

## VOWEL INTERACTION IN BIZCAYAN BASQUE\*<sup>1</sup>

Many varieties of Basque have rules affecting vowel sequences. The central importance of these rules for the functioning of the language is clear from the fact that the definite article is *a*, used as a postclitic to nominal stems, many of which end in a vowel. These rules, therefore, determine to a large extent the physiognomy of the different varieties of Basque. A thorough study of them will have to await the appearance of a linguistic atlas, so urgently needed for practically all aspects of Basque studies. Sad to say, the political conditions are, at present, hardly favorable to its preparation. Yet, thanks to the pioneering work of Prince Louis Lucien Bonaparte more than a century ago and taking into account the data provided by the efforts of more recent scholars,<sup>2</sup> an outline of the main facts can be given. Such an outline will be presented below with the aim of encouraging further studies in the field, and showing at the same time the importance of Generative Phonology with its concept of ordered rules (developed by Noam Chomsky and Morris Halle) for Basque dialectology. The Bizcayan dialect area, the area west of the Deva river, is particularly interesting for our purposes, and will thus claim most of our attention.

Basque has a five vowel system:<sup>3</sup> *i, e, a, o, u*, with, apparently, no distinctions of tenseness or length. The Suletin dialect and some neighbouring varieties of Low-Navarrese have acquired a sixth vowel *ü*, but this has happened nowhere near the area we are considering here.<sup>4</sup>

---

\* *FLV* II-5 (1970), 149-167.

<sup>1</sup> As on so many occasions, I have greatly benefitted from the helpful comments of Prof. Dr. Luis Michelena, who kindly read a first version of this article and provided a great deal of pertinent information and expert advice. I am very grateful to Dr. Ambrosio Zatarain for his most generous assistance in contacting reliable informants. I want to thank also Prof. E. Wayles Browne, whose insightful suggestions lead to considerable improvements. In a more general and pervasive way, I am thoroughly indebted to my teachers of Generative Phonology, Professors Dr. Noam Chomsky, Dr. Morris Halle and Dr. Paul Kiparsky (Massachusetts Institute of Technology).

<sup>2</sup> In particular, Azkue (1926), Ormaetxea (1925) and Michelena: *Fonética histórica vasca* (to be abbreviated henceforth as *FHV*). After a first draft of this article had been written, I also read the valuable contribution *Apuntes vizcaínos* by N. M. Holmer and V. A. de Holmer, where I found several data of great interest and relevance.

<sup>3</sup> See: *FHV*, Chapter I.

<sup>4</sup> See: René Lafon: "Sur la voyelle *ü* en basque". *BSLP* 57 (1962), 83-102. And also: *FHV*, § 1.5, page 51 ff.

We will use a system of binary features as designed by Roman Jakobson, Morris Halle and Noam Chomsky to characterize classes of vowels. High vowels are said to be *+hi*, low vowels are said to be *-hi*. Back vowels are called *+back*, front vowels are called *-back*. Rounded vowels are *+round*, unrounded vowels are *-round*.

We thus have for Basque the following table:

	<i>a</i>	<i>o</i>	<i>e</i>	<i>i</i>	<i>u</i>	( <i>ü</i> )
<i>hi</i>	-	-	-	+	+	+
<i>back</i>	+	+	-	-	+	-
<i>round</i>	-	+	-	-	+	+

$V$  will be used as an informal designation of the class of vowels, and may be substituted for by the conjunction of the features *+syllabic* and *-consonantal*.

We will now list the rules first, and then discuss their interrelations. Only the first rule, which we will call *Raa*, is common to the whole area; local varieties differ as to which of the other rules operate. Using the notational conventions of Generative Phonology, the first rule can be written as follows:

$$Raa: \left[ \begin{array}{c} V \\ -hi \\ -round \end{array} \right] \rightarrow \left| \begin{array}{c} -back \end{array} \right| \left| \begin{array}{c} / \\ - \end{array} \right| + \left[ \begin{array}{c} V \\ -hi \\ -round \\ +back \end{array} \right]$$

In words: low unrounded vowels (i.e.: *a, e*) are fronted before a low unrounded back vowel (i.e.: *a*), that is:  $a + a \rightarrow ea$ .

The rule accounts for alternations such as:

<i>alaba bat:</i>	"one daughter"	<i>alabea:</i>	"the daughter"
<i>neska bat:</i>	"one girl"	<i>neskea:</i>	"the girl"
<i>eleixa bat:</i>	"one church"	<i>eleixea:</i>	"the church" <sup>5</sup>
<i>erreka bat:</i>	"one brook"	<i>errekea:</i>	"the brook"
<i>isera bat:</i>	"one sheet"	<i>iserea:</i>	"the sheet"

In many localities of the Bizcayan dialect area the rule has exactly the form given above. In other localities of the same area, however, the rule has been simplified in that the feature *-round* no longer appears in the environment of the rule. The result is a more general rule, which we will call *Rao*.

$$Rao: \left[ \begin{array}{c} V \\ -hi \\ -round \end{array} \right] \rightarrow \left| \begin{array}{c} -back \end{array} \right| \left| \begin{array}{c} / \\ - \end{array} \right| + \left[ \begin{array}{c} V \\ -hi \\ +back \end{array} \right]$$

<sup>5</sup> *x* denotes the same sound as the digraph *sh* in English, and *tx* is the corresponding affricate, spelled *ch* in English.

In words: low unrounded vowels (i.e.: *a*, *e*) are fronted before a low back vowel (i.e.: *a*, *o*), that is,  $a + a \rightarrow e a$  and  $a + o \rightarrow e o$ . This gives rise to alternations of the following type:

<i>alaba bat</i> :	“one daughter”	<i>alabe ori</i> :	“that daughter”
<i>neska bat</i> :	“one girl”	<i>neske orrek</i> :	“that girl” (erg case)

If we try to delimit geographically the *Rao* area against the *Raa* area, no clear picture seems to arise. Oñate has *Raa*, but Placencia has *Rao*. Aulestia has *Rao*, but Elanchove and Marquina have *Raa*. For the latter town, we can cite Moguel, *Peru Abarca* (in the second dialogue: *ule trenza ori* “that braid”) and Rollo, *The Basque Dialect of Marquina*, where we read: *elixa orren barruan* “in the interior of that church” (page 8T).

