Surface Structure, Intonation, and "Focus"*

Mark Steedman

University of Pennsylvania

Abstract

The paper briefly reviews a theory of intonational prosody and its relation syntax, and to certain oppositions of discourse meaning that have variously been called "topic and comment", "theme and rheme", "given and new", or "presupposition and focus." The theory, which is based on Combinatory Categorial Grammar, is presented in full elsewhere. The present paper examines its implications for the semantics of "focus".

*Paper given to the ESPRIT worshop on Language and Speech, Brussels, Nov 1991. A revised version appears in E. Klein and F. Veltmann, (eds.) *Natural Language and Speech: Proceedings of the Symposium, ESPRIT Conference, Brussels, Nov. 1991*, Springer Verlag, Berlin. 21-38. The research was supported in part by NSF grant nos IRI90-18513 and CISE IIP, CDA 88-22719 , DARPA grant no. N00014-90-J-1863, and ARO grant no. DAAL03-89-C0031.

The structural units of phrasal intonation are frequently orthogonal to the syntactic constituent boundaries that are recognised by traditional grammar and embodied in most current theories of syntax. As a result, much recent work on the relation of intonation to discourse context and information structure has either eschewed syntax entirely (cf. [4], [10], [12], [5]), or has supplemented traditional syntax with entirely non-syntactic string-related principles (cf. [7]). Recently, Selkirk [23] and others have postulated an autonomous level of "intonational structure" for spoken language, distinct from syntactic structure. Structures at this level are plausibly claimed to be related to discourse-related notions, such as "focus". However, the involvement of two apparently uncoupled levels of structure in Natural Language grammar appears to complicate the path from speech to interpretation unreasonably, and to thereby threaten the feasibility of computational speech recognition and speech synthesis.

In [28] and [30], I argue that the notion of intonational structure formalised by Pierrehumbert, Selkirk, and others, can be subsumed under a rather different notion of syntactic surface structure, that emerges from the "Combinatory Categorial" theory of grammar [26], [27]. This theory engenders surface structure constituents corresponding directly to phonological phrase structure. Moreover, the grammar assigns to these constituents interpretations that directly correspond to what is here called "information structure" – that is, the aspects of discourse-meaning that have variously been termed "topic" and "comment", "theme" and "rheme", "given" and "new" information, and/or "presupposition" and "focus".

Certain syntactic categories and constructions, such as particles like *only* and *even* also "associate with focus" in the sense that they conspire with intonation in utterances like the following to yield a fixed information structure, carrying presuppositions about the background context.

(1) Harry only introduced MARY to Alice.

Here the effect is not only to make *Mary* "new" or the center of attention, and the fact that *Harry introduced someone to Alice* the background, but also, as Rooth [22] and von Stechow [24] have pointed out, to entail that Harry did not introduce anyone *else* to Alice. The present paper reviews the intonational theory and examines its applications to this problem.

§1 The Combinatory Grammar of Intonation

¶1.1 The Problem

One common intonation for an answer to the question (a) intuitively imposes the structure indicated by the brackets in b, below (stress, marked in this case by raised pitch, is indicated by capitals):

(2) a. I know that Alice likes velvet. But what does MAry prefer?
b. (MA-ry prefers) (CORduroy).

Such a grouping is orthogonal to the traditional syntactic structure of the sentence.

Intonational structure nevertheless remains strongly constrained by meaning. For example, contours imposing bracketings like the following are not allowed:

(3) #(Three cats)(in ten prefer corduroy)

Halliday [13] observed that this constraint, which Selkirk [23] has called the "Sense Unit Condition", seems to follow from the *function* of phrasal intonation, which is to convey what will here be called "information structure" – that is, distinctions of focus, presupposition, and propositional attitude towards entities in the discourse model. These discourse entities are more diverse than mere nounphrase or propositional referents, but they do not include such non-concepts as "in ten prefer corduroy."

Among the categories that they *do* include are what Wilson and Sperber and E. Prince [21] have termed "open propositions". One way of introducing an open proposition into the discourse context is by asking a Wh-question. For example, the question in 2, *What does Mary prefer?* introduces an open proposition. As Jackendoff [14] pointed out, it is natural to think of this open proposition as a functional *abstraction*, and to express it as follows, using the notation of the λ-calculus:

(4) $\lambda x\ [(\mathit{prefer}'\ x)\ \mathit{mary}']$

(Primes indicate semantic interpretations whose detailed nature is of no direct concern here.) When this function or concept is supplied with an argument *corduroy*$'$, it *reduces* to give a proposition, with the same function argument relations as the canonical sentence:

(5) $(\mathit{prefer}'\ \mathit{corduroy}')\ \mathit{mary}'$

It is the presence of the above open proposition rather than some other that makes the intonation contour in 2b felicitous. (That is not to say that its presence uniquely *determines* this response, nor that its explicit mention is necessary for interpreting the response.)

These observations have led linguists such as Selkirk to postulate a level of "intonational structure", independent of syntactic structure and related to information structure. The involvement of two apparently uncoupled levels of structure in natural language grammar appears to complicate the path from speech to interpretation unreasonably, and to thereby threaten a number of computational applications in speech recognition and speech synthesis.

