in origin. The typological viability of the semantic extension LOCATIVE → INSTRUMENTAL has been defended by a number of authors in recent works. On a theoretical level, it has been defended by Heine *et alii* (1991: 159). Palancar (2002: 125 ff.), in turn, mentions some particular cases of such a process. Likewise, Narrog (2010) compiles several descriptions of concrete cases in Indo-European and Pama-Nyungan languages, among others.

54 In some varieties, the *-(e)ta* morpheme of indef.. and pl. of local cases has spread to the instr. (instr. pl. *etxe-eta-z* vs. *etxe-ez*), but this is a very recent phenomenon. This encroachment of *-(e)ta* on the instr. has most probably been favored by the fact that in personal pronouns and demonstratives the instr. pl. does have *-(e)ta* (*zu-taz* 'of you', *hori-etaz* 'of those'). In old texts, however, we only find *-zaz* for personal pronouns and *-ezaz* or *-ez* for demonstratives (*zu-zaz*, *hori-ez(az)*).

the article, according to Trask (ibid.). The absence of the article in the sg. paradigm, along with the presence of the morpheme *-eta-*, causes this author to believe that the distinction between sg. and pl. that we find in the local cases must be recent and inherited from the grammatical cases. Hence, Trask defends the view that there was a period in which local case markers entailed no number opposition, and were attached directly to the root. The morpheme *-eta-* would simply be a collective suffix which took on pl. value when the local case markers were attached.

6.4.2.1. The inessive

The inessive, formed by the suffix *-n*, can express both spatial and temporal notions. Thus, *mendi-an* means 'on the mountain', but *arratsalde-an* 'in the afternoon'. This case is one of the most studied by Bascologists, and that is not surprising, due to a peculiarity that will be immediately set forth. First, let us take a look at the following examples put forward by Manterola (2009b: 93):

(5)   a.   *mendi + a = mendia*   *mendi-a + n = mendian*

           mountain + ART   the mountain + INES

           'The mountain'   'On the mountain'

      b.   *lan + a = lana*   *lan-a + n = lanEan (**lanan)*

           work + ART   the work + INES

           'The work'   'At work'

If we look at (5a), we notice that the inessive marker seems to be attached to the root with mediation of the article (cf. Lafon 1999 [1970]: 169). This would be the only local case with such mediation, since, as I have said, the rest of them are attached directly to the root. But let us now consider example (5b). Consonant-ending roots generally take the article in the same manner as vowel-ending roots, but not in the inessive. If we attached the inessive marker to the /root-article/ complex, as we find in *mendi-a-n* (5a), we would expect to find **lan-an, but this is ungrammatical in any dialect of any period. Instead, we find *lan-ean* (or phonetic evolutions of this). Where does this *-e-* come from? It is undoubtedly an epenthesis. If we compare Tables 6.1 and 6.2, we can observe that, leaving the inessive aside, in order to form the sg. of local cases, when the root ends in a consonant, an epenthetic *-e-* is inserted. This occurs because all local case markers begin with a consonant. The epenthesis of *-e-* prevents the clash of the root-final and the case-initial consonants:

(6) a. *mendi-ra*   *mendi-tik*

     b. *jolas-e-ra*   *jolas-e-tik*

However, the inessive inserts that epenthesis between the root and (what at first sight seems to be) the article. In view of this unexpected phenomenon, Jacobsen (1977) opts for considering this *-a-* to be not the article, but a part of the inessive marker itself, which would be *-an*. In this view, he would later be followed, with some modifications, by de Rijk (1998 [1981]) and Trask (*HB*). Let us look into this explanation in more detail.

As has already been mentioned, the context for the epenthesis is C# _ #C. But here we find C# _ #V instead (cf. *jolas-e-an* 'in the game'). In order to couple the inessive with the structure of the rest of the local cases, Jacobsen suggests that the original form of the inessive was **-Can*. If we start from this case form, the expected inessive of a consonant-ending root would be **jolas-e-Can*, *-e-* being the epenthetic element which prevents the clash between two consonants, exactly as in the other local cases.

In favor of this hypothesis, Jacobsen sets forth a number of arguments. The first is the following. In Biscayan (also in some linguistically Guipuscoan valleys), Alavese and in the area of Burunda (western Navarre), *-a*-ending roots undergo a dissimilation when a case marker beginning with *-a* is attached to them. This dissimilation can be represented in the following way: *-a + -a > -ea*.[55] Thus, for instance, while in the rest of varieties a word like *proba* 'proof' remains *proba* when taking the article (*-a*),[56] in southern Guipuscoa, Biscay and Alava we have *probea*. The same holds for *hanka* 'leg' vs. *hankea* 'the leg', *gona* 'skirt' vs. *gonea* 'the skirt', etc. This occurs with all of the grammatical cases, because they all contain the article *-a-*, but not with the inessive, even if on the surface it also begins with *-a*: *gona* 'skirt', and *gonan* 'in the skirt' (but never ***gonean*).[57] The reason suggested by Jacobsen is that during the period in which the process *-a + -a > -ea* was operative, the inessive marker still maintained the structure **-Can*, which prevented the two *a*-s from being in contact. The original form, then, would have been **gona-Can*, where *-a-* would not be the article. Thus, as occurs with the other local cases, the inessive would not be built upon the article.

---

55 This form *-ea* can undergo subsequent dialectal changes and evolve into *-ia*, *-ie*, or *-i*.

56 In fact, as already mentioned some varieties display an accentual opposition between the root form and the root plus article form, and in Salazarese the root plus article form undergoes an epenthesis of *-r-* with *-a*-ending roots, i.e. *-a + -a > -ara* (cf. Sal. *alaba* 'daughter' vs. *alabara* 'the daughter').

57 In Old Biscayan and Alavese, this *-a + -an* juncture sometimes appears as *-aan*, as in Laz. *plazaan* 'in the square'.

This explanation can account for the inessive *-ean* of consonant-ending nouns, since *-e-* would be, as in the rest of the local cases, an epenthetic segment preventing the clash of two consonants: *lan# + #-Can > lan-e-Can* (cf. (5b)). But how can we get from **-Can* to the historical inessive *-an*? Apparently, by the dropping out of that C in intervocalic position. This is not surprising in Basque phonetics, provided that the C is voiced. In particular, the consonants that can drop out between vowels are *b, d, g, r, n* (Jacobsen 1977: 164-5). Of these, de Rijk (1998 [1981]: 217) recalls that *g* is the one that most easily drops out, from a typological perspective, and also in Basque. Furthermore, *g* is also the best candidate for an etymological explanation for **-Can*. In fact, if we postulate that *-gan* is the source of ines. *-an*, this would not require any asterisk according to Jacobsen, for it is attested in the animate paradigm of local cases (cf. Table 6.4.). If the initial consonant of the morpheme is not lost in this paradigm, it is because here *-gan* is sometimes preceded by the genitive *-ren*, so that the voiced stop is not between vowels (*lagun-aren-gan* 'in the friend'; cf. Table 6.4). We would have, then, an inessive morpheme *-gan*. This theory has been commonly accepted (Trask *HB*: 202-3) until recently, although, as we shall see, there do exist (in my opinion more reasonable) alternatives.

An apparent advantage of the theory initiated by Jacobsen is that if we assume it, a plausible relative chronology can be established. By the time **-g-* had dropped out intervocalically, the process *-a + -a > -ea* was no longer operative. As a result, when it dropped out, the two *a*-s did not dissimilate, but coalesced (**gona-gan > gonaan > gonan* vs. ***gonean*). The same would have occurred when attaching the inessive marker to the morpheme *-(e)ta* of the pl. and indef. local paradigm: **mendi-(e)ta-gan > *mendi(e)ta-an > mendi(e)tan* vs. ***mendi(e)tean* (ibid.: 164).

As for the origin of *-gan*, Jacobsen (ibid.: 166-7) believes that it was an independent word with locative meaning. According to him, it might have something to do with the common *gain* or its B. variant *gan(e)* 'high place, top'.[58] Whether that lexical origin is correct or not, *-gan* would have been a locative postposition, later grammaticalized as inessive.

De Rijk (1998 [1981]: 217) recalls the locative postpositions used in Basque today:

---

58 This comparison, however, would be extremely problematic. From a formal point of view, the *-i-* of the common form would be difficult to explain. From a semantic point of view, it would remain unclear as to how a locative marker of such a generic meaning like *-gan* has become as specific as the lexical item just described. Last but not least, the change **-ain > -an* is dialectal and late.

(7)   a.   *mendi-(a-ren)*   *atze / aurre / gain / azpi...-an*

            mountain-(ART-GEN)   behind / in front / on / under...-INES

            'Behind / in front of / on / under ... (lit. of) the mountain'

      b.   *lagun-a-ren*   *atze / aurre / gain / azpi...-an*

            friend-ART-GEN   behind / in front / on / under...-INES

            'Behind / in front of / on / under ... (lit. of) the friend'

As can be seen in (7), the presence of the genitive (with article) is optional on inanimate nouns, but obligatory on animate nouns (cf. **lagun atzean).[59] If we consider *-gan* to be a postposition *per se*, we should expect to find the same behavior, and in fact we do find it. With the inanimate nouns of Figure 6.4.a) —the first root being vowel-ending, the second one consonant-ending and the third *a*-ending—, the postposition is attached directly to the root, just as it is optionally in the examples in (7a). In Figure 6.4.b), with animate examples (*abokatu* 'lawyer', *lagun* 'friend', *turista* 'turist'), *-gan* maintains the initial consonant because this is not in intervocalic position.

      a.   **mendi-gan*   >   *mendi-an*

            *jolas-e-gan   >   *jolas-e-an*

            *gona-gan   >   *gonaan*   >   *gonan*

      b.   *abokatu-aren-gan*

            lagun-aren-gan

            turista-ren-gan[60]

Figure 6.4.   Evolution of -V-gan *as opposed to* -(ar)en-gan *according to de Rijk*

According to de Rijk, the forms without the genitive but with an article in the animate paradigm (*abokatuagan*, *lagunagan*, *turistagan*), perfectly possible today, must be later innovations, because they maintain the intervocalic voiced velar fricative. As will be shown below, if we consider the text tradition, this view proves untenable, since the type ines. sg. *lagun-aren-gan* does not appear in the oldest texts.

---

59   However, Michelena (1985 [1976]: 381) notices an exceptional case where the genitive is not overtly marked: Lç. (*Mk* IX, 24) *haur aitá-k* 'the child's father (erg. indef.)'. As he observes, it is not easy to establish whether this is a simple slip of the regular inflected *haurr-aren aitá-k*, or it represents a remnant of a stage at which such a construction was more productive.

60   In this case, the root-final *-a* and the article have coalesced.

In Jacobsen's view (1977: 166-7), the inessive marker -n found in proper nouns of place (cf. Table 6.3) would be secondary. After the loss of -g- between vowels, speakers interpreted the -a- of the inessive as a marker of sg. article. As a result, -n was reinterpreted as the inessive marker proper. Proper nouns of place would use this form "without article" due to their semantic determined nature.

