

## Some Theories of the Interpretation of Accent Placement

Nirit Kadmon, Tel Aviv University

September 2009

### Contents:

1. Introduction
2. Question-answer pairs and pitch accent placement
3. Structural principles governing pitch accent placement?
  - 3.1 A hard look at some of the basic data
  - 3.2 Syntactic distinctions do not play a decisive role here
4. Selkirk's theory (without reference to argument structure):  
A recursively projected feature E; the maximal E-marked constituent interpreted as focus; absence/presence of E interpreted as 'given'/'new'
5. Schwarzschild's theory:  
A freely but sparingly distributed feature R; no focus interpretation; absence of R interpreted as GIVEN, with definition of GIVEN itself making reference to R
6. A Selkirk-like theory with given:  
A recursively projected feature E; the maximal E-marked constituent interpreted as focus; absence/presence of E interpreted as given/new, but with given applied only to focus-internal constituents
7. Another look at the 'givenness' associated with accent placement
8. A new theory with RECOVERABLE:  
No recursively projected or interpreted syntactic feature; no focus interpretation; an unaccented/accented word interpreted as RECOVERABLE/not RECOVERABLE
9. In the realm of RECOVERABILITY
  - 9.1 The Flexibility of RECOVERABILITY:  
RECOVERABILITY comes in degrees and gets manipulated by the speaker
  - 9.2 The determinants of RECOVERABILITY
  - 9.3 The 'her mother' example
10. Beyond the realm of RECOVERABILITY
  - 10.1 'Highlighting'
  - 10.2 Arbitrary patterns in pitch accent placement
11. Conclusion

### 1. Introduction

This paper is concerned with the way in which discourse congruence depends upon the location of pitch accents. For the sake of brevity, I will concentrate on the congruence of question-answer pairs, although every claim or proposal below either has been or can (I think) be applied to other sorts of pieces of discourse as well.

I will first make some claims and suggestions about the data, and then argue that the interpretation of pitch accent placement (or a corresponding

syntactic feature) is not sensitive to syntactic distinctions like 'heads and arguments vs. modifiers and adjuncts', or 'head vs. argument'. After that, I will move on to a number of alternative theories of the interpretation of pitch accent placement and its effect on discourse congruence.

Going through a variety of different theories in search of the best one, I will, in effect, be investigating the following questions. One, what is the nature of the semantic/pragmatic interpretation of accent? It is fairly widely agreed – roughly – that presence vs. absence of a pitch accent (or a corresponding syntactic feature) means 'new' vs. 'given'. But what sense of 'given'/'new', exactly? And to what extent is the interpretation consistent, uniform, straightforward? Two, is there a need to mark various syntactic phrases with a syntactic feature corresponding to accent, or is it good enough to interpret pitch accent placement just at the level of the word? Three, should the theory employ a notion of 'focus', as a distinct theoretical entity?

Schwarzschild 1999 offers a novel and interesting theory of the effect of pitch accent placement on discourse congruence, which differs quite radically from the earlier theories of Selkirk 1984, 1996. In particular, it formulates a precise, recursively defined notion of 'given', and it involves no reference to 'focus'. I will show some problems with Schwarzschild's theory. I will then propose an amended version of the theory of Selkirk 1996 combined with a simpler notion of 'given' (offered but dismissed by Schwarzschild), and argue that this more conservative theory is at least as successful as Schwarzschild's theory is. Finally, I will propose yet another theory of the effect of pitch accent placement on discourse congruence. This theory, like Schwarzschild's, will make no reference to 'focus'. It will employ yet another notion of 'given', which harks back to Kuno's 'predictable' and Halliday's 'recoverable'. It will differ from both Selkirk's and Schwarzschild's theories in that pitch accent placement will be interpreted at the level of the word alone, leaving no use for either recursive projection or recursive interpretation of a syntactic feature corresponding to accent. I will argue that my new theory is more successful than the previous two.

## **2. Question-answer pairs and pitch accent placement**

Given an intonation-less (printed on paper, say) question-answer pair, we would like to be able to predict which distribution of pitch accents in the answer will make for a felicitous, congruent answer, and – to the extent that that depends on further context – in what sort of context.

We want to be able to predict, for example, that the pairs on the left are felicitous, while the ones on the right are not:

(3) What did you do to Bill?  
 (5-a) I invited Bill  
           H\* L L%

(3) What did you do to Bill?  
 (5-b) I invited Bill  
           H\* L L%

(4) Who did you invite?  
 (5-b) I invited Bill  
           H\* L L%

(4) Who did you invite?  
 (5-a) I invited Bill  
           H\* L L%

In the above dialogs, we see that that word in the answer which corresponds to the wh-word in the question must be accented. The following dialog – being infelicitous – suggests, in addition, that the other words in the answer must be **unaccented**.

(3) What did you do to Bill?  
 (5-c) I invited Bill  
           H\* H\* L L%

Such (and further) data have led, of course, to the postulation of the theoretical notion 'focus', as conceived in Halliday 1967, Chomsky 1971, and Jackendoff 1972. The **focused constituent**, or simply **focus**, is that constituent which constitutes 'the (short) answer to the question'. It is customary to assume that the focus carries a syntactic feature F which gets interpreted both by the semantics/pragmatics and by the phonology.

Consider now the following example, adapted from Selkirk 1984. All three versions of (2) can answer the question in (1).

(1) What did she do on Monday?

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

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

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

But the different versions of (2) will make appropriate answers to (1) in different contexts. Selkirk 1984 notes that while (2-a) is a natural 'out of the blue' response to (1), (2-c) can answer (1) only when reference to certain sketches is expected – e.g., in a context where Jane's job is illustrating books, and we've been talking about the sketches that she has recently prepared. Similarly, (2-b) can answer (1) only in the right context – say, when it is known or natural 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.

Of course, it could still be the case that pitch accent placement, besides being linked to the informational status of words, is also to some extent determined by structural principles.

### **3. Structural principles governing pitch accent placement?**

#### **3.1 A hard look at some of the basic data**

Consider the following examples. Much of the literature would suggest that questions (1) and (2) can be answered freely by both (5-c) and (5-b), and cannot be answered by (5-a) at all.

- (1) What happened?
- (2) What did you do?

(5-c) I invited Bill  
H\* H\* L L%

(5-b) I invited Bill  
H\* L L%

(5-a) I invited Bill  
H\* L L%

Based on this sort of data, it was thought at first that there was some structural constraint which determined – roughly – that the right-most (content) word within a focus must be accented (Halliday 1967, Jackendoff 1972). But, as argued extensively in the literature, this cannot be right (Bolinger 1972, Schmerling 1976, Ladd 1980, Selkirk 1984, Gussenhoven 1984). The right-most (content) word in

a focus is not always accented – counterexamples abound. (For instance: (2-b) above; *My nose is itching*; *There were crawling things all around* (Bolinger1972).)  
H\* L L% H\* L L%

Well then, perhaps there is a more sophisticated structural principle involved? I think that one has to be very careful drawing conclusions about that from the traditional judgments on (1)/(2)+(5).

First, consider replying to (1) or (2) with (5-a). Do such exchanges show, perhaps, that it is impossible, within a focus, to accent a transitive verb without accenting its object? No, they do not. Further inspection of the data clearly shows that that is not the case. For example, the following dialog, from Ladd 1980, is OK.

- (16) A: Has John read Slaughterhouse-Five?  
B: No, John doesn't read books.  
H\* L L%

In fact, (5-a) itself is a perfectly good response in the right context. I think it cannot be denied that dialog (7), for instance, is felicitous.

- (7) A: Hey, Bill is back in town. But if he shows up at the party he will start quarrelling with Sally and it will all be a mess.  
B: Don't worry, I have thought about that already.  
A: What did you do?  
B: I invited Bill. That way he won't feel unwanted, and will be  
H\* L L%  
less quarrelsome. And I'm having Sally's brother stick to her the whole evening and smooth things out... [and so on]

Next, consider replying to (1) or (2) with (5-b). The traditional judgment is unequivocal: this response is felicitous. Based on such judgments, it is assumed in much of the literature that in an all-new transitive VP, the main verb need not be accented. Which suggests some structural principle determining that lack of accent on the transitive verb does not get interpreted. However, I think that just as in the case of replying with (5-a), further exploration of the data must lead us to revise our assumptions about the facts. I submit that replying to (1) or (2) with (5-b) is in fact felicitous only in contexts in which it is rather natural or expected that one would talk about inviting people. For instance, if the speakers are planning a party. In contexts where it is not at all natural or expected that the answer should involve inviting someone, (5-b) is not a good response. Surely, replying with (5-c) would be a much more natural response in such a context. (Out of the blue, the judgment that (5-b) is not a good response may be clearer in the case of answering (1).)

If I am right about the judgments, then it seems quite clear that the absence of pitch accent on *invited* is not due to structural principles, but is rather something that gets interpreted (meaning 'expected', perhaps). When the interpretation it receives is uncalled for, the pitch-accent-impoverished example seems infelicitous.

But **am** I right about the judgments, you may ask. What about those traditional judgments? I think that the case for the judgments as I have presented them is quite compelling. I think there are reasons why the traditional judgments that (5-b) can freely answer (2) whereas (5-a) cannot answer it at all have been commonly assumed.

First let us look at (5-a). As I said, I think it cannot be denied that dialog (7) above is felicitous. Many informants have confirmed this judgment for me. And why is it traditionally assumed that (5-a) cannot answer (2)? Well, precisely because we need a relatively complex context to see that it can.

Here are some more details about that: Perceptually, the nuclear pitch accent in a sequence of H\* tones cannot be mistaken, so we can tell for sure that the nuclear accent of (5-a) is on *invited*. Therefore, we can also tell for sure that *Bill* carries no pitch accent at all. So we must unambiguously interpret *Bill* as expected in the context. But names of salient referents are usually replaced with pronouns. It is hard to reconcile the idea that reference to Bill is expected in the context with the fact that the speaker has not chosen to use a pronoun instead, because one does not easily imagine a relatively complex context such as the one in (7), where reference to Bill is indeed expected and yet can naturally be made using *Bill* rather than *him*. So we judge the sequence (2)+(5-a) to be infelicitous.

Now consider (5-b). I am claiming that (5-b) cannot freely answer (1)/(2). I am claiming that (5-b) can only answer (1)/(2) in contexts where inviting people is natural or expected. This judgment goes against what many of us have been used to assume, but it too has been confirmed for me by quite a few informants.

I believe that in general, it is not as easy to have large foci with very few pitch accents in them as is sometimes assumed in the literature. I am in total agreement about that with Dwight Bolinger. I think, for instance, that contrary to what is sometimes assumed in the literature, for (6) to have a broad focus, a number of pitch accents must be spread throughout the utterance. (cf. Bolinger 1986, p.100.)

(6) He was warned to look out for an ex-convict with a red shirt.

More specifically, Bolinger gives the following example, where his judgment is analogous to my judgment concerning (5-b):

Suppose, for example, that the verb is deaccented in the sentence

The Chicágo police confiscated a cáche of drúgs.

- The effect is to presuppose confiscation, by prior mention, by obviousness under the circumstances, or whatever. (Bolinger 1986, p. 100)

And what about the traditional judgment, that (5-b) is a felicitous response to (2) in the absence of further context? I think it is highly instructive to take a look at this judgment and at how it might have come about. Obviously, people made such judgments relative to some actual (though not necessarily naturally-occurring) utterance of (2)+(5) (or similar examples) 'out of the blue'. I would like to argue that such judgments could easily have come about through misconstrual of the example being judged. I would venture to claim that either (i) **it wasn't really (5-b)** that we were judging, or (ii) **it wasn't really in the absence of further context** that we were making our judgment. Let me elaborate on each of these options in turn.

Starting with number (i), the thing is, it is not always easy to tell whether we are hearing (5-b) or (5-c). That is, it is not always easy to hear if *invited* carries a pitch accent or not. We know for sure that the nuclear accent is on *Bill*, but that in itself is not enough to tell us whether or not there is a pitch accent on *invited*. Also, if there does happen to be a pitch accent on *invited*, it would of course be much less prominent than the nuclear accent. Therefore, a H\* pitch accent on *invited* could easily go unnoticed. (Perhaps even phonetically neutralized altogether.) But that means that it could easily have happened that we believed we'd heard (5-b) answering (2), when in fact it was (5-c) that had been pronounced. And so, since (2)+(5-c) is OK, we mistakenly concluded that (2)+(5-b) was OK too.

Regarding the potential difficulty in identifying a pre-nuclear H\* accent, Pierrehumbert 1980, section 2.2, already notes that the typical dipping between two H\* pitch accents often disappears in a fast sequence of H\*s. And I'd like to cite a few passages from Beckman 1996, pertaining to the 'flat-hat' contour, which consists of a sequence of one or more H\* accents followed by a L phrase accent and L% boundary tone (and is the archetypal American English citation-form intonation). Beckman notes not only that pre-nuclear H\*s may be hard to identify:

[...] the most consistent and by far the most prominent event in the contour will be the large pitch fall from the nuclear pitch accent to the L phrase accent. (p.35)

[...] there may be a small rise from a neutral pitch at the onset of phonation to the high pitch target of the first H\*, giving some phonetic salience to the first accent. However, if the initial accent is on the first syllable in the phrase [...], there may not be even the phonetic rise to the initial H\* target. Moreover, for stressed syllables between this initial rise and the nuclear pitch accent, it will be difficult to differentiate between a H\* target specification and a phonologically unspecified point in the phonetic interpolation between the two H\* targets, so that accentuation in this part before the nuclear syllable is inherently ambiguous. (p.35)

but also that, as I have just suggest, this fact may have interfered with our impression of what accent-placements were consistent with which foci:

We might ask whether the ambiguous scope of focus for utterances with late nuclear accent (as in the utterance of *Legumes are a good source of vitamins* with nuclear accent on *vitamins*) is in some sense an artifact of the inherent ambiguities in parsing pre-nuclear accent placement in the flat-hat tune. (p.36)

As for number (ii): People intended to judge (2)+(5-b) (or similar examples) in the absence of further context. However, when presented with (2)+(5-b) out of the blue, one is very likely to automatically accommodate further context, possibly without realizing that. In which case one would certainly be justified in judging the sequence felicitous – but, of course, one would **not** be judging it in the absence of further context.