It is therefore interesting to observe that all natural languages include syntactic constructions whose semantics is also reminiscent of functional abstraction. The most obvious and tractable class are Wh-constructions themselves, in which some 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 problematic class of fragments results from coordinate constructions. It is striking that the residues of wh-movement and conjunction reduction are also subject to something like a "sense unit condition". For example, strings like "in ten prefer corduroy" are as resistant to coordination as they are to being intonational phrases.[1]

(6) *Three cats in twenty like velvet, and in ten prefer corduroy.

Since coordinate constructions constitute another major source of complexity for theories of natural language grammar, and also offer serious obstacles to computational applications, the earlier papers suggest that this conspiracy

[1]I do not claim that such coordinations are absolutely excluded, just that if they are allowed at all then: a) extremely strong and unusual contexts are required, and b) that such contexts will tend to support 3 as well.

between syntax and prosody should be interpreted as evidence for a unified notion of structure that is somewhat different from traditional surface constituency, based on Combinatory Grammar.

¶1.2 Combinatory Grammars.

Combinatory Categorial Grammar (CCG, [26]) is an extension of Categorial Grammar (CG). Elements like verbs are associated with a syntactic "category" which identifies them as *functions*, and specifies the type and directionality of their arguments and the type of their result. We use a notation in which a rightward-combining functor over a domain β into a range α are written α/β, while the corresponding leftward-combining functor is written $\alpha\backslash\beta$. α and β may themselves be function categories. For example, a transitive verb is a function from (object) NPs into predicates – that is, into functions from (subject) NPs into S:

(7) $\textit{prefers} := (S\backslash NP)/NP : \textit{prefer}'$

Such categories can be regarded as encoding the semantic type of their translation, which in the notation used here is identified by the expression to the right of the colon. Such functions can combine with arguments of the appropriate type and position by functional application:

(8)
```
Mary   prefers   corduroy
----  ---------  --------
 NP   (S\NP)/NP     NP
      ---------------->
            S\NP
   -------------<
         S
```

The syntactic types are identical to semantic types, apart from the addition of directional information. The derivation can therefore also be regarded as building a compositional interpretation, ($\textit{prefer}'$ $\textit{corduroy}'$) $\textit{mary}'$, and of course such a "pure" categorial grammar is context free.

Coordination might be included in CG via the following rule, allowing constituents of like type to conjoin to yield a single constituent of the same type:

(9) $X \quad conj \quad X \;\Rightarrow\; X$

(10)
```
I    loath   and   detest   velvet
-- --------- ---- --------- ------
NP (S\NP)/NP conj (S\NP)/NP   NP
   ------------------------&
            (S\NP)/NP
```

(The rest of the derivation is omitted, being the same as in 8.) In order to allow coordination of contiguous strings that do not constitute constituents, CCG generalises the grammar to allow certain operations on functions related to Curry's combinators [9]. For example, functions may nondeterministically *compose*, as well as apply, under the following rule:

(11) *Forward Composition:* (>B)
$$X/Y : F \quad Y/Z : G \;\Rightarrow\; X/Z : \lambda x\ F(Gx)$$

The most important single property of combinatory rules like this is that they have an invariant semantics. This one composes the interpretations of the functions that it applies to, as is apparent from the right hand side of the rule.[2] Thus sentences like *I suggested, and would prefer, corduroy* can be accepted, via the following composition of two verbs (indexed as **B**, following Curry's nomenclature) to yield a composite of the same category as a transitive verb. Crucially, composition also yields the appropriate interpretation for the composite verb *would prefer*:

(12)
```
... suggested  and   would   prefer ...
    --------- ---- --------- ------
    (S\NP)/NP conj (S\NP)/VP VP/NP
                   --------------->B
                     (S\NP)/NP
    -------------------------&
         (S\NP)/NP
```

[2]The rule uses the notation of the λ-calculus in the semantics, for clarity. This should not obscure the fact that it is functional composition itself that is the primitive, not the λ operator.

Combinatory grammars also include type-raising rules, which turn arguments into functions over functions-over-such-arguments. These rules allow arguments to compose, and thereby take part in coordinations like *I dislike, and Mary prefers, corduroy.* They too have an invariant compositional semantics which ensures that the result has an appropriate interpretation. For example, the following rule allows the conjuncts to form as below (again, the remainder of the derivation is omitted):

(13) *Subject Type-raising:* (>T)
$NP : y \;\; \Rightarrow \;\; S/(S\backslash NP) : \lambda F \; Fy$

(14)
```
    I     dislike   and   Mary    prefers  ...
-------- --------- ---- -------- ---------
   NP    (S\NP)/NP conj    NP    (S\NP)/NP
-------->T              -------->T
S/(S\NP)                S/(S\NP)
------------------->B    ------------------->B
        S/NP                   S/NP
        ---------------------------&
                    S/NP
```

This apparatus has been applied to a wide variety of coordination phenomena, including "left node raising" [11], "backward gapping" in Germanic languages, including verb-raising constructions [25], and gapping, [27]. For example, the following analysis is proposed by Dowty [11] for the first of these:

(15)
```
   give          Mary            corduroy   and        Harry           velvet
---------- ------------------<T --------<T ---- ------------------<T --------<T
(VP/NP)/NP (VP/NP)\((VP/NP)/NP) VP\(VP/NP) conj (VP/NP)\((VP/NP)/NP) VP\(VP/NP)
           -------------------------------<B    -------------------------------<B
                    VP\((VP/NP)/NP)                         VP\((VP/NP)/NP)
                    -------------------------------------------------------<&>
                                 VP\((VP/NP)/NP)
------------------------------------------------<
                      VP
```

The important feature of this analysis is that it uses "backward" rules of type-raising <T and composition <B that are the exact mirror-image of the

two “forward” versions introduced as examples 11 and 13. It is therefore a prediction of the theory that such a construction can exist in English, and its inclusion in the grammar requires no additional mechanism whatsoever. The earlier papers show that no *other* non-constituent coordinations of dative-accusative NP sequences are allowed in any language with the English verb categories, given the assumptions of CCG. Thus the following are ruled out in principle, rather than by stipulation:

(16) a. *Harry velvet and give Mary corduroy
b. *give corduroy Mary and velvet Harry

A number of related well-known cross-linguistic generalisations concerning the dependency of so-called “gapping” upon lexical word-order are also captured (see Dowty [11] and others [25], [27]).

¶1.3 Intonation, Parsing, and Context

Examples like the above show that combinatory grammars embody a view of surface structure according to which strings like *Mary prefers* are constituents. It follows, according to this view, that they must also be possible constituents of non-coordinate sentences like *Mary prefers corduroy*, as in the following derivation:

(17)
```
   Mary    prefers   corduroy
 -------- ---------  --------
    NP    (S\NP)/NP     NP
 -------->T
 S/(S\NP)
 ------------------>B
        S/NP
        ------------------>
                S
```

An entirely unconstrained combinatory grammar would in fact allow any bracketing on a sentence, although the grammars we actually write for configurational languages like English are heavily constrained by local conditions. (An example might be a condition on the composition rule that is tacitly assumed below, forbidding the variable Y in the composition rule to be instantiated as NP, thus excluding constituents like $*[ate\ the]_{VP/N}$). It nevertheless follows that, for each semantically distinct analysis of a sentence, the involvement of the combinatory operation of functional composition engenders an equivalence class of derivations, which impose different constituent structures but are guaranteed to yield identical interpretations. In more complex sentences than the above, there will be many semantically equivalent derivations for each distinct interpretation.

Such additional non-determinism in grammar, over and above the non-determinism that is usually recognised, creates obvious problems for the parser, and has on occasion been referred to as "spurious" ambiguity. This term is very misleading. Whether or not the present theory is correct, the non-determinism is *there*, in the competence grammar of coordinate constructions, and any parser that actually covers this range of constructions will have to deal with it. It is only the comparitive neglect of these constructions by the parsing community that has led them to ignore this perfectly genuine source of nondeterminism. The papers [30] and [31] discuss the complexity of this problem in the worst case. in [8] it is suggested that the evaluation of partial, incomplete, interpretations with respect to a discourse model including a representation of discourse information plays a crucial role.

However the parsing problem is resolved, the interest of such non-standard structures for present purposes should be obvious. The claim is simply that the non-standard surface structures that are induced by the combinatory grammar to explain coordination in English subsume the intonational structures that are postulated by Pierrehumbert *et al.* to explain the possible intonation contours for sentences of English. The claim is that 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 discourse focus among the concepts and open propositions that the speaker has in mind.

The proof of this claim lies in showing that the rules of combinatory grammar can be made sensitive to intonation contour, which limit their application in spoken discourse. We must also show that the major constituents of intonated utterances like 2b, under the analyses that are permitted by any given intonation, correspond to the information structure of the context to which the intonation is appropriate, as in (a) in the example 2 with which the proposal begins. This demonstration will be quite simple, once we have established the following notation for intonation contours.

We will use a notation which is based on the theory of Pierrehumbert [18], as modified in more recent work by Selkirk [23], Beckman and Pierrehumbert [2], [19], and Pierrehumbert and Hirschberg [20], and as explicated in the chapter by Pierrehumbert in the present volume. The theory proposed below is in fact compatible with any of the standard descriptive accounts of phrasal intonation. However, a crucial feature of Pierrehumbert's theory for present purposes is that it distinguishes two subcomponents of the prosodic phrase, the *pitch accent* and the *boundary*.[3] The first of these tones or tone-sequences coincides with the perceived major stress or stresses of the prosodic phrase, while the second marks the righthand boundary of the phrase. These two components are essentially invariant, and all other parts of the intonational tune are interpolated. Pierrehumbert's theory thus captures in a very natural way the intuition that the same tune can be spread over longer or shorter strings, in order to mark the corresponding constituents for the particular distinction of focus and propositional attitude that the melody denotes. It will help the exposition to augment Pierrehumbert's notation with explicit prosodic phrase boundaries, using brackets. These do not change her theory in any way: all the information is implicit in the original notation.

Consider for example the prosody of the sentence *Mary prefers corduroy* in the following pair of discourse settings, which are adapted from Jackendoff [14, pp. 260]:

(18) Q: Well, what about the CORduroy? Who prefers THAT?

```
     A:  (MARy )   (prefers CORduroy).
           H* L              L+H* LH%
```

[3] For the purposes of this proposal, the distinction between the intonational phrase proper, and what Pierrehumbert and her colleagues call the "intermediate" phrase is ignored. These categories differ in respect of boundary tone-sequences – see the chapter by Pierrehumbert in the present volume.

