The communicative significance of primary and secondary accents

David Beaver and Dan Velleman
The University of Texas at Austin

Submitted for review April 2010

Abstract

We offer an account of the function and meaning of primary and secondary accenting in English, including archetypal examples of so-called "second occurrence focus". The model is based on two factors, first *predictability*, which is a generalization of Given/New, and second *pragmatic importance*, which includes various phenomena for which semanticists typically invoke so-called *f-marking*, e.g. question-answering constituents and contrastive constituents. We show that relative acoustic prominence can be predicted on the basis of a summation of unpredictability and importance, and argue that the model is empirically superior both to the quite similar model proposed by Selkirk (2002,2007), and to the architecturally quite different models of Büring (2008) and Rooth (2010), models in which relative acoustic prominence is predicted on the basis of *domain size*.

KEYWORDS: *phonology, semantics, pragmatics, intonation, prosody, focus, accent, stress, second occurrence focus, secondary accent, nuclear accent, predictability, given/new*

1 Introduction

In English, expressions are made acoustically prominent using a mixture of intonational cues, notably pitch accents. But what is the semantic and pragmatic function of these accents? Much of the semantic literature on English intonation models that function using the notion of *focus* marking, with some semanticists also invoking additional notions of *contrast* or *topic* marking. Such categorizations seem to make cross-linguistic sense, because some other languages, e.g. Hungarian and Catalan (Vallduví and Engdahl, 1996), mark something like *focus* and *topic* or *contrastive topic* explicitly in the syntax.

And yet, while there is a compelling parallel between intonational features of English and certain syntactic features of other languages, the parallel casts light on only a tiny portion of intonational meaning in English. For the analogy with languages that syntactically mark focus and topic would suggest that utterances of English sentences typically contain only one or two acoustically prominent expressions. But studies of corpora of spontaneous speech reveal that it is common for even relatively short sentences to contain many more than two accented elements. In 31 telephone conversations drawn from the Switchboard corpus (Godfrey et al., 1992), Calhoun et al. (pear) found that 40.9% of all words were accented. So it seems unlikely that the distribution of accents in English could be reduced to the same factors that predict syntactic focus and topic marking in languages like Hungarian and Catalan.[1]

The notion of *focus* has long been viewed as covering a broad range of phenomena: Rooth (1992) takes focus to subsume marking of answers to questions, marking of contrast, marking of expressions that associate with focus sensitive particles, and marking of alternatives made salient to generate scalar implicatures. Furthermore, some scholars (e.g. Krifka 2008) argue that the notion of focus at least partly subsumes the notion of topic, based on the idea that both involve *alternatives*: a topic is a thing being talked about, and a topic is marked when it represents an alternative to something else being talked about.

We will say more about semantics based on alternatives below, but for the moment let us ignore how exactly the details could be worked out, and just allow that focus is a broad notion that covers not only question-answers, contrast, association with focus, and scalar implicatures triggered by focus, but also topics. Is that broad enough? Well, even if focus is taken to cover all these information structural categories, it still remains difficult to account adequately for the entirety of accenting that is found in English utterances. At an intuitive level, much material is accented because it is in some sense *new*, or, as we prefer, *unpredictable*: one might speculate that in such cases accenting makes comprehension easier in part by drawing the hearer's attention to the pieces of the utterance which require the most processing.

Can something like an existing notion of focus, say that of Rooth (1992), cover the way in which the newness or unpredictability of material relates

[1]This is not to claim that in Hungarian or Catalan information structure is marked only via syntactic position. Hungarian, for example, is well-known to have (post-verbal) prosodic foci in addition to the better studied (preverbal) positional focus.

to accenting, or is something else needed in addition to standard focus? We can divide scholars who've provided answers to that question into two camps, *lumpers* and *splitters*. Lumpers include Rooth 2010 and Büring 2008: both Rooth and Büring claim that if the notion of focus is spelt out properly, then the theory of accent will not need to make any further reference to newness of information. Selkirk (2002, 2007) and Beaver et al. (2007) are examples of splitters, presenting explanations of accenting which make separate reference to a standard notion of focus, and to newness or givenness of information. Below, we present data that invalidates a few details of Selkirk's approach, but that confirms her basic premise — that a two-pronged approach to the phenomenon is required.[2]

The crucial examples, for Selkirk's two factor account and the rival one-factor accounts alike, have to do with phenomena that we will collectively label *secondary accent*, and which include what has been widely discussed under the heading *second-occurrence focus*. Secondary accent is what happens when (i) a group of words form a single intonational unit (in simple cases, an entire sentence), (ii) one word is the most prominent, and carries a so-called *nuclear accent*, which we could also term the *primary phrasal stress*, but (iii) there is at least one more word which while less prominent than the one which carries primary phrasal stress, is nonetheless more acoustically prominent than the remaining words in the sentence. What is standardly termed *second-occurrence focus* involves secondary accent falling on a focused word or phrase — as in the classic example (1), adapted from Partee (1999).[3] It has been considered especially puzzling because foci almost always bear primary accent, and describing the precise conditions under which it occurs is a major goal for modern theories of focus prosody.

(1) **A:** Everyone already knew that Mary only eats VEGETABLES.
B: If even PAUL knew that Mary only eats vegetables // then he should have suggested a different RESTAURANT.

[2]Much of the literature is based on the use of syntactic markers. Let us note in passing a distinction: when one talks of focus marking, one might refer to the way in which intonation marks information status, or one might refer to some type of marker in an abstract level of representation, e.g. a syntactic marker. Much of the literature on focus invokes abstract syntactic focus markers. In this sense, we are not arguing for a two-*marker* account, but for a two-*factor* account — one which takes both predictability and importance into account, and which allows them to interact as Selkirk's account does. However, in Section 3 we will present an implementation of our account which, like most other accounts, is defined in terms of syntactic markers.

[3]Note here that achieving terminological neutrality is not entirely trivial. For lumpers, focus marking is the sole source of accenting, and thus the terms *secondary focus* and *secondary accent* could be used to describe a similar range of phenomena. But for us, the term *secondary accent* is more general that *secondary focus*, since we do not attribute all accenting (primary or secondary) to focus-marking.

2 Introducing predictability and importance

Our view is that Selkirk is correct in her suggestion that two factors are required to predict English prosody, but slightly mistaken about what the proper factors are. In particular, where she uses a straightforward form of Prince's (1981) given/new distinction as her second factor, we use the *predictability* of a constituent. Predictability is related to givenness, insofar as givenness in Prince's sense of the word can contribute to a constituent's predictability, but it is not exactly the same thing. Take the juicy piece of second-hand gossip in (2) — the juiciest piece of which was unfortunately inaudible to our crack team of eavesdropping field workers. All five partygoers are given after the first sentence of the discourse. And yet this does not seem to enable us to guess at the identity of the drunkest partier. If we had heard the drunkest partier's name, whoever it was, it would have been unpredictable to us — just as unpredictable, in fact, as the exact nature of the unspeakable act performed by Larry in (3). The point here is that while often something unpredictable is discourse new in Prince's sense (as "threw up on the hostess" would be if filling in the blank in (3)), sometimes something unpredictable is discourse given (as *Larry* would be if filling in the blank in (2)).[4]

(2) Gary, Larry, Harry, Barry and Mary all showed up at the party. And you *won't believe* who got the drunkest. It was _____!

(3) Gary, Larry, Harry, Barry and Mary all showed up at the party. And you *won't believe* what Larry did. He _____!

Of course, the above just demonstrates that predictability is different from givenness. It does not provide a precise definition of predictability, or a reason to believe that it will lead to better results than givenness as a factor in an account of prosody. But we will get to that soon enough. The point for now is that givenness and predictability are not identical; the choice between Selkirk's theory which uses givenness and ours which uses predictability is therefore nontrivial.

The roles which predictability and focus play in our theory can be understood in terms of the following principles:

(4) **Prominence Principle:** If one expression is more *communicatively significant* than another expression, then the first should be more *surface prominent* than the second.

(5) **Predictability and focus:** An expression's communicative significance is based on a number of different factors. We group these factors into two broad constellations: those having to do with the *predictability* of the expression, and those which determine whether or not it is *focused*.

[4]The question "How easily could a reader/hearer fill in this blank?" will be familiar to some readers from the cloze test — used in language learning to gauge the readability of a passage, and in psycholinguistics to ensure that sets of experimental stimuli will be equally surprising or unsurprising for subjects.

(6) **Competition for prominence:** Some forms of prominence are restricted in their distribution. Primary accent, for instance, can appear only once in an English intonational phrase.[5] It follows from this (with the important caveat that principles which determine phrasing must also be spelt out) and the two preceding principles that expressions within an intonational phrase will have to compete for primary accent. This sort of competition is what determines how informationally complex sentences will be realized.

We assume that predictability and focus contribute equal amounts of communicative significance to an expression, and that they are cumulative in their effect. You can think of unpredictability as conferring one point, and focus as conferring one point, such that expressions which are unpredictable and focused have a cumulative score of two.

It follows from these principles that a two-point expression will always bear primary accent. On the other hand, a one-point expression can only bear primary accent if it doesn't have any two-point competitors; it will fall back to secondary accent if it does. In particular, a focused-but-predictable expression (one point!) can be forced down to secondary accent if it has a focused-and-unpredictable competitor (two points!). This, we argue, is the cause of second-occurrence focus.

There is one slightly odd thing about the principles as we have stated them. It seems intuitively clear that an unpredictable constituent ought to count as more communicatively significant than a predictable one. An unpredictable constituent conveys more information than a predictable one; and the point (or at least *a* point) of communication is to convey information. But it is less clear why focus, out of all semantic or pragmatic phenomena, should be singled out as a source of communicative significance. Why not stipulate that definite nouns are more significant than indefinites, or that present-tense verbs are more significant than past-tense verbs? Is there something *ad hoc*, or even circular, about the claim that focus-marked constituents are communicatively significant?

The point here is that while in the simple formalization below we will simply assume focus-marking, there are in fact underlying reasons why constituents are focus-marked. For all the cases we will consider in the main body of the paper, the focus-marked constituents are thus marked because they play one of two discourse roles: answering a question, or marking parallelism. These functions we take to be of *pragmatic importance* in helping the speaker reach his goals and communicate effectively. Thus, more generally, rather than

[5]We are making the following hopefully uncontroversial assumptions here about English prosody: Utterances are split into intonational phrases. Each intonational phrase contains a primary (or *nuclear*) accent, which is realized with some combination of pitch accent, increased intensity and increased duration on the word that bears it. Intonational phrases may also contain secondary (or *non-nuclear*) accents. A secondary accent that falls before the primary one in its phrase (a *pre-nuclear* accent) may be realized as a minor pitch accent. One which falls after the primary accent in its phrase (a *post-nuclear* accent) does not exhibit a major pitch movement, and is typically realized with an increase in intensity and/or duration.

stating that a constituent is communicatively significant if it is unpredictable or focused. we want to say that a constituent is communicatively significant if it is unpredictable or (pragmatically) important.

Answers are important for obvious reasons: when the speaker answers a question which has been raised in the current discourse, then we can take answering the question to have been one of his communicative goals.[6] The answering constituent is instrumental in meeting this inferred goal; we can thus assign it more importance than it would have in a context where it did not answer any questions.

