Information-Structural Semantics for English Intonation*

Mark Steedman
University of Edinburgh
steedman@cogsci.ed.ac.uk

Draft 2.2, July 30, 2003

Selkirk (1984), Hirschberg and Pierrehumbert (1986), Pierrehumbert and Hirschberg (1990), and the present author, have offered different but related accounts of intonation structure in English and some other languages. These accounts share the assumption that the system of tones identified by Pierrehumbert (1980), as modified by Pierrehumbert and Beckman (1988) and Silverman et al. (1992), has as transparent and type-driven a semantics in these languages as do their words and phrases. While the semantics of intonation in English concerns information structure and propositional attitude, rather than the predicate-argument relations and operator-scope relations that are familiar from standard semantics in the spirit of the papers collected as Montague 1974, this information-structural semantics is fully compositional, and can be regarded as a component of the same semantic system.

The present paper builds on Steedman (1991) and Steedman (2000a) to develop a new semantics for intonation structure, which shares with the earlier versions the property of being fully integrated into Combinatory Categorial Grammar (CCG, see Steedman 2000b, hereafter *SP*). This grammar integrates intonation structure into surface derivational structure and the associated Montague-style compositional semantics, even when the intonation structure departs from the restrictions of traditional surface structure. Many of the diverse discourse meanings that have been attributed to intonational tunes are shown to arise via conversational implicature from more primitive literal meanings distinguished along the three dimensions of information structure, speaker/hearer commitment, and contentiousness.

1 Tones and Information Structure

It is standard to assume, following Bolinger (1958, 1961) and Halliday (1963, 1967a,b), that pitch-accents, high or low, simple or compound, are in the first place properties of the words that they fall on, and that they mark the interpretations of those words as contributing to the distinction between the speaker's actual utterance and other things that they might be expected to have said in the context to hand, as in the "Alternative Semantics" of Kartunnen (1976),

*Thanks to Betina Braun, Daniel Büring, Klaus von Heusinger, Stephen Isard, Alex Lascarides, and Bonnie Webber for comments on the draft. An earlier version of some parts of the paper appears as Steedman (2002). The work was supported in part by EPSRC grants GR/M96889 and GR/R02450, and EU FET grant MAGICSTER.

Karttunen and Peters (1979), Rooth (1985, 1992), and Büring (1997a,b).[1] In this sense, all pitch accents are *contrastive*. For example, in response to the question "Which finger did he bite?", the word that contributes to distinguishing the following answer from other possible answers via reference is the deictic "this", so the following intonation is appropriate.[2]

(1) He bit THIS one .
H* LL%

It is important to be clear from the start that the set of alternative utterances from which the actual utterance is distinguished by the tune is in no sense the set of all possible utterances appropriate to this context, a set which includes infinitely many things like "Mind your own business," "That was no finger," "What are you talking about?" and "Lovely weather we're having." Rather, the presupposed set of (presumably, ten) alternative utterances is *accommodated* by the hearer in the sense of Lewis (1979) and Thomason (1990), like any speaker presupposition that is not actually inconsistent with their beliefs. This does not imply that such alternative sets are confined to things that have been mentioned, or that they are mentally enumerated by the participants—or indeed that they are even finite.

In terms of Halliday's given/new distinction pitch-accents are markers of "new" information, although the words that receive pitch-accents may have been recently mentioned, and it might be better to call them markers of "not given" information. That seems a little cumbersome, so I will use the term "kontrast" from Vallduví and Vilkuna 1998 for this property of English words bearing pitch-accents, spelling the corresponding verb "k-contrast".[3]

I'll further attempt to argue that there are just two independent semantic binary-valued dimensions along which the literal meanings of the various pitch-accent types are further distinguished. The first of these dimensions has been identified in the literature under various names, and distinguishes between what I'll continue to call "theme" and "rheme" components of the utterance, using these terms in the sense of Bolinger (1958, 1961) rather than Halliday. Theme can be thought of informally as the part of the sentence corresponding to a question or topic that is presupposed by the speaker, and rheme is the part of the utterance that constitutes the speaker's novel contribution on that question or topic. However, it will become clear below that the notion of theme differs from that of topic as defined by, for example, Gundel (1974); Gundel and Fretheim (2001) in being speaker-defined rather than text-based.

A great deal of the huge and ramifying literature on information structure can be summarized as distinguishing two dimensions corresponding to the given/kontrast and theme/rheme distinctions, although the consensus has tended to be obscured by the very different nomenclatures that have been applied. (See discussion by Steedman and Kruijff-Korbayová (2001), which summa-

[1] The term "pitch-accent" is here restricted to what Ladd (1996) calls "primary" pitch-accents, sometimes called "nuclear" pitch accents (although there may be more than one in a sentence). Ladd follows Bolinger and many others in distinguishing primary accents from certain other accents that arise from the interaction of lexical stress with metrical the metrical grid. While there is still no objective measure to distinguish the two varieties, it is the primary accents that are perceived as emphatic or contrastive.

[2] The notation for tunes is Pierrehumbert's, see Pierrehumbert and Hirschberg 1990 for details including characteristic pitch-contours.

[3] In Steedman 2000a and earlier work I called this property "focus", following the "narrow" sense of Selkirk (1984). However this term invites confusion with the "broad" sense intended by Hajičová and Sgall (1988) and Vallduví (1990), which is closer to the term "rheme" as used in the present system, and in Steedman 2000a and Vallduví and Vilkuna 1998.

rizes the terminology and its lines of descent, along with some contiguous semantic influences.)

However, there is a further dimension of discourse meaning along which the pitch-accent types are distinguished which has not usually been identified in this literature. It concerns whether or not the particular theme or rheme to hand is mutually agreed–that is, uncontentious. This notion is related to various notions of Mutual Belief or Common Ground proposed by Lewis (1969), Cohen (1978), Clark and Marshall (1981) and Clark (1996).[4]

Both of these components of meaning are projected by the process of grammatical derivation from the words that carry the pitch-accent to the prosodic phrase corresponding to these information units, following Steedman 2000a, along lines briefly summarized in section 5.

