

16. Sentence Prosody: Intonation, Stress, and Phrasing

ELISABETH SELKIRK

Subject	Theoretical Linguistics » Phonology
DOI:	10.1111/b.9780631201267.1996.00018.x

0 Introduction

The term sentence prosody encompasses three distinct aspects of the phonological representation of the sentence: intonation, phrasal rhythmic patterning, and prosodic phrasing. This article examines the relations between these aspects of sentence prosody as well as the relation of sentence prosody to the meaning of the sentence. Drawing on earlier work by Schmerling (1976), Gussenhoven (1984), Selkirk (1984), and Rochemont (1986) on the relation between intonation and meaning, I intend to show that the distribution of intonational pitch accents in the sentence in English directly constrains the focus structure of the sentence, which in turn constrains the range of meanings available for the sentence (section 1). I will also argue that the distribution of pitch accents places constraints on the possible rhythmic stress patterns of a sentence in English (section 2). The view of the organization of the grammar of English that emerges is one that gives the tonal representation of the sentence – its intonation – a pivotal position, crucially mediating the meaning–sentence prosody relation. In so doing it denies to the phrasal stress pattern of the sentence any such central role.¹

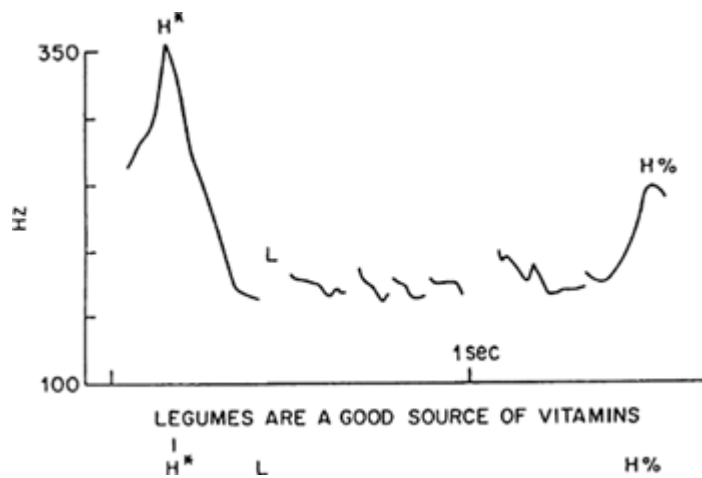
1 The Grammar of Intonation

1.1 The Phonological Representation of Intonation

The term *pitch accent* has been used in the description of tone languages in which the distribution of tone within words is highly restricted. For example, in standard Japanese, the paradigm case of a “pitch accent language” (McCawley 1968), the phonological representation of a word contains at most one locus of tonal activity, a HL pitch fall. In nouns the location of that fall must be specified lexically, in verbs and adjectives it is positioned on the penult syllable. A pitch accent, then, is simply a tonal entity with a restricted distribution within a word. In her groundbreaking work on the phonology and phonetics of intonation in English, Pierrehumbert (1980) shows that the English intonational contour must be analyzed as a sequence of one or more pitch accents, an analysis earlier suggested in the work of Bolinger (1958). In English a pitch accent associates to a stress–prominent syllable in a word (typically the main word stress). Pierrehumbert shows that a variety of pitch accent types, both monotonal (H* or L*) and bitonal (e.g., H*+L), must be recognized for English. (The * marks the tone that is realized on the stressed syllable, the other tone is realized on the adjacent syllable.) This variety accounts for the comparative richness in intonative shapes that characterizes English (see Kingdon 1958; O'Connor and Arnold 1973; Palmer 1938; etc.). Following Liberman (1975), Pierrehumbert isolates a further tonal entity, the *boundary tone*, which is positioned at the edge of a phrasal constituent. The end of an English sentence is necessarily marked by a final H% or L% boundary tone (a “%” following a tone indicates that the tone associates with the final (or initial) syllable of the phrase). Beckman and Pierrehumbert (1986), building on Pierrehumbert (1980), distinguish two types of boundary tone, one associated with the edge of an intonational phrase and another with a smaller, intermediate–level phrase.

Consider the figure in (1),² which graphs a fundamental frequency contour, the phonetic implementation of the tonal representation of the sentence:

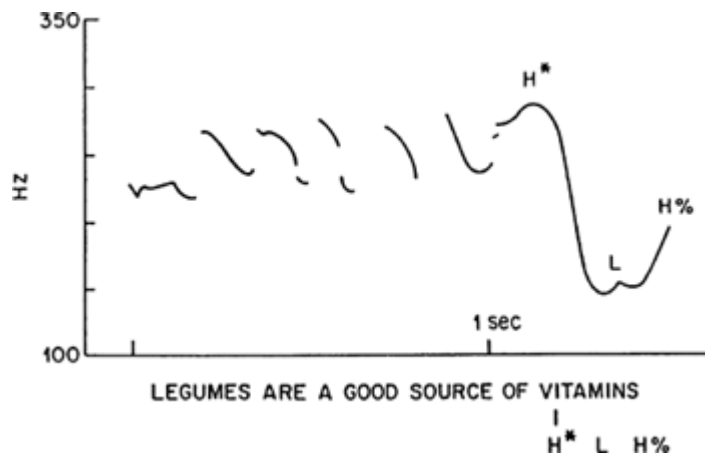
(1)



This intonational contour has a high peak at the stressed syllable of *legumes*; this is followed by a fall to a low pitch, which is maintained, plateau-like, up to the final syllable, where there is an abrupt rise in pitch to another peak. According to the Pierrehumbert (1980) theory of phonetic implementation, H and L tones in the phonological representation mark the targets of the movements in fundamental frequency. In (1) the H tone associated to the stressed syllable *le-* is a pitch accent; the H tone appearing on the sentence-final syllable is a boundary tone. As for the L tone responsible for the fall in pitch following *le-*, I will assume here that it is the second half of a bitonal pitch accent whose first element is the preceding H.³ This utterance of *Legumes are a good source of vitamins* would be appropriate only in certain discourse contexts. The absence of pitch accent on *vitamins* indicates that the predicate noun phrase is "given" in the discourse (see sec. 1.2). The H boundary tone at the end indicates that the sentence is not a simple declarative assertion, which in English would end in a L% boundary tone. Rather, the fall-rise contour of (1) would make it appropriate as a contradiction of the assertion *Nothing in this cupboard is a good source of vitamins* (see, e.g., Ward and Hirschberg 1985).

