Chapter 1

Introduction

1.1 Overview

This thesis can be summarized in the following two objectives: (i) to develop a new version of the Nuclear Stress Rule (NSR) which overcomes certain problems found with previous versions (see, among others, Chomsky and Halle 1968, Halle and Vergnaud 1987, Cinque 1993, Zubizarreta 1998), and (ii) to explain the so-called preverbal focus position in Basque in terms of the new NSR and prosodic principles imposed on focused and *wh*-phrases.

With respect to the NSR, I argue certain generalizations about stress above the word level can be reduced to two basic syntactic properties of phrases: headedness and branching. The proposal is based on certain crucial insights found in previous work on the topic (Chomsky, Halle, and Lukoff 1956, Chomsky and Halle 1968, Liberman 1975, Liberman and Prince 1977, Halle and Vergnaud 1987, Cinque 1993, Zubizarreta 1998). The work reported in this thesis puts these insights together, resulting in a new version of the NSR, within the formalism of the metrical grid, which makes explicit reference to syntactic structure.

With respect to the preverbal focus position in Basque, I argue, contra much previous work on the topic, that it is not a syntactically defined position. Rather, it is to be explained in terms of certain prosodic conditions imposed on focused phrases. More specifically, focused phrases need to have primary stress in the sentence. The analysis is based on insights about the relation between syntax and discourse found in Cinque 1993, Zubizarreta 1998 and Reinhart 1995. The basic idea is that,

given certain independently motivated hypotheses about Basque syntax, the NSR proposed in this thesis predicts that, in many cases, sentence stress is on the preverbal constituent. Since focused phrases need to have sentence stress, it follows that they must be in the preverbal position. In some cases, the analysis correctly predicts that the focused phrase is not the one preceding the verb, but one containing the verb. I argue that this provides further evidence in favor of this analysis, and against analyses in which the preverbal focus position is defined syntactically.

In the remainder of this chapter, I provide some important background for the analyses to be developed in later chapters, and briefly summarize the main results obtained in this thesis. In §1.2, I give some background on the Ondarroa dialect of Basque, on which the analysis is based. §1.3 is an overview of the formalism for representing stress adopted in this thesis, i.e. the metrical grid. Finally, §§1.4–1.7 provide an overview of the main conclusions reached in later chapters.

1.2 A Note on the Data

One of the most striking properties of Basque is the enormous variety of accentual dialects to be found in it. For instance, Hualde (1997) describes four major dialects, and provides detailed descriptions of twenty-four distinct subvarieties. Since the analysis of the syntax-discourse interface provided in this thesis relies strongly on stress, to provide a detailed account for all dialects would involve work which is well beyond the scope of the present thesis. On the other hand, the main features of Basque syntax, as described in descriptive grammars (see, among others, Saltarelli 1988, Laka 1996), remain the same throughout all dialects. This means that, despite the variety in the stress systems, the main properties of stress at the sentence level are expected to be the same in all of them. That is why I have decided to concentrate on a single dialect, the one spoken in the town of Ondarroa.[1] This allows us to examine in detail

[1]The main reason for not choosing *Batua*, the current standard dialect, is that it does not have an established accentual system. This standard dialect, which is not based on any specific local variety, was created and has been regulated by *Euskaltzaindia* (The Academy of the Basque Language) since the sixties. Therefore, at this point, most speakers of this dialect do not speak it natively. To the best of my knowledge, there has been no systematic attempt to describe the accentual system of

the predictions about Basque syntax and phonology made in this thesis.

Ondarroa is the easternmost coastal town in the western province of Biscay. As in many other towns in this province, its accentual system can be described as a 'pitch accent' system: words and phrases contain certain tone contours whose distribution depends on the placement of stress. The accentual data from this dialect discussed in chapter 2 (and summarized in §1.4 below) is taken mainly from Hualde 1991a, 1996, Elordieta 1997a, and from my own field work (see the introduction to chapter 2 for a more extensive list of references).

To the extent that the syntax of other dialects is the same as Ondarroa Basque, the analysis of sentence stress and focus presented in this thesis can be applied to all of them without major changes. Due to the great variety in accentual systems in Basque, some variation is expected, but the main predictions of the analysis are the same for all dialects. Given time and space limitations, a more detailed account of the facts in other dialects in terms of the theory presented in this thesis is left for future work.

1.3 The Metrical Grid

Throughout this thesis, I shall assume that the metrical grid is the correct formalism to describe the placement of stress (see, among others, Liberman 1975, Liberman and Prince 1977, Prince 1983 and Halle and Vergnaud 1987). In particular, I adopt Idsardi's (1992) version of the metrical grid (see also Halle and Idsardi 1995, Halle 1998, Purnell 1997). In this section, I sketch the main features of this theory.

