Information structure in discourse:
Towards an integrated formal theory of pragmatics[1]

Craige Roberts
The Ohio State University

Abstract A framework for pragmatic analysis is proposed which treats discourse as a game, with context as a scoreboard organized around the *questions under discussion* by the interlocutors. The framework is intended to be coordinated with a dynamic compositional semantics. Accordingly, the context of utterance is modeled as a tuple of different types of information, and the questions therein — modeled, as is usual in formal semantics, as *alternative sets* of propositions — constrain the felicitous flow of discourse. A requirement of *Relevance* is satisfied by an utterance (whether an assertion, a question or a suggestion) *iff* it addresses the question under discussion. Finally, it is argued that the prosodic *focus* of an utterance canonically serves to reflect the question under discussion (at least in English), placing additional constraints on felicity in context.

Keywords: discourse, pragmatics, context of utterance, question under discussion, relevance, focus

---

For many linguists interested in pragmatics, including the Prague School theorists, Halliday (1967) and Vallduví (1993), *information structure* is a sentence level structure. It is generally characterized as a variation of sentential structure along certain parameters to modulate the presentation of the information imparted by the sentence in such a way as to relate that information to prior context. The factors in that relationship are characterized in terms of primitive functional roles such as theme/rheme, focus/(back)ground, topic/link, old/new, etc. These primitives and the correlated information structure of sentences are used to explain the roles of intonational focus and of particular syntactic focus constructions, of topicalization and other displacement transformations and of a fairly wide range of other phenomena, including word order in relatively free word order languages such as German, the functions of distinguished structural positions in Catalan or Hungarian, definiteness in Slavic languages, specificity in Turkish, use of specific affixes such as Japanese *-wa*, etc.

Most researchers who work on focus, topic, etc., talk about their relationship to felicity and hence to the context of utterance. But few look beyond the sentence to examine in any detail the range of kinds of contexts in which an utterance with a given focus is actually felicitous.[2] Kadmon & Roberts (1986) argued that this methodology was crucial to understanding the role of focus in interpretation. They claimed that at its most abstract, this role is that of constraining (via conventional presuppositions) what they called the *information structure of contexts* in which the utterance would be felicitous. This perspective on information structure stems ultimately from the work of Carlson (1982), though he himself realized it with a text grammar.[3] The conception I will develop here is one in which information structure is literally a structure on information — on the inquiry pursued in discourse and the information which that inquiry yields — and not on the utterances or sequences of utterances used to present it. All that is given at the sentential level, conventionally, are certain sorts of presuppositions about the place and function of the utterance in the information structure of the discourse in which it occurs.

---

2 Ellen Prince and her students are generally exceptions to this criticism.

3 After the ideas related here were worked out, I received a copy of Ginzburg's (1996) article. He also follows Carlson and proposes something related to the conception of information structure related here (though he doesn't call it such) and takes very different turns at a number of points. His intended application seems to be more narrowly to characterize well-formed exchanges in discourse.

I first define, in section 1, the intended notion of the *information structure of a discourse* in an extension of Stalnaker's (1978) *possible worlds view of context*, an extension which takes into account questions as well as assertions and imposes a pragmatically-motivated structure on well-formed exchanges. Then in section 2, to illustrate how this structure provides an account for the types of phenomena classically pursued in theories of information structure, I will outline a theory of a phenomenon very directly related to information structure — the conventional implicatures associated with prosodic focus in English utterances — and suggest how this account improves on others currently in the literature.

I will assume that information structure is a universal of human discourse, not dependent on the specific syntactic structures or other conventions which are developed in a given language to fulfill the associated functions. But the account of the meanings associated with prosodic focus in English is very language-specific; we might expect that other languages would use very different means to achieve some of the same ends or would use similar means to encode other kinds of information. The point of combining the two accounts in one discussion is to use the second to illustrate how the first enters into the interpretation of particular utterances in particular languages. But I think that one of the virtues of the type of information structure I have in mind is that it has much broader implications for formal pragmatic theory, providing a unified foundation for accounts of phenomena otherwise considered to be quite diverse and independent of each other. In section 3, I will very briefly note some of its other potential applications and the relationship of this proposal to a number of others in artificial intelligence, linguistics and philosophy.

1   Information structure and questions in a language game

Discourse is organized around a series of *conversational goals* and the plans or *strategies* which conversational participants develop to achieve them. This should not be surprising, given the central role granted to intentions in communication and meaning, e.g., by Grice (1989), Lewis (1969), Grosz & Sidner (1986), Planning Theorists in artificial intelligence, Sperber & Wilson (1986) and Thomason (1990); but an integrated theory of this intentional structure and its ramifications for pragmatic theory has not, to my knowledge, been developed previously.

Following Stalnaker (1978), I assume that the primary goal of discourse is communal inquiry — the attempt to discover and share with the other interlocutors "the way things are," i.e., to share information about our world. But we must develop strategies for achieving this goal, and these strategies involve subinquiries. As in a game, some strategies may be better, some worse; this is largely a matter of the rationality of the participants and not of linguistic competence *per se.* Whether strategies are effective involves, as well, an element of luck, as in any inquiry. To understand better what I have in mind, let us pursue the analogy with games.[4]

The principal elements of a game are its *goals*, the *rules* which players abide by, the *moves* they may make towards achieving the goals and the *strategies* they may pursue in making their moves, the last generally constrained by the first three and, above all, by rational considerations. We take the aims or goals of a language game, as already noted, to be to come to agree on the way things are in the world. Using Stalnaker's notion of the *common ground* (the set of propositions which the interlocutors in a discourse behave as if they all hold to be true, with a proposition realized technically as a set of possible worlds) and related *context set* (the intersection of the common ground, the set of worlds where all the propositions in the common ground are true), our goal is to reduce the context set to a singleton set, the actual world. There are two types of rules, both viewed as constraints on the interlocutors' linguistic behavior: *conventional rules* (syntactic, compositional semantic, etc.) and *conversational rules* (e.g., the Gricean maxims). The latter are not properly linguistic but are given by rational considerations in view of (i) the goal of the game (e.g., the Maxim of Cooperation follows from the fact that playing the language game is a coordination problem, *à la* Lewis; Quality from the fact that truth is the ultimate goal; Quantity 1 from the desire to maximize the payoff of a move in view of commitment to the ultimate goal) and (ii) human cognitive limitations (cf. Sperber & Wilson's (1986) discussion of Relevance and Quantity 2 in this light).[5] There are two types of moves which players may make, linguistic behaviors which fall under the kinds of acceptable behavior defined by the rules and are classified on the basis of

---

4 In this, I follow Carlson (1982), though I have developed the idea somewhat differently here. Of course, ultimately the comparison of discourse to a game goes back to Wittgenstein 1953 and has been picked up by others, including Lewis (1969) and, of course, Carlson's mentor Hintikka (1972, 1981). (See also Hintikka & Saarinen 1979.)

5 Here and below, I capitalize the Gricean notion of *Relevance* and a related notion (which I will formally define below) to distinguish them from the ordinary English term.

their relationship to the goals of the game (I ignore here the general class of imperatives): what Carlson (1982) calls *setup moves*, which are *questions*, and what he calls *payoff moves*, which are *assertions*, the answers to questions. Note that moves, on the interpretation I will give them, are not speech acts but the semantic objects which are used in speech acts: a speech act is the act of proffering a move. I will return to discuss strategies of inquiry just below.

I assume that there are two aspects to the interpretation of any given move: its *presupposed content* and its *proffered content*. I use the term *proffered* as a cover for what is asserted in an assertion and for the nonpresupposed content of questions and commands. Lewis (1969) treats questions as a type of imperative; this strikes me as correct in that a question, if accepted, dictates that the interlocutors choose among the alternatives which it proffers.[6] Most contemporary semantic analyses regard a question as denoting or determining the set of propositions which are the possible answers (or correct answers, in some theories) to the question; these are the proffered alternatives. If a question is accepted by the interlocutors, this commits them to a common goal, finding the answer; like the commitment to a goal in Planning Theory, this is a particularly strong type of commitment, one which persists until the goal is satisfied or shown to be unsatisfiable. The accepted question becomes the immediate topic of discussion, which I will also call the *immediate question under discussion*, often abbreviated as the *question under discussion*.

Note that Stalnaker's goal of discourse can itself be viewed as a question, the Big Question, *What is the way things are?*, whose corresponding set of alternatives is the set of all singleton sets of worlds in the context set at a given point in discourse. This suggests another way of viewing the set of alternatives proffered by a question: a question sets up a partition on the context set at the point of utterance, each cell the set of worlds in which one complete answer to the question is true (cf. the use of partitions to characterize answer sets in Groenendijk & Stokhof 1984). Then we can view the context set itself as representing the ultimate set of alternatives, for it is the selection of a unique ("actual") world which is our ultimate goal.

---

6 There are also reasons one might want to treat questions and imperatives as distinct types of speech acts. Since I'm not addressing imperatives here, I won't try to argue one way or the other.

When interlocutors accept a question, they form an intention to answer it, which intention is entered into the common ground.[7] If a cooperative interlocutor knows of this intention, she is committed to it, i.e., she herself (ostensibly) has an intention to answer the question. Then Relevance, an organizing principle of discourse which supports coherence and hence facilitates the processing and storage of information, will lead her to attempt to answer it as soon as possible after it is asked. Grice's Quantity 1 maxim, in view of the goals of discourse, makes a complete answer preferable to a partial one.

Assertions are, as for Stalnaker, choices among alternatives. If accepted, they are added to the common ground and thereby shrink the context set. In order for discourse to be coherent (obey Relevance), it must be clear what alternatives (corresponding to cells in a partition on the context set) a given assertion selects among. The relevant alternatives are those proffered by the question, or topic, under discussion. That's the sense in which assertions are payoff moves — they choose among the alternatives proffered by a setup move/question, and in so doing they further the goals of the game. Non-sequiturs are assertions which don't bear on the question under discussion; even though informative in the abstract, they reflect poor strategy and a lack of commitment to the immediate goals of the discourse, i.e., a lack of cooperation. Non-sequiturs also fail to maximize payoff; good strategists make assertions with a view to optimizing the number of relevant inferences that they will trigger, and it seems reasonable to assume that such inferences are facilitated by the discourse segmentation induced by the plan structure of the discourse (see Grosz & Sidner 1986, Sperber & Wilson 1986, though the latter work doesn't use the term *discourse segmentation*).

Strategies of inquiry are sequences of setup moves, or questions, designed to (at least partially) satisfy the aims of the game while obeying the game's constraints. Given that the main goal is to answer the Big Question, a reasonable strategy will involve a plan to do this by developing subgoals which are easier to achieve and are logically related to each other in such a way as to facilitate achieving the main goal. We can define an entailment relation

---

7 This is in distinction to Carlson's (1982) epistemic desideratum of a question, which has to do with increasing the knowledge of the questioner, or with the related views of Ginzburg (1994). On the present view, it is the common ground, not the speaker, that's "informed," and it is mutual-belief behavior, and not knowledge, that's sought. This permits a generalization over rhetorical questions, quiz questions, etc., which are problems for more solipsistic views of information in discourse.

on questions, following Groenendijk & Stokhof (1984: 16): *one interrogative q_1 entails another q_2 iff every proposition that answers q_1 answers q_2 as well.* (This presupposes that we're talking about complete answers, for otherwise the entailments can actually go the other way around.) For example: *What do you like?* entails *What food do you like?*. An answer to the Big Question, *What is the way things are?*, entails the answer to any other possible question. We might call q_1 in such a relation the *superquestion*, and any q_2 which it entails we might call a *subquestion*. On the other hand, if we can answer enough subquestions, we have the answer to the superquestion. Given the ultimate aim of discourse and the rationality of the participants, these types of relations are the principal factors that structure our moves.

Of course, questions in discourse are generally more specific, and hence more manageable, than the Big Question. Besides the discourse goal of inquiry in its most general sense, we usually have goals in the real world, things we want to achieve quite apart from inquiry, *domain goals*. And our domain goals, in the form of deontic priorities, generally direct the type of inquiry which we conduct in conversation, the way we approach the question of how things are. We are, naturally, most likely to inquire first about those matters that directly concern the achievement of our domain goals.

Hence, domain goals tend to dictate which subquestions of the Big Question we take up at any given point. But, as discussed above, once we've committed ourselves to a given question, i.e., once we plan to answer it, then we pursue it until and unless it is either answered or it becomes clear that it isn't presently answerable. The interlocutors' strategy in this pursuit may, however, include the decision to pursue answers to subquestions, i.e., a series of related questions may realize a strategy to get at the answer to the most general, logically strongest question among them. Hence, a strategy of inquiry will have a hierarchical structure, a set of questions partially ordered by entailment. Things are actually more complex than this, as questions in an actual strategy may be only logically related in view of certain contextual entailments. But this is the basic nature of strategies, and in what follows I will assume that they have this idealized logical structure, relativized to context.

Of course, there are many sequences of questions in discourse which are apparent counterexamples to the constrained relation posited here. For example, at a given point an interlocutor might utter one after the other a number of relevant subquestions to some superquestion, without yet answering any of them. But this is just enumeration, suggesting a plan

for how to attack the superquestion; none of these questions has yet been proffered or accepted. We can only address such subquestions one at a time, and when one is being addressed, we stick to it (or can complain to interlocutors who don't and thereby "change the subject"). In another type of case, someone might have an epiphany at some point, realize how some information just discussed bears on another, now dormant question and introduce that. This can derail the conversation, and either we later try to get back on track with the strategy we were pursuing or else we just give it up and move on. But these are violations of the rules which only underline their general force. See also the brief discussion of the questions under discussion in corrective contexts in section 2.1, where these are a type of metaquestion about the discourse itself.

One of Carlson's (1982) central insights is that dialogues are functionally organized by question/answer relations, though the questions are often only implicit, inferred on the basis of other cues. We will see in the following section one way in which interlocutors can retrieve questions under discussion which are only implicit, never explicitly asked: I will argue that prosodic focus in English presupposes the type of question under discussion, a presupposition which enables the hearer, with some other contextually given clues, to reconstruct that question and its relation to the strategy being pursued. This is just one example of the general case, which is modeled more abstractly in Planning Theory via Plan Inferencing Rules that permit us to infer interlocutors' plans from other information in the common ground plus what is actually said. Similarly, sometimes answers which are obviously entailed in a given context are not explicitly uttered but are nonetheless entered into the common ground. These cases involve *accommodation* in Lewis's (1979) sense and are quite normal in discourse: if it is clear that an interlocutor presupposes a question or assertion ϕ which is not yet commonly agreed upon but the others have no objection, then they behave as if the common ground contained ϕ all along. Hence, the notion of a move in a discourse game is essentially semantic. A question is not necessarily realized by a speech act but is only a question denotation in the technical sense that it proffers a set of relevant alternatives which the interlocutors commit themselves to addressing:[8] it tells you what the discourse is "about" at that

---

8 One of the nice consequences of this semantic realization of the notion of questions under discussion, as opposed to Carlson's assumption that implicit questions are actual parts of the text of a discourse, is that these more abstract questions needn't carry the types of presuppositions, especially existential presuppositions which are sometimes said to be associated with actual interrogative speech acts.

point in the discourse, and further, if we look at the strategy of questions in which it participates, it tells us where the discourse is going.

In the remainder of this section, I will first, as a preliminary to developing a view of discourse as founded on questions, sketch a semantics for questions in section 1.1. Then I will turn to a formal characterization of the notion of information structure that I have in mind in section 1.2. Finally in section 1.3, I will return to a consideration of the pragmatics of questions and answers, viewed as a corollary of their roles in information structure.

1.1 A semantics for questions

I adopt a semantics for questions which borrows various elements from the earlier work of Hamblin (1973), Groenendijk & Stokhof (1984) and von Stechow (1991), though it is different from each of these in important respects. As in Hamblin's account, a question will denote a set of alternatives, which I will call the *q-alternative set* of the question. I borrow from Groenendijk & Stokhof the use of the partition established over a set of worlds by the q-alternative set to define the notions of *complete* and *partial answer* to a question. And I adopt von Stechow's assumption that the q-alternatives of a question are all asked by the interlocutors and that this fact influences the common ground, though I make this part of the pragmatics of what it is to accept a question instead of part of its semantic denotation as in von Stechow 1991. This account also differs from von Stechow's and from Groenendijk & Stokhof's in that I eschew a structured semantics for questions, nor do I cogenerate question/answer pairs. In this account, the intimate relation between an assertion and the question it addresses, and hence (partially) answers, is generally reflected in prosodic focal presuppositions of the assertion. The latter are argued in section 2 below to be presuppositions about the role of the assertion with respect to the question under discussion in the information structure of the discourse. Throughout, I ignore the important question of how to analyze embedded questions. I assume that their semantics is different from that of main clause questions, though, of course, the two are closely related; cf. Jacobson 1995, Ginzburg 1995a,b and Higginbotham 1996 for some relevant discussion.

In what follows, for any constituent α, $|\alpha|$ is the regular denotation of α, arrived at by recursive, compositional rules in the usual fashion. The logical form of a question is assumed to include a wide-scope interrogation operator, '?'. Within its scope, the *wh*-clause in a question like *Who arrived?* is assumed

to be *who(λx . *x arrived*), so that the *wh*-element is in prenex position, with the main clause interpreted as an abstract on its trace. For convenience, I will confine the *wh*-elements considered to *who* and *what* since they are of the same semantic type. For a simple *wh*-question, the q-alternative set denoted by the question is the set of propositions which can be derived by abstracting on its *wh*-phrase, then applying this abstract in turn to all the things in the model (entities, functions, whatever) which are of the same type as that denoted by the *wh*-phrase. This can be generalized as follows:

(1)   The *q-alternatives* corresponding to utterance of a clause α:

$$\text{q-alt}(\alpha) = \left\{ p : \exists u^{i-1}, \dots, u^{i-n} \in D \left[p = |\beta| \left(u^{i-1}\right) \dots \left(u^{i-n}\right) \right] \right\}$$

where α has the logical form $wh_{i-1}, \dots, wh_{i-n}(\beta)$, with $\{wh_{i-1}, \dots, wh_{i-n}\}$ the (possibly empty) set of *wh*-elements in α, and

where D is the domain of the model for the language, suitably sortally restricted, e.g., to humans for *who*, nonhumans for *what*.