I will argue below that the presence of a pitch accent on *legumes* implies that the word is focused (in a sense to be made precise). Different distributions of pitch accents imply different focus structures. So, in contrast to (1), the intonational contour in (2) calls for the focusing of *vitamins* and would be appropriate as a contradiction of the assertion *Legumes aren't good for anything*.

(2)



Moreover, if both *legumes* and *vitamins*, or any other combination of words in the sentence, were pitch accented, yet different focus structures and different appropriateness conditions would be defined for the utterance.

It has been shown that in Dutch (Gussenhoven 1984) and German (Selkirk 1984; Uhmman 1987; Féry 1989) pitch accents have the same relation to focus and meaning that they do in English. Pitch accents have also been shown to play a role in the phonological and phonetic characterization of intonation in the Romance languages (see Avesani 1990 on Italian and Sosa 1991 on Spanish) and appear to correlate with the focus structure of sentence as well. In so-called “pitch accent languages” like Swedish, Serbo-Croatian, or Japanese, by contrast, pitch accents have a radically different role in the grammar. In Serbo-Croatian and Japanese, accent is exploited for lexical contrast and makes no more contribution to meaning than, say, the phonological features [labial] and [voice].⁴ In Swedish (see Bruce 1977) pitch accent is not an idiosyncratic property of morphemes, rather it is a redundant property, located on the first (stressed) syllable of every word in the sentence.

Boundary tones appear to be more uniformly exploited crosslinguistically. In English, and in languages as diverse as Chinese, Japanese, Serbo-Croatian, Hungarian, and Igbo, a sentence-final boundary tone may mark whether a sentence is an assertion, a contradiction, a question, etc. Medial boundary tones, such as the H% of the “continuation rise” in English, may set off topics, focus, parenthetical expressions, nonrestrictive relative clauses, and other syntactically definable classes of constituents.⁵

1.2 Intonation and focus in English

Consider the two different utterances in (3), homophonous except for their prosody.

- (3) (a) TRESpassers will be prosecuted.
 (b) Trespassers will be PROsecuted.

(Capitals indicate the presence of a H* pitch accent on a syllable.) Only the first is an appropriate answer to the question “Who will be prosecuted?”. Only the second is a possible answer to “What will be done to trespassers?”. A *wh*-expression focuses a constituent, and an appropriate answer to a *wh*-question must focus the same constituent. I will use the term *Focus* to designate the constituent that is conventionally referred to as “the focus of the sentence.” Generative theories of the semantics of Focus from Jackendoff (1972) to Rooth (1992)⁶ have assumed that in the surface structure of a sentence a Focus constituent is *F-marked*,⁸ thus [[TRESpassers]_F will be prosecuted] vs. [Trespassers will be [PROsecuted]_{F}] for (3a) and (3b), respectively, when uttered as answers to the *wh*-questions above. Retained in the Logical Form of the sentence (LF in the Chomskyan sense [see May 1985]), this F-marking figures crucially in the input to the semantic component. The basic function of Focus is to introduce a set of alternatives into the discourse, which are put to use in various ways in the semantics.⁷ These examples show there is a relation between pitch accent and Focus in English. Given the notion of F-marking, we can understand this as a relation between pitch accent and the F-}

marking of syntactic constituents.

The central problem in the characterization of the prosody–focus relation is the question of what principles govern the relation between pitch accent and F–marking. Call this the problem of *focus projection*. Consider, for example, the different F–markings that are possible when the sentence *Mary bought a book about bats* is pronounced with a pitch accent on *bats*:

- (4) Mary bought a book about BATS.
- (a) Mary bought a book about [BATS]_F.
(What did Mary buy a book about?)
 - (b) Mary bought a book [about BATS]_F.
(What kind of book did Mary buy?)
 - (c) Mary bought [a book about BATS]_F.
(What did Mary buy?)
 - (d) Mary [bought a book about BATS]_F.
(What did Mary do?)
 - (e) [Mary bought a book about BATS]_F.
(What's been happening?)

This same sentence is an appropriate answer to the set of *wh*– questions listed at the right, meaning that the same location of pitch accent is consistent with the corresponding range of F–markings. The sentence *MARY bought a book about bats*, on the other hand, is an appropriate answer to just one question, “Who bought a book about bats?”, which means that only the subject noun phrase in this sentence may be F–marked:

- (5) MARY bought a book about bats.
[MARY]_F bought a book about bats. (Who bought a book about bats?)

The earliest generative solutions to the problem of focus projection (Chomsky 1971; Jackendoff 1972) hypothesized that the location of the intonational accent within the F–marked constituent was determined by the location of the main phrase stress within that constituent. A formulation in the spirit of the Jackendoff (1972) proposal would read as follows:

- (6) Accent to Main Stress within Focus

$$[\dots \sigma_{\text{main stress}} \dots]_F \quad \rightarrow \quad [\dots \underset{\substack{| \\ \text{pitch accent}}}{\sigma_{\text{main stress}}} \dots]_F$$

According to the Nuclear Stress Rule (NSR) of Chomsky and Halle (1968), assumed in earlier accounts, the rightmost word–level stress of a phrase would carry the main stress of the phrase, therefore *bats* would carry the main stress within any of the F–marked constituents in (4) and hence would receive an accent. This NSR–based theory also correctly predicts that an accent on *Mary* in the same sentence would only be possible if it were the subject noun phrase that was F–marked, since only in the subject noun phrase would *Mary* carry the main phrase stress. However, the examination of a broader range of cases – in English and in other languages like German and Dutch – has shown that the Nuclear Stress Rule does not predict the right location of pitch accent within a focused constituent.⁸

Gussenhoven (1984), Selkirk (1984), and Rochemont (1986) – building on initial insights in Schmerling (1976) – argue that the theory of focus projection is instead framed in terms of syntactic notions and in particular makes appeal to the *argument structure* of the sentence. As Gussenhoven (1984) and Selkirk (1984) show, an argument–structure–based account of focus projection in English can be carried over intact to Dutch and German. The theory of focus projection laid out in Gussenhoven (1984, 1991), while different in conception, covers much the same empirical ground as

theory to be laid out here. The Selkirk (1984) theory of focus projection, refined in Rochemont (1986), consists of a set of principles for the *licensing* of F-marking. The Basic Focus Rule states that the assignment of a pitch accent to a word entails the F-marking of the word:

(7) Basic Focus Rule

An accented word is F-marked.

