# Chapter 5

Structure and Intonation

Take care of the sense, and the sounds will take care of themselves.
Lewis Carroll, *Alice in Wonderland*

In chapter 2, (13), we considered the following exchange, in which intonation imposes a perceptual grouping of words in the spoken utterance into fragments that are inconsistent with traditional linguistic notions of syntactic constituency.[1]

(1) a. I know that Alice married Alan. But who did ANNA marry?
    b. (ANNA married) (MANNY).

The prosody informally indicated in (1b) by capitals (for stress) and parentheses (for intonational phrase boundaries) is one possibility (among others that we will come to later) for an answer to the question (1a). It consists in not only marking the focused information in the answer by the use of high pitch on the stressed first syllable of the word *MANNY*, but also in stressing the first syllable of *Anna*, using a high pitch-accent, and placing a final rise at the end of *married*, with lower pitch interpolated in between. This utterance might give rise to a pitch contour something like the sketch in figure 5.1, which is an idealized version of the actual pitch track shown in 5.2. This contour conveys the contrast between the previous topic concerning Alice's marriage and the new one concerning Anna's, and it imposes the perceptual grouping indicated by the brackets. Such a grouping cuts across the traditional syntactic analysis of the sentence as a subject and a predicate VP.

Many authorities, such as Chomsky (1971), Jackendoff (1972), Cooper and Paccia-Cooper (1980), Gee and Grosjean (1983), Kaisse (1985), and Croft (1995) have continued to argue, nevertheless, that intonation can be driven directly from Surface Structure. There is an immediate intuitive appeal to the idea, as noted in chapter 2, for it is hard to see why intonation should depart from the Constituent Condition on Rules in any language. However, the apparent complexities engendered by examples like (1) have led many others, such

(ANNA   married)(MANNY)
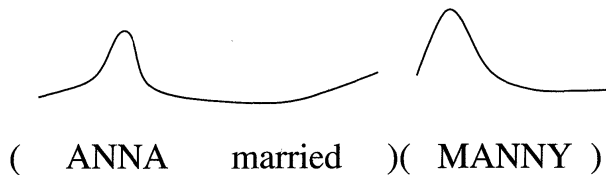
Figure 5.1
Idealized pitch contour for (1b)
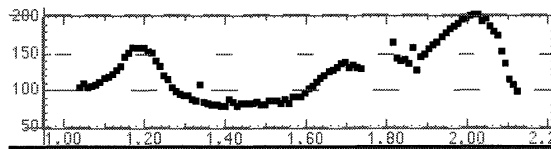
Figure 5.2
Actual pitch contour for (1b)

as Liberman (1975), Goldsmith (1976), Pierrehumbert (1980), Selkirk (1984), and Nespor and Vogel (1986), to postulate an autonomous level of "Intonation Structure" independent of Surface Structure and related only indirectly to Logical Form or function/argument structure, via Information Structure.

However compelling the logic of this argument may appear, we noted in chapter 2 that the involvement of two apparently autonomous levels of structure, related to two autonomous levels of meaning representation, complicates the theory considerably. The picture becomes even bleaker when it is realized that the two levels of structure must communicate, because of the presence of certain focusing constructions and operators, such as the English topicalization construction or the focusing particle *only*, the latter exemplified in the following sentence:

(2) John introduced only BILL to Sue.

Such constructions and particles, which have been discussed by Rooth (1985), von Stechow (1991), Hoeksema and Zwarts (1991), and others, have effects in both domains. These observations have seemed to demand the quite complex theoretical architecture shown in figure 2.3. Such a theoretical architecture offers a view of sentence structure as having an "autosegmental" topology which Halle influentially likened to that of a spiral-bound notebook (cf. Halle and Vergnaud 1987 p78-79). This notebook has phonetic segments arranged along the spine, and different autonomous levels of structure—prosodic, syntactic, and others—written on different leaves, each of which may refer to descrip-

tions on other pages. As Zwicky and Pullum (1987, 4) have pointed out, such theories are potentially very unconstrained, in the absence of a principled statement about which of the pages may cross-refer, and why. The simplest possible constraint upon such a theory would be a demonstration that certain communicating levels involve isomorphic structural descriptions, for those levels at least could be combined upon a single page of the notebook.

However, a strong hint that a simplification might be possible seems to be provided by the observation that Intonation Structure is, despite its apparent partial independence from syntax, nonetheless constrained by meaning, and in particular by distinctions of focus, information, and propositional attitude toward concepts and entities in the discourse model. The intonation contour in the response in (1) seems to divide the utterance into a topic or theme to do with *Anna marrying*, and a comment or rheme *Manny*.[2] These terms will be defined formally below, but informally the theme can be thought of as denoting what the speaker assumes to be *the question under discussion* and the rheme can be thought of as what the speaker believes to be *the part of the utterance that advances the discussion*. Even in advance of a more formal definition, it will be convenient to refer to such partitions of the information in the proposition as the "Information Structure" of an utterance.

A theme in the present sense of the term can be more concretely exemplified as *that which is introduced into the discourse context by a* wh-*question*. By now we are familiar with the idea that such an entity can be expressed as a functional abstraction, as Jackendoff (1972) and Sag (1976) point out, equivalent in this case to the following λ-term:[3]

(3) $\lambda x.marry' x\ anna'$

Establishing a theme with content (3) in the context via a *wh*-question such as *Who did Anna marry?* is one way to make the intonation contour in (1) felicitous. (Of course, it is not claimed that an explicit mention, via a *wh*-question* or otherwise, is necessary for interpreting the response. Nor does this *wh*-question uniquely determine this response—for example, for reasons that we will come to later, it is possible to answer the same question with the "fall-rise" contour confined to *Anna* and the boundary before the VP *married Manny*. There is also no claim that intonation contours in general determine Information Structure unambiguously. We return to these points in section 5.5.2 below.)

The close relation in English of Intonation Structure to Information Structure, first proposed by Halliday (1967a), has recently been endorsed by Selkirk

(1984, 284) as "The Sense Unit Condition" on intonational constituency, which says in essence that intonational constituents must have coherent translations at Information Structure.

However, we have seen in previous chapters that natural languages include a number of other constructions whose semantics is also reminiscent of functional abstraction. The most obvious and theoretically tractable class are Wh-constructions, in which many of the same fragments that can be delineated by a single intonation contour appear as the residue of the subordinate clause. Another and much more diverse class are the fragments resulting from the coordinate constructions discussed in previous chapters. The latter constructions are doubly interesting, because they and certain other sentence-fragmenting constructions, such as parentheticals, interact very strongly with intonation, on occasion making intonation breaks obligatory, rather than optional, as Downing (1970) and Bing (1979), among others, have noted. For example, the intonation indicated on the following ambiguous sentence forces one syntactic analysis with an absurd reading, and leaves the sensible analysis quite inaccessible (the example is from Pierrehumbert (1980), who attributes it to Mark Liberman):

(4) *(Harry likes the NUTS) (and bolts APPROACH).

It is therefore tempting to think that the nonstandard concept of Surface Structure and constituency that has been developed in earlier chapters in order to explain coordination and unbounded dependency might directly provide the notion of structure that is required to account for intonational prosody. If this conclusion is correct, then both the camps identified earlier are in a sense correct. Intonation can indeed be specified directly from surface syntactic structure, without the mediation of an autonomous Intonation Structure. However, the syntactic structure in question corresponds to Information Structure rather than traditional Surface Structure, and hence directly subsumes the Intonation Structure of English.

5.1 Surface Structure and Intonation Structure

According to the combinatory theory, conjoinable strings like *Anna married* and even *a policeman a flower* correspond to constituents in their own right, without syntactic operations of deletion or movement. It follows that they must also be possible constituents of simple noncoordinate sentences like *Give a policeman a flower* and (1b), *Anna married Manny*, as well. Such sentences have several Surface Structures, corresponding to different sequences of com-

position and application. As we have seen, the nonstandard derivation (5) is allowed for the latter sentence, as well as the traditional derivation (6):

$$
\begin{array}{ccc}
(5) & \text{Anna} & \text{married} \qquad \text{Manny} \\
\hline
& NP : anna' & (S\backslash NP)/NP : marry' \quad NP^{\uparrow} : manny' \\
\hline
& S/(S\backslash NP) : \lambda f.f\, anna' & \\
& \hline \qquad\qquad\qquad >\text{B} \\
& S/NP : \lambda x.marry'x\, anna' & \\
& \hline \qquad\qquad\qquad\qquad\qquad\qquad < \\
& S : marry'\, manny'\, anna' &
\end{array}
$$

(5)

Anna	married	Manny

$NP : anna'$ $(S\backslash NP)/NP : marry'$ $NP^{\uparrow} : manny'$

$\overline{S/(S\backslash NP) : \lambda f.f\, anna'}>\text{T}$

$\overline{\qquad\qquad\qquad S/NP : \lambda x.marry'x\, anna'}>\text{B}$

$\overline{\qquad\qquad\qquad\qquad\qquad S : marry'\, manny'\, anna'}<$

(6)

Anna	married	Manny

$NP : anna'$ $(S\backslash NP)/NP : marry'$ $NP^{\uparrow} : manny'$

$\overline{S/(S\backslash NP) : \lambda f.f\, anna'}>\text{T}$

$\overline{\qquad\qquad\qquad S\backslash NP : \lambda y.marry'\, manny'\, y}<$

$\overline{\qquad\qquad\qquad S : marry'\, manny'\, anna'}>$

Such families of derivations form equivalence classes, for the semantics of the combinatory rules guarantees that all such derivations will deliver an interpretation determining the same function-argument relations—in this case, $marry'\, manny'\, anna'$. Moreover, the interpretation of the nonstandard constituent *Anna married* of type S/NP bears an interpretation equivalent to the abstraction (3).

It is therefore tempting to believe that these semantically equivalent derivations convey distinctions of discourse information and that they are on occasion distinguished by intonational markers in spoken language.[4]

For example, the following bracketings correspond to alternative CCG Surface Structures, arising out of different sequences of compositions and applications, each of which corresponds directly to a possible intonation contour:

(7) a. (I)(want to begin to try to write a play).
   b. (I want)(to begin to try to write a play).
   c. (I want to begin)(to try to write a play).
   d. (I want to begin to try)(to write a play).
   e. (I want to begin to try to write)(a play).

The leftmost element is in every case a fragment that can be coordinated. For example:

(8) *I wanted, and you expected,* to write a play.

Conversely, the following are at least as strange (and pragmatically demanding) as coordinations as they are as intonational phrases:

(9) a. ?(I want to BEGIN to), (try to write a PLAY).
     b. ?I wanted to, and you actually expected to, try to write a play.

(Examples like (7) and (9a) are used by Selkirk (1984, 294) to motivate a definition of the Sense Unit Condition in terms of a relation over the heads of constituents.) A stronger example emerges from comparing the following examples, in which the string *three mathematicians* is as hard to make an intonational phrase as it is to coordinate. (The unacceptability of (10a) is also used by Selkirk as evidence for the Sense Unit Condition.)[5]

(10) a. ?(Three MATHEMATICIANS) (in ten prefer MARGARINE).
      b. ?Three mathematicians, in ten prefer margarine, and in a hundred can cook a passable soufflé.

