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

Table of contents

_____

0. Introduction.

       There are two main kinds of grammatical studies on phrasal prosody that bear on the prosody-syntax relation: one that deals with prosodic phrasing and another one that deals with the relation between prosodic prominence and focus. Here we will address mainly the second type of issue; the first one will be discussed only briefly.

       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 two-fold. 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. Many notions and much terminology have been brought to bear on this issue (for a review, see Vallduví and Engdahl 1995), giving rise to a lot 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, Rochemont 1986, Rooth 1985, and Szabolsci 1981, 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*.[3] This is illustrated by the examples below, where the scope of 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)   In a sentence $S = \alpha\ X\ \beta$, X is a possible focus just in case S would count as a felicitous answer to the question $Q = wh\text{-}X\ \alpha\beta$.

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)   Q. C'est qui qui a écrit un livre sur les rats?
                    'It is who that wrote a book about rats?'

        A. C'est [DP le chat] qui a écrit un livre sur les rats.
           'It is the cat that wrote a book about rats.'

(7)   Q. C'est quoi que le chat a écrit?
                    'It is what that the cat wrote?'

        A. C'est [DP un livre sur les rats] que le chat a écrit.
                    '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)   Q. Who wrote a book about rats?

        A. [DP The cat] wrote a book about rats, and [DP the bat] did too.

(9)   Q. C'est qui qui a écrit un livre sur les rats?
                    'It is who that wrote a book about cats?'

        A.*C'est [DP le chat] qui a écrit un livre sur les rats, et aussi [DP la chauve-souris].   'It is the cat that wrote a book about rats, and also the bat.'

Vergnaud & Zubizarreta 2000 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 i-focus in Germanic.

2.1. The Classic View (Chomsky 1971, Jackendoff 1972).

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, the VP, or the TP as the scope of the focus. 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 hàt has written a book about ráts.

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

Based on such observations, the above-cited authors proposed a constraint of the following sort:[4]

(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 says 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 reapplies 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 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 α the *primary stress* within α. 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. To illustrate, consider the structure in (13), where *t*, *u*, *v*, *w* are words and α, β and γ phrases:[5]

(13)   (((t u) 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) $t^4 \, u^3 \, v^2 \, w^1$

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) "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 β and a subordinate stress within another phrase α.

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):[6]

(16)   Assign primary stress to a primary-stressed word w^1 in the context

$w^1 \ldots \quad \ldots]$

where α 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 α. 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 theses 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 surface constituent structure are really two competing ways of analyzing the sentence; see Chomsky 1994, Epstein 1994. Indeed, one can take this hypothesis to be one of the central insights in these works. For concreteness, we shall postulate the following (see Freidin & Vergnaud 1999):

(19)   The structural description of a sentence ω is an abstract structural analysis Ω which can be interpreted either as a derivational history, i.e., a transformation-marker in the sense of Chomsky 1955, or as a phrase-marker.

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 the nonderivational implementation of (12). 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:[7]

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

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

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

Quite obviously, the "head-of" and "projection-of" relations defined in (22) are transitive in the following sense:

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

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

(24)   Given a sequence $\alpha_1, \alpha_2, ... , \alpha_i, \alpha_{i+1}, ..., \alpha_n$ such that, for $1 \quad i \quad n - 1$, α_{i+1} is the metrical head of α_i in the sense of definition (22), $\alpha_{i+1} = H_m(\alpha_i)$, then α_n is defined to be a *metrical head* of α_1 and, correspondingly, α_1 is defined to be a *metrical projection* of α_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)   John^2 liked^3 ET^1

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).[8] 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.[9] 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. [10] 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 (29), 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 lìnguists' dógs
         b. Nìneteen hundred linguists' dógs

(28)   a. Nineteen hundred lìnguists sáng
         b. Nìneteen 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 (29b). Again this is probably due to the fact that a longer VP calls for a break immediately following the subject.

(29)   a. *Nìneteen thousand linguists sang the Marseilláise.
         b. Nineteen thousand lìnguists sang the Marseilláise.

      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 to 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.

(30)   a. Nineteen hundred línguists sàng
        b. Níneteen hundred linguists sàng

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

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.

(31)   Q. Who has written a book about <u>ráts</u>?

        A. [The cat in the blue hát] has written a book about <u>ràts</u>.

Note furthermore that:

(32)   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.

(33)   Q. Who has written a <u>BOOK</u> about rats? (not a squib)

        A. [The cat with the blue hát] has written a <u>bòok</u> 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 1978, 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 Búch 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 ràts (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 Papíer.
         b. Hans is working on a páper.

