Phrasal stress, Focus, and Syntax*
Maria Luisa Zubizarreta & Jean-Roger Vergnaud
U.S.C., November 2000

Table of contents

0. Introduction.
1. The relation between phrasal prominence and the focus structure of the sentence.
2. Phrasal prominence and informational focus (i-focus) in Germanic.
   2.2. Some general remarks on the Nuclear Stress Rule (NSR).
      2.2.1. The algorithm.
      2.2.2. Secondary stress and eurhythmicity.
   2.3. Cases not covered by the Classical View: in search of a new generalization.
      2.3.1. Wide i-focus structures.
      2.3.2. The narrow i-focus structures.
      2.5.1. Some assumptions and a revision of the NSR.
      2.5.2. The basic examples.
      2.5.3. Beyond the basic examples: the cases of narrow i-focus.
3. A typologically different way of identifying the intonational nucleus: Bengali and Romance.
4. Where does the NSR apply?
5. Some final remarks

0. Introduction.

It has been observed that in many languages of the world there is a correlation between prosodic prominence and the informational structure of the sentence, and much current work attempts to determine the laws that govern this correlation. Although we have made some progress, a lot remains unknown factually and therefore our understanding of the phenomenon is yet precarious.

On the factual side, the difficulty lies on the nature of the data itself. Much of it is based on introspective judgments. Although this is common practice in linguistics, it is not without problems especially in the area under discussion, which consists of native speakers’ intuitions about the contextual appropriateness of a sentence with a certain prominence pattern. The delicate nature of the judgments is twofold. On the one hand, within impressionistic phonetics, the level of stress of a vowel in a given utterance can be a rather elusive feature. It is certainly a difficult feature to identify in complex utterances. It has no unique correlate, such as pitch level, vowel duration, amplitude, or a combination of these.1 And while speakers are generally good at identifying the syllable with the highest pitch, they often lack linguistic awareness of the other acoustic properties. As it turns out, the perception of relative prominence is really bound to the perception of the groupings determined by the surface structure of the utterance.2 The particular difficulty in the case of the association of stress with focus is that the
information structure of the utterance is only partially governed by the rules of syntax. Thus, the determination of stress as a function of the information structure of the utterance or vice-versa does not have the automatic character that its determination as a function of constituent structure has (see note 2).

The second problem with judgments concerning the appropriateness of a certain prominence pattern is that it is very hard to control for context. Strictly speaking, there is no context neutral sentence because speakers tend to (un)consciously add background information. In other words, speakers tend to provide a context for isolated sentences because this is the way that they are naturally used. Furthermore, we need to elucidate which aspect of the relation between prosodic prominence and informational structure is grammatically relevant. Numerous disparate notions and much terminology have been brought to bear on this issue (for a review, see Vallduví and Engdahl 1995), sometimes giving thes area an aura of confusion.

Despite the above shortcomings and while awaiting new experimental tools to better elucidate the facts and better define the object of study, there are some preliminary, core data which have given rise to interesting questions and theoretical debates. The present paper intends to discuss some of these.

1. The relation between phrasal prominence and the focus structure of the sentence.

Let us begin by briefly defining the notion of focus. (on the notion of focus, see Chomsky 1971, Jackendoff 1972, Szabolcsi 1981, Rooth 1985, Rochemont 1986, and Kiss 1998, among others). According to Chomsky and Jackendoff op.cit, the sentence is partitioned into two complementary parts, the presupposition and the focus. The focus is then understood to be the non-presupposed part of the sentence. The most reliable (although not infallible) test that we have at our disposal to determine the partitioning of the sentence into focus and presupposition is the question/answer pair.

In a question like What happened?, the presupposition is that there is an x, x an unspecified event, such that x happened. Roughly speaking, the wh-phrase in the question stands for the focus-variable and the corresponding answer provides a value for that variable. This is illustrated by the examples below, where the focus is the material contained within brackets. In (1), the focus is the TP, in (2), the focus is the VP, in (3), it is the DP object, and in (4) it is the DP subject. We refer as wide focus to the largest focus-structure, namely (1) below, and as narrow focus to all others, namely (2)-(4) below.

(1) Q. What happened?
   A. [TP The cat in the blue hat has written a book about rats].

(2) Q. What did the cat in the blue hat do?
   A. The cat in the blue hat [VP has written a book about rats].

(3) Q. What did the cat in the blue hat write?
   A. The cat in the blue hat has written [DP a book about rats].

(4) Q. Who has written a book about rats?
   A. [DP The cat in the blue hat] has written a book about rats.
In general, then, one can adopt the following operational definition:

\[(5) \text{In a sentence } S = \alpha X \beta, X \text{ is a possible focus just in case } S \text{ would count as a felicitous answer to the question } Q = \text{wh-X } \alpha \beta.\]

One must distinguish two types of focus: *informational focus* (*i-focus*) and *contrastive focus* (*c-focus*). It is the property of exhaustivity that distinguishes the two: an *i-focus* is non-exhaustive, while a *c-focus* is exhaustive. The above wh-fronted questions allow for answers with *i-focus* interpretation, while the clefted questions in French give rise unambiguously to a *c-focus* type.

\[(6) \text{Q. C’est qui qui a écrit un livre sur les rats?} \]
\text{‘It is who that wrote a book about rats?’}
\text{A. C’est [DP le chat] qui a écrit un livre sur les rats.}
\text{‘It is the cat that wrote a book about rats.’}

\[(7) \text{Q. C’est quoi que le chat a écrit?} \]
\text{‘It is what that the cat wrote?’}
\text{A. C’est [DP un livre sur les rats] que le chat a écrit.}
\text{‘It is a book about rats that the cat wrote.’}

The difference between the two types of focus with respect to exhaustivity can be seen in the following contrast:

\[(8) \text{Q. Who wrote a book about rats?} \]
\text{A. [DP The cat] wrote a book about rats, and [DP the bat] did too.}

\[(9) \text{Q. C’est qui qui a écrit un livre sur les rats?} \]
\text{‘It is who that wrote a book about cats?’}
\text{A.*C’est [DP le chat] qui a écrit un livre sur les rats, et c’est aussi [DP la chauve-souris].}
\text{‘It is the cat that wrote a book about rats, and also the bat.’}

Vergnaud & Zubizarreta (to appear) suggest that because contrastive focus is exhaustive in nature, it requires uniqueness of description. And it is precisely this requirement that is violated in the answer form in (9), since the focus-value is given in two separate (conjoined) clauses (the second of which has undergone conjunction reduction). On the logical structures of *i-focus* and *c-focus*, see Vergnaud & Zubizarreta op.cit.

2. Phrasal prominence and informational focus (*i-focus*) in Germanic.


From the earliest work on focus, researchers have noticed that there is a correlation between the locus of main prosodic prominence and the focus structure of the sentence; cf. Chomsky 1971, Jackendoff 1972. Thus, a transitive sentence like (10a), with main prominence on the object, can have the DP object, the VP, or the TP as its
focus, but not the DP subject. In other words, it is an appropriate prominence pattern in the discourse context (1)-(3), but not in context (4). The sentence in (10b), on the other hand, is only appropriate in the discourse context (4), i.e. the DP subject is unambiguously the focus in that sentence.

(10)  a. The cat in the blue hat has written a book about rats.

 b. The cat in the blue hát has written a book about ràts.

From such observations, Chomsky 1971 and Jackendoff 1972 concluded that a constraint like (11) was at work:

(11) The focused constituent must contain the intonational nucleus of the intonational phrase, where the intonational nucleus is identified as the syllable that bears main phrasal prominence.

In the above-mentioned works, it was assumed that main phrasal stress is assigned by the Nuclear Stress Rule (NSR), as formulated by Chomsky and Halle 1968. This version of the NSR basically states that:

(12) The rightmost word-level stress of a phrase bears main stress within that phrase.

If a bottom up, cyclical application of the NSR is assumed, then the rule can account for both primary and secondary phrasal stress in examples like (10a). In this example, the NSR assigns phrasal stress to hat and rats at the DP cycle. It reapplyes to the VP and TP cycles, assigning primary phrasal stress to rats and demoting the stress level on hat to secondary prominence (by convention). In the case of (10b), the constraint in (11) requires that the prosodic pattern be reinterpreted at the TP cycle, shifting main prominence to the DP subject and demoting NS on the DP object to secondary status.

The formulation in (12) raises a number of questions, which we consider briefly in the next section before continuing the discussion of (11).

2.2. Some general remarks on the Nuclear Stress Rule (NSR).

2.2.1. The algorithm.

For the purpose of analyzing the formulation in (12), it will be convenient to adopt the terminology in Chomsky & Halle 1968. Accordingly, we shall call the main stress within some phrase $\alpha$ the primary stress within $\alpha$. The next level of stress below the primary stress, we shall call secondary stress, etc. The formulation in (12) can be viewed as a rule applying to the surface structure of an expression and assigning relative prominence within all the phrases. Within the derivational framework assumed in Chomsky & Halle 1968, this could be done in one of two ways. The prominence assignment could be a one-step process, with the rule (12) applying to all the phrases in a simultaneous fashion. Call this the “one-step” algorithm. Alternatively, the rule (12) could be defined as a “cyclical” rule applying in a bottom-up fashion, first to the smallest phrases in the structure, then to the next bigger ones and so on and so forth (see Chomsky & Halle 1968, Chapter 2 and Chapter 8, pp.340-350 for a discussion). Call this the “cyclical algorithm.” It is clear that the one-step algorithm will yield the right result for subordinate stresses only if, for each subordinate stress, it includes a device that registers the number of domains in which subordination takes place, some function of the depth of
embedding. To illustrate, consider the structure in (13), where \( t, u, v, w \) are words and \( \alpha, \beta \) and \( \gamma \) phrases:

(13) \((\alpha(\beta(\gamma tu)v)w)\)

The neutral stress contour for (13) is the one in (14), where stress levels are indicated by numerals, with “1” representing primary stress, “2” representing secondary stress, etc. (following the convention adopted in Chomsky & Halle 1968):

(14) \( t^4u^3v^2w^1 \)

To derive the contour in (14) by the one-step algorithm, it is necessary to have a counter that indicates the number of phrases containing each subordinate stress in (13). The formal difficulty arises from the conjunction of the two facts in (15):

(15) (i) A “level of stress” is a feature (of words and, ultimately, of vowels or syllables).

(ii) A level of stress is relative to a phrasal domain. Thus, a given word-stress may at the same time be the main stress within a phrase \( \beta \) and a subordinate stress within another phrase \( \alpha \).

The cyclical algorithm provides an elegant resolution of the tension between (15i) and (15ii). We give in (16)-(17) below the formulation put forth in Chomsky & Halle 1968 (see Chomsky & Halle 1968, Chapter 2, pp. 16-18):

(16) Assign primary stress to a primary-stressed word \( w^j \) in the context

\[ w^j \ldots [\alpha \ldots ] \]

where \( \alpha \) is a nonlexical category

(17) When primary stress is placed in a certain position, then all other stresses in the string under consideration at that point are automatically weakened by one.

