Basque from a typological, dialectological and diachronic point of view

Georges Rebuschi

Georges Rebuschi. 2004.

Basque from a typological, dialectological and diachronic point of view.

In Thorsten ROELCKE (ed.), *Variation Typology. Variationstypologie*,
Berlin & New York : W. de Gruyter, 837-865.

Contents

Basque from a typological, dialectological and diachronic point of view

Introduction

Basque (*euskara*) is the only pre-Indo-European language that has survived in Western Europe, and there is as yet no compelling evidence that it is genetically related to any known language — except Aquitanian, its more or less direct ancestor (Michelena 1964, Gorrochategui 1995); it has been compared to every known language families and language isolates; a review of many of those hypotheses is provided by Trask 1997; see also Michelena 1968 for the B/Caucasian hypothesis, and my own review of Morvan 1996 for the putative Ural-Altaic/B connection.

Around the beginning of our era, as place-names and various inscriptions show, B was spoken throughout Caesar's *Novempopulania* and, on the southern side of the Pyrenees, as far as the Ebro plain; many place-names also indicate that it must have been spoken well into today's Catalonia; but the extent to which it was spoken to the west of Navarre's western boundary is still a matter of debate; in any case, the toponymy of the southernmost province of the Autonomous Community or *Euskadi* in modern Spain, Alava (*Araba*), which is almost exclusively Spanish-speaking today, was clearly Basque-speaking in the Middle Ages. Finally, there is no evidence that B was ever spoken to the west of Bilbao, the capital of the western province of Biscaye (*Bizkaia*).

The first book ever printed in B dates back to 1545, so that the real history of the language is fairly short. However, onomastic data, hand-written annotations or yet a few travellers' vocabulary notes have been discovered and published (Michelena 1964, Sarasola 1983, Satrustegi 1987, Orpustan 1999, among others). Such documents enable us to draw a fairly secure picture of many aspects of the language as it was spoken in the Middle Ages, and even earlier, as far as the phonology is concerned.

Even the oldest documents are dialectally differentiated. The founder of B dialectology, L.-L. Bonaparte, basing his classification on verbal morphology, identified eight major dialects, as appears in the full title of his 1869 book. Among those, he recognized four literary dialects, namely, from east to west: Souletin (*zuberoera*, henceforth Zu) and Labourdin (*lapurdiera*, La), spoken in France, and Guipuzcoan (*gipuzkera*, Gi) and Biscayan (*bizkaiera*, Bi), spoken in Spain. Table (A) gives a rough idea of the geographical distribution of those eight dialects (new abbreviations: 'WLN', Western Lower Navarrese; 'ELN', Eastern Lower Navarrese; 'NHN', Northern Higher

Navarrese; 'SHN', Southern Higher Navarrese, 'Ae', Aezkoan, 'Sa', Salazarese, and 'Ro', Roncalese; the last three were considered to be southern subvarieties of the corresponding dialects spoken in France across the border):

Table A. *The B dialects according to Bonaparte 1869*

	West			*East*		
North			La	WLN	ELN	Zu
South	Bi	Gi	NHN	(Ae)	(Sa)	(Ro.)
			SHN			

The dialect spoken in Alava (if there ever was such an independent variety of B, already extinct when Bonaparte drew his map), would have occupied the empty box on the left-hand side of the bottom line.

In his latest study to date, Zuazo 1998 reduces the sum total of dialects still alive to six: NL (*nafar-lapurtera*, French *Navarro-labourdin*) and Zu correspond to the first line of Table A; corresponding to the second line, we find Bi (labelled 'Western'), Gi ('Central') and *nafarrera* (i.e. Higher Navarrese), now taken to include Ae, and finally *ekialdeko nafarrera* or Eastern [H-]Navarrese (Ro being now extinct too).

As far as the written language is concerned, two main varieties have emerged in the past decades. In France, the NL macro-dialect obtained quasi-official recognition with the publication of Lafitte's *Grammaire [...]* in 1944. Across the border, in the thirties, Euskaltzaindia (the Academy of the Basque language) first promoted an "enriched" variety of Gi (*gipuzkera osotua*) which has now developed into "unified" B (*euskara batua*, henceforth EB), the language used today (together with Spanish) in education and administration in the Autonomous Community. Given the demographical weight of the "Southern" Basques (there are roughly 20 times as many B speakers in Spain as there are in France), it is EB as it is written (and spoken today on television) that will be described in the next section.

1. The main typological characteristics of standard Basque

1.1. Sound structure

Owing to the essentially written character of standard Basque, I will start from its spelling, which is a fair, if not perfect, approximation of the phonemic system of *euskara batua* as it is pronounced in Guipuzcoa.

1.1.1. The segmental level.

1.1.1.1. The vocalic system has a quite ordinary five unit system, written *a, e, i, o, u*. There also are five falling diphthongs, spelt *ai, ei, oi, au, eu* (conversely, orthographic sequences like *ia* or *ua* are typically bisyllabic); it follows that the syllable boundary insequences like *aia* or *eua* is after, rather than before, the closed vowel.

1.1.1.2.1. The standard spelling consonants are those in Table (B).

Table B *Standard Basque consonants*

	labial	*alveolar*		*pre-palatal*	*velar*	*post-velar*
plosives [-voiced]	p	t		(tt)	k	
plosives [+voiced]	b	d		(dd)	g	
affricates		tz	ts	tx		
fricatives	f	z	s	x		h
nasals	m	n		ñ		
liquids		l		ll		
rhotics		r / rr				
glide				j		

Comments. (a) The [±voiced] opposition only functions with the plosives, the [+voiced] ones often being realized as voiced continuants or fricatives ([β, δ, γ]). (b) The phonemic status of *tt* and *dd* is only (marginally) clear in the northern dialects, whereas the sounds these digraphs represent are mere allophones of *t* and *d* in the southern ones after [±syllabic] *i*. (c) The fricatives and affricates are all voiceless, cf. (a) above; *z* and *tz* are lamino- or dorso-alveolar, *s* and *ts* are apico-(post-)alveolar, and *x* and *tx* are prepalatal: there is a larger number of fricatives than of plosives. (d) EB has imported *h* into its spelling: this letter is only pronounced in the north. (d) The nasals are neutralized as orthographic *n* before consonants (but regressive assimilation is the rule), and at the end of words. (e) The palatal nasal and liquid are spelt *ñ* and *ll* only at the beginning of words; owing to a very general process of palatalizing *n* and *l* after a syllabic or non-syllabic *i*, the same sounds occur as allophones of the full phonemes *n* and *l*, but if the *-i-* is dropped, as is often the case, minimal pairs are brought about which do not contrast in the number of segments, but in their quality, as in *oina* 'the foot', pronounced [ojna] in the north but [oña] in the south, where it contrasts with *ona* '(the) good (one)'. (f) The two rhotics (which do not appear initially except in recent borrowings) are just one flap in the case of simple *r*, and a trilled rhotic in the case of *rr* – but the opposition only works intervocalically: in other contexts, the grapheme *r* is used, but is realised trilled. (g) Finally, the letter *j* whose recommended pronunciation is a simple yod is generally realised like a post-velar continuant in Guipuzcoa (as it is in Spanish).

1.1.1.2.2. Consonant clusters are normally absent in word-initial position: ancient borrowings have an epenthetic vowel in nouns like *giristino* 'Christian' or have simplified the initial group, as in *lore* 'flower' (< Lat. *florem*). In simple words, intervocalic clusters consist of a continuant (a fricative, nasal, liquid or the trilled rhotic) followed by a plosive or an affricate – but fricative + affricate sequences are not tolerated: *enbor* '[-mb-] (tree-)trunk'; *zortzi* 'eight'; a subset of those clusters is also admitted word finally: *beltz* 'black'. In complex words, more elaborate clusters are allowed, e.g. *da.tor.z.ki.t* ' 3rd.p.-come-plural-DF-me', 'they're coming to me', but many simplifications also occur. For instance, a sequence /affricate-plosive/ results in /fricative-plosive/, but *progressive* assimilation in voice/voicelessness also takes place, as in *irispide* 'way of achieving something' <*irits-* 'arrive' + *bide* 'way', with the simultaneous affricate > fricative change.

1.1.2. At the supra segmental level, EB can be characterized as follows (after Hualde 1997, but the analysis is partly my own) for young speakers who have made their studies in Basque: (a) two (or more) syllable-long words have an intensity stress; (b) the length and phonological structure of the *root* is relevant.

1.1.2.1. When the accentual unit is the word, (i) two and three syllable-long words whose root is monosyllabic are stressed on the first syllable: *lúrra* <*lur*+*a* 'the earth', *lúrrari* 'to the earth'; *béltzago* (<*beltz*+*ago*) 'more black'; (ii) when the root is bisyllabic and ends in a vowel, the root pronounced in isolation is also stressed on the first syllable: *áte* 'door', *méndi* 'mountain'; (iii) in all other cases, polysyllabic words have a primary stress on the second syllable, counting *from left to right*, plus a final secondary stress if the word has at least four syllables: *gizón* 'man', *gizóna* 'the-man', *gizónarì* 'to-the-man', *gizónarentzàt* 'for-the-man'. Words described under (i) are no longer irregular when they have at least four syllables: *lurrárentzakò* 'for the earth'; finally, words whose root has two syllables, the last of which is open, follow the general rule as soon as they are at least three syllables long: *mendía* 'the mountain', *atéarentzakò* 'for the door'.

1.1.2.2. The accentual unit is longer that the word when a noun is followed by an attributive adjective, or an adjective is preceded by an adverb, or yet when the copula cliticizes to the predicative phrase. In such cases, the general rule applies, as in: [*mendí zabalà*] 'a/the big mountain', [*alárguna dà*] '(she) is a widow', or even: [*osó zabala dà*] 'it's very wide', with no stress on the lexical word.

1.1.2.3. At the utterance level, the main stress falls on the accented syllable of the last word preceding the verb complex (either a lexical verb directly inflected – except the copula, see above – or a verbal participle followed by an inflected auxiliary): this position has been identified as the focus position since at least Altube 1929.

1.1.2.4. Intonation. Assertive utterances have an falling contour; exclamatory utterances

have a lower than usual final tone; yes-no questions have a raising contour, at least if there is no morpheme distinguishing them from the corresponding assertions, whereas yes-no questions segmentally marked as such (by the insertion of the particle *al* or *ote* immediately to the left of the inflected verb form) and wh-questions pattern with assertive utterances.

1.2. Morphology

1.2.1.1. B is typically an agglutinative language with by a vast amount of suffixes, and very few prefixes. The case system has generally been described as absolutive/ergative, and the conjugation is pluripersonal.

1.2.1.2. Most verbs cannot be inflected for tense and agreement (but are rather aspectually marked, and combined with a tensed auxiliary), and nouns and adjectives are easily transformed into verbs; the divide between these parts of speech is thus not as clear as is sometimes assumed:

(1)	a	ezkondu da	's/he has married'
	a'	ezkontzen da	's/he marries / is getting married'
	b	gizendu da	's/he has fattened/put on weight'
	b'	gizentzen da	's/he is getting fat'
	c	gizondu da	's/he has become a man'
	c'	gizontzen da	's/he becomes a man'

The difference may however be established by dropping the perfective *-du* or the imperfective/iterative *-tzen* ending and replacing it with *-a*:

(2)	a	*ezkona	
	b	gizena	'the fat one'
	c	gizona	'the man'

(In fact, *ezkon-a* was artificially created a hundred years ago with the meaning 'spouse', but just as soon dropped out of use.) Semantically, *gizena* does not denote an entity *per se*, but is understood anaphorically, whereas the reference of an item like *gizona* is not felt to rely on ellipsis interpretation. Moreover, the relative word order is telling:

(3)	a	gizon gizena	'the fat man'
	b	*gizen gizona	

Other word classes are postpositions, which are either declined nouns (*gainean* 'on, on top of', *gainetik* 'from above') or unanalysable (*aurka* 'against'), adverbs which may also be analysable (*labur-ki* 'in short') or not (*gaur* 'today') and conjunctions (*eta* 'and', *ala* 'or').

Finally, note that Basque has a very powerful means of deriving adnominal elements from about everything: the suffix *-ko* can be attached not only to ablative or allative case

endings (see 1.2.2.1) but also to inflected verb forms, like the ancient *balizko* 'fictitious', from *(ba-)litz* '(if) it were', or to complement clauses, as in (4b), derived from (4a):

(4) a [ezkondu dela] esaten dute
married aux-that saying aux
'they say he's got married'
b ezkondu dela-ko gizona
'the man who they say has got married'

1.2.2. Nominal morphology

1.2.2.1. B has neither nominal gender nor any classifier system.

1.2.2.2. The determiner+case system concerns nominal phrases, not nouns as such, since the number and case suffixes appear on the *last* word of the NP. The Number/Det subsystem is tripartite: (a) singular (*-a-*), (b) plural (*-e-*), and (c) "indeterminate" (Ø). The indeterminateness of zero concerns both number and definiteness, since no Det is allowed to be suffixed to the NP after interrogatives like *zein* 'which', or *zenbat* 'how many' – which both precede the head noun – or quantifiers like *bat* 'one' and *asko* 'many', which follow the noun. See Table (C).