De Rijk (1998 [1981]: 217) disagrees on this point, and believes that the inessive -n of proper nouns of place is old. He dismisses the interpretation of -an as article + -n because there is no trace of the expected forms prior to that reinterpretation (like e.g. **Iruri-an 'in Iruri'). In fact, the oldest attestation of a proper place name in the inessive dates back to the 11th century, and in that attestation, as well as ever since then, the inessive marker has been -n (with epenthetic -e- on consonant-ending roots), with -a- never showing up. Moreover, de Rijk notes that to consider -a- to be the article would run contrary to the whole argumentation of Jacobsen himself to prove that there was an stop (*-Can) previous to that -a- which, precisely, would prevent that process. I believe that de Rijk is right on this particular point.

By contrast, de Rijk believes that while -gan was the case form to mark definite nouns or NPs, -n was the one to mark indefinite ones. He splits -gan, then, into -ga + -n, whereby the latter would be the inessive marker and the former a morpheme *-ga (ibid.: 218).

Where would this *-ga- have come from? In de Rijk's opinion, it would be the locative morpheme -aga-, from which the plural article would have grammaticalized.[61] Hence, in de Rijk's view this suffix has little to do with the notion of plurality, and it would be a morpheme with the function of making locative those nouns which are not locative *per se*, i.e. all nouns except proper nouns of place, which are intrinsically locative. A place name like *Iruri* would take the inessive case with no need to mark its locative quality (*Iruri-n*). By contrast, common nouns, non-locative in themselves, would mark their extrinsic locative nature by means of this morpheme (*etxe-(g)a-n*, *lagun-aren-ga-n*, etc.).

A third view, within the hypothesis that the inessive -an comes from -gan, is the one set forth by Trask. Let us recall (cf. §6.3.2) that, according to him, in a certain period the local cases lacked morphologized number opposition. By contrast, the grammatical cases did have this opposition, and they would have passed this feature on to the local ones. Trask (*HB*: 204), then, sets out the following chronology:

---

61 Let us recall that he proposed the evolution *-ag-a > *-ag > -ak*, even if Artiagoitia, Lakarra or Martínez-Areta have argued for the unfeasibility of this segmentation (cf. §6.3.1).

1) Before the grammaticalization of the article and the consequent number distinction, the general locative morpheme was *-n*. By then, the grammatical cases did already have a morphologized number opposition.

2) A new locative is built, by means of the postposition *-gan*, to denote the locative notion on nouns for which this notion is not intrinsic (all except proper nouns of place).

3) On animate NPs, number is marked by the article (+ genitive) (cf. sg. *lagun-a-ren-gan*). In local cases of animates, the element *-gan* moves on to be reinterpreted as sg., and *-g-* drops out (**mendi-gan > mendi-an* 'on the mountain').

4) What in origin was a collective suffix *-eta- + -n* moves on to be interpreted as a pl. locative.

5) The *-e-* of *-eta-* is reanalyzed as the same *-e-* as that of the pl. of grammatical cases (cf. *lagun-e-i* 'to the friends'). As a result, *-ta-* turns into a marker of indefinite forms.[62]

There are, however, other views which dispense with the assumption that the ines. *-a-n* comes from **-ga-n*. In §6.3.2 it was shown that Manterola (2009b) argued that article formation consisted in the 3rd grade demonstrative, already declined, being attached to the root. Manterola also casts doubt on the view that the initial consonant of the inessive marker was *g-*. In his view, the consonant of the inessive **-Can* would be the aspiration inherent to the demonstrative. In other words, the morpheme *-han* would consist of the 3rd grade demonstrative plus the inessive morpheme *-n* (*ha-n*). As we have seen, the aspiration later dropped out. In order to explain the inessive of consonant-ending roots, Manterola's analysis can be represented as in Figure 6.5.

a.	*lan ha*	>	*lan-ha*	>	*lan-a*
b.	*lan ha-n*	>	*lan-e-ha-n*	>	*lan-e-a-n*

Figure 6.5.   Development of C-ending roots according to Manterola

Figure 6.5.a) shows the grammaticalization of the absolutive grammatical case with the word *lan* 'work', while Figure 6.5.b) does the same with the inessive grammatical case. As I have said, the grammaticalization of grammatical cases most probably took place earlier than that of local cases (Michelena 1987 [1981]: 48-9, Trask *HB*: 202). Figure 6.5 displays one difference between the two sets. Grammatical cases on consonant-ending roots (cf. Figure 6.5.a)) are attached in their aspirated form, and later the aspiration drops out, thus giving

---

62   As we can see, in this account Trask ignores the western variants erg. pl. *lagun-a-k* and dat. pl. *lagun-a-i*, which present *-a-* instead of *-e-*.

rise to the current form. The place names referred to by Manterola (cf. §6.3.1) seem to attest this fact. That is why he dispenses with the asterisk to represent the form *lan-ha*. In local cases (cf. Figure 6.5.b)), by contrast, when *-ha-* is attached to the root, an epenthetic *-e-* has to be inserted to prevent a clash between two consonants (in this case *n-h*), just as in the other local cases, and in the locative itself according to the hypothesis starting from *-gan*.

If Manterola is right, a) and b) of Figure 6.5 correspond to two different grammaticalization periods. a) would have taken place in an earlier period, where the clash *-n-h-* did not trigger epenthesis of *-e-*. b), by contrast, would correspond to a later period, in which for some reason *-n-h-* did trigger epenthesis of *-e-*.[63]

Manterola's account of the inessive *-an* is actually a return to the views held previous to Jacobsen (1977), according to which the *-a* of *-an* is nothing but the article ($< 3^{rd}$ grade demonstrative). It was already defended by Schuchardt (1947 [1923]), Lafon (1999 [1970]) and Castaños-Garay (1979: 139). According to these authors, the *-a-* of the inessive (cf. *iturri-a-n* 'in the fountain') would simply be the article, incorporated into the inessive marker *-n* at a late stage by analogy with the grammatical cases, which do have *-a-*.[64] The main advantage of this approach is that it allows us to dispense with the obligation of reconstructing any morpheme or phoneme. In fact, this account has some other advantages, which I shall later detail. In Manterola's account, however, the grammaticalization of the ines. sg. would have taken place when the element attached to *lan-* still presented the form *ha-*.

The assumption that the *-a-* of ines. sg. *-a-n* is simply the article and not a remnant of a morpheme *-ga-*, allows us to dispose of a problem which Jacobsen and his followers, and particularly de Rijk, systematically overlook. It is not easy to understand why the intervocalic *-g-* should systematically drop out in **etxe-gan > etxe-an* 'in the house', but instead should always remain untouched

---

63 Why should this be the case? Manterola (2012) offers two different possible explanations: "An answer for the differences between absolutive and locative definite phrases has already been given by Michelena (1987 [1981a]: 48-49) [...]: he suggests that absolutives attached to the noun earlier than locatives (and plurals), and this difference in time would explain some differences in their respective behaviors". And two paragraphs below: "[...] another possible answer coming to my mind is that both demonstratives, the absolutive *ha* 'that' and the locative *han* 'in that' simply followed different grammaticalization paths, so maybe we don't really need to explain both on the same grounds. This question, anyway, remains open".

64 Such encroachments of morphological features between grammatical and local cases are not unparalleled within the Basque declensional system. As we saw in §6.4.1.4, we have recent attestations of the instr. pl. as *etxe-etaz*, instead of the common *etxe-ez*. Here, the pluralizing *-eta-* of local cases has spread to a grammatical case.

in the sg. forms of local cases with animates like *gizon-a-gan* 'in the man'. To explain this, as we have seen, de Rijk (1999 [1981]: 217) resorts to the argument that this occurred because the variant *gizon-aren-gan* (with *-ga-* attached to the genitive) regularly preserved the *-g-*, and analogically influenced *gizon-a-gan*. However, the *gizon-aren-gan* type appears to be a late innovation. It already appears in Ber. (*Jaungoiko-aren-gan-a* 'towards the Lord'), and, in the plural and the indef., in Pouv. (*gizon-en-gan-a* 'to(wards) the men', *zein emazte-ren-gan-ik?* 'from what woman?').[65] But in both western and eastern old texts of the 16th century we only find the *gizon-a-gan* type.

An apparent advantage of deriving ines. sg. *-an* from **-ga-n* was that it provided an explanation for the fact that in Old Biscayan and Alavese, *-a*-ending roots are not dissimilated when they receive the ines. sg. case (Laz. *ene plaza-an* 'in my square'), unlike in the other grammatical cases (abs. sg. *plazea*, erg. sg. *plazeak*, etc.). The dropping out of *-g-* would have occurred when the dissimilating rule was no longer operative, hence **plaza-gan > plaza-an*. However, this hardly entails any advantage over the interpretation of the *-a-* of *-an* as an analogical extension of the article, since within this explanation we can likewise assume that this extension took place when the dissimilating rule no longer applied.

---

65 In this author, however, the singular of such forms is still exclusively formed by attaching *-ga-* to the abs. sg.: *Iainko-a-gan-ik* 'from God', *Espiritu Saindu-a-gan-ik* 'from the Holy Spirit', *bere seme-a-gan-a* 'to(wards) his son' etc. The *-(r)en-ga-* structure (*-ga-* upon the genitive), then, appears to have entered the noun declension through the plural and indefinite. It is important to bear in mind that this structure was already found in demonstratives (the 3rd grade forms of which are also 3rd person pronouns), even in the singular: Pouv. *har-en-gan-ik* 'from him', *har-en-gan-a* 'to(wards) him'. This already occurs in Lç. (*har-en-gan-a* 'to(wards) him', etc.), who, in the sg. of the noun declension, only makes use of the *-a-ga-* structure (*-ga-* upon the absolutive, as in *senharr-a-gan-ik* 'from the husband', *Iainko-a-gan-a* 'to(wards) God', etc.). In the plural, this author forms the local cases of animates not by means of *-en-ga-*, but by simply using the usual declension of inanimates: Lç. *gizon-etarik* 'from the men', *Profet-etan* 'in the Prophets', etc. This is most probably the old pattern for the formation of the plural of animates, in which case the *-en-ga-* structure must be an innovation. The following quotation, taken from Creissels & Mounole (2011: 177-8), summarizes how this grammaticalization may have occurred: "Given the tendencies generally observed in phenomena sensitive to animacy hierarchy, it is not surprising that, before generalizing to all animate NPs, the use of *gan* or *baita* in contexts triggering the use of spatial cases was already more systematic with pronouns that with nouns. The fact that the use of *gan* or *baita* was already regular with nouns in the definite singular at stage of the evolution when direct affixation was still widespread with plural or indefinite nouns is not surprising either. (...) [T]he use of *gan* or *baita* became general in the allative and ablative cases spreading to the locative".