Rule (16) assigns primary stress to a primary-stressed vowel which is preceded by another primary-stressed vowel in a phrase \( \alpha \). Both rule (16) and convention (17) apply cyclically. The cyclical application of (16)-(17) to the structure in (13) yields a natural derivation of the contour in (14). Facts of this kind constitute empirical evidence in support of the principle of the transformational cycle, stated as follows in Chomsky & Halle 1968:

(18) “... it is natural to suppose that in general the phonetic shape of a complex unit (a phrase) will be determined by the inherent properties of its parts and the manner in which these parts are combined, and that similar rules will apply to units of different levels of complexity.” (see Chomsky & Halle 1968, Chapter 2, p.15)

The conjunction of (16)-(17) and of the principle of the transformational cycle constitutes a particular formalization of rule (12). Recent work within the framework of the Minimalist Program suggests that the derivational history of a sentence and its constituent structure are really two alternative ways of analyzing the sentence; see
Chomsky 1994, Epstein 1999. For concreteness, we shall postulate the following (see Freidin & Vergnaud 1999):

(19) The structural description of a sentence $\omega$ is an abstract analysis $\Omega$ which can be interpreted either as a derivational history or as a phrase-marker.\(^8\)

Assuming (19), we are led to conjecture that there should be a nonderivational correspondent to the cyclical derivational implementation of (12) above. In fact, the proposal in Liberman 1975 appears to be just that. The latter proposal requires that the notion of stress be generalized so that not only words (and phonemes), but also every lexical or phrasal category, will be marked for stress:

(20) “Level of stress” is a feature of phonemes, words and categories.

Given the more abstract notion in (20), the formulation in (12) can be replaced by that in (21), which is essentially the formulation of the NSR in Liberman 1975:

(21) The rightmost immediate constituent of a phrase bears main stress within that phrase.

The formulation in (21) should be construed as a well-formedness condition on surface constituent structure. Quite obviously, that condition must be supplemented by a convention that relates the level of stress of a category to those of its constituents, since stress is ultimately realized on words and, as a phonetic feature, on vowels (or syllables). Such a convention is provided in Halle & Vergnaud 1987, where it is defined in terms of the X-bar theoretic constructs “head” and “projection.” Specifically, one may postulate the following correspondence between the notion “main stress” and the X-bar notion of head:\(^9\)

(22) Given a category $\alpha$, let $H_m(\alpha)$ be the category or word immediately dominated by $\alpha$ that has main stress within $\alpha$. The interpretation of the abstract feature “level of stress” is governed by the following convention:

For any category $\beta$ that dominates $\alpha$, the levels of stress of $\alpha$ and of $H_m(\alpha)$ within $\beta$ are defined to be equal.

Correspondingly, $H_m(\alpha)$ is defined as the metrical head of $\alpha$ and $\alpha$, as the metrical projection of $H_m(\alpha)$.

The “head-of” and “projection-of” relations defined in (22) are transitive in the following sense:

(23) Let $\alpha$ be a category with metrical head $H_m(\alpha) = \delta$. If $\delta$ is itself a category, then, by definition, the metrical head of $\delta$ $H_m(\delta) = H_m(H_m(\alpha))$ has the same level of stress as $\alpha$ in any category that dominates $\alpha$.

Accordingly, the definitions of “metrical head” and of “metrical projection” can be generalized by transitivity as in (24):

(24) Given a sequence $\alpha_1, \alpha_2, \ldots, \alpha_i, \alpha_{i+1}, \ldots, \alpha_n$ such that, for $1 \leq i \leq n - 1$, $\alpha_{i+1}$ is the metrical head of $\alpha_i$ in the sense of definition (22), $\alpha_{i+1} = H_m(\alpha_i)$, then $\alpha_n$ is defined to be a metrical head of $\alpha_1$ and, correspondingly, $\alpha_1$ is defined to be a metrical projection of $\alpha_n$. 
The conventions in (21)-(23), which account for the location of the main stress in each constituent, will in general fail to determine the hierarchy of subordinate stresses in a direct or complete fashion. It is instructive, in that respect, to compare the way these conventions apply to a left-branching structure such as that in (13) with the way they apply to a right-branching structure such as that in (25):

(25) John liked ET.

In the case of the structure in (13), the conventions in (21)-(23) fully determine the stress levels of all the words, yielding the contour in (14). But in the case of (25), these conventions, while assigning a subordinate stress to each of the two words John and liked, fail to define any relative ranking between them. Note that the transformational algorithm in Chomsky & Halle 1968 (see (16)-(17)) generates the contour in (26) below for (25), assigning more stress to John than to liked (see, for example, Chomsky & Halle 1968, Chapter 3 pp. 89-91 and Halle & Vergnaud 1987, Chapter 7, p. 265).

(26) John2 liked3 ET1

It is of course an empirical issue whether a unique stress contour is manifested in the right-branching structure in (25) and, if so, whether that contour is the one in (26).10 Since this article is concerned with the location of main stress within the sentence, in particular in relation to the information structure of the sentence, we will not consider any further the question of stress subordination.11 But a few remarks on the nature of secondary stress are in order.

2.2.2. Secondary stress and eurhythmicity.

Secondary stress (with the exception of “echo accent” discussed below) is fundamentally different from primary stress in that it is sensitive to factors such as length, a phenomenon known as eurhythmicity, rather than to discourse-related factors such as focus.12 As an illustration of this phenomenon, see the examples below (due to Bruce Hayes, p.c.). There are two possible locus for the secondary stress: it may surface on the rightmost edge of the DP subject or specifier (as predicted by the classical NSR) or it may appear at the left-most edge of the DP subject or specifier (contrary to what is predicted by the NSR). In the case of (27) the second option is preferred, while in the case of (28) the first option is preferred. It seems that the difference in preference between the two cases is due to the presence/absence of a break following the DP in question. More precisely, it is more natural to have secondary stress on the right edge of the DP if there is a bit of break after it, but not otherwise. And it seems that it is more natural to have a bit of a break between the verb and its nominative subject than between the noun and its genitive specifier.

(27) a. Nineteen hundred linguists’ dógs
   b. Nineteen hundred linguists’ dógs

(28) a. Nineteen hundred linguists sáng
   b. Nineteen hundred linguists sáng

Note furthermore that a heavier VP has the effect of blocking secondary stress on the right-edge of the DP subject, as shown in (29a). Again this is probably due to the fact that a longer VP calls for an obligatory break immediately following the subject.
Main phrasal prominence, on the other hand, is never affected by factors such as length; it is only sensitive to discourse-related factors. If main phrasal stress falls within the DP subject (as is necessarily the case if the subject is the focus of the sentence), then phrasal stress within the DP is obligatorily where the NSR predicts it should be, namely at the rightmost-edge of the phrase. In effect, while both options in (28) are possible as an answer to the question *what happened?*, only (30a) is possible as an answer to the question *who sang?*. (30b) is only possible if it is understood as introducing a contrast with some other number of linguists.

For this reason, we will ignore secondary stress in this paper, except when a secondary stress is of the “echo-accent” type.

In certain cases, secondary stress is an “echo” of the most prominent pitch accented syllable in the context-question, and for this reason it is referred to as *echo accent* by Zubizarreta 1998. Secondary stress in (10b) as an answer to (4) (repeated in (31)) is a case in point. (We underline words with an echo accent, as well as their source in the context question.) The focus in the answer-statement in (31) is the DP subject; therefore main phrasal prominence falls within the subject. Secondary stress within the defocalized VP is in this case a remnant (or an echo) of primary stress in the context question.

Q. Who has written a book about *råts*?

A. *[The cat in the blue hát] has written a book about *råts*.

Note furthermore that:

A word that bears an echo stress within a phrase is rhythmically subordinate to the word that bears a nonecho stress within the same phrase.

In fact (32) follows from a more general intonational principle to be discussed below (see (49)). We can see that the echo-accent in the answer part in (31) originates in the NS of the context question because its displacement does affect the focus-structure of the sentence. Thus, secondary stress in the answer part in (33) signals the c-focus status of *book*, as it does in the context-question where it originates.

Q. Who has written a *BOOK* about rats? (not a squib)

A. *[The cat with the blue hát] has written a *bòok* about rats.

2.3. Cases not covered by the Classical View: in search of a new generalization.

While the classical formulation of the NSR accounts for the prominence patterns in (10), it fails to account for several others. As shown by Schmerling 1976, Gussenhoven 1984, Selkirk 1984, 1995, in Germanic (including English) not all cases of widest focus take a rightmost stress. We will review many of these cases in 2.3. In fact, the version of the NSR in (12) even fails to account for some basic cases in German (and
Dutch), namely verb final transitive structures where main stress falls on the direct object, rather than on the final verb:

(34) Hans hat ein Buch gelesen.
    Hans has a book read

Note that if we can come up with a successful version of that rule, then we should be able to deduce the scope of the i-focus solely from the locus of the primary nuclear stress. Consequently, we would not need to annotate the syntactic structure with a focus feature. This would indeed be a desirable result to the extent that such a feature does not have any independent justification in the grammar. We will return to this point in the last section.

An NSR-based theory intends to account only for the relation between prominence and i-focus. Therefore, this theory must assume the existence of an independent phrasal stress rule that places main prominence on the c-focus constituent. This rule, generally referred to as emphatic or contrastive, not only serves to identify the c-focus, but also has metalinguistic functions such as correction; cf. *I said CONFirmation, not Affirmation*. The contention that c-focus is or can be identified by emphatic stress (generated by the emphatic stress rule rather than by the NSR) is justifiable on several grounds. Function words are invisible for the computation of the NSR (as we shall see in 2.3) and NS never surfaces on a function word. On the other hand, contrastive stress can surface on a function word; cf. John *DID leave*. Furthermore, contrastive stress is always associated with an audibly higher pitch level than nuclear stress and it is strictly narrow in scope. Compare (4), with NS on *hat*, and (35), with contrastive stress on *blue* (as indicated by capitals).

(35) The cat in the [ADJ BLUE] hat wrote a book about rats (not the one in the RED hat).

In what follows, we will center the discussion only on nuclear stress and its relation to the i-focus structure of the sentence.

2.3.1 Wide i-focus structures.

Schmerling 1976 noted that phrasal prominence in Germanic is sensitive to argument structure. This point was further reinforced in work by Gussenhoven 1984 and Selkirk 1984, 1995. Typically, in non V-final structures in German and in English, NS goes on the rightmost word irrespective of transitivity. See the examples below:

(36) a. Hans arbeitet an einem Papier.
    b. Hans is working on a paper.

(37) a. Hans arbeitet an einem kleinen Tisch.
    b. Hans is working on a small table.

(38) a. Hans arbeitet an einem Papier in seinem Zimmer.
    b. Hans is working on a paper in his room.

(39) a. Das Taxi kommt spät.
    The taxi is-coming late
    b. The mail arrived late.
On the other hand, in V-final structures, an approximate generalization that seems to emerge is that phrasal prominence is sensitive to the transitive/intransitive distinction and that within the intransitive class, the unaccusatives behave differently from unergatives. In German, in a transitive structure like the one in (34a), repeated below, NS goes on the direct object. (This form is nonexistent in Modern English since this language is head initial.)