One of the major insights guiding modern work on stress is Liberman's (1975) idea that stress is not a phonetic feature of segments, but the reflection of a grouping of syllables (or stress bearing units) into higher units called *feet*. In the notation of the metrical grid, this grouping is formalized as follows. Certain segments in the string

native speakers of this dialect. Surprisingly, as noted by Hualde (1997, §5.8), many of those who learn Batua as a second language seem to follow a very specific and regular accentual pattern, even though, generally, no explicit instruction is given on what accentual system they should use. This is true both for native speakers of other dialects, and for those whose first language is not Basque. In any case, at present, there is not sufficient data available on which an analysis could be based.

are designated as *stress bearing units*, i.e. segments which can in principle bear stress. The particular choice of stress bearing units varies from language to language. In the two languages discussed in this thesis, English and Basque, these are all and only the vowels that are syllable nuclei. Stress bearing units are distinguished formally from other segments by projecting onto a separate autosegmental *metrical plane* in the form of *grid elements* (represented as asterisks '$*$'). The metrical plane contains several lines (numbered 0, 1, ...) containing one or more grid elements, each corresponding to some segment in the string of phonemes. The resulting representation has the shape of a grid, and can accordingly be called a *metrical grid*. The level of prominence (stress) assigned to a given segment is a function of the number of lines in which there is a grid element linked to the segment: the higher the line a given segment projects onto, the greater its prominence. In the following representation of the English word *execution*,

```
(1)        *        line 2
        *    *       line 1
        * * * *      line 0
        execution
```

the segment *u* has more prominence than the segment *e* in the first syllable, which in turn is more prominent than the other two syllable nuclei.

Halle and Vergnaud (1987) (H&V) propose that projection to lines higher than 0 is a reflection of the grouping of grid elements into feet. Within each foot, one grid element is designated as the "metrical head" of the foot, which means that it is projected onto the next higher line in the grid. The boundaries of feet are represented with parentheses—')' for the right boundary, and '(' for the left boundary. The example given above could then be represented as:

```
(2)        *        line 2
        (*    *)     line 1
        (* *)(*  *)  line 0
         execution
```

On line 0, there are two left-headed feet. Line 1, on the other hand, contains a single

right-headed foot.[2]

H&V provide evidence for the proposal that grid elements are grouped into feet from the fact that it (correctly) predicts where stress is shifted when a stressed vowel is deleted (or rendered unstressable by some other means). In a left-headed foot, it is shifted to the right, and in a right-headed foot, it is shifted to the left (see Hayes 1995, §3.8.1, for relevant examples). If the grid did not contain feet, as in the simpler representation exemplified in (1), the direction of the shift would have to be stipulated.

The main innovation introduced by Idsardi (1992) (see also Halle and Idsardi 1995) into this formalism is the idea that only one parenthesis is necessary to delimit a foot. A left parenthesis groups into a foot all the grid elements to its right up to the next boundary or end of the string, and a right parenthesis groups into a foot all grid elements to its left up to the next boundary or end of the string. For instance, the example above could be represented as follows:[3]

```
(3)        *        line 2
      (*   *        line 1
       * *)(*) *    line 0
       execution
```

In this representation, there are two feet that are delimited by a single parenthesis: the leftmost foot on line 0 is delimited only by a right parenthesis, and the single foot on line 1 is delimited only by a left parenthesis.

In this formalism, parentheses, rather than feet, are primitives. Thus, the following three representations are different, even though they all result in a single foot containing the same number of grid elements:

```
(4)  a.  (* * **                    b.  * * **)

     c.  (* * **)
```

Of these, only the last one makes sense in H&V's formalism. Halle 1998 gives evidence for the richer set of representations provided by Idsardi's notation, drawing on data

[2]The details of English word level stress are not important for present purposes. See H&V (§7) for details.

[3]See Halle 1998 for a detailed analysis of English stress using this formalism.

from the Leka dialect of Russian. As argued by Halle, certain rules are sensitive to the distinctions that are representable in Idsardi's formalism but not in other theories (see Halle 1998, p. 545–547 for details). Further evidence for the need of this richer formalism is given in Purnell 1997.

So far, we have seen how parentheses have an effect on the projection of the grid, but not much has been said about how parentheses themselves are inserted. Idsardi (1992) and Halle and Idsardi (1995) identify three mechanisms. One is the insertion of a parenthesis to mark some syllable with special properties. For instance, a parenthesis can be inserted to the left or right of a grid element linked to a heavy syllable. This is a way of accounting for the effects of syllable weight on stress placement. Another possibility is that a morphemes can be lexically specified as having a parenthesis inserted to the left or right of some specific syllable within it. This accounts for cases in which stress placement is (partially) lexically determined.

Another mechanism proposed by the authors cited above is insertion of parentheses at the edges of the string. These are termed *edge-marking* rules. These rules have the following general form:

(5) Insert a left/right parenthesis to the left/right of the left-/rightmost grid element.

