Checking Domains in Basque and Breton

JON ORTIZ DE URBINA
University of Deusto

This article presents an analysis of some restrictions on verb initial word orders in Basque. Following Rivero's (1993a) analysis of similar facts in Breton, I will claim that the morphological feature for tense in these languages must occur in a configuration with a licenser in the appropriate structural position. Languages may differ as to licensing domain requirements, and I will claim that while Breton uses an internal domain, tense licensing in Basque takes place in the canonical head checking-domain. Where the relevant licenser is not available in that domain, last resort operations will apply to provide the correct licensing configuration. In particular, long-head movement in Breton creates the appropriate internal domain for tense, while ba-insertion produces the correct checking configuration for Basque. The article is organized as follows. The introductory section provides the relevant background information on word-order and verb first and verb second patterns in Basque. Section 1 deals with some restrictions on verb first structures in Basque, while Section 2 introduces similar facts from Breton, pointing out both the similarities and differences between the two languages. After showing in Section 3 that the verb initial patterns under consideration in Basque contain a tensed verbal form in C, Section 4 draws the analysis of the data. I claim that the morphological feature for tense must be licensed in the overt syntax in Basque, much like other morphological features. Such licensing is carried out in the canonical configurations for head checking as defined in Chomsky (1992). The article closes with a look at 'tenseless' finite constructions like imperatives, where, as expected, the restrictions on initial verbs do not hold.

0. Introduction: word order patterns in Basque

Although word-order is extremely flexible in Basque, it is often claimed that SOV sequences like the one in (1) represent the neutral, less marked pattern:

(1) Jonek liburua irakurri du.
John book read has
'John has read the book.'

(1) A version of this paper was presented at the Lund meeting of Group 3 (Complementation and Subordination) of the Eurotyp Project, in April 1993. I would like to thank G. Rebuschi and, especially, M. Rivero for their helpful comments. Usual disclaimers apply.
This claim is reconstructed as indicating a rather consistent right peripheral position of syntactic heads. Assuming the VP-internal subject hypothesis, an simplified initial representation of (1) would be as in (2), with final V and Infl (the inflected auxiliary):

(2) \[ [[\text{Jonek liburua irakurri}]_{\text{VP}} \text{dul}]]_{\text{IP}}

However, there are many apparent cases of "residual" (in Rizzi's 1991 terms) V2 phenomena, where the verbal element(s) occur systematically displaced to a second position, to the right of the operator, in the left periphery of the clause:

(3) Zer irakurri du Jonek? (4) LIBURUA irakurri du Jonek.
   'What has John read?' 'It is the book that John has read.'

Similar patterns, at least with interrogative clauses, are customarily analyzed as produced by movement of some verbal heads to a left Comp or Infl head in head-first languages. Still such patterns are in a certain way unexpected in a head-last language like Basque, and it seems that some head must precede its complement to account for these structures in familiar ways. I will assume here that the relevant left headed phrase is CP, resulting in the following basic configuration for the main categories:

(5) CP
    \[ \text{foc} \]
    \[ \text{wh} \]
    \[ C' \]
    \[ C \]
    \[ \text{IP} \]
    \[ \text{VP} \]
    \[ I \]
    \[ V \]

That is, I will assume throughout that wh-words and foci move to Spec, CP, with some residual head movements to the Comp head (see Uriagereka 1992 for a different view). This accounts, in broad terms, for the patterns in (3) and (4). Although there are some overt independent complementizers to the left of their complements, the most typical ones, -(e)n, bait- and -(e)la are clitic-like elements which appear amalgamated to the inflected element. I will offer more details later. Suffice it to say here that unless independent principles force movement of some head material to the C head, the clitic complementizer will have to lower to be provided with a morphological base.

1. Verb initial patterns

Verb initial orders like the ones in (6) are also possible in Basque:

    'Bought has John book' 'Died has father
    'John HAS bought the book' 'Father HAS died.'
Such V1 patterns correspond pragmatically to verbal focalization patterns, and are described as such by Basque grammars. Now, not all verbs may occur in clause initial position. The verbs in (6) are periphrastic verbal forms made up by combining a lexical tenseless verb with the inflected auxiliaries. Synthetic verbal forms, where the lexical verb itself is inflected, are ruled out clause initially:

(7) a. * Dator Jon orain.  
comes John now
knows John that

Exactly the same patterns are reproduced in yes/no question formation. A typical yes/no question on the declarative statements in (8) is given in (9), again with a V1 pattern:

John book bought has  
'John has bought the book.'

(9) a. Erosi du Jonek liburua?  
bought has John book  
'Has John bought the book?'

However, the corresponding yes/no questions of the declarative sentences in (10), which contain tensed amalgamated verbs, are not acceptable:

(10) a. Nik hori nekien.  
I that knew

(11) a. *Nekien nik hori?

All of the starred sentences above become acceptable if the particle ha appears immediately to the left of the tensed verb:

(12) a. BA-dator Jon orain (=7a)  
comes J. now
b. BA-nekien nik hori? (=11a)  
knew I that  
'Did I know that?'

(3) A few counterexamples occur, since forms like da 'is', nago and dauka: 'I think' sometimes occur initially in old texts (G. Rebuschi, p.c.; see also Oyarzabal 1984). These examples, very marked at present, usually involve heavy sentential complements or complex NPs. Thus, in the following example from Joannes d'Etcheberri's 1712 Latin grammar, the answer to the question 'What is a verb?' is given in the order verb-complement, with a tensed initial verb:

(i)' Da perpausaren parte bat ceina declinatzen baita moduetatik, eta demboretatik,  
is clause part one which decline aux mood and tense  
etta sinificatzen bairosu iñatea, eguztia, edo sofritcea  
and mean aux being, doing or experiencing  
'(It) is the part of the clause which is declined for mood and tense and which means being, doing or experiencing.  
Mitxelena (1981:70) considers similar examples in old christian doctrines as 'abbreviations' with elliptical subjects.
I will refer to this phenomenon as *ba*-support and in section 5 I will try to relate it to licensing conditions on Infl (actually for Tense).

Let’s turn now to embedded contexts, beginning with indirect questions. A clear difference emerges between direct and indirect yes/no questions. Unlike in direct yes/no questions, verb initial patterns are very marked and *ba*-support is excluded in embedded yes/no questions:

(13) Ez dakit (*ba-)*dakizu-n (zuk) hori.
    neg know know-comp (you) that
    ‘I don’t know whether you know that.’