(40) Hans hat ein Buch gelesen.
    Hans has a book read

Within the class of intransitives, we find that unaccusatives behave differently from unergatives. In unaccusative structures, NS falls unambiguously on the subject. Compare (41) with (39a).

(41) Das Taxi kommt.
    The taxi is-coming

Although there are German speakers that have a preference for the verb in V-final structures to be unstressed (see more on this below), it appears that in the case of unergatives NS may go either on the subject or on the verb (in wide focus structures):

(42) a. Es heisst, dass ein Junge gelacht hat.
    It is-said that a boy laughed has.

    b. Es heisst, dass ein Junge gelacht hat.

Compare (42) with the unaccusative example in (43), where NS falls unambiguously on the subject in wide focus contexts:

(43) Es heisst, dass ein Junge kommt.
    It is-said that a boy came.

Similarly, in English, Schmerling op.cit notes that unergatives (in V-final contexts) allow NS on the subject. But as noted by Selkirk and Gussenhoven, NS may also fall on the verb. Of course, this does not mean that both prosodic patterns are used in the same context. In effect, it would appear that some speakers have a preference for using the pattern in (44) in a surprise-context or in order to highlight a certain aspect of the information being conveyed, namely, the subject of the action. The point is that both options are possible as an answer to What is the matter?, indicating that both prominence patterns are compatible with a wide focus structure. This shows that prominence has multiple discourse and grammatical functions, one of which is to prosodically identify the scope of the focus.

(44) a. The baby is crying.
    b. The baby is laughing

(45) a. The baby is crying.
    b. The baby is laughing.

On the other hand, in comparable structures with an unaccusative verb, NS falls unambiguously on the subject:\(^13\)
(46)  a. The mail arrived.
    b. The sun came out.
    c. My bag has disappeared.

Suppose we were to put aside the pattern of unergatives with initial stress (as in (44)). To account for the unambiguous initial stress in (46), we might capitalize on the fact that the surface subject in unaccusative structures is an object at a deeper level of representation. But the data is in fact more complex, as noted in Gussenhoven op.cit. While NS is unambiguously on the subject in (46c), the presence of an adverb between the subject and the verb pulls the NS onto the verb:

(47) The dog mysteriously disappeared.

Another grammatical distinction that is relevant with respect to the computation of the NSR is the complement-adjunct distinction in V-final structures; cf. Gussenhoven op.cit., Selkirk op.cit. and also Krifka 1984. This is illustrated by the following examples (from Krifka 1984, cited in Truckenbrodt 1993), in which a complement left adjacent to the verb attracts NS (see (48a)), but not an adjunct in the same position (compare (48b) with (37b)):

(48)  a. Peter hat an einem Papier gearbeitet.
       Peter has on a paper worked.
       ‘Peter worked on a paper’

       b. Peter hat an einem kleinen Tisch gearbeitet.
       Peter has on a small table worked
       ‘Peter has worked on a small table’

As can be easily appreciated, the NSR in (12) does not cover the distinctions discussed above. That formulation of the NSR wrongly predicts main phrasal stress on the verb across V-final structures. Another problematic case for the NSR is that of wide focus structures that contain anaphoric material. Anaphoric material is generally pronounced with a reduced pitch range, an intonational property that is often referred to as “deaccenting”. Furthermore, we have the general intonational property in (49) (due to Selkirk 1984, Gussenhoven 1984).

(49) Deaccented material is perceived as less prominent than accented material.

It follows from (49) that main phrasal prominence cannot fall on anaphoric material in Germanic (unless it is contrastive). A few examples are given below, but see Gussenhoven 1984 for many more (deaccented material is highlighted with italics):

(50)  Q: Mary walked in. And then what happened?
    A: John kissed her.

(51)  Remember John? Have you heard from him?
(52) A bill was sent to Congress today by President Carter which would require peanut butter sandwiches to be served at all government functions. At a press conference today, a group of Senators led by Republican Barry Goldwater of Arizona denounced the measure. (Ladd 1980).

Like anaphoric elements, defocalized material is also “deaccented” in Germanic. Therefore (32) also follows from (49). In effect, since echo accents are contained within defocalized constituents and given that defocalized material is pronounced with a reduced pitch range (i.e. it is deaccented), then it follows from (49) that an echo accent will be perceived as subordinate to an accent contained within the focused phrase.

In (50)-(52), as well as in (31), the NSR would have to appeal to an extra mechanism that shifts main phrasal stress to the left sister node. The problem is that this solution misses an important generalization. As we will see in the next section, patterns with deaccented material by and large fall under the same generalization as other patterns. Appending the NSR with a mechanism that shifts stress to the left fails to account for this observation.

2.3.2 The narrow i-focus structures.

The prosodic pattern of narrow i-focus structures follow the same generalization as the wide i-focus patterns. To illustrate this point, we will examine some examples from German (discussed by Höhle 1981). In the answer in (53), the DP subject Karl is defocalized; it is therefore deaccented (as well as the auxiliary hat, given its status as a function word.) This example with narrow focus on the fronted DP object and the verb (den Húnd geschlagen) follows the typical pattern for transitives discussed in the previous section: main phrasal stress falls on the DP object.

(53) Q: What did Karl do?

A: Den Húnd hat Karl geschlagen
The dog (Acc) has Karl (Nom) beaten

Of particular interest are the examples where the object is defocalized, as illustrated in (54). The pattern we see in the answer form A₁, with main prominence on the subject, suggests that transitive structures with a defocalized object behave like intransitives (more precisely, like unergatives). In effect, for speakers that dislike main prominence on the verb when it is in final position, this is the only pattern available (these are the same speakers that dislike (42b)). Such speakers allow prominence on the final verb only if the subject is anaphoric (i.e. if it has been previously mentioned) and is therefore itself deaccented. But interestingly, there are speakers that lack this restriction. These speakers allow the pattern in (42b) alongside (42a) and they also accept the prosodic pattern A₂ in (54). Thus, although there are idiolectal differences, the data seems to be coherent.

(54) Q: Why is Hans happy?

A₁: Weil ihn/Hans ein Verwandter angerufen hat.
A₂: Weil ihn/Hans ein Verwandter angeredufen hat.

The following empirical generalization seems to emerge:
(55) Transitive structures with a defocalized object behave like unergative structures and not like unaccusative structures.

The above generalization is also supported by the following paradigm. Although some speakers strongly prefer for defocalized objects to undergo scrambling, others lack any strong preference. The latter allow the word order in (56), where the object is defocalized, and NS is either on the subject or on the verb:

(56) Q: What happened to the book?

   A1: Ich glaube, dass ein Junge das Buch genommen hat.
       I believe that a boy the book taken has

   A2: Ich glaube, dass ein Junge das Buch genommen hat.

Note that the classical NSR (even if amended with a stress-shifting mechanism that moves NS to the left when it falls on deaccented material) cannot account for the difference in behavior between the different verb classes in wide focus contexts, nor can it account for the generalization in (55).


Cinque 1993 puts forth a revised NSR that is both conceptually and empirically different from the classical NSR. Like the classical NSR, Cinque’s theory assumes that the syntactic tree is the object on which the NS is computed. But his theory of phrasal stress is even more syntax-dependent than the original NSR. It is designed to capture the following basic generalization:

(57) NS falls on the most embedded element on the recursive side of the tree.

Cinque builds upon the metrical version of the classical NSR put forth by Halle & Vergnaud 1987. In a nutshell, like the classical NSR, the metrical NSR applies to the syntactic tree in a bottom-up, cyclical fashion and derives a metrical grid as output by assigning an asterisk to each syntactic node.
The most prominent word within a domain is the one that is dominated by the greatest number of asterisks. Like the classical NSR, *cat* is assigned NS within the DP subject domain and *rats* is assigned NS within the VP domain. We know that *rats* is more prominent than *cat* and that this is achieved by the classical NSR by assigning main prominence to the rightmost category (this time at the TP cycle).

Cinque’s theory, which is sensitive to depth of embedding (as encoded by the number of asterisks stacked up on a word), and not to linear order, needs yet another ingredient in order to account for the following generalization: irrespective of the degree of embeddedness within a specifier, main prominence is contained within its sister category the structures; cf. *The cat that chased the ràt won a prize*. To address this issue, Cinque introduces a distinction between “major” and “minor” paths of embedding and proposes that the algorithm be sensitive to that distinction. Informally, the “major path of embedding” within some tree $U$ is the tree obtained by stripping $U$ off all the specifiers and adjuncts in it. Thus, the “major path of embedding” within $U$ is the connected set of categories which includes the head of $U$ and is such that, if it includes a head $H$, it also includes the projections of $H$ and the categories governed by $H$ (in the sense of GB theory) together with their heads, but excludes the nongoverned categories. Formally, the notion “major vs. minor path of embedding” is defined relatively to a given maximal projection $XP$ (or, equivalently, relatively to a given head $X$). To make explicit the relative character of the notion, we reformulate the definition in Cinque 1993 as in (59), where the clause in (i) defines the primitive notion and that in (ii) extends it by transitivity:  

```
(58)

\[ \text{TP*} \]

\[ \text{DP*} \]

\[ \text{NP*} \]

\[ \text{D} \quad \text{N*} \]

\[ \text{VP*} \]

\[ \text{DP*} \]

\[ \text{NP*} \]

\[ \text{V*} \quad \text{D} \quad \text{N*} \quad \text{P} \quad \text{N*} \]

The cat wrote a book about rats
```
(59) (i) Given some head category \( X \), the major path of embedding with respect to \( X \) is primitively constituted by \( X \) and the projections of \( X \) together with the head of the complement of \( X \) and the projections of that head. The minor path of embedding with respect to \( X \) is primitively constituted by the head of the specifier of \( X' \) and the projections of that head.

(ii) Given heads \( X, Y \) and \( Z \) such that \( X \) belongs to the major path of embedding with respect to \( Y \), whatever status \( Y \) has with respect to \( Z \) is extended to \( X \): if \( Y \) belongs to the major (resp. minor) path of embedding with respect to \( Z \), then, by extension, \( X \) is defined to belong to the major (resp. minor) path of embedding with respect to \( Z \).

When there is no ambiguity, we shall refer to “the major (resp. minor) path of embedding with respect to \( XP \)” simply as “the major (resp. minor) path of embedding.” The convention that Cinque 1993 postulates to account for the metrical asymmetry between the specifier and its sister is the following:

(60) **Convention**

When the minor path joins the major path, the cycle corresponding to the minor path is visible in the form of one asterisk.

Cinque’s algorithm delivers one very important result. Unlike the classical NSR, it is able to account for the presence of NS on the object in German V-final transitive structures. In effect, the head of the complement phrase is more deeply embedded than the head of the verb phrase and therefore it is dominated by a bigger number of asterisks. But as it stands, Cinque’s theory, like the classical NSR, fails to account for the data discussed in 2.3.