They are accordingly called LLL, RRR, LRL, RLR, etc.:

(6) a. RRR: Ø →) / ∗__]
∗∗∗∗∗)

b. LLL: Ø → (/ [__∗
(∗∗∗∗∗

c. RLR: Ø →) / __∗]
∗∗∗∗)∗

d. LRL: Ø → / [∗__
∗(∗∗∗∗

In §2, I provide several cases of this type of rule, and discuss their effect on stress placement.[4]

[4]Finally, parentheses can also be inserted iteratively, forming several (binary) feet in a word. Iterative foot construction results in the rhythmic patterns characteristic of languages like English. Since none of the analyses proposed in this thesis need iterative foot construction, I do not provide any relevant examples here. See Idsardi 1992 and Halle and Idsardi 1995 for details and discussion.

To summarize, the stress rules in a given language generate a metrical grid for each string (word, phrase, etc.) by specifying, for each line in the grid, where parentheses are inserted, and what element in each foot (the leftmost or rightmost) projects to the next line.

1.4 Chapter 2: Stress in Ondarroa Basque

In §2, I provide an analysis of the basic facts of the accentual system of Ondarroa Basque within the framework sketched in the previous section. Certain aspects of the data and analysis discussed in this chapter are crucial in order to understand the effects of the NSR in this dialect. Since Ondarroa Basque is a pitch accent language, it is not clear how the notion of 'stress' can be applied in a meaningful way to describe the facts. In fact, Hualde (1991a, §6.1), in the first detailed generative analysis of the pitch accent dialects of Basque, proposes a purely tonal analysis. Following much subsequent work in the literature on Basque phonology (see, among others, Hualde 1991b, Elordieta 1997a), I propose a metrical analysis of the facts, in which the surface distribution of tones is the result of metrical rules and rules that assign tone contours by making reference to the metrical grid. The analysis is based on similar ones proposed by Purnell (1997) for other so-called 'tone' and 'pitch accent' languages.

The hypothesis that the distribution of tones in this language is the reflection of stress and the metrical grid is crucial in the explanation of sentence level prosodic facts, and of their relation to focus. If there were no such thing as stress in this language, the NSR could not apply to it, and notions such as 'sentence stress' would not be relevant. The fact that, as shown in chapters 4–5, the NSR plays a crucial role in explaining the relevant prosodic, syntactic and semantic facts discussed in this thesis provides strong support for this hypothesis.

In Ondarroa Basque, as in other pitch accent languages, there are two types of words. Accented words always contain a drop in pitch starting on some specific syllable, and unaccented words contain a drop in pitch only in certain syntactic environments. The basic facts can be summarized as follows:

(7) a. An accented word contains a drop in pitch beginning on its penultimate syllable.

b. A phrase contains a drop in pitch beginning on its penultimate syllable.

c. A word contains a rise in pitch in phrase initial position and when following an accented word within the same phrase.

d. All syllables between a rise and a fall in pitch are linked to a high tone.

The following are some relevant examples:

(8) a. *Isolated unaccented word:*

barbero + antzako bar beruntza ko
barber BEN.SG

b. *Isolated accented word:*

belarri + antzako be larrixantza ko
ear BEN.SG

c. *Unaccented + unaccented*

gixon andi+a gi xon andi xe
man big.A.SG

d. *Accented + unaccented*

belarri andi+a be la rri an di xe
ear big+A.SG

e. *Unaccented + unaccented + unaccented:*

gixon andi+an ixen+a gi xon andixan ixe na
man big+G.SG name.A.SG

f. *Unaccented + accented + unaccented:*

nire auma+an ixen+a ni re au man i xe na
my grandmother+G.SG name.A.SG

g. *Unaccented+unaccented+accented+unaccented:*

nire ama+an lagun+an txakur+a
my mother+G.SG friend+G.PL dog+A.SG

ni re aman lagu nan txa ku rre

In order to account for these facts, I propose that a rise in pitch corresponds to a left parenthesis on line 0 in the grid, and a fall in pitch corresponds to a right parenthesis on line 0:

(9) $\underline{\sigma}\boxed{\sigma\sigma\ldots\sigma\sigma}\underline{\sigma}$

$$\begin{array}{ccccccc} * & (& * * \ldots & * * &) & * & \text{line 0} \\ \sigma & \sigma & \sigma \ldots & \sigma & \sigma & \sigma & \\ | & | & & | & & | & \\ \text{L} & \text{H} & & \text{H} & & \text{L} & \end{array}$$

The metrical gird for each phrase is computed in two separate cycles: the word and the phrase. In accented words, the right parenthesis is inserted at the word level. In phrases, it is inserted at the phrase level. This accounts for all the falls in pitch that can be seen in the data above. With respect to the rise in pitch appearing at the beginning of some words, it is not clear that there is a uniform way of characterizing them: they are the ones appearing in phrase initial position, and those appearing after an accented word. However, as shown in Hualde 1991a (§6.1), those that do *not* have an initial rise can be characterized in a uniform manner: they are the ones that appear preceded by an unaccented word. Thus, I propose that, at the word level, all words contain a left parenthesis, and that this parenthesis is deleted at the phrase level when the word is preceded by an unaccented word.