I'll also try to argue that the intonational boundaries such as those sometimes referred to as "continuation rises," which delimit the prosodic phrase, fall into two classes respectively distinguishing the speaker or the hearer as responsible for, or (in terms of the related accounts of Gussenhoven (1983, p.201) and Gunlogson (2001, 2002)) **committed to**, the corresponding information unit.[5]

I'll assume that the speaker's knowledge can be thought of as a database or set of propositions in a logic (second-order, since themes etc. may be functions), divided into two subdomains, namely: a set S of information units that the speaker claims to be committed to, and a set H of information units which the speaker claims the hearer to be committed to. Information units are further distinguished on a dimension $\pm$*AGREED* according to whether the speaker claims them to be uncontentious or contentious. The set of $+$*AGREED* information units is not merely the intersection of S and H: the speaker may *attribute* uncontentiousness to an information unit and responsibility for it to the hearer whilst knowing that in fact they do not regard themselves as so committed. In Steedman 2000a, S and H are treated as modalities $[S]$ and $[H]$ of a modal logic, and Stone (1998) has proposed a similar modality for mutual belief. In the present paper we will combine the feature $\pm$*AGREED* with the speaker/hearer modalities, writing it as a superscript $\pm$, as in $[H^{+}]$.

These classifications can be set out diagrammatically as in the tables 1 and 2, in which θ signifies theme, ρ signifies rheme, $+$ indicates $+$*AGREED*, $-$ indicates $-$*AGREED*, and $[S]$ and $[H]$ respectively denote speaker and hearer commitment. If a theme or a rheme is marked as

	$+$	$-$
θ	L+H*	L*+H
ρ	H*, (H*+L)	L*, (H+L*)

Table 1: The Meanings of the Pitch-accents

$[S]$	L, LL%, HL%
$[H]$	H, HH%, LH%

Table 2: The Meanings of the Boundaries

[4] Hobbs (1990), who proposes a very different revision of Pierrehumbert and Hirschberg (1990) to the present one, also gives a central role to Mutual Belief.

[5] In Steedman 2000a, I called this dimension "ownership".

agreed, then it's in *AGREED*, whoever is explicitly claimed to be committed to it. If it is not so marked, then it is not in *AGREED*, *even if speaker and hearer in fact both believe it*. This last possibility arises because *H* is only the speaker's attribution of commitment to the hearer, not the hearer's actual belief. It follows that a theme or rheme may be believed by the speaker, and asserted by the speaker to be something that the hearer is committed to, without the hearer's actually agreeing to it. We will come to a case of this later on.

At first glance, this proposal might appear to miss the point entirely. Where are notions like "topic continuation" (Brown, Currie and Kenworthy 1980) and "evaluation with respect to subsequent material" (Pierrehumbert and Hirschberg 1990), or the latter authors' scales of commitment and belief? I'm going to argue that many of the effects that have been associated with intonational tunes arrise as conversational implicatures from the interaction with context of literal meanings made up of the above simple components. To consider this claim we need some examples.

2 An Example: Pitch-accents

The first example commemorates Miles Davis' response to Dave Brubeck's question concerning his reason for playing *E*♮ as the final note of *In Your Own Sweet Way*, in place of *E*♭ as written by Brubeck:[6]

(2) DB: Why did you play *E*-natural?
MD: (Why didn't YOU) (WRITE *E*-natural ?)
L+H* LH% H* LL%
background kontrast *kontrast background*
theme *rheme*

The LH% boundary splits the utterance into two intonational phrases and two information units. The L+H* accent marks the first of these units as theme (L*+H would also be appropriate). It falls on the word *you* because its referent (Brubeck) is the element that distinguishes *this* theme from the other themes that are available. (Lambrecht and Michaelis (1998) in a related approach call such "marked" or contrastive themes "ratified topics". Ratification certainly presupposes some alternative. However, the example to hand suggests that ratification is only one of many things that you can do with a contrastive theme or topic.)

The set of available themes, which we will call the "Theme Alternative Set" (ThAS) is presupposed by Davis and accommodated by Brubeck as including just two possible themes. These can be thought of informally as "Why did/didn't Davis do *x*?" and "Why did/didn't Brubeck do *x*?" More formally we can think of the Theme Alternative Set as a set of λ terms, which for this context is as follows, in which ± stands for polarity:

(3) $\left\{ \begin{array}{l} \lambda vp.\lambda reason.cause'\,reason(\pm do'\,vp\ brubeck') \\ \lambda vp.\lambda reason.cause'\,reason(\pm do'\,vp\ davis') \end{array} \right\}$

(It's assumed here that the fragment *Why didn't you* is assigned a meaning which is a function from VP interpretations to *why*-question interpretations—the latter being themselves functions

[6]The story comes from Dave Brubeck. Miles was of course absolutely right. The tones shown in the example remain conjectural, however, given his complete lack of any trackable F_0.

from adverbial interpretations to causal propositions. This is in fact what the CCG grammars outlined in *SP* and below actually deliver, given an appropriate lexicon.)

Other themes and Theme Alternative Sets are possible. For example, a further L+H* pitch-accent on *didn't* is possible:

(4) (Why DIDN'T YOU) (WRITE *E*-natural ?)
L+H* L+H* LH% H* LL%

By saying (4), Davis presupposes, and Brubeck accommodates, a Theme Alternative Set which informally can be thought of as "Why did Davis do *x*" and "Why didn't Brubeck do *x*", and can be written as:

(5) $\left\{ \begin{array}{l} \lambda vp.\lambda reason.cause'\,reason(-do'\,vp\ brubeck') \\ \lambda vp.\lambda reason.cause'\,reason(+do'\,vp\ davis') \end{array} \right\}$

In both cases, words whose interpretation distinguishes the intended theme from the others—which is how "k-contrasted" or "not given" is defined in the present system—bear pitch-accents, while those that do not contribute to the distinction—which is how we define "background" or "given"—do not. (See Prevost and Steedman 1994; Prevost 1995 for further detail on the determination of pitch-accent placement in sentence generation.)

