



July 1990

# Structure and Intonation

Mark Steedman  
*University of Pennsylvania*

Follow this and additional works at: [http://repository.upenn.edu/cis\\_reports](http://repository.upenn.edu/cis_reports)

---

## Recommended Citation

Mark Steedman, "Structure and Intonation", . July 1990.

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-90-45.

This paper is posted at ScholarlyCommons. [http://repository.upenn.edu/cis\\_reports/571](http://repository.upenn.edu/cis_reports/571)  
For more information, please contact [libraryrepository@pobox.upenn.edu](mailto:libraryrepository@pobox.upenn.edu).

---

# Structure and Intonation

## **Abstract**

Rules for assigning phrasal intonation to sentences are often assumed to require an autonomous level of "intonational structure", distinct from what is usually thought of as surface syntactic structure. The present paper argues that the requisite notion of structure can be subsumed under the generalised notion of surface structure that emerges from the combinatory extension of Categorical Grammar. According to this theory, the syntactic structures and the intonational structures of English are one, and can be captured in a single unified grammar. The interpretations that the grammar provides for such constituents correspond to the entities and open propositions that are concerned in certain discourse-related aspects of intonational meaning that have variously been described as "theme" and "rheme", "given" and "new", or "presupposition" and "focus".

## **Comments**

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-90-45.

**Structure and Intonation**

**MS-CIS-90-45  
LINC LAB 174**

**Mark Steedman**

**Department of Computer and Information Science  
School of Engineering and Applied Science  
University of Pennsylvania  
Philadelphia, PA 19104**

**July 1990**

# STRUCTURE AND INTONATION\*

MARK STEEDMAN

*University of Pennsylvania*

Rules for assigning phrasal intonation to sentences are often assumed to require an autonomous level of “intonational structure”, distinct from what is usually thought of as surface syntactic structure. The present paper argues that the requisite notion of structure can be subsumed under the generalised notion of surface structure that emerges from the combinatory extension of Categorical Grammar. According to this theory, the syntactic structures and the intonational structures of English are one, and can be captured in a single unified grammar. The interpretations that the grammar provides for such constituents correspond to the entities and open propositions that are concerned in certain discourse-related aspects of intonational meaning that have variously been described as “theme” and “rheme”, “given” and “new”, or “presupposition” and “focus”.

## THE PROBLEM

1. It is well-known that phrasal intonation organises the perceived grouping of words in spoken utterances in ways which are, on occasion, inconsistent with traditional linguistic notions of syntactic constituency. For example, consider the following exchange:

- (1) a. I know that Alice likes velvet. But what does MARY prefer?
- b. (MARY prefe-rs) (CORDUROY).

---

\*I am grateful to Steven Bird, Dwight Bolinger, Elisabet Engdahl, Ellen Hays, Julia Hirschberg, Jack Hoeksema, Stephen Isard, Aravind Joshi, Ewan Klein, Bob Ladd, Mark Liberman, Mitch Marcus, Dick Oehrle, Donna Jo Napoli, Michael Niv, Janet Pierrehumbert, Henry Thompson, Bonnie Lynn Webber, and three anonymous referees for comments, advice, and moral support at various stages. The research was supported in part by NSF Grant CISE IIP, CDA 88-22719, DARPA grant no. N00014-90-J-1863, and ARO grant no. DAAL03-89-C0031 to CIS, U.Penn. To appear in slightly revised form in *Language*, 1991.

One normal prosody for the answer (b) to the question (a) consists in not only marking the new information in the answer by the use of high pitch on the stressed first syllable of the word *corduroy*, but also in stressing the first syllable of *Mary*, using a high pitch-accent, and placing a final rise at the end of *prefers*, with lower pitch interpolated in between. This intonation contour, which conveys the contrast between the previous topic concerning Alice and the new one concerning Mary, imposes the perceptual grouping indicated by the brackets (stress is indicated by capitals).<sup>1</sup> Such a grouping cuts across the traditional syntactic analysis of the sentence as a subject and a predicate VP. The presence of two apparently uncoupled levels of structure in natural language grammar appears to complicate the path from speech to interpretation unreasonably. Such a theory seems likely to be very difficult to apply in the form of computer programs for automatic speech synthesis or recognition.

Despite its apparent independence from syntax, it is widely accepted that intonational structure is, nonetheless, strongly constrained by meaning, and in particular by distinctions of focus, information, and propositional attitude towards concepts and entities in the discourse model. For example, the intonation contour in the above example seems to divide the utterance into what Ellen Prince 1986, following Wilson & Sperber 1979, calls an “open proposition”, and its complement *corduroy*. (A similar partition is embodied in Cresswell’s (1973) and von Stechow’s (1989) notion of a “structured proposition”.) It will be convenient to refer such partitions of the information in the proposition as the “information structure” of an utterance.

Open propositions are most simply exemplified as that which is introduced into the discourse context by a Wh-question. Such an entity can be thought of as a proposition with a “hole”, which the discourse must “fill in”. So for example the question in 1, *What does Mary prefer?* introduces an open proposition which we might informally write as follows:

(2) Mary prefers ...

More formally, it is natural to think of open propositions as functional abstractions, as Jackendoff 1972 and Sag 1976 pointed out, and to write them

---

<sup>1</sup>The intuition of structure imposed by intonation is very compelling. A common initial problem in teaching formal syntax is to persuade students that this is NOT the notion of structure to which they are to attend.

using the notation of the  $\lambda$ -calculus, in which the place of “...” is taken by a variable, whose scope is defined using the operator  $\lambda$ :

(3)  $\lambda x [(prefer' x) mary']$

(Primes indicate interpretations whose detailed semantics is of no direct concern here. The  $\lambda$  notation will be used from time to time below to identify interpretations. The reader may safely ignore this aspect of the notation, if they are willing to take the semantics on trust.) When this function or concept is supplied with an argument, say *corduroy'*, it yields a proposition, with the same function-argument relations as the canonical sentence:

(4)  $(prefer' corduroy') mary'$

It is the presence of the open proposition 3 in the context that makes the intonation contour in 1 felicitous. (Of course, it is not claimed that the presence of the open proposition uniquely determines this response, nor that its explicit prior utterance or mention is necessary for interpreting the response. There is also no claim that intonation contour determines the corresponding open proposition uniquely. We return to this point in section 3.4 below.) The relation of intonational structure to information structure, first proposed by Halliday (1967a), has recently been enshrined by Selkirk (1984:286) in “The Sense Unit Condition” on intonational constituency, which says in essence that intonational constituents must have coherent translations at information structure.

Many authorities, such as Chomsky (1970), Jackendoff (1972), Kaisse (1985), and Cooper and Cooper (1980), have continued to argue that intonation can be driven directly from surface structure. However, the apparent complexities of these proposals when faced with examples like the above have led many others such as Liberman (1975), Goldsmith (1976), Pierrehumbert (1980), Selkirk (1984), and Nespor and Vogel (1986) to postulate an autonomous level of “intonational 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, 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 realised that the two levels

of structure must communicate, because of the presence of certain focussing constructions and operators, such as the English topicalisation construction, or the focussing particle *only*, exemplified in the following sentence:

(5) John introduced only BILL to Sue

Such constructions and particles, which have recently been discussed by Rooth (1985) and von Stechow (1989), have effects in both domains. These observations seem to suggest a theoretical architecture shown in Figure 1 for these components of grammar.

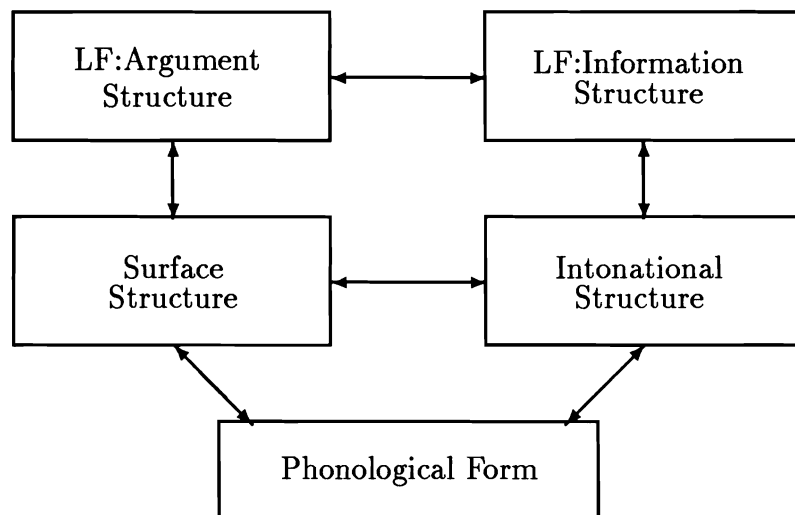


FIGURE 1. Architecture of a Standard Prosody.

Such a theoretical architecture offers a view of sentence structure as having an “autosegmental” topology which Halle in recent lectures has likened to that of a spiral-bound notebook. This notebook has phonetic segments arranged along the spine, and different autonomous levels of structure – prosodic, syntactic, and others – written on different leaves of the notebook, each of which may make reference to descriptions on other pages. As Zwicky and Pullum have pointed out (1987:4), such theories are potentially very unconstrained, in the absence of a principled statement as to 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.

A strong hint that such a simplification might be possible is provided by the observation that the syntax of all natural languages includes 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 that result from coördinate constructions. The latter constructions are doubly interesting, because they and certain other sentence-fragmenting constructions such as parentheticals interact very strongly with intonation, and on occasion make 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 1990):

(6) \*(Harry likes the NUTS) (and bolts APPROACH)

Coördinate constructions currently represent one of the least adequately accounted for, and hence most controversial, problem areas in the study of formal natural language grammar. It is therefore tempting to think that the conspiracy between syntax and prosody noted above might point to a unified notion of structure, somewhat different from traditional surface constituency, which will account for both intonational and coördinate structure.

#### COMBINATORY GRAMMARS.

2. Combinatory Categorical Grammar (hereafter “CCG”, cf. Ades & Steedman, 1982; Steedman, 1987; Szabolcsi 1987a,b) is an extension of Categorical Grammar (hereafter “CG”, cf. Ajdukiewicz 1935; Bar-Hillel 1953). That is to say that elements like verbs are associated with a syntactic “category” which identifies them as FUNCTIONS, and specifies the type and directionality of their argument(s) and the type of their result. This paper uses a notation in which the argument or domain category always appears to the right of the slash, and the result or range category to the left. A forward



slash / means that the argument in question must appear on the right, while a backward slash \ means that the argument must appear on the left.<sup>2</sup>

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

The category  $(S \backslash NP) / NP$  could be regarded as both a syntactic and a semantic object, as in the unification-based categorial grammars of Karttunen 1986, Uszkoreit 1986, Wittenburg 1986, and Zeevat et al. 1986. (See Steedman 1990a for an explicit expression of combinatory categorial grammars in unification-based terms, uniting syntax and semantics in this way.) However, it will be convenient for present purposes to separate the semantic and syntactic types in the notation. In this paper, an expression identifying the translation of a category appears to its right, separated by a colon. It is of course the translation which determines the grammatical or functional role of the first argument to be that of the object, and the second to be the subject.