While this sentence is marked for reasons I will turn to below, it differs from the clearly ungrammatical verb initial direct questions in (11). Moreover, in sharp contrast again with direct questions, *ba* insertion not only does not improve the sentence, but actually renders it ungrammatical.

As for embedded declarative complements, the situation is slightly more complex: tensed verb-initial orders without *ba* are possible, but *ba*-support is required if the verb is emphasized. This can be observed in the following examples:

(14) a. BA-dakit [datorr-ela] (Cfr. *Dakit datorrela)
    -know comes-that
    ‘I know that (he) is coming.’

b. Esan didate [BA-datorr-ela]
    told have comes-that
    ‘They told me that he IS coming.’

In (14a), the initial matrix lexical verb must be preceded by *ba* if tensed (14a), but the embedded verb may but need not be preceded by *ba*. It will only if emphatic, as in (14b). Finally, the sentence initial periphrastic verb in (14b) is perfectly acceptable, as usual. I will return to these asymmetries below. To summarize, verb fronting in root clauses (interpreted here as movement to left-headed C) cannot leave a tensed form in initial position. A language specific process of *ba*-support must apply in these circumstances. This rule cannot apply in embedded questions, but may do so in embedded declaratives.

2. Verb initial patterns in Breton

The restrictions on verb-initial orders examined in the preceding section are reminiscent of the constraints on verb placement in Breton discussed in Rivero (1993a). In this VSO language, tensed lexical verbs may not occur in matrix initial position, though they are acceptable in embedded clause initial position:

(15) a. *En deus lavaret [he deus desket he c’henteliou]
    3S have said 3S have learned her lessons
    ‘He has said that she has learned her lessons.’

b. *Lenn Anna al levr.
    reads Anna the book

(4) Examples like (14a), with *ba*-less embedded initial tensed forms are possible where a focalized constituent appears in the matrix clause (in this example, the matrix verb itself).
In (15a) the tensed element is the auxiliary part of the periphrastic verbal form. The order Aux-V in (15a) is acceptable in the embedded clause, but the existence of the same order in the matrix clause is the offending factor rendering the sentence ungrammatical. In (15b) the synthetic tensed form without auxiliary is also ruled out in the same matrix initial position.

All of the ungrammatical Basque examples above contain “synthetic” tensed verbs like (15b) rather than inflected auxiliaries like (15a). This is so because, Basque being a head-final language, the sequential order of elements in periphrastic verbal forms is V-Aux in both matrix and embedded contexts. This means it is difficult in most dialects to find a clause initial auxiliary. However, some dialects admit Aux-to-C raising in questions and focalizations, leaving behind the lexical verb and producing V2 contexts where the second element is the auxiliary head. This dialectal pattern is represented in (16), where the sentences parallel the common pattern in (3) and (4) above, repeated here for convenience:

(3) Zer irakurri du Jonek?
What has John
‘What has John read?’

(4) LIBURUA irakurri du Jonek.
book ‘read has John
‘It is the book that John has read.’

(16) a. Zer du Jonek irakurri?
what has Jon read
b. LIBURUA du Jonek irakurri.
the book has Jon read

We also find dialectal V1 patterns similar to the ones under consideration, with a tensed initial auxiliary; this is possible provided the auxiliary is preceded by ba:5

(17) a. *(ba)dut (hori) ikusi!
have that seen
‘I have seen that.’
b. *(ba)duzu (hori) ikusi?
have that seen
‘Have you seen that?’

Examples like the ones in (17) are the counterpart of the Breton (15a), where a tensed auxiliary initiates the clause. In both languages, a last resort rule must apply, Long Head Movement (LHM; see Lema and Rivero 1989) in Breton and ba-support in Basque. LHM of the tenseless lexical verb over the inflection produces an ‘inverted’ order which leaves the tenseless form in initial position and the tensed one in ‘second’ position, as observed in (15a) above, repeated here as (18):

(18) Lavaret en deus [he deus desket he c’henteliou]
said 3S have 3S have learned her lessons
‘He has said that she has learned her lessons.’

(5) At present, though, the patterns in (17) are found mainly as residual structures in dialects different from those where the productive types in (16) are found.
Here, the V-Aux order in the matrix is produced by LHM, and contrasts with the Aux-V order in the embedded clause. Both ba-support and LHM place the offending tensed element in second position.

Actually, the similarity between the two languages extends further in significant ways: what counts as a valid 'first' element also coincides to a large extent, as described in the following sections.

2.1 Negation

The negative particles *ne and *ez count as first elements:

(19) a. *dator Jon comes John
b. Ez dator Jon. 'John is not coming.'
c. *lenn Anna al levr reads Anna the book
d. Ne lenn ket Anna al levr. 'Anna does not read the book.'

The ungrammatical tense-initial orders in (19a,c) are perfectly acceptable if preceded by the negative element (19b,d). The examples in (20) show the same pattern with periphrastic verbs:

(20) a. *duJonek liburua irakurri has John book read
b. Ez duJonek liburua irakurri. 'John has not read the book.'
c. *en de us lennet Tom al levr has read Tom the book
d. N'en deus ket lennet Tom al levr. 'Tom has not read the book.'

No dialect admits tensed auxiliaries in initial position, but all display them dislocated to the left when preceded by the negative particle, as in (20b). This is parallel in the relevant aspect to the Breton (20d).

2.2. Illocutionary particles

Basque and Breton possess a yes/no question particle, *al and *ha, respectively. Although placed immediately preceding the tensed verb, this interrogative particle fails to shield the inflected element from the verbal position, and the usual last resort mechanisms must apply to salvage the sentences:

(21) a. *Ha a raint soñjal er bleuniou? Q aux think the flowers
     'Will they think of the flowers?'
b. Ha soñjan a raint er bleuniou? (LHM)

(22) a. *Al daki Jonek? Q knows John
     b. Ba al daki Jonek? (ba-support) 'Does John know?'