The F-marking of higher constituents is projected from the F-marking of words through two principles, grouped under the heading of Focus Projection:

(8) Focus Projection

F-marking of the *head* of a phrase licenses the F-marking of the phrase.

F-marking of an *internal argument* of a head licenses the F-marking of the head.

This theory of focus projection hypothesizes that the F-marking of the Focus of a sentence is licensed by a chain of F-marked constituents at the bottom end of which is the word bearing the pitch accent. I propose that the Focus of a sentence (FOC) is defined as an F-marked constituent not dominated by any other F-marked constituent. Given this theory, the focus structures compatible with the utterance *Mary bought a book about BATS* will read roughly as in (9) (rather than as in (4)), where FOC marks the highest F-marked constituent.

(9) Mary bought a book about BATS.

- (a) Mary bought a book about _{FOC}[[BATS]_F]_{FOC}.
 (b) Mary bought a book _{FOC}[[about]_F [[BATS]_F]_F]_{FOC}.
 (c) Mary bought _{FOC}[a [book]_F [[about]_F [[BATS]_F]_F]_F]_{FOC}.
 (d) Mary _{FOC}[[bought]_F [a [book]_F [[about]_F [[BATS]_F]_F]_F]_F]_{FOC}.
 (e) _{FOC}[Mary [[bought]_F [a [book]_F [[about]_F [[BATS]_F]_F]_F]_F]_{FOC}.

(For the sake of simplicity, these representations ignore functional projections like IP and DP.) The F-marking of the noun *bats* licenses the F-marking of the NP *bats*, which in turn license the F-marking of its head the preposition *about*. The F-marking of *about* licenses the F-marking of the prepositional phrase, and so on. The F-marking of the direct object *a book about bats* licenses the F-marking of the verb and so of the verb phrase, which via licensing of the various intervening inflectional heads (Pollock 1989), licenses F-marking of the sentence. The syntax-based theory of Focus Projection thus correctly predicts the variety of Focus representations that are possible on the basis of the presence of an accent on BATS. As for the utterance *MARY bought a book about bats*, the F-marking of the accented subject NP cannot license the F-marking of the verb, since the subject is not a sister argument to the verb here, and hence F-marking cannot be projected any further than the subject noun phrase itself. Since there is no F-marked constituent dominating *Mary* in this sentence, *Mary* qualifies as the Focus of the sentence:

(10) MARY bought a book about bats.

_{FOC}[[MARY]_F]_{FOC} bought a book about bats.

This theory thus coincides with the “standard” phrase-stress-based theory of focus projection in its predictions about the accent-Focus relation in the utterances *Mary bought a book about BATS* and *MARY bought a book about bats*. But the two theories differ significantly in their ability to account for (i) the relation between accent and the “given-new” structure of the sentence, (ii) the lack of focus projection from accent in positions where main phrase stress would be predicted, and (iii) the projection of focus from accent in positions that current theories would not define as positions of main phrase stress.

The F-marking that is posited in the theory of focus projection proposed above has a role to play in defining the given–new articulation (the “information. structure”) of the sentence. It appears to be the case that 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. For example, every constituent making up the verb phrase in (10) is interpreted as given in the discourse, while the Focus subject *Mary* may be either given or new. Compare the dual-accented sentence *MARy bought a book about BATS*, uttered in response to the question “What’s been happening?”, which calls for the entire sentence to be the Focus. Accented *Mary* must be interpreted as new in this discourse. This would follow from the F-marking of *Mary* since it is not itself a Focus in this context.

(11) *MARy bought a book about BATS.*

$_{FOC}[[MARY]_F [[bought]_F [a [book]_F [[about]_F [BATS]_{F^1F^1F^1}]]]_{FOC}]_{FOC}$.

(Compare (9e), also a case of sentence Focus, but where unaccented *Mary* must be assumed to be “old information” in the discourse.) Note that in (11) the nonFocus F-marking which our theory says must be a property of *about*, *book*, *bought* and the phrases they head – if sentence Focus is to be projected from *bats* – also calls for the interpretation of these constituents as “new,” and that is the right result.

With this understanding of focus projection and of the given–new interpretation of nonFocus F-marking, we can also explain the phenomenon infelicitously referred to as “deaccenting,” studied in Ladd (1980), Gussenhoven (1984), Selkirk (1984), Rochemont (1986), and exemplified in the utterance *MARy bought a BOOK about bats*. A sentence such as this, which lacks accent on *about bats*, would be appropriately uttered in a discourse where (*about*) *bats* is “given.” (The presence of an accent on *about bats* would entail F-marking and therefore an interpretation as “new.”) The “deaccenting” puzzle is that a higher phrase, even the sentence, may be a Focus with this utterance, despite the absence of accent on *about bats*. This projection of focus is explained by the present theory, since focus may project from a head, in this case *book*:

(12) *MARy bought a book about bats.*

$_{FOC}[[[Mary]_{F^1F^1} [[bought]_F [a [BOOK]_F [[about][bats]_{F^1F^1}]]]_{F^1F^1}]_{FOC}]_{FOC}$.

Indeed, the syntax-based theory of focus projection predicts that if the sentence is the Focus and if *about bats* (but not *book*) is given, then the accent must fall on *book*, since this is the only location of accent in the verb phrase that will both project F-marking to the sentence and at the same time ensure the nonFocus F-marking that the newness of the remaining constituents in the verb phrase requires. That the presence of pitch accent on *book* may indeed license sentence Focus here is shown by the fact that (12) could be uttered in a context where sentence Focus is called for. For example, *only* may associate to the Focus of an embedded sentence in a configuration such as “I only thought that $_{FOC}[\dots]_{FOC}$.” Embedding (12) would give (13a). The fact that sentence (13a) may be appropriately uttered when the alternatives under consideration are entire sentences, such as in (13b)–(13d), rather than ones where nouns like *game* or *poster* substitute for *book*, shows that the accenting of *book* here may license sentence Focus.

(13) (a) I was only thinking that $_{FOC}[MARY bought a BOOK about bats]_{FOC}$.