After all, it is very easy to imagine a simple context where (5-b) would felicitously answer (2), such as the party-planning context. (Recall that with (5-a) we needed a special context in which *Bill* would be expected and nevertheless not replaced by a pronoun. There is no parallel difficulty here, since verbs are less often replaced with anaphoric expressions, even if expected in the context.) Craige Roberts (p.c.) notes, in addition: Of the two internal arguments of *invite* – the invitee and the event to which they're invited – the second one is omitted, which means that (5-b) presupposes some context rich enough to supply it. This strengthens our tendency to automatically assume that inviting is salient in the context.

The idea that examples uttered/presented out of the blue are not always really judged in the absence of further context is not new. Bolinger 1986, pp. 100-101 states it very clearly precisely in the context of pitch accent placement:

One way of testing for a neutral sentence [i.e., roughly, one with sentential focus – NK] is to see whether it is a normal answer to the general question *What happened?* This is just a way of determining whether the sentence is one that could come out of the blue. But it is not an entirely reliable test, because there can be presuppositions and “givens” even then. Modifying our last example slightly we might get

What happened? – The Chicágo police found a cáche of drúgs.

With *found* deaccented. Finding is presupposed on the strength of the mere mention of police and drugs. [...]

Unless the speakers themselves come out of the blue, they are bound to share varying amounts of information and expectations and then may deaccent the referring words. One might thus greet a friend with

Hi there, Jack – I see you're not góing to the office these days.

with *to the office these days* deaccented, when there is prior knowledge of Jack'e office-going habit.

Finally, note also that we can see why the intuition that (1)+(5-b) is not a good response seems clearer when the question being answered is (1). We are not as likely to imagine that (1)+(5-b) is uttered in something like the party-planning context, because (1) is typically asked in contexts where it is not even expected that the addressee has done anything, let alone invited someone.

### **3.2 Syntactic Distinctions Do not play a decisive role here**

Having discussed at some length a few alleged facts which suggest that pitch accent placement is partially governed by structural principles, I will now go on to briefly look at a number of specific theories which do in fact link accent placement to syntactic distinctions – in particular, to argument structure.

Selkirk 1984 argues that the theory of pitch accent placement and interpretation should make reference to the distinction between heads and arguments on the one hand and modifiers and adjuncts on the other. This position is based on the observation that a lone pitch accent on a modifier or adjunct does not seem to allow for a broad focus on a larger constituent. For instance, while it is possible to answer (8) with both (9-c) and (9-b), it seems impossible to answer (8) with (9-a).

(8) What bit you?

(9-c) A red ant bit me.  
H\* H\* L L%

(9-b) A red ant bit me.  
H\* L L%

(9-a) A red ant bit me.  
H\* L L%



- (17) John bought a red tie. (Selkirk 1984, her (5.48))  
pa pa

In (16), *read* is accented, while it is evidently 'given' in the context. In (17), *bought* and *red* are unaccented, while they can, it seems, be 'new' or unexpected in the context (as when a surprising piece of news is reported out of the blue). So it looks like accent placement on these words does not get interpreted.

I should like to note, however, that (18) is an example parallel to (16) (the accented *compliments* being quite as 'given' as *read* was), which the above proposal leaves unexplained. Does the presence or absence of a pitch accent on an argument not get interpreted either?

- (18) A: Why don't you go and give her a compliment?  
B: Sorry, I can't. I don't give compliments. It's a strict policy.  
pa

Further, as already indicated in the preceding section, I do not believe that unaccented words can (in English) really be 'new' or unexpected in the context. (An exception is a word within a compound, since accent placement within a compound is determined by the stress pattern, as it is within a word.) Regarding (17), I would dispute the claim that it can really be felicitous, with the entire sentence focused, in a context where *bought* and *red* are 'new' or unexpected. Suppose there was no talk of shopping or colors at all, and I report the events of the day 'out of the blue'. Would I pronounce (17) as marked? I think not. I think the most natural pronunciation is with four pitch accents, one on each content word. If shopping is frequent or expected, then I may drop the pitch accent on *bought*. But I would still not drop the pitch accent on *red*. (Our sentence is much more likely I think to be pronounced with just the two pitch accents indicated if we replace *red* with *new*, for the obvious reason that in the context of shopping, 'red' is likely not to be expected, while 'new' probably is expected.)

A different proposal was made in Rochemont 1986, and later adopted in Selkirk 1996, to the effect that a head is allowed to get interpreted as 'new' if it has an accented internal argument. The idea is to allow unaccented heads, as in (17), to be interpreted as 'new', while holding on to the view that unaccented arguments are always 'given'. But this approach does not seem to me to be right either, as I've just argued that *bought* in (17) cannot in fact be interpreted as 'new' or unexpected. Note also that since the new approach continues to assume that any word carrying a pitch accent is invariably 'new', it has nothing to say about why either (18) or (16) is felicitous.

I would conclude that pitch accent placement and interpretation do not (in English) depend on syntactic argument structure. Not in any absolute or very

general way, at any rate. (A similar conclusion is independently reached in Schwarzschild 1999. For more discussion of the effect of syntactic or structural factors, see Kadmon 2001, section 13.4.4.1 (on Ladd 1980) and section 13.4.5.)

An anonymous referee suggested that examples like (101)+(102-a) show that, contrary to my claim about the facts, a verb can be both unaccented and 'new' after all:

- (101) Hi, what's up?  
(102-a) My mother is coming.  
          H\*   L       L%

I don't think so. Note that, out of the blue, the version in (102-b) is **not** a natural response to (101).

- (102-b) My mother is singing.  
          H\*   L       L%

The contrast between (102-a) and (102-b) is clearly due to the fact that while mothers are not in general expected to sing, the mention (and accenting) of 'my mother' is enough to make 'coming' (i.e., coming to visit), 'rather highly predictable' (Bolinger 1986, p.105). So I simply think that given the right notion of 'new' – something like 'unexpected'/'unpredictable' – the verb in (102a) is **not** 'new'. (For more concerning the notion of 'new', the reader is referred of course to the rest of this paper.) Note, moreover, that in the right context, it is possible for *mother* too to be part of the focus but unaccented:

- () A: My parents are here, and a lot of other guests.  
   B: What's happening?  
   A: My mother is singing. (Kadmon 2001, p.269)  
          H\* L L%

Hence, the ability to be unaccented in the context of (101) does not distinguish the verb from its argument, and cannot provide evidence for the role of the head vs. argument distinction in accent placement.

#### 4. Selkirk's theory (without reference to argument structure):

**A recursively projected feature E; the maximal E-marked constituent interpreted as focus; absence/presence of E interpreted as 'given'/'new'**

Given the conclusions of the preceding section, I should like to consider a version of the theory of Selkirk 1984, 1996 which is stripped off of its reference to argument structure. Such a version is given in (19). Note: Selkirk uses a

syntactic feature *F*. Since my personal practice is to reserve the use of *F* to mark the focus (i.e., 'the answer to the question'), I have replaced Selkirk's *F* with *E* (for *Elisabeth*).

(19) **(A) The syntactic feature *E* and its projection:**

- (i) A word is *E*-marked iff it contains a pitch accent.
- (ii) A constituent may be *E*-marked if it contains at least one sub-constituent which is *E*-marked.

Other constituents may not be *E*-marked.

**(B) Interpretation of the feature *E* (Selkirk 1996):**

- (i) A maximal (= non-embedded, 'highest') *E*-marked constituent is defined as a focus, to be interpreted by the semantic/pragmatic theory of focus.
- (ii) An embedded *E*-marked constituent is interpreted as 'new' in the discourse, while a constituent which is not *E*-marked is interpreted as 'given'.

The recursive clause (A-ii) allows the feature *E* to 'project upward' from an accented word to higher, containing constituents. (A) determines that (20), for instance, is compatible with the two representations in (21) and (22).

(20) He introduced his mother to us.

H\* L L%

(21) He [introduced [his [mother]<sub>E</sub>]<sub>E</sub> to us ]<sub>E</sub>

*pa*

(22) He [introduced [his [mother]<sub>E</sub>]<sub>E</sub> to us ]

*pa*

Such representations are interpreted by the clauses in (B). Concentrating for the moment on (B-i), let us assume that the semantics/pragmatics interprets the maximal *E*-marked constituent as the (short) answer to the last question under discussion in the discourse. (For a formal theory of focus along these lines, see Roberts 1996.) (B-i) predicts, correctly, that (20) can answer both *What did he do?* and *Who did he introduce to us?:* in (21), the maximal *E*-marked constituent is the VP, so the VP is the focus = the answer to the question; in (22), the maximal *E*-marked constituent is the NP *his mother*, so this NP is the focus = the answer to the question.

A major claim of Selkirk 1984 is that the particular pitch accent placement inside the focus indicates the informational status of sub-parts of the focus, along the lines already sketched in section 2 above. It is this suggestion that is represented by (B-ii).

Consider for instance the examples in (23), from Selkirk 1984.

- (23) a. She sent the book to Mary.  
          pa      pa
- b. She sent the book to Mary.  
          pa

Selkirk notes that the two versions can answer *What did she do next?*. This is correctly predicted by her theory, which declares the following two representations to be well-formed.

- (24) a. She [ sent [the [book]<sub>E</sub>]<sub>E</sub> [to [Mary]<sub>E</sub>]<sub>E</sub> ]<sub>E</sub>  
          pa                  pa
- b. She [ sent [the book] [to [Mary]<sub>E</sub>]<sub>E</sub> ]<sub>E</sub>  
                                  pa

Selkirk also notes that while (23a) is a natural 'out of the blue' response to *What did she do next?*, (23b) can answer such a question only when reference to a certain book is expected – e.g., in a context where Jane's job is illustrating books, and we've been talking about her current book. This too is explained on the basis of the representations in (24): in (23a), represented as in (24a), both direct object and indirect object are E-marked, and hence 'new'; in (23b), represented as (24b), the direct object is not E-marked, and hence 'given'. (*Note*: I believe that *sent* should also be accented, unless sending is expected in the context.)

Selkirk does not provide a definition or explicit characterization of the notions 'given' and 'new'. How to explicate such notions is, of course, a difficulty often encountered and much discussed in the semantic/pragmatic literature.

Schwarzschild 1999 briefly considers the idea of giving some official content to Selkirk's 'given' and 'new' by interpreting them as the **given** and **new** informally defined in (25).

- (25) An expression B is **given** iff there is an expression A previously mentioned in the discourse such that A is salient and A entails B.

Where 'entails' applies to expressions of various semantic types in the obvious way (*green apple* entails *apple*, *ran* entails *moved*, etc.).

An expression is **new** iff it is not **given**.

Schwarzschild shows that this idea immediately encounters empirical difficulties. To see the difficulty, let us consider how Selkirk's theory, with 'new'/'given' interpreted as new/given, handles dialog (2)+(5-c).

- (2) What did you do?  
 (5-c) I invited Bill  
           H\*   H\* L L%

The theory of E-projection allows the following representation of (5-c).

- (26) I [ [invited]<sub>E</sub> [Bill]<sub>E</sub> ]<sub>E</sub>  
           pa           pa

This almost succeeds in predicting that in the absence of further context, (5-c) can felicitously follow (2). The maximal E-marked constituent is the VP, so by (B-i) the VP should be interpreted as the focus = the answer to the question. That is exactly what's compatible with (2) being the preceding question in the discourse. *Invited* is E-marked, so by (B-ii) it is interpreted as 'new', and the same goes for *Bill*. And indeed, they are each new in the context of (2). So far so good. But now consider the entire S. The S is not E-marked (nor could it be, in the context of (2), since it does not constitute the short answer to (2)), so (B-ii) requires it to be 'given'. But in the context of (2), it is not given. So (26) is not compatible with (2) being the preceding question. There is no other representation that would be compatible with (2) (the reader can verify that for herself). So in the end, our theory incorrectly predicts that (5-c) cannot follow (2).

The problem just illustrated is a general one: (B-ii) requires any larger constituent containing the focus to be 'given'; but as a matter of empirical fact, focus-containing constituents can certainly be new.

Schwarzschild also brings up the following example, which Selkirk's theory, with 'new'/'given' interpreted as new/given, cannot handle. Both (28-a) and (28-b) are felicitous responses to (27).

- (27) John's mother saw Bill.  
       And then what happened?
- (28-a)           She saw her mother.  
   pa
- (28-b)           She saw her mother.  
   pa

But the theory under consideration incorrectly rules them both out. The accented word is required to be E-marked, and hence must be interpreted as either **new** or the answer to the preceding question. However, it is neither. (It is the entire S that constitutes the answer to the preceding question.)

### 5. Schwarzschild's Theory:

**A freely but sparingly distributed feature R; no focus interpretation; absence of R interpreted as GIVEN, with definition of GIVEN itself making reference to R**

Schwarzschild 1999 offers an interesting alternative theory of the effect of pitch accent placement on question-answer congruence, which differs quite radically from Selkirk's theory. Schwarzschild's theory, which includes a new notion of 'givenness' more sophisticated than the **given** presented above, makes no use of recursive projection of a syntactic feature, and makes no use of any theoretical notion of 'focus'. In this section, I will present Schwarzschild's theory, and argue that it quickly runs into empirical problems.