Múgica has *Rao*, for in two local folktales we notice the form *elíxi ori*, derived from \**elixa ori* by means of rule *Rao*, followed by a rule *Rea* to be discussed later on. These folktales can be found in Barandiaran II, page 11 and 26.

For the valley of Arratia, the forms *alabe onek* and *alaba onek* “this daughter” (erg. case sing.) are equally acceptable according to Father Lino Aquesolo (oral communication). This means that whereas the original rule *Raa* is always obligatory, the simplified rule *Rao* can be optional.

In *Barandiaran III*, we find some folktales taken down in Dima, a town belonging to Arratia. The same informant, Claudio Pujana, uses forms where *Rao* has applied and forms where it has not. On page 116, *alabe ori* occurs, but on page 138 *alaba orrek*. On the same page, we find both *neske orrek* and *neska orrek*. On page 139 we find again *alaba orrek*.

For Ceánuri, which is also in the valley of Arratia, we find only forms where *Rao* has not applied: *elixa ori* and *isera ori* (*Barandiaran II*, page 22 and 57).

In some localities the combination  $a + o$  never occurs, so that *Raa* is indistinguishable from *Rao*. This happens, for instance, in Baquio, a town about 10 kms. west of Bermeo, in which the demonstrative pronouns always precede the noun instead of following it: *ori alabie*, “that daughter”.

Of course, this small amount of data does not permit to draw any final conclusions. But we do get the impression that rule *Rao* shows a scattered distribution, that is, it does not form a connected sub-area inside the *Raa* area. If this impression is borne out by further observations, it confirms Kiparsky's view to the effect that rule simplification is a spontaneous process that will arise independently in different points of the same dialect area. (See Kiparsky 1968).

I have put an obligatory morpheme boundary (+) in the environment of rule *Raa*. This is to prevent it from applying to stems containing an  $a + a$  sequence. Such stems are extremely rare in Bizcayan. An example is *z̄aar* “old”, which may simplify to *zar*, but never turns into *zear*.

The correctness of this procedure, however, is open to question. In nearly all of Bizcaya,<sup>6</sup> we find the alternation: *z̄ara*, “you are”, *z̄arean*, “that you are” (relative

<sup>6</sup> As to the exceptions, Azkue mentions them in the following terms: «Al ir a consultar de nuevo al pueblo para redactar el Flexionario del dialecto vizcaíno, vi que en varios pueblos territorialmente gipuzkoanos y algún alabés, como también hacia Plencia, dice el pueblo *garala*, *z̄arala*, *z̄aralaw*». (Azkue 1926: 35).

clause) and *zareala*, "that you are", (verbal complement). Likewise: *gara*, "we are", *garean*, *gareala*; *dira*, "they are", *direan*, *direala*. The Guipuzcoan dialect shows no alternation: *zera*, *zeran*, *zerala*; *gera*, *geran*, *gerala*; *dira*, *diran*, *dirala*.

As I have established elsewhere (De Rijk 1969), there is in Basque a rule of Vowel Truncation, and the underlying form of the relativizer and complementizer suffix are *-n* and *-la*, respectively. The alternations above are then explained as follows: *zara* from underlying *zaraa* by Vowel Truncation; *zarean*, *zareala*, from underlying *zaraa* + *n*, *zaraa* + *la* by rule *Raa*. Guipuzcoan does not have rule *Raa*, and so we get *zeran*, *zerala* from *zeraa* + *n*, *zeraa* + *la* by Vowel simplification.

Yet, in the form *zaraa* + *la*, where we need rule *Raa* to apply, the two *a*'s are not separated by any morpheme boundary.

Several tentative solutions come to mind. We may try one of the two proposals discussed by Nancy Woo in a slightly different context: retention of an intervocalic *b* in the synchronic phonology, or, introduction of a distinction between *aa* and *aa* (with the second *a* non-syllabic), rule *Raa* applying only to the latter combination (cf. Woo 1968). We might also conceive of a proposal to the effect that rules like *Raa* never apply unless they create alternations. Since, at present, I have no criteria for preferring one possibility above the others, I must leave this question open.

Rule *Raa* is subject to various other restrictions; e. g. it does not apply to compounds: *neba* "brother" and *arriba* "sister" combine into *nebarrebak* "siblings", but not into *\*nebearrebak*.

Details may vary from locality to locality: with the suffix *-ago* "more", an informant from Placencia accepts both *ozperago* and *ozpereago*, derived from the adjective *ozpera*, "sensitive to cold", an informant from Vergara, 6 kms. south of Placencia, only has a form of the latter type: *ozperiau*, whereas in most regions (in Marquina for instance) rule *Raa* does not apply to the suffix *-ago*.

Common to the whole area is a curious morphological condition, whose raison-d'être I do not claim to understand. As Prince L. L. Bonaparte already noticed (Bonaparte 1862: 29), *Raa* does not apply to the forms of the plural, even though, there too, the definite article *a* can be suffixed to a stem ending in *a*. Note the following paradigms:

## SINGULAR

<i>sagar bat</i>	"one apple"	<i>alaba bat</i>	"one daughter"
<i>sagarra</i>	"the apple"	<i>alabea</i>	"the daughter"
<i>sagarrak</i>	"the apple" (erg.)	<i>alabeak</i>	"the daughter" (erg.)
<i>sagarrari</i>	"to the apple"	<i>alabeari</i>	"to the daughter"

## PLURAL

<i>sagar batzuk</i>	"a few apples"	<i>alaba batzuk</i>	"a few daughters"
<i>sagarrak</i>	"the apples"	<i>alabak</i>	"the daughters"
<i>sagarrak</i>	"the apples" (erg.)	<i>alabak</i>	"the daughters" (erg.)
<i>sagarrari</i>	"to the apples"	<i>alabai</i>	"to the daughters"

The forms *alabak*, "the daughters" (erg.) and *alabai*, "to the daughters" result from *alaba* + *a* + *k* and *alaba* + *a* + *i* by Vowel simplification. In a few Southern

regions (e. g., Arechavaleta, Orozco, Llodio) Vowel simplification does not take place and we find *alabaak* and *alabai* (cf. Ormaetxea 1925).