This rough sketch of the analysis developed in more detail in chapter 2 is sufficient to account for all the relevant facts. However, nothing has been said so far as to what constitutes a 'phrase' for the stress rules. As shown above, a phrase is necessarily delimited by a rise in pitch at the beginning, and by a fall at the end. Since accented words introduce rises and falls that do not necessarily coincide with phrase edges, the relevant examples which can be used to determine what counts as a phrase for the stress rules are those which contain only unaccented words, i.e. (8a,c,e), repeated below in (10)

(10) a. *Isolated unaccented word:*

barbero + antzako $\underline{\text{bar}}\boxed{\text{beruntza}}\underline{\text{ko}}$
barber BEN.SG

b. *Unaccented + unaccented*

gixon andi+a gixon andixe
man big.A.SG

c. *Unaccented + unaccented + unaccented:*

gixon andi+an ixen+a gixon andixan ixena
man big+G.SG name.A.SG

In these three examples, all of which are DPs, the internal structure of the DP seems to be irrelevant for stress. Thus, the AP *andixe* 'big' in (10b) does not start with a rise, and the genitive DP *gixon andixan* 'man big' in (10c) does not end with a fall. Thus, a DP counts as a phrase for the stress rules, and any phrase contained in it (even other DPs) is irrelevant for stress. In this thesis, I will not have much to say about why Ondarroa Basque (and other dialects) imposes this particular restriction on the application of phrase level stress rules. Nevertheless, as shown in later chapters (specifically, §§4.7, 5.5), this restriction plays an important role in the description and analysis of sentence stress and its relation to focus in this dialect.

1.5 Chapter 3: The Nuclear Stress Rule

Since Chomsky, Halle, and Lukoff's (1956) seminal work on English stress, it is well-known that syntactic structure influences the distribution of primary and lower levels of stress in a sentence. This is reflected in the use of the cycle in the computation of stress. However, within each cycle, stress placement is determined by rules which are based on linear order. For instance, in Chomsky and Halle 1968 (SPE), primary stress within each cycle is assigned to the leftmost or rightmost peak (in compounds and phrases, respectively). This assumption is adopted in some form or another in later proposals which use different notations (e.g. the labeled tree notation in Liberman 1975, Liberman and Prince 1977, and the grid in Prince 1983, Halle and Vergnaud 1987).[5]

[5]In some of these works, some reference is made to syntactic structure in the stress rules, but linear order is still crucial. For instance, in Liberman and Prince's (1977) Compound Stress Rule, primary stress is assigned to the rightmost constituent only if it branches.

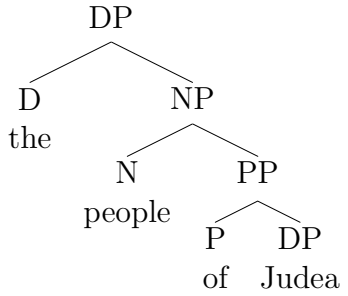
In chapter §3, I argue that, *within each cycle*, the NSR makes direct reference to syntactic structure. In particular, I propose a new version of the NSR in which prominence depends primarily on two aspects of syntactic structure: headedness and branching. The analysis is based on Cinque's (1993) crucial observation that, in many cases, linear order is not necessary in determining primary stress in each cycle. He shows that whether stress is leftmost or rightmost in a phrase is predictable given independent syntactic properties of the phrase. The two main generalizations that he discusses are the following:

(11) a. In a head-complement structure, the complement is more prominent than the head.

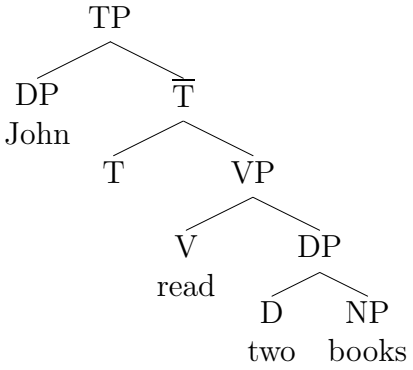
b. In a specifier-$\overline{\text{X}}$ structure, $\overline{\text{X}}$ is more prominent than the specifier.

The following English examples illustrate these generalizations:[6]

(12) a. the people of **Judea**



b. John read two **books**



In (12a), *Judea* has primary stress in the DP. Within the PP, *Judea* is more prominent than *of* because English prepositions do not bear stress when followed by a DP. According to (11a), the complement PP *of Judea* is more prominent than the head noun. Furthermore, this NP is the complement of D, which makes the former more prominent within the DP. As a result, *Judea* is the most prominent word in the DP. Similarly, in the VP in (12b), *books* has primary stress, since this NP is the complement of D, and the DP that contains both is the complement of the verb.

[6]Primary stress in the phrase is represented by using bold face in these examples.

