CONTRASTIVE TOPICS AND THE FOCAL STRUCTURE OF QUESTIONS

Nirit Kadmon
Tel Aviv University
kadmon@post.tau.ac.il

1. Introduction

This paper is concerned with prosodically-marked contrastive topics (in English). For instance, *Larry* in (1C). Using the terminology introduced in Kadmon 2001, we may also refer to the contrastive topic as a 'TOPIC-focus' (or TF, or just 'topic', for short). To an ordinary focus,[1] such as *Nina* in (1c), we may refer as a 'FOCUS-focus' (or FF, or just 'focus', for short).

(1) A: Who kissed who?
B: Well, who did Larry kiss?
C: Larry kissed Nina
(L+)H* L H% (L+)H* L L%
[rise] [fall]
TOPIC-focus FOCUS-focus

For the purposes of this paper, let me represent (1C) as (2). | marks the boundary between two Intonation Phrases.

(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

Two (independently developed) theories of the interpretation of prosodically-marked contrastive topics are proposed in Kadmon 2001 (chapter 21) and Büring 2003. Both these theories make use of the 'topic semantic values' proposed in Büring 1999. I think it is also fair to say that these two theories share the same basic insight, viz., that TF+FF configurations serve as **a means of encoding information about the preceding two moves in the discourse**.

Let me briefly illustrate this idea. We can have a coherent piece of discourse composed of (b)+(c).

(b) Who did Larry kiss?
(c) Larry kissed Nina

[1] In the sense of Jackendoff 1972, von Stechow 1981, Rooth 1982, and their many followers.

Suppose now that (c) is pronounced with a single pitch accent, on *Nina*, and an ordinary declarative fall. We assume, as is customary, that in that case *Nina* is the focus (a FOCUS-focus).

(3) Larry kissed $[\text{Nina}]_F$.

I would say that pronouncing (c) in this way is a way of recording (or at least hinting at) the discourse history of (c), because it actually indicates (or at least suggests) that the preceding move in the discourse is (b). Even if you hear (c) pronounced this way out of the blue, you can immediately tell that the question being addressed is (b) (and that it couldn't possibly be, say, *who kissed Nina*…). Indeed, that is the gist of the treatment of (ordinary) focus proposed in Roberts 1996, where (3) is taken to presuppose precisely this – that the last question under discussion in the discourse is (b).

We can also have a coherent piece of discourse composed of (a)+(b)+(c), or of (a')+(b)+(c).

(a) Who kissed who?
(b) Who did Larry kiss?
(c) Larry kissed Nina

(a') For each individual, who did that individual kiss?
(b) Who did Larry kiss?
(c) Larry kissed Nina

Suppose now that (c) is pronounced as in (1C) above, indicating the TF+FF configuration given in (2). I would say that pronouncing (c) in this way is a way of recording the discourse history of (c), simultaneously indicating that the immediately preceding move is (b) **and** that the move before that is (a), or (a').

(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

If you hear (c) pronounced this way, even out of the blue, you can tell that *Nina* is 'the (short) answer to the question', which means that the preceding move in the discourse is (b). You can **also** tell that at an earlier point in the discourse, *Larry* was an element picked out of a set of alternatives: we could have asked who Bill kissed, who John kissed, who Kim kissed, etc. – but we chose, for the moment, to ask about Larry: *who did Larry kiss?*. But that means that earlier in the discourse, we were considering a collection of questions, the one that can be expressed by (a'). Hence, the last move in the discourse can be taken to be (a'), or (a). In short, (2) indicates that the preceding two moves in the discourse were: first (a')/(a) and then (b). And that is the central intuition behind both the theory of Kadmon 2001 and the theory of Büring 2003.

The above two theories of contrastive topics differ, however, on two major issues:

(A) The Treatment of Focus

In Kadmon 2001, TF and FF are treated on a par. Both are taken to be devices that constrain preceding moves in the discourse.[2] In contrast, in Büring 2003, only the contrastive topic (TF) is treated in that manner. Focus (ordinary focus, FF) is taken, instead, to have a rather different function, that of indicating whether a given item is 'given' or 'new' in the discourse.

Büring 2003 does not seem to consider the treatment of ordiary focus central to his immediate concern, that of developing a theory of contrastive topics. He remains non-commital about the precise theory of givenness that should be employed by the proposed treatment of focus, which seems to me to be somewhat problematic. Moreover, in one place (fn.12), Büring says that he needn't commit himself concerning the relation between F-marking and pitch accents, because he doesn't consider examples with broad foci. That I think is even more problematic. What if examples with broad foci actually teach us something significant about focus and about contrastive topics?

I do wish to discuss Büring's proposed treatment of FOCUS-focus. The goal of research on contrastive topics is to make the right predictions about the congruence of TF+FF configurations in different contexts. But it's only the joined forces of our theory of TF and our theory of FF that would enable us to do that! It is therefore problematic to take the attitude that our main concern now is contrastive topics, and hence we need not go too deeply into the theory of ordinary focus.

Moreover, I believe that Büring's theory as it stands yields some incorrect predictions, due to his treatment of focus – and that, even if we confine our attention to examples with narrow foci and narrow contrastive topics. I will discuss these predictions below.

(B) Contrastive Topics and the Focal Structure of Questions

In Kadmon 2001, I observe that there is a systematic relation between contrastive topics and the **focal structure** of the question that immediately precedes them in the discourse. That means two things: (i) that, indeed, it seems that the theory of focus is going to play an important role when we try to construct a theory of topic, as suggested above; (ii) that it is important in the context of treating TF to have an adequate theory of the focal structure of **questions**. The focal structure of questions and its relation to contrastive topics, however, are issues that Büring 2003 does not touch upon.

In the present paper, I would like to present once more the theory of contrastive topics of Kadmon 2001, which supplements Roberts 1996's theory of focus. I will argue that my theory is an elegant theory, and will show how it is better suited to deal with

[2] Following Jackendoff 1972, I take both the focus and the contrastive topic to be focal elements; that is what my 2001 terminology is meant to reflect.

examples that are problematic for the Büring 2003 theory of topics. I would like to focus in particular on my treatment of the focal structure of questions. I will recapitulate my 2001 treatment of the focal structure of questions and its role in making the right predictions about the effect of TF+FF configurations on discourse congruence, and will point out recent evidence that my treatment is independently motivated by other phenomena (besides the effects of TF on discourse congruence).

2. Assumptions that I would Like to Adopt

2.1 Regarding the analysis of FOCUS-focus (ordinary focus), I would like to adopt the four elements summarized in (4) below.

(4) (i) from Rooth 1985, 1992: focus semantic values
(ii) from Hamblin 1973: A question denotes the set of possible answers to it (true and false ones alike).
(iii) from Roberts 1996: the theory of information structure in discourse:
-- A context includes a push-down store of (as yet unanswered) **questions under discussion**.
-- Each move in the discourse must be **relevant** to the last QUD: it must either give a (partial) answer to it, or be a subquestion of it. (A is **subquestion** of B iff the complete answer to A entails a partial answer to A.)
(iv) from Roberts 1966: the QUD constraint on focus

First, I would like to use the standard focus semantic values proposed in Rooth's work. For instance, the focus semantic value of (5), written as $[\![(5)]\!]^f$, is as in (6) – the set of propositions saying about some individual d that Nina introduced d to Sue.

(5) Nina introduced $[\text{Bill}]_F$ to Sue.
(6) $[\![(5)]\!]^f = \{[\![I(n,x,s)]\!]^g : g \text{ assignment}\}$
i.e.,{'Nina intorduced Larry to Sue', 'Nina introduced Bill to Sue', 'Nina introduced Martha to Sue'...}

We may assume that focus semantic values are derived by substitution into the focus position(s), as in the following definition.[3,4]

[3] I assume a logical language where names are treated extensionally (the possible denotations of type e are individuals), but other types are intensional, since the possible denotations of type t are propositions (=sets of possible worlds). See Kadmon 2001, section 14.5.2.
[4] A more formal version of this definition is given in Kadmon 2001, section 14.5.4.

(7) $[\![\varphi]\!]^{f\text{-}M,g}$, the **focus semantic value** of a formula φ relative to a model M and an assignment function g, is the set of all ordinary semantic values that are obtainable by steps (1) and (2).

1. Replace each F-marked subformula in φ with a distinct F-marked variable matching that subformula in type.
2. Interpret the result of (1) relative to some assignment g' which is identical to g except that it may assign different values to the F-marked variables.

Secondly, let us assume Hamblin's theory of questions. A question denotes the set of all possible answers to it. For instance, the denotation of (8) is as in (9). To give a partial answer to a question is to supply the truth value of one of its members; to give a complete answer to a question is to supply the truth values of all of its members.

(8) Who did Nina introduce to Sue?

(9) $[\![(8)]\!] = \{[\![I(n,x,s)]\!]^g : g \text{ assignment}\}$
i.e.,{'Nina intorduced Larry to Sue', 'Nina introduced Bill to Sue', 'Nina introduced Martha to Sue'...}

Thirdly, I would like to adopt Roberts' theory of information structure in discourse. I assume that the context of utterance, besides including a set of propositions called the 'common ground', also includes a **stack of questions**. When a question is (explicitly or implicitly) raised in the discourse, it is called a 'question under discussion' (QUD), and is added to the top of the stack. It stays in the stack as long as it hasn't been answered. In addition, there is a **relevance** condition: each move in the discourse must be **relevant** to the last QUD (to the top-most question in the stack). If it's a statement, it must answer it; if it's a question, it must be a subquestion of it. This condition applies to explicit moves as well as implicit moves.[5]

For instance, look at (1) again. Once (1A) is uttered, it[6] becomes a QUD. (1B) is felicitous as the discourse move following (1A), because it is relevant to (1A), being a subquestion of it. Once (1B) is uttered, it becomes a QUD, and gets added to the top of the question stack. The third move in this discourse, (1C), is relevant to its preceding move (1B), because it answers it. Once answered, (1B) is popped off of the question stack.

The forth thing I would like to adopt is Roberts' constraint on focus, given in (10).

(10) **Roberts' Question-Under-Discussion constraint on FOCUS-focus:**
The focus semantic value of an utterance must be identical to the last QUD obtaining at the time of utterance.

---

[5] For more detail and relevant discussion, see Roberts 1996, Kadmon 2001.

[6] Or, more precisely, $[\![(1A)]\!]$. But I will let myself be a bit sloppy about that.

The semantic/pragmatic theory of focus, as envisaged by Rooth 1992, Roberts 1996, Schwarzschild 1999 and others, consists of two parts: (i) the definition of focus semantic values, and (ii) some linguistic principle which is supposed to capture **the function of focus**, thereby explaining why it is that foci produce their familiar effects (discourse congruence effects, 'association with focus' effects, etc.). My own theory of topic is based on, and supplements, Roberts' particular proposal, viz., that the linguistic priciple in question is the one in (10). In short, I'd like to assume that it is (10) that captures the role of focus – the role of focus as a device for regulating the flow of discourse. I favor this proposal both because it is embedded in an independently motivated theory of information structure in discourse, and because I think it allows for an elegant account of contrastive topics, as I hope to show below.

The resulting theory of focus predicts the usual facts of question-answer congruence. We predict that (8)+(5) is felicitous. The QUD constraint says that the focus semantic value of (5) must be identical to question (8). And so it is. Look at (6) and (9) above – they are exactly the same.

(8) Who did Nina introduce to Sue?
(5) Nina introduced $[\text{Bill}]_F$ to Sue.