Definition (1) yields a set of alternatives for any utterance, even nonquestions.[9] We abstract over any *wh*-elements there may be in the utterance and then permit the variables of abstraction to vary freely over entities of the appropriate sort in the model.[10] Yes/no questions contain no *wh*-elements, so by (1) the set of q-alternatives to a yes/no question ?(α) will be just the singleton set $\{|\alpha|\}$.

The semantics for questions is quite simple. The denotation of a question is its q-alternative set:[11]

---

9 Note that this set differs from the alternatives determined by a question for von Stechow (1991), since his alternatives vary simultaneously over both the *wh*-elements in the question *and* its focused constituents, while I have separated out these two types of variation, as will become clear in section 2 below.

10 It is not crucial here that the *wh*-elements be treated as operators. In the logical forms assumed here, the role of the prenex *wh*-elements is primarily to mark their scope. Otherwise, I might have assumed that the *wh*-elements themselves are distinguished variables, following Ginzburg 1995a,b, with a special null operator, like his *QUANT-CLOSURE* operator, unselectively binding all the free *wh*-variables at the appropriate level. The facts about scoping of *wh*-elements, especially relative to scope of other types of operators, remain unclear to me, as does the empirical content of positing null operators. Hence, I leave this issue unaddressed here.

11 I adopt this semantics without argument here. It is just the simplest type of denotation for questions consistent with the alternative semantic account of focus and questions which I

(2)   The *interpretation of a question* $?(\alpha)$:

$$|\,?(\alpha)\,| = \text{q-alt}(\alpha)$$

For a question like *Who did Mary invite?*, if we take *who* to be of type *e*, the q-alternative set described would not include the answer *Nobody*, though this is presumably felicitous. So the q-alternative set isn't necessarily the set of *all* possible answers to the question. Of course, if we assumed that *who* is of a higher type, say a generalized quantifier of type $\langle\langle e, t\rangle, t\rangle$, then *nobody* would be of the appropriate type to give as value for the corresponding abstraction. But as we will shortly see, we won't have to assume anything about the type of *who* in order to permit such an answer.

To answer a question, we must evaluate all the elements of its q-alternative set to see which, if any, are true. Answerhood is defined as follows:

(3)   a. A *partial answer* to a question q is a proposition which contextually entails the evaluation — either true or false — of at least one element of q-alt(q).

        b. A *complete answer* is a proposition which contextually entails an evaluation for each element of q-alt(q).

Suppose we have only two propositions in the q-alternative set for some question, p and p'. The set of complete answers to the question are those which entail one member of the set in (4):

(4) $\{(p \wedge p'), (p \wedge \neg p'), (\neg p \wedge p'), (\neg p \wedge \neg p')\}$

Note that this set establishes a partition on any given set of possible worlds, including the context set for a given utterance; for in any given world, one and only one of these four possible formulae can be true, and in an interesting set of worlds, there will be some in which each of these formulae is true. Each complete answer corresponds to a cell in the partition and its acceptance leads to discarding all the other cells, as in the informal discussion in the preceding section.

To illustrate, consider (5) and its denotation, derived in (6):

(5)   Who did Mary invite?

(6) $\big|\,?(\text{who}(\lambda x \, . \, \text{Mary invited } x))\,\big| = \text{q-alt}(\text{who}(\lambda x \, . \, \text{Mary invited } x))$

---

develop below. Nothing in the proposals about information structure and focus hinges on (2), so that one might substitute another type of interpretation for questions.

$$= \{p : \exists u \in D\,[\,p = \mid \lambda x \,.\, \text{Mary invited } x \mid (u)\,]\}$$
$$= \{\text{Mary invited } u : \ u \in D\}$$

If the proffered (5) is accepted by the interlocutors, they will then be committed to answering the question, i.e., evaluating all the alternatives it proffers. Suppose that there are only three people in the model, so that $D = \{\text{Mary, Alice, Grace}\}$; then the denotation of (5) will be the set in (7) (assuming the relation denoted by *invite* is irreflexive and ignoring joins of individuals in the model):

(7)   {Mary invited Alice, Mary invited Grace}

If we take *Mary invited Alice* to be p and *Mary invited Grace* to be p', then (5) corresponds to the partition suggested by (4). If the partition is established on a reasonably rich context set, the set of complete answers will include *Mary invited Alice and Grace*, *Mary invited Alice but not Grace*, *Mary invited Grace but not Alice* and *Mary invited nobody*, the last of these picking out the cell corresponding to the formula $\neg p \wedge \neg p'$. A partial answer would be an utterance whose truth rules out at least one of the cells of the partition, e.g., *Mary didn't invite Grace*, ruling out both of the cells in which p' is true but leaving open the question of whether Mary invited Alice. All complete answers are partial answers but not *vice versa*.

Given the way we have defined answerhood in (3), we can define some other useful terms for discussing questions:

(8)   A question q_1 *entails* another question q_2 *iff* answering (i.e., giving an answer to) q_1 yields a complete answer to q_2.

(cf. Groenendijk & Stokhof 1984: 16)

(9)   A question q_1 *contextually entails* another q_2 *iff* answering q_1 in a discourse context with common ground C (a set of propositions) is such that $C \cup \text{Ans}(q_1)$ entails a complete answer to q_2.

I will assume without argument that an utterance may partially answer a question either by directly asserting a partial answer, by contextually entailing a partial answer or by presupposing or conversationally implicating something which contextually entails a partial answer to the question.

The semantics for questions I have given in this section is essentially static. It will serve the purposes of the static view of information structure presented in the following section. However, in section 1.3 I will consider the question of how presuppositions are projected in questions, and this will

necessitate a dynamic view of questions and of the information structure in which they play a role. Hence, I will present a *context change potential* for questions in dynamic information structure. However, though defining an update function over a set of q-alternatives is rather more complex than giving static semantic interpretations, the simple semantics presented here still represents the heart of the proposal, along with the notions of answerhood and question entailment.

1.2 A formal theory of information structure

In this section, I will explore a formal definition of the *information structure* of a discourse, taking it to be the set of moves in the discourse, both questions and answers, explicit and implicit (pragmatically retrieved), along with various functions and relations on them, including the structure induced by the strategy or strategies which order the questions. The question/answer relation gives us pairs of moves, the pairs in turn partially ordered by the strategic relation on the questions. Note that just as questions may be only indirectly "asked," answers may in fact be provided by a sequence of one or more assertions; deriving an answer may even require inferences drawn on the basis of such sequences, so that the structure is more complex than that of a simple question/answer dialogue, though I cannot investigate such matters in detail here.

There are a couple of ways I could imagine defining information structure formally. Here is one:

(10)   The *information structure* for a discourse $\mathcal{D}$ is a tuple:[12]

$$\text{InfoStr}_{\mathcal{D}} = \langle M, Q, A, <, Acc, \text{CG}, \text{QUD} \rangle, \text{ where:}$$

   a. M is the set of (setup and payoff) moves in the discourse.
   b. Q is the set of questions (setup moves) in M, where a question is a set of propositions.

---

12 One might want to define a discourse as the set of explicit moves made in a period of time between a set of interlocutors. In that case, there might be different sets of possible implicit interpolated moves that the interlocutors could "agree" on, with different resulting information structures. Hence, one would talk not about *the* information structure for the discourse but about *an* information structure for the discourse. In what follows, I assume that a discourse consists of all the moves made within it, implicit as well as explicit, though this isn't crucial. In either case, one might also add a set *Exp* to the tuple, the subset of *M* which consists of all the explicit moves in the discourse.

c. A is the set of assertions (payoff moves) in M, where an assertion is a set of possible worlds.

d. $<$ is the precedence relation, a total order on M:

$$m_i < m_k \text{ iff } m_i \text{ is made/uttered before } m_k \text{ in } \mathcal{D}$$

The order of any two elements under $<$ will be reflected in the natural order on their indices, where for all m_i, $i \in \mathbb{N}$.

e. Acc is the set of accepted moves in M.

f. CG is a function from M to sets of propositions, yielding for each $m \in M$ the common ground of $\mathcal{D}$ just prior to the utterance of m. Further, we require the following:

   i. For all $m_k \in M$, $\mathrm{CG}(m_k) \supseteq \bigcup_{i<k}(\mathrm{CG}(m_i))$.

   ii. For all $m_k \in M$, $\mathrm{CG}(m_k) \supseteq \{m_i \colon i < k \text{ and } m_i \in Acc \setminus Q\}$.

   iii. For all $m_k, m_i \in M$, $i < k$:

   1. The proposition that $m_i \in M$ is in $\mathrm{CG}(m_k)$.
   2. If $m_i \in Q$, the proposition that $m_i \in Q$ is in $\mathrm{CG}(m_k)$.
   3. If $m_i \in A$, the proposition that $m_i \in A$ is in $\mathrm{CG}(m_k)$.
   4. If $m_i \in Acc$, the proposition that $m_i \in Acc$ is in $\mathrm{CG}(m_k)$.
   5. For all propositions $p \in \mathrm{CG}(m_i)$, the proposition that $p \in \mathrm{CG}(m_i)$ is in $\mathrm{CG}(m_k)$.
   6. Whatever the value of $\mathrm{QUD}(m_i)$, the proposition that that is the value of $\mathrm{QUD}(m_i)$ is in $\mathrm{CG}(m_k)$.

g. QUD, the *questions-under-discussion stack*, is a function from M (the moves in the discourse)[13] to ordered subsets of $Q \cap Acc$ such that for all $m \in M$:

   i. For all $q \in Q \cap Acc$, $q \in \mathrm{QUD}(m)$ *iff*

   1. $q < m$ (i.e., neither m nor any subsequent questions are included), and
   2. $\mathrm{CG}(m)$ fails to entail an answer to q and q has not been determined to be practically unanswerable.

   ii. $\mathrm{QUD}(m)$ is (totally) ordered by $<$.

---

13 Even assertions or questions which are rejected by the participants can be determined to have been felicitous (or not) in terms of their relation to the information structure of the discourse in which they occur. Hence, the domain of QUD isn't restricted to the accepted moves.

      iii. For all $q, q' \in \text{QUD}(m)$, if $q < q'$, then the complete answer to q' contextually entails a partial answer to q.

Since we are only considering questions and assertions, I will assume that the set of assertions in $\mathcal{D}$ is the complement of Q in M.

Note that not all moves in a discourse are necessarily accepted, so that *Acc* will generally be a proper subset of the set M. Presumably, we keep track of proposed questions and assertions, even if they're rejected, an accounting which is crucial, e.g., in explaining denials and corrections. But unlike Carlson's (1982) view of discourse, the moves that we keep track of are semantic entities, the information expressed by the utterances in the discourse, and not structural analyses of those utterances. This seems desirable in view of the psycholinguistic literature which suggests a fairly rapid loss of structural information about preceding discourse.

The constraint on the value of CG for any move, making it a superset of the common ground for any previous move, guarantees that the common ground will be monotonic, preserving information contributed earlier. Of course, discourse isn't always monotonic in this sense, but I will ignore this problem here. We also require that the common ground include all those previous accepted moves which are not questions, i.e., the previous accepted assertions, in keeping with Stalnaker's (1978) characterization of what it is to accept an assertion. By making the common ground a possibly proper superset of previous common grounds and of the set of accepted assertions, we leave open the possibility that additional information is added not only by accepted payoff moves/assertions but also possibly by accommodated implicatures, inferences, common perceptual experience, etc. Clause (iii) of (10f) is intended to capture the fact that at any given point in the discourse, the interlocutors have complete information about the information structure itself, including what moves have been made, which were questions and which assertions, which were accepted, what was in the common ground at the point a given move was made and what questions were under discussion at that point.

The set of questions under discussion at a given point in a discourse is modelled using a push-down store, which I call QUD, the *questions-under-discussion stack*. Intuitively, QUD yields the ordered set of all as-yet unanswered but answerable, accepted questions in Q at the time of utterance of q. When we accept a question, we add it to the top of the stack. Its relationship to any question previously on top will be guaranteed by a combination of Relevance, entailing a commitment to answering prior questions, and logical

constraints on the way that the stack is composed. If we decide to pursue an accepted question by asking a subquestion, we may add the subquestion to the stack, so that the stack reflects (part of) a strategy of questions. When a question is answered or determined to be practically unanswerable, it is popped off the stack, revealing any below it. At any point in discourse, the question on top of the stack is the (immediate) question under discussion.

I will discuss directly below the motivation for clause (iii) in the definition of QUD in (10g). But first, so that this discussion is not entirely abstract, let me illustrate how the QUD stack works with a very simple and rather excessively explicit discourse. This example presupposes a model with only two individuals, Hilary and Robin, and two kinds of foods, bagels and tofu. I assume that each of the questions in the discourse ($\mathcal{D}_0$) is accepted by the interlocutors:

($\mathcal{D}_0$) 1. Who ate what?

        a. What did Hilary eat?

            i. Did Hilary eat bagels?
               $\mathrm{Ans}(a_i) = $ yes

          ii. Did Hilary eat tofu?
               $\mathrm{Ans}(a_{ii}) = $ yes

        b. What did Robin eat?

            i. Did Robin eat bagels?
               $\mathrm{Ans}(b_i) = $ yes

          ii. Did Robin eat tofu?
               $\mathrm{Ans}(b_{ii}) = $ yes

This entire discourse realizes a strategy to answer the first question, move 1; here, the questions involved stand in simple entailment relations. For each question in ($\mathcal{D}_0$), the set of questions which it entails is given below:

$$\vDash (1) = \{a, a_i, a_{ii}, b, b_i, b_{ii}\}$$
$$\vDash (a) = \{a_i, a_{ii}\}$$
$$\vDash (b) = \{b_i, b_{ii}\}$$

(These relations are also reflected in the indentation hierarchy in the layout of the discourse.) Note also that the following facts hold in the model described, where the model's domain is $\{\mathrm{Hilary}, \mathrm{Robin}, \mathrm{bagels}, \mathrm{tofu}\}$:

(11) a. $\mathrm{Ans}(a_i) \cap \mathrm{Ans}(a_{ii}) = \mathrm{Ans}(a)$, i.e., giving complete answers to a_i and a_{ii} yields a complete answer to a. Answering a_i hence yields a partial answer to a, and similarly for a_{ii}.

  b. $\mathrm{Ans}(b_i) \cap \mathrm{Ans}(b_{ii}) = \mathrm{Ans}(b)$, i.e., giving complete answers to b_i and b_{ii} yields a complete answer to b. Answering b_i hence yields a partial answer to b, and similarly for b_{ii}.

  c. $\mathrm{Ans}(a) \cap \mathrm{Ans}(b) = \mathrm{Ans}(1)$, i.e., giving complete answers to a and b yields a complete answer to 1. Answering a hence yields a partial answer to 1, and similarly for b.

  d. Since partial answerhood is transitive under clause (iii) of (10g), answering a_i, a_{ii}, b_i or b_{ii} yields a partial answer to 1.

Intuitively, it is partly because of these four facts that $(\mathcal{D}_0)$ realizes a success-ful strategy for answering 1. As each question in $(\mathcal{D}_0)$ is asked, it is added to the QUD stack, with 1 on the bottom. When a subquestion is answered, e.g., a_i, that question is popped from the stack and the answer added to the common ground. When a_i and a_{ii} have been answered, the common ground then entails the answer for a, as stated in the first fact in (11), and then a is popped as well. When both a and b have been answered in this way, the common ground yields the answer for a, as reflected in the third fact in (11); then 1 is popped as well, leaving the QUD stack empty (insofar as this discourse comes out of the blue).

   The logical relations among the questions in $(\mathcal{D}_0)$ are also what guarantee that the QUD stack for $(\mathcal{D}_0)$ satisfies clause (iii) in the definition of QUD in (10g), which basically requires that higher questions on the stack be subquestions of the lower, previously accepted questions on the stack. The ordering function $<$ in $\mathrm{InfoStr}_{\mathcal{D}_0}$ yields the total order:

$$\langle 1, a, a_i, \mathrm{Ans}(a_i), a_{ii}, \mathrm{Ans}(a_{ii}), b, b_i, \mathrm{Ans}(b_i), b_{ii}, \mathrm{Ans}(b_{ii}) \rangle$$

and the QUD function is as follows:

$$QUD(1) = \varnothing$$
$$QUD(a) = \langle 1 \rangle$$
$$QUD(a_i) = \langle 1, a \rangle$$
$$QUD(Ans(a_i)) = \langle 1, a, a_i \rangle$$
$$QUD(a_{ii}) = \langle 1, a \rangle$$
$$QUD(Ans(a_{ii})) = \langle 1, a, a_{ii} \rangle$$
$$QUD(b) = \langle 1 \rangle$$
$$QUD(b_i) = \langle 1, b \rangle$$
$$QUD(Ans(b_i)) = \langle 1, b, b_i \rangle$$
$$QUD(b_{ii}) = \langle 1, b \rangle$$
$$QUD(Ans(b_{ii})) = \langle 1, b, b_{ii} \rangle$$

To satisfy clause (iii) of the definition of QUD in (10g), in each of the ordered sets just listed, each element must be such that its complete answer entails a partial answer to any element to its left. So, given $\langle 1, b, b_{ii} \rangle$, answering b_{ii} must entail a partial answer to b as well as to 1 and answering b must entail a partial answer to 1, etc. This is exactly what the facts in (11) tell us.

