

## When Semantics meets Phonetics:

### *Acoustical studies of Second Occurrence Focus\**

David I. Beaver, *Stanford University*

Brady Z. Clark, *Northwestern University*

Edward S. Flemming, *Massachusetts Institute of Technology*

T. Florian Jaeger, *Stanford University*

Maria K. Wolters, *University of Edinburgh*

Contact: dib@stanford.edu

December 2004

#### **Abstract.**

A second occurrence (SO) focus is the semantic focus of a focus sensitive operator (e.g. *only*), but is a repeat of an earlier focused occurrence. SO foci are intonationally distinct from the original occurrence of the material, and are often claimed to lack any intonational marking, e.g. pitch accent (Partee 1999). This apparent dissociation of semantic and intonational focus is commonly used as an argument against theories which make strong assumptions about the representation of focus at the syntax-semantics interface, e.g. Chomsky (1976), Jacobs (1983), Krifka (1992), Rooth (1985) and von Stechow (1989). Contrarily, the phenomenon has been argued to support pragmatic theories of focus, e.g. Roberts (1996), Schwarzschild (1997).

We report on the first systematic production and perception experiments to show that SO foci occurring after a nuclear accent are, as Rooth (1996b) has claimed, prosodically marked. We find (i) there is no mean pitch rise on SO foci; (ii) SO foci are marked by longer duration and greater intensity; and (iii) listeners are able to detect the difference between SO foci and non-foci. On the basis of these results, we argue that SO focus is compatible with theories of focus interpretation that it has been claimed to contradict. We also show that the phenomenon is of significance for intonational phonology, specifically as regards prosodic prominence in the post-nuclear domain (e.g. Ladd 1996).

**Keywords:** focus, intonation, prosody

---

\* This work was partially funded by grants from the a Stanford University Office of Technology Licensing and by Scottish Enterprise, the latter via the project *Sounds of Discourse*. The paper has benefited from commentary of Language reviewers and editors, from discussions with Katy Carlson, Bob Ladd and Mark Steedman, from audiences at conference presentations at the Linguistic Society of America annual meetings at Atlanta in 2003 and Boston 2004, the Phonologie und Phonetik Workshop at Potsdam in 2004, and the Texas Linguistic Forum at Austin in 2004, and also from seminar audiences at Cornell University, Edinburgh University, the Massachusetts Institute of Technology, New York University, Northwestern University, Stanford University, and the University of California at Santa Cruz. Our greatest intellectual debt is to Mats Rooth, whose insightful observations motivated us to perform an experimental study.

## 1. Introduction

Phoneticians and phonologists use semantic criteria to provide independent evidence for distinctions they make. If two productions are consistently associated with distinct meanings, then the phonetic difference between those productions must be phonologically significant. Semanticists, by contrast, tend to rely on the written word for much of their study, and so take phonetics and phonology for granted. There is at least one area, however, where semanticists cannot rely on the written word, and that is the study of focus, as well as related information structural phenomena like topic and contrast. Writing systems do not provide any consistent representation of the prosodic features which English, like many other languages, uses to mark these information structural phenomena. Consequently it seems that the study of information structure is an area where semanticists are dependent on the work of phoneticians and phonologists to provide a clear account of what types of marking speakers use. If only it were that simple!

Despite a great deal of progress in intonational phonology in the last few decades, Ladd (1996) argues that the development of phonological theories of intonation is hindered by an absence of clear evidence as to what the underlying semantic distinctions are: “For the present, proposals about intonational meaning are not a reliable source of evidence on intonational phonology.” (Ladd, 1996, p.102)<sup>1</sup> This suggests that progress in the phonological study of intonation and in the semantic study of information structure may benefit from collaborative work between semanticists and phoneticians/phonologists.

It has been known that intonation affects truth-conditional meaning since at least the work of Paul (1888:312ff). Consider our examples in (1), where italics indicate prosodic prominence (on either *money* or *Bill*), and this emphasis is standardly described as marking focus. In English, focus is typically marked by a nuclear pitch accent; i.e., the last pitch accent in a phonological phrase — see Cohan (2000) and Ladd (1996:45-46). The two different prosodic realizations of the same string of words yield truth-conditionally different semantic interpretations. In a situation where Jan gives Bill and

---

<sup>1</sup> The reader should set Ladd’s comments against a background of significant advances in the study of intonation in the previous twenty years, both as regards theoretical innovation and as regards empirical work using elicited and naturally occurring data. These developments are most clearly visible in the development of transcription standards for English intonation, e.g. the ToBI system developed from Pierrehumbert (1980) (as well as for comparable work on other languages). A goal of such work has been to analyze speech corpora, and several corpora now have intonational transcriptions, e.g. the BRN corpus (Ostendorf et al., 1995). Anyone who has studied intonational marking in such corpora, and who has also studied the literature on the interpretation of focus, will realize that there remains a gulf between the two. Pierrehumbert and Hirschberg (1990) provide the best known attempt to bridge the gulf, although their model does not incorporate the phenomena which will concern us here.

Malachi money, and gives nobody anything else, (1a) is true while (1b) is false.

- (1) a. Jan only gave Bill *money*.  
 b. Jan only gave *Bill* money.

Many standard approaches to grammar separate phonology and semantics into components which cannot exchange information directly, as in Figure 1.<sup>2</sup>

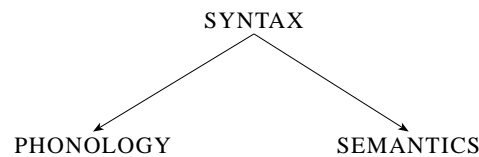


Figure 1. Generic depiction of core grammar architecture

In this standardly assumed architecture, the relationship between intonation and meaning is mediated by syntax. Therefore, many authors from Halliday (1967) and Chomsky (1972) on have concluded that there are syntactic features which encode prosodic prominence. Chomsky (1976), for example, suggested that all focused phrases move outside their base position in syntax. Related ideas are present in recent work in the Principles and Parameters framework; see e.g. Kayne (1998) and Tancredi (1990). Many other authors also postulate some syntactic effect of intonational focus, though they might use features or derivations in a quite different way; see e.g. Rooth (1985), von Stechow (1989), Krifka (1992), and Steedman (2000).

What all these accounts have in common is that they accept the premise that there is a grammatically mediated interaction between focus and meaning, albeit that this linkage may involve several stages. Yet recently some have suggested that the interaction between focus and meaning is not grammatically mediated at all; see Büring (1997), Dryer (1994), Kadmon (2001), Roberts (1996), Schwarzschild (1997), Rooth (1992), and Williams (1997). These authors have suggested instead that many effects of focus on meaning might not involve syntax at all, but instead should be explained in terms of pragmatics. This line of work strikes such a radically discordant note with the approaches listed earlier that it is worth exploring whether the claims

<sup>2</sup> The architecture assumed in the Minimalist Program, e.g. Chomsky (2000), diverges from that in prior work, as does much of the terminology. It remains the case that the phonological processes at *PF* (*phonetic form*) and the semantic processes at *LF* (*logical form*) do not exchange information directly. Furthermore, if we take *syntax* to refer collectively to all the processes occurring between the *Lexicon* and *LF* (rather than to a distinct level of representation), the architecture assumed by Minimalists can still be taken as a special case of that in Figure 1.

of the pragmaticists are justified. The current paper concerns one particular claim which the pragmaticists have used as the cornerstone of their argument. Specifically, they have claimed that in certain environments the interpretational effects normally attributed to focus are found even without the intonational prominence that marks focus, and thus that those interpretational effects need to be accounted for by a mechanism that is independent of such focus marking.

Let us examine a little more closely how the phenomenon we will study here bears on the nature of grammatical interfaces, since this an issue of potential interest across several linguistic subfields. In traditional architectures, interfaces between modules are presumed to be tightly constrained: we may think of interfaces metaphorically as narrow channels through which only certain types of information may pass. Specifically, it is common to think of the function of the morpho-syntax/phonology interface primarily in terms of segmental information, the information that determines which morphemes are present and how those morphemes are grouped. The place of supra-segmental information in a modular grammar is unclear and rarely discussed. Pragmatic theories of focus could be seen as part of an enterprise of keeping the interfaces between phonology, morpho-syntax and semantics as narrowly circumscribed as possible. These pragmatic theories lean on extra-grammatical mechanisms to make up the shortfall, much as in the model of Grice (1989).<sup>3</sup> If the interpretive effects of focus could be explained pragmatically, then the phenomenon of focus would provide us with little insight into how supra-segmental information is represented at the morpho-syntax/phonology and semantics/morpho-syntax interfaces. But if there is a grammaticized connection between focus and meaning, then that places minimal constraints on the supra-segmental information that must be represented at the morpho-syntax/phonology interface. It also places a lower limit on what information must be passed between syntax and semantics. What the current paper will show is that the major argument which has been put forward in favor of pragmatic theories of focus is empirically flawed. As far as we can tell, there is no evidence that supra-segmental information like that marking focus is treated differently than segmental information as regards the architecture of the grammar: both contribute to meaning, and there is no *a priori* reason to view them as contributing in substantially different ways.

The specific phenomenon at issue in this paper has been studied extensively in the semantic literature and involves expressions which are *focus sensitive*. An expression is focus sensitive if its interpretation is dependent on the placement of focus, keeping in mind that the linguistic realization

---

<sup>3</sup> But note that it is controversial to what extent pragmatics may itself be conventionalized and possibly grammaticized — see e.g. Levinson (2000) for discussion.

of focus (e.g., by prosody, syntactic position or morphology) varies cross-linguistically.

All theories of focus sensitivity agree that expressions like *only* and focus interact in linguistic contexts like that exemplified by (1). However, as a special case of the controversy as to whether the interpretation of focus is mediated via syntax, there is disagreement as to how grammaticized the relationship between *only* and its associated focus is (Partee 1999:215ff). Does the lexical entry of *only* stipulate association with a focused constituent in its syntactic scope, or are there contexts in which the interpretational effects illustrated by (1) occur without focus being present either phonologically or syntactically? This question provides the main distinction between different contemporary theories of focus sensitivity.

A number of examples have been cited in the literature involving an apparent dislocation between the interpretation of *only* and the positioning of prosodic prominence. Most of these examples fall under the rubric of *second occurrence focus*, where a repeated focused item apparently lacks a pitch accent.<sup>4</sup> The example in (2), adapted from Partee (1999:215), illustrates second occurrence focus. The two sentences are to be read as a dialogue. The convention of indicating the focused item with a subscripted “F” is widespread in the syntactic and semantic literature, and, as will be discussed shortly, is usually taken to indicate marking at a syntactic level, while our marking of the second occurrence focus with “SOF” allows us to remain (temporarily) agnostic as to whether the item described as a second occurrence focus is in fact focused at a syntactic level or any other.

(2) *Me*: Everyone already knew that Mary only eats [vegetables]<sub>F</sub>.

*You*: If even [Paul]<sub>F</sub> knew that Mary only eats [vegetables]<sub>SOF</sub>, then he should have suggested a different restaurant.

Partee (1999) summarizes the problem succinctly. If *only* is a focus sensitive operator (i.e., needs an intonationally prominent element in its scope) then the two occurrences of *only eats vegetables* in (2) should have the same analysis. However, if there is no phonological reflex of focus in the second occurrence of *vegetables* then this leads to the notion of “phonologically invisible focus”. The notion of inaudible foci “at best would force the recognition of a multiplicity of different notions of ‘focus’ and at worst might lead to a fundamentally incoherent notion of focus” (Partee 1999:215–216).

Following up on observational evidence presented by Rooth (1996b), in this paper we present experimental data concerning precisely the cases where the pragmaticists have claimed interpretational effects without intonational

<sup>4</sup> See Gussenhoven (1984), Hajičová (1973), Hajičová (1984), and Taglicht (1984) for related discussion.

marking. We will show that in these cases the claim that there is no intonational marking is wrong. Specifically, we present the results of multi-subject production and perception experiments in which we examine the acoustic correlates of second occurrence focus in the scope of operators *only* and *always* which are both assumed to be focus sensitive in the bulk of prior literature.<sup>5</sup> Our results confirm that there are acoustic correlates involving duration and intensity of the focused item, and show that the acoustic correlates are perceptible. These conclusions make clear that the phenomenon of second occurrence focus does not provide an argument that current syntactic and semantic theories of focus are “fundamentally incoherent”. Indeed, the results could be taken as an embarrassment for pragmatic theories of focus, since it is unclear on these accounts why speakers would mark foci in contexts where that marking is pragmatically unnecessary.