It is irrelevant to the present purpose to ask *how* sentences like (10b) might be excluded, or even to ask whether what is wrong with them is a matter of syntax, semantics, or pragmatics. The important point for present purposes is that the *same* constraint applies in syntactic and prosodic domains. That is, the Sense Unit Condition on prosodic constituents simply boils down to the Constituent Condition on Rules of grammar. This result is a very reasonable one, for what else but a constituent could we expect to be subject to the requirement of being a semantic unit?

   It follows that we predict the strongest possible conspiracy between prosodic constituency and coordinate structure. Noncoordinate sentences typically have many equivalent combinatory derivations, because composition is optional and associative. These analyses can give rise to many different intonation contours. On the other hand, coordinate sentences, like relative clauses, have fewer equivalent analyses, because only analyses that make the conjuncts into constituents are allowed. Two predictions follow. First, we must expect that any substring that can constitute a prosodic constituent will also be able to coordinate. Second, of all the intonational tunes that distinguish alternative prosodic constituencies in noncoordinate sentences, we predict that only the ones that are consistent with the constituents demanded by the coordination rule will be allowed in coordinate sentences. Intonation contours that are appropriate to the alternative constituencies are syntactically ruled out. So for example, there are many prosodic constituencies for the example (7), *I want to begin to try to write a play*, realized by a variety of intonational contours. However, there

are many fewer possible intonation contours for the following coordinate sentence, and they seem intuitively to be closely related to the ones that impose the corresponding bracketing (7e) in the simpler sentence:

(11) I want to begin to try to write, and you hope to produce, a musical based on the life of Sir Stafford Cripps.

Observations like the above make it seem likely intonation often determines which of the many possible bracketings permitted by the combinatory syntax of English is intended, and that the interpretations of the constituents are related to distinctions of information-structural significance among the concepts that the speaker has in mind. Thus, whatever problems for parsing written text arise from the profusion of equivalent alternative Surface Structures engendered by this theory, these "spurious" ambiguities seem to be to some extent resolved by prosody in spoken language. The theory therefore offers the possibility that prosody and syntax are one system, and that speech processing and parsing can be unified in a single process.[6]

This section and the next show that the combinatory rules of syntax that have been proposed in order to explain coordination and unbounded dependency in English do indeed induce Surface Structures that subsume the structures that have been proposed by Selkirk (1984) and others in order to explain the possible intonation contours for all sentences of English. The proof of this claim depends upon two results. First, it must be shown that the rules of combinatory grammar can be made sensitive to intonation contour, so as to limit the permissible derivations for spoken sentences like (1b). Second, it must be shown that the interpretations of the principal constituents of these derivations correspond to the Information Structure established by the context to which they are appropriate, such as (1a).

5.2 Two Intonation Contours and Their Functions

I will use a notation for intonation contours that is based on the theory of Pierrehumbert (1980), itself a development of proposals by Bruce (1977), Liberman (1975), and Goldsmith (1976). The version used here is roughly as presented in Selkirk 1984, Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, and Pierrehumbert and Hirschberg 1990, although it will become clear below that I have extended Pierrehumbert's notation in a couple of minor respects. I have tried as far as possible to take my examples and the associated intonational annotations from those authors.

The advantage of this theory is that it specifies intonation contour independently of the string, in terms of just two kinds of fixed points or "tones." The contour between tones can be determined by interpolation. The first group of tones are the "pitch-accents," which are the substantial local pitch maxima and minima that coincide with perceived contrastive emphasis. The other group of tones are the "boundaries," that mark the right-hand edge of a prosodic phrase. I follow Pierrehumbert in assuming that intonation contours can be described in terms of two abstract or relative pitch levels, H and L, denoting high or low abstract pitch.

Of Pierrehumbert's six pitch accent tones, I will consider only two here, H* and L+H*.[7] The phonetic or acoustic realization of pitch accents is a complex matter. Roughly speaking, the L+H* pitch accent that is extensively discussed below in the context of the L+H* LH% melody generally appears as a maximum that is preceded by a distinctive low level and peaks later than the corresponding H* pitch accent when the same sequence is spoken with the H* LL% melody. (See Silverman 1988 for discussion. Nothing in the combinatory theory hinges on the precise identities of the pitch accent types. All that matters is that the two complete melodies are distinct, a matter on which all theories agree.)

The intonational constituents of interest here are made up of one or more pitch accents (possibly preceded by other material), followed by a boundary. In recent versions of the theory, Pierrehumbert and her colleagues distinguish two distinct levels of prosodic phrase: the intonational phrase proper and the "intermediate phrase." The intermediate phrase boundary is a bare phrasal tone, either L or H.[8] Intonational phrase boundaries are L or H phrasal tones plus an boundary tone written H% or L%. We will principally be concerned here with the intonational phrase boundaries that are written LH%, and the boundary L or LL%.

The intermediate phrase is distinguished in Pierrehumbert's theory as defining the domain of a phenomenon known as "downstep." If more than one pitch accent occurs without an intervening boundary—that is, within an intermediate phrase—then the entire pitch range of each successive pitch accent is shifted downward from its predecessor. At the intermediate phrase boundary (and therefore at any higher-level boundary including the intonational phrase boundary), the pitch levels are reset to the normal level. Although this aspect of the pitch contour is completely rule-governed, so that in Pierrehumbert's own notation, downstepped pitch accents are not distinguished, it is sometimes useful to include such a notation. On such occasions will use "!" as a prefix to

the pitch accent type, a notation that originates with Ladd (1980) and has been included in the ToBI conventions, writing such sequences as follows:

(12) blueberries, bayberries, raspberries, mulberries, and brambleberries
      H*   !H*   !H*   !H*   !H* LL%

I have followed Beckman and Pierrehumbert in regarding boundaries of all kinds as confined to the right-hand end of the prosodic phrase. However, the position and nature of the phrasal tone is one of the more controversial details of the theory (Pierrehumbert and Beckman 1988, 236–237). The influence of, say, an L or LL% boundary on a preceding H* pitch accent is apparent immediately after the maximum, no matter how distant the right-hand boundary is. Pierrehumbert and Beckman point out that this influence may be apparent by the end of the word bearing the pitch accent preceding the boundary. Indeed, in the framework of the British School, the event corresponding to the phrasal tone component L or H of the boundary is considered to be part of the pitch accent, rather than part of the boundary event.

For all other regions of the prosodic phrase, notably the region before the (first) pitch accent, the regions between pitch accents, and the region between pitch accent and boundary, the pitch contour is merely interpolated. In Pierrehumbert's notation, such substrings therefore bear no indication of abstract tone whatsoever. It is sometimes convenient to regard such elements as bearing a "null" tone.[9]

Thus, according to this theory, the shape of a given pitch accent in a prosodic phrase, and of its phrase accent and the associated right-hand boundary, is essentially invariant. If the constituent is very short—say, a monosyllabic NP— then the whole intonation contour may be packed into that one syllable. If the constituent is longer, then the pitch accent will appear further to the left of the phrasal tone and boundary tone at the right-hand edge. The intervening pitch contour will merely be interpolated, as will any part of the contour preceding the pitch accent(s). In this way, the same tune can be spread over longer or shorter strings, in order to mark the corresponding constituents for the particular distinction of information and propositional attitude that the melody denotes.

Consider the prosody of the sentence *Anna married Manny* in the following pair of discourse settings, which are adapted from Jackendoff (1972, 260) and Steedman (1991a). To aid the exposition, words bearing pitch accents are printed in capitals, and prosodic phrase boundaries are explicitly marked in the sentences, using parentheses. (These devices are not part of Pierrehumbert's

notation.)

(13) Q: Well, what about MANNY? Who married HIM?
     A: (ANNA) (married MANNY).
        H* L   L+H* LH%

(14) Q: Well, what about ANNA? Who did SHE marry?
     A: (ANNA married) (MANNY).
      L+H*   LH%   H* LL%

In these contexts the stressed syllables on both *Anna* and *Manny* receive a pitch accent, but a different one. In answer (13A) there is a prosodic phrase on *Anna* made up of the sharply rising pitch accent that Pierrehumbert calls H*, immediately followed by an L boundary, perceived as a rapid fall to low pitch. There is another prosodic phrase having the somewhat later- and lower-rising pitch accent called L+H* on *Manny*, preceded by null tone (and therefore interpolated low pitch) on the word *married*, and immediately followed by a rising "continuation" boundary, written LH%. (See Pierrehumbert and Hirschberg 1990, (33), for discussion of a similar example.)[10] In answer (14A) the two tunes are reversed (see figures 5.1 and 5.2): this time the tune with pitch accent L+H* and boundary LH% is spread across a prosodic phrase *Anna married*, and the other tune with pitch accent H* and boundary LL% is carried by the prosodic phrase *Manny*, again starting with an interpolated or null tone.[11]

The intuition that there is some systematic distinction in meaning between these tunes seems to be very compelling, though it has in the past proved hard to formalize. The tunes have been associated with such factors as social attitude (O'Connor and Arnold 1961; Merin 1983; Bartels 1997), illocutionary acts (Liberman and Sag 1974; Sag and Liberman 1975; Liberman 1975), propositional attitudes (Ward and Hirschberg 1985), maintenance of mutual belief (Pierrehumbert and Hirschberg 1990), and Information Structure (Halliday 1967a; Jackendoff 1972; Schmerling 1976; Ladd 1980; Gussenhoven 1983; Selkirk 1984; Terken 1984, Cormack 1992, Terken and Hirschberg 1994; Morel 1995; Rochemont 1986, 1998; Rochemont and Culicover 1990; Steedman 1991a,b; Zubizarreta 1998).

The present chapter concentrates on certain aspects of intonation that primarily concern Information Structure, in the sense of that term proposed by Valldoví (1990), and Steedman (1991a), although these proposals differ in detail (see Valldoví and Engdahl (1996) for a survey). These theorists follow Halliday (1967b, 1970) in assuming that there are two independent dimensions to

Information Structure that are relevant to intonation. The first corresponds to the distinction, informally introduced at the start of the chapter, between theme and rheme. In English we will see that this dimension of Information Structure determines the overall shape of the intonational tune or tunes imposed upon an utterance. The second dimension is one of salience or contrast. In English this dimension is reflected in the position of pitch accents on particular words. The presence of a pitch accent of any shape is generally agreed to assign salience or contrast independently of the particular shape or contour of the pitch accent or overall phrasal melody (see Pierrehumbert and Hirschberg 1990, 288-289). The next sections consider these two dimensions in turn.