We can now define the notion of a strategy of inquiry relative to some topic, or question under discussion, in terms of the QUD function, as follows:

(12)   The *strategy of inquiry* which aims at answering q, $Strat(q)$:

For any question $q \in Q \cap Acc$, $Strat(q)$ is the ordered pair $\langle q, S \rangle$, where S is the set such that:

If there are no $q' \in Q$ such that $QUD(q') = \langle \ldots q \rangle$, then $S = \varnothing$.

Otherwise, for all $q' \in Q$, $QUD(q') = \langle \ldots q \rangle$ *iff* $Strat(q') \in S$.

We might read the ordered pair which Strat yields for a given question q, $\langle q, S \rangle$, as 'the strategy to answer q by conducting the set of subinquiries in

S.' For ($\mathcal{D}_0$), Strat yields:

$$\text{Strat}(a_i) = \langle a_i, \varnothing \rangle$$
$$\text{Strat}(a_{ii}) = \langle a_{ii}, \varnothing \rangle$$
$$\text{Strat}(a) = \langle a, \{\langle a_i, \varnothing \rangle, \langle a_{ii}, \varnothing \rangle\}\rangle$$
$$\text{Strat}(b_i) = \langle b_i, \varnothing \rangle$$
$$\text{Strat}(b_{ii}) = \langle b_{ii}, \varnothing \rangle$$
$$\text{Strat}(b) = \langle b, \{\langle b_i, \varnothing \rangle, \langle b_{ii}, \varnothing \rangle\}\rangle$$
$$\text{Strat}(1) = \langle 1, \{\langle a, \{\langle a_i, \varnothing \rangle, \langle a_{ii}, \varnothing \rangle\}\rangle, \langle b, \{\langle b_i, \varnothing \rangle, \langle b_{ii}, \varnothing \rangle\}\rangle\}\rangle$$

The final line tells us that $\mathcal{D}_0$ involves a strategy to answer 1 by conducting two subinquiries, that of answering a by answering a_i and a_{ii} and that of answering b by answering b_i and b_{ii}. Because of the answerhood relation required by clause (iii) in the definition of QUD in (10g) and the way that Strat is defined in terms of QUD in (12), we are assured that any question q is part of a strategy to answer a question q' only if a complete answer to q will contextually entail a partial answer to q'. It may be that there are more complex types of strategies whose characterization requires ordering the second element in the ordered pair $\langle q, S \rangle$, but I won't investigate that here. Basically, rational considerations, in connection with the information available in the common ground (and hence, what kinds of contextual inferences may potentially be drawn) will determine whether a strategy is well formed.

Note that we wouldn't want to strengthen clause (iii) in (10g) with the requirement that all questions on a QUD stack entail those higher on the stack. This is illustrated by discourses like that in (13) given that:

$$\text{CG(13a)} \supseteq \left\{ \begin{array}{l} \text{John is allergic to clams, One won't eat any-} \\ \text{thing one is allergic to, One will eat some-} \\ \text{thing unless one has some reason not to} \end{array} \right\}$$

(13)   a. What kinds of seafood will John eat?

b. Isn't John allergic to clams?

Here, (13a) doesn't entail an answer to (13b). There might be reasons why John won't eat clams, even if he isn't allergic to them; e.g., John might keep kosher. If the answer to (13b) is *Yes, he is*, then this will contextually entail, in combination with the common ground for the discourse, that John won't eat clams, yielding a partial answer to (13a). However, if the answer to (13b) is *No, he isn't*, that won't by itself tell us that John *will* eat clams. *Prima facie*

this suggests that the requirement in clause (iii) of (10g) may be itself too strong. But I think this apparent problem is resolved if we see that an answer to (13b) is a partial answer to the question (14):

(14)   What reasons would John have for not eating clams?

In turn, (14) is a way of exploring whether John will eat clams, given the assumption that one eats whatever one has no reason not to; hence an answer to (14) contextually entails a partial answer to (13a). I take it that (13b) implicitly assumes a bridging question like (14), making the resulting discourse well formed under clause (iii) of (10g).

As defined, the QUD function does not, and should not, require that it order a continuous sequence of questions. At any given time, QUD may not represent the entire strategy of inquiry with respect to the question on the "bottom" of the stack (i.e., the superquestion): some subquestions may have already been answered, i.e., parts of the strategy already realized, so that they are no longer on the stack. For example, in $(\mathcal{D}_0)$ after answering a via answering a_i and a_{ii}, all these questions pop off the stack when we turn to address b, but they are still part of the strategy for answering 1. Also, in defining QUD, I assumed that a question is popped off the stack if it is answered or determined to be (practically) unanswerable. A question may also be popped off the stack if some question lower in the stack is answered, even if the answered question doesn't directly entail the "higher" questions popped. This makes sense if we see the QUD stack as a fragment of a larger strategy. If we are committed to each question in turn, then each subsequent question must help to answer those already on the stack, even if indirectly. But our commitment is really to the overall strategy, so that commitment to individual questions is relative to our commitment to the earlier questions on the stack below. Then answering a lower question q relieves us of our commitment to any higher questions in $\text{Strat}(q)$. For example, in (13), if *John keeps kosher* follows (13b), then the interlocutors may drop the unanswered question of whether John is allergic to clams, since they are most interested in the larger question (13a), to which they now have a partial answer.

In view of what I said earlier about how questions structure discourse, we also want to guarantee that all of the nonquestion moves, i.e., the assertions, in a discourse are at least partial answers to accepted questions and that in fact each is a (partial) answer to the question under discussion at the time of utterance. This will follow from the way that Relevance is defined within the Information Structure framework. Suppose we define the (immediate)

question under discussion at the time of a move m to be last(QUD(m)), i.e., the last question in the ordered set QUD(m). We can now characterize the notion of Relevance in terms of the question under discussion at a given time (cf. Grice's relativization of his Maxim of Relation to "the purposes of the discussion") and what it is to address such a question:[14]

(15)   A move m is *Relevant* to the question under discussion q, i.e., to last(QUD(m)), *iff* m either introduces a partial answer to q (m is an assertion) or is part of a strategy to answer q (m is a question).

Cooperative speakers strive to make their utterances Relevant, in keeping with the goals of the game. From this it follows that each move in a felicitous discourse, one with a proper information structure, will be Relevant to the question under discussion at the time of its utterance. So, each assertion in such a discourse will provide at least a partial answer to the question under discussion, as desired. Further, (15), in connection with clause (i) of the definition of QUD in (10g), captures the commitment involved in accepting a question: given (i), a question cannot be removed from the set of questions under discussion unless it's answered or unanswerable; given Relevance, we know that a question can only be accepted if it furthers answering those to which the interlocutors are already committed, thereby perpetuating that commitment.

Instead of the static characterization of the information structure of a discourse which I have developed above, one could define information structure in context change terms, a view which I plan to develop in subsequent work. One advantage of the static characterization is that it offers a more global view, facilitating discussion of properties of information structure *per se*, apart from issues of context change; however, I believe that a more dynamic view yields other types of insight, especially pertaining to presupposition projection in questions.

---

14 This calls out for comparison with Sperber & Wilson's (1986) notion of Relevance. A detailed comparison is not possible here, but I will note two significant differences between their notion and that given in (15): First, Sperber & Wilson's Relevance reflects their reductionist program, since it is intended (so far as I can understand) to play the role of all of the original Gricean conversational maxims, whereas (15) is not reductionist; e.g., it is not intended to account for Quantity implicatures. Second, Sperber & Wilson do not relativize their notion to the interlocutors' immediate intentions or goals (and in fact, they deny the very possibility of a common ground), so that the maximization of informativeness while minimizing processing cost is calculated absolutely. But the Relevance defined in (15) is, crucially, relativized to the question under discussion by the interlocutors, and hence, given the pragmatic function of questions in information structure, to the interlocutors' goals.

1.3 Pragmatics of questions in information structure

We want to characterize how posing a question affects the context of the discourse in which it occurs, in view of the question's denotation. The theory of Information Structure outlined above suggests an answer. Recall Stalnaker's (1978) pragmatics of assertion, paraphrased in (16):

(16)   *Pragmatics of assertion*:

If an assertion is accepted by the interlocutors in a discourse, it is added to the common ground at that point in the discourse; i.e., in discourse $\mathcal{D}$ with InfoStr$_\mathcal{D}$, for payoff move m_i and the subsequent move m_{i+1}, if $m_i \in Acc_\mathcal{D}(m_{i+1})$, then $m_i \in CG_\mathcal{D}(m_{i+1})$.

The principle in (16) is the pragmatics for a payoff move. Then for setup moves, we adopt the parallel (17):

(17)   *Pragmatics of questions*:

a.   If a question is accepted by the interlocutors in a discourse, then it is added to the set of questions under discussion, i.e., in discourse $\mathcal{D}$ with InfoStr$_\mathcal{D}$, for question move m_i and the subsequent move m_{i+1}, if $m_i \in Acc_\mathcal{D}(m_{i+1})$, then $m_i \in QUD_\mathcal{D}(m_{i+1})$.

b.   A member of the set of questions under discussion in a discourse $\mathcal{D}$ is removed from that set *iff* it is either answered or determined to be unanswerable, i.e., in discourse $\mathcal{D}$ with InfoStr$_\mathcal{D}$, for moves m_i, m_k and m_{k+1} such that $m_i < m_k < m_{k+1}$ and $m_i \in QUD_\mathcal{D}(m_k)$:

$$m_i \notin QUD_\mathcal{D}(m_{k+1}) \text{ } iff \text{ } \bigcap(CG(m_{k+1})) \subseteq \alpha \text{ or } \pi$$

where α is a complete answer for m_i, and π is the proposition that m_i is unanswerable.

Addition to the QUD stack entails a strong commitment to answering the question. If a question is accepted by the interlocutors, they are committed to answering it; unless it is determined to be unanswerable, it will remain on the stack until answered. Since the InfoStr containing the QUD stack is reflected in the common ground, the fact of this commitment is reflected there as well. Finally, given the way entailment among questions and answerhood are defined, in case a question is removed from the QUD stack, any questions it entails are removed from the QUD stack, as well. Note that (16) and (17) are only the principal pragmatic effects of questions and assertions; there are

others, as well. For example, if a question is asked, the fact that it is asked is entered into the common ground, whether or not it is accepted, this by virtue of the fact that the asking is a speech act performed in full knowledge of all the interlocutors and that such (nonlinguistic) shared information is also represented in the common ground. And if the question is accepted, then the interpretation of the question and the fact that it was added to the set of questions under discussion at that point also becomes part of the common ground, by virtue of the way that the character of the changing InfoStr is continuously reflected in the common ground.

There are a number of other important aspects of the pragmatics of questions which I do not have the space to discuss in detail here. See Groenendijk & Stokhof 1984 and Ginzburg 1995a,b for discussions of some of these, and see Roberts in preparation for an extended examination within the present framework. One of these aspects, not resolved elsewhere in the literature so far as I know, is part of the motivation for the way in which the semantics of questions is formulated in the present account and is directly relevant to the relationship between presupposed and proffered meaning posited in the Information Structure framework: this is the question of how presuppositions project in questions and, in particular, of how it is that questions are holes to presupposition in the sense of Karttunen (1973), i.e., how it is that a question with the logical form $?(\alpha)$ has all the presuppositions of α. Given the semantics for questions I proposed in section 1.1, we might expect that all of the q-alternatives of $?(\alpha)$, i.e., the instantiations of α, would carry the presuppositions of α, but that doesn't explain why the entire question inherits them as well.

By presupposition, I have in mind the phenomenon described by Stalnaker (1978) and Heim (1992): an utterance of ϕ presupposes proposition p *iff* ϕ is felicitous in any given context C only when C entails p. In order for the context to entail p, the context set should be a subset of p.

For Heim (1982, 1983, 1992), presuppositional felicity for assertions follows from their context change potentials: for example, in order to calculate the context change for an assertion with negation, you first have to temporarily update the context set at the time of utterance with the material under the scope of negation, then subtract the resulting context set from the actual context set:

(18) Heim's (1992) *context change potential* for *negation*:

C + not(ϕ) is defined just in case C + ϕ is, in which case:

$$C + \text{not}(\phi) = C \setminus (C + \phi)$$

But the function determining update of contexts, and hence $C + \phi$, is only defined when all of the presuppositions of the assertion to be added, ϕ, are entailed by the context set. So the very possibility of updating the context with the negative assertion in her technical sense requires the satisfaction of any presuppositions of material under the scope of negation. One can argue that by making the facts about presupposition projection follow from the nature of the update operation and the standard semantics for negation, Heim has explained the presupposition projection properties of negative assertions. This is the type of explanation we want for the presupposition properties of questions, as well. I can't see a nonstipulative way to incorporate such an account into von Stechow's (1991) semantics. But a dynamic version of the Information Structure framework has the potential to offer such an account, although it will have to await further work.

The account sketched here and in section 1.1 derives the same end result as that given by von Stechow, though it does so by drawing the line between semantics and pragmatics differently. If we were to adopt the central features of his account apart from structured propositions, the semantics for questions would be as in (19):

(19) Interpretation of a question ?(α), following von Stechow 1991:

$$|\,?(\alpha)\,| = \{w : \text{every } p \in \text{q-alt}(\alpha) \text{ is asked in } w\}$$

For von Stechow, the denotation of a question is a proposition: the set of worlds where every proposition in the corresponding set of q-alternatives is asked. If a question is accepted by the interlocutors, this means that all the propositions in its q-alternative set are asked. Under (19), the denotation of (5) would be as follows:

$|\,?(\textit{who did Mary invite})\,|$

$$= \{w : \text{every } p \in \text{q-alt}(\textit{who did Mary invite}) \text{ is asked in } w\}$$
$$= \{w : \text{every } p \in \{\textit{Mary invited } u : u \in D\} \text{ is asked in } w\}$$

With $D = \{\text{Mary, Alice, Grace}\}$, as above, the question would have as denotation the set of worlds in (20) (assuming again that the relation denoted by *invite* is irreflexive and ignoring joins of individuals in the model):

(20)　{w : *Mary invited Alice* and *Mary invited Grace* are both asked in w}

We might reasonably assume that von Stechow's questions, like other propositions, would contribute directly to the common ground, so that the prior context set would be reduced to include only those worlds in which all the q-alternatives of α are asked. This has the virtue of relating the acceptance of a question to the common ground. But what is it for a proposition to be "asked" in a world? von Stechow doesn't say, but perhaps it's this: if a proposition is asked, the interlocutors attempt to evaluate its truth. On this view of what asking is, (19) would entail (17). If a question is accepted, the common ground will be updated with the information that the interlocutors are committed to determining which of its q-alternatives, if any, are true. But notice that this same incremental effect on the common ground, while not semantic, will still follow directly from the pragmatics for questions in (17) in conjunction with the semantics for questions in (2), the definition of an answer in (3) and the relationship between information structure and common ground discussed in section 1.2. Once a question is accepted, so that by (17) its q-alternative set under (2) is added to the QUD stack of the information structure, the interlocutors are committed to answering it (until and unless it's determined to be unanswerable). Answering a question entails giving an evaluation of all of the propositions in its q-alternative set; hence a commitment to answering involves asking each of those propositions, in the sense just defined. Recall that I assumed that the information structure of a discourse is known to all interlocutors, so that propositional information about it is in fact part of the common ground (see the clauses in (iii) of the definition of CG in (10f)); hence, the common ground will also entail that the question's q-alternatives are asked. So the present account yields the same results as von Stechow's but separates out pragmatic from semantic aspects of the contribution of questions to discourse. Note that the presupposition projection properties of questions do not automatically follow from the sort of treatment exemplified by the semantics in (19), even though in it the proffered content of the question is directly added to the common ground. The proffered content itself is only indirectly related to one of the question's q-alternatives. It doesn't automatically carry along the presuppositions of those q-alternatives, because there is nothing in his semantics that would prevent the interrogative operator itself from being a plug to presuppositions, blocking them from projecting to the entire question.

　　Recall that a question is, in some respects at least, like an imperative, and this proposal for the semantics for questions, like von Stechow's, should be

judged partly in that light. If accepted, an imperative commits the hearer to trying to make the corresponding assertion true; i.e., it commits the hearer to a certain domain goal. If accepted, a question commits the hearer to trying to add its answer to the common ground; i.e., it commits the hearer to a certain discourse goal. Saying that the acceptance of a question involves a commitment on the part of the interlocutors would suggest that we could respond to a question with a refusal, just as we can refuse to accept an imperative. That seems right and is compatible with both the treatment in (2)/(17) and that in (19). However, if we say that the proffered content of a question is a proposition, as in (19), this would also suggest that we could deny its truth, as with assertions. But, though we can answer imperatives with *No*, it doesn't seem quite right to me to refuse a question by denying its truth, as in, *No, it's not true that I'm committed to discussing that question*; this argues in favor of the more pragmatic account developed here, involving (2) and (17), over the more semantic accounts in (19) or von Stechow 1991. It is possible that careful consideration of the facts about indirect questions will bear on this issue as well, but that remains to be seen. In addition, the treatment of the pragmatics of questions in the present account parallels the elegant treatment of assertion in Stalnaker 1978 while avoiding the complications of structured propositions. Whatever we make of these considerations, I do think that utterance of a question, when it is accepted by the other interlocutors, at least has von Stechow's effect indirectly. Knowing what a question is, and what it is to accept it, the interlocutors know that this is the question under discussion.

2 Focus and information structure

2.1 The presuppositions of prosodic focus in English

Skillful interlocutors in discourse expend a good deal of effort making sure that at any given point all the participants are clear about what the common ground is like, including what they are talking about (the question under discussion) and how what they're talking about relates to the rest of the information in the common ground (the strategy of inquiry they're following and its general relationship to the information structure of the discourse). As Mike Calcagno (personal communication) pointed out, in a competitive game players hide their strategies; in a cooperative game they make their strategies explicit. Language is cooperative, even when we're using it to compete.