We also predict that (11)+(5) is not felicitous: the denotation of (11) is given in (12), and it is not identical to the focus semantic value of (5).

(11) Who introduced Bill to Sue?
(5) Nina introduced $[\text{Bill}]_F$ to Sue.

(12) $[\![(11)]\!] = \{[\![I(x,b,s)]\!]^g : g \text{ assignment}\}$
i.e.,{'Nina introduced Bill to Sue', 'Lyn introduced Bill to Sue', 'Martha introduced Bill to Sue',...}

In short, Roberts' theory determines that (5) **presupposes** that its last QUD is (8).

2.2 Regarding the analysis of the TOPIC-focus + FOCUS-focus configuration, I would like to take over two of the proposals made in Büring 1999. We may add these to the list of assumptions that I would like to adopt, which got started in (4) above:

(4) (cont'd)	(v)	from Büring 1999:	the focus semantic value of examples with TOPIC-focus + FOCUS-focus
	(vi)	from Büring 1999:	topic semantic values

Let me start with item (v). In the literature, we find two conflicting positions regarding the focus semantic value of examples with the TF+FF configuration. According to Roberts 1996, the focus semantic value is derived by substitution into the

TF **and** FF positions. So on her view, the focus semantic value of (2) is as in (13) – which can be described (a bit sloppily) as the set of propositions of the "form" 'x kissed y'.

(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.
(13) $[\![(2)]\!]^f = \{ [\![\text{kissed}(x,y)]\!]^g : g \text{ assignment}\}$ **(Roberts 1996)**
i.e., {'Bill kissed Sue', 'Larry kissed Nina', 'Lisa kissed Jo'...}

According to Büring 1999, the focus semantic value is derived by substitution into the FF position only. So for him, the focus semantic value of (2) is as in (14) - the set of propositions of the "form" 'Larry kissed x'.

(14) $[\![(2)]\!]^f = \{ [\![\text{kissed}(l,y)]\!]^g : g \text{ assignment}\}$ **(Büring 1999)**
i.e.,{'Larry kissed Sue', 'Larry kissed Mary', 'Larry kissed Lisa',...}

I side with Büring 1999 on this matter. Note first that empirically, (15)+(2) and (16)+(2) below are both felicitous question-answer pairs. Now, given the QUD constraint on focus, Roberts' position determines that the last QUD of (2) is invariably question (15), whereas Büring 1999's position determines that the last QUD of (2) is invariably question (16).

(15) Who kissed who?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

(16) Who did Larry kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

As argued in Kadmon 2001, it is much better to take the latter position. Note, first, that intuitively, of the two questions (15) and (16), the one that (2) answers more directly is (16). Secondly, Roberts' theory incorrrectly predicts that (16)+(2) should **not** be felicitous. The denotation of (16) is as in (17). But if the focus semantic value of (2) is (13), we get a QUD constraint violation. (13) is not the same as (17).

(17) $[\![(16)]\!] = \{ [\![\text{kissed}(l,y)]\!]^g : g \text{ assignment}\}$
i.e.,{'Larry kissed Sue', 'Larry kissed Mary', 'Larry kissed Lisa',...

Could Roberts say the following? The last QUD obtaining when (2) is uttered is always (15), even in the case of (16)+(2); (16) is not really the last QUD here, because we have (15) implicitly present in between (16) and (2). Well, no, Roberts could not say that. That's because in that case (15) would be required to be relevant to (16), which it is not, because it is not a subquestion of (15). So, indeed, Roberts incorrectly rules (16)+(2) out. If we adopt Büring 1999's position, on the other hand, we immediately allow (15)+(2), of course, and we can also allow (16)+(2), by saying that in this dialog, (15) is implicitly

present in between (16) and (2) (as already seen above, a discourse consisting of moves (16)+(15)+(2) satsifies the relevance constraint). Thirdly, note also that (18)+(2) below is infelicitous. The problem is that Roberts cannot rule this out.

(18) I know that some boy kissed some girl, but I forget who the boy and girl actually were. Who kissed who?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

Let us now move on to item (4-vi). The semantic/pragmatic theory of contrastive topics, as envisaged by Büring 1999, Kadmon 2001 and Büring 2003, consists of two parts: (i) defining yet another semantic value, called the 'topic semantic value', and (ii) some linguistic principle which is supposed to capture **the function of topic**. Item (4-vi) concerns the first part. It is in Büring 1999 that topic semantic values are proposed, and his 1999 definition of these is adopted later by both Büring 2003 and Kadmon 2001.

According to Büring's proposal, the topic semantic value of (2), for instance, is as in (19).

(19) $[\![(2)]\!]^t$ = {{'Larry kissed Sue', 'Larry kissed Mary', 'Larry kissed Lisa',...},
{'Bill kissed Sue', 'Bill kissed Mary', 'Bill kissed Lisa',...},
{'John kissed Sue', 'John kissed Mary', 'John kissed Lisa',...},...}

Note that (19) is in fact a bunch of focus semantic values – each one of them with a different kisser. We have here the focus semantic value of *Larry kissed* $[\text{Nina}]_F$, then the focus semantic value of *Bill kissed* $[\text{Nina}]_F$, etc. But a bunch of focus semantic values is the same as a bunch of questions: the set in (19) is the same as the set in (20).

(20) $[\![(2)]\!]^t$ = {'Who did Larry kiss?', 'Who did Bill kiss?', 'Who did John kiss?',...}

As suggested in Kadmon 2001, the topic semantic value of (2) can be specified more formally as in (21),

(21) $[\![(2)]\!]^t$ = { { $[\![\text{kiss}(x,y)]\!]^{g'}$: g' is identical to g except that it may assign a different value to y} : g assignment }

and topic semantic values can be defined as follows.

(22) $[\![\varphi]\!]^{t\text{-}M,g}$, the **topic semantic value** of a formula φ relative to a model M and an assignment function g, is the set of all those sets of ordinary semantic values that are obtainable by steps (1), (3) and (4).
1. Replace each F-marked subformula in φ with a distinct F-marked variable matching that subformula in type.

3. In the result of step (1), replace the T-marked subformula with a T-marked variable matching it in type.
4. For each assignment g' which is identical to g except that it may assign a different value to the T-marked variable, form the set of ordinary semantic values obtainable by interpreting the result of step (3) relative to all assignments g" s.t. g" is identical to g' except that it may assign different values to the F-marked variables.

3. Büring 1999's Theory of Contrastive Topics

Büring 1999 proposes a theory of contrastive topics based on adopting all of the assumptions specified in section 2 above. The first part of this theory is the definition of topic semantic values, discussed above. And the second part of this theory is the principle of question-Answer congruence, given in (23). It is this principle which is to capture the role of a contrastive topic – to capture the role of topic as a device for regulating the flow of discourse.

(23) **Büring 1999's Principle of Question-Answer Congruence**
Statement A is a felicitous answer to question Q iff $[\![Q]\!] \in [\![A]\!]^t$.

Büring 1999's theory makes some nice predictions. One: unlike Roberts, Büring does predict that (16)+(2) is felicitous. According to the principle in (23), the denotation of (16) is required to be a member of the topic semantic value of (2). And so it is. Just look at (20) – it includes question (16) as a member.

(16) Who did Larry kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

Two: As Büring 1999 points out, he also predicts, correctly, that (24)+(2) is felicitous.

(24) Who did Bill kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

(If you are not sure that (24)+(2) is felicitous, consider the version in (25), and also Büring's own example, in (26). These sound perfectly natural, and I think that with a little pause, (24)+(2) can sound OK as well.)

(25) A: Who did Bill kiss?
B: I don't know. $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

(26) A: Do you think Fritz would buy this suit?
B: Well, $[\text{I}]_T$ | certainly $[\text{wouldn't}]_F$.

At the same time, as shown in Kadmon 2001, the theory of Büring 1999 also has some problems. One problem is this: Intuitively, there is a difference between (16)+(2) and (24)+(2). (2) seems like a very direct answer to (16). But it is not such a direct answer to (24) – that is why (25) seems more natural. But Büring 1999 can't explain the difference. The two question-answer pairs satisfy his constraint in (23) equally well. Another problem is this: As already noted in section 2, (15)+(2) is felicitous. It is certainly no less felicitous than (24)+(2) is!

(15) Who kissed who?
(2) (Well, let's see...) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

But Büring's principle (23) predicts that (15)+(2) is **not** felicitous, since the denotation of (15) is not a member of the topic semantic value of (2).

4. Truly Direct Answers

I have already mentioned cases where one answer is intutively a more **direct** answer than another. Regarding that, I would like to repeat the following remarks, from Kadmon 2001, chapter 13 (section 13.3).

Consider the well known and widely accepted generalization that the focus in the answer correlates with the questioned position in the question. It is based on contrasts like the following.

(83) A: Who did you introduce to Sue?
B: I introduced Bill to Sue. FELICITOUS
H* L L%

(84) A: Who did you introduce to Sue?
B: I introduced Bill to Sue. INFELICITOUS
H* L L%

Obviously, by 'answer' we mean an utterance that explicitly expresses an answer to the question, excluding other felicitous responses. So the replies in (85) and (86), for instance, are not relevant to the above generalization. ((86B) may well implicate an answer to (86A) or even contextually entail one, but it does not explicitly express one.)

(85) A: Who hit you?
B: I don't know.

(86) A: Was Smith officially invited?
B: Jones was officially invited.

What I would like to do now is bring up a further point, viz., the following. By 'answer' we actually mean what might be called a **truly direct answer**. The intuitions here are more subtle, but let me try to clarify what I mean.

Consider (88)-(91) as replies to (87). All these replies explicitly express an answer to (87). And yet, there is an intuitive difference, I think, between (88) and the rest: (88) is a truly direct answer, in a way that (89)-(91) are not.

(87) Was Smith officially invited?

(88) He was officially invited.
H* L L%

(89) Everybody was officially invited.
H* L L%

(90) He was officially invited three months in advance.
L* H* L L%

(91) He was officially invited...
L*+H L H%

(88) seems to be the only reply that does nothing more and nothing less than directly answer question (87). (89) and (90) clearly provide more information than was requested; I think they each seem to be concerned with a somewhat broader topic/question than the one raised in (87). (91) turns the discussion to the variety of manners of invitation, suggesting that one other than 'officially' is of greater interest.

In sum, the utterance-pairs that display our generalization are pairs of a question and a truly direct answer to it. It is the focus of a truly direct answer – but not of other replies – that must correspond to the questioned position in the question. So, the intuitions about what constitutes a truly direct answer to a given question might be subtle, but they are crucial to the identification of focus and to studying its empirical behavior.

Most of the time, when we identify foci or study them, we don't talk about the notion of a truly direct answer. Why is that? Well, most of the time, we automatically pick responses that are clearly intended as truly direct answers. Also, we often deliberately pick examples where the focus is easy to identify on the basis of prosodic clues (in particular, we often use non-sentence-final narrow foci, as in (83)).

Take for instance the case of (83) and (84) above. It is very clear, intuitively, that (83B) is a truly direct answer to (83A). Also, the prosody strongly suggests that *Bill* is the focus. So we can safely conclude that the focus is on *Bill*, and don't even give it a second thought. We also note that 'the focus in the answer corresponds to the questioned

position in the question', and don't give that a second thought. The prosody of (84B) strongly suggests that *Bill* is not the focus. So we note that 'in the answer to (84A), the focus cannot fail to correspond to the questioned position'. Without thinking about it, we are talking here about a truly direct answer. Note that (84B) expresses the very same proposition as (83B), and is clearly intended as a truly direct answer as well.