5.3 Theme and Rheme

The λ-abstraction operator is closely related to the existential quantifier $\exists$. It is therefore natural to associate the notion of theme with the set of propositions among all those supported by the conversational context that could possibly satisfy the corresponding existential proposition. In the case of the exchange in (14) it is the following, in which $\diamond$ indicates possibility:

(15) $\exists x.\diamond marry' x\ anna'$

This might be a set like the following:

(16) $\left\{ \begin{array}{l} \diamond marry'\ alan'\ anna' \\ \diamond marry'\ fred'\ anna' \\ \diamond marry'\ manny'\ anna' \\ \quad \ldots \end{array} \right\}$

This extensional interpretation of the notion theme resembles the Alternative Semantics approach to presupposition and focus of Karttunen and Peters (1979), Rooth (1985, 1992), and Kratzer (1991), and the related analysis of German intonational meaning of Büring (1995, 1997). Specifically, the alternative set in question is the one that Rooth and Büring call C, the "contextual" alternative set. Since all alternative sets are contextual, I will refer to it here as the "rheme" alternative set.

Alternative sets are of course in many cases not exhaustively known to hearers, and in practice one would want to compute with something more like the quantified expression (15) or the λ-term itself, as in the structured-meanings approach of Cresswell (1985) and others. However, alternative sets are easy to grasp and are used here for reasons of exposition.

In semantic terms the theme and rheme can therefore be characterized as follows:

(17) a. The Theme *presupposes* the rheme alternative set.
      b. The Rheme *restricts* the rheme alternative set.

The sense in which a theme "presupposes" the rheme alternative set is much the same as that in which a definite expression presupposes the existence of its referent. That is to say, there is a pragmatic presupposition that the relevant alternative set is available in the contextual "mental model" (Johnson-Laird 1983) or database. The presupposition may be "accommodated" in the sense of Lewis (1979)—that is, be added by the hearer after the fact of utterance to a contextual model that is consistent with it.

5.3.1 Update Semantics for Theme and Rheme

One way of making such referents available is to think of the theme of an utterance as *updating* or having side-effects on the context or discourse model.[12] Following Jacobs (1991) and Krifka (1991), it can be characterized as in general causing one or more existing referents or "facts" such as $(\theta'\lambda x.marry'x\ anna')$, where θ marks the λ-term as a theme, to be *retracted* or removed from the context model, and causing a new thematic referent or fact to be *asserted* or added. If the theme is unmarked by any accent, then it will simply be the corresponding thematic referent that is retracted and asserted. Unless a fact of the appropriate form is already present in (or is at least consistent with) the context, the first of these effects will cause the discourse to fail. Otherwise, the thematic referent will be reasserted.

The rheme should also be thought of as updating the context with a similar type of referent, which may become the theme of a subsequent utterance. However, the rheme does not require a preexisting referent or cause any existing thematic referents to be retracted (although we will see that it may have other effects on the database, via the entailments and implicatures discussed above).

The exact form of the retracted and/or asserted informational referents in all of the above examples depends upon the location of focus and pitch accents in the utterance and is determined in a manner discussed in section 5.4.

Noncompositional, procedural notions like assertion and retraction must eventually be declarativized, if we are to be able to prove anything about the expressive power of this theory. However, procedural descriptions can be very transparent, probably because they remain very close to what is actually going on in our heads, and for the moment it will be helpful to think of the problem

in these terms.[13]

The claim that the L+H* LH% tune when present marks the theme in one or the other of these closely related senses is implicit in the accounts of Jackendoff (1972), Ladd (1980), and others, but it remains controversial. Pierrehumbert and Hirschberg (1990, 294–297) propose a compositional semantics for intonational tunes that is based on scalar values on dimensions of propositional attitude such as certainty concerning relevance and degree of commitment to belief revision. According to their account, the L+H* pitch accent is used "to convey that the accented item—and not some alternative related item—should be mutually believed" (p. 296).

As an example, Pierrehumbert and Hirschberg discuss the following dialogue (adapted from Jackendoff (1972, 258–265), and also discussed by Ladd (1980, 157–159), and Steedman (1991a)), which is isomorphic to (13):

(18) Q: What about the beans? Who ate THEM?

    A: Fred ate the BEANS.
      H* L   L+H*LH%

In support of their claim that the L+H* pitch accent evokes a set of alternatives besides the accented item, they correctly observe that the utterance implicates the possibility that other people may have eaten other things. However, this particular alternative set has already been introduced into the context by the question, and in the absence of such a question (or some other utterance establishing a context that supports or is at least consistent with this theme), the intonation contour is inappropriate. The example therefore does not exclude the possibility that the L+H* LH% tune evokes this set of alternatives by marking a part of the theme.

The following minimal pair of dialogues will be helpful in deciding between these claims, because it appears at first glance to raise problems for both.

(19) Q: Does Mary like corduroy?

    A: Mary likes BOMBAZINE.
               H* LL%

(20) Q: Does Mary like corduroy?

    A: Mary likes BOMBAZINE.
        L+H* LH%

In (19), the entire response is marked with the H*LL% tune that we have identified as marking the rheme, constituting what the speaker believes the hearer

needs. Depending on the context, the speaker may thereby be committed by the usual Gricean principles to a number of conversational implicatures. For example, if liking bombazine entails hating corduroy, then this response implicates denial. If on the other hand liking bombazine entails liking corduroy, then the response implicates affirmation. Either way, the speaker's intonation commits them only to the claim that *bombazine* is the rheme—that is, that it restricts the set of alternatives to just one—rather than to a particular change in belief.

More specifically in both cases, the rheme of the yes-no question adds a theme $theme'(like'corduroy'mary')$ to the facts making up the respondent's context. (The alternative set here is confined to the proposition and its negation.) The respondent then constructs the corresponding query and evaluates it with respect to the context. If the query immediately succeeds, or fails altogether, then it is appropriate to respond with a direct yes or no. If the query succeeds but a step of inference involving the respondent's rule that *Everyone who likes bombazine likes corduroy* and the respondent's knowledge that *Mary likes bombazine* is needed to establish the answer, then one of the following cases may apply. If the respondent's discourse model implies that the questioner knows neither the rule nor the truth of the premise, then the respondent should state them both, as in the extended example (23). On the other hand, if the discourse model implies that the respondent knows the rule, but not the premise, then the response should be either as in (19) or as in (20). If there is reason to believe that this is the only relevant difference between the questioner's knowledge and the respondent's own, then stating the premise as a rheme, as in (19), is appropriate, since the respondent can sincerely claim that it is everything the questioner needs. But if the respondent has reason to suspect that there may be other differences and therefore cannot sincerely claim that the questioner can make this inference, then the respondent should mark the premise as a theme, as in (20), and leave the questioner to derive the rheme or not, as the case may be.

As is often the case, the respondent may for reasons of politeness or other pragmatic footwork use an utterance of an isolated theme to conversationally implicate lack of willingness to commit to the adequacy of their information, simultaneously being perfectly certain of the outcome. Nearly all speech acts like the response in (20) have the smell of indirection about them, and we should not expect to capture them in terms of literal meaning alone.

Example (20), which is of a kind discussed by Jackendoff (1972, 264–265), Liberman and Sag (1974), and Ladd (1980, 146–148), appears at first glance

to be almost equivalent. In particular, the possibilities for conversational implicature of either affirmation or denial seem identical. Any difference seems to lie in the degree of conviction that the utterance constitutes an answer to the question.

Since in other respects the two utterances seem similar, there is a temptation to believe that the L+H* LH% tune in this case might mark a rheme, rather than a theme, differing from the standard H*LL% rheme tune in terms of the degree of commitment to whether it does in fact restrict the set of rheme alternatives sufficiently.

However, it is also possible to believe that this utterance is in fact a theme and that what the respondent has actually done is to offer a new set of alternatives, without stating a rheme at all, leaving the other party to supply the rheme. This would certainly explain the lack of commitment to whether the utterance restricts the rheme alternative set, since that is exactly what a theme does not do. It is also likely that the effect of not taking responsibility for a rheme in this utterance will be that of *conversationally implicating* a lack of confidence in either the relevance of the theme or the certainty of the inference that might be drawn. But that would not be a matter of literal or conventional meaning of the utterance or the intonation contour itself.

This is essentially the analysis proposed by Ladd (1980, 153-156), who relates "fall-rise" contours to the function of evoking a set of alternatives established by the preceding context—a notion I have identified with the notion of theme and have interpreted above in terms of the Alternative Semantics of Karttunen and Peters (1979) and Rooth (1985).

In the case of answer (20A) the new theme is simply the following:

(21) *like′ bombazine′ mary′*

Since this is a fully saturated proposition, with no λ-bound variables, the corresponding rheme alternative set is a singleton:

(22) $\{$*like′ bombazine′ mary′*$\}$

Since it contains only one member, it also entails an answer to the question via exactly the same chain of inference from shared beliefs as (19).

Further support for the claim that the L+H* LH% tune marks theme in (20), establishing a new set of alternatives, and that any effect of lack of commitment arises by conversational implicature, can be found in the fact that this intonation remains appropriate when the step of inference that generates the rheme itself is explicitly spelled out, as in the following deliberately exaggerated ex-

ample, in which *liking bombazine* is necessarily distinct from the rheme:

(23) Q: Does Mary like corduroy?
      A: Well, she likes   BOMBAZINE,
                                  L+H* LH%
         And people who like   BOMBAZINE like CORDUROY.
                                  L+H* LH%   H*LL%
      So I am sure that Mary likes CORDUROY.
                                H*LL%

(Note that *likes bombazine* in the first conjunct could equally well be uttered with an H*L% rheme accent, but in the second it really must be marked as a theme. Under most circumstances the first and third conjuncts could be omitted entirely, as being implicated by the second.)

Still more evidence for the claim that L+H* LH% invariably marks a theme can be found in the fact that when a similarly implicative reply states a law from which the conclusion necessarily follows, so that there is no plurality of alternatives, then only the rheme tune is felicitous, as in the following minimal pair:

(24) Q: Will Mary come to the meeting on time?
      A: Mary is ALWAYS on time.
               H*   LL%

(25) Q: Will Mary come to the meeting on time?
      A: #Mary is ALWAYS on time.
              L+H*   LH%

Conversely, if the content of the response necessarily implicates a plurality of alternatives, then the position is reversed: only the theme tune is felicitous:

(26) Q: Will Mary come to the meeting on time?
      A: Mary is USUALLY on time.
              L+H*   LH%

(27) Q: Will Mary come to the meeting on time?
      A: #Mary is USUALLY on time.
              H*   LL%

It is clear that the L+H*LH% tune in the latter example indicates no lack of commitment to the probabilistic claim about Mary's punctuality; rather, it indicates a lack of commitment to the adequacy of this information to fully answer the question.

5.3.2 Unmarked Themes

Extravagantly informative intonation contours like those in (13) and (14) are the exception. It is only appropriate to mark the theme with an L+H* pitch accent when it stands in contrast to a preceding different theme. If the rheme alternative set that a theme presupposes is unambiguously established in the context, it is common to find that the theme is deaccented throughout—in Pierrehumbert's terms, without any pitch accent or obvious boundary, as in the following exchange:

(28) Q: Who did Anna marry?