Following is a version of Schwarzschild's theory. Schwarzschild too uses a syntactic feature F, and I have replaced his F with R (for *Roger*). I have only included an informal definition of GIVEN – for formal details I refer the reader to Schwarzschild's paper. As with Selkirk's theory, I have eliminated reference to syntactic argument structure (Schwarzschild's constraint HeadArg). I have also left out Schwarzschild's constraint FOC, which I will briefly introduce below.

#### (29) Schwarzschild's Theory (minus HeadArg and FOC):

##### (A) The syntactic feature R:

- (i) If a word contains a pitch accent it is R-marked.
- (ii) Any constituent may be R-marked.

##### (B) Definition of GIVEN:

An expression B is **GIVEN** iff there is an expression A previously mentioned in the discourse such that A is salient and A entails the result of replacing with a variable any (top-most) R-marked constituent within B.

Where 'entails' applies to variable-containing expressions in the obvious way (*Bill smiled* entails *Bill X-ed*, *a green apple* entails *an X apple*, etc.).

##### (C) Constraints:

- (i) GIVENness: A constituent that is not R-marked must be GIVEN.

- (ii) AvoidR: The representation must contain the smallest number of R features possible.

It is assumed that a *wh*-question makes salient the corresponding existential statement. Hence, in the context of *Who smiled?*, *X smiled* is GIVEN, in the context of *Who did Bill invite?*, *invited X* is GIVEN, etc.

Let us look at some data. Of the three versions of (5), (5-a) is the only one that can felicitously answer (3). This is correctly predicted by Schwarzschild's theory, as follows.

(3) What did you do to Bill?

(5-a) I invited Bill  
H\* L L%

(5-b) I invited Bill  
H\* L L%

(5-c) I invited Bill  
H\* H\* L L%

We need to find out which representation(s), of all possible representations of (5), will satisfy (C-ii), that is, AvoidR. To do that, let us consider each version of (5) in turn, to see how it might be represented, in the context of (3), with the smallest number of Rs.

For (5-a), the representation compatible with (3) with the smallest number of Rs is (30).

(30) I [invited]<sub>R</sub> Bill  
pa

*Invited* must be R-marked both because it is accented (see (A-i)) and because in the context of (3), it is not GIVEN (see (C-i)). None of the other constituents need to be R-marked. Note, in particular, that the VP is GIVEN: *did something to Bill*, which has been made salient by question (3), entails *Xed Bill*, which is the result of replacing the R-marked *invite* in the VP with a variable.

For (5-c), the representation compatible with (3) with the smallest number of Rs is (31).

(31) I [invited]<sub>R</sub> [Bill]<sub>R</sub>  
pa pa

*Invited* must be R-marked, as before, both because it is accented and because it is not GIVEN. And this time *Bill* too must be R-marked, because it is accented. (Note: the VP is GIVEN, since *did something to Bill* entails *Xed Y*.)

For (5-b), the representation compatible with (3) with the smallest number of Rs is (32).

(32) I [invited]<sub>R</sub> [Bill]<sub>R</sub>  
pa

*Invited* must be R-marked because it is not GIVEN. *Bill* must be R-marked because it is accented.

Given all of the above, it is correctly predicted that (5-a) is the only felicitous response to (3). Out of (30)-(32), AvoidR forces the speaker to choose (30) – i.e., the representation of (5-a) – since it is (30) that has the smallest number of Rs.

Schwarzschild can also handle the **her mother** example, which was a problem for Selkirk's theory.

(27) John's mother saw Bill.  
And then what happened?

(28-a) She saw her mother. (27)+(28-a) GOOD Schw: predicted  
pa (Selk: not predicted)

(28-b) She saw her mother. (27)+(28-b) GOOD Schw: predicted  
pa (Selk: not predicted)

Schwarzschild's theory correctly predicts that both (28-a) and (28-b) are felicitous answers to (27). That is because they both require two R features and no more than that, and therefore AvoidR does not determine a choice between them.

(28-a) can be represented as in (33).

(33) She [ saw [ [her]<sub>R</sub> mother ] ]<sub>R</sub>  
pa

The object NP is GIVEN, because *John's mother* entails *X's mother*. But the VP is not GIVEN, because in the context of (27), nothing entails *saw X's mother*. The S is GIVEN again, because *John's mother saw someone* entails *she Xed*.

(28-b) can be represented as in (34).

(34) She [ saw [ her [mother]<sub>R</sub> ]<sub>R</sub> ]  
pa

The object NP is not GIVEN, because in the context of (27), nothing entails *her X*. The VP is GIVEN, because *saw someone* entails *saw X*. The S too is GIVEN, because *John's mother saw someone* entails *she saw X*.

It is interesting to note that on Schwarzschild's theory, the relation between pitch accents, the posited syntactic feature, and semantic/pragmatic interpretation is rather more intricate than it was on Selkirk's theory. There is no longer a 1-1 correspondence between accented words and 'new' words, since a word can be GIVEN with or without a pitch accent. There is no 1-1 correspondence between R-marking and 'new', since an R-marked constituent can be either GIVEN or not GIVEN. There isn't any simple, straightforward way of saying what R - or lack thereof - actually means, semantically/pragmatically. It is true that the **absence** of R is uniformly interpreted as GIVEN. But there is much more than that. For one thing, due to AvoidR, there is a **tendency** for constituents that **are** R-marked **not** to be GIVEN. In addition, the R feature plays a crucial role in the recursive definition of GIVEN. Due to that, the presence of R on a constituent  $\alpha$  means that  $\alpha$  being 'new' (i.e., not GIVEN) cannot prevent a constituent containing  $\alpha$  from being, in a sense, 'old' (i.e., GIVEN).

I will now show that Schwarzschild's theory quickly runs into empirical problems. We will see what might be learned from these problems.

Problem type 1: failure to force certain words that are 'new' to be accented

Let us look at more data. In response to question (35), sentence (5) must bear a pitch accent both on *invited* and on *Bill*. In other words, (5-c) is a felicitous answer to (35), whereas (5-a) and (5-b) are not.

(35) What did you do to who?

(5-a)	I invited Bill H* L L%	(35)+(5-a) BAD	Schw: not predicted (predicted with FOC)
(5-b)	I invited Bill H* L L%	(35)+(5-b) BAD	Schw: not predicted (predicted with FOC)
(5-c)	I invited Bill H* H* L L%	(35)+(5-c) GOOD	Schw: predicted

But Schwarzschild's theory, as given above, incorrectly predicts that all three versions of (5) should be felicitous answers to (35). In the context of (35), both *invited* and *Bill* must be R-marked, since they are not GIVEN. So at least two Rs are required. No further R feature is required, since the VP is GIVEN. But all that is independent of which, if any, of *invited* and *Bill* are accented. Hence, all three versions of (5) equally allow a representation with just two Rs. So AvoidR does not prefer (5-c) to the other versions.

Schwarzschild adds to his theory the constraint in (36), which can take care of this case.

- (36) (C-iii) FOC: An R-marked constituent which is not immediately dominated by another R-marked constituent must contain a pitch accent.

Given FOC, to allow the R-marking on the unaccented non-GIVEN word in (5-a) or (5-b), the VP or S will also have to be R-marked, making three Rs in all. Therefore, AvoidR will prefer (5-c), which allows for just two Rs.

However, exactly the same problem re-emerges in examples where FOC is of no help. Consider once more our old friend question (2). Which of the versions of (5) can answer (2), in the absence of any further context? Here too only (5-c) is good. It is agreed in the literature that (2)+(5-a) is not felicitous (out of the blue), and I have claimed in this paper that careful consideration of the judgments reveals that (2)+(5-b) is not felicitous either (out of the blue).

(2) What did you do?

- |                                    |                                    |  |
|------------------------------------|------------------------------------|--|
| (5-a) I invited Bill<br>H* L L%    | (2)+(5-a) BAD out of the blue      | Schw: not predicted<br>(not even with FOC) |
| (5-b) I invited Bill<br>H* L L%    | (2)+(5-b) Kad: BAD out of the blue | Schw: not predicted<br>(not even with FOC) |
|                                    | [trad'l: GOOD out of the blue]     | [Schw: predicted]                          |
| (5-c) I invited Bill<br>H* H* L L% | (2)+(5-c) GOOD out of the blue     | Schw: predicted                            |

What Schwarzschild's theory predicts, incorrectly, is that all three versions are felicitous responses to (2), out of the blue. *Invited* and *Bill* must be R-marked because of their not being GIVEN. The VP is not GIVEN either, because *did something*, made salient by (2), does not entail the result of replacing the two R-marked parts of the VP with variables, which is *Xed Y* ('did something to someone'). So the VP must bear a third R. No further Rs are needed. (The S is

GIVEN, because *I did something* entails the result of replacing the R-marked VP with a variable, viz., *I Xed.*) But all that is independent of which, if any, of *invited* and *Bill* are accented. Hence, all three versions of (5) equally require exactly three Rs. Therefore, AvoidR does not determine a choice between them, and Schwarzschild predicts that they should all be possible responses.

We saw then that in the absence of further context, (5-a) and (5-b) are not felicitous responses to either (35) or (2), while Schwarzschild's theory (without FOC) incorrectly predicts that they are. The theory fails to force *invited* and *Bill* to both be accented.

(35) What did you do to who?

(2) What did you do?

(5-a) I invited Bill  
H\* L L%

(5-b) I invited Bill  
H\* L L%

What is the source of the problem? Well, in a context consisting of just (35)/(2), both *invited* and *Bill* are 'new'. So it seems that we would have been better off with a theory that forced 'new' words to be accented.

The fact is that Schwarzschild's theory allows words that are both not GIVEN and unaccented. In each of (5-a) and (5-b), there is just such a word. It seems that we would have been better off with a theory that did not allow a word to be both 'new' and unaccented, which would have directly ruled out these two responses. Note also that the effect of FOC is precisely to force 'new' constituents to contain a pitch accent – though only in certain cases. This suggests to me that it is problematic that Schwarzschild's theory allows 'new' constituents to contain no pitch accent at all in the first place, and that we need more of a 1-1 correspondence between accent and 'novelty'.

- conclusion 1: Allowing a 'new' constituent to contain no pitch accent at all is problematic. I think we need more of a 1-1 correspondence between accent and 'new'.

Problem type 2: failure to force the word providing the (short) answer to the question under discussion to be accented, when it occurs in preceding discourse

Next, I would like to consider an example taken from Schwarzschild 1999, namely, responding to question (39) with the sentence in (40).



But let us move on to the case where (40) answers question (39). Here the situation is more difficult, because not even the use of 'rhetorical relations' can yield the right prediction. Schwarzschild naturally holds that when (40) answers (39) it bears a 'rhetorical relation' to the whole of (39). But that means that all of (39) is salient, and therefore *borrowed* is GIVEN. This forces an incorrect prediction. We saw that on the assumption that *borrowed* is GIVEN, both (40-a) and (40-b) can be represented with just one R, as in (41) and (42) above. Therefore, AvoidR allows both (40-a) and (40-b) as answers to (39), and provides no means of predicting that in fact only (40-a) is felicitous.

This problem is acknowledged by Schwarzschild. His solution is that, 'apparently', (40-a) is chosen because it unambiguously answers question (39), while (40-b) could be used either to answer question (39) or to correct the assumption that Max purchased the book. However, this solution is not going to work. Take example (44). Here too, it is not possible to utter (40-b). But Schwarzschild's theory fails to predict that. The particular accent placement of (40-a) is not needed for making clear that the speaker is answering (39), since the speaker is saying so, explicitly.

(44) I'll answer your question: Max borrowed it.

I think that a very intuitive explanation of why the pitch accent must be on *Max* and not on *borrowed* is that it is *Max*, but not *borrowed*, that constitutes the answer to question (39). This suggests to me that the source of the problem that Schwarzschild's theory encounters here is that it fails to capture part of the role that the question under discussion in the discourse plays in constraining pitch accent placement.

I should note that Schwarzschild's theory does allow **some** role to be played by the question under discussion. It does correctly predict that various words are **not** accented based on the fact that the question under discussion renders them GIVEN (which allows for an absence of R). The part that Schwarzschild's theory does not seem to capture is that some words **are** accented because they provide the answer to the question under discussion, even if they too are rendered GIVEN by that question.

- conclusion 2: I think Schwarzschild does not capture in full the role of the question under discussion. I think that, simply: if a word constitutes the (short) answer to the question under discussion, it will be accented (even if it is GIVEN).

The idea that a word might be accented (even when GIVEN) if it is the answer to the question would work for the case of (40) being used to correct the

assumption the the book was purchased, as well. Consider (45). When the response to (39) is (45), the intended reading of (40) is disambiguated.

- (45) Ah, wait, but you know how Max really got this book?  
Max borrowed it.

Rather than say that *borrowed* is accented here because it is no longer salient, we can say that *borrowed* is accented because it is the answer to the question in (45). Which would be more intuitive, I think. Rather than say that (40) bears a rhetorical relation to the object NP (or relative clause) in (39), we can say that (40) addresses the question in (45). It is also rather intuitive, I think, to say that when (40) directly follows (39), it still addresses the question in (45), even though that question is not explicitly mentioned.

#### 6. A Selkirk-like theory with given:

**A recursively projected feature E; the maximal E-marked constituent interpreted as focus; absence/presence of E interpreted as given/new, but with given applied only to focus-internal constituents**

Given the problems with Schwarzschild's theory, it seems worthwhile to reconsider the combination of Selkirk's theory with the notion of given, as discussed in section 4. Recall that Selkirk interprets E-marking as 'new' and its absence as 'given', and that Schwarzschild 1999 briefly considers equating her 'given' and 'new' with the notions given and new, which he defines as in (25).

- (25) An expression B is given iff there is an expression A previously mentioned in the discourse such that A is salient and A entails B.

Where 'entails' applies to expressions of various semantic types in the obvious way (*green apple* entails *apple*, *ran* entails *moved*, etc.)

An expression is new iff it is not given.