Note here that Beaver and Clark (2008), building on Roberts (1996) and others, analyze association with focus in terms of question answering. Suppose that some accented expression associates with a focus sensitive particle. In such a case, Beaver and Clark analyze the accented expression as being the answering constituent for the Question Under Discussion, and argue that that the focus sensitive particle, as part of its conventionalized meaning, provides a comment on that question. For example, in "Mary only ate CHEESE", the Question Under Discussion would be paraphrasable as "What did Mary eat?", and the exclusive "only" would rule out every answer to the question except the answer "Mary ate cheese". Returning to the current paper, the relevant consequence of the Beaver and Clark model is that all expressions that associate with (conventionally) focus sensitive particles are important, because they provide answers to the Question Under Discussion.

Let us turn briefly to contrast. We do not believe that there is any obligation to mark contrasts — but when a speaker does emphasize a contrast, or make a correction, we take this to have been one of his communicative goals, leading the contrasted constituents to be more important than they would have been otherwise. It should be noted here that contrast-marking could be analyzed in terms of predictability rather than importance. The point would be that a contrastive element is usually the least predictable element within some string. For example, in "Jane likes MARY and she likes JOAN", the word "JOAN" is the least predictable in "she likes JOAN". There are, in fact, at least three obvious ways in which "JOAN" might be communicatively significant: it is the single semantically non-parallel part of a parallel structure, it might be part of the answer to a Question under Discussion, and it is unpredictable. More generally, contrastive constituents will often be both important and unpredictable.

We acknowledge that there may be discourse goals, even quite general and widely held ones,[7] beyond those we have specified (answering questions, and marking contrast/parallelism. We only claim that these goals can be reliably imputed — that for instance when someone gives a direct answer to a question,

[6]Note that this does not require us to assume cooperativity on the part of all speakers at all times. If the speaker does answer, we infer that the answer furthers *some* goal of his. Perhaps the goal is to help the hearer; but perhaps the goal is to decieve or annoy, or to insult the hearer by giving a blatantly unhelpful answer — and since we only make these inferences in cases where the speaker *does* answer the question, this leaves outright uncooperative refusal to answer as another possibility.

[7]"Be polite," or at least "be likeable," is a likely candidate, but it is unclear that this ever confers importance on a specific constituent.

we can infer that answering that question was a goal of theirs. These reliably imputed goals give us a leg to stand on in talking about the general role of importance in language without worrying about goals which vary from one speaker to the next.

We can now restate (5) in terms of importance, as follows:

(7) **Predictability and importance:** An expression's communicative significance is based on a number of different factors. We group these factors into two broad constellations: those having to do with the *predictability* of the expression, and those having to do with its *importance* in meeting the speaker's goals.

If we restrict ourselves to considering the sort of discourse goals mentioned above, this principle will generate the same predictions as (5). If we branch out to consider a wider set of goals, it will predict additional prominence for some constituents, based on their importance to some real-world goal or preference of the speaker's. There is, for what it's worth, some evidence that this occurs. Watson et al. (2008) show in an experiment based on a game of tic-tac-toe that when game-winning moves are called out by speakers, these moves bear a special form of acoustic prominence. This could plausibly be argued to be the result of the importance of these moves in meeting a real-world goal of the speaker's — namely, winning the game of tic-tac-toe. It is also suggestive that the emotional weight a speaker attaches to different expressions can affect their prosodic realization. One could imagine discussing the prosodic effects of emotion in terms of some notion of emotional importance. For now we will restrict ourselves to the sort of discourse importance that is essentially equivalent to focus — but we will take up these more speculative subjects again towards the end of the paper.

There is similar variability in what can be called predictable, once again including a large number of idiosyncratic factors. Indeed, whenever speaker and hearer have any common ground whatsoever — any special mutual knowledge which isn't simply shared by all competent speakers and hearers — there will be idiosyncratic factors arising from it that affect what they will consider predictable. Once again, though, there are some factors which we maintain will reliably affect predictability; because of these, we can talk about the role of predictability in a general way, without getting hung up on interpersonal differences. In particular, we maintain the following:

(8) **New referents are unpredictable:** This is fairly straightforward — referring expressions whose referents are new in Prince's (1981) sense of the word are unpredictable.

(9) **Old referents in new roles are unpredictable:** Referring expressions with given referents may still be unpredictable if there is a lack of parallelism between occurrences. For instance, if *John loves Mary* is in the context (and nothing else is), then the occurrence of *John* in *John loves somebody* is predictable, but its occurrence in *Somebody loves John* is un-

predictable.

(9) captures the unpredictability of the last word in (2). For while we've introduced all the partygoers as discourse referents, we haven't talked at all about the question of who got the drunkest. Thus, anyone filling that role will be doing it for the first time, and will be equally unpredictable in it. Compare (10), in which it seems like a sure bet that we can fill in the blank with Gary's name. Certainly there are other factors here that guide us towards this conclusion, but one factor — and the one that will prove to matter most in our discussion of secondary accent — is that Gary has filled this role before. Anyone else whose name occurred in the blank would be an old referent in a new role, and hence less predictable than Gary.

(10) Gary, Larry, Harry, Barry and Mary all showed up at the Halloween party, and out of all of them, Gary got the drunkest. They all showed up at the Christmas party too, and once again Gary was the most out of control. Well, now they're all here for Easter dinner, and guess who's gotten roaring drunk? That's right, it's _____.

In Section 3 we will consider how these generalizations can be made more formal. But even with the informal versions we have stated here, we can make some worthwhile predictions. In the remainder of this section, we will demonstrate the reasoning behind a few of these predictions; this will also let us illustrate the general shape of our account before we delve into specifics.[8]

Let's return to the classic example of secondary accent, adapted from Partee (1999).

(1) **A:** Everyone already knew that Mary only eats VEGETABLES.
B: If even PAUL knew that Mary only eats vegetables // then he should have suggested a different RESTAURANT.

When B utters the word *vegetables*, it is focused, but unlike most foci it bears secondary accent. (We adopt the convention of writing primary-accented words in SMALL CAPS and secondary-accented words with underlines. Intonational phrase breaks are indicated by double slashes.)

The explanation is simple. Consider B's utterance in (1). The word *Paul* is unpredictable, because it is new in the immediate discourse context. But furthermore, *Paul* is important, because it is contrastive: *two points!* By comparison, the occurrence of *vegetables* in B's utterance is discourse old and highly predictable, the VP that contains it being string identical to one in the previous utterance. Still, *vegetables* is important because of its association with the exclusive particle *only*. We award it one point. Since *Paul* and *vegetables* are in the

[8]There is, too, some merit to the informal versions. In particular, if we express them in terms of psychological traits that can be indirectly observed — what a hearer can and can't predict; what a speaker will and won't do to meet goals he has been set — it is easier to incorporate them into psycholinguistic experiments.

same intonational phrase, they must compete for prominence.[9] *Paul* wins with two points, and receives primary accent. *Vegetables*, with one point, still needs to be more prominent than the scoreless words around it. We give it secondary accent as a sort of consolation prize.

Rooth's (1992) "rice farmer" example has gotten a lot of attention in the literature on focus recently.

(11) People who GROW rice generally only EAT rice.

In some ways, it is similar to (1). It, too, has a repeated constituent *rice*, which is associated with *only* in its second occurrence, but which bears only secondary accent. But as we will see in Section 4, there are several interesting differences between the two examples, which make it problematic for accounts of second-occurrence focus which handle (1) just fine. (Here's a sneak preview: in (1), the word which intervenes between *only* and *vegetables* is totally without communicative significance — unfocused, unimportant, highly predictable. In (11), the word which intervenes between *only* and *rice* is contrastively focused. This difference makes a difference for other accounts.) But on our account, we can handle this new example in exactly the same way, despite its difference in structure. We note that the second occurrence of *rice* is predictable. This is because it does not have a discourse-new referent, and because it is not newly occupying a particular slot in a previously used expression. We note that *eat*, on the other hand, is unpredictable. As before, both competitors are important — *rice* because of its association with *only*, *eat* because it is contrasted with *grow*. And as before, the unpredictable-and-important competitor beats the predictable-but-important one. The winner is *eat*.

[9] Astute readers will note that we have not explained *why* both expressions fall in the same intonational phrase. And this is currently the most serious gap in our account, for we do not attempt to predict phrasing. Here, we observe that the most natural pronunciation of B's contribution is phrased as in (1) but we offer no explanation of *why* it is phrased this way.

Here is one relevant observation, though. It is possible to phrase B's contribution differently — and when it is phrased in such a way as to separate *Paul* and *vegetables*, *vegetables* becomes eligible for primary accent once more. Consider the following.

(i) If even PAUL // KNEW // that Mary only eats VEGETABLES // then he should have suggested a different RESTAURANT.

We find that it gives an exasperated tone to the whole *if*-clause, and singles out the word *knew* as an especially strong source of exasperation or even outrage. But in a situation where such outrage is warranted, we find it to be perfectly felicitous. Note that pauses before and after *knew* are called for, suggesting that there are phrase boundaries here. And note too that, safe from competition on the far side of these phrase boundaries, *vegetables* can get its pitch accent back.

We take this as some evidence that the intonational phrase is the proper domain for competition for prominence. But we are still merely observing the locations of the phrase boundaries, and not making any predictions about their locations. This is no substitute for a proper theory of phrasing — indeed, it only reinforces the need for such a theory, by raising the question of why *outrage over Paul's epistemic state* should license phrase boundaries around the word *knew* — and clearly more work is required here.

3 Formalizing predictability and importance

The next step is to formalize the account we have laid out in the last section. We propose a formalization which is based on two sets of syntactic marks. As in Selkirk's account, the first set of marks is identical to the F-marking used in Rooth (1992) to capture focus. The second set of marks captures the notion of predictability we have used in the previous section. As in the last section, we assign prominence based on competition between the constituents within an intonational phrase.

It should be born in mimd here that we will not present empirical evidence that the phenomenon of acoustic prominence should be mediated via syntactic marking, and there is no direct evidence for the syntactic markers we will assume. But our use of syntactic markers follows what has become the norm in the field, and we hope it will facilitate comparison to other models which use syntactic marking, e.g. Rooth (1992).

The system of marking we use to capture the predictability of an expression is derived from Schwarzschild (1999), and provides us with a rule for *N-marking* those constituents that are unpredictable in the current context.[10] The core principle driving Schwarzschild's approach is this: if you take any constituent in a felicitous sentence and cut out the N-marked parts, the semantic value of whatever remains ought to be old news. This principle guides the application of N-marking, and the goal is to N-mark as few constituents as possible while still adhering to it.

Here is how Schwarzschild formalizes that principle. Let us define two operations: *N-closure*, which "cuts out" the N-marked parts of a constituent's semantic value, and *Existential type-shifting*, which turns the semantic value of any constituent into a proposition. Let us then call any constituent *Given* iff its N-closure is entailed by some salient antecedent — where for the purposes of calculating givenness, we use existential type-shifting in order to turn non-propositions into propositions so that we can check if they entail one another, and we count one expression of type e as entailing another if the two expressions corefer.

(12) **N-closure:** The N-closure of U is defined as the result of replacing N-marked phrases in U with existentially quantified variables.