Table C. *The Det+Case system; examples in the dative.*

sg	gizon.a.ri man.sg.dat 'to the man'	gizon gazte.a.ri man young.sg.dat 'to the young man'
pl	gizon.e.i man.pl.dat 'to the men'	gizon gazte.e.i man young pl.dat 'to the young men'
indet	gizon bat.Ø.i man one.Ø.dat 'to a man'	gizon gazte bat.Ø.i man young one.Ø.dat 'to a young man'
indet	zein gizon.Ø.i which man.Ø.dat 'to which man/men?'	zein gizon gazte.Ø.ri man young.sg.instr 'to which young man/men?'
indet	gizon asko.Ø.ri man many.Ø.dat 'to many men'	gizon gazte asko.Ø.ri man young many.Ø.dat 'to many young men'

Note the introduction of an epenthetic *-r-* between a vowel and the case suffix (save in the dative plural, a clear exception). Another property of the Number/Det subsystem is that when it qualifies a direct object NP, the sg. suffix *-a* is used ambiguously with ±definite meaning, as in *ogiA jan dut*, either 'I've eaten *the* bread' of 'I've eaten bread', whereas it is unambiguously definite when it specifies a subject NP.

The absolutive, besides being the dictionary form of nouns, is used with

"unaccusative" intransitive subjects and with direct objects, whereas the ergative is used with the subjects of transitive verbs. Morphologically, the abs. pl. ending is irregular: its plural form is the (synchronically) synthetic *-ak*, which is identical to the erg. sg. suffix.

Scholars do not agree on the exact nomber of cases in B, for two main reasons. (a) The divide between case and determiner is unclear in the case of the partitive suffix *-ik*; for reasons discussed in 1.3.2, I regard it as a contextual variant of of the indet. abs. (b) The divide between case suffixes and postpositions is unclear too. Thus, the "prolative" case *-entzat* could be analysed as a postposition *-tzat* 'for' governing the genitive (*-en*). Tables (D.i-iii) summarize the main cases (the segments in brackets in the indet. column are epenthetic; irregular forms are in italics). The morphologically complex cases, and the locative ones, are set apart.

(D.i) *Simple case suffixes*

	sg	*pl*	*indet*
absolutive	-a.Ø	*-ak*	Ø≈-(r)ik
ergative	-a.k	-e.k	-(e)k
dative	-a.ri	-e.i	-(r)i
genitive	-a.ren	-e.n	-(r)en
instrumental	-a.z	-e.z	-(e)z

(D.ii) *Complex case suffixes*

	sg	*pl*	*indet*
comitative	-a.re+kin	-e+kin	-(r)e+kin
prolative	-a.ren+zat ≈ -a.ren.dako	-en+tzat ≈ -en+dako	-(r)en+tzat ≈ -(r)en+tzako

(D.iii) *Locative cases*

	sg	*pl*	*indet*
locative/inessive	*-ea*.n	-e-*ta*-n	-(e)*ta*-n
ablative	*-Ø*-tik	-e-*ta*-tik	-(e)*ta*-tik
allative	*-Ø*-ra	-e-*ta*-ra	-(e)*ta*-ra
adnominal loc	*-Ø*-ko	-e-*ta*-ko	-(e)*ta*-ko

Notes. (a) The suffix *-tzat* of the prolative may also be directly suffixed to the root. The difference in meaning between the two uses is as in (5):

(5) a seme.a.ren.tzat hartu dut
son.sg.gen.tzat taken I-have-3sg
'I've taken it (e.g. a book) for my son'
b seme.Ø.tzat hartu dut
'I've (mis)taken him for my son'

(b) The locative cases are peculiar in several ways. (i) The sg. suffix *-a* does not appear.

(ii) The pl. and indet. numbers trigger the apppearance of the "morph" *-ta-*. (iii) When head nouns are [+animate], the case morphemes generally cannot be directly suffixed to the last word in the NP: a postposition is used, which may be declined, as in *gizona(ren)gan* 'in the man', *bi gizon(en)gandik* 'from two men' etc. In the plural, however, such forms as the following *can* be used: *lagun.e.ta.ra joan da* '(he) has gone to (meet) (his) friends'. (iv) The "adnominal locative" or "locative genitive" in *-ko* has many uses (1.2.1.2). First, if used directly, it transforms an adverbial phrase in the locative case into an adnominal adjunct, as in:

(6) a Donostia.n diren/dauden elizak
San Sebastian.loc which-are churches
'the churches that are in San Sebastian'
b Donostia.ko elizak
'The churches in San Sebastian'

(**Donostian elizak* is unacceptable). Second, any case from the instr. down to the allative in Tables (D) can also be made into an adnominal adjunct or complement:

(7) a *urre.z* '(made) of gold' (instr) => *urre.z.ko eraztuna* 'a gold ring'
(**urre eraztuna* and **urrez eratztuna* are impossible)
b arrotz herri.e.kila.ko harremanak
foreign country.pl.comit.*ko* exchanges
'exchanges with foreign countries'
c Donosti.ra.ko bidea
San Sebastian.allat.*ko* road
'the road to S.S.'

(*-rako* can also be used as yet another allomorph of the (ordinary) prolative case with inanimates).

1.2.2.2. Although B is a "generalized pro-drop language", its has both unmarked and marked (emphatic or strong), pronouns. Table (E) gives their forms in the absolutive and the genitive (see 1.3.3 for their distribution).

Table E *The absolutive, and weak and strong genitive pronouns of standard Basque*

Person	*weak absolutive*	*emphatic abs.*	*weak genitive*	*strong genitive*
I sg	ni	neu	nire	neure
I pl	gu	geu	gure	geure
II sg	hi	heu	hire	heure
II pol	zu	zeu	zure	zeure
II PLUR	zuek	zeuek	zuen	zeuen
III sg	hura	bera	haren	bere
III pl	haiek	berak	heien	beren

The pronouns are strictly personal only in the 1st and 2nd persons; 3rd p. pronouns as such do not exist, but any demonstrative (with the distal *hura* as the default) or the

"emphatic" *bera* can be used to refer to human beings.

There are no reflexive or reciprocal pronouns: reflexivity is expressed either by a medio-passive form, or by a full NP consisting of a possessive (in the genitive) followed by *buru-a* lit. 'head-sg' which always functions in verbal cross-reference like a 3rd p. item; likewise, reciprocity can be expressed either by the medio-passive voice or by the item *elkar* which, even when its antecedent is 1st or 2nd p., is cross-referenced as 3rd p. sg.

1.2.3. Verbal morphology

1.2.3.1. Synthetic and periphrastic conjugations. A handful of verbs can be directly inflected, but most of them only surface as participles (or a suffixless radical form), the tense and agreement affixes being carried by an auxiliary that *follows* the participle (in positive sentences – see 1.3.1 on negation).

Transitive verbs are always associated with the aux. °*edun* (a reconstructed form), 'have', whereas intransitive (unaccusative) verbs are followed by *izan* 'be': contrary to what happens in a language like English, there is thus no association between a specific aux. and a specific participle (compare [*have*+*-en, be*+*-ing*]): the participles, which vary for aspect, select the aux. on the basis of their own argument structure.

There are three participles, perfective (or "past"), imperfective-iterative (or "present"), and prospective (or "future"). The perf. part. is also the citation form of verbs; it is characterized by various suffixes (*-i, -n*, zero, and the productive *-tu*); the prospective part. is formed by suffixing *-en* or *-ko*, i.e., one of the two genitive endings, to the perf. part; the impf. part. is obtained by dropping the perf. part. suffix and adding *-t(z)en* – in fact, this ending consists of *-t(z)e*, the gerundive suffix, followed by the loc. case ending *-n*. There are three basic "tenses", two of which correspond to time reference (present and past), and one to *irrealis* modality: the prefixes when all arguments are 3rd p. are respectively *d-, z-* and *l-* (*irrealis* forms come in pairs, one to be found in the protasis of conditional sentences, the other in the apodosis). The basic combinations appear in table (F):

Table F *Basic participle + auxiliary combinations*

	perf part.	*imperf part.*	*prosp. part.*
present	erori da 'he has fallen'	erortzen da 'he falls'	eroriko da 'he'll fall'
	irakurri du 'he has read'	irakurtzen du 'he reads'	irakurriko du 'he'll read'
past	erori zen he fell / had fallen'	erortzen zen 'he fell/used to fall'	eroriko zen 'he was about to fall / would have fallen'
	irakurri zuen 'he read/had read'	irakurtzen zuen 'he read/used to read'	irakurriko zuen ha was about to read / would have read'
irrealis *(hypo-thetical)*	erori ba.litz 'if he'd fallen'	erortzen ba.litz 'if he fell'	eroriko ba.litz 'if he fell'
	irakurri ba.lu 'if he'd read'	irakurtzen ba.lu 'if he read'	irakurriko ba.lu 'if he read'
irrealis *(conse-quential)*	erori litza(te)ke (no longer used)	erortzen litza(te)ke 'he would fall'	eroriko litza(te)ke 'he would fall'
	irakurri luke (no longer used)	irakurtzen luke 'he would read'	irakurriko luke 'he would read'

When a verb is directly inflected, one of its values is "progressive"; the opposition between the synthetic and the periphrastic conjugations is therefore the reverse of the English one:

(8) a etortzen da 'he comes'
b dator 'he's coming'

The other value of directly inflected verbs is aoristic; their periphrastic counterpart consists of the verb *root* followed by a suppletive aux.; these forms correspond to the subjunctive mood: present (Gi) or past (Gii), the imperative (iii) with the prefix *b-*, and yet another series of *irrealis* forms which may take on an additional potential value:

Table G *Tense and modality in the aoristic conjugation*

i	eror dadin	'that he fall'	irakur dezan	'that he eat'
ii	eror zedin	'that he fall'	irakur zezan	'that he read'
iii	eror bedi	'let him fall'	irakur beza	'let him read'
iv	eror ba.dadi	'if he falls'	irakur ba.deza	'if he reads'
v	eror daiteke	'he can fall'	irakur dezake	'he can read'
vi	eror baledi	'if he fell'	irakur baleza	'if he read'
vii	eror laiteke	'he could fall'	irakur lezake	'he could read'

The forms in (F) and (G) can be unified if we postulate the existence of a zero aspectual suffix: when the conjugation is periphrastic, then, the choice will be between one of the four options below:

Table H		*intransitive*		*transitive*	
	a	eror.Ø	-adi-	irakur(r).Ø	-eza-
	b	eror.i	-a-	irakurr.i	-u-
	c	eror.tzen	-a-	irakur.tzen	-u-
	d	eror.iko	-a-	irakurr.iko	-u-

Since the past tense is marked by the suffix *-(e)n*, and since erg. agreement is marked by suffixes (but see 1.2.3.2.3), the four options in (H) and the synthetic conjugation can be conflated as in (9), where the brackets represent an option: both an aspectual suffix and an aux. will either be selected together, or left out.

(9) VC => root (aspect+aux) Infl

The suffix *-ke* which appears in (Gv,vii) and in the consequential *irrealis* forms of Table (F) has a life of its own, because it can in fact be combined with *any* inflected form, as shown in (10):

(10) a irakurri du*ke* 'he must have read'
b irakurtzen du*ke* 'he must read' [probability]
c irakurriko zu*ke*en 'he would (probably) have read'

Now this *-ke-*, although suffixed to the inflected verb form (IVF), is *followed* by the erg. and tense morphemes; hence (11), a revision of (9):

(11) VC => root (aspect+aux)(*-ke-*) Infl

The aux. may be iterated, with a past-in-the-past or iterative value (as (b) shows, the aspectual choices are marked on the aux., not the lexical verb):

(12) a irakurri izan du 'he had read / he used to read'
b irakurri izango zuen 'he would have read'

The final formula must therefore be:

(13) VC => root ((perf.+aux) aspect+aux)(*-ke-*) Infl

To derive the real forms, one must select the obligatory morphemes (root and Infl), and may choose the optional ones. The morphological adjustments are minimal: any affix selected must be suffixed to the verb to its left. Thus, if the inner optional [perf.+aux] sequence is not selected, but the rest is, we get (14a) from the general formula, and (b) as the morphological output (where square brackets indicate word boundaries – see the next section for the missing prefix):

(14) a root aspect+aux *ke* Infl
a' irakur- tzen+-*u*- *ke* Infl
b [irakur-tzen] [...-*u*-*ke*-Infl]

If we compare (13) to Chomsky's famous formula for English verb complexes, slightly adapted here:

(15) VC => Infl (Mod) (aux+*en*) (aux+*ing*) root

we note that, other things being equal (English has a whole class of modal auxiliaries, whereas Basque only has one modal affix, etc.), the two formulae are linearly mirror images of each other, but exhibit exactly the same relative distance of the main grammatical morphemes with respect to the lexical verb: the aspectual morphemes are closest to it, followed by modality material, and finally, farthest from the verb, we find Infl, thereby exemplifying the *same* hierarchy. (A more general study, which also takes the data described in the next section into account, can be found in Rebuschi (1999)).