     A: (Anna married) (MANNY).
                       H* LL%

We would be missing an important semantic generalization if we failed to note that examples (14) and (28) are identical in Information Structure as far as the theme-rheme division goes. We will therefore need to distinguish the "marked" theme in the former from the "unmarked" theme in the latter. Unmarked intonation, unlike the marked variety, is always ambiguous with respect to Information Structure. In the following context the same contour will have the Information Structure of (13):

(29) Q: What do you know about Anna?

     A: (Anna) (married MANNY).
                   H* LL%

In these terms it is clear that the context-establishing questions in (13) and (14) can also be analyzed in terms of a theme and a rheme. In both cases, as Prevost (1995) points out, the *wh*-item constitutes the theme. Usually such themes are prosodically unmarked, but they may also bear the marked L+H*LH% theme tune. In either case the phrase *who*, *what*, or *which man* translates as an abstraction such as the following:

(30) $\lambda p.\lambda x.man'x \wedge px$

Such a theme defines a very unspecific set of alternatives, namely, the set of all contextually supported predicates applying to persons, things, men, or whatever. However, the remainder of the *wh*-question, which must bear the H*LL% rheme tune, restricts this set to one particular predicate. It is this predicate that typically becomes the theme of the answer.

5.3.3 Multiple Themes and Rhemes

It is quite possible for an utterance to mark more than one disjoint segment as either theme or rheme. An example of a multiple or discontinuous theme can be found in the following exchange:

(31) Q: I know which team Mary EXPECTS to LOSE. But which one does she
           WANT to WIN?
        A: (Mary   WANTS)   (IPSWICH) (to   WIN.)
                L+H* LH%   H* L   L+H* LH%

The theme established by the question is *Which one Mary wants to win.* That is:[14]

(32) $\exists x.\Diamond *want'(*win'(ana'x))x\ mary'$

We may assume that the rheme alternative set includes propositions like the following:

(33) $\left\{ \begin{array}{l} \Diamond want'(win'(ana'watford'))watford'mary' \\ \Diamond want'(win'(ana'ipswich'))ipswich'mary' \\ \Diamond want'(win'(ana'sunderland'))sunderland'mary' \end{array} \right\}$

In the answer the words *wants* and *win* get L+H* pitch accents, because the theme alternative set includes the previous theme, *Which one Mary expects to lose*, or $\exists x.\Diamond want'(lose'(ana'x))x\ mary'$. Since elements of the theme are separated by the rheme *Ipswich* (which of course has its own H* pitch accent and boundary), there are two L+H* LH% theme tunes. These fragments work independently to have the effect of a "discontinuous theme." The first presupposes that the rheme alternative set consists entirely of propositions of the form $want'x\ y\ mary'$; the second presupposes that it consists of propositions of the form $p(win'(ana'y))y\ z$. Both presuppositions are compatible with the same rheme alternative set, so together they require that it consists of propositions of the form $want'(win'(ana'x))x\ mary'$, just as if they constituted a single discontinuous theme.

5.4 Focus and Background

The possibility of such unmarked themes, lacking any pitch accent, draws attention to a second independent dimension to discourse Information Structure that affects intonational tune. In (14) the L+H* LH% tune is spread across the entire substring of the sentence corresponding to the theme in the above

sense—that is, over the substring *Anna married*.[15] In (13) the same tune L+H* LH% is confined to the object of the theme *married Manny*, because the intonation of the original question indicates that marrying Manny *as opposed to someone else* is the new topic or theme. In (28) and (29) there is no L+H* LH% tune at all.

The position of the pitch accent in the phrase has to do with a further dimension of Information Structure *within both theme and rheme*, corresponding to a distinction between *the interesting part(s)* of either information unit, and the rest. Halliday (1967b), who was probably the first to identify the orthogonal nature of these two dimensions, called it "new" information, in contrast to "given" information (cf. Brown 1983). The term "new" is not entirely helpful, since (as Halliday was aware) the relevant part of the theme need not be novel to the discourse, as in the examples to hand. Here I will follow the phonological literature and Prevost (1995) in calling the information marked by the pitch accent the "focus," distinguishing theme focus and rheme focus where necessary, and use the term "background" for the part unmarked by pitch accent or boundary. Again there are a number of other taxonomies, most of which are fairly straightforwardly compatible with the present proposal.[16]

The following example serves to illustrate the full range of possibilities for the distribution of focus and background within the theme and the rheme.

(34) Q: I know that Mary envies the man who wrote the musical. But who does she ADMIRE?

A: (Mary   ADMIRES) (the woman who DIRECTED the musical)
             $L+H*LH\%$ $H*$ $LL\%$

Background	*Focus*	*Background*	*Focus*	*Background*

Theme   *Rheme*

Here the theme is something that I will call *Mary admires*, as an informal shorthand referring to the translation of that part of the utterance as the abstract proposition $\lambda x.admires'x\ mary'$. Only the word *admires* is emphasized, because the previous theme was also about Mary. The presence of pitch accents in themes like that in (34) is marked by distinguishing the corresponding constant in the translation *admires'* with an asterisk:

(35) $\exists x.\diamond *admires'x\ mary'$

Unless a compatible prior theme—that is, one that matches (35) when *admires'* is replaced by some other constant, as in (36)—can be retrieved or accommodated, the utterance is simply infelicitous, and the analysis will fail at this point:

(36) $\exists x.\diamond like' x \, mary'$

The set of alternative themes in this case is the following:

(37) $\left\{ \begin{array}{l} \exists x.\diamond admires' x \, mary' \\ \exists x.\diamond like' x \, mary' \end{array} \right\}$

The set of alternative themes is closely related to Büring's (1995) Q, or "question" alternative set. Here I will call it the "theme" alternative set, in contrast to the earlier rheme alternative set.

The rheme alternative set presupposed by the theme is therefore a set of propositions about Mary admiring various people. The rheme is *the woman who directed the musical*, where only the word *directed* is contrasted.

It is important to note that it is all and only the material marked by the pitch accent(s) that is contrasted. This applies when there is more than one pitch accent, as the reader can verify by observing the effect of adding a further pitch accent on the word *musical*. Anything not so marked, including the material between the pitch accent(s) and the boundary, is background. Examples like this suggest that the focusing property of pitch accents applies at the level of words and their interpretations, not at the level of larger constituents, unlike the theme/rheme distinction.

However, there is an asymmetry between the "prenuclear" background material *the woman who* ... that precedes the pitch accent on *directed*, and the background material that succeeds it (*the musical*). The fact that there is no pitch accent on the latter seems to demand that all individuals in the context have the property of having something to do with this particular musical. It would actually be wrong in this context to have a pitch accent. However, the lack of accent on the former does not seem similarly to demand that all the individuals that we are restricting over are women, and in fact in the example they are not. The implication is that in this context the property of directing the musical is sufficient to distinguish the individual uniquely—the fact that this individual is *also* unique by virtue of being a woman need not be marked.[17]

Why does this asymmetry hold? Could it work the other way around? Prevost (1995) has proposed that pitch accents are assigned to words corresponding to properties that successively limit the extension of an NP to the desired referent. If we assume that the order in which these predicates are successively evaluated is "bottom up" with respect to some kind of interpretation or predicate-argument structure, then we predict that if a modifier like the relative clause *who directed the musical* completely specifies the extension, then words corresponding to predicates higher up the predicate-argument structure, such

as the head noun, need not be stressed.[18] Thus, the pragmatic difference between prenuclear and postnuclear background material arises from the relation of word order to predicate-argument structure in English.

5.5 Grammar and Information Structure

What is the relation of such Intonation Structures to syntax and semantics, and how is Information Structure computed? Many of the intonational constituents that we have just been examining—such as the string *Anna married*—do not correspond to traditional syntactic constituents. Jacobs (1991) and Krifka (1991, sec. 4.8) have pointed out the problems that ensue for the semantics of focus.

Such "nonconstituent" intonational units are very widespread and can cooccur with other intonational tunes, including the H*+L rheme tune considered here. Consider the following utterance (adapted from Ladd 1980—see below), uttered in the context of a question like *I know that Harry keeps up with the newspapers, but has he read* War and Peace*?*:

(38) Harry doesn't READ BOOKS!
                 H*L L+H* LH%

Here the theme seems to be *books*, marked because the concept stands in contrast to *newspapers*. The rheme seems to be *Harry not reading something*, with the usual final H*L tune on *read*. The theme can also be unmarked, even in this context, as in Ladd's original example:[19]

(39) Harry doesn't READ books.
                 H*   LL%

The interest of such derivations for present purposes will be obvious. The claim is simply that the nonstandard Surface Structures that are induced by the combinatory grammar to explain coordination in English subsume the Intonation Structures that are postulated by Pierrehumbert and her colleagues to explain the possible intonation contours for sentences of English. The claim is that in spoken utterances, intonation helps to determine which of the many possible bracketings permitted by the combinatory syntax of English is intended, and that the interpretations of the constituents that arise from these derivations, far from being "spurious," are related to distinctions of Information Structure and discourse focus among the topics that the speaker has in mind and the comments that the speaker has to contribute.

The proof of this claim lies in showing that the rules of combinatory grammar can be made sensitive to intonation contour, which limits their application in spoken discourse. We must also show that the major constituents of intonated utterances like (14), under the analyses that are permitted by any given intonation, correspond to the Information Structure required by the context to which the intonation is appropriate. A preliminary proposal, to which the next section proposes a revision and an extension, was sketched in Steedman 1991a,b, and Prevost and Steedman 1994.

5.5.1 Combinatory Prosody

The papers just cited followed Pierrehumbert and colleagues in taking an "autosegmental" approach to the tones, in the sense of the term introduced by Goldsmith (1976), assigning a distinct prosodic CG category to all elements of the sentence, as well as a grammatical one. Like grammatical categories, prosodic categories could be either functions or arguments and could apply to or compose with one another. Syntactic combination was made subject to intonation contour by the assumption of a "Prosodic Constituent Condition", which only permitted combination of syntactic categories via a syntactic combinatory rule if their result was also a prosodic constituent.

The present version of the theory takes a different approach, integrating prosodic information with the standard grammatical categories to more directly capture Intonation Structure, together with its interpretation as Information Structure, in CCG.

We have already noted that the focus-marking property of pitch accents seems to belong at the level of the word, whereas the theme/rheme-marking property seems to belong at the level of phrasal constituents. We therefore begin by assuming that pitch accents both mark (some element of) the interpretation of the words they occur on for focus or contrast, and mark the syntactic category in a way that "projects" theme-rheme status to elements with which the word combines. Although eventually we will certainly want to do this by morphological rule, for present purposes we will regard this compiled out into distinct lexical entries like the following categories for the verb *ate* bearing the two pitch accents under discussion here. before syntax gets to work on them:[20]

(40) ate   := $(S_\theta \backslash NP_\theta)/NP_\theta : *ate'$
      L+H*

(41) ate := $(S_\rho \backslash NP_\rho)/NP_\rho : *ate'$
      H*

The subscript symbols θ and ρ are mnemonic for theme and rheme respectively, and are a shorthand for a value on a feature of the whole category that I will call *INFORMATION*. A category like (40) ensures that any argument that combines with it must be phonologically compatible with being part of a theme.