In this section, I will argue that the theory given in (46) below – that is, Selkirk's theory with given plus one amendment (italicized) to clause (B-ii) – is a viable option, at least as empirically adequate as Schwarzschild's theory.

#### (46) A Selkirk-like theory with given:

##### (A) The syntactic feature E and its projection:

- (i) A word is E-marked iff it contains a pitch accent.
  - (ii) A constituent may be E-marked if it contains at least one sub-constituent which is E-marked.
- Other constituents may not be E-marked.

**(B) Interpretation of the feature E:**

- (i) A maximal (= non-embedded, 'highest') E-marked constituent is defined as a focus, to be interpreted by the semantic/pragmatic theory of focus.
- (ii) An embedded E-marked constituent is interpreted as **new**, while a constituent which *is embedded in an E-marked constituent but* is not itself E-marked is interpreted as **given**.

Recall the problem that Schwarzschild 1999 points out with regards to Selkirk's theory combined with **given**: that theory requires any larger constituent containing the focus to be **given**, while empirically, focus-containing constituents can certainly be **new**. My amendment to clause (B-ii) is a way of solving that problem. I have simply limited interpreting a constituent without the E feature as **given** to constituents (properly) contained within a focus.

Rather than almost succeed, the present theory really succeeds in predicting that in the absence of further context, (5-c) can felicitously follow (2). (5-c) can be represented as (26).

(2) What did you do?

(5-c) I invited Bill (2)+(5-c) GOOD out of the blue Schw: predicted  
           H\* H\* L L% Selk-like: predicted

(26) I [ [invited]<sub>E</sub> [Bill]<sub>E</sub> ]<sub>E</sub>  
           pa pa

As seen in section 4, the three E features are compatible with the context of (2). So is the lack of E on *I*. And given the new (B-ii), the absence of E on the entire S is also compatible with (2), because this absence does not get interpreted. (As shown in section 5, Schwarzschild's theory too makes the correct prediction here.)

In fact, the present theory correctly predicts all of the data discussed in the previous section, with the exception of one example (the **her mother** example).

The present theory correctly predicts that of the three versions of (5), (5-a) is the only felicitous response to (3). (As was also correctly predicted by Schwarzschild's theory.)

(3) What did you do to Bill?

(5-a)	I invited Bill H* L L%	(3)+(5-a) GOOD	Schw: predicted Selk-like: predicted
(5-b)	I invited Bill H* L L%	(3)+(5-b) BAD	Schw: predicted Selk-like: predicted
(5-c)	I invited Bill H* H* L L%	(3)+(5-c) BAD	Schw: predicted Selk-like: predicted

(3)+(5-a) is predicted to be felicitous, because (5-a) can be represented as in (47), which is fully compatible with (3) being the context. The theory requires *invited* to be interpreted as the answer to the question, and indeed, it is the answer to (3).

(47) I [invited]<sub>E</sub> Bill  
pa

Both (5-b) and (5-c) are predicted to be bad responses to (3), because they don't have any representation which is compatible with (3). In both cases we have this: *Bill* is accented, so it must be E-marked; therefore, it must be interpreted either as **new** or as the focus; but that is incompatible with (3), since *Bill* is neither **new** in the context of (3) nor the answer to (3).

The present theory correctly predicts that of the three versions of (5), (5-c) is the only felicitous response to (35) – a prediction that Schwarzschild's theory was only able to make with the added constraint FOC.

(35) What did you do to who?

(5-a)	I invited Bill H* L L%	(35)+(5-a) BAD	Schw: not predicted (predicted with FOC) Selk-like: predicted
(5-b)	I invited Bill H* L L%	(35)+(5-b) BAD	Schw: not predicted (predicted with FOC) Selk-like: predicted
(5-c)	I invited Bill H* H* L L%	(35)+(5-c) GOOD	Schw: predicted Selk-like: predicted

(35)+(5-c) is predicted to be felicitous, because (5-c) can be represented as in (90), which is fully compatible with (35) being the context. *Invited* and *Bill* are interpreted as two separate foci, which is exactly what's required in the context of question (35).

(90) I [invited]<sub>E</sub> [Bill]<sub>E</sub>  
       pa       pa

(5-a) and (5-b) are predicted **not** to be good responses to (35), because the unaccented word – *Bill* or *invited* – cannot be E-marked, while it should, in the context of (35), be interpreted as a focus (one of two separate foci answering 'what' and 'who').

Now let us return to answering question (2). Note, first, that the present theory correctly predicts that (2)+(5-b) is felicitous in a context where a party is being planned and people are being invited. (48) below is a representation of (5-b) which is fully compatible with such a context. Given question (2), the VP must be the focus, so it must be E-marked. So *invited* is embedded in an E-marked constituent. But *invited* itself cannot be E-marked, since it is not accented. Hence, by clause (B-ii), *invited* must be interpreted as **given**. Is it in fact **given**? Yes, in the party context it is. (This prediction was also made by Schwarzschild's theory.)

(2) What did you do?

(5-b) I invited Bill	(2)+(5-b)	GOOD in context (party)	Schw: predicted
H* L L%			Selk-like: predicted
	(2)+(5-b)	Kad: BAD out of the blue	Schw: not predicted
			Selk-like: predicted
(5-a) I invited Bill	(2)+(5-a)	BAD out of the blue	Schw: not predicted
H* L L%			Selk-like: predicted

(48) I [ invited [Bill]<sub>E</sub> ]<sub>E</sub>  
       pa

Now consider answering (2) in the absence of any further context. The present theory correctly predicts that in that case, (5-b), and likewise (5-a), are **not** good responses to (2) – predictions that Schwarzschild's theory was not able to make at all. Take (2)+(5-b) first. I just showed that *invited* must be interpreted as **given**. But now that becomes a problem, because if we are not in the context of inviting people, the word *invited* is in fact **new**. Therefore, there is no representation of (5-b) that is compatible with (2) being the sole context. Similarly with (2)+(5-a). Here we have exactly the same problem, but with *Bill* – the unaccented *Bill* must be interpreted as **given**, but without further context, it is in fact **new**.

The present theory correctly predicts that in (44), (40) must be pronounced as (40-a), while in (45), it must be pronounced as (40-b). This is another advantage over Schwarzschild's theory, since, as shown in section 5,

Schwarzschild's theory can predict that (40) in (45) must be pronounced as (40-b) only on the assumption that *borrowed* is no longer salient, and it is completely unable to predict that (40) in (44) must be pronounced as (40-a).

- (39) Who borrowed the book that Max had purchased?  
 (44) I'll answer your question: Max borrowed it.  
 (45) Ah, wait, but you know how Max really got this book?  
 Max borrowed it.

(40-a)	Max borrowed it. pa	in (39)+(44): GOOD  in (39)+(45): BAD	Schw: predicted Selk-like: predicted Schw: predicted? only assuming <i>borrowed</i> is no longer salient Selk-like: predicted
(40-b)	Max borrowed it. pa	in (39)+(44): BAD  in (39)+(45): GOOD	Schw: not predicted Selk-like: predicted Schw: predicted Selk-like: predicted

Take the good cases first. Representation (49) is compatible with the context of (39)+(44), and (50) is compatible with the context of (39)+(45). In each case, the maximal E-marked constituent is indeed the answer to the question under discussion.

(49) [Max]<sub>E</sub> borrowed it.  
pa

(50) Max [borrowed]<sub>E</sub> it.  
pa

Now the bad cases. There is no representation of (40-b) which is compatible with the context of (39)+(44): since *borrowed* is accented, it must be E-marked, and hence be either **new** or the focus; but in fact, in the context of (39)+(44), it is neither **new** nor the answer to the last question under discussion, viz., (39). Similarly, there is no representation of (40-a) which is compatible with the context of (39)+(45), because *Max*, being accented, must be E-marked, while in the context of (39)+(44) it is in fact neither **new** nor the answer to the last question under discussion, viz., how Max got the book.

What we see up till now is that the Selkirk-like theory considered in this section enjoys some empirical advantage over Schwarzschild's theory, as it is

able to handle both types of examples that were seen in section 5 to be problematic for Schwarzschild's theory.

Recall that I said in section 5 that the problems with Schwarzschild's theory suggest (i) that we need more of a 1-1 correspondence between accent and 'novelty', and (ii) that we need to let the question under discussion play a bigger role in constraining pitch accent placement. It seems that these two things are indeed what lends the present theory its empirical advantage.

There is a closer correspondence between accent and 'novelty' in that the present theory requires focus-internal unaccented words to be 'given'. That is why it succeeds in predicting that in the absence of further context, (2)+(5-b) is infelicitous.

- (2) What did you do?  
 (5-b) I invited Bill  
           H\* L L%

The present theory also brings in the notion of 'focus', forcing the maximal E-marked constituent to constitute the answer to the question under discussion. That is what allows it to predict that in (44), it must be *Max*, and not *borrowed*, that bears the pitch accent.

- (39) Who borrowed the book that Max had purchased?  
 (44) I'll answer your question: Max borrowed it.

At the same time, it must be admitted that the theory of this section is at a disadvantage compared to Schwarzschild's theory in that it is unable to handle the **her mother** example – that it, to predict the fact that both (28-a) and (28-b) are felicitous responses to (27).

- (27) John's mother saw Bill.  
       And then what happened?

- |                                  |                  |   |
|----------------------------------|------------------|---|
| (28-a) She saw her mother.<br>pa | (27)+(28-a) GOOD | Schw: predicted<br>Selk-like: still not predicted |
| (28-b) She saw her mother.<br>pa | (27)+(28-b) GOOD | Schw: predicted<br>Selk-like: still not predicted |

As already noted in section 4, the accented word is required to be E-marked, and hence must be interpreted as either **new** or the answer to the preceding question. However, it is neither.

All in all, I would say that the Selkirk-like theory is preferable to Schwarzschild's theory. It seems fair to say that the Selkirk-like theory is at least as empirically adequate as Schwarzschild's theory. As for explanatory adequacy, the Selkirk-like theory is definitely simpler (see section 5 regarding the complexity of Schwarzschild's theory), so it seems superior in that regard. Note also that my amendment to clause (B-ii) is very natural: Selkirk's original clause restricted interpreting the E feature (as marking **novelty**) to constituents contained within the focus; my clause restricts to constituents contained within the focus both interpreting the E feature (as marking **novelty**) and interpreting its absence (as marking **givenness**), so that the presence and absence of E are treated on a par.

## 7. Another look at the 'givenness' associated with accent placement

The notion **given** is akin to the 'anaphoric' of Kuno 1972, 'given' of Clark and Haviland 1977 and 'familiar' of Heim 1982 (that is, the  $given_k$  – 'shared knowledge' – of Prince 1981). Like them, it is a conception of 'given' as being in some sense 'already present in prior discourse', which is naturally formulated in terms of being **anaphoric** (requiring an antecedent): an expression is **given** iff it has an antecedent in prior context which entails it, so that the content of the expression must already be present in the context. I think it is fair to say that Schwarzschild's GIVEN is a refinement of the same idea: on Schwarzschild's theory, instead of an expression being 'given' only if there is an antecedent that entails it just as it is, it is enough that there be an antecedent that entails it as revised by replacing its R-marked constituents with variables.

(I should note the following. The way **given** and GIVEN are formulated by Schwarzschild, the antecedent is necessarily linguistic – an expression actually mentioned in prior discourse. But of course an approach often taken in the literature on anaphora is to only require an antecedent in the interlocutors' representation of the context of utterance, thus allowing antecedents that are not expressed linguistically. We should probably assume that the definition of **given** is relaxed in this manner. For instance, an expression will count as 'given' if there is an object that is salient in the context which can serve as its antecedent, or even if the interlocutors have accommodated a discourse referent which can serve as its antecedent.)

In this section, I would like to take a look at some further data, which are not covered by the theories discussed in the preceding sections, and which call into question interpreting the absence of accent as  $given_k$  (that is, as 'already in prior discourse', 'anaphoric', 'familiar', or some such thing).

Consider the following example. In the context of (58), (59-a) is felicitous. *Anna* is accented, despite the fact that it is **given**. In fact, *Anna* **must** be

accented in this context – again despite being **given**: (59-b) is not a felicitous response to (58).

(58) John and Anna walked in.  
And then what happened?

(59-a) Anna yawned. (58)+(59-a) GOOD  
pa pa

(59-b) Anna yawned. (58)+(59-b) BAD out of the blue  
pa

At this point two notes are in order. One: it can hardly be denied that *Anna* is salient and therefore **given** when (59) is uttered in the context of (58), given the felicity of (60)+(61) below.

(60) Anna walked in.  
And then what happened?

(61) She yawned.  
pa

Two: to be precise, (59-b) is not a felicitous response **out of the blue**. I think that (59-b) **can** felicitously answer (58) if it is clear in the context that attention is centered around Anna, and it is therefore expected that the answer will concern her. Still, I believe that in the absence of further context which would guarantee that, (59-b) cannot felicitously answer (58).

Schwarzschild's theory gets the wrong predictions entirely. (59-a) requires at least two R features, one on each accented word. But, in the context of (58), (59-b) can be represented with just one R. (59-b) only needs an R on the accented *yawned*. No R is required on the entire S, because the S is GIVEN: *John and Anna walked in* entails *Anna Xed*. But that means that AvoidR prefers (59-b) to (59-a). Hence, it is incorrectly predicted both that (59-b) is a felicitous response to (58) (in the absence of further context), and that (59-a) is not.

The Selkirk-like theory with **given** yields the same two incorrect predictions. This theory allows a representation of (59-b) which is compatible with a context consisting of (58) alone, namely, (62). So it predicts that (59-b) should be a felicitous response.