We do not need to think of the Theme Alternative Sets as closed under terms that are already in play in the conversation. A more general representation of the ThAS for (2) reminiscent of the "Structured Meaning" approach of Cresswell (1973, 1985) and von Stechow (1981) can be obtained by abstracting over the element(s) corresponding to accented words, thus:

(6) $\lambda subj.\lambda vp.\lambda reason.cause'\,reason(\pm do'\,vp\ subj)$

Similarly, the ThAS for (4) can be written as follows:

(7) $\lambda polarity.\lambda subj.\lambda vp.\lambda reason.cause'\,reason(polarity(do'\,vp\ subj))$

Of course, themes including this one may not, and in fact usually do not, bear any pitch-accent at all, as in:

(8) (Why didn't you) (WRITE *E*-natural ?)
H* LL%

Such noncontrastive or "unmarked themes" presuppose or are accommodated to a singleton ThAS—in this case the following:

(9) $\{\ \lambda vp.\lambda reason.cause'\,reason(-do'\,vp\ brubeck')\ \}$

Thus according to the present theory, as Halliday and Brown insisted, what is "new", "not given," or k-contrasted vs. what is "given" or background is in part determined by the speaker, not a property of a text or context alone (Brown 1983:67). By the same token, the notion of theme is also partly speaker-determined, not text-based as is the notion of topic of Gundel (1974); Gundel and Fretheim (2001).

Similar considerations govern the effect of the rheme-tune in (2) and (4). The H* accent marks the second information unit as a rheme, and it falls on the word *write* because it is the interpretation of that word that distinguishes *this* rheme from the others that the context affords. This set of available rhemes, which we will call the "Rheme Alternative Set" (RhAS) is, again presupposed/accommodated by the participants to include only doing things to $E\natural$. In this par-

ticular case we can think of the RhAS as being closed under the things that have actually been mentioned—that is as

(10) $\left\{ \begin{array}{l} \lambda x.play' e \natural' x \\ \lambda x.write' e \natural' x \end{array} \right\}$

Again, we can again think of the RhAS more generally by abstracting over the transitive predicate in structured-meaning style:

(11) $\lambda tv.\lambda x.tv\ e \natural' x$

We have so far passed over the role of the particular boundary tones in (2) and (4). Earlier we identified this role as assigning responsibility for theme/rheme status to either speaker or hearer. Thus the claim must be that in the above examples the theme is marked by Miles Davis as Brubeck's responsibility, whereas the rheme is marked as his own. To see what this means, and to understand the implication of table 1 that both are "agreed", we must look more broadly at the function of the boundary tones.

3 An Example: Boundaries

Brown (1980:30) identifies the role of high boundaries as indicating that there is more to come on the current topic from some participant. Pierrehumbert and Hirschberg (1990:304-308), from whom the following example is adapted, make a related claim concerning interpretation with respect to succeeding material (again, this may come from either participant):

(12) a. Attach the jumper cables to the car that's running,
L+H* L+H* LH%

b. Attach them to the car you want to start,
L+H* L+H* LH%

c. Try the ignition,
L+H* L+H* LH%

d. If you're lucky,
L+H* LH%

e. You've started your car.
H* LL%

Pierrehumbert and Hirschberg don't actually specify the pitch-accent types for this example, but L+H* seems appropriate for all accents except those in the last clause—in fact, H* accents sound quite odd, for reasons we'll come to. In present terms, this means that the earlier clauses are all themes, and illustrates the fact that multiple themes, and in fact isolated themes without any rheme, are all possible.

It is interesting to consider the effect of replacing the LH% boundaries by LL% boundaries, retaining the L+H* accents. This manipulation does not affect the coherence of the example very much. The main effect is to make the speaker's prescription seem somewhat abrupt and discouraging of any interruption, and to be generally unconcerned with whether or not it is making any sense to the hearer. In comparison, the original (12) seems more attentive, and to invite the hearer to take control of the discourse if they want to.

I'm going to claim that in both cases the forward motion of the discourse is the same, and is brought about, not by the inclusion of high boundaries as such, but by the rheme-expectation stemming from the theme-marking L+H* pitch-accents. The specific "kinder, gentler" effect of the version with LH% boundaries arises from their primary meaning of marking hearer-commitment. By marking the themes as, in the speaker's view, the hearer's responsibility (although in fact they may be completely new to the hearer), the possibility of the latter taking control of the discourse is maintained at every turn.

These claims are borne out by considering the effect of substituting H* rheme accents for L+H* in both high- and low- boundary versions. With high boundaries, the instructions become quite irritating, and seem to imply that the hearer knows all this already. With low boundaries, the effect is again abrupt and not hearer-oriented. In both cases, coherence (though inferable from world knowledge) is reduced.

I'm further going to claim that all the related effects of high boundaries, which have been variously described in the descriptive literature as "other-directed", "turn-yielding", "discourse-structuring," or "continuation" are similarly indirect implicatures that follow from the basic sense of high boundaries, which is to identify the hearer as in the speaker's view committed to the relevant information unit.

4 The Full System

We are now ready to look at the entire system laid out in tables 1 and 2, via some simpler minimal pairs of examples in which tones including the L* pitch-accents and boundaries are systematically varied across the same text.

If we limit ourselves for the sake of simplicity to tunes with a single pitch-accent, assume that H*+L and H+L* are not distinct from H* and L*, and take LL% and LH% as representative of the two classes of boundary then the classification in tables 1 and 2 allow eight tunes which exemplify the $2^3 = 8$ possible combinations of these three binary features. It is instructive to consider the effect of these tunes when applied to the same sentence "I'm a millionaire," uttered in response to various prompts. It's important to realize that all these responses are indirect, and their force depends on whether the participants regard being a millionaire as counting as being rich.

(13) H: You appear to be rich.
S: I'm a MILLIONAIRE.
H* LL%
$[S^+]\rho$ *millionaire′me′* (S committed to an agreed rheme.)

(14) H: You appear to be poor
S: I'm a MILLIONAIRE.
L* LL%
$[S^-]\rho$ *millionaire′me′* (S committed to a non-agreed rheme.)

(15) H: Congratulations. You're a millionaire.
S: I'm a MILLIONAIRE?
H* LH%
$[H^+]\rho$ *millionaire′me′* (H committed to an agreed rheme.)

(16) H: Congratulations. You're a millionaire.
S: I'm a MILLIONAIRE?
L* LH%
$[H^-]\rho$ *millionaire′me′* (H committed to a non-agreed rheme.)

The above four responses can be assumed to consist of a single rheme.[7] The ones involving an L* pitch-accent mark the rheme as being not agreed. However, the pitch-accent itself does not distinguish who the opposition is coming from. This is not an ambiguity in the pitch-accent itself. Rather, the identification of the source of the conflict and the entire illocutionary force of the response depends on inference on the basis of what else is known about the participants' beliefs. Thus, in (14), the one who appears to doubt the proposition in the second utterance is the hearer, but in (16) it is the speaker. In different contexts, the difference could be reversed or eliminated.