(13) **Existential type-shifting:** Existentially type-shifting A is defined as existentially closing off any open variables in A.

[10] We do deviate from Schwarzschild on one point: *terminology*. There are now simply too many forms of "F-marking" in the literature, often with radically different meanings. We have already followed Selkirk and Rooth in using "F-marking" to indicate important constituents. But Schwarzschild uses "F-marking" to indicate unpredictable constituents; and we obviously cannot use the same mark for both, since it is crucial for our account that predictability and importance be marked separately and that the same constituent be able to be marked for both. We will go on using "F-marking" for important constituents, and we will adopt "N-marking" — N stands for New — for unpredictable ones. If this is confusing, remember: our N-marking is just Schwarzschild's F-marking under a new name.

(14) **Givenness:** An utterance U counts as Given iff it has a valid salient antecedent A, where
   a. if U is an individual, then A is a valid antecedent for it if A and U corefer;
   b. if U is not an individual, then A is a valid antecedent for it if A entails the N-closure of U modulo existential type-shifting.

Given these definitions, we apply as little N-marking as is required to satisfy the following constraint:

(15) **Givenness constraint:** A constituent that is not N-marked must be Given.

This constraint tells us when we have correctly N-marked a sentence; but it does not give us a recipe for arriving at the correct N-marking. Still, there are a few rules of thumb that can be useful here. Here we will give a few examples illustrating these rules of thumb — and illustrating also the sort of reasoning which our system of F- and N-marking lets us do.

First of all, entities, events and properties that are being mentioned for the first time trigger N-marking on constituents that refer to them. This means that noun phrases with brand-new referents according to Prince's (1981) taxonomy get N-marked. And while there is no similarly well-established taxonomy of givenness for verb or adjective meanings, we can apply similar criteria to them. If a verb or adjective has not been uttered in the current discourse, has not had any synonyms or hyponyms uttered, and does not have its meaning entailed by anything else that's been said so far, we must N-mark it. We might say that constituents N-marked according to this rule of thumb are N-marked "for their own sake."

Note that things which are entailed by the context do not get N-marked for their own sake even if they are being explicitly mentioned for the first time. As a result there are some lexical categories that are rarely or never N-marked for their own sake, because they are entailed by every context. For instance, Schwarzschild points out that this is true of universal quantifiers. $\exists P\, \exists Q\, [\text{all } P\ Q]$ is entailed by any context — for instance, consider the witnesses $P = Q = [\text{purple}]$, so long as there are purple things in the universe — and so *all*, *every*, *each* and so on are never N-marked for their own sake. We observe that the same is true of the exclusive particle *only* and the scalar particle *even*: any context at all will entail $\exists P\, \exists Q\, [\text{only } P\ Q]$, and likewise for *even*, and so these words are never N-marked for their own sake.

But even if a constituent is not N-marked for its own sake, it is possible that it be N-marked for the sake of another, larger constituent. While it is difficult to set down with complete generality the precise conditions under which this occurs, there are some cases where it occurs that are easily described. In particular, if one utters an expression that *would be* synonymous with one which has been used before, *except* that a single constituent has been added or changed, then the added or changed constituent must be N-marked for the sake of the

larger expression *even* if it is not N-marked for its own sake.

A few examples ought to make this clear. (16-b) has been raised as a puzzling example in some discussions of focus because *John*, despite being Given itself, is accented.

(16) a. Who does John's mother love?
b. She loves John$_N$.

But note that on Schwarzschild's account, it is still N-marked — not for its own sake, but for the sake of the larger expression. For while we have an antecedent for the word *John*, we don't have an antecedent for the phrase *loves John* — or, more precisely, for its existentially type-shifted N-closure $\exists x$ [x loves John]. But crucially, we *do* have an antecedent for the existentially type-shifted N-closure of *loves John*$_N$, which takes the form $\exists x\, \exists y[x$ loves $y]$; this is entailed by any context where the question in (16-a) can be asked without presupposition failure. So *loves John* without N-marking would violate the Givenness constraint, *loves John*$_N$ with N-marking would not violate it, and thus we N-mark *John* "for the sake of" the phrase *loves John*. This is how Schwarzschild explains the fact that the accent falls on *John* in this sentence.

Note that on our account, both F-marking and N-marking must be assigned to *John* in (16-b). Not only must it be N-marked for the sake of the verb phrase it appears in, it is also the answering constituent in a question. Examples like these help explain the enormous mileage that can be gotten out of single-factor accounts like Schwarzschild's, despite their being, or so we argue, ultimately incorrect. In many cases, F- and N-marking coincide perfectly — and in these cases, one can get perfectly good results by ignoring one form of marking so long as one correctly applies the other. It is only when F- and N-marking fail to coincide that the single-factor accounts break down.

We can see the same phenomenon in (17-b), *John's mother* does not need to be N-marked for its own sake, because it is entailed by the earlier mention of John. (Everybody has a mother.) But the larger constituent *he loves his mother* is not entailed by anything in the context. (Some people don't love their mothers, and for all we know, John is such a person.) And the easiest way to "rescue" *he loves his mother* from violating the Givenness constraint is to N-mark *his mother* — so we N-mark it for the sake of the clause.

(17) a. John loves Mary...
b. ...and he loves [his mother]$_N$.

Once again, a case could be made for F-marking *his mother* as well as N-marking it, on the grounds that *Mary* and *his mother* form a contrastive pair. We do not believe that a speaker is required to mark this contrast. But on accounts like Rooth's and Büring's which only use F-marking, the correct prosody can be predicted for clauses like (17-b) by assuming that the speaker does mark the contrast. Once again, the frequent coincidence of F- and N-marking saves the day for a single-factor account.

One last example: it is interesting to note that words like *all* and *only*, which

are never N-marked for their own sake, can be N-marked for the sake of a larger constituent. In (18), while *only* is Given, the larger expression *only eats vegetables* is not — $\exists x$ [x only eats vegetables] is a stronger claim than $\exists x$ [x eats vegetables], and the latter is all that we have an antecedent for. The easiest way to remedy this is to N-mark *only*.

(18) a. Mary eats vegetables?
b. That's not all.
She only$_N$ eats vegetables.

If we add F-marking to this last example, you can see in a nutshell how our approach to secondary accent works. *Only* must be F-marked in addition to being N-marked, because it is being contrasted with what has come before. *Vegetables* must be F-marked as well, because it is associated with *only*. The result is the marking in (19-a), which places greater communicative importance on *only* than on *vegetables*. Thus, *only* wins the competition for prominence, leading to the pronunciation in (19-b).

(19) a. She only$_{F,N}$ eats vegetables$_F$.
b. She ONLY eats vegetables.

Here we have a case where F- and N-marking fail to coincide. And here the single-factor accounts begin to struggle. For those that only use F-marking, the challenge is explaining why *vegetables* does not bear primary accent. (The recent versions of Rooth's and Büring's accounts make a valiant effort at explaining this, and both indeed give the right prediction on this particular example, although as we will see both accounts have problems with other examples.) For those that only use N-marking, the challenge is explaining why it does bear secondary accent. (We are not aware of any N-marking-only account that has risen to this challenge.)

One last general point before we start delving into specifics. We have indicated a few cases in which one must N-mark a constituent. But sometimes it is useful to show the opposite — to demonstrate that a constituent *isn't* N-marked. In order to do this, one must show that the constituent is "Given all the way up." That is, one must show that it is Given, *and* that all larger constituents containing it are Given. This notion of Givenness-all-the-way-up will turn out to be a useful one in discussing examples of secondary accent, since showing that a constituent deserves secondary accent often involves showing that it is not N-marked.

We are now ready to show how our formal account handles some classic examples of secondary accent. Consider B's contribution to (1), repeated here.

(1) **A:** Everyone already knew that Mary only eats VEGETABLES.
B: If even PAUL knew that Mary only eats vegetables then he should have suggested a different restaurant.

First we observe that the Givenness constraint requires us to N-mark *Paul*.

Paul refers to an individual; so for it to count as Given, we would need an antecedent in the discourse that refers to the same individual. Since there is no such antecedent, the only way to satisfy the requirement that every un-N-marked constituent be Given is to N-mark it.

Second, we observe that we are not required to N-mark *vegetables*, as it is Given all the way up. First of all, the word *vegetables* itself is Given. And second, provided we have already N-marked *Paul*, all the larger constituents that contain *vegetables* are Given whether or not *vegetables* is N-marked. (We can confirm this by checking that every proposition in (20) is entailed by A's contribution to (1).) Because our goal is to apply as little N-marking as possible while remaining consistent with the Givenness constraint, we do not N-mark *vegetables*.

(20) a. $\exists P\,[P(\text{vegetables})]$
b. $\exists x\,\exists y\,[x \text{ eats } y]$
c. $\exists x\,[x \text{ eats vegetables}]$
d. $\exists x\,[x \text{ only eats vegetables}]$
e. $\exists P\,[P(\text{Mary})]$
f. Mary only eats vegetables.
g. $\exists x\,\exists \phi\,[x \text{ knew } \phi]$
h. $\exists x\,[x \text{ knew that Mary only eats vegetables}]$
i. $\exists x\,\exists \phi\,[\text{If } x \text{ knew that Mary only eats vegetables, } \phi]$

Finally, we observe that both *Paul* and *vegetables* must be F-marked — *Paul* because it is associated with the focus-sensitive operator *even*, and *vegetables* because it is associated with the focus-sensitive operator *only*.[11]

So now when we assign accents, we find that *Paul*, being F- and N-marked, is more communicatively significant than *vegetables*, which is only N-marked. We can only assign one primary accent; based on the Prominence Principle, we must assign it to *Paul*. We can assign as many secondary accents as we like, but the only place where the Prominence Principle requires one is on *vegetables* — since it is F-marked, it must be more prominent than the words around it which bear neither F- nor N-marking.

Our formal account of the rice-farmer example proceeds in much the same

[11] Is this *all* the marking that's required? Well, the consequent of the conditional — *then he should have suggested a different restaurant* — will require some N-marking. But it falls in a different intonational phrase than the antecedent, and so we do not have to consider it in deciding where in the antecedent to place the primary accent.

It is also possible that the phrase *even Paul* will require N-marking in addition to the word *Paul*. Whether we N-mark it depends on what we think is asserted by a use of *even*, and what we think is merely implicated or presupposed. If we analyse the meaning of *even* in such a way that A's contribution to (1) *entails* the propositions in (i), then we do not need to N-mark *even Paul*; if not, then we do need to N-mark it. But whether or not we N-mark *even Paul* does not affect the argumentation that follows.

(i) a. $\exists P\,\exists x\,[P(\text{even } x)]$
b. $\exists x\,\exists \phi\,[\text{even } x \text{ knew } \phi]$
c. $\exists x\,[\text{even } x \text{ knew that Mary only eats vegetables}]$

way. Here we have no explicit context given as part of the example, so the placement of N-marking depends on the context we imagine it uttered in. Out of the blue, *people*, *grow*, the first occurrence of *rice*, *generally* and *eat* all call for N-marking. Let us assume for now that this is the case.