This work is an attempt to meet descriptive adequacy, while keeping alive the basic insight in Cinque’s theory: NS falls on the “lowest” element in the (visible part of the) syntactic tree. We begin by presenting some basic assumptions about this analysis.

2.5.1. Some assumptions and a revision of the NSR.

In the spirit of Liberman 1975, the NSR applies to sister nodes and it assigns relative strength to them (i.e. Weak vs. Strong); see (20)-(22) above.

NSR applies to a metrical interpretation of the syntactic tree, rather than to the syntactic tree itself. In other words, metrical sisterhood and syntactic sisterhood do not always coincide. The reason is that anaphoric, defocalized, and functional categories are pronounced with a reduced pitch range, an intonational property, which, as mentioned earlier, is referred to as “deaccented”. And as stated in (49), it is a fact that deaccented material is always perceived as less prominent than any accented material. Consequently, the algorithm that determines which element is the most prominent in the string can only apply to accented material. It therefore follows that:

(61) A syntactic category that fully dominates anaphoric or defocalized material, a functional word, or an empty category is invisible to the application of the NSR.
To illustrate, consider the following tree, where the parenthesized node (D) is metrically invisible because it dominates deaccented material. We then say that in this structure, C and F are metrically non-distinct. Therefore, B and F, as well as B and C, count as metrical sisters.

(62)  
```
   A
  /  \
 B  C
   / \            (D)
  /   \    F
```

Based on the data discussed in section 2.3, the following approximate generalization is assumed. NS in V-final and non V-final structures in Germanic follow two distinct patterns (putting aside the cases of intransitives with NS on the verb, such as (45) and (47)):

(63)  
- In V-final structures NS falls on the right-most selected visible element in the tree.
- In non-V final structures, NS falls on the right-most visible element in the tree.

Suppose we adopt the essential ideas on phrase structure put forth by Kayne 1994. First, syntactic trees are universally right-branching, i.e., they conform to the X-bar form \( [X_{P}Z_{P}[X'X_{0}Y_{P}]] \). Second, languages universally obey the Linear Correspondence Principle (LCA), a weaker version of which is given in Chomsky 1995 and reads as follows:

(64)  
- a. Given two constituents A and B, if A asymmetrically c-commands B, then every terminal node that A dominates precedes every terminal that B dominates.
- b. A asymmetrically c-commands B iff. A c-commands B and B does not c-command A.\(^{17}\)

It then follows that the “right-most” (visible) element in the tree is also the “lowest” one. As mentioned at the beginning of this section, Zubizarreta’s analysis preserves the basic insight of Cinque’s algorithm: NS goes on the “lowest” (metrically visible) element in the tree. Embeddedness and ordering coincide, and for the present purposes they are indistinguishable.

The ordering relevant for the notion “lowest” may be defined along two distinct dimensions of the tree: either with respect to the subpart that defines selectional relations (giving rise to selectional ordering) or with respect to the constituent structure irrespective of selectional considerations (giving rise to asymmetric c-command ordering). For ease of exposition, we depart with respect to a number of details from the analysis in Zubizarreta 1998.\(^{18}\) The notion of asymmetric c-command ordering is defined in (64). Selectional ordering relative to some head \( H \) is then defined as the restriction of the asymmetric c-command ordering to the set of constituents directly or indirectly selected by \( H \):
Let $H$ be a head and let $K_H$ be the set of constituents selected by $H$ or a projection of $H$. Given two constituents $C_i$ and $C_j$ in $K_H$ such that $C_i$ asymmetrically c-commands $C_j$, then $C_j$ is lower than $C_i$ in the selectional ordering.

The NSR is then modularized in terms of these two types of orderings:

Revised NSR:
- **S-NSR**: Given two metrical sister categories $C_i$ and $C_k$ in which one is selectionally dependent on the other, assign main prominence to the dependent category if it is the lowest in the selectional ordering.
- **C-NSR**: Given two metrical sister categories $C_i$ and $C_k$, assign main prominence to the category that is lowest in the asymmetric c-command ordering.

Note that the context to which the S-NSR applies is a proper subset of the context to which the C-NSR applies. The S-NSR will then get a chance to apply only if it is ordered before the C-NSR (cf. the “Elsewhere Condition” on the application of phonological rules put forth by Kiparsky 1973). It is therefore assumed that the order of application is as stated in (66).

We have two tree structures: a metrical tree (which is the syntactic tree with the deaccented nodes omitted) and the full-fledged syntactic tree; see (62) above. When applying the NSR, one of the two structures, the metrical tree or the full-fledged syntactic tree is selected and the ordering relations are assessed within the selected structure, and solely within that structure. It is suggested below that the cases of stress ambiguity (like (44)-(45)) are due precisely to this option.

In the next section, we illustrate how the rule works with some of the basic examples discussed in 2.3. As we shall see, the analysis sketched above accounts not only for the cases that fall under the approximate generalization given in (63) above, but also those that fall outside of it (such as examples (45) and (47)).

### 2.5.2. The basic examples.

Recall that in non V-final structures, such as (36)-(39), NS falls systematically on the right-most element. Let us examine (39). As we have seen in 2.3, the position of NS in unaccusative structures varies when other material such as adverbs is added on. The English example in (39) may be assumed to have roughly the structure below, where the parenthesized nodes are metrically invisible because they dominate deaccented or phonologically empty material. (The asterisks encodes main prominence among metrical sisters, as assigned by the revised NSR. It is equivalent to the “S” label in Liberman 1975.)
Consider a step by step (top-down) application of the rule (a bottom-up application would give the same result):

a) D in line 1 and N in line 2 are metrically non-distinct. Therefore any rule application to D in line 1 extends to N in line 2. D and T are sisters but not selectionally related. Therefore, C-NSR applies, instead of the S-NSR. Since D asymmetrically c-commands T, T is assigned main prominence.

b) T in line 1 and V in line 2 are metrically non-distinct. Consequently, D in line 1 and V in line 2 count as metrical sisters. There is no selectional relation between them; therefore S-NSR fails to apply to these metrical sisters. Instead C-NSR applies and assigns main prominence to V in line 2, given that D in line 1 asymmetrically c-commands V in line 2.

c) V and Adv in line 3 are (metrical) sisters. There is no selectional relation, therefore S-NSR does not apply. Instead, C-NSR applies and assigns main prominence to the Adv since V asymmetrically c-commands Adv.

d) Since the Adv node in line 3 is metrically non-distinct from the Adv node in lines 4 and 5, V in line 3 counts as metrical sisters to Advs 4 and 5. Again, S-NSR fails to apply in each of these cases; C-NSR applies instead to the lowest node in the asymmetric c-command ordering, namely to Adv in lines 4 and 5.

Consider next the unaccusative structures without the adverb in final position. In this V-final unaccusative structure, NS falls on the subject (cf. (41)). The structure is as follows:

Consider next the unaccusative structures without the adverb in final position. In this V-final unaccusative structure, NS falls on the subject (cf. (41)). The structure is as follows:
As before, D in line 1 and N in line 2 are metrically non-distinct. T and D in line 1 (and therefore N) are (metrical) sisters. Since T in line 1 and the Vs in lines 2 and 3 are metrically non-distinct, D in line 1 (and the N it dominates) function as a metrical sister of the Vs as well. Since these are selectionally related, the S-NSR applies and assigns main prominence to D (and N).

We now compare the above unaccusative structure with one which contains an adverb between the subject and the verb (cf. (47)):

(69)

1. \(D_i\) T
2. (D) \(N\) (T) V*
3. Adv V*
4. V* (D)
5. The dog mysteriously disappeared e_i

This metrical structure differs with respect to the previous one in a crucial way. The D in line 1 and the Vs in lines 3 and 4 are not metrical sisters. In this structure, D in line 1 and V in line 2 are metrical sisters but they are not selectionally related. Therefore, the C-NSR applies and assigns main prominence to V in line 2. Adv in line 3 is a metrical sister of the Vs in line 3 and in line 4 (the two Vs being metrically non-distinct). Again, since there is no selectional relation between the Adv and the Vs, the C-NSR applies and assigns main prominence to V.

The last two examples are very relevant to evaluate Cinque’s system. The case of (68) is problematic because in that system we do not expect the minor path of embedding to receive NS. Recall that within that system the non-recursive side of the tree is visible only in the form of one asterisk. We could amend his system to account for this case by appealing to the presence of the subject’s trace within the VP. In other words we could say that the subject receives NS not by virtue of its surface position but by virtue of being a copy of the DP in object position, which is the most embedded node in the tree. But, (69) shows that this cannot be the case. In effect, if the trace of the displaced object plays a role in determining the location of NS in (68), we would not expect the presence of a preverbal adverb to pull NS to the right, as is the case in (69). We conclude therefore that the notion of minor vs. major path of embedding is not the right notion to compute NS.

We turn next to the unergative structures. These are particularly challenging because NS may appear either on the subject or on the verb with a wide focus interpretation (e.g. (44) and (45)). Let us assume, following Hale & Keyser 1991, that unergatives are covert transitives. They have a cognate object which may be syntactically incorporated into the verb. Cf. John slept vs. John slept a long sleep. We show below that the NSR generates two distinct patterns in the case of unergatives depending on whether the full-fledge syntactic tree or the metrical tree (which is a subpart of the syntactic tree) is used to compute ordering.
If "ordering" is computed on the basis of the full-fledge syntactic tree, the trace of the cognate object that has been incorporated into the verb counts for the computation, more specifically it counts for the computation of selectional ordering. In this case, the pattern in (45) is obtained: NS falls on the verb.

(70)  
1. \[D \quad T^*\]  
2. \[(D) \quad N \quad (T) \quad V^*\]  
3. \[V^* \quad (D_i)\]  
4. *The baby is crying +D\_i e*

The computation in (70) proceeds as follows. As in the structures examined earlier, D and T in line 1 are sisters; C-NSR applies and assigns main prominence to T. T in line 1 is metrically non-distinct from the V nodes in lines 2 and 3. Therefore D in line 1 and the V nodes in lines 2 and 3 are metrical sisters. While D in line 1 and the Vs are selectionally related, D in line 1 is not the lowest element in the selectional ordering; D in line 3 is the lowest element in the selectional ordering. Therefore NSR fails to apply; C-NSR applies instead and assigns main prominence to the V nodes in lines 2 and 3, since these nodes are asymmetrically c-commanded by D in line 1.

If ordering is computed solely on the basis of metrically visible nodes, then the trace of the incorporated object does not count for the computation. The pattern in (44) is obtained: NS falls on the subject (as in the unaccusative structure in (68)).

(71)  
1. \[D^* \quad T\]  
2. \[(D) \quad N^* \quad (T) \quad V\]  
3. \[V \quad (D_i)\]  
4. *The baby is crying +D\_i e*

In (71), as in (70), T in line 1 and the Vs in lines 2 and 3 are metrically non-distinct; D in line 1 and the V nodes are therefore metrical sisters. The two set of nodes are selectionally related. Given that the silent object is invisible to the computation, D in line 1 counts as the lowest selected node in the tree. The S-NSR therefore applies and assigns NS to it.