The "null tone," which I will follow Pierrehumbert in leaving without any annotation in strings, does not affect the interpretation of a word that carries it, and leaves the syntactic category unspecified as to the value of the feature *INFORMATION*. It can therefore conveniently be written without any annotation, as before:

(42) ate := $(S \backslash NP)/NP : ate'$

Since the value of *INFORMATION* is unspecified, this category can combine with either θ, ρ, or unmarked categories. However, it is important to remember that the unspecified values on arguments and result are the *same* unspecified value of the same attribute *INFORMATION*. In the first two cases, this *INFORMATION* value becomes specified for all arguments and the result, by the usual unification mechanism, and subsequent combinations must be compatible with that value.[21]

Prosodically annotated categories of this kind allow the influence of the pitch accent to spread over arbitrarily large constituents. For example, in analyzing the first two words of the sentence *Fred ate the beans*, uttered in response to a question like *I know what Harry ate. But what did FRED eat?*, the following partial derivation can occur:

(43)

$$\begin{array}{ccc}
\text{FRED} & \text{ate} & \text{the BEANS} \\
\text{L+H*} & & \text{LH\%} \quad \text{H*} \quad \text{LL\%}
\end{array}$$

$$\frac{\overline{S_\theta/(S_\theta \backslash NP_\theta) : \lambda p.p * fred'} \quad \overline{(S \backslash NP)/NP : \lambda x.\lambda y.ate' xy}}{S_\theta/NP_\theta : \lambda x.ate' x * fred'} >\mathbf{B}$$

The L+H* pitch accent on *FRED* marks all elements of the raised subject category as θ on the *INFORMATION* feature. The verb bears the null tone, but when the subject composes, all occurrences of the verb's own *INFORMATION* feature come to bear the value θ by the unification mechanism. Hence the object in the category that results from composition also bears the *INFORMATION* value θ.

In contrast to the version in Steedman 1991a, the present theory assumes that boundary tones, unlike pitch accents, are not associated with words, but are elements of the string in their own right, much like the punctuation marks

that, on occasion, represent them in the orthography. Like the pitch accents, the boundary tones affect both the syntactic and the semantic components of categories. The grammatical category of a boundary is that of a function from categories marked as θ, ρ, or unspecified, into phonological phrasal categories, distinguished by a value ϕ.[22]

The boundary tones must also mark the informational units at the level of the interpretation, so that the combination of a constituent bearing a boundary tone with another including a pitch accent semantically defines the major informational elements such as the theme and the rheme.

For present purposes, the full categories for the three boundary tones under discussion here are written in full using the following notation.[23]

(44) L, LL%, LH% := $S\$_\phi \backslash S\$_\eta : \lambda f.\eta' f$

The variable $S\$$ ranges as usual over a set $\{S\$\}$ of categories including S and all functions into members of $\{S\$\}$—that is, it includes S, S/NP, and all verbs and type-raised arguments of verbs, but not nouns and the like. The subscript η, which can be thought of as a variable ranging over the two *INFORMATION* values θ and ρ, further specifies it as ranging over correspondingly marked categories S_θ, S_θ/NP_θ, $(S_\rho \backslash NP_\rho)/NP_\rho)$, etc. When it combines with such a function, it has the effect of replacing its θ or ρ marking with a distinct marker ϕ (for "phrasal"), which can only unify with itself. Such a category can only combine with other ϕ-marked prosodically phrasal categories.

Semantically, the boundary categories apply a corresponding thematic or rhematic function θ' or ρ' to the interpretation of the category with which they combine, via a corresponding variable category η'. In terms of the Logical Form, θ' or ρ' are identity functions that effectively vanish from the predicate-argument structure when they apply. However, they are assumed to cause the appropriate alternative set to be evoked from the database, and to be accompanied by the updates discussed in section 5.3.1 Until they do apply, they block any further reduction of the interpretation to the canonical predicate-argument structure.

This specification of boundaries allows them to combine with either pitch accent, consistent with Pierrehumbert's own system, in which all pitch accents and boundaries can combine freely. For example, low boundaries can combine with L+H* pitch accents, as in the following sentence, from Ladd 1996, 96-7:

(45) THAT's the whole POINT of the exercise!
      H* L   L+H*   LL%

The implicit claim that the boundaries also project θ or ρ marking from the pitch accent, so that the phrase *the whole point of the exercise* is a theme in the above example (as it would be if it bore an LH% boundary), is more controversial, and will not be discussed here. Nevertheless, we will see later that at least one variety of theme, the unmarked variety, does bear low boundaries.

To say that the boundary projects the category of the pitch accent is not to exclude a more active role in the semantics for the boundary, analogous to a specifier such as a determiner, contributing distinctively to information structural content along the lines suggested by Pierrehumbert and Hirschberg (1990).

Since they bear exactly the same category as L boundaries, LL% boundaries are free in the present system to occur utterance-medially, in contrast to earlier versions of the theory which were criticized on this point by Beckman (1996, 63-64). Utterance medial LL% boundaries do not figure much in the examples discussed in the present chapter (although see (75) below and the discussion of example (60) in chapter 7). In particular it does not appear to be possible to substitute them freely for L intermediate phrase boundaries in examples like (13) (a fact upon which Beckman does not comment). By the same token, the present system allows L boundaries to occur utterance-finally, which is impossible. We will assume for present purposes that these details are to do with finer distinctions between the boundaries, and in particular with the distinction between intermediate and intonational phrases. The question of whether LL% boundaries are or are not categorically distinct from intermediate phrase L boundaries is a matter of some dispute among phonologists, and we will continue to pass over it here.

The following example, which completes the derivation of the theme of the earlier sentence *FRED ate the BEANS*, demonstrates the effect of the boundary tone:

(46)

$$
\begin{array}{llll}
\text{FRED} & \text{ate} & & \text{the BEANS} \\
\text{L}+\text{H}* & & \text{LH\%} & \quad\ \text{H}* \quad \text{LL\%} \\
\hline
S_\theta/(S_\theta\backslash NP_\theta) & (S\backslash NP)/NP & S\$_\phi\backslash S\$_\eta \\
: \lambda p.p \ *fred' & : \lambda x.\lambda y.ate'xy & : \lambda f.\theta'f
\end{array}
$$

$$
\cfrac{\cfrac{\quad\qquad\qquad\qquad}{S_\theta/NP_\theta : \lambda x.ate'x *fred'}\ ^{>\mathbf{B}}}{S_\phi/NP_\phi : \theta'(\lambda x.ate'x *fred')}\ ^{<}
$$

The second prosodic phrase in (46) bears the H*LL% rheme tune, parallel to (14) above. The complete derivation is as follows:

(47)   FRED   ate   the   BEANS
           L+H*   LH%   H*   LL%
           ─────────   ───────────   ──────────   ─────────   ─────   ──────────
           $S_\theta/(S_\theta\backslash NP_\theta)$ $(S\backslash NP)/NP$ $S\$_\phi\backslash S\$_\eta$ $NP^\uparrow/N$ N_ρ $S\$_\phi\backslash S\$_\eta$
           $:\lambda p.p *fred'$ $:\lambda x.\lambda y.ate'xy$ $:\lambda f.\theta'f$ $:\lambda x.\lambda p.p(the'x)$ $:*beans$ $:\lambda f.\rho'f$
           ─────────────────────>**B**   ──────────────────────>
           $S_\theta/NP_\theta : \lambda x.ate'x *fred'$ $NP^\uparrow_\rho : \lambda p.p(the'*beans')$
           ──────────────────────<   ─────────────────────────<
           $S_\phi/NP_\phi : \theta'(\lambda x.ate'x *fred')$ $NP^\uparrow_\phi : \rho'(\lambda p.p(the'*beans'))$
                                                                                   ──────────<
                      $S_\phi : \rho'\lambda p.p(the'*beans')(\theta'\lambda x.ate'x *fred')$
           ───
                      $S_\phi : ate'(the'*beans')*fred'$

No *other* division into a theme and a rheme is possible for this intonation contour.[24]

It is only once the functions θ' and ρ' have applied that the final semantic reduction or normalization can take place, to yield the canonical predicate-argument structure. As far as the interpretation goes, these are just the identity function, but we have assumed that they are accompanied by side effects of assertion or retraction on the database. The reduction of θ' and ρ' can occur at any point in a derivation.

The division of the utterance into a property constituting the theme and an argument constituting the rheme is appropriate to a context parallel to that established in (14)—say, by the *wh*-question *What did FRED eat?* uttered in the context of a discussion of what somebody else (say, Harry) ate and a prior theme such as the following:

(48) $\theta'\lambda x.ate'x harry'$

The theme in derivation (47), $\theta'\lambda x.ate'x *fred'$, is a member of the theme alternative set in the earlier sense.[25] The new theme presupposes a rheme alternative set of propositions about Fred eating things, which the rheme $\rho'\lambda p.p(the'*beans')$ reduces to a single proposition, $ate'(the'beans')fred'$.

Since categories bearing the null tone can compose either with others bearing null tone, or with those bearing pitch accents, intonational phrasal tunes like L+H* LH% can spread unboundedly across any sequence that forms a grammatical constituent according to the combinatory grammar. For example, if the answer to *What did FRED eat?* is *MARY says he ate BEANS*, then the tune will typically be spread over *Mary says he ate ...* as in the (incomplete) derivation (49), in which the semantics has been suppressed in the interests of brevity:

(49) ˙MARY   says   he   ate   . . .
      L+H∗   LH%

$$\underline{S_\theta/(S_\theta\backslash NP_\theta)} \quad \underline{(S\backslash NP)/S} \quad \underline{S/(S\backslash NP)} \quad \underline{(S\backslash NP)/NP} \quad \underline{S\$_\phi\backslash S\$_\eta}$$

$$\underline{}\text{>}\mathbf{B}$$
$$S_\theta/S_\theta$$
$$\underline{}\text{>}\mathbf{B}$$
$$S_\theta/(S_\theta\backslash NP_\theta)$$
$$\underline{}\text{>}\mathbf{B}$$
$$S_\theta/NP_\theta$$
$$\underline{}\text{<}$$
$$S_\phi/NP_\phi$$

A number of derivations are permitted for more complex rheme NPs like *the green beans* with a final pitch accent, represented by the following bracketings:

(50) a. (the green) (BEANS)
                       H*   LL%

      b. (the) (green BEANS)
                  H*   LL%

The same derivational ambiguity is characteristic of a further intonation contour, in which the H* pitch accent is applied to the word *green*, making that word alone contrasted:

(51) Q: Did Fred eat the green beans or the yellow beans?