Redundancy helps to assure that the cooperation in sharing information is effective.[15]

One of the tools we use to realize desirable redundancy is presupposition, which also serves to help give coherence to discourse, as we will see. Recall that utterance of a linguistic expression ϕ is only felicitous if the context of utterance entails the presuppositions of ϕ. This means that when an utterance has a presupposition, this is a conventional redundancy — the presupposition is already at least entailed by the context, if not explicitly part of the common ground. But this built-in redundancy of presuppositions can be used, as has often been noted, to actually introduce information which was only implicitly assumed previously or even not previously introduced at all. If an utterance has a conventionally expressed presupposition, it is often quite clear what the context should be like in order for it to be felicitous. And if the context isn't quite like that but the hearer is cooperative and has no objection to the truth of the presupposed information, she will accommodate it — behave as if the context included that information all along and hence as if the utterance which triggered the presupposition were felicitous. Consider the gossip's use of factive verbs to introduce scandalous information with impunity — *Isn't it shocking that Mary has run away from home?*

I will argue here that intonational focus in English is presuppositional, giving information about the type of information structure in which the utterance associated with it occurs and its role in that structure. It is thereby redundant, given the fact that the information structure is in principle common information. And interlocutors use this redundancy to indirectly convey information about the information structure they intend. The main idea is that assertions, like questions, are conventionally associated with a set of alternatives, although these alternatives are presupposed by the prosody rather than proffered as are the q-alternatives of questions. Focal alternatives are calculated on the basis of the placement of prosodic focus in the utterance, along lines similar to those developed in the theories of Rooth (1992a) and von Stechow (1991), among others. It has long been noted (see, e.g., Jackendoff 1972) that in question/answer pairs, the prosody of the answer constrains the questions which it can answer; in this theory, as in the others just cited, this relationship will be reflected in a condition on the relationship between the alternatives associated with each member in the pair, the q-alternatives of the question and the focal alternatives of the answer. Also,

---

15 For an enlightening discussion of the value and role of informational redundancy in discourse, see Walker 1993.

it has been occasionally noted in other connections (cf., e.g., the discussion of yes/no vs. alternative questions in von Stechow 1991) that the set of alternatives proffered by a question may depend in part on its prosody. This is reflected in the present theory in the assumption that prosody plays a role in determining a set of focal alternatives for questions as well as for assertions. In a question these alternatives constrain the type of superquestion it may be related to in the strategy of inquiry. Hence, this is an alternative semantics account of prosodic focus. However, it differs from other recent accounts along these lines (Rooth 1985, von Stechow 1991, Krifka 1992) in that it does not assume any direct focus sensitivity for particles like English *only* or *even* or for the other operators and particles discussed in the extensive linguistics literature on focus. In this respect, it is in the same vein as the recent work of Rooth (1992a), Schwarzschild 1994a,b and von Fintel (2004), who also attempt to derive certain focus effects pragmatically, though they develop different mechanisms.

I am explicitly confining myself here to discussion of the role of English prosodic focus and hence am not making claims about some universal linguistic phenomenon, Focus. With Rooth (1996), I would argue that instead we need to give detailed analyses of particular conventional elements in particular languages. However, I do assume that information structure is a universal of human discourse. And the persistent intuitions on the part of researchers that there are universals of information structure (topic, focus, theme/rheme, etc.), plus the persistent linkage of focus with the question/answer paradigm, need to be explained. If we assume that discourse has a functional structure of the sort I'm proposing, then we can go quite a bit of the way towards explaining these intuitions while leaving open the possibility that syntactic constructions and other conventional factors in various languages contribute to fulfilling these functions in somewhat different ways.

Let me briefly sketch what I mean by English intonational focus. Then I will turn to a more detailed proposal of how the presuppositions arise and what their content is. I will make the following, somewhat simplified assumptions about the prosodic phonology of focus:[16]

---

16 Selkirk (1984) provides a fairly detailed exploration of the issues and discusses many valuable examples; my assumptions about prosody are adopted from her work, except where noted. Selkirk's work, in turn, relied heavily on the work of J. Pierrehumbert (1980) on the phonology of tone, including pitch accents; see also Liberman & J. Pierrehumbert 1984.

(22) The *phonology of focus*:

    a. There is at least one intonation phrase per sentential (or sentential-fragment) utterance.

    b. There is at least one focused subconstituent (possibly nonproper) per intonation phrase. This *focused constituent* is marked with the feature F in what follows.

    c. There is at least one pitch accent per focused constituent, associated with a subconstituent.

    d. Every pitch accent must be associated with material in a focused constituent.

    e. There is one phrase accent (H– or L–) and one boundary tone (H% or L%) per intonation phrase.[17]

    f. The string-final pitch accent in the focused constituent is assigned the most prominent stress in the intonation phrase (*Nuclear Stress Rule*).

All of the following prosodic factors are probably pragmatically (and hence potentially semantically) significant:

(23) *Pragmatically-significant prosodic factors*:

    a. the choice of intonation phrase constituent(s) (which I tentatively assume are always correlated with syntactic constituents);

    b. within the intonation phrase:

        i. the choice of focused constituent,

        ii. the placement of pitch accent(s), and

        iii. the choice of pitch accent(s) and of phrase accent and boundary tone; and

    c. the relative prominence of different intonation phrases, both within an utterance and across utterances.

Though I will concentrate in this section on the significance of the selection of focused constituent, the other factors are surely important to a full understanding of the phenomena usually grouped under the term *Focus*,

---

17 In Liberman & J. Pierrehumbert 1984, boundary tones are realized at the end of their associated intonation phrase, while phrase accents effect a transition between the last pitch accent in the intonation phrase and the boundary tone. See Beckman & Ayers 1997 for a somewhat different account, involving intermediate intonation phrases, as well.

and in section 2.2.2.2 I will note the crucial role in certain examples of (23a) and (23c), the choice of intonation phrase constituents and their relative prominence within an utterance.[18]

The above assumptions reflect the fact that all English utterances bear prosodic focus (tonal and rhythmic prominence) on at least one constituent, though that focus may be broad or narrow (see Ladd 1980 and examples below). In keeping with this and with the methodological principle of semantic compositionality, the theory I offer aims to give a unified account of the semantics of English prosodic focus across all utterance types, whereas most theories address only one or two contexts of use and some authors even claim that no unified account is possible. Note also that the assumptions about the placement of pitch accents and stress within focused constituents assure that the feature F is invariably realized in the surface form of the utterance by rhythmic and tonal prominence. The prosodic realization of focus is not universally assumed by those working on the semantics of focus. See Partee 1991 and Krifka 1992 for examples where foci crucially are *not* so realized. There are no abstract foci in the present analysis of English; for the time being, my critique of other approaches will have to remain implicit in the way the present approach is developed.

I want to bring special attention to (22d), the assumption that all pitch accents must be within focused constituents (cf. Selkirk's (1984: 282) *Focus Domination of Pitch Accent*). It entails that if an utterance bears a single narrow (often called "contrastive") focus on some constituent x, the remainder of the utterance bears no intonational prominence, so that its contour sounds flat to the ear of the native speaker. Selkirk talks about pitch accent placement in terms of "old" and "new" information. Unaccented elements of an utterance (or at least NPs and other arguments which would otherwise nor-

---

18 I must ignore the issue of the significance of multiple intonation phrases per utterance, and I will only comment briefly on the significance of various combinations of phrase accent and boundary tone, though I feel certain that these matters warrant careful investigation. And there are a number of interesting questions about pitch accent placement within focused constituents (see Selkirk 1984) and about the choice of pitch accents (see J. Pierrehumbert & Hirschberg 1990) which I must also ignore. Another issue that needs to be addressed in more detail is the question of where in the grammar (e.g., at what level of representation) these conventional elements should be encoded. Selkirk encodes them at SS ("surface structure") in a Government-and-Binding style grammar, both to capture the fact that the relationship between pitch accent placement and F-marking appears to be syntactically constrained and so that this information is available to both PF ("phonological form") and LF ("logical form"). I have not seen a careful exploration of these issues in a constraint-based grammatical framework.

mally bear accent) are taken to be "old information," accented constituents are taken to be "new information," but Selkirk doesn't make precise the sense she would give to those terms. Within the present account, we can offer a more precise sense for those terms — they are taken to be part of the question under discussion.[19] I take the expression *old* to be synonymous with *given by the question under discussion* and *new* (in the sense relevant here) to mean *not given by the question under discussion*.

The focal alternatives for a given utterance are defined as follows:[20]

(24) The *focus alternative set* corresponding to a constituent β:

$\|\beta\|$ is the set of all interpretations obtained by replacing all the F-marked (focused) constituents in β with variables and then interpreting the result relative to each member of the set of all assignment functions which vary at most in the values they assign to those variables.[21]

Given (24), both questions and their answers, assertions, have corresponding alternative sets, the q-alternatives of the questions and the focal alternatives of their answers. This observation is the basis of von Stechow's (1991: 36) notion of question/answer congruence, modified and generalized in (25):

(25) Move β is *congruent* to a question $?(\alpha)$ *iff* its focal alternatives $\|\beta\|$ are the q-alternatives determined by $?(\alpha)$, i.e., *iff* $\|\beta\| = $ q-alt(α).

Given (24) and (25), the assertion in (5b) is a congruent answer to the question in (5a):

---

19 Selkirk points out that it appears that unaccented NPs sometimes occur even phrase-finally within a focused constituent, and still they seem to be old information in the sense intended here. This suggests to me that we may need to reconsider the phonological representation of focus in some respects, but I can't explore that in the present paper.

20 Rooth (1985) and von Stechow (1991) don't define alternatives in this way. Instead, as in Krifka 1992, alternative sets are cancelled out when the constituent which determines them is argument to a focus-sensitive operator. That is not the case with the definition I offer here, and the difference should be borne in mind in comparing this account with those of Rooth, von Stechow or Krifka. For simple cases, (24) yields the same results as the assumption that focused constituents are raised at LF ("logical form") and that such raising is island insensitive (see Rooth 1992b). However, see Rochemont & Culicover 1990 for arguments that association with focus is bounded, i.e., sensitive to islands; I don't yet see how to cope with the problem this raises in defining the focus alternative set, but it doesn't seem that, e.g., Rooth or von Stechow have taken this type of data into account either.

21 We might want to require that the replacement variables be novel in the sense of Heim (1982) and constrain the set of assignment functions to include only those in the satisfaction set of the file/context at the time of utterance of β. This would guarantee that any definite (or familiar) NPs in β would be appropriately bound. But I won't pursue this refinement here.

(5) a. Who did Mary invite?

b. Mary invited nobody_F.

The set of focal alternatives for (5b) (with the empirically correct focus in this context) is the same as the q-alternative set for (5a). Congruence doesn't require that the answer be among the corresponding set of alternatives (and in this respect is like the answerhood relation given in the previous section) but only that the answer and question evoke the same set of alternatives. Further, note that this theory of questions and answers gives a question like (5a) no existential presuppositions. The propositions in its q-alternative set do all entail the existence of someone invited, but these propositions are just asked, not asserted; the question/answer relationship as defined above does not require that the answer be one of the q-alternatives. They may all be false, as entailed by (5b), in which case there is no existential implication.

Consider the following presupposition of prosodic focus for assertions, generalized from the requirement for question/answer pairs in von Stechow 1991 and Rooth 1992a:

(26) *Presupposition of prosodic focus in an assertion β*:

β is a congruent answer to the question under discussion at the time of its utterance.

Given (26), the prosodic focus on an assertion presupposes that the current goal of the discourse is to choose among the alternatives in $\|\beta\|$. (See Kadmon 2001: 341-343 for discussion and illustration of the empirical predictions which the congruence presupposition makes.)

This works fine for question/answer pairs like (5a) and (5b). But questions also bear prosodic focus, and we want a more general principle than (26), one which tells us the meaning of prosodic focus in an utterance of any mood. Recall that questions can be logically related to each other, and in effective strategies of inquiry they generally are. To motivate this on empirical grounds, consider the multiple *wh*-question (27a) and the simple *wh*-question (27b):

(27) a. Who invited whom?

b. Who did [Mary]_F invite?

Note that (27b) can be asked more or less immediately following (27a), so that ⟨(27a), {(27b), ∅}⟩ seems to be a felicitous strategy of inquiry. It seems that the prosody in (27b) is crucial to this felicity; e.g., the same string but with focus on *invite* instead would be infelicitous. In order to capture this,

we need to say something about the presuppositions of prosodic focus in questions.

It would be desirable to generalize over the presuppositions of prosodic focus for the two types of speech act considered, as in (28) below. Assume that each utterance has a logical form which involves a mood operator. $*\beta$ is an utterance of β with mood variable $*$ ranging over $\{?, \cdot\}$, interrogative and assertional operators, respectively:

(28)   *Presupposition of prosodic focus in an utterance $*\beta$:*

   β is congruent to the question under discussion at the time of utterance.

(28) entails the subgeneralization in (26). However, given the definition of congruence in (25) and, especially, the definition of the focus alternative set in (24), (28) would fail to explain the felicity of $\langle$(27a), (27b)$\rangle$, since in (27b) the *wh*-word, *who*, is not prosodically focused (compare the echo question *WHO did Mary invite?*, where *who* is plainly focused). However, the generalization will hold if we change (24) to obtain the set ranging not only over the F-marked constituents in β but over any *wh*-elements as well. Then we would obtain (29):[22]

(29)   The *focus alternative set* corresponding to a constituent β (revised):

   $\|\beta\|$ is the set of all interpretations obtained by replacing all the F-marked (focused) and *wh-* constituents in β with variables and then interpreting the result relative to each member of the set of all assignment functions which vary at most in the values they assign to those variables.

Now we will apply the analyses in (25), (28) and (29) to the analysis of the sequence of utterances in (27). Given the semantics for questions in section 1.2, the relevant interpretations are those in (27a$'$) and (27b$'$):

(27)   a$'$. $|\,?(\text{who invited whom})\,| = \{u \text{ invited } u': u, u' \in D\}$

So if $D = \{\text{Mary, Alice, Grace}\}$, then the question will have the following q-alternative set (ignoring again the possibility of collective arguments):

$$\left\{ \begin{array}{l} \textit{Mary invited Alice, Mary invited Grace, Alice invited Grace,} \\ \textit{Alice invited Mary, Grace invited Mary, Grace invited Alice} \end{array} \right\}$$

---

22 This would not give the correct results for utterances containing indirect questions.

In the same model, given the derivation of this example in (6):

(27)   b'. $\big|\, ?(\text{who did } [\text{Mary}]_F \text{ invite})\,\big| = \left\{ \begin{array}{l} \textit{Mary invited Alice,} \\ \textit{Mary invited Grace} \end{array} \right\}$

Clearly, the q-alternative set of *Who did Mary$_F$ invite?* is a subset of that for *Who invited whom?*. Hence, (27a) entails (27b), since a complete answer to (27a) will yield a valuation for all the alternatives in (27b) as well. (27b) is a subquestion of (27a), so that (27) represents a felicitous strategy of inquiry.

Calculating the focal presuppositions of (27b) as in (29), we find that $\|(27b)\| = \{u \text{ invited } u' : u, u' \in D\}$. But this is just the set of propositions which is the q-alternative set of (27a); i.e., $\|(27b)\| = \text{q-alt}(27a) = \{u \text{ invited } u' : u, u' \in D\}$. Hence (28) is satisfied.

Note that with broad focus on (27a), the question which it is congruent to is trivial: it is just the Big Question. However, I think this is misleading in this example: usually *invite* takes not only an agent and a patient, but also a goal, the event to which the patient is invited. Since no goal is given in (27a), it strikes me as elliptic and this ellipsis suggests a connection between this question and an on-going prior discussion of some event. I won't go into this in any further detail here but just want to note that ellipsis, like destressing/deaccenting within a focused constituent, may influence the determination of the question to which an utterance, either question or assertion, is presupposed to be congruent. See Rooth 1992b for discussion of the relationship between ellipsis and "anaphoric destressing."

Principle (28) not only explains the felicity of the discourses in $\langle (5a), (5b) \rangle$ and $\langle (27a), (27b) \rangle$, repeated below, but also the infelicity of sequences like those in $\langle (5a), (5c) \rangle$ and $\langle (5a), (5d) \rangle$ and in $\langle (27a), (27c) \rangle$ and $\langle (27a), (27d) \rangle$, since in none of the last four cases will the focus alternative set of the second element equal the q-alternative set of the first, the question under discussion:

(5)   a. Who did Mary invite?

       b. Mary invited [nobody]$_F$.

       c. Mary [invited]$_F$ nobody.

       d. [Mary]$_F$ invited nobody.

(27)   a. Who invited whom?

       b. Who did [Mary]$_F$ invite?

       c. Who did Mary [invite]$_F$?

       d. [Who]$_F$ did Mary invite?