Notice, in particular, that the ending *-ak* triggers *Raa* when it represents ARTICLE + ERGATIVE, but not when it represents ARTICLE + PLURAL. I can find no phonological motivation for this difference in behaviour. Nils Holmer has detected an interesting difference in accentuation between the singular and the plural forms of nouns: *gixónak* "the man" (erg.) versus *gixonak* "the men", *auntzák* "the goat" (erg.) versus *áuntzak* "the goats" (cf. Holmer, § 9.9-9.13).

Yes, it is not at all clear whether, and if so, how, this is connected with the application of rule *Raa*.

Even more mysterious is the fact that *Raa* does not apply to the locative (inessive) case in the singular.<sup>a)</sup> We have:

<i>etxe bat:</i>	"one house"	<i>etxea:</i>	"the house"	<i>etxean:</i>	"in the house"
<i>gona bat:</i>	"one skirt"	<i>gonea:</i>	"the skirt"	<i>gonan:</i>	"in the skirt"
<i>anka bat:</i>	"one leg"	<i>ankea:</i>	"the leg"	<i>ankan:</i>	"in the leg"

The forms we would expect *\*gonean* and *\*ankean* do not exist anywhere (cf. Azkue 1926: 34).

Note, finally, that rule *Raa* represents a process of dissimilation, which, like most dissimilations in Basque,<sup>7</sup> is regressive. This means that the articulation of a preceding sound is modified by that of a following one. The contrary is true of assimilation, which, in Basque, is practically always progressive; this being one of the most salient differences between the phonology of Basque and that of the neighbouring Romance languages. (See Holmer, § 6.13).<sup>b)</sup>

Our next rule, *Rij*, can be written as follows:

$$Rij: \quad \emptyset \quad \rightarrow \quad \left[ \begin{array}{l} - \text{syll} \\ - \text{cons} \\ - \text{back} \end{array} \right] \quad / \quad \left[ \begin{array}{l} \checkmark \\ + \text{bi} \\ - \text{back} \end{array} \right] \quad \text{---} \quad \left[ \begin{array}{l} \checkmark \\ - \text{bi} \end{array} \right]$$

<sup>7</sup> Liquids form a notorious exception to this generalization about dissimilation being always regressive. From an original *\*ark-ar* (he-him, according to C. C. Uhlenbeck) we find the following forms: *alkar* (Sul.), *alkar* (Bizc., Guip., Ronc.), *elgar* (L. Nav. Sul.), *elkar* (H. Nav., L. Nav., Guip., Lab.) and *arkal* (Bizc.: Guemica, Orozco, Zomoza), all meaning each other.

The onomatopoeic noun *dardar* (H. Nav., L. Nav., Bizc., Guip., Lab.) «tremble» has the variants: *daldal* (Sul., Ronc.), *daldar* (H. Nav., Lab., Ronc.) and *dardal* (Guip.) (Cf. Azkue 1905).

However, the forms *arkal* and *dardal* may be due not to progressive dissimilation but to secondary metathesis, a process abundantly attested for Basque liquids. (See *FHV*, Chapter 17).

Note that it is always the first vowel that changes in the *alkar* case, never the second one.

<sup>a)</sup> This mystery was later solved by W. H. Jacobsen in his article "The Basque locative suffix", published in *Anglo-American contributions to Basque Studies*, pp. 163-168.

<sup>b)</sup> In all likelihood, this rule first came about as an external sandhi phenomenon: the insertion of a palatal glide (j) between two non-high vowels belonging to two different morphemes. This process can still be observed even now in the usage of some *bertsolaris* and other singers. It is reflected in the form *ateraia* (*DGV* III, 173), now evolved into *atera*, the definite form of the perfect participle *atera*. In the same way, the definite form of the participle *bota* is *botea*. Given that these forms are common to all of Guipuzcoa, rule *Raa* seems to have been operative at one time in the Guipuzcoan dialect also. See also *Mezea* "the Mass", used in Aitaun.

This rule inserts a glide *y* after a high front vowel (*i*), whenever this vowel is followed by a low vowel (*a, e, o*). So we get: *erri bat*, "one village"; *erriya*, "the village"; *erriyen*, "of the village"; *erriyok*, "these villages".

Usually, there will be a morpheme boundary after the first vowel *i*, but, if we want, we can have the rule apply also within a single morpheme, to account for the fact that just those dialects that have rule *Rijy* show *iya* where other dialects show *ia*: e. g. *biyar* versus *biar*, "tomorrow", *iya* versus *ia*, "almost".

In various regions, especially in the Eastern area of Bizcayan, this epenthetic *y* turns into a true consonant: *ř*, *ž* or *ǰ* (*dž*). Thus, in Placencia and Eibar, the forms given above are pronounced: *erriřa*, *erriřen*, *erriřok*, *biřar*, *iřa*. And in Guernica: *erriže*, *errižen*, *errižok*, *bižer*, *iže*, thus giving rise to a sound *ž* that does not otherwise exist in the language. We even have a minimal pair in *oriže* (from *ori* + *a*), "the yellow one", and *oriše* (spelled *orixe*), "that same one".

Rule *Ruw* applies after high back vowels, like *Rijy* after high front vowels:

$$Ruw: \quad \emptyset \quad \rightarrow \quad \left[ \begin{array}{c} - \text{syll} \\ - \text{cons} \\ + \text{back} \end{array} \right] / \left[ \begin{array}{c} V \\ + \text{hi} \\ + \text{back} \end{array} \right] \text{---} \left[ \begin{array}{c} V \\ - \text{hi} \end{array} \right]$$

This rule inserts a glide *w* after a high back vowel (*u*), whenever this vowel is followed by a low vowel (*a, e, o*). It is commonly spelled *b*, as it is homophonous with the intervocalic allophone of the voiced labial stop *b*. Bonaparte asserts: "le son de ce *b* est identique ou du moins fort ressemblant à celui du *w* anglais" (Bonaparte 1862: 33). Examples: *buru bat*, "one head"; *buruwa*, "the head"; *buruwen*, "of the heads"; *buruwok*, "these heads".