1.2.3.2. The polypersonal conjugation

1.2.3.2.1. The IVFs agree with up to three arguments: absolutive, ergative and dative. Some examples are given below, in the present tense; the roots are in italics; comments follow.

(16) a na.iz 'I am'
I-*iz*
b za.to.z (<za.tor.z) 'you [polite] are coming'
you-*tor*-pl
c ga.u.de (<ga.go.de) 'we are (somewhere)/we are staying'
we-*go*-pl
d joan na.tzai.o 'I have gone to him'
gone I-*tzai*-him
e etorri ga.tzai.z.ki.zu 'we have come to you'
come we-*tzai*-pl.DF.you

(17) a ikusi za.it.u.t 'I have seen you'
seen you-pl-*u*-I
b ikusi ga.it.u.zu 'you have seen us'
seen us-pl-*u*-you
c eman d.i.da.zue 'you have given it to me'
given it-*i*-me-you
d saldu d.i.z.ki.zu.gu 'we have sold them to you'
sold it-*i*-pl.DF.you.we

The prefixes are absolutive: they correspond to the unique argument of the intransitive verb and the direct object of a transitive one. The abs. pl. morpheme (present with the polite 2nd pers. forms, which are in fact referentially sg. today) is distinct from the personal affix; it follows the root (16b,c,e), (17d), except in the case of the transitive

aux. *-u-* ,(17a,b).

The suffixes are dative and ergative in that order, if both occur; note the dative flag (DF) *-ki-* which precedes the dative personal suffix (it could be argued that it is always present, but regularly disappears for morpho-phonological reasons in (16d) or (17c), cf. Rebuschi 1984).

Ergative and dative plurality is not marked for the 1st p. pl. or for the 2nd p. polite, historically a plural. Two suffixes (*-e*, *-te*) mark plurality for the 3rd p. (Ø for erg. sg., *-o* for dative sg.): *du* 'he has it' / *dute* 'they have it', *natzaio* 'I am to him', *natzaie* 'I am to them', and are also used to build a general 2nd p. plural form out of the 2nd p. polite form: *zatoz* 'you (pol.) are coming', *zatoz.te* 'you (pl.) are coming'.

As was said earlier, nouns and pronouns are not marked for gender. However, the familiar 2nd p. sg. *suffixes* display a gender difference; the four translations of ergative "you" are given in (19) – where the first three lines correspond to a single addressee:

(18)	a	ikusi na.u.k	'you (male) have seen me'	[familiar]
	b	ikusi na.u.n	'you (female) have seen me'	[familiar]
	c	ikusi na.u.zu	'you (polite) have seen me'	[polite]
	d	ikusi na.u.zu.e	'you (neutral) have seen me'	[plural]

1.2.3.2.2. The DF *-ki-* is not functionally necessary, as the aux. roots are distinct for each case frame:

(19) Case frames and aux. roots

a	absolutive:	*-iz, -ra*	c	abs.+ergative:	*-u-*
b	abs+dative:	*-(t)zai*	d	abs.+dat.+erg:	*-i-*

Moreover, synthetically inflected transitive verbs resist valency reduction (the "anti-causative transformation"): *dakar* for instance is always 's/he's carrying it/him/her', *never* 's/he/it is (being) carried', in spite of the absence of a visible ergative suffix, and contrary to what one may observe in the periphrastic conjugation, where *ekartzen du* 's/he carries it/him/her' may reduce to medio-passive *ekartzen da* 's/he/it is carried'. Likewise, unaccusative verbs may not be synthetically inflected and undergo a causative transformation: *joan da* 'he's gone' has a causative counterpart (*joan du* 'he took it away'), but the strong form *doa* does not. Thus, even though the erg. and dat. suffixes are identical in the 1st and 2nd p., both sg. and pl. (see table I), the lexical or aux. verbal root always indicates whether they are to be interpreted as dative or ergative:

Table I. *Dative and ergative suffixes*		
	dative	ergative
1sg	-t / -da-	
2sg masc.	-k / -(k)a-	
2sg fem.	-n / -na-	
3sg	-o	-Ø
1pl	-gu(-…)	
2pol	-zu(-…)	
2pl	-zu.e(-…)	
3pl	-e [=-Ø-e?]	-Ø-te

1.2.3.2.3. In the past and *irrealis* tenses, the form of the 1st and 2nd p. prefixes changes, somehow incorporating a copy of the past suffix *-n*:

Table J	*Present abs. prefixes*	*Past/irrealis abs. prefixes*
1sg	na-	nind-
2sg (familiar)	ha-	hind-
1pl	ga-	gind- / gint-u-
2 polite	za-	zind- / zint-u-
2 plural	za-…-e/te	zind- /zint-u- … -e/te

When all the arguments are 3rd. p., the following four affixes contrast (thereby indicating that they are *not* personal prefixes, as discussed in Rebuschi [1983] 1997: 139-154, among others):

(20) a d- (present) c l- (*irrealis*)
b z- (past) d b- (imperative)

However, when the abs. argument is 3rd p., but the erg. one is 1st or 2nd p., this latter is cross-referenced by a *prefix* (but the pl. morpheme of the abs. is retained):

Table K	*Ergative prefixes (past &* irrealis*)*
1st sg.	n(e)-
1st pl.	gen(e)-
2nd p. polite	zen(e)-
2nd p. pl.	zen-(e) … -te

1.2.3.2.4. Basque IVF may contain yet another personal affix, the so-called "allocutive" one. Examples are given in (21):

(21) a ikusi di.*na*.t 'I've seen it/him' [*addressing a woman*]
b jan di.*(k)a*.gu 'we've eaten it' [*addressing a man*]
c etorri zitzaida.*(k)a*.n 'he came to me' [*addressing a man*]
d emango zizkio.*na*.gu 'we'll give it to him' [*addressing a woman*]

Such forms, which contain a post-dative 2nd p. sg. suffix, are automatically used if the register is familiar rather than polite, and the hearer is *not* the referent of an argument of the lexical verb, a choice which is conspicuous when *familiar* 2nd p. pronouns or IVFs are used elsewhere in the utterance:

(22) a Zu etortzen bazara, ni joango naiz [*polite*]
b Hi etortzen bahaiz, ni joango nauK [*familiar: male hearer*]
b' *Hi etortzen bahaiz, ni joango naiz
c Uste diNAt nekatua haizela [*familiar: female hearer*]
c' *Uste dut nekatua haizela

(b') is out because of a clash in register, made visible by the constrast between the familiar IVF of the protasis and the polite form in the apodosis; likewise, (c') is ill-formed because the matrix IVF is [-alloc.], hence [-fam.], whereas the subordinate form is [+fam.].

In some cases, the same form may be interpreted either as a normal transitive or ditransitive aux. with a 2nd p. sg. erg. or dat. suffix, or as the allocutive counterpart of an intransitive or monotransitive IVF:

(23) a hilko nauk
(i) you will kill me [-alloc: erg. [+fam, +masc]]
(ii) I will die [+alloc, +masc]
b eman dinat
(i) I've given it to you [-alloc: dative [+fam, +fem]]
(ii) I've given it (away) [+alloc, +fem]

Besides, allocutive IVFs are restricted to root sentences; thus, (24a) is unacceptable because the subordinate form *nau(k)ala* can receive neither a transitive interpretation (due to the argument structure of *etorri* 'come') nor an allocutive interpretation, whilst (24b) is acceptable, since the IVF in the complement clause contains an affix that can be read as cross-referencing an argument (the suffix *-(e)la* marks subordination):

(24) a *esan di.*(K)A*.t [etorriko nau.*(k)a*.la]
a' esan di.*(K)A*.t etorriko naizela
(i) 'I've told you [+fam, +masc] that I will come'
(ii) 'I've said [+alloc, +masc] that I will come'
b esan di.*(K)A*.t [hilko nau.*(k)a*.la]
(i) 'I've told you/I've said that you'll kill me'
(ii) *'I've told you/I've said that I'll die'

The examples above illustrate the replacement of one aux. root by another – the one that

corresponds to increased addicity. When such a process is not possible (i.e. with the tri-personal aux. and with synthetic non-auxiliary verbs), a general process of pre-radical *-i-* insertion, sometimes leading to spirantisation, takes place: see (59b) in 3.2.1.3.

To conclude this section, it must be underlined that many traditional descriptions of the morpho-syntax of allocutivity have been marred by the syncretism between these forms proper and what has been dubbed "implicative" forms in the present tense in root clauses (Rebuschi 1984): these latter can be described as the use of forms of 'have' with an 2nd p. ergative agreement morpheme instead of the copula. What distinguishes implicative constructions from allocutive forms is (i) that the former can be used in embedded clauses, (ii) that they are always restricted to stative predicates, and (iii) that polite and plural suffixes are possible, as in (25):

(25) ederra dun / dun / duzu / duzue 'it is beautiful', lit. 'you have it beautful'

1.3. Aspects of Basque syntax

1.3.1. Word order. Greenberg (1966) classified B as an SOV language, but it has also been characterized as a free word order (or non-configurational) language, and as a Verb-second or (Topic-)Focus-Inflected Verb language. There is some truth in each of those characterisations, depending on the viewpoint adopted. Consider a transitive sentence with an ergative subject, an absolutive object and a transitive verb. The six word orders which are theoretically possible (SVO, SOV, OVS, OSV, VOS and VSO) are all empirically attested, but they do not have the same status.

Out of context, or as a reply to *Zer gertatu zen?* 'What happened?', a normal SOV sentence would be used in the answer:

(26) Jonek Miren jo zuen 'Jon hit Miren'
Jon.erg Miren.abs hit aux

However, totally rhematic sentences are not the norm: animate subjects tend to be discourse topics, leaving the sole VP as rhematic. Furthermore, the object left-adjacent to the VC can, but need not, be interpreted as constrastively focused; thus, a written sentence like (26) may also be interpreted with the S as a topic, and the whole VP as rhematic (in which case the prosodic prominence of the O relative to the VC is weak), or with the sole O as rhematic (the deaccenting of the VC being more marked: Elordieta 1997). But when it is the subject that is adjacent to the VC, it is normally interpreted as focused. Thus, if there is a focused NP, it must immediately precede the VC, but if there is none, the O-VC order obtains. A particular subtype of focus consists of interrogative pronouns: they occupy the focus position, just as the elements provided by the answer do; when there are several *wh-* phrases in the question, the XPs in the reply will

necessarily occupy the same positions, as in:

(27) a Nork zer erosi du?
who.erg what.abs bought has
'Who has bought what?'
b Jonek liburua erosi du, eta Mirenek arrosak (erosi ditu)
Jon.erg book.sg bought has and Miren.erg roses bought aux
'Jon bought a/the book and Miren bought (the) roses'

Last but not least, negation in root sentences triggers a different word order: the negative particle *ez* immediately precedes the auxiliary, which now precedes the main verb (and possibly intermediary material):

(28) Jonek ez du Miren jo 'Jon has not hit Miren'
Jon.erg neg has Miren-*Ø* hit

As explained in de Rijk 1996, a sentence like this has two interpretations: it can be an reply to 'Who hasn't hit Miren?' (*Jonek* being the focus), or the denial of the focus of a preceding assertion like (26) in which *Miren* would have been the focus.

1.3.2. Up to now, ergativity has been taken for granted. It must be emphasized, though, that the facts are not crystal-clear. In Dixon's 1994 model, for instance, there are three syntactic primitives, S, A and O: S is the unique argument of an intransitive verb, A is the agentive subject of a transitive verb, and O its object. In such a framework, if we take into consideration the fact that *Miren* has a zero suffix in (26) above just as in (29) below, but has the *-k* suffix in (27b), we have detected ergative morphology in Basque.