Another advantage of this account is that, unlike the interpretation of *-an* < **-gan*, it provides an explanation for the forms with the inessive in *-n* residually attested in Souletin, such as *etxe-n* 'at home', *gañ-e-n* 'on top', *soñ-e-n* 'on the body', instead of the common *etxe-a-n*, *gañ-e-a-n*, *soñ-e-a-n*, and also for B. *itaurre-n* 'leading oxen', *bat-e-n* 'in one', RS. *orbide-n* 'in the journey', RS. *zaarza gatx-en* 'in bad old age', instead of the corresponding ones with *-e-an*. That these are archaisms not reached by the extension of the article, perhaps because of their adverbial nature, was already defended by Lafon (1999 [1970]: 169). It should also be noted that they belong to marginal dialects, which are typically the last to be reached by innovating tendencies. We have, however, at least two contexts in which the archaic inessive *-n* has not been reached by the extension of the article and which appear in all dialects: 1) the imperfective suffix *-t(z)e-n* (as in west. *esa-te-n*, east. *erra(i)-te-n* 'to say (ipfv.)'), formed upon the verbal nominalizer, and 2) the proper nouns of place (as in *Elorrio-n* 'in Elorrio', *Gasteiz-e-n* 'in Gasteiz', etc.). In both cases, the non-extension of the article is expected. In the former, this is because the imperfective form is not found in a noun paradigm within which *-a* can spread.[66] In the latter, it is because proper nouns are definite *per se*, and hence never take the article.

Those who defend, by contrast, that the *-a-* of ines. sg. *-an* is not the article are obliged to explain these data as innovations. For the pair S. *étxe-n / etxí-an*, Trask (*HB*: 204) ventures to propose an influence from Spanish, where two forms like *en casa* and *en la casa* are differentiated, with the former containing a possessive, even affective meaning, implying 'in my/his/her house'. This explanation, however, is clearly not valid for some of the data above.

A final point to mention concerns the ines. sg. of consonant-ending nouns, which points to *-e-an* in all dialects. Even if Manterola's explanation in Figure 6.5.b) above is plausible, another possibility would be that the extension of the article occurred after the insertion of the epenthetic *-e-*. This would have been inserted to prevent not the clash *lan-* + *-han*, but the clash *lan-* + *-n*, hence → **lan-e-n*, and later → *lan-e-a-n*. This account has the advantage that the stage **lan-e-n* is actually attested in S. *gañ-e-n*, B. *gatx-e-n*, etc. An additional advantage is that we would not require any explanation as to why *-e-* is inserted in ines. *lan-han*, but not in abs. *lan-ha*.

Obviously, the difference between the two options implies a difference in chronology. While in the development of Figure 6.5.b), ines. sg. *lan-e-(h)an*

---

66   As a matter of fact, in some western varieties *-t(z)ean* is found as the imperfective suffix (as already attested in the Chant of *The Burning of Arrasate-Mondragón*; cf. *TAV*: 81-88). To be more exact, then, we should say, instead, that it is not as expected in this context as it would be within a noun paradigm.

would have arisen simultaneously with the creation of the noun paradigm, the one hinted at in this section would have taken place later, due to an analogical extension of *-a* within a noun paradigm that was already formed.

6.4.2.2. The allative

The allative case, typically employed to mark the target of a movement (cf. *mendi-ra* 'to the mountain'), displays some degree of polymorphism. The most common variant is *-ra*, but in some eastern varieties we find a variant with final stop, *-rat*, and also *-ara*, and in still more eastern varieties, we have *-la(t)* and *-ala(t)*. Contemporary forms such as *bidi-la(t)* (*bide* 'path') come from *bidi-ala(t)*. As is usual in morpho-phonological processes in which a noun declensional suffix is attached to a vowel-ending root, the change *-e-* > *-i-* has been brought about by the juncture *bide-* + *-ala(t)* > *bidi-ala(t)*, with the *-a-* finally dropping out.

Some Bascologists like Gavel (1929: II, 28-9) have stated that there was originally a semantic differentiation between the forms with final stop and those without. Thus, according to Michelena (1979c: 577; with examples from Lç., Ax. and Bp.) and Igartua (1986; with examples from Lç.), *etxe-ra* would be 'to the house', but *etxe-rat* 'towards the house'. Some other authors like Lafitte, Gèze and Txillardegi assign the variant with final stop a more terminative than directional sense, or ascribe a briefer duration to the variant without a stop (apud Igartua 1986). While the issue has not been thoroughly researched, I think that nowadays few or no speakers make that distinction by means of the stop. At least since Schuchardt (1947 [1923]: 56), though, it is clear that it is an independent morpheme.

In the varieties which make use of the forms with *-l-*, these appear in complementary distribution with the forms with simple rhotic. As with the inessive (cf. §6.4.2.1), it is proper nouns of place that behave differently: they form the allative with *-ra(t)*, whereas common names use *-la(t)* (cf. *eliza-lat* 'to the church' vs. *Maule-(r)at* 'to Maule').[67]

De Rijk (1998 [1981]: 219) believes that, if the morpheme **-ga* of the inessive is employed to endow common nouns with locative meaning (cf. §6.4.2.1), the other local cases —e.g. the allative— must have it too. According to him, the

---

[67] It is interesting to note that, in Souletin, the variant with rhotic is usually used for the word *etxe* 'house', even though it is not a proper noun of place. Furthermore, the distinction *etxe-rat / etxi-alat* entails the same semantic distinction as the one that I mentioned (§6.4.2.1) in relation to *etxe-n* 'at home' / *etxí-an* 'in the house', in the inessive. Thus, *etxe-rat* means 'homeward', whereas *etxi-alat* means 'to the house'.

-a- that appears as the first segment of some variants, but never on place names, corresponds to this morpheme. To support this view, he puts forward data from Roncalese. In (8a), the locative morpheme is not necessary because we have a place name, and hence the allative case marker is directly attached to the root, with epenthesis of *-e-* to prevent a clash between two consonants. In (8b), the example is a common noun and thus the alleged locative marker would be necessary. The velar fricative drops out in intervocalic position, as usual, and the epenthesis would dissimilate by the closing of *-e-* into *-i-*, before *-a-*. (8c) shows a similar case, this time in Souletin, where the variant of the suffix has *-l-*:

(8)   a.   *Jerusalem-e-ra*

            Jerusalem-EP-ALL

            'To Jerusalem'

       b.   *lurr-i-a-ra*   (< *lurr-e-ga-ra*)

            earth-EP-LCV-ALL

            'To the earth'

       c.   *oihan-i-a-la*

            forest-EP-LCV-ALL

            'To the forest'

     (8b) and (8c) differ in that (8b) has *-r-*, whereas (8c) has *-l-*. De Rijk (1998 [1981]) explains this difference on phonetic grounds. In Basque, there is a common development whereby $l > r$ / V _ V (*FHV*: 311-2).[68] It can be noticed in loan words such as *zeru* 'sky' (< Lat. *caelum*), or *borondate* (< Lat. *voluntātem*). Thus, the variant with the rhotic in (8b) might be the result of *-ala > -ara*. To explain away the variant in (8c), where the change has not occurred, in that morpheme de Rijk proposes the presence of an old fortis **-L-*, which would naturally evolve into *-l-*. Hence, we would have something like **-(g)aLa*. The origin of that fortis **-L-* might be searched for in the juncture of two consonants, which according to de Rijk are *n* and *l*. Thus, if for the *lurriara* of (8b) he reconstructs **lurr-e-ga-la > *lurr-e-ga-ra*, for *lurriala* he proposes **lurr-e-ga-n-la > *lurr-e-ga-La*. That is, the locative marker **-ga-* would have been reinforced by the inessive marker *-n*, and this occurred only dialectally in the varieties with *-l-* in the allative.

---

68   Michelena even mentions the possibility, with reservations, that all simple rhotics are secondary (ibid.: 327-8).

Trask (*HB*: 205-6) has another view of the facts. Apart from the variants mentioned above, the allative case has another one: *-a*. This appears on place names —which are typically conservative— ending in a consonant, such as *Zarautz-a* 'to Zarautz' and *Irun-a* 'to Irun' (cf. Table 6.3), on adverbials derived from demonstratives such as *hon-a* 'hither', *har-a* 'thither', and after the morpheme *-gan-* of local cases of animates (cf. Table 6.4): *lagun(-aren)-gan-a*, etc. In Trask's opinion, these are conservative contexts, and hence *-ra* is an innovation. When the article was grammaticalized, a form such as *mendi-a* became ambiguous, with *-a* now having two possible interpretations: as the article or as the allative morpheme. In other words, *mendi-a* can be understood as 'the mountain' or as 'to the mountain'. This ambiguity could be removed by inserting an epenthetic *-r-*,[69] but demonstratives and proper nouns of place, which did not take the article, preserved the old form. A factor which may have contributed to the change all. *-a → -ra* is the adverbial *hara* 'thither', in origin *har-a*, but perhaps subject to reinterpretation as *ha-ra*. In the forms like *mendi-ala* 'to the mountain' (vs. *mendi-ra*), Trask, unlike de Rijk, does not see the morpheme **-ga-*, and considers them, as I do, to be innovations in which the article has been introduced. As for *-la*, he suggests that it might be an unknown morpheme, secondarily used as allative.

As I have said, in eastern varieties there are variants of the allative with final *-t*. This appears to imply a semantic difference. We have already seen (§6.3.1) that an old stop in word-final position is not possible, and that the ones which exist historically in grammatical morphemes can be derived from a structure DV, after apocope and devoicing of the stop. If this is the case, *-t* can be reconstructed as **-da*.

The origin of this morpheme is a mystery (*HB*: 206). Preposed or postponed, it generates a prolific set of etymological combinations, beginning with the *da-* of the verbal morphology. De Rijk (1995a) reconstructs a morpheme **-dan*, which he postulates to be an old patrimonial morpheme, and assigns it a meaning of 'now'. It would also appear in adverbials like *jada* 'already now', (*ja* 'already then' + **-da* 'now'), and, redundantly and with the inessive *-n*, in adverbials such as *orain-da-n-ik* 'since now' (ibid.: 364). It also appears by analogy with this —but in other contexts due to their loss of meaning— in other adverbials such as *betidanik* 'since always' (ibid.: 369), or *noizdanik* 'since when?' (see §9.4.1.1.1 for further details on this morpheme).

Although the formal equivalence with the *-t* of *-rat* would be perfect, de Rijk does not mention this morpheme in the set of etymological combinations in

---

69 Just as it occurs in Salazarese with the article when it is attached to *a*-ending roots (*alaba- + -a > alaba-r-a* 'the daughter'; cf. §6.4.2.1).

which *da* takes part, and actually the semantic idea behind an alleged develop-
ment 'now' → 'towards' is not easy to work out. However, Lakarra (2008a:
459) does include the *-t* of *-rat* within the set of morphemes which could be re-
lated to the *da-* of the verbal morphology, although his interpretation of this *da-*
is different.