Note also that we needn't really worry about whether (84B) is a truly direct answer or not, since (84B) is an entirely infelicitous response, and hence certainly not a possible truly direct answer. Compare with (89), which is a felicitous response to (87): to maintain the generalization that focus must correspond to the questioned position, we must make sure that while (89) does answer (87), it is not a truly direct answer to it.

It is worth keeping in mind that in many examples identifying the focus may be a matter of debate, and it is worth keeping in mind that such a debate may often involve the question of whether a given answer is a truly direct one or not. I think that ignoring speakers' intuitions about what constitutes a truly direct answer can easily lead us astray in our efforts to identify foci and study their behavior. I will have occasion to resort to such intuitions in what follows.

Certainly, there will be cases where speakers can't intuitively tell if a reply is a truly direct answer or not. Just as there are cases where truth value judgments or grammaticality judgments are unclear. As is the usual practice, we will rely on the clear(er) cases as our evidence, and let the resulting theory decide the status of the unclear cases.

5. My Theory of Contrastive Topics and the Focal Structure of Questions

In Kadmon 2001, I propose a new theory of contrastive topics, which, like Büring 1999's theory, is also based on adopting all of the assumptions specified in section 2 above. I will present it here again, with considerable changes in exposition, and only minor changes in the theory itself.

My point of departure is the observation that there is a systematic relation between examples with contrastive topics and the **focal structure** of the question that immediately precedes them in the discourse.

I note that the congruence of responding to a certain uttered question with (2) depends on the focal structure of that question. Look again at our three-utterance conversation:

(15) Who kissed who?
(16) (Well,) who did Larry kiss? <--- *Larry* must be focused!
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

The fact is, that for this conversation to be congruent, (16) must have narrow focus on *Larry*. That is, in the above conversation (16) must actually be (27).

(27) Who did $[\text{Larry}]_F$ kiss?

Similarly, in (26) below, question (16) must also have narrow focus on *Larrry*.

(26) A: I know who Bill kissed, but what about Larry? Who did Larry kiss?
B: $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

Generally, it seems that for a reply to be both congruent and a truly direct answer (in the sense of section 4), the contrastive topic position in the answer must correspond to the focus position in the question, as in (27)+(2).

(27) Who did $[\text{Larry}]_F$ kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. VERY CONGRUENT; DIRECT ANSWER

We clearly have a congruent truly direct answer here. (2) is a very natural response to (27). The intuition is that (2) is a very direct answer to (27), nothing less and nothing more. Compare with the following question-answer pair. (28)+(2) is congruent, but (2) here is not just a 'straight' answer to (28).

(28) $[\text{Who did Larry kiss}]_F$?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. CONGRUENT; IMPLICATION (as for Bill...)

As noted in Roberts 1996 and Büring 1999, 2003, in this sort of example, besides answering the question raised, the response also implies that alternative questions are relevant or might be worth discussing. In (28)+(2), the response tells us who Larry kissed, but at the same time, it also seems to hint that the kissings of some other guy might be of greater interest… (that is, it implies alternative questions such as 'who did Bill kiss?'). Now consider the following two pairs. These are less congruent, aren't they?

(29) Who did Larry $[\text{kiss}]_F$?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. NOT REALLY CONGRUENT; IMPLICATION

(30) Who $[\text{did}]_F$ Larry kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. NOT REALLY CONGRUENT; IMPLICATION

In these pairs, the answer seems not to be quite to the point. It would be more to the point to answer (29) with (31), and to answer (30) with (32).

(31) He $[\text{kissed}]_T$ | $[\text{Nina}]_F$
(32) He $[\text{did}]_T$ kiss | $[\text{Nina}]_F$.

We may note also that if a speaker does choose to respond to (29) or (30) with (2), she seems to imply that she would prefer to discuss the kissings of some other guy, just as in dialog (28)+(2).

To account for the above facts, we must have an adequate theory of focus in questions. Contra Roberts 1966, my position is that the wh-word position is not relevant to the formation of the focus semantic value of a question. I propose the following.

(33) **The focus semantic value of a question** is the set of question-denotations obtainable by substitution into the FOCUS-focus positions.

This can be executed by simply letting our definition of focus semantics values in (7) above apply not only to formulas, but also to questions, as follows.

(7') Let β be a formula or a question.[7]
$[\![\beta]\!]^{f\text{-}M,g}$, the **focus semantic value** of β relative to a model M and an assignment function g, is the set of all ordinary semantic values that are obtainable by steps (1) and (2).
1. Replace each F-marked subformula in β with a distinct F-marked variable matching that subformula in type.
2. Interpret the result of (1) relative to some assignment g' which is identical to g except that it may assign different values to the F-marked variables.

For instance, the focus semantic value of (27) is as in (34).

(27) Who did $[\text{Larry}]_F$ kiss?

(34) a. $[\![(27)]\!]^f$ = { $\{[\![\text{kiss}(x,y)]\!]^{g'}$: g' is identical to g except that it may assign a different value to y} : g assignment }

b. i.e., {{'Larry kissed Sue', 'Larry kissed Mary', 'Larry kissed Lisa',...},
{'Bill kissed Sue', 'Bill kissed Mary', 'Bill kissed Lisa',...},
{'John kissed Sue', 'John kissed Mary', 'John kissed Lisa',...},...}

c. = {'Who did Larry kiss?', 'Who did Bill kiss?', 'Who did John kiss?',...}

As you can see, the focus semantic value of (27) is a set of questions, which differ on the value filled in for the position of the focus.

---

[7] We may assume that the logical translation of questions is as illustrated in (i).
(i) *Who arrived*? translates as ?[arrived(who_1)]
Who invited who? translates as ?[invited(who_1,who_2)]
Is Sue pretty? translates as ?[pretty(Sue)]

I believe that this treatment of foci in questions yields an insightful overall view of focus, as it captures the intuitive use of foci in questions and makes for a uniform conception of the role of focus in declarative and interrogative utterances.

I think that intuitively, the focus on *Larry* in (27) signals that the speaker is choosing the question 'who did Larry kiss?' out of a set of alternative questions. Further, I think that the relevant alternative questions are understood to be precisely the questions of the "form" 'who did x kiss?', viz., the members of $[\![(27)]\!]^f$.

Given my treatment of foci in questions, FOCUS-focus in general can be characterized as 'the last element to be filled in'. In the case of a declarative utterance, the FOCUS-focus, when 'filled in', determines a choice of **proposition** from among all members of the focus semantic value, which are in this case propositions. (These members are identical to the possible answers to the last QUD, and that is why the FOCUS-focus of a declarative can be described as 'the answer to the question being addressed'.) In the case of an interrogative utterance, the FOCUS-focus, when 'filled in', determines a choice of **question** from among all members of the focus semantic value, which are in this case questions.

Let us now move on to the analysis of TOPIC-focus. The first part of my theory is the definition of topic semantic values. As already noted in section 2 above, I am taking over Büring 1999's proposal concerning these, which can be formulated as in (22), repeated below.

(22) $[\![\varphi]\!]^{t\text{-}M,g}$, the **topic semantic value** of a formula φ relative to a model M and an assignment function g, is the set of all those sets of ordinary semantic values that are obtainable by steps (1), (3) and (4).

1. Replace each F-marked subformula in φ with a distinct F-marked variable matching that subformula in type.
3. In the result of step (1), replace the T-marked subformula with a T-marked variable matching it in type.
4. For each assignment g' which is identical to g except that it may assign a different value to the T-marked variable, form the set of ordinary semantic values obtainable by interpreting the result of step (3) relative to all assignments g'' s.t. g'' is identical to g' except that it may assign different values to the F-marked variables.

The second part of the theory should presumably be a principle determining the function of topic. As we saw in section 3, Büring 1999's theory runs into problems. So I think we should not adopt his principle of Question-Answer congruence (given in (23), section 3). Instead, I propose that we supplement Roberts 1996's QUD constraint on focus (in (10), repeated below) with the QUD constraint on topic stated in (35).

(10) **Roberts' Question-Under-Discussion constraint on FOCUS-focus:**
The focus semantic value of an utterance must be identical to the last QUD obtaining at the time of utterance.

(35) **Kadmon's Question-Under-Discussion constraint on TOPIC-focus:**
The topic semantic value of an utterance must be identical to the focus semantic value of the last QUD obtaining at the time of utterance.

The idea is that just as an utterance presupposes that its focus semantic values is identical to (the ordinary denotation of) its last QUD, an utterance also presuppose that its topic semantic value is identical to the focus semantic value of its last QUD.

This idea requires, of course, the assumption that a discourse 'move' doesn't merely consist of a denotation, but also has a focal structure. So let us use the following definition.

(36) A **move** in the discourse is an ordered pair whose first member is an ordinary semantic value (a proposition or a question) and whose second member is a focus semantic value (a set of propositions or a set of questions). The move corresponding to a logical expression β is $\langle [\![\beta]\!], [\![\beta]\!]^f \rangle$.

Now let us see what my theory predicts regarding our example (2). Recall that given the assumptions of section 2, we have the values specified in (37) and (38).

(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

(37) $[\![(2)]\!]^f = \{[\![\text{kissed}(l,y)]\!]^g : g \text{ assignment}\}$
i.e.,{'Larry kissed Sue', 'Larry kissed Mary', 'Larry kissed Lisa',...

(38) $[\![(2)]\!]^t = \{ \{[\![\text{kiss}(x,y)]\!]^{g'} : g' \text{ is identical to g except that it may assign a different value to y}\} : g \text{ assignment} \}$
i.e., {{'Larry kissed Sue', 'Larry kissed Mary', 'Larry kissed Lisa',...},
{'Bill kissed Sue', 'Bill kissed Mary', 'Bill kissed Lisa',...},
{'John kissed Sue', 'John kissed Mary', 'John kissed Lisa',...},...}
= {'Who did Larry kiss?', 'Who did Bill kiss?', 'Who did John kiss?',...}

Given these values, and given the two QUD constraints, I predict that the last QUD when (2) is uttered must invariably be (27). (Or, in other words, that (2) presupposes that its last QUD is (27).)

(27) Who did $[\text{Larry}]_F$ kiss?

Here is why. The QUD constraint on FOCUS-focus requires the denotation of the last QUD when (2) is uttered to be identical to the focus semantic value of (2), in (37). But according to Hamblin's theory, the set in (37) is none other than the denotation of (27). The QUD constraint on TOPIC-focus requires the focus semantic value of the last QUD of (2) to be identical to the topic semantic value of (2), in (38). But the set in (38) is precisely $[\![(27)]\!]^f$, which was given in (34) above. So, the last QUD when (2) is uttered must be (27) – both denotation-wise and focus-wise. I think that this prediction is a good basis for explaining the various empirical facts.

We said about dialog (27)+(2), that the intuition is that here the reply just directly answers the question, no more and no less. This is immediately explained, because according to my theory, in this dialog, (2) actually follows the very question that it addresses.

(27) Who did $[\text{Larry}]_F$ kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. DIRECT ANSWER; NO IMPLICATION

Of course we know that (2) is perfectly felicitous in other contexts too, for instance, following question (28). How come?

(28) $[\text{Who did Larry kiss}]_F$?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. CONGRUENT; IMPLICATION (as for Bill...)