For semantics and pragmatics, our results contribute to a long running debate. But the phenomenon of second occurrence focus is not well known in phonetics and phonology. The results are of relevance to these fields because they shed light on the acoustic realization of focus, and the acoustic properties of the *post-nuclear* domain, the part of an intonational phrase that follows the nuclear accent. As will be explained, the results (i) imply increased role for phrasal stress in marking focus, while showing that pitch accents are not required, and (ii) suggest that although post-nuclear deaccenting restricts the appearance of pitch accents, phrasal stress distinctions must be permitted in the post-nuclear stretch. This latter point buttresses arguments Ladd (1996) has put forward for the existence of post-nuclear prominence, although, our conclusions differ in detail from his.

The paper is organized around three hypotheses, the *Marking Hypothesis*, the *Non-Standard Marking Hypothesis*, and the *Perceptibility Hypothesis*, all to be detailed in section 3. The Marking Hypothesis says that second occurrence foci are prosodically distinguished from non-foci; i.e. second occurrence focus is prosodically marked. The Non-Standard Marking Hypothesis concerns the manner in which second occurrence focus is marked, and, specifically, how it is distinct from ordinary focus marking. The Perceptibility Hypothesis says that second occurrence focus marking is perceptible.

Before presenting the hypotheses in full (section 3), we introduce the technical background of the paper as regards the syntax, semantics, pragmatics, and phonetics/phonology of focus (section 2). The production and perception experiments are presented in one section each (sections 4 and 5, respectively), together with their associated results and discussion of the interpretation of

---

<sup>5</sup> But c.f. discussion in section 5: work conducted since our original production study indicates that *only* may be more prototypical as a focus sensitive operator than *always*. To look ahead, the experiments reported here did not yield significant effects in support of this distinction, but this may be because our experimental design was not intended to reveal a distinction between the two.

those results. There follows a more general discussion connecting the results to wider theoretical concerns (section 6) and, finally, some concluding remarks and suggestions for future work (section 7).

## 2. Background

The following discussion will involve a range of phonetic, phonological, syntactic, and semantic notions of focus. These notions include *prominence* (a psychoacoustic notion), *phonological focus marking* (which in prior literature is sometimes referred to as *prosodic* or *intonational* focus), *F-marking* (which is syntactic), and *semantic focus* (an aspect of the representation of meaning).<sup>6</sup> Figure 2 illustrates the five-way distinction we employ in the remainder of the paper.

We must now explain why the absence or presence of certain acoustic properties comes to be a decisive factor for theories of focus. Section 2.1 provides an introduction to the notion of semantic focus, specifically focus sensitivity. We discuss syntactic F-marking in section 2.2 and discuss competing theoretical approaches towards F-marking in section 2.3. Section 2.4 links the absence or presence of F-marking to the phonology of focus (i.e. focus marking). Section 2.5 summarizes the *argument from second occurrence focus*, i.e. the theoretical claim that motivated the experiments presented in this paper. Section 2.6 then provides the necessary background on phonological focus marking and its phonetic correlates. Finally, section 2.7 summarizes what was known about the phonetics and phonology of second occurrence focus at the onset of our research. This sets the ground for the formulation of our hypotheses with regards to the nature of second occurrence focus marking in section 3.

---

<sup>6</sup> We are not committed to all the above notions being necessary components of a theory of focus: we merely seek to avoid terminological and conceptual confusion in an investigation which spans multiple sub-disciplines of linguistics. To avoid one further potential source of confusion, it is important to note that the term *focus* has been used to refer to a variety of semantic phenomena in addition to the association with a focus sensitive operator (as in (1) above), including contrastive focus (a), focus in question-answer pairs (b), and focus in clefts (c), or left dislocations.

- a. I said I wanted [waffles]<sub>F</sub> not [pancakes]<sub>F</sub>.
- b. A: What would you like for breakfast?  
B: I would like [waffles]<sub>F</sub>, please.
- c. It was [Karl]<sub>F</sub>, who ate all the waffles.

These foci have sometimes been assumed to share phonological features. This is, however, far from obvious (cf. Bartels and Kingston 1984, Hedberg and Sosa 2001, and Hedberg 2003).

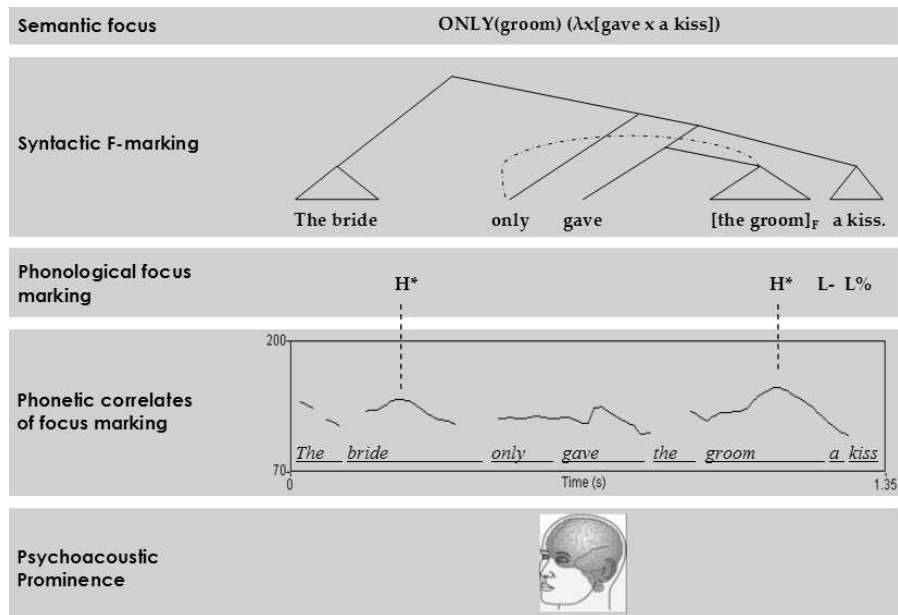


Figure 2. Focus at different levels of representation

### 2.1. THE PHENOMENON OF FOCUS SENSITIVITY

Focus sensitivity can be thought of as a mechanism that allows two place operators to select their arguments. Consider the English exclusive *only*, which is sometimes treated as a universal quantifier. The parallel with the canonical universal *every* is seen in examples like (3a,b), which involves *only* within an NP (hence “NP-*only*”). (4) relates the semantic interpretation of (3a) to (3b).

- (3) a. Only felines are immortal.  
 b. Every immortal is feline.
- (4) ONLY(FELINE)(IMMORTAL)  $\equiv$  EVERY(IMMORTAL)(FELINE)

In (4), ONLY represents the meaning of *only*. Let us call the first argument of ONLY its *restrictor*, i.e. FELINE, and the second its *scope*, i.e. IMMORTAL. Then *only* is just like *every* except that the operator we use to represent its meaning has its arguments reversed.<sup>7</sup> In the case of NP-*only*, configuration

<sup>7</sup> For arguments that *only* and *every* are related in this way semantically, see e.g. de Mey (1991) and Horn (1996). Note that on some approaches, the focus of *only* would appear in *only*'s specifier at some point in the syntactic derivation (Kayne, 1998; Krifka, 2004). In this case, NP-*only* might combine first with the argument we call here its scope, and second with the argument we call here its restrictor, making our terminology less than transparent for



fully determines the restrictor and scope, but, as we will see, the situation is more complex for *VP-only*.

The standard mechanisms for argument selection are (a) syntactic configuration and (b) morphology (e.g. case marking). In English, the burden of selection falls primarily to configuration. However, focus sensitive operators like *VP-only* are unusual in that their arguments are not determined configurationally, at least not in surface form. Indeed, many analyses do not even assume that the arguments of focus sensitive operators need to be continuous surface constituents. Consider (5a,b) (= (1a,b) with explicit F-marking), and their intended meanings (6a,b).

- (5) a. Jan only gave Bill [money]<sub>F</sub>.  
 b. Jan only gave [Bill]<sub>F</sub> money.
- (6) a. ONLY(MONEY)( $\lambda x$ [JAN GAVE BILL  $x$ ])  
 i.e. everything Jan gave Bill was money.  
 b. ONLY(BILL)( $\lambda x$ [JAN GAVE  $x$  MONEY])  
 i.e. everyone Jan gave money was Bill.

If (5b) is analyzed as in (6b), then the restrictor is the interpretation of *Bill*, while the scope is derived by interpretation of *Jan gave [...] money*, which is discontinuous.<sup>8</sup> Common approaches to argument selection neither allow for discontinuous arguments, nor explain how focus could play any role. The puzzle of focus sensitivity is this: what mechanisms explain the correlation between the differing placement of focus in (5a,b) and the differing interpretations in (6a,b)?

A wide range of English expressions have been analyzed as focus sensitive, including exclusives (*only*, as in the above examples, but also *just*, *merely*, *solely*), additives (*too*, *also*), scalar additives (*even*), particularizers (*in particular*, *for example*), intensifiers (*really*, *totally*), quantificational adverbs (*usually*, *always*), certain quantificational determiners (*many*, *most*), comparatives, certain sentence connectives (*because*, counterfactual *if-then*), negation, questions, and certain emotive predicates (*regret*, *be glad that*) (Rooth, 1996a; Hajičová et al., 1998).<sup>9</sup>

adherents of these approaches. It is presumably because of the non-standard way in which *only* combines with its arguments that terminology has not become standardized. Here we stick to a terminology that takes the analogy between NP-*only* and *every* as basic, and labels arguments in other cases by extrapolation from this case.

<sup>8</sup> It might be argued that JAN should be kept out of the argument of *only* in (6a,b), in which case we could think of the meaning of e.g. (5a) as underlyingly consisting of separate NP and VP meanings combined conjunctively: ONLY(MONEY)( $\lambda x$ [SUBJ GAVE BILL  $x$ ])  $\wedge$  SUBJ = JAN. The interpretation of the subject in sentences with VP focus sensitive operators will not be a concern of this paper. See Beaver and Clark (2003) for an analysis.

<sup>9</sup> It must be stressed that this is not a phenomenon peculiar to English. Indeed, many languages appear to have an even richer array of focus sensitive constructions than English:

Clearly not all focus sensitive expressions are quantifiers, which makes it inappropriate to label the arguments of those expressions as *restrictor* and *scope*. However, all the expressions can be thought of semantically as multi-argument operators, such that one argument position is filled by a focused expression. We will term this argument the *semantic focus* of the operator.<sup>10</sup> Thus, according to the terminology we introduced previously, the semantic focus of *only* is its restrictor, and the puzzle of focus sensitivity may be stated more broadly than above: how do language users associate phonologically focused constituents with the semantic foci of focus sensitive operators?

Let us summarize just what is in need of explanation. First, sometimes an operator can select for an argument (its semantic focus) even when the argument is separated from the operator. Second, there is no morphological case marking or agreement to facilitate the link between operator and argument, and neither is there any canonical configurational relationship in surface form that determines what the argument of the operator is. But there is a correlation between the selection of one of the operator's arguments and the placement of focus marking.

## 2.2. GRAMMATICAL MECHANISMS FOR FOCUS SENSITIVITY

It is natural to think of focus as having a function analogous to that of a morphological case marker: it signals which argument goes where. Thus, many theorists have assumed that a focused constituent is marked at the level of syntax, normally with a special "F" feature, and then postulated that a further grammatical mechanism explains focus sensitivity. The details of the mechanism depend on whether the F-marked constituent is moved or interpreted *in situ*.

Early proposals suggesting that F-marking triggers movement were those of Anderson (1972) and Chomsky (1976). In work within the Minimalist Program (Kayne, 1998), the focus sensitive operator would attract the F-marked *Bill* in (5b) to the specifier of a phrase headed by *only*.<sup>11</sup>

---

witness, e.g., the many Dutch particles lacking simple translation equivalents in English (Foolen, 1993). However, the claims we are attacking in this paper were made for English, and the paper is restricted to English alone.

<sup>10</sup> Note that *semantic focus* is logically distinct from the phonological notion of *focus marking*, although the experiments reported here suggest that a large class of semantic foci must be focus marked. See the introduction to section 2.

<sup>11</sup> In Kayne (1998), the entire VP *gave Bill a book* moves into the Spec of *only*. It is unclear what Kayne's preferred analysis of the semantic association with a focus within the VP would be. Note that Krifka's (2004) hybrid association with focus account is comparable to Kayne's but assumes that a separate semantic mechanism of focus association operates in tandem with the syntax.

*In situ* approaches became dominant in the semantics community after Rooth (1985) observed a range of problems with movement accounts.<sup>12</sup> But in order to achieve *in situ* interpretation, the semantics must depart quite radically from the now canonical compositional account of Montague (1974). Loosely, we may say that *in situ* approaches, while they do not move the F-marked constituent itself, do move the meaning of the F-marked constituent, passing that meaning from node to node on the syntactic tree until it becomes accessible to a focus sensitive expression like *only*. For example, Rooth's (1985, 1992) *alternative semantics* adds to every single constituent a second semantic value for the focus in addition to the standard one, while yet another modification to Montague Grammar is made in the Structured Semantics of von Stechow (1989) and Krifka (1992).<sup>13</sup> For current purposes, the details of the modified semantics are not important. What is important is that in some way the modifications mean that the semantics encodes which parts of the meaning come from focused constituents, and which parts of the meaning come from non-focused constituents. A focus sensitive operator like *only* may then be defined so as to act on this enriched encoding.