This example also illustrates the generalization about specifiers (11b). The subject *John* is the specifier of the TP, which, by (11b), means that it is less prominent than its sister $\overline{\text{T}}$. Within $\overline{\text{T}}$, VP is the only constituent that contains overt material, so the word *book* within it is the one that has primary stress in the sentence.

The analysis I propose in §3 accounts for these generalizations by making explicit reference to headedness and branching in the syntactic structures involved. I compare this analysis to the one proposed by Cinque (1993), in which it is claimed that the placement of primary stress correlates with depth of embedding.

Cinque argues that the generalization about complements (11a) is a direct consequence of cyclic rule application. In particular, he shows that simple cyclic rule application of stress rules at the phrase level derives the following generalization:[7]

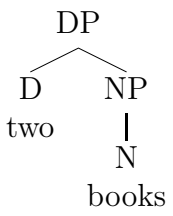
(13) Primary stress in a phrase is on the most deeply embedded word in the phrase.

In cases such as (12a), this derives the generalization about complements. Within the PP, *Judea* is the only word with stress. Within the NP, *Judea* is more deeply embedded than *people*, since the former is dominated by more (phrasal) nodes than the latter. Thus, *Judea* has primary stress in the DP.

In (12b), however, a potential problem arises, since there appears to be no 'most deeply embedded word'. Although both *two* and *books* are more deeply embedded than other words in the sentence, they seem to be as deeply embedded as the other. Standard $\overline{\text{X}}$-theory provides a solution to this problem. In this theory, it is stipulated that everything in a structure that is not a head is a phrase. Thus, the complement of a head is always a phrase. For instance, the structure for *two books* is as follows in this theory:

[7] As noted in the text, this is a generalization that Cinque derives from deeper principles. See §3 and Cinque 1993 for details.

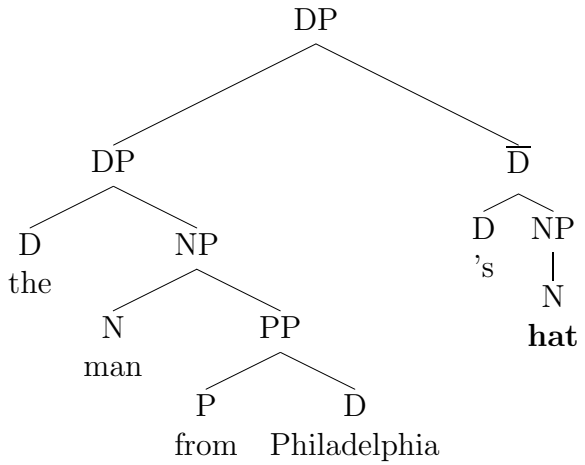
(14)



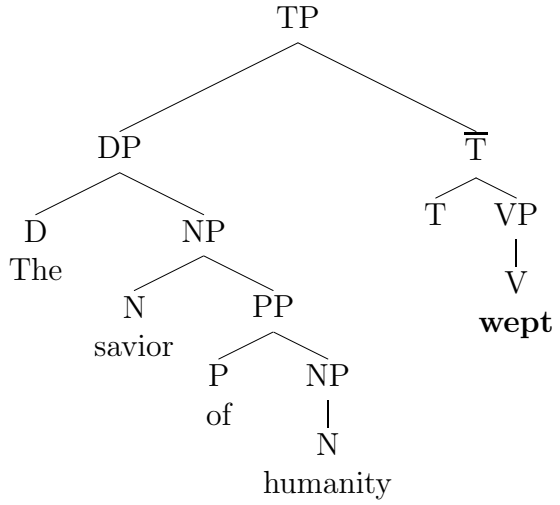
In this structure, the N *books* is more deeply embedded than the D *two*: the former is dominated by NP, which does not dominate the latter. Thus, the complement *books* is more prominent than the head *two*. Further embedding of the DP in head-complement structures, as in (12b), also results in primary stress on *books*, as desired.

As noted by Cinque himself, this creates a problem for the generalization in (11b). A specifier is always less prominent than its sister, even in cases in which it contains the most deeply embedded word in the structure. In particular, in both examples in (15), primary stress is not in the specifier, even though it does contain the most deeply embedded word:

(15) a. the man from Philadelphia's **hat**



b. The savior of humanity **wept**.



In both examples, the most deeply embedded word is contained in the specifier (*Philadelphia* and *humanity*, respectively). However, as expected given generalization (11b), these words do not have primary stress. Rather, primary stress is on the complement. Thus, Cinque's generalization cannot account for the fact that specifiers do not, in general, attract primary stress.

Cinque proposes a solution to this problem by assuming that the internal structure of a specifier is, in effect, 'invisible' for the computation of primary stress in higher cycles. Thus, it is stipulated that depth of embedding within them does not 'count'. To summarize so far, Cinque accounts for both generalizations in (12) by adopting two assumptions: (i) complements are phrasal, and (ii) the internal structure of a specifier is 'invisible' for the computation of stress in higher phrases.