Note that on this conception of how prosodic focus works, it triggers a presupposition but does not uniquely determine the nature of the set of alternatives with which the utterance must be compatible/congruent, following Kadmon & Roberts 1986. Because of potential operator scope ambiguities, we also require, in many examples, fairly rich information about the structure of the preceding discourse in order to determine what the question under discussion actually is and hence, indirectly, what the truth conditions of the utterance itself should be. Consider the example from Kadmon & Roberts 1986, with the alternative scopes represented by the possible questions under discussion below it, where δ is a variable of type $\langle\langle e, t\rangle, \langle\langle e, t\rangle, t\rangle\rangle$ and Δ is the set of determiner denotations:[23]

(30)   He doesn't like [most]$_F$ of the songs.

    a. most($\neg$) : q-alt(30) = $\{\,|\,\delta\,\text{songs}(\lambda x \, . \, \neg\,\text{likes}(\text{he}, x))\,|\, : \, |\,\delta\,| \in \Delta\,\}$

    b. $\neg$(most) : q-alt(30) = $\{\,|\,\neg\delta\,\text{songs}(\lambda x \, . \, \text{likes}(\text{he}, x))\,|\, : \, |\,\delta\,| \in \Delta\,\}$

These sets of alternatives correspond with different questions. For (30a) the question is, roughly, *What is the proportion of songs that he doesn't like?*. For (30b) the question is *What is the proportion of songs that I deny that he likes?*. The focus on the utterance makes it in principle congruent with either question, and so the utterance is ambiguous in its presupposition. The actual context of utterance will generally disambiguate. If it doesn't, i.e., a hearer encounters such an utterance without adequate context, then, as discussed by Kadmon & Roberts (1986), the simplest context one can evoke will be the default assumption; here, this will be the nondenial context, in which *most* has wide scope. Setting up the other possible reading involves a more complex context, involving an assertion, its denial and a correction, so we are less likely to retrieve it when hearing the utterance out of the blue.

In summary, the prosodic focus on an utterance gives rise to the presupposition that the utterance, whether assertion or question, is congruent to the question under discussion. Such a mechanism will obviously help to give coherence to discourse and assure relevance. And it is at least to some extent redundant: for example, in a question/assertion sequence, if the speakers are playing the language game correctly, i.e., *inter alia* observing

---

23 Kadmon & Roberts (1986) give an analysis of the full prosodic contour of the example, including the pitch accents, their placement and the phrase and boundary tones; this is the same for each of the noted readings. They also give a detailed account of the kinds of context in which each of these readings can arise. The reader is referred to their paper for details.

Relevance and the imperative to address the goal represented by the question under discussion, then the assertion should be an answer to the preceding question. If the assertion's prosody also presupposes that it plays this role, this is a redundant confirmation of the assertion's role and hence indirectly of the question being investigated. But of course, once the rules of a game are set, we often develop strategies for more efficient play. In the case of an assertion, if the prosody presupposes the type of question it addresses, then it may not be necessary to actually ask the question explicitly, especially if the relevance of the question to the ongoing strategy of inquiry is clear. Hence, hearers can use the presuppositions of English prosodic structure to accommodate portions of the information structure assumed by the speaker.

I will now turn to briefly review several types of examples from the literature on focus in light of this proposal.

2.2 English focus phenomena

In section 2.2.1, I will illustrate the application of the theory in section 2.1 to one of the most discussed phenomena involving focus in English: association with focus. Then in section 2.2.2 I will briefly consider extensions to deal with uses of focus, often called *contrastive*, in contrastive topics, contrastive focused pairs and yes/no questions interpreted as alternatives.

2.2.1 Association with focus as a conversational implicature[24]

Some of the most interesting work in the formal literature on focus in the past decade is that which addresses what Jackendoff (1972) called *association with focus*: the tendency for the interpretation of certain operators, including negation, *only*, *even*, modals and adverbs of quantification, to depend partly on the prosodic focal structure of the utterance in which they occur. Rooth (1985) argued that this sensitivity to focus is part of the lexical semantics of the operators he considered and that it bears on the determination of their domain restriction. Basically, he argued that the domain is (a subset of) the focus-determined alternative set corresponding to the constituent which is the operator's scope at logical form, this being specified as part of the lexical meaning of an operator like *only*. More recently, von Stechow

---

24 Thanks to Nirit Kadmon and Paul Portner, whose thoughtful comments on an earlier draft led to what is, I hope, a considerable improvement of this section. They should not, however, be taken to necessarily agree with the view proposed here.

1991 and Krifka 1992 (*inter alios*) have adopted a similar approach within the structured propositions framework. In their work, a structured meaning reflects the focal structure of the relevant utterance; it is an ordered pair, dividing the utterance's meaning into focus and ground. The lexical semantics of one of the focus-sensitive operators makes reference to both aspects of the structured meaning of the constituent which is its scope. Both of these approaches involving focus-sensitive operators place a heavy burden on focus in the architecture of the resulting semantic theory; association with focus is taken, by itself, to motivate a significant complication of the semantics for a language (see Rooth 1996 for some relevant discussion). Further, it is clear that focus doesn't fully determine the domain of these operators; other contextual factors plus presuppositions of the constituent in the operator's scope also play important roles in its determination (see Roberts 1995). It would seem desirable to say something about the relationship between these various ways in which the domain is determined, but the focus-sensitive approach, being basically semantic, cannot generalize to include both the semantic focus sensitivity and the clearly pragmatic factors. Simply making focus sensitivity part of the lexical semantics for the relevant operators fails to capture what it is that makes this phenomenon so pervasive in domain restriction. But it also seems plain that focus plays a regular and significant role in the interpretation of utterances containing these operators. We need to account for this role, something which, e.g., Vallduví (1993) and Vallduví & Zacharski 1994 cannot do in a systematic way.

Rooth (1992a), von Fintel (1994, 2004) and Schwarzschild 1994a,b have recently attempted more pragmatic accounts of the various association with focus phenomena, accounts which are either anaphoric, as in Rooth and von Fintel's work, or involve entailments from principles for the interpretation of focus as contrast, as in Schwarzschild's. The Information Structure framework presented here suggests another kind of pragmatic account for association with focus effects, one which follows straightforwardly from independently motivated principles defined in previous sections of this paper, without necessitating any additional stipulations. The account is like the earlier account in Rooth 1985 in assuming that association with focus is essentially a fact about how domain restriction of operators is influenced by prosodic focus. But I assume no special lexical sensitivity of the relevant operators; and unlike von Stechow and Krifka, I would argue that the focal structure of an utterance is transparently given by its prosodic structure, without the need to specify an independent partition on its meaning. In this

section, I will outline how the Information Structure framework from section 1 and the assumptions about prosodic focus in the previous section explain the role of prosodic focus in restricting the domain of *only*. See Calcagno 1996 for a related view of domain restriction in adverbs of quantification.

Utterance of a construction where *only* is a VP modifier, *x only VPs*, is often taken by theorists to presuppose that *x* VPs, but is sometimes assumed to only entail it. I will assume the presuppositional view, but this doesn't make a difference to the account I will offer. If *x VPs* is presupposed, then just as prosodic focus, in presupposing a question under discussion, can actually be used to introduce that question, so *only* can be used to introduce the presupposition that *x* VPs, even when it isn't already known to the interlocutors. The two cases are parallel in that the presupposition in each is quite explicit, so that it is clear what must be accommodated in order to repair the context. Unless they have grounds to object, cooperative interlocutors will accommodate these presuppositions as necessary.

Assume that the meaning of VP-adjunct *only* is as in (31):

(31) Interpretation of VP-adjunct *only*:

Presupposed content: the subject has the property denoted by the VP.

Proffered content: the subject has no properties apart from that denoted by the VP.

Given this, the analysis of association with focus, following Rooth, involves determining which class of properties is the intended domain of quantification for the operator *no* in the proffered content of only. In Roberts 1995 I argued that this couldn't be determined solely on the basis of an algorithm based on prosodic focus but that it should be the set of properties which are relevant in the context at that point in the discourse. Now we can be more precise about what that means. The answer will be seen to follow from the requirement of Relevance as defined in (15) and the presupposition of prosodic focus in (28), repeated here (along with (25)):

(15) A move m is *Relevant* to the question under discussion q, i.e., to last(QUD(m)), *iff* m either introduces a partial answer to q (m is an assertion) or is part of a strategy to answer q (m is a question).

(25) Move β is *congruent* to a question ?(α) *iff* its focal alternatives $\|\beta\|$ are the q-alternatives determined by ?(α), i.e., *iff* $\|\beta\|$ = q-alt(α).

(28)   *Presupposition of prosodic focus in an utterance $*\beta$:*

β is congruent to the question under discussion at the time of utterance.

To see how this domain restriction works, consider the example in (32):

(32)   Mary only invited [Lyn]$_F$ for dinner.

Given (28), (32) presupposes the question in (33):

(33)   Which individual(s) is/are such that Mary has no properties apart from having invited that/those individual(s) for dinner?

In order to understand what a speaker intends by uttering (32), we must determine the intended domain for *only*. We already know that there is an independent requirement that the utterance, and the question it presupposes, be Relevant to the prior context, i.e., that they address the question under discussion. If (32) is uttered out of the blue, this would mean that the presupposed question (33) would have to form part of a strategy to answer the Big Question; all of Mary's properties would be Relevant in such a case, as any information about her having or not having them would address the Big Question. But, under the assumption that Mary exists, utterance of (32) out of the blue would make it false in any reasonable model, since Mary surely has, at the very least, the property of self-identity (and presumably a number of others), as well as perhaps that of having invited Lyn for dinner. Hence, on the assumption that the speaker of (32) is competent and is observing the Gricean maxim of Quality, i.e., trying to assert something truthful, this can't be the intended context of utterance for (32). As I have argued elsewhere (Roberts 1995, 1996a), domain restriction is always constrained by cooperativity: the cooperative hearer assumes that the speaker is cooperative as well (and competent) and on this basis seeks to resolve any apparent failure of cooperativity, e.g., failure to satisfy the Gricean maxims of Quality or Relevance, presupposition failure, etc., by restricting the domain in such a way as to make the contribution cooperative after all.

Now imagine instead that (32) is uttered following (34):

(34)   Who did Mary invite for dinner?

(32)   Mary only invited [Lyn]$_F$ for dinner.

Of course, as we have already seen, (32) prosodically presupposes that the question under discussion is (33), which is not the same question as the

explicit (34). We can only felicitously accommodate the presupposed (33) if the newly accommodated question is itself Relevant to the accepted question (34) so that they form a felicitous strategy of inquiry and the resulting information structure is hence well formed. This requirement is straightforwardly met just in case we assume that the intended domain of *only* is (a possibly nonproper subset of) the set of properties which Rooth's (1985)/von Stechow's (1991)/Krifka's (1992) theories would fix conventionally as the domain of *only* in (32). This is because, given the semantics of questions, only with this domain restriction will all of the answers to (33) be answers to (34) as well, so that (33) is Relevant to (34). Here's why:

The only way to address (34) is to determine the truth of a proposition of the form | *m invited α* |, where $α$ is either a rigid designator denoting an individual in the model or else *nobody*.[25] But then this means that the set of Relevant properties is just the set which, when predicated of Mary, yields one of those propositions, i.e., a property of the form | *invited α* |, where $α$ is either a rigid designator denoting an individual in the model or else *nobody*. This set of properties is then the set which must restrict the domain of *only* in (33) if it is to be Relevant to (34), i.e., to form part of a strategy of inquiry for addressing (34). But this is just the set of alternative properties which is retrieved in other accounts, e.g., in Rooth 1985 via the calculation of the p-set of the VP in (32). In those accounts, that p-set is referred to by the lexical semantics of *only*. But on the present account, this isn't necessary.

When we take the domain of *only* (the set of relevant properties) in (32) to be the set of properties of inviting someone or other to dinner, then (33) is logically equivalent to (35):

(35)   Which individual(s) x is/are such that of all the properties of inviting someone or other for dinner, Mary has none apart from having invited x for dinner, i.e., which individual(s) is/are such that Mary invited nobody else for dinner?

(34) and (35) are very close in meaning, though they are not identical: their q-alternative sets are distinct. The difference is also reflected in the fact that all of the direct answers to (35) are complete, each of them also a complete answer to (34); if we answer (35) with *Alice and Gertrude*, then this entails that nobody else was invited, including Grace. But (34) has partial direct answers

---

25 Again, I'm ignoring the possibility that Mary invited more than one individual, which would make Relevant as partial answers propositions like | *Mary invited few of the students* |. This is merely in the interest of simplicity and is inessential to the argument I give.

which are not complete, so that we can answer it with *Alice and Gertrude* without precluding Mary's also having invited Grace. However, though these two utterances don't denote the same question, in any given model, all the complete answers to (34) are complete answers to (35) and *vice versa*, so that they logically entail each other in the sense defined by Groenendijk & Stokhof.[26] Because they have the same set of complete answers, (34) and (35) would establish the same partition on the context set at a given time of utterance. So with the domain fixed as in (35), the presupposed question (33) is Relevant to (34), in the technical sense defined in (15). But then in order to be Relevant to (33)/(35), and hence indirectly to (34), the domain of *only* in (32) must be set in the same fashion.

As Paul Portner pointed out (personal communication), this won't guarantee that the domain isn't a proper subset of Rooth's properties. But, of course, Rooth also admitted the possibility of additional, contextually given restrictions on the domain of *only* as given by his conventional calculation of p-sets and focus-sensitive lexical semantics for *only*. This can be motivated for discourses such as ⟨(34), (32)⟩. Consider a scenario where the interlocutors already know that Mary invited a visiting colleague to dinner and wanted to get someone else in her department to accompany them. Then the alternative set of properties should involve all the other people in her department, including Lyn but not the visitor. Then utterance of (32) in that context would not entail that Mary invited nobody else in the world except Lyn to dinner and would be consistent with the speaker being invited as well.

But what other kinds of questions might (32) Relevantly address? It turns out that this class is quite restricted and depends in part on context. The prosodically presupposed question (33) sets up a partition on the context set; the cells, possible complete answers to the question, differ from each other in the identity of the sole invitee. Besides the relevant properties considered above, corresponding to Rooth's p-set for the VP, any other properties that Mary might have are one of two types, which I will call the *contextually*

---

26 Recall that questions and answers are semantic entities, not to be confused with the utterances which denote them. If the complete answer to (34) is that suggested by replying *nobody*, this entails that the answer to (35) must be that denoted by the reply *everyone* if there are people in the model or by *nobody* if there are no people; the converse entailments hold as well. The complete answer to (34) denoted by *everyone* is similarly correlated with the complete answer to (35) given by replying *all the people*. In the less extreme cases, complete answers to the two questions could be expressed identically, so that answering (34) by *Alice, Gertrude and Grace* would entail and be entailed by the answer denoted by the same reply to (33) and so forth.

entailed restrictions and the *logically independent properties.* The entailed restrictions are entailed by the context set, because they are known by the interlocutors to be Mary's properties; these include the property of inviting the visitor in the described context. Entailed restrictions serve to implicitly restrict the domain of *only* in a negative fashion, because there is no world in the context set in which Mary does not have all these properties; so we have the implicit commitative restriction *with the visitor* on the VPs in (32) in the situation described. But an entailed restriction is never Relevant because it doesn't distinguish between cells in the partition established by (34).

The other type of non-p-set properties are those which so far as the interlocutors know may or may not hold of Mary and hence are in principle completely independent of the question under discussion. For example, if (32) is uttered following a question like *Did Mary brush her teeth this morning?*, it will be irrelevant to that question in most contexts. That having dinner with someone is not usually Relevant to the question of whether she brushed her teeth is reflected in the fact that in any reasonably realistic model, the question of whether Mary has the logically independent tooth-brushing property establishes a distinct partition from that established by (34). Questions based on such independent properties do not bear on the choice of a (partial) answer to the question under discussion (33) because they crosscut the cells in the partition it has established; determining that Mary has one of them will hence not lead to removal of a cell from the partition. The discourse would fail on those grounds alone, never mind that the domain of *only* in such a context would be as in the case of utterance out of the blue. But now suppose that the context entails that Mary was going to have dinner with one or more of her friends. Further, assume that Mary's friend Lyn is also her dentist and always nags her to watch her dental hygiene but only when the two are alone. Finally, suppose that Lyn's reminders to Mary are generally effective so that if, out of the group of her friends, Mary only dined with Lyn, then (Lyn having nagged her) she would remember to brush her teeth the next morning. In such a case, the question presupposed by (32), (33), under the entailed restriction that the only people she was considering dining with were the aforementioned group of friends, *would* be part of a reasonable strategy to answer the question of whether she brushed her teeth, for (32) would contextually entail the answer to that question. Relevance is essentially context dependent. But still, (32) could only entail the answer to *Did Mary brush her teeth?* on the assumption that the correct domain, i.e., the contextually salient group of her friends, was chosen to restrict the

interpretation of *no* in the interpretation of *only*. So the logically independent properties may sometimes be Relevant, but this is entirely dependent on context.

Rooth's (1985) theory of association with focus and the subsequent theories of von Stechow (1991) and Krifka (1992) stipulated, as part of the lexical meaning of *only*, that its domain should be calculated directly on the basis of the prosodic focus in its scope. There is a type of context in which such theories systematically make the wrong predictions, as has been pointed out by Vallduví (1990), Partee (1991), von Fintel (1994), and Roberts (1995). In the present theory, I take their examples to illustrate how prosody, though it makes a regular, conventional (because presuppositional) contribution to the determination of intended context, and hence of intended domain restriction, cannot by itself give the intended domain restriction for *only*. Nirit Kadmon (personal communication) offers the following illustration of this for the example in (32). Suppose we are talking about some things that we had been afraid that Mary was going to do today that we disapprove of. These included Mary's having the properties listed in (36):

(36)   a. inviting Lyn for dinner
        b. inviting Bill for dinner
        c. staining the tablecloth at lunch
        d. smoking before dinner

In this context, consider (37):

(37)   a. Mary wasn't so bad after all. Of all the things we were afraid she might do, she only [invited Bill for dinner]$_F$
        b. You got the person wrong. She only invited [Lyn]$_F$ for dinner. But it's true that she did only one of those terrible things she could have done.

In this case, we want the domain of *only* to be the set of properties in (36) and not the set of properties of inviting someone or other for dinner, which latter would be the set given compositionally by, e.g., Rooth (1985). Conventionally fixing the domain yields the wrong truth conditions.

Here is how we can account for such examples in the framework of Information Structure, given the presupposition of prosodic focus in (28). (37a) presupposes the question (38):

(38)   Which property is such that out of all of the properties in (36), Mary had no other properties apart from that?