My position is that (27) is present here implicitly, in between (28) and (2). For that to be possible, (27) must be relevant to (28), to satisfy Roberts' relevance contsraint. And indeed it is, since (27) is certainly a subquestion of (28). So the dialog is congruent.

The nice thing is that my theory also explains the **contrast** between the last two dialogs. As discussed above, in (28)+(2), the speaker of (2) implies that she wishes to discuss another matter (for instance, 'Who did Bill kiss?'), whereas in (27)+(2), there isn't a similar implication. Büring 1999 can't explain the contrast. On my theory, it can be explained as follows. In (27)+(2), (2) is a direct answer to (27), and there's nothing more to it. And in (28)+(2)? Here the speaker of (2) has chosen not to directly answer (28), but rather to – implicitly – directly answer (27). The hearer must wonder why; after all, this speaker could have directly answered (28), by uttering (39).

(39) Larry kissed $[\text{Nina}]_F$.

Why then did she choose to address (27) rather than (28)? Well, as far as the ordinary denotations go, (27) and (28) are the same. So the decision to address (27) must have to do with the focal structure of (27). So the hearer can conclude that the speaker wanted to draw attention to members of the focus semantic value of (27). There was no need to draw attention to 'Who did Larry kiss?', which has just been explicitly mentioned. So she

must have meant to draw attention to other members of $[\![(25)]\!]^f$, such as 'Who did Bill kiss?'. And that is the implication.

Let us now consider the relation between (2) and (15).

(15) Who kissed who?

As discussed in Roberts 1996, **(2) can form part of a strategy aimed at answering (15)**. This is intuitively the case in the following two pieces of discourse, (15)+(27)+(2), and (15)+(2).

(15) Who kissed who?
(27) Who did $[\text{Larry}]_F$ kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. CONGRUENT

(15) Who kissed who?
(2) (Well, let's see...) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. CONGRUENT

Note, in this connection, that Büring 1999's theory does not link (2) to (15) at all; recall, in particular, that it fails to predict that (15)+(2) is felicitous. Given my theory, we can easily account for the relation between (2) and (15). We need not officially define a 'strategy of inquiry' (cf. discussion in Roberts 1966 and Kadmon 2001). We can simply show that pieces of discourse such as (15)+(2) and (15)+(27)+(2) are felicitous, because they are each the explicit manifestation of a sequence of (explicit and implicit) discourse 'moves' which satisfies the two QUD constraints as well as the relevance constraint.

Take (15)+(27)+(2) first. What is the last QUD of (2)? It is (27), of course, which means that (2) meets both QUD constraints. And how about (15)+(2)? This pair is not as direct as (27)+(2), but it **is** perfectly felicitous, even without the part in parentheses. How come? Of course, my position is that (27) is implicit here too, in between (15) and (2). (27) **can** be implicit here, because it is relevant to (15): technically, it's a subquestion of (15); intuitively, it is clear that (27) is relevant in that it can serve as part of a strategy aimed at answering (15).

In general, (2) will be felicitous only in contexts where to implicitly introduce question (27) is a relevant and cooperative thing to do. In a context where (27) is irrelevant, we predict that (2) will be infelicitous. Take for example (40)+(2). It's infelicitous.

(40) Who Kissed Nina?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. INFELICITOUS

This is correctly predicted: (2) implicitly introduces (27); but (27) is not relevant to (40), so it shouldn't be implicit here. Technically, (27) cannot be implicit in between (40) and (2), because it is not a subquestion of (40).

Next, consider the following dialog.

(24) Who did Bill kiss?
(2) (I don't know.) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. CHANGE OF TOPIC

As noted in section 3 above, this dialog is not as congruent as some of the other ones. Without the sentence in parentheses intervening, (24)+(2) doesn't sound quite natural. With or without the part in partheses, the intuition is that the speaker of (2) is changing the subject. (That is precisely the reason why the response sounds more natrual with the part in parentheses.) My theory predict this intuition. (2) is implicitly introducing (27), but (27) is not relevant to (24) – technically, it's not a subquestion of (24); intuitively, it wouldn't be of any help in trying to answer (24). Of course there $[\textbf{is}]_F$ $[\textbf{some}]_T$ connection between (24) and (27): they are both members of $[\![(27)]\!]^f$. And it is this connection that makes (2) a possible response to (24) – it's a change of topic, but **within reason**.

And what about the following pair, (18)+(2)? Why is it not felicitous?

(18) I know that some boy kissed some girl, but I forget who the boy and girl actually were. Who kissed who?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. INFELICTIOUS

As usual, (2) addresses question (27). Now, (27) is in fact technically relevant to the explicit question just asked, *who kissed who*. However, implicitly introducing (27) in this context is not a cooperative thing to do. Given the context in (18), it is clear that (2) actually provides the complete answer to the question just asked, *who kissed how* – it gives the one pair of boy and girl. But there is no point in implicitly introducing (27), when you're actually already providing the complete answer to *who kissed who*. Why introduce the extra question? It would be more to the point to say (41), which directly addresses *who kissed who*.

(41) $[\text{Larry}]_F$ kissed $[\text{Nina}]_F$.

So, it is not enough for (27) to be technically relevant in its context. There must also be some good reason to implicitly introduce it, to implicitly introduce an extra question. If we can show that there isn't a good reason to do that, we predict incongruence.

Finally, consider the following dialogs, (29)+(2) and (30)+(2). As noted earlier, these pairs don't feel smooth. The intuition is that the answer seems not to be quite to the

point. It would be more to the point to answer (29) with (31), and to answer (30) with (32).

(29) Who did Larry [kiss]$_{FOCUS}$?
(2) [Larry]$_T$ | kissed [Nina]$_F$. NOT REALLY CONGRUENT; IMPLICATION

(30) Who [did]$_{FOCUS}$ Larry kiss?
(2) [Larry]$_T$ | kissed [Nina]$_F$. NOT REALLY CONGRUENT; IMPLICATION

(31) He [kissed]$_T$ | [Nina]$_F$
(32) He [did]$_T$ kiss | [Nina]$_F$.

Well, of course that would be more the point. On my theory, (31) directly addresses (29), and (32) directly addresses (30). (2), in contrast, does not address the explicit question just asked; it directly addresses a different question, namely, (27).

So far so good. But why can't (27) be implicit in between (29) and (2) or (30) and (2)? Well, first of all, that is not absolutely prohibited. I think it is possible, empirically, to respond to (29)/(30) with (2). It is not really congruent, but not quite terrible either. As noted above, if a speaker does choose to respond to (29) with (2), she implies that she would prefer to discuss the kissings of some other guy, just as in dialog (28)+(2). The explanation being that the only possible purpose of saying (2) and thereby implicitly introducing an extra question, namely (27) (the question that (2) directly addresses), would be to draw attention to other members of the focus semantic value of that question.

Secondly, there is the narrow focus in (29)/(30). Take (29) for example. The narrow focus on *kiss* means that the focus semantic value of (29) is the set of questions of the "form" 'Who did Larry R?'. To ignore this and utter (2), suddenly introducing the set of questions of the "form" 'Who did x kiss?', is a bit much; it makes the resultant change of topic not quite 'within reason'. Compare that with the situation in the case of (24)+(2).

(24) Who did Bill kiss?
(2) (I don't know.) [Larry]$_T$ | kissed [Nina]$_F$. CHANGE OF TOPIC

Since there is nothing in (24) to suggest otherwise, it is easy to imagine one of two thing: either (i) that (24) has focus on the entire question, and no questions besides (24) are currently under discussion in the discourse, or (ii) that the collection of questions of the "form" 'Who did x kiss?' is already relevant in the discourse. In either case, drawing attention to questions of the "form" 'Who did x kiss?' via the focus semantic value of the implicit (27) is not unreasonable. The speaker may fail to answer (24), but at least she is not ignoring the collection of questions which has been suggested by the focus semantic value of that question.

6. Questions and Collections of Questions

6.1 Questions and Collections of Questions Roberts' QUD constraint on FOCUS-focus and my analysis of the focal structure of questions put together have one consequence that may seem a little surprising. What is the last QUD obtaining in the discourse just before (27)? The focus semantic value of (27) is (34); therefore, the (ordinary denotation of) the last QUD of (27) must also be (34). But, what kind of question is (34)? (34) is not a set of propositions; it's a set of questions.

(27) Who did $[\text{Larry}]_F$ kiss?

(34) a. $[\![(27)]\!]^f = \{ \{[\![\text{kiss}(x,y)]\!]^{g'} : g' \text{ is identical to g except that it may assign a different value to y}\} : g \text{ assignment} \}$

b. i.e., {{'Larry kissed Sue', 'Larry kissed Mary', 'Larry kissed Lisa',...},
{'Bill kissed Sue', 'Bill kissed Mary', 'Bill kissed Lisa',...},
{'John kissed Sue', 'John kissed Mary', 'John kissed Lisa',...},...}

c. = {'Who did Larry kiss?', 'Who did Bill kiss?', 'Who did John kiss?',...}

Well, I think that (34) is actually a very useful question-denotation to have. Consider questions (44) and (45). What do such questions denote? Surely, they denote precisely a set of questions like the one in (34). (What else would they denote?...)

(44) Who did each individual kiss?
(45) For each individual, who did that individual kiss?

Note also that there is an empirical difference between the above two questions and question (15).

(15) Who kissed who?

Given (15), it is legitimate to list in reply kisser-kissee pairs arranged in any order – you can use the order of kissings in time, the order in which they come to mind, or whatever; but given (44)/(45), the kisser-kissee pairs must be ordered by the kissers: all the kissings of kisser 1, all the kissings of kisser 2, all the kissings of kisser 3, and so on. This contrast is explained, if we assume that (15) has the ordinary question-denotation in (46), whereas (44) and (45) denote a collection of questions, as in (34).

(46) $[\![(15)]\!] = \{[\![\text{kissed}(x,y)]\!]^g : g \text{ assignment}\}$
i.e., {'Bill kissed Sue', 'Larry kissed Nina', 'Lisa kissed Jo'...}

Returning now to question (27) and its last QUD, consider again our 3-utterance conversation:

(15) Who kissed who?
(27) Who did [Larry]$_F$ kiss?
(2) [Larry]$_T$ | kissed [Nina]$_F$.

Intuitively, what is the purpose of following (15) with question (27)? Obviously, (27) is part of a strategy for answering (15). But more specifically, it amounts to **a decision to list the kissings by the kissers**. The fact that (27) implies a decision to go by the kissers is represented nicely, if we recognize that the last QUD of (27) is indeed (34). Our theory determines that in this conversation, (34) must be implicitly present in between (15) and (27). But (34) must be answered by the kissers, of course, because we have to answer each question in the collection of question separately. Finally, note that (34) **can** be implicit in between (15) and (27), because (34) is relevant to (15). Indeed, it is very relevant to (15), because a complete answer to (27) provides not just a partial answer to (15), but a complete answer to (15).

My theory determines, of course, that question (34) must be explicitly or implicitly present in every discourse containing (27), as well as in every discourse containing (2). After all, (2) presupposes that its last QUD is (27), and (27) presupposes that its last QUD is (34). For instance, we get that the two-utterance discourse (15)+(2) is in fact the explicit manifestation of the following series of moves (implicit moves italicized).

(15) [Who kissed who]$_F$? (denoting question (46))
(45) For each individual, who did that individual kiss? (denoting question (34))
(27) Who did [Larry]$_F$ kiss?
(2) [Larry]$_T$ | kissed [Nina]$_F$.