Such functions can combine with arguments of the appropriate type and position by rules of functional application, written as follows:

(8) The Functional Application Rules:

- a.  $X/Y : F \quad Y : y \quad \Rightarrow \quad X : Fy \quad (>)$
- b.  $Y : y \quad X \backslash Y : F \quad \Rightarrow \quad X : Fy \quad (<)$

Such rules are both syntactically and semantically rules of functional application, as is indicated by the interpretations that appear to the right of the colon for each category.<sup>3</sup> They allow derivations like the following:

(9) 

Mary	prefers	corduroy	
----	-----	-----	
NP	$(S \backslash NP) / NP$	NP	
	----->		
	$S \backslash NP$		
	-----<		
	S		

---

<sup>2</sup>The reader is warned that theories deriving from the work of Lambek 1958 use a different convention – cf. Moortgat 1988b.

<sup>3</sup>Again, it is helpful for expository purposes to explicitly identify the semantics, but the semantic annotations are, strictly speaking, redundant, since the categories themselves can be regarded as both syntactic and semantic objects.

Coördination might be included in CG via the following rule, allowing any constituents of like type, including functions, to form a single constituent of the same type, and thereby to take part in derivations exactly analogous to the above:<sup>4</sup>

(11) I loath and detest corduroy

NP (S\NP)/NP conj (S\NP)/NP NP

(S\NP)/NP

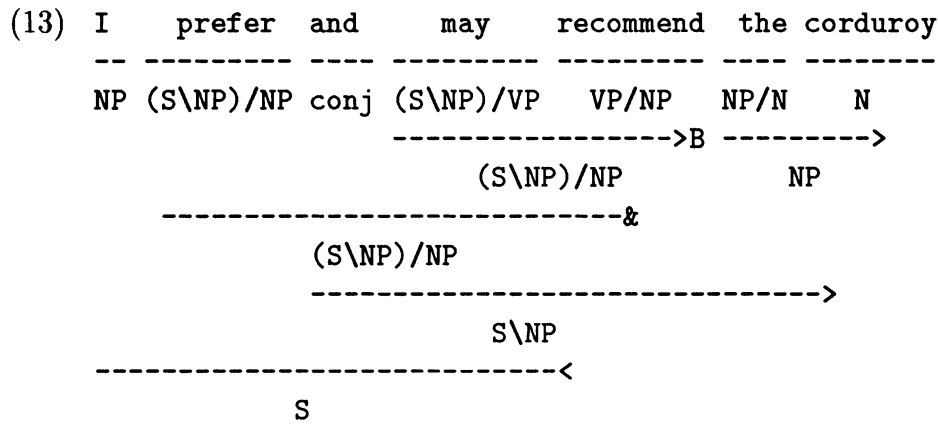
S\NP

S

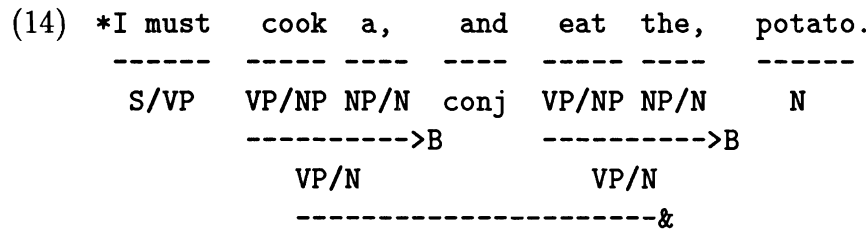
(12) Forward Composition:  
 $X/Y : F \quad Y/Z : G \Rightarrow X/Z : \lambda x [F(Gx)]$

<sup>4</sup>Such a rule is in fact a simplification. Steedman 1990a presents an alternative which captures the fact that conjunctions in English are proclitic, and associate to the rightmost conjunct.

of the functions that it applies to, a detail that is made explicit in the semantic annotations that appear to the right of the colons:<sup>5</sup> Thus sentences like *I prefer, and may recommend, the corduroy* can be accepted, via the following composition of two verbs (indexed as  $>B$ , following Curry's use of the identifier **B** for the composition combinator) to yield a composite of the same category as a transitive verb. Crucially, composition also yields the appropriate interpretation, assuming that a semantics is also provided for the coördination rule.



The forward composition rule will potentially allow certain non-conjoinable sequences to compose and coördinate. For example, since determiners are  $NP/N$ , the following derivation is potentially allowed, by the composition of the transitive verbs with the determiners:



The ways in which such examples might be excluded whilst still allowing examples like the following are discussed briefly in the earlier papers, and

<sup>5</sup>Again, the explicit identification of the semantics of this and all subsequent rules can be ignored by all but the specialist. Such specialists will realise that the notation of the  $\lambda$ -calculus is only used for expository clarity. The categories themselves are complete syntactic and semantic entities. It is functional composition itself, embodied in the rule relating the syntactic categories, that is the primitive of the theory, not the  $\lambda$  operator.

more extensively by Szabolcsi & Zwarts, 1990, where it is suggested that the solution involves semantics, as well as syntax.

(15) I will cook two, and eat three, potatoes.

It is interesting in the present context to note that the restriction, whatever its origin, also shows up in prosody. Such strings as *cook a* are hard to make into intonational phrases, for two presumably related reasons. The first is that it is hard to establish the appropriate open proposition in the context. (For example, you cannot do so with a *Wh*-question). The second is that such elements are not normally stressed, as Oehrle (1988) and Moortgat (1988a,b) have pointed out, and cannot bear a pitch accent or a boundary tone.

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 coördinations like *I dislike, and Mary prefers, the corduroy*. They too have an invariant compositional semantics corresponding to the combinator known to (some) combinatory logicians as **T**, which ensures that the result has an appropriate interpretation. This semantics is as usual made explicit in the rule, via the annotations to the right of the colons.<sup>6</sup> For example, the following rule, indexed >T, allows the conjuncts to form as below (again, the remainder of the derivation is omitted):

(16) Subject Type-raising:

$$NP : x \Rightarrow S/(S \backslash NP) : \lambda P [Px]$$

(17)

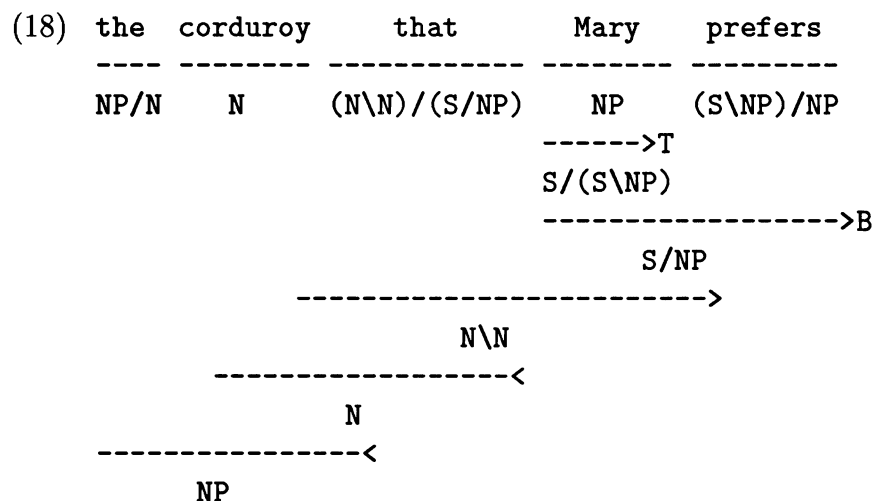
I	dislike	and	Mary	prefers	...
NP	(S\NP)/NP	conj	NP	(S\NP)/NP	
----->T			----->T		
S/(S\NP)			S/(S\NP)		
----->B			----->B		
S/NP			S/NP		
-----&					
S/NP					

---

<sup>6</sup>As usual, this detail can be safely be taken on trust by the non-specialist, and as usual the specialist will realise that its inclusion is purely mnemonic. The **T** combinator is also known as **C\***.

The introduction of type-raising may appear at first glance to be a very arbitrary move, lacking in linguistic motivation. In fact, the contrary is the case. The operation of turning a subject NP into something that will only combine with a predicate is simply the traditional effect of nominative case. We shall see below that this analogy between English NPs and other argument categories with cased categories in a language like Latin is quite general.

The earlier papers show that the addition of type-raising and composition to the theory of grammar immediately provides an account of leftward extractions in relative clauses, on the further assumption that relative pronouns bear a lexical category  $(N \setminus N)/(S/NP)$  – a function from fragments like *Mary prefers* into noun modifiers which is itself closely related to a type-raised category:



It should be similarly obvious that the theory immediately predicts that leftward and rightward extraction will be unbounded:<sup>7</sup>

- (19) a. I think that Mary prefers, and I know that you dislike, corduroy.  
b. The corduroy which I think that Mary prefers.

The subject type-raising rule is a special case of a more general rule which can be written as follows:

---

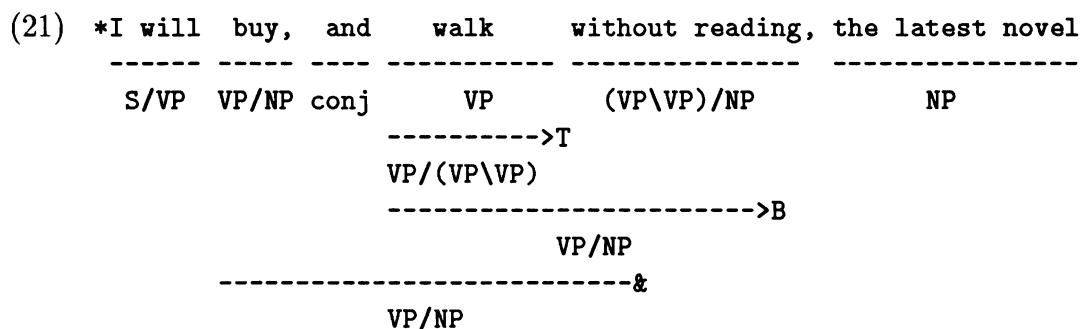
<sup>7</sup>See the earlier papers and Szabolsci, 1987a,b for details, including remarks concerning ECP and Coördinate Structure constraints, and on pied piping.

(20) Forward Type-Raising:

$$X : x \Rightarrow T/(T \backslash X) : \lambda P [Py]$$

The symbol  $T$  is a special category that can for the moment be thought of as matching any category that the grammar permits.

As in the case of the forward composition rule 12, such a free type-raising rule threatens to overgeneralise. Most obviously, it must be restricted to “basic” types, like  $NP$  and  $PP$ , for if we allowed it to apply to categories resulting from raising, we would engender an infinite regress in the rules, and lose decidability. Even when restricted in this way, it potentially permits VPs to raise over adjunct categories, to allow adjunct island violations like the following:



The prevention of such overgeneralisations is again discussed in the earlier papers, which argue that a natural way of limiting the rule is to make type-raising subject to limitations arising ultimately from semantics and general knowledge, which forbid verbs like *walk* from becoming in effect subcategorised for modifiers like *without reading “Ulysses”*. Such a move would explain the notoriously ambiguous status of adverbials as between arguments and modifiers (Cf. McConnell-Ginet 1982). It would also in principle permit certain marginally acceptable related examples discussed by Chomsky (1982:72), such as the following, on the assumption that phrasal verbs like *go to England* can acquire arguments like *without reading Fodor’s Guide* in this way.

(22) ? Which guidebook did you go to England without reading.

Such type-raising could in principle be done “off-line” in the lexicon.