Lakarra's (2006a, 2006b, 2008a) view of the allative is highly interesting.
The key for exploring it is his criticism of de Rijk's conception of *da(n)* as an
element meaning 'now'. According to him, there are several problems with this
assumption. To begin with, on phonetic grounds we cannot have an independent
morpheme beginning with *d-* at a very old period, since as Michelena (*FHV*:
257) observed, the natural tendency of *d-* is to become *l-*, so that we would ex-
pect **la-tor*, instead of the historical *da-tor* '(s)he comes, is coming'. On the
other hand, on semantic grounds Lakarra (2006a: 598) observes that as late as
the 16th century *da-kar* is not '(s)he is bringing' but rather '(s)he will bring' in
plenty of examples already put forward by Lafon. Hence, following Trask
(1977), he ventures to propose that *da-* (~ *-da*) is an old imperfective marker,
which only in a recent period has undergone the usual semantic development
imperfective → present. As for its lexical source, Lakarra proposes the Proto-
Basque root *daR* 'to sit', unproblematically present as the root of the participle
jarri 'to put, to sit', < *e-daR-i*. As Newman (ed.; 2002) and Heine & Kuteva
(2002: 276-278) show, this lexeme is a common source of both locative markers
and imperfective (~ progressive) markers.

According to Lakarra, in addition to the preposed *da-* of the verbal mor-
phology, there are two more *da*. All three of them would come from *daR* 'to
sit', but the last two would have ended up as postpositions, and finally as suffix-
es. The first of these *da*, of local meaning, would be the one which grammatical-
izes as allative. In order to explain its postpositional nature, and at the same time
its phonetic development, Lakarra (2006b: 602) resorts to a morpho-syntactic
explanation. Starting from a period in which the verb preceded the object (cf.
§10.4.3), we would first have a structure [*da*-V O], which after getting reversed
would become [O *da*-V]. At this point the change *d-* > *l-* applied, hence > [O
la-V], but then the reconstructed morpheme underwent an "excorporation"
from the verb, → [O *la* V], and later *la* was incorporated as a postposition, and
finally as a suffix to the O, hence → [O-*la* V]. From this construction, the all.
sg. form *-ra* would be phonetically regular. The variant S. *-ala* could be ex-
plained —as de Rijk did (cf. §6.4.2.2)— by assuming that, in the structure [O-
la V] which grammaticalized, the O was in the ines. sg., so that the juncture

-an-la would have generated a context from which a fortis arose, > *-alla-* (i.e. *-aLa-*), and hence > *-ala*.[70]

The other **da* which has ended up as a postposition would be that of *etxe-rat* (< **-ra-da*). This would have grammaticalized as imperfective, just like the *da-* of the verbal morphology, and thus would assign the allative suffix an imperfective, durative semantics. This would explain the difference between *-ra* 'to' and *-rat* 'towards'. How the transformation of a prefix into a suffix has taken place in this case, whether through excorporation and later incorporation, or through "migration" from the word-initial position to the word-final one (as proposed in Lakarra 2006a: 602-606), remains to be clarified. Be it as it may, Lakarra insists that the transformation of a prefix into a suffix is anything but unparalleled in Basque, as happens with *gi-* → *-gi* / *-ki* 'thing, stuff, piece', **de-* / *le-* → *-de* (verbal pluralizer and potential marker, whether different or not in origin), *za-* → *-za* '~ abundantial', and probably some others (cf. Lakarra 2006a: 599-606). The phenomenon is also paralleled in Munda, in several Australian languages, etc.

6.4.2.3. The ablative (and prosecutive)

This local case, which today mainly implies 'from', i.e. 'place from which' in relation to a movement, is also rather polymorphic. Nowadays, the most widespread form is *-tik*, but there also exist other forms, such as *-ti*, *-tika*, *-tikan*, and *-(r)ik*. In its formation, the case marker is directly attached to the root, with no article. From de Rijk's point of view (1998 [1981]: 219), this would be the only local case for which the locative morpheme **-ga* can never be reconstructed for any of the variants, since here the case marker is always attached to the root, with no *-a-* between the two. In Old Biscayan and Alavese texts, another ablative suffix, namely *-rean*, is well attested, together with its pleonastic variant *-reanik*.

Let us begin with the most perceptible difference, which is that between *-rean* and the other morphemes. As Lakarra (1984a) shows, in western texts

---

70 The power of such a morphosyntactic explanation, however, must be taken with care, since in any case it would not be equally applicable to the whole declensional system. The origin of the cases (whether they come from postpositions, serial verbs, or any other element), then, must be studied individually. Nonetheless, precisely this impossibility of deriving all cases from the same construction can help us to establish a relative chronology of their grammaticalization.

from the 17th century until the disappearance of *-rean* in the 18th century,[71] the two ablative morphemes appear in complementary distribution. Thus, in the oldest texts in western dialects until the 18th century, the ablative case is expressed either by *-rean* or by *-(r)ik*,[72] but never by *-ti(k)*, as the latter does not have an ablative, but a prosecutive meaning. It implies 'through', in relation to a movement. Hence, *mendi-rik* or *mendi-rean* meant 'from the mountain', but *mendi-ti* 'through the mountain'. In fact, in Old Biscayan and Lazarraga's Alavese, the prosecutive —which also inflects for plural under the form *-etati*—[73] is a fully-fledged case different from the ablative.

It is from the 18th century onwards that *-ti(k)* begins to lose its original meaning and encroach on the realm of *-(r)ik*, which will progressively disappear with this sense. We can establish the relative chronology shown in Figure 6.6.

*Stage 1	Stage 2	Stage 3	Stage 4
Abl. *-(r)ik*	Abl. *-rean* / *-(r)ik*	Abl. *-rean* / *-ti(k)*	Abl. *-ti(k)*
Pros. *-ti(k)*	Pros. *-ti(k)*	Pros. *-ti(k)*	Pros. *-ti(k)*

Figure 6.6.   Chronology of the grammaticalization of the ablative in the west

Stage 1 is not attested, but it is my assumption, partially based on Lakarra (ibid.: 189-190), that *-rean* is an innovation competing with *-(r)ik* from Stage 2 onwards, even if Michelena (1987 [1981]: 44) believed that it was common Basque. Two reasons for these assumptions are the fact that *-(r)ik* appears in place names, which are typically archaizing, and the limited geographical distribution of the morpheme *-rean*.

In 16th and 17th century texts (those belonging to Stage 2), *-rean* is the usual ablative in the west, except on animates, which use *-(gan)ik*. If we leave aside Cap., *-rean* does not appear when the ablative has a temporal meaning either, as in *one-rik aurre-ra* 'from now on', nor when the ablative is attached to a proper noun of place, as in Mic. *Flandes-ik* 'from Flanders'. In the 17th century, a new pleonastic form appears: *-rean-ik*. Already in the 18th century (Stage 3), *-rean* loses ground, and at around the middle of the century it disappears. *-reanik* will continue to survive for one more century.

Also at Stage 3, *-(r)ik* is progressively replaced by the prosecutive morpheme, which keeps its original meaning but, in addition to that, it takes over the

---

71 From then onwards, it has only survived in archaizing sources such as ballads, which typically maintain structures that have disappeared from the spoken language; *-rean* appears in some of these ballads gathered at the beginning of the 20th century.

72 This is the form still used today for the partitive.

73 As in Cap.: pros. sg. *pensamentu-ti* 'through the thought', pl. *esku-etati* 'through the hands'; Laz.: pros. sg. *bioz-e-ti* 'through the heart', pl. *odei-etati* 'through the clouds'.

ablative function, thus becoming an ablative-prosecutive morpheme. It appears that, on one hand within animate roots, and on the other hand within the plural of any root, *-(r)ik* better withstood the invasion of the prosecutive morpheme. Thus, in 1802, Moguel still preserves forms with *-(r)ik* for these paradigms, while the old prosecutive morpheme, now turned into an ablative-prosecutive one, occupies the whole inanimate sg. paradigm. Gavel (1929: I, 55) believes that the pl. *-etarik* may have arisen as a result of the dissimilation of two stops (*-etatik > -etarik*). However, this is hardly plausible, if we consider that in this case the plural appears to have been more conservative, judging from Moguel's attestations. In the other battle, fought between *-rean* and *-(r)ik*, we should further note that *-rean* lost its scope earlier in the plural than in the singular.

Finally, Stage 4 is the one found nowadays in the west. In this phase, the old prosecutive morpheme *-tik*, which over the course of time has become ablative-prosecutive, is the only option available to mark the ablative. As in other regions of the Basque Country, prosecutive notions can also be indicated by means of postpositions after the root in the inessive case, as in *mendi-an zehar* 'through the mountain'.

In the east, the situation is somewhat different. We would originally have had an ablative in *-(r)ik*. When *-ti* enters the declensional system, it does not only acquire a prosecutive value, as in the west, but it also replaces *-(r)ik* in its ablative function, and relegates it —in this case, as in the west— to the plural paradigm (*etxe-etarik* 'from the houses'), place names (S. *Maulé-(r)ik* 'from Maule') and some residual items (Tt. *bekhatu-rik kanpo-an* 'out of sin'). Perhaps the existence of the ablative *-rean* in the west, alongside *-(r)ik*, could slow down the supplantation of *-(r)ik* by *-ti(k)*, which first assumed a prosecutive function (cf. Figure 6.6), contrarily to what happened in the east. It should be recalled that, while in the western RS. *-ti* is only prosecutive and the appearances of *-(r)ik* are ablative (obviously, in addition to those of *-rean*), *-ti* is ablative in old eastern attestations such as Oih. *azi-ti* '~ from the seed' (cf. (9) below) or Tt. *gaiñ-e-ti behe-ra* 'from top to bottom'.

There is great controversy as to the origin and grammaticalization process of these morphemes. Although Uhlenbeck (1909/10 [1903]: 107) postulated an evolution *-tik > -ti*, Lafon (1999 [1948]: 199) is right when he states that there are no grounds for such an assumption. Michelena (*FHV*: 236) is of the same opinion as Lafon, and argues that *-(r)ik* has no variant without a final stop (as expected if the alternation between *-tik* and *-ti* were motivated on phonetic grounds).

Schuchardt (1947 [1923]: 51), in turn, proposed that *-tik* was the juncture of *-ti* plus *-(r)ik*. He related the former element to the *-ti* denoting 'place' in words like *mahasti* 'vineyard' (from *mahats* 'grape'), which would come from the suf-

fix *-tegi*, similar in meaning. As for the latter element, he related it to the erga-tive *-k*, without giving further details. This would not account for the variant *-tika*, already present in Dch. (*zeru-tika* 'from heaven') and Laz. (*Gaztela-tika* 'from Castile'), nor for the variant G. *-tikan*, derived from it.

In fact, the oldest forms appear to be *-ti* (the pure abl.(/pros.) suffix) and *-tika* (*-ti* with the enlargement *-ka*, probably an adverbial suffix, see below), *-tik* being an apocopated variant of the latter. In addition to this morphological ar-gument, *-tik* must be secondary, also on phonological grounds, since a stop in the coda of a syllable is necessarily extrametrical, and hence impossible in Pro-to-Basque (Artiagoitia 1990).