The analysis sketched above for English readily extends to German V-final and non V-final structures. Recall that not all German speakers accept the unergative
structure with NS on the final V node; see the discussion with respect to (42a) and (42b).

Given the above analysis, we can account for this idiolect by assuming that for such speakers ordering relations is unambiguously determined in terms of the metrical tree.

An interesting contrast that surfaces in German V-final structures and which is unavailable in English is the one in (48), repeated below.

(72) a. Peter hat an einem Papíer gearbeitet.
    Peter has on a paper worked.
    ‘Peter worked on a paper’

b. Peter hat an einem kleinen Tisch geárbeitet.
    Peter has on a small table worked
    ‘Peter has worked on a small table’

In (72a) the PP immediately preceding the lexical verb is selected by the verb and therefore the S-NSR assigns it main prominence. On the other hand, in (72b) the PP that immediately precedes the lexical verb is an adverb which bears no selectional relation with the verb. Consequently, the S-NSR does not apply; the C-NSR applies instead assigning main prominence to the verb, the lowest node in the asymmetric c-command ordering.19

2.5.3. Beyond the basic examples: the cases of narrow i-focus.

The case of (53), repeated below, is straightforward.

(73) Q: What did Karl do?
    A: Den Húnd hat Karl geschlagen
    The dog (Acc) has Karl(Nom) beaten

The subject that intervenes between the fronted object and the verb is metrically invisible. Therefore the fronted object and the verb are metrical sisters. The S-NSR applies and assigns main prominence to the object.

The case of (54) and (56) are more challenging. We repeat (54) below.

(74) Q: Why is Hans happy?

In transitive structures where the object is defocalized, NS falls either on the subject or on the verb, as in the unergative structures examined earlier. Why should this be the case?

The analysis of unergatives suggested above provides a straightforward answer. Unergatives are covert transitive structures. And, as in the case of unergatives, if ordering is assessed on the full-fledge syntactic tree (including the defocalized object and its trace), then the S-NSR fails because the subject is not the lowest selected element in the tree. The C-NSR applies and assigns main prominence to the verb. Alternatively, if ordering is assessed with respect to the metrical nodes (which exclude the defocalized object and its trace), then the S-NSR applies. The subject is the lowest selected element in the metrical tree and is therefore assigned main prominence.
Recall that there is an idiolect of German that dislikes NS on the verb in cases like (74). As mentioned earlier, these seem to be the same speakers that reject NS on the verb in unergative structures like (42b) (unless the subject itself is defocalized). The same explanation applies in both cases. These speakers compute ordering relations solely on metrically visible nodes; therefore the subject will inexorably be assigned main prominence by the C-NSR in such structures.

3. A typologically different way of identifying the intonational nucleus: Bengali and Romance.

We know that main phrasal stress plays a crucial role in identifying the intonational nucleus of the intonational phrase (or I-phrase) and that the intonational nucleus is the center around which the intonational contour is organized. The studies that assume the classical NSR, as well as some of those that have attempted to revise it (e.g. Cinque 1993, Zubizarreta 1998), have recognized that syntactic information plays a crucial role in the computation of main phrasal prominence (and therefore of the intonational nucleus) in the Germanic languages. But this does not seem to be universally true. A case in point is that of Bengali, described by Hayes and Lahiri 1991.

In Bengali, as in Germanic, the locus of NS plays a role in determining the possible scope of the focus. But the syntax does not play any direct role in the computation of NS (and therefore of the intonational nucleus). Instead, NS in Bengali is computed in terms of prosodic phrasing, which itself is constrained by the syntax (see further below).

In Bengali, there are two types of prosodic phrases, as identified by Beckman and Pierrehumbert 1986 for Japanese and English, the intermediate prosodic phrase (also referred to as phonological phrase or P-phrase) and the intonational phrase or I-phrase, mentioned earlier. The I-phrase is composed of one or more P-phrases. Hayes and Lahiri (p. 57) identify the following phrasal stress rules:

(75) **Bengali P-phrase Stress**

Within the P-phrase, the leftmost non-clitic word is the strongest.

(76) **Bengali I-phrase Stress**

a. A P-phrase bearing narrow focus receives the strongest stress of its I-phrase.

b. Under wide focus, the rightmost P-phrase within the I-phrase is the strongest.

Furthermore, Hayes and Lahiri show that:

(77) In Bengali a p-phrase boundary must be inserted at the right edge of the focused constituent.

Given (77), we may rephrase (76) as follows:

(78) Within the I-phrase, NS falls on the right-most P-phrase, the right edge of which must coincide with the right-edge of the focused constituent.

In the case of wide-focus structures and in the case of a narrow focused constituent in final position, NS will fall within the last P-phrase of the I-phrase. In these cases, the last phrasal stress will be the strongest. In cases where the narrow focused constituent is internal within the I-phrase, then NS will not be rightmost.
To repeat, in Bengali, the position of nuclear stress is not sensitive to syntactic structure; instead it is sensitive to prosodic phrasing. On the other hand, syntactic structure is relevant in constraining the formation of p-phrases. Hayes & Lahiri (p. 91) report that:

(79)  
a. Two constituents X and Y may be grouped into a single P-phrase only if X c-commands Y.

b. The verb, which is in final position, is phrased separately.

We may think of Romance languages as belonging to the same typological class as Bengali. But unlike Bengali, the p-phrase in Romance is not relevant for the computation of the NS. More precisely,

(80)  
In Romance, NS falls on the rightmost accented word within the I-phrase

Zubizarreta 1998 shows that languages like Spanish and Italian do not deaccent anaphoric and defocalized constituents. On the other hand, these languages (like Germanic and Bengali) comply with the focus/prosody correspondence principle in (11), which we repeat below:

(81)  
The focused constituent must contain the intonational nucleus of the intonational phrase, where the intonational nucleus is identified as the syllable that bears main phrasal prominence.

Zubizarreta 1998 argues that Spanish and Italian resolve cases in which (80) and (81) clash by movement (which she terms p-movement, meaning that it is a type of movement that is prosodically motivated). More precisely, when the last constituent in the I-phrase is defocalized, the defocalized constituent is fronted in order for NS to be able to fall on the right-most constituent within the I-phrase, while complying with the requirement in (81). An example is provided below (where the focused constituent is the direct object). The canonical word order SVOPP is not acceptable in this context, as shown in A1 and A2. The focused direct object must appear at the right-most edge of the I-phrase to receive NS. For this reason the defocalized PP must be fronted, as shown in A3. Alternatively, the PP may be right-dislocated, in which case it is outside of the I-phrase, as shown in A4.20

(82)  
Q: Qué puso Juan sobre la mesa?

   What put Juan on the table
   ‘What did Juan put on the table?’

*A1: Juan puso el pán sobre la mesa.
   Juan put the bread on the table.
*A2: Juan puso el pan sobre la mesa.
   A3: Juan puso sobre la mesa el pán.
   A4: Juan puso el pán # sobre la mesa.

Note that the Germanic NS algorithm share something with the Bengali type system at a more abstract level. In Bengali and Romance it is the linear notion “right-most” within the I-phrase that counts. In Germanic, it is “being the lowest” within the syntactic tree or a subpart of it that counts. Interestingly, the LCA (stated in (64)) establishes that the phonological notion of “right-most” and the syntactic notion of
“lowest” coincide. In fact, in Romance the “rightmost” node in the I-phrase and the “lowest” node in the asymmetric c-command ordering are indistinguishable. But why do Germanic and Bengali differ? Recall that in Bengali the focused constituent is obligatorily flanked by a p-phrase boundary. On the other hand, in English, as shown by Beckman & Pierrehumbert 1986, the focused constituent is not systematically followed by a p-phrase boundary. If that turns out to be the case for German as well, then we could advance the hypothesis that therein lies the source for the divergence between the two language types. More precisely,

(83) A system in which a focused phrase is not systematically flanked by a prosodic boundary computes prominence in terms of the syntactic notion of “lowest”, rather than in terms of its phonological counterpart, namely the notion of “rightmost”.

Future research on prosodic prominence, prosodic phrasing, and their relation to syntactic structure should say whether or not (83) is on the right track. If (83) turns out to be correct, it would lend support to the hypothesis that “rightmost” and “lowest” are the two sides of the same coin, as suggested by the LCA.

4. Where does the NSR apply?

It has been suggested above that there are typologically two distinct classes of languages: one (like Germanic) in which the metrical interpretation of the syntactic tree is relevant in computing NS and another one (like Bengali and Romance) in which NS is computed strictly in terms of prosodic phrases. What can we infer from this conclusion regarding the locus of application of the NSR? Does it mean that NS applies within the syntax in Germanic and within the phonological component (or PF) in Bengali and Romance? While tempting, this conclusion cannot be correct under the standard assumption that PF and Logical Form (LF) are systematically mediated by syntax.

First, as we have seen, prominence plays a fundamental role in identifying the focus domain of a sentence in both language types (cf. the principle stated in (11), repeated in (81)). To the extent that focus is part of the logical representation of a sentence (see Vergnaud & Zubizarreta (to appear)), it follows that information regarding prominence must be available at LF in both types of languages (i.e. in Germanic as well as in Bengali and Romance).

Second, Zubizarreta 1998 has argued that the formulation of p-movement in Romance (see section above) is intimately related to the NSR; therefore, they must both apply at the same point in the derivation. (For details on the formulation of p-movement, see Zubizarreta op.cit.; section 3.5.1.) Furthermore, Suñer 2000 has shown that p-movement in Romance affects binding relations. If binding relations are determined in the LF component of the grammar, then p-movement must feed LF, and so must the NSR.

Third, it was shown above that NS in Bengali (and in Romance) is computed in terms of prosodic boundaries, while in English NS is computed on the basis of metrically interpreted syntactic structures. Nevertheless, even in English, a sentence may be composed of more than one intonational phrase, with each intonational phrase containing a NS (as required by the principle in (11)). See for example, the famous example from Chomsky and Halle 1968:
(84) (This is the cát) (that chased the rát) (that stole the chéese)

The implication is that information regarding prosodic boundaries must be available at the point where the NSR applies (even in English).

Our conclusion is that prosodic information (such as prosodic boundaries and main phrasal prominence) must be present at the stage in the derivation where LF is constructed. That conclusion appears to be broadly in line with the standard derivational architecture. Within that architecture, PF interpretation (“Spell-Out”) applies at the level of S-Structure, from which LF is also derived (see, Chomsky and Lasnik 1993, Zubizarreta 1998). Zubizarreta 1998 develops a particular version of the standard model in which prosodic information is incorporated into S-Structure. In particular, prosodic structure in Germanic-like languages is presented as an annotation of S-Structure, which is produced by interpretive principles that apply to local configurations of the constituent structure (see 2.5. above). But we saw earlier that a constituent structure is a derivation in a different guise (cf. (19) and note 8 in section 2.2.1.). So, the account in Zubizarreta 1998 can presumably be recast and be embedded in an architecture in which PF interpretation is cyclical (see Chomsky 1999, Frampton and Gutmann 1999, Uriagereka 1999). We sketch this alternative rendering below.