(b) I was thinking that they (bats) eat mosQUItoes.

(c) I was thinking that we should reMOVE them from South COLLEge.

(d) I was thinking that ROBERt never found the one that bit his CAT.

Embedded sentence Focus would moreover be possible with the utterance *I was only thinking that MARY had BOUGHT a book about bats*, where *a book about bats* is given in the discourse. Note, finally, that this theory allows us to understand why sentence Focus is possible through the accenting of the auxiliary *do*:

(14) $_{FOC}[[Mary] [[DID]_F [buy a book about bats]]_{F^1F^1}]_{FOC}$.

As a head of an inflectional projection (assume it is the affirmation/negation projection of Laka 1990), accented *DID* licenses sentence Focus. Employing the accented auxiliary *do* solves the problem of how to focus a sentence when everything in it is old information.

We have seen that an array of F-markings are possible internal to the Focus of a sentence. NonFocus F-marking is interpreted as “new.” The correlation of given and new with the absence and presence of accent is captured by the theory of Focus Projection proposed here. By contrast, the phrase-stress-based theory of the accent-focus relation, proposed by Jackendoff (1972) (and revised by Cinque 1993), gives no recognition to the given-new articulation that is possible within the Focus phrase. It wrongly predicts that sentence Focus should be impossible in cases of “deaccenting” like (12). And it has no means of characterizing the fact that under the same Focus condition systematic contrasts in the discourse-appropriateness of utterances are possible, depending on the distribution of accents. *MARY* bought a *BOOK* about bats; *MARy* bought a book about *BATS*; *Mary* bought a book about *BATS*; etc., differ in their given-new articulation, but are all possible as cases of sentence Focus.

As Gussenhoven (1984) and Selkirk (1984) point out, a non-argument does not project focus. It is well known that an accented pre-head modifier does not project focus. The utterance *I saw my OLDER sister* has Focus restricted to the adjective phrase, and so sets up alternatives like *I saw my YOUNGER sister* or *I saw my MIDDLE sister* which all presuppose that “I saw my x sister” is given in the discourse. Accented adjunct phrases in post-head position also fail to project focus. For example, focus does not project from an adjunct locative PP in *He smoked in the TENT*. For such a sentence to have VP or sentence Focus, there must be an accent on the verb: *He SMOKED in the TENT* (or *He SMOKED in the tent*). (Experimental work (Gussenhoven 1983) confirms this distinction.) Placement of Focus-sensitive *only* before the VP brings out the difference:

- (15) (a) He only smoked_{FOC} [in the TENT]_{FOC}.
 *He only_{FOC} [smoked in the TENT]_{FOC}.
 (b) He only_{FOC} [SMOKED in the TENT]_{FOC}.

The alternatives implied by the pronunciation in (15a) include *He smoked by the campfire*, *He smoked in the outhouse*, and other sentences where the locative PP is substituted for. (15b), on the other hand, may evoke alternatives like *He left trash on the trail*, *He picked the endangered toadflax*, or other such infractions of sound eco-policy. Compare the case where a post-verbal PP is an argument; here accenting of the prepositional object licenses focus on the VP:

- (16) He only_{FOC} [looked at the GARDen]_{FOC}.

(16) is a possible answer to the question “What did he do?”. Alternatives could be *He put the sprinkler on*, *He refilled the bird feeder*, and so on, showing VP Focus and hence projection of F-marking from accented *garden*.

There is a further broad range of counterexamples to the phrase-stress-based theory of focus projection, ones which are readily accounted for with a minor revision of the theory of Focus Projection given in (8). These involve projection of focus from an argument which is neither phrase-final nor sister to its head in surface structure. It is widely recognized that accent on the subject of an intransitive verb may license sentence Focus:

- (17) FOC [JOHNSON died]_{FOC}.
 FOC [The SKY is falling]_{FOC}.
 FOC [The SUN came out]_{FOC}.
 FOC [The BABY's crying]_{FOC}.

These utterances are appropriate “out-of-the-blue” responses to “What's been happening?,” or may be embedded in the sentence frame “I was only thinking that FOC [.....]_{FOC}.” Note that alternative pronunciations with additional accent on the verb, e.g., *JOHNson DIED*, are equally possible, with

identical F-marking possibilities. The interesting point is that the accent on the verb is not necessary. The utterance *JOHNson died* is appropriate in a context where both *Johnson* and *dying* are new in the discourse and where the sentence is the Focus.

Current syntactic analysis understands the surface subject in the sentences in (17) to be generated within the verb phrase in deep structure (see Diesing 1992 and references therein). Only a small amendment is required for the syntax-based theory of Focus Projection to be able to account for the accent Focus relation in these cases, namely the provision that the F-marking of a constituent licenses the F-marking of its *trace*. The Focus status of the sentence *JOHNson died* is therefore licensed in virtue of the F-marking chain given in (18):

(18) ${}_{\text{FOC}}[[[\text{JOHNson}]_{\text{F}_1}]_{\text{F}_2} \text{VP}[[\text{t}]_{\text{F}_3} [\text{died}]_{\text{F}_4}]_{\text{F}_5}]_{\text{FOC}}$.

An alternative to this appeal to trace would be to assume that Focus Projection held at the level of Logical Form, where the raised subject NP would be restored to its deep structure position through reconstruction. The trace-based account is adopted here.

The notion that focus projection from an argument to its head proceeds through the mediation of a trace allows us to explain other classic puzzles. For example, Gussenhoven (1992) observes that while accenting *clock* in (19a), a small clause structure, allows for *tick* to be interpreted as new information and for the entire sentence to be Focus, in (19b) accenting *clock* does not allow the accentless *tick* to be interpreted as new, and the sentence is not a Focus:

(19) (a) I heard a CLOCK tick.
(b) I forced the CLOCK to tick.

These facts follow, given the proposed amendment to Focus Projection (as they do in the Gussenhoven 1984/1991 theory). The surface syntactic representation of (19b) involves a control structure, with a PRO in the VP-internal subject position, whereas the surface representation of (19a) has a trace in that position, the subject having been moved out of the VP by NP-movement:

(20) (a) ${}_{\text{FOC}}[\text{I heard } [[\text{a } [\text{CLOCK}]_{\text{F}}]_{\text{F}} \text{VP}[[\text{t}]_{\text{F}} [\text{tick}]_{\text{F}}]]]_{\text{FOC}}$
(b) $[\text{I forced } [\text{the } [\text{CLOCK}]_{\text{F}}]_{\text{FOC}} \text{ to } \text{VP}[\text{PRO } [\text{tick}]]]]$.