A similar pattern can be observed for the theme pitch-accents:

(17) H: You appear to be rich.
S: I'm a MILLIONAIRE.
L+H* LL%
$[S^+]\theta$ *millionaire′me′* (S committed to an agreed theme.)

(18) H: You appear to be poor.
S: I'm a MILLIONAIRE.
L*+H LL%
$[S^-]\theta$ *millionaire′me′* (S committed to a non-agreed theme.)

(19) H: You appear to be a complete jerk.
S: I'm a MILLIONAIRE.
L+H* LH%
$[H^+]\theta$ *millionaire′me′* (H committed to an agreed theme.)

(20) H: You appear to be a complete jerk.
S: I'm a MILLIONAIRE.
L*+H LH%
$[H^-]\theta$ *millionaire′me′* (H committed to a non-agreed theme.)

At first encounter, it may appear that these tunes must mark rhemes, like those in (13) to (16). However in Steedman 2000a, I show that these are in fact isolated themes, of the kind we have already noticed in connection with example (12). These isolated themes achieve the effect of a response (as well as various other implicatures of impatience, diffidence, incompleteness, etc.) via the indirect speech act of leaving the hearer to generate the rheme for themselves.

As before, the tunes involving L*+H accents imply disagreement or absence from mutual belief. Once again, the source of the disagreement can only be identified from the full discourse context. In the case of (19) and (20), it is important to remember that the speaker's LH% boundary means only that the speaker *views the hearer as* committed to these themes. As far as the hearer is concerned, that is not the same as an actual commitments. Thus the L*+H in (20) sim-

[7] Under the proposal in Steedman 2000a, they could also be analyzed as an unmarked theme "I'm" and a rheme "a millionaire". In this particular context it makes very little difference, and we'll ignore these readings.

ply has the effect of correctly excluding from the mutual belief set *AGREED* this theme which the boundary marks as in *H*, in spite of the fact that can also be inferred to be in the speaker's own beliefs *S*. This is the possibility that was noticed in the discussion of tables 1 and 2: it seems a fundamental property of the system that there is a distinction between a proposition merely being in both *S* and *H* and it actually being in *AGREED*. The former amounts to a claim by the speaker that both participants *ought* to be committed to it. The latter is a claim by the speaker that both actually *are* committed.

Example (20) is identical in information structural terms to the following example, extensively discussed by Ward and Hirschberg (1985) (see Pierrehumbert and Hirschberg 1990:295, (26)):

(21) H:Harry's such a klutz.
S: He's a good BADMINTON player.
L*+H LH%

In terms of the present theory, the response is an isolated theme, which achieves its effect of contradiction by: a) claiming via an LH% boundary that the hearer is committed to the proposition (even though in fact they may not be); b) claiming via the L*+H pitch-accent that the theme is not (yet) mutually agreed (even though the hearer may in fact believe its content already); and c) leaving the hearer to infer for themselves on the basis of their world knowledge about badminton players the implicated rheme, that Harry is not in fact a total klutz. The contradiction is particularly effective, because a and b between them further implicate that H's original remark was pretty stupid, and thereby force the hearer to infer this intended further conclusion for themselves, without the speaker needing to explicitly uttering it. However, this effect of the utterance is an indirect speech-act or conversational implicature, not part of the literal meaning of the words or the tones.

As an aside, it is striking that within the present theory, such conversational implicatures can be analyzed solely in terms of knowledge and modality, without appealing explicitly to notions of cooperation, flouting, or to speech-act types and illocutionary force recognition. Many of the examples discussed by Grice (1975) and Searle (1975) seem to be susceptible to similar knowledge-based analysis, making Speech-act-theoretic analyses merely emergent, as in Steedman and Johnson-Laird (1980) and Cohen and Levesque (1990).

For example, consider Grice's famous analysis of the sarcastic or ironic conversational implicature achieved by saying "You're a fine friend!" in a situation where the hearer has actually done the speaker a disservice. His analysis requires the hearer to detect that the speaker has flouted a conversational maxim (Quality), to assume that the speaker is still cooperating and therefore (by a step that is not quite clear), to infer that the speaker must mean the opposite of what they said. It is interesting however, to observe that one intonation contour with which such sarcastic comments are characteristically uttered is the following:

(22) You're a FINE FRIEND !
L* L* LH%

This all-rheme utterance is marked by the L* pitch-accents as not agreed or in Mutual Belief, and by the LH% boundary as being something the hearer is in the speaker's view committed to. It is the latter marking that makes the hearer compare the speaker's proposition with their own beliefs, and identify the Rheme Alternative Set as something like the following:

(23) $\left\{ \begin{array}{l} (-\mathit{fine}'(\mathit{friend}'\,\mathit{self}')) \\ (+\mathit{false}'(\mathit{friend}'\,\mathit{self}')) \end{array} \right\}$

At this point, the speaker has achieved their goal of making the hearer aware of their own misdeed, and the indirect speech-act is complete, without any appeal to cooperation, maxims, or rules explicitly associating maxim-violating utterances with their negation. Indeed the effectiveness of the indirect accusation is greatly increased by the fact that the speaker has, so to speak, got under the hearer's guard, forcing them into coming up with this thought for themselves, rather than stating it as a speaker commitment, which the hearer might reject. We as linguists may identify this as illocutionary uptake of an act of sarcasm, but the participants don't need to know about any of this.

5 Intonation in Combinatory Categorial Grammar

CCG is a form of lexicalized grammar in which grammatical *categories* are made up of a syntactic type defining valency and order of combination, and a logical form. For example, the English intransitive verb *walks* has the following category, which identifies it as a function from (subject) NPs (which the backward slash identifies as on the left, and the feature-value indicated by subscript *SG* identifies as bearing singular agreement) into sentences S:

(24) walks := $S\backslash NP_{SG} : \lambda x.\mathit{walk}'x$

Its interpretation is written as a λ-term associated with the syntactic category by the operator ":". The transitive verb *admires* has the category of a function from (object) nounphrases (which the forward slash identifies as on the right) into predicates or intransitive verbs:

(25) admires := $(S\backslash NP_{SG})/NP : \lambda x.\lambda y.\mathit{admire}'xy$

In this case the syntactic type is simply the SVO directional form of the semantic type. (Juxtaposition of function and argument symbols in logical forms as in $\mathit{admire}'x$ indicates function application. A convention of left associativity holds, according to which $\mathit{admire}'xy$ is equivalent to $(\mathit{admire}'x)y$).