In (21b) the last resort mechanism is LHM, which produces the inverted pattern Aux-V. In (22b) the mechanism is ba- insertion.
2.3. XPs in Spec of CP

There are two types of phrasal elements which appear in the specifier position of CP: a) wh-words and b) topicalized/focalized XPs. Beginning with wh-words, wh-questions show the inflected element in second position, after the wh-word:

(23) Piv en deus lennet al levr?  
who has read the book
(24) Nork daki hori?  
who knows that

The ‘second’ position is occupied by the auxiliary-like element in (23) and the synthetic verbal form in (24). Where the verb is periphrastic, most dialects of Basque place the lexical verb-auxiliary sequence to the right of the wh-word, behaving like Spanish in this area (see (3) above). However, I have shown above that some (northern) dialects also admit an ‘English-type’ of question where the lexical tenseless verb is left behind and the tensed auxiliary appears in second position. No ba-support is then required:

(25) a. Nork ikusi du hori?  
who seen has that
b. Nork du hori ikusi?  
has that seen

Non-interrogative elements may also occupy the Spec, CP position, and they uniformly count as “first” for tense licensing. Thus, Borsley, Rivero and Stephens (1992) analyze topicalized NP’s as occupying that position; there they can be sequentially followed by an inflected element:

(26) Al levr en deus lennet Tom.  
the book has read Tom

In Basque, focalized constituents are often claimed to occupy that position, accounting in this way for their behavior, parallel to wh-words. Much like the topicalized constituent in (26), focalized constituents in Basque can (indeed, must) occur before the tensed verb:

(27) JONEK daki hori  
JOHN knows that  
(Cfr. neutral Jonek hori daki)

With periphrastic verbs, the pattern is also identical to that of wh-questions: all dialects can have the lexical verb-auxiliary sequence following the wh-word, while some dialects also admit as a marked option a pattern where the auxiliary follows the focalized constituent and the lexical verb is left behind:

(28) a. JONEK ikusi du hori  
JOHN seen has that
b. JONEK du hori ikusi  
has that seen

In (27) and (28b), therefore, a tensed verbal form (auxiliary or inflected verb) is licensed (in a way to be made precise) by the preceding NP in Spec of CP.

2.4. Dislocated constituents

In contrast with XPs in Spec of CP, dislocated constituents external to CP do not count as first elements. Thus, Breton (26) differs from (29), where Yann has been dislocated:
Rivero (1993a) uses this sentence to show that left-dislocation is compatible with participle preposing (via LHM). If preposing does not take place, the tensed verbal form would occur to the right of the dislocated constituent, creating an ungrammatical pattern:

(30) *Yann, meus roet al lev dzhan
have given

Similarly, Basque constituents preposed and separated intonationally from the rest by a pause, referred to as topics in Basque grammar, do not count as first elements, and cannot be immediately followed by tensed constituents. Since Basque is an extended pro-drop language, no pronominal copy of the dislocated element is apparent:

(31) a. *Jonek, daki hori b. *Joni, dakar-kio-t hori
John knows that John-dat carry-3D-1E that
‘To John, I bring that.’

As usual, ba-support must apply to salvage the sentence:

‘As for John, he does know that.’

b. Joni, badakarkiot hori.
‘As for John, I AM bringing this to him.’

Functionally, the preposed elements are topics, and the clause must contain a focalized element (the verb itself in (32)). A special subtype of this case involves repetition of a participial copy of the verb, while the clause-internal verb receives prominence and moves to the left-peripheral position. Again, where as in (33) the verbal form thus emphasized is tensed, ba-support is required:

(33) a. Ikusi, ikusten dugu hori. b. *Jakin, dakigu hori
see see aux that know, know that
‘As for seeing, we do see that.’

c. Jakin, badakigu hori.
‘As for knowing, we do know that.’

2.5. A licensing approach

Rivero’s (1993a) analysis is based on the idea that the finite feature of Infl, like other morphological features, must be licensed (Chomsky 1992). This licensing takes place within local X-bar theoretic domains, in particular, in the case of Breton, within the internal domain of the c-commanding head immediately above Infl, that is, C. Finite Infl in Breton, then, will be licensed when IP is the complement of C. When an element is hosted within CP, the C head is generated and finite Infl can then be licensed in its complement. X will ‘count as first’ to which Infl can be
'second'. X could be the negative head moved to C (2.1), a phrasal constituent in the specifier of CP (a topic or a wh-phrase as in 2.3), or a participle moved by LHM to C as a last resort process to create the appropriate domain for licensing (2.2 and 2.4).

The analysis of Basque VI phenomena I will outline below shares Rivero's approach. V or Aux raising to I would not check successfully all of the I features, and the finite feature would have to be licensed by some other element. I will claim that the finite feature is licensed within its checking domain, defined as in Chomsky (1992):

$$\begin{align*}
\text{XP}_1 & \sim \text{WP} \sim \text{XP}_2 \\
\text{ZP} & \sim X' \\
X_1 & \sim YP \\
H & \sim X_2
\end{align*}$$

Where X=finite Infl (or TNS), it is checked and licensed by a head adjoined to it (H) or by an agreement element in its specifier (ZP). I will be discussing these two subdomains in sections 4 and 5, respectively.

The divergence in checking domains is one of the differences between the two languages, but there are other not less important. Due to the predominantly head-final nature of Basque projections, the structural analysis of the Basque structures above is far less similar to that of their Breton counterparts than the sequential parallelisms might suggest at first sight. In this connection, there are two main differences I would like to point out. First, since Basque VP and IP are head-final, an 'initial' order like [V - X] or [Infl - X] does not contain a verbal or inflected element in situ. Rather, these seem to be derived patterns where such heads have been removed from their canonical position: V - [X - t], Infl - [X - t]. A second important difference worth remembering is that while participle preposing by LHM occurs in neutral contexts in Breton, ba-support only occurs in question or focalization contexts. Therefore, we may assume that in Breton tensed Infl remains in situ and is licensed in that position too. On the other hand, the Basque counterparts contain displaced V or Infl, the result of movement to a left-headed C position, prompted perhaps by something like Rizzi's wh-criterion. Ba-support is a last resort rule licensing a finite form in C, while participle preposing by LHM licenses Infl in situ.

These differences between Basque and Breton are exemplified in (35) for the main patterns examined above in the analyses under consideration:

(35) This might be different in northern Basque dialects, where, according to Rebuschi (1982, 1983), ba- is sometimes unmarked.