The domain restriction for *only* is explicit and refers to the principal question under discussion, i.e., which of the properties in (36) Mary has. Then (37a) asserts that Mary has the property of being such that out of all of the Relevant properties in (36) she only has the property of having invited Bill for dinner. The question presupposed by (37b) is (33), just as it was for (32). But here the Relevant properties will be different, since the context to which (37b)'s utterance must be Relevant is different. (37b) makes it explicit that it is offering a correction to (37a); even if it hadn't, in order to explain the Relevance of (37b) to (38), the question addressed by the immediately preceding utterance (37a), especially given their parallel forms but contrasting prosodic focus, we would assume that (37b) is offering a correction of the assertion made by (37a). As in corrections generally, the corrector addresses a different question in (37b) than that addressed by the corrected interlocutor in (37a), as we see by the distinct prosodic pattern employed in (37a); but still the correcting utterance (37b), by virtue of what it is to correct, gives *de facto* a distinct alternative answer to the question addressed by the corrected (37a). The questions addressed by corrective utterances are metaquestions. Here, (33) is logically equivalent to:

(39)   Which individual(s) is/are such that of all of the properties in (36), Mary has none apart from having invited that/those individual(s) for dinner?

Of course, this is not the same question as (35). The correcting (37b) asserts that it is Lyn who is the individual such that of all the Relevant properties in (36), Mary only has the property of inviting that person for dinner. This is the correct interpretation and in fact also gives a complete answer to the question (37a) addressed, (38), so that it is also contextually Relevant.

For similar examples, Partee (1991) claimed that there were embedded foci with no prosodic reflexes; the present account seems to improve on hers by avoiding the assumption of abstract foci with no surface reflexes. Vallduví (1990) took similar examples to discredit Rooth's general approach but offered no general explanation for the association with focus effects. In the present framework, the contribution of focus *is* conventional, albeit presuppositional instead of being proffered as in Rooth 1985, and so we can both account for the classic examples of association with focus and allow for examples like (37).

Another of the virtues of the present account is that it can explain association with focus effects in questions, as well as in assertions. So consider (40) and (41):

(40) Did Mary only invite [Lyn]$_F$?

(41) [Did Mary only invite Lyn]$_F$?

By (28), (40) is presupposed to address the question (42):

(42) Who did Mary only invite?

What (42) means can only be given in view of the context of utterance for (40), i.e., the QUD stack and the strategy which it (partially) embodies. One possibility is that it is the same as the question (35) discussed above. In that case, in order to be Relevant, and hence address (35)/(42), (40) must be part of a strategy to answer (35). But we have seen that (35) establishes a partition in which only certain of Mary's properties are Relevant. In order to form part of a strategy to address (35), answering (40) must entail at least a partial answer to (35) and hence pertain to which of the Relevant properties Mary has. This yields the correct domain restriction on *only* in (40), as it did for (32). (41) has broad focus and so presupposes only the Big Question, which its own answer, of course, will be part of a strategy to answer. But the Big Question doesn't give us any clues about the domain restriction of *only*. Here, only additional contextual factors can help, which seems appropriate. As we saw in the previous example, prosodic focus doesn't always give the intended domain restriction for the relevant operators.

Finally, notice that on this account the selection of the intended domain of *only* turns out to be purely pragmatic, a conversational implicature: only if the correct domain is assumed will the utterance be Relevant. Assuming that the speaker is cooperative, then, we must assume that that domain is intended. It might be objected that this type of approach to the problem is too weak to account for the robust facts about association with focus. What, for example, of implicature cancellation? Wouldn't this implicative account predict that association with focus is cancelable? Though I cannot go into this in detail here, let me briefly rebut this objection. Following Welker 1994, I believe that implicature cancellation is generally misunderstood. Given a particular context, conversational implicature cancellation is no more possible than the cancellation of conventional implicatures (in the sense of Karttunen & Peters 1979), i.e., presuppositions. What happens in the classic cases of so-called implicature cancellation (see, e.g., the papers on implicature in Grice 1989 or chapter 3 of Levinson 1983) is that the speaker clarifies the intended context for the utterance and often, in particular, corrects a misapprehension of the question being addressed and/or its role in the strategy of inquiry. So "implicature cancellation" might more aptly be called "*post hoc* clarification

(by the speaker) and revision (by the hearer) of intended context." Sometimes this is necessitated by breeches of conversational competence, while in other cases the speaker may actually set up the implicature on purpose, hastening to pretend it was not intended after the fact to avoid taking responsibility for it (a conversational implicature parallel with the gossip case in presuppositions). In the present framework, cancellation is not an issue. The theory of Information Structure makes it very precise both what a "context" of utterance is and what is required for an utterance to be Relevant in that context, so that the predictions are straightforward and unambiguous. "Cancellation" is a species of miscommunication, extratheoretical.

This illustrates the methodological advantages offered by the type of framework for pragmatic analysis proposed here. Though it is common to talk about the context of utterance of a given example, this framework, and the prosodic presupposition of congruence to a question in particular, force us to look at concrete facets of the utterance's context and predict how they directly influence its interpretation. If we assume something like InfoStr, in which the roles and relations between questions and answers are well defined, then determining the question under discussion gives us information about what is Relevant, puts constraints on logical aspects of interpretation like domain restriction and hence begins to illuminate the hitherto rather mysterious way in which context influences interpretation. This permits us to make falsifiable predictions about these pragmatic aspects of interpretation (cf. also Kadmon & Roberts 1986), legitimating the use of pragmatic explanations as alternatives to the elaboration of the semantic architecture of our theories. For association with focus, the domain of operators like *only* is pragmatically given, assuming only the independently motivated principles in (15), (25) and (28). No additional stipulations about lexical semantics need be made, structured propositions are not required and no special anaphoric elements associated with focus need be assumed to annotate surface structures. Insofar as Information Structure is independently motivated as a general framework for pragmatic explanation, the account carries no additional cost at all.

2.2.2   Contrast and alternatives

2.2.2.1   Focal presuppositions in utterances with contrastive topics

Jackendoff (1972) discussed an interesting phenomenon wherein multiple foci in a single utterance may bear two distinct intonation contours, which he called the *A*- and *B-contours*. In terms of the analysis of contours of the sort developed by J. Pierrehumbert (1980) and her associates, each of Jackendoff's contours appears to involve an entire intonation phrase, with a L+H* pitch accent on the focused subconstituent,[27] a L– phrase accent and distinct boundary tones: L% for A-contours, H% for B-contours. (Native speakers, at least, should be able to reconstruct what this sounds like on the basis of the question/answer pairings below and their consequences for prosody in the answer.) Jackendoff called the focused constituent in the B-contoured phrase the *independent focus*, that in the A-contoured phrase the *dependent focus*, and showed how the use of these accents correlates with certain question/answer pairings. In addition, the B-contour is not generally used alone but assumes the existence of an A-contour as well. For example, consider Jackendoff's (43a) and (43b):

(43)   a. [John]$_B$ ate [beans]$_A$
     b. [John]$_A$ ate [beans]$_B$

Along the lines of the analysis of similar examples in J. Pierrehumbert & Hirschberg 1990, these examples would be represented as in (44):

(44)   a.   [John]$_F$ [ate beans]$_F$
        L–H%   L–L%
     b.   [John]$_F$ [ate beans]$_F$
        L–L%   L–H%

Each of these has two intonation phrases, the bracketed constituents annotated with phrase accent/boundary tone sequences, each of which contains a focused constituent (cf. the principles of the phonology of focus in (22)). In what follows I will sometimes use the A- and B-accent notations for simplicity, but I assume that something like (44) is the more accurate transcription.

Jackendoff pointed out that (43a) would answer a question like *What about John — what did he eat?*, whereas (43b) would answer a question like *What about beans — who ate them?*. But the question/answer pairings couldn't be

---

27 Sometimes this sounds to me like a simple H* pitch accent, but I follow J. Pierrehumbert & Hirschberg (1990) in assuming L+H*.

switched: (43a) couldn't answer *What about beans — who ate them?*. I would propose the following account of these contours and their distribution within the present framework; the account is based on unpublished joint work with Nirit Kadmon (who might not, however, subscribe to the present framework within which it is developed).

Ignoring intonation phrasing, both (43a) and (43b) should presuppose that the question under discussion is:

(45) $\{u$ ate $u'\colon u, u' \in D\}$, i.e., *Who ate what?*

However, intuitively they aren't direct answers to (45) but respond instead to the distinct subquestions (46a) and (46b):

(46)   a.   What did [John]$_F$ eat?
           cf. *What about John — what did he eat?*
       b.   Who ate [beans]$_F$?
           cf. *What about beans — who ate them?*

The fact that (43a) and (43b) presuppose (45), as well as one of the subquestions in (46a) or (46b), is related to what's going on in the following type of dialogue, with no prior discourse, at least on a related subject:

(47)   a.   [When are you going to China]$_F$?
       b.   Well, I'm going to [China]$_B$ in [April]$_A$.

(47b) answers the question in (47a), but it does more than that. Its prosodic focal structure presupposes that the question under discussion isn't (47a) but the superquestion in (48):

(48)   When are you going to which place?
       i.e., $\{$you are going to u at $t\colon u$ a place, t a time$\}$

It is generally assumed that alternative sets, like operators' domains, are nonsingleton, nonempty. This question, then, implicates that there is more than one place that the speaker of (47b) is planning to visit. Of course, in this context, the superquestion hasn't already been accepted by the speaker of (47a). But she is likely to accommodate it out of curiosity and ask (49), i.e., ask for the rest of the information which would complete (47b)'s answer to the presupposed superquestion.

(49)   Oh? Where else are you going, and when?

In this case, the overall strategy of inquiry is $\langle(48), \{\langle(47a), \varnothing\rangle, \langle(49), \varnothing\rangle\}\rangle$, wherein the superquestion (48) has been accommodated.

Similarly, the B-accents in (43) tell us that the superquestion (45) is under discussion, that there is also under discussion a subquestion of one of the forms in (46a) or (46b) and further that (43a) could only answer (46a) while (43b) could only answer (46b). We might say that (43a) presupposes the strategy of inquiry $\langle(45), \{\langle(46a), \varnothing\rangle\}\rangle$, while (43b) presupposes the strategy $\langle(45), \{\langle(46b), \varnothing\rangle\}\rangle$. (43a) is an infelicitous move in response to the strategy $\langle(45), \{\langle(46b), \varnothing\rangle\}\rangle$, while (43b) is infelicitous in response to $\langle(45), \{\langle(46a), \varnothing\rangle\}\rangle$. Of course, the subquestions don't have to be explicitly asked any more than (45) does, but that isn't an issue given the semantic character of the presupposed questions in this theory and the abstract nature of the information structure of discourse. The goals and assumptions presupposed by the examples in (43) appear to be correctly captured by these correlated strategies.

In order to capture these facts, I assume that both A- and B-accented noun phrases are foci, so that the question in (45) is presupposed; this will follow from the F-marking on the utterances in (43).[28] Then how is the assumption of the immediate subquestions in (45) triggered? The crucial difference between the two accents in such examples appears to be the boundary tone, H% (or the sequence L–H%, should it turn out that in general that phrase+boundary tones form a morphological unit). The (L–)H% accent on the B-contoured intonation phrase marks its focused constituent as the independent argument to the abstract corresponding to (45), i.e., the one chosen first, whose choice then determines, or at least narrows, the choice of an alternative from the set corresponding to the other, A-contoured focus (hence, Jackendoff's term *dependent focus*). '(L–)H%' means 'answer still in progress — another focus to fill,' which explains why it cannot occur by itself (but can iterate, so long as there's an A-contoured focus around somewhere). If you choose this independent argument from one of the alternative sets associated with the question in (45), then the presupposed question under discussion will be either (50a) or (50b).

(50) a. $\{u \text{ likes } u': u, u' \in D \wedge u = \text{John}\}$

     b. $\{u \text{ likes } u': u, u' \in D \wedge u' = \text{beans}\}$

But (50a) is equivalent to (51a), which is the interpretation of (46a), and (50b) is equivalent to (51b), which is the interpretation of (46b):

---

28 Vallduví doesn't assume that contrastive topics, among which I take him to include B-accented elements in examples like Jackendoff's, are foci. But so far as I can see, this leaves him without an explanation for facts of the sort discussed here.

(51) a. $\{j \text{ likes } u' : u' \in D,\ j = |\,\text{John}\,|\}$

  b. $\{u \text{ likes beans} : u \in D\}$

So, (43a) and (43b) both presuppose (45) by virtue of identical placement of prosodic focus. But the location of a L–H% boundary sequence indicates a sequence of selection from the contrast sets for the foci and hence presupposes a subquestion to the prosodically presupposed question. Because this boundary sequence is located differently in (43a) and (43b), they presuppose distinct subquestions. The fact that for (43a), the presupposed questions (45) and (46a) form a question/subquestion strategy, means that (43a) itself presupposes such a strategy; similarly for (43b) and the strategy $\langle (45), \{\langle (46b), \varnothing \rangle\} \rangle$. Hence, on this account utterances which contain both independent and dependent foci presuppose not just a question under discussion but a possibly complex strategy of questions.

  We would of course need to look at more data on the occurrence of the L–H% boundary sequence to see if this is a plausible account for its effects across the board. Hence at present, this can only be a provisional proposal for the treatment of examples like those in (43). But note that the boundary tone sequence on B-accented phrases isn't that of question intonation, which involves a H–H% phrase+boundary tone sequence. In the other types of examples involving L–H% which I've reviewed informally, its proposed contribution to meaning is at least consistent with its apparent role here.

  Note that in these accommodationally enriched discourses, for example, $\langle (45), (46a), (43a) \rangle$, the answer, here (43a), is congruent to both the immediate question under discussion, (46a), and the superquestion, (45). But this does not by itself tell us that the entire strategy $\langle (45), \{\langle (46a), \varnothing \rangle\} \rangle$ is presupposed. Presumably, the latter fact should follow from presuppositions associated with the choice of the B-accent. This illustrates how the presupposition of focus formulated in (28), while necessary, is not yet sufficient to capture all the presuppositions of the prosodic structure of an English sentence.

(28) *Presupposition of prosodic focus in an utterance* $*\beta$:

  β is congruent to the question under discussion at the time of utterance.

To give a fully adequate account, we would need to look not only at the selection of focused constituents but also at both the intonation phrasing and the types of phrase accents and boundary tones associated with the

chosen intonation phrases (and most likely at the particular pitch accents chosen, as well). The complexity and richness of prosodic structure is too often overlooked in semantic/pragmatic studies of focus, and this sometimes leads to crucial misanalyses of certain types of examples. We will see one of these in the following section.

2.2.2.2 Contrastive focus within an utterance

There are a number of types of examples of what has often been called *contrastive focus* which seem to follow fairly straightforwardly from the framework and principles proposed here. I'll consider first examples that involve intersentential (or interclausal) examples like that in (52) below:[29]

(52)   Mary called Sue a Republican, and then [she]$_B$ insulted [her]$_A$.

Note that there is a B-accent on *she* in the second clause.

It does not suffice to explain the intended interpretation of the second clause in (52) to simply say that it offers an answer to (53):

(53)   Who insulted whom?

This is because (52) seems to have two additional presuppositions:

(54)   a. The first clause also reports an insult.
        b. The insulter and insultee roles are reversed in the second clause relative to the first.

However, I will argue that these presuppositions follow fairly straightfor-wardly from the view of prosodic focus sketched earlier in this section, so that this type of example doesn't motivate any extra principles pertaining to the interpretation of specifically contrastive focus.

First, notice that in uttering (52), *insulted* is "deaccented" (see Ladd 1980). Selkirk (1984) objects to the term *deaccenting* for the prosodic phenomenon in question, as it suggests that the constituent in question was initially accented and then lost its accent. Especially when the constituent is a verb, it's not clear that it would "normally" be prosodically prominent, let alone receive nuclear stress. But in (52), *insulted* receives a rendering which is noticeably flat when compared to its realization in the utterance of *Mary*

---

29 Rooth (1992b) discusses a similar example. It has been suggested to me that this type of example was first noticed by George Lakoff, but I don't know where. I'd appreciate any references that readers might point me to.

insulted Sue after *What happened next?* or even in the second clause of (52) when the implicature that being called a Republican is insulting is not intended. This type of deaccenting presupposes that the insult relation is already under discussion, as would be the case if (53) is the question under discussion. I would maintain that a conjoined structure is generally offered as a complex (partial) answer to a single question; this would follow from the theory of Information Structure in section 1, in particular from the requirement that any given move be Relevant and hence address the question under discussion.[30] The destressing on *insulted* in the second conjunct of (52) presupposes that the question is (53). If the first conjunct is then to be part of the answer to this question, it must report an insult as well, yielding the implication in (54a).

The presupposition (54b) arises because the only way that both (54a) can hold and the pronouns take *Mary* and *Sue* as their antecedents and yet the second clause still be informative, is for the insulter/insultee roles to be reversed. The B-accent on *she* tells us that not only does the clause address (53) but also a subquestion, one of the two questions which would result if we replace *she* by one of the salient women in the context, yielding either *Who (of Mary and Sue) did Mary insult?* or *Who (of Mary and Sue) did Sue insult?*. But the first conjunct has already told us that Mary insulted Sue, so the only subquestion which can be nonredundantly addressed by *she insulted her* is the second, i.e., *Who (of Mary and Sue) did Sue insult?*. Then the nonreflexive pronominal object must denote Mary and not Sue.[31] Hence, the presupposed question is contextually restricted in such a way as to presuppose the reversal of roles in the answer. No special stipulations about contrast between pairs of NPs is required.

A similar story can be told about examples like (55) from Schwarzschild 1994a:

(55) a. John was a victim of suicide(, but)
      b. THE MOB killed SAM.