But crucially, the second occurrence of *rice* does not call for N-marking. It is itself Given, with the first occurrence of *rice* as its antecedent. And the larger consituents that is is part of are also already Given, based on N-marking which we have performed for independent reasons. For instance once we have N-marked *eat*, which we had to do anyway because we have no antecedent that entails any sort of eating went on, the phrase eat_N *rice* is Given. Its N-closed, existentially type raised form is (21), which is entailed by the existentially type raised form of the earlier utterance of *grow rice* (22).

(21) $\exists x \exists R\,[R(x)(\text{rice})]$

(22) $\exists x\,[x$ grows rice$]$.

Meanwhile, F-mark *eat* and *grow* to show contrastive emphasis, and the second occurrence of *rice* due to its association with *only*. The result is as follows:

(23) People$_N$ who grow$_{F,N}$ rice$_N$ // generally$_N$ only eat$_{F,N}$ rice$_F$.

In the first intonational phrase, *grow* wins the competition for prominence and gets primary accent. In the second, *eat* wins and gets primary accent. The other N-marked words get secondary accent, including the first occurrence of *rice*; and so does the second occurrence of *rice*, which is F-marked but now no longer new.

4 It's not size that counts

In this section, we will introduce the notion of a focus domain, from Jacobs (1991) and Rooth (1992), show how this is used by Büring (2008) and Rooth (2010) to account for secondary accent data, and detail some shortcomings of the approach.

4.1 Focus Domains and Secondary Foci

Rooth's (1992) alternative semantics for focus provides two meanings for every expression, a regular meaning, and a focus meaning. The focus meaning for a constituent is just like the regular meaning, except that the meaning of focused elements is replaced with alternatives of the same type. Thus e.g. the focal meaning for "ate a bagel$_F$" is $\{$ ‖ate a bagel‖, ‖ate a sandwich‖, ‖ate a bicycle‖$\ldots\}$. Simply stating that these are the focal meanings does not in and of itself lead to any testable predictions. The theory gains empirical bite only when additional constraints are placed on the alternatives. Such constraints are introduced via Rooth's variant of the notion of a *domain* of focus.

In Rooth's account, focal domains determine which sets of alternatives matter. Rooth marks out focus domains using a syntactic operator which he notates $\sim$; we will use corner brackets $\lfloor \ldots \rceil$. Thus in the configuration "Fred only $\lfloor$ate a bagel$_{\mathrm{F}}\rceil$", the focus domain is the VP "ate a bagel$_{\mathrm{F}}$", and the set of alternatives is supplied to "only", with the semantic effect that the sentence means that none of the alternatives is true of John, except the alternative "ate a bagel". To take a case of contrast, consider "Mary washed $\lfloor$a red$_{\mathrm{F}}$ car$\rceil$and $\lfloor$a blue$_{\mathrm{F}}$ car$\rceil$." Here there are two focus domains: the phrase "$\lfloor$a red$_{\mathrm{F}}$ car$\rceil$" requires that there's some "a X car" present, with X$\neq$"red", and similarly for "$\lfloor$a blue$_{\mathrm{F}}$ car$\rceil$".

Both Büring (2008) and Rooth (2010) account for secondary accent in terms of domain size, with variants of a simple generalization which says that more prominent foci have bigger focus domains. But before seeing how the domain size account is developed, let us consider its motivation.

Büring considers examples that are structurally comparable to the dialogue in (24), which has the following property: both occurrences of "Beaver" in the (b) sentence are Given in the traditional sense of referring to a previously mentioned entity, and both occurrences are semantic foci of focus sensitive operators, "even" and "only" respectively. And yet one occurrence is more prominent than the other.

(24) a. We can't accept this: the author only cites Beaver.
b. Well then we know who wrote it: only BEAVER only cites Beaver.

Based examples similar to these, Büring concludes: "The difference between primary focus and [second occurrence focus] can not generally be reduced to the latter being focussed and Given, while the former is focussed and non-Given." We agree completely, at least if "Given" means "referring to a previously mentioned entity." However, and as should be clear, if by "Given" Büring were to have meant "predictable from context", we would disagree completely: we will return to this important distinction shortly. But first, let us take a look at the alternative Büring offers to explaining secondary accent in terms of the combination of focus and Givenness.

Having rejected a two-component analysis of secondary accent, Büring proposes instead to explain the phenomenon in terms of the relative size of focus domains, using two main principles:

(25) a. **Büring's Domain Theory of Primacy**: Among two foci in a sentence, the primary focus is the focus whose domain contains the domain of the other.
b. **Büring's Focus Prominence Principle**: If P is the domain of a focus sensitive operator O, the most prominent element in P is a focus of O.

Consider the final clause of (24)b. The effect of the above two principles can be seen if a syntactic structure like that in (26-a) is assumed, and the focal domains of the two foci are as in (26-b). Büring's Focus Prominence Principle

implies that the second "Beaver" must be the most prominent element in the string "cites Beaver". But since the first "Beaver" is the focus of a domain (domain 1) which strictly includes the domain of the second "Beaver" (domain 2), Büring's Domain Theory of Primacy implies that the first "Beaver" must be more prominent than the second. Therefore the first "Beaver" gets nuclear stress. The second is then predicted to receive a secondary stress, which in this post-nuclear region is commonly realized through slight increases in duration and intensity.

(26) a. (only)(Beaver only cites Beaver)
b. $(\text{only}_1)(\lfloor \text{Beaver}_{\text{F}1}\ \text{only}_2\ \lfloor \text{cites Beaver}_{\text{F}2}\ \rceil_2 \rceil_1)$

The analysis of (25) depends upon assumptions regarding the focus domains needed for focus sensitive operators. Although the syntactic analysis might be questioned (i.e. the assumption that at least one "only" is underlyingly a sentential operator), the focus domains are semantically well motivated. Based on a consideration of the truth conditions, it is clear that the alternatives which must be excluded by the first "only" are of the form "X only cites Beaver", and thus the focal domain associated with the first "only" and its focus must include the second "only" and its focus.

It is not always the case that truth conditions determine how big the focus domains are. Consider (27-a). This can be analyzed as in (27-b) in order to predict that "told" gains primary stress, and "vegetables" secondary stress. But here "told" is not associated with an overt focus sensitive operator, so the truth conditions are not directly affected by the position of focus domain 1. In such a case, Büring analyzes the expression that receives primary stress as an instance of *free focus*, with focal domains taking scope over the entire sentence. In fact, Büring motivates this using precisely the model of givenness discussed above, Schwarzschild's. In our terms, whenever N-marking of an expression is needed to make the sentence Given, Büring assumes a sentential focal domain for that expression. In the case at hand, the focus domain for "told" must be sentential because if there were no N-marking on "told", then the entire sentence "I told you she only eats vegetables" would fail to be Given.

(27) a. (You SEE!) I TOLD you she only eats vegetables.
b. $\lfloor \text{I told}_{\text{F}1}\ \text{you she only}_2\ \lfloor \text{eats vegetables}_{\text{F}2} \rceil_2 \rceil_1$

4.2 Focus intervention structures

It is not obvious how to intone the string in (28-a) and (28-b) such that it successfully answers the question of what John ate in Paris and nowhere else. Büring (2008) describes this answer as *ineffable*, although he is quite clear that he only means by this that the string in (28-a) and (28-b), the answer form most straightforwardly obtained from the given question form, cannot represent the desired meaning. He recognizes that other forms might successfully answer the question: example (29) has such a form. The issue that Büring considers,

though, is why various renditions of the string form in (28-a) and (28-b) cannot be used to answer the question.

(28) What did John only eat in PAris?
   a. #John only ate crêpes in PAris.
   b. #John only ate CRÊpes in Paris.
      (Büring 2008, cited there as Schwarzschild p.c.; judgment as given by Büring)

(29) What John only ate in PAris was CRÊpes.

Structurally, the important property (28-b) has is that a focus intervenes between an exclusive and the focus associate of the exclusive. That is, the example contains a string with an Exclusive, a Question answering constituent, and the Associate of the exclusive, in that order: we will say that such cases have an E-Q-A structure. In an utterance of (28-b), "crêpes" should have a domain with sentential scope, since it is the question answering constituent, and the entire sentence serves as the answer to a prior question. On the other hand, "Paris", the associate of the exclusive "only", should have a smaller domain, namely the VP *ate crêpes in Paris*. Büring's Domain Theory of Primacy then implies that "crêpes" must be more prominent than "Paris". But Büring's Focus Prominence Principle requires that the focus of "only", i.e. "Paris", be the most prominent element in its domain. Since that domain also includes "crêpes", it follows that "Paris" must be more prominent than "crêpes", and we reach a contradiction. And this, Büring concludes, is why (28-b) is ineffable.

The major problem with this account of the ineffability of (28-b) is that it is based on a false premise: it is not the case that sentences with an E-Q-A structure are in general ineffable. Rooth (2010) points out that variants of (28-b) with extra material around either of the two foci are ameliorated, and also points out that (28-b) becomes felicitous if uttered with "a rising intonation indicating a partial answer on the first focus crepes." We agree with these judgments, and have verified them with several consultants. Furthermore, we were able to find naturally occurring examples with the E-Q-A structure on the web. Example (30-a) is the last sentence of the lead paragraph in a discussion thread, and (30-b-e) are extracted from answers provided by various respondents, all with the desired form:[12]

(30) a. What do you only eat out at a restaurant and, vice versa, what do you only eat at home?
   b. I pretty much only eat non-European food (Indian, Japanese, Malaysian, Middle Eastern, etc.) out because I don't have most of the ingredients or tools to make them in my pantry.
   c. I pretty much only eat omlettes at home (and most eggy breakfast) because I'm always afraid that restaurant versions will be too greasy.

[12]Example collected from `http://www.thekitchn.com/thekitchn/what-do-you-only-eat-in-restaurants-103998`, 2-17-2010.

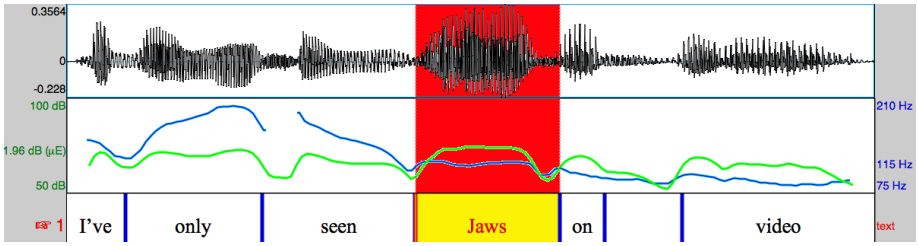

Figure 1: Sound profile for (31-b), with F_0 in blue, intensity in green.

d. I only eat fried chicken at home, I do not like eatting with my hands when I go out.
e. I only eat mashed potatoes at home - I don't want to be judged by how much butter or salt I put into mine. The other restaurant-goers would be scarred, probably.

Although the above examples were collected in written rather than spoken form, they remain felicitous when read. Furthermore, when we asked consultants to read E-Q-A structure sentences, we found little evidence of any ineffability effect at all. After we read the context setting question to them, they either immediately produced the examples in spoken form without a problem, or else requested further clarification about the intended context and then produced them, not reporting any infelicity. As an illustrative example, and though we do not wish to make any generalizations as regards intonational phonology based on this or other recordings we made, consider the speech sample rendered using the Praat acoustic analysis software, in figure 1. This depicts the production that resulted when a consultant asked to produce (31-b,c) in answer to (31-a) read (31-b), with pitch tracks (blue) and intensity (green). For this particular speaker, intensity, or total energy, plays at least as strong a role in prominence marking as pitch. Audibly, the question answer "Jaws" is the most prominent element in this production (as judged by the current authors), but the focus of "only", i.e. "video", is also prominent, and indeed is accompanied by both a significant intensity peak and a small pitch movement on the stressed syllable.