Which I think is just as it should be. I think this is a good representation of what actually happens when we respond to (15) with (2). Intuitively, when we respond to (15) with (2), we simultaneously decide (i) to seek the answer to (15) by considering each kisser in turn and finding out who that kisser kissed, and (ii) to start with the kisser Larry.

6.2 Additional support for allowing a move in the discourse to be a collection of questions comes from Sevi 2005's treatment of the projection of scalar implicatures.

Consider example (67). We consider only the reading(s) where *every boy* has wide scope relative to *three girls*. In many contexts, sentences (67) is understood as conveying (68).

(67) Every boy kissed [three girls]$_F$ equivalent to: $\forall$ boy $\exists$3 girls that he kissed (i.e., at least 3)

(68) Every boy kissed exactly 3 girls

How can this fact be predicted? It is rather widely accepted that as far as the semantics goes, (67) is truth conditionally equivalent to 'Every boy kissed at least three girls'. It is customary to derive the 'exactly' reading of *three girls* via a scalar implicature. For instance, *Larry kissed three girls* is taken to have, in many contexts, the scalar implicature that Larry kissed no more than three girls, yielding the interpretation (I mean, content+implicature) that larry kissed exactly 3 girls. But how are we to derive the appropriate scalar implicature for (67)? Assuming Horn 1972's definition of a scalar implicature, we have to use the scale <…5 girls, 4 girls, 3 girls, 2 girls…>, and we get the implicature that of all the statements that can be formed by substituting for *three girls* in (67) one of the items on the scale, all the statements stronger than (67) are false. That is, we get the implicature in (69).

(69) It is not the case that every boy kissed 4 girls

But that is the wrong implicature; (69) is too weak – it is compatible with some boys kissing more than 3 girls. The scalar implicature which we really **want** to derive is (70).

(70) For every boy, he kissed no more than 3 girls and no non-girls

The problem is that we tried to derive the scalar implicature 'globally', with negation taking scope over the entire sentence. What examples like (67) show is that in some cases the attested scalar implicature is 'embedded', and we must compute it 'locally'. How do we compute an implicature for (69) locally, then?

Let's see. It is well known (cf. Rooth 1992) that scalar implicatures depend on focal structure. In fact, as is revealed in Sevi 2005, they depend on the particular quesiton being addressed. Empirically, (67) will not have implicature (70) as a response to just any question. It may well have it, though, when it is a response to (71).[8]

(71) For every boy, who did that boy kiss?
i.e., For every boy x, what (sg. or plural) individual y is s.t. x kissed y?

Of course, (71) denotes not a single question, but a collection of questions! For each boy $b \in D$, we have to answer the question 'who did b kiss?'. In the context of (71), we have to interpret the content of (67) as being roughly as in (72). That is, (67) actually expresses a collection of all answers of the form 'b kissed three girls', for all boys $b \in D$.

(72) For every boy x, an answer is: x kissed 3 girls.

[8] It seems clear to me that (71) is often expressed explicitly by uttering (*).
(*) Who did every boy kiss?
Of course I use (71) in the text because (*) is ambiguous while (71) is not.

Evidently, what we need to do in order to get the implicature in (70), is to compute a scalar implicature for each of these answers separately. Fomrally, we need to compute the implicature 'locally', based not on the entire sentence but rather on *x kissed 3 girls*.

I think the above is already enough to suggest rather strongly that we must recognize that a move in the discourse can be a collection of questions, providing independent support for my position on the matter.

And yet let me strengthen my position even more, by referring to Sevi's comprehensive and fully worked-out theory of the so-called 'projection' of implicatures, and showing how the correct implicature of (67) can be derived, making reference to the collection of questions in (71).

Sevi 2005 makes a very strong case against a 'Hornian' theory of scalar and clausal implicatures, and offers a new theory of these implicatures, which derives them based on a semantic operation of **exhaustivization** (*exh*), and which is by far more complete and more satisfying than any of the previously existing theories. The *exh* operation is sensitive to the question being addressed (it is defined in such a way that the question addressed enters into calculating the result of an *exh* application), thereby capturing the way in which scalar implicatures depend on focal structure. Sevi assumes that *exh* is available in the grammar, and can optionally be applied to any clause which addresses some QUD in the discourse. Let us take a simple example:

(59) A: Who smiled?
B: $[NP]_F$ smiled.
where the focused NP denotes a generalized quantifier, call it GQ.

We asssume that smile is a set of individuals. Let Q also be a set of individuals. What an application of *exh* adds to the semantic content of (59B) can be stated roughly as in (60), or, more precisely, as in (61).

(60) $\forall Q [[Q \subseteq \text{smile} \wedge GQ(Q)] \rightarrow \text{pred.cont.}(Q)]$
'Every set of individuals which consists of smilers and is a member of the generalized quantifier satisfies the predicative content of the focused NP.'

(61) $\forall Q [[Q \subseteq \text{smile} \wedge GQ(Q)] \rightarrow GQ^2 (\sqcup Q)]$
where GQ^2 is a predicate derived (usually by the *BE* operation) from the predicative content of the focused NP, and $\sqcup Q$ is the maximal element of Q (i.e., the sum of Q).

A hearer would usually choose the exhaustive interpretation of (59B) (i.e., would interpret it with an implicit *exh*), based on Gricean considerations of cooperation, which

direct speakers to **answer a question by giving a complete answer**. (Applying *exh* will often turn the answer into a complete answer to the question addressed.)

To see how Sevi's account of scalar implicatures works, let us take a couple of examples of what the focused NP (=the answer to the question) in (59B) might be. First, suppose that the answer is (62).

(62) $[\text{John and Mary}]_F$ smiled.

What an application of *exh* adds to the content of (62) is as in (63),

(63) Every set Q of smilers that **contains** J and M, its sum **is** 'John and Mary'
(i.e. Q={John, Mary})

which predicts the correct scalar implicature for (62), viz., the implicature that nobody else (besides John and Mary) smiled. Note that the exhaustive interpretation of (62) is a complete answer to the question addressed (to (59A)). Now suppose, intstead, that the answer is (63).

(63) $[\text{John or Mary}]_F$ smiled.

What an application of *exh* adds to the content of (63) is as in (64),

(64) For every set Q of smilers that contains J or contains M,
$\lambda x[x=\text{John} \vee x=\text{Mary}](\sqcup Q)$
i.e., Every set Q of smilers that contains J or contains M,
its sum is 'John' or its sum is 'Mary'
(i.e. Q={John} or Q={Mary})

which predicts the correct scalar implicature for (63), viz., the implicature that only John smiled or only Mary smiled. Note that the exhaustive interpretation of (63) yields the interpretation of *or* as exclusive 'or'. Note also that it yields a stronger proposition, getting us closer to a complete answer to the quesiton.

Now let us move on to the implicatures of complex sentence. I think it is fair to say that Sevi 2005 offers the first viable theory of so-called 'implicature projection'. At any rate, Sevi's theory is more complete and better motivated than the few previously existing accounts. Sevi argues convincingly against theories where scalar implicatures are generated 'locally' and then 'inherited up'. His general principle of implicature 'projection' is simply this:

(65) All scalar implicatures are generated directly, by an *exh* application which yields an exhaustive interpretation of some clause.

While Sevi does not really state this, I think the major insight which enables him to predict the range of facts regarding complex sentences may be stated as follows.

(66) An *exh* application may be found wherever there is a clause/subformula which is taken to answer a quesion under discussion (QUD) in the discourse. In many cases, it is a whole complex sentence that's taken to answer a (complex) QUD, so we will find an exhaustive interpretation of the matrix clause, and hence a 'global' scalar implicature. In other cases, an embedded clause/subformula is taken to answer its own QUD, so we get an exhaustive interpretation of that clause/subformula, and hence a 'local' scalar implicature corresponding to that clause/subformula.

Back to example (67), Sevi derives the scalar implicature in (70) by applying *exh* 'locally', to the subformula *x kissed three girls*. It is possible to apply *exh* in this way, if the last QUD of (68) happens to be something like the collection of questions given in (71), repeated below. Given (71), *x kissed three girls* is taken to be answering its own 'local' QUD (viz., 'who did x kiss?'), so *exh* can apply to it.

(71) For every boy x, what (sg. or plural) individual y is s.t. x kissed y?

What applying *exh* to *x kissed 3 girls* adds to the content of *x kissed 3 girls* is roughly as in (75).

(75) Every set Q of individuals who were kissed by x that **contains** three girls, its sum **is** 'three girls'

Note that to **be** 'three girls' is to be a set of **exactly** three girls. Hence, (75) yields the desired scalar implicature, the one in (70) above: every boy kissed no more than 3 girls and no non-girls.

7. On the Contrastive Topic as part of a Uniform System of Focal Elements

I would like to argue that in addition to the empirical advantages pointed out above, my theory of TOPIC-focus and the focal structure of questions offers a coherent and uniform view of all focal elements – including the focus of a statement, the focus of a question, and the contrastive topic.

My treatment of foci in questions has the conceptual advantage that, as already noted in section 5, it allows for a general characterization of FOCUS-focus as 'the last element to be filled in'. To take some examples: When you say *[Nina]$_F$ ate it*, it's like you already had *x ate it*, and you are now filling in the value for x.

[Nina]$_F$ ate it. x ate it x=Nina

This value, when 'filled in', determines a choice from among the members of the focus semantic value of your utterance. (These members are identical to the possible answers to the last QUD, and that is why the FOCUS-focus of a declarative can be described as 'the (short) answer to the question being addressed'.) Same thing in our example (2): when you say *[Larry]$_T$ | kissed [Nina]$_F$*, it's like you already had *Larry kissed x* (which corresponds to the question 'Who did Larry kiss?'), and you are now filling in the value for x.

[Larry]$_T$ | kissed [Nina]$_F$ Larry kissed x x=Nina

And it's the same thing with a question: When you say *Who did [Larry]$_F$ kiss?*, it's like you already had *Who did x kiss?*, and you are now filling in the value for x – you have decided to ask about *Larry*.

Who did [Larry]$_F$ kiss? Who did x kiss? x=Larry

In the case of an interrogative utterance, the FOCUS-focus, when 'filled in', determines a choice of **question** from among all members of the focus semantic value, which are in this case questions.

In addition, my theory of contrastive topics yields a coherent conception of FOCUS-focus and TOPIC-focus as two closely-related components of the same system of focal structure. TOPIC-focus is in fact very much like a FOCUS-focus, except that it operates on a 'deeper' level of a two-layered focal structure. I will try to explain what I mean by that.

The simplest thing to note is that just as a FOCUS-focus is 'the last element to be filled in', a TOPIC-focus is 'the penultimate element to be filled in', as is already clear in Jackendoff 1972. When you say example (2), *[Larry]$_T$ | kissed [Nina]$_F$*, that suggests that at some point you had just the skeleton *y kissed x*. Then first you decided that you are interested in who Larry kissed, so you first filled in the value for y. And after that you filled in the girl that was kissed, i.e., the value for x.

[Larry]$_T$ | kissed [Nina]$_F$ y kissed x y=Larry
x=Nina

But let us note also that according to my theory, there is a level at which a TOPIC-focus too is 'the **last** element to be filled in': a TOPIC-focus is the **last** element to be filled in **in the last QUD**! (Indeed, I would say that the TOPIC-focus is the 'penultimate element' in the current utterance precisely because it represents the 'last element' in the preceding move in the same discourse.) I said that when you say (2), *[Larry]$_T$ | kissed [Nina]$_F$*, it's like you aleardy had *Larry kissed x*, and you are now filling in the value for x.

