4. Phonetics and Phonology

Ander Egurtzegi
(University of the Basque Country, UPV/EHU)

4.1. Introduction[*]

Fifty years after its original release, *Fonética Histórica Vasca* (= *FHV*, Michelena 1977 [1961]) continues to be the main reference insofar as Basque historical phonology is concerned. Although the seminal work of Michelena is, undoubtedly, the most significant milestone in the history of Basque diachronical phonology (maybe even within Basque diachrony itself), an update of the latest research postdating this work represents indeed an interesting endeavor.

There are, however, different ways to undertake such an investment of time and effort. In this work, I have chosen to go beyond a simple compilation of the articles on Basque historical phonology written in the last fifty years, by producing a new survey. Thus, this chapter is a summary of the most significant work on the history of the Basque language, but it also introduces comments and discussion on it.

The chapter is composed of seven sections including this introduction. Section §4.2 presents the three main stages of the language usually mentioned in the bibliography (namely Old Proto-Basque, Proto-Basque and the modern language) as well as their (reconstructed) phonemic inventories. Section §4.3 discusses the vowels and the processes that have affected them, while Section §4.4 focuses on the consonantal inventory in a similar way. Section §4.5 describes the phonotactics of the language and the possible consonant clusters, and Section §4.6 describes the different kinds of accentuation found in the modern dialects, proposing a possible relative chronology of the accentuation types and

their variants. Eventually, a brief Section §4.7 with some general conclusions is provided.

4.2. The Basque phonemic inventory: now, then, and before then

4.2.1. Modern Basque dialects

Modern Basque dialects vary to some extent in their phonemic inventory. Nevertheless, a more or less common inventory would include the phonemes usually linked to Standard Basque. These segments are presented in Table 4.1 and Table 4.2, together with some other added sounds, which correspond to those found in only a limited number of dialects (or even in one single dialect). The latter are represented in parentheses.[1]

From a typological point of view, an unusual distinction within the Basque consonantal inventory is the phonemic opposition between six different voiceless sibilants: an apical, a laminal and a post-alveolar, along with their affricated counterparts. In Souletin and some Low Navarrese varieties, there is also a series of voiced fricative sibilants —and even an apico-alveolar affricate— which are phonemically distinct (cf. Hualde 1993b).

When carefully studying the segments which are present in only a few varieties, the historical linguist must distinguish between the segments which represent an innovation and those present in an older stage of the language but which are only peripherally present today after having become lost in most of the modern dialects.

---

1   Hualde (*GB*: 15) offers a simpler table containing only the segments found in Standard Basque. In addition, he includes a table with the phonemic inventory of the easternmost dialect, Souletin (*GB*: 18), usually described as the most deviant variety of the Basque language.

Table 4.1.   Consonantal inventory of modern Standard Basque

		stop		fricative		affricate		nasal	lateral	tap	trill	glides
		voice-less	voi-ced	voice-less	voi-ced	voice-less	voi-ced					
bilabial	labial	p (pʰ)	b					m				
labio-dental				f								
apico-dental	apical	t (tʰ)	d									
apico-alvelar				ṣ	(ẓ)	ţ͡ṣ	(ḓ͡ẓ)	n	l	ɾ	r	
	laminal (alveolar)			s̺	(z̺)	t͡s̺						
palato-alveolar	predorsal		ɟ		(ʒ)							
palatal		c (cʰ)	g		(j)			ɲ	λ			j
	postdorsal	k (kʰ)		x								w
	laryngeal			(h)	(h̃)							

Table 4.2.   Vowel inventory of modern Standard Basque

	front		central	back
high	i (ĩ)	(y) (ỹ)		u (ũ)
mid	e (ẽ)			o
low			a (ã)	

4.2.2. Michelenian stages

In fact, as we go back in time, we find a very different scenario: the reconstructed Michelenian Common Basque (cf. Michelena 1987 [1981a]) from the beginning of the Middle Ages (5th-6th centuries AD), and Proto-Basque (cf. *FHV*), dating to around the 3rd century BC. After separating the wheat from the chaff, fewer than half of the segments found today had a phonemic status at those early stages, and most of these only after having suffered significant changes (cf. Martinet 1974 [1950], *FHV*, 1988 [1951b]). As a matter of fact, the PB. system was not grounded on the opposition between voiced and voiceless stops or fricative and affricate sibilants, but on a more general fortis versus lenis distinction which involved almost the entire consonantal inventory (apud *FHV*, although some authors, such as Hualde 1999 or Egurtzegi, in this paper, are skeptical about this reconstruction).

   As Table 4.5 shows, the reconstructed vowel inventory contains the same vowels as in modern times. The sixth vowel from Souletin (the front high rounded vowel /y/) as well as the nasalized series are both regarded as innovations, although they have a very different chronology. The consonant inventory in Tables 4.3 and 4.4, on the other hand, looks very different shape in the stages of the language reconstructed by Michelena from the modern one, with an opposition dividing the system into two groups —the aforementioned fortes and lenes— in an almost regular way. However, this is still incomplete given the lack of the labial fortis in the inherited lexicon, as discussed by Michelena (*FHV*: 261, §4.4.2.2).

   The first proposal for such an opposition comes from Martinet (1974 [1950]), who hypothesized that Proto-Basque may have had a system analogous to that of the Danish language. He proposed an opposition between two series —namely fortis and lenis— that had different phonetic realizations according to the position of the word in which they were produced. The fortis series was produced as aspirated and voiceless in a prominent position and as plain voiceless preceding an unstressed vowel, while the "soft" or lenis stops were realized as devoiced stops in a prominent word-initial position and as voiced fricatives in an

unstressed position. Table 4.6 illustrates the hypothesized phonetic realization of the stops, according to Martinet (1974 [1950]: 533).

Table 4.3.   Consonants involved in the fortis/lenis opposition[2]

	stop			sibilant		sonorant		
	labial	dental	velar	apical	laminal	nasal	lateral	rhotic
fortis	(P)	T	K	tṣ̌	tṣ̱	N	L	R
lenis	p	t	k	ṣ̱	ṣ̌	n	l	r

Table 4.4.   Consonants absent from the fortis/lenis opposition

	labial	palatal	post-dorsal	laryngeal
fricative	(f)			h
nasal	(m)			
glide		(j)	(w)	

Table 4.5.   Vowels reconstructed for Proto-Basque

	front	central	back
high	i		u
mid	e		o
low		a	

Table 4.6.   Hypothesized phonetic realization of stops

Phonemes	Initially (stressed)	Intervocalically (unstressed)
/p/	[pʰ-]	[-b̥-]
/b/	[b̥-]	[-β̞̊-]
/t/	[tʰ-]	[-d̥-]
/d/	[d̥-]	[-ð̞̊-]
/k/	[kʰ-]	[-g̊-]
/g/	[g̊-]	[-ɣ̞̊-]

Taking this system as a starting point, Martinet proposes a word-initial lenitive process involving the fricativization of the aspirated stops. This is followed

---

2   Here we follow Michelena's notation (*FHV* and elsewhere), using uppercase letters to denote the fortis and lowercase for the lenis series.

by a subsequent debuccalization that could have resulted in the complete loss of the segment —i.e. /t^h-/ > /θ-/ > /h-/ (> Ø)—, along with an initial voicing of the lenis series, which the author attributes to neighboring Romance pressure, since "in Latin pronunciation, the articulatory energy was equitably distributed along the chain" (ibid.: 54 [my translation, AE]). Michelena (i.a. 1988 [1951b], *FHV*) accepts and expands this system with the addition of the sonorant segments to which, interestingly enough, he does not attribute any kind of contextual divergence in their phonetic realizations.

To the consonants in Table 4.6, two different groups of consonants may be added. Firstly, the aspiration, which was most probably part of that stage of the language, even though it did not form part of the opposition proposed by Martinet (1974 [1950]) and expanded by Michelena (1988 [1951b], *FHV*). Secondly, the segments /m/ and /f/ along with the glides *yod* and *wau* and the aforementioned /P/ were not as usual as the other segments and were probably of later introduction. While the former was almost undoubtedly part of the previous stages of the language, the latter labial segments seem to be secondary in older times (some maybe even marginal) and it is possible that they were mostly present in the borrowed lexicon.

Michelena (*FHV*: 374) cites the adaptation of Latin and Romance medial voiced stop clusters as evidence of a tense/lax contrast in opposition to the modern voiced/voiceless contrast. The fact was that all geminates and heterorganic voiced stop groups were adapted to voiceless stops with the place of articulation of the prevocalic segment, as in Lat. *abbas* > *apaiz* 'priest' or Old Sp. *cobdiçia* > *gutizia* 'desire, whim'. An alternative account, however, would be that medial voiced stops were realized as approximants —as they are in the modern language and already were in the 12th century (cf. §4.4.2.1)— and that voiced stop clusters could be perceived as the only existing voiceless stops, rather than word-internal voiced approximants. Along the same lines, he also cites the regular adaptation of Latin /f/ as Basque /b/. However, the latter may have been the only labial consonant available in the language and thus the only suitable option, given the lack of /p/, as proposed by Michelena himself.

4.2.3. Lakarran Old Proto-Basque

If we move into even older times, some centuries before the time estimated by Michelena for PB (approx. 3rd century BC) we encounter a very different state of the Basque language which we can label Old Proto-Basque. This stage has mainly been reconstructed by Lakarra (1995, 2005, 2011b, etc.), and is a direct consequence of his research on the old canonical root structure which he found

to be a monosyllabic CVC, lacking any consonantal clusters or diphthongs (cf. Lakarra 1995 and elsewhere).

Thus, Old Proto-Basque has much simpler phonotactics than the following stages of the language, allowing only two possible consonantal positions (i.e. C_1VC_2) instead of the initial/medial/final distinction made in both Proto-Basque and modern Basque. Thus, we distinguish between onset and coda consonants in Old Proto-Basque, as illustrated in Table 4.7.

Table 4.7.   Consonants in Old Proto-Basque according to their syllabic position

	stop			fricative			sonorant		
	lab	dent	vel	apical	laminal	glottal	nasal	lateral	rhotic
onset	b	(t), d	(k), g	s̺	s̻	h	n	-	-
coda	-	-	-	s̺	s̻	-	n	l	(r, ɾ)

It is uncertain what the total number of proposed stops and rhotics should be, or what was the nature of the plosive series. The reconstructed monosyllabic roots involving voiceless stops (such as *tor* or *kar*) are scarce, but cannot be derived from other segments either.

Little research has been carried out on the Proto-Basque vocalic system, and thus it is even harder to know the shape of that system in previous stages of the language. Until new results are found, a 5-vowel inventory is reconstructed by default, since it is the one found in the modern stages of the language and in any written document, except for those written in the Souletin dialect, with a front high rounded vowel as well as a nasalized series (§4.3.2 and §4.3.3).

4.3. Vowels

4.3.1. Main vowels

Basque vowels do not seem to have changed to a great extent over time. The five vowels divided into three different degrees of openness are found in the majority of the modern dialects —/a, e, i, o, u/— directly correspond to the vowels reconstructed for the Michelenian Proto-Basque.[3] An uncertain point for this 5-vowel reconstruction was the alternation between /i/ and /u/ and the possibility

---

3   To this common inventory, certain (mostly local) innovative segments may be added, namely, the Souletin high front labialized /y/, the nasalized series from Souletin and Roncalese, and the Biscayan "sensibly palatal, relaxed and slightly labialized" *a* from Gernika, as described by Navarro-Tomás (1923: 50).

of such an alternation being the consequence of an older phoneme, an option which had already been discarded by Michelena (*FHV*: 73, §4.3.4.1).

Some varieties maintain a vocalic gemination or even a length distinction (cf. Michelena *FHV*: 111-112), which originated from the encounter of two similar vocalic segments in derivation, in composition, or after the loss of an intervening consonant. This distinction is especially frequent in the Biscayan dialect, although not necessarily limited to it. Examples of this include Old B. *errekaak* 'rivers, streams', *domeekaak* (< Lat. *dominica*) 'Sundays', *arabaar* 'Alavese' and B. *haar > har* 'worm'. Contraction is the most usual outcome in other dialects.

Vowels are characterized by two seemingly contradictory tendencies (apud *FHV*: 50): they tend to be affected by neighboring sounds but they have not undergone any deep or generalized changes.

Following the observations of Gavel (1920) and Navarro-Tomás (i.a. 1923), Michelena points out that most vocalic permutations occur either between front vowels (i.e. /a, e, i/) or back vowels (/o, u/), although there are some other processes which affect vowels grouped according to their height[4] (*FHV*: 59). Likewise, he notes that the ease for vowel interaction increases in proportion to the number of syllables present in the word (*FHV*: 70).

4.3.2. Nasalized Vowels

The Basque vocalic phonological inventory has historically possessed a series of phonemically contrastive nasalized vowels. These principally resulted from the loss of the nasal environment of the vowels surrounding an intervocalic /n/, which would have become a nasalized aspiration at around the beginning of the Middle Ages (cf. *FHV*: 300-303, Igartua 2011, §4.4.1.3.1 and §4.4.3.1.2).

Today, these nasalized vowels are still found in the modern Souletin dialect exactly as they were found in the defunct Roncalese variety, which disappeared in the 20th century. However, according to Michelena (*FHV*: 48), they represent an "archaic particularity on its way to extinction" even in Souletin, due to their unstable condition. Souletin lacks the nasalized /õ/, since this systematically became /ũ/ after the development of the rounded high front vowel /y/ (*FHV*: 48). Souletin nasalized vowels appear in two different environments: surrounding a nasal (or nasalized) consonant, as in *khatiña* [kʰatĩɲã] 'chain', *gazna* [gaz̃nã] 'cheese', *ama* [ãmã] 'mother' and *ahate* [ãh̃ãte] 'duck'; or in a stressed final position, after the apocope of an old /-n/ (Hualde *GB*: 31): *salu* [s̺aˈlũ] (cf. Sp.

---

4   See, e.g., reciprocal metatheses involving vowels (cf. Egurtzegi 2011, §4.3.4.6)

salón) 'living room', *bedezi* [beǯe'ʂĩ] (< Fr. *medecin*) 'physician', *ardu* [ar'ǫ̃ũ] (< *ardano*) 'wine'. Although the former context may be a coarticulatory outcome with no phonemic consequences, the latter clearly contrasts with word-final non-nasalized vowels in other oxytones.

Historically, nasalized vowels are undoubtedly found in Biscayan until at least the 16th century, as documented by various contemporary authors, and they are present in Biscayan texts such as *Refranes y Sentencias* despite the lack of any form of notation to represent nasal phonemes in several words (cf. *FHV*: 48-49). They are also explicitly attested in texts such as those of Madariaga and the manuscript from Garibay.

Michelena assumes that nasalized vowels were present in further varieties of the language (*FHV*: 50), noting that they can be attested as far back as the 12th century in a word list included in *Le guide du pélerin de Saint-Jacques de Compostelle*,[5] in words such as *ardum* 'uīnum' and *arraign* 'piscem'.

As for the loss of this nasalized series in most dialects, this was probably the consequence of a natural internal development: a reinterpretation of those nasalized vowels as a /Vn/ sequence; an example of sound change by CHANCE (Blevins 2004) or a hypercorrection *à la* Ohala (1993: 250, 255), as in *arrãi*, *artzãi* becoming *arrain* 'fish' and *artzain* 'shepherd' in Biscay (cf. Egurtzegi 2011). Nevertheless, this process may have been conditioned by neighboring languages, given that varieties in contact with Spanish —which does not possess nasalized vowels— lost their nasalization at around the end of the Middle Ages, while the easternmost Souletin[6] —which is in close contact with the Gascon dialect of Occitan— is the only dialect that has maintained the opposition until today. However, it is worth noting that the other continental dialects (i.e. Labourdin and Low Navarrese) have lost the opposition, in spite of them all being in contact with French. It should also be considered that the aforementioned Biscayan and the peninsular dialects may have undergone different developments (and perhaps even chronologies) regarding this series.

4.3.3. Distribution of the Souletin /y/

Out of all the modern Basque dialects, only one, Souletin, has developed a vocalic system which diverges to such an extent from the reconstructed Proto-Basque vowels that it includes a sixth item. This item is the high front rounded

---

5   Text edited by Jeanne Viellard in 1938, *Le guide du pélerin de Saint-Jacques de Compostelle, texte latin du XIIe siècle* (Aimery Picaud, *Liber Sancti Jacobi*, *Codex Calixtinus*) (cf. *TAV*: 49-51).

6   Alongside its close but now extinct relative, the Roncalese.

vowel /y/. According to Michelena (*FHV*: 52), the realization of the S. /y/ is closer to the French /ø/ than it is to the French /y/. Hualde (1993b: 290) also describes it as being an intermediate sound between French /y/ and /ø/ and further notes, referring to Larrasquet (1932) and Lafon (1999 [1958]), that this fact is consistent with the fact that the other two high vowels of Souletin Basque (/i/ and /u/) are also more open than in French (Hualde 1993b: 290-291).