(19) Q: Well, what about MARy? What does SHE prefer?

A: (MARy prefers) (CORduroy).
L+H* LH% H* LL%

In these contexts, the main stressed syllables on both *Mary* and *corduroy* receive a pitch accent, but a different one. In the former example, 18, there is a prosodic phrase on *Mary* made up of the pitch accent which Pierrehumbert calls H*, immediately followed by an L boundary. There is another prosodic phrase having the pitch accent called L+H* on *corduroy*, preceded by null or interpolated tone on the words *prefers*, and immediately followed by a boundary which is written LH%. (I base these annotations on Pierrehumbert and Hirschberg's [20, ex. 33] discussion of a similar example.)[4] In the second example 19 above, the two tunes are reversed: this time the tune with pitch accent L+H* and boundary LH% is spread across a prosodic phrase *Mary prefers*, while the other tune with pitch accent H* and boundary LL% is carried by the prosodic phrase *corduroy* (again starting with an interpolated or null tone).[5]

The meaning that these tunes convey in these contexts is intuitively very obvious.[6] As Pierrehumbert and Hirschberg point out, the latter tune seems to be used to mark some or all of that part of the sentence expressing information that the speaker believes to be *novel to the hearer*. In traditional terms, it marks the "comment" – more precisely, what Halliday called the "rheme". In contrast, the L+H* LH% tune seems to be used to mark some or all of that part of the sentence which expresses information which in traditional terms is the "topic" – in Halliday's terms, the "theme". For present purposes, a theme can be thought of as conveying *what the speaker assumes to be the subject of mutual interest*, and this particular tune marks a theme as *novel to the conversation as a whole*, and as standing in a contrastive relation to the previous theme. (If the theme is not novel in this sense, it receives *no*

[4] We continue to gloss over Pierrehumbert's distinction between "intermediate" and "intonational" phrases.

[5] The reason for notating the latter boundary as LL%, rather than L is again to do with the distinction between intonational and intermediate phrases.

[6] I do not of course intend to claim that these are the *only* meanings that these tunes can convey.

tone in Pierrehumbert's terms, and may even be left out altogether.)[7] Thus in 19, the L+H* LH% phrase including this accent is spread across the phrase *Mary prefers*.[8] Similarly, in 18, the same tune is confined to the object of the open proposition *prefers corduroy*, because the intonation of the original question indicates that prefering corduroy *as opposed to some other stuff* is the new topic or theme.

It follows that the position of the pitch accent in the phrase has to do with a further orthogonal dimension of information structure *within both theme and rheme*, corresponding to *the interesting bit* of either information unit. It is tempting to call this part the "focus" of the theme or rheme, but I shall avoid the term and follow Halliday in calling it "new" information, in contrast to the "given" information accompanied by the null tone. The term "new" is perhaps not entirely satisfactory, but it is intended to suggest that the part of the theme that is marked in this way is *new to the speaker*, while the part of the rheme that is so marked is believed by the speaker to be *new to the hearer*. This usage is illustrated in the following example:

(20) Q: I know that Mary's FIRST degree is in PHYSICS.
But what is the subject of her DOCTORATE?

A: (Mary's	DOCTORATE)	(is in	CHEMISTRY)
	L+H* LH%		H* LL%
Given	*New*	*Given*	*New*
Theme		*Rheme*	

Here the theme is *Mary's doctorate*, where the head noun is emphasised because it stands in contrast to another of her qualifications. The rheme is that it *is in chemistry*, where chemistry is emphasised in contrast to another subject.

¶1.4 Combinatory Prosody

[7]Here I depart slightly from Halliday's definition. The present proposal also follows Lyons [16] in rejecting Hallidays' claim that the theme must necessarily be sentence-initial.

[8]An alternative prosody, in which the contrastive tune is confined to *Mary*, seems equally coherent, and may be the one intended by Jackendoff. I believe that this alternative is informationally distinct, and arises from an ambiguity as to whether the topic of this discourse is *Mary* or *What Mary prefers*. It too is accepted by the rules below.

The L+H* LH% intonational melody in example 19 belongs to a phrase *Mary prefers ...* which corresponds under the combinatory theory of grammar to a grammatical constituent, complete with a translation equivalent to the open proposition $\lambda x[(prefer'\ x)\ mary']$ introduced by the question in that example *What does Mary prefer?* The combinatory theory thus offers a way to derive such intonational phrases, using only the independently motivated rules of combinatory grammar, entirely under the control of appropriate intonation contours like L+H* LH%.

In fact we might think of doing this in one of two ways. We might either think of annotating combinatory rules, for example imposing a restriction forbidding composition across an intonational phrase boundary.[9] Or one might exploit the advantages of categorial grammar more directly, as follows. [10]

¶1.5 Constituency and Intonation.