(7) This is similar to the differences between V2 patterns in Germanic and in English, that is, between the regular and residual types: ba-support only applies in residual contexts, while LHM applies in basic ones.
While finite Infl is differently located, I will claim that its licensers (XPs, participles and negative elements) occupy the same position in the two languages. As indicated above, Rivero's analysis is based on the idea that finite inflection in Breton may be licensed in internal domains, usually as complement of C. Where the C head is not generated independently, a last resort rule of LHM may take the tenseless V form to C skipping over Infl, thus creating the internal domain where tensed Infl may be licensed. The analysis sketched below for Basque follows Rivero's analysis in considering ba-support a last resort rule required to salvage structures where some feature of Infl (TNS, actually) has not been successfully satisfied. However, I will claim that Basque Infl is licensed in two ways: within a head-domain or by SPEC-head agreement with a lexical element. The latter licensing takes place within CP.

The basic similarity between Breton and Basque lies in the morphologically weak nature of tensed items in the two languages. A weak finite feature or TNS head must be licensed by another element in the appropriate configuration. This gives rise to Wackernagel-like second position effects, analyzed in this approach as stemming from the checking requirements of some morphological features. 8

3. V/Infl movement to C in Basque

Before proceeding with the analysis, I will support the claim made above that in the VI contexts where ba-support takes place the Verb and/or tensed auxiliary occupy a left-peripheral C position. This has been assumed up to now on the basis of the head-final character of VP and IP. In this section, however, I will support this claim by comparing VI and V2 orders. I will show that VI is in fact a type of V2 pattern where the first position is occupied by an empty operator, either an interrogative yes/no operator or an null emphatic operator.

3.1. Residual V2 and V1

The V1 patterns exemplified above are found in 'verb focalization' sentences, where a positive assertion is emphasized, and in yes/no questions. These two are precisely the clause types where (residual) V2 patterns are found in Basque. Thus, as

(8) The resemblance of the Basque constraints to Wackernagel effects had already been explicitly asserted by Mitxelena (1957), who mentions "the enclitic nature of finite verbal forms, at least auxiliaries. That is, this should be interpreted in the form proposed by Wackernagel for similar Indoeuropean facts" (Mitxelena 1957: 177, ft. 32).
indicated above, focalized or questioned constituents occur at the left periphery, where the first position is occupied by the operator and the verbal element occupies the second position:

(36) a. Zer irakurri du Jonek? b. LIBURUA irakurri du Jonek
 what read auxJon book read auxJon
 'What did Jon read?' Jon has read THE BOOK

Nothing may intervene between the operator and the verb. The VI patterns above fall under the same generalization if we follow the assumption that they contain a null operator in first position, a yes/no question operator and an emphatic operator:

(37) a. Q irakurri du Jonek liburua? b. ! irakurri du Jonek liburua
 read auxJon book Jon HAS read the book
 'Has Jon read the book?'

Following Rizzi (1991), Brody (1990) and Ortiz de Urbina (in press), among others, we may assume that the operator feature of wh-words and foci must be matched by a corresponding feature in the C head. This independent feature must be supplied by some head provided with that feature and moving to C. Rizzi locates the [wh] feature in Infl., and we can follow Horvath (in press) and Tuller (1992), and assume that the syntactic feature [Focus] can be hosted by some functional head, among them Infl itself. Thus, if Spec of CP is occupied by an element bearing an operator feature (including the null operators in (37)), C will have to possess that feature to agree with its specifier, and movement of the functional head hosting the feature will supply it.

These hypotheses are not enough by themselves, since the head moving to C in Basque V2 sentences like (36) and VI sentences like (37) is not Infl alone, but Infl and V, whether amalgamated in synthetic forms or as the complex head of periphrastic verbs. In effect, something similar happens in the Romance languages, as in the Spanish (38), with a periphrastic verb:

(38) a. ¿A quién ha visto María? b.*¿A quién ha María visto?
 who has seen Mary
 'Who has Mary seen?'

The sequence Wh-aux-participle in (38a) may arise either by head-to-head movement from V to C through Infl or, alternatively, by a more limited Infl-to-C movement plus displacement of the intervening subject. In this second analysis the auxiliary is located in C, but V remains within VP. Rizzi (1991) rejects the first option at least for Italian, on the grounds that intervening elements (adverbs) may appear in between the verb and inflection. This seems to exclude extended head-to-head raising in Italian. However, subject postposing with Infl-to-C, the alternative proposed for Italian, does not go through in Basque. Given a left head CP but a right headed IP/VP as in (5), I-to-C would remove the auxiliary or Infl from the right edge of the clause to the left periphery where C is located. This would give us an Infl-X-V order, with the wrong sequence Aux-V even if we found reasons to eliminate any possible argument and adjunct intervening between the two. It seems then that the extended head-to-head analysis is more appropriate in Basque, and this
squares well with the fact that the unit verb-inflection, in that order, is not breakable in Basque, unlike Italian.

Therefore, since the extended head movement alternative (V-to-I- to-C) is preferable, we have to find a reason for V to move to Infl first; once there, further movement to C will be prompted by the Operator Criterion. For this purpose, I will adopt here the analysis of Focus phenomena developed in Tuller (1992) and Horvath (in press). The functional head hosting the syntactic features [wh] and [Focus] must be “lexicalized” in some languages, that is, a lexical head must move to it. This is achieved in Basque by head-to-head movement of V to Infl, both for synthetic and periphrastic verbs. Once the head Infl is thus lexicalized, the new complex V/I unit must move to C, where specifier-head agreement with the feature-bearing operator will satisfy the Operator Criterion. All this is represented in (39):

(39)

The need to lexicalize a feature bearing Infl head is open to parametric variation, and the dialectal difference sketched above, where northern dialects admit, in a marked register, the restricted Infl-to-C movement, might be related to this.

3.2. V/Infl-to-C and ba-support

After this cursory account of the residual V2 (and V1) phenomena in Basque, let us return to the task of showing that ba-support affects synthetic verbs in C. Not all cases of question formation or focalization prompt verb movement to C: some embedded contexts do not trigger movements of V or Infl to C, and it can be shown that it is precisely in those cases where this residual movement does not take place that ba-support does not apply either. Conversely, those contexts where V-to-C movement takes place involve ba-support if V is synthetic. This shows that ba-support affects a synthetic verb in C position.