In other cases categories may "wrap" arguments into the logical form, as in the analysis of Bach (1979, 1980), Dowty (1982), and Jacobson (1992). For example, the following is the category of the English ditransitive verb *showed*, which reverses the dominance/command relation of indirect and direct object x and y between syntactic derivation and the logical form:[8]

(26) showed := $((S\backslash NP)/NP)/NP : \lambda x.\lambda y.\lambda z.\mathit{show}'yxz$

(The reason for doing this is to capture at the level of logical form the binding theory and its dependence on the c-command hierarchy in which subject outscopes direct object, which outscopes indirect (dative) object, which outscopes more oblique arguments—see Steedman 1996 for discussion).

The syntactic operations of CCG by which such interpretations are assembled are distinguished by being strictly type-dependent, rather than structure-dependent. For present purposes

[8]The present analysis differs from that of Bach and colleagues in making Wrap a *lexical* combinatory operation, rather than a syntactic combinatory rule. One advantage of this analysis, which is discussed further in Steedman 1996, is that phenomena depending on Wrap, such as anaphor binding and control, are immediately predicted to be *bounded* phenomena.

they can be regarded as limited to operations of *type-raising* (corresponding to the combinator **T**) and *composition* (corresponding to the combinator **B**).

Type-raising turns argument categories such as *NP* into functions over the functions that take them as arguments, such as the verbs above, into the results of such functions. Thus NPs like *Harry* can take on categories such as the following:

(27) a. $S/(S\backslash NP_{SG}) : \lambda p.p\ harry'$
b. $S\backslash(S/NP) : \lambda p.p\ harry'$
c. $(S\backslash NP)\backslash((S\backslash NP)/NP : \lambda p.p\ harry'$
d. etc.

This operation has to be strictly limited to argument categories. One way to do so is to specify it in the lexicon, in the categories for proper names, determiners, and the like.

The inclusion of composition rules like the following as well as simple functional application and lexicalized type-raising engenders a potentially very freely "reordering and rebracketing" calculus, engendering a generalized notion of surface or derivational constituency.

(28) *Forward composition* (>**B**)
$X/Y : f \quad Y/Z : g \quad \Rightarrow_{\mathbf{B}} \quad X/Z : \lambda x.f(gx)$

For example, the simple transitive sentence of English has *two* equally valid surface constituent derivations, each yielding the same logical form:

(29)
$$\dfrac{\dfrac{\dfrac{\text{Harry}}{S/(S\backslash NP_{SG}) : \lambda f.f\ harry'}{>\mathbf{T}} \quad \dfrac{\text{admires}}{(S\backslash NP_{SG})/NP : \lambda x.\lambda y.admire'xy}}{S/NP : \lambda x.admire'x\ harry'}{>\mathbf{B}} \quad \dfrac{\text{Louise}}{S\backslash(S/NP) : \lambda p.p\ louise'}{<\mathbf{T}}}{S : admire'louise'harry'}{<}$$

(30)
$$\dfrac{\dfrac{\text{Harry}}{S/(S\backslash NP_{SG}) : \lambda f.f\ harry'}{>\mathbf{T}} \quad \dfrac{\dfrac{\text{admires}}{(S\backslash NP_{SG})/NP : \lambda x.\lambda y.admire'xy} \quad \dfrac{\text{Louise}}{(S\backslash NP)\backslash((S\backslash NP)/NP) : \lambda p.p\ louise'}{<\mathbf{T}}}{S\backslash NP_{SG} : \lambda y.admire'louise'y}{<}}{S : admire'louise'harry'}{>}$$

In the first of these, *Harry* and *admires* compose as indicated by the annotation >**B** to form a non-standard constituent of type S/NP. In the second, there is a more traditional derivation involving a verbphrase of type $S\backslash NP$. Both yield identical logical forms, and both are legal surface or derivational constituent structures. More complex sentences may have many semantically equivalent derivations, a fact whose implications for processing are discussed in *SP*.

This theory has been applied to the linguistic analysis of coordination, relativization, and intonational structure in English and many other languages. For example, since substrings like *Harry admires* are now fully interpreted derivational *constituents*, they can undergo coordination via the schematised rule (31), allowing a movement- and deletion- free account of right node raising, as in (32):

(31) *Simplified coordination rule* (<Φ>)
$X \quad CONJ \quad X' \Rightarrow \quad X''$

(32) [Harry admires] and [Louise detests] a saxophonist

$\frac{\frac{\frac{\text{[Harry admires]}}{S/NP}\,{>}\mathbf{B} \quad \frac{\text{and}}{CONJ} \quad \frac{\text{[Louise detests]}}{S/NP}\,{>}\mathbf{B}}{S/NP}\,{<}\Phi{>} \quad \frac{\text{a saxophonist}}{S\backslash(S/NP)}\,{<}\mathbf{T}}{S}\,{<}$

This type-dependent account of extraction, as opposed to the standard account using structure-dependent rules, makes the across-the-board condition on extractions from coordinate structures a prediction or theorem, rather than a stipulation, as consideration of the types involved in the following examples will reveal:

(33) a. A saxophonist [that$_{(N\backslash N)/(S/NP)}$ [[Harry admires]$_{S/NP}$ and [Lousie detests]$_{S/NP}$]$_{S/NP}$]$_{N\backslash N}$
b. A saxophonist that$_{(N\backslash N)/(S/NP)}$ *[[Harry admires]$_{S/NP}$ and [Lousie detests him]$_S$]]
c. A saxophonist that$_{(N\backslash N)/(S/NP)}$ *[[Harry admires him]$_S$ and [Lousie detests]$_{S/NP}$]