This vowel, which usually appears as <ü> in modern orthography,[7] developed from an older /u/ common to the other dialects and, although it can still easily be tracked back to that segment by means of cross-dialectal comparison, it is present in the very oldest documents written in this dialect.[8] The development of this vowel did not involve the loss of the older /u/ in Souletin, given that, although the /u/ fronting was almost context-free, it may have been hindered by certain consonants or clusters following the etymological /u/ (cf. infra). In addition to this, a later /u/ was also developed from /o/ in certain conditions, e.g. before a tautosyllabic nasal (cf. *FHV*: 54, and §4.3.4.2 here), after the fronting of the etymological /u/ was already complete.

The main difference between the /u/ fronting in Souletin Basque and the analogous processes in the neighboring Gallo-Romance languages is the context dependency which the Souletin process seems to exhibit. The contexts that block the change are not only far from being homogeneous or unifiable based on some phonological feature —this is not even the case when using a group of them— but they also seem to be very similar to other contexts in which the sound change systematically occurs. Michelena (*FHV*: 52-53) focuses on these seemingly similar sequences with different outcomes, previously compiled by Uhlenbeck (1909/10 [1903]), Gavel (1920) and Lafon (1999 [1937], 1999 [1958], 1999 [1962a], 1999 [1962b]). The following list is based on Oñederra's (2009b: 671) summary of the contexts mentioned by Michelena, who focused on those in which the sound change failed to take place (here reproduced under a.). Michelena also mentions some seemingly similar sequences of following consonants which did not hinder the process (under b.):

---

7   And as <u> in older ones, in which /u/ will usually be reproduced as <ou>.

8   It is present in Sauguis (Urquijo 1908/09) and Oihenart, who describes this vowel (Arnaud d'Oihenart, 1657, *Les Proverbes Basques recueillis par le Sr. d'Ohenart, plus les poésies basques du mesme auteur*, Paris), and represents it as <u>. It is also present in the first book written in Souletin (Jean de Tartas, 1666, *Onsa hilceco bidia*, Orthez), which is also consistent with this tradition (cf. *FHV*: 52).

a. contexts that do not yield fronting (*u > u*):[9]

a.1 / _ɾ (before tap): *gu(r)e* 'our', *gu(r)i* 'us (dat.)', *barur* 'fasting', *hirur (hiu)* 'three', *hur* 'water', *u(r)in* 'fat, grease'.

a.2 / _rð, rt, rtʰ (before rhotic-dental cluster): *urde* 'pig', *urdin* 'cloudy (water)', *urthe* 'year'.

a.3 / _s̠ (before apical fricative sibilant): *uste* 'supposition', *ikhusi* 'see', *busta*, *busti* 'wet'. But see also *üstel* 'rotten', *eüskara* 'Basque' and *deüs* 'something, nothing (when negated)'.

b. similar contexts that yield the change (*u > y*):

b.1 / _r (before trill): *hür* 'hazelnut', *ürrin* 'smell'.

b.2 / _rp(ʰ), rβ, rk(ʰ), rγ (before rhotic-labial or rhotic-velar clusters): *ürpho* 'manure pile', *bürkhi* 'birch', *ürgaiztü* 'help'.

/ _rl (before rhotic-lateral clusters): *khürlo* 'crane'.

/ _rs̠, rs̺, rʃ (before rhotic-sibilant fricative clusters): *ürsañ* 'sneeze', *ürzo* 'dove', *ürxantx* 'squirrel'.

/ _rh (before rhotic-aspiration clusters): *ürhe* 'gold', *ürhats* 'step'.

b.3 / _s̺, ts̠, ts̺, (before laminal fricative sibilants and both apical and laminal affricates): *güzi* 'all', *hüts* 'void', *hütz* 'flatulence'.

As Oñederra (2009b: 666) observes, it is a very difficult task to find any feature that would unify those environments, and even more so to find any that would also give reasons for the blocking of the fronting process. Nevertheless, it seems rather plausible that the sound change has developed within the Basque language, given not only its unusual context, not present in any neighboring language, but also the fact that it does not show any affinity for loanwords (either from Gascon or from any other Gallo-Romance language) and thus cannot be straightforwardly explained as a borrowed development.

However, in the literature, the process has mainly been described as a result of the borrowing of this vowel from Occitan dialects, as proposed by Gavel (1920), Lafon (1999 [1962a], 1999 [1962b]) and Zarabozo (1972), among others. Hualde also speaks of it as undoubtedly resulting from its contact with Gascon (1993b: 290), but does not further specify as to its origin. Oñederra (2009a: 56-65, 2009b) takes a different stand and proposes an origin within the language itself, hypothesizing a rule telescoping. According to her, some gradual steps would have taken place between the /u/ and its actual outcome, in which a diphthongization and a subsequent monophthongization would have occurred

---

9   Lafon (1999 [1937]: 88) noted that the fronting also failed to apply before /k/ and /nk/ in the LN. dialect (variety of Mixe), and additionally before /g/ and /n/ in the variety from Bardoze (*FHV*: 53).

(2009b: 674).[10] However, she does not formulate any strong hypothesis —given the difficulty in linking the aforementioned contexts to the proposed development— nor does she explain the need for a diphthongization.[11] Against this hypothesis, crucially, stands the fact that there is no diphthong which is present in the other dialects but missing in Souletin, as may have been expected if they had yielded /y/.

The neighboring Gascon may in fact be relevant for the development of a split of the /u/ into /u/ and /y/, given that multilingualism or very close linguistic contact makes speakers more sensitive to subtle sub-phonemic contrasts, and that these contrasts may be prone to phonemicize over time (Blevins, p.c.). In this possible scenario, two different allophonic variants could have been present in the language, with back rounded vowels occurring before /ɾ/, /rð, rt, rtʰ/ and /s̺/, and their fronted allophones realized elsewhere.

When attempting to explain the contexts in which the different allophones are produced, certain details seem to be particularly pertinent, especially the division between rhotics.[12] Therein, the tap blocks the process, while the trill appears not to do so, with the exception of the trill/dental stop cluster, which is the only case in which the process is also blocked. Here, we could imagine a place assimilation which would make the coda rhotic a dental,[13] and we can wonder whether anything distinguishes this type of trill from the others. Indeed, Ladefoged and Maddieson point out[14] that trills usually have a lowered third formant, in contrast to uvular and dental *r*-sounds, which have a relatively high F3 (1996: 244). This could in fact be the phonetic cue favoring the non-fronting of /u/, given that the /u/ > /y/ change acoustically involves a lowering in the third formant along with the rise in F2.[15]

4.3.4. *Vocalic processes*

4.3.4.1. Alternations between /i/ and /u/, /e/ and /o/

As mentioned above, the only reason behind proposing a larger vocalic inventory for older stages of the language is based on the alternation between the high

---

10   Here, she follows Donegan (1978: 96).

11   If not by means of a phonetically more natural explanation, as understood within Natural Phonology.

12   Recall also that the block of the fronting preceding an apico-alveolar sibilant does not seem to be exhaustive.

13   As Michelena does (cf. *FHV*: 473).

14   Following Fant (1968).

15   This is a proposal which I intend to develop further in future works.

vowels. However, neither the quantity of the affected words nor their geographic distribution seem large enough to accept this hypothesis (*FHV*: 73).

There is a group of words with a general initial /i-/ which show an /u-/ in Biscayan and some other peninsular dialects instead of the /i-/: *igeri : uger* 'to swim', *ile : ule* 'hair', *(h)iri : uri* 'village', *iltze : ultze* 'nail'. Michelena (*FHV*: 280) suggests that the unprotected /u-/ may become /i-/ in disyllabic words when followed by an anterior vowel, and specifies that some intervening segments may block the assimilation, such as a trill or a rhotic-consonant cluster. There are some instances of this alternation which do not necessarily occur in word-initial position and which are also found in eastern dialects: *mutil* > R. S. *mit(h)il* 'boy'; *zubi* > R. *zibi* 'bridge'; *zumel* > HN. *zimel* 'withered'; *puxika* > LN. *bixika* 'bladder'.

The Roncalese dialect has regularly assimilated /i/ to /u/ when followed by /u/ in the following syllable, while Souletin has done so with /y/ instead, leaving the instances of /u/ unaffected (*FHV*: 79): *tipula > *tupula > tupla* 'onion', *ilun > ulun, ülhün* 'dark', *irun > urun, ürün* 'spin (thread)', Sp. *hechura > itxura > ütxüra* 'appearance', *ekuzi > ikuzi > üküzi*[16] 'wash'. The reverse perseverative assimilation (/i/ to /u/ when preceded by /u/) also occurs in the same dialects, but in a sporadic manner: *burdina > burruña, bürdüña* 'iron', *urritz > urrutx, ürrütz* 'hazelnut tree'.

Alternations between /e/ and /o/ are not as usual as those involving high vowels, but Michelena also provides examples of some of these, especially in eastern dialects (*FHV*: 82-83): *sendo* > R. *sonto* 'strong', *behor* > S. *bohor* 'mare', *txekor* > HN. *txokor* 'young bull', *oren* > HN. *oron* 'for an hour'. According to Hualde, a similar assimilation is also found in the Biscayan variety of Bermeo (*GB*: 56): *etorri > otorri* 'come', *emon > omon* 'give', *erosi > orosi* 'buy', *eterno > otorno* 'funeral'.[17]

4.3.4.2. Different vowel risings

In addition to the diphthongization of mid vowels discussed in §4.3.5, there are some other, more common processes which result in a syllabic vowel becoming higher than before. These processes can affect any of the three non-high vowels in the Basque phonemic inventory and can occur not only by assimilation to a high vowel but also by the addition of a following vocalic segment.

---

16   The assimilation occurs after the mid vowel rising /e/ > /i/ (§4.3.4.2), as in the examples from Michelena (*FHV*: 79).

17   Hualde refers to Pérez-Bilbao (1991) in relation to this last example.

When preceded by a high vowel, /a/ has been raised to /e/ in multiple areas of the Basque Country: *ogia > ogie* 'the bread', *txakurra > txakurre* 'the dog', *katua > katue* 'the cat', *doguna > dogune* 'that we have', *izan > izen* 'to be', *ura > ure* 'water', *indar > inder* 'strength'. More precisely, the low vowel has been perseveratively assimilated to become a variant of /e/, which, according to Michelena (*FHV*: 47, 63, who mentions Navarro-Tomás 1923: 50), would be slightly more open than the standard. This process was already present in the 17th century.[18]

The high vowels have also anticipatorily assimilated a preceding mid vowel, especially in trisyllabic words. In such cases, /e/ and /o/ from the first syllable rise to /i/ and /u/: *hegi > higitu* 'move', *epistiko > ipiztiko* 'bishop', *mediku > midiku* 'doctor', Old Sp. *cobdiçia > gutizia* 'desire, whim' (cf. however, the vowel in **cupiditia* < Lat. *cupiditas, -ātis*), *ekusi > ikusi* 'see', etc. This process is more general and thus older than the progressive assimilation described above (apud *FHV*: 64).

A rising of the /o/ to /u/ occurs throughout the continental dialects, and is at its greatest strength in the Souletin dialect. There, this rising has developed regularly before a tautosyllabic /n/, being sporadic in the other dialects on the French side of the border: *gizon > gizun* 'man', *hon > hun* 'good', *untza* (cf. Sp. *onza*) 'ounce', *ezkondu > ezkuntü* 'married', *ontzi > untzi* 'container, ship'. This change is also present in some other barely specifiable contexts: *kozka > kuzka* 'hit', *molde > mulde* 'manner', *motz > mutz* 'short', *nor > nur* 'who', *nola > nula* 'how'.[19] Michelena disregards any connection to the similar development found in Bearnese, especially due to the fact that it is most present in the inherited lexicon and absent in certain borrowings which would be susceptible of developing the process, such as S. *botz* 'voice', in contrast to Bearn. *bouts* (*FHV*: 55).

Vowel rising develops regularly when any vowel is placed after a non-high vowel in inflection. This can occur both in nominal stems (with the addition of the article *-a*) and verbal stems (before the relative suffix *-en*, among others). The given contexts imply a late chronology for this process, since such vocalic clusters were only created after the loss of aspiration in the definite article during the Middle Ages.

The low vowel /a/ is raised to /e/ in Biscayan, Alavese and some Guipuscoan varieties under the aforementioned circumstances (*FHV*: 114-115):

---

18 As shown by the Biscayan text *Viva Jesus* (cf. Ulibarri 2010).

19 However, the vowel change in the interrogative pronouns could be a consequence of an analogy with the contextually conditioned *non > nun* 'where'.

neska-a > neskea 'the girl', *alaba-a > alabea* 'the daughter', *doa-an > doean* 'that goes', *daroa-anean > daroeanean* 'when (he/she/it) carries'.

The mid vowels also undergo rising when in stem-final position, after a following vowel is added to the stem in derivative processes. This rising is regular in many modern varieties of Basque and affects either /e/ or both /e/ and /o/ (apud Hualde *GB*: 47), as shown by the following examples: *etxe-a > etxia* (> *etxie*, cf. the aforementioned low vowel assimilation) 'the house', *gazte-a > gaztia* 'the young', *kafe-a > kafia* 'the coffee', *asto-a > astua* 'the donkey', *ard(a)o-a > ard(a)ua* 'the wine'.

Finally, there are some sporadic instances of /e/ > /i/ rising before /n/ or a sibilant which are mentioned by Michelena (*FHV*: 67-68), such as **e-nor > inor* 'nobody', *etxe > itxe* 'house', *etxola > itxola* 'shed', *(h)ertsi > itsi, itxi* 'close' and *pezu > pizu, pisu* 'weight'.

4.3.4.3. Vowel lowerings

A following trill, located either in the coda or at the onset of the following syllable, has opened /e/ to /a/.[20] This has occurred mainly in the western dialects, but only when /e/ was not at the very beginning of the word: *berri > barri* 'new', *txerri > txarri* 'pig', *gernu > garnu* 'urine', *piper > pipar* 'pepper', *maizter > maiztar, maistar* 'teacher' (< Lat. *magister*), *izter > iztar* 'thigh', *musker > muskar* 'lizard'. Nevertheless, there are also some examples of this lowering in the eastern dialects, such as *at(h)erbe > at(h)arbe* 'lodging', *pert(h)ika > part(h)ika* 'pole', *txert(h)atu > xart(h)atu* 'insert'.

Likewise, a tautosyllabic trill has sometimes turned /i/ into /e/ as well as /u/ into /o/ in the western dialects, or, at least, has produced some variation among these: *irten > erten* 'exit, leave', *kirten > kerten* 'handle', *bi(h)urtu > biortu* 'crooked', *agiri > ageri* 'manifest, evident'.

Finally, there are some variations between /a/ and /e/ or /e/ and /i/ before a tautosyllabic /l/, but there seems to be no such variation when they are located in the following syllable onset, as in common *beltz* > western *baltz* 'black', common *beldur* : western *bildur* 'fear' or eastern *elki* : *ilki* 'go/take out'.

4.3.4.4. Vowel addition

Prothesis is regular in loanwords beginning with a trill or a cluster with an initial sibilant. The prothetic vowel is usually an /e-/, but some instances of /a-/ are

---

20 Michelena notes that this occurs especially when /e/ is followed by a high vowel in the following syllable (*FHV*: 60).

found before /ra-/ or /ro-/. Any other vowel is the result of an assimilation (such as the /i-/ from /e-/ when preceding an /i/ in trisyllabic words, see §4.3.4.2). Examples of this process are: Lat. *rege* > *errege* 'king'; Lat. *ratione* > B. *errazoe*, G. *arrazoi* 'reason'; Sp. *raza* > *arraza* 'race', Sp. *ropa* > B. G. *erropa*, comm. *arropa* 'clothes', Lat. *spatha* > *ezpata* 'sword'; Lat. *spiritu(m)* (> *ezpiritu*) > *izpiritu* 'spirit, ghost'.

Other tautosyllabic clusters such as those composed of a stop and a liquid (present in Latin and Romance borrowings) have also been adapted to the phonotactic restrictions of the language by means of the addition of a vocalic segment. In this case, the epenthesized vowel is not a predetermined one, but its quality is predictable from the vowel found next to the "ill-formed" cluster: Sp. *Inglaterra* > *Ingalaterra* 'England'; Lat. *apricus* > *apirila* 'April'; Lat. *fronte* > *boronde* 'front'; Lat. *libru(m)* > *liburu* 'book'. Modern Basque is less restrictive insofar as these clusters are concerned.

In Guipuscoan, there are some instances of the addition of a palatal glide which is inserted between a vowel and a hetorosyllabic cluster formed by either a sibilant or a nasal and a following stop: *azken* > *aizken* 'last', *laster* > *laister* 'soon', *angeru* > *aingeru* 'angel', *aingira* (cf. Sp. *anguila*) 'eel'.