At this point, it is natural to ask what further rules are permitted by the theory, and whether the degrees of freedom that they imply are required

elsewhere in the grammar of English and other languages. This question has been discussed in Steedman 1987 and 1990a, where it is argued that the possible rules are limited by three Principles called Adjacency, Directional Consistency, and Directional Inheritance. The first of these simply amounts to the assumption that purely local combinatory rules, as opposed to long-range rules of movement, abstraction, or indexing over variables, will in fact do the job. The second principle, that of Directional Consistency, prohibits combinatory rules which contradict the directional specifications of the lexicon, such as the following version of functional application:

$$(23) \quad * \quad X \backslash Y : F \quad Y : y \Rightarrow X : Fy$$

Finally, the Principle of Directional Inheritance forbids rules which change the directionality of an argument (such as  $Z$  in the following version of functional composition) from the left-hand side to the right hand side of the rule:

$$(24) \quad * \quad X/Y : F \quad Y/Z : G \Rightarrow X \backslash Z : \lambda x [F(Gx)]$$

It is claimed in the earlier papers that these three principles allow all and only the rules that are required to capture a wide range of generalisations concerning long-range dependency and coördination in a number of languages. In particular, it is claimed that suitably restricted versions of all four possible rules of functional composition are implicated in the grammar of English. For example, the mirror-image rules of functional composition and type-raising to those given at 12 and 20 are permitted under these principles, and were shown by Dowty to give rise to the English “Left Node Raising” construction in the following kind of derivation, to which we shall return below:

$$\begin{array}{ccccccc}
 (25) & \text{give} & & \text{George} & & \text{a book} & \text{and} & & \text{Martha} & & \text{a record} \\
 & \text{-----} & & \text{-----} & & \text{-----} & \text{-----} & & \text{-----} & & \text{-----} \\
 & (VP/NP)/NP & & (VP/NP) \backslash ((VP/NP)/NP) & & VP \backslash (VP/NP) & \text{conj} & & (VP/NP) \backslash ((VP/NP)/NP) & & VP \backslash (VP/NP) \\
 & & & \text{-----} & & \text{-----} & & & \text{-----} & & \text{-----} \\
 & & & & & & & & & & & \text{-----} & \text{<B} \\
 & & & & & & & & & & & & \text{VP} \backslash ((VP/NP)/NP) & & \text{VP} \backslash ((VP/NP)/NP) \\
 & & & & & & & & & & & & \text{-----} & \text{&} \\
 & & & & & & & & & & & & & \text{VP} \backslash ((VP/NP)/NP) \\
 & & & & & & & & & & & & \text{-----} & \text{<} \\
 & & & & & & & & & & & & & \text{VP}
 \end{array}$$

The rules in question are the following:

- (26) Backward Composition (<B):  

$$Y \backslash Z : G \quad X \backslash Y : F \Rightarrow X \backslash Z : \lambda x F(Gx)$$
- (27) Backward Type-raising (<T):  

$$X : x \Rightarrow T \backslash (T/X) : \lambda P P x$$

The backward type-raising rule allows the indirect object to turn into a function from ditransitive to transitive verbs, and the direct object to turn into a function from transitive to intransitive verbs. (Not surprisingly, given the earlier analogy between subject type raising and nominative case, these categories correspond exactly to the categories that dative and accusative NPs would bear in a cased SVO language.) The three principles, as expanded in the earlier papers, will not permit such sequences of verb complements to combine in any other order, given the English lexicon and the fact that it is a configurational language. The earlier papers generalise this observation to show that certain well-known universal laws concerning the dependency of the direction of gapping upon basic clause constituent order are necessary corollaries of the fact that the possible combinatory rules are limited by the three principles of Adjacency, Consistency and Inheritance. (For example, I argue in Steedman 1990a that it is a corollary of this theory, rather than a stipulation, that English and all other SVO languages must gap on the right, like VSO languages, not on the left, like SOV ones, cf. Ross 1970).<sup>8</sup> As in the case of earlier non-standard constituents, the result is guaranteed by the combinatory semantics to yield appropriate function/argument relations.

### SURFACE STRUCTURE AND INTONATIONAL STRUCTURE

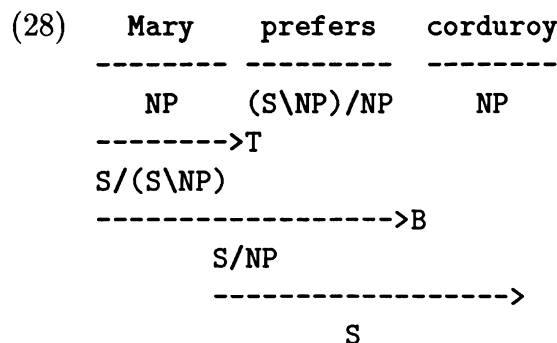
**3.** It will be clear from the above examples that CCG embodies a very strong adherence to what has been termed the “Constituent Condition” on rules. This condition simply says that grammatical rules should be limited to operations upon grammatical constituents. The assumption is widespread in work within the generative tradition, although it is rarely pursued to this extreme. For according to the combinatory theory, conjoinable strings like *Mary prefers* and even *a policeman a flower* correspond to constituents in their own right, without deletion or “gaps”. It follows, according to this

---

<sup>8</sup>See the earlier paper for some discussion of the well-known exceptions to this generalisation that arise in the SOV languages, due to the involvement of case and consequent less strict order.



view, that they must also be possible constituents of simple non-coördinate sentences like *give a policeman a flower*, and example 9, *Mary prefers corduroy*, as well. It follows that such sentences must have SEVERAL surface structures, corresponding to different sequences of composition, type-raising and application.<sup>9</sup> For example, the following derivation is also allowed for the latter sentence, as well as the traditional derivation given earlier at 9:



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.

This proliferation of surface analyses creates obvious problems for the parsing of written text, because it compounds the already grave problems of local and global ambiguity in parsing by introducing numerous semantically equivalent potential derivations. The problem is acute: while it clearly does not matter which member of any equivalence class the parser finds, it does matter that it find SOME member of EVERY semantically distinct class of analyses. The danger is that the entire forest of possible analyses will have to be examined in order to ensure that all semantically distinct analyses have been found. This problem has been referred to (misleadingly, as will become apparent) as the problem of “spurious” ambiguity.<sup>10</sup> However, I shall argue not only that these semantically equivalent derivations are functionally

---

<sup>9</sup>An entirely unconstrained CCG using all the rules allowed by the Principles identified above, would allow any bracketing on a sentence. However, it will be recalled that the CCG for any given language will restrict such rules as to the categories to which they apply, or even exclude some rules entirely.

<sup>10</sup>Cf. Wittenberg 1986, and Hepple & Morrill 1989. The interested reader is referred to Weir & Joshi 1988 and Vijay-Shankar & Weir 1990 for results on automata-theoretic power and a polynomial worst-case parsing complexity result.

significant, in that they convey distinctions of discourse information, but also that the extra structural ambiguity that they engender is to some extent resolved by intonation in spoken language.

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:

- (29) 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 coördinated – for example:

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

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

- (31) 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 29 and 31a above 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).<sup>11</sup> A stronger example emerges from comparison of the following examples, in which the string *three mathematicians* is as hard to make an intonational phrase as it is to coördinate. (The unacceptability of (a) is also used by Selkirk as evidence for the Sense Unit Condition):<sup>12</sup>

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

<sup>11</sup>See Jacobson 1990 for a discussion of this and related constructions in categorial terms.

<sup>12</sup>Dwight Bolinger and Julia Hirschberg have at least half convinced me (p.c.) that there are circumstances under which one or the other is allowed. However, the only claim I make is that if such circumstances exist, they are such as to make BOTH more felicitous.

It is irrelevant to the present purpose to ask HOW sentences like (b) might be excluded, or even to ask whether what is wrong with them is a matter of syntax, semantics or pragmatics.<sup>13</sup> 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 coördinate structure. Non-coördinate 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, coördinate sentences, like relative clauses, have fewer equivalent analyses, because only analyses which 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 coördinate. Second, of all the intonational tunes that distinguish alternative prosodic constituencies in non-coördinate sentences, we predict that only the ones which are consistent with the constituents demanded by the coördination rule will be allowed in coördinate sentences. Intonation contours which are appropriate to the alternative constituencies are syntactically ruled out. So for example, there are many prosodic constituencies for the example 29, *I want to begin to try to write a play*, realised by a variety of intonational contours. However, there are many fewer possible intonation contours for the following coördinate sentence, and they seem intuitively to be closely related to the ones which impose the corresponding bracketing 29e in the simpler sentence:

- (33) I want to begin to try to write, and you hope to produce,  
       a musical based on the life of Denis Thatcher.

Observations like the above make it seem likely that in spoken utterance, 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 focus among the

---

<sup>13</sup>As in the earlier discussion of adjunct island constraint violations, the squishiness of this constraint suggests that its source lies in the lexicon, and ultimately in lexical semantics.

concepts and open propositions 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 phonology and syntax are one system, and that speech processing and parsing can be merged into a single unitary process.<sup>14</sup>

This and the next section of the paper will show that the combinatory rules of syntax that have been proposed in order to explain coördination and unbounded dependency in English do indeed induce surface structures that are isomorphic to the structures that have been proposed by Selkirk 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 also 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.

**3.1. TWO INTONATION CONTOURS AND THEIR FUNCTIONS.** I shall use a notation for intonation contours which is based on the theory of Pierrehumbert 1980, itself a development of proposals by Liberman 1975 and Goldsmith 1976. The version used here is roughly as presented in recent work by Selkirk 1984, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1989, and Pierrehumbert & Hirschberg 1990, although it will become clear below that I have departed from this theory 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.

I follow Pierrehumbert in assuming that intonation contours can be described in terms of two abstract pitch levels, and three types of tones. There are two phrasal tones, written H and L, denoting high or low “simple” tones

---

<sup>14</sup>The conspiracy between prosodic phenomena and the notion of constituency that emerges from related grammars including associative operations is noted by Moortgat (1987, 1988b), and by Oehrle (1988), and in earlier work in the combinatory framework (Steedman 1985:540). Related points concerning “bracketing paradoxes” in morphology are made by Moortgat (1985, 1988b), and Hoeksema (1984, 1988). See also the categorial analyses of Wheeler (1981, 1988), and Schmerling (1981).

— that is, level functions of pitch against time. There are also two boundary tones, written H% and L%, denoting an intonational phrase-final rise or fall. There are a number of types of pitch accent, or tones corresponding to the stressed syllable(s) in the prosodic phrase. Of Pierrehumberts six pitch accent tones, I shall only consider two, the H\* accent and the L+H\*.<sup>15</sup> The phonetic or acoustic realisation 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 which 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, which is the other melody considered below. (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 phrasal tone, and optionally a boundary tone. In recent versions of the theory, Pierrehumbert and her colleagues distinguish two distinct levels of such prosodic phrases. They are the intonational phrase proper, and the “intermediate phrase”. Both end in a phrasal tone, but only intonational phrases have additional boundary tones H% and L%. Intermediate phrases are bounded on the right by their phrasal tone alone, and do not appear to be acoustically characterised in  $F_0$  by the same kind of final rise or fall that is characteristic of true intonational phrases. The distinction does not play an active role in the present account, but I shall follow the more recent notation for prosodic phrase boundaries in the examples, without further comment on the distinction.<sup>16</sup>

---

<sup>15</sup>However, I believe that the account given below generalises to the other pitch accent tones.