*Rijy* and *Ruw* can be collapsed into one rule:

$$\emptyset \quad \rightarrow \quad \left[ \begin{array}{c} - \text{syll} \\ - \text{cons} \\ \alpha \text{ back} \end{array} \right] / \left[ \begin{array}{c} V \\ + \text{hi} \\ \alpha \text{ back} \end{array} \right] \text{---} \left[ \begin{array}{c} V \\ - \text{hi} \end{array} \right]$$

The reason why I have listed the two rules separately, is that they do not always co-occur. According to data supplied in 1861 by Prince Bonaparte's highly reliable informant José Antonio Uriarte, Orozco in the South has *Ruw* but not *Rijy*, the central region, including e. g. Bermeo and Guernica, has *Rijy* but not *Ruw*, Marquina and Ochandiano in the East have both, while the valley of Arratia, containing e. g. the towns of Yurre. Ceberio and Ceánuri, has neither rule.

The next rule is again one of dissimilation:

$$Rea: \quad \left[ \begin{array}{c} V \\ - \text{back} \end{array} \right] \quad \rightarrow \quad \left[ + \text{hi} \right] / \text{---} \left[ \begin{array}{c} V \\ - \text{hi} \\ + \text{back} \end{array} \right]$$

This rule changes a front vowel (*e, i*) to *i*, before a low back vowel (*a, o*). Examples: *seme bat*, "one son"; *semia*, "the son"; *semi ori*, "that son"; *semiok*, "these sons".<sup>c)</sup>

Like *Rjy*, *Rea* also applies within stems, at least as a diachronic process: morphemes that have *ea* or *eo* in Basque varieties lacking rule *Rea*, present *ia* and *io* in the *Rea* area. So we find *biar*, "work", "need", and *bior*, "mare", whose original forms *bear* and *beor* are still found in the valley of Arratia and in many regions of the Guipúzcoan dialect area.

The specification +back is necessary in the environment of the rule, as, according to Bonaparte, *e+e* is not changed to *ie*: "L'e qui précède un autre *e* ne se transforme en *i* que dans la variété d'Orozco et de Barambio et de quelques autres localités limitrophes jusqu'aux environs de Bilbao. Cette règle toutefois ne s'applique pas au génitif pluriel, car *semeen*, 'des fils', ne se change pas en *semien*, quoique *deutsee*, 'ils leur ont' se transforme en *deutsie*" (Bonaparte 1862: 31).

According to Holmer's recent data (see Holmer 1968: § 10.8), *ee* turns into *ie* also in the area around Bermeo, specifically in Mundaca (*etxiatar*, "to the houses"; from *etxé + eta + ra*) and in Frúniz (*etxiatan*, "in the houses"; from *etxé + eta + n*). This is again an instance of simplification: the feature +back has been dropped from the environment of the rule.

Next on our list is rule *Roa*:

$$\text{Roa: } \left[ \begin{array}{c} \checkmark \\ + \text{back} \\ + \text{round} \end{array} \right] \rightarrow \left[ + \text{hi} \right] \quad / \quad \left[ \begin{array}{c} \checkmark \\ - \text{hi} \end{array} \right]$$

This rule turns *o* into *u* before a low vowel (*a, o, e*). Examples: *asto bat*, "one donkey"; *astua*, "the donkey"; *astuen*, "of the donkeys"; *astu ori*, "that donkey"; *astuok*, "these donkeys".

Intuitively, *Rea* and *Roa* appear to be instances of the same process of regressive dissimilation, and it seems natural to collapse the two rules into something like:

$$\left[ \begin{array}{c} \checkmark \\ \alpha \text{ back} \\ \alpha \text{ round} \end{array} \right] \rightarrow \left[ + \text{hi} \right] \quad / \quad \left[ \begin{array}{c} \checkmark \\ - \text{hi} \end{array} \right]$$

Then, however, we will need a later rule to turn *ie* back into *ee* (in those areas where *ee* does not turn into *ie*), which is feasible only where *Rjy* has eliminated all other *ie*'s. Curiously enough, it seems indeed to be true that all areas that have both *Rea* and *Roa* also have *Rjy*.

Furthermore, *Rea* can exist without *Roa*, e. g. in the central area around Guernica (including Frúniz, but not Bermeo and Mundaca), but in Bonaparte's rather abundant material there are no instances of a locality, either inside or outside the Bizcayan area, that has *Roa* but not *Rea*.

<sup>c)</sup> Like *Raa*, *Rea* does not apply between the members of a compound: *asteazkena* "Wednesday", never \**astiazkena*.

A quite different type of rule is *Rui*:

$$Rui: \left[ \begin{array}{c} \check{V} \\ -bi \\ -round \end{array} \right] \rightarrow \left[ -back \right] / \left[ \begin{array}{c} \check{V} \\ +bi \end{array} \right] C_0 \text{ ---}$$

Here  $C_0$  stands for any number of consonants. In actual fact, this number will be 0, 1, 2 or 3.

The rule states that *a* gets fronted, that is, turns into *e*, whenever there is a high vowel (*u*, *i*) in the preceding syllable. Examples: the numeral "one" is *bat* (often used as an indefinite article), the definite article is *a*: *akats bat*, "a defect"; *akatsa*, "the defect"; *mozkor bat*, "a drunkard"; *mozkorra*, "the drunkard". But, *egun bet*, "a day"; *egune*, "the day"; *erri bet*, "a village"; *errie*, "the village"; *intxaur bet*, "a nut"; *intxaurre*, "the nut"; *oin bet*, "a foot"; *oiñe*, "the foot".

Rule *Rui* applies also morpheme-internally, that is, it can be viewed as a morpheme structure condition: *inder*, "strength"; *iketx*, "coal"; *uketu*, "to deny"; *buzten*, "tail", correspond to the more general forms *indar*, *ikatx*, *ukatu*, *buztan* (See Rollo 1925: Vocabulary). As L. Michelena has brought to my attention, rule *Rui* is already evident in a 17<sup>th</sup> century Bizcayan Catechism known as the "Viva Jesús".

Like most of the rules discussed, *Rui* is also found outside the Bizcayan dialect area. So e. g. in Azpeitia (Guipúzcoa) and in Leiza (Navarra) but not in Areso, only 4 kms. to the West of Leiza. It is interesting to note, however, that in Leiza the rule does not apply within the same morpheme. Although we have: *egun bet*, *egune*, *erri bet*, *errie*, *intxaur bet*, *intxaurre*, *oin bet*, *oiñe*, just like in Marquina, we find: *indar*, *ikatx*, *ukatu*, *buztan*, and not the Marquinese forms: *inder*, *iketx*, *uketu*, *buzten*. Thus, in Leiza, *Rui* only applies when  $C_0$  contains a morpheme boundary. This means that *Rui* cannot be interpreted here as a morpheme structure condition, but only as a purely synchronic rule, restricted to cases where it creates alternations.