Much more common is the dialectal insertion of an epenthetic segment to break the hiatus between a high vowel and its following vowel. We find an instance of the development of an anaptyctic palatal glide between an /u/ and its following vowel in derived forms in the eastern dialects, as in *thüya* ['tʰyja] 'the spit' and *süya* ['s̺yja] 'the fire'. However, in Souletin and some varieties of Roncalese, the vowel sequence —[yja] and [wja], respectively— has later been simplified to -*ia* (cf. *FHV*: 121).

According to Bonaparte (1862: 33-34), the Biscayan variety of Orozko underwent an epenthetic process involving the insertion of /m/ between a stem-final /-o/ and a following vowel, which lasted until at least the 19th century. Examples of this include *ollo-a* > *olloma* 'the chicken' and *arto-a* > *artoma* 'the corn' (Hualde *GB*: 49).

A paragogic /-e/ is often developed in place names in the continental varieties as a consequence of a reanalysis of the vowel in locative suffixes, such as inessive -*(e)n*, allative -*(e)rat*, ablative -*(e)tik*, etc. as the end of the root. We find instances of this final -*e* in *Parise*, 'Paris' or *Miarritze* 'Biarritz', but also in words like *aurre* 'front' and *atze* 'back'.

In the Biscayan variety of Zeberio, a paragogic echo vowel is often added after a word-final -*k,* as reported by Etxebarria (1991). This is a copy-vowel, qualitatively similar to the preceding one, as shown in the examples: *zuk* > *zuku* 'you (erg.)', *nik* > *niki* 'I (erg.)', *gixonak* > *gixonaka* 'the man (erg.)'.

4.3.4.5. Vowel deletion

Apheresis is rare in Basque, although it can be found under certain specific circumstances. Firstly, there is a process of word-initial /e-/ deletion which occurs in different regions of Navarre, according to Hualde (*GB*: 57): *eman > man* 'give', *ekarri > karri* 'bring', *emazte > mazte* '(married) woman'. It can also result from morphological loss, principally of Latin prefixes, as stated by Michelena (*FHV*: 157, 511) and exemplified by Lat. *exemptus > sendo* 'robust'. Alternatively, it can result from a palatalization with a fusion of the /i/ with the new palatal segment, as in the modern *(t)xukatu* 'wiped away', whose stem Leiçarraga wrote as *ichuca*. However, this desegmentalization is mainly found in medial position and affects glides (*FHV*: 103), as shown by the much more usual examples: *zein > zeñ* 'who', *baino > baño* 'than', *(h)aize > axe* 'wind', *naiz > nax* '(I) am'.

Vowel deletion is particularly common in hiatus (see §4.3.5). The combination of /a/ followed by any other vocalic element often leads to the loss of the /a/: *da+en > den* '(that it) is'; *za+en > zen* '(that it) was'; Old B. *Bilbao > modern Bilbo* 'Bilbo'; *hamabortz > L. amortz* 'fifteen'; **baga-uso >* Al. *baguso* 'ring dove', **aldairi > aldiri* 'proximity'. For the sequence of a high vowel followed by any other vowel in medial position, the latter (if any) will be deleted: *amuarrain >* B. *amurrai(n)* 'trout', *luebaki > lubaki* 'trinchera', *bariaku > bariku* 'Friday', *azienda > azinda* 'livestock'. When two mid vowels are combined, /o/ remains while /e/ is usually dropped, as in *moeta > mota* 'kind' and *adiskideok >* B. *adiskidok* '(these) friends'. The deletion of the second vowel in a /VjV/ sequence is very common: *baieta > baita* 'too', **bai-ez-ik < baizik* 'but rather', *jaiotegun > jaitegun* 'day of birth'.

In some dialects, an intense process of syncope is found. It is particularly widespread in the eastern regions of Salazar and Roncal —but not in Souletin—, although it also occurs in High Navarrese, Low Navarrese and Aescoan (*FHV*: 160). This syncope develops in the following contexts, as described by Michelena (ibid.):

Contexts that yield syncope:
a. / T_ɾ, l (between a stop and a tap or a lateral, mainly in R and Sal but present in other dialects): comm. *aingeru > aingru* 'angel'; comm. *denbora >* R. *tenpra*, Sal. *tenbra* 'time'; S. *nabela >* Sal. *nabla*, R. *ñabla* 'knife'; LN. *gira >* R. Sal. *gra* 'we are', comm. *andere >* HN. LN. R. Sal. *andre* 'lady'
b. / S _ɾ (between a sibilant and a tap, only in R and Sal): LN. *zira >* R. Sal. *zra* 'you are', LN. *ziren >* R. Sal. *zren* 'they were'.
c. / ɾ, r_O (between a rhotic and an obstruent): L. LN. *eriden >* R. Sal. *erden* 'find'; comm. *eraman >* Sal. *erman* 'carry'; comm. *baratze >* R. *bartze*

'vegetable garden'; comm. *-arena* > R. Sal. *-arna* 'the one from'; comm. *-garrena* > R. Sal. *-garna* 'the Xth'.

d. / O_O (between two obstruents): Aesc. *zitu* > R. Sal. *ztu* 'has (...) you' (cf. also the plural R. *ztei*, Sal. *ztie*); Aesc. *tuzu, tuzie* > Sal. *tzu, tzie* 'you have (sg./pl.)'; comm. *ganibeta* (> *gabineta*) > Sal. *gaminta* 'knife'.

These contexts include virtually any outcome which the phonotactics of the dialect could handle: e.g., no Basque dialect favors nasal/tap clusters and thus these are absent, while tautosyllabic clusters involving a sibilant only appear in the varieties from Roncal and Salazar.

There is a final syncope process which results in the loss of an /e/ in the second syllable and after a trill. This is a general development which is not so dialectally biased as the one described above. Examples of this syncope are mainly found in loanwords that have already developed a prothetic /e-/, such as *erreloju* > *erloju* 'clock', *errelijio* > *erlijio* 'religion', *errelazio* > *erlazio* 'relation', Sp. *reyerta* > *errie(r)ta* 'brawl', etc.

Apocope is a common device in composition, in which it is systematically applied to the first member of any trisyllabic word ending in a vowel. Examples of this process are Ax. *burdin gori* 'red-hot iron' (cf. *burdina* 'iron') and *(h)euskaldun* 'Basque speaker' (cf. *euskara* 'Basque'). For disyllabic words, /-i/ is regularly lost while /-u/ is only dropped in eastern dialects, and non-high final vowels are systematically neutralized to /-a/: *hardi* 'stony area', *zaldun* 'knight', *burhezur* 'skull', *urdaki* 'pork fat' or *gogatsu* 'enthusiastic' from, respectively, *harri* 'stone', *zaldi* 'horse', *buru* 'head', *urde* 'pig' and *gogo* 'will, mind' (*FHV*: 125-127).

Besides the ones involved in composition, there are some words which can undergo apocope. In Guipuscoa, some words lose their final /-a/ (which corresponded to the root) after another vowel, as a consequence of its reinterpretation as the definite article: *burnia* > *burni* 'iron', *katea* > *kate* (cf. Lat. *catēna*) 'chain', *koro(i)a* > *koro(i)* (cf. Lat. *corōna*) 'crown'. Some other words such as *barda* > *bart* 'last night' or *giltza* > *giltz* 'key' also occur in different varieties, as well as the widespread verbal forms such as *naiza* > *naiz* '(I) am'.

Finally, we find a (very modern) process of apocope in the Biscayan varieties surrounding Markina and Getxo. In this process, the second vowel (the *-a* or *-e* from the definite article) of an inflected nominal stem is regularly dropped by young speakers, even in final open syllables.[21] Examples of this include: *neskie*,

---

21 Hualde (*GB*: 50) reports that this second vowel loss is much more geographically widespread in closed syllables than it is in open syllables.

neskea[22] > *neski*, *neske* 'the girl'; *alabie*, *alabea* > *alabi*, *alabe* 'the daughter'; *kafie*, *kafea* > *kafi*, *kafe* 'the coffee'; *astue*, *astoa* > *astu*, *asto* 'the donkey'.

4.3.4.6. Vocalic reciprocal metathesis

Vocalic metatheses are not rare in the history of the language, but the process is conditioned by a common phonological feature shared by both of the reordered segments, so that only certain possibilities of transposition are attested. The relevant feature which constrains two-vowel non-local metatheses is [+/- high], which divides the 5-vowel inventory into two groups of segments. A given segment can change position with another one only if both are members of the same group, i.e. reciprocal metatheses only occur among segments that agree in height (Egurtzegi 2011). Examples of metathesis of non-high segments include: *atera* : *etara* 'come out, take out', *alkandora* : *alkondara* 'shirt', and *(h)odei* : *(h)edoi* 'cloud', while *ikutu* : *ukitu* 'touch' and *ingude* : *ungide* (< Lat. *incūde*) 'anvil' exemplify metatheses involving a pair of high vowels.

4.3.5. Diphthongs

The diphthongs which Michelena proposes for Proto-Basque are /au̯/, /eu̯/, /ai̯/, /ei̯/ and /oi̯/ (*FHV*: 87). Regardless of their common status in modern Basque, diphthongs seem to be secondary (probably in their entirety), and a significant rise in their frequency can be traced back. Diphthongs usually originate from the loss of an approximant, aspiration[23] or tap,[24] etymologically located between the two vowels. Given that most of these processes probably occurred some centuries after the Michelenian Proto-Basque, it is very likely that only some of the aforementioned five diphthongs were known at this stage, if any.

Diphthongs can also be nasal in the varieties in which the intervening source of the nasality is lost and a vocalic nasality contrast is still productive, as in Roncalese or in 16th and 17th century Biscayan. Although falling diphthongs (/ai̯/, /au̯/, etc.) are more widespread than their rising counterparts (/ja/, /wa/, etc.), some of the latter may be older than the former, at least those involving a word-initial yod.

One of the reasons behind the older status of rising diphthongs is that their word-initial position, in an Old Proto-Basque mainly consisting of roots and

---

22 The variety from Getxo does not rise the prevocalic mid vowels as in Markina (cf. §4.3.4.2).

23 Aspiration which was a former nasal, cf. §4.4.3.1.2.

24 In more modern stages of the language than the others.

their prefixes, was peripheral to the root as well as to the stress. Thus, word initial segments were the most susceptible to suffering different kinds of deletion, including those deletions yielding consecutive vowels that would later become diphthongs. These kinds of yod-based vocalic clusters are usually found in verbal forms and they originated from the combination of an initial /e/ and a following vowel, which were most probably separated by a coronal approximant in older stages of the language. Examples of this process include: *e-da-kin > jakin* 'know' and *eoan > joan* 'go' (according to Michelena *FHV*: 119).

4.3.5.1. Falling diphthongs

Falling diphthongs are formed by the encounter of a non-high and a following high vowel, either in composition or derivation, and after the occurrence of a phonological process affecting an intervening consonant. /a/ may combine with either of the glides, although these are frequently reduced to a simple high vowel in composition (see *FHV*: 116), as in *basa-urde > basurde* 'wild boar' and *garaunak > garunak* 'brains'. Examples of this combination include those created after the loss of the aspiration, as in *ahuntz > auntz* [au̯nts̻] 'goat', and *nahi > nai* [nai̯] 'will', which was still disyllabic in RS.

The mid vowels are normally combined so that diphthongs are formed by anterior/posterior or posterior/anterior pairs (cf. ibid.), as in *ehun > eun* [eu̯n] 'hundred' and *ohi > oi* [oi̯] 'habit'. The combination of /a/ and a mid vowel may also result in diphthongization (*apaez > apaiz* 'priest', *hama-eka > hamaika* 'eleven'), but the most usual outcome would involve a contraction favoring the latter element.

The general tendency noted by Michelena (*FHV*: 486) implies that word-initial falling diphthongs usually lose their second element. He also emphasizes that any diphthong ending in a palatal will lose its second element before a rhotic, as in *sorburu* 'back' from *soin* 'shoulder' and Lç. *ohorgoa* 'larceny, theft' from *ohoin* 'thief' (cf. *jaurgoa* 'empire' from *jaun* 'Sir').

Along the same lines, a palatal off-glide is often lost after the palatalization of a following consonant. This is the consequence of a reinterpretation of the source of this feature as originating in the consonant instead of the etymologically palatal vowel. Examples of this are mainly found in the Biscayan dialect —as well as in peninsular areas— but can be found in any palatalizing variety: *naiz > nax* '(I) am', *noiz > nox* 'when', *baina > baña* 'but', etc.

There are some geographically widespread cases of the dissimilation of an off-glide, when followed by its equivalent vowel in the following syllable. /w/ can dissimilate to /j/, as in LN. S. *haizu* < Lat. *ausus (sum)* 'to be lawful, to dare'; *errekaitu* 'provision, food' < Sp. *recaudo*, but they are usually dropped:

auntzume > B. G. *antzume* 'kid (goat)'; *itaundu* > B. *itandu* 'to ask'; *jaiki, jeiki* > *jaki, jeki* 'stand up'; *-gaitik* > *-gatik* 'because of, due to'.

Syncope has spontaneously affected /au̯/, simplifying the diphthong after the loss of the off-glide. The frequency of this process increases with the length of the word, and is particularly usual in initial position and among verbal forms: *aulki* > *alki* 'chair', *laurogei* > *larogei* 'eighty', as well as in Lat. *audītum* > *aditu* 'listened, understood', provided that the etymology is correct. Other simplifications are rare, at least in initial position (cf. *FHV*: 491).

/ai̯/ is relatively stable, although the low vowel can sometimes be assimilated to /e/, even resulting in its loss (cf. infra for /ei̯/ > /i/): *gai* > *gei* 'matter', *jaiki* > *jeiki* 'wake up', *egitai* > *igitei* > *igiti* 'sickle'. In Souletin and Roncalese, an older **au* has regularly become /ai̯/, except when followed by /ɾ/, /r/, /s̺/ or /ts̺/: *gau* > *gai* 'night', *gauza* > *gaiza* 'thing', *laudatu* > *laidatü* 'to laud'.

/eu̯/ is much more unstable than /au̯/, and different processes have affected it so that nowadays it is rare. It has been simplified by the loss of either of its elements —usually /e/, as in *heuragi* > *ugari* 'abundant' and G. B. *euli* > HN. LN. L. *uli* 'fly' (but cf. also the diminutive *eltxo* 'mosquito')— or dismantled by means of excrescence, as in *leun* > *legun* 'soft' and *neurri* > *negurri* 'measure'.

The stability of /ei̯/ is largely dependent on its position within the word. In initial position, apheresis has regularly occurred, with the initial /ei̯-/ being practically absent from modern Basque: **eizan* > *izan* 'to be', *eihar* > *ihar* 'dry', although in the latter the semivowel may also be epenthetic (cf. Lakarra 2009c).[25]

/oi̯/ is usually maintained, although it sometimes alternates with the rarer /wi/ (< /ui̯/), as in the collective suffix *-doi, -dui, -di* from toponymy. This diphthong sometimes comes from /ei̯/, as in *hogei* > LN. L. *hogoi* 'twenty' or Aesc. *oreitu* > *oroitu* 'remember'.

4.3.5.2. Rising diphthongs

Rising diphthongs develop in almost the same circumstances as their falling counterparts, but these are particularly habitual in the nominal forms of certain verbs. In these cases, as mentioned above, an initial **eðV-* becomes *jV-* after the loss of the intervening consonant. Such forms are usually followed by /a/ due to the high frequency of an older **e-da-*,[26] as in *jakin* 'know', *jausi* 'to fall', *jaun* 'sir', etc. It is also frequent before /o/ (*jo* 'hit', *josi* 'sew', *joan* 'go') and much rarer before /i/ (*jin* 'come'). The examples of /je-/ are much more modern, and

---

25   On the other hand, in medial and final position, /ei̯/ is more stable, although it can also be dropped there.

26   On this second morpheme, see Lakarra (2006a, 2006c).

they continue —perhaps in their entirety— from an older *ea- > *ja-, given that an etymological *e-e- may have yielded e-. Examples of this include: *jasan* > *jesan* 'borrow', *jaitsi* > *jeitsi* 'go down' or *jasarri* > *jesarri* 'sit down'. This initial yod is absent before /u/, maybe as a result of the formation of the falling diphthong from *eu- sequences, instead of the raising diphthongs seen in the former vowel combinations (i.e., *eu- > eu- and not > **ju-). We can reconstruct an example of this: *e-dun-tz-i* > *eutsi* 'hold'.[27]

Rising diphthongs involving /w/ have usually been simplified, yielding different results: /wa/ has been replaced by /o/ (especially after a velar stop), as in *kodaña* (cf. Sp. *guadaña*) 'scythe' and *gorde* (cf. Sp. *guardar*) 'save', and /we/ has been syncopated to /e/, as in Sp. *ruego* > *erregu* 'request' and perhaps in *lueco* > *leku* 'place' (*FHV*: 167-168).

4.4. Consonantism

4.4.1. Sonorants

4.4.1.1. Rhotics

Most modern Basque dialects oppose an apical tap to a trill, also apical. Nevertheless, the opposition among rhotics is only appreciable between vowels, given that it is absent from initial position and has been neutralized to the trill in final (in every dialect except Roncalese) and in pre-consonantal position.