(29) Miren joan/etorri da 'Miren has gone/come'
Miren.Ø gone/come is

But is the category "S" basic? In the late seventies, Permutter showed that there were two types of (superficially) intransitive verbs, which he (somewhat unfortunately) labelled "unaccusative" and "unergative". The verbs in (29), but also *erori* 'fall' or *hil* 'die' belong to the first category: cross-linguistically, they allow 'be' auxiliaries (vs. 'have' for unergatives, as in French or Italian), and impersonal constructions of the type *there came two men* in English or *il est venu deux personnes* in French, etc. Now, Basque has a curious property: almost all the items that are traditionally used as instances of intransitive verbs in the description of its ergativity really are unaccusative ones. Moreover, most of its few unergative verbs are borrowings, and their only argument is in the so-called ergative case, the auxiliary being 'have' rather than 'be', as shown in (30); finally, most of the unergative verbs found in other languages must be translated into a complex expression, with the verb *egin* 'do, make' associated with a determinerless noun, as in (31).

(30) a Mirenek kantatu du 'Miren has sung'

Miren-*k* sung has
a' *Miren kantatu da
b gerlak luzaro iraun zuen 'the war lasted for a long time'
war.sg.-*k* long lasted aux[+tr]
b' *Gerla luzaro iraun zen

(31) a Jonek lan egiten du 'Jon works'
Jon-*k* work doing has
b Mirenek eztul egin du 'Miren has coughed'
Miren-*k* cough.Ø done has

Since using auxiliary selection might lead to circularity, another test can be used that shows that the "erg" and "abs" Ss are syntactically different: it is possible to negatively quantify over "abs" Ss by using the partitive ending *-(r)ik*, but it is not possible to do so with "erg" Ss, as shown in (32) (Levin 1989).

(32) a ez da gizonik etorri 'no man has come'
neg. is man-*ik* come
b *ez du/da gizonik kantatu
neg. has/is man-*ik* sung

Fillmore's Case Grammar does not fare any better; here, there are no Ss: the language is ergative if the A associated with a bivalent verb receives a special treatment, the remaining O, just as an A or an O used alone, being morphologically identical. What the handful of "unergative" subjects in B show is plainly corroborated in this model: when there is no explicit object, as in (33), the sentences are ambiguous, since the object can be either understood as definite (contextually given) or really implicit, whereas the sentences in (34) are ungrammatical:

(33) a Jonek jan du 'Jon has eaten it / Jon has eaten'
Jon.erg eaten has
b Mirenek irakurtzen zuen 'Miren was reading it / Miren was reading'
Miren.erg reading has

(34) a *Jon jan da b *Miren irakurtzen zen
Jon.Ø eaten is Miren.Ø reading was

The case-marking and verbal agreement of A arguments remain the same, independently of the presence of an O argument: *in Fillmore's terms, Basque would then be an "active" language, not an "ergative" one.*

The only argument that remains to stick to the "erg/abs" labels (which I will do, for readability purposes) is to consider that the suffixless argument is unmarked, since it is obligatorily represented in the IVF, whereas the argument that carries a *-k* suffix is marked, because it is not compulsory, as indicated by the fact that an implicit, archetypal A is not projected, leading to medio-passive readings:

(35) a ogia jaten da 'bread is eaten'
bread.Ø eating is
b liburuak irakurtzen ziren 'one used to read books'
books.pl.Ø reading were

1.3.3. Subject/object (a-)symmetries. Whatever the proper characterization of the Ø/*-k* opposition in B, its morphology is not nominative/accusative. But it remains to be seen if its morphology determines its syntactic properties or not (semantically, recall the (in-)definiteness data concerning the sg. ending *-a* mentioned in 1.2.2).

1.3.3.1. First, abs. arguments do not have the same interpretation when they are used in isolation and when they are syntagmatically opposed to a erg one. We just saw with (33) that a silent object could be interpreted either anaphorically or archetypally (one eats edible stuff, or reads readable things). But what happens if there is no A, i.e. with unaccusative verbs? In root sentences, the silent unique argument must be interpreted anaphorically:

(36) a sartzen da 's/he comes in'/*'one comes in'
entering is
b erori da's/he has fallen'/*'someone has fallen'
fallen is

A phonetically empty 3rd p. sg. A is also anaphoric:

(37) a ogia jaten du
bread-*Ø* eating has
's/he eats bread'/*'one is eating bread' (compare (35a))
b liburua irakurri zuen
book-*Ø* read had
's/he read a book'/*'a/the book was read'

1.3.3.2. In control structures, abs. arguments of intransitive verbs behave like erg., rather than abs., arguments of transitive verbs. If the syntax were ergative, one would expect the abs. argument of *jaten* below to behave like the abs. argument of *joaten* 'going', but the data indicate otherwise: the argument which is "understood" in the lower clause (*PRO*) corresponds to the subject of nominative-accusative languages:

(38) a Nik haurrak ikusi ditut [*PRO* joaten]
I-*k* children-*Ø* seen I-have-them going
'I saw the children going'
b Nik haurrak ikusi ditut [*PRO* sagarra jaten]
I-*k* children-*Ø* seen I-have-them apple.sg.-*Ø* eating
'I saw the children eating an/the apple'
c *Nik sagarrak ikusi ditut [*PRO* haurrek jaten]
I-*k* apples-*Ø* seen I-have-them children-*k* eating

1.3.3.3. Another argument that may be given in favour of morphologically-independent subjects and objects is provided by the distribution of *elkar* 'each other'. On the one hand, *elkar* may never be in the erg. case; on the other hand, it can be in the abs. case, but only when the V has a direct object; thus, if *elkar* is in the abs. case, its antecedent must be erg., and cannot be dative (for instance):

(39) a haurrek elkar ikusi dute
children-*k elkar-Ø* seen they-have
'the children have seen each other'
b *elkarrek haurrak / haurrak elkarrek ikusi ditu
elkar-k children-*Ø*
c haurrak elkarri mintzatu zaizkio
children-*Ø elkar*-dat. spoken they-are-to-him
'the children have spoken to each other'
d *elkar haurrei / haurrei elkar mintzatu zaie
elkar-Ø children-dat.

1.3.3.4. B does not, however, exhibit as many subject/object asymmetries as nominative-accusative languages. I will only consider one example here (see Rebuschi 1989). Standard B as it is promoted by the Basque Academy has two series of genitive pronouns, which, for want of a better word, I shall call "weak" and "strong", because in the unmarked series, there are no diphthongs, whereas the marked series exhibits diphthongs – see table (E). This system, and the specific distribution of the two genitive forms, are borrowed from classical La texts. Roughly speaking, the weak forms are used when there is no coreferent pronoun or NP in the same clause: there may, but need not, be such an antecedent in a higher clause; the strong forms are "reflexive", but in a highly specific way: *any* co-argument of the NP that contains a genitive pronoun will serve as an antecedent and thus trigger or licence the strong form. Examples in (40) summarize the basic facts:

(40) a Nik neure / *nire liburua irakurri dut
I-*k* my[+str]/[-str] book.sg read I-have
'I've read my book'
b Nik uste dut [Patxik nire / *neure liburua irakurri duela]
I-*k* belief I-have P.-*k* my[-str.]/[+str] book read he-has-that
'I think that Patxi has read my book'
c Bere anaiak Iosephi mintzatu [zitzaizkionean]
his[+str] brothers-*Ø* Joseph.dat spoken aux.comp.loc
'When his$_i$ brothers talked to Joseph$_i$…' (Axular 1643)
d Elisabet [h]eure emaztea erdiren zaik seme batez.
E. thy[+str] wife-*Ø* bear[intr.].fut. aux-to-you son one-instr
(Luke 1:13, Liçarrague 1571)
'Thy$_i$ wife Elizabeth shall bear thee$_i$ a son.'
d' Zeure emazte Elisabetek semea emango dizu (*id.*, BEB 1994)

your[+str] wife E.-*k* son.sg-*Ø* give-fut s/he-has-to-you
'Your$_i$ wife Elizabeth will give you$_i$ a son'

From a typological viewpoint, the locality constraint exemplified in (40b) is well-attested, but the utter lack of functional yield in this system when 1st and 2nd persons are concerned must be noticed. More intriguing, though, is the case illustrated in (c), due to the fact that the strong genitive pronoun is within the *subject* NP, and linked to a non-subject argument. The (d,d') examples have been added that show that the same lack of argument asymmetry that existed in the Northern dialects four centuries ago indeed represents the norm today in carefully written EB.

1.3.4. The complex sentence. B has three distinct complementizers, which are all affixed to the finite verb form: the prefix *bait-*, and the suffixes *-(e)n* and *-(e)la*, and, as we saw in section 1.2.2, there are tense, mood and aspect oppositions both in matrix and in embedded clauses; besides, B also uses quite a few distinct forms of tenseless clauses.

1.3.4.1. Finite subordinate clauses. The (obligatory) conplementizer *-(e)la* is typical of clauses governed by verbs of saying; the indicative forms are factual, and the subjunctive forms indicate a (mental) goal:

(41) a Patxik esan dit [Jon joan dela]
Patxi-*k* said he-has-to-me Jon-*Ø* gone he-is-*la*
'Patxi has told me that Jon has gone'
b Mirenek agindu du [kanta dezadala]
Miren-*(e)k* ordered has sing I-do[subjn]-*la*
'Miren has ordered me to sing' *lit.* '…that I sing'

Complement clauses governed by a predicate of wish and indirect questions have *-(e)n*:

(42) a Patxik nahi du [Miren joan dadin]
Patxi-*k* want has M.-*Ø* go[root] is[subjn]
'Patxi wants Miren to go' *lit.* '…that she go'
b Patxik galdatzen du [nor joango den]
Patxi-*k* asking he-has who go.fut is-*(e)n*
'Patxi asks who will leave'

-(e)n is also used in relative clauses, which precede their antecedents:

(43) a joan de*n* gizona 'the man who has left'
gone is-comp man.sg
b irakurtzen duzu*n* liburua 'the book that you're reading'
reading you-have-*n* book.sg

When the head noun of the NP is absent, free relatives are created:

(44) a joan den.Ø.a [cp. (43a)]
'the one that≈he who has left'/'what has gone'
b irakurtzen duzun.Ø.a [cp. (43b)]

'the one you read / what you read'

Such headless relative clauses are used to build adverbial clauses:

(45) a joan den.etik [cp. (44a)]
gone he-is-*n*.abl
'(ever) since he left' [also: 'from the one that's gone']
b irakurtzen duzun.ean [cp. (44b)]
reading you-have-*n*.loc
'when you read' [also: 'in the one you read']

When the matrix clause is negative, *-(e)n* plus the partitive suffix *-ik* are used instead of *-(e)la*:

(46) Ez dut uste [joan de*nik*] [cp. (41a)]
neg I-have belief gone s/he-is-*n-ik*
'I don't think s/he's left'

The third complementizer, *bait-*, is mainly causal in EB as it is used in Spain, and thus serves as a paraphrase for a construction of a new type, in which it is now the first comp., *-(e)la*, which carries a suffix:

(47) a joan baita [= bait-da]
gone *bait*-aux 'because s/he's left'
b joan delako
gone aux-*la.-ko* *id.*

Let us also recall the existence of the prefix *ba-* 'if', used in the protasis of conditional sentences:

(48) Zu etortzen ba zara, ni joango naiz 'If you come, I'll go'
you coming if you-are I go-fut I-am

1.3.4.2. Non-finite subordinate clauses. Both the perf. part. and the gerund can head tenseless subordinate clauses. The former is used with no suffix as a complement of *behar °edun* 'have to' and *nahi °edun* 'want', but it can also receive various case suffixes or be followed by postpositions, in which case the clause that contains it is typically adverbial:

(49) a joan behar dut 'I must go'
go *behar* I-have
b irakurri nahi du 's/he wants to read'
read *nahi* s/he-has
c joanez gero 'after leaving'
gone.instr then
d irakurria gatik 'in spite of reading (it)'
read-sg *ga*-from

The gerund is obtained by deleting the final *-n* of impf. participles. It can also be

followed by case endings or postpositions:

(50) a irakurtze.ra joango da
reading.all go.fut he-is
'he will go and/to read' [complement clause]
b irakurtze.a.n / irakurtze.a.rekin
reading.sg.loc / reading.sg.comit.
'while reading' [adjunct clause]
c irakurtzeko
'in order to read' [complement or adjunct clause]
d irakurtzea gatik / irakurtze arren
'in order to read' [adjunct clause]

The most interesting feature of those structures is that *the subject can be explicit.* This is not, however, an instance of "exceptional case marking" by the matrix verb, since the subject's case is determined by the lower predicate; furthermore, as shown by (c) (from Ortiz de Urbina 1989), the subject of a subject clause may also be explicit:

(51) a Jonek erabaki du [Mirenek irakurtzea]
Jon.erg decided he-has Miren.erg reading.sg
'Jon has decided that Miren should read'
b Jonek erabaki du [Miren etortzea]
Jon.erg decided he-has Miren-*Ø* coming.sg
'Jon has decided that Miren should come'
c [semeak atzerrian ibiltze.a.k] kezkatzen du ama
sons-*Ø* abroad walking.sg.-*k* worrying it-has mother.sg.-*Ø*
'it worries (their) mother that (her) sons are abroad'

1.3.5. Other aspects of Basque syntax

1.3.5.1. Basque has many properties typical of *head-final* languages, i.e. its case-assigning heads generally follow the elements they govern: as we saw in 1.3.1, objects often precede verbs; more systematically, Basque has postpositions (which are often declined nouns), but no prepositions:

(52) a ni baitan 'in me'
b zu atzean 'behind you'
c ohearen azpian 'under the bed'
d gerla aurretik 'from before the war'

The internal structure of NPs is typically head final, since genitive complements and adjuncts, ordinal and cardinal numbers (except *bat* 'one'), and relative clauses, precede the head noun. However, attributive adjectives (except geographical/ethnic ones, which appear on either side of the noun) and determiners *follow* the head:

(53) a Donostiako nire hiru lagun kutunekin
San Sebastian-*ko* my three friend dear.pl.with
'with my three dear friends in/from San Sebastian'

b [zuk irakurri zenuen] haren liburu gotor hau
you.erg read you-had-*n* his book thick this
'this thick book of his which you read yesterday'

1.3.5.2. Case duplication. Three subcases must be distinguished. First, when a quantifier follows a demonstrative, both are number+case-marked:

(54) liburu heiekin guztiekin 'with all those books'
book dem.pl.with. all.pl.with

Second, as we saw in 1.2.2.2, some case endings are complex: the prolative ending *-entzat* may be analysed as the concatenation of two distinct case endings, the genitive and the (simple) prolative but it is also possible to regard *-tzat* as a postposition that cliticized to the preceding word at the phonological level.

Third, what has been dubbed "superdeclension", as in *gizonarentzakoaz*, simply does not exist: such words must not be analysed as in (55b) below, which supposes that the head of the expression is *gizon* 'man', followed by a string of concatenated suffixes, but as in (c), i.e. as a complex NP with a null head, and a postpositional phrase transformed into a adnominal adjunct by *-ko*:

(55) a gizonarentzakoaz
b *[[[[[gizon-a]-(r)en-tzat]-ko]-a]-z]
c [NP [ADJ.P [PP gizon.a.(r)en.tzat]-ko].Ø.-a]-z
man.sg. gen. for *-ko-* sg. about
'about the one (*-Ø-*) for the man'

2. Phonetic & phonological variation

2.1. Segments

2.1.1. Michelena (1961, 1977), using Aquitanian data, internal reconstruction and the phonetic shape of the borrowings from Vulgar Latin and, later, from Romance languages, has been able to reconstruct the B phonological system at the beginning of this era. The vowels seem to have been identical to those used in EB today. The consonants were possibly limited to 16, with a fortis-lenis opposition: *(p)/b, t/d, k/g, tz/z, ts/s, N/n, L/l* and *R/r*, later reinterpreted as a voiceless/voiced opposition for the plosives. Somewhat later, lenis *l* underwent rhotacisation and lenis *n* was lost (compare *zeru* 'sky' < *cœlu(m)*, *ahate* 'duck' < *anate(m)*), the fortis segments being the ancestors of today's *l* and *n*. Since this opposition only existed intervocalically, and since "geminate sequences of Latin are invariably borrowed as fortis consonants", the fortis segments may well have been simple geminates, thereby reducing the "Pre-Basque consonant

system to the surprisingly small number of only eight" units (Trask 1997); note however that the status of initial *h* in words like *herri* 'people/country/village' remains unclear under these reconstructions.

2.1.2. The only dialect that has six vowels is Souletin (and the adjoining LN varieties): it has a closed front rounded vowel *ü* (<*u*: this process has taken place everywhere except before *-s*, flapped *-r*, and the sequence *-nk-*); this *ü* has even evolved into *i* in certain contexts, as in *die* <*düe*, EB *dute* 'they have (it)'. Zu also has nasalized simple vowels and diphthongs, as in *ãhãte*, 'duck', and *ardãuÙ* 'wine'.

Zu and LN also have (and La had) aspirated consonants; the aspirated plosives seem to be a reflex of a former stress assigned to the first or second syllable (no plosive being aspirated in a later syllable). Initial *h* has phonemic status in all the varieties of B spoken in France except coastal La (where it has been lost): *hari* 'to-him', *ari* 'busy'.

As noted in 1.1.1.2, the orthographic letter *j* has several realisations, among which, alonside those mentioned there, that of a voiced prepalatal continuant [J] (Zu, Bi). Note also that intervocalic voiced plosives and flapped *r* are very easily dropped.

2.2. Rules

2.2.1. A very general process across dialects consists in closing the mid vowels when they are followed by open *a*; thus, *etxe-a* 'the house' is often realised like *etxia*, and *etxeko-a* 'the (one) from/in the house' as *etxekua*.

2.2.2. Quite common in the southern dialects is the *progressive* semi-closure of *a* into *-e* if it follows a closed vowel (with any number of intervening consonants): *ur-a* 'the water' is thus pronounced *ure* and *hil-a*, 'the dead (one), *(h)ile*. The foregoing rules apply, according to the sub-dialects, either in bleeding order, or in feeding order; in the latter case, *etxekoa* is pronounced *etxekue*.

2.2.3. In Gi and Bi again, *l, n, t* and *d* are palatalised after *i*. This process is "word-internal", but it must be remembered that some words, e.g. the copula or intr. aux. *da*, cliticize to the preceding one; as a consequence, the written sequence *hil da* 'he's died' can be pronounced *illdde*. This rule, which was operative in 17th and 18th c. La (since words like *zein* 'which' were spelt *zeiñ*, etc.), has been lost in this dialect, perhaps due to the influence of neighbouring LN, which does not seem to have ever had it, and which has clearly been gaining in importance in the last two centuries.

2.2.4. A typically western (Bi) process is the introduction of [ß] (written *b*) between the closed back vowel *u* and a following vowel, and of a [J] or a [S] (spelt *j* and *x*) between an *i* and another vowel; those continuants are obvious evolutions of intervocalic glides, giving rise to realisations like [bejSa] for written *behia* 'a/the cow'. Another feature of

spoken Bi is the absence of opposition between *(t)z* and *(t)s*: the remaining fricative is *s* and the surviving affricate is *tz*.

2.2.5. At least one process is becoming obsolete everywhere (probably partly because of the standard spelling, and partly because of the absence of such processes in Romance): it is the progressive assimilation in voicelessness which has apparently always applied between the negative particle *ez* [es], and a following IVF beginning in *d-*, as in *ez da*; this sequence, traditionally spelled *ezta*, phonetically [esta], is more and more often heard [ezda]. Another trend among young southern speakers is to neutralise of the *ts* ≈ *tx* opposition with the former segment being the general realization; moreover, in the subdialects which palatalise *t* after *i*, the younger generation quite often substitute *tx* [tS] for *tt* [t'].

2.3. Accentuation

There are basically two main accentual types, one a stress system, the other a pitch system. The former type has two basic subtypes, stress being assigned either from right to left, as in the eastern part of the B Country (and some parts of Biscaye), or from left to right, as in its central part (Guipuzcoa); since an example of [+2] rightward stress assignment was given in 1.1.2.1, I will concentrate here on the other two options.

2.3.1. Souletin B provides the best studied case of leftward stress assignment (Michelena 1957-58, Hualde 1991, 1997). The general rule is for the last syllable but one to be salient. It is useful to distinguish between uninflected simple, derived, and compound words, and their inflected counterparts. Exceptions to the general [-2] rule with simple uninflected words are oxytones: (i) words ending in diphthongs (*ardã;uÙ* 'wine'); (ii) a few lexically marked items. Derivational suffixes may be unmarked, in which case the [-2] rule applies to the new word, as in *ogéndün* 'guilty' (cp. *ógen* 'guilt'), or marked as stressed: *aitañí* 'grand-father' (cp. *áita* 'father'); proparoxytones may also be created by suffixes like *-tegi*, as in *háustegi* 'ash-pan' (cf. *hauts* 'ash'). Compound words normally have their main stress on the second one.

When words are inflected, the [-2] rule generally applies to the inflected item; from *gízun* 'man', we thus obtain *gizúna* 'the man'; likewise, 'the guilty one' will be *obendü;na.* An important fact is that sequences consisting of a root-final *-a* followed by the sg. suffix *-a* reduce to a single stressed *-á* segment – obviously as a result of the stress having been assigned before the *a*+*a* sequence is reduced; the functional yield of stress position is clear: *alhába* 'daughter', *alhabá* 'the daughter', *bi alhábak* 'two daughters [erg]', *bi alhabák* 'the two daughters'. With non-abs. pl. forms, a stressed *-é-* appears in all words: *gizunék* 'the men [erg]', *gizunér* 'to the men [dative]', *alhabék* '(the) daughters [erg]',

alhabén 'of (the) daughters'; this has been used as an argument for reconstructing non-abs. pl. endings as having been suffixed to an independent plural form, e.g. **gizun-a(g)-ek* (Michelena 1957-58), but, as Hualde 1991 insists, this *-é-* morpheme must be synchronically regarded as inherently stressed.

Roughly speaking, all the eastern dialects of table (A) and the adjacent ones (La, HN) are of the [-2] type, but with important differences, the most conspicuous of which being that *-a+a(k)* sequences are reduced to a single segment before the stress assignment rule applies outside of Zu: as a consequence, stress has no functional yield in these dialects. Note too that in Roncalese, now extinct but closely related to Zu, the [-2] rule only applied to the root, all the suffixes being extra-metrical.

2.3.2. In northern Biscayan, the situation is totally different: polysyllabic words typically start with a low tone, and the second syllable has a high tone which may, or may not, be sustained on the following syllables.

2.3.2.1. A typical example is Guernica B, where singular and indefinite nouns have a high tone throughout, whereas in the plural forms, "there is a sharp final drop in pitch from the first syllable containing a plural suffix" (Hualde 1991). Here are a few illustrations, which show that pitch has a partially distinctive function – in this section, accute accents will indicate high tone (H), and grave accents low tone (L):

(56)	*indef*	*sg*	*pl*
abs	gìxón 'man'	gìxóná	gìxónàk
erg	gìxónék	gìxónák	gìxónàk
dative	gìxónérí	gìxónárí	gìxónàrì

When the roots are monosyllabic, the words have a H tone if uninflected or in the plural, but otherwise follow the general pattern:

(57)	*indef*	*sg*	*pl*
abs	ár 'worm'	àrrá	árràk
erg	àrrák	àrrák	árràk
dative	àrrérí	àrrárí	árràrì

Guernica B then apparently contradicts the morpho-syntactic result obtained in 1.3.5.2, namely, that there is no "superdeclension": even if a plural morpheme belongs to a left adjunct, it triggers the presence of a L on every syllable following it. Thus, 'the one (Ø) of the men' [+erg], EB *[[gizonen][-Ø-a]-k]*, will have a final L (in spite of its being sg), because the adjunct to the left of the empty head is plural, and consequently has a L: *gìxónànàk*; compare: *gìxónának* 'the one of the man' and *gìxónánàk* 'the ones of the man'. The contradiction is only apparent, though: if the assignment proceeds from left to

right, the first low syllable following a high one will be the ultimate phonologically relevant one and the low tone(s) to its right is/are a simple case of L tone spreading. This dialect also has lexical exceptions: certain items inherently possess a L tone which can fall on just *any* syllable (Hualde 1997, *contra* Hualde 1991). Verbs, like nouns, are generally unmarked for stress, but, here again, exceptions do exist. What is more, the participial endings may themselves be unmarked (like the perf. endings *-i*, *-tu*), or carry a L tone (impf. *-t(z)èn*, prospective *-kò*). However, the properties of markedness and strength do not coincide; thus, in composition, an unmarked second word may erase the inherent L tone the root of the first word carries: *txístù* 'flute' is marked, but the derived *txìstúlá(r)íJ é* 'flautist' is not.

2.3.2.2. We owe Hualde/Elordieta/Elordieta 1994 a remarkable description of a broadly similar system, that of Lequeitio, which contains a lot of information on units larger than the word. The description is in terms of accent: an accented syllable has a H*L tone, the H part of which spreads lefwards as far as the second syllable of the tonal unit, which may be much larger than the word. The basic default rule is that when no item is lexically marked, the *last* syllable of the group receives this complex tone: in a sentence like *nire laguna dator*, 'my friend is coming' which contains no marked element, the accent is assigned to the very last syllable, and the H tone spreads as far as the second syllable of the first word: *nìré lágúná dátôr*.