[Larry]$_T$ | kissed [Nina]$_F$ Larry kissed x x=Nina

Now according to my theory, it is also the case that this utterance, (2), presupposes that its last QUD is not simply *Who did Larry kiss?* (question (16)), but in fact that question with a specific focal structure, viz., *Who did [Larry]$_F$ kiss?* (question (27)). But when you utter the latter question, we said, it's like you already had *Who did x kiss?*, and you are now filling in the value for x.

Who did [Larry]$_F$ kiss? Who did x kiss? x=Larry

So, in the last QUD of (2), that is, in question (27), the TOPIC-focus of (2), *Larry*, was the last value to be filled in.

In short, TOPIC-focus is just like FOCUS-focus – the last thing to be filled in – but in the move preceding the current utterance. I think this is an elegant result.

8. Büring 2003's theory of focus and contrastive topics

Büring 2003 offers his own theory of contrastive topics. Like my theory, this theory sees contrastive topics as a device for encoding information about the two preceding moves in the discourse. I think it is a nice theory; it does not suffer from the problems that afflict Büring 1999 (see section 3), it explicates the role of contrastive topics in a simple and intuitive manner, and it gets many of the predictions discussed in section 6. Nonetheless, I would argue that my theory is preferable, for reasons to be discussed below.

8.1 The treatment of (Jackendoffian) 'focus', and pitch accent placement Unlike Büring 1999, Büring 2003 does not adopt any of the assumptions about FOCUS-focus specified in section 2 above. Instead, he adopts the view that FOCUS-focus should be treated in terms of 'givenness'. He suggests that we should follow Schwarzschild 1999 in assuming that (47) holds (p.527).

(47) **Givenness**: Every constituent which is not Given needs to be marked.

'Given' in (47) is to be understood more or less as 'previously mentioned in the discourse', corresponding to the 'c-construable' of Rochemont 1986.[9] Büring does not offer a precise definition of 'givenness', and refers his readers to Schwarzschild 1999 for such a definition. Büring suggests that 'marked' in (47) is to be understood as 'F- or T-marked'.[10]

[9] And, I might add, to the 'anaphoric' of Kuno 1972, 'given' of Clark and Haviland 1977, and 'familiar' of Heim 1982; that is, to the given$_k$ – 'shared knowledge' – of Prince 1981.

[10] He notes (in fn.12) that (47) does not require every non-Given constituent to carry a pitch accent, since a broad FOCUS-focus or TOPIC-focus is non-Given (hence the F- or T-marking), but may contain words that do not carry a pitch accent.

Büring stipulates in addition that an item can be rendered Given only by something in an **explicit** prior move. This last assumption is crucial, he says, for predicting that T-marking is obligatory when the subquestion being answered is implicit. For instance, for predicting that while (15)+(27)+(39) is felicitous, (15)+(39) is not.

(15) Who kissed who?
(27) Who did $[\text{Larry}]_F$ kiss?
(39) Larry kissed $[\text{Nina}]_F$. CONGRUENT

(15) Who kissed who?
(39) Larry kissed $[\text{Nina}]_F$. INCONGRUENT

In my opinion, it is a mistake to abandon the treatment of FOCUS-focus in terms of Rooth's focus semantics values and Roberts' QUD constraint on focus and to replace it with a treatment of FOCUS-focus in terms of givenness. And that is despite the fact that I firmly believe that what **carrying a pitch accent** marks is in fact some sort of 'givenness'.[11]

I should like to note, first of all, that Schwarzschild 1999 does not purport to be supplying a theory of 'focus'; he is explicitly concerned with the interpretation of pitch accent placement, and is in fact explicitly providing a theory of it that does not involve the Jackendoffian notion of 'focus' at all. (Schwarzschild does mark constituents with F, but that F does not mean 'focus'. His F is a feature which is related in a systematic way to carrying a pitch accent and to not occuring in preceding discourse, but **not** to being a Jackendoffian 'focus'.)[12]

Büring 2003 actually says very little about his reasons for abandoning the Rooth-Roberts treatment of FOCUS-focus. One reason is presumably the wish to explain the contrast between (15)+(27)+(39) and (15)+(39), as discussed above. 'Givenness' is something that can be generalized to apply to FOCUS-focus and TOPIC-focus alike, thereby explaining that contrast.[13]

---

[11] See discussion below. For more detail, see Kadmon 2001, chapter 13, and Kadmon (ms.).

[12] Compare Schwarzschild 1999's system with that of Selkirk 1996: both allow structures with small F-marked constituents embedded within bigger F-marked constituents. However, while for Selkirk 1996 the top-most F-marked constituent received its own special interpretation as a 'focus', for Schwarzschild 1999, the top-most F-marked constituent is not interpreted any differently from the embedded F-marked constituents.

[13] Another reason is stated in his fn.13 (p.528), and involves examples like (*)+(2).
(*) Who did the guys kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.
On both my theory and Büring 2003's theory (see below), the last QUD of (2) is the iplicit question (27) (*Who did* $[\text{Larry}]_F$ *kiss?*). Büring notes that if we adopt Roberts 1996's treatment of the focus semantic value of questions (see beginning of section 6), the last QUD of the iplicit (27) would have to be *Who kissesd who?*. But this is a problem, since in fact, it can also be question (*). Well, I certainly agree that

Can my theory, which does adopt the Rooth-Roberts treatment of FOCUS-focus, explain this contrast too? Well, no, I think that the theory of section 5, on its own, can't explain it. The focus semantic value of (39) is as in (48).

(48) $[\![(39)]\!]^f = \{[\![\text{kissed}(l,y)]\!]^g : g \text{ assignment}\}$
i.e.,{'Larry kissed Sue', 'Larry kissed Mary', 'Larry kissed Lisa',...}

Therefore, the QUD constraint on focus determines that the last QUD of (39) is (16).

(16) Who did Larry kiss?

Since there is no TF in (39), we have no information about the focus semantic value of its last QUD (if you like, we can assume that by default it is the set of all question-denotations). All we know is that it's (16). Therefore, the last QUD of (16), in turn, is not subject to any QUD constraint; it only has to meet the relevance requirement. Since (16) is relevant to the explicit (15), my theory allows (15) to be the last QUD of (16), and does not rule out the dialog (15)+(39).

As it happens, I am actually in full agreement with Büring 2003 regarding the idea that the incongruence of (15)+(39) is due to the fact that *Larry* in (39) is non-Given (or, at any rate, 'new' in whatever the right sense is). (I also agree with him that *Larry* in (39) is not rendered Given by the fact that it occurs in the last QUD of (39), as this last QUD, i.e., (27), is only implicit.) However, I do not think that the non-Givenness of *Larry* directly forces it to be a TF or FF.

I think that the non-Givenness of *Larry* directly forces it to carry a pitch accent. As noted above, I also believe that pitch accent placement should be treated in terms of some sort of 'givenness'. But I do not believe that the placement of pitch accents directly marks what the focused constituent (in Jackendoff's sense) is – that is, I do not believe that it directly marks 'the answer to the question'. I believe that, as suggested in Kamdon 2001, the placement of pitch accents helps us identify the focused constituent(s) indirectly, via the fact that carrying a pitch accent marks an item[14] as 'new', and without there being any grammatical principle which directly relates an item's carrying a pitch accent to its being (part of) the focus.

To just finish my story about (15)+(39): Since *Larry* in (39) is accented, it is 'new'. (I believe there is a 1-1 correspondence between carrying a pitch accent and being

---

Roberts' treatment gets the wrong predictions, but then, as shown above, we can adopt all the assumptions of section 2 without adopting her treatment of the focus semantic value of questions. My own theory allows (*)+(2), because it determines that the last QUD of (27) is *For each individual, who did that individual kiss?*, and given the right domain selection for 'individual', (*) may be taken to express just that question.

[14] Usually a word, in my opinion.

'new' – see Kadmon (ms.).) Hence, *Larry* can't be part of a previously established question under discussion in the discourse (that is because if an item is part of what is previously known to be the question being addressed, then that item is definitely 'given' and not 'new'). So either it is a focus (i.e., the answer to the question addressed), or a contrastive topic (i.e., a piece of news about what the question being addressed actually is). And, indeed, it can be either one: both (15)+(39) and (15)+(41) are felicitous.[15]

(15) Who kissed who?
(41) $[\text{Larry}]_F$ kissed $[\text{Nina}]_F$. CONGRUENT

Returning to the discussion of pitch accents, 'givenness' and focus, it seems to me that Büring 2003 is conflating two separate things: (i) our theory of the interpretation of pitch accent placement, and (ii) our theory of the notion 'focus', in the sense of Jackendoff 1972.

To illustrate some of the things I've just said, I'd like to consider the following example, adapted from Selkirk 1984. All three versions of (50) can felicitously answer the question in (49). Now what is the 'focus' of (50), in Jackendoff's sense? Well, it is undoubtedly the VP. No other constituent, and in particular no constituent properly embedded within the VP, is a 'focus' in this sense. That is the fairly standard view, taken by Jackendoff, Rooth, Roberts, Selkirk 1996, and others. Those who believe in a feature F which 'projects upward' from single words to larger and larger constituents, take the largest ('top-most') constituent to be the 'focus'.

(49) What did she do on Monday?

(50-a) She sent her sketches to the publisher.
(H*) H* H* L L%

(50-b) She sent her sketches to the publisher.
(H*) H* L L%

(50-c) She sent her sketches to the publisher.
(H*) H* L L%

[15] Büring seems to think that examples like (15)+(41) are not congruent – his example is this:
A: What did the pop stars wear?
B: The $[\text{female}]_F$ pop stars wore $[\text{caftans}]_F$
But I think that both his and my examples are congruent. First, note that (15)+(41) is clearly just fine when it is assumed that there is just one kisser-kissee pair. Secondly, even when we assume many kisser–kissee pairs or pop-star-subgroup–item-of-clothing pairs, there is nothing wrong with having two equal-status foci in the answer, per se. With regards to both examples, if we imagine that B ends in a declarative fall, then yes, we judge it bad – but that's presumably becauses B fails to give a complete answer while sounding so final. If we give a complete answer by listing all the pairs (in random order), most pairs should end with a continuation rise, and only the last pair (the one completing the list) should end with a fall.

Now, the different versions of (50) will make appropriate answers to (49) in different contexts. Selkirk 1984 notes that while (50-a) is a natural 'out of the blue' response to (49), (50-c) can answer (49) only in certain contexts, e.g., if Jane's job is illustrating books, and we've been talking about the sketches that she has recently prepared. Similarly, (50-b) can answer (49) only in the right context – say, when it is known that when one sends out one's sketches, it is to the publisher that one sends them.

In terms of the notion of 'focus', what we see here is that whether or not an example allows for a broad focus may depend on what's given vs. new or expected vs. unexpected in the context (Schmerling 1976, Ladd 1980, Selkirk 1984). What we can learn from this type of example is that the presence or absence of pitch accents within a broad focus gets systematically interpreted – it is used as a device for indicating the informational status of parts of the focus. As argued in Selkirk 1984, that in turn shows something about the connection between prosody and focus: we see that pitch accent placement plays a role that is distinct from just indicating what the focus (='the (short) answer to the question') is, and that our theory should explain how, via that role, pitch accent placement succeeds in constraining what the focus of the utterance might be – i.e., in constraining the questions that the utterance can felicitously answer. In Kadmon 2001, chapter 13, I sketch just such a theory, which does not include any rule of the grammar relating accent to 'focus'.