(31) a. What movies have you only seen on video?
b. I've only seen Jaws on video.
c. I've never seen it in the theatre.

While the phonetic and phonological analysis of (31-b) and the other productions we have collected would in itself be a worthy enterprise, strong conclusions about how such examples are produced must await a systematic experimental study. For our current purposes, it suffices to note that there is no question at all of sentences with E-Q-A structure being ineffable. The fact that some speakers have difficulty producing the original crêpe example as a com-

plete answer to a question remains in need of explanation, but whatever the explanation, it is not the explanation Büring gave in his discussion of the original crêpe example. For Büring's explanation would incorrectly predict that many robustly acceptable sentences, including naturally occurring examples we have observed, were infelicitous. On the contrary, a theory of secondary accent must account for the fact that E-Q-A structures are possible. Thus, rather than supporting his model, examples with E-Q-A structures rule out Büring's proposal in its original form. We will turn shortly to a modification of Büring's system which remedies this problem, but first we consider a further example of focus intervention, the *rice-grower* example from Rooth (1992) that we already considered above.

(32) People who *grow* rice generally only *eat* rice.

The rice-grower example is structurally comparable to the crêpe example, except that a Contrastive element, i.e. "eat", intervenes between the exclusive and its associate, instead of a question answering constituent intervening. Accordingly, we can label the structure E-C-A instead of E-Q-A. Crucially, in a natural production of this example the most prominent element in the VP "eat rice" is "eat", not "rice", thus falsifying Büring's Focus Prominence Principle.

Büring (2008) is in fact well aware that the rice-grower example constitutes a *prima facie* counter-example to a version of his account like that presented above. And in fact, for this and other reasons he rejects various of the assumptions we have made above. Most notably, he advocates for a notion of focus domain that is marked at a phonological level rather than a syntactic level. Phonological focus domains, unlike Rooth's syntactically marked focus domains, do not interface directly with the semantics. This allows Büring to save his Focus Prominence Principle by specifying a narrow focus domain for the focus "rice" in (32), e.g. the domain containing "rice" and nothing else.[13] However, the move to a phonological notion of focus domain results in a problematically under-constrained theory. Whereas syntactic focus domains can be motivated, at least in part, in terms of truth-conditional effects, Büring provides no way of connecting phonological focus domains to the semantics. It seems natural, then, to suppose that phonological focus domains will be motivated in some other way, and the obvious way to motivate them would be in terms of the acoustic realization of the utterances in question. However, Büring does not assume that phonological focus domains are necessarily marked acoustically. He does suggest that the placement of phrasal boundary tones (whether tones from from the boundaries of full *intonational phrases*, or from the boundaries of smaller *intermediate phrases*) can sometimes indicate the extent of a focus domain. But he also allows that sometimes focus domains correspond to

[13]Note that in the syntactic version of Büring's proposal, a separate principle "Domain of a Focus/an Operator" constrains the focus domain for the associate of a focus sensitive particle. This principle requires that the domain of a focus on an expression associating with an overt operator has to be the largest constituent containing the focus but not containing the operator. Absent syntactic transformations, for a VP-modifying exclusive, this domain will be the entire VP sister to the exclusive, e.g. "eat rice" in (32).

smaller underlying units that may not be marked acoustically. Thus phonological focus domains are highly abstract, not motivated semantically, and at best lightly constrained by phonetics and standard phonology. Rather than arguing against phonological focus domains, we merely note that at present the notion is not sufficiently tightly specified to provide us with a predictive theory of secondary accent. Accordingly we will not consider them further here, but instead turn to an alternative reaction to focus intervention examples, a modification of Büring's proposal considered by Rooth (2010).

Rooth proposes a principle he terms *Relatavized Stress-F* which amounts to the following modification (the italicized part) of Büring's Focus Prominence Principle.

(33) **Modified Focus Prominence Principle**: If P is the domain of a focus sensitive operator O, the most prominent element in P is a focus of O *excepting any other foci occurring within P which happen to have a domain strictly bigger than P*.

Three issues arise: (i) a lack of motivation, (ii) the complication the principle brings to Rooth's original alternative semantics, and (iii) empirical shortcomings.

Regarding the first point, the Modified Focus Prominence Principle is little more than a description of the data it was designed to account for, i.e. Büring's principle plus a clause that deals with a set of cases Büring's principle gets wrong. But we recognize that one linguist's *ad hoc* is another's *sine qua non*, so we will not press the issue of motivation any further.

The second issue concerns a complication that the Modified Focus Prominence Principle introduces, although we concede that this particular complication may reflect an inevitable consequence of the data, rather than demonstrating a genuine problem with the principle. The complication stems from the fact that the Modified Focus Prominence Principle allows nested focus domain structures as in (34) (with $O_{i,j}$ focus sensitive operators). Note that such structures are ruled out on Büring's model, because his Focus Prominence Principle would require each of the two foci to be more prominent than the other.

(34) $O_i \lfloor_i \ldots \lfloor_j \ldots F_i \ldots F_j \ldots \rceil_j \ldots \rceil_i$.

In the model of Rooth (1992), each expression is associated with exactly two components, the ordinary semantics, and a set of alternatives. But the intended interpretation of the structure in (34) requires two separate sets of alternatives to be associated with some expressions. The focal meaning needed for focus domain j involves plugging in alternatives to F_j while holding F_i fixed, while the focal meaning needed for focus domain i involves plugging in alternatives to F_j while holding F_i fixed. For example "eat$_{F1}$ rice$_{F2}$" in the rice-grower example (32) would pick out both the set of alternative properties of the form "eat X", and the set of alternative properties of the form "X rice".

More generally, the Modified Focus Prominence Principle allows embeddings like that in (34), but with arbitrarily many focus domains, so the principle

implies that expressions may have to be associated with arbitrarily many sets of alternative meanings. This complication to the syntax-semantics interface is an unattractive consequence of the Modified Focus Prominence Principle. But the fact that this consequence is unattractive does not make the principle inconsistent: in fact, Rooth (2010) shows in detail how a system which embodies the principle can be implemented. Furthermore, maybe the felicity of examples with focus intervention structures implies that something like the complication that the Modified Focus Prominence Principle introduces is inevitable. If so, this complication obviously cannot be taken to be a counterargument to the principle.

The third issue we wish to raise as regards the Modified Focus Prominence Principle stems from the fact that it implies (like Büring's original combination of principles) that if the focus domain for one focus is nested inside the focus domain for another focus, then the focus with the bigger domain size will be more prominent. It seems to us that the next two examples we will consider can plausibly be analyzed as counterexamples to this property. Consider (35-a), which, like the original rice-grower example, is naturally uttered with contrastive accents on "grow" and "eat", and with the final "rice" produced with secondary stress, and thus no pitch accent. Based on the analysis of contrast of Rooth (1992), we take it that he would consider an analysis like that in (35-b) to be possible.

(35) a. People who grow rice usually only want to eat rice.
b. People who $\lfloor$grow rice$_{\mathrm{F1}}\rceil_1$ usually only$_2$ $\lfloor$want to $\lfloor$eat$_{\mathrm{F2}}$ rice$_{\mathrm{F3}}\rceil_3\rceil_2$

The thing to note about the analysis in (35-b) is that the focus domain associated with "eat" (i.e. focus domain 3) is smaller than that associated with the final "rice". Thus in this case the most prominent element has a smaller focus domain than does a less prominent element.

The same holds for the final sentence of the discourse in (36), which can be produced with nuclear stress on "candy", a focused element associated with an occurrence of "only". But in this example, unlike in (35-a), there is a second overt focus sensitive operator, "also", for which the associate is "her sister". Crucially, the relative scopes of the two operators clearly imply that the focus domain for "her sister" strictly contains the focus domain for "candy", and thus the Modified Focus Prominence Principle implies that "her sister" should be more prominent than "candy". Since an utterance with exactly the reverse pattern of prominence is felicitous, our analysis of (36), like our analysis of (35), implies that the modified principle is incorrect.

(36) a. You're visiting Jane and her sister at christmas? I thought you hated them!
b. Yeah, well, I decided I'm only going to give a DIME BAR to Jane, and
I also decided I'm only going to give a crappy mass-produced CANDY bar to her sister.

There is one obvious way to reanalyze (36) and (35) such that they do not constitute counterexamples to the Modified Focus Prominence Principle. The reanalysis would involve marking the most prominent expressions in the examples with additional free foci, i.e. foci that have sentential scope focus domains. Given such a reanalysis, it would of course be the case that the most prominent elements were exactly those associated with the widest focus domains. And yet...such a move does not seem entirely satisfactory. To the extent that there is nothing preventing postulation of a covert wide domain operator whenever the Büring-Rooth model gets in trouble, the Modified Focus Prominence Principle above would appear to have no more empirical bite than the following variant:

(37) **An Unfalsifiable Focus Prominence Principle**: If P is the domain of a focus sensitive operator O, the most prominent element in P is a focus of O, *except when it isn't*.

Our own conclusion is that while we concede that it may be possible to base an analysis of prominence on domain size, current models seem to us to be either incorrect or vacuous. The most natural way to rescue them from vacuity would be to follow the strategy that Büring in fact suggests (before later dropping it when he advocates for a phonological notion of focus domain): motivate focus domains for free foci in terms of Schwarzschild Givenness. While it is not entirely clear to us how this would work for (36) and (37), we are not averse to such a possibility. But just suppose that a domain-size based theory could be developed which correctly accounted for all the data we have seen, and that in this model all focus domains were motivated either by association with overt operators, or using Schwarzschild Givenness. In that case, the resultant model would have exactly the components in it which we advocate as the basis for what we have termed Communicative Significance, but with slightly different syntactic marking. We have no major objection to such marking if it can be shown to work, but for the moment we will stick with the model we have already discussed, a model which is not so dependent as domain-size models on the details of syntactic configuration.

5 Comparing the splitting approaches

We have already described Selkirk's account as an important antecedent for our splitting approach to prosody. But there are several differences between our account and hers that we have glossed over in doing so. Having pointed out the flaws in the lumpers' approach, we turn to these internal differences within the splitters' camp, and argue that our account represents an improvement over Selkirk's.

First of all, while our account is based on the principle of competition for prominence, Selkirk bases hers on the Focus Prominence Principle with a few further constraints added. We can think of this difference in terms of the flexi-

bility or inflexibility of our respective principles. Competition for prominence is somewhat flexible, in that it allows a focused constituent to be less prominent than its neighbors (so long as those neighbors have won their greater prominence fair and square). The Focus Prominence Principle, as we saw in the last section, is inflexible on this point; a focused constituent must be the most prominent thing in its domain, and if there is no way to satisfy this requirement, ineffability results. We argue that the flexibility of our account is a virtue here, for it allows us to make correct predictions about focus intervention structures that FPP-based accounts — including Selkirk's — cannot make.