The absence of an F-marked argument or its trace as sister to the accentless *tick* in (20b) accounts for the lack of F-marking on *tick*, and therefore its interpretation as given and failure to project focus to higher phrases.

Two other intriguing sorts of contrast receive a similar account. Gussenhoven observes that when a sentence like *Trespassers will be prosecuted* is uttered, out of the blue, as a description of an event (as in the mouth of the town crier or a headline in the *Gazette*), it may have a pronunciation with accent only on the subject, as in (21a), or with accent on both the subject and the verb, as in (21b):

(21) (a) RESpassers will be prosecuted.
(b) TRESpassers will be PROSecuted.

In both cases sentence Focus is possible and in neither case is *prosecuted* interpreted as given. However, when the sentence does not describe an event but rather has the status of a generic-like statement about the properties of the subject, as it would when posted as a sign on the edge of a neighbor's woods, then accent is required on the VP if it is to be interpreted as new and the sentence is to be in Focus. Diesing (1992) argues that generic sentences involve control structures, where the surface subject controls a VP-internal PRO. As Diesing points out, if we assume this analysis, it explains why projection of focus from the subject (via F-marking of verb, verb phrase, etc.) is not possible in generic sentences, and thus why such sentences require accent within the predicate if they are to qualify as cases of sentence Focus. In the case of the event reading in (21a), where the subject receives an existential interpretation ("There are trespassers who will be prosecuted"), Diesing argues that the subject is raised from a deep structure position within the VP. In this case the trace(s) left by

NP-movement provide the route along which F-marking of the surface subject can ultimately license Focus of the sentence, and hence accent within the predicate is not required.

Gussenhoven (1984, 1992) notes further that the difference between an individual level and a stage-level predicate (Kratzer 1988; Diesing 1992) is reflected in the accentual properties of the sentence under sentence Focus, citing the examples in (22)

- (22) (a) Your EYES are red. or Your EYES are RED.
Your EYES are BLUE. *not* *Your EYES are blue.

A stage-level predicate like redness, when predicated of eyes, does not require accent in an utterance with sentence Focus, while an individual-level predicate like *blue*, when predicated of eyes, does. The pair in (23), discussed by Diesing (1992), have the same possibilities for sentence Focus:

- (23) (a) FIREMEN are available. or FIREMEN are AVAILABLE.
FIREMEN are ALTRUISTIC. *not* *FIREMEN are altruistic.

Altruistic is an individual-level predicate. *Available* is stage-level. Diesing posits a raising from VP-internal position (leaving a trace) for the stage-level case and a control structure, with VP-internal PRO for the individual-level case. She points out that the focus projection facts follow, assuming the Focus Projection theory.

A final puzzle which the appeal to traces in focus projection allows us to explain involves *wh*-movement. The question *Which BOOK did Helen review?* can be appropriately uttered in a discourse where *Helen* is given in the discourse, but *reviewing* is not. That *Helen* is interpreted as old follows from its lack of accent, since the subject NP could not have F-marking licensed by anything else. But the accentless verb *review* can get F-marked, through the trace of the *wh*-moved NP:

- (24) [_{FOC}[What BOOKS]_{FOC} has [[Helen]]_{VP}[[reviewed]_F [t]_F]_F?

This particular sort of example is one of a set of puzzles for the phrase-stress-based theory of focus projection noted by Bresnan (1971, 1972) which the present theory of focus projection solves.⁹

We are now in a position to formulate the necessary revision of our syntax-based theory of Focus Projection:¹⁰

- (25) Focus Projection (Selkirk 1984; Rochemont 1986)
(a) F-marking of the head of a phrase licenses the F-marking of the phrase.
(b) F-marking of an internal argument of a head licenses the F-marking of the head.
(c) F-marking of the antecedent of a trace left by NP- or *wh*-movement licenses the F-marking of the trace.

These three principles of Focus Projection combine with the language-particular Basic Focus Rule ("A word with a pitch accent is F-marked") to define the relation between accent and focus in intonational languages like English, Dutch, and German.

Work by Ladd (1980, 1983), Ward and Hirschberg (1985), Pierrehumbert and Hirschberg (1990), McLemore (1991), and others has sought to show that different choices of pitch accent and boundary tone in a sentence correlate with different conditions of appropriateness in a discourse. To the extent to which different choices of tonal entity—e.g., a H% versus a L% boundary tone or a H* pitch accent versus a L*+H pitch accent—entail differences in pragmatic or semantic interpretation, further support is provided for construing these tonal entities as morphemic in character, varieties of "floating" tonal particle which must occupy a position in surface structure.¹¹ It may turn out that only certain of the pitch accents of English license the projection of focus. Should this be the case, the Basic Focus Rule, which as currently stated makes no distinction between the different pitch accents, would have to be reformulated. However, the theory of Focus Projection, which makes no appeal to accent, would be predicted to remain intact.

2 Phrase Stress

The relative prominence of a syllable within a sentence has been attributed to three different factors: (1) the presence or absence of a pitch accent on the syllable (the *accent* factor), (2) the position of the syllable within a constituent structure (the *phrasing* factor), and (3) the presence or absence of other prominent syllables in the immediate vicinity of the syllable (the *rhythm* factor). A fuller role for phrasing-based constraints on prominence than is conventionally assumed might remove motivation for the rhythmic component of a phonological theory of phrase stress.

2.1 Pitch Accents and Nuclear Stress

Informally stated, the Nuclear Stress Rule reads as follows:

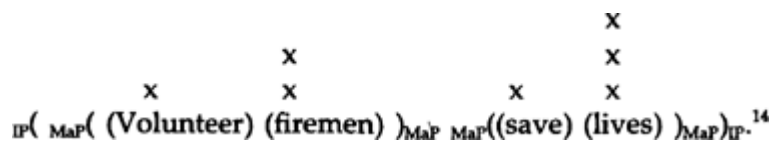
(26) Nuclear Stress Rule

The most prominent syllable of the rightmost constituent in a phrase P is the most prominent syllable of P.