In the native lexicon, rhotics do not occur in initial position, and a prothetic vowel is systematically added in loanword adaptations. This vowel may be either /e/ or /a/ (see §4.3.4.4 on vowel addition).

A straightforward consequence of this may be the removal of the onset rhotic from the oldest stages of the language that we are currently able to reconstruct, i.e. we can hypothesize that a monosyllabic Basque root had only word-final —maybe non-contrasting— rhotics. Later, alveolar taps would result from the intervocalic /l/ in a regular manner (cf. §4.4.1.2), while coda rhotics would slowly evolve towards the trill, creating an intervocalic opposition after the loss of the consonant which followed the coda trill, and its subsequent resyllabification. Pre-consonantal rhotics have systematically yielded a trill until very recent times, and it is possible that this was the case in Proto-Basque. Modern examples of these are found after a resyllabification caused by the loss of a following onset, as in the dissimilations or metatheses in the Biscayan names *Fernando* > *Ferrando* and *Gernika* > *Gerrinka* or *Zornotza* > *Zorrontza* (cf. Egurtzegi 2011),

---

27   Cf. Lakarra (2006a) on the -(n)tzi > -tsi development proposed by himself.

and the loss of the aspiration on the Labourdin coast: *nerhabe > nerrabe* 'single man', *orhoitu > orroitu* 'remember' and *erhaztun > erraztun* 'ring' (cf. the un-aspirated common forms *nerabe*, *oroitu* and *eraztun*).

When they represent the second element of an onset cluster, rhotics seem to be historically perceived as a tap. Given that Proto-Basque disallowed such tautosyllabic clusters, we find evidence for this assumption in the adaptation of Latin borrowings such as *libru(m) > liburu* 'book', *escribe(re) > izkiri(b)atu* 'write' or *fronte(m) > boronde* 'front' or *lucru(m) > lukuru* 'profit, gain'.

As stated above, the distinction between rhotics is not made in the coda, as most Basque varieties only produce a trill in this position. However, as may be inferred from historical evidence, this opposition was in fact effective before an aspiration,[28] although this can no longer be discerned since the aspirating dialects have lost their rhotic opposition due to independent reasons. However, it is most probable that this contrast was lost even before the rhotic neutralization.

As a consequence of its intervocalic tap loss, modern Souletin has only a single rhotic (*FHV*: 328-330). However, this is only the case for recent times, since the tap has been lost in the course of the last two centuries and has played a role in many phonological changes. The remaining rhotic is produced as an alveolar trill, although it has fewer vibrations than the common Basque trill (Hualde *GB*: 30-31).

Also in very recent times, in most of the varieties from Labourd and Low Navarre, the apical articulations have merged into a posterior realization. Nevertheless, Gavel reported that, in sporadic cases, some speakers have conserved an apical tap along with the uvular trill (1920: 239). The trill first became uvular, and then the tap began to be produced likewise, giving rise to their merger in most cases. The rest of the dialects[29] maintain both rhotic segments intervocalically, although /ɾ/ tends to drop sporadically in many varieties (cf. *FHV*: 330).

The process of neutralization in final position seems to be complete nowadays, after the disappearance of Roncalese, the only dialect which discerned between the two apical rhotics word-finally (cf. Azkue 1931: 218-219). However, there is still a small group of words which possess an underlying word-final tap that surfaces in derivation: *nor* 'who', *zer* 'what' (alongside *inor* 'nobody' and *ezer* 'nothing'), *(h)ur* 'water', *zur* 'wood' and *hor* 'dog'.[30] The non-fronted high rounded vowels in Souletin *hau* 'this', *hiru* 'three' and *lau* 'four' denote that the

---

28 Hualde only mentions the example of Souletin as possessing a contrast before the aspiration, but Leiçarraga also shows a systematic differentiation between rhotics in this position. He writes *urrhets* 'step', *bur-hezur* 'skull', etc.

29 Those located in Spanish territory.

30 A couple of recent loanwords increase the number of words in this group: *paper* 'paper' and *plater* 'dish'.

final *-r reconstructed for those forms (cf. the older *haur*, *hirur*, *laur*) was in fact a tap.[31] Some of these words, especially *(h)ura* 'the water', also have dialectal variants involving palatal realizations which are very different to both the etymological tap and the otherwise extended word-final trill. Such examples include the voiced affricate and approximant in [ud̥ᶻa] and [uʝa] from the Guipuscoan varieties of Etxarri-Aranatz and Goierri, respectively (cf. *FHV*: 556).

There are some instances of compounds in which an etymological final tap is found as an aspiration in the first member of the compounded form. This process seems to develop in old monosyllabic words, such as *hur* 'water', *zur* 'wood' and *hor* 'dog': *uhalde* 'shore, riverside', *uharte* 'island', *ühatx* 'torrent', *zuhaitz*, *zühaintze*, *zühañ* 'tree', *zuhandor* 'shrub, bush', *zübihotz* 'center of the oak', *ohara* 'dog in heat', *ohalano* 'mastiff', etc. This process seems to be common, and thus prior to dialectal developments such as the vowel fronting found in the Souletin variants above, which would not have occurred before a tap.

4.4.1.2. Laterals

Common to all modern Basque dialects are the apico-alveolar and palatal[32] laterals, but another one —usually transcribed as *L*, following the Michelenian tradition— is postulated for older stages of the language. Although the most widespread realization of /l/ is apical, Michelena (*FHV*: 311) reports the presence of velar productions in coda position in some Souletin territories, as well as isolated retroflex productions (*FHV*: 550).

During the Middle Ages —before the 11th century, since this evolution was already completed in the first written documents— the alveolar lateral yielded /ɾ/ intervocalically, much in the same way that /n/ had previously changed to /h/ (cf. §4.4.1.3.1 and §4.4.3.1.2). Although some exceptions can be found, this process occurred in a regular way: Lat. *angelu(s)* > *aingeru* 'angel', Lat. *gula* > *gura* 'hunger, desire', Lat. *voluntate(m)* > *borondate* 'will'. Most exceptions either correspond to recent borrowings or show dialectal variation (e.g. B. *solo* : comm. *soro* 'field, meadow' (cf. Lat. *solum*, Sp. *suelo*) and comm. *-(a)ra* : S. Sal. *-ala* 'to', although the latter could be derived from two suffixes: *-an-la > *-alla > -ala*).

This process created an alternation between a name and its composed or derived forms, after the loss of the final vowel[33] had bled the phonological envi-

---

31   See §4.3.3 on the Souletin /y/.
32   On the development and hypocoristic usage of /ʎ/, see §4.4.4.1.
33   Which is systematic for trisyllabic words in composition, cf. §4.3.4.5.

ronment when followed by an initial consonant in the second member of the compound: *euskara* 'Basque' > *euskaldun* 'Basque speaker', *erdara* 'foreign language' > *erdaldun* 'foreign speaker', etc.

The fortis **L*, a segment reconstructed by Michelena for Proto-Basque, parallels the rest of the sonorant fortes *R* and *N* (cf. §4.4.1.1 and §4.4.1.3) respectively. The primary reasons for the reconstruction of this segment stem from the high frequency of the modern intervocalic /l/. Recall that **l* systematically evolved to /ɾ/ in this position, as well as the observation that Latin -*Vl.lV*- had regularly yielded Basque -*V.lV*- (cf. *ballaena* > *balea* 'whale', *cella* > *gela* 'room' and *castellum* > *gaztelu* 'castle'). Thus, Michelena proposes a lateral segment that would ultimately become the modern intervocalic /l/ for all the reconstructions which share the outcome of this development, present in Latin loanwords.

Although it remains uncertain as to whether this segment had a different realization[34] or was solely a longer production of the alveolar /l/ (i.e. a geminate), what seems to be clear is that any /l/ which was not found in an intervocalic position at the time of the **l > r* change would remain the same after a late alteration of its etymological phonological context. Consequently, any lateral historically located before an aspiration appears as an /l/ in all modern dialects[35] (cf. *alhaba* > *alaba* 'daughter') as does any lateral which etymologically belonged to a cluster which may be absent today (cf. **erlana* > **erlaha* > *elae* > *elai* 'swallow (bird)').[36]

This segment appears in the *Reja de San Millán* (*CSM*, 1025) as both -*lh*- and -*ll*-, and, according to Michelena (*FHV*: 321), this seems to be related to the restrictions applied to the aspiration within the boundaries of the word, i.e. /Vl.hV/ occurs between the initial and second syllables and /Vl.lV/ elsewhere. It is also emphasized by Michelena that in order to justify the high number of modern word-final laterals, the lateral segments may have been limited to **L* word-finally, since they never become taps before the definite article (e.g. *Kessalla* > *gezal* 'salted water' and *Zaballa* > *zabal* 'broad', cf. *FHV*: 322).

However, there may be another explanation for this phenomenon, since the definite article, which is reconstructed as **-(h)a(r)* and present as -*ha* in San Millán (cf. Manterola 2006: 674), together with the distribution of the /lh/ and /ll/ within the frame of the word mentioned above, lead us to think that heterosyllabic -*l.h*- clusters may have yielded -*l.l*- (which would later simplify to -.*l*-)

---

34 Michelena (*FHV*: 320-321) suggests that **L* was a segment —and not a cluster— in the first centuries of the Middle Ages.

35 -*lh*- in the aspirating dialects and -*l*- in the dialects that no longer possess it.

36 See Ariztimuño & Egurtzegi (2011) on this etymology.

after the loss of aspiration of the definite article. This explanation does not necessarily rely on the dropping of the aspiration after the second syllable (cf. §4.6.4), since the process of grammaticalization itself would explain the loss of the /h/. This explanation, in contrast to that of final or boundary lengthening, would not hinder reconstructions involving the *-l > r* change in suffixed verbal forms which add an *-i* to a root ending in a lateral, such as **e-dol-i > jori* 'abundant, plenty' or **e-lol-i / *e-ra-dol-i > erori* 'fall' (cf. also the related reduplicated form *odol* 'blood', apud Lakarra 2011b).

There are some word-initial laterals that came from an older /d-/ in words such as Lat. *thēca* > Basq. *leka* 'pod', **doni* > modern *lohi* 'body', etc. This process may have been systematic in non-verbal words, given that most initial /d/s are found in auxiliary and synthetic verb forms, along with recent loanwords in modern Basque (cf. §4.4.2.1).

4.4.1.3. Nasals

In modern Basque, an alveolar /n/, a bilabial /m/ and a palatal nasal /ɲ/ contrast prevocalically, but only the alveolar and the palatal segments contrast word-finally.

When in word-medial coda, nasals are systematically assimilated to the following consonant and are realized with the same place of articulation. Examples of this assimilation are: *enbor* [embor] 'trunk, log',[37] *handi* [ha̠ndi] 'big', *onddo* [oɲɟo] 'fungus', *aingura* [ai̠ŋguɾa] 'anchor', *anfibio* [aɱfibjo] 'amphibian', etc.

Along with the nasal stops, a nasalized glottal fricative /h̃/ is also present in modern Souletin (see §4.4.3.1.2).

4.4.1.3.1. /n/

Although the modern language possesses three contrasting nasal obstruents in all of its varieties, historical and, to a greater extent, reconstructive evidence suggests that only the alveolar obstruent was present in the oldest stages we can reach. However, not every alveolar nasal has remained unchanged until today.

As a matter of fact, Proto-Basque /n/ was regularly affected by a rhinoglottophilia process (cf. Matisoff 1975, Igartua 2008) which altered it to the nasalized aspiration /h̃/ in intervocalic position, probably its most usual location (cf. *FHV*: 142-153, 215, 300-303; Igartua 2011 and §4.4.3.1.2). Since an older etymological /h/ was also present between vowels, both aspirations contrasted in this position, and perhaps even prevocalically to a lesser extent (cf. §4.4.5.3 on

---

37 Which does not yield ***emor*, cf. §4.4.1.3.2.

consonantal metathesis). However, /h̃/ only remains in the Souletin dialect now-adays, since it has lost its nasalization —and thus merged with /h/— in the rest of the aspirating varieties.

After the complete loss of the aspiration in the western dialects, the nasalization was kept in the vowels etymologically surrounding the nasal segment for some centuries, so that it still appears in some Biscayan documents from the 16th and 17th centuries (cf. §4.3.2 on nasalized vowels). This nasalization was often reanalyzed as a syllable-final obstruent, which was subsequently restored after the diphthong: *artzani > *artzah̃i > artzãĩ > artzain 'shepherd', Lat. *lūcānica* > *lukah̃ika > *lukãĩka > lukainka 'spicy pork sausage', *arrani > *arrah̃i > arrãĩ > arrain 'fish', Aquit. seni- > seh̃i > Old B. sẽĩ > modern sein 'child'.

The reestablishment of the nasal stop is very common after an /i/, but a pala-tal nasal is reinstated instead of the etymological apico-alveolar: comm. *katea* : S. *kathiña* 'chain', Al. *erregia* : B. G. *erregiña* 'queen', L. *ihes* : B. *iñes*, etc.

An intervocalic /n/ usually yields /ɲ/ when following an /i/ or a yod, as de-scribed in more depth in §4.4.4.2. In many cases, the palatal off-glide merges with the following nasal, so that the etymological context which produced the change is regularly lost in some varieties and thus creates a contrast between the two nasal segments intervocalically.

Along the same lines, there are also some isolated examples of the restitu-tion of an intervocalic /m/ after an /u/: L. LN. *zuhai* : HN. *zumai* 'forage'; L. LN. *zuhar* : HN. G. B. *zumar* 'elm'. An /n/ can also yield /m/ after an /u/, as in *un(h)il* > L. *umil* 'funnel'.

4.4.1.3.2. The development of /m/

Despite its modern high frequency, the Proto-Basque consonantal system lacked a labial nasal, as was made apparent by Michelena (*FHV*: 270) and accepted by Trask (*HB*), who also pointed out its lack of morphological value. The segment was later introduced by means of borrowing from Latin and Romance languages and also developed within the language itself.

Two different assimilatory processes which led to the nasalization of a labial (oral) stop are reconstructed with different chronologies. Common to all of the dialects is the transformation of a nasal and a labial stop sequence into a labial nasal: /nb/ [mb] > /mm/ (> *m*). This development would have occurred in the first centuries of our era, since *Sembe-* (cf. modern *seme* 'son'), *Ombeco* and

Ombexonis (usually linked to *ume* 'child'), attested in Aquitanian, do not show the assimilation of the labial stop.[38]

The second process of assimilation, the similarly widespread regressive assimilation of nasality, has been productive within the modern dialects themselves. In this latter process, a non-local assimilatory nasalization of /b/ occurs when followed by a nasal obstruent (cf. §4.4.5.1), as in *bendekatu > mendekatu* (cf. Lat. *uindicāre*) 'avenge', *bagiña > magi(ñ)a* (cf. Lat. *uāgīna*) 'scabbard, sheath' or *Mendekoste* (< Sp. *Pentecostés*) 'Pentecost'. Although we can extend its productivity to some centuries onwards, the latter assimilation also occurred before the rhinoglottophilia process had apparently ended. This is shown by the chronology necessary in multiple standard reconstructions such as **bene > *mene > mehe* 'thin' or **bini > *mini > mihi* 'tongue', although the assimilation of nasality may also have originated from a nasalized aspiration.

There are nevertheless some phonological contexts which block the assimilation of nasality (cf. Egurtzegi 2011, §4.4.5.1), which would probably involve the presence of an intervening segment such as a coda /r/ between the /b/ and the following /n/. This is exemplified by the form *tramena* 'tavern', which shows an assimilation, in contrast with other variants such as *taberna*, *taferna*, etc., from which it is consistently absent. The apparent lack of ***mVrn* sequences backs this hypothesis. The biggest problem for identifying the aforementioned contexts lies in the large number of examples in which a voiced labial stop has become nasal even when it was not followed by a nasalized environment. Examples of this /b/ - /m/ alternation are: *makulu* (< Lat. *bacculus*) 'walking stick', *makar < bekar* 'sleep, rheum' and *maru* (< **balu* < Lat. *pālus*) 'pole, post'.

As a consequence of its marginal status in older stages of the language, /m/ has had an expressive value similar to that of the palatal segments (cf. §4.4.4.1), and it appears in the second component of reduplicated forms such as *zaldiko-maldiko* 'carousel', *zirimiri* 'drizzle' and *xistmist* (> *tximist*) '(lightning) bolt'.

4.4.2. Stops

All modern Basque dialects have at least two series of stops (voiced and voiceless, alongside aspirated voiceless stops in the eastern varieties) which are produced at four different points of articulation: labial, dental, palatal and velar. Palatal stops are less frequent than the others and often serve a very specific affective purpose (cf. §4.4.4.1).

---

38 Unlike other Aquitanian forms such as *Ummesahar* (cf. also modern *ume* 'child') from Lerga, but this form also shows the expected vocalism, in contrast to *Ombeco* and *Ombexonis* mentioned above.