### 2.3. STRONG THEORIES OF FOCUS SENSITIVITY

The *in situ* and movement models described above explain the interpretational effects which we term focus sensitivity in terms of grammatical mechanisms conditioned on syntactic F-marking. Two questions arise as regards the predictions of a model of this sort. First, if a constituent in an appropriate position relative to a focus sensitive operator is F-marked, does it necessarily become the semantic focus of that operator? Second, and more relevantly to the current paper, is the semantic focus of a focus sensitive operator necessarily F-marked?

It is hard to find explicit discussions of these questions in the early literature on focus sensitivity. Nonetheless, we are not aware of any syntactic proposal that would allow material with no associated F-marking to undergo focus movement. As regards *in situ* approaches, it seems fair to say that prior to Rooth (1992), it had been widely assumed both that F-marked constituents were obligatorily interpreted as semantic foci, and that semantic foci were obligatorily F-marked. And such assumptions seem hardly surprising: if there are some environments where the semantic focus is not what is F-marked,

<sup>12</sup> Rooth's (1985) empirical arguments against movement-based accounts include the interaction of focus sensitive expressions with multiple foci and syntactic islands. Krifka (2004) suggest that focus sensitive operators associate through movement of the syntactic island that contains the associated focus, thus side-stepping some of Rooth's objections to movement-based accounts.

<sup>13</sup> Note that the analysis of focus given by Steedman (2000), although it involves Structured Semantics, does not depend on syntactic F-marking. For Steedman, intonation controls derivation rather than syntactic form.

then yet another mechanism must be postulated to explain focus sensitivity. The very fact that models of focus like those described above require significant revisions to syntax and/or semantics makes it unattractive to suppose that yet further mechanisms are used in some cases.<sup>14</sup>

It was Rooth (1992) who introduced a more nuanced view, one which allows for F-marking to be dissociated from the semantic focus. In a broadly methodological discussion, Rooth introduces a spectrum of possible theories of focus, ranging from *weak* to *strong*. The idea is that *weak* theories are those that involve the most stipulation and are thus only weakly explanatory. Specifically, a weak theory would involve stipulating item by item which words are members of the class of focus sensitive operators, and item by item exactly how each member of the class operates on the focus. Both the movement-based and *in situ* models sketched above would be classified by Rooth as weak, since no general principle constrains which expressions can be focus sensitive.

A strong theory, in Rooth's sense, would not involve any grammaticized mechanism relating focus to argument selection or semantic interpretation. Instead, the theory would have to rely on general principles of pragmatics and interpretation to derive its predictions as to the effects of focus. If, in spite of this handicap, and in spite of not allowing lexical specifications to make any reference to focus, a theory could still predict focus sensitive behavior, the theory would clearly deserve to be called *strong*.<sup>15</sup> But is a *strong* theory in this sense really a possibility? In fact, a number of candidate proposals for strong theories have been put forward, although none of them have been applied to a wide range of focus sensitive expressions.

The general template Rooth presented for strong theories allows that operators which have been termed *focus sensitive* have a free variable; the argument that we have been calling the semantic focus is introduced as a free variable lexically. An effect of focus is to make certain discourse objects salient, or to show that an already salient discourse object is to remain so. For example, *I only like [cheese]<sub>F</sub>* would be appropriate when the set of things the speaker likes is under discussion, and hence this set is salient. In a strong theory, there is no grammatical rule forcing such a salient discourse entity to become associated with the free variable, but this can optionally happen. On

<sup>14</sup> Of course, it is clear that in written language F-marking usually cannot be derived from surface form alone, which might lead us to wonder to what extent language users really need a special grammatical mechanism to mediate focus sensitivity. But then again, there are many obligatory distinctions in spoken language that are not available in standard orthography. An example is the stress pattern on some bisyllabic words which are ambiguous between noun and verb in standard written form (*to invite* versus *an invite*).

<sup>15</sup> Rooth also considers the possibility of what he terms an *intermediate* theory, one which allows for lexical stipulation, but only of a very limited sort. Von Stechow's (1994, 1995) account of quantificational adverbs may well be an intermediate theory, though it could also be taken as strong.

such a view, free variables involved in focus sensitivity are resolved pragmatically in much the same way as a pronoun optionally becomes associated with a salient discourse entity: Martí (2003) explores this parallel. The tricky task that proponents of strong theories set themselves is then to make sure that the right discourse object gets tied to the right free variable at the right time. We will not discuss the details of how pragmatic models might achieve this.<sup>16</sup> For current purposes, suffice it to say that all such models would predict the possibility that an F-marked constituent would not become the semantic focus, and that sometimes the semantic focus could be a constituent that is not F-marked.

It should now be broadly clear how the the strong/weak split relates to the question of whether there are environments in which semantic foci are not syntactically F-marked. Weak theories, in the form they have been stated, would lead us to expect obligatoriness of F-marking.<sup>17</sup> Thus they predict that there are no environments in which semantic foci are not syntactically F-marked. To complicate matters, it is possible to conceive of a theory in which the relationship between F-marking and interpretation was grammaticized but optional. We will not attempt to describe such a model in detail. It suffices to say that the more evidence we find for non-F-marked semantic foci, the more plausible strong, pragmatic theories are going to become, while contrarily, evidence of a compulsory link between F-marking and interpretation provides support for grammaticized models, despite Rooth's disparaging description of them as *weak*.

#### 2.4. RELATING THE STRONG/WEAK DISTINCTION TO PHONOLOGICALLY UNMARKED SEMANTIC FOCI

We are almost at the point where we can tie this discussion in with the main empirical phenomenon to be examined in this paper. But first, there is one further subtle distinction which we need to amplify: that between phonological focus marking and syntactic F-marking. We have described above how models may vary according to whether the link between interpretation and F-marking varies according to the environment, or is compulsory, but we have not said much about how F-marking relates to phonological focus marking (and, ultimately, to psycho-acoustic features such as prominence). The relationship is not straightforward. Often, a phonologically focus marked constituent appears to be F-marked, and that constituent may become the semantic focus of a focus sensitive operator. But it is also possible

<sup>16</sup> For pragmatic models which Rooth would classify as *strong*, the reader is referred to Vallduví (1990), Roberts (1995), Roberts (1996), Schwarzschild (1997), Kadmon (2001), Martí (2003) and Geurts and van der Sandt (2004).

<sup>17</sup> Beaver and Clark (2003) provide syntactic and semantic criteria for obligatoriness of F-marking, and show that F-marking is obligatory in the scope of some expressions which are usually taken to be focus sensitive.

for the F-marked constituent not to correspond directly to expressions bearing phonological focus marking. For example, the VP *have a dog* may be F-marked if *dog* is the only word which is phonologically focus marked.

Relating phonological focus to F-marking are theories of *focus projection*, which include the models of Selkirk (1995) and Truckenbrodt (1995). The question that we care about here is this: is the relationship between F-marking and phonological focus optional or compulsory? If there was considerable optionality in this relationship, then it would be very hard to ever get data on whether the relationship between F-marking and semantic focus was itself optional or compulsory. In effect, all we would know would be that somewhere between phonological focus and semantic focus, some optionality crept in, but we could not tell whether this should be thought of as occurring in the phonetics-phonology map, the phonology-syntax map or the syntax-semantics map. Focus projection remains an area rife with disagreement, but it suffices for current purposes that in models with which we are familiar, the following principle is adhered to:

**Focus Projection** An F-marked constituent must contain at least one phonologically focus marked sub-part.<sup>18</sup>

Given that F-marked constituents contain a phonological focus, we can now say that many weak, grammaticized theories of focus should lead us to expect not only a compulsory relationship between F-marking and interpretation, but a compulsory relationship between phonological focus marking and interpretation. Specifically, such theories predict that the semantic focus of a focus sensitive operator should correspond to a constituent that contains a phonological focus (Partee 1999:217). Of all the predictions made by theories of focus, it is this prediction that has come under heaviest attack in the semantics literature, and by far the most common basis for the attack is the phenomenon known as *second occurrence focus*.

## 2.5. THE ARGUMENT FROM SECOND OCCURRENCE FOCUS

We have already seen (an adaption of) an example of *second occurrence focus* from Partee:

- (2) A: Everyone already knew that Mary only eats [vegetables]<sub>F</sub>.  
 B: If even [Paul]<sub>F</sub> knew that Mary only eats [vegetables]<sub>SOF</sub>, then he should have suggested a different restaurant.

<sup>18</sup> Note that what we call F-marking corresponds to what others would describe as undominated F-marking. Selkirk (1995) distinguishes F-marked constituents that are not dominated by any other F-marked constituent by the label *Foc*. Schwarzschild (1999) introduces a constraint FOC (“A Foc marked phrase contains an accent”), which corresponds to the principle termed *Focus Projection*, above.

Let us repeat the analysis of second occurrence focus examples given in section 1, but using some of the concepts we have introduced. Sentence B contains two focus sensitive operators, *even* and *only*. The semantic focus of *even* is *Paul*, which is also a phonological focus. The semantic focus of *only* is *vegetables*. But Partee and others note that *vegetables* in its second occurrence, i.e. the occurrence in sentence B, can be felicitously uttered without any pitch movement, and hence, they suppose, it cannot contain a phonological focus. By focus projection, it follows that this occurrence of *vegetables* is not syntactically F-marked. Thus we appear to have a case of a dissociation between semantic focus and F-marking, contra what is expected on the movement and *in situ* accounts of focus sketched above.

It is this type of analysis which has allowed theorists to argue that second occurrence focus data counter-exemplifies weak theories of focus interpretation. As mentioned previously, many authors hold that second occurrence focus data provides evidence against the major *in situ* focus accounts (alternative semantics and structured meaning semantics). These include Valluví (1990), Rooth (1992), Dryer (1994), Roberts (1996), Büring (1997), Schwarzschild (1997), Williams (1997), Kadmon (2001), and Martí (2003), all of whom interpret second occurrence focus data as favoring strong accounts of association with focus phenomena over weak accounts.

We now present in a maximally explicit form the argument from second occurrence focus data against grammaticized theories of focus sensitivity — most arguments in the literature from second occurrence focus have basically the same structure:<sup>19</sup>

1. Weak, grammaticized theories of focus sensitivity require the semantic focus to be F-marked.
2. By focus projection, these theories also predict that the semantic focus should contain a phonological focus.
3. (a) Second occurrence foci are semantic foci,  
(b) but second occurrence foci contain no phonological focus marking.
4. Therefore we should prefer a strong, pragmatic account of focus.

This paper hinges on step 3b of this argument, which we will show to be unjustified empirically, and we now turn to a crucial assumption on which 3b rests. Most semanticists (a notable exception being Rooth (1996b), discussed below) have assumed that absence of pitch movement correlates with absence of phonological focus. But acoustically based observations of second occurrence focus data cannot directly record the absence of phonological focus

<sup>19</sup> See e.g. Rooth (1996b:206) for a similarly organized argument.

marking, so it is crucial to consider in more detail how phonological focus marking corresponds to the acoustic signal.

## 2.6. THE PHONETICS AND PHONOLOGY OF FOCUS IN ENGLISH

According to standard accounts of English intonation, focus marking typically involves the alignment of a pitch accent with the primary stressed syllable of the focussed item. A pitch accent is “a local feature of a pitch contour — usually but not invariably a pitch change, and often involving a local maximum or minimum” (Ladd 1996:45-46) — cf. Bolinger (1958) and Pierrehumbert (1980). In the widely adopted framework of Pierrehumbert (1980), pitch accents are represented phonologically as high or low tones (H\*, L\*) or combinations of tones (e.g. L+H\*), and are annotated with an asterisk to differentiate them from boundary tones, which associate to the edges of phonological phrases. A phonological representation of this kind is illustrated in Figure 2 above.

Although focus is most frequently marked by an H\* accent (e.g. Hedberg and Sosa 2001; Hedberg 2003), there are many exceptions to this generalization. For example, in yes-no questions, focus is often marked by an L+H\* accent (Hedberg and Sosa 2002). More specifically, focus is usually marked by a nuclear pitch accent, i.e. the last pitch accent in a phonological phrase — see Cohan (2000) and Ladd (1996:225ff). The nuclear accent is perceived as more prominent than preceding accents in the same phrase. For example, in an utterance of (5b), repeated below, there could be a pitch accent on *Jan*, but this first accent is perceived as less prominent than the nuclear accent that marks the focus, *Bill*.