I hope that it is now clear why I think that if Büring wishes to treat focusing as indicating that the focused constituent is non-Given (new_k), that is incompatible with the standard Jackendoffian sense of 'focus'. At the same time, Büring's intention seems to be just that – treating Jackendoffian focus in terms of 'givenness'. If so, I think he is wrong. And I think it is a mistake to focus exclusively on single-word narrow foci and ignore broader foci, when we try to decide what our theory of focus is going to be. The examples with broad foci show that various subparts of the focus can be marked as 'new' or 'given' by the absence or presence of pitch accents. So if we claim that the 'focus' is just interpreted as 'new' in the same sense, and nothing else, then we do not have a theory of what 'focus' (in Jackendoff's sense) is; there is nothing in our theory to distinguish 'focus' from items that are 'new' but are not foci. I think we should hold on to the theory that focus induces focus semantic values, which allows it to play its special discourse regulating role (possibly via a principle like Roberts' QUD constraint, as I would like to assume in this paper).

Another option, which is in my opinion worth exploring, is that we don't need the Jackendoffian notion of 'focus' in our theory of grammar at all. As shown by Schwarzschild, we can explain a wide array of facts about the effect of pitch accent placement on discourse congruence without using any notion of 'focus' in our theory. Büring 2003 thinks so too. And so do I. In fact, in Kadmon (ms.), I develop my own theory of the interpretation of pitch accent placment, which differs from

Schwarzschild's theory in many respects,[16] but just like Schwarzschild's theory, does not use any notion of 'focus'. Perhaps, then, we don't need the Jackendoffian notion of 'focus' at all? Well, Büring does not explicitly propose such a thing. But perhaps we could interpret what he says as suggesting just that. We don't need focus semantic values, we don't need focus.

Surprisingly perhaps, I would be sympathetic to this line of thought. In fact, Aldo Sevi and myself are currently exploring together the possibility that we don't need Jackendoffian 'focus' as part of the grammar, or if we do need it, that perhaps we don't need to have both focus semantic values and a principle determining the role of focus in our grammar. However, what we cannot do is dismiss Jackedoffian 'focus' off-hand. It is a whole research project to first of all find out if there are examples where the role of pa placement will not be enough to get the congruence predictions, and secondly to study all the different well known effects of focusing ('association with focus' etc.), and try to see whether they can be adequately explained without recourse to a notion of 'focus' as part of our grammar.

Back to more concrete empirical facts, I would like to point out that Büring 2003's treatment of focus yields incorrect predictions. In the principle of givenness that Büring wishes to adopt, that is, (47'), there is an inherent asymmetry;

(47') **Givenness**: Every constituent which is not Given needs to be T- or F-marked.

we have that non-Given ==> F-marked, but F-marked =/=> non-Given. (This asymmetry is taken over from Schwarzschild's work.) But that is problematic, because it means that the resulting theory fails to rule out infelicitous two-utterance dialogs such as the following.

(40) Who Kissed Nina?
(51) Larry kissed Nina
H* (L) H* L L%

(40) Who Kissed Nina?
(52) Did Larry kiss Nina?
H* H* L L%
or H* H* H H%

The obvious hypothesis is that these dialogs are ruled out by the fact that *Nina* in (51) and (52) is accented despite the fact that it is 'given'. However, given the

[16] Among other things, my theory employs a completely different sort of 'givenness', which I call RECOVERABILity. While Selkirk 1984, 1996 and Schwarzschild 1999's theories use notions akin to givenness_K, I argue in detail that much better suited for the purpose of interpreting the presence or absence of a pitch accent is a notion of 'givenness' akin to the idea of 'predictable' of Kuno 1972 and 'recoverable' of Halliday 1967 (the given_p of Prince 1981).

asymmetry just noted, Büring's theory of focus is unable to use this fact to rule out the dialogs. Büring is non-commital about the relation between accenting and F-marking. Well, if the pitch accent on *Nina* in (51) and (52) does not even force any F-marking, then there certainly isn't anything in the theory to rule out the above dialogs. So let us assume, as does Schwarzschild, that an accented word must be F-marked. Actually, let us be more Büringian, and interpret this as meaning that an accented word must be F- or T-marked. Does this help us rule out the above dialogs? No, it does not, because even if *Nina* is F- or T-marked, that does not require it to be 'new'. There is nothing in the theory to disallow (40)+(51') or (40)+(52').[17]

(40) Who Kissed Nina?
(51') [Larry]$_F$ kissed [Nina]$_F$.

(40) Who Kissed Nina?
(52') Did [Larry]$_F$ kiss [Nina]$_F$?

8.2 The treatment of contrastive topics Büring 2003 offers a new theory of contrastive topics. Like Büring 1999 and Kadmon 2001, Büring 2003 treats TOPIC-focus as a discourse regulating device. According to the Büring 2003 theory, the raison d'être of a contrastive topic is to indicate that a certain answering-strategy is being used in the discourse.

Following Roberts 1996, Büring 2003 assumes a representation of the context of utterance which includes discourse moves of two kinds, viz., propositions and QUDs, and also assumes a relevance requirement. In addition, Büring proposes that the context includes a hierarchical structure of discourse moves, which is represented by what Büring calls **d(iscourse)-trees**. Given a node in the tree which is a question, its daugher nodes must be relevant to it, i.e., either answers to it, or subquestions of it. The grammar generates the set of admissible d-trees, and a sequence of utterances forms a felicitous piece of discourse if these utterances can map onto the moves in some d-tree, preserving linear order. Crucially, some of the moves in the tree may be only implicit, i.e., not corresponding to any utterance.

On to the theory of topic: the first part of the theory is the familiar topic semantic values, taken over from Büring 1999. (We may continues to assume the definition of topic semantic values given in section 2 above.) The second part of the theory is a new contsraint meant to capture the role of contrastive topic in discourse, viz., the following constraint.

[17] Wait, and what if a pitch accent doesn't require a word to be F-marked, but only to be contained within an F-Marked constituent? Well, there is nothing in the theory to disallow (40)+(51") or (40)+(52"), either.
(51") [Larry kissed Nina]$_F$.
(52") Did [Larry kiss Nina]$_F$?

(52) **Büring 2003's Principle of Contrastive Topic Congruence**

An utterance U containing a contrastive topic can map onto a move M_U within a d-tree D only if there is a non-singleton set of questions **Q** s.t. for each Q∈**Q**, (i) Q is identical to or a sister of the question that immediately dominates M_U; and (ii) $[\![Q]\!] \in [\![U]\!]^t$.

I think it is fair to say that what this amounts to is the constraint in (53).

(53) The last QUD of the current utterance plus at least one sister of it (under a superquestion 'mother') must be members of the current utterance's $[\![\ ]\!]^t$.

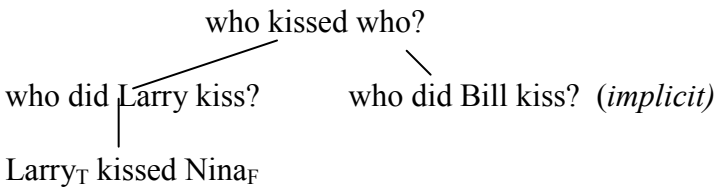
As long as a given explicit piece of discourse can be interpreted as being part of a tree that satisfies the above constraint – possibly by imagining that certain implicit questions are present – this piece of discourse will be judged felicitous.

Let us go through some predictions. Let us start with our three-utterance conversation:

(15) Who kissed who?
(27) Who did $[\text{Larry}]_F$ kiss?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. CONGRUENT

(15)+(27)+(2) can be mapped onto the d-tree in (54).

(54)



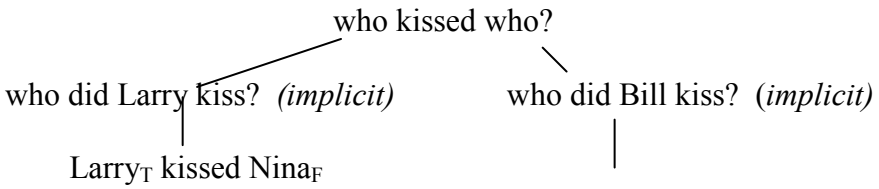
This satisfies Contrastive Topic Congruence: The 'mother' (superquestion, last QUD) of (2) is simply the preceding question in the discourse, (27). (27) is a member of $[\![(2)]\!]^t$, so that much is fine. Now (27) must have a sister in the tree, also a member of $[\![(2)]\!]^t$. Well, (27) has a a 'mother' in our discourse, *who kissed who*, and, intuitively, it is obviously part of a strategy aimed at answering its mother – a strategy that can only be completed by asking (and answering) not only (27) istelf, but further subquestions of *who kissed who*. Therefore, the hearer of (15)+(27)+(2) can tell that such a subquestion – a sister of (27) – is already on our mind (or at least will be so in a minute), allowing the hearer to assume that such a subquestion is implicitly present. But the sister is also required to be a member of $[\![(2)]\!]^t$, so it must be of the "form" 'who did x kiss?'. Hence we may assume that a question of that "form" (e.g. 'who did Bill kiss') is a sister of (27). So, our three-utterance dialog is correctly predicted to be congruent, and it is also correctly predicted that we intuitively expect the next question addressed to be of the form 'who did x kiss?'.

And how about the following question-answer pair?

(15) Who kissed who?
(2) (Well, let's see...) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. CONGRUENT

(15)+(2) can also be mapped onto the same d-tree as in (54), the only difference being that now the last QUD of (2) is implicit, as noted in (55):

(55)



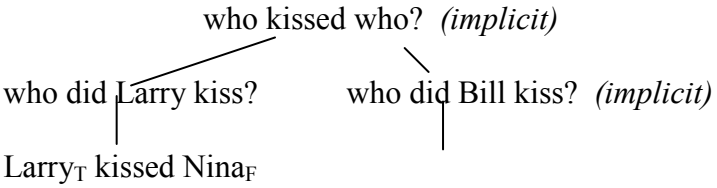
When we hear (15)+(2), we must find a question which is (i) answered by (2), (ii) a member of $[\![(2)]\!]^t$, and (iii) a subquesiton of (15) – and that leads us to posit *who did Larry kiss* as an implicit move.

And this pair?

(16) $[\text{Who did Larry kiss}]_F$?
(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. CONGRUENT; IMPLICATION (as for Bill...)

(15)+(2) can also be mapped onto the same d-tree as in (54), the situation differing from the previous examples only in which moves are explicit and which are implicit:

(56)



The last QUD of (2) is simply the explicit (16). (16) is a member or $[\![(2)]\!]^t$, so that much is fine. There must also be a sister of (2)'s in the tree. So when we hear (16)+(2) we must imagine a mother question, which could have further daughters that are also member of $[\![(2)]\!]^t$. We therefore imagine that implicit is *who kissed who*, and that that's the superquestion of (16) in our discourse. And we know that there must be at least one other daughter of *who kissed who* (which is also a member of $[\![(2)]\!]^t$) in our discourse. If we already know that such a question, e.g., *who did Bill kiss*, is in our discourse – fine. If not, we must conclude that there is in fact such a question currently "under discussion" (i.e. at issue, waiting to be answered) in our discourse, or at least that there is immediately going to be such a question under discussion in our

discourse. That explains why, empirically, saying (2) in this context implies that there is another question, also a member of $[\![(2)]\!]^t$, which is of interest.