<sup>16</sup>I have talked here of the phrasal tone as occurring at the right hand end of the prosodic phrase, and as marking the right-hand boundary, together with the boundary tone. While this is consistent with Pierrehumbert’s account, and notationally convenient for present purposes, it should be remarked that the position and nature of the phrasal tone seems to be one of the more controversial details of her theory (cf. Pierrehumbert & Beckman 1989:236-237). The influence of, say, an L phrasal tone on an H\* pitch accent is apparent immediately, no matter how distant the righthand boundary is. (Pierrehumbert and Beckmann suggest that this influence may be apparent by the end of the word bearing the nuclear pitch accent.) Indeed, in the framework of the British school, the event cor-

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 phrasal tone, the fundamental frequency is merely interpolated. In Pierrehumbert's notation, such substrings therefore bear no indication of abstract tone whatsoever.<sup>17</sup>

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, are essentially invariant. If the constituent is very short – say, a monosyllabic nounphrase – then the whole intonational contour may be squeezed onto 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 for example the prosody of the sentence *Fred ate the beans* in the following pair of discourse settings, which are adapted from Jackendoff 1972:260.<sup>18</sup> (To help the reader, the prosodic phrase boundaries that are implicit in Pierrehumbert's notation are indicated by parentheses in the string):

---

responding to Pierrehumbert's phrasal tone is considered to be part of the pitch accent, rather than part of the boundary event. While nothing in the present account hinges on this detail, it is a point at which the notation might change in future.

<sup>17</sup>Neither Pierrehumbert's theory nor its combinatory expression below should be taken as implying that the null tone corresponds to an absence of the fundamental frequency. Nor does either version imply that an element bearing the null tone is always realised with the same intonation contour. They merely imply that the intonation is independently specified. It follows that the null tone may carry information about what pitch accents and other tones are downstream of it. It also follows that a processor might make use of this information.

<sup>18</sup>Jackendoff's choice of lexical items in this example is not entirely helpful. Words including post-tonic syllables make the different intonational tunes easier to perceive. Those who have access to pitch-tracking facilities will find materials consisting entirely of vowels nasals liquids and glides, such as *Lenora will marry Emanuel* give the best results. However, I have retained Jackendoff's example throughout the extended discussion below, both to facilitate comparison with his account, and because it keeps the derivations small and the contexts simple. The reader is encouraged to make their own mental substitution of lexical items in any cases they find doubtful.

(34) Q: Well, what about the BEANS? Who ate THEM?

A: (FRED) (ate the BEA-NS).  
           H\*L                      L+H\*LH%

(35) Q: Well, what about FRED? What did HE eat?

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

Each answer is uttered as a sequence of two prosodic phrases delimited by the two tunes that appear in Pierrehumbert's notation as L+H\* LH% and H\* LL%. (The difference between H\* LL% and H\* L is not relevant here). I base these annotations on Pierrehumbert & Hirschberg's 1990:296 discussion of this example.<sup>19</sup> The fact that these two tunes are spread across different sections of the sentence has the effect of partitioning the tune differently in the two cases. In the first case, the sentence is divided into a subject and a predicate, a division which happens to coincide with traditional surface structure. In the second, the division is orthogonal to traditional surface structure.

The two tunes do more than delimit the intonational constituents of the two sentences. As Pierrehumbert & Hirschberg 1990 point out, , they also identify discourse functions. It seems as if at least one function of the tune L+H\* LH% is to mark a constituent whose translation corresponds to the open proposition established by the question. It may thus be thought of as marking WHAT THE UTTERANCE IS ABOUT. In the first case, what it is about is the open proposition  $\lambda x \text{ eat}' x \text{ beans}'$  – roughly, *eating the beans*. In the second case it is about the open proposition  $\lambda x \text{ eat}' \text{ fred}' x$  – roughly, *Fred eating*.<sup>20</sup> However, the tune does something more. The presence of a pitch accent also marks some or all of the open proposition as emphasised or contrasted with something mentioned or regarded by the speaker as implicated by the previous discourse and/or context. It marks, as it were, THE

---

<sup>19</sup>Again, nothing in the present account hinges on these precise identifications, apart from the basic claim that there is a distinction between the two entire tunes notated here as L+H\* LH%, and H\* L or H\* LL%.

<sup>20</sup>An alternative prosody, in which the same tune is confined to *Fred*, 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 or theme of this discourse is *Fred* or *What Fred ate*. It is accepted by the present rules, and is discussed at example 52 below.

INTERESTING PART of the open proposition. In the former case, the beans are marked in this way as standing in contrast to some other comestible. In the second case, the entire open proposition is contrasted with some earlier one.

The other phrasal tune, H\* LL%, conveys a quite different function. It can be thought of as marking WHAT THE SPEAKER HAS TO SAY about what the utterance is about – that is, the argument that satisfies the open proposition. In the first case, this argument is *Fred*. In the second, it is *the beans*. Again, the position of the pitch accent marks the interesting part of the argument.

These two tunes, L+H\* LH% and H\* LL%, thus distinguish two informational units within the sentence. Their functional roles of defining “what the utterance is about”, versus “what the speaker says about it”, have been discussed under a bewildering variety of nomenclatures, most of them unformalised. The functions in question are distinguished by Chomsky (1970) and Jackendoff (1972) as “presupposition” and “focus”, by Hajičová and Sgall (1987, 1988) as “topic” and “focus”, and in older traditions as “topic” vs. “comment”, and “theme” vs. “rheme”.

There are problems with all of these terminologies. While the theory presented here is in some ways close to that of Chomsky and Jackendoff, the term presupposition has engendered considerable confusion with the semantic and pragmatic notions of the same name, as Schmerling (1976) and Rochemont and Culicover (1990:21-22) have pointed out. Consider for example the presupposition (in this latter sense) that is associated with the relative clause in cleft constructions, as in the following sentence:

(36) It was Harry who taught me how to Tango.

If we compare the function of such a sentence used as an answer to the question “I know that Mary taught you the Lambada. But who taught you how to Tango?” with its use as an answer to “Why are you so fond of Harry?”, then it is clear, as Delin 1989 has pointed out, that such a presupposition may either be assumed, or used to supply novel information about Harry. In this case, the presupposition will typically be “accommodated” by the hearer – that is, will cause their discourse model to be updated to match the presupposition. Under these circumstances, the presupposition will typically be uttered with an H\* LL% intonation, marking novel information.

There are also problems with the term “focus”, as it is used by Chomsky and Jackendoff, and by Hajičová and Sgall. Their respective dichotomies



obscure the fact that intonation contour, and in particular the position of the pitch accent, is also used to further divide the two major information units into background and foreground. For example, in 34, the pitch accent occurs on the word *beans*, while the rest of the predicate bears the null tone. Such an utterance, which contrasts eating beans with eating something else, is clearly pragmatically different from the corresponding utterance in which the L+H\* LH% tune is spread across the entire verb-phrase, and the contrast is between eating beans and doing something else to them. The trouble is that this distinction WITHIN the major informational units is also often referred to as “focus”, particularly by phonologists (and also by Lyons (1977:500-511) – see below). The problem is compounded by the fact that it is much harder to give an intuitive description for this latter function than for the primary dichotomy, since accent can be used to emphasise things like syllables, as well as more obviously referential entities, as in utterances like *My name is FishMAN, not FishFACE*

Halliday (1967b, 1990:160-164, 1976:325-326), following the Prague linguists, has proposed the terminology that comes closest to fulfilling our needs. His system distinguishes two independent dimensions of informational structure. The primary informational units are called the “theme” (which he himself identifies as “what the utterance is about”) and the “rheme” (which was earlier identified as “what the speaker says about the theme”). For present purposes, we shall assume that these notions generalise to non-standard constituents like *Mary prefers*, as well as the categories explicitly considered by Halliday. Within theme and rheme alike, a further distinction is drawn between what he calls “new” information (which a speaker makes salient by the use of accent) and “given” information (which a speaker does not make salient). This second dimension is clearly the one considered in the last paragraph, which is dependent upon the position of the pitch accent within the theme or rheme (cf. Halliday 1970, p. 163, ex. 33).

A minor problem in applying Halliday’s analysis to the combinatory theory is that he repeatedly insists that in English the theme is sentence initial, and precedes the rheme. Inspection of the examples 34 and 35 shows that this condition jeopardises the straightforward association of the terms theme and rheme with the two intonational tunes and the two discourse functions that we have identified. It is probably true that there is a strong tendency in written English text for the theme to precede the rheme. However, it is surely no more than a statistical tendency that follows from the nature of

the functions concerned. In the spoken language, where intonation is available to mark discourse function explicitly, this tendency seems to be much less marked. In fact, many who have attempted to apply Halliday's insights to discourse function, such as Lyons (1977:509), and especially to intonation, such as Bolinger, (1989:389) have regarded this detail as a minor aberration, and have used the terms *theme* and *rheme* to refer to the functional categories exemplified above, without regard to linear order.<sup>21</sup> I shall follow Bolinger and Lyons in using these terms to describe the function of the intonational/intermediate phrase. I shall follow what I take to be Halliday's original intention in using the term "new" to refer to that part of the theme or rheme which is made salient by accent, and "given" for the rest. This usage is illustrated in the following example:

(37) 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.

**3.2. CONSTITUENCY AND INTONATION.** The L+H\* LH% intonational melody in example 35 belongs to a phrase *Fred ate ...* which corresponds under the combinatory theory of grammar to a grammatical constituent. What is more, this constituent comes supplied with a translation equivalent to the open proposition  $\lambda x[(eat' x) fred']$ . This translation corresponds to the open proposition introduced by the question in 35, *What did Fred eat?* The theory thus offers a way to assign contours like L+H\* LH% to such novel constituents, entirely under the control of independently motivated rules of grammar.

---

<sup>21</sup>Those who have attempted to adhere more closely to the letter of Halliday's writings have tended to abandon the notion of theme and rheme as discourse categories, and to make given and new do all the work. However, such a tactic seems to lose one of the most distinctive and useful contributions of this theory.

The combinatory apparatus itself offers an extremely natural way to do this. We will begin by defining the two pitch accents as functions of the following types:

$$(38) \quad \begin{aligned} L+H^* &:= Theme/Bh \\ H^* &:= (Utterance|Theme)/Bl \end{aligned}$$

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  $Utterance|Theme$  from themes into utterances.<sup>22</sup> The vertical slash indicates a function which can combine with its argument in either direction. It will be convenient to refer to this category as “the rheme category”.

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

$$(39) \quad \begin{aligned} LH\% &:= Bh \\ LL\% &:= Bl \\ L &:= Bl \end{aligned}$$

(As before, we ignore for present purposes the distinction between intermediate- and intonational- phrase boundaries.)<sup>23</sup> 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  $\emptyset$ :

$$(40) \quad \emptyset := X/X$$

---

<sup>22</sup>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.

<sup>23</sup>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).

$X$  is a category that can match any category, importantly including  $X/X$ .<sup>24</sup> 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:

(41) Forward Prosodic Functional Composition:

$$X/Y \ Y/Z \Rightarrow X/Z$$

where  $Y \in \{Bh, Bl\}$

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.<sup>25</sup>

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

(42) 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).

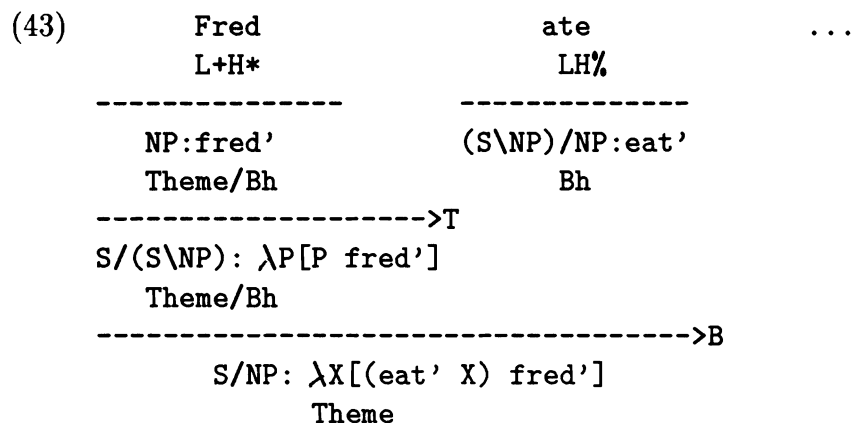
This principle has the sole effect of excluding certain derivations for spoken utterances that would be allowed for the equivalent written sentences. For example, consider the derivations that this principle permits for example 35 above. The rule of forward composition is allowed to apply to the words *Fred ate ...*, because *Fred* is prosodically a function, and *ate* is its argument.<sup>26</sup>

---

<sup>24</sup>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.