- (5) a. Jan only gave Bill [money]<sub>F</sub>.  
 b. Jan only gave [Bill]<sub>F</sub> money.

This observation corresponds to Jackendoff’s (1972:229) generalization that focus is marked by “the main stress [...] in the sentence”. Jackendoff refers primarily to phrasal stress in the sense of Chomsky and Halle (1968), but it is clear that these stresses correspond to nuclear accents, as is discussed further below.

At the phonetic level, a pitch accent is primarily a pitch event, as the name suggests. Pitch is properly speaking an auditory property, but it corresponds fairly directly to fundamental frequency (f<sub>0</sub>), which is a measurable property of the speech signal. So the main acoustic correlates of a pitch accent are to be found in the f<sub>0</sub> contour. For example, the H\* pitch accent marking the focus in the example given in Figure 2 is realized by a local maximum in f<sub>0</sub>. However, it is well established that accented syllables also differ from lexically stressed but unaccented syllables in having greater duration and intensity (a



key acoustic correlate of perceived loudness) (Sluijter and van Heuven 1996a; Turk and White 1999). There is also evidence that accented vowels tend to have more extreme formant frequencies, and thus better differentiated vowel quality (de Jong, 1995; Harrington et al., 2000).

It is not clear that the greater duration and intensity of accented syllables are correlates of the pitch accent per se, instead they may be correlates of phrasal stress. Many models of intonation analyze the prosody of a phrase as involving two interacting components: the intonational melody and phrasal stress (e.g., Pierrehumbert 1980, Selkirk 1995). The intonational melody consists of a sequence of pitch accents and boundary tones, while the stress pattern involves the relative prominence of syllables in the phrase. However, stress and intonation are intimately related because pitch accents can only be assigned to the syllables that receive the strongest stresses in the phrase, and the nuclear accent falls on the main stress (Pierrehumbert, 1980). These constraints imply that pitch accented syllables are also stressed, but not vice versa. Given such a model it is possible to construe the greater duration and intensity of accented syllables as a correlate of the phrasal stress that necessarily accompanies the pitch accent, and thus treat pitch accents as purely pitch-based. This analysis is consistent with evidence that stress is marked by greater duration and intensity in the absence of pitch accent (Beckman 1986; Sluijter and van Heuven 1996a, 1996b).

Other analysts, such as Bolinger (1958), have argued that phrasal prominence is marked by pitch accent alone, so there is no role for a separate notion of phrasal stress (see Ladd 1996:45ff. for a detailed discussion of this dispute). In section 6, we will argue that the distinction between phrasal stress and pitch accent is helpful in analyzing the differences between the marking of Second Occurrence Focus and regular focus, but for now all that matters is that pitch accents are generally accompanied by greater duration and intensity, whether these properties are direct correlates of the pitch accent, or of the accompanying phrasal stress. Accordingly, duration and intensity can also play a role in marking focus. Ladd (1996:226f) has gone further, suggesting that certain foci (e.g. in the answer to a question) can be marked without a pitch accent, so that non-pitch measures may become the primary indicator of focus. It is thus natural to examine the role of non-pitch measures in second occurrence focus, and, as will now be discussed, we are not the first to do so.

## 2.7. THE PHONETICS AND PHONOLOGY OF SECOND OCCURRENCE FOCUS

We now turn from the general issue of how focus is marked to the more specific issue of how second occurrence focus is marked. It is generally accepted that second occurrence foci do not contain a pitch movement, although we know of only one systematic experimental study of pitch movement in

second occurrence focus — Bartels (1997), discussed below. But it remains possible that a phonological correlate of second occurrence focus corresponds acoustically to other measures discussed above, such as duration, intensity and vowel quality.

We presented in section 2.5 an explicit argument from the phenomenon of second occurrence focus, an argument which would be undermined if it could be demonstrated that second occurrence focus has phonological correlates. Rooth (1996b) presented suggestive evidence indicating just that, and Rooth's observations provided the impetus for the more systematic experimental studies which we report here.

Rooth recorded himself uttering two minimal pairs of SOF dialogues like (7). The VP containing the second occurrence focus was textually identical in all the examples he discussed. This allowed Rooth to compare second occurrence focus NPs (e.g., *Manny* in (7b)) to controls where the same NPs were not in focus at all (e.g., *Manny* in (7a)), but in which other factors are held constant.

- (7) a. A: Do you want Sue to only [name]<sub>F</sub> Manny today?  
       B: No. I only want [Eva]<sub>F</sub> to only [name]<sub>SOF</sub> Manny today.
- b. A: Do you want Sue to only name [Manny]<sub>F</sub> today?  
       B: No. I only want [Eva]<sub>F</sub> to only name [Manny]<sub>SOF</sub> today.

Rooth found that the pitch track was flat in second occurrence focus position for examples like (7), but that he could auditorily detect the marking of second occurrence focus in his own productions. Furthermore, he noted that this marking of second occurrence focus is visible in the small sample of his productions for which he presents waveforms and spectrograms. In these productions, although no major pitch movement is visible on second occurrence focus expressions, these expressions have greater duration and absolute intensity than their non-focussed counterparts.

In related work, Bartels (1997) studied non-pitch correlates of prominence (relative syllable lengthening and amplitude) on second occurrence expressions in a multi-subject production study. She found that second occurrence foci differed from ordinary foci not only with regard to pitch, but also amplitude and duration. While Bartels demonstrated that a systematic experimental approach to second occurrence focus phenomena was possible, her conclusions concerned the different realization of second occurrence focus from regular focus. She did not add controls in which the test words were not in focus at all, i.e. in which the test words were neither foci nor second occurrence foci (Bartels 1997:24). Thus her experiments do not determine whether second occurrence focus is marked prosodically or what its correlates might be, but merely establish that if second occurrence focus is marked, then, on several acoustic dimensions, it is marked differently from ordinary focus.

### 3. Three Hypotheses

Are there, as Rooth suggested, acoustic correlates of focus in second occurrence focus position? Rooth's work suggests the hypothesis in (8). If this hypothesis is correct, and under the reasonable assumption that phonological focus marking can be operationalized in terms of these acoustic correlates, the argument from second occurrence focus unravels. In Section 4, we report on a production study that confirms that second occurrence foci are indeed acoustically distinguished from non-foci.

- (8) MARKING HYPOTHESIS: Second occurrence foci are acoustically distinguished from non-foci.

In discussions of second occurrence focus, it has been generally assumed that lack of pitch accent implies lack of focus marking. A possible explanation for the intuition of some researchers that second occurrence foci are *not* prosodically marked is that second occurrence focus is not marked in the same way as ordinary focus. As noted above, focus could be marked by other means than pitch accent such as duration. This hypothesis is stated in (9). As noted above, Rooth (1996b) and Bartels (1997) provided preliminary evidence that the acoustic realization of second occurrence foci differs from that of ordinary foci. In Section 4, we confirm that duration and intensity rather than pitch are correlates of second occurrence focus marking.

- (9) NON-STANDARD MARKING HYPOTHESIS: Second occurrence focus marking is distinct from ordinary focus marking.

While Section 4 provides experimental evidence that second occurrence focus correlates with some phonological feature which is acoustically distinguished, this leaves open the question whether the phonetic difference between second occurrence foci and non-foci is perceptible. In Section 5, we present results from a perception study that demonstrates that the prosodic correlates of second occurrence focus are strong enough to be perceptible. This hypothesis is stated in (10).

- (10) PERCEPTIBILITY HYPOTHESIS: Hearers can distinguish between the prosodic realizations of second occurrence foci and that of non-foci.

### 4. A Production Experiment

The primary goal of the first experiment concerns the MARKING HYPOTHESIS that second occurrence foci are acoustically distinguished from non-foci.

This hypothesis competes with the null hypothesis implicitly assumed in the argument from second occurrence focus (cf. section 2.5) that second occurrence focus is not prosodically marked. The experiment tests this null hypothesis.

Earlier, we alluded to a possible explanation for the strong intuition of some researchers that second occurrence focus is *not* focus marked: second occurrence focus is not realized acoustically in the same way as ordinary focus. This intuition was formulated in the NON-STANDARD MARKING HYPOTHESIS. A secondary goal of the production experiment was to confirm the non-standard marking hypothesis by singling out respects in which second occurrence foci differ acoustically both from non-focussed expressions and from ordinary foci.

#### 4.1. DESCRIPTION OF THE EXPERIMENT

We ran a production experiment in which naive subjects read preprepared written materials.

**Subjects:** 20 US born native speakers of English (10 female, 10 male) were recruited, none with any training in linguistics.

**Stimuli:** An example of a minimal pair of discourses used as stimuli is given in (11) and (12). These three-sentence discourses are designed to probe second occurrence focus effects in the scope of a focus sensitive operator in the third sentence. In this case, the relevant operator is *only*, while in other test pairs the operator was *always*. Note that another focus sensitive operator *even* is also present. The presence of this operator in the subject NP of the third sentence leads the sentence to be uttered with a nuclear accent in the initial NP (here on *state prosecutor*). On the basis of prior literature, we know that this creates a condition for reportedly unaccented realization of the second occurrence focus.

- (11) a. Both Sid and his accomplices should have been named in this morning's court session.  
 b. But the defendant only named Sid in court today.  
 c. Even the state prosecutor only named Sid in court today.
- (12) a. Defense and Prosecution had agreed to implicate Sid both in court and on television.  
 b. Still, the defense attorney only named Sid in court today.  
 c. Even the state prosecutor only named Sid in court today.

In examples (11) and (12), the relevant potential foci are *Sid* and *court*. For all the stimuli we used, the material following the second focus sensitive

operator in the third sentence (e.g., (11c) and (12c)) does not differ between the two members of the discourse pair. Thus, the phonological context for the segment of text containing the the two potential foci should not differ between the two elements of the minimal pair. We can therefore attribute acoustic differences between the potential foci in the two pairs to second occurrence focus effects. In all the minimal pairs, it is variations in the content of the first two sentences of the discourse (e.g., (11a,b) and (12a,b)) which are used to set up context and produce the effects to be measured.

In (11b), *Sid* is the semantic focus of the focus sensitive operator *only*, whereas in (12b), *court* is the semantic focus, and *Sid* is non-focal. Sentences (11c) and (12c) contain a textually identical VP (*only named Sid in court today*) to the (b) sentences, but the VP in (11c) and (12c) follows an earlier phrase *even the state prosecutor* containing what we take to be the nuclear accent of the sentence (presumably on *state prosecutor*). According to the discussion in section 2, this is a configuration in which second occurrence foci, such as the focus in the VP in each of the (c) sentences above, have been reported to be unmarked prosodically.

Crucially, the focus in each of the (b) sentences in (11) and (12) was different, so the textually identical VPs in the (c) sentences should differ only in so far as they contain different expressions which are second occurrence foci. In (11c), *Sid* is the second occurrence of a focus, and *court* is non-focal, whereas in (12c), *Sid* is non-focal, and *court* is the second occurrence of a focus. Thus the hypotheses above can be operationalized in terms of the acoustic differences between the realizations of *Sid* in each of the two final sentences, and the acoustic differences between the realizations of *court* in each of the two final sentences.

If an expression is the second occurrence of a focus in the final sentence of a discourse, we will say that the expression is in the *second occurrence focus* condition, and otherwise we will say it is in the *non-focal* condition.

Note that our stimuli are designed so that each minimal pair of discourses gives us two distinct probes of the hypotheses, since the conditions are reversed for the two potential second occurrence foci *Sid* and *court* in the two discourses.

We used a total of 14 discourse stimuli, made up of 7 minimal pairs like those above. In each case, the repeated focus sensitive operator was *only* or *always*, while the focus sensitive operator *even* was used to induce the nuclear accent in the subject of the final sentence of the discourse. In all cases the potential second occurrence foci in the final sentence are not sentence final, this being contrived by the use of an additional adverb; e.g., *today* in (11c) and (12c). The reason for adding an adverb is to prevent features of a potential focus expression from being combined with, and perhaps masked by, pitch movements marking the end of an intonational phrase.

**Presentation and procedure:** Following standard procedure, the 28 experimental stimuli were arranged in a pseudo-random order and intermingled with 16 unrelated filler stimuli, also discourses.<sup>20</sup> To control for recency effects which might occur when a subject encounters both members of a minimal pair of discourses, we ensured that paired discourses were always separated by at least four other discourses, either other experimental stimuli, or fillers. In this way, four different stimuli lists were constructed. Each participant in the study received one of these lists and list was included as a factor in the statistical analysis. The 28 experimental discourses consisted of seven minimal pairs repeated twice (for an example of a minimal pair, see (11) and (12) above). After recording all the speakers, the word boundaries were marked on the two probes taken to be potential foci in the target sentence (i.e., the third sentence of each discourse); e.g., *Sid* and *court* in (11c) and (12c) above. This procedure yielded  $28 \text{ (discourses)} \times 20 \text{ (subjects)} \times 2 \text{ (probes per target sentence)} = 1120 \text{ probes}$ .