      A: (FRED   ate) (the GREEN beans.)
         L+H* LH%   H*   LL%

It also applies when both words are marked for rheme accent with H* pitch accents, in a context in which both elements of the noun group *green beans* are contrasted or informative:

(52) Q: Did Fred eat the green beans or the yellow squash?

      A. (FRED   ate) (the GREEN BEANS.)
         L+H* LH%   H*   !H* LL%

Example (52) needs further comment. First, the rheme includes more than one H* pitch accent. These pitch accents belonging to the same intonational/intermediate phrase would be seen in a pitch track to be downstepped—that is, to have successively lower peaks and baselines for approximately the same pitch range. The ToBI conventions (Silverman et al. 1992) again offer a convenient notation for downstep, extending Pierrehumbert's system with the prefix "!". However, as in Pierrehumbert's original theory, words bearing downstepped pitch accents bear identical categories to non-downstepped accents. The !-notation is redundant, and the effect on pitch-contour can be determined automatically.

Downstepped pitch accents can therefore compose, to allow derivations like the following:

(53)

$$
\begin{array}{l}
\text{FRED ate} \qquad \text{the} \quad \text{GREEN} \quad \text{BEANS} \\
\text{L+H*} \qquad \text{LH\%} \qquad \qquad \text{H*} \qquad \text{!H*} \qquad \text{LL\%}
\end{array}
$$

$$
\cfrac{
\cfrac{
\cfrac{
\cfrac{
\cfrac{
\cfrac{
\begin{array}{c} S_\phi/NP_\phi \\ : \theta'(\lambda x.ate'x \ast fred') \end{array} \quad
\cfrac{
\begin{array}{c} NP/N \\ : the' \end{array} \quad
\begin{array}{c} N_\rho/N_\rho \\ : \ast green' \end{array}
}{
\begin{array}{c} NP_\rho/N_\rho \\ : \lambda x.the'(\ast green'x) \end{array}
}{\scriptstyle >\mathbf{B}}
\quad
\begin{array}{c} N_\rho \\ : \ast beans' \end{array}
}{
NP_\rho : (the'(\ast green' \ast beans'))
}{\scriptstyle <}
}{
\begin{array}{c} S_\rho\backslash(S_\rho/NP_\rho) \\ : \lambda f.f(the'(\ast green' \ast beans')) \end{array}
}{\scriptstyle <\mathbf{T}}
\quad \begin{array}{c} S\$_\phi\backslash S\$_\eta \\ : \lambda f.\rho'f \end{array}
}{
\begin{array}{c} S_\phi\backslash(S_\phi/NP_\phi) \\ : \rho'(\lambda f.f(the'(\ast green' \ast beans'))) \end{array}
}{\scriptstyle <}
}{
S_\phi : \rho'\lambda f.f(the'(\ast green' \ast beans'))(\theta'\lambda x.ate'x \ast fred')
}{\scriptstyle <}
}{
S_\phi : ate'(the'(\ast green' \ast beans'))\ast fred'
}
$$

This time, there is another derivation for the rheme in (53). In fact, both derivations illustrated in (50) are permitted for all three stress patterns. The sentences differ only in the elements of the interpretation that are marked for contrast. These observations reinforce the earlier suggestion that the effect of the pitch accents applies at the level of words words and their interpretations, rather than at higher levels of derivation, unlike the effect of boundary tones.

Many impossible intonation contours correctly remain excluded by the fact that CCG conflates prosodic structure and syntactic structure, including examples of the kind that motivated Selkirk's Sense Unit Condition. For example, the following are disallowed because their Intonation Structure is not compatible with any syntactic analysis, owing to island constraints:

(54) a. *(FRED ate the green) (BEANS.)
          L+H*   LH% H* LL%

      b. *(My OLDER) (sister ate the green BEANS.)
          L+H* LH%   H* LL%

5.5.2 Unmarked Themes

So far I have only considered sentences that include a theme and rheme that both include words marked for contrast by pitch accents. Such sentences are relatively unambiguous with regard to their Information Structure. However, sentences like the following, which in Pierrehumbert's terms consist of a single intonational phrase, are much more common:

(55) (Mary wrote a book about BATS.)
         H* LL%

Such sentences are notoriously ambiguous with respect to the theme they presuppose (cf. Selkirk 1995). For example, (55) seems equally appropriate as a response to any of the following questions:

(56) a. What did Mary write a book about?
     b. What did Mary write?
     c. What did Mary do?
     d. What's new?

Such questions could in more contrastive contexts give rise to themes marked by the L+H* LH% tune, bracketing the sentence as follows:

(57) a. (Mary wrote a book about)*Theme*(BATS.)*Rheme*
     b. (Mary wrote)*Theme*(a book about BATS.)*Rheme*
     c. (Mary)*Theme*(wrote a book about BATS.)*Rheme*
     d. (Mary wrote a book about BATS.)*Rheme*

It is therefore a virtue in the grammar as it stands that it already allows all of the implicit derivations for the sentence in which the theme is unmarked, while the various possible rhemes are marked as such in the derivation.

Such unmarked themes can be made explicit in the theory as follows. The boundary categories (44) are already defined so as to allow them to combine with unmarked categories, on the assumption that an unspecified *INFORMATION* value can unify with or match the variable η in the boundary category.

If we further assume that an L boundary is phonetically indistinguishable from the null tone, then such a boundary tone may be postulated anywhere there is null tone (and low pitch). Such a tactic nondeterministically allows all of the derivations indicated in (57). For example:

(58)

Mary wrote a book about		BATS	
	L	H*	LL%

$$\frac{S/NP}{:\lambda x.write'(a'(book'(about'x)))mary'} \quad \frac{S\$_\phi \backslash S\$_\eta}{:\lambda f.\eta'f} \quad \frac{S_\phi \backslash (S_\phi/NP_\phi)}{:\rho'(\lambda p.p * bats')}$$

$$\frac{S_\phi/NP_\phi}{:\eta'(\lambda x.write'(a'(book'(about'x)))mary')} <$$

$$\frac{S_\phi : \rho'(\lambda p.p * bats')(\eta'(\lambda x.write'(a'(book'(about'x)))mary'))}{} <$$

$$S_\phi : write'(a'(book'(about' * bats')))mary'$$

On the reasonable assumption that an unspecified η' has the same effect as a value of θ' in the interpretation, apart from necessarily applying to an argument lacking any $*$ marker of contrast, then the representation of theme and rheme in the interpretation is exactly as in the earlier examples.

As Steedman 1991a points out, this nondeterminism can be eliminated for processing purposes by taking advantage of the fact that the unmarked theme is exclusively used when the hearer is assumed to already know the theme. Thus, the appropriateness of applying the rule to a given category can be directly decided by referring to the discourse model to see whether it supports the presupposition that the corresponding referent is theme, background, or whatever. (See Straub 1997 for experimental evidence for the systematic omission of explicit prosodic boundaries by speakers when alternative sources of disambiguating information, including contextual, are present.)

The ambiguity of such contours with respect to Information Structure appears to be correctly constrained by the assumption that Information Structure and syntactic structure must coincide. That is, the following do not appear to be possible Information Structures, because, like the related examples in (54), they are not possible syntactic structures.[26]

(59) a. $*$(Fred ate the green)$_{Theme}$(BEANS)$_{Rheme}$

 b. $*$(My older)$_{Theme}$(sister ate the green BEANS)$_{Rheme}$

The trick of nondeterministically assuming an invisible boundary to null themes may seem unnecessarily clumsy, and to compromise the theory by introducing phonological entities that have no phonetic realization, as Croft (1995) and Ladd (1996) have suggested. I will return to this point below. However, the device captures an important generalization concerning the relation of these unmarked themes to the corresponding marked ones, and to another variety of unmarked theme which *does* have an explicit boundary.

In English (as opposed to many other languages—see Ladd 1996; Zubizarreta 1998) unmarked themes can occur utterance-finally, and when they do, they end with an LL% boundary, as in the following example:[27]

(60) Q: Who ate the beans?

 A: (FRED) (ate the beans).
 H$*$ L   LL%

If the rheme *FRED* is to be a well-formed intonational phrase distinct from the unmarked theme, it must end in an L intermediate phrase boundary. Again postulating such a boundary introduces a nondeterminism—but again this nonde-

terminism arises only in contexts where the theme in question is entirely given, or background, and hence is recoverable by the hearer. (When evaluated, such themes must by definition yield a set of alternative propositions identical to the background set. I will return to this point.)

However, an intonational-phrase-final LL% tone cannot always be analyzed in this way. It may just be background information in the rheme. To take an old example from Schmerling (1976), one may announce the death of Nixon in the absence of any prior discourse by saying the following:

(61) NIXON died.
     H*   LL%

The second word is then part of the rheme, which of course is allowed by the grammar, and the utterance is felicitous just in case dying is a background possibility for the individual in question. (If not, as Schmerling points out, one has to say something like *Nixon DIED*.)

The other, apparently phonetically indistinguishable analysis for this sentence, with an unmarked theme, is of course still available and is appropriate to a situation where the question is *Who died?*, as Ladd (1980, 53) points out. It is this analysis that is at work in Ladd's own example (39), repeated here in the revised notation, uttered in the context of a question *Has Harry read* War and Peace*?*:

(62) Harry doesn't READ books.
             H* L LL%

Although the analysis proposed here is quite different from Ladd's, he supports a view of the Information Structure according to which the utterance is "about a book" (Ladd 1980, 52)—in present terms, where *books* is the theme.

5.6 Intonation and the Simplex Clause

Jackendoff (1972) exhaustively examines the effect of all possible assignments of the two tunes considered here to a simple transitive clause, *Fred ate the beans*, and it is instructive to do the same within the present framework.

In contrast to the intonation in the derivation (47), the intonation contour on (18) prevents the composition of subject and verb, because under the forward prosodic composition rule the subject is not allowed to combine with the verb. It follows that a derivation parallel to the earlier one (and the formation of the corresponding theme) is not allowed. On the other hand, the following

derivation *is* allowed for (18):

(63)

| Fred | | ate | the | beans | |
| H* | L | | | L+H* | LH% |

$$S_\rho/(S_\rho\backslash NP_\rho) \quad S\$_\phi\backslash S\$_\eta \quad (S\backslash NP)/NP \quad NP^\uparrow/N \quad N_\theta \quad S\$_\phi\backslash S\$_\eta$$
$$: \lambda p.p *fred' \quad : \lambda f : \rho'f \quad : \lambda x.\lambda y.ate'xy \quad : \lambda x.(the'x) \quad : *beans' \quad : \lambda f : \theta'f$$

$$\underline{\qquad\qquad}_<$$

$$S_\phi/(S_\phi\backslash NP_\phi) : \rho'\lambda p.p *fred' \qquad\qquad NP_\theta^\uparrow : the'*beans'$$

$$\underline{\qquad\qquad}_>$$

$$S_\phi : \rho'\lambda p.p *fred'(\theta'\lambda y : ate'(the'*beans')y)$$

$$S_\phi : ate'(the'*beans')*fred$$

Again no other analysis is allowed, and again the division into rheme and theme, and the associated interpretations, are consistent with the context given in (18).