<sup>25</sup>Two null tones of type  $X/X$  can therefore only combine by application, not composition.

<sup>26</sup>Again, the semantic annotations simply identify interpretations that are implicit in the categories themselves.



It is assumed here that the input to the system is now the speech wave, rather than written text, and that words together with identifications of the tune-fragments that they are uttered with, can be extracted from the speech wave, and provide the input to the rules considered here.<sup>27</sup>

Given the category  $X/X$  for elements not bearing pitch accents or boundaries, the Prosodic Constituent Condition 42 similarly allows the theme tune L+H\* LH% to “spread” across any sequence that can be assembled by repeated applications of the syntactic forward composition rule (including ones crossing  $S$  boundaries). For example, if the reply to the same question *What did Fred eat?* is *FRED must have been eating the BEANS*, then the tune will typically be spread over *Fred must have been eating ....* Such a prosodic constituent is accepted as in the following derivation, in which much of the syntactic and semantic detail has been suppressed in the interests of brevity:

---

<sup>27</sup>The way in which this might be done is briefly discussed in the concluding sections. There is no assumption here that such processes of word-recognition will segment the speech-wave unambiguously.

(44)

Fred	must	have	been	eating	...
L+H*				LH%	
-----					
NP	(S\NP)/VP	VP/VPen	VPen/VPing	VPing/NP	
Theme/Bh	X/X	X/X	X/X	Bh	
----->T					
Theme/Bh					
----->B					
Theme/Bh					
----->B					
Theme/Bh					
----->B					
Theme/Bh					
----->B					
Theme					

Of course, other derivations for this substring are permitted.

In both cases, the presence of a boundary on the main verb completes the prosodic constituent of type *Theme*. In both cases, the interpretation of this category is a function from object interpretations into interpretations of propositions.

The derivation of 35 continues to completion as follows:

(45)	Fred	ate	the	beans
	L+H*	LH%		H* LL%
	-----	-----	-----	-----
	NP:fred'	(S\NP)/NP:eat'	NP/N: the'	N:beans'
	Theme/Bh	Bh	X/X	Utterance Theme
	----->T			
	S/(S\NP): $\lambda P$ [P fred']			
	Theme/Bh			
	----->B			
	S/NP: $\lambda X$ [(eat' X) fred']		NP:the' beans'	
	Theme		Utterance Theme	
	----->			
	S: eat' (the' beans') fred'			
	Utterance			

First the null tone  $X/X$  combines with the noun, which carries the rheme

category *Utterance|Theme* resulting from the combination of H\* and LL% on the word *beans*. (This latter combination is not shown but is no different to any other). This category can combine by backward prosodic functional application with the theme, to yield a complete utterance, whose interpretation embodies the appropriate function-argument relations.

More importantly, the division under the above derivation into open proposition marked as theme and argument marked as rheme is contextually appropriate. An alternative derivation, under which the verb applies to its object, which would yield a pragmatically inappropriate information structure, is correctly excluded by this intonation contour, because a boundary *Bh* cannot combine to its right with a rheme, or with the null tone *X/X*. Repeated application of the composition rule, as in 44, would allow the L+H\* LH% contour to spread further, as in 44 (*FRED must have eaten*) (*the BEANS*). However, the prosodic categories would still only permit one partition of the sentence at the highest level into theme and rheme.

In contrast, the intonation contour on 34 prevents the composition of subject and verb, because the subject is not allowed under the forward prosodic composition rule to combine with the verb. It follows that a derivation parallel to the above (and the formation of the corresponding open proposition) is not allowed. On the other hand, the following derivation of 34 is allowed:

(46)	Fred	ate	the	beans
	H* L			L+H* LH%
	-----	-----	-----	-----
	NP:fred'	(S\NP)/NP:eat'	NP/N: the'	N:beans'
	Utterance Theme	X/X	X/X	Theme
	----->T		----->	
	S/(S\NP):λP[P fred']		NP:the' beans'	
	Utterance Theme		Theme	
		----->		
		S\NP:eat'(the' beans')		
		Theme		
	----->			
	S: eat'(the' beans') fred'			
	Utterance			

Here forward functional application makes *the beans* into a theme, via application of the null tone category *X/X* to *beans* marked as theme. The verb

*ate* which also bears the null tone category can also apply, to yield a VP, also marked as theme. This prosodic category can combine with the subject bearing the rheme category. 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 34.

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. The finer grain information that, within the theme, the position of the pitch accent defines the object *the beans* as new or salient information, while the verb, and the associated concept of eating, is given information, is not made explicit within the present set of rules.<sup>28</sup>

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 left as an exercise, and the symbol *Rheme* is used as an abbreviation for the full category *Utterance|Theme*). The first two yield derivations parallel to 46, 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):

- (47) What about FRED? What did HE do to the beans?

(FRE-ED)	(ATE the beans)
L+H* LH%	H* LL%
<i>Theme</i>	<i>Rheme</i>

- (48) I know who COOKED the beans. But then, who ATE them?

(FRED)	(ATE the beans)
H* L	L+H* LH%
<i>Rheme</i>	<i>Theme</i>

The other two cases considered by Jackendoff yield derivations parallel to 45, in which the fundamental division of the sentence is orthogonal to the traditional subject-predicate structure:

- (49) I know what Fred COOKED. But then, what did he EAT?

(Fred	A-ATE)	(the BEANS)
L+H* LH%		H* LL%
<i>Theme</i>		<i>Rheme</i>

---

<sup>28</sup>We return briefly to this question in the concluding sections.



(50) Well, what about the BEANS? What did Fred do with *them*?

(Fred	ATE)	(the	BEA-NS)
	H* L		L+H* LH%
<i>Rheme</i>		<i>Theme</i>	

In the case of 49 at least, it seems obvious that the open proposition established by the context is indeed the one corresponding to the bracketing. In the case of 50 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 50 is also allowed, as is a parallel derivation for 49:

(51) Well, what about the BEANS? What did Fred do with *THEM*?

(Fred)	((ATE)	(the	BEA-NS))
	H* L		L+H* LH%
$X/X$		<i>Utterance</i>	
<hr style="width: 100%;"/> <i>Utterance</i>			

Since the word *ate* can be treated as an entire rheme, bearing the category *Utterance*|*Theme*, it can combine with *the beans* to yield a prosodic utterance. The subject *Fred* bearing the null tone  $X/X$  then applies to it, to yield a prosodic utterance. This alternative analysis would be easy enough to exclude, via restrictions on  $X$  in the null tone category. However, it seems likely that it should be allowed. If so, the question of the discourse function of the subject *Fred* arises. The most reasonable suggestion would seem to be that it is an unmarked theme, of the kind discussed in section 3.4 below, to which the entire sequence *ate the beans* is the rheme. Such a suggestion would in turn entail replacing the category of prosodic utterance with the rheme category, a suggestion which is reminiscent of Ladd's 1986 proposal for a recursive prosodic structure. However, we will pass over this possibility here.<sup>29</sup>

Two further cases, which are parallel to 34 and 35 but with the H\*L and L+H\* LH% tunes exchanged, are also accepted, again yielding unique, contextually appropriate analyses. The first is the following:

---

<sup>29</sup>See section 4.2 below.

- (52) I know that ALICE read a BOOK.  
 But what about FRED? What did HE do?  
 (FRE-ED) (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 35. As remarked in the footnote to the discussion of that example, this possibility appears to arise from an ambiguity in that context in that example, and may be the contour intended by Jackendoff. However, the converse does not apply: the intonation on the response in 35 is not felicitous in the above context, as the following example shows:<sup>30</sup>

- (53) I know that ALICE read a BOOK.  
 But what about FRED? What did HE do?  
 ?(FRED ate) (the BEANS)  
 L+H\* LH% H\* LL%  
*Theme Rheme*

The final possibility is not considered by Jackendoff, and 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, as discussed in example 56 in the next section, would also improve the example. Its place in the scheme of things will become clearer in the section below on relatives.

- (54) Well, what about the BEANS? What happened to THEM?  
 (FRED ate) (the BEA-NS)  
 H\* L L+H\* LH%  
*Rheme Theme*

It should be noted at this point that the association of tones with words at the lowest level of the derivation does not mean that they are associated with them in the lexicon. The tones are properties of prosodic phrases, whose

---

<sup>30</sup>The multiple pitch accent on the verb phrase in example 57 in the next section is an even more appropriate response, and also serves to distinguish these contexts, since it is not appropriate to 35.

extent and limits they define. It will be clear from these remarks, and from the derivations above, that the phonological categories define an autonomous or “autosegmental” level of intonational structure, in much the same sense of the term introduced by Goldsmith (1976). However, the Prosodic Constituent Condition 42 constitutes the strongest possible constraint on the syntactic and intonational levels or “tiers”. The constraint simply expresses the fact that the structures are isomorphic, and can therefore be considered as annotations to a single structure.

**3.3. PHRASES WITH MULTIPLE PITCH-ACCENTS.** If more than one element of an intonational/intermediate phrase is new information, say for reasons of contrast, then they may all be marked by pitch accents in Pierrehumbert’s terms, the intervening contour being interpolated. A very common pattern is where all pitch accents in such a phrase are of the same type.<sup>31</sup> Phrases of this kind can be brought into the grammar by adding a second “endotypic” prosodic category to the two pitch accents under consideration here, as follows:

$$\begin{aligned}
 (55) \quad L+H^* &:= Theme/Bh \\
 &:= Bh/Bh \\
 H^* &:= (Utterance|Theme)/Bl \\
 &:= Bl/Bl
 \end{aligned}$$

Such a modification will allow derivations for examples like the following, which are as usual given with the kind of context that facilitates the contours in question, and as usual abbreviating the category *Utterance|Theme* as *Rheme*. (The restriction on the forward prosodic composition rule 41 allows pitch accents to prosodically compose with one another, as well as apply.)

$$\begin{array}{rcl}
 (56) \quad \text{Well, what about the beans? What happened to THEM?} & & \\
 \quad (FRED \quad ATE) \quad (the \quad BEA-NS) & & \\
 \quad H^* \quad H^* L \quad L+H^* LH\% & & \\
 \quad Rheme \quad Theme & &
 \end{array}$$

---

<sup>31</sup>Pierrehumbert’s theory also allows them to be different – cf. Selkirk 1984:438, note 29. Such a possibility could be permitted under the combinatory theory by giving pitch accents a more general second category *B/B*, where *B* is any boundary.

(57) Well, what about FRED? What did HE do?

( FRE-ED)	(ATE the BEANS)
L+H* LH%	H* H* LL%
<i>Theme</i>	<i>Rheme</i>

(See examples 52 and 54, above.) The derivations, and the generation of contexts for the two parallel examples with multiple L+H\* pitch accents, are left as an exercise. The remainder of the paper will mostly be concerned with phrases with only a single pitch accent.