With regard to the origin of *-ti*, Lafon (1999 [1948]: 200-2) observed that although *-ti* could be phonetically derived from *-tegi*, this derivation was not jus-tified on semantic grounds. He preferred to consider *-ti* to be the same deriva-tional morpheme as that which can be seen on adjectives such as *beldur-ti* 'fear-ful' (from *beldur* 'fear'), which eventually entered the declensional paradigm. But how? Lafon proposes that this may have occurred after the conversion of an apposition of the subject into a complement of the verb. In fact, he even suggests that this conversion is under way in some attestations, such as in Oih.'s Proverb 44 (9a):

(9)   a.   *Azi-ti   bihi-a*   (Oih., *Prov.* 44)

           seed-'ti'   grain-ART

           'The grain comes just as the seed is (= just as the seed is, so will the grain be)'

      b.   *Itsu-a   da   bahe-ti   ez-takusa-n-a*   (Oih., *Prov.* 288)

           blind-ART   is   sieve-'ti'   not-sees-RLV-ART

           'Blind is the one who does not see through the sieve'

A 21st century Basque speaker would translate (9a) simply as 'From the seed (comes) the grain'. But Oihenart's French translation is the following: 'Le grain vient *tel qu'estoit* la semence' [the italics are mine, ES]. What the Souletin au-thor still interprets as an apposition of the subject is apparently taking on predic-ative prominence, until eventually it has become a verbal complement, semanti-cally ablative (here, of an implicit 'comes'). This is the syntactic path through which the derivational *-ti* would have entered the declensional system, according to Lafon, followed by Michelena (*FHV*: 236), but not by Trask (*HB*: 205). In (9b), by contrast, Oihenart's translation is: 'Celuy la est aveugle, qui ne voit pas au trauers d'vn crible'. In this case, then, *bahe-ti* has already become a verbal complement, now of prosecutive meaning, but Lafon (1999 [1948]: 203) be-

lieves that it must have been, in origin, a derived adjective —just like *beldur-ti*— in apposition to *ez-takusana* (the subject).[74]

It is important to note that, in Lafon's account, once *-ti* had entered the noun paradigm, it adapted to the morpho-phonological processes of this paradigm. Thus, when abl. sg. *-ti* is attached to a consonant-ending root, an *-e-* is inserted: *lurr-e-ti* 'from the earth', just as all. sg. *lurr-e-ra* (as opposed to *beldur-ti*, where *-ti* is derivational).

As for the *-k* present in the variant *-tik*, Lafon believes that it has been inserted into the ablative by analogy with the variant *-(r)ik*. Michelena (*FHV*: 237), by contrast, considers *-k* to be an apocopated variant of *-ka*, this being the same morpheme employed to form iterative adverbs such as *irri-ka* 'laughing' or *korri-ka* 'running'.

If a reasonable source of abl.-pros. *-ti* is the derivational *-ti* (*beldur-ti*, etc.), as the next step we can wonder where this derivational *-ti* might come from. Lakarra (2012b) proposes as a source the verbal root **din* 'to come' (east. *jin* < **e-din* 'to come'), which could also mean 'to become' if we consider the verbal root of subjunctive intransitive auxiliaries *na-din*, *da-din*, etc. Thus, **din* with the meaning 'to come' might have grammaticalized as the partitive/ablative (cf. §6.6.3), whereas **din* with the meaning 'to become' might have grammaticalized as the derivational *-ti*. In fact, the two meanings 'to come' and 'to become' are related in both Romance (cf. Sp. *venir ~ devenir*) and Germanic languages (cf. Eng. *to come ~ to become*), and also in many more languages beyond European borders.

Lakarra (2012b) also believes that the derivational *-ti* coming from the root **din* would still be attested with a voiced stop, hence in a *fase sparita*, in forms with the structure CVC-*di*, as in *han-di* 'big', *hor-di* 'drunk', etc.,[75] which would literally be 'who has become an animal' and 'who has become a dog' (*hor* was 'dog' in old continental varieties, Lcc. and RS, and most probably also in CB). Furthermore, Michelena (apud Lakarra 2006a: 574 and 2012b) had already noted several decades ago that words like *ber-din* 'equal' (cf. *ber* 'same') and *gor-din* 'raw' (cf. **goR-i > gorri* 'red') still have the suffix *-din* with the meaning 'to become'. This can be interpreted as an even more conservative variant of **din* 'to become', or alternatively as a relative clause (*di-* + relative *-n*): *ber-din* 'which has become the same', etc. The (eventually prevailing) variant *-ti* would have been generalized from contexts in which *-d-* devoiced after a sibilant.

---

74 The original meaning of (9b), then, would have been something like 'The one who does not see "sievous" is blind'.

75 Lakarra here applies a principle taken from Trask (1995b), according to which the variant with voiced stop is earlier than the one with voiceless stop.

The morpheme *-rean* can be analyzed in two ways. It could either be the old form of the genitive *-(r)e* plus the ines. sg. *-an*, as it is according to Michelena (1987 [1981]) and Trask (*HB*: 205; in this case, < *-ga-n*), or the new form of the genitive *-(r)en*, to which the article has spread in a somewhat unorthodox manner, as Martínez-Areta (p.c.) suggests to me. In either case, the *-r-*, which, attached directly to the root, appears only with proper nouns or in the indef., and in any case only on vowel-ending roots, has morphologized as a part of this ablative (*etxe-rean*, *lurr-e-rean*). It appears, then, that once this abl. *-re(a)(n)* was formed, in order to get into the paradigm it acquired a pl. and an indef. form like the rest of the local cases, by means of the morpheme *-(e)ta*: Cap. *arerio-etarean* 'from the enemies' (he also uses *-etarik*, the distribution being unpredictable; cf. Lakarra 1984a: 178), *zeru-etarean* 'from the heavens' in some verses in the *Doctrina Christiana* by Zubia (1691) (known thanks to the copy made by Lezamis in 1699; cf. *TAV*: 138-139, 162).

In the latter analysis, the *-a-* of *-rean* would be an article of late incorporation, similar to that of the allative in Souletin, and most probably to the ines. *-an* of all dialects. Nonetheless, this ablative is attested as *-ren* in Lcc. (*one-ren* 'from here', *orre-ren* 'from there'; cf. Agud & Michelena 1958: 34), in some posters in verse exhibited in Tolosa in 1619 (*esku-ren* 'from the hand'; cf. *TAV*: 127-131), several times in a text from the beginning of the 17th century from Azkoitia (Guipuscoa), along with *-rean* and *-tik* (cf. Urquizu 2008-09), and even as an archaism in western forms like *liburu-ren bat* 'some book'. The lack of *-a-* in Lcc.'s attestations may be due to their adverbial nature. The *esku-ren* of Tolosa (in Guipuscoa), by contrast, is within the noun paradigm. If it is not an archaism without an article, it might be a dropping out of *-a-* (*-rean* > *-ren*), or even some kind of metrical license.

All this would imply that the source of the western ablative was the genitive. In fact, identical marking for the genitive and the ablative is not unusual, as the Romance languages show.

6.5. Secondary cases

In this section, I shall discuss some cases which were not included in the paradigms because they are created by being built upon other cases.

6.5.1. Cases built upon the allative

There are two cases which are built upon the allative. One is the directional allative, and the other is the terminative allative. The former can be roughly translated as 'towards', and the latter as 'up to'.

6.5.1.1. The directional allative

The directional allative has already been partially discussed (cf. §6.4.2.2), where it was pointed out that the variant *-rat* of the simple allative may have had precisely this meaning. But there are other options for expressing it. In many areas, this semantic nuance of proximity is expressed by postpositions in the allative case after the bare root, as in *mendi alde-ra* 'towards the mountain' (*alde* 'area'), or in some continental varieties with the postposition *buru-z* in instr. indef. attached to the root in all. or dat., as in *mendi-ra/-ari buru-z* 'towards the mountain' (*buru* 'head'). Another option, actually very recent, is to express the proximate idea by means of the iterative suffix *-ka* (mentioned in §6.4.2.3; cf. *korrika* 'running', *irri-ka* 'laughing') attached to the allative case (*mendi-ra-ka* 'towards the mountain').

In the west, the center and in part of Navarre, there is another suffix which covers the same function, and which has the variants *-rantz*, *-rontz*, and *-runtz*. Further variants, such as the one present in *mendi-rantza* 'towards the mountain', seem to be innovations which are formed by pleonastically attaching the allative *-a* to the proximate *-rantz*. Azkue (*MV*: 333), quoting the poet Nicolás Ormaetxea "Orixe" as the source of the idea, suggested that *-rantz* is simply all. *-ra + antz* 'face', and recalls that the development 'face' → 'towards' is common in the surrounding Romance languages (cf. Sp. *hacia* < Lat. *faciēs* 'face', or even High Aragonese expressions such as *voy cara al monte* 'I am going towards the mountain').[76] As de Rijk (1998 [1972a]: 121-123) observed, this explanation ignores the variants with *-o-* and *-u-*, and is at odds with the fact that *antz* with the meaning of 'face' is not attested in any dialect.

More recently, Gómez (2005) has reconstructed **-raontz*, from which the three attested variants could be derived.[77] It would be an agglutination of the allative *-ra* plus an old noun **-ontz*, which according to Gómez would be semantically close to Lat. *versus* 'facing → towards', and would have undergone the same grammaticalization. Hence, the Basque reconstruction might be a loan

---

76 In fact, directional expressions like *cara al monte* are also heard in some areas of Navarre.

77 Gómez mentioned more variants, but the others are clearly secondary.

translation of the Latin *ad versum* (ibid.: 276). The Basque word *ontzi* 'recipient (of a liquid)', in turn, would be a morphological derivation of that **-ontz*.

Lakarra (2008a) suggests that *ontzi* is a loan word (possibly from Lat. *conchea* 'shell-like, concave'), and prefers instead **-ra + dontz* (ibid.: 484), or even **-ra + *-da-non-tz*, as the source of the directional allative. This would be formed by the variant **-rada* (> *-rat*) of the allative (cf. §6.4.2.2) + the Proto-Basque root **non* + the derivational suffix *-tz*. The root **non* would have a meaning of movement which would also be present in *joan* 'to go' (< **e-da-non*) (cf. Lakarra 2006a: 583-4).

6.5.1.2. The terminative allative

The terminative allative appears as *-raño*, *-radino*, *-raino*, *-raindo*, *-d(a)rano*, *-giño*, etc. All of these have final *-o*. De Rijk (1995a: 297-302) believes that the most conservative of these variants is *-raindo*. In fact, forms attested in Navarre such as *bat-e-o* 'at the same time' (from *bat* 'one') could point to the old existence of a consonant which would justify the epenthesis of *-e-*. This consonant would be, according to de Rijk, the *-d-* which is found in *-raindo*, but also in words such as *oraindo* 'until now' (it also appears as *oraino*). **-do* would be, then, an old morpheme to express 'until, up to'. In the rest of the morpheme, in de Rijk's view we would have the word *gain* 'top', which, according to Jacobsen and Trask, is also present in the inessive (cf. §6.4.2.1). Consequently, we would start from the grammaticalization of an independent element **gaindo* (lit. 'up to the top'). As *gain* lost its lexical content, its initial consonant dropped out, thus giving rise to one of the attested secondary cases from the allative: **-ragaindo > *-raaindo > -raindo*. If we start from **-ragaindo*, all the variants mentioned above are easy to explain, except for two.