Verb movement to C is obligatory in root yes/no questions and “verb focalization” clauses. In the latter, after V has lexicalized the functional head, it must move on to supply C with the feature required to match the wh-feature in Spec,CP, as dictated by the Operator-Criterion. Ba-support is obligatory here with synthetic verbs. As for embedded contexts, the situation is different: as indicated above, ba-support does not take place in indirect yes/no questions, while it may occur in embedded declaratives:
Uste dut [badatorrela Jon]
'I think Jon IS coming.'

(40) a. Ez dakit [(Jon) (*ba-)datorr-|en]
    neg know comes-whether
'I don’t know whether Jon is coming.'
b. Ba dakit [(Jon) datorr-ela] c. Uste dut [badatorrela Jon]
    know Jon coming-that I think Jon IS coming
'I do know that Jon is coming.'

This correlates with the incidence of (residual) V2 effects in these structures. Thus, the unmarked order in embedded indirect questions is not V2 by verb final, suggesting that V(Inf) movement to C does not take place there; correspondingly, ba-support does not take place, as (40a) indicates. On the other hand, embedded declaratives show V2 orders only when a constituent is being focalized, as in (40c). Otherwise, we find the neutral verb-final pattern in (40b). Not surprisingly, ba surfaces only where the verb is being emphasized; otherwise it remains final and ba-less.9 What these correlations suggest, therefore, is that V-to-C does not take place in indirect questions, while it does in embedded emphatic declaratives. This can be analyzed along the lines of Rizzi. In indirect yes/no questions, the matrix verb subcategorizes for a [wh] complementizer, and this feature matches the [wh] feature of the covert yes/no operator in Spec of CP. No raising to C is then required to apply, and ba-support will not occur. Embedded declaratives differ from indirect questions in that there is no focalization subcategorization: verbs may subcategorize for interrogative complements, but not for focalized complements. Where an empty emphatic operator is generated in SPEC of CP, Infl, lexicalized by V, must move to C to provide it with the relevant feature, and ba-support will apply if the verb in C happens to be synthetic. Otherwise, V remains final in V (or in Infl). Verb movement to C will not take place, and hence ba-will not appear. As a conclusion, ba-support only applies to license Infl in C, since it only takes place where the Operator Criterion requires Infl to move to that position. How this licensing is done is discussed in the following section.

4. Infl (TNS) licensing

In this section, I address the question of Infl licensing, and will claim that Infl (the T head or a finite feature) is licensed in its checking-domain, in the sense of Chomsky (1992). The licenser in that checking-domain may be a head adjoined to Infl (head-licensing) or its specifier (specifier-licensing). I will discuss them in turn.

4.1. Head-licensing

Ba-support only applies to license synthetic verbal forms as a last resort rule, as shown again in (41a). Where the tense-bearing element is an auxiliary as in (41b), licensing seems direct and this rule does not apply:10

9 If the subject in (40a,b) is pro rather than Jon, the verb would also be ‘initial’, but even so it would appear without ba, since this order would still not be the result of movement to initial C. Embedded questions where the verb is the only overt element, like the one in (40a), sound better if the tag ais ez ‘or not’ is added to them.

10 The parenthesized question mark indicates that the same constraints hold true of verb initial emphatic declaratives and questions.
(41) a. *Dator lonek?  
   is coming Jon  

b. Irakurri du jonek liburuak?  
   read aux Jon book  
   'Has John read the book?/  
   John HAS read the book.'

Notice that the head moving up to C in (41b) is not only Infl, but also V, as discussed in section 3.1. I assume that V raises to I first, adjoining and licensing it. Licensing is achieved by establishing a relationship with the licenser within a particular domain. In the case of (41b), the domain where Infl(TNS) is licensed is the head-domain of V. The head-checking domain can be represented as in Chomsky and Rivero:

(42) Head-checking domain for licensing:

\[
\begin{array}{c}
H \\
 Y \\
 H
\end{array}
\]

As (42) indicates, a head H is licensed by another head Y adjoined to it. In our sentences H=Infl and Y=V. Further movement to C is prompted by the Operator Criterion. Thus, a pattern like (41b) would differ from the apparently similar Breton case in (43a). The structures are represented below:

(43) a. Lennet en de us Tom al levr.  
   read aux Tom the book  
   'Tom has read the book.'

b. Irakurri du Jonek liburuak (=41b)  
   read aux Jon book  
   'John HAS read the book'

c. cp[OP read+hasj IP[Jon; VP[t; the book t; t;]] (Basque)

d. cp[ readj IP[has VP[Tom t; al levr]] (Breton)

The sequence Participle+auxiliary is an inverted pattern in Breton, different from the order found in embedded clauses. On the other hand, that is the standard sequence in both matrix and embedded contexts in Basque. The participle in the Basque sentence (43b) has moved first to Infl to license the [TNS] feature of this head and then the unit has moved to C for independent reasons. In contrast, the participle in the Breton example (43a) appears in C as a result of Long Head Movement of V over Infl to C. In Rivero’s analysis, this movement to C is required to establish the licensing domain in Breton: a complement domain where [TNS] occurs within the complement of a head. By moving the participle there, the C head is activated, as it were, and Infl occurs within its complement domain. The participle in Basque also appears in C, but this is due to independent reasons, namely to the Operator Criterion.

Since the two sentences are so similar, one might want to argue that LHM takes place in Basque too, and that Infl moves to C first by head-to-head movement, followed by V skipping the original Infl position by LHM. In so doing, V would both license [TNS] (and lexicalize it). However, this analysis seems unwarranted for Basque. Thus, the participle (plus inflection) will also move to C in other contexts where Infl is being licensed independently. For instance, in wh-questions like (24) above, repeated here, the synthetic form daki ‘knows’ does not require ba- (actually cannot take it):
(24) Nork daki hori?
who knows that

This means that [TNS] is being licensed independently. But if the verb is periphrastic the participle, in most dialects, must move along with the auxiliary:

(44) Nork jakin du hori?
who known aux that
‘Who has found out about that?’

That is, even though Infl can be licensed independently, the participle also appears in C. This would be hard to understand if participles moved to C by LHM to salvage TNS in Basque as they do in Breton. Rather, it seems that V-to-I serves that purpose, while other principles (like the wh-criterion) prompt V+I raising to C.