Chomsky op.cit. outlines a particular cyclical architecture, in which the cycles are called phases. To illustrate the workings of a derivation by phase, let us consider the structure in (85):

(85) [Jane saw John]

We assume that that structure has five primitive constituents, namely those given in (86), where $C$ is a (silent) complementizer:

(86) \{Jane, John, see, Past, C\}

This is an oversimplification, but it will do for our purpose. The first cycle may be represented as the derivation-marker in (87):

(87)

```
   W
   \ /
  Jane see John
```

That is, Merge first applies to the pair \{John, see\}, then to the pair \{Jane, W\}, where $W$ is the product of the application of Merge to \{John, see\}. Constituents can be described as sets and Merge can be defined as set union, so that $W$ is the set \{John, see\}, formed from the union of \{John\} and \{see\}. And the outcome of the complete derivation in (87) is the set \{Jane John, see\}. Of course, that set could also be generated by the derivation in (88), since set union is associative:
The two derivations in (87) and (88) define different constituent analyses based on the same array of syntactic primitives.

A syntactic set created by Merge has a very special feature: among its elements, it counts one which may be taken to stand for the whole set. That element is called the head of the set. In the particular derivation in (87), the head of the first constituent created by Merge is see, which is also the head of the final outcome. Then, the description of the derivation in (87) can be simplified as follows: Merge first applies to John and see, and then it applies to Jane and see. This is represented in (89):

(89)

The derivation marker in (89) constitutes the first phase. Spell-Out applies to that first cycle. Following Chomsky 1999, we assume that the interpretively relevant part of the phase does not include the highest specifier in it. In the case of (89), it is then the submarker in (90):

(90)

One may distinguish three functions or components of Spell-Out. The first component (SO^I) defines the primitive domains of PF interpretation. These are the minimal syntactic collocations which can be translated into assemblages of morphosyntactic and, ultimately, PF features. Call these primitive domains morphosyntactic domains (MSDs). An MSD is a “local” object: it may be a syntactic head or a short continuous sequence thereof, but nothing bigger than that. The second component of Spell-Out (SO^II) determines how the MSDs are strung together. In particular, SO^II delimits the prosodic domains and assigns relative prominence to the various MSDs within these domains. The third component (SO^III) actually effects the translation of MSDs into morphosyntactic and, ultimately, PF representations. The theory of SO^III, we assume, is Distributed Morphology (DM) (see Halle and Marantz 1993). SO^III treats the MSDs created by SO^I
as independent domains and applies to these domains in an unordered fashion. From the point of view of the cyclical architecture, the critical parts of Spell-Out are then SO\(^I\) and SO\(^{II}\).

The output of SO\(^I\) is not an independent “PF derivation marker,” but merely an annotated version of the syntactic derivation marker, in which MSDs are identified. SO\(^I\) converts the derivation marker in (90) into the tree in (91), where \(<...>\) represents an MSD:

(91)

```
     see
    /   \
see    <John>
```

A verbal head can be identified as an MSD only in combination with the \(T(ense)\) head. This is why there is no MSD corresponding to see in (91). The SO\(^I\) interpretation of the first phase in (89) then reduces to the marking of John as an MSD (the constituent John is assigned Case in the first phase in (89), so it is a complete MSD by the end of that phase).

The array for the next phase includes the primitive element Past together with the elements at the “margin” of the previous phase, i.e., the phasal head and its specifier. The second phase has the structure in (92):

(92)

```
     Past
    /    \
Past    Past
|     /  \nJane  see
```

The interpretive domain in (92) is the marker in (93):

(93)

```
     Past
    /    \
Past    see
```

SO\(^I\) transforms this marker into the tree in (94):

(94)

```
     Past
    /    \
<Past>  <see>
```

In isolation, each of the MSDs \(<Past>\) and \(<see>\) would surface as did and see, respectively -more exactly, as the PF representations corresponding to did and to see.-
On the other hand, since $<\text{Past}>$ and $<\text{see}>$ are adjacent in the structure in (94), a more “economical” interpretation is to fuse the two MSDs into a single one, ultimately giving rise to the PF corresponding to saw. So the image of (93) by SO$^1$ is the tree in (95):

(95)       
        Past
         \<Past + --- see\>

The transformation of (94) into (95) is the rule of Affix-Hopping.

The array for the third and final phase includes the primitive element $C$ together with the head of the second phase and its specifier. The final phase has the structure in (96):

(96)    
        C
         \Past
        \C \Jane \Past

The interpretive domain in (96) is the whole marker. The application of SO$^1$ to (96) yields the tree in (97):

(97)       
        Past
         \<\text{Jane}\> \Past

The complementizer $C$ has no morphosyntactic translation and. Furthermore, a functional head such as $\text{Past}$ cannot be identified as an MSD unless it is combined with some verbal projection (this is the converse of the situation in (91)). The SO$^1$ interpretation of the final phase in (96) then reduces to the marking of $\text{Jane}$ as an MSD.

Note that we are assuming that the derivation of the structure in (85) involves three phases, a departure from Chomsky 1999. The latter postulates two phases, a first phase associated with the verb domain and a second one, with the complementizer domain. The difference between the two analyses rests on the status of the functional category $T$.

Implicit in our proposal is the assumption that $T$ is a “complementizer-like” category and $C$ a higher complementizer. Hence the additional phase. The question at this point is how the various phases are formally linked together. How does one know that, e.g., $\text{see}$ in the phase in (89) is the same element as $\text{see}$ in the phase in (92)? A natural approach is to try to incorporate all the phases into a single connected derivation marker. To that effect, we capitalize on the fact that a syntactic head may represent several levels of constituency at the same time. For example, $\text{see}$ in (92) can represent the head in the phase in (89) or it can represent any projection of that head. The derivation markers in
(89), (92) and (96) were defined with the first interpretation in mind. If, on the other hand, we assume the second interpretation, it becomes possible to join the three markers into a single one describing the whole derivation, as shown in (98) (note that the two overlapping markers in (96) an (92) have been “unified” in the sense that the common subtree is represented only once):

(98)

```
+--- C ---+  
    |      |   
   +--- Past ---+  
     |     |    
    C     Past   
     |     |    
    Jane   Past   see 
           |       |   
           see     
               |   |   
              Jane see John 
```

There remains a problem, though, common to the new integrated representation in (98) and to the disconnected one displayed in (89), (92) and (96). Both fail to identify the multiplicity of occurrences of *Jane* as a *chain*, i.e., as the same constituent partaking of separate phases. A chain is the association of a constituent $\alpha$ with a list of contexts $K_\alpha$ for $\alpha$ (see Chomsky 1995). For any given $\alpha$, the derivation may then be equated with the generation of $K_\alpha$. Under the simplest definition, a context for $\alpha$ identifies only one occurrence of $\alpha$ (thus, $X \rightarrow U \rightarrow W$ is not a context in that restricted grammatical sense). The pair of derivation markers in (99), where $\Delta$ is a fixed “dummy symbol,” is a description that identifies the chain of occurrences of *Jane* so defined:
We note that there is a second constituent that shows up in two contexts, namely the structure in (100):

(100)

The verbal structure in (100) has two distinct grammatical roles in the sentence in (85), *predicate of* the constituent *Jane* and *complement of* the functional head *Past*. The two occurrences of (100) in (99) then also constitute a chain. Accordingly, we revise (99) to (101), where \( \Delta \) and \( \Delta' \) are two distinct dummy symbols:
The whole description in (101) can be legitimately identified as the LF of the sentence in (85). In some sense, we are back to the two phase theory in Chomsky 1999, with the diagram in (101i) representing the first phase and that in (101ii) the second phase. There are differences, though. The derivation described in (101) goes like this. First, a small clause is assembled, with a verb, a subject and an object. The small clause is then separated into two parts, its subject (Jane) and the predicate cycle (see John). At this point, the derivation starts afresh, using the parts from the disassembled first phase as well as new material. The cycle [see John] from the first phase is plugged in the second phase at the stage where the functional category T is merged with its verbal complement. Following this, the subject Jane is merged as the specifier of T. In general, the LF of a sentence will have the following structure:

(102) An LF representation is a partially ordered set of derivation markers, where the ordering is defined in (i):

(i) The derivation marker \( p \) precedes the derivation marker \( q = \text{def} \) the lexical items in \( p \) form a subset of those in \( q \)

Note that one effect of disassembling the first phase before starting the second one in an LF such as (101) is what has been called *trace deletion* (see Chomsky 1995, for example). Assuming this as a general mechanism, no phase will ever include any trace (at least, of A-movement).\(^2\) A “trace” is now viewed as a derived construct: a constituent in a phase \( p_i \) is a trace to the extent that it partakes of a more inclusive phase \( p_j \).

Within the above execution of phase theory, it is not necessary for Spell-Out to go by the same phasal “clock” as LF. In particular, a single “LF phase” might correspond to a multiplicity of “PF phases” (see Megerdoomian 2000 for a discussion of this issue).
We will assume that this is indeed the case. For example, we will maintain the PF analysis of the structure in (101) (the description of the sentence in (85)) in three phases, displayed in (91), (95), and (97). Linking these SO\textsuperscript{I} phases together yields the tree in (103):

\[(103)\]

```
Past
  
  <Jane>
  
  Past
  
  <Past + see>
  
  see
  
  <John>
```

The PF linear sequence of MSDs mirrors the derivational ordering.\textsuperscript{22} That is, the sequence of MSDs is that in (104), where "^" is the concatenation operator:

\[(104)\] \textless Jane\textgreater ^\textless Past+see\textgreater ^\textless John\textgreater

Determining the string of MSDs from the SO\textsuperscript{I} derivation marker is one of the roles of SO\textsuperscript{II}. In addition, SO\textsuperscript{II} defines prosodic domains, assigning stress at the boundaries of these domains. Concretely, we shall assume the following rule:

\[(105)\] Let \(\mu\) be a sequence of MSDs (possibly reduced to one) such that \(\mu\) corresponds to an LF phase. A prosodic boundary (denoted by \#) is inserted after \(\mu\), unless one already exists within or after \(\mu\) (having been inserted in a prior phase).

Furthermore, we assume that any DP is an LF phase by itself (on par with CP).\textsuperscript{23} In the derivation in (103), SO\textsuperscript{II} will then insert \# after the MSDs \textless John\textgreater and \textless Jane\textgreater:

\[(106)\]

```
Past
  
  <Jane>#
  
  Past
  
  <Past + see>
  
  see
  
  <John>#
```
The only other string in the representation in (106) that corresponds to an LF phase is the whole string itself (neither \textit{[see John]}, nor \textit{[Past see John]} are phases in the analysis in (101)). The string of MSDs is then parsed as in (107):
(107) $<\text{Jane}>#^{\Delta}<\text{Past+see}>^{\Delta}<\text{John}>#$

We assume that some principle subordinates $<\text{Jane}>#$ to $<\text{John}>#$. Thus, main prominence is assigned at the last (rightmost) boundary within the sequence.