I propose a new version of the NSR which provides a unified account of the two generalizations in (11), and which does not need the additional assumptions about syntactic structure needed by Cinque's version. In particular, I adopt Chomsky's (1995) Bare Phrase Structure, in which the concepts *complement* and *specifier* are derived from the more basic ones *headedness* and *branching*:

(16) *Complements and Specifiers in Bare Phrase Structure*

a. A complement is the sister of a non-branching head.

b. A specifier is the sister of a branching head.

Once we take this into account, a simple algorithm that can derive the two generalizations in (11) comes to mind. These generalizations are repeated below, for ease of exposition:

(17) a. In a head-complement structure, the complement is more prominent than the head.

b. In a specifier-$\overline{\text{X}}$ structure, $\overline{\text{X}}$ is more prominent than the specifier.

Taking (16) into account, these generalizations can be reduced to the following statement:

(18) In a structure of the form $[_{\gamma}\ \alpha\ \beta]$ (order irrelevant), where α is the head of γ, α is more prominent than β iff α is branching.

In a specifier-$\overline{\text{X}}$ structure, $\overline{\text{X}}$ is the head and is branching, so $\overline{\text{X}}$ is more prominent than its sister (cf. 17b). In a head-complement structure, the head is not branching, so the complement is more prominent than the head (cf. 17a). Thus, in §3, I propose that the NSR makes crucial reference to syntactic headedness and branching, and formalize it in terms of the metrical grid (cf. §1.3). This new version does not need the added stipulations assumed in Cinque 1993. Furthermore, as I argue in §3, this new version of the NSR accounts for the stress pattern of different types of phrases in both English and German, and for stress in English compounds.[8]

1.6 Chapter 4: The NSR in Basque

Chapter 4 brings together the results from the previous two chapters in order to account for the distribution of primary stress within sentences in Ondarroa Basque. In neutral contexts, a transitive sentence has the order SOV, and sentence stress is on the object:[9]

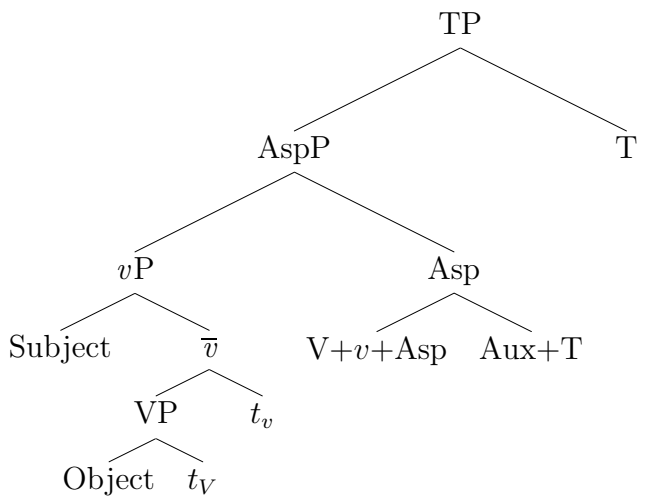
[8] §3 also discusses Zubizarreta's (1998) approach to nuclear stress, which is based on asymmetric c-command. As I show there, this alternative can account for the generalization about specifiers (17b), but not for the generalization about complements (17a). Thus, some additional stipulation, similar to Cinque's (1993), is needed.

[9] By 'neutral context', I mean a context in which the sentence can be understood as an answer to questions like *What happened?*, and in which no consituent is understood as given. See chapter 5 for details.

(19) Mirenek **Jón** ikusi rau.
Miren.E Jon.A see.PRF Aux.PR
Miren has seen Jon.

Following Laka 1990 and Arregi 2000, I assume that Basque sentences have the following basic structure:[10]

(20)



In this structure, the V+v complex moves to Asp, in order to form a *participle*. Furthermore, an auxiliary verb is inserted in T. At PF, the Aux+T complex is lowered to Asp. Following standard terminology in the literature on Basque, I refer to the resulting complex head as the *verb complex*.

The NSR proposed in chapter 4 correctly predicts that sentence stress in cases like this is on the object. First, within $\overline{v}$, the object is the most prominent constituent, since it is the only overt one. Within vP, $\overline{v}$ is more prominent than the subject, since the former is the branching head of vP. Within AspP, vP is more prominent than the verb complex in Asp, since the former is the complement of the latter.[11] Finally, within TP, AspP is the most prominent constituent, since it is the only one containing overt material. The result, as desired, is that sentence stress is on the object.

Since the NSR is dependent on syntactic structure, this analysis predicts that movement can have an effect on the placement of sentence stress. In this chapter, I

[10]This structure corresponds to transitive sentences with compound tenses. See chapter 4 for other types of sentences and for sentences with simple tenses.