We will begin by defining the two pitch accents as functions of the following types:[11]

(21)
$$\begin{array}{lll} \text{L+H*} & := & Theme/Bh \\ \text{H*} & := & (Utterance/Theme)/bl \\ \text{H*} & := & (Utterance\backslash Theme)/Bl \end{array}$$

These categories define the two pitch accents as functions over boundary tones into the two major informational types, *Theme* and *Rheme*, where the latter category is itself a function from themes into utterances.[12] In the case of the H* pitch accent, there are two categories, one combining with an intermediate phrase boundary to yield a *forward* category of the latter kind, the other combining with an intonational phrase boundary to yield a

[9]This is in fact what I advocated in [29].

[10]This section is a summary and an extension of the proposals in [28] and [30].

[11]The following rules constitute a minor revision of [30], prompted by Steven Bird's [3] observation of an egregious overgeneralisation arising from the conflation of intonational and intermediate phrases in the earlier account.

[12]The choice of the rheme rather than the theme as the "head" of the prosodic utterance is to some extent arbitrary, but is motivated by the observation that the rheme is the obligatory member of the pair.

backward version. It will be convenient to refer to the backward and forward category collectively as "the rheme category".

We define the various boundary tones as arguments to these functions, as follows:[13]

(22) LH% := Bh
LL% := Bl
L := bl

Finally, we accomplish the effect of interpolation of other parts of the tune by assigning the following category to all elements bearing no tone specification, which we will represent as the null tone Ø:[14]

(23) Ø := X/X

X is a category that can match any category, importantly including X/X.[15] It will therefore introduce a considerable amount of non-determinism to the prosodic side of the grammar. However, this will turn out to be strictly necessary: the null tone is very ambiguous in present terms.

The prosodic combinatory rules include forward and backward functional application. They also include the following very restricted version of forward functional composition:

(24) *Forward Prosodic Functional Composition:*
$X/Y \quad Y/Z \quad \Rightarrow \quad X/Z$
where $Y \in \{Bh, Bl, bl\}$

[13] An alternative grammar, which would be closer to Pierrehumbert & Beckman 1989, and which might also be more directly compatible with Pierrehumbert and Hirschberg's proposals for the compositional assembly of discourse meanings from more primitive elements of meaning carried by each individual tone, would be obtained by assigning pitch accents the category of functions from PHRASAL tones into intermediate phrases marked as theme, rheme etc, and assigning the boundary tones the category of functions from intermediate to intonational phrases, similarly marked. However, the precise details of such an alternative depend on some imponderables in the original theory concerning the precise position and nature of the phrasal tone itself (see Pierrehumbert & Beckman 1989:236-237 and an earlier footnote).

[14] I am grateful to Steven Bird for first suggesting a related category for the null tone.

[15] In a unification-based realisation such as that sketched in Steedman 1990a, X would be realised as the equivalent of a unique Prolog variable, distinct from that in any other instance of the null tone.

The restriction is required because the whole point of the prosodic categories is to PREVENT composition across the theme/rheme boundary. It will become apparent below that, given the categories chosen above, the only occasion on which composition is required is when X/Y is a pitch accent – that is, a function over a boundary. Another important consequence is that two null tones of type X/X can only combine by application, not composition.

Syntactic combination can now be made subject to intonation contour by the following simple restriction:

(25) THE PROSODIC CONSTITUENT CONDITION: Combination of two syntactic categories via a syntactic combinatory rule is only allowed if their prosodic categories can also combine (and *vice versa*).

(The prosodic and syntactic combinatory rules need not, and usually will not, be the same).

This principle has the sole effect of excluding certain derivations for spoken utterances that would be allowed for the equivalent written sentences. For example, consider the derivations that it permits for example 19 above. The rule of forward composition is allowed to apply to the words *Mary* and *ate*, because the prosodic categories can combine (by functional application):

(26)
```
       Mary                 prefers       ...
      L+H*                    LH%
---------------      -----------------
   NP:mary'          (S\NP)/NP:prefer'
   Theme/Bh                  Bh
------------------->T
S/(S\NP): \P[P mary']
   Theme/Bh
------------------------------------->B
       S/NP: \X[(prefer' X) mary']
              Theme
```

The category X/X of the null tone allows intonational phrasal tunes like L+H* LH% tune to spread across any sequence that forms a grammatical constituent according to the combinatory grammar. For example, if the reply to the same question *What does Mary prefer?* is *MARY says she prefers CORduroy*, then the tune will typically be spread over *Mary says she prefers*

... as in the following (incomplete) derivation, in which much of the syntactic and semantic detail has been omitted in the interests of brevity:

(27)
```
   Mary     says        she       prefers  ...
  L+H*                              LH%
-------->T --------   -------->T ---------
S/(S\NP)   (S\NP)/S   S/(S\NP)   (S\NP)/NP
Theme/Bh      X/X       X/X         Bh
-------------------->B
     Theme/Bh
     ------------------------->B
            Theme/Bh
            ------------------------------>B
                         Theme
```

Such non-standard constituents, marked as Theme, can take part in such derivations as the following:

(28)
```
      Mary               prefers             the           green         corduroy
     L+H*                 LH%                                             H* LL%
---------------   -----------------   ----------    ----------   -------------
   NP:mary'       (S\NP)/NP:prefer'   NP/N: the'    N/N:green'    N:corduroy'
  Theme/Bh               Bh              X/X           X/X      Utterance\Theme
---------------->T                    ----------------------->B
S/(S\NP):\P[P mary']                  NP/N:\Y[the'(green'(Y))]
  Theme/Bh                                      X/X
---------------------------------->B  ----------------------------------------->
       S/NP: \X[(eat' X) mary']              NP:the'(green' corduroy')
                Theme                           Utterance\Theme
       ------------------------------------------------------------->
                  S: eat' (the'(green' corduroy')) mary'
                             Utterance
```

The division of the utterance into an open proposition constituting the theme and an argument constituting the rheme is appropriate to the context established in 19. Moreover, the theory permits no *other* division into a theme and a rheme for this intonation contour.