The deaccenting of the verb in (55b) presupposes that it's part of the question under discussion. Whether it's conjoined with (55a) or simply follows it, the

---

30 Of course, much more needs to be said to support this claim, but here it would take me too far afield. This is ultimately a question about the rhetorical roles played by complex clausal structures. See section 3 for suggestions about how the theory of Information Structure may relate to rhetoric.

31 We could replace the second conjunct by *and then SHE insulted HERSELF*, so that Sue insults Sue, with the same implicature that calling someone a Republican is insulting.

content of the two utterances plus the prosody in (55b) will suggest that both are partial answers to the same question, i.e., *Who killed whom?*. Notice the difference between the deaccented predicate in (55b) and the realization of the predicate in (55c):

(55)   c.   The mob poisoned Sam.

For me, when (55c) follows (55a), *mob, Sam* and *poisoned* are all accented. I believe that the NPs both have B-accents, the verb an A-accent, suggesting that the question under discussion is *How was who killed by whom?*. The difference between ⟨(55a), (55b)⟩ and ⟨(55a), (55c)⟩ seems to be that John's having committed suicide in (55a) entails his having been killed, albeit by himself, whereas it doesn't entail his having been poisoned. Hence, the killing in (55b) is "old" information, whereas the poisoning in (55c) is not.

The principle which Schwarzschild offers to account for such examples follows from the theory offered here:

(56)   Contrast constraint (Schwarzschild 1994a):

If *B* is uttered in contrast with *A*, then:

a.   *A* is not identical in meaning to *B*.
b.   The meaning of *A* can be gotten by replacing focussed elements of *B* with elements of the same semantic type and computing the meaning of the resulting expression.

However, this principle can be shown to follow from the present account plus general Gricean principles. The nonidentity in meaning, (56a), follows from Grice's Quantity 1 maxim, requiring that one be (as) informative (as possible), and the replacement clause (56b) from the assumption that both the clauses in these examples address the same question, plus congruence.

But there is a type of example involving intraclausal contrast which poses *prima facie* problems for the account proposed here. Consider the following example from Rooth 1992a:

(57)   An [American]$_F$ farmer was talking to a [Canadian]$_F$ farmer…

As transcribed by Rooth, (57) is predicted by my account to presuppose a question under discussion like (58a) or (58b):

(58)   a.   What kind of farmer was talking to what kind of farmer?
        b.   What kinds of farmers were talking (to each other)?

But somehow this doesn't seem quite right; those native English-speaking linguists with whom I've discussed this generally seem to agree that (59) may be felicitously uttered out of the blue:

(59)   An AMERICAN farmer was talking to a CANADIAN farmer.

(59) is a fairly crude impressionistic transcription, which only notes the primary accented elements in the utterance. If Rooth has correctly analyzed (59) as (57), then the theory of prosodic focus proposed in this paper makes an incorrect prediction and offers no account of the contrast in (59). Since the contrast is intraclausal, we cannot address the contrast by claiming that the contrasted constituents suggest alternative (partial) answers to a single question under discussion.

However, there is reason to question whether (57) is the correct analysis of the impressionistically transcribed (59). Recall the phonological principle (22d), repeated here:

(22)   The phonology of focus:

  a. There is at least one intonation phrase per sentential (or sentential-fragment) utterance.

  b. There is at least one focused subconstituent (possibly nonproper) per intonation phrase. This *focused constituent* is marked with the feature F in what follows.

  c. There is at least one pitch accent per focused constituent, associated with a subconstituent.

  d. Every pitch accent must be associated with material in a focused constituent.

  e. There is one phrase accent (H– or L–) and one boundary tone (H% or L%) per intonation phrase.

  f. The string-final pitch accent in the focused constituent is assigned the most prominent stress in the intonation phrase (Nuclear Stress Rule).

This predicts that all the material in (57)/(59) except the two adjectives should be deaccented. However, this is only a felicitous prosodic contour for (59) when it is in fact the answer to one of the questions in (58) and not when it is uttered out of the blue. In the latter, out-of-the-blue type of context, both the first noun *farmer* and the verb *talking* are accented, although they do not bear nuclear stress; they do not have the flat, deaccented contour. But how can we

reflect the sense that *American* and *Canadian* are the primarily contrasted elements of (59) while avoiding the transcription in (57) and hence incorrectly predicting a lack of pitch accents on other elements of the utterance? One possibility is reflected in (60) below. In it, intermediate intonational phrases (see Beckman & J. B. Pierrehumbert 1986) are marked with '–', full intonation phrases with '%'; as is standard in the ToBI transcription system for English (see Beckman & Ayers 1997), starred tones are pitch accents (aligned with syntactic constituents), T– (where T is either a high or a low tone) is a phrase marker (aligned with an intermediate phrase) and T% is a boundary tone (aligned with a full intonational phrase):

(60) [[An American] [farmer]] [[was talking] [to a Canadian farmer]]
          H*   L–   H* L–L%   H*   L–   H*   L–L%

The existence of two full intonational phrases is indicated by a slight pause and lengthening at the end of the first, after *farmer*, as well as the falling phrase accent and boundary tone. In addition, each full intonation phrase involves two intermediate phrases, with intermediate phrase boundaries between *American* and *farmer* and again between *talking* and *to*. In this structure, we might claim that each of the intermediate phrases is focused (i.e., each has maximally broad focus). This is consistent with the pitch accent placement, which in the first three intermediate phrases falls on the final element, i.e., on *American*, *farmer* and *talking*. Consistent with the informants' realization of (59), and unlike the realization of Rooth's (57), the only deaccented element in the string is the second token of *farmer*. This lack of accent may be accounted for by the anaphoric deaccenting argued for in detail by Ladd (1980) and Selkirk (1984), with the antecedent in (59)/(60) being the first token of the noun; hence, the deaccented final element is still consistent with assuming that the final intermediate intonation phrase has broadest focus.

The utterance in (60) would be congruent with the question:

(61) Who was doing what?

Unlike the questions in (58), (61) is compatible with the intuition that (59) may be uttered out of the blue, if we take the latter to mean that the utterance addresses a very general question.

But even if this is correct, what can we say about the placement of primary stress on the adjectives? That does plainly seem to be motivated by the intended contrast between their values. David Dowty (personal communication) pointed out a related type of example, inspired by earlier examples due to

Horn (1985, 1989):[32]

(62)   A PROactive farmer was talking to a REactive farmer.

Example (62) may also be uttered more or less out of the blue; the prosodic facts in such an utterance — lack of deaccenting on the first adjective stem, noun and verb — parallel those for (59). This suggests that the focal structure of (62) is more like that in (60) than that in (57). In fact, it doesn't seem that this type of example can be understood to involve narrow focus on the contrasted prefixes, insofar as focus is taken to correlate with syntactic constituents. What kind of question under discussion could such narrow focus indicate: *What kind of -active farmer was talking to what kind of -active farmer?* Since the prefix *re-* in *reactive* isn't productive in contemporary English, such a narrow focus would be not only sublexical but submorpho-logical, and this is surely related to the infelicity of such a question. Several other pairs of adjectives would work in the same way, such as *UNinterested* and *DISinterested*, so the phenomenon is fairly general.

It seems that in examples like (62), as in (59) under the analysis in (60), a contrast is superimposed on an otherwise broadly focused prosodic structure. With respect to (60), note that, whatever physical measure(s) actually correlate with perceived relative stress, none of the principles for the phonology of English prosody in (22) bear on the comparative value of multiple intonation phrases or intermediate intonation phrases with respect to these measures. In fact, of course, we know that final phrases are not always stronger in this respect than those earlier in an utterance; witness the *sotto* voice often used for intraclausal parenthetical elements or asides. Hence, what I take (60) to suggest is that there is a mechanism for determining relative weight of multiple intermediate phrases which goes beyond the rules for intonation phrasing and placement of accent and focus, as given in (22).

Of course, determining whether (60) is, in fact, the correct analysis of the out-of-the-blue utterance of (59) would require phonetic analysis which I am not prepared to undertake. I do think it obvious that (57) is *not* the correct analysis, given the way focus placement relates to deaccenting and the implications of the latter. Perhaps the clear moral to be drawn is that we must be more careful about what we claim about the focal structure of utterances, as there are more matters to be considered than the impression of primary stress or contrast. In addition, the discussion raises other important

---

32 Horn gives examples combining contrastive focus with metalinguistic negation. He credits Bolinger (1961) and Carlson (1982) with earlier examples of that type (Horn 1989: 434-435).

questions about the pragmatics of prosody, including especially the function of the intonational phrasing which also seems to play a role in conveying the contrast. My suspicion that in examples like these the contrast involves a metalinguistic use of prosody is partly inspired and strengthened by the role that focus plays in Horn's metalinguistic negation and in so-called *alternative questions*, a phenomenon to which I turn now.

2.2.2.3 Yes/no vs. alternative questions

The string in (63), von Stechow notes, is ambiguous between a yes/no-question reading and an alternative-question reading. Though he doesn't discuss prosody, I believe these readings correspond to the alternate realizations roughly indicated in (63a) and (63b), respectively. (I ignore pitch accents here and give only the phrase accents, foci and boundary tones):

(63) Do you want coffee or tea?

    a. [Do you want coffee or tea]$_F$?
                          H–H%

    b. [Do you want [coffee]$_F$] [or [tea]$_F$] ?
                    H–H%   L–L%

These should have the interpretations given in (64), where *coffee* ∨ *tea* is the meet of | *coffee* | and | *tea* |:

(64) a. {you want coffee ∨ tea}
     b. {you want u: $u \in$ {coffee, tea}} = {you want coffee, you want tea}

The interpretation for (63a) in (64a) is arrived at straightforwardly. The simple yes/no question will set up a partition on the context set which has only two cells, one containing those worlds in which the single proposition in the set (64a) is true and one containing those worlds in which it is false. The issue is how we arrive at the interpretation in (64b). Note that this is equivalent to the question denoted by the conjoined *Do you want coffee and do you want tea?*. The interpretation of (63) in (64b) thus involves a semantic version of conjunction reduction, converting the object language disjunction into metalanguage conjunction. von Stechow proposes that *or* in questions generally behaves as set union: at the content level it joins sets of worlds, but at the alternative level it joins sets of propositions. In his theory, (64b) is the straightforward reading, while (64a) must be explained by some

additional mechanism (he basically uses quantifying-in). The account says nothing about the correlation between prosody and interpretation.

I would propose that *or* ordinarily (even in questions) has the standard Boolean interpretation, so that (63a)/(64a) is quite straightforward. But *or* sometimes has a metalinguistic use, whose function is to indicate contrasting alternatives expressed by the independently focused disjuncts. Compare Horn's (1985, 1989) metalinguistic negation, which is also generally accompanied by narrow focus on the corrected or cancelled constituent. Horn (1989: 379ff) argues that other operators, including *or*, also have metalinguistic interpretations and offers examples like (65):

(65)   Is the conductor Bernst[í^y]n or Bernst[á^y]n?

This is a question about the pronunciation of the conductor's name and hence is clearly metalinguistic.

I would argue that in its metalinguistic use, *or* functions as suggested by von Stechow, yielding the union of the alternatives given by the disjuncts, which set then gives alternative values for a single focused argument in the utterance in question. Hence, (63b) presupposes that the question under discussion is *What do you want?*, {you want $u: u \in D$}. The particular alternatives provided suggest the natural class which we take to restrict D, so that it must include at least coffee and tea. The default final phrase/boundary tone sequence in (63b) is L–L%, which suggests finality, i.e., that the alternatives given are the full set. Thus we arrive at the question {you want $u: u \in D$ and $u \in$ {coffee, tea}}. This is the set of propositions *want(you, coffee)* and *want(you, tea)* in (64b), as desired, equivalent to asking two conjoined yes/no questions. (I take it that the exclusive sense of the disjunction arises from a conversational implicature, an assumption which is now fairly standard in the literature.) A final H–H% sequence is also possible (like the H–H% generally present at the end of the intonation phrase including *coffee*), which would suggest that there may be other beverages available as well.[33] Using metalinguistic disjunction, the desired result is obtained without conjunction reduction, and (63a) is not more complicated than it

---

33 That is, the alternative question (64b) may also be asked with the focal structure in (i):

(i)   Do you want [coffee]$_F$ or [tea]$_F$?
                     H–H%   H–H%

Like (63b), the prosody in (i) seems to implicate that the hearer does want something; this is due to the fact that it is presupposed to be part of a strategy to answer the question *What do you want?*. (ii) also denotes (64b), but it doesn't presuppose such a superquestion.

ought to be.

This view of how metalinguistic disjunction works also appears to be extendable to explain examples like Horn's (65). We do ask questions like, *Is the conductor Bernst[ʔʸ]n?*, meaning, 'Is his name pronounced like this?' and the present view would make (65) denote a conjunction of two such questions. This is comforting, but of course it doesn't mean that what I have suggested here constitutes a theory of the phenomena in question. I have only argued that these phenomena are compatible with the theory of prosodic presuppositions outlined in this section.

3 Further applications of the theory of Information Structure

The theory sketched in section 1 and section 2 seems to capture what's essential about the type of alternative semantics promoted by Rooth (1992a) but without any special anaphoric mechanism or focus sensitivity. And it also promises to lead to a general, presuppositional theory of domain restriction for operators in natural language, along lines sketched in Rooth 1992a, Roberts 1995, Calcagno 1996 and von Fintel 2004. But I think that it potentially has much broader applications. Here I will briefly outline some of the connections I see between this theory and other work in the pragmatic literature (from various fields) on a variety of issues. The thread that connects them is the notion of a plan or intention. Probably not everyone whose work I mention here would be comfortable with translating those notions into that of the goal, or question under discussion in a discourse, with the information structure in which that goal plays a part reflected in the common ground. However, if we are willing to do so, information structure turns out to be a central facet of pragmatic theory, unifying a number of apparently disparate issues and opening up some intriguing possibilities. The central claim is that the analysis of all these various pragmatic phenomena requires a richer notion of context than that generally given, one of the sort sketched in section 1. Then many of the pragmatic phenomena in question can be seen as presuppositions about the structure of the context of utterance, its information structure, while others involve the generation of

---

(ii)   [Do you want coffee]$_F$ or [tea]$_F$?
                        H–H%   H–H%

Both (i) and (ii) clearly ask two yes/no questions, each marked by the H– phrase accent and H% boundary tone. All of this, of course, requires further investigation. See J. Pierrehumbert & Hirschberg 1990: 302ff for relevant discussion.

inferences which are partly triggered by constraints on how an utterance fits into that information structure. The task of analyzing those presuppositions or determining how those inferences are drawn is considerably lightened when we have a clearer notion of the organization of the flow of information in discourse.

Thomason (1990) argues that the common ground, plans (as developed in Planning Theory in artificial intelligence) and accommodation should be central components of a theory of pragmatics. He sketches how the interlocutors' discourse plans, information in the common ground and plan inferencing play a crucial role in the calculation of the Gricean conversational implicatures intended by speakers; this is quite in keeping with Grice's own consistent relativization of implicature (and in fact, of nonnatural meaning itself) to the speaker's intentions.[34] Further, both plans and conversational implicatures are accommodated to the common ground. In a sense, the present proposal can be seen as an extension of Thomason's vision of a pragmatic theory, using semantic questions and strategies of inquiry to capture the notions of *goal* and *plan*. As Sperber & Wilson (1986) argue in detail, one problem with the Gricean approach to date has been the vagueness in his definitions of the conversational maxims and hence the indeterminacy about their applicability. Besides Relation, Grice defined both maxims of Quantity relative to the "purposes of the conversation." It doesn't seem unreasonable to hope that their senses, as well as the often-cited problem of determining how to balance their conflicting requirements, might also be clarified through exploring their role in a framework based on information structure. Welker (1994) takes up the basic idea sketched in Thomason 1990 and develops a theory of plans in the common ground, in which conversational implicatures are generated by applying plan inference rules to contextual information (including already available information about the interlocutors' plans) plus the content of the utterance itself. Then, like other contextual entailments, implicatures are accommodated to the common ground. Her approach explores several types of implicature which are generally overlooked in less systematic approaches and reveals the close relationship of conversational implicatures to other kinds of contextual entailments. In view of the work of Thomason and Welker, I speculate that we can use the same structures motivated here for focus in the generation of conversational implicatures, characterizing discourse plans in terms of

---

34 I believe that Thomason's student McCafferty (1987) takes a similar approach to conversational implicature, though I'm not directly familiar with his work.

strategies of inquiry. Roberts (1996c) is a first sketch of how this would work.

Some who work in Planning Theory (see Perrault 1990 and references therein) have recognized its utility in explaining and motivating the characteristics of various kinds of speech acts in discourse, characteristics which are otherwise simply stipulated, claimed to be primitive. In these theories, plan inferencing mechanisms may be used to draw conclusions about what speech act a speaker intended, even though the speech act type need not, contra Searle (1969) *inter alios*, be part of the conventional content of the utterance. In terms of the present theory, this suggests that we can infer speech act information (is that a threat or a promise, an assertion or a warning?) partly from the relationship of an utterance to the strategy of inquiry in which it plays a role and partly from the relationship of that strategy of inquiry to the domain plan(s) which it serves.

There is also a good deal of work on discourse cohesion and discourse segmentation and the role of the latter in anaphora and inference, which I think could be related to the perspective developed here. One example is the work of Grosz & Sidner (1986) on the use of what they call the Intentional Structure of discourse to identify discourse segments, which in turn, they claim, play a constraining role on the possibility of anaphora in discourse. Sperber & Wilson (1986) attempt to reduce Gricean maxims of conversation to entailments derived from principles of rational behavior and facts about the limitations of human cognitive processing;[35] from the facts about cognitive processing follow coherence constraints such as Relevance (which they give a definition quite different from that in (15) above), while from the principles of rational behavior follow the development of rational strategies for achieving one's ends. They also discuss how Relevance plays a role in limiting the domains for inferencing in discourse by restricting access to information which, although in the common ground, is not relevant to the issue at hand. The latter connection is particularly promising, suggesting that the strategies we develop limit the kinds of inferences we're likely to draw from the information we're given access to, given the common ground.