To explain *-giño*, residually attested in B, we would have to propose that **-gaindo* has been attached directly to the root, with no allative suffix. The second one, *-d(a)rano*, consists of *-da-* + **-ragaindo > -raindo > -raino > -raño > -rano*. This *-da-* would be, again, the same archaic morpheme discussed above (cf. §6.4.2.2), with the meaning of 'now', although, by the period of the grammaticalization of the whole suffix, it would already have lost this meaning (ibid.: 362). The desemantization of the morpheme *-da-*, as well as of *-gain-*, further gave rise to a high number of pleonastic forms in many authors, perhaps the most intricate one being Tt. (S, 17th century) *egün-ialadrano* 'until the day'. In this form, we have two already known allatives, *-ala-* and *-ra-*, with which the terminative postposition has agglutinated.

Let us conclude by considering a form which proves problematic for de Rijk's analysis. Dechepare (LN, 1545) has *oraidano* (today's spelling), which

like *oraino* means 'until now'. If the nasal comes from the erosion of *gain*, then this form should be reconstructed as **orai-da-gain-do*, where the second morpheme would be *-da-* 'now'. We would have, then, a redundant word consisting of the old and modern forms for 'now', just as we had in the above mentioned adverbial *orain-da-n-ik*. Notwithstanding, in the latter form, de Rijk considered the nasal to be the inessive marker suffixed to *-da-*, not to the erosion of *gain*. In order to explain Dechepare's form, I believe that it would be more economical for him to apply the same procedure, thus starting from something close to **orai-da-n-do*.[78]

Lakarra (2008a: 484 and 2012b) prefers to reconstruct *-ra* + **din* + **non*, with a serial verb construction after the all. suffix. The first would be the monosyllabic root **din* 'to (be)come' (also present in east. *jin* 'to come', < **e-din*), and the second the root **non*, implying both movement and a locative idea. The juncture **-din-non-* would create a fortis environment for the nasal which would prevent it from disappearing. If this were the case, we would be obliged to propose an antihiatic function for the velar of the variant *-ragiño*.

6.5.2. *Cases built upon the genitive*

This is the case upon which the greatest number of postpositions have grammaticalized. However, we have to divide these subcases into those built upon the old genitive *-(ar)e* (cf. §6.4.1.2), and those built upon the more recent genitive *-(ar)en*, with nasal. The former ones should logically be older, although it is equally possible that those built upon the old genitive have replaced this form with the modern genitive.

6.5.2.1. The sociative

In the first group, built upon the old genitive, we find the sociative *-(ar)e-kin*, as in *gizon-are-kin* 'with the man'. Two variants of the same suffix are *-(ar)e-ki* and *-(ar)e-ki-la(n)*.[79] Schuchardt (1947 [1923]: 61) reconstructed this case suffix as **-(ar)e-kide-n*, which would be the grammaticalization of the postposition *kide* 'companion' with the inessive *-n*, upon the genitive (lit. 'in the company

---

78   Schuchardt (1947 [1923]: 83) believed that what underlied *oraidano* was: 1) *orai* 'now'; plus 2) the relative form of the verb 'to be' in 3rd p. sg. (*da* 'is' + rlv. *-n*); plus 3) a morpheme meaning 'until, up to', which he reconstructed as **-no* (which in turn could also develop into *-o* if the nasal was intervocalic, cf. *FHV*: 299). Hence, we would have **orai-da-n-no*.

79   The genitive is optional on common nouns, but obligatory on pronouns.

of').[80] This etymology was later accepted by all Bascologists (Lafon 1999 [1965]: 196, Michelena 1985 [1972b]: 297, Trask *HB*: 201).[81] The development would be *-(ar)e-kide-n > -(ar)e-kie-n > -(ar)e-ki-n*. The intermediate step does not require an asterisk, as it is attested e.g. in E.Cib. *lagun-are-kien* 'with the companion'.

In Trask's opinion (*HB*: 202), the form without the inessive *-(ar)e-ki* would be an innovation by analogy with forms with the local genitive, such as *lagun-eki-ko* (as in *lagun-eki-ko festa* 'the party with the friends'), which have no nasal either. However, I believe, as do Gómez (2005: 275) and Michelena (apud ibid.), that the forms without the inessive —including *lagun-e-ki-ko*— are older. The variant *-(ar)e-ki-la(n)* seems to be an innovation, which added the morpheme *-la* (*HB*: 201), and, in some cases, also the ines. *-n*. The morpheme *-la* could be the same as the one found on pronouns with adverbial meaning like *nola* 'how', *horrela* 'that way', etc.

In western varieties, there is another way to form the sociative. For this, the postposition is not *-ki(n)*, but *-gaz*, and is attached not to the genitive, but to the root in the abs. case. Hence, we have e.g. *gizon-a-gaz* 'with the man' in the singular, but *gizon-akaz* 'with the men' in the plural. The pl. form must be more recent according to Trask (*HB*: 202), and possibly formed by analogy with that of the sg., since in some areas in which the sg. sociative is *-gaz*, the plural is still *-kin*. Morphologically, this pl. sociative is formed by attaching the postposition *-gaz* to the root in abs. pl.: **lagun-ak-gaz > lagun-akaz*. As a matter of fact, in Guipuscoa and adjacent regions, where the other sociative —the one with *-kin*— is used, the postposition has also been attached to the abs. pl. of the root: *lagun-a-kin* 'with the friends' (vs. *lagun-e-kin* in eastern varieties). We find the etymological spelling in Old G. authors such as Ub. *Apostolu-ak-kin* 'with the Apostles'.

The western *-gaz* may have been an independent postposition, later grammaticalized. In this case form, Trask (*HB*: 203) identifies, on one hand, the already known morpheme *-gan*, which in Biscayan would have entered the realm of the grammatical cases, and, on the other hand, the instrumental *-z*. Hence, he proposes the following grammaticalization: **-ganez > *-gahez > *-gaez > *-gaiz > -gaz*. As we have seen above, Trask believes that *-gan* comes from *gain*. Consequently, he has to resort to this evolution to get rid of the nasal. I agree, however, with de Rijk (cf. §6.4.2.1), that if we assume that *-gan* is nothing but the

---

80 In Schuchardt's formulation, the genitive *-(ar)e* had lost the *-n*.

81 In fact, Gómez (2005: 275) recalls that this etymology had already been suggested by van Eys (1879: 75-6).

suffix -*ga* + the inessive -*n*, then the reconstruction is as simple as -*ga* plus the instrumental -*z*, as Michelena (1987 [1981]: 44) had already anticipated.

Following Lakarra (cf. §6.4.1.1), I believe that, in origin, the **-ga* of the postposition -*gaz* is the morpheme which ends up becoming ergative, and also the same morpheme which allows animates to take local cases. Morphologically, the instrumental is not a local case (it has -*a*- in the sg., but no -*eta*- in the pl.), but syntactically it is close to them. Bearing this in mind, it appears that the western dialect has formed a sociative suffix, different from the periphrastic one in the rest of the dialects (-*(r)e-kin*), by activating the common instr. -*z* for the local inflection of animates by the same procedure used in the local cases: the insertion of -*ga*. Hence, just as we have ines. *etxe-a-n* 'in the house' but *gizon-a-ga-n* 'in the man', we would have instr. *etxe-a-z* 'with/of the house' but (west.) soc. *gizon-a-ga-z* 'with the man'.

In old western texts and in more recent oral sources, the sociative, both the one in -*gaz* and the one in -*(r)ekin*, sometimes appears with a function which is not sociative, but coordinative, in structures of the type X-*gaz/-(r)ekin* Y 'X and Y'. Here are two examples, one corresponding to Old Biscayan, and the other to Old Guipuscoan: RS. (246) *Ogi-a-gaz hur-a* 'Bread and water (lit. with bread water)', (G, 1705) *asto-arekin idi-a* 'the donkey and the ox (lit. with the donkey the ox)' (in a Christmas song; apud Lakarra 1996a: 193). In addition to Old B and Old G, this coordinative structure also appears in Laz., and there are also traces of it further to the east, as in a Navarrese song gathered by Azkue in his *Cancionero Popular Vasco* (apud Lakarra 2008b: 91), in which we can read *ze-ru-arekin lurr-a / zaure-ak dituzu* 'heaven and earth (lit. with heaven earth) are yours'. Even in the continent, we have found some traces of it, as in the recently published text *K&B* (~ West.-LN, +/-1750; cf. Padilla-Moyano 2011), and Lakarra (2008b) suggests that it may once have been common to the whole territory.

Similar sociative constructions with coordinative function exist in many languages. Stassen (2000) even makes a binary typological (and geographical) classification into WITH-languages and AND-languages. Lakarra (2008b: 92 ff.) applies Stassen's generalizations to Basque and recalls that languages with coordination of the WITH type usually employ serial verbs, which supports the existence of these in Proto-Basque, and probably also the origin of some case markers from the grammaticalization of such serial verbs. Likewise, the fact that WITH-languages tend to lack case systems and *vice versa* (Stassen 2000: 44) argues in favor of an evolution of Basque in which the coordination of the WITH type is increasingly reduced, and the case system —actually not very old— increasingly larger.

6.5.2.2. The motivative

Among the cases built upon the modern form of the genitive, we have the motivative case. It is commonly represented as *-(aren)gati(k)*, and translated as 'because of', as in *su-a-ren-gatik* 'because of the fire'. Since the postposition has the variant *-gaiti(k)* in western dialects, it was proposed by Michelena (*FHV*: 92), and has been believed since then, that this postposition is a loan translation of Latin expressions such as *hominis causā* 'on behalf of (lit. through the cause of) the man', the second element being in the ablative case. Since *gai* is 'element, thing' in Basque, *-gai-ti(k)* would be the already known ablative-prosecutive morpheme suffixed to *gai*, and the whole postposition added to the lexical item.

In the oldest texts, the postposition is added to the stem plus article in the singular, and directly to the stem on proper nouns. The possibility of adding it to the genitive form, which is optional in the modern standard language, is a later innovation. In the plural, in 16th century texts two main patterns are found: 1) the postposition added to the old form of the genitive (e.g. Dch. *gure gaizki-e-gatik* 'due to our evil deeds', Laz. *an eozen arzai-ae-gati* 'due to the shepherds that were there'), and 2) the postposition added to the abs. pl. form (e.g. RS. *Butroeko meni-ak-gaiti* 'due to the truces of Butroe', Lç. *gure bekatu-ak-gatik* 'on behalf of our sins'). However, the possibility of adding *-gatik* to the modern form of the gen. pl. is already present on some nouns in Lç. (*gizon-en-gatik* 'due to the men'), and in the following century, in Ber. (HN) it is the only option. While the details deserve more research, it seems sound to suggest that the possibility of forming the motivative upon the genitive has entered the noun declension through the plural, and geographically it has spread from the central dialects.