All the other possibilities for combining these two contours in a simple sentence are shown elsewhere [30] to yield similarly unique and contextually appropriate interpretations.

Sentences like the above, including marked theme and rheme expressed as two distinct intonational/intermediate phrases are by that token unambiguous as to their information structure. However, sentences like the following, which in Pierrehumbert's' terms bear a single intonational phrase, are much more ambiguous as to the division that they convey between theme and rheme:

(29) (I read a book about CORduroy)
H* LL%

Such a sentence is notoriously ambiguous as to the open proposition it presupposes, for it seems equally appropriate as a response to any of the following questions:

(30) a. What did you read a book about?
b. What did you read?
c. What did you do?

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

(31) a. (I read a book about)(CORduroy)
b. (I read)(a book about CORduroy)
c. (I)(read a book about CORduroy)

It seems that we shall miss a generalisation concerning the relation of intonation to discourse information unless we extend Pierrehumbert's theory very slightly, to allow prosodic constituents resembling *null* intermediate phrases, without pitch accents, expressing unmarked themes. Since the boundaries of such intermediate phrases are not explicitly marked, we shall immediately allow all of the above analyses for 29. Such a modification to the theory can be introduced by the following rule, which nondeterministically allows constituents bearing the null tone to become a theme:

(32) *Null Theme Promotion Rule:*
$X/X \quad \Rightarrow \quad Theme$

The rule is nondeterministic, so it correctly continues to allow a further analysis of the entire sentence as a single Intonational Phrase conveying the

Rheme. Such an utterance is the appropriate response to yet another open-proposition establishing question, *What happened?*.)

The grammar exemplified in the derivation 28, above, implicitly embodies an explanation of the second dimension of information structure identified above, Halliday's contrast between "given" and "new" information. The derivation that is given there for the rheme *the green corduroy* is only one of two alternatives permitted by that intonation contour, represented by the following bracketings:

(33) a. (the green)(CORduroy)
b. (the) (green CORduroy)

However, this ambiguity, like the ambiguity inherent in the null theme, is also a genuine one. The intonation contour in question does not distinguish between the information structures appropriate to responses to the following two questions:

(34) a. Does Mary like the green corduroy or the green velvet?
b. Does Mary like the green corduroy or the red velvet?

That is, it does not distinguish whether *green* is part of the given or part of the new. This is a subtle ambiguity, but its reality is clear from the existence of the intonation contour exhibited in the following exchange, for which the grammar only permits the second of these derivations, a fact that is consistent with the intuition that here it is the entire N-phrase *green corduroy* that is marked as being new and interesting:

(35) Q. Does Mary like the green corduroy or the red cordurory?
A. (MARY prefers)(the (GREEN corduroy))

It follows from this analysis that the related utterance with an unmarked theme is extremely ambiguous as to its information structure, since there are several ways of splitting it up into a theme and a rheme, and there may be several ways of further dividing each of these into given and new:

(36)	(Mary prefers the green	CORDUROY).
		H* LL%

However, this ambiguity appears to be correctly constrained by the grammar. That is, the following does not appear to be a possible information structure, anymore than it is a possible syntactic structure. (Relativisation and Coordination reveal that the sentence is in violation of an NP constraint.)

(37) *(Mary prefers the green)$_{Theme}$(CORDUROY)$_{Rheme}$

With the generalisation implicit in the above rules, we are therefore in a position to make the following claim:

(38) The structures demanded by the theory of intonation and its relation to contextual information are the same as the surface syntactic structures permitted by the combinatory grammar.

Because constructions like relativisation and coordination are more limited in the derivations they require, often forcing composition, rather than permitting it, a number of corollaries follow, such as the following:

(39) Anything which can coordinate can be an intonational constituent, and *vice versa*.

and

(40) Anything which can be the residue of relativisation can be an intonational constituent.

These claims are discussed further in [30].

Under the present theory, the pathway between the speech-wave and the sort of logical form that can be used to interrogate a database is as in Figure 1. Such an architecture is considerably simpler than the one that is implicit in the standard theories. Phonetic form now maps via the rules of metrical phonology directly onto surface structure, whose highest level constituents correspond to intonational/intermediate phrases, annotated as to discourse functions like Theme and Rheme. The finer structure of those constituents in turn determines further distinctions of discourse meaning, such as that between Given and New information. Surface structure is therefore isomorphic to intonational structure, and subsumes information structure.