At the word level, Lequeitio B behaves like Guernica B, except that its marked items always carry their accent on the *penultimate* syllable; likewise, the plural morpheme triggers the presence of a H*L tone on the last syllable but one, independently of its own concrete position in the word; consequently, the counterpart of the dative plural *gìxónàrì* of Guernica B (see (56) above) will be *gìxónârì*, and rightward stress displacement will also be observed when a noun is lexically marked, as in *lau libûru* 'four books [indet.]', *liburûa* 'the book', *liburuâri* 'to the book'.

Morpho-syntactic words and prosodic words do not always coincide; in particular, although verbal participles and tensed auxiliaries can be separated by various particles, and undergo linear inversion in negative sentences (cf. 1.3.1), they form one prosodic word. If focus is assigned to the preverbal constituent, and the aspectual suffix is lexically accented (as in Guernica B, this is true of imperfective and prospective endings), the preverbal constituent will contain the main stress, as in: *lagunâ etorrikó-da* 'the friend will come' or *lagûnak etorriko-díra* 'the friends will come'. But if the positive polarity of the sentence is emphasized, it is the lexical verb that carries the main accent: *lagúnak etorrîko-dira*, 'the friends *will* come'.

When whole phrases that contain no lexically marked item are focused, the prominent

accent will fall on the last syllable of this phrase, in compliance with the general default rule; in such a case, it is impossible to know, out of context, which particular word is really marked for focus; thus, in *etxe barriJâ ikusi-dot* 'I saw the new house', the contrast may be between a house (*etxe*) and another building, of between the *new* house and the *old* one. On the contrary, when there are lexically marked words in non-final position, a contrast in accent obtains, as in *libûru barríJa ekarri dot* 'I brought the new *book*' vs. *libúru barriJâ ekarri dot* 'I brought the *new* book'.

3. Morphological variation

3.1. The nominal domain

3.1.1. Determiners and pronouns. The sg. and pl. suffixes *-a(k)* seem to have evolved from a former "distal" demonstrative, as evidenced by (i) the fact that the theme for non-abs. cases of *hura* 'that (one)' is *(h)a-* (erg. *har(e)k*, gen. *haren*...), and (ii) the fact that Bi still uses *a* as a distance III demonstrative (and personal pronoun). The southern dialects: Bi, Gi (and La until the 18th century) also have an inclusive plural det. *-ok* as in *euskaldunok ...gara* '*we* Basques are…', vs. *italiarrak ...dira* 'Italians are…'. 18th and 19th century La also had a typologically very rare pronoun, *haina*, best translated by 'all such (people)', which could not take specific or definite antecedents, but rather functioned as a lexicalisation of "E-type anaphora" (Rebuschi 1998).

Bi, contrary to the other dialects, can put the demonstrative *before* the noun, both items being case-marked: *á gizona* (=common B *gizon hura*) 'that man'. Bi is also the only dialect that has the numeral *bi* 'two' after the noun; the other numerals follow the general pattern and precede it.

An important feature that distinguishes NL and Zu from Gi and Bi is the use of *hura* cited above as a unmarked personal pronoun in the first two: in the southern dialects, this use of *hura* is literary, the would-be emphatic pronoun *bera* being normally used to refer to persons, *hura* often being restricted to non-human referents in spontaneous speech.

As far as the strong and weak forms of genitive pronouns are concerned, it must be noted that the standard system described in 1.3.3.5 does not correspond to spontaneous use: everywhere, the weak forms are unmarked – and the strong forms have almost completely disappeared in La, the dialect whose 17th century variety has inspired that normative system. In general, the strong forms are used emphatically, when the possessor's identity is focused, but they also frequently occur as the genitive in the *X-en*

burua 'X's head' phrase when its interpretation is reflexive rather than literal (cf. 1.2.2.2); another fairly general use of those forms in earlier texts is "stylistic": when the addressee's status was high, he was typically referred to in vocative NPs as *neure Jauna* 'my Lord' rather than *nire/ene Jauna* (*ene* is the typical eastern form of *nire* 'my').

3.1.2. Case.

3.1.2.1. Diachronically, the case system has been fairly stable. Note however that the current ablative suffix *-tik* has evolved from an earlier *-ti*, attested in old Bi and contemporary Zu; likewise, there is some evidence that the comitative ending *-(re)kin* formerly was *-(re)ki* (which it still is in Zu). Old Bi also had and ablative in *-ean* (today's sg locative): suffixed to the demonstrative *a*, old Bi thus had an equivalent of French *en* or Italian *ne*, *arean* 'of it'.

3.1.2.2. The non-abs. plural case-endings are characterized by *-e-* in standard B (1.2.2.2), which is in fact is typical of the eastern half of the domain (NL, Zu, and HN) only: in Bi and Gi, *-a-* is used, just as in the sg. However, differences in pitch (§ 2.3) or stress (Jacobsen 1972) may play the same differentiating role.

3.1.2.3. Another dialectal variation worth mentioning is the existence of an allative suffix*-rat* in LN, which results from the amalgamation of all. *-ra* plus a final *-at*, since its Zu variant *-lát* is stressed (see 2.3.1). (i) In Gi, the same *-at* is a postposition that governs the ablative, yielding a static external localization (cp. *etxetik dator* 'he's coming out of the house', but *etxetik at dago* 'he is out of the house'). (ii) In older NL, *-ra* and *-rat* contrasted in meaning, the former simply indicating movement to a place, and the latter implying that the stay was to last for some time, as in *tabernara joanen da* 'he'll go to the pub', vs. *zerurat joanen da* 'he'll go to heaven'.

3.2. Variation in the verbal domain

3.2.1. Old Basque.

3.2.1.1. The most obvious diachronic piece of data is the dramatic reduction of the number of "strong" verbs which can be directly conjugated; Lafon 1944 has compiled the inflected forms of 6 distinct auxiliaries, 12 intransitive verbs, and 40 transitive ones for 16th century B. Today, apart from the aux., hardly a dozen verbs are spontaneously used so. It is not, however, possible to to say that *all* B verbs could be directly inflected in the past: only those whose perfective part. begins in *e-*, *i-* and ending in *-Ø*, *-n* or *-i* were, and there is no evidence that a V like *sartu* 'enter' has ever been conjugated. On the other hand, as shown by the oldest texts, it is clear that the aoristic conjugation described in 1.2.3.1 once was quite common, and not retricted to either imperative, potential or subjunctive moods. Liçarrague's (1571) translation of the *NT* thus contains

past aorist forms, e.g. *sar zedin* 'he entered', which are typically narrative: the periphrastic forms used today in narrations, like *sartu zen* (dit.), are restricted to one value of the English pluperfect, i.e. to refer to events relevant for the current state of affairs. However, even in the earliest known proverbs, which reflect an archaic form of the language, non-subordinate non-potential present aoristic forms are absent.

3.2.1.2. Another property of the ancient past tense is that, in two non-adjacent dialects, the suffix *-(e)n* was optional; see for instance the following proverb (# 386) from the 1596 *Refranes [...]* :

(58) Artzaiok arri zitea, gaztaeok agir zitea
herdsmen quarrel aux-past cheeses appear aux-past
'the shepherds quarrelled and the cheeses turned up'

Such forms also appear in early 17th century HN poems first published in Michelena 1964 – and SHN still has these suffixless past IVFs.

Recall now (a) that when all arguments are 3rd p., the IVF prefixes code tense or modality, as indicated in (20) (in fact, Bi has always had Ø- rather than *z-* in past transitive IVFs), and (b) that abs. prefixes in the past incorporate a copy of the past affix (1.2.3.2.3). A natural question then is to ask whether this *-n-* element has always been final (as it is synchronically). Suppose it originally was a prefix: the copy would then be the suffix, and the role of this prefixal *-n-* would have been to help disambiguate 1st and 2nd agreement markers which possibly did not distinguish between absolutive and ergative case/function, contrary to the *d-/z-* opposition that appeared with 3rd p. arguments. (Of course, the reason why a copy of it was made of it at the end of the IVF is just as mysterious as why a copy should have been made from a suffix into a prefix.)

3.2.1.3. The ergative prefixes of table K (1.2.3.2.3) have also raised discussions. As noted there, some scholars have proposed that the ergative suffix was moved towards the initial position at some remote date; however, this "ergative displacement" might just as well have been the other way round: no conclusion can be drawn from the pair *dakigu* 'we know' ≈ *genekien* 'we knew', or from the fact that certain dialects have *genekigun* for the latter, with both a prefix and a suffix cross-referencing a 2st p. pl. argument. Recall however that the allocutive affix *-k(a)-/-n(a)* (1.2.3.2.4) occurs between the dative and the erg. suffixes in the present tense, and that it is restricted to root clauses (contrary to the "ethical datives" of the neighbouring Romance languages). This latter restriction suggests that its presence has something to do with the structure of the utterance as such (i.e. "CP"), rather than with the internal structure of the sentence(s) that constitute(s) it ("IP"). In other words, the fact that the erg. suffix follows the allocutive one is surprising, in particular under the hypothesis that morphological structure mirrors syntactic structure: the erg. morpheme should be closer to the head of the IVF than the

alloc. one. But that is a possibility left open by the internal structure of *past* tense forms. Consider for instance the IVFs of (59):

(59)	a	eman diogu	'we have given it to him'	[*non-allocutive*]
	b	eman zionagu	*id.*	[+*alloc.*,+*fem.*]
	c	eman genion	'we had given/gave it to him'	[*non-allocutive*]
	d	eman genionan	*id.*	[+*alloc.*,+*fem.*]

In (b), whatever the structural relation between the prefix and the root *-i-*, the allocutive morpheme *-na-* is closer to *-i-* than the ergative suffix *-gu*. However, in (d), *ge-* or *gen-* (if *-n-* is not synchronically analysed as a tense marker) may be construed as closer to the root than *-na-*, as in the following bracketed representations:

(60)	a	[[[[ge- [n- [-i-]]] -o] -na] -n]
		1pl past root dat.3sg alloc past
	b	[[[[gen- [-i-]] -o] -na] -n]
		1pl-past root dat.3sg alloc past

The foregoing data clearly point towards a diachronic analysis according to which both the erg. agreement marker and the past tense affix were originally prefixes: only in such places could they ever be structurally closer to the root than the allocutive suffix. Needless to say, something more should be said concerning the suffixes when 1st or 2nd p. abs. prefixes occur; but the plural suffixes are phonemically transparent (*-gu, -zu* are the exact form of the pronouns themselves), thereby indicating that their appearance must be more recent than that of the prefixes.

3.2.2. Dialects.

3.2.2.1. The three macro-dialectal zones recognized in 2.3: East (Zu, NL, HN), Central (Gi), and Western (Bi) may again be recognized on the basis on isoglosses concerning IVFs and participles:

(61) The main conjugation isoglosses from a dialectal viewpoint

	East	*Gi*	*Bi*	
a	dut	det	dot	'I have (it)'
b	du	du	dau	's/he has (it)'
c	daut	dit	deust	's/he has (it) to me'
d	z-uen	z-uan	Ø-euen	's/he had (it)'
e	Ø -eza-	tu -eza-	tu -agi-	subjunctive
f	hilen	hilko	hilko	prospective part.
g	t(z)en	t(z)en	ketan/tu+ten	imperf. part.

The first three lines do not require any comments; (61d) illustrates the fact that Bi has a zero prefix in the past instead of *z-*; (61e) shows the formation of the subjunctive and related moods: in the eastern dialects, just as in EB, it is the root of the lexical verb

which is used (cf. table G), but in Gi and Bi, it is the perfective participle; moreover, Gi and the eastern dialects share the same suppletive aux. whose root is *-eza-*, whereas Bi uses another, *-(e)gi-*. (61f) shows the prosp. part. formation (but NL also has *-ko* after vowels). Finally, (g) indicates the formation of the imperfective participle. To summarize, the eastern dialects share three or four forms – depending on one's judgment on the pair *zuen≈zuan* in (d) – with Gi out of seven, whereas Gi and Bi only share one.