The word boundaries were labeled by a subset of the authors (DB, BC, and MW), based primarily on examination of spectrograms. For each word, boundaries were marked at acoustic landmarks (such as stop closures) near the word onset and offset. The landmarks were selected (i) to be consistently identifiable across utterances of a word, and (ii) to include the vowel of the test word, since this is the expected locus of pitch and duration effects. All analyses are based on paired comparisons of the same words in non-focal and re-focused conditions, so consistency across conditions is of the greatest importance.

Having hand-labeled all the relevant word-boundaries, seven phonetic measures were taken for each target word. The measures were duration of the target word (in seconds), root-mean-square intensity of the word (in decibels (dB); henceforth rms intensity),  $f_0$  range, and maximum, minimum, and mean  $f_0$  (all in Hz).<sup>21</sup> A pseudo-energy measure was created by multiplying the standardized rms intensity value (standardized to the mean and standard de-

<sup>20</sup> Fillers were drawn from a separate experiment, reported in Wolters and Beaver (2001).

<sup>21</sup> Intensity and  $f_0$  information were extracted using standard procedures in Boersma and Weenink's *Praat* program. All pitch measures and the intensity measure were normalized using standard deviation from mean  $f_0$  in that single utterance and therefore are analogous to z-scores. This was done to prevent confounding effects from differences in the speakers' characteristic pitch ranges. Distributional normality of test measures can be affected by such operations, and is addressed in the last appendix.

Note also that all  $f_0$  measures are across a section strictly including the vowel. Thus there is potential interference from obstruent perturbations. This could have been avoided by analyzing only a subpart of the vowel. However, (i) we performed an analysis of foci (involving the same focused words) in the second sentences of the discourses using the same methodology, and found that pitch accents were easily detectable, so there is reason to think that our approach is sensitive to  $f_0$  effects; (ii) obstruent effects, if present, would have been revealed by our item analysis, but we observed no such effects; (iii) we hand-corrected pitch measures for a subset of the data, and found no change in results. Furthermore, analyzing only a subpart of the vowel

viation of the overall utterance) by the duration of the target. These measures were selected based on previous research concerning the acoustic correlates of focus, reviewed in section 2.6.

We were interested in minimal differences in the phonetic realization of a phrase depending on whether it was repeated and a second occurrence focus or repeated and not focused. In the experimental items (the seven different types of discourses), the areas of interest were both the first and the second post-verbal NP (in most cases, the direct and the indirect object) of the last sentence in a discourse.

For each phonetic measure taken on the target items we conducted both a by-subjects (henceforth F1) and a by-items (henceforth F2) analysis. Each ANOVA corresponded to a two-way crossed  $2 \times 2$  repeated measures design of **second occurrence focus** (whether the preceding context identified the target phrase as second occurrence focus or not) and **word position** (whether the target phrase was the direct or the indirect object). Following Clark (1973) and Raaijmakers et al. (1999), a minimum F was calculated from the F1 and F2 analyses. This maximally conservative approach was taken because the results reported below refute the long-standing opinion that second occurrence focus is not phonetically marked. Although, we report F1, F2, and minF values below, our conclusions are solely based on the conservative minF analysis. If not mentioned otherwise, F-values are reported as significant if  $p < 0.05$ , and as marginal if  $p < 0.1$ . Word position was included in the design in order to test for interactions between **word position** and **second occurrence focus**. Table I provides a summary of all independent and dependent variables considered in the experiment.

A blind review process was performed prior to the statistical analysis. Overall, 0.7% of the data were used for the analysis of duration, intensity, and energy and 2.8% of the data were excluded from the pitch analysis (for a detailed description of the exclusion criteria, see appendix B.1). The higher data loss for the pitch analysis is due to the well-known problems of pitch-tracking algorithms.<sup>22</sup>

---

might have lead to other problems, for example the risk of missing late pitch accents, which do occur in our data.

A related issue comes up with respect to variation of intrinsic  $f_0$  in the stimuli according to vowel height. We did not control explicitly for vowel height, although the target words in the stimuli have a range of vowel heights. We can make a similar comment here as with respect to segmental effects of obstruents: if intrinsic  $f_0$  effects due to vowel height had systematically affected second occurrence focus expressions differently from non-focal expressions, then that would have shown up in our item analysis, but it did not.

<sup>22</sup> Most exclusions were due to problems tracking  $f_0$  in creaky voice. This could be problematic in an experimental study of prosody, since creaky voice is presumably correlated with prosodic factors. E.g., creaky voice is less likely to obscure the  $f_0$  of a nuclear accent than to obscure the  $f_0$  of material in the post-nuclear tail. Since  $f_0$  tracking problems affected less

Table I. Summary of independent and dependent variables

Type	Variable name	Variable description
Independent vars.		
<i>(within subjects)</i>	<b>second occurrence focus</b>	Whether the probe is a second occurrence focus expression
	<b>word position</b>	Whether the probe is the first or second of the two marked NPs in the target sentence
<i>(between subjects)</i>	<b>list</b>	Which of the four lists did the subject read?
Dependent vars.		
	<b>duration</b>	Duration of the probe in msec
	<b>intensity</b>	R.m.s. intensity of the probe in dB
	<b>energy</b>	Relativized overall energy over the probe
	<b>maximum <math>f_0</math></b>	Maximum $f_0$ over the probe in Hz
	<b>mean <math>f_0</math></b>	Mean $f_0$ over the probe in Hz
	<b>minimum <math>f_0</math></b>	Minimum $f_0$ over the probe in Hz
	<b><math>f_0</math> range</b>	Range of $f_0$ over the probe in Hz

## 4.2. PREDICTIONS

The MARKING HYPOTHESIS predicts that subjects will distinguish prosodically between second occurrence foci and non-foci. Given earlier experimental studies of second occurrence focus (cf. section 2.7), we expect duration and intensity measures to distinguish second occurrence focus from non-focal expressions. Similarly, the NON-STANDARD MARKING HYPOTHESIS predicts that subjects will distinguish prosodically between second occurrence foci and their focal counterparts. In coming to these predictions, we made two assumptions. First, we assumed that the subjects understood the task and produced the stimuli naturally, and second, we assumed that the discourses were unambiguous in which focus assignments they elicited. So if significant main effects of the predicted type are observed, this provides support both for theoretical generalizations that would explain the predictions and for the assumptions in our experimental design.<sup>23</sup>

than 5% of the data, we did not attempt to build creaky voice in as a factor in our statistical analysis.

<sup>23</sup> Another general issue is the prescriptivist dogma that expressions such as exclusives should not be separated from what they modify, and thus in ‘good English’ are not focus



### 4.3. RESULTS

The analyses revealed a significant main effect of **second occurrence focus** on duration ( $F(1,19)= 6.9$ ;  $F(1,6)= 15.3$ ;  $\min F(1,24)= 4.8$ ) and energy ( $F(1,19)= 8.2$ ;  $F(1,6)= 17$ ;  $\min F(1,24)= 5.5$ ). Second occurrence focus phrases were on average 6 msec longer than non-focused phrases and were realized with 3.4% more energy. There also was a marginal effect of **second occurrence focus** on standardized rms intensity ( $F(1,19)= 7.1$ ;  $F(1,6)= 8$ ;  $\min F(1,19)= 3.8$ ), the standardized  $f_0$  range ( $F(1,19)= 8.8$ ;  $F(1,6)= 8$ ;  $\min F(1,17)= 4.2$ ) and the standardized minimum  $f_0$  ( $F(1,19)= 5.2$ ;  $F(1,6)= 9.1$ ;  $\min F(1,23)= 3.3$ ). Second occurrence focus phrases were realized with 0.31 dB higher intensity, a 10.5% higher relative  $f_0$  range (38.3 vs. 34.9 Hz), and a 6.9% lower relative minimum pitch (121.4 vs. 124.5 Hz). No main effect of **second occurrence focus** was found for maximum  $f_0$  and mean  $f_0$ . The means for all measures in both conditions of **second occurrence focus** and the mean differences between the two conditions of **second occurrence focus** are given in Table II. Significant results are marked by ‘\*’, marginal results by ‘+’.

Unsurprisingly, **word position** had a main effect on rms intensity, energy, and all standardized  $f_0$  measures. All of these main effects were in the F1 analyses and only the two main effects of standardized mean  $f_0$  ( $F(1,19)= 11$ ;  $F(1,6)= 5.7$ ;  $\min F(1,13)= 3.8$ ) and standardized minimum  $f_0$  ( $F(1,19)= 17.1$ ;  $F(1,6)= 11.4$ ;  $\min F(1,15)= 6.8$ ) were significant in the minF analyses.

For reasons addressed in section 4.1 (under the heading “Presentation and Procedure”), we also tested for interaction effects between **second occurrence focus** and **word position**. We found a marginal interaction effect of **second occurrence focus** and **word position** on the standardized mean  $f_0$  in

---

sensitive at all. We know of no basis for such a claim, since focus sensitivity is not only widespread, but also not a historically recent development. For detailed historical consideration of the focus sensitivity of English exclusives in particular, see Nevalainen (1991). Even allowing that the prescriptivists’ rules are baseless, there is a possibility that these rules may have affected how subjects in our experiments reacted to the stimuli we used, which involve exclusives that are separated from what they modify. While, logically, this is a possibility, we cannot see how it would produce any consistent effect. Furthermore, the use of three sentence discourses in which only the third sentence provides our main dataset means that we have a baseline measure of how subjects produce sentences with focus sensitive expressions. In particular, if prescriptive rules had any effect, we would expect that effect to arise with regard to both the second and the third sentences of our discourses. However, on the basis of informal sampling, the huge majority of the second sentences in the stimuli were produced with the type of intonation that would be expected in the theories of focus we consider in this paper. Thus it seems that prescriptive rules did not interfere with the subjects’ ability to produce foci naturalistically.

Table II. Summary of means and mean differences for non-focal and second occurrence focus expressions

Dependent variable	N	Non-focal ( $\mu_{nf}$ )	SOF ( $\mu_{sof}$ )	Mean difference $\Delta(\mu_{sof}-\mu_{nf})$
Duration (s)	521	0.267	0.273	*0.006
Std. r.m.s. intensity (dB)	521	-27.28	-26.97	+0.31
Rel. energy	521	0.0106	0.0117	*0.0011
Max $f_0$ (Hz)	483	162.2	163.0	0.8
Mean $f_0$ (Hz)	440	143.1	142.5	-0.6
Minimum $f_0$ (Hz)	473	127.5	124.1	+3.4
Range $f_0$ (Hz)	457	34.6	38.7	+4.1

both the F1 and the F2 analyses. Whereas the relative mean  $f_0$  was higher for second occurrence focus than for non-focused indirect objects, this effect was reversed for the direct object. The effect, however, turned out to be neither significant nor marginal for the minF analysis (F1(1,19)= 4.4; F2(1,6)= 5.3; minF(1,20)= 2.4). We also found a marginal interaction effect in the F1 analysis for standardized rms intensity (F1(1,19)= 3.5; F2(1,6)= n.s.; minF= n.s.) and in the F2 analysis for the standardized minimum  $f_0$  (F1(1,19)= n.s.; F2(1,6)= 6.7; minF= n.s.). Whereas the standardized minimum  $f_0$  was the same for second occurrence focused and for non-focused indirect objects, it was lower for second occurrence focused than for non-focused direct objects. The inverse was observed for intensity. Relative rms intensity on direct objects was the same in both **second occurrence focus** conditions, but was higher for second occurrence focused than non-focused indirect objects. This means that the main effect of **second occurrence focus** on intensity is only due to intensity differences on the indirect object. The main effect on minimal  $f_0$  is only due to  $f_0$  differences on the direct object. The means and standard deviations of all measures that were somehow affected by the interaction of **second occurrence focus** and **word position** are given in Table III.

The between-subject factor **list** (i.e. which of the four lists with different stimuli orders was read by the subject) was not significant in any of the analyses.

#### 4.4. DISCUSSION

As expected on the basis of prior literature, we found no evidence for systematic pitch marking of second occurrence focus. Instead, we found that a small

Table III. Summary of mean differences between second occurrence focus and non-focal expressions listed by **word position**

Dependent variable	1st NP	2nd NP
	$\Delta(\mu_{sof} - \mu_{nf})$	$\Delta(\mu_{sof} - \mu_{nf})$
Std. r.m.s. intensity ( <i>dB</i> )	0.03	0.58
Mean $f_0$ ( <i>Hz</i> )	-2.8	1.5
Minimum $f_0$ ( <i>Hz</i> )	-6.1	-0.6

but significant lengthening of second occurrence foci above that found for the same expression in the same sentential context but lacking focus. Across all trials, this lengthening averaged 6ms. There is also a statistically significant increase in energy, as well as marginally significant increases in intensity and the  $f_0$ -range in the second occurrence focus condition. Finally, there was a marginally significant drop in minimum  $f_0$  in the second occurrence focus condition. The nature of interaction effects between **word position** and **second occurrence focus** for minimum  $f_0$  and intensity suggests a more complex relation between these measures and second occurrence focus marking.