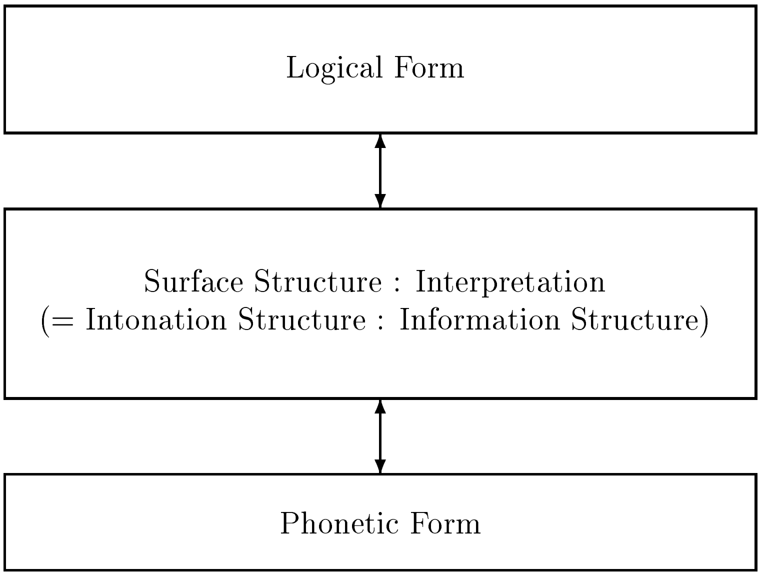

Figure 1: Architecture of a CCG-based Prosody

The major constituents carrying the theme (if any) and rheme in turn reduce via functional application to yield the interpretation of the root S node. This object makes explicit the traditional function argument relations of the proposition that is common to all these information structural derivations.[16] The canonical function-argument structure is neutral with respect to distinctions of quantifier scope. It is presumably therefore further mapped onto a level of "logical form", at which such distinctions are drawn, a matter about which we have nothing to say here.

§2 Information Structure and "Focus"

Now that we have made explicit in the grammar the finer distinction within theme and rheme between given and new information, we are in aposition

[16]This particular interpretation resembles a traditional "deep structure" or (less exactly) a GB "S-structure" in this respect. In particular, it is in this representation that relations such as *c*-command can be defined in pretty much their traditional form.

to address a wider range of questions in discourse information that have been identified with the notion of "focus" (cf. Jackendoff [14]; Chomsky [?]; [22]). In present terms, "focus" seems to correspond to the notion "the new information within the rheme". We are therefore in a position to ask whether this grammar captures phenomena of "focus" that have been identified in semantic accounts of the focussing particles. If so, we should have some strong predictions concerning sentences like the following:

(41) a. (John only introduced)(BILL to Sue)
b. = (John introduced)(only BILL to Sue)
c. (John only introduced Bill)(to SUE)
d. (John only introduced)(BILL to SUE)

It will also examine some apparent counterexamples noted by Rooth, where it appears that information structure of this kind can violate some well-known constraints on syntactic structure, thus threatening the claim of isomorphism between syntax, intonation structure and information structure. These apparent counterexamples concern sentences like the following:

(42) They only asked whether you knew the woman who chairs the ZONING board.

The presupposition here seems to be that *they only asked whether you knew the woman who chairs something.* However, this cannot be a constituent of syntax, because it is in violation of the complex NP constraint:

(43) *Which board did they ask whether you knew the woman who chairs?

Such examples were used by Roth to argue agains a "movement" account of focus. However, it looks at first glance as though they are equally telling against the present theory of discourse information.

It certainly is thie case that the present theory does not allow the sentence to be split into *the zoning board* as rheme and the rest of the sentence as theme. If it did, then the following intonation, in which this illegal constituent is marked as theme by the theme tune would wrongly be allowed:

(44) *(They only asked whether you know the woman who chairs)$_{L+H*LH\%}$ (the ZONING board)$_{H*LL\%}$

The paper will argue that the present theory already accounts correctly for 42. The theory implies there are actually *two* givens involved, one belonging to the open proposition or theme, and one stemming from its complement, the rheme. Both givens are marked (or rather, unmarked) by the null tone.[17] Both are presupposed, and therefore affect the p-set, or set of related propositions that are denied. The following example is in fact only one of the information structures that the grammar given earlier will permit.

(45) (They only asked whether you knew)(the woman who chairs the ZONING board).

The evidence is as follows. First, we know that *only* does not simply associate with *the zoning board*, because 42 does not mean the same as the following:

(46) They asked whether you knew the woman who chaired only the ZONING board.

That is, 42 does not entail that they did not ask whether you knew the woman who chaired the zoning board and the parking permit committee, as this analysis would imply. In contrast, 42 *can* mean the same as the following example, which is provided with a contextual question motivating the division concerned:

(47) Which women did they ask whether I knew?
(They asked whether you knew)(only the woman who chairs the ZONING board)

We can tell that this is so by *marking* part of the theme hypothesised in 45 as new, using the theme tune L+H* LH%:

(48) (They only asked whether you (KNEW)$_{L+H*LH\%}$)(the woman who chairs (the ZONING board).

In both cases, the p-set or set of negative entailments includes the following, just as it would if the grammar were able to build a monolithic open proposition **They asked whether you knew the woman who chaired ...*:

(49) a. They didn't ask whether you knew the woman who chairs the Parks Committee.
b. They didn't ask whether you knew the man who (co-)chairs the Zoning Board.
(etc.)

[17] This detail is still not included in the formal grammar, but is implicit.

The difference is that part each of the entailment stems from the open proposition or theme, and part of it stems from the rheme.