(62) [ Anna [yawned]<sub>E</sub> ]<sub>E</sub>  
pa

On the other hand, it allows no representation of (59-a) which is compatible with



equally good. So, the two responses are predicted to be felicitous. That's fine for (66-a), but not for (68-a). For the Selkirk-like theory, we have the following. The unaccented *singing* is definitely not E-marked. So, being contained inside a broad focus, it must be interpreted as **given**. But we said that it was **new**. So the two responses are predicted to be bad. That's correct for (68-a), but not for (66-a).

In short, Schwarzschild's theory and the Selkirk-like theory each make one incorrect prediction, and neither can explain the contrast between (66-a) and (68-a).

But wait, what if *singing* should, after all, count as **given** and GIVEN? It could be argued that the subject NP makes singing salient, and so the verb **is** entailed by something salient in the discourse. Well, maybe. But *singing* is just as salient in (68-a) as it is in (66-a), so it would have to be **given** and GIVEN in both responses. So again we'd get that the two older theories will not be able to explain the contrast between them; again each of these two theories will make one incorrect prediction. On Schwarzschild's theory, *singing*, being GIVEN, will have to carry an R only if it is accented. So AvoidR will actually prefer a response where it is not accented. Hence in this case too the two responses will be predicted to be good. On the Selkirk-like theory, *singing* still can't be E-marked, so it must still be interpreted as **given**. Which is compatible with it's being in fact **given**. So the two responses will now be predicted to be good.

#### Summary:

Assuming that *singing* is **new** and not GIVEN in both replies:

(65)+(66-a) GOOD	Schw: ✓
	Selk-like: ✗
(65)+(68-a) BAD	Schw: ✗
	Selk-like: ✓

Assuming that *singing* is **given** and GIVEN in both replies:

(65)+(66-a) GOOD	Schw: ✓
	Selk-like: ✓
(65)+(68-a) BAD	Schw: ✗
	Selk-like: ✗

And what is the moral of the story? Once more, the data suggest that when interpreting pitch accent placement, we need to replace  $\text{given}_k$  with a different notion of 'novelty'. Now, what other notion? The last examples are very suggestive. It is clear, intuitively, why it is that *singing* should be unaccented in (66) but not in (68). We know that Pavarotti is a famous singer, and therefore, if Pavarotti is on the show tonight, he might well be expected to sing. But if a

publisher of vocal music is on the show, we do not expect that she will be singing.

To my mind, this suggests rather strongly that what we need for the purpose of interpreting lack of accent is a notion of 'given' akin to the idea of 'predictable' of Kuno 1972 and 'recoverable' of Halliday 1967 (the given<sub>p</sub> of Prince 1981). The notion of 'predictable' or 'recoverable' is quite different from 'familiar' (given<sub>k</sub>), in that it concerns not the expression itself, but rather its place in the utterance. An expression is 'given' in this sense – given<sub>p</sub> – not when the expression itself (or a revision of it) is inferrable from prior context, but rather when its figuring in a certain role/position in the utterance is inferrable from prior context. And that is just what we see in (66)-(68): we don't need to know here if *singing* is a familiar or known thing. Rather, we need to know if *singing* is likely to be the action that the **subject of the sentence** will perform on the show. I suggest that what the lack of accent on *singing* in (66) means, roughly, is that if the hearer were faced with *Pavarotti is Xing*, she would be able to tell, at least as a reasonable guess, that X is supposed to be *singing*.<sup>1</sup>

A given/new<sub>p</sub> approach would seem to cover other cases quite nicely, too. Recall the case of (58)+(59-a) (the *Anna Yawned* example). *Anna* is accented, but clearly not because it is new<sub>k</sub>. Why is it accented then? I think it is accented because it is not given<sub>p</sub>. If you were faced, in the context of (58) alone, with *X yawned*, you would have no way of guessing that X is supposed to be *Anna*, since X would be at least as likely to be *John*, or *they* (i.e., John and Anna).

Further, the given/new<sub>p</sub> approach (i) immediately predicts the fact that expressions that are part of the question being addressed can never be accented, and (ii) has the advantage that it allows for a uniform analysis of the import of placing an accent on a word which is itself the focus (=the answer to the question) and on a word contained within a focus. Regarding (i), obviously, if you have asked me who ate the beans, then when I answer *Fred ate the beans*, *ate the beans* is as given<sub>p</sub> as it gets – even if I just answer *Fred*, you can't fail to figure out that I meant 'Fred ate the beans'. Regarding (ii), the claim is that in both of these cases, the presence of accent means the same thing: the word is new<sub>p</sub>. If the word itself is the focus – well, obviously, the focus itself is invariably (presented as) new<sub>p</sub>; we ask a question precisely because we don't know what the answer is. A word contained within a focus, on the other hand, can be either given<sub>p</sub> or new<sub>p</sub>; it is the way we place our pitch accents that indicates which it is (as in *Pavarotti is singing* vs., say, *Clinton is singing*).

pa

pa

pa

---

<sup>1</sup> I may add that Bolinger often informally attributes lack of accent to the word being "predictable", or to its containing "information already known, or information that can easily be guessed at". (see e.g. Bolinger 1986, pp.51-52)

What Halliday 1967, p.204, actually suggests is that **the focus** is 'new', in the sense that 'the speaker presents it as not being recoverable from the preceding discourse'. Note that while this **is** the sense of 'new' that I am after, I am not proposing to associate it with being a focus, but rather with bearing a pitch accent.<sup>2</sup> I also wish to emphasize the 'the speaker presents it as' bit, which I take very seriously. Accented words are being **presented** as given<sub>p</sub>.

#### **8. A new theory with RECOVERABLE:**

**No recursively projected or interpreted syntactic feature; no focus interpretation; an unaccented/accented word interpreted as RECOVERABLE/not RECOVERABLE**

It is now time for me to offer an alternative theory of the interpretation of pitch accent placement. I will propose a theory that differs from the theories of both Selkirk and Schwarzschild in two central respects. First of all, it is based – in the light of the previous section – on an alternative notion of 'givenness', corresponding not to given<sub>k</sub> but to given<sub>p</sub>. 'Givenness' in this sense does not involve an antecedent in prior discourse, and is not checked in relation to the context alone, but rather to the context and the rest of the utterance. Secondly, my theory does not involve any recursively projected or recursively interpreted syntactic feature; instead, accent placement is only interpreted at the level of the word.

Regarding the second point: On Selkirk's theory, there is a syntactic feature corresponding to accent which gets recursively 'projected upward'. On Schwarzschild's theory, the syntactic feature is not 'projected upward', but it does play a crucial role in the recursive definition of GIVEN, which serves to interpret (the absence of) that very feature. I am claiming now that it is also possible to have a viable theory of the interpretation of pitch accent placement which does not involve either the recursive projection or the recursive interpretation of a syntactic feature. The theory proposed below makes no use of any recursive clauses.

I say that accent placement is only interpreted at the level of the word. Well, I don't really mean 'word' too seriously. What I have in mind is more like 'word or other small item'. It can certainly be a syllable, or a compound. Maybe a direct quote, sometimes. (I briefly go back to this issue at the end of this section.)

I propose the theory given in (51).

---

<sup>2</sup> As discussed in Kadmon 2001, section 13.4, I believe that pitch accent placement does not serve to directly identify the focused constituent; rather, it helps to identify the focus indirectly, via the role of pitch accents in indicating the informational status of words. I am now adding the suggestion that the informational status in question is that of being new<sub>p</sub> vs. given<sub>p</sub>.

(51) **A new theory with RECOVERABLE:**

**(A) Interpretation of pitch accent placement:**

A word is interpreted as RECOVERABLE iff it is unaccented.

**(B) Definition of RECOVERABLE:**

An expression B is RECOVERABLE in utterance U iff the following holds.

Presented with the result of replacing B in U with a variable, it would be possible for the hearer to infer on the basis of prior context that in the actual utterance, the position of that variable should be occupied by B.

Recall that Halliday 1967, p.204, suggests that the focus is 'new' in the sense that 'the speaker presents it as not being recoverable from the preceding discourse'. It is this sense of 'new' that I am attempting to capture by my notion of 'not RECOVERABLE'. But, as already noted above, on my proposal, this sense of 'new' is applied not to the focus, but rather to accented words. Indeed, my current theory makes no use at all of the notion of 'focus' – in that respect, I am following Schwarzschild 1999.

The definition of RECOVERABLE in (B) talks of being able to **infer** that the position of the variable should be occupied by a certain expression. By 'infer' here I do not mean 'infer with absolute certainty'. Making an educated guess would do too. As will become clear below, this is quite crucial. It is good enough if the hearer can figure out that the position in question is **likely** to be occupied by that particular expression, rather than any other expression.

Let us look at the data, and see how the theory works. Consider (3)+(5).

(3) What did you do to Bill?

(5-a)	I invited Bill H* L L%	(3)+(5-a) GOOD	Schw: ✓ Selk-like: ✓ Kadmon: ✓
(5-b)	I invited Bill H* L L%	(3)+(5-b) BAD	Schw: ✓ Selk-like: ✓ Kadmon: ✓
(5-c)	I invited Bill H* H* L L%	(3)+(5-c) BAD	Schw: ✓ Selk-like: ✓ Kadmon: ✓

My theory correctly predicts that dialog (3)+(5-a) is felicitous. That is because in the context of (3), *I* and *Bill* are RECOVERABLE in (5), whereas

*invited* is not. To see that *I* is RECOVERABLE in (5), consider (52).

(52) X invited Bill.

A hearer faced with (52) in the context of (3) would be able to infer with great confidence that the position of X in the actual utterance should be occupied by *I*. This inference is based on the assumption that the speaker is cooperative, and intends to answer the last question under discussion in the discourse, viz., (3). For the actual utterance of (5) to answer (3), the position of X must be occupied by *I*. Similarly, *Bill* is RECOVERABLE in (5), because a hearer faced with (53) would be able to confidently infer that here X should be *Bill*.

(53) I invited X.

*Invited*, on the other hand, is not RECOVERABLE in (5). That is because the context we are assuming contains (besides general assumptions and world knowledge) only question (3), and there is nothing in this context that could tell the hearer that the position of X in (54) should be occupied by *invite*.

(54) I Xed Bill.

Indeed, (5) is addressed at the very person who has asked (3) – who has presumably asked it precisely because she doesn't know what X in (54) should be.

The remaining options, dialogs (3)+(5-b) and (3)+(5-c), are correctly predicted to be infelicitous: in the context of (3), *Bill* is RECOVERABLE in (5), as shown above – but that is incompatible with *Bill* being accented.

Now let us turn to (2)+(5).

(2) What did you do?

(5-c)	I invited Bill H* H* L L%	(2)+(5-c)	GOOD out of the blue	Schw: ✓ Selk-like: ✓ Kadmon: ✓
(5-b)	I invited Bill H* L L%	(2)+(5-b)	Kad: BAD out of the blue	Schw: ✗ Selk-like: ✓ Kadmon: ✓
			GOOD in context (party)	Schw: ✓ Selk-like: ✓ Kadmon: ✓
(5-a)	I invited Bill H* L L%	(2)+(5-a)	BAD out of the blue	Schw: ✗ Selk-like: ✓ Kadmon: ✓

(2)+(5-c) is correctly predicted to be felicitous in the absence of further context, because in the context of (2), *I* is RECOVERABLE in (5), whereas both *invited* and *Bill* are not. (5-b) and (5-a) are correctly predicted to be infelicitous out of the blue, because following just question (2) without any further context, we have here an unaccented word – *Bill*, or *invited* – which is not RECOVERABLE. If we take (5-b) in the context of party planning, on the other hand, the word *invited* becomes RECOVERABLE: if the hearer were faced with (54) above (*I Xed Bill*) it would be quite easy for her to infer that Xed should be *invited*. So (5-b) is correctly predicted to be good in such a context.

As for (35)+(5), the reader can verify for herself that the new theory again makes the correct predictions, in exactly the same manner.

(35) What did you do to who?

(5-a)	I invited Bill H* L L%	(35)+(5-a) BAD	Schw: ✗ (predicted with Foc) Selk-like: ✓ Kadmon: ✓
(5-b)	I invited Bill H* L L%	(35)+(5-b) BAD	Schw: ✗ (predicted with Foc) Selk-like: ✓ Kadmon: ✓
(5-c)	I invited Bill H* H* L L%	(35)+(5-c) GOOD	Schw: ✓ Selk-like: ✓ Kadmon: ✓

I go on to the **book-borrowing** example. My theory makes the right predictions here too.

(39) Who borrowed the book that Max had purchased?

(44) I'll answer your question: Max borrowed it.

(45) Ah, wait, but you know how Max really got this book?  
Max borrowed it.

(40-a)	Max borrowed it. pa	in (39)+(44): GOOD	Schw: ✓ Selk-like: ✓ Kadmon: ✓
		in (39)+(45): BAD	Schw: ✗ (✓assuming <i>borrowed</i> is no longer salient) Selk-like: ✓ Kadmon: ✓

(40-b)	Max borrowed it. pa	in (39)+(44): BAD	Schw: ✗ Selk-like: ✓ Kadmon: ✓
		in (39)+(45): GOOD	Schw: ✓ Selk-like: ✓ Kadmon: ✓

In the context of (39)+(44), *borrowed* and *it* are RECOVERABLE in (40), while *Max* is not. A hearer faced with (55) will confidently infer that X should be *it* (i.e., the book that Max had purchased), because otherwise (40) would fail to answer question (39), in contradiction with the speaker's pledge in (44).

(55) Max borrowed X

Similarly, a hearer faced with (56) will confidently infer that X should be *borrow*.

(56) Max Xed it.

But a hearer faced with (57) will have no way of telling that X should be *Max*. The whole purpose of (40) here is to inform the hearer of what X in (57) should be – that is what it is to answer (39).

(57) X borrowed it.

Hence, in the context of (39)+(44), (40-a), with just *Max* accented, is the only version of (40) which is felicitous.