Rule *Rui* is interesting in yet another respect: The fronting of the *a* under influence of a preceding high vowel is clearly a case of assimilation in aperture, and, as such, easily understandable from an articulatory point of view. However, the way we formalized the rule, following standard practice in Generative Phonology, the feature composition of the rule does not immediately show this assimilatory character, since *bi* and *back* are separate features. It remains to be seen whether or not this should be considered a shortcoming of the present feature system as such.

We now have six rules: *Raa* (and its variant *Rao*), *Rij*, *Ruw*, *Rea*, *Roa*, *Rui*. Inspecting this sequence of rules, we see that their order of application is not indifferent. There is a critical ordering relation between *Raa* and *Rea*; *Rij* and *Rea*; *Ruw* and *Roa*; *Rea* and *Rui*; *Roa* and *Rui*. Thus, e. g., if *Rea* applies before *Rui*, we get *begie*, "the eye", from *begi* + *a*, and *semie*, "the son", from *seme* + *a*. But if *Rea* applies after *Rui*, we will have *semia* from *seme* + *a*, while still getting *begie* from *begi* + *a*.

Historically, rule *Raa* is a very old rule. As far back as our sources of knowledge reach, it has been a constant characteristic of the entire Western area of Vasconia, i. e., an area that includes the whole Bizcayan dialect and the Westmost part of the Guipúzcoan dialect as well. The antiquity of the rule is reflected in its position in

the synchronic phonology. Except for the Vowel Truncation rule, it is very rare for a rule to precede *Raa*: *Rea* nearly always treats *ea* derived from *aa* in just the same way as it does an underlying *ea*.

The other five rules, however, are more or less recent innovations (see *FHV*, § 5.1, p. 109 ff). In some cases, we can even see the rules gain new territory in quite modern times. A good example of this is Marquina. For this East-Bizcayan town we have the testimony of José Antonio Uriarte, who lived there from 1840 to 1860. In two of his letters to Prince Bonaparte (June 4 and October 28, 1861) he cites for Marquina forms like the following:<sup>8</sup> *alabia*, *semia*, *mendiya*, *astua*, *buruba* (definite forms of *alaba*, "daughter"; *seme*, "son"; *mendi*, "mountain"; *asto*, "donkey", and *buru*, "head", respectively), while giving the forms *alabie*, *semie*, *mendiye*, *astoa*, *burue* for the central subdialect. There is thus no doubt that in Uriarte's time, rule *Rui*, while prevalent in the centre, had not yet reached Marquina.

In the early nineteen-twenties, however, a Scottish linguist, William Rollo, spent two summers in Marquina and published a doctoral thesis *The Basque Dialect of Marquina* (Amsterdam, H. J. Paris, 1925). From his description, as well as from the numerous texts that accompany it, we learn that *Rui* had become an obligatory rule in the whole area: forms like *alabie*, *semie*, *mendixe*, *astue*, *buru(b)e* are the only ones used. See also his phrase: *udie nitzet*, *negue itzet* (page 40), "summer for me, winter for you", from underlying *uda* + *a ni* + *tzat*, *negu* + *a i* + *tzat*. Azkue, too, confirms these data and adds that Bermeo also has the forms ending in *-e*. Apparently unable to believe in linguistic change, he then denies the authenticity of Bonaparte's data (Azkue 1926: 25).

It happens to be a fact —noticed already by Bonaparte (op. cit.)— that when a locality has both rule *Rij* and rule *Rea*, the former always precedes the latter. That is to say, we get *begiya* from *begi* + *a*, and *semia*, but not *semiya* from *seme* + *a*.

Why is this? We cannot attribute the absence of the ordering 1 *Rea*, 2 *Rij* to a presumed tendency to ensure that distinct underlying forms also receive different phonetic representations, for some regions in Bizcaya and many outside it, have rule *Rea* but lack rule *Rij*, and thus do not distinguish *semia* from *begia*.

The rules *Rij* and *Rea*, at least in Bizcaya, seem to have spread in the same general direction. The synchronic order 1 *Rij*, 2 *Rea* may thus simply reflect chronological order. The Western areas (Llodio and Munguía, according to Bonaparte) that show *Rea* but not *Rij*, indicate that rule *Rea* kept spreading westward after rule *Rij* had already ceased to do so.

Notice, finally, that the non-existent order 1 *Rea*, 2 *Rij* is precisely the "feeding order", that is, the unmarked order in Kiparsky's sense.

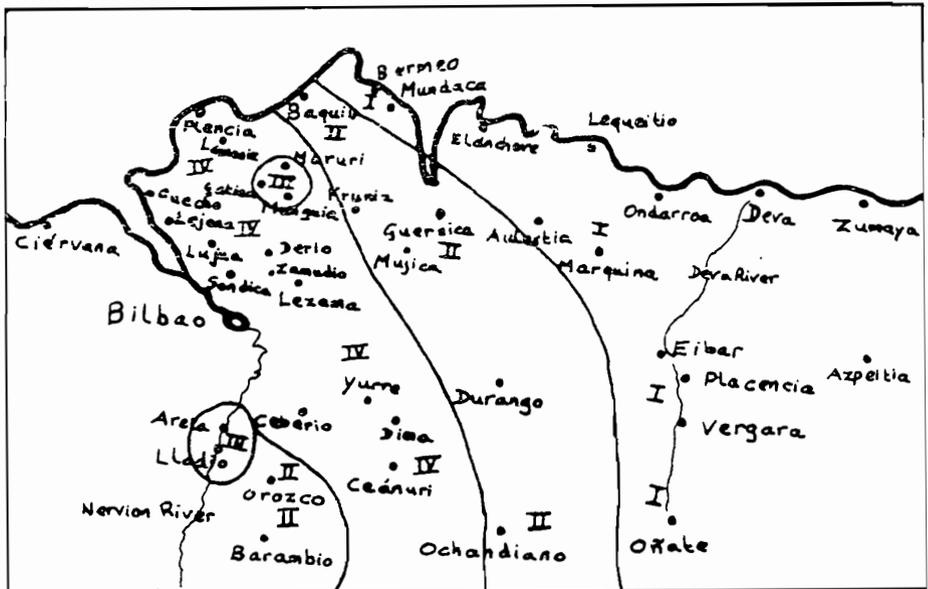
After the preceding section had been written, I learned that Nils Holmer obtained the forms *iya* (from underlying *ea*, "whether") and *etxiyen* (from underlying *etxe* + *a* + *n*, "in the house") from his Mundaca informant Mrs. de Bilbao (See Holmer 1968: § 6.11). Here, obviously, rule *Rij* has applied after rule *Rea*. What