In the longer version of this paper I shall provide a more complete grammar and examine the information structures that arise in some further derivations that are permitted, and are here left as an exercise.

References

[1] Altmann, Gerry and Mark Steedman: 1988, 'Interaction with Context During Human Sentence Processing' *Cognition*, 30, 191-238.

[2] Beckman, Mary and Janet Pierrehumbert: 1986, 'Intonational Structure in Japanese and English', *Phonology Yearbook*, 3, 255-310.

[3] Bird, Steven: 1991, 'The Intonational Phrase in Sign-based Grammar', Workshop on Declarative Persectives on the Syntax-Prosody Interface, Utrecht, May 1991.

[4] Bolinger, Dwight: 1972, 'Accent is Predictable (If You're a Mind Reader)', *Language*, 48, 633-644.

[5] Brown, Gillian, and George Yule: 1983, *Discourse Analysis*, Cambridge, University Press.

[6] Chomsky, Noam: 1970, 'Deep Structure, Surface Structure, and Semantic Interpretation', in D. Steinberg and L. Jakobovits, *Semantics*, CUP, Cambridge, 1971, 183-216.

[7] Cooper, William and Julia Paccia-Cooper: 1980, *Syntax and Speech*, Harvard University Press, Cambridge MA.

[8] Crain, Stephen and Mark Steedman: 1985, 'On not being led up the garden path: the use of context by the psychological parser', in D. Dowty, L. Kartunnen, and A. Zwicky, (eds.), *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, ACL Studies in Natural Language Processing, Cambridge University Press, 320-358.

[9] Curry, Haskell and Robert Feys: 1958, *Combinatory Logic*, North Holland, Amsterdam.

[10] Cutler, Anne, and Stephen Isard: 1980, 'The Production of Prosody', in Brian Butterworth, (ed.), *Language Production, Vol. 1*, New York, Wiley, 246-269.

[11] Dowty, David: 1988, Type raising, functional composition, and non-constituent coordination, in Richard T. Oehrle, E. Bach and D. Wheeler, (eds), *Categorial Grammars and Natural Language Structures*, Reidel, Dordrecht, 153–198.

[12] Gussenhoven, Carlos: 1983, *On the Grammar and Semantics of Sentence Accent*, Dordrecht, Foris.

[13] Halliday, Michael: 1967, *Intonation and Grammar in British English*, Mouton, The Hague.

[14] Jackendoff, Ray: 1972, *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge MA.

[15] Joshi, Aravind, Bonnie Lynn Webber, and Ivan Sag (Eds.): 1981, *Elements of Discourse Understanding*, Cambridge, University Press.

[16] Lyons, John: 1977. *Semantics, vol. II*, Cambridge University Press.

[17] Markel, John, and Augustine Gray: 1976, *Linear Prediction of Speech*, Springer-Verlag, Berlin.

[18] Pierrehumbert, Janet: 1980, *The Phonology and Phonetics of English Intonation*, Ph.D dissertation, MIT. (Dist. by Indiana University Linguistics Club, Bloomington, IN.)

[19] Pierrehumbert, Janet, and Mary Beckman: 1989, *Japanese Tone Structure*, MIT Press, Cambridge MA.

[20] Pierrehumbert, Janet, and Julia Hirschberg, 1987, 'The Meaning of Intonational Contours in the Interpretation of Discourse', ms. Bell Labs.

[21] Prince, Ellen F. 1986. On the syntactic marking of presupposed open propositions. Papers from the Parasession on Pragmatics and Grammatical Theory at the 22nd Regional Meeting of the Chicago Linguistic Society, 208-222.

[22] Rooth, Mats: 1985, *Association with Focus*, unpublished PhD dissertation, University of Massachusetts, Amherst.

[23] Selkirk, Elisabeth: 1984, *Phonology and Syntax*, MIT Press, Cambridge MA.

[24] von Stechow, Arnim: 1989, 'Focussing and Backgrounding Operators', Fachgruppe Sprachwissenschaft der Universität Konstanz, Arbeitspapier Nr. 6.

[25] Steedman, Mark: 1985a. Dependency and Coordination in the Grammar of Dutch and English, *Language* 61.523-568.

[26] Steedman, Mark: 1987. Combinatory grammars and parasitic gaps. *Natural Language & Linguistic Theory*, 5, 403-439.

[27] Steedman, Mark: 1990. 'Gapping as Constituent Coordination', *Linguistics & Philosophy*, 13, 207-263.

[28] Steedman, Mark: 1990, Structure and Intonation in Spoken Language Understanding, *Proceedings of the 28th Annual Conference of the ACL*, Pittsburgh, June 1990, 9-17.

[29] Steedman, Mark: 1991a, Parsing Spoken Language with Combinatory Grammars, in Masaru Tomita (ed.) *Current Issues in Parsing Technology*, Proceedingd of the International Workshop on Parsing Technologies, Pittsburg, August 1989, Kluwer, Boston. 113-126.

[30] Steedman, Mark: 1991, Structure and Intonation, *Language*, 68, 260-296.

[31] Vijay-Shankar, K and David Weir: 1990, 'Polynomial Time Parsing of Combinatory Categorial Grammars', *Proceedings of the 28th Annual Conference of the ACL*, Pittsburgh, June 1990.