In the context of (39)+(45), *Max* and *it* are RECOVERABLE in (40), while *borrowed* is not. A hearer faced with (55) above will know that X should be the book, because otherwise (40) would fail to answer the question that has just been raised – the question in (45). Similarly, a hearer faced with (57) will know that X should be *Max*. But a hearer faced with (56) will not be able to tell that X should be *borrow*. At least, that's what the speaker of (45) wishes to imply, or she wouldn't have raised the question. In sum, the speaker of (45) is presenting *borrowed* as the only thing that is not RECOVERABLE in (40). The only pronunciation compatible with that is (40-b).

What I have shown so far is that the new theory covers all the data that the Selkirk-like theory with **given** has been shown to cover.

Note (i) that on the new theory, there is a perfect 1-1 correspondence between accent and 'novelty', and (ii) that without bringing in a theoretical entity 'focus', the new theory does seem to capture the role that the question under

discussion plays in constraining pitch accent placement – in particular, in the case of (39)+(40).

Let us now turn to the new data presented in section 7 above. Recall dialog (58)+(59). In the context of (58), (59-a) is felicitous. In the same context, and in the absence of any further context (in particular, if it doesn't happen to be clear in the context that attention is centered around Anna), (59-b) is not felicitous. That is, the word *Anna* is accented – indeed, must be accented – despite the fact that it is clearly **given** (cf. (60)+(61) below).

(58) John and Anna walked in.  
And then what happened?

(59-a)	Anna yawned. pa pa	(58)+(59-a) GOOD	Schw: ✗ Selk-like: ✗ Kadmon: ✓
--------	-----------------------	------------------	--------------------------------------

(59-b)	Anna yawned. pa	(58)+(59-b) BAD out of the blue	Schw: ✗ Selk-like: ✗ Kadmon: ✓
--------	--------------------	---------------------------------	--------------------------------------

(58)+(59-b) GOOD in context where  
attention is clearly  
centered around Anna

(60) Anna walked in.  
And then what happened?

(61) She yawned.  
pa

As discussed in section 7, both Schwarzschild's theory and the Selkirk-like theory with **given** get the wrong predictions entirely. The following was also noted in section 7: the Selkirk-like theory was able to handle words that were accented despite being **given** in case the word was itself a focus (a maximal E-marked constituent), as in the **book-borrowing** example, by claiming that a pitch accent on a focus marks a different kind of 'novelty', and is 'exempt' from being interpreted as **new**; but this sort of explanation doesn't extend to examples like the present (58)+(59), where the accented but **given** word is clearly not a focus.

On the new theory given in this section, there is no reference to 'focus', and there are no words that are 'exempt' from being interpreted as 'new'. Instead, we have changed our approach to 'new' – we have replaced new<sub>k</sub> (**new**) with new<sub>p</sub> (not RECOVERABLE). This allows the new theory to handle not just (39)+(40), but also (58)+(59).

In the context of (58) alone, both *Anna* and *yawned* are not RECOVERABLE in (59), and it is therefore correctly predicted by the new theory that both these words can and must be accented (=that (58)+(59-a) is felicitous but (58)+(59-b) is not). That *yawned* is not RECOVERABLE is obvious – there is no way of telling that X in (63) should be *yawn*.

(63) Anna Xed.

*Anna* too is not RECOVERABLE, even though it has occurred in (58). Given (58) alone, there is no way of telling – not even as a reasonable guess – that X in (64) should be *Anna*.

(64) X yawned.

The question under discussion, *What happened?*, obviously does not give a clue to what X should be. How about the fact that *Anna* is **given** – that is, salient in the context? That too does not tell the hearer that X is *Anna*, or even that X is more likely to be *Anna* than anybody else. After all, X could equally well be *they* ('John and Anna') or *John*, since John is just as salient in the context as Anna is. In fact, the default case would be to retain the sentence-topic of the preceding statement – in our case, 'John and Anna'.

If, on the other hand, it happens to be clear in the context that attention is centered around Anna, and it is therefore expected that the answer will concern her, the new theory correctly predicts that *Anna* should **not** be accented: in such a context, *Anna* is RECOVERABLE.

Next, recall the following examples. If the TV show in question is a varied one (rather than devoted to singing), then (66-a) is a good response to (65), whereas (68-a) is not. ((68), but not (66), seems to require a pitch accent on *singing*.)

(65) What's happening on your TV show tonight?

(66-a) Pavarotti is singing.            (65)+(66-a) GOOD  
pa

(68-a) A very famous publisher of vocal music is singing.    (65)+(68-a) BAD  
pa pa pa pa pa

As shown in section 7, whether we consider *singing* **new** and not GIVEN or **given** and GIVEN, Schwarzschild's theory and the Selkirk-like theory with **given** are both unable to explain the contrast between these two responses, each making one incorrect prediction:

### Summary

Assuming that *singing* is **new** and not GIVEN in both replies:

(65)+(66-a) GOOD	Schw: ✓
	Selk-like: ✗
(65)+(68-a) BAD	Schw: ✗
	Selk-like: ✓

Assuming that *singing* is **given** and GIVEN in both replies:

(65)+(66-a) GOOD	Schw: ✓
	Selk-like: ✓
(65)+(68-a) BAD	Schw: ✗
	Selk-like: ✗

Now let us consider the RECOVERABILITY-based theory. It is easy to see that we get exactly the right predictions. If Pavarotti is on the show, we expect that he will be singing. But if a publisher of vocal music is on the show, we do not expect that she will be singing. Therefore, *singing* is RECOVERABLE in (66) but not in (68), and that is why it can be unaccented in (66) but not in (68).

I think that this pair of responses brings out quite nicely the advantage of having defined RECOVERABLE as a relation between a word, its context, and its utterance as a whole. As noted in section 7: we don't need to know if *singing* is a familiar or known thing; we need to know if it's likely to be the action that the **subject of the sentence** will perform.

In this section we have seen some empirical advantages of the RECOVERABILITY-based theory: my theory gets the right predictions with

- the **NP is singing** example, which both of the competing theories are unable to handle,
- the **Anna yawned** example, which again both of the competing theories are unable to handle,
- the **book-borrowing** example, which only the Selkirk-like theory but not Schwarzschild's theory can deal with, and
- (35)+(5) and (2)+(5), which again the Selkirk-like theory but not Schwarzschild's can deal with.

I think we can see already that my theory is more empirically adequate than the competing theories. There is one remaining example studied above that we haven't tried yet,

- the **her mother** example, which Schwarzschild's theory, but not the Selkirk-like theory can deal with.

In section 9.3 below, I will argue that my theory can explain this sort of example as well.

Note: Concerning my claim that accent placement is only interpreted at the level of the word, 'or other small item', an anonymous referee argued that it must be a phrase, sometimes, based on the possibility of responding to (59) with (--), with no pitch accents on *John and Anna*.

- (58) John and Anna walked in.  
And then what happened?  
(--) John and Anna yawned. (58)+(--)*GOOD* out of the blue  
pa

The referee said that the sentence topic of (--), and hence the RECOVERABLE part, is *John and Anna*, which is a phrase and not a single word. Well, I agree that the NP *John and Anna* is RECOVERABLE, and indeed I have defined RECOVERABLE for expressions, not words. Does that mean that the absence of pitch accents marks it as such? I don't know. Is there actually anything wrong with saying that the absence of accent on *John* marks *John* as RECOVERABLE and the absence of accent on *Anna* marks *Anna* as RECOVERABLE? So far, I haven't found an example that convinced me that there are cases where we must assume that accent placement marks a phrase (rather than a word or a compound).

## 9. In the realm of RECOVERABILITY

### 9.1 The Flexibility of RECOVERABILITY:

#### *RECOVERABILITY comes in degrees and gets manipulated by the speaker*

I say that if it would be reasonable for a hearer to guess that a certain word should indeed occur in the position that it in fact occupies in the utterance, then that word is RECOVERABLE; otherwise, it is not. But what is a reasonably good guess? How likely should the word be to occur in its position in the utterance, to count as RECOVERABLE? The extent to which a hearer would be able to tell that a particular position should be occupied by a particular word will vary, of course. So **RECOVERABILITY comes in degrees**.

This, I think, fits very nicely with the data. Consider our TV show again. Remember that it is by no means devoted to singing. We said that in response to question (65), it was very natural to utter (66-a). In contrast, (67-a) and (68-a) are both not good. And what about (69-a)? I think that (69-a) is a much more natural reply than (67-a) or (68-a) are, but also considerably less natural than (66-a) is. This can be explained based on the different degrees to which *singing* is RECOVERABLE in these examples.

- (65) What's happening on your TV show tonight?

(66-a) Pavarotti is singing. (65)+(66-a) GOOD  
pa

(67-a) Clinton is singing. (65)+(67-a) BAD  
pa

(68-a) A very famous publisher of vocal music is singing. (65)+(68-a) BAD  
pa pa pa pa pa

(69-a) Fred Astaire is singing. (65)+(69-a) SO-SO (MEDIUM)  
pa pa

Faced with (70), would a hearer be able to tell that X should be *sing*? Well, X is much more likely to be *sing* in (70) than in (71) or (72).

(70) Fred Astaire is Xing

(71) Clinton is Xing

(72) A very famous publisher of vocal music is Xing

And yet Fred Astaire is at least as famous as a dancer as he is as a singer. Therefore, one cannot tell that X in (70) should be *sing* rather than *dance*. I think that is the reason that (69-a) is not quite as natural as (66-a). Pavarotti is famous first and foremost as a singer. So one can reasonably guess – especially in the context of a show – that X in (73) should be *sing* rather than anything else.

(73) Pavarotti is Xing.

I have already had occasion to stress that linguistic material gets **presented** (by the speaker) as RECOVERABLE or not RECOVERABLE. But that means that the following is also a characteristic of RECOVERABILITY: when the status of a word as RECOVERABLE or not RECOVERABLE is not clear-cut (as, e.g., in the Fred Astaire case), **the speaker can choose how she would prefer to present it**, and place pitch accents accordingly.

I can choose to say (69-a), for instance, if I feel that dancing is not likely in the context of my TV show, or if I wish to create that impression. I can choose (66-c) in preference to (66-a), if I believe that my audience cannot be counted on to know that Pavarotti is a famous singer, or if I wish to imply that they can't, or if I think that it is not obvious that Pavarotti will be performing rather than just getting interviewed.

(66-c) Pavarotti is singing.  
pa pa

I can choose to say (67-a) if I wish to imply that on my show, everyone sings, or if I wish to imply that Clinton is in fact famous for singing. And so on and so forth.

We have seen then that there is some indeterminacy and flexibility to the notion RECOVERABLE, and I think we are also entitled to conclude that this is in fact a welcome feature of the proposed theory. The fact that I have left 'the hearer would be able to infer' vague enough to allow RECOVERABILITY to come **in degrees**, plus the idea that words are **presented** as RECOVERABLE or not RECOVERABLE, together account for the wealth of options that are in fact open to the speaker, and for subtle intuitive differences among these options. The theories based on **given** or GIVEN, in contrast, do not possess the necessary flexibility.

## **9.2 The determinants of RECOVERABILITY**

As should be clear from section 8, all sorts of factors go into determining RECOVERABILITY. These seem to fall into two main categories:

- world knowledge
- general principles that govern discourse

In this subsection I would like to explicitly list some of these factors, and to some extent also give justifications for them and clarify the relations among them.

As seen in the discussion in section 8, one central assumption that I am making is that when a word is part of the question being addressed, it is definitely RECOVERABLE – one can infer with full confidence that it, indeed, should be the word that occurs in its position in the utterance. This assumption allows us to account for a well known pattern, namely, that there are no pitch accents outside the focus, without having to have a special rule for that, and without even having to assume a theoretical notion of 'focus'.

From the speaker's point of view, she knows what question she is addressing. So all she has to make sure of is that no words in her utterance which correspond to that question will be accented. The hearer, on the other hand, relies on various clues to figure out what the question being addressed is. If the question has just been mentioned, well that's certainly a good clue. Assuming of course that general principles that govern discourse tell us that an utterance is likely to address a question that has just been uttered – which I think we are certainly entitled to assume. Another very useful clue is precisely the knowledge that no words that correspond to the question being addressed are accented. Further, world knowledge will often allow us, in a given context, to have a pretty good idea of which questions are likely to be addressed in the context and which are not.

When a word is **not** part of the question being addressed (=when it is part

of the focus, if you like), it can be either RECOVERABLE or not. World knowledge plays a central role in determining which it will be. For instance, I think it cannot be denied that it is world knowledge (knowing that Pavarotti is a famous singer) that is responsible for the contrast between (66-a) and (68-a). World knowledge tells us that if Pavarotti is on the show, it is likely and even expected that he will be singing. (We may note also that singing is salient in the context of the subject NP of (68-a) as well, but that is not enough to make singing RECOVERABLE there.)

(65) What's happening on your TV show tonight?

(66-a) Pavarotti is singing. (65)+(66-a) GOOD  
pa

(68-a) A very famous publisher of vocal music is singing. (65)+(68-a) BAD  
pa pa pa pa pa

Often, it is information given in the immediate context of the utterance (be it linguistically expressed or not) that plays the same role. For instance, it is already noted in Schmerling 1976 that (i) can have broad focus on the entire predicate NP, if spoken in the context of a hospital or medical convention, while at a high society cocktail party, (ii) would have to be used instead. What I am adding is that what the context determines is whether *doctor* is RECOVERABLE or not.

(i) This is the doctor I was telling you about.  
pa

(ii) This is the doctor I was telling you about.  
pa

Certainly world-knowledge and principles of discourse are different types of factors, but as they both affect accent placement, one may hope that their effects fall under a uniform principle that governs accent placement. I am suggesting here that there is such a uniform principle, that of RECOVERABILITY.

When discussing (58)+(59-b), repeated below, I said that in *Anna yawned*, the word *Anna* is RECOVERABLE, because there is nothing to hint to the hearer that X in (64) should be *Anna*,

(58) John and Anna walked in.  
And then what happened?