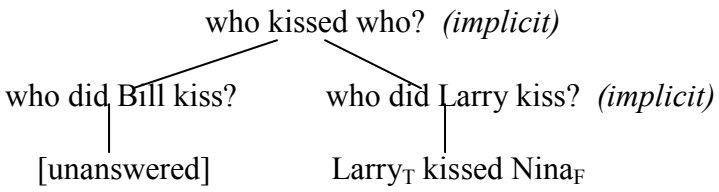
As you can see, both Büring 2003 and myself force the 'mother' (superquestion, last QUD) of sentence (2) to be a question of the "form" 'who did x kiss?' – Büring by stipulating that its mother must be a member of $[\![(2)]\!]^t$, and me by stipulating that $[\![(2)]\!]^t$ must be identical to the $[\![\ ]\!]^f$ of its mother. As you can also see, we both force (2) to have a 'grandmother' in the discourse. Büring does it sort of indirectly and without specifying what the grandmother is, by requiring the last QUD of (2) to have at least one 'sister' which is also a member of $[\![(2)]\!]^t$, which means in turn that it must also have an appropriate mother – a mother which can have other daughters that are members of $[\![(2)]\!]^t$. I do it through the combination of (i) the requirment that (2)'s mother must itself have a focus semantic value identical to $[\![(2)]\!]^t$, and (ii) the QUD constraint on focus, which directly requires that the grandmother question (=the last QUD of (2)'s last QUD) be identical to the focus semantic value of (2)'s mother (=the last QUD of (2)), and hence that it be the collection of questions of the "form" 'who did x kiss?'.

And how about this case?

(20) Who did Bill kiss?
(12) (I don't know.) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$. FELICITOUS; CHANGE OF TOPIC

A hearer can tell that (20) can't be the 'mother' (last QUD) of (12), because that would violate the relevance requirement. (20) can only be the 'aunt' of (12) inside a d-tree, as in (57).

(57)



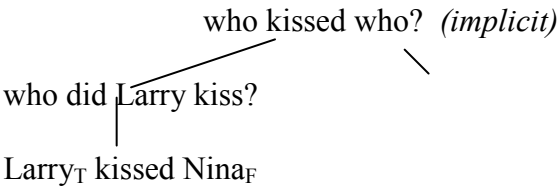
The hearer may need some help in figuring out that (20) is not being answered, because a cooperative interlocutor will usually aim to answer. That is why (20)+(12) is more natural with the part in parentheses. But certainly the speaker of (12) can have her reasons not to answer it. So presumably (57) is a possible d-tree to imagine. And this d-tree satisfies Contrastive Topic Congruence, so (20)+(12) is judged OK, albeit a change of topic.

Finally, take the following pair.

(18) I know that some boy kissed some girl, but I forget who the boy and girl actually were. Who kissed who?
(2) [Larry]$_T$ | kissed [Nina]$_F$. INFELICTIOUS

Büring 2003 predicts the infelicitly via the requirement that the last QUD of (2) must have at least one sister: The hearer can form the tree in (58),

(58)



but the context in (18) makes it clear that there isn't any other daughter of *who kissed who* that could be of interest in our discourse.

So far so good. As I said in section 1, both Büring 2003 and myself see the TF+FF configuration as a means of encoding information about the preceding two moves in the discourse. And, not surprisingly, we make some similar predictions. However, Büring 2003 does not note the connection between an example with the TF+FF configuration and the focal structure of its last QUD or refer to it in any way, and for that reason his theory fails to explain certain empirical facts that my theory does explain.

In section 5, I discussed the following two question-answer pairs, and noted the contrast between them. The theory of Büring 2003 cannot explain this contrast.

(16) [Who did Larry kiss]$_F$?
(2) [Larry]$_T$ | kissed [Nina]$_F$. CONGRUENT; IMPLICATION (as for Bill...)

(27) Who did [Larry]$_F$ kiss?
(2) [Larry]$_T$ | kissed [Nina]$_F$. DIRECT ANSWER; NO IMPLICATION

More specifically, Büring 2003 cannot explain the following intuitions:

(i) that (2) is not a truly direct answer to (16);
(ii) that (2) is a truly direct answer to (27);
(iii) that (2) as a response to (27) does not have the implication that the kissings of some other guy are (or might be) of (possibly greater) interest.

Büring 2003 has no means of explaining the difference in how direct an answer (2) is felt to be in response to (16) and in response to (27). Contrastive Topic Congruence requires that the last QUD of (2) be of the "form" who did x kiss', and combined with the relevance constraint, i.e., assuming that it must answer its last QUD, we get that the last QUD of (2) must in fact be *who did Larry kiss*. But there are no

requirements at all concerning the focal structure of this last QUD, or anything else that will make (2) less of a direct answer to (16) than it is to (27).

As shown above, Büring 2003 predicts that when (2) is uttered as a response to the question *who did Larry kiss*, the hearer is forced to assume that another question of the "form" 'who did x kiss?' is (or is immediately going to be) under discussion in the discourse. And that is quite independent of what the focal structure of *who did Larry kiss* might be. Therefore, Büring 2003 has no means of explaining why the empirical fact is that responding to (27) with (2) does not force the hearer to make that assumption.

8.3 The theory as a whole, as compared with my theory In the previous subsections, I pointed out several disadvantages of the overall theory of Büring 2003.

I have argued that Büring's treatment of FOCUS-focus in terms of 'givenness' is unsatisfactory for two reasons: (i) it is actually inconsistent with the Jackendoffian notion of 'focus' which it apparently intends to capture (while not explicitly making or trying to support the claim that that notion of 'focus' can be dispensed with); and (ii) because Büring doesn't require F-marked constituents to be non-Given, he incorrectly predicts that an items can carry a pitch accent without being 'new' – for instance, he fails to rule out the accent on *Nina* in the second utterance of the following pairs.[18]

(40) Who Kissed Nina?
(51) Larry kissed Nina
H* (L) H* L L%

(40) Who Kissed Nina?
(52) Did Larry kiss Nina?
H* H* L L%
or H* H* H H%

In addition, I argued that Büring's 2003 treatment of TOPIC-focus has the following empirical disadvantage: because he fails to recognize the connection between an example with the TF+FF configuration and the focal structure of its last QUD, he fails to predict that (2) is more of a direct answer to (27) than it is to (16), and that as a response to (27) it does not imply that the kissings of some other guy might be of (possibly greater) interest.

(2) $[\text{Larry}]_T$ | kissed $[\text{Nina}]_F$.

(27) Who did $[\text{Larry}]_F$ kiss?

[18] As a matter of fact, this *Nina* is neither new_k nor new_p, and the theory should make that fact rule out its being accented.

(16) [Who did Larry kiss]$_F$?

I think that the above considerations taken together point to the conclusion that the overall success of Büring 2003's theory, though considerable, has significant limitations, and that it is not as empirically successful as my own theory.

In section 7, I argued that my theory has a conceptual advantage, viz., that of allowing for a coherent and uniform conception of focal elements, including FOCUS-focus in declaratives, FOCUS-focus in questions, and TOPIC-focus. I said that I considered that an elegant result. Büring 2003, on the other hand, treats FOCUS-foci and TOPIC-foci quite differently from each other, attributing to them discourse-roles of two different sorts. I leave it to the reader to form their own judgment on the import (if any) of this difference regarding the relative elegance of the two theories.

Aknowledgements

My 2001 theory of the TF+FF configuration was presented at SALT 10, Cornell, 2001 (but for health and family reasons did not make it into the proceedings), and at a number of graduate seminars at Tel Aviv University. I thank the audiences for their comments. I am grateful to Craige Roberts for detailed and helpful comments and discussion concerning the same material (as well as for all our discussions and joint explorations of this and related phenomena over many years). I am grateful to Aldo sevi for our discussion of Büring 2003, and for further input concerning the present paper.

References

Büring, Daniel 1999: 'Topic'. In P. Bosch and R. van der Sandt (eds.), *Focus – Linguistic, Cognitive, and Computational Perspectives*, Cambridge University Press, 142-165.

Büring, Daniel 2003: 'On D-Trees, Beans, and B-Accents'. *Linguistics & Philosophy* 26:5. 511-545.

Clark, H. and S. Haviland 1977: 'Comprehension and the given-new contrast'. In R. Freedle (ed.), *Discourse Production and Comprehension*, Ablex, 1-40.

Halliday, M.A.K. 1967: 'Notes on transitivity and theme in English, part 2'. *Journal of Linguistics* 3.2, 199-244.

Hamblin, C.L. 1973: 'Questions in Montague Grammar'. *Foundations of Language*, 41-53. Reprinted in B. Partee (ed.), *Montague Grammar*, Academic Press, 1976.

Heim, Irene 1982: *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. dissertation, University of Massachusetts, Amherst. Reproduced by the GLSA, Linguistics Department, UMass, Amherst.

Horn, Laurence 1972: *On The Semantic Properties of Logical Operators in English*, Ph.D. dissertation, UCLA. Reproduced by the Indiana University Linguistics Club, 1976.

Jackendoff, Ray S. 1972: *Semantic Interpretation in Generative Grammar*, MIT Press.

Kadmon, Nirit 2001: *Formal Pragmatics: Semantics, Pragmatics, Presupposition, and Focus*, Blackwell

Kadmon, Nirit (ms.) 'Some Theories of the Interpretation of Accent Placement'. A 2000 ms. was submitted to NALS and refereed, but for health and family reasons was not immediately revised for publication. It was then given at CSSP 2003 (invited talk). The final version is now ready (summer 2009), and being sent to NALS.

Kuno, Susumu 1972: 'Functional sentence perspective: a case study from Japanese and English'. *Linguistic Inquiry* 3, 269-320.

Ladd, D. Robert 1980: *The Structure of Intonational Meaning: Evidence from English*, Indiana University Press.

Prince, Ellen 1981: 'Towards a taxonomy of given-new information'. In P. Cole (ed.), *Radical Pragmatics*, Academic Press, 223-256.

Roberts, Craige 1996: 'Information structure in discourse: towards an integrated formal theory of pragmatics'. In J. Hak Yoon and A. Kathol (eds.), *Ohio State Univsersity Working Papers in Linguistics, volume 49*.

Rochemont, Michael S. 1986: *Focus in Generative Grammar*, John Benjamins.

Rooth, Mats 1985: *Association with Focus*, Ph.D. dissertation, University of Massachusetts, Amherst. Reproduced by the GLSA, Linguistics Department, UMass, Amherst.

Rooth, Mats 1992: 'A theory of focus interpretation'. *Natural Language Semantics* 1.1, 75-116.

Schmerling, Susan F. 1976: *Aspects of English Sentence Stress*, University of Texas Press.

Schwarzschild, Roger 1999: 'GIVENness, AvoidF and other Constraints on the Placement of Accent', Natural Language Semantics 7.2, 141-177

Selkirk, Elisabeth O. 1984: *Phonology and Syntax: The Relation between Sound and Structure*, MIT Press.

Selkirk, Elisabeth O. 1996: 'Sentence prosody: intonation, stress, and phrasing'. In J. Goldsmith (ed.), *The Handbook of Phonological Theory*, Blackwell.

von Stechow, Arnim 1981: 'Topic, focus and local relevance'. In W. Klein and W. Levelt (eds.), *Crossing the Boundaries in Linguistics*, Reidel.

Sevi, Aldo 2005: *Exhaustivity: A Semantic Account of 'Quantity' Implicatures*, PhD dissertation, Tel Aviv University. Reproduced by VDM Verlag, May 2009