The availability of fully interpreted nonstandard derivational constituents corresponding to substrings like *Harry admires* was originally motivated by their participation in constructions like relativization and coordination and the desire to capture those constructions with a grammar obeying a very strict form of the Constituent Condition on Rules (*SP*, chapter 1). However, a theory that allows alternative derivations like (29) and (30) is clearly immediately able to capture the fact that prosody can make exactly the same non-standard constituents into intonational phrases, as in (34a), as easily as the standard consituents in (34b):

(34) a. HARRY admires LOUISE
L+H* LH% H* LL%
b. HARRY admires LOUISE
H* L L+H* LH%

The way that CCG derivation is made sensitive to the presence of tones is as follows (adapted from Steedman 1999). The presence of a pitch-accent on a word infects its whole category with themehood or rhemehood, via a pair of feature-values θ/ρ and $\pm$*AGREE*, the latter here abbreviated as superscript +/-. For example the transitive verb *admires* bearing an H* pitch-accent has the following category:[9]

(35) admires := $(S_{\rho^+}\backslash NP_{\rho^+})/NP_{\rho^+} : \lambda x.\lambda y.{*}admire'xy$
H*

The feature ρ ensures that a verb so marked can only combine with arguments that are compatible with rheme marking—that is, which do not bear the theme marking feature value θ—and marks its result as rheme marked as well. The element in the logical form correponding to the accented word itself is marked for k-contrast with the asterisk operator.

Boundaries, by contrast are not properties of words or phrases, but independent string elements in their own right. They bear a category which, by mechanisms parallel to those discussed in more detail in *SP*, "freezes" $\theta^{\pm}/\rho^{\pm}$-marked constituents as complete information-/intonation-structural units, making them unable to combine further with anything except similarly complete prosodic units.

[9]Number agreement is suppressed in the interests of reducing formal clutter.

For example, the hearer-responsibility signalling LH% boundary bears the following category:

(36) LH% := $S\$_\phi \backslash S\$_{\eta^\pm} : \lambda f.[H^\pm]\eta' f$

—where $S\$$ is a variable ranging over S and syntactic function categories into S, η is a variable ranging over syntactic features θ/ρ, η' ranges over the corresponding semantic translation θ'/ρ' defined in terms of the alternative semantics discussed in section 2, superscript $\pm$ is a variable ranging over $\pm AGREE$, and ϕ marks the result as a complete phonological phrase.

The derivation of (34a) then appears as follows:

(37)

Harry	admires	LH%	Louise	LL%
L+H*			H*	
$S_{\theta^+}/(S_{\theta^+}\backslash NP_{\theta^+})$: $\lambda f.f\ *harry'$ (>T)	$(S\backslash NP)/NP$: $\lambda x.\lambda y.admire'xy$	$S\$_\phi\backslash S\$_{\eta^\pm}$: $\lambda f.[H^\pm]\eta' f$	$S_{\rho^+}\backslash(S_{\rho^+}/NP_{\rho^+})$: $\lambda p.p\ *louise'$ (<T)	$S\$_\phi\backslash S\$_{\eta^\pm}$: $\lambda g.[S^\pm]\eta' g$
$S_{\theta^+}/NP_{\theta^+} : \lambda x.admire'x\ *harry'$ (>B)				
$S_\phi/NP_\phi : [H^+]\theta'(\lambda x.admire'x\ *harry')$ (<)			$S_\phi\backslash(S_\phi/NP_\phi) : [S^+]\rho'(\lambda p.p\ *louise')$ (<)	
$S_\phi : [S^+]\rho'(\lambda p.p\ *louise'))([H^+]\theta'(\lambda x.admire'x\ *harry'))$ (<)				
$S : admire'louise'harry'$				

In the last step of the derivation, the markers of speaker/hearer commitment, agreement/disagreement, and theme/rheme are evaluated with respect to the database, to check that the associated presuppositions hold or can be accomodated. In the latter case this includes support or accommodation for the relevant alternative sets, and will include updates corresponding to the new theme and rheme. If any of these presuppositions fails, then processing will block and incomprehension will result. If it succeeds, then the two core λ-terms can β-reduce to give the canonical proposition as the result of the derivation.

6 Empirical Issues

The present paper has laid a considerable burden of mening on the distinction between pitch-accent types, and in particular that between H* and L+H*, which according to the present theory are respectively the most frequent rheme accent and theme accent. It might therefore appear to be an embrassment that there is controversy in the literature over the reality of this distinction.

Part of this controversy stems from the fact that trained ToBI annotators show quite low interannotator reliabiliity in drawing this particular distinction (John Pitrelli, p.c.). When the characteristics of the actual pitch-accents annotated by them as H* and L+H* are plotted in terms of objective TILT parameters, there is very considerable overlap between the two categories (Taylor 2000).

However, this seems to be a problem with the definitions of the relevant pitch contours that are provided in the ToBI annotation conventions (Beckman and Hirschberg 1999). The distinguishing characteristic of the L+H* accent is that the rise to the pitch maximum is *late*, typically beginning no earlier than onset of the vowel in the accented syllable. H* accents typically begin to rise earlier, in many cases much earlier. The definition of L+H* in the manual as "a high peak target on the accented syllable which is immediately preceded by relatively sharp rise from a valley in the lowest part of the speaker's pitch range" does not make this entirely clear. Indeed

it is likely that the distinction can only be drawn reliably if syllable boundary alignment is taken into account, and this information is not provided in the ToBI annotation system.

It is also important to recall in using ToBI-annotated material that the manual explicitly instructs the annotator to use H* as the "default" accent type, explicitly instancing L+H* accents as examples that when in doubt should be annotated as H*.[10]

These characteristics of the ToBI annotation scheme mean that, useful though it is for other purposes, extreme caution has to be exercised in drawing strong conclusions concerning the reality of the H*/L+H* distinction from ToBI annotated corpora. In particular, while Taylors conclusion that the H*/L+H* distinction *as drawn in the annotation to the relevant section of the Boston News Corpus* is not phonetically real, it does not follow that the pitch-accent types themselves are not distinct.