Another difference between our accounts, which we hinted at earlier, is that we use the notion of predictability as our second factor, formalized above in terms of N-marking, whereas Selkirk uses a modified notion of Givenness as the second factor in her account. We argue that this represents a basic misunderstanding about the conditions under which second-occurrence focus can be found. Selkirk has suggested that second-occurrence foci are simply foci which are Given, on her modified definition of Givenness. We show that this is not the case: that there are Selkirk-Given foci which are not second-occurrence foci, and second-occurrence foci which are not Selkirk-Given.

5.1 Selkirk and focus intervention structures

We can think of Selkirk's account, like Rooth's and Büring's, as an attempt to rescue the Focus Prominence principle. Alone, the Focus Prominence Principle cannot distinguish between primary and secondary foci; to rescue it, therefore, one needs to add a criterion that will make this distinction. Rooth and Büring, as we saw, use domain size as a distinguishing criterion; Selkirk instead uses the givenness/newness of foci. She does this by adding to the original formulation of the Focus Prominence Principle a second principle proposed by Féry and Samek-Lodovici (2006) which requires that Given constituents be as low in prominence as they can be without violating the Focus Prominence Principle. In practice, this amounts to a guarantee that Given constituents will never bear primary accent.

(38) **The good old Focus Prominence Principle again**:
If P is the domain of a focus sensitive operator O, the most prominent element in P is a focus of O.

(39) **Destress Given**:
A Given constituent does not bear primary accent.

In addition, Selkirk makes explicit the requirement that every intonational phrase contain at least one primary accent.

The result of the interaction between these principles is as follows. If an expression is focused but not Given, the FPP places a lower bound on its prominence, but there is no upper bound on how prominent it may be. Therefore, it is eligible to receive primary accent. On the other hand, if an expression is focused and Given, we have both a lower bound and an upper bound on its

prominence. It must be prominent to some degree, but it may not bear primary accent. Therefore, secondary accent is assigned to it.

Selkirk maintains the convention of marking focused phrases with an F, and additionally adopts the convention of marking Given phrases with a G. This results in four possible sets of markings, and we may summarize the possibilities and her predictions as follows.

(40) **Marking options on Selkirk's theory**

a. F- and G- marked: "Secondary focus" — these constituents must bear secondary accent.
b. F-marked: "Contrastive focus" — these constituents can bear primary accent.
c. unmarked: "Information focus" — these constituents can bear primary accent.
d. G-marked: Unfocused — these constituents bear no accent.

Now, it is clear that G-marking must be somehow constrained if it is going to be used in this way. For if we apply G-marking without constraint to every discourse-given constituent, we will predict secondary focus where it does not in fact occur. Selkirk gives the following example:

(41) A: Anscombe has been feuding with her colleagues.
B: Wittgenstein brought a glass if wine over to Anscombe. But not to the others. Presumably it was an act of reconciliation.

Suppose that B is emphasizing the contrast between *Anscombe* and *the others*, meaning that both these constituents bear F-marking. If we now G-mark every discourse given constituent, we will have to G-mark *Anscombe*. This leads to the false prediction that *Anscombe* will bear secondary accent, and does not account for the pronunciation in (43).

(42) Wittgenstein brought a glass of wine over $\lfloor$to Anscombe$_{F1,G}\rceil_1$. But not $\lfloor$to the others$_{F2}\rceil_2$.

(43) Wittgenstein brought a glass of wine over to ANSCOMBE. But not to the OTHERS.

Selkirk's solution is to suggest that focused constituents must meet a higher standard of givenness than ordinary unfocused constituents if they are to be G-marked. Recall that focused constituents, in addition to their ordinary meaning, are assigned a focal meaning in alternative semantics, consisting of a set of alternatives of the same type. Her proposal builds on this fact: she suggests that constituents should be G-marked when *every element* of their meaning is given in the discourse. For unfocused constituents, this means their ordinary meaning must be given. But for focused constituents, it means that their ordinary *and* their focal meanings must be given.

(44) **Selkirk's G-marking Condition:**

a. An F-marked constituent α will be G-marked iff the phrasal scope ϕ of the focus operator corresponding to it has an antecedent in the discourse for its focus semantic value $[\phi]^f$.
b. Otherwise, a constituent α will be G-marked if it has an antecedent in the discourse for its ordinary semantic value $[\alpha]^o$.

Under this condition, *Anscombe* is not G-marked in B's contribution to (41). In order to G-mark it, we would need an antecedent for the alternative set given in (45), which is the focus semantic value of *brought a glass of wine over to Anscombe*$_F$. And no such antecedent is available. This means that under Selkirk's G-marking condition, we get the following marking on B's contribution. This correctly predicts that both *Anscombe* and *others* will bear primary accent, as shown in (43)

(45) { ‖to Anscombe ‖, ‖to Geach ‖, ‖to Frege ‖, ‖to Russell ‖,... }

(46) Wittgenstein brought a glass of wine over ⌊to Anscombe$_{F1}$⌉$_1$. But not ⌊to the others$_{F2}$⌉$_2$.

With the new, stricter G-marking condition in place, Selkirk can correctly predict some instances of secondary focus. Consider (1) again. We have already seen that *Paul* and *vegetables* must be F-marked. According to Selkirk's G-marking condition, *Paul* may not be G-marked. But *vegetables* may, because the alternative set it evokes (looking perhaps something like (47)) has an antecedent in the discourse.

(47) { ‖eats vegetables‖, ‖eats dairy‖, ‖eats eggs‖, ‖eats meat‖... }

The same alternative set was evoked when A used the phrase *only eats vegetables*. *Mary* and *eats* are also G-marked — because these words are not F-marked, we do not need to find alternative sets to serve as antecedents for them, but only ordinary semantic values, and so the mere mentions of *Mary* and *eat* in A's contribution are good enough. Therefore, we end up with the crucial clause decorated as in (48). From this decoration follows the correct predictions about accent. *Vegetables* must be the most prominent thing in its domain, but it must also remain as low in prominence as possible. Secondary accent satisfies these constraints. *Paul*, meanwhile, must be the most prominent thing in *its* domain, but there is no upper bound on how prominent it may be. This means it is eligible to bear primary accent — and since, after all, something must bear primary accent, and no better candidate is available, we assign it to *Paul*.

(48) If even$_1$ ⌊Paul$_{F1}$ knew that Mary$_G$ only$_2$ ⌊eats$_G$ vegetables$_{F2,G}$⌉$_2$ ⌉$_1$...

(49) If even PAUL knew that Mary only eats vegetables....

But focus intervention structures pose just as much of a problem for Selkirk's account as they do for the domain-size accounts of Rooth and Büring. In fact, Selkirk gives a strikingly similar analysis of these examples to that of Büring,

including the claim of ineffability with which we (like Rooth 2010) have taken issue. For instance, her analysis of the rice-grower sentence must be (50). And this gives rise to the familiar set of conflicting requirements due to the Focus Prominence Principle: *eat* must be more prominent than *rice*, but *rice* must also be more prominent than *eat*, and there is no way to make this happen.

(50) People who $\lfloor \text{grow}_{F1}\ \text{rice} \rceil_1$ generally $\text{only}_2\ \lfloor \lfloor \text{eat}_{F3}\ \text{rice}_{F2}\ \rceil_2\ \rceil_3$.

Note that G-marking does not offer us a way out of this dilemma. Suppose that Selkirk had assigned the structure in (51) to the phrase *only eat rice*, with both F- and G-marking on the word *rice*. The same conflict would still arise: the Focus Prominence Principle would still require *rice* to be the most prominent constituent in its domain, and that would still be incompatible with the requirement that *eat* be the most prominent constituent in *its* domain. Above, we mentioned the inflexibility of the (unmodified) Focus Prominence Principle, and here we have an example of that inflexibility. Despite the fact that $rice_{F,G}$ is (to slip back into our own terminlology) less Communicatively Significant than eat_F, the Focus Prominence Principle will not permit it to be a runner-up for prominence in its own domain, or even to be tied for first place. Rather, it must be *the most prominent* constituent bar none, and we have seen that this is impossible.

(51) $\ldots \text{only}_2\ \lfloor \lfloor \text{eat}_{F3}\ \text{rice}_{F2,G}\ \rceil_2\ \rceil_3$

Nevertheless, it is interesting that Selkirk fails to G-mark the second occurrence of *rice* — that she produces the analysis in (50) and not the one in (51). For this means that, even if she had a more flexible principle for assigning accents, she would still be unable to produce the correct prediction for the rice-grower sentence. Here is the problem: Because *rice* is a focus, we may only G-mark it on Selkirk's account if we find antecedents for its ordinary meaning *and* its focal meaning. The ordinary meaning is easy to find an antecedent for. But its focal meaning — an alternative set like the one in (52), consisting of meanings of the form "eat ___" — has no antecedent. The only other focal meaning in the sentence which might have been able to serve as an antecedent is the one evoked by the phrase $\lfloor grow_{F1}\ rice \rceil_1$. And while $\lfloor grow_{F1}\ rice \rceil_1$ evokes an alternative set, it evokes the wrong one: not (52), but (53), consisting of meanings of the form "___ rice."

(52) {‖eat rice‖, ‖eat beans‖, ‖eat corn‖, ‖eat wheat‖…}

(53) {‖grow rice‖, ‖eat rice‖, ‖buy rice‖, ‖sell rice‖, ‖fill beanbags with rice‖…}

What has gone wrong here? It appears that Selkirk has misunderstood the circumstances under which second occurrence focus occurs. Her G-marking condition leads to secondary accent on a constituent under the following conditions:

(54) a. The constituent is discourse-old and focused, and

b. It was a focus in its previous appearance as well as its current one.

The flip-side of the G-marking condition is thus that a primary accent is produced when a new constituent is focused, *or* when a constituent that has been used before without focus now appears as a focus for the first time. The trouble is that (54-a) and (54-b) are neither necessary nor a sufficient conditions for second occurrence focus. The rice-grower example shows that (54-b) in particular is not a necessary condition. For in this example, *rice* fails to satisfy (54-b) — it is only appearing as a focus now for the very first time — and yet it receives secondary accent anyway. We can show that (54-a) and (54-b) are not sufficient conditions for second occurrence focus by considering (55). In the last line of this example, *vegetables* satisfies both conditions: it is a discourse-old constituent, it is focused, and it was focused in a previous appearance as well. And yet we find it quite natural to give it primary accent, and rather unnatural to give it the pronunciation which Selkirk predicts, found in (56).[14]

(55) A: Mary only eats VEGETABLES.
B: Are you sure? I swear I saw her car in the lot at that steakhouse she used to visit...
A: Don't be ridiculous. I trust her. I'm telling you, she only eats VEGETABLES.

(56) I'm TELLING you, she only eats vegetables.

So if (54-a) and (54-b) above are the wrong conditions, what are the right ones? Well, unsurprisingly, we think our account gives them. In place of Selkirk's requirement that secondary foci must be Given, we make the claim that they must be *predictable* — which on our formal version of the account, is just a shorthand for the requirement that they be Given-all-the-way-up. Selkirk found that she needed to constrain the notion of Givenness in order to prevent her theory from wildly overgeneralizing. But we find that having opted for the stronger notion of Givenness-all-the-way-up, there is no need for any additional constraint. If we mark foci as important, apply N-marking to unpredictable con-

[14]Here is a context where we do find the accent Selkirk predicts natural:

(i) A: Mary only eats VEGETABLES.
B: Oh, I don't know about that. I've barely met her. How could I possibly know what she eats?
A: I'm TELLING you she only eats vegetables.