**3.4. SENTENCES WITH UNMARKED THEME.** The above variations on Jackendoff's sentence, in which an H\*-based tune marks the rheme, and an L+H\*-based tune marks a theme standing in some sort of contrast to a previous one, has the advantage for present purposes of being easy to provide with relatively unambiguous contexts.

However, such examples should not be taken to imply that ALL information structure is phonetically marked by events like boundary tones. When the theme is not contrasted, but is nonetheless expressed for some reason, it often bears the null tone. In such utterances, we must assume that there are information-structural boundaries which are NOT so marked. For example, consider the following answer to a question about some kind of vegetable, uttered with only an H\* LL% tune on the last word:

(58) (They are a good source of VITAMINS).  
H\* LL%

In Pierrehumbert's terms, such an utterance constitutes a single intonational phrase, since it contains no internal intonational/intermediate phrase boundaries. However, by the same token, the combinatory theory allows a number of different analyses. In fact, such an intonation contour is compatible with ALL the analyses that the unannotated CCG would allow, because the associativity of the category  $X/X$ , which can apply to itself, parallels that of syntactic functional composition. Thus we get the following analyses:

- (59) a. (They are a good source of) (VITAMINS)  
H\* LL%
- b. (They are) (a good source of VITAMINS)  
H\* LL%
- c. (They) (are a good source of VITAMINS)  
H\* LL%

This is as it should be, if we are to preserve the identity between surface structure and information structure. The ambiguity is a genuine one, and such an utterance is indeed compatible with a large number of contextual open propositions. For example, it is a reasonable response to any of the following:

- (60) a. What are legumes a good source of?  
 b. What are legumes?  
 c. What about legumes?

These contexts are close relatives of ones that would induce the same bracketings with more marked intonation contours and explicit prosodic boundaries. The ambiguity of intonation with respect to such distinctions of focus and information is well-known (cf. Chomsky 1970), and it would simply be incorrect not to permit it. However, we can no longer simply identify such information-structural partitions at surface structure with intonational structure as defined by Pierrehumbert, for in the original terms of her theory, there is simply one prosodic analysis of sentence 58, as a single intonational phrase.

At this point, two options are available. The system as it stands embodies a direct equivalence between Pierrehumbert's specification of intonation structure and information structure, implying that such tunes delimit a single rheme, and that such an isolated rheme constitutes a well-formed utterance. Such isolated rhemes undoubtedly occur, as for example when the entire theme is simply omitted, or when an entire clause is marked as new information, via multiple H\* accents, as in the following:<sup>32</sup>

- (61) (YOU are STANDING ON my FOOT)!  
 H\* H\* H\* H\* LL%  
*Rheme*

However, to take this line in the case of 58 would be to do violence to our intuitions concerning information structure. In terms of theme and information, it seems that we miss a generalisation if we do not rather distinguish a number of different partitions of the sentence, just as if these constituents

---

<sup>32</sup>As usual, the symbol *Rheme* is merely a shorthand for the category *Utterance*|*Theme*. There is an implicit assumption here either that this category is an alternative start-symbol to *Utterance* in the phonological tier, or that it can find its argument in the context.

were explicitly delimited by boundary tones. The claim that there are a number of phonetically indistinguishable but informationally distinct analyses of the sentence into two syntactic/phonological constituents seems an extremely natural generalisation of Pierrehumbert’s proposal. Its sole effect is to bring the domains of prosody and its discourse interpretation more exactly into line by distinguishing underlying phonological boundaries from their phonetic realisation. The boundaries represented here by brackets are of the former kind, whereas the tones are the latter.

We include null intermediate phrases in competence grammar by adding the following rule, allowing constituents marked with the null category  $X/X$  to non-deterministically turn into themes.<sup>33</sup>

(62)  $X/X \Rightarrow Theme$

This rule in effect allows the processor to freely postulate a “virtual” intermediate phrase boundary at any constituent boundary where there is the null tone. (Such null themes might also be seen as “phonological phrases”, at a level below the intonational/intermediate phrase – cf. Selkirk, 1984:29. Indeed, since the CCG categories make no explicit reference to phonological levels, they are entirely compatible with frameworks in which discourse functions like theme and rheme are associated with this level, as is proposed by Nespor & Vogel 1986). We return to the question of how the structural ambiguity of such sentences is resolved in processing in a later section. However, a broad hint may be gained from the observation that the null tone is only used in this way when the theme is entirely given – that is, when hearers are in a position to decide the open proposition for themselves.

#### INTONATION IN COMPLEX CONSTRUCTIONS.

4. The number of possible intonational 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 (such as reduced coördinate sentences and relative clauses),

---

<sup>33</sup>An alternative would be to double up on the rheme categories, allowing them to subcategorise for  $X/X$  as well as for *Theme*.

rather than merely an alternative.

**4.1. COÖRDINATE SENTENCES.** Since the coördinate sentence (a) below necessarily involves composition of the (type-raised) subject with the verb, while (b) necessarily does not, it is predicted that the intonation contours that they permit will be more restricted than for the non coördinate sentence (c):

- (63) 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 following pair of intonation contours to be possible for (a) above. (The example assumes the mechanism for multiple pitch accents of section 3.3. Example (a) is a possible answer to the question *What did Bill and Fred do with the beans?*, while (b) is one possible answer to *What did Bill and Fred cook and eat?*):

- (64) a. (Bill COOKED and Fred ATE) (the BEA-NS)  
                                   H\*                                  H\*L                                  L+H\* LH%  
       b. (Bill COOKED and Fred ATE) (the BEA-NS)  
                                   L+H\*                                  L+H\*LH%                                  H\* LL%

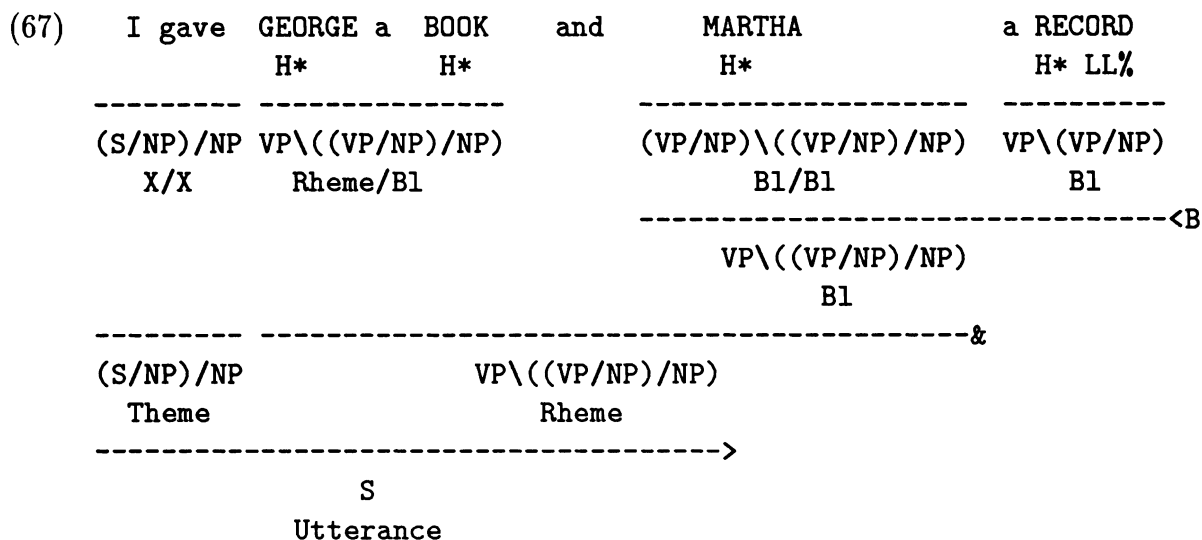
By contrast, intonational tunes which assign categories that are not consistent with the crucial syntactic compositions under the Prosodic Constituency Condition block derivation:

- (65) a. \*(Bill cooked and FRED) (ate the BEA-NS)  
   H\*L  L+H\* LH%  
       b. \*(Bill cooked and FRED) (ate the BEA-NS)  
   L+H\*LH%  H\* LL%

Similarly, garden paths can be forced under the same principle, as was seen earlier:

- (66) \*(Harry likes the NUTS) (and bolts APPROACH)  
   L+H\*LH%  H\* LL%

Another coördinate construction in whose derivation composition and type-raising are obligatory, rather than optional, is the “left node raising” construction illustrated by example 25 above. As in the above examples, the theory predicts that intonation will parallel syntactic derivation in such sentences, and it does. For example, the following is allowed:



**4.2. PROSODIC COÖRDINATION.** There are a number of further intonation contours possible for the previous example 67, which are not permitted by the fragment of prosodic grammar given so far. They could be brought within its scope by the addition of one further rule.

It seems to be possible to utter the sentence with a complete intonational phrasal tune on *George a book* and *Martha a record* – say by putting H\* LL% on *book* and *record*.<sup>34</sup> It also seems possible to put an entire intonational phrase on each noun – say by putting L+H\* LH% on *George* and *Martha* as well.

The theory as given so far does not allow these examples. However, if we were to include a version of the coordination rule 10 which allowed constituents of like syntactic type, when both were marked as rheme or ut-

<sup>34</sup>Selkirk (1984:292) claims that sequences like *Martha a record* in related ditransitive examples cannot by themselves form a single intonational phrase, and suggests that this result follows from the Sense Unit Condition. Selkirk does not discuss coördinate sentences, or the rather exotic contexts that would be required to force the corresponding intonation in simplex sentences. But no such restriction can in the terms of the combinatory theory follow from the Sense Unit Condition. Such sequences can constitute constituents, complete with senses or interpretations. The unacceptability that Selkirk observes seems rather to be related to the similar unacceptability of Right Node Raising for such non-standard constituents, especially when they include proper names, as in

(i) ? I offered, and you gave, Martha a record.

Abbott 1976 has shown that such anomalies do not lie in the realm of grammar.



terance, to combine, then such examples would be allowed, together with certain relatives of earlier examples like 51 and 58, such as the following:

(68) Well, what about the BEANS and the BEER?

What did Fred do with THEM?

(Fred)	((ATE)	(the	BEA-NS))	and	((DRANK)	(the	BEER))
	H* L		L+H* LH%		H* L		L+H* LH%
X/X		Utterance			Utterance		

(69) (They are) (a good source of VITAMINS), and (a cure for SCURVY)

		H* LL%			H* LL%
X/X		Rheme			Rheme

Such a rule is simple enough to specify, but to deal with it here would require attention to some much larger questions concerning related intonational phenomena in other constructions which in some sense involve more than one utterance, notably the parentheticals (cf. Levelt 1989), and some of the coördinate constructions considered by Ladd (1986, 1988), who concludes that prosodic structure is recursive, as opposed to the fixed hierarchies assumed by most authors.

**4.3. RELATIVE CLAUSES.** Since relative clauses, like the coördinate structures of section 4.1, 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 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. Thus we get:

(70) 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%, the following discourses show one way of making the first two alternatives felicitous:

- (71) a. FRED didn't eat the POTATOES. HARRY ate THEM.

(It was the BEANS) (that FRED ate.)

H\* L L+H\* LH%

- (72) b. Fred didn't eat the POTATOES. He threw THEM AWAY.

(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 relatives, so the following discourses may serve to make the remaining two cases felicitous. (I have assumed an analysis with an unmarked theme, but this detail is not crucial):

- (73) c. It wasn't the beans that HARRY ate that looked so delicious.

(It was) (the beans that FRED ate.)