<sup>8</sup> The same forms are also used by Juan Antonio de Moguel in his famous book *Peru Abarca*, written in about 1800 but first published in 1881. Moguel was born in Eibar, but lived in Marquina nearly all his life.

does this indicate? Does it mean that Mundaca is the only town in Bizcaya that was reached by rule *Rea* before it acquired rule *Rij*? If this is true, the exceptional status of Mundaca must have escaped the attention of all observers during the last hundred years, including that of Azkue. An extremely unlikely assumption for anyone acquainted with the enormous labor in Basque dialectology carried out by Azkue, who was, moreover, a native speaker of Bizcayan and a careful observer of its local variations. It is much more likely that these Mundaca forms are innovations due to a beginning tendency to invert the order of the rules. According to Kiparsky's general theory, the order 1 *Rij*, 2 *Rea*, where a later rule creates new strings that obey the structural description of an earlier rule, is a highly marked one and hence unstable. A diachronic process of reordering is therefore expected to occur. We seem to be witnessing here the very beginning of the process: in Mrs. de Bilbao's speech, forms generated by the original order of the rules are still by far the most frequent. Holmer reports the following forms from this same informant: *atia* (§ 7.2), *maistria* (§ 8.5), *aldiano* (§ 9.6), *egun guztien* (§ 9.16), *atiari* (§ 10.6), *etxietara* (§ 10.8). In all these examples rule *Rij* has not applied to the result of rule *Rea*.

The next and last case of critical ordering we will consider is that of *Rea* and *Rui*. Here the geographical distribution clearly shows that the two rules have travelled in different, almost opposite, directions. *Rea* spread from North to South and from East to West, whereas *Rui* seems to have been propagated from South to North and from West to East.

Assuming now that synchronic order corresponds to chronological order, we should expect to find four areas: I. *Rea* but not *Rui*; II. 1 *Rea*, 2 *Rui*; III. 1 *Rui*, 2 *Rea*; IV. *Rui* but not *Rea*. Interestingly, this is exactly what Bonaparte's data show us:



*Schematic representation of the geographical distribution in 1860*

Area I: Bermeo and Marquina.<sup>9</sup>

Area II: Ochandiano and the Central Subdialect (e. g., Guernica).

Area III: Munguía and Llodio (Alava).

Area IV: The Arratia valley (towns: Ceánuri, Ceberio, Dima, Yurre).

Area II, a rather larga one, has the "feeding order" 1 *Rea*, 2 *Rui*. Here rule *Rea* creates new instances for rule *Rui* to apply to.

The opposite order is found in area III, which represents all or part of the region that was first touched by rule *Rui* and only afterwards by rule *Rea*. The area is small and discontinuous, Llodio and Munguía, being about 35 kms. apart, and separated by a large piece of area IV, characterized by the absence of rule *Rea*.

It is worthwhile to quote Uriarte's own words<sup>10</sup> on Munguía. He writes to Bonaparte (July 7, 1861):

He estado en Munguía y sus inmediaciones. En Munguía son las eufonías de este modo: *alabia, semia, mendie, artoa, burue*. A Munguía siguen sólo dos pueblos, que son Gatica y Jatabe. En todos los otros pueblos, a saber: Lezama, Zamudio, Derio, Lujua, Sondica y todos los pueblos hasta Plencia con sus alrededores, son iguales a Arratia, a saber: *alabea, semea, mendie, artoa, burue*.

(I have been in Munguía, and its surroundings. In Munguía, the euphonic laws are as follows: *alabia, semia, mendie, artoa, burue*. Only two villages, those being Gatica and Jatabe, follow Munguía. In all the other villages, to wit: Lezama, Zamudio, Derio, Lujua, Sondica and all the villages up to Plencia and its neighbourhood, they are the same as in Arratia, to wit: *alabea, semea, mendie, artoa, burue*).

Since the time that Uriarte collected these data more than a hundred years have passed. Social interaction among the peasant population of Northern Bizcaya, like almost everywhere else, has greatly increased. Changes in the geography of the phonological rules are to be a priori expected. This author, therefore, decided to make a trip (in August 1969) to area III in order to investigate whatever changes have occurred.

One change is a rather sad one. The town of Llodio and its twin Areta no longer contain native speakers. They have become "erdaldun", Spanish-speaking. It is likely that in the surrounding country area some farmhouses still remain «euskaldun», Basque-speaking, but I had no opportunities to investigate this.

However, in the region of Munguía,<sup>11</sup> some interesting developments have taken place. First of all, in the countryside immediately East of Munguía, the situation described by Uriarte still subsist. A woman of about 50 years old born and raised on a farm named *Belako*, supplied these forms: *alabia, semia, mendie, astoa, burue*. That means she has the rules *Rui* and *Rea* in this order.

<sup>9</sup> We have seen before that Bermeo and Marquina had already acquired rule *Rui* around 1920. However, according to Holmer's data, Mundaca, a town 2 kms. East of Bermeo, remains even now virtually untouched by this rule.

<sup>10</sup> I am quoting directly from the original letters, which the library of the Bizcayan Provincial Deputation in Bilbao was kind enough to put at my disposition. The published version of Uriarte's correspondence (See Larrinaga) contains no less than three major errors in this small passage.

<sup>11</sup> I am greatly indebted to Don Paulino Solozábal for his eager help in contacting local informants.

The town of Munguía itself, however, has the following forms: *alabea* (freely alternating with *alabia*),<sup>12</sup> *semea* (but not *semia*), *mendie*, *astoa*, *burue*.