The statistical significance of the duration and energy results provide strong evidence that second occurrence foci are focus marked at some level of linguistic representation, since we have been careful to ensure that there is no difference between the second occurrence focus and non-focussed condition except for the position of focus. This corroborates what Rooth observed in his own productions, and is sufficient to undermine any argument for a strong (pragmatic) theory of focus that relies crucially on second occurrence foci not being formally marked. As far as second occurrence focus phenomena are concerned, weak (syntactic/semantic) theories of focus are quite defensible.

Accented words are known to differ from unaccented words in duration and intensity, as discussed in section 2.6. The results we have presented show that second occurrence focus is marked by the same cues. On the other hand, fundamental frequency plays a central role in the realization of pitch accent, but appears to play only a minor role in the marking of second occurrence focus the only identifiable effect on  $f_0$  measures is a marginally significant lowering of minimum  $f_0$  in second occurrence focus expressions in direct object position. We have no convincing explanation for this limited effect. It does not correspond to any local pitch movement in the examples that we have examined, and in the few cases where second occurrence focus is marked by a clear pitch accent, the accent is high ( $H^*$ ), which would lead us to expect an increase in  $f_0$  measures (at least the maximum and minimum  $f_0$  measures) on the second occurrence focus expression (see section 6.1 for further dis-

cussion).<sup>24</sup> In any case will see that the perception experiment reported in section 5 provides evidence that this  $f_0$  difference does not play a significant role in the perception of the marking of second occurrence focus. So we will conclude that greater duration and intensity are the primary correlates of second occurrence focus (section 6.1).

While the existence of a difference between second occurrence foci and their non-focal counterparts in production indicates that speakers are making a distinction between the two, the mean difference between second occurrence foci and their non-focal counterparts on the measured parameters are small. This raises the question of whether the difference is actually perceptible. We address this question in the experiment reported in the next section.

## 5. A perception study

While the production study presented in the previous section shows that the absence or presence of second occurrence focus correlates with certain phonetic properties, it leaves open whether the phonetic differences between second occurrence focus expressions and non-focal expressions are actually perceptible. As mentioned in the introduction, in order to address the assumption of the argument from second occurrence focus that second occurrence focus is *not* marked by perceptible prosodic features (see section 2.5), it is not sufficient to only show that second occurrence focus has prosodic correlates.

Thus we conducted a forced-choice perception experiment. Our hypothesis is formulated as the PERCEPTIBILITY HYPOTHESIS (cf. section 3, repeated below):

---

<sup>24</sup> Katy Carlson (p.c.) suggests that some of our recordings should be transcribed with a low pitch accent (L\*) on the second occurrence focus. If correct, this might explain the observed  $f_0$  effect. We did not identify any L\* accents in our own transcriptions of a subset of the data, but it can be difficult to distinguish an L\* accent from absence of accent in a context where  $f_0$  is expected to be low in any case. In support of the unaccented interpretation, we argue below (section 6.2) that analyzing the usual realization of second occurrence focus as unaccented enables us to relate the occurrence of this otherwise anomalous pattern of focus marking to the phenomenon of post-nuclear deaccenting. On the other hand, current generalizations about the use of L\* accents (e.g. Pierrehumbert and Hirschberg 1990) do not provide any basis for expecting an L\* accent in second occurrence focus environments, and in fact the few cases in our data where second occurrence focus unambiguously receives a pitch accent involve high accents (H\*). Note also that Hedberg and Sosa (2001) provide corpus-based distributional evidence that, for ordinary focus marking, H\* accents are much more common (61.9%) than L\* accents (19.0%). Nothing we currently know about H\* and L\* accents would lead us to expect that the relative frequencies of H\*s and L\*s are reversed in the case of second occurrence focus marking.

- (13) PERCEPTIBILITY HYPOTHESIS: Hearers can distinguish between the prosodic realizations of second occurrence foci and that of non-foci.

The purpose of the experiment was to provide evidence for the validity of the PERCEPTIBILITY HYPOTHESIS by refuting the null hypothesis that second occurrence focus-marking (henceforth, SOF-marking) is not perceptible.

Section 5.1 describes the methodology, materials, and the design of the experiment. Section 5.2 summarizes the predictions that follow from PERCEPTIBILITY HYPOTHESIS. Section 5.3 summarizes the results of the experiment. Finally, section 5.4 contains a discussion of the results and section 5.5 the summary of the study's implications.

### 5.1. DESCRIPTION OF THE EXPERIMENT

**Subjects:** 14 native speakers of English (10 female, 4 male; age range: 23-28) were recruited as follows. 10 linguistically naive subjects (8 female, 2 male; age range: 19-26) in the study, all without prior experience in linguistic perception studies participated. 4 subjects (2 female, 2 male; age range: 24-28) were linguistically trained.

**Stimuli:** 40 minimal discourse pairs of stimuli were quasi-randomly selected from the production experiment described in the previous section.<sup>25</sup> Three (out of the four available) minimal **discourse pairs** with the focus sensitive operator *only* were selected for use.<sup>26</sup> This choice was made for two reasons. First, *only* is standardly used in second occurrence focus examples in the literature. Indeed, Beaver and Clark (2003) provide evidence independent of the studies presented here that *only* is a more prototypical focus sensitive operator than *always*. Second, in order to keep the number of stimuli for the perception experiment low, a decision had to be made to either not have stimuli by all speakers from the production experiment or to limit the number of distinct pairs. Since we are interested in generalizations across speakers and listeners,

<sup>25</sup> The random selection was generated using the program Shuffle by Pallier (1997).

<sup>26</sup> The overall 40 minimal discourse pairs of stimuli were split up as follows. The stimuli consisted of 16 tokens each of the pairs in (i) and (ii), and eight tokens of the pair in (iii).

- (i) (a) Even [the state prosecutor]<sub>F</sub> only named [Sid]<sub>SOF</sub> in court today.
- (b) Even [the state prosecutor]<sub>F</sub> only named Sid [in court]<sub>SOF</sub> today.
- (ii) (a) Even [the doctor]<sub>F</sub> only gave [Pete]<sub>SOF</sub> a pill today.
- (b) Even [the doctor]<sub>F</sub> only gave Pete [a pill]<sub>SOF</sub> today.
- (iii) (a) Even [the kids]<sub>F</sub> would only let [the cat]<sub>SOF</sub> in the tent today.
- (b) Even [the kids]<sub>F</sub> only let the cat [in the tent]<sub>SOF</sub> today.

we decided to reduce the number of stimuli by excluding discourses with *always*.

Discourses that contained stuttering, unusually long pauses, or other signs of ‘reading effects’ (recall that the production experiment elicited the stimuli in a reading task) were excluded from the experiment in a blind review process prior to the random selection of the discourse pairs. Of the selected discourses, only the final sentence was used. Both parts of a pair always were taken from the same three sentence discourse pair and were uttered by the same speaker. The two elements of each pair differed only in their (second occurrence) focus assignment. This is illustrated by the examples given in (14) and (15). The sentences in (14b) and (15b) form a stimulus pair for the perception study. The preceding discourse from the production study, given in (14a) and (15a) are not used in the perception experiment. In the examples, second occurrence foci are indicated by ‘SOF’ but subjects were presented acoustic stimuli and therefore had no access to any artificial annotation.

- (14) a. Both Sid and his accomplices should have been named in this morning’s court session. But the defendant only named Sid in court today.
- b. Even the state prosecutor only named [Sid]<sub>SOF</sub> in court today. in court today.
- (15) a. Defense and Prosecution had agreed to implicate Sid both in court and on television. Still, the defense attorney only named Sid in court today.
- b. Even the state prosecutor only named Sid [in court]<sub>SOF</sub> today.

The list of stimulus pairs was balanced to meet the following conditions. Two pairs each from all 20 speakers of the production experiment were used. For each of the three minimal discourse pairs (see footnote 26), there were an equal number of instances where the first NP was supposed to be SOF-marked in the first sentence of a pair and instances in which the second NP was supposed to be SOF-marked in the first sentence.<sup>27</sup> Since, the two sentences in a pair only differed in their focus assignment, the second NP was SOF-marked if and only if the first NP was not SOF-marked.

Two versions of the experiment with inverse order were produced. 7 subjects took version 1 and 7 subjects took version 2 of the experiment. The following order constraints were observed when the list of stimuli was constructed: (i) the same internal order of a pair (the first sentence has a SOF-marked first NP or a SOF-marked second NP) was not allowed to occur more

<sup>27</sup> Subjects in the production study read every stimulus twice (see section 4). The stimuli used in this perception experiment were balanced for repetition. Both parts of a pair were from the same repetition.

than three times in a row; (ii) the same pair was not allowed to occur twice in a row.

**Presentation and procedure:** The acoustic stimuli were incorporated into an HTML page. Subjects first read through the instructions and then were allowed to ask the experimenter questions before they started the experiment. Subjects were asked to judge “in which of the two sentences (A or B) the speaker wished to make the second word (which will be given in bold face, cf. Figure 2) more prominent than the first.” Subjects were not allowed to make any changes once they reached a decision for a pair.

Each pair of acoustic stimuli was accompanied by two words that were displayed in the same row. This is exemplified below for two pairs (pair 4 and 5 out of 40):

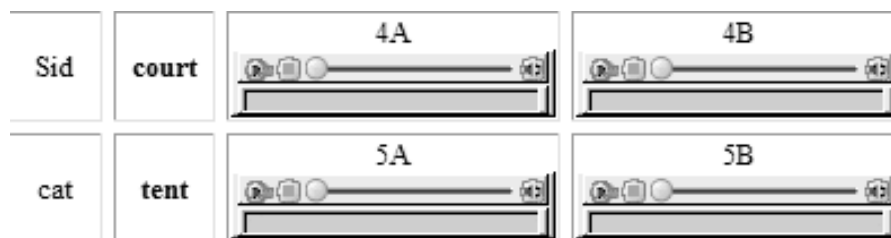


Figure 3. Excerpt of the HTML page displayed to subjects during the experiment

10 pairs of stimuli were displayed at a time. After every 10th pair of sentences, subjects were asked to take a break of 30 seconds to 2 minutes, which they could time themselves, before they continued to the next page. Subjects were asked to proceed from pair to pair and always listen to the first stimulus in a pair first. After that, they were allowed to listen to each of the stimuli as often as they wanted to. All subjects required between 15-25 minutes for the experiment and approximately 2 minutes to read the instructions.

To test whether subjects understood the instructions, after the experiment all subjects were asked to repeat in their own words what they had been asked to do.

Table IV provides a summary of all independent and dependent variables considered in the experiment.

Table IV. Summary of independent and dependent variables

Type	Variable name	Variable description
Independent vars.		
<i>(within subjects)</i>	<b>second occurrence</b>	Whether the first or the second NP was the second occurrence focus expression
	<b>focus location</b>	From which discourse pair is the stimulus pair taken?
	<b>discourse pair</b>	When in the list was the stimulus presented?
<i>(between subjects)</i>	<b>order</b>	Which of the two lists did the subject read?
	<b>list</b>	Gender of participant
	<b>gender</b>	Whether the participant was a linguist or not
	<b>linguistic education</b>	
Dependent vars.		
	<b>answer</b>	Was the answer given the expected one in light of the <b>second occurrence focus location</b>

## 5.2. PREDICTIONS

It follows from the PERCEPTIBILITY HYPOTHESIS that subjects should perform above chance in selecting from each pair the stimulus that was uttered with the intention of marking the second NP as second occurrence focus. Even if participants in the production study did not always produce natural stimuli, e.g. due to confounding factors like reading effects and the artificial situation created in a laboratory experiment, participants in the perception study should still overall perform above chance.

## 5.3. RESULTS

All subjects performed above chance (mean performance = 63% correct answers; range = 52.5-77.5%). The one-sample t-test against an expected mean of 20 correct answers out of 40 reveals that subjects on average perform significantly above chance ( $t = 7.7$ ;  $df = 13$ ;  $p < 0.001$ ) and that subjects performed 'alike' (i.e. there was no subject effect;  $\chi^2 = 9.6$ ,  $df = 13$ ;  $p > 0.7$ ). The average effect size is small to medium ( $\omega = 0.28$ ). All subjects showed clear understanding of the task in the post-experimental interview.<sup>28</sup>

<sup>28</sup> Based on the introspective data from post-experimental interviews, the precise strategy chosen to solve the task differed from subject to subject. We will not attempt to quantify this variation here, or to suggest a connection to our quantitative results.