What's the difference? In (55), A is re-asserting the claim that Mary is a vegetarian; here, he is asserting his own status as the source of that claim.

Selkirk correctly F-marks *telling*, because it contrasts with B's claim that he has no source of information about Mary's diet. But as in (55), she both F- and G-marks *vegetables* — and this time around, the resulting prediction is that *vegetables* should bear secondary accent is correct. The problem is that she does not provide a relevant way to distinguish between (i) and (55).

Similarly, (27-b) is unproblematic for Selkirk. Here too, the speaker is not reasserting the claim that Mary eats vegetables, but identifying himself — in fact, in this case, re-identifying himself — as the source of the claim.

stituents, and apply the principle of Competition for Prominence, we will arrive at the correct predictions.

5.2 An analysis of accent in intervention structures

In the model we are proposing, focus intervention structures do not present any greater difficulty than all the other examples of secondary accent we have discussed. We have already presented one analysis of a focus intervention structure, when we discussed the prosody of the orignial rice-grower example. The other examples we have raised all amenable to precisely the same analysis. Any time we see the structure in (57) (where the parentheses delimit an intonational phrase), we predict secondary accent on Y. Other constituents which bear a single mark of communicative significance, either F- or N-marking, may appear anywhere in (57), so long as X is the only constituent with *both* F- and N-marking.

(57) $(\ldots X_{F,N} \ldots Y_{F} \ldots)_{ip}$

Thus from our point of view, there are no prosodically important difference between any of the focus intervention structures in (58) and the classic intervention-free examples of secondary accent in (59).

(58) a. (usually only eat$_{F,N}$ rice$_{F}$)$_{ip}$
b. (usually only want$_{N}$ to eat$_{F,N}$ rice$_{F}$)$_{ip}$
c. (I only eat [fried chicken]$_{F,N}$ at home$_{F}$)$_{ip}$
d. (I'm only going to give a crappy$_{N}$ mass-produced$_{N}$ candy$_{F,N}$ bar to her sister$_{F}$)$_{ip}$

(59) a. (if even John$_{F,N}$ knew that Mary only eats vegetables$_{F}$)$_{ip}$
b. (only Beaver$_{F,N}$ only cites Beaver$_{F}$)$_{ip}$

Of these, (59-b) calls for some additional explanation. We have already observed that the two mentions of Beaver in (59-b) are both references to a discourse-old referent. But of the two, only one is treated as predictable on our model, and one is treated as unpredictable. This is a consequence of Schwarzschild's definition of Givenness, on which not only a NP but all larger constituents containing it must have antecedents in the discourse for the NP to be certain to avoid N-marking. In this particular example, we must N-mark the first occurrence of *Beaver* in order to allow the larger phrase *Beaver only cites Beaver* to satisfy the Givenness constraint. (Without N-marking, this phrase needs an antecedent entailing (60-a), which we do not have. With N-marking, thanks to the principle of N-closure, it only needs an antecedent entailing (60-b), which we do have.) And once the first occurrence of *Beaver* is N-marked, we can avoid N-marking the second, because we also have antecedents entailing all of (61).

(60) a. Beaver only cites Beaver
b. $\exists x$ [x only cites Beaver]

(61) a. $\exists x\,[x = \text{Beaver}]$
b. $\exists x\,[x \text{ cites Beaver}]$
c. $\exists x\,\exists y\,[x \text{ cites } y]$
d. $\exists x\,[x \text{ only cites Beaver}]$

So contrary to Büring's claim, the difference between the two mentions of Beaver here can be described in terms of Givenness *so long as you use the right definition of Givenness*. Schwarzschild's requirement that *all* constituents in a sentence be Given, even the ones above the NP level, is what makes his definition the right one here, because it is this requirement that lets us distinguish between the two mentions. Both are Given nouns; but only one is "Given all the way up" — a Given noun in a Given syntactic slot in a Given verb phrase in a Given sentence. The other mention is appearing for the first time in a new syntactic slot; for we've never before discussed the question of who Beaver *cites*, only the question of who he's cited *by*. This makes a difference in N-marking, and it explains why we arrive at the structure in (59-b) which is parallel to all the other examples of secondary accent.

In short, we have arrived at a model that attributes the same structure to all the examples of secondary accent attested in the literature, and that provides semantic or pragmatic motivation for each part of the structure wherever it occurs. We feel that this represents progress over the domain size accounts, which attribute unnecessary structural differences to prosodically identical sentences, and which must use arbitrary, semantically unmotivated manipulations of focus domain size in order to capture all the data.

Nevertheless, there are some problems with the formal version of our model. These problems pertain not to secondary accent, but to a much older set of data that has not been widely discussed in recent years; and these same problems arise in other formal theories of focus we have discussed in this paper — although, interestingly, they do *not* arise in the informal model we discussed in section 2. This suggests that in moving to a more formal model, we have thrown the baby out with the bathwater, where by "baby" we mean "the useful notions of predictability and importance, as understood in psychological rather than purely semantic terms." We now consider these problematic examples, and comment on our prospects for handling them — which look quite bright to us, so long as we are willing to go back to our psychological understanding of predictability and importance.

6 Discussion: extending the notions of importance and predictability

We began this paper talking about "predictability" and "importance" quite generally. But since it is difficult to say in general what a speaker will find predictable or important, we quickly turned to a set of well-behaved special cases in which only a few forms of predictability and importance can be found. In the short run, this was a useful move, since it let us formalize our account of

focus using existing techniques and concepts. But in the long run a theory of focus will need to be able to cope with the full generality of human discourse.

We think the account we have sketched here can be extended into a suitably general one. To that end, let us reconsider the psychological notions we offered in section 2, where "predictability" meant predictability *for a human audience*, based on their own expectations and knowledge as well as the contents of the context, and where "importance" meant importance *to a particular speaker*, based on that speaker's own goals. Building a proper theory on these notions, rather than the informal sketch we pursued in section 2 of this paper, will require borrowing some techniques from experimental psychology, and statistical techniques of the sort that are the mainstay of modern computational linguistics. Here we would like to briefly discuss how we think this could be done, and why we think it would be worth doing.

We said a moment ago that we had been restricting ourselves to special cases. For instance, in the examples considered in this paper, there are no discourse referents from the middle of Prince's information status taxonomy. Everything is either given or brand-new; nothing must have its existence inferred based on world knowledge. In fact, basically all the information speakers wish to convey, in the examples we've been considering, is strictly entailed by their utterances. So not only is it unnecessary to infer the existence of discourse referents, it is unnecessary to make educated guesses about the events and situations they participate in. This all meant that predictability was quite tractable in these cases, for we excluded any situation in which a constituent might be predictable based on world knowledge, probablistic "common sense," or unspoken common knowledge between conversation partners. The role of importance in these cases was similarly tractable, for we considered only a small set of discourse goals, formalized using Rooth's alternative-semantic account of F-marking, and not any private or personal goals or preferences that the speaker might have.

We observe that we are not alone in restricting our attention to cases like these. They dominate the recent literature on focus. The papers by Büring, Rooth and Selkirk that we discuss here mention no examples where real-world knowledge of any sort would be needed in order to make the right predictions about prosody, and none where private information or personal goals need to be taken into account.

Obviously there is something attractive about this restricted set of examples. Their formal analysis does not require particularly sophisticated mathematical techniques, even if it can be difficult to get the details right. For instance, Schwarzschild's Givenness constraint, which is at the core of our account of N-marking, is based on the idea that constituents are Given only if they are strictly entailed by what has come before. This constraint worked as well as it did in part because we did not consider any examples where a constituent could be considered predictable based on anything less than strict entailment.

The trouble is that plenty of examples exist which are not so well behaved.

(62) a. I have a POINT to make.
b. I have a POINT to EMPHASIZE.

(63) a. I've got to go SEE a guy.
b. I've got to go SEE a FRIEND.

(64) [Calling upstairs to a roommate]
a. The PHONE is ringing.
b. The PHONE is BROKEN.

(65) [Walking out of a phone booth in disgust] The PHONE is broken.

(66) [Nightclub owner, explaining why the place is a mess]
a. The POLICE came and raided us last night.
b. The POLICE came and PARTIED here last night.

These examples are adapted from Bolinger (1972), which includes many further examples along the same lines. We freely admit that the formal version of our theory in section 3 gives the wrong predictions on them. For if we N-mark constituents based on what is strictly entailed by the context, then we must N-mark the (a) and (b) sentences in each pair identically. Meanwhile, there is no reason to F-mark any constituents in these sentences at all. (Assume that none are given in answer to a specific question, or are intended contrastively. It is clear that none contain any exclusive particles.) And this means we cannot predict any of the differences in intonation shown above.

For instance, both *friend* in (63-b) and *guy* in (63-a) must be N-marked, because neither is entailed by what has gone before. We have $\exists x$ [I've got to go see x], but x need not be a friend or a guy.[15] There is no other difference between the sentences, and no F-marking whatsoever. Consequently we predict no difference in prosody — and this is a mistake.

It is, moreover, a mistake that we share with the other contemporary accounts of focus that we've been discussing. On Selkirk's account, when sentences have no F-marking, it is their G-marking that determines their accent — and (63-b) and (63-a) are identically G-marked, since both *friend* and *guy* count as discourse new. Thus, she too predicts that they should be accented identically. Büring and Rooth fare just as badly, since in sentences with no F-marking, they use a default rule based on intonational phrase structure to predict prosody. With no F-marking here, and no principled reasons we can see not to phrase each pair of sentences identically, there is no way for them to make the right predictions either.

Bolinger's own account of these contrasts is instructive. He maintains that *friend* is accented, and *guy* is not, precisely because *friend* is informative under the circumstances in a way that *guy* is not. He does not go so far as to say that the meaning of *guy* is logically necessary in this context — and as we've seen,

[15] This is true even if we restrict ourselves to the "meet with people" sense of *see* and the gender-neutral sense of *a guy* on which it essentially means "a person." One could have to go see a branch of one's family, a half-dozen students, or a jury of one's peers, and none of these is *a guy* on any honest reading of the phrase.

that claim would be untenable if he'd made it — but he does insist that there is a useful sense in which it is just not news.

Likewise, there is some sense in which *raided* is predictable in a context containing the word *police* and a trashed nightclub, although a notion of predictability based on entailment will not capture this because it is logically possible that the police have come for some other purpose. Suppose, in our first example with the poor beleaguered nightclub owner, he had simply shaken his head and said *The police...* before trailing off. On the one hand, if you had to guess what happened — or, equivalently, to predict how he would finish his sentence if pressed to do so — you might be more likely to predict a raid than a wild Fraternal Order of Police party. And on the other hand, if you do find yourself more likely to predict the party than the raid, we suspect you will have found something odd about the pronunciations given in (66). The same point can be made another way by comparing (66) and (67). Apparently, a difference in stereotypes about policemen and heavy metal musicians corresponds to a difference in intonation patterns on these sentences. We do not expect that any account that handles information structure purely in terms of entailments or strict antecedence will be able to predict this.[16]

(67) [Nightclub owner, the next week, explaining why it's a mess again]
MÖTORHEAD came and partied here last night.