Now consider the sentence in (108) (cf. the discussion of (68) above):
(108) Jóhn disappeared

The neutral stress contour is the one indicated, with main stress on John. The LF description of the sentence is given in (109):

(109) (i) (1\textsuperscript{st} phase)

\[
\Delta
\]

(ii) (2\textsuperscript{nd} phase)

\[
\begin{array}{c}
\Delta' \\
\Delta
\end{array}
\]

(iii) (3\textsuperscript{rd} phase)

\[
\begin{array}{c}
C \\
\begin{array}{c}
C \\
\Delta
\end{array} & \begin{array}{c}
\text{Past} \\
\Delta'
\end{array}
\end{array}
\]

(iv) (identification of the dummy symbols)

\[
\Delta = \text{John}, \quad \Delta' = \text{disappear}
\]

One difference between the structure in (109) and that in (101) is that, whereas the object of the verb in (101) gets Case in the second phase, the object in (109) does not get Case until the last phase. As a consequence, the object in (109) is not interpreted as an MSD by SO\textsuperscript{I} until that final phase, once it has “moved” to subject position. The prosodic boundary $#$ is then inserted after the subject. We display in (110) the image of (109) by SO\textsuperscript{I}, together with the prosodic boundary inserted by SO\textsuperscript{II}:
The corresponding string is that in (111):
(111)  \langle \text{John}\rangle \#^{\langle \text{Past}+\text{disappear}\rangle}

Compare next the sentence in (112) (cf. the discussion of (69) above):
(112)  \text{John mysteriously disappeared}

In the sentence in (112), the neutral intonation is with main stress on \textit{disappear}. The essential difference between (112) and (108), of course, is the presence of an additional constituent in (112), namely the adverb adjoined to the verb phrase. We analyze the structure as in (113):
(113)  (i) (1\textsuperscript{st} phase)
\[
\Delta
\]

(ii) (2\textsuperscript{nd} phase)
\[
\begin{array}{c}
\Delta' \\
(\text{mysteriously})
\end{array}
\begin{array}{c}
\Delta' \\
\Delta
\end{array}
\]

(iii) (3\textsuperscript{rd} phase)
\[
\begin{array}{c}
C \\
\Delta
\end{array}
\begin{array}{c}
\text{Past} \\
\Delta'
\end{array}
\begin{array}{c}
\text{Past} \\
\Delta
\end{array}
\]

(iv) (identification of the dummy symbols)
\[
\Delta = \text{John}, \quad \Delta' = \text{disappear}
\]
The parentheses around the adverb *mysteriously* in (113ii) indicate that the second phase really is an abstract construct made up of two actual phases: one phase has the adverb *mysteriously* in it and the other lacks it. In other words, the second phase is a pair of “parallel” derivations, a “short” one, represented in (114), and a “long” one, represented in (115):

(114) Short derivation

```
               disappear
                  \      /
            disappear   John
```

(115) Long derivation

```
               disappear
                  \      /
                        disappear
                           \  /
            mysteriously   disappear   John
```

The two derivations in (114)-(115) have different outcomes, the short one yielding the set \{disappear, John\} and the long one, the set \{disappear, John, mysteriously\}. That difference affects later cycles. For example, consider the stage at which *Past* is brought into each derivation. The two parallel cycles that constitute the second phase give rise to the two parallel derivations in (116)-(117):

(116) Short derivation

```
        Past
           \   /
              \ 
       disappear
          \     /
      Past   disappear   John
```

The two derivations in (114)-(115) have different outcomes, the short one yielding the set \{disappear, John\} and the long one, the set \{disappear, John, mysteriously\}. That difference affects later cycles. For example, consider the stage at which *Past* is brought into each derivation. The two parallel cycles that constitute the second phase give rise to the two parallel derivations in (116)-(117):

(116) Short derivation

```
        Past
           \   /
              \ 
       disappear
          \     /
      Past   disappear   John
```
(117) Long derivation

```
    Past
     \     /     \
      |  |     |
    disappear
       \   /   /
      Past  disappear
             \ /  /
        mysteriously  disappear  John
```

The two parallel derivations above may be combined into a single complex marker:

(118)

```
    Past
     \     /     \
      |  |     |
    disappear
       \   /   /
      Past  disappear
             \ /  /
        mysteriously  disappear  John
```

The double connection between Past and disappear in (118) reflects the constituent analysis distinction between category and segment introduced in May 1985. We take the derivational ambiguity above to be characteristic of adjunction.

The diagram in (118) is the LF representation generated by the derivation up to the stage where Past is merged with the outcome of the second phase in (113ii). The SO image of that stretch of the derivation is as in (119):

(119)

```
    Past
     \     /     \
      |  |     |
    disappear
       \   /   /
      <Past ------------ + ------------ disappear>
             \ /  /
        <mysteriously>  disappear  John
```
One issue is how the PF ordering of *<mysteriously>* and of *<Past+disappear>* is determined. One may again appeal to the principle that the PF ordering mirrors the derivational one (cf. the structure in (103) and the comment that follows it). Specifically, we will assume that the PF ordering of the MSDs displayed in (119) is determined at the combined derivational stage represented in (120):

(120)

```
Past
    / \                       /   \\
disappear          <Past -------------------- + -------------------- disappear>  \\
     /  \            John                        John
<mysteriously>    disappear
```

The diagram in (120) represents the combination of the following cycles from the two parallel derivations: the cycle in the long derivation in which the verb *disappear* is merged with *mysteriously*, and the cycle in the short derivation in which it is merged with *Past*. Each cycle immediately follows that in which *disappear* is merged for the first time in each derivation. At the combined stage depicted in (120), *disappear* both precedes *mysteriously* and *Past* in the derivational ordering, while *mysteriously* and *Past* are unordered with respect to each other. Since *Past* is incorporated with *disappear* into a single MSD, the unique PF ordering compatible with the derivational ordering defined by (120) is that in (121), where *<mysteriously>* precedes *<Past+disappear>*:

(121)  *<mysteriously>*^<Past+disappear>*

We see that using the complex cycle in (120) to determine the PF ordering of MSDs has the same effect as having *Affix-Hopping* “lower” the inflection onto the verb. If we now assume that the combination of an adverb with the verb phrase it modifies constitutes an LF phase, a prosodic boundary will be inserted after *disappeared*. The result is the form in (122):

(122)  *<John>*^<mysteriously>*^<Past+disappear>^#

We have illustrated the cyclical framework in some detail to show that, within such an architecture, PF interpretation raises issues which are conceptually analogous, or even identical, to those arising within the standard architecture. The solutions proposed in the literature can often be translated quite directly into the cyclical format. Obviously, there are differences. Indeed, the cyclical theory allows for a more felicitous definition of “depth of embedding” in the sense of Cinque 1993 or the revised sense of Zubizarreta 1998. A maximally embedded constituent within some structure *K* is one that constitutes a primordial LF phase in the cycle corresponding to *K*. As a consequence of the derivational approach, the interplay of structural and selectional ordering in Zubizarreta
1998 may have to be revised, possibly to be replaced by some interplay of structural and linear ordering. This would leave the logic of the parametric account unaffected, though.

5. Some final remarks.
While we have chosen to focus on a certain view of the world in this paper, it is certainly not the only one around. The view assumed here and captured by the principle in (11) (repeated in (81)) is that the possible scopes of the focus are determined by the location of the intonational nucleus. There is an alternative view, the pitch-accent-based theory of focus, put forth by Gussenhoven 1984 and Selkirk 1984, 1995. These authors propose that in Germanic pitch accent assignment and predicate-argument relations is mediated by focus projection rules (or focus-domain formation rules). Selkirk 1995 proposes the following rules, where the feature F is a diacritic for annotating syntactic structure: (This is a non-technical rephrasing of her rules.)

(123) a. F-marking rule: an accented word must be marked in the syntactic structure with the feature F.
    b. F-projection rule: F-marking may project upward within the syntactic structure from a head or its internal argument.

A late rhythmic rule ensures that the last accented word receives main prominence, but it plays no role in determining the possible F-structures of the sentence:

(124) Nuclear Stress Rule
The most prominent syllable of the rightmost constituent in a phrase P is the most prominent syllable of P.

The F-marked structures are related to the informational structure of the sentence by the following rules:

(125) a. The Focus of a sentence (FOC) is defined as an F-marked constituent not dominated by any other F-marked constituent.
    b. F-marked constituents which are not a Focus are interpreted as new in the discourse, while a constituent without F-marking is interpreted as given. A Focus constituent, on the other hand, may be interpreted as either given or new in the discourse.

We will not attempt to evaluate here the empirical success of the pitch-based theories of focus. (For some discussion, see Zubizarreta 1998; section 2.5 and also Ladd 1996.). We simply note here that this theory must make use of an ad-hoc feature F to annotate the syntactic structure. This feature is undesirable because it lacks independent justification. It is only needed to establish the relation between pitch accent and the informational structure of the sentence. An NSR-based theory of focus does not require such a feature. In this theory, focus is directly identified by prosodic notions such as relative prominence. Other grammatical means for focus identification are also available. For example, contrastive focus may be associated with logical operators such as only. Some languages may make use of a bona fide syntactic position or it may make use of a
particular morphological marking. But these are all independently justified notions. F-marking is not.

While some progress has been made in our understanding of the relation between prosody, focus, and syntax, much remains to be done. In effect, we have only scratched the surface of it. More typologically different languages need to be studied. In particular, it would be useful to learn the exact correlation between focus and prosody in tone languages, in particular those in which pitch is distinctive. It has been claimed that certain African languages (such as Chadic; see Tuller 1992) resort to morphology and syntactic position rather than to pitch to mark the focus of the sentence. It may be premature to draw such a conclusion before we have more in-depth studies of the phonology/phonetics of such languages. As we said in the introduction, prominence is a rather abstract notion for which there is no unique phonetic correlate. In fact, there may be relevant phonetic cues other than pitch accent (such as other aspects of the intonational contour). It is precisely this type of languages that we need to investigate further in order to better understand the relation between prosody, focus, and syntax.

References.


**Notes**

* This work was completed in August 2000 and it does not reflect work that has been done on the topic since then.
The most consistent acoustic correlate in a language like French or English may be the amplitude and sharpness of fundamental frequency ($F_0$) variation associated with a given syllable (see Martin 1975, 1982).

The following quote from Chomsky & Halle 1968 is of relevance here:

(i) “We do not doubt that the stress contours and other phonetic facts that are recorded by careful phoneticians and that we will study here constitute some sort of perceptual reality for those who know the language in question. In fact we are suggesting a principled explanation for this conclusion. A person who knows the language should “hear” the predicted phonetic shapes. In particular, the careful and sophisticated impressionistic phonetician who knows the language should be able to bring this perceptual reality to the level of awareness, and there is ample evidence that phoneticians are capable of doing this. We take for granted, then, that phonetic representations describe perceptual reality. Our problem is to provide an explanation for this fact. Note, however, that there is nothing to suggest that these phonetic representations also describe a physical or acoustic reality in any detail. For example, there is little reason to suppose that the perceived stress contour must represent some physical property of the utterance in a point-by-point fashion; a speaker who utilizes the principle of the transformational cycle and the Compound and Nuclear Stress Rules should “hear” the stress contour of the utterance that he perceives and understands, whether or not it is physically present in any detail. In fact, there is no evidence from experimental phonetics to suggest that these contours are actually present as physical properties of utterances in anything like the detail with which they are perceived. Accordingly, there seems to be no reason to suppose that a well-trained phonetician could detect such contours with any reliability or precision in a language that he does not know, a language for which he cannot determine the surface structure of utterances.