I assume that the constituent structure relevant to the NSR is the *prosodic structure*¹² of the sentence, and that for each constituent in a prosodic structure there is a most prominent syllable. I

assume that prominence relations are represented with a *metrical grid*,¹³ where each constituent-prominence is marked with an x. The NSR would predict the grid prominences in (27) for the sentence *Volunteer firemen save lives*:

(27)



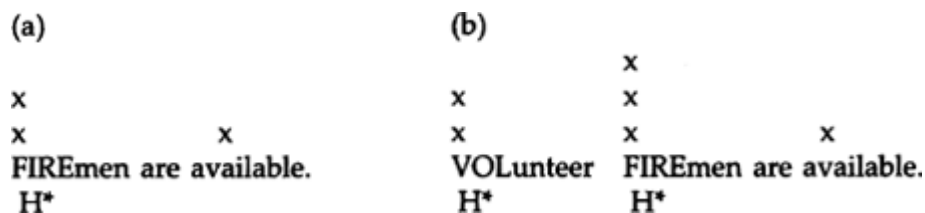
But violations of the NSR are commonplace, as examples discussed in the previous section and other works cited there show. In cases like *FIREmen are available*, *I heard a CLOCK tick*, or *Mary wrote a BOOK about bats*, the most prominent syllable of the sentence is not the main-stressed syllable of the final word or phrase, but rather the pitch accented syllable. The generalization holds that a pitch-accented syllable is always more prominent than a syllable which lacks a pitch accent. Assuming that prominence is not merely a matter of intonation, but also of stress, this generalization reflects that the grammar of English gives place to a constraint on the relation between pitch accent and stress prominence, which Selkirk (1984) formulates as the Pitch Accent Prominence Rule:

(28) Pitch Accent Prominence Rule (PAPR)

A syllable associated to a pitch accent has greater stress prominence than a syllable which is not associated to a pitch accent.

The Pitch Accent Prominence Rule is understood to take precedence over the Nuclear Stress Rule. In *FIREman are available*, the NSR calls for greatest prominence on *-vail-*, while the PAPR calls for greatest prominence on *FIRE-*. The PAPR wins,¹⁵ as shown in (29a). Where the NSR and the PAPR are not in conflict, however, the NSR would be predicted to be in effect. Consider the combined effect of the NSR and the PAPR in (29b), where the accented syllables of both *volunteer* and *firemen* are more prominent than *-vail-*, by virtue of the PAPR, and the NSR assures the greatest prominence of the final pitch accented syllable *Fire-*.

(29)



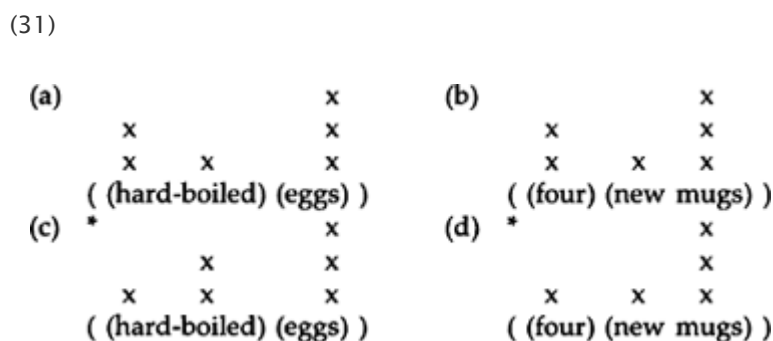
Given the Pitch Accent Prominence Rule, the Nuclear Stress Rule has the status of a default principle, applicable only when "all else is equal." It will indeed be responsible for the phrasal stress pattern in (27), but only when all the words are accented (as in an all-new utterance), or when all are unaccented (because given in the discourse):

- (30) (a) VOLunteer FIREmen SAVE LIVES.
 (b) I already TOLD you that volunteer firemen save lives.

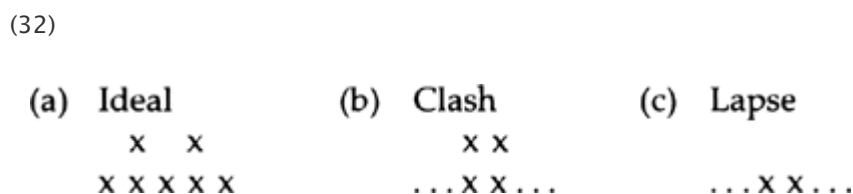
Any study of patterns of phrase stress clearly must control for the presence/ absence of pitch accents. In what follows, assume that the stressed syllables are all pitch accented.

2.2 Rhythmic Alternation Versus Edge Prominence

Consider now the stress pattern of the right-branching and left-branching phrase structures in (31a, b):



The presence of the initial prominence in (31a) and (31b) is not predicted by the NSR, which, instead, predicts the patterns in (31c) and (31d). A standard explanation for the presence of the initial secondary prominence has relied on the idea, due originally to Liberman (1975), that principles of rhythmic alternation complement, and sometimes even override, constituency-based principles like the NSR in defining the stress pattern of a sentence. The ideal arrangement of prominences, the story goes, is one in which beats alternate, one in which both *clash* and *lapse* are avoided.¹⁶



When the rhythmic principles Avoid Clash and Avoid Lapse are violated, as in (31c, d), it has been assumed that these violations may be repaired through operations moving or adding rhythmic prominences (= beats), resulting in rhythmically well-formed surface representations like those in (31a, b).