[11]The fact that the head Asp is branching is irrelevant, since structure below the word is invisible to the NSR.

discuss two such movements in Basque: left and right dislocation. These are illustrated with movement of the object in the following two examples:

(21) a. [$_{TP}$ Jon [$_{TP}$ **Mirének** *t* ikusi rau]].
Jon.A Miren.E see.PRF Aux.PR
Jon, Miren has seen.

b. [$_{TP}$[$_{TP}$ **Mirének** *t* ikusi rau] Jon].
Miren.E see.PRF Aux.PR Jon.A
Miren has seen Jon.

As shown in these examples, I assume that both movements result in adjunction to TP. The NSR proposed in this thesis correctly predicts that, in both cases, sentence stress is on the subject.

In all the cases seen so far, sentence stress is on the constituent preceding the verb. In sentences in which both subject and object are dislocated, the NSR correctly predicts that sentence stress is on the verbal complex:

(22) a. [$_{TP}$[$_{TP}$[$_{TP}$ *t* *t* **Ikusí** **rau**] Jon] Mirenek].
see.PRF Aux.PR Jon.A Miren.E
Miren has seen Jon.

b. [$_{TP}$[$_{TP}$ Mirenek [$_{TP}$ *t* *t* **ikusí** **rau**]] Jon].
Miren.E see.PRF Aux.PR Jon.A
Miren has seen Jon.

Thus, the interaction between movement and the placement of sentence stress in Ondarroa Basque provides further support for the version of the NSR proposed in this thesis.

A crucial hypothesis in the present analysis is that phrases appearing to the right of the verb are right dislocated (cf. 21b, 22). However, this hypothesis has been challenged in the literature. In particular, Elordieta (2001) claims that there is no rightward movement in Basque. In this analysis, the appearance of phrases to the right of the verb in examples like (21b) above is the result of leftward movement of the verbal complex:

(23) [Subject [[V-Aux] [Object *t*]]]

If this were the right structure, the NSR would predict that sentence stress is on the object, which is contrary to fact. This would force us to propose a language-particular version of the NSR based on linear order rather than syntactic structure. This NSR would state, roughly, that sentence stress is on the constituent preceding the verb. However, this would be little more than a mere description of the facts. On the other hand, the NSR proposed in this thesis explains why sentence stress is on the preverbal constituent. Furthermore, it also explains in a unified way the placement of sentence stress in both Basque and English. Surface differences between the two languages are seen as the result of independently motivated syntactic differences between them. Since the hypothesis that Basque has right dislocation is crucial in accounting for the sentence stress facts, the NSR provides strong support for this hypothesis.

1.7 Chapter 5: Basque Movements and Focus

One of the most studied properties of Basque syntax is its preverbal focus position. In this language, a *wh* or focused phrase (*wh/f*-phrase) must be left-adjacent to the verbal complex. This is exemplified in the question-answer pairs in (24–25).

(24) Q: Jon **señek** *t* ikusi rau?
Jon.A who.E see.PRF Aux.PR
Who saw Jon?

A: Jon **Mirenek** *t* ikusi rau.
Jon.A Miren.E see.PRF Aux.PR
MIREN saw Jon.

(25) Q: **Señek** *t* ikusi rau Jon?
who.E see.PRF Aux.PR Jon.A
Who saw Jon?

A: **Mirenek** *t* ikusi rau Jon.
Miren.E see.PRF Aux.PR Jon.A
MIREN saw Jon.

In (24), the subject is left-adjacent to the verbal complex as a result of left dislocation of the object. In (25), the subject is left-adjacent to the verbal complex as a result of right dislocation of the object. If, on the other hand, there is no left or right dislocation of the object, the subject cannot be a *wh/f*-phrase, but the object can:

(26) Q: *__Señek__ Jon ikusi rau?
who.E Jon.A see.PRF Aux.PR
Who saw Jon?

Q': Mirenek **sein** ikusi rau?
Miren.E who.A see.PRF Aux.PR
Who did Miren see?

A: Mirenek **Jon** ikusi rau.
Miren.E Jon.A see.PRF Aux.PR
Miren saw JON.

In this chapter, I argue that this condition is derived from the following principle:

(27) The F-marked phrase in a sentence must contain the primary stress in that sentence.[12]

(See Chomsky 1971, Jackendoff 1972, Truckenbrodt 1995, Zubizarreta 1998, Reinhart 1995.)

As argued for in chapter 4 (see the summary in §1.6), the NSR predicts that sentence stress in Basque is assigned to the constituent immediately preceding the verbal complex. In (24-25), where the subject is the *wh/f*-phrase, this condition is satisfied by moving the object from its preverbal position. In (26), the object receives sentence stress, so the object, not the subject can be a *wh/f*-phrase. This analysis follows recent works on the syntax of focus in several languages, including Vallduví 1992, Zubizarreta 1998 and Reinhart 1995.

One of the main advantages of the this analysis is that it can account for *focus projection* facts in a unified way in both Basque and English:

[12] As we will see in chapter 5, this principle needs to be slightly modified in order to accommodate certain restrictions on the NSR discussed in chapter 4.