The availability of head licensing in general accounts for the fact that (45a), but not (45b) is a well-formed sentence in Basque:

(45) a. pro etorri da
come has
(S)he has come

Basque is an extended pro-drop language, where any of the arguments reflected in the inflection (subject, direct object and indirect object) may be pro-dropped. If the subject is dropped in a SV sequence like the ones in (45), the result is acceptable only if V is a periphrastic verb, but not is it is synthetic. If V-to-Infl takes place in normal clauses, this would be accounted, since the adjoined V head creates the head-checking domain in (45a) while the amalgamated head does not in (45b).11

4.2. Negation

V’s are not the only heads that create the head-checking domain that licenses TNS in Basque. Just like negative heads can function as first elements in Breton, so can they in Basque, as seen in (20) above, repeated here:

(20) a. *du Jonek liburu irakurri
has John book read
b. Ez du Jonek liburu irakurri.
‘John has not read the book.’

Negative structures in Basque are interesting in this respect because they represent the only case where auxiliaries are separated from the lexical verb in all dialects, thus eliminating the latter as a possible head-licenser. Negative formation, with both synthetic and periphrastic verb forms, is presented in (46):

(11) One might be tempted to interpret ha-support as a rule providing a head (ha-) creating a domain for head-licensing in synthetic forms. I will claim later, though, that this is not the case and that ha- occurs in SPEC and is then not a head.
The negative element occurs in a left peripheral position (although subjects often precede it), and immediately preceding the tense-bearing element. In (46) the inflected element is a synthetic verb, while in (47) it is the auxiliary of a periphrastic form. In the latter case, the lexical verb appears to the right. I will assume that negation heads a phrase in Basque placed above VP but under Infl:

\[
(48) \quad \text{IP} \\
\quad \text{NegP} \quad \text{I} \\
\quad \text{VP} \quad \text{neg} \\
\quad \text{V}
\]

Let us see how this analysis accounts for both negative formation in general and Infl licensing in particular. Beginning with periphrastic structures like (47), Neg proceeds by head-to-head movement to Infl, adjoining to it and thereby creating the head-checking domain where the latter is licensed. A further movement, required perhaps by scopal properties of negation, will take this unit to the next head up, left-peripheral C. This account of (47) is represented in (49):

\[
(49) \quad \text{CP} \\
\quad \text{C} \quad \text{IP} \\
\quad \text{ez-du} \quad \text{Jek} \quad \text{I} \\
\quad \text{NegP} \quad \text{I} \\
\quad \text{VP} \quad \text{t} \quad \text{t} \\
\quad \text{NP} \quad \text{V} \\
\quad \text{liburua} \quad \text{irakurri} \\
\quad \text{book} \quad \text{read}
\]

(12) This assumption is based on the existence of patterns where the negative element does not move to C, such as relative clauses, and where the order of heads is as in (48):

(i) liburua irakurri ez d-en gizona
    book read neg aux-that man
    'the man that has not read the book'

See Laka (1990) for an alternative view of negation in Basque, involving a left-headed Sigma Phrase projection above IP.
A similar process takes care of the synthetic examples in (46), with the difference that here V is a morphemic root that must be licensed itself by Infl. To satisfy its own morphological requirements, V must amalgamate with Infl. In order to do so, it must skip over Neg to Infl, producing the tensed form _dator_. This is then licensed by Neg-movement to V/I which adjoins the negative element to the tensed amalgamated form. Movement of V to Infl proceeds over the intervening Neg head, without ungrammatical results. Similar movements have been proposed independently for different languages (see, for instance, Belletti 1990). LHM itself falls under this type of ‘skipping’ head-movements which seem to violate ECP, from which the Head-Movement Constraint is usually derived. However, their availability can be explained if, following Roberts (1992), we assume that Relativized Minimality is sensitive to the distinction in L-relatedness (see Chomsky and Lasnik (forthcoming)) between V/Infl on one hand and Neg/C on the other. In this approach, movement of V over Neg (non L-related) to Infl (L-related) does not violate ECP because the type of the target head is different from that of the skipped head.

The complex head resulting from this head-movement operation will move to C to satisfy scopal properties of negation. The analysis is very similar to the one proposed for Breton by Rivero, as shown in the following representations:

\[
\text{(50) a. CP} \quad \begin{array}{c}
C \\
\text{IP}
\end{array} \quad \begin{array}{c}
\text{NegP}
\end{array} \\
\text{lenn}_i \\
\text{ket}
\]

\[
\text{b. CP} \quad \begin{array}{c}
C \\
\text{IP}
\end{array} \quad \begin{array}{c}
eg \text{NegP}
\end{array} \\
\text{Jon.}
\text{neg-comes}
\text{Neg'} \\
\text{VP}
\text{t}_i \text{ Anna t}_j \text{ al levr}
\text{etxetik t}_i \\
\text{from home}
\]

In both languages, V moves first over Neg to I. The second movement differs in the two languages: in Breton, the Neg head skips over I to C, for scopal reasons. This movement is again possible due to the differences in L-relatedness between the heads involved in it. Non L-related Neg skips over L-related Infl to land in non L-related C. In Basque, on the other hand, the same derivation discussed above for periphrastic verbs obtains here: the negative head adjoins to the synthetic amalgamated form and the neg+verb unit moves on to C. Why should Breton and Basque part company in the second part of the movement, after proceeding alike in the first one? The answer could lie in the type of Infl licensing difference proposed here: Breton requires the establishment of a complement domain for Infl. By moving to C over I, negation both satisfies its own scopal requirements and licenses Infl. Basque...
Infl requires a head domain, and neg must proceed by head-to-head movement to Infl. A further movement of ez-V/Infl to C is required to establish the scopal domain of the neg head. In negative wh-questions the same process takes place: neg licenses Infl and since the latter contains the [wh] feature, neg+Infl moves to the head C:

(50) Zer ez du Jonek irakurri?  
What neg aux Jon read  
What hasn’t Jon read?

Notice that in negative questions, the lexical verb remains behind and does not raise to C along with the other heads. This indicates that the negative head can lexicalize/license Infl.