The claim that rhythm-based principles like Avoid Clash and Avoid Lapse are central to an explanation of the distribution of secondary prominences in examples like (31a, b) is the subject of some debate (see Bolinger 1958, 1965; Hayes 1984; Kager and Visch 1988; Gussenhoven 1991). Bolinger argues that what is at issue is a propensity for pitch accents to be placed closer to the left edge of a constituent. A tendency to “early accent,” even in the absence of stress clash, is amply documented in recent work by Shattuck-Hufnagel, Ostendorff, and Ross (to appear), as well as in work by Beckman and her colleagues (Beckman et al. 1987, 1991). I want to argue that the tendency to early placement of accent is to be ascribed, in part at least, to a constraint calling for stress prominence at the left periphery of a constituent, rather than to a constraint calling directly for accent placement at the left edge. If we assume a peripheral stress prominence constraint we can indeed account for the presence of early accent in standard “inversion” cases like *fôurtèen múgs*, where the first pitch accent is left-peripheral rather than in the otherwise final position of *fourtéer*: in phrasal contexts where the phrase edge prominence constraint calls for word-initial stress prominence, the Pitch Accent Prominence Rule (PAPR) will guarantee that a pitch accent in such a word be realized on the syllable with that initial stress prominence. Yet, more than this, a phrase edge prominence constraint will also explain prominence patterns in collocations like *fôur nèw múgs*, in the case where each of the words is pitch accented. The intuition is that the phrase-initial *four* has greater prominence than *new*, a fact that cannot be explained by the distribution of pitch accents, since they are both accented in this case. Rather, *four* must have greater stress, a fact predicted by assuming the stress-based account of early accent sketched here.

Note that it is a recurrent fact about patterns of word stress that the syllables heading feet at both the right and left periphery of a word are more prominent than those in nonperipheral feet. We see this in English words with three feet, e.g., *Áppalàchicóla, chîmpànzéé*. Of the two edge prominences, the Main Word Stress Rule declares that it is the rightmost one which is most prominent. This word-level patterning is arguably replicated at the phrase-level in English. *Fôur nèw múgs, Fârrah Fàwcett-Májors, côlorless grèen idéas, hârd-bòiled éggs, ôver-dòne stéak, thîrty-nîne stéps* all show peripheral prominences, with the greatest prominence of all on the right, guaranteed by the Nuclear Stress Rule. We assume, then, a role for a Phrase Edge Prominence constraint.

(33) Phrase Edge Prominence

The most prominent syllable of an edge constituent is more prominent than that of a constituent not located at an edge.

(Both Phrase Edge Prominence and the Nuclear Stress Rule are End Rules in the sense of Prince 1983 or edge-alignment constraints in the sense of McCarthy and Prince 1993.)

In the case of *fôur nèw múgs*, (31b), Phrase Edge Prominence and the NSR complement each other. The case of (31a) is more interesting. Satisfying Phrase Edge Prominence with respect to the higher level phrase *hard-boiled eggs* entails violating the NSR within the phrase *hard-boiled*. This sort of constraint interaction, where the violation of one constraint is the consequence of the satisfaction of another, is formalized within optimality theory (Prince and Smolensky, to appear). Note further the stress pattern of the sentence in (34) (cited in Hayes 1991):

(34)

				x
	x			x
	x		x	x
	x	x	x	x

$IP_{(MaP)}((Nineteen\ thousand)(linguists))_{MaP}MaP((sing))_{MaP}IP$

Within the Intonational Phrase the NSR provides for greatest prominence in the sentence on *sing*. Satisfying Phrase Edge Prominence at the level of the IP has the consequence that within the lefthand Major Phrase the prominence on *nineteen* is greater than the prominence on *linguists*, in violation of

the NSR within that MaP. Quite generally, satisfying Phrase Edge Prominence on a higher level phrase can be seen as responsible for the failure to respect the NSR within a lower phrase, rendering superfluous an appeal to rhythmic constraints in many cases.

A theory of early accent based on Phrase Edge Prominence provides a straightforward explanation for a classic puzzle for rhythm-based accounts of early accent—the different locations of main word stress in the medial words of near minimal pairs of the type in (35):¹⁷

(35)

- | | |
|---------------------------------|---------------------------------|
| (a) | (b) |
| ((nôrthern) (Càlifornia wínes)) | ((Nôrthern Califòrnia) (wínes)) |
| ((nôrthern) (Jàpanese béetles)) | ((Nôrthern Japanèse) (áccent)) |
| ((grêen) (bàmboo tábles)) | ((grêen bambòo) (tábles)) |

In the right-branching (a) cases, main stress on the medial words has “moved” to (i.e., appears in) initial position in the embedded phrase. This locus of main word stress allows the embedded phrases to satisfy Phrase Edge Prominence. A comparable positioning of main word stress in the left-branching (b) cases is not to be expected, since it would purchase no better-formedness with respect to Phrase Edge Prominence. Viewed in purely rhythmic terms, the retention of final word stress in the medial words in the (b) cases has no explanation, since it creates a clash with main stress of the following word, just as it does in the (a) cases.

3 Intonational Phrasing

The term *intonational phrase* is usually applied to spans of the utterance which are delimited by boundary tones (cf. sec. 1.1):

(36)

- | | | | | | |
|-----|--------------|----------------------------------|----------------------------|--|----|
| | H% | | H% | | L% |
| (a) | IP(Fred,)IP | IP(who's a volunteer fireman,)IP | IP(teaches third grade)IP. | | |
| | H% | L% | L% | | |
| (b) | IP(Alice,)IP | IP(this is Mary,)IP | IP(my sister)IP. | | |

More than a unit with respect to which the distribution of boundary tones is defined, the intonational phrase (IP) appears to constitute a domain relevant to various aspects of the phonetic implementation of the sentence, including timing effects like constituent-final lengthening,¹⁸ and the setting of register for the realization of tone.¹⁹ Ladd (1986) claims that intonational phrasing is relevant to the phenomenon of “rhythmic inversion” examined in 2.2. In an example like (37a), where the subject is set off as a separate IP, there is no alteration in the locus of main stress in *hotel*, but there can be in (37b), where the subject and what follows occupy the same IP:²⁰

(37)

- | | | | |
|-----|-----------------------------|-----------------|------------------------------|
| (a) | IP(The hoTEL's)IP | IP(TERRible)IP. | [But the beach is great.] |
| (b) | IP(The HOtel's TERRible)IP. | | [It' not even on the beach.] |

The prediction, which seems to be correct, is that a backwards shift in stress away from the edge of a constituent that ends in a boundary tone – the mark of an IP edge – is not possible.

At present, the principles governing intonational phrasing are not well understood. Certain syntactic constructions – vocatives, appositives, parentheticals, preposed clauses, nonrestrictive relative clauses – are necessarily set off in separated IPs.²¹ In other cases, like the examples in (37), there are options in phrasing. There are impossible phrasings. Compare the set of examples in (38), due to Mark Liberman:

(38)

- (a) $IP_{(There\ mathematicians\ in\ ten)} IP_{(derive\ a\ lemma)}$.
- (b) $*IP_{(There\ mathematicians)} IP_{(in\ ten\ derive\ a\ lemma)}$.
- (c) $IP_{(There\ mathematicians)} IP_{(intend\ to\ rival\ Emma)}$.