H\* LL%

- (74) d. It wasn't the beans that Fred COOKED that looked so delicious.

(It was) (the beans that Fred ATE.)

H\*LL%

The converse also holds. Tone sequences which violate the Prosodic Constituent Condition 42 are forbidden from appearing on the relative clause. Thus we predict the following, because forward composition cannot combine the theme or rheme on the left with the verb, since the latter bears the null tone, and neither phonological application nor composition can apply.

- (75) 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%

#### REMARKS ON PROCESSING

5. The problem of so-called "spurious" ambiguity engendered by combinatory grammars now appears in a different light. While the semantic properties of the rules (notably the associativity of functional composition) do 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 coördinate constructions considered here must implicate exactly the same non-determinism. It is simply THERE.

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.

(76) Put the block in the box on the table

Moreover, in the discussion in section 3.4 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 as to the nature of a mechanism for resolving not only the nondeterminism inherent in the null tone, but other structural ambiguities such as PP-attachment, as well.

The null tone is found on the theme precisely when the corresponding open proposition is entirely given 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 that this particular ambiguity is only permitted when the theme or open proposition is already in the listener’s model of the discourse. In the case of the earlier example 58, this means that at successive positions in a left-to-right analysis of the string *They are a good source of VITAMINS*, the open proposition corresponding to *They*, *They are*, and *They are a good source of*, 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 combinatorial grammar allows more or less any leftmost substring to be treated as a constituent, complete with an interpretation, the parser that will permit this analysis is extremely simple, amounting to little more than a “reduce-first” shift-reduce categorial parser (cf. Ades and Steedman 1982, Steedman 1990b, 1990c).<sup>35</sup>

---

<sup>35</sup>It seems likely that the similar nondeterminism engendered by the null tone in the

The possibility that so-called spurious ambiguity is resolved by appeal to context in processing spoken language is made more likely by the increasing amount of evidence that the same is true for processing attachment ambiguities like 76 in the written language, as Winograd 1972 originally suggested. Crain & Steedman 1985, Altmann 1988, and Altmann & Steedman 1988 have shown that certain famous “garden path” effects, arising from attachment ambiguities of which PP-attachment is representative, are substantially under the control of referential context, and have argued that these effects show that the human parser resolves syntactic non-determinism in the manner sketched above.<sup>36</sup>

This proposal stands in contrast to that of Kimball (1973), Frazier (1978), and others, who ascribe garden path effects to purely structural preferences arising from parsing strategies such as “Minimal Attachment”. The interpretation of the experimental evidence is still in dispute (see the exchange between Clifton & Ferreira 1989 and Steedman & Altmann 1989). However, if our position is correct, it provides further evidence not only that the problem of “spurious” ambiguity has been misleadingly named, but also that its negative implications for the parsability and psychological reality of combinatory grammars have been greatly exaggerated.

## CONCLUSION.

6. According to the combinatory theory of grammar, the pathway between spoken language and its interpretation is more direct than is implied by the standard theories that were summarised in Figure 1. Intonation structure and surface structure are in fact isomorphic. They can therefore be merged, together with their interpretations, into a system with the architecture shown in Figure 2.

---

other, given/new, dimension of information structure should be treated in the same way.

<sup>36</sup>Note that such processors are also very directly compatible with the proposal to treat some island constraints discussed in connection with examples 14 and 21 under the same kind of semantic control.

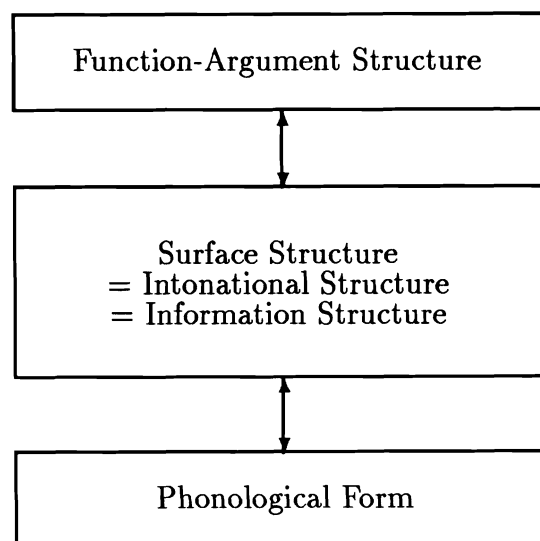


FIGURE 2. Architecture of a Combinatory Prosody.

According to this theory, phonological form maps directly onto surface structure, via the rules of combinatory grammar, subject to the prosodic constituent condition 42. The grammar assigns each constituent of surface structure a discourse function according to the intonational tune that it bears. Surface structure under this new definition therefore subsumes Pierrehumbert's and Selkirk's notion of intonational structure. Moreover, each surface constituent bears an interpretation, which may be a function or an argument, and which in the case of the major information units Theme and Rheme corresponds to such discourse entities as the open proposition and its complement. It follows that surface structure is also isomorphic to what Selkirk called focus structure, here called information structure. Such structures, in which focussed and backgrounded entities and open propositions are represented by functional abstractions and arguments, reduce by functional application, to yield canonical function-argument structures, the traditional basis of logical form.<sup>37</sup>

---

<sup>37</sup>It is assumed here that certain ambiguities of meaning such as quantifier scope are not distinguished at this level, and are derived from these representations later on in the

The combinatory proposal thus represents a return to the “annotated surface structures” proposed by Chomsky (1970) and Jackendoff (1970). It is true that the concept of surface structure has undergone a radical change in a way that allows the autosegmentalist insight concerning intonational structure to be captured. But the structure associated with intonation contour really is *ONLY* surface structure in this new sense, supplemented by annotations which do no more than indicate the information structural status and intonational tune of surface constituents in the extended combinatory sense of the term. This alternative may go some way towards answering the objections that Hajičová and Sgall (1987, 1988) raise against Chomsky’s original proposal, concerning the limitations of his notion of surface constituency when faced with the full range of entities that can as a matter of fact be focussed. This freer notion of surface structure may also explain some of the examples which Bolinger (1985) has used to argue for an entirely autonomous, lexically-oriented account of accent assignment, and which Gussenhoven (1985) has used to argue for a similarly autonomous focus-based account. It may also allow us to eliminate some of the non-syntactic string-based rules and “performance structures” that Cooper and Cooper (1980) and Grosjean and Gee (1983) have proposed to add to the syntax-driven model. In the terms of Halle’s metaphor, the spiral-bound notebook of structural descriptions may be quite a slim volume after all.

Much further work remains to be done. 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 all the metrical theories mentioned in the Introduction.) Serious difficulties still attend the automatic identification of prosodic boundaries in speech. The phonetic realisation of elements such as pitch accents and boundary tones are subject to coarticulation effects, like all phonological segments, and are hard to recognise. In fact, it is highly likely that their identification cannot be carried out in isolation from the recognition of the words that carry them. This observation might seem daunting, since current techniques for word recognition, while improving dramatically, are nonetheless not very good. However, it is likely that the task of recognising words and intonation together will turn out to be easier than doing either task in isolation. One reason for the success of

---

comprehension process.

the stochastic techniques which are currently in vogue seems to be that they fairly directly model at least some aspects of word-level prosody, particularly timing. However, by the same token these techniques as they are currently applied are extremely vulnerable to the variations in prosody that are induced by the phenomena discussed here. It is therefore likely that these two tasks may facilitate each other, as Pierrehumbert 1990 points out.

The most significant practical benefit of the combinatory theory therefore seems likely to be the following. 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. Now they can be seen to be in complete harmony. Processors may more easily be devised which use all these sources of information at once, potentially simplifying both problems. In particular, 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 open propositions 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 bottom-up processing at lower levels.

#### REFERENCES

- ABBOTT, BARBARA. 1976. Right Node Raising as a Test for Constituenthood. *Linguistic Inquiry*. 7.639-642.
- ADES, ANTHONY AND MARK STEEDMAN. 1982. On the Order of Words. *Linguistics and Philosophy*. 4.517-558.
- AJDUKIEWICZ, KAZIMIERZ. 1935. Die syntaktische Konnexität. *Studia Philosophica*. 1.1-27. English translation in *Polish Logic 1920-1939*, ed. by Storrs McCall, 207-231. Oxford University Press.
- ALTMANN, GERRY. 1988. Ambiguity, Parsing Strategies, and Computational Models Language and Cognitive Processes, 3.73-98.
- ALTMANN, GERRY AND MARK STEEDMAN. 1988. Interaction with Context During Human Sentence Processing Cognition. 30.191-238
- BAR-HILLEL, YEHOShUA. 1953. A Quasi-arithmetical Notation for Syntactic Description. *Language*. 29.47-58.
- BAR-HILLEL, YEHOShUA, C. GAFFMAN AND E. SHAMIR. 1960. On Categorical and Phrase Structure Grammars. *The Bulletin of the Research Council of Israel*. 9F.1-16.

- BECKMAN, MARY. 1986. Stress and Non-Stress Accent. Dordrecht: Foris.
- BECKMAN, MARY, AND JANET PIERREHUMBERT. 1986. Intonational structure in Japanese and English. *Phonology Yearbook*. 3.255-309.
- BING, JANET. 1979. Aspects of English Prosody. Ph D dissertation, University of Massachusetts at Amherst. New York: Garland Publishing.
- BOLINGER, DWIGHT. 1985. Two views of Accent. *Journal of Linguistics*. 21.79-124.
- BOLINGER, DWIGHT. 1989. *Intonation and its Uses*. Stanford CA: Stanford University Press.
- CHOMSKY, NOAM. 1970. Deep Structure, Surface Structure, and Semantic Interpretation. *Semantics*, ed. by D. Steinberg and L. Jakobovits. 1971. 183-216. Cambridge: Cambridge University Press.
- CHOMSKY, NOAM. 1982. Concepts and Consequences of the Theory of Government and Binding. Cambridge MA: MIT Press.
- CHOMSKY, NOAM AND MORRIS HALLE. 1968. *The Sound Pattern of English*. New York: Harper and Row.
- CLIFTON, CHARLES AND FERNANDA FERREIRA. 1989. Ambiguity in Context. *Language and Cognitive Processes*. 4.77-104.
- COOPER, WILLIAM AND J. PACCIA-COOPER. 1980. *Syntax and Speech*. Cambridge MA: Harvard University Press.
- CRAIN, STEPHEN AND MARK STEEDMAN. 1985. On not being led up the garden path: the use of context by the psychological parser. *Natural Language Parsing*. Ed. by David Dowty, L. Karttunen, and A. Zwicky, 320-358. Cambridge: Cambridge University Press.
- CRESSWELL, MAX. 1973. *Logics and Languages*. Methuen.
- CURRY, HASKELL AND ROBERT FEYS. 1958. *Combinatory Logic*. Amsterdam: North Holland.
- DELIN, JUDY. 1989. *Cleft Constructions*. unpublished Ph.D dissertation. University of Edinburgh.
- DOWNING, BRUCE. 1970. *Syntactic Structure and Phonological Phrasing in English*, Ph.D dissertation, University of Texas at Austin.
- DOWTY, DAVID. 1988. Type raising, functional composition, and non-constituent coordination. *Categorial Grammars and Natural Language Structures*, ed. by Richard T. Oehrle, E. Bach and D. Wheeler, 153-198. Dordrecht: Reidel.
- FRAZIER, LYN. 1978. *On Comprehending Sentences*. PhD dissertation. University of Connecticut. distributed by Indiana University Linguistics Club.
- GEE, JAMES, AND FRANÇOIS GROSJEAN. 1983. Performance Structures: a Psycholinguistic and Linguistic Appraisal. *Cognitive Psychology*. 15.411-458.