(37)   a. Hans arbeitet an einem kleinen Tísch.
         b. Hans is working on a small táble.

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

(39)   a. Das Taxi kommt spáet.
              The taxi is-coming láte
            b. The mail arrived láte.

   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 Búch 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 Táxi 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 Júnge gelacht hat.
                        It is-said that a boy laughed has.

          b. Es heisst, dass ein Junge gelácht 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 Júnge 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 báby is crying.
          b. The báby is laughing

(45)   a. The baby is cr_ing.
          b. The baby is láughing.

On the other hand, in comparable structures with an unaccusative verb, NS falls unambiguously on the subject:

(46)   a. The máil arrived.
                        b. The sún came out.
                        c. The máil has arrived.
                        d. My bág 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 disappéared.

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 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'

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".[11] 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. 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 kíssed *her*.

(51)   Remember John? Have you heard fróm *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 denóunced *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 A$_1$, 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 A$_2$ in (54). Thus, although there are idiolectal differences, the data seems to be coherent.

(54)   Q: Why is Hans happy?

        A$_1$: Weil *ihn/Hans* ein Verwándter angerufen *hat*.
        A$_2$: Weil *ihn/Hans* ein Verwandter ángerufen *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?

        A$_1$: Ich glaube, dass *ein* Júnge *das Buch* genommen *hat*.
                I believe   that a boy   the book taken has

        A$_2$: Ich glaube, dass *ein* Junge *das Buch* genómmen *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).

2.4. Cinque 1993.

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.

(58)
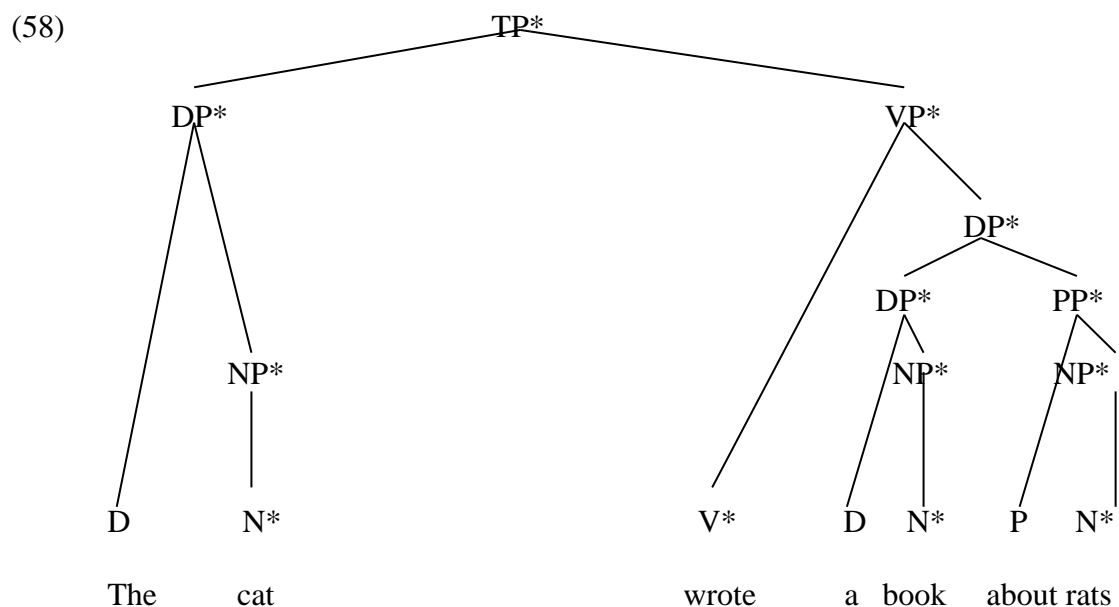
The most prominent word within a domain is the one that is dominated by the greatest number of asterisks. Like the classical NSR, c*at* is assigned NS within the DP subject domain and *rats* is assigned NS within the VP domain. We know that *rats* is more prominent than c*at* and that this is achieved 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 the specifier, main prominence is contained within the sister category in such structures: cf. *The cat that chased the ràt won a príze*. 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 T is a maximal set of categories P within T such that, if P includes a head H, it also includes the categories governed by H (in the sense of GB theory), 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:[12]

(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.