6.5.2.3. The destinative

Another case built upon the modern form of the genitive is the destinative, formally *-(r)en-tzat*, as in *ama-ren-tzat* 'for the mother', and also *-(r)en-dako*, as in *ama-ren-dako* 'idem'. Morphologically, *-(r)en-tzat* is formed by the postposition *-tzat* attached to the modern form of the genitive, and *-(r)en-dako* by the indef. form of the local genitive agglutinated to the modern form of the genitive (the deep form would be *-tako*, which surfaces as *-dako* after a nasal). When *-tzat* or *-tako* are attached directly to the stem, then what we get is the essive case (cf. §6.6.2).

Both forms of this case appear in all dialects from the earliest texts (RS. *idi-en-tzat* 'for the ox', Lç. *bere senharr-aren-tzat* 'for her husband'; Cap. *lagun-en-*

dako 'for the fellow man', Lç. *Iainko-aren-dako* 'for God', etc.), although many varieties have generalized one of them. They are commonly employed with animates, since for inanimates another destinative suffix is used, formed by the juncture of the allative plus the local genitive: *-rako* (cf. §6.6.1), as in RS. *gudura-ko* 'for the war', Lç. *su-ra-ko* 'for the fire'.[82]

Especially in old texts, the destinative function is also sometimes expressed by the simple (either possessive or local) genitive: Tt. *gu-ré oro-ren* 'for all of us', RS. *gibel(-e)-ko* 'for the liver', Oih. *hortz(-e)-ko* 'for the tooth'.

In addition to all these grammaticalized cases or subcases built upon the genitive, there exist some other postpositions which also require the genitive but which have not reached a grammaticalized stage and are consequently written separately, such as *-(ar)en inguruan* 'around', *-(ar)en esperoan* 'waiting for', etc. Some other postpositions require the dative or other cases. Since these have remained at their periphrastic stage, their study would go beyond the scope of noun morphology proper, and consequently they will be omitted here.

6.6. The pseudo-flectional morphemes (or non-cases)

There are a number of morphemes which, although usually considered as part of the declensional paradigm, show special behavior in several respects. In fact, I agree with Trask (*HB*) that in a linguistic analysis they cannot be treated as if they were cases in the normal sense of the word. Consequently, I will analyze them in this separate section.

6.6.1. The local genitive

One of these morphemes is the local genitive *-ko*. The reason for not considering it to be a proper case is that it can be attached to nearly everything, turning the element to which it is attached into an adnominal modifier of the NP head. That element can be, among other things, a case-inflected NP, as in (10a, b). In fact, any case-inflected NP can be turned into an adnominal modifier by *-ko*, unless that case is the ergative or the dative.[83] It can also be any other adverbial, as in

---

82 In fact, the animate benefactive sometimes appears on inanimates as well, as in Mic. *osasun-enzako* 'for the health'.

83 NPs in the inessive case made dependent by *-ko* are not common in modern dialects either. But they do exist in old texts, as in the famous Chant of *The Burning of Arrasate-Mondragón* collected by Garibay in the 16th century (cf. *TAV*: 81 ff.), where we find *oñeztar barru-an-go-ak* 'the Oñacins of inside', in which the inessive adverbial phrase

(10c, d). Moreover, attached to finite verbs plus *-la* they can form dependent clauses, i.e. clauses dependent on an NP head, as in (10e), or adverbial clauses meaning 'because', as in (10f):

(10) a. *lagun-ekin* → *lagun-eki-ko* *festa*

friend-SOC.PL friend-SOC.PL-L.GEN party

'With the friends' 'The party with the friends'

b. *txokolate-z* → *txokolate-z-ko* *tarta*

chocolate-INSTR.INDEF chocolate-INSTR.INDEF-L.GEN cake

'With chocolate' 'The chocolate cake'

c. *gaur* → *gaur-ko* *festa*

today today-L.GEN party

'Today' 'Today's party'

d. *horre-la* → *horre-la-ko* *auto* *bat*

that-ADV.SUFF that-ADV.SUFF-L.GEN car a

'That way' 'A car like that'

e. *Etorri-ko* *dire-la-ko* *sinismen-a*

come-PROSP AUX-CPL-L.GEN belief-ART

'The belief that they will come'

f. *Pozik* *nago,* *etorri-ko* *dire-lako*

happy am come-PROSP AUX-CS.SUFF

'I am happy, because they will come'

In addition to these and some other functions (*HB*: 101-3), the prototypical function of the local genitive is to form adnominal genitive modifiers of NP heads, which —unlike the genitive *-en*— are associated to local notions and typically express origin. In the examples in (11), the *-ko* phrases are modifiers of *jauna* and *Milia*. (11b) is a proper noun attested in a dirge collected by Garibay in the 16th century (cf. *TAV*: 75 ff.). It is most probably archaic, and shows that,

---

barru-an 'inside' is turned into a dependent of the head *oñeztar* by *-ko*, which becomes *-ko* after *n*. Inessive *-n* + *-ko* structures also survive residually in archaic forms such as S. *etxe-n-ko* 'of home'.

in the past, the modifier in the local genitive case, as in relative clauses (§6.4.1.2), followed the head modified by it.

(11) a. *etxe-ko*   *jaun-a*

    house-L.GEN lord-ART

    'The lord of the house'

  b. *Milia*   *Lastur-ko*

    p.n. (of a woman) p.n. (of a place)-L.GEN

    'Milia from Lastur'

Lafon suggested that the suffix *-ko* was derivational, even adjectival, in origin, and only later entered the declension (1999 [1948]: 207, [1965]: 182). In the declensional paradigm, it came to function as a kind of counterpart of the possessive genitive *-en*, with the distribution of semantic scopes already described. He also noted the interesting fact that both suffixes have served, when attached to non-finite verbs, to express future aspect. Broadly speaking, western dialects chose *-ko* (*etorri-ko naiz* 'I will come'), whereas eastern dialects opted for *-en* (*jin-en niz* 'I will come') to express the same function (but see §8.3.3 for the precise distribution of both morphemes).

By contrast, Michelena (1985 [1972b]: 301) noted that, unlike *-en* —which can appear on both animates and inanimates—, *-ko* cannot appear on animate nouns (*seme-aren* vs. **seme-ko 'of the son'). As a matter of fact, there is another *-ko* which appears quite abundantly in the scant corpus of the Aquitanian inscriptions (as in Aquit. *SENI-CCO* 'little child', (Soria) *SESEN-CO* 'little bull'), and survives in medieval proper nouns (such as *Enne-co*, perhaps attached to *ene-* 'mine', > Rom. *Íñigo*), and in historical Basque in a few nouns (as in *muti-ko* 'little boy', *zati-ko* 'little piece'). This is clearly a diminutive suffix, hence derivational, and can appear on both animates and inanimates. Taking this into account, as well as the fact that the local genitive does not appear on animates in historical Basque, Michelena seems to suggest that the reason for this gap may be that *-ko* soon lost its function in the declensional paradigm, and was specialized in its derivational nature, diminutive in meaning. Although both elements could be the same, the path of development proposed by Michelena to link them (from declension to derivation) is unexpected according to the principles of typology. Moreover, the derivational *-ko* appears as a productive element in Aquitanian, but is marginal in historical Basque. Thus, it seems to have been on the wane in the prehistoric centuries.

As for its morphological formation, the local genitive functions like the other local cases. It has no article in the sg. (*mendi-ko* 'of the mountain'), but it pre-

sents *-eta* in the pl. (*mendi-eta-ko* 'of the mountains'), and *-ta* in the indefinite of vowel-ending roots (*mendi-ta-ko* 'of certain mountains'). As I have said, *-ko* can be attached to an NP in any case except in the erg., in the dat., and —old attestations and fossilized expressions left aside— in the inessive. Since the inessive is, then, the only local case which presents this restriction, Michelena (1985 [1972b]: 302) proposed that the local genitive was built precisely upon this case. Thus, the underlying form of, say, *mendi-ko*, would be **mendi-an-ko*.[84]

When *-ko* is attached to the allative to form the destinative (cf. §6.3.2), or when it is attached to temporal adverbs (as in (10c)), or to a finite verb plus *-la* to form an adverbial subordinate clause of cause, in some eastern varieties *-ko* appears as *-ko(t)z*. I believe that this added morpheme is the instrumental, which sometimes surfaces as an affricate in word-final position. Not all uses of *-ko*, however, present the allomorph *-ko(t)z*: cf. *etxe-ko* 'of home' vs. ***etxe-ko(t)z*, but *daude-la-ko(tz)* 'because they are', *-lako(tz)* being a complementizer with *-ko(tz)* as a second element.

6.6.2. The essive

Another non-case is the essive *-tzat*, with the possible variant *-tako* in the eastern half of the country. These are used to form object complements, as in the sentence *tontu-tzat hartu nau* '(s)he has taken me for a fool'. It is not considered a case because it can only appear attached directly to the root. Hence, it has neither sg. nor pl.

The etymology of these elements is unclear. If we apply the phonotactic rules that we can assume for Old Basque, the final stop of *-tzat* must have been **-da*. The resulting **-tza-da* could then be split into two morphemes. The first one is formally similar to the instrumental suffix *-za*, perhaps also related to the derivational one of place names, or even to the verbal nominalizer (Lakarra 2008a: 484, and §6.4.1.4), and the second one to the **-da* of the variant *-rat* of the allative, perhaps derived from **daR* 'to sit' (§6.4.2.2). In both morphemes, however, it is very difficult to find a semantic connection with the essive idea.

The variant *-tako* is clearly the local genitive suffix, in its indefinite form. Thus, it appears to have come about as a specialization of the local genitive. 16th century classical authors only use *-tzat*, although the essive *-tako* is attested from at least 1626 in the wedding formula *Nai nauzu zure senarr-etako?* 'Dost thou want me as thy husband?', to be located in the surroundings of Orreaga-Roncesvalles (northern Navarre; cf. *ContTAV*: 105).

---

84   In fact, the need for an asterisk would be doubtful in view of old attestations such as *bar-ruan-go-ak* 'the ones of inside', mentioned above.

6.6.3. The partitive

Let us conclude this subsection with the partitive *-(r)ik*. As I have already said (§6.4.2.3), the form of the old ablative is the same as that of the partitive. Lafon (1999 [1948]: 206) believed that both the ablative and the partitive came from the same morpheme abl.-ptv. *-(r)ik*, and that the former has become independent as the derivational *-ti* of adjectives incorporated into it. De Rijk (1998 [1996]: 436) and Trask (*HB*: 205) also state that the partitive derives from the old ablative. In fact, marking ablative and partitive notions by means of the same morpheme is a common phenomenon, as can be observed in Romance (Sp. *dame algo de* ₐbl. / ₚtv. *esto* 'give me some of this'), Finnish, Kerongo, etc. (Heine & Kuteva 2002: 32-3).