3.2.2.2. All the dialects can refer to the state (of the subject, object, or situation) resulting from an achieved action (itself indicated by the use of the perf. part., cf. 1.2.3.1), but the way this is encoded is not uniform; (i) the perf. part. itself may take on one of three suffixes: partitive *-(r)ik*, number agreement with the absolutive argument, or the suffix *-ta*, a contraction of the conjunction *eta* 'and'; (ii) the aux. may either remain unchanged, or be replaced by *egon* 'be (somewhere/in some state)' if the predicate is intransitive, or *e(d)uki* 'possess, hold', if it is transitive. This yields six theoretical possibilities, which are all realised in one dialect or another for *(nik) liburuak hartuak ditut* 'I have received the books':

Table L. *Formation and distribution of the resultative state pattern*

(i)	(nik) liburuak hartuAK ditut	NL, Gi, Bi
(ii)	(nik) liburuak hartuRIK ditut	Zu
(iii)	(nik) liburuak hartuTA ditut	Gi (rare)
(iv)	(nik) liburuak hartuAK dauzkat	Gi, Bi
(v)	(nik) liburuak haurtuRIK dauzkat	La, Gi, Bi
(vi)	(nik) liburuak hartuTA dauzkat	Gi, Bi

The same combinations are used to derive a "passive" and an "antipassive" voice: if the subject is demoted (it may then either retain the erg. ending or receive instead the instr. case, but is no longer cross-referenced in the IVF), a passive obtains; if it is the object that is demoted, both it and the agent are in the abs., but the IFV only agrees with this agentive abs. term, and if it is number that is marked on the participle, it will be the agent's number:

(62) a liburuak (nik) hartuak daude
books.abs I.erg received-pl are
'the books have been received (by me)'
b ni liburuak hartua nago
I.abs books.abs received-sg am
lit. 'I am received the books'

(Note that the agent is optional in the passive, but that the patient is retained in the antipassive, which is much rarer than the passive.)

3.3. Allocutive forms and the speaker/hearer relationship

We saw 1.2.3.2.4 that B has special IVFs to indicate high familiarity between the speaker and the hearer. Although the system illustrated there in (22) is typical (it is attested in the 1571 translation of the *NT*, and very much alive today in many places), other systems exist.

3.3.1. Some purists at the turn of the 20th century decided that allocutive forms were either "child talk" (*not* motherese!) – or even "Gipsy talk" (Azkue 1923-25). This has led to the development of two poorer systems, one in which all alloc. forms are banned, leading to the following readjustments of (22) in acceptability: (22a,b',c') are now good, but (23b,c) is out (this system is taught in some urban *ikastola* or B schools). Others have reduced their system to (22a), all uses of the 2nd p. sg. *hi* and the corresponding IVFs being banned as "rude".

3.3.2. The easternmost dialects (Zu, eastern LN, Sal and Ro) have, on the other hand, developed a richer system, in which there are two levels of allocutive forms, with the polite suffix *-zu* contrasting with the highly familiar ones *-k(a)* and *-n(a)*. Besides the neutral form *joanen n(a)iz* 'I'll go' and the masc. and fem. familiar ones *joanen n(a)uk / n(a)un*, we then also find *joanen nuzu*, with intermediary status between the first one and the second ones from the point of view of respect or distance between the speaker and the addressee.

3.3.2.1. Some varieties of Zu, ELN, and the Sa dialect across the border have even done away with the neutral forms in the syntactic contexts where allocutive forms are simply licit elsewhere; in this case, the system is another two register system, with the familiar *joanen nuk / nun* and the polite *joanen nuzu* contrasting in root clauses, and neutralized by the compulsory use of non-allocutive forms in subordinate clauses: *joanen nizela* 'that I will go'.

3.3.2.2. Eastern LN has grammaticalized a frequent process of "hypochoristic palatalisation": the counterparts of the foregoing forms are as in (63a-c) for 'I'll go' in root clauses; when the addressee is an argument: 'you'll go', the forms in (d-f) are used:

(63) a joanen niz [*polite*]
b joanen nuxu [*intermediate*]
c joanen nuk≈nun [*familiar*]
d joanen zira [*polite*]
e joanen xira [*intermediate*]
f joanen hiz [*familiar*]

3.3.3. Although it is not the morphology proper which is at stake, a final fact must be mentioned: in the Spanish Basque Country, a deferential way of referring to the

addressee consists in using the 3rd p. intensive pronoun *Berori* (made up of the intensive/reflexive root *ber-* 'self' followed by the distance II deictic *(h)ori*); the agreement morphemes on the IVFs are then 3rd p.: *Berori joango da* 'you'll go'. No allocutive flexions correspond to this use, but when *Berori* is used, the number of registers is again three if familiar forms are used, or at least two if only *zu* and *Berori* are used to refer to the hearer.

4. Syntactic variation

4.1. Word order and focus

4.1.1. If question words and phrases must precede the verb complex (§1.3.1), they did not have to be *adjacent* to it in the old language: *all* interrogatives (but especially sentential adjuncts like *zergatik* or *zertako* 'why') could appear sentence intially and be separated from the VC by various arguments and adverbial phrases. What is more, it is not really true that all rhematic material must precede it: when the rheme is not contrastive, it very often occurs to the right of the IVF, especially in the eastern dialects; in the central and western ones, the puristic tendencies that developed in the 19th century and reached their peak with Altube's 1929 book have influenced the linguistic behaviour of many Gi and Bi speakers; it is thus significant that Ubillos, who translated a catechism into Gi in 1785, and insisted in his preface that the rules of B word order were all too often violated in written texts, clearly possessed a rule of "end focus", as illustrated by the first sentences of his 9th lesson:

(64) Israeldarrak$_1$, Kanaango$_2$ bazterretan$_3$ sartu$_4$ ezkero$_5$, lendabizian$_6$ Juezak$_7$ ta$_8$ gero$_9$ Erregeak$_{10}$ izandu zitu[zt]en$_{11}$ agintari$_{12}$. Lenengo$_{13}$ Erregea$_{14}$ izandu zan$_{15}$ Saul$_{16}$, ta$_{17}$ urrengoa$_{18}$ David$_{19}$. Hau$_{20}$ zan$_{21}$, Juda$_{22}$ zeritzan$_{23}$ Jakobren$_{24}$ seme$_{25}$ baten$_{26}$ odoletik$_{27}$ ta$_{28}$ etxetik$_{29}$ zetorren$_{30}$ gizon$_{31}$ handi$_{32}$ bat$_{33}$.
'The Hebrews$_1$, after$_5$ entering$_4$ the-regions$_3$ of-Canaan$_2$, first$_6$ had$_{11}$ the-Judges$_7$, and$_8$ then$_9$ the-Kings$_{10}$, (as) leaders$_{12}$. The-first$_{13}$ King$_{14}$ was$_{15}$ Saul$_{16}$, and$_{17}$ the-next-one$_{18}$ David$_{19}$. This(-latter)$_{20}$ was$_{21}$ a$_{33}$ great$_{32}$ man$_{31}$ that-came$_{30}$ from-the-race$_{27}$ and$_{28}$ house$_{29}$ of-a$_{26}$ son$_{25}$ of-Jacob$_{24}$ whose name was (that-was-named$_{23}$) Jude$_{22}$'.

Note in particular the position of the copula 15 with respect to the proper noun 16, or yet again the copula 21, followed by a comma (i.e. a pause, which is heard today in the northern dialects when the copula precedes the predicative phrase), which is in turn followed by the long predicative phrase 22-33. Standard B would expect the perfective *izan(du) zan* and imperfective *zan* to occur after *Saul* in the first case, and at the very

end of the sentence in the second one.

4.1.2. The southern dialects can focalize the verb itself by putting it left-adjacent to a dummy verb, *egin* 'do, make'. This use of B *egin* and that of English *do* yield opposite effects, since the accented lexical verb (which tends to appear sentence initially) corresponds to emphasis on the positive character of the assertion, whereas *egin* never does that:

(65) a Jonek Miren jo zuen
Jon-erg M. hit aux 'Jon hit Miren'
b Jonek "Miren jo zuen 'Jon hit *Miren*'
b' Jonek Miren "jo egin zuen
Jon-erg M. hit do aux 'Jon *hit* Miren' (he didn't *kiss* her)
c "Jo zuen Jonek Miren 'Jon *did* hit Miren'

When it is associated with *egin*, the lexical verb is in the perfective form (cf. (61e) in 3.2.2): it is *egin* itself that is aspectually marked, as more clearly shown by (66a) with a prospective part.:

(66) a Jonek Miren "jo egingo du 'Jon will *hit* Miren'
b *Jonek Miren joko egin du

When the conjugation is synthetic, some Bi subdialects may also focalise the verb, but in this case, it is a copy of it that is inflected, as in:

(67) ekarri zekarren
bring s/he-brought-it's/he *brought* it' (rather than take it away)

The use of *egin* to focalise the lexical verb is, on the one hand, quite ancient, since an anonymous Bi Catechism published in 1596, known as *Betolazako dotrina*, contains examples of it – e.g. in the Decalogue, where *egin* is nominalized as *e(g)itea* (the whole sentences are to be understood as identificational or equative):

(68) bostgarrena iñor hil ez *eitea*
the-fifth-one nobody kill neg doing
lit. 'the fifth one is not to kill anyone'

(Note however that a (non-referential) object precedes the lexical verb.) It is remarkable, however, that up to the 19th century, the occurrences of focusing *egin*, even in Biscayan, were extremely rare: today, both in Bi and Gi, its appearance when the lexical verb is the only rhematic element in the utterance is almost automatic.

4.1.3. The eastern dialects (La, LN and Zu) have worked out another focusing device: by inverting the canonical participle + auxiliary word order, stronger emphasis is given to the element that immediately precedes the aux. The main difference between these dialects and the southern ones then lies in the identification of the element which, in its turns, identifies the focus: it is the item left adjacent to the IVF in Gi and Bi, and the IVF

itself in NL and Zu. Intuitively, this difference may be linked to another, namely, that yes/no questions in Gi and Bi are typically formed by inserting the particle *al* between the participle and the inflected aux., as in *etorri al da?* lit. 'come *al* s/he-is?', 'has s/he come?' whereas LN typically uses the morpheme *-a* suffixed tothe IVF: *jin de(i)a?* (id.), lit. 'come s/he-is-*a* ': in Bi and Gi, the question particle precedes the IVF, whereas in the northern dialect, it is a *part* of it. Possibly also correlated to that parameter is the difference in use of the "positive assertion particle" *ba-* prefixed to synthetic IVFs; as (69)-(70) show, the shades of meaning conveyed by its presence or absence are different according to the dialects (the bracketed items represent the focus of the sentence):

(69) a gizon batek bi seme zituen
man one.erg two son he-had-them
b 'a man [had two sons]' (Gi, Bi)
c 'a man had [two sons]' (eastern dialects)

(70) a gizon batek bi seme *ba*-zituen
man one.erg two son *ba*-he-had-them
b 'a man [did] have two sons (Gi, Bi)
c 'a man [had two sons]' (eastern dialects)

4.2. Valency and case marking

4.2.1. Owing to the development of the passive voice (3.2.2.2), the medio-passive use of intransitive structures has been decreasing in the past centuries; for instance, 'you'll be saved' in Liçarrague's *NT* (1571) is regularly *salbaturen* [=*salbatuko*] *zara*, whereas the periphrastic passive *salbatua izanen zara* (with the 'be' aux. in the prospective) becomes the rule in the 18th century. More intriguing are 16th century patterns translating a Latin passive by a medio-passive associated with the agent in the instrumental, as in (71) (on the aoristic past, see 3.2.1.1):

(71) Orduan Iesus eraman zedin Spirituaz desertura (Liç 1571: Mt 4.1)
then Jesus.abs take aux[intr.] Spirit.instr desert-allat.
'Then Jesus was led up by the Spirit into the wilderness'

Later translations first used a passive (18th and 19th centuries), and next the active voice (20th c.), with *Spiritua* in the erg. case.