2.5. Zubizarreta 1998.
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)
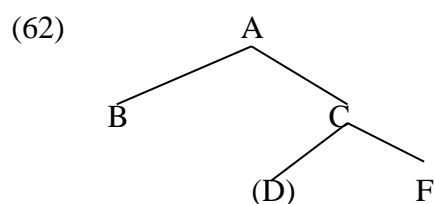
Based on the data discussed in section 2.3, the following approximate generalization is assumed. NS in V-final and non V-final structures in Germanic follow two distinct patterns (putting aside the cases of intransitives with NS on the verb, such as (45) and (47)):

(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 $[_{XP} ZP [_{X'} X^0 YP]]$. 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.[13]

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.[14] 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*:

(65)   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:

(66)   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-fledge 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)).
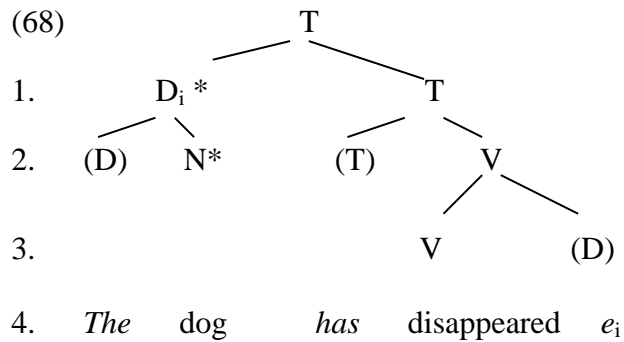
2.5.2. The basic examples.

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.)

(67)

1.

2.

3.

4.

5.

6. *The*   taxi   arrived e_i late
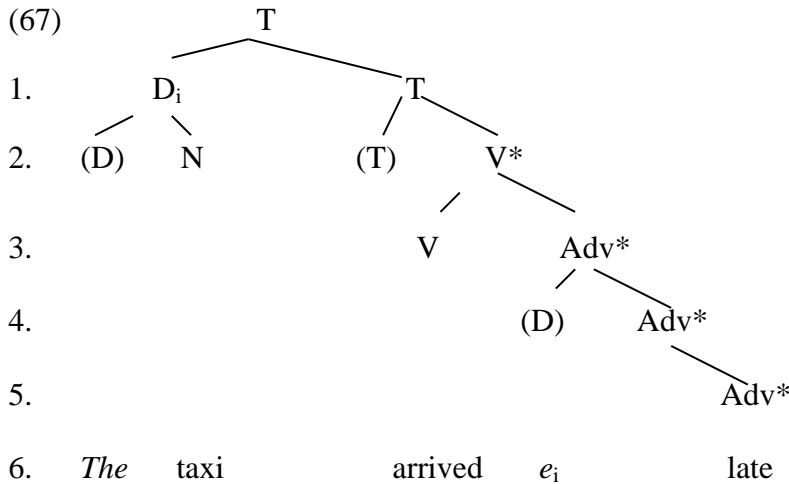
Consider a step by step (top-down) application of the rule (a bottom-up application would give the same result):

> 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:

(68)

```
1.          D_i *                    T
                                    /    \
2.    (D)    N*          (T)          V
                                     /    \
3.                         V          (D)
```