None of the between-subject factors (i.e. **gender**, **list**, **linguistic education**) had a significant effect. Neither did **order** have a significant effect (i.e. the task was not too tiresome for our subjects). Crucially for pooling the data from all subjects, the level of linguistics education did not matter. Even though linguists performed slightly better than linguistically naive subjects (67% correct answers vs. 62%) this difference was not significant ( $\chi^2(1) = 1.2$ ;  $df = 1$ ;  $p > 0.25$ ). Furthermore, all effects that hold for all subjects together also hold for linguistically-naive subjects alone. There also was an item effect ( $\chi^2 = 163.9$ ,  $df = 39$ ;  $p < 0.001$ ). That is, subjects did not perform equally well on all sentence pairs. Interestingly, there were items that nobody answered as expected and items that everyone answered as expected by the PERCEPTIBILITY HYPOTHESIS. A planned comparison of the **discourse pairs** used in the experiment revealed a significant effect ( $\chi^2(2) = 10.9$ ;  $p < 0.01$ ): one of the three minimal **discourse pairs** (pair (iii) in footnote 26) on average lead to better performance than the other two pairs (stimuli of type (i) lead to correct answers in 59.4%, type (ii) lead to 60.7% correct answers, whereas second occurrence focus was correctly detected in 76.8% for discourse pair (iii)). It is important for the current purpose that *all* stimuli of all types performed above chance.

#### 5.4. DISCUSSION

The results confirm the PERCEPTIBILITY HYPOTHESIS. Even though the phonetic correlates of second occurrence focus observed in the production experiment are very small, subjects are able to identify SOF-marked expressions correctly in 63% of all cases in the experiment. This shows that subjects are able to identify SOF-marking without contextual clues (recall that the sentence pairs were spliced from their preceding discourse), but it also raises the question of why people did not perform even better (or in other words, why the effect size is relatively small). For one thing, the overall performance is decreased because of 4 items that were almost always judged contrary to the predictions of the MARKING HYPOTHESIS (judged correctly by only 0-3 subjects out of the 14 overall subjects), and 4 more items that performed below chance (correctly judged by 4-6 subjects). Figure 4 shows the percentage of stimuli for which the SOF was identified as most prominent expression (a 'correct' response) for any given number of subjects. For example, it can be seen that stimuli for which 14 out of 14 subjects responded correctly amounted to 12% of all stimuli. It can also be seen that a clear majority of stimuli (66%) were judged correctly by more than 7 subjects, a rate above chance: the corresponding area of the pi-chart is shaded black.

Note that the existence of items judged below chance is neither what is expected according to the MARKING HYPOTHESIS, nor what would be expected in the absence of prosodic cues to second occurrence focus: since

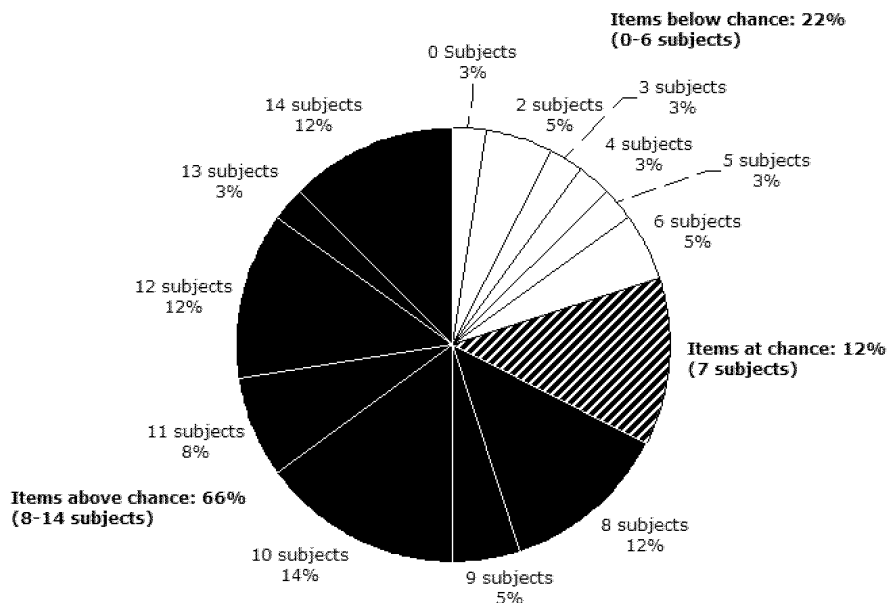


Figure 4. Which proportion of all items is judged correctly by how many of the subjects?

subjects should perform at chance in the absence of phonetic cues, these answers should have been judged correctly by 50% of all subjects. It is likely that, compared to naturally occurring speech, the artificial experimental scenario of the production experiment resulted in an above-average amount of intonationally ill-formed stimuli. If one excludes the four worst items, overall performance increases by 6% to 69%. If one also excludes the additional 4 items that performed below chance, the overall performance increases to 73%, so it is worth considering whether ill-formedness was the reason for poor performance on some items.

We performed an informal examination of the intonational well-formedness of a subset of data gathered from production experiment, concentrating in particular on items that yielded poor performance. Prior to the perception experiment, we had already examined the intonation on the second sentences in the read discourses, which contained one ordinary focus expression and no second occurrence focus expression, as in the final sentences in (14a) and (15a), repeated below.

- (14) a. Both Sid and his accomplices should have been named in this morning's court session. But the defendant only named [Sid]<sub>F</sub> in court today.
- b. Even [the state prosecutor]<sub>F</sub> only named [Sid]<sub>SOF</sub> in court today.

- (15) a. Defense and Prosecution had agreed to implicate Sid both in court and on television. Still, the defense attorney only named Sid in [court]<sub>F</sub> today.
- b. Even [the state prosecutor]<sub>F</sub> only named Sid in [court]<sub>SOF</sub> today.

For a sentence of that type (i.e. a sentence involving a single focus and no second occurrence focus) it is often easy to judge whether an utterance is intonationally well-formed. While the majority of the stimuli were well-formed in their second sentence, as judged by the authors in an informal listening experiment, the rate of intonationally ill-formed sentences was high enough to explain why some stimuli in the perception study were consistently judged counter to expectations. In several stimuli, the second sentence (e.g. in (14a) and (15a)) contained a strong pitch accent on the unfocused expression but not on the focused expression indicating that the subject had problems reading the stimuli (either because the stimulus was textually unnatural or because the subject was tired or not paying sufficient attention). The wrong placement of the pitch accent sometimes seemed to carry over to the third sentence containing the second occurrence focus expression. The decision not to exclude those supposedly ill-formed stimuli *a priori* was made because this may have seriously confounded the study. However, it is clear, and, indeed, unsurprising, that the existence of occasional unnatural productions in the original reading task is one of the factors limiting performance on the perception task.

Finally, it is significant that the perception task which subjects were asked to perform referred to intended prominence rather than directly to SOF-marking. The main purpose of the perception experiment was to establish whether second occurrence focus-marking is perceptible. In that sense, participants were used as (very intelligent) minimal difference detectors. That they performed significantly above chance means only that the difference between the phonetic realization of second occurrence focus and that of unfocused repeated expressions is humanly detectable, not that it is reliably detectable in natural situations. In addition, it is impossible to know from this experiment whether listeners actually use phonetic marking to infer second occurrence focus: it is logically compatible with our findings that hearer's disregard the phonetic differences we have observed, and rely entirely on context to identify second occurrence foci.

With regard to the latter issue, it should be borne in mind that the point of this paper is not to falsify pragmatic theories of focus or to show that pragmatic and discourse clues are irrelevant to the determination of focus assignments. Rather, our goal is to show that, contrary to claims in the literature, phonetic clues are available to infer the focus assignments of a sentence even in the case of second occurrence focus. This goal has been achieved.

We can only offer a plausibility argument as to whether phonetic information is used in interpretation of second occurrence focus. It is conceivable, though we take no position on this, that natural discourses are typically rich enough semantically to provide unambiguous clues as to whether a constituent is in second occurrence focus. Even if so, why should listeners not simultaneously use all information available to them? As more and more psycholinguistic research shows, the default in language processing is massively parallel constraint satisfaction, where the constraints come from different levels of linguistic competence as well as from non-linguistic processes (such as eye-movement; e.g. Tanenhaus et al. 1995). So, given that speakers mark second occurrence focus and that hearers are able to detect that marking, it would be surprising if hearers did not use the marking in interpretation. But this, of course, is not a proof that they do use that marking.

#### 5.5. CONCLUSIONS

The perception experiment presented above provides strong support for the PERCEPTIBILITY HYPOTHESIS, since it shows that hearers in principle have access to prosodic cues when identifying second occurrence focus: the marking of second occurrence focus is perceptible. This provides the final seal on our case that the argument from second occurrence focus (see section 2.5) is specious.

### 6. General Discussion

The experimental results reported in sections 4 and 5 support the MARKING, NON-STANDARD MARKING and PERCEPTIBILITY hypotheses. This is sufficient to undermine the argument from second occurrence focus described in section 2.5, but these results raise further questions about the prosodic marking of focus. That is, previous descriptions of second occurrence focus were not mistaken in claiming that the realization of words in second occurrence focus differs from the realization of normal unrepeated foci, they were only mistaken in claiming that second occurrence focus is not differentiated from unfocused material. So we are left with the following questions:

1. What is NON-STANDARD about second occurrence focus marking – What is the nature of the difference between the realizations of second occurrence focus and regular focus?
2. Why is second occurrence focus marking NON-STANDARD – Why is second occurrence focus realized differently from regular focus?
3. In what other environments should we expect similar effects to occur?

Our experiments do not allow us to answer these questions conclusively, but we will argue for tentative answers that situate the phenomenon of second occurrence focus within the broader theory of the realization of focus, and which bring order to a collection of related phenomena discussed previously by semanticists. Briefly, we suggest that second occurrence focus is made prominent by greater duration and intensity (and maybe other prosodic correlates that we did not consider in this study), but it is usually not pitch accented. In short, second occurrence focus marking is an example of phrasal prominence without pitch accent. This realization can be understood as resulting from the fact that items in second occurrence focus occur after the nuclear accent of their phrase, the environment of post-nuclear deaccenting.

Below we address (i) and (ii) in detail (section 6.1 and 6.2 respectively) and show that this interpretation has implications for the general theory of phrasal prominence (stress and accent) and the marking of focus. We then bring some of the ideas we have developed to bear on prior discussion in the semantics literature of example types closely related to the classic second occurrence focus paradigm (section 6.3).

#### 6.1. WHAT IS NON-STANDARD ABOUT SECOND OCCURRENCE FOCUS MARKING?

We conducted a comparison between second occurrence focus marking and the marking of ordinary, non-repeated focus.<sup>29</sup> Specifically, we investigated two questions that derive from our NON-STANDARD MARKING HYPOTHESIS.

- (16) PITCH-SUPPRESSION: Does standard second occurrence focus marking lack a pitch accent?
- (17) PITCH-OPTIONALITY: Can second occurrence foci optionally be marked using (ordinary) focus marking, complete with pitch accent?

As discussed in section 2, focus is usually marked by a pitch accent, perhaps specifically a nuclear pitch accent, and the primary correlates of pitch accent are to be found in the fundamental frequency contour. So a rather striking feature of second occurrence focus marking is the minor role played by fundamental frequency ( $f_0$ ). The production study revealed no significant main effect for  $f_0$  differences between words in second occurrence focus and unfocused words. Auditory transcription of a subset of the data indicates that second occurrence focus expressions are occasionally marked with high pitch accents ( $H^*$ ), but in most cases there is no perceptual indication of a pitch

<sup>29</sup> See Jaeger (2004, section 4) for further details.

accent on the second occurrence focus expression. For cases without a pitch accent on the second occurrence focus expression, two representative pitch tracks along with their transcriptions are shown in Figure 5 and 6. The focus expression associated with *even* is marked by a clear f<sub>0</sub> peak, indicating that the focus of *even* is marked by a high pitch accent. Then the fundamental frequency falls to a low level, and remains relatively low, or slightly declining, to the end of the utterance. That is, there is no evidence for a pitch accent on any words after the focus expression associated with *even*.<sup>30</sup>

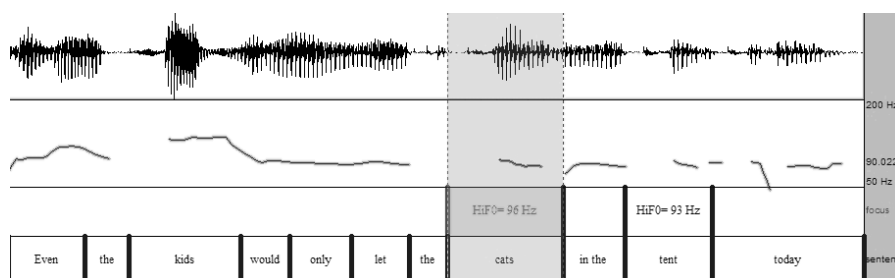


Figure 5. First sample with a second occurrence focus lacking any pitch accent.