It is not impossible that this state of affairs existed already in Uriarte's time. We know that, on Bonaparte's advice, he preferred rural informants:

... he averiguado con todo cuidado las eufonías de este pueblo de Bermeo, y el subdialecto a que se extiende, valiéndome para esto, no de Eclesiásticos ni personas instruidas, sino de personas rústicas, y entre éstas, las que no tienen noticia de otros subdialectos... (Letter to Bonaparte, Bermeo, June 4, 1861).

(...I have investigated very carefully the euphonic laws of this town of Bermeo and of the subdialect to which it belongs making use, not of clergymen or educated people, but of rural people, and among these, those who have no knowledge of other subdialects...).

Yet it is strange that Uriarte with his usual concern for preciseness should not have reported this important difference between the town itself and some of its surroundings. Could it be that in his time no such difference existed, and that the town of Munguía has since dropped role *Rea*? If so, this seems to have been by virtue of an internal development rather than by the influence of the neighbouring area IV, since in the latter case *semia* and *alabia* should both have been reduced to *semea*, *alabea*. The reason, however, that *semia* but not *alabia* has been rejected, is that *semia* could be compared with *seme*, but there was no motive to prefer *alabea* over *alabia*, as *\*alabe* does not exist, and both forms have a stem final vowel distinct from that of *alaba*.

In a synchronic grammar of the town of Munguía, we must complicate rule *Raa* and add to it an optional feature:

$$\left[ \begin{array}{c} V \\ - bi \\ - round \end{array} \right] \rightarrow \left[ \begin{array}{c} - back \\ (+ bi) \end{array} \right] / \text{---} + \left[ \begin{array}{c} V \\ - bi \\ - round \\ + back \end{array} \right]$$

And, since *alabie* does not occur here, rule *Rui* has to precede this modified rule *Raa*.

It now remains to consider the area West of Munguía. Here we encounter an interesting rule of vowel assimilation that can be stated as follows:

$$\text{Ras: } \left[ \begin{array}{c} - bi \\ - round \end{array} \right] \rightarrow \left[ \begin{array}{c} \alpha bi \\ \beta round \\ \gamma back \end{array} \right] / \text{---} \left[ \begin{array}{c} \alpha bi \\ \beta round \\ \gamma back \end{array} \right] \text{---}$$

This rule is not mentioned by Uriarte or Bonaparte, but Azkue is well aware of it. Talking about a subdialect of Southern Guipúzcoa, he writes:

<sup>12</sup> This form *alabia* is in accordance with the data of the Triple Questionnaire. In Munguía, the form *arrebía* was given for "the sister", from *arriba* + *a* (See Ormaetxea 1925: 18).

La *a*, lo mismo cuando es artículo (que es lo general) como cuando forma parte del núcleo del verbo *yoan* (y en Legazpia aun de *eroan*), se dice entera tras las vocales *e* o, cuando no le sigue ningún sufijo: *semea*, *besoa*, *noa*, *oa*, *daroa*; mas si le sigue cualquier sufijo, se deja asimilar por la vocal precedente: *semeeke*, *besook*, *noonean*, “cuando voy”; *daroot*, “lo llevo”... etc. (1).

(1) En varias zonas del dialecto Bizcaino (comarca de Plencia) se hacen estas asimilaciones, aun cuando la vocal *a* no esté apoyada en ninguna consonante: *neu noo*, «yo voy»; *orrek daroo*, «ése lo lleva»; *neure semee*, «el hijo mío»; *besoo*, «el brazo» ... etc. (Azkue 1926: 23).

(The *a*, whether it is an article (as it generally is) or whether it is part of the stem of the verb *yoan* (= “to go”) (and in Legazpia also of *eroan* (= “to take away”)) is pronounced unchanged after the vowels *e* o, when no suffix follows: *semea*, *besoa*, *noa*, *oa*, *daroa*; but if any suffix follows, it is assimilated by the preceding vowel: *semeeke*, *besook*, *noonean*, “when I go”; *daroot*, “I take it away” ... etc. (1).

(1) In various areas of the Bizcayan dialect (Region of Plencia), these assimilations are made even when the vowel *a* is not checked by any consonant: *neu noo*, “I go”; *orrek daroo*, “he takes it away”; *neure semee*, “my son”; *besoo*, “the arm” ... etc.).

The more precise data supplied by the answers to the “Triple Questionnaire” show that rule *Ras* exists in the whole area North of Bilbao and West of Munguía, as the following villages are explicitly mentioned: Erandio, Lejona, Berango, Guecho, Barrica, Urdúliz, Plencia, Górliz, Lemóniz, Maruri (= Jatabe), Gatica, Lauquíniz. Moreover, the rule also exists in a small coastal area North of Guernica, with the villages Elanchove and Ibarranguelua, and nowadays, also in the town Ondárroa.

Accepting Uriarte’s testimony as valid, we know that around 1860 Gatica and Maruri had rules *Rui* and *Rea*, but lacked rule *Ras*. According to my own data, Gatica and Maruri now have rules *Rui* and *Ras*, but lack rule *Rea*. E. g. (Maruri): *nire semé gaur etor da*,<sup>13</sup> “My son has come today”, *nire alabé bere bai etor da*, “My daughter has also come”.<sup>d</sup> The difference between the definite and the indefinite forms, as I have heard them, is not so much in the length of the final vowel, but in the fact that it is stressed: we have *séme*, “son”, but *semé*, “the son”. That vowel contractions show up phonetically as stress is a well known characteristic of the Suletin and Roncalese dialects as well: *alhába*, “daughter”; *alhabá*, “the daughter” (See *FHV*, § 20.11).

Although the assimilated forms are preferred, we also find unassimilated ones used by the same speakers: *alabé* alternates freely with *alabea*, *semé* with *semea*, *asté* (“the week”) with *astea*. Forms with *i*, however, such as *alabia* or *semia* are rarely used and considered as foreign to the village.

That in former days Maruri did have rule *Rea*, can be inferred —apart from Uriarte’s testimony— from the fact that the form for “work” is *bear* and that for

<sup>13</sup> Notice the form *etor da* instead of the more extended form *etorri da*. *Etor da* is the usual form in Chori-erri, and according to Holmer’s data (§ 19.7) it also occurs in Bermeo.