The etymology of this *-(r)ik* is not clear. At first sight, the pure marker appears to be *-ik*, *-r-* being an epenthesis to separate vowel-ending roots from the case marker, just as in the indef. of grammatical cases (with which *-(r)ik* shares precisely the indef. semantics; cf. §6.4.1.2). In turn, *-k* may be compared to the *-k* or *-ka* of abl. *-ti(k)(a)*, perhaps deriving from the iterative suffix *-ka* (§6.4.2.3). *-(r)ika* is actually attested, although perhaps favored *metri causa*, in Port.'s poem of 1610 (*pobreza-rika irten-ika* 'once come out of poverty';[85] cf. *TAV*: 118-122). If *-ka* were the original morpheme, then, the succeeding *-k* would be an apocopated variant. What is left to us is a simple *-i-*. This has been analyzed by Lakarra (2008a: 482) as derivable from **nin* 'to give'. This would imply a grammaticalization path 'to give' → ABL-PTV, which does not seem very compelling, although it has some parallels mentioned by Lakarra himself. Ariztimuño (2012b) prefers to think of **din* 'to come' as the source of the abl.-ptv. *-(r)ik*. The path 'to come' → ABL-PTV seems to be more semantically well-founded, although, as Ariztimuño himself admits, this development has the problem that it would imply that *-r-* is not an epenthesis but a part of the suffix (*-r-* coming from *-l-*, which would in turn come from a **d-* in initial position). If it is a part of the suffix, then we would expect to find it on consonant-ending roots as well (cf. *etxe-ra* 'to the house', *lurr-e-ra* 'to the earth', with the allative), but ***lurr-e-rik* does not exist. We find *lurr-ik* instead.

The difference between the ablative and the partitive, in their historically attested forms, is that the latter does not inflect for number, and thus some authors do not consider it to be a case. Moreover, it fits into the same contexts as the article when the NP is indefinite in both negative (as in (12b)) and interrogative (as in (12c)) clauses, and dialectally in some affirmative contexts (as in (12d);

---

85 There are actually two *-(r)ika* there. The abl. one is *pobreza-rika*, whereas *irten-ika* is an adverbial built upon the participle *irten* 'to come out' (see de Rijk's analysis below).

cf. de Rijk 1998 [1972b]). Thus, the partitive has been considered an indefinite determiner rather than a case marker by some analysts.[86]

(12) a. *Liburu-a*   *du-t*

        book-ART   have-1SG

        'I have the book'

    b. *Ez*   *du-t*   *liburu-rik*

        not   have-1SG   book-PTV

        'I have no book'

    c. *Liburu-rik*   *ba-du-zu?*

        book-PTV   FOCUS.PTC-have-2SG

        'Do you have any book?'

    d. *Ba-du-t*   *liburu-rik*   *(asko / ezer)*

        DUMMY.PTC-have-1SG   book-PTV   (much / something)

        'I have books (many books / some books)'

De Rijk (1998 [1996]: 435) identifies two homophonous morphemes of the form *-(r)ik*. One *-(r)ik* is found in predicative complements such as *gizon-a bizi-rik dago* 'the man is alive', or absolute constructions such as *hori esan-ik, ...* 'that said, ...'. The other *-(r)ik* is the partitive morpheme. According to de Rijk, this would be the old ablative, which evolved from being a case marker to its function as an indefinite determiner. Its case marking function is still preserved in certain varieties and in certain contexts. Apart from the cases given in §6.4.2.3, *-(r)ik* with ablative function still survives in fossilized expressions such as *mendi-rik mendi* 'from mountain to mountain' or *urte-rik urte* 'from year to year' (ibid.: 436),[87] and also in superlative NPs such as *emakume-rik ederr-en-a* 'the most beautiful woman' (ibid.: 437).[88] But the most interesting examples are

---

86 By contrast, an aspect pointed out by Ariztimuño (2012b) which brings the ptv. closer to the case system and away from determiners is that, while determiners can acquire case morphemes, the partitive cannot. On these issues, cf. Etxeberria (2010).

87 This construction is typical in peripheral areas. In more central areas, it is more typical to find the instr., as in *mendi-z mendi* 'from mountain to mountain', the abl. + all. construction, as in *urte-tik urte-ra* 'from year to year', or even the ines. + ines. one, as in Añ. *urte-an urte-an* 'idem'.

88 They can also appear in the bare root form, *emakume ederr-en-a*, or in pl. inessive, always inanimate, *emakume-etan ederr-en-a*.

the following ones, given by de Rijk, which today have virtually disappeared from the language:

(13) a. *Amets-ik*   *(asko)*   *sortu*   *zai-zki-gu*

          dream-PTV   (many)   create   AUX-3PL-1PL

          '(Plenty of) dreams have arisen to us'

      b. *(Ezer)*   *gauza*   *on-ik*   *ni-gan*   *ba-da*

          (something)   thing   good-PTV   I-INES   COND-there.is

          'If there is something good in me'

In all of these examples, the partitive always appears with an NP of unspecified value. In (13a), there is an indefinite quantifier, *asko*, which licenses the presence of the partitive on the quantified element, provided that this is of a generic, non-specific nature. In (13b), we find an indefinite quantifier, *ezer*, which licenses the presence of the partitive. In these examples, the partitive fulfills its usual function of linking a non-specific NP to its quantifier (ibid.: 439).[89] The interesting point in these examples is that the quantifier can —and for many current speakers, must— be omitted. We could consider it to have dropped out. After this loss, the functions of the quantifiers were taken over by the partitive, which thus moved from being a link between the quantifier and the quantified, to acquiring a function of indefinite determiner.

From a diachronic point of view, I agree with Ariztimuño that it is misleading to think of one casual partitive and another which is not, as different grammaticalizations of the same original morpheme. I believe that Ariztimuño is also right when he dates the split-off of the ptv. from the abl. —or rather, following Lafon (1999 [1948]), from a ptv.-abl.— to a period prior to the grammaticalization of number and definiteness into the local paradigm, as this would justify the absence of such parameters in the partitive. The ablative would have acquired them at a later stage, since it belonged to the local paradigm.

In Figure 6.7, I reproduce Ariztimuño's proposal. The whole process illustrated there must have been a common phenomenon. The innovation of the abl.-pros. by introducing *-ti* may have been motivated, or at least favored, by the need to distinguish the ptv. from the abl. sg. This abl.-pros. sg. *-ti* arose most

---

89 The relationship between ablative and partitive is still clear in examples like the following, where the ablative fulfills a partitive function:

    Liburu-etatik   *asko*   *txarr-ak*   *dira*

    book-ABL.PL   many   bad-ART.PL   are

    'Of the books, many are bad'

probably as a reinterpretation of the derivational *-ti* (§6.4.2.3), and came to replace abl. sg. *-(r)ik*, pushing *t*-less ablatives into archaizing contexts (the plural, place names, etc.).
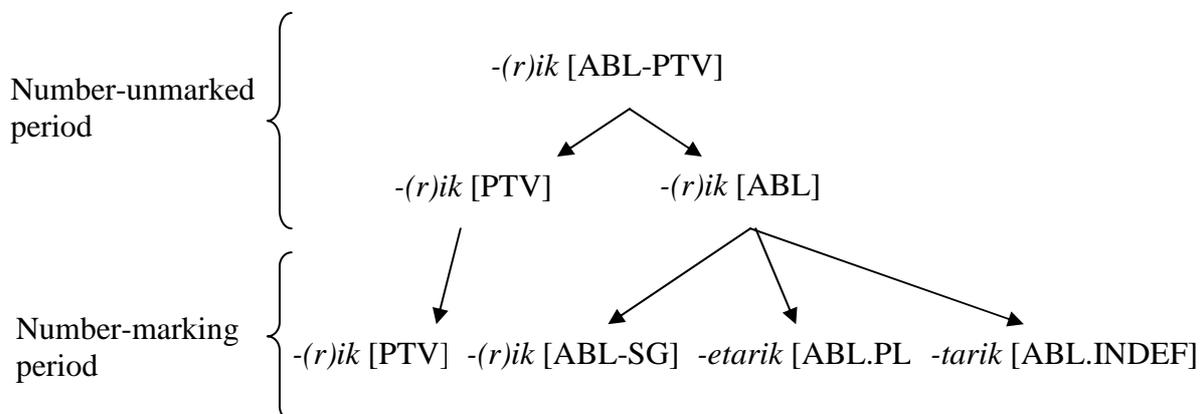
Number-unmarked
period
$\left\{\begin{array}{c} \end{array}\right.$

-(r)ik [ABL-PTV]

-(r)ik [PTV]   *-(r)ik* [ABL]

Number-marking
period
$\left\{\begin{array}{c} \end{array}\right.$

-(r)ik [PTV] *-(r)ik* [ABL-SG] *-etarik* [ABL.PL *-tarik* [ABL.INDEF]

Figure 6.7.   Grammaticalization of the ablative / partitive according to Ariztimuño (2012b)

6.7. Conclusion

In this chapter, the rise of definiteness and number as grammaticalized categories has been described, as well as that of the cases within the case paradigm, which, as we have seen, have been discussed for many years. The idea that emerges is that these grammaticalization processes correspond to a transition from a more isolating stage to a more agglutinating one. Even so, the grammaticalization of the declension system as a whole is not homogenous, since morphemes have arisen in different periods, through diverse and unsystematic paths.

Perhaps it is time to reconsider, from the current state of research, the grammaticalization of each morpheme, and jump from a case-by-case study of the grammaticalization of each morpheme to a more paradigmatic vision. After searching for explanations for individual morphemes, and discussing the ones proposed so far, we should further pursue a global interpretation of the rise of the paradigm. I believe that it is viable to establish a relative chronology of the Basque case paradigm as a whole. In this chronology, several phases along the path isolation → agglutination → fusion could be sorted out, once we have synchronically identified the different groups, such as the local cases, the plural of which, formed by means of the morpheme *-eta-*, is neatly agglutinating, unlike the plural of grammatical cases. An element which could help is the study of the chronology of intervocalic epenthetic consonants, which appear only in some cases (cf. gen. indef. *etxe-r-en* vs. pl. *etxe-en*), and hence, in certain periods.

A promising factor in this enterprise is that we own progressively more and more typological information as material for comparison, which allows us to prioritize the hypotheses which best fit into the most frequently attested grammaticalization paths. In spite of recent contributions such as Lakarra (2005, 2006a, 2006b, etc.), Manterola (2009b) and Martínez-Areta (2009a, b), which aim at reconstructions of the whole paradigm rather than of concrete cases, and employ typological comparison in their research, this aspect has not been sufficiently exploited. Comparing the Basque noun declension to that of other languages, and looking for similar patterns and paths of development is a method which must be taken advantage of, as far as linguistic typology will allow us.