Within the modern dialects, stops are mainly found in onset position, since they are limited to the borrowed vocabulary in preconsonantal position, and only /-t/ and /-k/ are (very infrequently) found word-finally. These word-final stops are the result of apocope followed by a final devoicing (cf. §4.4.2.2), and there is no reason to think that stops could have appeared in coda position in older times.

4.4.2.1. Voiced stop series

Basque voiced stops can be produced either as stops or as approximants. Stops are produced phrase-initially and in certain postconsonantal positions. This is systematic for any voiced stop located after a pause as well as after a nasal and, in the case of /d/, it is also realized as a stop after /l/. In those contexts in which the spirantization of the voiced stop does not occur, the place of articulation of the preceding segment is assimilated to that of the stop (cf. §4.4.1.3): *enbor* [embor] 'trunk, log', *handi* [ha̠ndi] 'big', *iñddar* [iɲɟar] 'strength', *aingura* [ai̯ŋguɾa] 'anchor', *alde* [a̠lde] 'side'. Approximants are found between two vowels, as well as between a heterosyllabic continuant consonant (/l/ —with the previously mentioned exception of [l̺d]—, /r/, /s̺/ and /s̻/) and a following vowel): *arreba* [areβ̞a] 'sister (of a boy)', *adostu* [a̠ð̞o̠s̺tu] 'agree', *lagun* [laɣ̞un] 'friend', *albo* [alβ̞a] 'side', *elgar* [elɣ̞ar] 'each other', *arbaso* [arβ̞a̠s̺o] 'ancestor', *ardo* [arð̞o] 'wine', *argi* [arɣ̞i] 'light', *ezberdin* [es̻β̞erdiɲ] 'different', *ezdeus* [es̻ð̞eu̯s̻] 'nothing', *desgogo* [des̻ɣ̞o̠ɣ̞o] 'reluctance', etc. Note that, as in Spanish, fricatives are voiced before a voiced consonant in modern compounds (cf. §4.4.3.3.2). However, stops have been regularly devoiced after sibilants in the native vocabulary as well as in old borrowings and compounds, so that contextually voiced sibilants are rare before stops.

This process of spirantization has sometimes even yielded the complete loss of the etymological voiced stop, although this has only occurred in intervocalic position and not after a continuant. Given that intervocalic approximant loss is found in onomastics dating from at least the 12th century, the spirantization of intervocalic voiced stops must have been even older. Old examples of this loss are usual in compound surnames such as: *Soracoiz* (< *sora-bakoitz*), *Sorauren* (< *sora-guren*), *Echarri*, *Echerri* (< *etxa-barri* / *-berri*), etc.

Instead of yielding deletion, some instances of intervocalic /d/ (more accurately, [ð̞]) have become taps in many Biscayan, Guipuscoan and High Navarrese areas. This process may appear to be almost systematic in some dialects, but [ð̞] has not merged with /ɾ/ anywhere (apud *FHV*: 228): *edo > ero* 'or', *edan > eran* 'drink', *odoi > oroi* 'cloud', *etorri re* (< comm. *etorri da*) 'has come', etc. As a matter of fact, this process has sometimes produced ultra-corrections

such as *idu* (< *hiru*) 'three', *aidean* (< *airean*) 'in the air' and *usadioa* (< *usari-oa*) 'the usage', *amodio* (< *amorio*) 'love'.

Although present in a certain number of loanwords, word-initial /d-/ is almost entirely absent from the native vocabulary; the only exception being represented by the auxiliary verbal system, where initial /d-/ is very common, contrasting with its scarce presence in lexical items. Alongside these common auxiliaries, some verbal forms in the subjunctive and conditional (*ledin*, *letorke*, *lioke*, etc.) have an initial *l-* which seems to correspond to the initial *d-* widely present in the indicative (*da* 'is', *dira* 'are', *dut* 'I have', etc.). As noted by Michelena (*FHV*: 257), some Latin and Romance borrowings also show an initial lateral instead of the expected voiced dental stop: *leka* (< **deka* < Lat. *thēca*) 'pod', *lanjer* (< Fr. *danger*) 'danger', *libersio* (cf. Sp. *diversión*) 'diversion', etc. These correspondences lead us to conclude that the initial /d-/ regularly became an alveolar lateral at some time within the Proto-Basque period, thus predating the development of the dialects. The maintenance of the initial /d-/ in the indicative may be a trace pointing towards the suspicion that such forms were preceded by more phonological material when that particular change happened.

Also word-initially, /b-/ has been lost before a back vowel. This process is almost systematic before /o/ —with *bost* 'five' being one of the few exceptions— and it is also usual before /u/: *otu* (< Lat. *uōtum*) 'request', *okela* (< Lat. *buccella*) 'piece, bit, mod. meat', *buztarina* > *uztarina* (< Lat. *postilēna*) 'cushion to take baggage on the rump of a horse', etc.

4.4.2.2. Voiceless stop series

Although the modern language shows no difference between the number of voiced and voiceless stops, older stages of the language lack the labial voiceless stop that would have been the counterpart of /b/, as was made apparent by Michelena (*FHV*: 261).

As a matter of fact, the labial voiceless stop /p/, which nowadays is as common as any other stop in the inventory, becomes more and more scarce the further we go back in time. The general absence of /p/ in patrimonial words —since most words containing /p/ may be analyzed either as loaned vocabulary or compounds— along with its widespread phono-symbolic use, make it very difficult to accept its presence in older stages of the language.

As mentioned above, the modern language has voiceless stops not only in onsets —both in initial and medial position— but also in some codas in word-final position. Some centuries ago, not only were the voiceless stops fewer than they are now in the modern dialects, but they were also more constrained as to the positions in which they could appear.

Voiceless stops were completely absent from the word-initial position, as evidenced by the adaptation of loanwords from Latin and Romance, in which the initial stops were systematically neutralized to voiced, as shown by the following examples: Lat. *pācem > bake* 'peace', Lat. *peccātum > bekatu* 'sin', Lat. *parcere > barkatu* 'forgive', Lat. *tempora > denbora* 'time', Sp. *torre > dorre* 'tower', Sp. *castaña > gaztaña* 'chestnut', Lat. *catēna > gatea* 'chain', etc. The instances of initial voiceless stops are either modern borrowings or result from a later assimilation, as in the evolution of the word *katea < gatea < *gatena* < Lat. *catēna* 'chain' (see §4.4.5.1).

It is also worth mentioning that no intervocalic voicing comparable to that of peninsular Romance languages developed within Basque, so that all intervocalic voiceless stops borrowed from Latin remained voiceless. However, another neutralization can be observed in the adaptation of Latin and Romance word-medial stop clusters and geminates, which were systematically borrowed as voiceless stops. This occurred not only with the voiceless clusters, but also with the etymologically voiced clusters: Lat. *abbas > apaiz* 'priest', Old Sp. *cobdiçia > gutizia* 'desire, whim', Lat. *sabbatum > zapatu* 'Saturday', Lat. *ad uallem > apal* 'downward, humble', Lat. *captīvum > gatibu* 'captive, prisoner', etc. A possible approximant production of the intervocalic voiced stops from early times (see §4.4.2.1 above) may have conditioned the perception of foreign medial stop clusters as stops rather than approximants, which may have been the only other available option. The heterorganic stop clusters (both voiced and voiceless) have been regularly adapted as taking the place of articulation of the second segment. This is probably a consequence of stop place cues being better perceived in prevocalic than in postvocalic position (Ohala 1993).

Coda stops, which are even today restricted to a reduced number of morphemes, are the result of an apocope and the devoicing of the final voiced stop, and have been introduced into the language in recent times (cf. Artiagoitia 1993). Examples of these are the development of the morpheme *-k* (< *-ga*, still present in toponymy), currently used as ergative marker (cf. Lakarra 2006b) as well as for pluralizing in nominal morphology, and the morpheme *-t* used for the first person in verbal morphology (< *-da*, cf. *dut* 'I have' versus *dudan* 'that I have'). Michelena mentions that Oihenart wrote *dud* for the modern *dut* 'I have' in *Notitia utriusque Vasconiae, tum ibericae, tum aquitaniae*, in 1638 (*FHV*: 236).

The voicing opposition was also partially neutralized after /l/ or nasals, but it favored the voiced series in this context. However, both of the easternmost dialects, Roncalese and Souletin, have kept the distinction: Lat. *tempora > denbora* (but R. *tenpra*) 'time', comm. *igande* (but R. S. *igante*) 'Sunday', comm. *sendo* (but R. S. *sent(h)o*) 'strong', etc. Another neutralization occurred after fricative

sibilants, where voiced stops became voiceless, as in *ez du* > *eztu* 'has not', *ez da* > *ezta* 'is not', *deskana* (< Sp. *desgana*) 'faint', etc. This process was probably systematic at some point and is traceable to at least the Middle Ages, as shown by place names such as *Erdozpea* (cf. *behea*) from 1143 and *Çuastuy* (cf. *-dui*) from 1284.

4.4.2.3. Aspirated voiceless stop series

Alongside the voiced and voiceless stop series, modern eastern dialects also possess a series of aspirated voiceless stops composed of $/p^h/$, $/t^h/$, $/k^h/$ and the not so usual palatal $/c^h/$.[39] This series does not seem to have an autonomous status in older stages of the language, since the aspirated stops mainly behave as allophones of the voiceless series, whose presence was probably conditioned by prosodic prominence. An evidence of their secondary status is the development of such segments in the place of etymological voiceless stops in Latin and Romance loanwords, such as S. *phena* (cf. Lat. *poena*, Sp. *pena*) 'sorrow', S. *merkhatü* (cf. Lat. *mercātus*, Sp. *mercado*) 'market' and S. *apho* (cf. Sp. *sapo*) 'toad'.

As a consequence of their prosodically conditioned allophonic status, their possible contexts are mostly limited to the onset of the first two syllables (cf. Michelena 1988 [1951a]), following the restrictions also met by aspiration (see §4.4.3.1.1 and §4.6.4). Further constraints applied to an aspirated stop are that it cannot appear in words containing an aspiration, and that it may only appear once per word, albeit this latter restriction is similar to the one applied to aspiration. When a word —especially a loanword— contains two different voiceless stops subject to becoming aspirated (i.e. in the onsets of the first two syllables), it is regularly the first syllable which is produced as the aspirated one, if any.[40] Examples of this restriction include *khorpitz* (< Lat. *corpus*) 'body', *phiper* 'pepper', *khatiña* (< Lat. *catēna*) 'chain' and examples like *tenta* 'tempt' with no aspiration, but structures like $**TVT^hV$ and $**T^hVT^hV$ —as well as $**hVT^hV$ and $**T^hVhV$— are unattested.

Nevertheless, in the Souletin dialect, there are some words in which a stop may be produced with aspiration even after the second syllable, if it still precedes the stress —usually located in the penultimate syllable, but not necessarily

---

39   See §4.4.4.1 on palatal segments and their special status within the language.

40   This contrasts with aspiration, since /h/ is regularly kept on the second syllable instead of the first, which is usually dissimilated in words etymologically containing two aspirations (cf. §4.4.3.1.1).

(cf. §4.6.3)— as in *boronthate* 'will', *ba(r)anthalla* 'february', *arrathũ* [ara'tʰũ] 'mouse', etc.

4.4.3. Fricatives

4.4.3.1. Different kinds of aspiration

4.4.3.1.1. Voiceless glottal fricative /h/

Although nowadays it is only present in eastern dialects, some centuries ago aspiration was common to all Basque varieties. Michelena (*FHV*: 203) describes it as being voiceless in word-initial position but usually voiced between two vowels, and nasal between two nasalized vowels.

According to Michelena, its loss could have begun in Navarre in the 11th century, but it was still present in Alava and La Rioja until at least the 13th century (*FHV*: 205). It is already absent in the Biscayan texts from the 16th century. Its presence was diminishing in modern Labourdin and was already lost in its coastal variety,[41] at least from the 19th century, according to Bonaparte's (1991 [1869]) descriptions. Its frequency is higher in Low Navarrese and at its highest in Souletin.

The aspiration seems to be limited to the (prevocalic) onset position (but see §4.4.1.1 about a hypothesized coda /h/), and it is absent from any kind of tautosyllabic cluster. It can appear in the following contexts: in initial position, between two vowels (or a diphthong and a vowel) and between a heterosyllabic sonorant (/n/, /l/, /ɾ/, /r/, /ɲ/ or /ʎ/, but never /m/ or a sibilant) and a vowel: *harri* 'stone', *haur* 'child', *ahizpa* 'sister (of a woman)', *oihan* 'forest', *senhar* 'groom', *uñhürri* 'ant', *elhe* 'word', *illhargi* 'moon', *erhi* 'finger', *urrhets* 'step', etc. Although the rhotics were neutralized in coda position, their contrast was preserved before an aspiration (apud *FHV*: 203). This was the case in Souletin until very recent times, when the opposition between rhotics was systematically lost in all environments after the loss of the alveolar tap (cf. §4.4.1.1).

A further limitation as to the location of the aspiration implies that it is restricted to the first two syllables of the word. As Michelena (1988 [1951b], 1988 [1957/58]) proposed, this may be due to a demarcative stress pattern, according to which every word was stressed in the second syllable from the left, hence the aspiration was limited to pre-stressed position (cf. §4.6.4). This constraint may have biased the metathesis of several secondary aspirations that fell from the

---

41 Although it was systematically used in written documents from the 17th century (see *FHV*: 204).

third syllable onwards (cf. §4.4.5.3), such as Lat. *arēna* > **areha* > modern *harea* 'arena' and Lat. *leōnem* > **leohe* > modern *lehoi* 'lion'.

There is yet another restriction which affects the aspiration in the modern dialects, although it was not present in medieval times: no Basque word contains two aspirations or two aspirated segments (since aspirated stops are also constrained by this restriction, cf. §4.4.2.3). Words etymologically containing two aspirations have been simplified following a development similar to Grassmann's Law of Greek and Indo-Iranian, i.e. an anticipative dissimilatory loss has affected every aspiration followed by another within the boundaries of the word (see §4.4.5.2).

A reduced number of word-initial aspirations were also introduced along with loanwords which had /f-/ in their Latin forms, such as *horma* (cf. Lat. *forma*) 'wall', *haro* (cf. Lat. *pharum*, Sp. *faro*) 'lighthouse', *haxe* (cf. Lat. *fascem*) 'beam, bundle' and *holla* (cf. Lat. *foliam*) 'leaf, sheet' (see §4.4.3.2). However, this development may have occurred before the words were borrowed, and thus within the Romance languages themselves.

There are also some instances of /h/ that appear in composition, which are derived from a former final tap in the first member of the compound in words, such as *uharte* 'island', *uhalde* 'shore, riverside' and *zuhandor* 'shrub, bush' (see §4.4.1.1 on these).

4.4.3.1.2. Nasalized /h̃/

Alongside the etymological /h/, we also find a secondary aspiration which derived intervocalically from an older alveolar nasal stop /n/ (cf. *FHV*: 300-303, Igartua 2011 and §4.4.1.3.1). This aspiration /h̃/ was historically nasalized and probably also voiced (hence it could be transcribed as /ɦ̃/), and this is still the case for the Souletin dialect. It has seemingly merged with the other glottal fricative in the remaining aspirating dialects, although a description of its voicing is yet to be produced.

As a consequence of its origin, the modern context of the /h̃/ is limited to the intervocalic position, where it contrasts with the non-nasalized aspiration /h/ in the Souletin dialect. It was most probably common to all dialects during a medieval stage of the language, since the loss of the intervocalic /n/ is regular in all of the varieties which no longer preserve the aspiration today.

Examples of this development are especially explicit in Latin borrowings such as Lat. *honōre* > L. *ohore*, S. *uh̃ure* 'honor', Lat. *anatem* > S. *ah̃ate* 'duck' and S. *ah̃abe* (cf. Sp. *anavia*) 'blueberry'. This aspiration was also lost (§4.4.3.1.1) or metathesized when located after the second syllable, as shown by

Lat. *corōna* > *k(h)oroa* 'crown', Lat. *capitāne(m)* > Old B. *kapitãẽ* > modern *kapitain* 'captain', Lat. *arēna* > *(h)area*, R. *ãria* 'sand'.

4.4.3.2. The emergence of the labio-dental fricative /f/

Although /f/ is mainly labio-dental in the modern Basque dialects, Hualde reports some bilabial realizations (*GB*: 15). The labial fricative is proposed to be absent in the proto-language (*FHV*: 262), and the fact that /f/ is mostly found in borrowings provides evidence to suggest that this segment is of recent introduction.