The effect of the above derivation is to annotate the entire predicate as theme, just as if the tune L+H* LH% had been spread across the whole constituent.

Other cases considered by Jackendoff are accepted under the same assumptions and in every case yield unique and contextually appropriate interpretations, as follows. (The derivations themselves are suggested as an exercise.) The first two yield derivations parallel to (63), in that the fundamental division of the sentence is into a traditional subject and predicate (again these are the only analyses that the rules permit):

(64)   Q: What about FRED? What did HE do to the beans?

   A: (FRED)   (ATE the beans.)
      L+H* LH%   H*   LL%
       Theme   *Rheme*

(65)   Q: I know who COOKED the beans. But then, who ATE them?

   A: (FRED) (ATE the beans.)
       H* L   L+H*   LH%
      Rheme   *Theme*

The other two cases considered by Jackendoff yield derivations parallel to (47), in which the fundamental division of the sentence is orthogonal to the traditional subject-predicate structure:

(66)   Q: I know what Fred COOKED. But then, what did he EAT?

   A: (Fred   ATE) (the BEANS.)
       L+H* LH%   H* LL%
       Theme   *Rheme*

(67) Q: Well, what about the BEANS? What did Fred do with THEM?

    A: (Fred ATE) (the   BEANS.)
           H* L   L+H* LH%
        Rheme   *Theme*

In the case of (66) at least, it seems obvious that the theme established by the context is indeed the one corresponding to the bracketing. In the case of (67) it is less obvious. However, the treatment of relative clauses below will show that this analysis must at least be available.

The following further derivation for (67) is also allowed, as is a parallel derivation for (66), in which *Fred* is a background component of a discontinuous theme, rather than a background part of the rheme:

(68) Q: Well, what about the BEANS? What did Fred do with THEM?

    A: (Fred) (ATE) (the   BEANS.)
         L   H* L   L+H* LH%
    Theme Rheme   *Theme*

There seems to be little to distinguish the alternatives on pragmatic or phonetic grounds. It is the context that determines which Information Structure is felicitous.

Two further cases, which are parallel to (63) and (47) except that the theme and rheme tunes are exchanged, are also accepted, again yielding unique, contextually appropriate analyses. The first is the following:

(69) Q: I know that ALICE read a BOOK. But what about FRED? What did HE do?

    A: (FRED)   (ate the BEANS.)
      L+H* LH%   H* LL%
      Theme   *Rheme*

The contour on the response here is also a coherent response in the context used in (43). This possibility (which may be the one intended in Jackendoff's 1972 discussion of the example) appears to arise from an ambiguity in the context itself. However, the converse does not apply: the intonation on the response in (64) is not felicitous in the above context, as the following example shows:[28]

(70) Q: I know that ALICE read a BOOK. But what about FRED? What did HE do?

    A: #(FRED ate) (the BEANS.)
      L+H* LH%   H* LL%
      Theme   *Rheme*

There is one final possibility, which Jackendoff does not distinguish from (63). It is intuitively less obvious than the others, because its discourse meaning is better expressed (at least in the written language) by a left dislocation, *As for the BEANS, FRED ate them*, or even a passive *The BEANS were eaten by FRED*, uttered with the same assignment of pitch accents to *the beans* and *Fred*. Again the use of a second pitch accent on the verb *ate* in the rheme would also improve the example. Its place in the scheme of things will become clearer in section 5.7.2.

(71) Q: Well, what about the BEANS? What happened to THEM?
      A: (FRED ate) (the BEANS.)
          H*   L   L+H* LH%
        Rheme   *Theme*

5.7 Intonation in Complex Constructions

The number of possible intonation contours for complex sentences is naturally even larger than those that have just been demonstrated for simple transitive sentences, and the contextual conditions that are required to make them felicitous are even more abstruse. The following sections are necessarily restricted to showing that the theory makes correct predictions concerning the complex constructions in which forward composition is necessarily implicated in syntax (in particular, reduced coordinate sentences and relative clauses), rather than merely allowing alternative derivations.

5.7.1 Coordinate Sentences

Since the coordinate sentence (72a) below necessarily involves composition of the (type-raised) subject with the verb, whereas (72b) necessarily does not, it is predicted that the intonation contours that they permit will be more restricted than for the non coordinate sentence (72c):

(72) a. Bill cooked, and Fred ate, the beans.
     b. Fred ate the beans, and drank the wine.
     c. Fred ate the beans.

For example, among other alternatives, we would expect the intonation contour (73) to be possible for (72a). (The example assumes the mechanism for multiple pitch accents of section 3.3. It is a possible answer to the question *What did Bill and Fred do with the beans?*)

(73) (Bill COOKED and Fred ATE) (the BEANS.)
                 H*   !H*L   L+H* LH%

By contrast, intonational tunes that assign categories that are not consistent with the crucial syntactic compositions block derivation:

(74) a. *(Bill cooked and   FRED)   (ate the BEANS.)
                              H*L   L+H* LH%
       b. *(Bill cooked and   FRED)   (ate the BEANS.)
                         L+H*LH%   H* LL%

This effect is sufficiently strong for garden paths to be forced under the same principle, as we saw earlier:

(75) *(Harry likes the NUTS) (and bolts APPROACH.)
                     H*LL%   H* LL%

5.7.2 Relative Clauses

Since relative clauses, like the coordinate structures of section 4.1, also force the involvement of functional composition, a similar conspiracy with intonation is predicted for them as well. And indeed, all the possible intonational tunes that appeared in Jackendoff's (1972) examples on the fragment *Fred ate*—that is, all those that allow syntactic composition under the Prosodic Constituent Condition—can also appear on the same fragment when it occurs as the residue in a relative clause:

(76)   the beans that Fred   ate
      a.   L+H*   LH%
      b.   L+H*LH%
      c.   H*   LL%
      d.   H*LL%

(The null tone is of course also allowed on the relative clause.) Each alternative conveys different presuppositions concerning the context. Since the cleft construction is often used with the *wh*-clause marked with the theme tune, L+H* LH%, (77) and (78) show one way of making the first two alternatives—(76a) and (76b) respectively—felicitous:

(77) Q: FRED didn't eat the POTATOES. HARRY ate THEM.
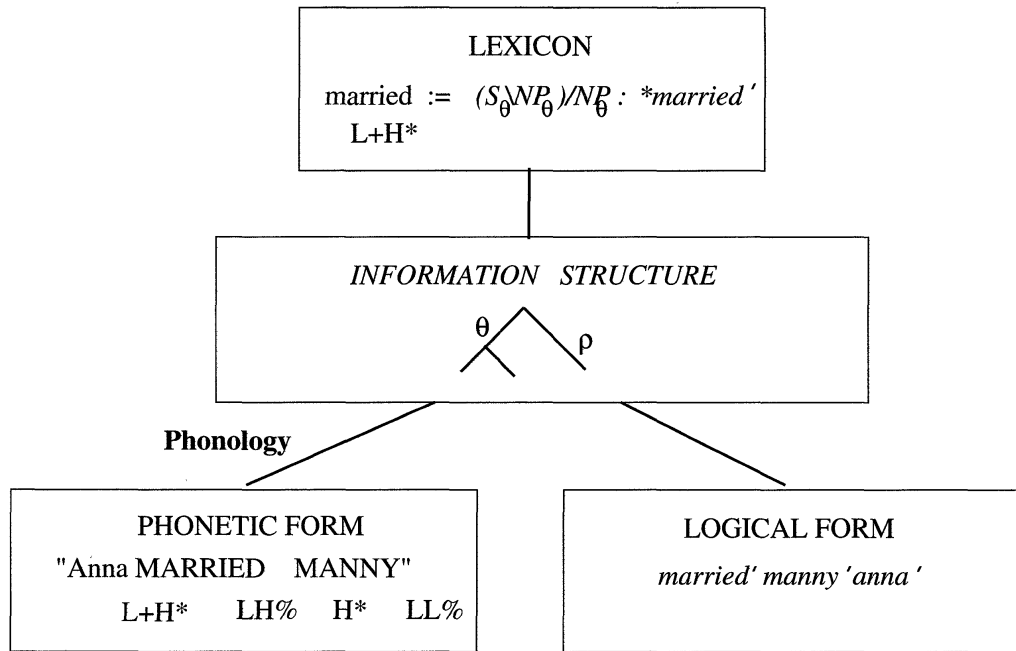   A: (It was the BEANS) (that FRED ate.)
                       H* L   L+H* LH%

(78) Q: Fred didn't eat the POTATOES. He threw THEM AWAY.
    A: (It was the BEANS) (that Fred   ATE.)
                 H* L   L+H*LH%

The H* LL%tune, which marks the rheme, is frequently used on restrictive rel-
atives, so (79) and (80) may serve to make the remaining two cases—(76c) and
(76d) respectively—felicitous. (I have assumed an analysis with an unmarked
theme, but this detail is not crucial.)

(79) Q: It wasn't the beans that HARRY ate that looked so delicious.
    A: (It was) (the beans that FRED ate.)
                        H* LL%

(80) Q: It wasn't the beans that Fred COOKED that looked so delicious.
    A: (It was) (the beans that Fred ATE.)
                     H*LL%

   The converse also holds. As in the case of coordination, tone sequences that
are consistent with no CCG derivation are forbidden from appearing on the
relative clause. Thus, we predict that (81a,b) are intonationally disallowed, for
the same reason that (81c) is ruled out:

(81) a. *(The beans that   FRED)   (ate were DELICIOUS.)
                          H*L   L+H* LH%
      b. *(The beans that   FRED)   (ate were DELICIOUS.)
                       L+H*LH%   H* LL%
      c. *The beans that I, and squash that you, cooked were delicious

Thus a condition akin to Selkirk's Sense Unit Condition emerges as a theorem
of the assumptions inherent in CCG, without independent stipulation.

5.8 Conclusion

At this point it should be clear that we can simply subsume both Intonation
Structure and Surface Structure under a single notion of Information Structure.
Such a view of Intonation Structure involves a richer structural representation
than the one invoked under that name by Pierrehumbert and others, since In-
formation Structure includes many constituent boundaries that do not have any
phonetic realization. This should not seem a strange conclusion to reach. It
simply means that information-structural boundaries are no more (and no less)
completely specified by tones than syntactic boundaries are specified by words.

This observation means that Ladd's (1996, 224) criticism of earlier version of this account as being "compromised by its dependence on entities whose presence cannot be independently [phonetically] verified" and Croft's (1995, 856) related suggestion that it "leaves unexplained the mismatch between prosody and information structure" both rather miss the point. One might as well criticize standard syntactic theories for using lexically unattested brackets. There *is* no "mismatch" between interpretable structure and its surface markers such as function words and tones. These markers are simply rather ambiguous and highly elliptical, just like everything else in natural language.