(28) Jonek **káfi** era ban.
Jon.E coffee.A.SG drink.PRF Aux.PST
Jon drank COFFEE.

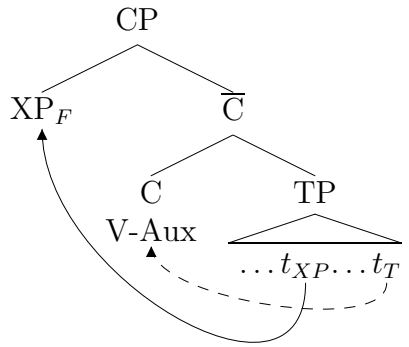
Possible focus readings: Obj_F, $[\text{Obj V}]_F$, $[\text{Sbj Obj V}]_F$

In both Basque and English, this sentence has three possible focus readings. According to (27), the focused phrase needs to contain sentence stress. Sentence stress is on the object, and, accordingly, each focus reading corresponds to some constituent containing the object: the object, object-verb, and subject-object-verb.

In this analysis of Basque focus, the focus projection facts are analyzed in the same way as in English, thus capturing an important crosslinguistic generalization: sentence stress on the object can yield three different focus readings in sentences of this type.

I compare this analysis with previous ones in which it is claimed that Basque has a syntactically defined overt focus position. In particular, Ortiz de Urbina (1989) (see also Elordieta 2001) proposes that the preverbal position is to be analyzed in terms of movement of the *wh/f*-phrase to [Spec, CP]. Adjacency with the verb is the result of movement of the verbal complex to the head of CP, which, by hypothesis, is left-headed in Basque:

(29)



In this analysis, the focus projection facts cannot be analyzed in a uniform way in both Basque and English. Since the hypothesis is that the focused phrase moves to [Spec, CP], it follows that different focus readings correspond to different syntactic structures (i.e. different constituents in [Spec, CP]). Since English does not have (overt) movement of the focused phrase, the facts in Basque and English are not

analyzed in a uniform manner. Thus, this analysis fails to capture an important generalization which is captured in the analysis defended in this thesis.

This chapter also discusses certain cases in which the PF condition introduced above seems to make wrong predictions:

(30) a. Maxe **Jónek** *t* apurtu rau.
table.A.SG Jon.E break.PRF Aux.PR
Jon has broken the table.

Possible focus readings: Sbj_F, $[\text{Sbj V}]_F$, but $*[\text{Obj Sbj V}]_F$

b. **Jónek** *t* apurtu rau maxe.
Jon.E break.PRF Aux.PR table.A.SG
Jon has broken the table.

Possible focus readings: Sbj_F, $[\text{Sbj V}]_F$, but $*[\text{Obj Sbj V}]_F$

These sentences have all the expected focus readings, except the one that includes the left or right dislocated object. It seems that movement of the object 'removes' it from the focus. The result is that the only available focus readings are those which would not be available had the movement not applied. Thus, it is tempting to account for these facts in terms of an economy condition that would restrict movement to cases in which the movement yields a new focus reading.

In this chapter, I argue that this economy condition is not necessary, and that, in the case of right dislocation, it makes wrong predictions. The basic idea is that these movements have certain discourse properties which can account for all the relevant facts:

(31) a. A left dislocated XP is interpreted as a topic.

b. A right dislocated XP is interpreted as given.

If a phrase is interpreted as a topic, it cannot be part of the focus. This explains the fact that the left dislocated object cannot be part of the focus in (30a). Thus, the economy condition is not necessary, at least with respect to left dislocation.

In the case of right dislocation, the present analysis in fact predicts that a right dislocated phrase can be interpreted as part of the focus *as long as it is also interpreted as given*. The following is a relevant example:

(32) Q: Jonek klasi amatxu te gero, se pasa san?
Jon.E class.A.SG finish.PRF and later what.A happen.PRF Aux.PST
After Jon finished the class, what happened?

A: **Jún** ein san Jon.
go.PRF do.PRF Aux.PST Jon.A
Jon left.

In the question, Jon is mentioned. Accordingly, the subject *Jon* is right dislocated in the answer, since it is given. Furthermore, the question is *What happened?*, which means that the answer is interpreted with focus on the whole sentence. Since the answer is felicitous, it follows that the right dislocated subject *is* part of the focus.

Therefore, in certain contexts, as predicted, right dislocated phrases can be understood as part of the focus. In these cases, the economy condition would clearly make wrong predictions.

This chapter concludes with two further issues in the syntax of focus in Basque. First, I discuss certain cases of long distance movement of *wh/f*-phrases, and propose an analysis in terms within the framework adopted in this thesis. The final section of this chapter examines certain predictions made by the current analysis with respect to the scope of left and right dislocated phrases. This section provides further evidence for the existence of right dislocation in Basque, and for the general approach to the syntax of focus in Basque defended in this thesis.