It is similarly unsafe to assess the present claim that L+H* is distinctively associated with theme by applying text-based criteria for identifying topics in free text such as those proposed by Gundel (1988). The only definition of a theme that is possible under the present proposal is in terms of contextually established or accomodatable alternative sets. While the definitions in Steedman 2000a would allow restricted contexts to be manipulated to control the available alternatives, and allow the predictions concerning tune to be tested, identifying themes in free discourse is not easy, because of the pervading involvement of accomodation and inference in human discourse. For example, as Hedberg notes in her paper in the present volume, some of the L+H* accents which she finds not to be associated with topics in Gundel's sense would be classified as isolated themes in the terms of the present theory (see Hedberg and Sosa 2001, note 3; Hedberg and Sosa 2002).[11]

7 Conclusion

The system proposed here reduces the literal meaning of the tones to just three semantically grounded binary oppositions. Crucially, it grammaticalizes a distinction between the beliefs that the speaker *claims* by their utterance that the speaker is committed to, and those that the hearer actually *is* committed to. It is only the latter set that includes Mutual Beliefs. It is therefore consistent for the speaker to claim and/or implicate that both they and the hearer are committed to a proposition, but that it is not mutually believed. This is a move in the present theory that is forced by examples like (21) and the minimal pairs in (13)-(20).

The theory places a correspondingly greater emphasis on the role of speaker-presupposition (and its dual, hearer-accommodation, and by inference and implicature. To that extent, the present theory follows the tradition of Halliday and Brown, in claiming that it is the speaker who, within the constraints imposed by the context and the participants' beliefs and intentions, determines what is theme and rheme, and what contrasts they embody, and not the text.

[10] "Implicit in our discussion of the five pitch-accents is the notion that H* is the 'default' accent type. So, if there is any uncertainty about how low the F0 is before the peak, as in some cases of possible L+H* near the beginning of an utterance, the transcriber should mark 'H*' rather than 'L+H*'." (Beckman and Hirschberg 1999).

[11] Similarly, the fact that non-native speakers often obliterate pitch-accent type distinctions, and yet manage to be understood, should no more lead one to conclude that the distinctions are not real than does the possibility of written communication.

References

Bach, Emmon. 1979. "Control in Montague Grammar." *Linguistic Inquiry*, 10, 513–531.

Bach, Emmon. 1980. "In Defense of Passive." *Linguistics and Philosophy*, 3, 297–341.

Beckman, Mary, and Julia Hirschberg. 1999. "The ToBI Annotation Conventions." URL `http://ling.ohio-state.edu/ tobi/ame_tobi/annotation_conventions.html`. Ms. Ohio State University.

Bolinger, Dwight. 1958. "A Theory of Pitch Accent in English." *Word*, 14, 109–149. Reprinted in Bolinger (1965), 17-56.

Bolinger, Dwight. 1961. "Contrastive Accent and Contrastive Stress." *Language*, 37, 83–96. Reprinted in Bolinger (1965), 101-117.

Bolinger, Dwight. 1965. *Forms of English*. Cambridge, MA: Harvard University Press.

Brown, Gillian. 1983. "Prosodic Structure and the Given/New Distinction." In Anne Cutler, D. Robert Ladd, and Gillian Brown, eds., *Prosody: Models and Measurements*, 67–77. Berlin: Springer-Verlag.

Brown, Gillian, Karen Currie, and Joanne Kenworthy. 1980. *Questions of Intonation*. London: Croom Helm.

Büring, Daniel. 1997a. "The Great Scope Inversion Conspiracy." *Linguistics and Philosophy*, 20, 175–194.

Büring, Daniel. 1997b. *The Meaning of Topic and Focus: The 59th Street Bridge Accent*. London: Routledge.

Clark, Herbert. 1996. *Using Language*. Cambridge: Cambridge University Press.

Clark, Herbert, and Catherine Marshall. 1981. "Definite Reference and Mutual Knowledge." In Aravind Joshi, Bonnie Webber, and Ivan Sag, eds., *Elements of Discourse Understanding*, 10–63. Cambridge: Cambridge University Press.

Cohen, Philip. 1978. *On Knowing What to Say: Planning Speech Acts*. Ph.D. thesis, University of Toronto.

Cohen, Philip, and Hector Levesque. 1990. "Rational Interaction as the Basis for Communication." In Philip Cohen, Jerry Morgan, and Martha Pollack, eds., *Intentions in Communication*, 221–255. Cambridge, MA: MIT Press.

Cresswell, M.J. 1973. *Logics and Languages*. London: Methuen.

Cresswell, M.J. 1985. *Structured Meanings*. Cambridge, MA: MIT Press.

Dowty, David. 1982. "Grammatical Relations and Montague Grammar." In Pauline Jacobson and Geoffrey K. Pullum, eds., *The Nature of Syntactic Representation*, 79–130. Dordrecht: Reidel.

Grice, Herbert. 1975. "Logic and Conversation." In Peter Cole and Jerry Morgan, eds., *Speech Acts*, vol. 3 of *Syntax and Semantics*, 41–58. New York: Seminar Press. Written in 1967.

Gundel, Janet. 1974. *The Role of Topic and Comment in Linguistic Theory*. Ph.D. thesis, University of Texas, Austin.

Gundel, Janet. 1988. "Universals of Topic-Comment Structure." In Michael Hammond, Edith Moravcsik, and Jessica Wirth, eds., *Syntactic Universals and Typology*, 209–242. Amsterdam: John Benjamins.

Gundel, Janet, and Torsten Fretheim. 2001. "Topic and Focus." In Laurence Horn and Gregory Ward, eds., *Handbook of Pragmatic Theory*. Oxford: Blackwell.

Gunlogson, Christine. 2001. *True to Form: Rising and Falling Declaratives in English*. Ph.D. thesis, University of California at Santa Cruz. In preparation.

Gunlogson, Christine. 2002. "Declarative Questions." In Brendan Jackson, ed., *Proceedings of* Semantics and Linguistics Theory XII, 144–163. Ithaca, NY: Cornell University. To appear.

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

Hajičová, Eva, and Petr Sgall. 1988. "Topic and Focus of a Sentence and the Patterning of a Text." In Jänos Petöfi, ed., *Text and Discourse Constitution*, 70–96. Berlin: de Gruyter.

Halliday, Michael. 1963. "The Tones of English." *Archivum Linguisticum*, 15, 1.

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

Halliday, Michael. 1967b. "Notes on Transitivity and Theme in English, Part II." *Journal of Linguistics*, 3, 199–244.

Hedberg, Nancy, and Juan Sosa. 2001. "The Prosodic Structure of Topic and Focus in Spontaneous English Dialogue." In Matthew Gordon, Daniel Büring, and Chungmin Lee, eds., *Proceedings of the Workshop on Topic and Focus*. Held at the LSA Summer Institute, Santa Barbara CA. July, (to appear).