4.2.2. The relationship between the case-frames of verbs and the theta-roles of their arguments may be considered an aspect of lexical variation, but it may also be considered to be syntactic. From this point of view, a verb like *urrikaldu* '(have) pity (on)' is worth mentioning, since until the 19th century, the same case-frame yielded two opposite readings:

(72) a urrikaldu natzaio

pity I-am-to-him
b 'I've had pity on him' (Gi, Bi)
c 'he's had pity on me' (northern dialects)

More generally, when an absolutive and a dative argument are associated, the subject is not automatically linked to the abs argument; compare:

(73) a Jon niri etorri zait 'Jon came (up) to me'
J.Ø to-me come he-is-to-me
b Joni dirua galdu zaio 'Jon has lost the money'
Jon.dat. money.Ø lost it-is-to-him
b' Jonek dirua galdu du (ditto)
Jon.erg money.Ø lost he-has-it

4.2.3. Finally, what Lafitte (1962) dubbed "the coastal solecism" is not only attested in 20th C. La, but also in the variety of HN spoken across the border (Irun, Fuenterrabia): in those areas, trivalent verbs (1.2.3.2), which normally accept only 3rd p. abs arguments, are associated with the bivalent aux. in *-u-* rather than with *-i-* or *-au-*, and the dative argument is cross-referenced in this aux. by an abs. prefix:

(74) a Jonek niri dirua eman dit / daut (EB / standard NL)
J.erg to-me money.Ø given he-has-it-to-me
'Jon has given me the money'
b dirua eman nau (coastal La, etc.)
money.Ø given he-has-me

4.3. Subordinate clauses

4.3.1. Relative clauses

4.3.1.1. Until the end of the 19th century, beside the strategy described in 1.3.4.1, B also formed relatives by using ordinary interrogative pronouns (except, crucially, *zein* 'which (one)' instead of *nor* 'who') as relative pronouns; the IVF then received either the prefix *bait-* (in the north: La, LN, Zu) or the suffix *-(e)n* (in the south: HN, Bi, Gi):

(75) gizona, *zeinak* ikusi *bait*-nau / nau-*en*, joan da
man.Ø which.erg seen *bai*--aux / aux-*en* gone is
'the man who saw me has gone'

It look *as if* the relative were appositive (since it follows a case-marked noun or NP), but the texts clearly show that, semantically, such relatives could just as well be interpreted as restrictive. Prescriptive grammarians then condemned this pattern as Romance, and it has fallen out of use; in the north, however, *bait-* subordinate clauses may often be interpreted as relative rather than causal (compare 1.3.4.1), thereby indicating that these dialects may in fact have substituted an empty operator for the explicit relative pronoun.

4.3.1.2. The standard strategy has always had an alternative, typical of the northern dialects, according to which the head noun may precede the relative clause (Oyharçabal 1985); in contradistinction to the mechanism described above, the head noun is *not* marked either for number or for case – the relevant morphemes are suffixed to the final complementizer *-(e)n*, as in the following excerpt from the 1571 *NT* (Mat 7,24):

(76) hura konparaturen dut [gizon xuhur [bere etxea gain batetan
he.abs I-will-compare man wise his house rock one-on
edifikatu duen]arekin]
built aux-*en*-sg.-comit.
'I will compare him to a/the wise man who has built his house on a rock'

The string *bere etxea...duen* would normally precede *gizon xuhur*, the latter adjective receiving the number and case suffix *-arekin*.

4.3.1.3. Free relative clauses were described in 1.3.4.1 too. To convey the meaning of generic or universal free relatives, the eastern dialects (the northern ones plus HN) also use *correlative sentences*, with a protasis introduced by an interrogative pronoun followed by *ere* 'ever/even', and a protasis often introduced by *eta* 'and' (in other contexts!) and containing a resumptive pronoun, as in:

(77) nork ere bekatu egin baitzuen, eta hura gaztigatua izan zen
who.erg ever sin made *bait*-aux, and him punished been aux
'Whosoever sinnned(, he) was punished'

Here again, the northern dialects use *bait-* in the protasis, whilst those spoken across the border use *-(e)n*. The presence of a pronoun (the special item *haina* mentioned in 3.1.1 was often used like *hura* in (77)), and the usual impossibility for the clause introduced by the interrogative to sit in an argumental position in the main clause clearly show that these structures cannot be reduced to English-type free relatives.

As far as (standard) free relatives are concerned, note finally that western B (Bi) *may* use the subjunctive mood instead of the indicative to trigger a generic (rather than specific) interpretation, as in Kerexeta's 1976 translation of Mat 5, 31-32, in which both options are used:

(78) bere emaztea itz*i* da*gi*anak, ezkontzauste-agiria
his wife-Ø leave aux[+subjn]-*an*-sg.erg certificate-of-divorce
emon dagiola. […] Bere emaztea iz*ten* d*au*anak […]
give let-him-to-her his wife-Ø leaving aux [+indic]-*an*-sg.erg
'Let the one who leaves [subjn] his wife give her a certificate.
The one who leaves [indicative] his wife…'

4.3.2. Tenseless subordinate clauses also exhibit isoglosses that essentially separate the northern dialects from Bi and Gi (with the other southern dialects somehow in-between). One example will suffice. In nominalised or gerundial clauses, the direct object is

generally in the genitive in the northern dialects, but remains in the absolutive in the southern ones:

(79) a egiaren erraiteko
truth.sg.gen say-*t(z)e*-for 'to tell the truth' [North]
b egia esateko
truth.sg.Ø say-*t(z)e*-for (ditto) [South]

Such variation in case-assignment may also be observed in the explicitly progressive construction with *ari* 'busy': this adjective governs the locative gerund in *-t(z)en*, which, in the northern dialects, differs from the present participle because here again it assigns the genitive to the object; all the options are listed below (note also the variant (d), typically northern, in which *ari* has grammaticalised into a pure aspectual particle):

(80) a Jonek liburua irakurtzen du
Jon.erg book.sg.Ø reading he-has(-it)
'John reads the book' (iterative, all dialects)
b Jon liburua irakurtzen ari da
Jon.Ø book.sg.Ø reading *ari* he-is
'John is reading the book' (progressive, south)
c Jon liburua*ren* irakurtzen ari da
Jon.Ø book.sg.gen reading *ari* he-is
'John is reading the book' (progressive, north)
d Jonek liburua irakurtzen ari du
Jon.erg book.sg.Ø reading *ari* he-has(-it)
'John is reading the book' (progressive, north)

5. Summary and final observations

At the segmental level, B phonology is fairly ordinary and homogeneous, except for the lamino-alveolar vs. apico-post-alveolar distinction between fricatives and affricates. At the supra-segmental level, there are three main accentual types, a [-2] and a [+2] stress systems, and a pitch-accent system, with an extra-metrical Low first syllable, and a High tone plateau that typically extends from the [+2] syllable to the [-1] or [-2] one. The fact that plural nouns systematically trigger anomalous stress or pitch contours indicates that some morphemes have always been inherently marked as irregular.

The case system marking the nuclear arguments is of the "active/inactive" type rather than ergative/absolutive, as is generally assumed. The conjugation is not only pluripersonal, mimicking the nominal case system: it may also incorporate morphemes referring to the hearer as such; those IVFs known as "allocutive", are restricted to root clauses, hence reminiscent of polite verb endings in Japanese.

Word order reflects a Topic - Focus – Verb order (which often coincides with Greenberg's SOV pattern), but post-verbal (non-wh) focus is also attested. In complex sentences, B has a wealth of complementizers and subordinate tenses and moods, but all the arguments of uninflected subordinate clauses may be explicit.

The unity of Basque is especially clear when seen from the outside: it is highly differentiated dialectally. The dialects fall into three main groups: the varieties spoken in France in the north-east, Gipuscoan in the middle, and Biscayan in the west. Interestingly, if Bi is prosodically and morphologically quite divergent from its Gi neighbour, the syntactic isoglosses would rather lead one to put these two together, and leave the northern dialects as more dissimilar.

References

Altube, Severo. Erderismos. Bermeo, 1929.

Azkue, Resurrección María: Morfología vasca. In: Euskera 1923-25.

Bonaparte, Louis-Lucien: Le verbe basque en tableaux, accompagné de notes grammaticales, selon les huit dialectes de l'euskara : le guipuscoan, le biscaïen, le haut-navarrais septentrional, le haut-navarrais méridional, le labourdin, le bas-navarrais occidental, le bas-navarrais oriental, et le souletin; avec les différences de leurs sous-dialectes et de leurs variétés. London, 1869.

Dixon, R.M.W.: Ergativity. Cambridge, 1994.

Elordieta, Gorka. Accent, tone and intonation in Lekeitio Basque. In: F. Martínez-Gil and A. Morales (eds.), Issues in the Phonology and Morphology of the Major Iberian Languages, Washington (DC), 1997, 3-78.

Gorrochategui, Joaquín: The Basque Language and its Neighbors in Antiquity. In: Hualde *et al.* 1995, 31-63.

Greenberg, Joseph H.: Some Universals of Grammar with particular Reference to the Order of Meaningful Elements. In: J. H. Greenberg (ed.), Universals of Language: Cambridge (Mass.), 1966, 73-113.

Hualde, José Ignacio. Basque Phonology. London/New York, 1991.

Hualde, José Ignacio. Euskararen azentuerak. San Sebastian/Bilbao, 1997.

Hualde, José Ignacio/Gorka Elordieta/Arantzazu Elordieta: The Basque Dialect of Lekeitio. Bilbao/San Sebastian, 1994.

Hualde, José Ignacio/Joseba Lakarra/R[obert] L. Trask (eds.): Towards a History of the Basque Language. Amsterdam/Philadelphia, 1995.

Jacobsen, William H. Jr.: Nominative-Ergative Syncretism in Basque. In: Anuario del Seminario de Filología Vasca "Julio de Urquijo" 6, 1972, 67-109.

Lafitte, Pierre: Grammaire basque (navarro-labourdin littéraire). Bayonne, [1944] 1962.

Lafon, René: Le système du verbe basque au XVI^e siècle. Bordeaux, 1944.
Levin, Beth: Lexical Structure in Basque. In: Marácz/Muysken 1989, 39-62.
Marácz, László/Pieter Muysken (eds.): Configurationality; The Typology of Asymmetries. Dordrecht, 1989.
Michelena, Luis: A propos de l'accent basque. In: Bulletin de la Société de linguistique 53, 1957-58, 204-233.
Michelena, Luis: Textos arcaicos vascos. Madrid, 1964.
Michelena, Luis: L'euskaro-caucasien. In: A. Martinet (ed.), Le langage, Paris, 1968, 1414-1437.
Michelena, Luis: Fonética histórica vasca. San Sebastian, 1977 [1961].
Morvan, Michel: Les origines linguistiques du basque. Talence, 1996. Review by G. Rebuschi in: Bulletin de la Société de linguistique 93.2, 1998, 324-325.
Orpustan, Jean-Baptiste: La langue basque au Moyen-Age (IX^e-XV^e siècles). Saint-Etienne-de-Baïgorry, 1999.
Ortiz de Urbina, Jon: Parameters in the Grammar of Basque. Dordrecht, 1989.
Oyharçabal, Beñat: Les relatives en basque. Paris, 1985.
Rebuschi, Georges: Structure de l'énoncé en basque. Paris, 1984.
Rebuschi, Georges: Is There a VP in Basque? In: Marácz/Muysken, 1989, 85-116.
Rebuschi, Georges: Essais de linguistique basque. San Sebastian/Vitoria, 1997.
Rebuschi, Georges: Nouvelles remarques sur "haina". In: Lapurdum 3, 1998, 53-75.
Rebuschi, Georges: Le complexe verbal basque; un regard «universaliste». In: Lapurdum 4, Hommage au professeur Jean Haritschelhar, 1999, 199-222.
de Rijk, Rudolf P.G.: Focus and Quasifocus [in] Basque Negative Statements. In: Revista Internacional de Estudios Vascos 41.1, 1996, 63-76.
Rotaetxe, Karmele: Estudio estructural del euskara de Ondárroa. Durango, 1978.
Sarasola, Ibon: Contribución al estudio y edición de textos antiguos vascos. In: Anuario del Seminario de Filología Vasca "Julio de Urquijo" 17, 1983, 69-212.
Satrustegi, Jose M.: Euskal Testu zaharrak, I. Pamplona, 1987.
Trask, R[obert] L.: The History of Basque. London/New York, 1997.
Zuazo, Koldo: Euskalkiak, gaur. In: Fontes Lingvæ Vasconvm 78, 1998, 191-233.

Abbreviations

A	agent
Abs	absolutive
Ae	Aezkoan
allat	allative
alloc	allocutive
aux	auxiliary
B	Basque
Bi	Biscayan
comit	comitative
dat	dative
dem	demonstrative
Det	determiner
DF	dative flag
EB	Euskara batua/unified B
ELN	Eastern LN
erg	ergative
fam	familiar
fem	feminine
Gi	Guipuzcoan
H	high (tone)
HN	Higher Navarrese
impf	imperfective
indet	indeterminate
infl	inflection
instr	instrumental

intr	intransitive	pl	plural
IVF	inflected verb form	prosp	prospective
L	low (tone)	Ro	Roncalese
La	Labourdin	S	subject
LN	Lower Navarrese	Sa	Salarese
loc	locative	SHN	Southern HN
neg	negation particle	str	strong
NHN	Northern HN	subjn	subjunctive
NL	Navarro-Labourdin	V	verb
NP	noun phrase	VC	verb complex
O	object	VP	verb phrase
p	person	WLN	Western LN
part.	participle	Zu	Souletin
perf	perfective		

Special phonetic symbols

[ß]	[J]
[δ]	[S]
[γ]	

Georges Rebuschi

Sorbonne nouvelle (Univ. Paris III)

Centre de recherche en Typologie grammaticale (TYGRe)

<rebuschi@ext.jussieu.fr>