4.   *The*   dog   *has*   disappeared e_i

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)          Adv*
                                     /    \
3.                         Adv          V*
                                       /    \
4.                             V*          (D)
```

5.   *The*   dog   mysteriously   disappeared e_i
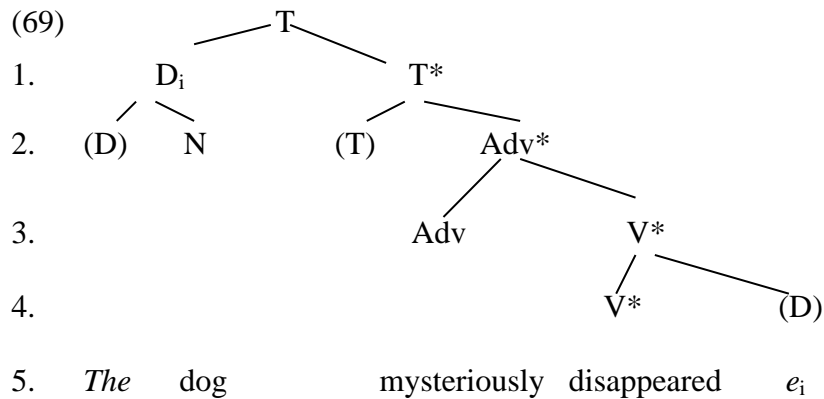
This metrical structure differs with respect to the previous one in a crucial way. The D in line 1 and the Vs in lines 3 and 4 are not metrical sisters. In this structure, D in line 1 and Adv in line 2 are metrical sisters but they are not selectionally related. Therefore, the C-NSR applies and assigns main prominence to Adv in line 2. Adv in line 3 is a metrical sister of the V in line 4 and in line 5 (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. (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)　　　　　　　T

1.　　　D　　　　　　T*

2.　(D)　N　　　(T)　　　V*

3.　　　　　　　　　V*　　　(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)　　　　　　　T

1.　　　D*　　　　　　T

2.　(D)　N*　　　(T)　　　V

3.　　　　　　　　　V　　　(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.

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?

        A$_1$: Weil *ihn/Hans* ein Verwándter angerufen *hat.*
        A$_2$: Weil *ihn/Hans* ein Verwandter ángerufen *hat*.

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 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. Take for example the case 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 repeated 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.[15]

(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 mésa.
         A3: Juan puso sobre la mesa el pán.
         A4: Juan puso el pán # sobre la mesa.
        Note that the Germanic NS algorithm share something in common 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. But recall that the LCA 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? Could it be that the NS system of Germanic evolved from a Bengali type system, and the notion of "lowest" evolved from the notion of "rightmost"? One important difference between Germanic and Bengali is that the focused constituent in Bengali is obligatorily followed by a p-phrase boundary, and therein probably lies the main cause for the divergence between the two systems. A system like Germanic in which a focused phrase is not flanked by a p-phrase boundary might then replace the phonological notion of "rightmost" with its syntactic counterpart, namely the notion of "lowest". It remains to be seen whether such speculations can be turned into a genuine object of research.

4. Some final remarks and conclusion.

It has been suggested above that there are typologically two distinct classes of languages: one (like Germanic) in which NS is computed on the basis of a metrical interpretation of the syntactic tree and another one (like Bengali and Romance) in which NS is computed in terms of prosodic phrases. What can we infer from this conclusion? 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 is incorrect. For if NS were to apply at PF in Romance, then p-movement would have to apply at PF. But this conclusion cannot be correct. Suñer 2000 has shown that p-movement in Romance affects binding relations. If binding relations are determined in the LF component of the grammar and if there is no direct grammatical connection between PF and LF, then p-movement must apply in the syntax. If p-movement applies in the syntax, then the NSR must also apply in the syntax. Our tentative conclusion is that certain prosodic information is available in the syntax, information such as phrasal stress, deaccenting, and prosodic phrasing. This information is crucial in many languages of the world to identify the scope of the focus within the sentence. And focus is a discourse-based notion that is part of the semantic representation of the sentence. We therefore conclude that some prosodic information within the syntactic component of the grammar must be available.

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 prosody principle in (11) (repeated in (81)) is that the possible scope of the focus is determined by the location of the intonational nucleus. There is an alternative view, the pitch-accent-based theory of focus, put forth by Gussenhoven op. cit. and Selkirk op.cit. 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.)

(83)   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 form 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:

(84)   *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:

(85)   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 theory of focus.[16] We simply note 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 overt or covert logical operators (e.g. *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 have 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.

Chomsky, N. 1971. Deep structure, surface structure, and semantic interpretation. In D. Steinberg and L. Jakobovits, eds., *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*. Cambridge: Cambridge University Press.