If Information Structure boundaries and surface syntactic boundaries coincide in this way, then there are a number of other prosodic effects that should depend upon the Surface Structures afforded by CCG in as direct a manner as English intonation contour. Some obvious candidates are such vowel-harmonic effects as French *liaison* (Selkirk 1972), American English flapping (Nespor and Vogel 1986), the Rhythm Rule (Gussenhoven 1983; Selkirk 1984), Bengali $/r/$ assimilation (Hayes and Lahiri 1991), and Italian *raddoppiamento syntattico* (Napoli and Nespor 1979; Kaisse 1985; Nespor and Vogel 1986; Vogel 1994). The last authors in particular show information-structural effects of focus that seem likely to be capturable in this way. These phenomena would then be brought under the category of "superficial" constraints of syntax on phonology called for by Pullum and Zwicky (1988), since Surface Structure now completely specifies the input to phonology and relation to the metrical grid (Liberman and Prince 1977; Prince 1983; Hayes 1995). In fact, in the sense in which Pullum and Zwicky intend the term, the present theory of syntax is "phonology-free," although in another sense CCG syntax actually subsumes prosodic structure.

More speculatively, it seems likely that many of the "end-based" effects of syntax upon phonology argued for by Selkirk (1986, 1990), Selkirk and Shen (1990), Hirst (1993), and Truckenbrodt (1995, 1999), according to which intonation-structural boundaries coincide with either left or right edges of syntactic constituents, but not necessarily both, are artifacts of the syntactic theories within which they are being framed. That is to say, English appears to be a left-edge language because a traditional right-branching account of its Surface Structure just doesn't offer phonologists enough right brackets to work with. The present theory simply claims that those right brackets are there in the syntax, in left-branching structures like (5). Under this interpretation of Surface Structure it is unnecessary to postulate an additional independent prosodic structure, as does Selkirk, and as do Nespor and Vogel (cf. Vogel and Kenesei

LEXICON

married := $(S_\theta \backslash NP_\theta)/NP_\theta$: *married'
L+H*

INFORMATION STRUCTURE

θ ρ

Phonology

PHONETIC FORM

"Anna MARRIED   MANNY"

L+H*   LH%   H*   LL%

LOGICAL FORM

married' manny 'anna '

Figure 5.3
Architecture of a combinatory grammar, II

1990; Zec and Inkelas 1990). We should instead adopt a more liberal notion of syntactic structure, one that is directly compatible with the boundaries that the phonologists observe.[29]

However, the status of Surface Structure in such a theory is very different from the status of the related concepts in GB and earlier theories such as the "annotated surface structures" of Chomsky (1971) and Jackendoff (1972). To understand this point, it will again be helpful to consider the architecture of the present theory of grammar in terms of the traditional "T" or "Y" diagram, as in figure 5.3, which includes an example of an object of the kind represented in each structural module for the following sentence:

(82) Anna MARRIED   MANNY.
          L+H*   LH%   H*   LL%

According to the present theory, the lexicon associates category-interpretation pairs with (the phonological specification of) each word of the language. Derived objects or constituents also pair (the phonological specification of) strings with category-interpretation pairs, which are projected in parallel from (ordered multisets of) lexical entries, via derivations using combinatory rules. In the case of both lexical items and derived ob-

jects, the category is, strictly speaking, redundant, since under the Principle of Categorial Type Transparency, it is presumed to be entirely predictable from (a) the type of the interpretation, (b) X-bar theory, and (c) a parametric description of the language specifying position of heads relative to complements. In effect, the CG category "compiles out" this information, in the sense that it represents explicitly information that could be derived. Therefore, the category-interpretation pairs really count as a single level of representation.[30]

Surface Structure does not figure at all as a level of representation in the theory. Although I have described the combinatory derivations that map phonological strings onto such category-interpretation pairs (and vice versa) in terms of structures, I have never predicated any rule or relation over such structures. They are merely a history or record of how an algorithm might get from the string to the interpretation (or vice versa). Although it is convenient to write such structures down and refer to them as Surface Structures, precisely in order to make the point that no rules of domination, deletion, or movement apply to those structures, they do not constitute a grammatical level of representation at all. No rule ever needs to know how a category that it applies to was derived.

It is the combinatory derivations that correspond to Intonation Structure in the extended sense of the term defined above, as well as capturing coordinate structure and the effects of relativization. Surface Structure or derivation in the present sense therefore subsumes some functions of S-Structure, and all those of Intonation Structure, together with some of the role of PF as it is understood in GB. Phonetic Form in present terms really is no more than an abstract specification of speech segments.

The interpretation the derivation associates with a constituent of category S (or any other derived constituent) directly reflects such information-structural distinctions as those between theme and rheme and between focus and ground.[31] Such information-structural elements are evaluated with respect to alternative sets in the contextual database, and they may be discontinuous.

The present realization of Surface Structure as Information Structure conspicuously fails to represent traditional notions of dominance and command, including c-command. However, relations of dominance and command *are* represented in the canonical predicate-argument structure that results from the trivial procedure of normalizing or "β-reducing" the alternative Information Structures yielded by the alternative derivations of a given proposition, as discussed in connection with examples (5) and (6), and as implicitly assumed in derivations throughout the chapter. It follows that all grammatical relations that depend upon c-command, notably including binding and control and such

related phenomena as crossover, must be treated as properties of predicate-argument structure, not Surface Structure, a suggestion consistent with the observations of Lasnik and Saito (1992).

By incorporating the finer distinction between focus and background within both theme and rheme, the present grammar opens up further possibilities of addressing a range of questions in semantics that have been explained in terms of various notions of focus (see Chomsky 1971; Jackendoff 1972; Rooth 1985; Rochemont 1986; von Stechow 1991; Jacobs 1991; Hoeksema 1991; Kratzer 1991; Krifka 1991). In particular, one may expect some light to fall on certain phenomena that have been identified in semantic accounts of particles like *only*, which are claimed to "associate with focus" and which, as Jacobs (1991) and Krifka (1991, sec. 4.8) have noted, interact with intonation in puzzling ways. They are exemplified in sentences like the following:

(83) a. Freeman even introduced HARDY$_{H*}$ to Willis$_{LL\%}$.
      b. Freeman only introduced HARDY$_{H*}$ to WILLIS$_{!H* \ LL\%}$.

One might have expected that availability of quantifier scope alternations in scope-ambiguous sentences like *Some boy admires every saxophonist* might be affected by intonation, since we saw in chapter 4 that scope alternations are limited by syntactic derivation, and we have seen in the present chapter that intonation may limit combinatory derivation. Such an expectation would be in line with the claims of the Prague School that Information Structure determines scope—see Hajičová, Partee and Sgall (1998) for discussion. However, the lexical mechanism for quantifier scope advanced in chapter 4, motivated by examples like (55), makes scope entirely independent of which combinatory derivation is involved, just as long as there is one.

This means that according to the theory of scope ambiguities sketched in section 4.4, the effects of intonation on availability of readings are essentially limited to changes in the relative preference or salience of the readings that the competence grammar makes syntactically and semantically available, a conclusion that appears to be consistent with the observations on both sides of the debate in Hajičová, Partee and Sgall (1998), although it remains somewhat unclear what the facts in this area actually are.

Much further work remains to be done to complete this picture of the interface between grammar and speech. Nothing has been said here about the way metrically related phenomena of rhythm, timing, and lengthening are to be accommodated. (It should be obvious nevertheless that the theory offered here is consistent with more or less any of the available theories.)

Serious difficulties also attend the automatic identification of prosodic boundaries in speech. The phonetic realizations of elements such as pitch accents and boundary tones are subject to coarticulation effects, like all phonological segments, and are hard to recognize. In fact, it is highly likely that the process of identifying them cannot be separated from that of recognizing the words that carry them. This observation might seem daunting, since current techniques for word recognition, although improving dramatically, are nonetheless not very good. However, it is likely that the task of recognizing words and intonation together will turn out to be easier than doing either task in isolation, as Pierrehumbert (1993) points out.

The problem of so-called spurious ambiguity engendered by combinatory grammars now appears in a different light. Although the semantic properties of the rules (notably the associativity of functional composition) indeed allow alternative analyses that are equivalent in terms of the function-argument structure to which their interpretations reduce, the corresponding distinctions in surface constituency are nonetheless meaning-bearing. To call them "spurious" is very misleading, for they are genuine ambiguities at the level of Information Structure. Any theory that actually addresses the range of prosodic phenomena and coordinate constructions considered here must implicate exactly the same nondeterminism.

However, the question remains, how does the parser cope with structural ambiguity in general, and with this kind in particular? Sometimes of course intonation uniquely specifies structure. But very often it does not. PP attachment ambiguities, of the kind exhibited in the following sentence, are not usually disambiguated by intonation:

(84) Put the block in the box on the table.

Moreover, in the discussion in section 5.5.2 of the null tone on unmarked themes, we saw that Information Structure boundaries need not be disambiguated by intonation either.

The pragmatic nature of sentences with unmarked themes actually provides a strong suggestion about the nature of a mechanism for resolving not only the nondeterminism inherent in the null tone, but also other structural ambiguities such as PP attachment.

The null tone is found on the theme precisely when the corresponding theme is entirely ground information—that is, when it is already established in the context and known to the listener, and when nothing else in the context stands in contrast to it. That is to say, this particular ambiguity is only permitted when

the theme is already in the listener's model of the discourse. In the case of (55) this means that at successive positions in a left-to-right analysis of the string *Mary wrote a book about BATS*, the property corresponding to *Mary*, *Mary wrote*, and *Mary wrote a book about* can be derived and can be compared with the one(s) present in the model, so that choices between syntactic alternatives such as composing or not composing can be made accordingly. What is more, since the combinatory grammar allows more or less any leftmost substring to be treated as a constituent, complete with an interpretation, there exist very simple parsing algorithms that will permit incremental analysis of this kind, consistent with the strict competence hypothesis.

This may be the most significant practical benefit of the combinatory theory. In the past, syntax and semantics on the one hand, and phonology and discourse information on the other, have appeared to demand conflicting structural analyses and to require processing more or less independently. The present theory shows them to be in harmony. Processors may more easily be devised that use all these sources of information at once, potentially simplifying both problems. The fact that the combinatory notion of syntactic structure and interpretation stands in the closest possible relation both to the prosodic structure of the signal itself and to the concepts, referents, and themes represented in the discourse context should make it easier to use all of these higher-level sources of information to filter out the ambiguities that will inevitably continue to arise from processing at lower levels.

I will return to this architecture and to the question of how to process these grammars in part III (to which the more psycholinguistically or computationally inclined reader might well turn directly). But first it is important to look more deeply into the linguistic justification for the grammars proposed here. Part II presents two related case studies, which examine in detail the extent to which the theory generalizes to more complex constructions, including further varieties of coordination in English and other languages, and to their interaction with quantifier scope and intonation.

PART II

Coordination and Word Order