(59-b) Anna yawned. (58)+(59-b) BAD out of the blue Schw: X  
pa Selk-like: X



moving on a computer screen and were asked to describe series of four events happening on the screen. The subjects produced series of utterances such as the following. Note that there is a noun in the fourth utterance which is 'given<sub>k</sub>' – in the examples below, *ball* – call it the target noun. The question was under what circumstances the subjects would put a pitch accent on this noun. I have marked a strong tendency to accent the target noun by **pa**, and a strong tendency not to accent it by - - -.

- (01) 1. The cone touches the ball.  
2. The cross touches the ball.  
3. The diamond touches the ball.  
4. The star touches the ball.  
- - -
- (02) 1. The cone pushes the rectangle against the ball.  
2. The cross pushes the line against the ball.  
3. The diamond pushes the triangle against the ball.  
4. The box pushes the star against the ball.  
- - -
- (03) 1. The cone touches the ball.  
2. The cross touches the ball.  
3. The diamond touches the ball.  
4. The ball touches the box.  
**pa**
- (04) 1. The ball touches the cone.  
2. The ball touches the cross.  
3. The ball touches the triangle.  
4. The star touches the ball.  
**pa**
- (05) 1. The ball touches the cone.  
2. The ball touches the cross.  
3. The ball touches the triangle.  
4. The box pushes the star against the ball.  
**pa**

As you can see, the findings clearly indicate some kind of a 'parallelism' effect. Terken and Hirschberg's interpretation is that two factors are at work here: (i) whether or not the grammatical function (subject / direct object / indirect object) of the NP containing the target noun is kept the same as in the preceding utterances, and (ii) whether or not the surface position of the target noun is the same as in the preceding utterances. In examples like (01)-(02), where both its grammatical function and surface position are preserved, the target noun usually does not carry a pitch accent; in examples like (03)-(05), where neither is

preserved, it usually does carry a pitch accent. Terken and Hirschberg have some evidence that preservation of grammatical function and of surface position are two independent factors: in (06)-(07) below, where the surface position of the target noun is preserved (final position) but its grammatical function is not, they found a tendency not to accent the target noun, but this tendency was significantly less strong than the strong tendency found in examples (01)-(03).

- (06) 1. The cone touches the ball.  
2. The cross touches the ball.  
3. The diamond touches the ball.  
4. The box pushes the star against the ball.
- (07) 1. The cone pushes the rectangle against the ball.  
2. The cross pushes the line against the ball.  
3. The diamond pushes the triangle against the ball.  
4. The box touches the ball.

Terken and Hirschberg have not attempted to explain why it is that preservation of grammatical function and/or surface position should be relevant to whether or not an item would get accented. What I would like to suggest is that when an item's grammatical function and/or surface position is the same as in the preceding utterance(s), that makes this item more RECOVERABLE. For instance, it seems quite clear to me that given the first three utterances in (01), a hearer faced with (08) would be able to guess that X should be *ball*, but a hearer faced with (09) would not be able to tell that X should be *ball*, not even as a reasonable guess.

(08) The star touches the X.

(09) The X touches the box.

At the very least, I can say with confidence that RECOVERABILITY (givenness<sub>p</sub>) is much better suited to explicating the effect of parallelism on accent placement than givenness<sub>k</sub> is. After all, the target occurrences of *ball* are all given<sub>k</sub>, so that's obviously not a property that can distinguish between the accented and unaccented occurrences.

Of course, one may ask further why it is, exactly, that parallelism affects RECOVERABILITY. As my referee put it: "what kind of an assumption about discourse lies behind it? that speakers like to repeat words? repeat them in the same spot?" I don't know that I have the full answer to these questions, but I think it's a matter of continuing a pattern. Our minds automatically identify patterns in the world around us. The first three utterances in the series of the Terken and Hirschberg experiment obviously make a certain pattern very salient

– that of repeatedly placing the relevant noun in exactly the same grammatical function and surface position time after time. I think that by not accenting the target noun, the speaker signals to the hearer that she is thinking of her fourth utterance as continuing the same pattern. That's because by not accenting the target noun she presents it as RECOVERABLE, but the only way the accommodating hearer can indeed accept it as RECOVERABLE is by assuming that the same pattern is being continued, since this pattern is the only potential clue here to what word should occupy the target position. (I may note also that I think sometimes a single preceding utterance is enough to be thought of as starting a pattern which is being continued by the current utterance.)

I would like to end this section with a short discussion of the RECOVERABILITY of pronouns and how it relates to the role that they play in discourse.

Here is what seems to be a problem for my theory: in the context of (010), all of (012), (013) and (011) seem to be natural continuations. And yet, it seems that the pronoun is not RECOVERABLE, since there is no way of telling whether X in (014) should be Mary, or Bill, or Mary and Bill.

(010) Mary gave the briefcase to Bill  
And then what happened?

(012) We arrested her.  
pa

(013) We arrested him.  
pa

(011) We arrested them.  
pa

(014) We arrested X.

But I think these examples are actually compatible with my theory. Let me sketch a suggestion about what is going on here.

It is well known not only that pronouns refer to entities which are salient in the immediate discourse context, but also that they have a strong tendency to refer to that entity which is currently the most salient/central focus of attention. (In terms of the theory of Grosz and Sidner 1986, pronouns tend to be the 'backward looking center' of the current utterance, anaphorically related to the 'preferred center' of the preceding utterance. See e.g. Nakatani 1997.) That means that if we know which entity is the most salient/central in the preceding utterance, that would make the pronoun RECOVERABLE for us.

Consider the discourse contexts given in (015) and (016) below. I think that in each of these contexts, only one of (012') and (013') is a natural continuation. Note that I didn't mark any pitch accents in (012') and (013'). I think that even given the printed intonation-less (012') and (013'), the judgments are as marked.

(015) By Monday, Bill had already recieved the briefcase. Mary was really anxious that he should get rid of it, but he was taking his time. And then what happened?

(013') We arrested him. (015)+(013') GOOD

(012') We arrested her. (015)+(012') BAD

(016) Mary stole the briefcase. She couldn't think of anybody more trustworthy to give it to, so she finally gave it to Bill. She was really anxious to get rid of it, you see. And then what happened?

(012') We arrested her. (016)+(012') GOOD

(013') We arrested him. (016)+(013') BAD

What we see here is that in (015), Bill is clearly the main focus of attention, so a pronoun in the following utterance is much more likely to refer to him, whereas in (016), Mary is clearly the main focus of attention, so a pronoun in the following utterance is much more likely to refer to her.

But that also means that in the context of (015), the pronoun in (013') is RECOVERABLE, and in the context of (016), the pronoun in (012') is RECOVERABLE. (Following (015), one can tell that in *We arrested X*, X should be *her* (i.e., Mary), and following (016), one can tell that in *We arrested X*, X should be *him* (i.e., Bill).) Which, in turn, predicts – correctly – that following (015), the pronoun in (013') should not be accented (as in (013)), and following (016), the pronoun in (012') should not be accented (as in (012)).

(013) We arrested him. (015)+(013) GOOD  
pa

(012) We arrested her. (016)+(012) GOOD  
pa

Going back to (010), I think we can now explain why (013) and (012) are

both possible continuations in this context. Unlike (015) or (016), (010) is not detailed enough to determine which entity is the main focus of attention, so it does not determine which pronoun is the most likely to occur in the following utterance. But that leaves the speaker the freedom to choose between (013) and (012). Each of the continuations **presents** a different pronoun as being RECOVERABLE in the context. By uttering (013), for instance, a speaker signals that she is taking the unaccented pronoun to be RECOVERABLE, and therefore, that she considers Bill to be the main center of attention in the on-going discourse.

Finally, I would like to note that in her detailed study of a naturally-occurring piece of monolog, Nakatani 1997 found that the vast majority of unaccented pronouns referred to the main focus of attention in the immediate discourse context, whereas the majority of accented pronouns did not. Nakatani suggests, in fact, that the main role accented pronouns play in discourse is precisely to signal a **shift** in the main focus of attention. As an example, Nakatani gives the following excerpt from her data.

(017) So Masson became the new curator.

He flies to London and, you know,  
he's already met Anna Freud and therefore  
he has access to the secret cupboard of Freudian letters  
and naturally Anna assumed the uh

**SHE [H\*]** was a brilliant woman too –  
she did more a lot of work in child psy- psychiatry and  
psychoanalysis

assumed that **HE [H\*]** would keep this information  
you know within the confines of the psychoanalytical group

Well, as Masson was studying these letters he realized...

According to Nakatani, the two accented pronouns in boldface capitals signal such a shift: in the first bit of discourse, the main focus of attention is Masson. Anna Freud is introduced in connection with him, but without becoming the main focus of attention. Then starts an aside (in an embedded discourse segment) about Anna Freud. The accented pronoun *she* is said to signal that Anna Freud is now becoming the main focus of attention. Having completed the aside, the speaker resumes the main theme, i.e., returns to Masson. The accented pronoun *he* is said to signal that Masson is now becoming the main focus of attention once more.

I think that Nakatani's findings fit quite nicely with the RECOVERABILITY theory of pitch accent placement. I would say that the H\* pitch accent simply

marks the pronoun in question as not RECOVERABLE. This in turn signals to the hearer that the pronoun does not refer to the entity which is the main focus of attention in the immediate discourse context. I think it is reasonable to hypothesize that what signals to the hearer that the actual referent of the pronoun is now becoming the main focus of attention is not so much the pitch accent, but rather the fact that our pronoun (i) occurs after a pause or interjection suggesting the beginning of a new discourse segment, and (ii) occurs in subject position, which tends to host the sentence-topic (or 'backward looking center').

### 9.3 The *her mother* example

Let us finally return to the **her mother** example. I think that given the discussion of pronouns in the last section, my theory can be shown to account for these examples.

Recall that both (28-a) and (28-b) are felicitous responses to (27).

(27) John's mother saw Bill.  
And then what happened?

(28-a) She saw her mother.  
pa

(28-b) She saw her mother.  
pa

As seen in sections 4 and 5, Schwarzschild's theory predicts this correctly, while the Selkirk-like **given**-based theory incorrectly rules out both responses. On the **given**-based theory, the accented word is required to be E-marked, and hence be interpreted as either **new** or the answer to the preceding question; but in fact, it is neither – hence the incorrect prediction.

And how would the RECOVERABILITY-based theory deal with these examples? To start with, I think that there is no problem with the first two words. *She* is RECOVERABLE – the hearer would be able to guess that X in (74) below should be *she* (i.e., John's mother) – because *John's mother* was presumably the sentence-topic of the preceding statement and the most salient focus of attention.

(74) X saw her mother

*Saw* is RECOVERABLE – the hearer would be able to guess that X in (75) below should be *see* – due to parallelism (*saw* occurred in the same position in a sentence with the same structure) and perhaps also because *see* is a verb that is

commonly used to introduce a new individual into the discourse.

(75) She Xed her mother

But what about *her* and *mother*? Consider (28-b) first. Is *her* RECOVERABLE? Would the hearer be able to tell that X in (76) should be *her*? This is similar to the case of (010)+(012)/(013) considered in the preceding section. It looks like *her* is not RECOVERABLE, since (77) below is just as likely an answer to (27) as (28-b) is.

(28-b) She saw her mother.  
pa

(76) She saw X's mother

(77) She saw his mother.  
pa

But I think the speaker is free to **present** the pronoun, be it *her* or *his*, as RECOVERABLE. That is because (27) is not detailed enough to force the hearer to assume that John's mother is a more salient/central focus of attention than Bill is (cf. contexts (015) and (016) above, which single out the main focus of attention much more clearly). By choosing to say (28-b), the speaker obviously keeps talking about John's mother, and not about Bill. By her choice not to accent *her* in this utterance, she is signaling to the hearer that she considers John's mother the main focus of attention throughout (27)+(28-b).

Is *mother* in (28-b) RECOVERABLE? Would the hearer be able to tell that X in (79) should be *mother*? The answer seems to be no – (80) below is just as likely an answer to (27) as (28-b) is. So the pitch accent on *mother* in (28-b) is correctly predicted to be appropriate.

(79) She saw her X.

(80) She saw her sister.  
pa

Now consider (28-a). How can *mother* count as RECOVERABLE? We just saw that in the given context it does not seem to be RECOVERABLE.

(28-a) She saw her mother.  
pa

I think it is reasonable to say that the speaker of (28-a) **presents** *mother* as

RECOVERABLE because she is thinking of (28) as continuing the mother theme. It's as if she was prefacing (28-a) with (81). Certainly, in the context of (27)+(81), *mother* is RECOVERABLE.

(81) Ah, don't think we are through with people's mothers...

From the point of view of the hearer: A cooperative, accommodating hearer faced with (28-a) is forced to think of a way in which the unaccented *mother* could count as RECOVERABLE. Noting what *her mother* and *John's mother* have in common, the hearer is able to conclude that the mother theme is being continued. This account seems to fit with the intuitive impression that when (28-a) is pronounced, *her mother* is being contrasted with *John's mother*.

Finally, how about the accented pronoun in (28-a)? On my theory, the pitch accent is supposed to signal that this pronoun is not RECOVERABLE. I think that, indeed, the speaker is presenting this pronoun as not RECOVERABLE. Which in turn plays a role in making (28-a) sound like it is offering a surprising piece of information, and in drawing attention to the contrast between Bill's mother and John's mother's mother.

I think that the above discussion of (27)+(28-b) sheds some interesting light on how we manipulate pitch accent placement when we want to draw a contrast. What this example seems to show is that, at least in this case, the notion of RECOVERABILITY allows us to give some explication of the intuition that a contrast is being drawn, without having to maintain that the pitch accent itself directly marks contrast, and without using any theoretical notion corresponding to contrast. It would be interesting to study this matter further, and try to see how this kind of story might be worked out more precisely, and whether this kind of story would work for other cases of contrast as well. Could all contrasts be explicated in terms of RECOVERABILITY?