Alongside its limited presence in native vocabulary, /f/ seems to vary to a great extent with other labial segments, generally /p/ or /b/, which could later become a nasal /m/. For the sake of argument, in Guipuscoan, /p/ is the usual outcome of /f/ in borrowings such as *paltsu* (< Lat. *falsum*) 'false', *peria* (< Lat. *fēria*) 'fair' or *inpernu* (< Lat. *infernum*) 'hell'. The development also works in reverse, with etymological voiceless labial stops becoming voiceless labial fricatives in examples such as *frogatu* (< Lat. *probātum*) 'proved, to prove', S. *foltsü* (cf. Lat. *pulsum*) 'pulse' or common *prakak* > B. *frakak* 'trousers'.

The Romance bilabial voiced fricative <v> /β/[42] is usually adapted to /b/, which is also produced as [β] between vowels. Nevertheless, some words have the voiceless fricative /f/ instead of the expected voiced stop in initial position: LN. *ferde* (cf. Sp. *verde*) 'green', HN. *fau(n)* (cf. Sp. *vano*) 'vain', etc.

As well as labials, the Latin /f/ can also result in a Basque aspiration, albeit this may have been after Romance (Gascon or Spanish) mediation. In such borrowings, some dialects show a labial while others display an aspiration (in eastern dialects) or zero (in western dialects).[43] Examples of this alternation are: L. *haro* (cf. Sp. *faro*) 'lighthouse', *haxe* (cf. Lat. *fascem*) 'beam, bundle', *holla* (cf. Lat. *foliam*) 'leaf, sheet' in the eastern dialects; and in the western dialects, *piku(o)*, *biko*, *fiku(o)* : B. *iko*, HN. *igu* (cf. Lat. *fīcum*) 'fig' and *biru*, *piru*, *firu* : B. *iru* (cf. Lat. *fīlum*) 'thread'.

Hualde has proposed that the word-medial /-f-/ arose within the native vocabulary from the combination of a labial —which could be a wau or an approximant [β]— and an aspiration (Hualde 1997a: 422-423). More specifically, Hualde suggests that *auhari* 'dinner' and *auher* 'lazy, idle' are the oldest variants in their dialectal cognate groups —cf. *afari*, *abari*, *apari*, *aufari*, *auhari*, *aihari*, *aigari*, etc., for the former; and *alfer*, *alper*, *aufer*, *aulfer*, *aurer*, *aulger*, *auher*, etc. for the latter— and that their medial clusters (i.e. -*βh*- or -*wh*-) would

---

42   See Penny (2006: 118-119).

43   Remember that /h/ was lost in western varieties (cf. §4.4.3.1.1).

yield /f/ in some dialects (cf. *afari*) and /b/ (as in *abari*) wherever the aspiration was already lost (cf. §4.4.3.1.1). While his hypothesis seems rather plausible, the proposal of an assimilation may be preferable to that of coalescence, since this would account for the variants which kept the /w/ (cf. *aufari* and *aulfer*). More-over, if the /l/ in *alfer* is etymological —which would appropriately relate the forms *abari* and *aurer* as results of an early loss of the aspiration— it could also permit the presence of an intervening segment such as that of **aulher > aulfer*, *alfer*.

4.4.3.3. Sibilants

4.4.3.3.1. Voiceless fricative and affricate sibilants

Until recent times, all Basque varieties have opposed a series of apico-alveolar voiceless sibilants to a lamino-alveolar series. These series consist of a fricative (/s̺/ or /s̻/) and an affricate sibilant (/t͡s̺/ or /t͡s̻/), which contrast in certain positions —e.g. intervocalically— but are (almost) absent from others. A series of the prepalatal sibilants /ʃ/ and /t͡ʃ/ is also present in all dialects, and, although they have historically been used as a device to develop affective variants (cf. §4.4.4.1), they are common in all modern dialects.

In most dialects, /t͡s̻/ and /t͡s̺/ are absent from word-initial position, where on-ly their fricative counterparts occur. They only appear in a few expressive words such as *tzar* 'evil' and *tzakur* 'big dog' and only in certain varieties (Hualde *GB*: 22). The distribution of the more modern prepalatal sibilants is the opposite, with only the initial /t͡ʃ/ present in most dialects, although both /ʃ-/ and /t͡ʃ-/ occur in Souletin and only the fricative occurs in Labourdin.

Word-medially, all of the sibilants contrast intervocalically, while only af-fricates appear after a nasal or a liquid in prevocalic position, and only fricatives appear after a stop or before a consonant: *hotz* 'cold' but *hoztu* 'cool, chill' or *beltz* 'black' but *belztu* 'blacken'. However, the stop/sibilant clusters are very modern and are usually found in very recent borrowings such as *akzio* 'deed' or *seksu* 'sex'. The absence of fricatives after liquids or nasals is a consequence of the affrication of sibilants in this context and the subsequent contextual neutrali-zation of the opposition it produced. This is appreciable in the modern Basque forms of Romance borrowings such as *pultso* (cf. Sp. *pulso*) 'pulse', *bertso* (cf. Sp. *verso*) 'verse, line' and *kontzientzia* (cf. Sp. *conciencia*) 'conscience'. This development was not thoroughly completed after rhotics, but it is systematic af-ter both nasals and laterals.

In word-final position, affricates are much more frequent than fricatives, which although still present, are significantly reduced in number. This is the re-

sult of an unfinished affrication process —similar to that which occurred after liquids and nasals— which has affected most final sibilants but which has not become systematic. In any case, the Latin final /-s/ has regularly affricated when adapted to Basque: Lat. *fortis > bortitz* 'strong', Lat. *corpus > gorputz* 'body', Lat. *opus > oputz* 'physical effort', etc.[44]

Insofar as loanword adaptation is concerned, the Latin /s/ has regularly been borrowed as the lamino-alveolar sibilant /s̺/: *sigillum > zigilu* '(postage) stamp', *sellam > zela* 'saddle', *paradisum > baradizu* 'paradise', *spatam > ezpata* 'sword'. Only more recent borrowings such as *musika* 'music' or *soinu* 'sound' have an apical corresponding to the Romance alveolar sibilant, and only in the western dialects that are in contact with Spanish, since French sibilants are still borrowed with /s̺/ (cf. Fr. *autobus > otobuz* 'bus'). In modern borrowings, the Spanish interdental fricative /θ/ is regularly adapted to the lamino-alveolar fricative, which may also be affricate in a limited number of examples: Sp. *plaza* [plaθa] > *plaza* [pla s̺a] 'square', *hacienda* [aθjen̪da] > *hazienda* [as̺jen̪da] 'Treasury, public funds', *pozo > putzu* 'well'.

In the modern Biscayan dialect, as well as in some Guipuscoan varieties, the two alveolar series of sibilants have merged so that every fricative is produced as an apical and every affricate as a laminal sibilant. This process was already complete in western Biscayan texts from the 17th century[45] onwards, and has gradually expanded westwards into Guipuscoa, first spreading through larger towns and later through the villages, so that it was already completed in Donostia-San Sebastián by the first half of the 18th century. In other dialects, sibilant fricatives have been biased towards the apical in certain contexts, although not exhaustively. Such contexts include preconsonantal and final positions (cf. *FHV*: 282).

4.4.3.3.2. *Voiced alveolar sibilants*

As in Spanish, voiced sibilants are only allophones of their voiceless counterparts in most Basque dialects, and they are produced when followed by a voiced consonant, as in *asmatu* [as̺matu] 'invent, guess correctly', *esnatu* [es̺natu] 'wake up' and *ezdeus* [es̺ðeu̯s̺] 'nothing'. However, modern Souletin and some Low Navarrese areas possess voiced alveolar sibilants (/z̺/, /z̻/ and /d͡z̺/) as phonemes, although they mainly appear in recent borrowings or as shandi effects: S. *arrazũ* [araz̺ũ] 'reason', *plazent* [plaz̺en̪t] 'pleasant', *etsenplü* [ed͡z̺emply] 'example' and *deüs ere > deüse* [deyz̺e] 'something, nothing (when negated)'.

---

44   There are, however, rare exceptions, as shown by Lat. *magis > maiz* 'often'.
45   Michelena mentions Micoleta and *Viva Jesus* (*FHV*: 282).

4.4.4. Palatal series

4.4.4.1. Palatal segments and their usage

An old series of palatal segments which is found in all of the modern dialects, contrasts with the non-palatalized consonants. It can be traced back to older stages of the language, but its use was probably marginal and banned from the standard inventory. They were almost certainly restricted to the creation of hypocoristic or affective variants, which were formed without derivative additions; as well as being present in the borrowed lexicon, e.g. *zikoña* 'stork' from Romance (cf. Lat. *cicōnia*). However, these segments have not themselves proven to be productive insofar as historical reconstruction is concerned, and most authors have systematically worked from their non-palatalized equivalents. The series of palatal segments is usually formed by /ɲ/, /ʎ/, /ʃ/, /ʧ/, /c/ and /ɟ/, along with the segments derived from /j/, which are dialectally divergent.

Although palatals were usually used as a morphological device, they could also arise from assimilation due to their phonological context, or even become incorporated through Romance borrowings. In such cases, they would have resulted from contact with a preceding palatal segment, thus being practically absent from word-initial position. Nevertheless, note that hypocoristic palatalization can occur in any given context, which means that affective variants are the only words with an initial palatal in all of the dialects in which the yod has become a non-palatal segment.[46]

Given that the creation of palatal from non-palatal segments in words with affective functions was undoubtedly sporadic, this section will focus on the palatalizations linked to phonological reasons.

4.4.4.2. Assimilations

Assimilation to a previous palatal segment (an /i/, a palatal off-glide or even a fronted consonant in very specific contexts) is regular in many varieties of Basque. There are a handful of segments which may undergo this kind of palatalization.

The sonorants are the consonants whose palatalization is most dialectally widespread. /ɲ/ and /ʎ/ appear regularly after a palatal segment in many Basque varieties: *iñor* 'nobody', *ora(i)ñ* 'now', *ba(i)ña* 'but', *illargi* 'moon' and *ille* 'hair'. /ɾ/ and /r/ are also palatalized to /ʎ/ in eastern dialects: *bero > bello* 'hot', *hori > holli* 'yellow', *hurran > hüllan* 'near'. This process mainly takes place after a palatal glide, but is also usual after a syllabic /i/. When an off-glide is in-

---

46   See §4.3.4.3 on this.

volved, the assimilation may result in an apparent coalescence with the following segment.

Less frequent, but also usual, is the palatalization of sibilants. This does not result in a "standard" palatal but rather a palato-alveolar /ʃ/ or its affricate counterpart /tʃ/. This is particularly common in Biscayan, and occurs most frequently with the etymologically laminal sibilant:[47] *(h)aize > axe* 'wind', *ixan* 'to be', *itsu > itxu* 'blind', *bakoitz > bakotx* 'each one', etc.

In some areas,[48] /t/ is palatalized to /c/ under the same circumstances (as in *aitta* 'father', *zaittut* 'I have (...) you' or *kitto* 'enough, even'), while in some Biscayan and Guipuscoan areas it has undergone further changes to /tʃ/ (*aitxa, zaitxut*, etc.).

The palatalization of /d/ to /ɟ/ (contextually spirantized to [ʝ] as in the rest of the voiced stop series) is not as extended as the aforementioned examples, but it is still present in some dialects. Guipuscoan and High Navarrese show examples of this: *fidatu > fiyatu* [fiʝatu] 'confident, trusting', *gidatu > giyatu* [giʝatu] 'guided'. It even appears after a palatal nasal or lateral (as in *iñddar* [iɲɟar] 'strength' and *billddur* [biʎɟur] 'fear'). There are also some instances of palatalization of the voiced velar obstruent: *igo > iyo* [iʝo] 'go up'. This segment can be produced both as a voiced palatal stop /ɟ/ and as a voiced palatal fricative /ʝ/, and it can even yield an affricate /dʒ/ in some Biscayan areas.

Insofar as affective palatalizations are concerned, these stops —not only the dental, but also the labial or velars— are usually palatalized to the more habitual affricate /tʃ/ instead of the less usual palatal stops, although the latter are also dialectally present.

4.4.4.3. Outcomes of *j-

The Basque yod (cf. §4.3.5.2) has developed a wide range of dialectal variants, some being more extended than others. These changes, which involve (pre-)palatal or even velar outcomes, appear to be diverging results along the same lines of evolution. They are illustrated in Figure 4.1.

Although the first group appears to have maintained the oldest state, the third may be the most conservative dialect, at least if it reflects the segment that was common to all Basque dialects, which may have been a fricative /j/, and whose evolution may have developed a wide range of fricatives, and even affricates.

---

47   Nowadays merged with the apical in this particular dialect, cf. §4.4.3.3.1.
48   Only in Spanish territory, according to Michelena (*FHV*: 185).

Segment	Proposed development	Approximate dialectal extension (apud *FHV*: 168-171)
[j]	j (> ʝ > j)[49]	North.-HN, most L. and some B. areas
[ʒ]	j > ʝ > ʒ	S, some northwestern B. areas
[ɟ]	j > ɟ	Some B., LN. and HN. areas
[ʝ]	j > ɟ > ʝ[50]	LN
[ʤ]	j > ɟ > ʒ > ʤ	Northwestern B
[ʃ]	j > ɟ > ʒ > ʃ	Aesc, Sal, R, most South.-HN. areas
[x]	j > ɟ > ʒ > ʃ > x	G, northwestern HN, eastern B

Figure 4.1.   Dialectal evolutions of the yod

As Michelena (*FHV*: 170) points out, the parallel evolution of Romance palato-alveolar fricatives appears to be crucial for the dialectal distribution of the modern Basque fricatives, since only the varieties in contact with Spanish have developed voiceless fricatives from **j-*.

4.4.4.4. Depalatalization

Depalatalization is common in Low Navarrese, Labourdin, and is even regular in the High Navarrese of Sakana. Where other varieties show /ɲ/ or /ʎ/, these dialects have a palatal off-glide in the previous syllable. This even occurs with the etymological palatals, as in *boteila*, *botila* 'bottle' from Sp. *botella* /boteʎa/.

4.4.5. *Some consonantal processes*

There are some processes which do not affect particular segments but rather specific phonological features, or even complex groups of features. Many of these deserve a special overview, but here we will focus on the analysis of three general kinds of processes: assimilations, dissimilations and metatheses.

4.4.5.1. Assimilation

Alongside the local assimilation of the palatality mentioned in §4.4.4.2, there are some other processes of local assimilation within the Basque language, although

---

49   At least for the Biscayan area, a round-trip development is reconstructed by Hualde & Elordieta, G. & Elordieta, A. (1994: 8-9).

50   Produced as an approximant in most contexts (§4.4.2.1).

they are by no means as widespread as the one produced by front high vowels. The most important among them is the local assimilation of labiality briefly mentioned in §4.4.1.3.1, which is also perseverative. In this process, a preceding /u/ labializes the following consonant, as in *un(h)il* > L. *umil* 'funnel' or **zunai* > HN. *zumai* 'forage', etc.

There are three main processes of distant consonantal assimilation within the history of the language, namely an assimilation of the nasal feature (cf. §4.4.1.3.2), an assimilatory process that occurs among stops (cf. §4.4.2.2), and an assimilation affecting the place of articulation of sibilants.

The first process involves the nasalization of a stop (usually labial) when followed by a nasal obstruent, and has been very productive, to the extent of greatly increasing the frequency of the labial nasal, previously absent, in the language.[51] This assimilation is particularly productive when only a vowel separates the segments involved in the process, which do not necessarily need to be tautosyllabic, as in *heben* > *hemen* 'here', *bendekatu* > *mendekatu* (cf. Lat. *uindicāre*) 'avenge', *menia* (< Lat. *uenia*) 'truce' and *bañu* (cf. Sp. *baño*) 'bath'. Along with these examples, there are also instances of such nasalization when there is a greater gap between the participating segments, as in *bagiña* > *magi(ñ)a* (cf. Lat. *uāgīna*) 'scabbard, sheath' and *baemana* > *maemana* 'already given'. Nevertheless, there seem to be certain environments in which this assimilation is blocked, since sequences with an intervening coda rhotic, such as *bVr.n*, never become ***mVr.n* (cf. Egurtzegi 2011, §4.4.1.3.2).

The second process involves the devoicing of a stop when followed by a voiceless stop in the next syllable. This process has been a very common mechanism for the development of initial voiceless stops in both the patrimonial and the old loaned vocabulary, since some stages of Proto-Basque lacked voiceless stops in word-initial position, and thus early Latin loanwords were not borrowed with word-initial voiceless stops (see §4.4.2.2). Examples of this process are *bihi gain* > **bit-gain* > *bikain* > *pikain* 'best, first', *behi gorotz* > **bet-gorotz* > *pekorotz* 'cow dung', **gurd(i)-bil* > *gurpil* > *kurpil* 'wheel' *gal-te* > *kalte* 'harm, damage' (cf. *galdu* 'lose'), Lat. *pācem* > *bake* > *pake* 'peace', Lat. *catēna* > *gat(h)ea* > *kat(h)ea* 'chain', *putre* (cf. Sp. *buitre*) 'vulture', etc.