Another desideratum might be to develop a theory of topics and of topicalization (syntactic preposing), based on information structure. The discussion in section 2.2.2 suggests a way of approaching the function of contrastive topics in terms of information structure, an approach which appears to be related to that of Büring 1999. However, topics, even when

---

35 Sperber & Wilson also attempt to reduce all the Gricean maxims to the single maxim of Relevance. However, I think this aspect of their argument is less successful.

topicalized, are not all necessarily contrastive; for example, sometimes they do not bear Jackendoff's B-accent (see Ward 1985 for lots of examples of topicalized constituents which are noncontrastive). Further, topics as defined by most authors aren't necessarily all preposed, so that it would seem desirable to give independent, though perhaps related, accounts of the functions of topichood and of topicalization. To my knowledge, no one has yet proposed a set of necessary and sufficient conditions for either topichood or for topicalization, so that it isn't entirely clear what the former notion is, what the latter's functional role is or what their relationship may be. Even worse, I think there is reason to be skeptical about whether there is such a universal function as Topic, or Vallduví's (1993) Link, or even whether there is a unified pragmatic function of topicalization/preposing; apart from contrastive topics, there is no consistent conventional indication of topic- or Linkhood in a language like English; see Culicover & Rochemont 1983 for suggestions that preposing may have multiple functions in English, and see McNally 1998 for more discussion of the problems in defining what is meant by *Topic*. Focus preposings would also fall under the theory of focus, but of course that wouldn't by itself explain their preposability. One thing which might be promising with respect to the utility of the present theory for the development of a theory of topics and topicalization is the possibility of explaining the contextual source of the salient partially ordered sets (*PoSets*) that Ward (1985) argues are presupposed by topicalization (though he doesn't explain why this should be or how these sets are contextually introduced in the general case).[36] If topicalized constituents are prosodically focused, which a preliminary review suggests that they are, these PoSets would be their corresponding alternative sets, and we would expect that they could be retrieved by studying the strategy of inquiry in which these constructions occur and the common ground to determine what is Relevant.

There has been a good deal of interest lately in the use of rhetorical relations to structure discourse, following the work of Mann & Thompson (1987). I believe that Aristotle (I don't know where) claimed that logic was the basis for a science of rhetoric. That would be an interesting point of departure in exploring the relationship of rhetorical structures to the information structure posited here, with particular types of rhetorical relations over stretches

---

36 This proposal in Ward 1985 is closely related to the assumption in Hirschberg 1985 that scalar implicature also requires salient PoSets, a requirement which should also fall out of the present approach, under the relation between implicature and prosodic focus noted in Rooth 1992a.

of discourse characterized as types of strategies. Rhetorical relations can often, at least, be characterized in terms of questions and answers: e.g., the use of a *why*-question and its answer to characterize explanations, etc. I suspect, however, that these relations often serve the goal, or question under discussion, in another respect — the goal of discourse is only partly to offer more information and partly to achieve consensus about the value of the information contributed. So some rhetorical structures are intended principally to convince one's hearers that the information offered is worth adding to the common ground, e.g., by showing how it follows from or explains other known facts, etc. Studying these might lead to a more subtle grasp of the nature of information structure and, in particular, of what it means to address a question under discussion.

Finally, if all these connections can be made, one of the advantages of this approach is that we only need to impose one primitive type of structure on the information in discourse, other than encyclopedic semantic networks and certain binary distinctions such as those between facts and default assumptions or between explicitly introduced information and that which was only implicated, etc.[37] This hope is the basis of the subtitle of this paper — *Towards an integrated theory of formal pragmatics*. If pragmatic explanations are ever to be theoretically reputable, they must be couched in terms of a theory of pragmatics which is explicit enough to be defeasible and sufficiently broad in scope to have something to say about the relationship between various kinds of purportedly pragmatic phenomena. The theory in section 1 aims at such explicitness, and the brief notes in this section are intended to suggest its potential breadth. The proposed application to the analysis of focus is intended to suggest how such a theory could take some of the burden off of semantic theory and ultimately I suspect off of syntactic theory as well, resulting in a simpler overall theory of interpretation while providing empirically superior accounts. Any theories which purport to offer pragmatic explanations should be gauged against such overarching desiderata and not solely on the basis of claims to address some particular linguistic phenomenon in isolation.

---

37 Louise McNally (personal communication) pointed out to me that Information Structure may not subsume the kind of relation between kinds of information sometimes referred to as the *figure/ground relation*, a relation sometimes apparently reflected in the use of frame adverbials, in the distinction between conjoined and subordinated clauses, etc. Though I suspect that this relationship should also be characterized presuppositionally, it may be that the kinds of presuppositions involved differ in interesting ways from those associated with prosodic focus.

References

Baüerle, Rainer, Urs Egli & Arnim von Stechow (eds.). 1979. *Semantics from different points of view* (Springer Series in Language and Communication 6). Berlin, Germany: Springer-Verlag.

Beckman, Mary E. & Gayle M. Ayers. 1997. Guidelines for ToBI labeling, *version 3.0.* http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf.

Beckman, Mary E. & Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309. http://www.jstor.org/stable/4615401.

Bolinger, Dwight. 1961. Contrastive accent and contrastive stress. *Language* 37(1). 83–96. http://dx.doi.org/10.2307/411252.

Büring, Daniel. 1999. Topic. In Peter Bosch & Rob van der Sandt (eds.), *Focus: linguistic, cognitive, and computational perspectives* (Studies in Natural Language Processing), 142–165. Cambridge, UK: Cambridge University Press.

Calcagno, Mike. 1996. Presupposition, congruence, and adverbs of quantification. In Jae-Hak Yoon & Andreas Kathol (eds.), *Papers in semantics* (Working Papers in Linguistics 49). The Ohio State University. http://linguistics.osu.edu/files/linguistics/workingpapers/osu_wpl_49.pdf.

Carlson, Lauri. 1982. *Dialogue games: an approach to discourse analysis* (Synthese Language Library 17). Dordrecht, The Netherlands: D. Reidel.

Culicover, Peter W. & Michael S. Rochemont. 1983. Stress and focus in English. *Language* 59(1). 123–165. http://dx.doi.org/10.2307/414063.

von Fintel, Kai. 1994. *Restrictions on quantifier domains.* Amherst, MA: University of Massachusetts Amherst PhD thesis. http://www.semanticsarchive.net/Archive/jA3N2IwN/fintel-1994-thesis.pdf.

von Fintel, Kai. 2004. A minimal theory of adverbial quantification. In Hans Kamp & Barbara H. Partee (eds.), *Context-dependence in the analysis of linguistic meaning* (Current Research in the Semantics/Pragmatics Interface 11), 137–175. Amsterdam, The Netherlands: Elsevier.

Ginzburg, Jonathan. 1994. An update semantics for dialogue. *First International Workshop on Computational Semantics (IWCS 1).* 111–120.

Ginzburg, Jonathan. 1995a. Resolving questions, I. *Linguistics and Philosophy* 18(5). 459–527. http://dx.doi.org/10.1007/BF00985365.

Ginzburg, Jonathan. 1995b. Resolving questions, II. *Linguistics and Philosophy* 18(6). 567–609. http://dx.doi.org/10.1007/BF00983299.

Ginzburg, Jonathan. 1996. Dynamics and the semantics of dialogue. In Jerry Seligman & Dag Westerståhl (eds.), *Logic, language and computation*, vol. 1 (CSLI Lecture Notes 58), 221–237. Stanford, CA: Center for the Study of Language & Information (CSLI), Stanford University.

Grice, Herbert Paul. 1989. *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Groenendijk, Jeroen & Martin Stokhof. 1984. *Studies on the semantics of questions and the pragmatics of answers*. Amsterdam, The Netherlands: Institute for Logic, Language & Computation (ILLC), University of Amsterdam PhD thesis. http://dare.uva.nl/record/123669.

Grosz, Barbara J. & Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3). 175–204. http://www.aclweb.org/anthology-new/J/J86/J86-3001.pdf.

Halliday, Michael Alexander Kirkwood. 1967. Notes on transitivity and theme in English: part 2. *Journal of Linguistics* 3(2). 199–244. http://dx.doi.org/10.1017/S0022226700016613.

Hamblin, Charles Leonard. 1973. Questions in Montague English. *Foundations of Language: International Journal of Language and Philosophy* 10(1). 41–53. Reprinted in Partee 1976: 247–259.

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst, MA: University of Massachusetts Amherst PhD thesis. http://semanticsarchive.net/Archive/jA2YTJmN/.

Heim, Irene. 1983. On the projection problem for presuppositions. *Second West Coast Conference on Formal Linguistics (WCCFL 2)*. 114–125.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221. http://dx.doi.org/10.1093/jos/9.3.183.

Higginbotham, James. 1996. The semantics of questions. In Shalom Lappin (ed.), *The handbook of contemporary semantic theory*, 361–383. Oxford, UK: Blackwell.

Hintikka, Jaakko. 1972. *Logic, language-games and information: kantian themes in the philosophy of logic*. Oxford, UK: Oxford University Press.

Hintikka, Jaakko. 1981. On the logic of an interrogative model of scientific inquiry. *Synthese* 47(1). 69–83. http://dx.doi.org/10.1007/BF01064266.

Hintikka, Jaakko & Esa Saarinen. 1979. Information-seeking dialogues: some of their logical properties. *Studia Logica: An International Journal for Symbolic Logic* 38(4). 355–363. http://dx.doi.org/10.1007/BF00370473.

Hirschberg, Julia. 1985. *A theory of scalar implicature*. Philadelphia, PA: University of Pennsylvania PhD thesis. http://repository.upenn.edu/dissertations/AAI8603648.

Horn, Laurence R. 1985. Metalinguistic negation and pragmatic ambiguity. *Language* 61(1). 121–174. http://dx.doi.org/10.2307/413423.

Horn, Laurence R. 1989. *A natural history of negation*. Stanford, CA: Center for the Study of Language & Information (CSLI), Stanford University.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar* (Current Studies in Linguistics 2). Cambridge, MA: The MIT Press.

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara H. Partee (eds.), *Quantification in natural languages*, vol. 2 (Studies in Linguistics and Philosophy 54), 451–486. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Kadmon, Nirit. 2001. *Formal pragmatics: semantics, pragmatics, presupposition, and focus*. Oxford, UK: Blackwell.

Kadmon, Nirit & Craige Roberts. 1986. Prosody and scope: the role of discourse structure. In Peter T. Farley Ann M. Farley & Karl-Erik McCullough (eds.), *Chicago Linguistics Society (CLS) 22(2): papers from the Parasession on Pragmatics and Grammatical Theory*, 16–28. Chicago, IL: University of Chicago.

Karttunen, Lauri. 1973. Presuppositions of compound sentences. *Linguistic Inquiry* 4(2). 169–193. http://www.jstor.org/stable/4177763.

Karttunen, Lauri & Stanley Peters. 1979. Conventional implicature. In *Syntax and Semantics Volume11: Presupposition*, 1–56.

Krifka, Manfred. 1992. A compositional semantics for multiple focus constructions. *Semantics and Linguistic Theory (SALT) I* (10). 127–158. http://elanguage.net/journals/salt/article/view/1.127/1558.

Ladd, D. Robert, Jr. 1980. *The structure of intonational meaning: evidence from English*. Bloomington, IN: Indiana University Press.

Levinson, Stephen C. 1983. *Pragmatics* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.

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(1). 339–359. http://dx.doi.org/10.1007/BF00258436. Reprinted in Baüerle, Egli & von Stechow 1979: 172--187.

Liberman, Mark & Janet Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In Mark Aronoff, Richard Oehrle, Frances Kelley & Bonnie Wilker Stephens (eds.), *Language sound structure*, 157–233. Cambridge, MA: The MIT Press.

Mann, William C. & Sandra A. Thompson. 1987. *Rhetorical structure theory: a theory of text organization* (Information Sciences Institute Research Report 87-190). Marina del Rey, CA: Information Sciences Institute.

McCafferty, Andrew S. 1987. *Reasoning about implicature: a plan-based approach*. Pittsburgh, PA: University of Pittsburgh PhD thesis.

McNally, Louise. 1998. On recent formal analyses of topic. *The Tbilisi Symposium on Language, Logic, and Computation: Selected Papers* (14). 147–160. http://www.upf.edu/pdi/louise-mcnally/_pdf/publications/tbilisi.pdf.

Partee, Barbara H. (ed.). 1976. *Montague Grammar*. New York, NY: Academic Press.

Partee, Barbara H. 1991. Topic, focus, and quantification. *Semantics and Linguistic Theory (SALT) I* (10). 159–188. http://elanguage.net/journals/index.php/salt/article/view/1.159/1557.

Perrault, C. Raymond. 1990. An application of default logic to speech act theory. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication* (System Development Foundation Benchmark Series), 161–185. Cambridge, MA: A Bradford Book, The MIT Press.

Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis. http://dspace.mit.edu/handle/1721.1/16065.

Pierrehumbert, Janet & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication* (System Development Foundation Benchmark Series), 271–311. Cambridge, MA: A Bradford Book, The MIT Press.

Roberts, Craige. 1995. Domain restriction in dynamic semantics. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara H. Partee (eds.), *Quantification in natural languages*, vol. 2 (Studies in Linguistics and Philosophy 54), 661–700. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Roberts, Craige. 1996a. Anaphora in intensional contexts. In Shalom Lappin (ed.), *Handbook of contemporary semantic theory*, 215–246. Oxford, UK: Blackwell.

Roberts, Craige. 1996b. Information structure in discourse: towards an integrated formal theory of pragmatics. In Jae-Hak Yoon & Andreas Kathol

(eds.), *Papers in semantics* (Working Papers in Linguistics 49). The Ohio State University. http://linguistics.osu.edu/files/linguistics/workingpapers/osu_wpl_49.pdf.

Roberts, Craige. 1996c. Information structure, plans, and implicature. Paper presented at the Association for the Advancement of Artificial Intelligence (AAAI) Spring Symposium on Computational Implicature: Computational Approaches to Interpreting and Generating Conversational Implicature. Stanford, CA: Stanford University.

Roberts, Craige. in preparation. The information structure of discourse. The Ohio State University.

Rochemont, Michael S. & Peter W. Culicover. 1990. *English focus constructions and the theory of grammar* (Cambridge Studies in Linguistics 52). Cambridge, UK: Cambridge University Press.

Rooth, Mats. 1985. *Association with focus.* Amherst, MA: University of Massachusetts Amherst PhD thesis. http://scholarworks.umass.edu/dissertations/AAI8509599.

Rooth, Mats. 1992a. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116. http://dx.doi.org/10.1007/BF02342617.

Rooth, Mats. 1992b. Ellipsis redundancy and reduction redundancy. *Stuttgart Ellipsis Workshop* (Arbeitspapiere 29).

Rooth, Mats. 1996. Focus. In Shalom Lappin (ed.), *The handbook of contemporary semantic theory*, 271–297. Oxford, UK: Blackwell.

Schwarzschild, Roger. 1994a. Association with focus: semantics or pragmatics. The Hebrew University of Jerusalem. Manuscript.

Schwarzschild, Roger. 1994b. The contrastiveness of associated foci. The Hebrew University of Jerusalem. Manuscript.

Searle, John R. 1969. *Speech acts: an essay in the philosophy of language.* Cambridge, UK: Cambridge University Press.

Selkirk, Elisabeth O. 1984. *Phonology and syntax: the relation between sound and structure* (Current Studies in Linguistics 10). Cambridge, MA: The MIT press.

Sperber, Dan & Deirdre Wilson. 1986. *Relevance: communication and cognition* (The Language and Thought Series). Cambridge, MA: Harvard University Press.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Pragmatics* (Syntax and Semantics 9), 315–332. New York, NY: Academic Press.

von Stechow, Arnim. 1991. Focusing and backgrounding operators. In Werner Abraham (ed.), *Discourse particles: descriptive and theoretical investiga-*

tions on the logical, syntactic and pragmatic properties of discourse parti-cles in German (Pragmatics & Beyond: New Series 12), 37–84. Amsterdam, The Netherlands: John Benjamins Publishing Company.

Thomason, Richmond H. 1990. Accommodation, meaning, and implicature: interdisciplinary foundations for pragmatics. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication* (System Development Foundation Benchmark Series), 325–363. Cambridge, MA: A Bradford Book, The MIT Press.

Vallduví, Enric. 1990. *The informational component.* Philadelphia, PA: University of Pennsylvania PhD thesis.

Vallduví, Enric. 1993. Information packaging: A survey. Prepared for the *Word Order, Prosody, and Information Structure* (*WOPIS*) initiative. http://groups.inf.ed.ac.uk/hcrc_publications/.

Vallduví, Enric & Ron Zacharski. 1994. Accenting phenomena, association with focus, and the recursiveness of focus-ground. *Ninth Amsterdam Colloquium.* 683–702.

Walker, Marilyn A. 1993. *Informational redundancy and resource bounds in dialogue.* Philadelphia, PA: University of Pennsylvania PhD thesis.

Ward, Gregory. 1985. *The semantics and pragmatics of preposing.* Philadelphia, PA: University of Pennsylvania PhD thesis. http://repository.upenn.edu/dissertations/AAI8523465.

Welker, Katherine. 1994. *Plans in the common ground: toward a generative account of conversational implicature.* Columbus, OH: Ohio State University PhD thesis. http://linguistics.osu.edu/research/publications/dissertations/welker1994.

Wittgenstein, Ludwig. 1953. *Philosophical investigations.* New York, NY: Macmillan.

Craige Roberts
Professor of Linguistics
The Ohio State University
210 Oxley Hall
Columbus, OH 43210
croberts@ling.ohio-state.edu