In other words, this is not a special failing of the account we are presenting in this paper. It is something that many of us approaching prosody on the basis of formal semantics and pragmatics need to work on. But while these examples are problematic for existing formal accounts, we believe they are not fatal to the basic, informal approach outlined in section 2 of this paper. That is, we can go on discussing prosodic prominence in terms of predictability, importance and the competition for prominence between constituents, *so long as* we are willing to look to a broader notion of predictability which takes into account ordinary commonsense reasoning as well as strict entailment.

And similarly, there are gains that could be made by moving back to a broader notion of importance. We have already acknowledged that our account says nothing about the prosidic effects of real-world goals. But it is clear that they do have prosodic effects — that when we're giving instruction, or offering an explanation, we prosodically highlight information that we con-

[16]There is one strategy available that we have not considered above: we could stipulate that *emphasize*, *friend* and so on are F-marked, and that there is no need to recognize an informational difference between sentence pairs.

The only way to do this, though, would be to claim contrary to our stipulation above that these words are intended contrastively by the speaker. (We think it is undeniable that these sentences can be pronounced as written above even if they are not answering an overt question, and that none contain any exclusives.) The question then arises, what are they being contrasted *with*? The only possible answer is that they are being contrasted with some unstated alternative or set of alternatives which the hearer must accommodate, for nothing has been explicitly said that they might contrast with. And at that point, we have snuck our world knowledge, stereotypes, and so on in through the back door, by providing the hearer with an implicit contrastive alternative for *the police partied here* and none for *Mötorhead partied here*.

sider crucial and background information that we consider secondary. Consider (68), which you might expect to hear from a host on a cooking show. We might ask what is conveyed by giving primary accent to *slowly* rather than *saucepan*. There is no clear difference in predictability between the two words — browning butter is a prototypically slow, saucepan-based activity, so neither word will be surprising to someone who cooks (and either or both could equally well be surprising to someone who does not). But it is plausible that accenting *slowly* emphasizes the importance of the word; that the host, by accenting it, shows us that he cares a great deal about keeping us from heating the butter too quickly and burning it, and cares somewhat less about the shape of the vessel we heat it in.

(68) Brown the butter SLOWLY in a saucepan, // and then toss in the GARLIC...

Of course, the prosody in (68) is not the only way to pronounce that sentence. One could just as easily give *saucepan* the primary accent in its phrase instead of *slowly*, and even stranger pronunciations like (69) are imaginable (though pragmatically odd). The fact that there is room for variation here emphasizes something interesting about importance — it is to some extent under the speaker's control, at least insofar as the speaker can choose which goals to pursue in an utterance and how to prioritize them. If (69) is odd, it is perhaps because it reflects an odd choice of goals to prioritize. Hearing it, one would wonder "Why is *in* so important? What is he getting at?", and perhaps the underlying question is "Why has he decided that his top priority is getting me to put the butter *in* the saucepan? Does he think that I'm likely to put it on the *outside* of the saucepan by mistake?"

(69) Brown the BUTTER slowly IN a saucepan...

On lumping, F-marking accounts of prosody, sentences like (68) and (69) are treated as examples of contrastive focus. And perhaps there is something contrastive here — we could see (69) as evoking the alternative set { ‖in the saucepan‖, ‖outside the saucepan‖, ‖under the saucepan‖, ‖all over the saucepan‖... }, and expressing the importance of choosing the correct alternative from this set. This gives as a nice way of accounting for the oddity of (69); it seems that there is only one remotely plausible alternative in this set, and so it is a strange set of alternatives to emphasize. Much more reasonable to emphasize the set { ‖brown the butter slowly‖, ‖brown the butter quickly‖, ‖brown the butter with the utmost haste‖... }, as (68) could be said to do.

If we adopt the idea that real-world importance, as expressed in (68) and (69), can be explained in terms of alternative sets, we will not need to broaden our *semantic* understanding any. But we *will* need a broader sense of the pragmatics of alternative sets — of what speakers and hearers use them fore. Previously we have talked about using alternative sets to guide the way a discourse unfolds, by ruling out possible states of affairs, answering questions and so on. Now we are suggesting a very different use, in which alternative sets guide

a hearer's attention in planning a real-world activity by warning him or her away from possible actions. Same semantics, different speech act, different consequences.

So we have suggested reasons why we might adopt broader notions of predictability and importance. But adopting them as concepts is not enough. We will also need some way to operationalize them, so that they can be incorporated into testable, falsifiable theories. Fortunately, there are existing experimental paradigms which could be adapted for this purpose, and while there will no doubt be challenges in doing so, we do not imagine they will be insurmountable.

We have already, for instance, been using in our argumentation one test — the cloze test of lexical surprise (Taylor, 1953) — which might be used to operationalize the notion of predictability. Let subjects be shown a sentence with one constituent blanked out, or played a sentence with one constituent masked by noise. The extent to which the subject is able to correctly guess the hidden constituent provides a measure of its predictability. We might relax the task by crediting subjects with correct predictions if they guess a synonym or hyponym of a hidden word, or make it more difficult by forcing the subject to make a prediction based only on the left-hand context, and there are other modifications to the procedure that might be made.

This sort of word prediction task is usually computationally modeled — indeed, when computational linguists talk about a "language model," what they mean is a system that is able to make work predictions given some amount of context — but there is some precedent for giving it as a task to human subjects, either as a way of balancing psycholinguistic stimuli by ensuring they are equal in practical predictability, or as a way of setting a baseline for the performance of computational language models' performance (Lesher et. al. 2002). There have been plenty of experiments in modeling prosody using computed predictability values (see for instance Pan and McKeown (1999), Pan and Hirschberg (2000), which discuss the task of computing a word's predictability given its immediate context for use as a feature in accent prediction), but none to our knowledge using actual tests of human predictability.

Similarly, it strikes us that importance could be operationalized using the techniques of experimental economics. If importance really is importance *as regards meeting a goal*, then we might expect speakers to place greater value on getting important constituents across to their hearer. By assigning subjects specific communicative goals, and incentivizing them with cash rewards for meeting those goals as experimental economists do, we should be able to determine how important any given constituent is in the speaker's eyes towards the fulfilment of the assigned goal. For instance, imagine an experiment in which participants worked together towards some cooperative goal by exchanging audio-recorded messages, and were rewarded in cash for timely completion of the goal. Suppose the experimenter occasionally intercedes, threatening to wipe out part of a recorded message unless its speaker forfeits a certain amount of cash. We can then compare, say, the prosodic qualities of expressions taken to be worth more than a quarter with the qualities of those taken to be worth

less.

Once again, there is some precedent for this sort of experiment in linguistics – the Watson et al. (2008) tic-tac-toe experiment, in which prosody is predicted based in part on what we might call a *computed* importance value: importance is assigned a priori to winning moves. But we are not aware of any attempts at predicting prosody based on speakers own *judgments* of an expression's economic value.

In conclusion, we have proposed an approach to intonational meaning which is designed to predict not only which constituents may be accented, but also the degree of accent. Whereas early studies of focus idealized by considering only a binary distinction (accented or not), we have considered data which motivates at least three levels of accenting (primary, secondary, or none). Our hypothesis is that the degree of accent an expression has relative to others is positively correlated with the degree of communicative significance it has relative to others. We analyzed communicative significance as a composite involving predictability and pragmatic importance, notions which are natural extensions of existing concepts from linguistic theory, i.e. Given/New and focus. Let us end by noting that our hypothesis of a correlation between degree of accent and degree of communicative significance does not require that either of these notions are categorical. In fact, we know of no evidence from production or perception studies which would imply that degree of prominence is inherently categorical, and we see no reason why communicative significance should be categorical. We suggest that in the future it might be useful to consider the possibility that both degree of prominence and communicative significance are continuous variables.

References

Beaver, D. and Clark, B. (2008). *Sense and Sensitivity: How Focus Determines Meaning*. Blackwell.

Beaver, D., Clark, B., Flemming, E., Jaeger, T. F., and Wolters, M. (2007). When semantics meets phonetics: Acoustical studies of second occurrence focus. *Language*, 83(2).

Bolinger, D. (1972). Accent is predictable (if you're a mind-reader). *Language*, 48(3):633–644.

Büring, D. (2008). Been there, marked that: A tentative theory of second occurrence focus. ms, UCLA, to appear in a volume edited by Kanazawa, Makoto & Christopher Tancredi.

Calhoun, S., Carletta, J., Brenier, J., Mayo, N., Jurafsky, D., Steedman, M., and Beaver, D. (to appear). The NXT-format Switchboard Corpus: A rich resource for investigating the syntax, semantics, pragmatics and prosody of dialogue. *Language Resources and Evaluation*.

Féry, C. and Samek-Lodovici, V. (2006). Focus projection and prosodic prominence in nested foci. *Language*, 82(1):131–150.

Godfrey, J., Holliman, E., and McDaniel, J. (1992). SWITCHBOARD: Telephone speech corpus for research and development. In *Proceedings of ICASSP-92*, pages 517–520.

Jacobs, J. (1991). Focus Ambiguities. *Journal of Semantics*, 8:1–36.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3):243–276.

Pan, S. and Hirschberg, J. (2000). Modeling local context for pitch accent prediction. In *Proceedings of ACL-00*.

Pan, S. and McKeown, K. (1999). Word informativeness and automatic pitch accent modeling. In *Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*.

Partee, B. H. (1999). Focus, Quantification, and Semantics-Pragmatics Issues. In Bosch, P. and van der Sandt, R., editors, *Focus: Linguistic, Cognitive, and Computational Perspectives*, pages 213–231. Cambridge University Press.

Prince, E. (1981). Toward a taxonomy of given-new information. In Cole, P., editor, *Radical pragmatics*, pages 223–256. Academic Press, New York, NY.

Roberts, C. (1996). Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. *OSU Working Papers in Linguistics*, 49. Papers in Semantics.

Rooth, M. (1992). A Theory of Focus Interpretation. *Natural Language Semantics*, 1:75–116.

Rooth, M. (2010). Second occurrence focus and relativized stress F. In Féry, C. and Zimmerman, M., editors, *Information structure: Theoretical, Typological and Experimental Perspectives*. Oxford University Press.

Schwarzschild, R. (1999). Givenness, avoidf and other constraints on the placement of accent. *Natural Language Semantics*, 7(2):141–177.

Selkirk, E. O. (2002). Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in english. In Bel, B. and Marlin, I., editors, *Speech Prosody 2002: Proceedings of the First International Prosody Conference*, pages 643–646, Aix-en-Provence. Laboratoire Parole et Langage.

Selkirk, E. O. (2007). Contrastive focus, givenness, and the unmarked status of discourse-new. In Féry, C., Fanselow, G., and Krifka, M., editors, *Interdisciplinary Studies on Information Structure 6*, pages 125–145. Universitätsverlag.

Taylor, W. (1953). Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, pages 415–453.

Vallduví, E. and Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics*, 34(3):459–520.

Watson, D., Arnold, J., and Tanenhaus, M. (2008). Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 106(3):1548–1557.