The final type of assimilation can be inferred from the high frequency of Basque words involving two homorganic sibilants, as noted by Michelena (*FHV*: 283), who explained most exceptions by analogy (e.g. with suffixes like *-tze*). Examples of this homorganicity include: *isats* 'tail', *itsaso* 'sea', *itsatsi* 'stick, glue', *itsusi* 'ugly', *sustrai* 'root', *zazpi* 'seven', *zezen* 'bull', *zorrotz* 'sharp', *zortzi* 'eight', *izotz* 'ice', etc; and examples of place assimilations are

---

51 Along with other processes and its presence in the borrowed vocabulary, see §4.4.1.3.2.

frantzes > frantses 'French' (cf. *Frantzia* 'France'), *sazoi > sasoi* (cf. Sp. *sazón*) 'seasoning' and *zin(h)etsi > sin(h)etsi* 'believe', which was still written with an initial laminal by Dechepare in the 16th century.

4.4.5.2. Dissimilation

Dissimilation occurs when a non-etymological reinterpretation results from an ambiguous sequence, including more than one segment with a particular phonological feature which contains an elongated phonetic cue (apud Ohala 1981 and 1993). It can affect either that particular feature (removing it from one of the segments which bears it) or the whole segment which possesses it (by removing the affected segment from the sequence). Segments affected by dissimilation include, among other phonological features, those bearing rhoticity, laterality, nasality, labiality and aspiration. However, the only systematic process of total dissimilation found in Basque takes place in a word which contains two aspirations.

In this process, usually known as Grassmann's Law in the Indo-European tradition, a sequence of two aspirations is systematically simplified by deleting the first of them. In Basque, it has affected both the oral /h/ and the nasalized aspiration /h̃/. Examples of this kind of dissimilation are mainly found in compounds such as *hil + hotz > ilhotz* 'corpse', *hil + herri > ilherri* 'graveyard', *hil + hargi > ilhargi* 'moon', etc.

We find further examples of dissimilations, which, as stated above, have only been sporadic. These include words such as *ardore > adore* 'fortitude, courage'; **artxeter > atxeter* (cf. Lat. *archiater*) 'physician'; *errierta > errieta* (cf. Sp. *reyerta*) 'quarrel', with the loss of a coda rhotic; *alkar, arkal < *hark-har* 'mutually'; *berar > bedar, belar* 'grass'; **berarri > belarri* 'ear'; *hortzadar > holtzadar* 'rainbow', with loss of the rhoticity; **eskinantzak > eskidantzak* (< Bearn. *esquinances*) 'tonsils'; **izenonbreak > izelonbreak* 'name and surname'; *puldamentu* (< Sp. *fundamento*) 'foundation'; with loss of the nasality; *baberruma > baberruna* 'bean', for labiality, etc.

4.4.5.3. Metatheses: reciprocal metathesis, perceptual metathesis and repercussion

There are two different kinds of metathesis which have been very productive throughout the history of the language (cf. Egurtzegi 2011), although neither of them seems to have been systematic in any particular context.

The first, namely reciprocal metathesis, affects two articulatorily similar segments located in the same syllabic position (i.e. both either on syllable onsets

or codas) in nearby syllables (usually consecutive). Some examples of these are the metathesis involving two coda liquids, as in *alper* : *arpel* 'lazy', *ergel* : *elger* 'stupid' and *sorbalda* : *solbarda* 'back (of the body)'; the metathesis of an onset rhotic and an onset approximant: *irago* : *igaro* 'go past', *irudi* : *iduri* 'image', *edarra* < *errada* (< Sp. *herrada*) 'wooden bucket' and *madarikatu* < *maradikatu* (< Lat. *maledīcere*) 'damn, curse'; the metathesis between an onset nasal and an onset lateral: *belena* (cf. Fr. *venelle*, Old Navarrese *benela*) 'gap between two houses', *il(l)unabar* > *iñulabar* 'dusk', *labana* < *nabala* (cf. Sp. *navaja*) 'knife'; and the metathesis of two onset obstruents agreeing in continuancy: *bage* > *gabe* 'without', *bilgor* > *gilbor* 'belly, paunch', *bekatari* > *betakari* 'sinner', *eskapatu* > *espakatu* 'escape, get away', *bezatu* : *zebatu* 'gotten used to', *ezabatu* (< Lat. *effaciāre*, cf. Occ. *esfaçar*) 'erase', *erakutsi* > *eratsuki* 'show' and *petxatu* (< Sp. *petacho*) 'patch'.

The second type of metathesis is directly related to dissimilation, as understood by Ohala (1981, 1993). As previously mentioned, features with elongated phonetic cues can create ambiguity in the string, which in some cases may be reinterpreted by the listener, thus yielding metathesis (Blevins & Garrett: 1998 and 2004), or even repercussion.

The features (and subsequent segments) susceptible to being metathesized are similar to those involved in dissimilation, and the instances of this process are also sporadic: *gurtze*, *kurtze* (< Lat. *crucem*) 'cross', *estrata* > *estarta* (< Lat. *strāta*) 'path, track' and *kruel* > *kurel* (< Sp. *cruel*) 'cruel', with metathesis of the rhoticity; *estalpe*, *estalbe* (cf. *establia*, < Sp. *establo*) 'stable' and *euli* > *ebli* > *elbi* 'fly' for laterality; and *harea* (< *areh̃a* < Lat. *arēna*) 'sand', *heuskara* (< *enuskara*) 'Basque' and *onherran* (< *honerran*) 'blessing' are examples of metathesis of the aspiration. Glides can also become relocated, as in *euskara* > *eskuara* 'Basque', *bakoitz* > *baikotz* 'each' and *hausin* > *hasuin* 'nettle'.

The case of repercussion (or copying of a segment) is similar to that of perceptual metathesis, but instead of reinterpreting the ambiguous consonant in a different location, the etymological position is maintained and a new segment is added in a different position. Examples of this process are: *altzairu* < *altzailu* < *atzailu* 'steel', *aitzina* > *aintzina* '(to the) front', *zentzu* > *zentzun* 'sense, consciousness'.

There are also some instances of feature metathesis, namely those of nasality and palatality: *holiñe* (< Sp. *hollín*) 'soot', *laño* (< Sp. *llano*) 'modest, affable', *señale* > *senalle* (< Sp. *señal*) 'signal', *artzani* > *artzãi* > *artzain* 'shepherd', *arrani* > *arrãi* > *arrain* 'fish', *lukanika* > *lukãika* > *lukainka* (cf. Vulgar Lat. *lucanica*) 'spicy pork sausage', Lat. *arēna* > *areã* > *ãrea* 'sand'.

4.5. Phonotactics

Given that there was a time when the language only had monosyllabic CVC roots (Lakarra 1995, etc.), every cluster must necessarily be considered secondary, both for consonants and vowels. However, various clustering and simplification processes must have played an important role in developing the modern language, in which certain clusters, but not any possible combination, are found in each word position.

Nuclei are always vocalic, since Basque lacks syllabic consonants.

4.5.1. Word and syllable initial positions

Not every segment was present word-initially in the Michelenian Proto-Basque. We reconstruct three voiced stops /b-/, /d-/ and /g-/, or, perhaps more precisely, a lenis stop series; /s̺-/, /s̻-/, /n-/ and /h/. This series of stops would later become voiced, as proposed by Martinet (1974 [1950]) and accepted by Michelena (*FHV*: 373-374). Initial /l-/ would later emerge instead of the coronal lenis stop /d-/ (as in *leka < *deka <* Lat. *thēca* 'pod', cf. §4.4.2.1), which was thereafter systematically lost from this position. Old Latin and Romance borrowings were, according to Martinet's theory, systematically adapted with initial voiced stops, given that Basque probably neutralized both series in favor of the voiced one word-initially, or even lacked a second series altogether in this position: *baradizu <* Lat. *paradīsum* 'paradise', *gurutze <* Lat. *crucem* 'cross', *leun < *deun <* Occ. *teun(e)* 'soft' (cf. however, the seemingly native *ke* 'smoke' and *kirats* 'stink').

Historically, there has never been any kind of cluster either in word initial or syllable-initial position until very recent times, when *muta cum liquida* types of clusters were not only allowed, but also produced. This is shown by examples such as the Romance loanwords *treku <* Sp. *terco* 'stubborn', *truku <* Sp. *turco* 'Turk' and *alpraketa <* Sp. *alpargata* 'espadrille', in which such clusters are produced by metathesis, or *andre < andere* 'lady', *eskla < ezkila* 'bell' and *mistrio <* Sp. *misterio* 'mystery', in which they are formed after a syncope.

In old borrowings, stop-liquid clusters were adapted by either of the following: anaptyxis, metathesis or deletion of the initial obstruent. Anaptyxis occurred mostly (although not exclusively) in clusters involving the rhotic, while deletion affected only obstruents which preceded a tautosyllabic lateral. We find instances of svarabhakti vowels, which are systematically copied from those following the cluster, in *liburu <* Lat. *librum* 'book', *gurutze <* Lat. *crucem* 'cross', *ferekatu <* Lat. *fricāre* 'caress, scrub', *garau̯(n) <* Lat. *grānum* 'grain' and *lukuru <*

Lat. *lucrum* 'interest, greed', and obstruent deletion in *laket* < Lat. *placet* 'like', *luma* < Lat. *plūma* 'feather', *loria* < Lat. *glōria* 'glory', *lore* < Lat. *flōrem* 'flower' and *eleiza* < Lat. *ecclēsia* 'church'. Finally, metathesis could affect clusters with either the rhotic or the lateral, and seems to be more dialectal and much less systematic than the other processes described above: B. *kurtze* < Lat. *crucem* 'cross', Al. *kurel* < Sp. *cruel* 'cruel', *Gregorio* > *Gergorio* 'Gregory', *estalbe*, *estalpe* < Sp. *establo* 'stable, barn' (but cf. also Rom. *estalvi*) and HN. G. *kolka* < *kloka* (cf. Sp. *clueca* < *clocar*) 'broody hen'.

See also Lakarra (2010a) on initial position and on the reconstruction of the left edge of the word.

4.5.2. Final position and coda tautosyllabic clusters

Only certain segments occur word-finally: sibilants, /n/, /r/ and less frequently, /l/. There are some instances of final stops in morphemes which can be found in both the declination and conjugation of the modern language, such as the ergative /-k/ and 1st person singular /-t/. However, these are secondarily formed from older voiced stop-vowel sequences (cf. *FHV*: 343 and Artiagoitia 1993, or §4.4.2.2), whose final stops lost voice after an apocope, as in **duda* > *dut* 'I have' or **niga* > *nik* 'I (erg.)' (cf. Lakarra 2006b). Certain modern dialects also contrast alveolar and palatal nasals and laterals in final position, but most do not. It is also possible that the Roncalese contrasted between rhotics word-finally (cf. Michelena 1988 [1953b], 1988 [1954a], Hualde *GB*: 34).

Final clusters, although probably secondary, are not rare in the modern language. These are formed by a sonorant /r/, /l/ or /n/, followed by a voiceless stop or an affricate, or a fricative sibilant followed by a voiceless stop. We find final clusters in *bost*, *bortz* 'five', *hartz* 'bear', *beltz* 'black', *ihintz*[52] 'dew', *bart* (< *barda*) 'last night', *galant* 'gallant'. Final stops and clusters are also usual in examples of phono-symbolism and onomatopoeia, as in *zart* 'smack', *urrup* 'slurp' and *zaunk* 'bark'.

4.5.3. Heterosyllabic clusters and medial position

The word-internal position is the only one in which most segments occur, since in the polysyllabic Michelenian Proto-Basque some segments, namely voiceless stops or taps, are only reconstructed between vowels.

---

52 This final laminal affricate has been proposed by Lakarra to be a suffix (1998a, 2003, etc.).

Medial clusters conserved in the modern dialects are usually composed of either a fricative sibilant or a sonorant (/r/, /l/ or a nasal) and a stop; or a sonorant and an affricate. These include: *asto* 'donkey', *oskol* 'shell', *ezti* 'honey', *azpi* 'down(side)', *erdi* 'half', *arto* 'corn', *argal* 'thin', *alde* 'side', *kalte* 'damage', *handi* 'big', *hanka* 'leg', *hertsi* 'close', *artzain* 'shepherd', *altxatu* 'stand up', *altzairu* 'steel', *intxaur* 'walnut' and *antzina* 'formerly'. Less frequently, we find clusters which contain a rhotic and either a lateral or a nasal, as in *erle* 'bee', *arnasa* 'breath' and *armiarma* 'spider'. Clusters containing a sibilant and a nasal are rare but attested: *asmatu* 'guess', *esne* 'milk' and *esnatu* 'wake up'.

However, modern clusters have not necessarily remained unchanged and, as mentioned in §4.4.2, most stops have spontaneously become voiced after a lateral or a nasal.[53] Clusters formed by a sibilant and a stop necessarily involve a fricative sibilant and a voiceless stop, since stops were devoiced after sibilants.

Although the aforementioned clusters have remained conserved in the modern dialects, most consonantal groups have been simplified, sometimes (but not necessarily) in a systematic way.

For clusters of two stops, these have been consistently adapted to a single voiceless stop which shares its place of articulation with the etymologically prevocalic stop (see §4.4.2.2), as in the following compounds: *begi* 'eye': **bet-gain > bekain* and **bet-buru > bepuru* 'eyebrow', **bet-belar > bepelar* 'eyelash', etc. Michelena (*FHV*: 346) suggests that a nasal in the second syllable may have had the same effect in some words, but only offers a couple of examples: *oroipen < *oroit-men* 'memory' and *zipotz* 'peg used to close a barrel', composed (as he himself proposed) of *zihi* 'wedge, little stick' (> *zit-*) and *motz* 'short'.

Clusters formed by a stop and a following prevocalic sibilant have usually yielded an affricate in composition: *artzain < *ard(i)-zani* 'shepherd', *betzain < *beh(i)-zani* 'cowherd', *betzain < *beg(i)-zain* 'inner membrane of the eye', *otsein < *og(i)-seni* 'servant', *betsein < *beg(i)-seni* 'pupil' (apud *FHV*: 346). The examples above may only correspond to clusters involving a coronal stop, since all stops —and even the aspiration— became /t/ in the final position of the first member of a compound (cf. *beg(i)* 'eye' + *ile* 'hair' > *betile* 'eyelash', etc.). However, this process also affected loanwords which included clusters with a non-coronal stop followed by a sibilant, as in *atsolutu* (cf. Lat. *absolūtum*) 'absolute', *etsamina* (cf. Lat. *exāmen*) 'exam', *etsenplu* (cf. Lat. *exemplum*) 'example'.

According to Hualde (*GB*: 23), in certain dialects the affrication of heterosyllabic stop+sibilant clusters can also occur across morpheme or even word

---

53 A neutralization developed in most dialects, with the exception of Roncalese and Souletin.

boundaries, across words with a high degree of syntactic cohesion: *bait zuen* > *baitzuen* 'because she/he had it', *onek zuen* > (Baztan) *onetzuen* 'this one had it'.

Two sibilants —either fricative or affricate— have regularly resulted in an affricate (*FHV*: 350-351). In the case of both sibilants being homorganic, the resulting affricate would also share their place of articulation (*ez-zen* > *etzen* 'it was not', *ez-zara* > *etzara* 'you are not', etc.), while a cluster formed by an apical followed by a laminal yields an apical affricate (**diots-zut* > *diotsut* 'I tell you', **ekus-zu* > *akutsu* 'see (you)', **sinhets-zazu* > *sinetsu* 'believe (you, imp.)', **hauts-tzaile* > *hautsaile* 'breaker'. However, Michelena (*FHV*: 351) finds no clear examples of two heterosyllabic sibilants yielding a laminal, but instead one which suggests the opposite: S. *egütseme* < **eguz-seme* 'godson' (cf. *egüzaita* 'godfather'); cf. also the etymologies **hortz-so* > *otso* 'wolf' and **hartz-so* > *atso* 'old woman', proposed by Lakarra. Any cluster involving a prepalatal sibilant would have a prepalatal outcome, regardless of its etymological position: **deretx-zun* > B. *deretxun* 'that you think', but also *gantxingor* < **gantz-txingor* 'crackling'.

Latin had some /mn/ nasal clusters that have been simplified in Basque. This simplification yielded either of the nasals involved: Lat. *damnum* > *damu* 'regret, damage', Lat. *autumnum* > B. *autono* 'autumn', Lat. *columna* > *koloma* 'column' or Lat. *domine* (> *dom'ne*) > *(jaun)done*, used with saint names.

Sequences of a nasal followed by a stop have sometimes been reduced to a single nasal, probably by assimilating the stop to the nasal (cf. **umbe* > *umme* —found in Lerga— > *ume* 'child' or *sembe* in Aquitanian anthroponyms > **semme* > *seme* 'son'). This process frequently affects labials, but is also found in alveolar clusters: *konbeni* > *komeni* 'to be appropriate, to be advisable', *konbentu* > *komentu* 'convent', *amizione* (cf. Lat. *ambitiōnem*) 'ambition', *gomitatu* (< Lat. *convitāre*) 'invite', *ganora*, *khanore* (< Lat. *candōrem*) 'skillfulness'.