(38) gives one possible phrasing of the sentence *Three mathematicians in ten derive a lemma*. The phrasing in (38b), which divides the subject noun phrase and groups its second half with the verb phrase, is ungrammatical, however. The nearly homophonous (38c) is fine with that phrasing, since the break falls at the end of the subject noun phrase.

It has been suggested that the material contained within an IP must constitute a “sense unit,”²² though this idea is difficult to implement. It has also been suggested that notions like Topic or Focus are relevant to a theory of intonational phrasing. Cross-linguistic examinations of the phonological and phonetic correlates of prosodic phrasing and of the relation of phrasing to notions like Topic and Focus have only recently begun to accumulate.²³ Together with investigations of the syntax and semantics of Topic and Focus that are “intonationally aware,” they will doubtless form the basis of a fuller understanding of the relation between sentence prosody, meaning, and discourse.

1 Jackendoff (1972), Liberman (1975), Bing (1979), Ladd (1980), Pierrehumbert (1980), Williams (1980), Culicover and Rochemont (1983), Halle and Vergnaud (1987), and Cinque (1993) are proponents of phrase-stress-based theories of the prosody-focus relation.

2 Thanks to Janet Pierrehumbert for these figures, which first appeared in Selkirk (1984).

3 The debate about whether the L tone in such a sentence is part of a bitonal pitch accent or has the status of an independent phrase accent assigned to the edge of the constituent *legumes* (see, e.g., Pierrehumbert 1980 and Ladd 1983) will be ignored here.

4 See Lehiste and Ivić (1986), Inkelas and Zec (1988) on Serbo-Croatian; see McCawley (1968), Poser (1984), and Pierrehumbert and Beckman (1988) on Japanese.

5 See, e.g., Bruce (1977) on the H tone marking Focus in Swedish; Hayes and Lahiri (1991) on focus-marking by H tone in Bengali; and Bing (1980), Ladd (1980), and Beckman and Pierrehumbert (1986) on medial boundary tones in English.

6 The list includes Williams (1980), Rooth (1985), Rochemont (1986), von Stechow (1989, 1991), and Kratzer (1991). See also Carlson (1983, 1984).

7 See Rooth (1985, 1992) for details as to how these alternatives are used in the semantics of various constructions.

8 See Bresnan (1971, 1972), Berman and Szamosi (1972), Bolinger (1972), Schmerling (1976), Ladd (1980), Gussenhoven (1984), Selkirk (1984), Rochemont (1986), and others. Cinque (1993) attempts to remedy the problems for a phrase-stress-based theory by modifying the theory of phrase stress. He rejects the Nuclear Stress Rule and proposes instead that main stress within a phrase is assigned to the “most deeply embedded “constituent” within that phrase. The “greater prominence” he ascribes to the main phrase stress within a Focus constituent derives from the presence of a pitch accent, which in his approach, like Jackendoff’s, could be assigned by the rule of Accent to Main Stress within Focus. Most of the counterexamples to Jackendoff’s NSR-based theory of the prosody-focus relation stand as counterexamples to Cinque’s theory of focus projection as well.

9 See Selkirk (1984) and Rochemont (1986) for further discussion of the Bresnan cases.

10 The cases examined involve traces left by NP-movement and *wh*-movement. A further case from Gussenhoven (1984, 1992) suggests that the traces left by quantified NPs do not project focus. Whereas the utterance *The PRISONERS escaped* allows the interpretation of *escaped* as new information and allows for VP or sentence Focus, the utterance *EVERYbody escaped* does not. Moreover, Angelika Kratzer (p.c.) points out that in German a noun phrase that has undergone Scrambling cannot project focus.

11 These would fall into the same class as the well-known and little understood discourse particles of German, Greek, and other languages.

12 See, e.g., Selkirk (1981, 1986), Nespor and Vogel (1986), Pierrehumbert and Beckman (1988), Selkirk and Tateishi (1988, 1990), Selkirk and Shen (1990), and the chapter by Inkelas and Zec, (chap. 15, this volume).

13 See chapters 10 and 11, this volume, as well as Liberman (1975), Liberman and Prince (1977), Prince (1983), and Selkirk (1984).

14 IP = Intonational Phrase, MaP = Major Phonological Phrase.

15 This sort of constraint interaction finds a welcome formalization in the context of optimality theory (Prince and Smolensky, to appear).

16 On the role of clash and lapse avoidance in stress patterning, see Liberman and Prince (1977), Nespor and Vogel (1979, 1989), Bing (1980), Prince (1983), Selkirk (1984) and numerous others.

17 See Prince (1983), Hammond (1984), Hayes (1984), Selkirk (1984), Kager and Visch (1988), Hayes (1991), and especially Gussenhoven (1991), by whom these pairs are inspired.

18 See, e.g., Selkirk (1984), Ladd (1986), Edwards and Beckman (1988), Beckman and Edwards (1989), Price et al. (1991), and Gussenhoven and Rietveld (1992).

19 See, e.g., Ladd 1988.

20 Vogel and Kenesei (1990a) report contrasts in inversion which they say are dependent on the focus structure of the sentence. These cases probably reduce to differences in intonational phrasing such as those examined here.

21 See especially Bing (1979).

22 See Selkirk (1984) and Steedman (1991), for example.

23 See Poser (1984) and Pierrehumbert and Beckman (1988) on Japanese, Kanerva (1989, 1990) on Chichewa, Vogel and Kenesei (1990a, 1990b), and Rosenthal (1992) on Hungarian, and Vallduví (1990) on Catalan.

Cite this article

SELKIRK, ELISABETH. "Sentence Prosody: Intonation, Stress, and Phrasing." *The Handbook of Phonological Theory*. Glodsmith, John A. Blackwell Publishing, 1996. Blackwell Reference Online. 31 December 2007 <http://www.blackwellreference.com/subscriber/tocnode?id=g9780631201267_chunk_g978063120126718>

Bibliographic Details

The Handbook of Phonological Theory

Edited by: John A. Glodsmith

eISBN: 9780631201267

Print publication date: 1996