Considerations of this sort lead us to suspect that the question of how highly differentiated the stress contours in a representation should be is of little significance. In a complex utterance with a rich surface structure, the rules outlined in the preceding section will lead to a stress contour of many levels. There may be no empirical sense to the question of whether the resulting representation is correct in full detail. Because of the completely impressionistic character of judgments of relative stress, decisions over a broad range are of little value. It is not at all surprising that there should be great difficulty, within impressionistic phonetics, in determining how many stress levels should be marked and how they are distributed in utterances that exceed a certain degree of complexity. The shape and the degree of differentiation of a stress contour are largely determined by obligatory rules and are therefore below the level of systematically significant representation. Once the speaker has selected a sentence with a particular syntactic structure and certain lexical items (largely or completely unmarked for stress, as we shall see), the choice of stress contour is not a matter subject to further independent decision. That is, he need not make a choice among various “stress phonemes” or select one or another “superfix.” With marginal exceptions, the choice of these is as completely determined as, for example, the degree of aspiration. Similarly, a hearer who has grasped the structure and morphemic constitution of an utterance from a rough sampling of the physical input need not attend to stress variation, to whatever extent this may actually be a physical property of utterances.

It is to be expected that determined phonetic features should be quite difficult for the user of the language to learn to identify, whether they involve stress or degree of aspiration (where undoubtedly there are many levels, predictable, at least roughly, by general rules). The apparent ease with which phoneticians trained in the same conventions can, to a large extent, agree on the assignment of four or five stresses in utterances may very well be traceable to their ability, as speakers of the language, to grasp the syntactic structure of utterances and to assign to them an “ideal” stress contour by the rules of the transformational cycle. Such an achievement may have little to do with any physical fact. This is, incidentally, a matter which should be subject to experimental investigation.” (Chomsky & Halle 1968, Chapter 2, pp. 25-26)
There is promising work coming out of Laboratory Phonetics that might pave the way towards remedying this state of affairs. See for example, XXX and references cited therein.

This test, of course, works only to the extent that the form of the answer respects the form of the question and the relation between the two is not mediated by any hidden inferences. It is precisely for this reason that the question-answer test is not an infallible one for determining the focus-structure of a sentence. See Zubizarreta 1998, note 6, p. 161.


For example, this could be the structure for the sentence John’s son’s baby cried.

Chomsky & Halle 1968 actually formulates the NSR as a rule assigning stress to vowels, with the structural description in (i):

\[(i) \quad V^1 \ldots \ldots \ldots \alpha \]

Thus, the text formulation is more abstract since it involves the notion of “stress of a word relative to other words” while stress is perceived as a phonetic property of vowels. See the text below. Nothing in our development hinges on the particular choice of formulation in terms of word stress or of vowel stress.

One may recall that in a constituent analysis (i.e., context-free) grammar of the standard type, a phrase-marker is a derivation (more exactly, a class of derivations); see Chomsky 1957, pp. 26-33. The equivalence postulated in (19) in the text can then be viewed as a generalization of this definition to derivations involving transformations. In a sense, the duality between constituent structure and derivation manifests the compositional nature of language. The duality would break down if language structure were to display properties of a global type. A full discussion of the duality would have to address the issue of cyclical interpretation. We briefly return to this question at the end of section 4. Note that compositionality holds even when the operation involve (say Merge, or Concatenate) is associative. That the operation is associative means that all constituent structures of a given string are equivalent.

The convention in (22) in the text is a partial reformulation of the mechanisms postulated in Halle & Vergnaud 1987; see in particular Halle & Vergnaud 1987, Chapter 7, pp. 263-266.

An account of the contour in (26) in the text could be given within the text formalization (20)-(23) by recording the depth of embedding. Thus, given a category \(\alpha\) with immediate metrical head \(H_m(\alpha)\), a sister of \(H_m(\alpha)\) has secondary stress within \(\alpha\). A sister of the immediate metrical head of \(H_m(\alpha)\) has tertiary stress within \(\alpha\), as does a sister of the immediate metrical head of a sister of \(H_m(\alpha)\). And so on and so forth. We encounter here the same unwieldiness that was noted in the text for the case of the one-step derivational algorithm, and for an obvious reason. Indeed, one should view the natural implementation of a “one-step procedure” to be an “interpretive system” such as that in (20)-(23) in the text, and not a transformational algorithm. The question is then whether the interpretive algorithm just defined and the cyclical transformational algorithm in (16)-(17) in the text are empirically distinguishable.

The question of stress subordination has two distinct, but related, aspects. First, we have the question of the preservation of stress contours under embedding. To illustrate, consider the structure in (i):

\[(i) \quad \text{John’s baby liked ET} \]

At issue here is whether the internal relations of stress in the phrases John’s baby and liked ET are the same in the sentence John’s baby liked ET as in isolation. We note that all three algorithms mentioned in this paper, the cyclical transformational algorithm of Chomsky & Halle 1968, the algorithm in Halle & Vergnaud 1987 and the interpretive algorithm discussed in note 8, assign the same stress contour to the structure in (i). Namely primary stress is assigned to ET, secondary stress, to baby and the same tertiary level of stress, to John and liked (see Chomsky & Halle 1968, Chapter 3, p. 90). Thus, the three algorithms presented here all imply that the internal relations of stress in the phrase John’s baby are the same in the sentence John’s baby liked ET as in isolation, while those of liked ET differ. As Chomsky & Halle 1968 notes (fn 43, p. 90), there is really no evidence bearing on the matter, however.
Which leads to the second general question concerning stress subordination, namely the question whether there exists any perceptual relation at all between the stress level of *John* and that of *liked* in a structure like (i). It may very well be the case that stress levels are comparable only to the extent that they independently “relate to the same main stress.” This is the case in (25) in the text, where *liked* and *John* are directly subordinated to the same main stress in distinct phrases. The text discussion in section 2.2.2. provides some indirect support for that view, as it suggests that the distribution of secondary stresses may be ruled by a different set of principles. If this turned out to be true, it would constitute evidence in favour of the interpretive approach.

12 For a detailed discussion of eurhythmicity in French, see Dell 1984.

13 Tokizaki 2000 argues that the stress pattern in (44) corresponds to a “thetive judgment” interpretation, while the stress pattern in (45) corresponds to a “categerical judgement” interpretation. While this may very well be the case, the point is that the unaccusatives in (46) give rise to an unambiguous stress pattern. Main stress on the verb gives rise to a contrastive focus interpretation.

14 Bresnan 1971 attempts to do just this by proposing that the NSR applies cyclically. For a critique of Bresnan’s data, see Bolinger 1972.

15 The term is a misnomer since deaccented words are in many cases associated with miniature but quite audible pitch accents.

16 The notion is defined as follows in Cinque 1993:

(i) a. The major path of embedding is constituted by nodes on the X-bar axis (X, X’, XP) and the nodes on the recursive side of the tree.

   b. The minor path of embedding is constituted by nodes on the nonrecursive side of the tree.

However, the notion “recursive side” used in that definition is somewhat ineflicitious, since both the complement of X and the specifier of X’ are recursive nodes. Furthermore, the definition should make explicit the fact that the distinction “major vs. minor path of embedding” is relative to a maximal projection XP (or, equivalently, relative to a head X). This is why we have replaced Cinque’s original formulation by the text one. Note that the notion of “major path of embedding” has a natural interpretation in the theory in Kayne 1994, which assumes that specifiers are adjuncts.

17 The definition of c-command proposed by Zubizarreta departs slightly from the standard one. It distinguishes between direct and indirect c-command:

(i) A directly c-commands B iff A and B are heads or maximal projections (excluding segments) and 
a) A and B are sisters or b) there exists a node C such that A and C are sisters and C dominates B.

(ii) If A directly c-commands B, then A indirectly c-commands C, C a projection of B that does not contain A.

According to the definition in (i), the Specifier ZP in the structure in (iii) directly c-commands the head X and its complement YP (as in the standard definition of c-command), but it does not directly c-command its sister X’ (because X’, being neither a head nor a maximal projection cannot be in a direct c-command relation with any category). According to the definition in (ii), ZP indirectly c-commands X’ by virtue of the fact that X’ is a projection of X and ZP directly c-commands X. It is this latter extension of the c-command relation that goes beyond the classical one.
In Zubizarreta 1998, selectional ordering was defined on a separate, more abstract tree. Here, we assume that the syntactic tree encodes both selectional ordering and constituent ordering in terms of asymmetric c-command.

We see that the interplay of C-NSR and S-NSR yields the effects of the English Compound Stress Rule (cf. Chomsky and Halle 1968), thus preserving an important result of Cinque 1993. To illustrate, consider the compound \([\text{kitchen} \ [\text{towel rack}]\)], with the bracketing indicated. In the inner constituent \([\text{towel rack}]\), rack is analyzed as a head taking towel for argument (heads are always on the right in words). Relative prominence is then determined by S-NSR, which assigns main stress to towel. In the higher constituent, no argument relation holds and C-NSR applies, assigning main prominence to the constituent \([\text{towel rack}]\).


Obviously, this imposes a contraint on the relation between phases and movement: movement may never be internal to a phase.

Of course, the derivational ordering is a partial one. Suppose, for example, that the subject of the sentence contains more than one MSDs. We will adopt the following rule:

(i) Let \(\ldots\ C \ldots\) be some structure, where \(C\) is a constituent ultimately interpreted as the sequence of MSDs \(\mu = (c \ m_1^\cdot m_2^\cdot \ldots^\cdot m_s^\cdot)\), \(\cdot\)the concatenation operator. Let \(h\) be the head of \(\ldots C \ldots\) and \(<h>\) be the MSD in \(\mu\) associated with \(h\) (\(<h>\) is some \(m_i\)). The PF ordering of \(\mu\) with respect to the context \(\ldots\) is the same as that of \(<h>\) in isolation.

The complex formal issues which arise in connection with the projection of the partial ordering of constituent structure onto the total ordering of the time axis are discussed in great depth in Kayne 1994.

The LF description in (101) in the text will then have to be revised accordingly. Note there must be some subordination principle that ensures that the boundaries introduced in previous phases do not interfere with current ones. For example, if John in the location of main stress within a complex DP must not affect the

The principle might be formulated in terms of a last vs. first (rightmost/leftmost) parameter or, alternatively, in terms of depth of embedding (as in Halle and Vergnaud 1987, Cinque 1993, or Zubizarreta 1998).