## 10. Beyond the realm of RECOVERABILITY

The theory that I proposed in section 8 claims that words are accented iff they are RECOVERABLE. The main purpose of the present section is just to come clean and admit that things are not quite that simple. I don't **really** believe that RECOVERABILITY can cover **all** the facts about pitch accent placement.

### 10.1 'highlighting'

Nobody would deny, would they, that pitch accents do regularly signal some sort of 'novelty'. I certainly believe that they do, and I do believe in particular that they regularly and systematically signal that the accented word is not RECOVERABLE. Yet it seems that they may also signal some other

thing(s). Dwight Bolinger's writings are very instructive in this regard. His work on the role(s) of prosodic prominences, over many years, strongly suggests that overall, the interpretation of pitch accent placement is not uniform, but rather varied and heterogeneous.

Bolinger 1986 regularly refers to the function of accenting as 'highlighting'. 'Highlighting' seems to me to be general enough to apply to all our (meaningful) choices regarding pitch accent placement. I think it makes good sense to say that whatever nuances it may express in addition, the function of accenting is always to 'highlight' – to attract the hearers' attention to something, to give (extra) emphasis to something, to signal that we attach importance to something.

One phenomenon that clearly does not fall under RECOVERABILITY is that of accenting a word not as a means of highlighting that word, but as a means of highlighting an entire sentence/utterance. Bolinger 1986 refers to this phenomenon with the term 'accents of **power**'. He discusses 'accents of power' of two main sorts: (i) a pitch accent placed at (or near) the edge of a sentence/utterance may serve to highlight that entire sentence/utterance; (ii) spreading a larger number of pitch accents throughout an utterance may also serve to highlight that entire utterance. Here are a few examples. (Note that sometimes the two above sorts are used simultaneously, other times just one of them. In example [4], we keep the final word unaccented to play down the utterance as a whole.)

[1] Suppose you ask me whether John swore on a particular occasion, and I reply

He not only swore...  
pa pa

with initial accent on *not*. This gives a matter-of-fact introduction to what I am about to add. But if I say

He not only swore...  
pa pa

you know that I am about to impart the information as something out of the ordinary. The accent on *he* cannot be there because of any special focus on the person (as in *Hé won't but i will*), but is there for the sake of the utterance as a whole. (Bolinger 1986, p.49)

[2] Suppose I ask you *How big is the portrait?* and you answer

It's life size.  
pa

This would be matter-of-fact giving of information. But now suppose that you do not know how big it is, and you look at it to see, and are surprised at how big it is

– you will probably say

It's life size. (Bolinger 1986, p.48)  
pa pa

[3] Excuse me! (Bolinger 1986, p.84)  
pa pa

[4] [...] a terminal accent may be inappropriate because it seems to give importance to something unimportant. Given the triviality of the state of one's beard, it probably sounds better to say

This is the most hair I've had on my face for a LONG time. (p.81)

Let us now return to placing a pitch accent on a word (or other small item) in order to highlight that word (or item) itself. Bolinger agrees with the view that when we accent an item to highlight it, we often do that in order to draw attention to new information, to something not previously in the hearer's awareness. He calls it 'accenting to show informativeness'. I think the theory I propose in this paper can be seen as a theory of what 'accenting to show informativeness' amounts to – in my opinion, we accent to draw attention to what is not RECOVERABLE. (And, indeed, Bolinger routinely uses the word 'recoverable' to describe those words that we choose not to accent.)

At the same time, Bolinger also firmly declares that showing informativeness is not the only reason to highlight a word – he says that it might be the reason in "the majority of instances", but not all of them. For that reason, Bolinger refers to cases of accenting an item to highlight that item as 'accents of **interest**' – a term general enough to cover various sorts of reasons for highlighting an item. Here are examples that Bolinger cites as 'accents of interest' not meant to show informativeness (or lack of accent not meant to show lack of informativeness).

[1] Raw fish is good for you, but after all, who likes raw FISH?

– Bolinger notes that the final word *fish* is certainly not 'new' or 'unexpected', and suggests that repeating and accenting this word "serves a kind of pointing function", akin to "holding one's hands up in mock horror at contemplating the object". (Bolinger 1986, p.90)

[2] John left early.

Why did he leave?  
pa

Why did he leave?  
pa            pa

The latter question shows a focus of interest on leaving (why did he leave, when I expected him to stay?), whereas the former implies 'That's all I need to know', [...]. (Bolinger 1986, p.96)

– We may well choose to accent the word *leave*, even though it is clearly not new<sub>p</sub> (or new<sub>k</sub>, either).

[3] I have to hurry along. My COUSIN'S having breakfast with us.

– The accent-less bit following the nuclear accent is not RECOVERABLE, or 'old' in any sense, but it is relatively unimportant: the hearer is expected "to accept the mere presence of *cousin* as explaining the haste; it could be any other action that puts him on the scene." (Bolinger 1986, p.114)

I agree then: not all 'accents of interest' fall under RECOVERABILITY.

## 10.2 Arbitrary patterns in pitch accent placement

Furthermore, pitch accent placement involves not only 'meaningful' choices. It also involves factors that don't have to do with (semantic or pragmatic) interpretation, but with other things entirely.

For one thing, considerations that have to do with rhythm also affect our choices in placing pitch accents, even though (in English, in prose) they tend to 'yield to' considerations of highlighting vs. playing down. (See Bolinger 1986, ch.5.) Everything else being equal, we tend to strive for rhythmic regularity. And we may sometimes even exploit rhythm for its own sake (as we regularly do in verse).

And there may even be structural constraints or structure-driven tendencies. Ladd 1980 does not posit any very general or very dominant structural constraint on pitch accent placement. At the same time, he allows for the possibility that alongside factors that involve the interpretation of pitch accent placement in context, there are also more arbitrary principles having to do with linguistic structure and/or the lexicon which also affect pitch accent placement; for example, the principle that (in English) nouns are more accentable than words of other syntactic categories, or the principle that makes the left-hand side of a compound more accentable than the right.

Allowing structural grammatical principles some role in pitch accent placement gives us one way of explaining the impression that certain dominant patterns and strong tendencies in pitch accent placement are observed which are



## 11. Conclusion

In this paper, I have been concerned with the interpretation of pitch accent placement and its effect on discourse congruence, and considered three main options for a theory: the theory of Schwarzschild 1999, a theory based on Selkirk 1996 and the notion of **given** (a notion offered but dismissed in Schwarzschild 1999), and a new theory of my own.

An advantage of Schwarzschild's theory is that it can automatically handle the **her mother** example (discussed in sections 4, 5, and 9.3). However, this theory runs into the two types of empirical problems discussed in section 4: failure to disallow unaccented words that are 'new', and failure to sometimes allow – indeed, require – a word mentioned in preceding discourse to be accented, because it is the focus (=the answer to the question).

It has been shown that the Selkirk-like theory with **given** of section 6, while unable to handle the **her mother** example, does solve the two types of problems just mentioned. We may therefore conclude, I think, that the Selkirk-like theory with **given** is at least as empirically adequate as Schwarzschild's theory.

The above results suggest that the intricate relation between pitch accents, the posited syntactic feature and semantic/pragmatic interpretation which is seen in Schwarzschild's theory is not necessarily advantageous. These results argue in favor of a theory which employs a simpler, more straightforward notion of 'given', and where there is more of a 1-1 correspondence between accent and 'novelty'.

It has been shown that the new theory proposed in section 8 covers all the data which the Selkirk-like theory was seen to cover (in section 6), and more. The new theory is better able to cover words that are unaccented despite being **new**: it allows *singing* to be unaccented in *Pavarotti is singing*. The new theory is also better able to cover words that are accented despite being **given**: of the three theories, it is the only one that can handle the **Anna yawned** example.

The new theory captures the way in which the informational status of a word is dependent on its relation to the rest of its utterance: of the three theories, it is the only one that can explain why while *singing* can easily be unaccented in *Pavarotti is singing*, it should be accented in *A very famous publisher of vocal music is singing*. Finally, the new theory is the only one of the three theories that is able to account for the actual wealth of options open to the speaker, and for subtle intuitive differences among these options (see section 9.1).

I think we may safely conclude that the new theory is the most empirically adequate of the three theories here considered.

I have also discussed some experimental findings regarding pitch accent placement. As shown in sections 9.2 and 9.3, these findings are compatible with my theory, and even support it. We saw that Terken and Hirschberg 1994's findings, which demonstrate a 'parallelism' effect influencing accent placement and suggest that this effect involves both the preservation of grammatical function and the preservation of surface position, can be explained in a uniform way if we make the natural assumption that such factors are determinants of RECOVERABILITY. We also looked at Nakatani 1997's finding that the vast majority of unaccented pronouns refer to the main focus of attention in the immediate discourse context, while the majority of accented pronouns do not. We saw that this finding enables us to explain why pronouns can be unaccented even when they seem, on the face of it, not to be RECOVERABLE (as in the **her mother** example): a speaker can sometimes choose to present a pronoun as RECOVERABLE in order to signal that its referent is the main focus of attention in the preceding (as well as the current) utterance.

The success of my new theory shows, first and foremost, that in treating the interpretation of accent placement, we should use a notion of 'given' corresponding to Kuno's 'predictable' and Halliday's 'recoverable' (=Prince's given<sub>p</sub>), rather than a notion more closely related to Prince's given<sub>k</sub> ('shared knowledge'). We need a notion of 'given' that measures how 'informative' a word is, and does it by relating it not just to preceding context, but also to the rest of its own utterance.

In addition, the success of the new theory supports Schwarzschild's idea that it is possible to have a viable theory of the interpretation of accent placement which does not make reference to a theoretical notion of 'focus'.<sup>3</sup>

The success of the new theory also shows that it is possible to have a viable theory of the interpretation of accent placement which does not employ either recursive projection or recursive interpretation of a syntactic feature corresponding to accent. In other words, we do not necessarily have to assume that the marking of a word by a pitch accent affects the interpretation of larger, containing constituents; it seems good enough to interpret accent placement just at the level of the word.<sup>4</sup>

---

<sup>3</sup> This does not necessarily mean that linguistic theory as a whole does not need to make reference to 'focus'. That, as noted in Schwarzschild 1999, would have to be determined by further research.

<sup>4</sup> And possibly also other small items – this would have to be studied further.

Finally, the present results suggest that a certain amount of vagueness and flexibility in the notion of 'given' (and in particular, the vagueness of 'the hearer would be able to infer' in the definition of RECOVERABLE) is a welcome feature in the interpretation of accent placement.

### Acknowledgements

An earlier version of this paper was submitted to NALS in 2000, and refereed. However, for health and family reasons, it was not immediately revised for publication. I have finally finished revising it now, and I would like to finally express my gratitude to the anonymous referee, for detailed and very helpful comments and suggestions. Versions of this paper were presented at CSSP 2003 (invited talk), and also at colloquia at the Ben Gurion University of the Negev and Ohio State University (2006), as well as at a number of graduate seminars held at Tel Aviv University. I am grateful to the audiences for their comments. I am also grateful to Craige Roberts for further comments and discussions, which were very helpful to me. Finally, I dedicate this paper to Ali Al-Azhari, Yasmin Azhari-Kadmon, and Ruth Ben-Gera.

### References

- Beckman, Mary E. 1996: 'The parsing of prosody'. *Language and Cognitive Processes* 11 (1/2), 17–67.
- Bolinger, Dwight 1972: 'Accent is predictable (if you are a mind-reader)'. *Language* 48.3, 633-644.
- Bolinger, Dwight 1986: *Intonation and Its Parts*, Stanford University Press.
- Chomsky, Noam 1971: 'Deep structure, surface structure and semantics interpretation'. In D. Steinberg and L. Jakobovits (eds.), *Semantics*, Cambridge University Press, 183-216.
- Clark, H. and S. Haviland 1977: 'Comprehension and the given-new contrast'. In R. Freedle (ed.), *Discourse Production and Comprehension*, Ablex, 1-40.
- Giora, Rachel 1983: 'Segmentation and segment cohesion: on the thematic organization of the text'. *Text* 3.2, 155-181.
- Grosz, Barbara J. and Candace L. Sidner 1986: 'Attention, intentions, and the structure of discourse'. *Computational Linguistics* 12.3, 175-205.

- Gussenhoven, Carlos 1984: 'Testing the reality of focus domain'. *Language and Speech* 26, 61-80.
- Halliday, M.A.K. 1967: 'Notes on transitivity and theme in English, part 2'. *Journal of Linguistics* 3.2, 199-244.
- 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.
- Jackendoff, Ray S. 1972: *Semantic Interpretation in Generative Grammar*, MIT Press.
- Kadmon, Nirit 2001: *Formal Pragmatics: Semantics, Pragmatics, Presupposition, and Focus*, Blackwell
- 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.
- Ladd, D. Robert 1990: 'Intonation: Emotion vs. grammar' (review of Bolinger 1989: *Intonation and its uses*, Stanford University Press), *Language* 66.4, 806-816.
- Nakatani, Christine 1997: 'Integrating prosodic and discourse modeling.' In Sagisaka, Campbell and Higuchi (eds.) *Computing Prosody: Computational Models for Processing Spontaneous Speech*, Springer-Verlag
- Pierrehumbert, Janet 1980: *The Phonology and Phonetics of English Intonation*, Ph.D. dissertation, MIT.
- 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 University Working Papers in Linguistics*, volume 49.
- Rochemont, Michael S. 1986: *Focus in Generative Grammar*, John Benjamins.

- 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.
- Terken, J. & Hirschberg, J. 1994: 'Deaccentuation of words representing given information: Effects of persistence of grammatical role and surface position', *Language and Speech* 37, 125-145.