- GOLDSMITH, JOHN. 1976. Autosegmental Phonology. Ph. D. dissertation, MIT. Published by Indiana University Linguistics Club.
- GUSSENHOVEN, CARLOS. 1985. Two Views of Accent: a Reply. *Journal of Linguistics*. 21.125-138.
- HAJIČOVÁ, EVA, AND PETR SGALL. 1987. The Ordering Principle. *Journal of Pragmatics*. 11.435-454.
- HAJIČOVÁ, EVA, AND PETR SGALL. 1988. Topic and Focus of a Sentence and the Patterning of a Text. *Text and Discourse Constitution*, ed. by János Petöfi, 70-96. Berlin: De Gruyter.
- HALLIDAY, MICHAEL. 1967. Intonation and Grammar in British English. The Hague: Mouton.
- HALLIDAY, MICHAEL. 1967. Notes on Transitivity and Theme in English. Part II. *Journal of Linguistics*. 3.199-244
- HALLIDAY, MICHAEL. 1970. "Language Structure and Language Function". in John Lyons (ed.). *New Horizons in Linguistics*. Penguin.
- HALLIDAY, MICHAEL, AND RUQAIYA HASAN. 1976. *Cohesion in English*. Longman.
- HEPPLE, MARK AND GLYN MORRILL. 1989. Parsing and Derivational Equivalence. *Proceedings of the 4th Conference of the European Association for Computational Linguistics*, Manchester, April 1989. 10-8.
- HIRSCHBERG, JULIA AND JANET PIERREHUMBERT. 1986. Intonational Structuring of Discourse. *Proceedings of the 24th Meeting of the Association for Computational Linguistics*, New York. 136-144.
- HOEKSEMA, JACK. 1984. *Categorial Morphology*, Ph.D thesis, Rijksuniversiteit Groningen. New York: Garland.
- HOEKSEMA, JACK AND RICHARD JANDA. 1988. Implications of Process-Morphology. *Categorial Grammars and Natural Language Structures*, ed. by Richard T. Oehrle, E. Bach and D. Wheeler, 199-248. Dordrecht: Reidel.
- JACKENDOFF, RAY. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge MA: MIT Press.
- JACOBSON, PAULINE. 1990. A Categorial Account of Raising. *Linguistics and Philosophy*, 6.XXX-XXX
- KAISSE, ELLEN. 1985. *Connected Speech: the Interaction of Syntax and Phonology*. New York: Academic Press.
- KARTTUNEN, LAURI. 1989. Radical Lexicalism. *Alternative Conceptions of Phrase Structure*, ed. by Mark Baltin and Tony Kroch, 43-65. Chicago: University of Chicago Press.
- KIMBALL, JOHN. 1973. Seven Principles of Surface Structure Parsing in Natural Language. *Cognition*. 2.15-47.

- LADD, D. ROBERT. 1986. Intonational Phrasing: the case for Recursive Prosodic Structure. *Phonology Yearbook*. 3.311-340.
- LADD, D. ROBERT. 1988. Metrical Representation of Pitch Register in John Kingston and Mary Beckman, (eds.), *Papers in Laboratory Phonology I*. Cambridge: Cambridge University Press (to appear).
- LAMBEK, JOACHIM. 1958. The Mathematics of Sentence Structure. *American Mathematical Monthly*. 65.154-170.
- LEVELT, WILLEM. 1989. *Speaking: From Intention to Articulation*, Cambridge MA: MIT Press.
- LIBERMAN, MARK. 1975. *The Intonation System of English*. Ph.D dissertation, New York: Garland.
- LIBERMAN, MARK AND ALAN PRINCE. 1977. On Stress and Linguistic Rhythm. *Linguistic Inquiry*. 8.249-336.
- LYONS, JOHN. 1977. *Semantics, Vol.II*, Cambridge: Cambridge University Press.
- MCCONNELL-GINET, SALLY. 1982. Adverbs and Logical Form. *Language*. 58.144-184.
- MOORTGAT, MICHAEL. 1987. Lambek Categorical Grammar and the Autonomy Thesis. INL Working paper 87-03, Leiden.
- MOORTGAT, MICHAEL. 1988a. Mixed Composition and Discontinuous Dependencies. *Categorical Grammars and Natural Language Structures*, ed. by Richard T. Oehrle, E. Bach and D. Wheeler, 319-418. Dordrecht: Reidel.
- MOORTGAT, MICHAEL. 1988b. *Categorical Investigations*, Dordrecht: Foris.
- NESPOR, MARIA, AND IRENE VOGEL. 1986. *Prosodic Phonology*, Dordrecht: Foris.
- OEHRLER, RICHARD Multi-dimensional Compositional Functions as a basis for Grammatical Analysis. *Categorical Grammars and Natural Language Structures*, ed. by Richard T. Oehrle, E. Bach and D. Wheeler, 349-390. Dordrecht: Reidel.
- PEREIRA, FERNANDO. 1985. A New Characterisation of Attachment Preferences. D. Dowty, L. Karttunen, and A. Zwicky, (eds.), *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, ACL Studies in Natural Language Processing, 307-319. Cambridge: Cambridge University Press.
- PIERREHUMBERT, JANET. 1980. *The Phonology and Phonetics of English Intonation*, Ph.D dissertation, MIT, distributed by Indiana Linguistics Club, Bloomington, IN.
- PIERREHUMBERT, JANET. 1990. *Intonation, Prosody and Speech Technology*. ms. Northwestern University. (Incomplete reference.)

- PIERREHUMBERT, JANET, AND MARY BECKMAN. 1989. *Japanese Tone Structure*, Cambridge MA: MIT Press.
- PIERREHUMBERT, JANET, AND JULIA HIRSCHBERG. 1990. The Meaning of International Contours in the Interpretation of Discourse. *Intentions in Communication*, ed. by Philip Cohen, Jerry Morgan and Martha Pollock, 271-312. Cambridge: MIT Press.
- PRINCE, ALAN. 1983. Relating to the Grid. *Linguistic Inquiry*. 14.19-100.
- 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.
- ROCHEMONT, MICHAEL, AND PETER CULICOVER. 1990. *English Focus Constructions and the Theory of Grammar*, Cambridge: Cambridge University Press.
- ROOTH, MATS. 1985. *Association with Focus*, unpublished PhD dissertation, University of Massachusetts, Amherst.
- ROSS, JOHN ROBERT. 1970. Gapping and the Order of Constituents. *Progress in Linguistics*, ed. by Manfred Bierwisch and M. Heidolph. The Hague: Mouton.
- SAG, IVAN. 1976. *Deletion and Logical Form*, Ph.D dissertation, MIT. New York: Garland Publishing.
- SCHMERLING, SUSAN. 1976. *Aspects of English Sentence Stress*, Austin TX: University of Texas Press.
- SCHMERLING, SUSAN. 1980. The Proper Treatment of the Relationship between Syntax and Phonology. paper presented at the 55th annual meeting of the LSA, December 1980. San Antonio TX.
- SELKIRK, ELISABETH. *Phonology and Syntax*, Cambridge MA: MIT Press.
- SILVERMAN, KIM. 1988. Utterance-internal Prosodic Boundaries. *Proceedings of the 2nd Australian International Conference on Speech Science and Technology*, Sydney, November 1988. 86-91.
- SMULLYAN, RAYMOND. 1985. *To Mock a Mockingbird*, New York: Knopf.
- VON STECHOW, ARNIM. 1989. Focussing and Backgrounding Operators. *Fachgruppe Sprachwissenschaft der Universität Konstanz, Arbeitspapier Nr. 6*.
- STEEDMAN, MARK. 1985. Dependency and Coordination in the Grammar of Dutch and English. *Language*, 61.523-568.
- STEEDMAN, MARK. 1987. Combinatory grammars and parasitic gaps. *Natural Language and Linguistic Theory*. 5.403-439.
- STEEDMAN, MARK. 1990a. Gapping as Constituent Coordination. *Linguistics and Philosophy*. 13.207-263.
- STEEDMAN, MARK. 1990b. *Grammars and Processors. Aspects of Computational Linguistics*, ed. by Hans Kamp and Christian Rohrer, Springer Verlag (to

- appear).
- STEEDMAN, MARK. 1990c. Type Raising and Directionality in Combinatory Grammar. ms, CIS, Univ. of Pennsylvania.
- STEEDMAN, MARK AND GERRY ALTMANN. 1989. Ambiguity in Context: a Reply. *Language and Cognitive Processes*. 4.105-122.
- SZABOLCSI, ANNA. 1987a. On Combinatory Categorical Grammar. *Proceedings of the Symposium on Logic and Language, Debrecen*, 151-162. Budapest: Akadémiai Kiadó.
- SZABOLCSI, ANNA. 1987b. Bound variables in Syntax: are there any?. *Proceedings of the 6th Amsterdam Colloquium on Logic and Linguistics*, ed. by Renate Bartsch, Foris:Dordrecht (to appear).
- SZABOLCSI, ANNA, AND FRANZ ZWARTS. 1990. Semantic Properties of Composed Functions and the Distribution of *Wh*-phrases. *Proceedings of the Seventh Amsterdam Colloquium*, ed. by Martin Stokhof and Torenvliet. 529-555. Amsterdam: ITLI.
- USZKOREIT, HANS. 1986. Categorical Unification Grammars. In *Proceedings of the 11th International Conference on Computational Linguistics, Bonn, August, 1986*, 187-194.
- VIJAY-SHANKER, K. AND DAVID WEIR. 1990. Polynomial Time Parsing of Combinatory Categorical Grammars. *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics, Pittsburgh, June 1990*. 1-8.
- WEIR, DAVID, AND ARAVIND JOSHI. 1988. Combinatory Categorical Grammars: Generative Power and Relation to Linear CF Rewriting Systems. *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics, Buffalo, June 1988*. 278-285.
- WHEELER, DEIRDRE. 1981. Aspects of a Categorical Theory of Phonology, Ph.D dissertation, University of Massachusetts, Amherst.
- WHEELER, DEIRDRE. 1988. Consequences of Some Categorially Motivated Phonological Assumptions. *Categorical Grammars and Natural Language Structures*, ed. by Richard T. Oehrle, E. Bach and D. Wheeler, 467-488. Dordrecht: Reidel.
- WILSON, DEIRDRE AND DAN SPERBER. 1979, Ordered Entailments. *Syntax and Semantics, XI, Presupposition*, ed. by Choon-Kyu Oh and David A. Dinneen, 299-323. New York: Academic Press.
- WINOGRAD, TERRY. 1972. *Understanding Natural Language*, Edinburgh: Edinburgh University Press.
- WITTENBURG, KENT. 1986. Natural Language Parsing with Combinatory Categorical Grammar in a Graph-Unification Based Formalism. D.Phil Dissertation, U. of Texas at Austin.

- ZEEVAT, HENK, EWAN KLEIN, AND JO CALDER. 1987. An Introduction to Unification Categorical Grammar. Edinburgh Working Papers in Cognitive Science, 1: Categorical Grammar, Unification Grammar, and Parsing, ed. by N. Haddock et al. Edinburgh University.
- ZWICKY, ARNOLD AND GEOFFREY K. PULLUM. 1987. The Syntax-Phonology Interface, course notes for LI219, LSA Summer Institute, Stanford 1987.

Department of Computer and Information Science,  
University of Pennsylvania,  
200 South 33rd Street,  
Philadelphia PA 19104-6389