The group of a sibilant followed by a sonorant —a cluster almost nonexistent in the modern language— has been simplified to a sonorant: *enaiz* < *ez naiz* 'I'm not', *igelenean* < **igaz-lenean* 'last year', *maialen* < *maiatz le(h)en* 'First of May', *mailasto* < *maiz-lasto* 'corn straw', *iraleku* < **iratz(e)-leku* 'bracken', etc. Michelena (*FHV*: 351) proposes that the coda sibilant would have become a voiced aspiration before being dropped.[54]

The clusters formed by a prevocalic lateral preceded by a sonorant have been simplified by deleting the first of the sonorants: *zala*, *zela* < *zan*, *zen* (plus the complementizer *-la*) 'that was (him)', *hola* < *hon(e)la* 'in this way, thus', B. *imillau(re)n* < **imin-lauren* 'measure for grain', B. *biloba* < *birloba* 'grand-

---

54 As mentioned in §4.4.3.3.2, sibilants are realized as voiced before a voiced consonant. Cf. also /-r/ > /-h/ in the first member of a compound, in §4.4.1.1.

child', B. *bialeku* < *bear-leku* 'workplace', or the place name *Goliz* (< *Gorliz*) in RS. However, /rl/ clusters are not absent from the modern language (cf. *arlo* 'task', *erle* 'bee', *gerla* 'war', *irla* 'island', etc.).

In consonantal clusters formed by three segments complex codas are simplified, while *muta cum liquida* onsets are much more common. In the case of a coda composed of a sonorant followed by a sibilant, the former is usually dropped: *belztu* > *beztu* 'blacken', *a(h)aztu* < *a(h)anztu* 'forget', *jazten* < *janzten* 'dressing', *ostu* < *onstu* 'stolen', *(h)osto* < *(h)orsto* 'leaf', *estu* < *herstu* 'tight', *ozpin* < *orzpin* 'lightning' and *asto* < *arsto* 'donkey'. The other case of complex coda described in §4.5.2 —formed by either a sonorant or a sibilant and a stop— would result in the stop at the end of the syllable being systematically lost, but only after devoicing the following stop (cf. the treatment of stop clusters supra): *ausarki* < *ausartki* 'bravely', *arkume* < *ard(i)-kume* 'lamb', *gurpegi* < *gurd(i)- begi* 'bucket' or *gurpil* < *gurd(i)-bil* 'wheel'. This intermediate stop is simply dropped before a *sonorant*, as in *izerleka* (from *izerd(i)* and *leka*) 'sweat that creates dirt' or *ipurñauska* (cf. *ipurdi*) 'jump, gambol'.

4.6. Accentuation

Modern Basque dialects have different types of accentuation, based on either stress or on pitch accent. However, there is a central area, including most of the High Navarrese, Low Navarrese and Labourdin varieties, which lacks any contrastive accentuation. In most High Navarrese varieties, stress automatically falls on the penultimate syllable, and it lacks any other accentuation pattern to contrast with, unlike Souletin, which has oxytonic words in addition to paroxytonic (cf. §4.6.3). According to Gavel (1920: 108), the Labourdin and Low Navarrese areas have no word-level stress (cf. Hualde 1993a: 31). Nevertheless, earlier types of accentuation may recently have been lost in the area (cf. Michelena 1987 [1972c]).[55]

The areas with distinctive accentuation surround this central territory and are located in both the western and eastern sides of the Basque Country (Hualde 1993a: 13). The most widespread accentual system is the central-western type, with peninitial stress, i.e., stress on the second syllable of the word.

---

55   Cf. also the Goizueta variety of High Navarrese (Hualde & Lujanbio 2008, and Hualde & Lujanbio & Torreira 2008), which still maintains a pitch accent system similar to the one found in Northern Biscayan (§4.6.2).

4.6.1. Central-western accentual type

Present in most of Guipuscoa along with some areas in southeastern Biscay and western Navarre, stress on the second syllable is the most common type of stress. It has also been proposed by Michelena to be common to all varieties in older stages of the language (1988 [1957/58], *FHV*: 405-410). In most modern varieties with this type of accentuation, disyllabic words are stressed on the initial syllable instead, thus avoiding final prominence: *emákumea* 'the woman', *gizóna* 'the man', *mutíla* 'the boy' but *néska* 'the girl'. There are also a number of marked words with initial stress, mostly borrowings and compounds (Hualde *GB*: 71): *básoa* 'the drinking glass' (< Sp. *vaso*, cf. *basóa* 'forest'), *léngusua* 'cousin', *béstea* 'the other', *égia* 'the truth', etc. In many varieties, plural forms also bear initial stress whenever they are either disyllabic or trisyllabic, but maintain post-initial accent when longer: *gízonak* 'the men', *gízonai* 'to the men', but *emákumeak* 'women', *emákumei* 'to the women'.

4.6.2. Northern Biscayan accentual type

In this region there is a lexical distinction between unaccented and accented words. This distinction is morphological, in the sense that roots and affixes are lexically specified as accented or unaccented. Most native roots and almost all singular affixes are unaccented, whereas most borrowings, compounds and plural affixes (as well as a few native roots and affixes) are accented. Unaccented words contain unaccented morphemes, and accented words contain one or more accented morphemes. Unaccented words do not bear prosodic prominence on any syllable, except when they occur in isolation or immediately preceding the verb, contexts in which they surface with final prominence in most local varieties. Accented words always display stress on a given syllable. In most subvarieties of Northern Biscayan Basque, an accent is assigned to the syllable preceding the accented morpheme,[56] and in a few subvarieties stress is assigned to the penultimate or antepenultimate syllable.

Examples of unaccented words are *semié* 'son', *mutillerí* 'to the boy', *lagunentzakó* 'for the friend', and examples of accented words are *béste* 'other', *léku* 'place', *léngusu* 'cousin', *egúzki* (or *éguzki*) 'sun', morphologically derived words such as *edérrago* 'more beautiful', *báltzegi* 'too black', *lóti* 'sleepy', or plural forms like *sagárratatik* 'from the apples', *itturríetara* 'to the fountains'.

---

56   When there is more than one accented morpheme, it is the syllable preceding the leftmost morpheme that bears stress.

Accent in Northern Biscayan Basque is realized as a H*+L falling contour tone on the accented syllable (Hualde 1989, Hualde & Elordieta, G. & Elordieta, A. 1994, Elordieta 1997, 1998, Gussenhoven 2004: 171, among others).

Recently, the High Navarrese variety of Goizueta has been found to be of the pitch-accent type, with similarities to Northern Biscayan Basque (cf. Hualde & Lujanbio 2008; Hualde & Lujanbio & Torreira 2008), showing that the prosodic system observed in Northern Biscayan Basque must have extended over all the western areas in older times (cf. Hualde 1993a, 1995, 2003a, 2006, 2007, Elordieta 2011a, 2011b).

4.6.3. Souletin accentual type

The Souletin dialect has a different accentual pattern, which is not far from that of the extinct Roncalese dialect. A similar pattern may also be hypothesized for older stages of Salazarese, although not for the modern dialect (*FHV*: 397-398). These dialects have their unmarked stress on the second syllable from the right (i.e., most words are paroxytones) as in Souletin *gízun* 'man', *gizúna* 'the man', etc., and there may have been a time when every word was stressed on the penultimate syllable.

Nevertheless, this pattern is not entirely systematic today. In those words in which a final vocalic coalescence has occurred, both modern Souletin and Roncalese have word-final stress: *alhába* 'daughter' + abs. sg. *-a > alhabá* 'the daughter', *néska + -a > neská* 'the girl'. This also occurs in non-absolutive plural forms (i.e., those bearing an *-e-*) such as S. *gizunék* 'the men (erg.)', *gizunén* 'of the men', *gizunér* 'to the men' (< **gizon-a(g)-ek*, **gizon-a(g)-en*, apud *FHV*: 394, etc.). Alongside these exceptions, oxytonic stress is found on some derivational suffixes (*izabár* 'from Izaba', *larrañtár* 'from Larraine', *aitañí* 'grandfather' < *áita* 'father'), compounds with a monosyllabic second member (*erháuts* 'dust', *ollár* 'rooster') and some other words, both in the borrowed vocabulary (*biarnés* 'Bearnese', *repattán* (cf. Pyrenean Aragonese *repatán*) 'young shepherd', *arrés* 'sheep', *animál* 'huge') and native vocabulary, mostly ending in diphthongs (*ãhái* 'ram' (< older *ãhári*), *exéi* 'vixen' (< *azeri*), *aihái* 'dinner' (< *aihari*), S. *ardṹ*, R. *ardáũ* 'wine' (< **ardano*), *orgã́* 'cart' (< **organa*), *etsái* 'foe', *eztéi* 'wedding', R. *idói* 'swamp', *odéi* 'cloud', etc.)[57] (*FHV*: 394-397,

---

57 Michelena (*FHV*: 397) notes the large number of stressed final diphthongs in Souletin and proposes that they originated from the loss of an intervening consonant, which he hypothesizes to be an aspiration, given its restriction to the first two syllables in the modern dialects.

Hualde *GB*: 72). Such words, at least those in the inherited lexicon, may also be the consequence of resyllabification after the disappearance of the final syllable.

Two main factors distinguish the Roncalese accentuation from that of Souletin (*FHV*: 395). Firstly, in Roncalese, declinational suffixes do not affect the location of the stress, as long as they do not imply any syllable loss: *gízon* 'man', *gízona* 'the man', *gízonaren* 'of the man', *gízonareki* 'with the man', but cf. *alába* 'daughter' vs. *alabá* 'the daughter'. Secondly, Roncalese adds some proparoxytones to the oxytones and paroxytones present in Souletin. This accentual retraction only seems to affect paroxytones —since words that have developed oxytonic stress tend to remain so— and applies not only to native words but also to paroxytonic loanwords. Examples from Izaba are *bézino* (cf. Sp. *vecino*) 'neighbor', *kósino* 'cousin' (*kosíno* in Uztarroze), *ézkapa* 'escape, run away', *zámari* (< Lat. *sagmārius*) 'horse', etc. Michelena (*FHV*: 397) relates these accentual differences to the lack of syncope, which is usual in Roncalese (cf. §4.3.4.5), and to the resyllabification of mid and high vowels (§4.3.4.2) in Roncalese and Salazarese, where *-ea* and *-oa* are monosyllabic.

4.6.4. Old Accentuation

As we have mentioned earlier in this work (in §4.4.3.1.1 and §4.4.2.3), Michelena (*FHV*: 407) proposed that stress was located on the second syllable of the word (as opposed to Martinet 1974 [1950], who proposed that it was on the first syllable) since this would account for the modern distribution of both the aspiration and the aspirated stops in the dialects that still maintain them. Aspiration is limited to the first two syllables in a word, and Michelena relates this to the presence of stress in the second syllable, so that the generalization would be that aspiration occurs on the stressed syllable or the one preceding it.

However, as Michelena himself stated (*FHV*: 418), assuming peninitial stress would run into the problem posed by the vowel neutralizations and apocopes found in the second syllables of disyllabic first members of compounds, in which final *-a*, *-e* and *-o* are neutralized to *-a* —probably after having been produced as a central unstressed schwa— and *-i* and *-u* are dropped (cf. also longer words, in which all final vowels are lost, see *FHV*: 410). In order to prevent this incompatibility, Michelena hypothesizes that peninitial stress was more modern than the system with initial stress proposed by Martinet. This would avoid any conflict with the neutralization or loss of a vowel that would have been stressed if the demarcative stress on the second syllable were already operative. The proposal of a late date for this stress pattern may also account for the aspirations found in the documents from San Millán originally written in the 11th century,

which appear on any syllable of the word. Lakarra (2005, 2006b) and Igartua (2001) also consider the period with initial stress to be posterior to the period with second-syllable stress.

Elordieta (2011a, 2011b) proposes two possible diachronic evolutions of Basque stress and explains the transition from one stage to the next. Elordieta (2011a, 2011b) begins with a mainly monosyllabic stage in which the first disyllabic words were created by means of reduplication. At that time, stress would have fallen on the last syllable, which was also the second. Elordieta subsequently provides support from tonal and intonational phonology to account for a later stage in which stress was word-initial (as proposed by Martinet 1974 [1950] and also assumed by Michelena in *FHV*) and for a further evolution into phrase-final stress, as in Northern Biscayan Basque. However, following an argument by Martínez-Areta (2004), Elordieta points out a problem for assuming word-initial stress, and that is the impossibility of explaining why marked words in many central varieties are characterized precisely by having initial stress (cf. Elordieta 2011a: 1006-1007). That is, if all Basque words were initially stressed at some point, it would be impossible to distinguish marked words on the basis of initial stress. Thus, Elordieta proposes a chronology without the conflicting stage, with the older word-final prominence becoming penultimate stress in eastern varieties and phrase-final in western dialects. Only Northern Biscayan Basque maintains the phrase-final stress system, and all other western stress systems derive from the old phrase-final stress system.

Hualde (2007) proposes that the accentuation pattern of Proto-Basque could have been similar to the pattern present in modern northern Biscayan, and that the pitch rise on the second syllable of the word may have been reinterpreted in eastern varieties as a demarcative non-contrastive stress on the second syllable. Hualde (2003a) claims that the influence of Spanish stress should be considered when understanding this reinterpretation, since this is also realized as a pitch rise. This hypothesis was confirmed empirically by Elordieta & Hualde (2003) by means of a perception test. The historical relation between the Northern Biscayan and the central-western accentual type is also favored by the fact that they share the difference between stressed and unstressed —or marked and unmarked— words, and the aforementioned divergence in the realization of accentuation (pitch accent vs. stress) is a typologically frequent change.

According to Michelena (*FHV*: 423-424), the transition from the central accentual system —reconstructed by him as the common accentuation— to the modern Souletin paroxytonic stress would have been the consequence of a reinterpretation of the position of the stress in trisyllabic words as counted from the right (instead of the left) edge of the word. This would be followed by a subse-

quent regularization of the new pattern in disyllabic words, which would have their stress systematically relocated to the initial syllable, i.e. the penultimate.

Let us roughly follow Hualde in order to accept the aforementioned chronological sequence, and propose that the diachronic evolution must have been phrase-final stress (nowadays only found in Northern Biscayan Basque) > peninitial stress (in central varieties) > penultimate stress (in eastern varieties). In this way, we can discard the reconstructed old second syllable stress proposed by Michelena and accepted by most authors. We must assume that the loss of the aspiration after the second syllable occurred not generally but only eastwards from the central territory, in an area which includes precisely those dialects in which the /h/ is maintained to this day. This would not be incompatible with the appearance of the aspiration on any syllable in medieval documents in the meridional area, given that such varieties probably followed the western accentual model. However, a reconstructed prefixing, mainly a monosyllabic stage with stress on the second and final syllable —occupied by the root— seems rather plausible (cf. Lakarra 1995, 2006b).

Regarding the proposal of demarcative stress on the first syllable (Martinet 1974 [1950]), I agree with Elordieta (2011a) that it is probably the hypothesis with the least empirical support.[58] Martínez-Areta (2004) rejects this hypothetical phase of the accentuation altogether. I can point to another weakness in Martinet's proposal, namely that the prosodic prominence granted to this position (in order to explain the consonantal processes affecting stops that occurred word-initially) can be attributed to word-initial position (cf. Smith 2005), without having to posit the presence of stress on that syllable.

4.7. Conclusion

Since the seminal work of Michelena (*FHV*) on the historical phonology of Basque, some aspects of the diachronic evolution of Basque phonology have received considerable attention by different researchers in the last two decades. These include the inventory of stop consonants in Old Basque (cf. §4.2.2) and the reconstruction of the accentuation system (cf. §4.6.4). Here we have provid-

---

58   The most frequently-mentioned fact which favors initial stress is that it would easily account for the vowel neutralizations and losses which occurred in the second and third syllables of the first members of compounds mentioned supra. However, Hualde (2006, 2007) gives a segmental explanation for such neutralizations: according to his hypothesis, the addition of an intervening laryngeal segment in composition would not only explain the vowel reductions but also the devoicing and neutralization of the stops, which become alveolar /t/, perceptually the most similar segment.

ed a revision of Michelena's work in the light of these recent advances, in an exercise that has tried to put this author's influential proposals into perspective. Revisiting Michelena's work while taking advantage of the progresses made by both typological approaches to phonology and phonetically-based phonology has proven fruitful. Without trying to make any significative incursion into the field of phonetic explanations for phonological processes, a couple of proposals have already surfaced in the development of this work, e.g. a hypothetical relation for the allegedly unconnected contexts in which the otherwise general /u/ > /y/ fronting fails to develop (cf. §4.3.3) and the reasons behind the metathesis of sounds bearing elongated phonetic cues (as understood by Ohala 1993), which does not apply to segments lacking such features (cf. §4.4.5.3). This can only encourage us to advance in the task of applying this type of framework to the Basque diachronical data.