Hedberg, Nancy, and Juan Sosa. 2002. "The Prosody of Questions in Natural Discourse." In *Proceedings of Speech Prosody, Aix en Provence, Aptil*. To appear.

Hirschberg, Julia, and Janet Pierrehumbert. 1986. "Intonational Structuring of Discourse." In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics, New York*, 136–144. San Francisco, CA: Morgan Kaufmann.

Hobbs, Jerry. 1990. "The Pierrehumbert-Hirschberg Theory of Intonational Meaning Made Simple: Comments on Pierrehumbert and Hirschberg." In Philip Cohen, Jerry Morgan, and Martha Pollack, eds., *Intentions in Communication*, 313–323. Cambridge, MA: MIT Press.

Jacobson, Pauline. 1992. "Flexible Categorial Grammars: Questions and Prospects." In Robert Levine, ed., *Formal Grammar*, 129–167. Oxford: Oxford University Press.

Karttunen, Lauri, and Stanley Peters. 1979. "Conventional Implicature." In Choon-Kyu Oh and David Dinneen, eds., *Syntax and Semantics 11: Presupposition*, 1–56. New York: Academic Press.

Kartunnen, Lauri. 1976. "Discourse Referents." In J. McCawley, ed., *Syntax and Semantics*, vol. 7, 363–385. Academic Press.

Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

Lambrecht, Knud, and Laura Michaelis. 1998. "Sentence Accent in Information Questions: Default and Projection." *Linguistics and Philosophy*, 477–544.

Lewis, David. 1969. *Convention: a Philosophical Study*. Cambridge MA: Harvard University Press.

Lewis, David. 1979. "Scorekeeping in a Language Game." *Journal of Philosophical Logic*, 8, 339–359.

Montague, Richard. 1974. *Formal Philosophy: Papers of Richard Montague*. New Haven, CT: Yale University Press. Richmond H. Thomason, ed.

Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. thesis, MIT. Distributed by Indiana University Linguistics Club, Bloomington.

Pierrehumbert, Janet, and Mary Beckman. 1988. *Japanese Tone Structure*. Cambridge, MA: MIT Press.

Pierrehumbert, Janet, and Julia Hirschberg. 1990. "The Meaning of Intonational Contours in the Interpretation of Discourse." In Philip Cohen, Jerry Morgan, and Martha Pollack, eds., *Intentions in Communication*, 271–312. Cambridge, MA: MIT Press.

Prevost, Scott. 1995. *A Semantics of Contrast and Information Structure for Specifying Intonation in Spoken Language Generation*. Ph.D. thesis, University of Pennsylvania.

Prevost, Scott, and Mark Steedman. 1994. "Specifying Intonation from Context for Speech Synthesis." *Speech Communication*, 15, 139–153.

Rooth, Mats. 1985. *Association with Focus*. Ph.D. thesis, University of Massachusetts, Amherst.

Rooth, Mats. 1992. "A Theory of Focus Interpretation." *Natural Language Semantics*, 1, 75–116.

Searle, John. 1975. "Indirect Speech Acts." In Peter Cole and Jerry Morgan, eds., *Speech Acts*, vol. 3 of *Syntax and Semantics*, 59–82. New York: Seminar Press.

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

Silverman, Kim, Mary Beckman, John Pitrelli, Marie Ostendorf, Colin Wightman, Patti Price, Janet Pierrehumbert, and Julia Hirschberg. 1992. "ToBI: A Standard for Labeling English Prosody." In *Proceedings of the International Conference on Spoken Language Processing, Banff, Alberta*, 867–870. Edmonton: University of Alberta.

von Stechow, Arnim. 1981. "Topic, Focus and Local Relevance." In Wolfgang Klein and Willem Levelt, eds., *Crossing the Boundaries in Linguistics*, 95–130. Dordrecht: Reidel.

Steedman, Mark. 1991. "Structure and Intonation." *Language*, 67, 262–296.

Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, MA: MIT Press.

Steedman, Mark. 1999. "Connectionist Sentence Processing in Perspective." *Cognitive Science*, 23, 615–634.

Steedman, Mark. 2000a. "Information Structure and the Syntax-Phonology Interface." *Linguistic Inquiry*, 34, 649–689.

Steedman, Mark. 2000b. *The Syntactic Process*. Cambridge, MA: MIT Press.

Steedman, Mark. 2002. "Towards a Compositional Semantics for English Intonation." URL `http://www.cogsci.ed.ac.uk/~steedman/papers.html`. Ms. University of Edinburgh.

Steedman, Mark, and Philip Johnson-Laird. 1980. "Utterances, Sentences, and Speech-Acts: Have Computers Anything to say?" In Brian Butterworth, ed., *Language Production 1: Speech and Talk*, 111–141. London: Academic Press.

Steedman, Mark, and Ivana Kruijff-Korbayová. 2001. "Two Dimensions of Information Structure in Relation to Discourse Semantics and Discourse Structure." *Journal of Logic, Language, and Information*. To appear: Introduction to the Special Issue on Information Structure, Discourse Semantics, and Discourse Structure.

Stone, Matthew. 1998. *Modality in Dialogue: Planning Pragmatics and Computation*. Ph.D. thesis, University of Pennsylvania.

Taylor, Paul. 2000. "Analysis and Synthesis of Intonation Using the Tilt Model." *Journal of the Acoustical Society of America*, 107, 1697–1714.

Thomason, Richmond. 1990. "Accomodation, Meaning, and Implicature." In Philip Cohen, Jerry Morgan, and Martha Pollack, eds., *Intentions in Communication*, 325–363. Cambridge, MA: MIT Press.

Vallduví, Enric. 1990. *The Information Component*. Ph.D. thesis, University of Pennsylvania.

Vallduví, Enric, and Maria Vilkuna. 1998. "On Rheme and Kontrast." In Peter Culicover and Louise McNally, eds., *Syntax and Semantics, Vol. 29: The Limits of Syntax*, 79–108. San Diego, CA: Academic Press.

Ward, Gregory, and Julia Hirschberg. 1985. "Implicating Uncertainty: the Pragmatics of Fall-Rise Intonation." *Language*, 61, 747–776.