Chomsky, N. and M. Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Chomsky, N. 1994. Bare phrase structure. MIT Occasional Papers in Linguistics 5. Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [Also in G. Webelhuth, ed., *Government and Binding Theory and the Minimalist Program.* Oxford: Blackwell, 1995].

Chomsky, N. 1995. *The Minimalist Program.* Cambridge, Mass.: MIT Press.

Cinque, G. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24, 239-298.

Culicover, P. and M. Rochemont. 1983. Stress and focus in English. *Language* 59, 123-165.

Epstein, S. 1994. The derivation of syntactic relations. Ms., Harvard University.

Freidin, R. and J.R. Vergnaud. 1999. A presentation of Minimalism. Ms., C.N.R.S. and Princeton University.

Gussenhoven, C. 1984. *On the grammar and semantics of sentence accents.* Dordrecht: Foris.

Halle, M. and J.R. Vergnaud. 1987. *An essay on stress.* Cambridge, Mass: MIT Press.

Hayes, B. and A. Lahiri, 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9, 47-98.

Höhle, T. 1981. Explication für "normale Betonung" und "normale Wortstellung". In W. Abraham, ed., *Erklärende syntax des Deutschen.* Tübingen: Narr.

Jackendoff, R. 1974. *Semantic interpretation in generative grammar.* Cambridge, Mass: MIT Press.

Kayne, R.S. 1994. *The antisymmetry of syntax.* Cambridge, Mass: MIT Press.

Kiparsky, P. 1973. "Elsewhere" in phonology, in *A Festschrift for Morris Halle*, eds., S.R. Anderson and P. Kiparsky, New York: Holt, Rinehart, and Winston, 93-106.

Krifka, M. 1984. Focus, Topic, syntaktische Struktur and semantische Interpretation. Ms., Universität Tübingen.

Ladd, D. R. 1980. *The structure of intonational meaning.* Bloomington: Indiana University Press.

Liberman, M. 1975. *The intonational system of English.* Doctoral dissertation, Department of Linguistics, MOT, Cambridge, Mass.

Martin, P. 1975. Analyse phonologique de la phrase francaise, *Linguistics* 146, 35-68.

Martin, P. 1982. Phonetic Realisations of Prosodic Contours in French, *Speech Communication* 1, 283-294.

Martin, P. 1987. Prosodic and rhythmic structures in French, *Linguistics* 25, 925-949.

Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. Doctoral dissertation, Department of Linguistics, MIT, Cambridge, Mass.

Rochemont, M. 1986. *Focus in generative grammar.* Amsterdam: John Benjamins.

Rooth, M. 1985. Association with focus. Doctoral dissertation, Department of Linguistics, University of Massachusetts, Amherst.

Tuller, L. 1992. The syntax of postverbal focus construction in Chadic. *Natural Language and Linguistic Theory* 10, 303-334.

Schmerling, S. 1976. *Aspects of English sentence stress*. Austin: University of Texas Press.

Selkirk, E. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, Mass: MIT Press.

Selkirk, E. 1995. Sentence prosody: Intonation, stress, and phrasing. In J. Goldsmith, ed., *The handbook of phonological theory*. Oxford: Blackwell.

Simpson, Andrew & Xiu-Zhi Zoe Wu 1999. IP-raising, tone-sandhi and the creation of particles: evidence for PF movement/cyclic Spell-Out. Ms., SOAS/USC.

Simpson, Andrew & Xiu-Zhi Zoe Wu 2000. Agreement, Shells and Focus. Ms., SOAS/USC.

Suñer, M. 2000. Object-shift: Comparing a Romance Language to Germanic. Probus 12.2.

Szabolsci, A. 1981. The semantics of topic-focus articulation. In T. Jansen and M. Stokhof, eds., *Formal methods in the study of language*. Amsterdam: Mathematisch Centrum.

Vallduvì, E. and E. Engdhal. 1995. The linguistic realisation of information packaging. Ms., Centre for Cognitive Science and Human Communication Research Centre, University of Edinburgh.

Vergnaud, JR. and M.L. Zubizarreta. 2000. Intervention effects in French wh-in-situ: syntax or interpretation? Ms. University of Southern California.

Zubizarreta, M.L. 1998. *Prosody, Focus, and Word Order*. Cambridge, Mass: MIT Press.

Notes

---

[1] 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).

[2] 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)

[3]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.

[4] See also Liberman 1975, Ladd 1980, Pierrehumbert 1980, Culicover and Rochemont 1983, Halle and Vergnaud 1987, Cinque 1993, Zubizarreta 1998.

[5] For example, this could be the structure for the sentence *John's son's baby cried*.

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

(i) $V^1 \ldots$ $\ldots$]

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.

[7] 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.

According to the definition in (i) above, Specifier ZP below c-commands the head X and its complement YP (as in the standard definition of c-command). According to the definition in (ii) above, ZP indirectly c-commands its sister 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.

```
                    XP

        ZP                  X'

            X        YP
```
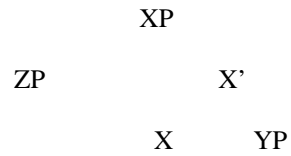
[14] In Zubizarreta 1998, selectional ordering was defined on a separate, more abstract tree. Here, we assume that the syntactic tree encodes both selectional ordering and constituent ordering in terms of asymmetric c-command.

[15] Arguably, a similar process is at play in the preposing of IP and VP with respect to selecting particle-type functional heads in Chinese. See Simpson & Wu 1999, 2000.

[16] For some discussion, we refer the interested reader to Zubizarreta 1998 (section 2.5).