4.3. Embedded patterns

As described above, ba-support is found in matrix focalization and yes/no questions. As for embedded contexts, ba-support does not take place in indirect questions, but it is found in embedded focalized verb constructions. I have shown above that the latter one is the only embedded structure of the ones we are considering where V/Infl move to C, and we have derived it from an extended version of Rizzi’s wh-criterion. If this is correct, we can still claim that ba-support applies whenever a synthetic form cooccurs with an empty operator in CP. I will turn to ba-support itself in the following section. Here, I will examine yet another instance of head-checking. A difference between matrix and embedded declarative patterns is that while a matrix clause may be composed of a single periphrastic verb but not a single synthetic verbs, both minimal clauses are acceptable in embedded contexts:

(51) a. pro etorri da  
come aux  
‘He has come.’  
b. *pro dator  
comes  
‘He is coming.’

(iii) Ez zuen inork ikusi.  
anymbody neg aux see  
‘Nobody saw it.’

This analysis of neg-in-C also accounts for the licensing of NPIs. As noted in Laka (1990), the asymmetry subject/object found in the English NPIs in (i) is not found in Basque:

(i) a. *Anybody didn’t see John  
‘He didn’t see anybody.’  
b. John didn’t see anybody.

(ii) a. Ez zuen inork ikusi.  
anymbody neg aux see  
‘Nobody saw it.’

Presumably neg in (ia) does not c-command the NPI and this structural relation must exist between the licenser and the NPI. In Basque, on the other hand, neg in C has scope over everything. Incidentally, an NPI may also precede the negation, as in (iii):

(14) There is a certain redundancy: TNS would be licensed by Neg, and also by the wh-word in Spec, CP. Nonetheless, none of the movements described here for negatives are prompted by the requirement to license TNS. Rather, licensing is a side-effect of independently motivated processes: lexicalization and the Operator Criterion.
c. Badakit \[ pro datorr-ela/etorri ci-ela \]
   I know (he) comes-that come aux-that
   ‘I know that he is coming/has come.’

d. Ez dakit \[ pro datorr-en /etorri ci-en \] \(^{15}\)
   ‘I don’t know whether he is coming/has come.’

This reminds one again of the Breton case, where V1 patterns with inflected verbal forms are grammatical in embedded clauses:

\[
(18) \text{Lavaret en deus [he deus desket he c’hentelioù]}
\]
\[
\text{said 3S have 3S have learned her lessons}
\]
\[
‘He has said that she has learned her lessons.’
\]

In (18), repeated here, the embedded clause maintains the aux-participle order, while the participle must be preposed in the matrix clause. Rivero accounts for this by showing that embedded Infl is always in the internal domain of C. One might try to claim that something similar happens in Basque. Remember, however, that the same synthetic forms are deviant in the embedded CP when the verb is emphasized:

\[
(52) Uste dut \[ *(ba) datorr-ela \]
\]
\[
\text{think aux comes-that}
\]
\[
\text{I think that (he) IS coming}
\]

Here the verb in C would be in the internal domain of the matrix verb, but still it is not licensed, indicating such licensing is not active in Basque. The analysis I will propose relies on head-checking here too. Before I begin the exposition, though, it is important to spell out some assumptions on Basque complementizers. As indicated in the introduction, most complementizers in Basque appear cliticized to the tense-bearing element. I assume that unless a verbal head moves to C due to independent reasons and attaches to the complementizer affix there, the latter will have to undergo lowering by affix-hopping to Infl.\(^{16}\) This produces two different structural analyses for the same sequence datorrela ‘that he comes’:

\[
(53) \begin{align*}
\text{a. CP} & \quad \text{b. IP} \\
\text{\hspace{2cm}raising} & \quad \text{\hspace{2cm}lowering}
\end{align*}
\]

If the head-checking domain is defined as in (42) above, it corresponds only to those cases of Infl which dominate Infl itself and the licenser adjoined to it, that is (53b) but not (53a). This means that Infl in situ is licensed by the lowered complementizer, but that the complementizer may not license Infl if the latter raises to C, as it

\(^{15}\) As with (40a), the sentence sounds better if \text{ala et} ‘or not’ is added to the embedded verb (as indicated also in note 9).

\(^{16}\) See Rivero (1993b) for a similar analysis of Bulgarian -li.
must in embedded focalizations. Thus, Infl in (52) must still be licensed by ba-support, while that in (51) is already licensed by C.17

5. Ba-support and specifier-head licensing

From everything we have seen up to now, it seems that while Infl in periphrastic forms can always be licensed by a head, synthetic forms are not, except in embedded contexts. However, we have seen that synthetic forms in second position are licensed by lexical elements occupying the Spec of CP position, and this is yet another similarity between Basque and Breton:

\[(54)\] Piv en deus lennet al levr?
who has read the book

who knows that knows Jon that
‘Does John know that?’

Rivero places the wh-word in Breton in SPEC,CP but assumes that Infl remains in situ. The latter is then licensed within the internal domain of C. Given the linearity problems discussed above for Basque, we must consider that V in (55) actually occurs in C:

\[(56)\] a. Breton b. Basque

\[
\begin{array}{c}
\text{CP} \\
\text{wh} & \text{C'} \\
\text{C} & \text{IP} \\
\text{Infl} & \text{V/I}
\end{array}
\]

We seem to require a different modality of licensing for Basque here: Infl-TNS may be licensed (in CP) not only by an adjoined head, but also by another element in its checking domain: a lexical specifier, so that (55a) is acceptable, while (55b) is not. If this is the case, we have a way of understanding ba-support: this would be a last resort language specific rule spelling out the null operator in Spec of CP to allow Infl licensing. It would correspond to the overt wh-word in (56b). The fact that this is a spell-out rule of a specific element accounts for the restriction of its availability only for Spec of CP. If it were a head inserted to provide a head-domain, we would expect it to be able to apply to Infl in situ. However, any verb with ba- prefixed to it is interpreted as emphasized or questioned, depending on the intonation, not as a

(17) Notice that given the lowering analysis, this type of licensing by head-checking is actually quite parallel to the internal-domain licensing from C to Infl in Breton.