<sup>d</sup> My data on the area around Munguía are nicely confirmed by Martin Olazar, from whose article “Mungia-aldeko Euskerea” (*Agur* III-28, jan. 15, 1972, p. 7) I quote: Some regularities in the north: In Jatabe, Gatica and Laukiniz they do not say *ogje*, but *ogi*, not *oilloa*, but *oillo*, not *basue*, but *basu* (“glass”), not *alabea*, but *alabe*. In the south, however: *alabia* (in Gamiz), *alabea* (in Munguía), *alabie* in Fruniz, Arrieta, Meñaca and Larrauri; similarly, *etxia*, *egije*, *oilloa*, *basue* (“glass”), *burue*. Going one step farther, in Bermeo, they say *oillue*, in Ondarrua *oillu*. (Translation mine, RdR.)

“mare” *bior*. In Arratia, those forms are *bear* and *beor* with *e*. We know that *Rea*, as a historical change, applied to non-alternating forms as well. Under the assumption that both forms once had an *i* in Maruri, we can easily explain why we now find *e* in one case and *i* in the other. The word *bear*, meaning both “work” and “need” is very frequent in everyday conversation, while the word *beor* “mare” is not. Therefore, when rule *Rea* was lost, *biar* was turned back to *bear* under the influence of the neighbouring Eastern area that never had rule *Rea*, but the rare word *bior* remained unchanged.<sup>e)</sup>

We see here that a marked order of rules can be disposed of in two ways: by reordering them, as in the case of Mundaca *etsiya*, or by simply dropping one of the rules, as Maruri has done with rule *Rea*.

Notice that the loss of rule *Rea* may have been caused by the addition of rule *Ras*. Indeed, if rule *Ras* was once obligatory, and if it was added before rule *Rea*, this rule would have had nothing to apply to.

In the costal area North of Guernica and in Ondárroa, rule *Ras* is ordered after the rules *Rea* and *Roa*. Thus in Elanchove, we find the alternations:

<i>alaba</i> : “daughter”	<i>alabí</i> : “the daughter”
<i>seme</i> : “son”	<i>semí</i> : “the son”
<i>beso</i> : “arm”	<i>besú</i> : “the arm” (cf. Ormaetxea 1925).

The corresponding derivations are: *alaba* + *a*  $\xrightarrow{Raa}$  *alabea*  $\xrightarrow{Rea}$  *alabia*  $\xrightarrow{Rai}$  *alabii*  $\rightarrow$  *alabí*.  
*Seme* + *a*  $\xrightarrow{Rea}$  *semia*  $\xrightarrow{Ras}$  *semii*  $\rightarrow$  *semí*. *Beso* + *a*  $\xrightarrow{Roa}$  *besua*  $\xrightarrow{Rai}$  *besuu*  $\rightarrow$  *besú*.

Despite the work of the past, some of which of a rather high quality, a lot of research remains to be done in Basque dialectology. We are still uninformed about a huge mass of important data, in Bizcaya as well as elsewhere in the Basque Country. The articles in which Ormaetxea reports the results of the Triple Questionnaire sent out in the early nineteen-twenties by the Basque Academy, are not as useful as they should have been. All too often, Ormaetxea writes statements like: “*noa*: in 47 villages, *noia*, *noie*, in 23 villages, *nua*, *nue*, in 23 villages”, without bothering to tell us what these villages were.

A linguistic atlas of Euskalerría is still very much needed. It is needed for the solution of many problems in the history of Basque that are at present insoluble. It will constitute an immense contribution to the full florescence of Bascology in its various branches. But, above all, the completion of a detailed and reliable linguistic atlas of Basque will be a great service to General Linguistics, as it will enable us to arrive at a better understanding of the mechanisms of linguistic change in general.

## References

- Azkue 1905: R. M. de Azkue, *Diccionario vasco-español-francés* (Bilbao, 1905).  
 Azkue 1926: R. M. de Azkue, “Observaciones acerca de la obra ‘Langue basque et langues finnoises’ del príncipe Luis Luciano Bonaparte”, *Euskera* 7 (1926) II-IV, 17-36.

<sup>e)</sup> In his noteworthy article «Butroiko euskara», Iñaki Gaminde likewise lists the forms *bearrá* “the work”, but *biorra* “the mare”. (*FLV* XIV-40, 403-460).

- Barandiarán: J. M. de Barandiarán, *El mundo en la mente popular vasca*, I, II, III, IV, San Sebastián.
- Bonaparte 1862: L. L. Bonaparte, *Langue basque et langues finnoises*, London, 1862.
- Holmer 1968: N. M. Holmer and V. A. de Holmer, "Apuntes vizcaínos", *ASJU* II (1968), 87-141.
- Kiparsky 1968: R. P. V. Kiparsky, "Linguistic Universals and Linguistic Change". In: E. Bach and R. T. Harms (eds.), *Universals in Linguistic Theory* (Holt, Rinehart and Winston).
- Lafon 1962: René Lafon, "Sur la voyelle ü en basque", *BSLP* 57 (1962), 83-102.
- Larrínaga: Fr. J. Ruiz de Larrínaga, O. F. M., "Cartas del P. Uriarte al Príncipe Luis Luciano Bonaparte", *BAP* 10 (1954), 231-302; 13 (1957), 220-239, 330-348, 429-452; 14 (1958), 396-443.
- Michelena, *FHV*: L. Michelena, *Fonética histórica vasca*. San Sebastián, 1961.
- Moguel 1881: J. A. de Moguel, *Peru Abarca*. 4<sup>th</sup> edit. Zarauz, 1956.
- Ormaetxea 1925: N. Ormaetxea, "Erixkizundi irukoitzari Bizkaiak erantzun diona (The Bizcayan Answers to the Triple Questionnaire)", *Euskera* 6 (1925), III-IV, pl. 16 ff.
- Rollo 1925: W. Rollo, *The Basque Dialect of Marquina*. Amsterdam, 1925.
- De Rijk 1969: R. P. G. de Rijk, "Is Basque an S.O.V. Language?", *FLV* 1 (1969), n.º 3.
- Woo 1968: Nancy Woo, "Tone in Northern Tepebuan". Paper read at the summer meeting of the L.S.A., Urbana, 1968.