(18) Again, perhaps not in northern dialects, as pointed out in note 6.
neural statement,\textsuperscript{18} which indicates that \textit{ba}- cannot be inserted unless an affective feature has prompted movement to \textit{C}. Moreover, we would still require this mode of ‘specifier licensing’ independently to account for (55a), where the licenser is the overt \textit{wh}-word. The same considerations extend without further modification to focalization patterns, where \text{Spec} of \text{CP} is occupied by either an overt focus or a null emphatic operator, which can then be spelled out as a last resort language specific rule to license \text{Infl}. Specifier licensing, then, can be represented as in (57):

\begin{equation}
\text{(57) Specifier licensing:}
\begin{align*}
\text{CP} & \quad \text{wh} \quad \text{C'} \\
& \quad \text{foc.} \quad \text{OP/BA} \quad \text{C} \quad \ldots
\end{align*}
\end{equation}

The specifier is occupied by a \textit{wh}-word, a focus or an empty interrogative or emphatic operator, spelled out as \textit{ba}.\textsuperscript{19}

5.1. \textit{Ba}-support with left-dislocation.

At the outset of this article, I pointed out that both in Basque and in Breton left-dislocated elements do not count as ‘first’ elements, as shown in (29) through (31), repeated here:

(29) \text{Yann, roet meus al levr dezhan.} \\
\text{Yann given have to him} \\
\text{‘Yann, I’ve given the book to him.’}

(30) \text{*Yann, meus roet al levr dezhan.} \\
\text{have given}

(31) a. \text{*Jonek, daki hori} \\
\text{John knows that} \\
b. \text{*Joni, dakar-kio-t hori} \\
\text{John-dat carry-3D-1E that} \\
\text{‘To John, I bring that.’}

This follows from the analysis, if, as assumed also for Breton, left-dislocated phrases are not in \text{SPEC},\text{CP} but adjoined to the highest functional projection of the clause or in the Spec position of a \text{TOP(IC)} phrase above \text{CP}, as in (58). This means, in the case of Basque, that the synthetic forms in (31) appear in \textit{C} but are not in a specifier-head relationship with the topic:

\begin{itemize}
\item (19) Although I don’t deal directly with \text{Infl} licensing in neutral contexts, similar mechanisms seem to obtain. Participles movement to \text{Infl} would head-license the latter. Synthetic verbs require to be specified by a lexical element. As discussed in connection with sentence (45b), non-lexical \textit{pro} does not serve this purpose. Since \textit{ba} is not available outside Spec of CP, such structure is ruled out.
\end{itemize}
It is noted in Basque grammars that whenever a clause contains a topicalized element, some element in the clause must be focalized. If no overt focus appears in these sentences, it then must mean that an empty emphatic operator is present, and this can be spelled out as *ba-, which licenses Infl, as in (32), repeated here again:

   'As for John, he does know that.'
   'As for John, I AM bringing this to him.'

5.2. Pre-Infl particles

The final parallelism between Breton and Basque noted above is the one found with certain particles, like the yes/no marker *ha which does not count as an Infl licensor when preceding the latter. The data are repeated here:

(21) a. *Ha a raint soñjal er bleuniou?
   Q aux think the flowers
   'Will they think of the flowers?'
   b. Ha soñjan a raint er bleuniou? (LHM)

   Q knows John
   'Does John know?'

Rivero places the Breton particle *ba in a functional projection dubbed Question Phrase above CP. Since Infl is in the complement domain of C, but not of Q, LHM must apply, taking the participle there, as in (21b). Then, C filled by the participle is functional and defines the internal domain where Infl is licensed. The situation in Basque, though apparently similar, is actually quite different. The particle al belongs to a set commonly referred to in Basque grammars as epistemic particles, including elements like rhetoric ote, quotative omen etc. These always appear with tensed verbal forms, and always precede them, in whichever position the latter might be:
The particle *omen* in the previous examples occurs in its basic position in (60a); it shows up moved along Infl and the negative element in (60b), and moved along Infl and participle to C in the interrogative sentence (60c). When preceding clause initial Infl, it behaves yet again like Infl itself and it must be licensed in the same fashion:

(61) a. Etorri omen dira.
   come ap. aux
   'Apparently, they have come.'

b. *Omen datoz.
   ap. come
   'Apparently they are coming.'

c. Ba omen datoz.
   'Apparently, they ARE coming.'

Although the exact position of these particles is far from clear (see Albizu 1991 for a discussion and a proposal), what should be stressed here is that inflectional particles behave just like another morpheme of the inflectional complex, and are licensed along with it. Notice for instance that a significant difference between Breton *ba* and Basque *al* is that the former precedes the participle moved to C by LHM and is therefore external to CP, while the latter follows *ba*, as expected if they are elements of Infl. For our purposes here, it suffices to say that they behave like Infl and are licensed along with the latter in the same fashion as discussed throughout this paper.

6. Tenseless finite clauses

Before concluding, it is worth pointing out that both in Breton and in Basque, what is being licensed seems to be the [+tense] value of Infl or T. Thus, finite forms which are arguably tenseless, such as imperatives, may occur clause initially, as in (62):

(62) Betor aita!
   come father
   'Let the father come!'
(63) a. erregela hori, diodan bidenabar, ...
   rule that say in passing
   'that rule, let me say in passing...'

   b. Demagun hipotesia onargarria dela.
   assume hypothesis acceptable is
   'Let's assume that the hypothesis is acceptable.'

All of the initial forms in (62) and (63) are synthetic, amalgamated verbs. Periphrastic imperatives contain the canonical sequence “participle+tensed auxiliary”, making the acceptability of tenseless finite initial patterns untestable for the most part. However, there are marked patterns where the auxiliary is preposed, eliminating the participle from the ‘first’ position, and producing a configuration where the finite auxiliary form is initial. As expected, the finite form can stand unlicensed in ‘first’ position, and ba- is not used:

(64) Dugun edan!
   aux drink
   'Let's drink!'

7. Conclusion

In this article, I have tried to show that the constraints on verb initial patterns in languages like Basque or Breton can be understood if, following Rivero (1993a); we assume that the morphological feature [tensed] must be licensed in certain specific domains. Tensed forms would be weak in these languages, requiring the existence of an external licenser which sanctions their appearance by occurring in a certain structural relation with them. I have argued that the structural relation between finite form and licenser in Basque can be characterized in terms of Chomsky’s (1992) “checking domain”. Some language specific rules can act as last resort processes to provide the morphological feature with the appropriate licensing domain, and I have shown that this assumption helps us understand the distribution of the preverbal particle ba- in Basque.

References


Horvath, J., in press, “Structural focus, structural case and the notion of feature-assignment”.