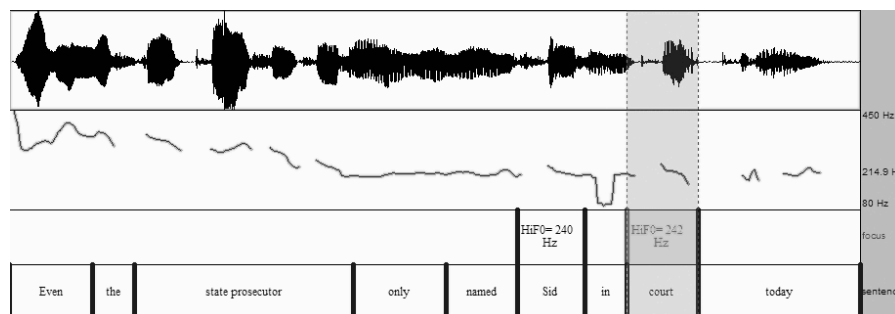


Figure 6. Second sample with a second occurrence focus lacking any pitch accent.

Figure 7 is a representative example of a second occurrence focus expression marked by an H\* pitch accent (though a comparatively subtle one).<sup>31</sup>

<sup>30</sup> The small perturbations mostly at the syllable onsets and syllable codas of some words are consonant effects and are not intonationally significant. For an overview of such effects, see e.g. Silverman (1987) or Beckman and Elam (1997). Note also that the pitch tracks in Figure 6 and 7 contain instances of pitch halving immediately preceding or following the target expression. That is, the pitch tracking algorithm erroneously halved the fundamental frequency for a short period (c.f. Ladefoged 2003:83–84). The actual f<sub>0</sub> is more or less level.

<sup>31</sup> In the utterance pictured in figure 7, *doctor*, the focus of *even*, is highly prominent auditorily, and this is reflected in the amplitude track in the figure. However, it is notable that (aside from onset effects) there is no pitch movement on *doctor*. It is conceivable that this is related to the fact that a pitch movement does occur on the later second occurrence focus *Pete*, but we have not studied this possibility further.

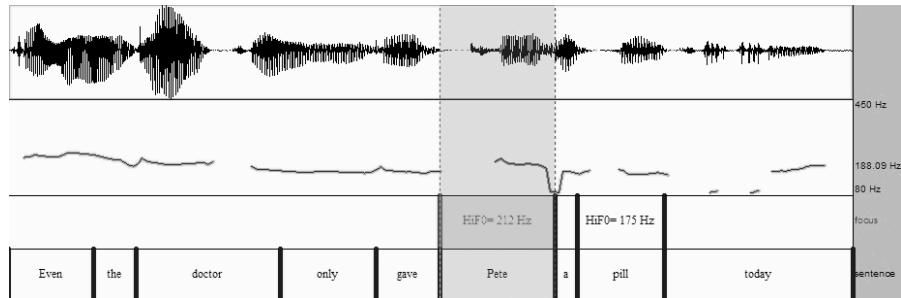


Figure 7. Sample with a pitch accented second occurrence focus.

So the results of the production study show not only that second occurrence focus is consistently marked by greater duration and intensity, but also that it is occasionally marked by  $f_0$ . To put it in the phonological terms we used to set up this issue: second occurrence foci can optionally be marked with a pitch accent. Given the existence of optional pitch accents, it is logically possible that the perception subjects' ability to distinguish second occurrence focus from lack of focus is based on a few stimuli in which second occurrence focus is marked by a pitch accent. We now turn to a brief analysis of the acoustic properties of the stimuli used in the perception experiment indicates that this is not the case.<sup>32</sup>

A logistic regression model was used to predict the subjects responses in the perception task based on measurements of the acoustic properties of the stimuli. Note that this model predicted the subjects' actual responses rather than the 'correct' answers, in order to try to identify the stimulus parameters that listeners attended to in making their judgments. The best predictors of subjects' judgments were measures based on the duration and intensity of the target words, and the best single predictor was the energy measure derived by multiplying duration by intensity. Adding fundamental frequency information did not significantly improve the fit of a model based on the energy factor. This indicates that subjects relied primarily on duration and intensity in identifying the more prominent word, and fundamental frequency did not play an important role.

In addition to the logistic regression analysis, we examined the 'best stimuli' from the perception test, i.e. those stimuli for which all subjects gave the same response. We would expect these stimuli to exhibit a particularly clear distinction between second occurrence focus and unfocused items. Careful inspection of the pitch measurements in the target words, and the overall pitch contours revealed no apparent way in which  $f_0$  could have been used to distinguish the prominence of the words.

<sup>32</sup> More details are given in Jaeger 2004, sections 4.4 and 4.5.

Based on both the logistic regression analysis and the sample of ‘best stimuli’, we conclude that the primary correlates of second occurrence focus are duration and intensity in the sense that these are the measures that most reliably distinguish second occurrence focus from lack of focus in production, and that they are the properties that listeners attended to in the perceptual task. Fundamental frequency is only optionally used in marking second occurrence focus and is not required for perception of the prominence of second occurrence focus. Of course if a speaker chose to mark second occurrence focus with a pitch accent, it is quite possible that the associated  $f_0$ -related cues would be more significant than duration and intensity in giving rise to a percept of prominence, but this is not the typical realization of second occurrence focus, nor is a pitch accent necessary for listeners to detect the prominence associated with second occurrence focus.

## 6.2. WHY IS SECOND OCCURRENCE FOCUS MARKING NON-STANDARD?

Phonologically, we interpret the confirmation that pitch suppression occurs in terms of a distinction between phrasal stress and pitch accent: focus is usually marked by phrasal stress and a nuclear pitch accent, but second occurrence focus is marked by phrasal stress only. As discussed in section 2.6, theories that distinguish stress and pitch accent generally identify greater duration and intensity as primary correlates of stress, while pitch accents are primarily realized in terms of the fundamental frequency contour. So the greater duration and intensity of SOF are attributed to phrasal stress, and the absence of  $f_0$  differences between SOF and non-focused material is attributed to the absence of pitch accents in both conditions (with a few exceptions in the case of SOF, as noted above).

Positing that second occurrence focus is marked by stress but not accent may help explain why second occurrence focus is realized differently from regular focus, by relating second occurrence focus to the well-established phenomenon of post-nuclear deaccenting. As mentioned in section 2.6, a nuclear accent is the last pitch accent in a phrase, so placing a nuclear accent early in a phrase, as in (18), implies that all following words must be unaccented. This is referred to as *deaccenting*, the reason being exemplified by in (18), where the word *Sandy* would be accented in a more neutral (or broad focus) reading of the sentence.

(18) Pat gave a [book]<sub>F</sub> to Sandy.

In all of the examples used in our experimental study, the second occurrence focus is preceded by an expression which we expect to be realized with a nuclear pitch accent, marking the focus associated with *even*. So, as we have indicated, the absence of a pitch accent on the second occurrence focus could be regarded as an instance of post-nuclear deaccenting (c.f. Bartels



1997:12 and Rooth 1992). That is, the second occurrence focus could not be accented without supplanting the accent of the focused subject as nuclear accent, or initiating its own phrase, in which case the sentence would contain two phrases, each with a nuclear accent.

We may conjecture that a two-phrase realization is employed in some of those occasional cases in which speakers choose to place a pitch accent on a second occurrence focus, but that the sentence containing the second occurrence focus is more usually produced as a single phrase.<sup>33</sup> The predominance of this phrasing is in turn plausibly related to the fact that the material following the subject (e.g., *only named Sid in court today* in Figure 6) is *given* information (in the sense of Prince 1981) and structurally identical to a portion of the preceding sentence (i.e. the same words bear the same grammatical functions and are in the same syntactic positions). These are both factors that have been shown to favor unaccented realizations (Ladd 1996:166, Terken and Hirschberg 1994), so speakers may be disinclined to place this material in its own phrase where it must receive a nuclear accent.

The picture we have arrived at is one according to which the focus of a focus sensitive operator should receive the strongest phrasal stress in the scope of the operator (cf. Jackendoff 1972), which is preferentially achieved by assignment of a nuclear accent to the focus. In the canonical second occurrence focus examples this conflicts with a dispreference for nuclear accents on repeated material, and an independent preference for placing the nuclear accent on a separate focussed constituent which occurs earlier in the sentence. This conflict is generally resolved in favor of deaccenting, but the second occurrence focus can still be marked by the strongest stress in the scope of the focus sensitive item with which it is associated.

### 6.3. WHAT ENVIRONMENTS YIELD PITCH ACCENT-LESS FOCI?

Our results demonstrate that in one specific type of environment focus is marked using duration and intensity, and the earlier discussion in this section addressed features of the environment that might be relevant in producing such realizations. The semantics literature includes discussions of a number of variants in the basic second occurrence focus paradigm. In this section we will consider these variants in the light of what we have learned.<sup>34</sup>

<sup>33</sup> We have not performed a direct study to establish the plausibility of the two phrase analysis, as might be marked by boundary tones and duration effects on both stressed and unstressed syllables, and by breaks in the speech signal, but see Jaeger (2004, section 5.3) for further discussion of issues connected with phrasing.

<sup>34</sup> In the discussion in this section we concentrate on the importance of the relative positioning of second occurrence (or other) foci and nuclear accented foci. The role of discourse status or textual repetition has also come up in prior literature, specifically in the work of Krifka (1997:270–271). Krifka considers variants of second occurrence focus examples in which the second occurrence focus is pronominalized, and argues that such examples require

In the experimental materials, the second occurrence focus is preceded by a nuclear pitch accent, marking the focus associated with *even*. In looking through the literature on apparent dissociation between focus and pitch accenting, we observe that the great majority of examples involve an expression which semantically one would expect to be accented, but which occurs in post-nuclear position.

Consider the following classic example:

- (19) People who *grow* rice generally only *eat* rice. (Rooth 1992:109)

Like second occurrence focus examples, this case involves an apparent mismatch between the intonational focus of a VP following a focus sensitive operator (*only*), and the semantic focus of the operator. The intonational focus, at least as regards pitch movement, is *eat*, but the semantic focus of *only* is *rice* (as evidenced by the meaning of the sentence). As for the second occurrence data we have examined, the semantic focus follows a nuclear accent, so neutralization of pitch movement is to be expected. It is natural to wonder whether the semantic focus of *only* is lengthened in the same way as we found for de-accented second occurrence foci. We leave examination of this question for further work.

A more complex type of example, due to Roberts (1996), is given in (21). Roberts imagines a conversation where we are discussing a number of things we would rather Mary did not do, listed in (20):

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

In this context, Roberts' considers the following dialogue:

- (21) A: Mary wasn't so bad after all. Of all the things we were afraid she might do, she only [invited Bill for dinner]<sub>F</sub>.  
 B: You got the person wrong. She only invited [Lyn]<sub>F</sub> for dinner. But it's true that she did only one of those terrible things she could have done.

On the desired reading of the penultimate sentence of Bs reply, the domain of *only* is the set of properties in (20). However, given that there is <sup>\_\_\_\_\_</sup>pitch accents on the pronominal foci. If so, then textual identity of the second occurrence focus with the first occurrence may be crucial for pitch suppression of foci. This accords with Ladd's (1980) observation that there is a strong tendency not to accent repeated material in English, mentioned in section 6.2, above.

focal prominence on *Lyn*, standard theories of focus sensitivity, like Rooth's (1985) weak theory, would lead us to expect that all the alternatives under consideration have the form *invited X for dinner*.

Unlike other examples we have considered, in this case we are considering a VP-focus, rather than focus on an argument. Thus to make sense of the example it is necessary to consider how VP-focus is normally marked. This issue is addressed in accounts of *focus projection* mentioned earlier, such as Gussenhoven (1984) and Selkirk (1995). Without going into detail, it suffices here to point out that in most cases focus on a VP can be marked by placing a pitch accent on the final argument NP, here *dinner*.

Having identified an accent pattern we might expect for such a case of VP focus, we can now say in exactly what sense the penultimate sentence in example (21) shows an unexpected pattern: *dinner* does not receive a pitch accent. Having observed this, it becomes clear that the lack of accent may be correlated with the fact that *dinner* occurs in post-nuclear position, where the nuclear accent is the contrastively stressed *Lyn*. (Presumably the fact that *Lyn* rather than *dinner* is chosen by the speaker to be the nuclear accent relates to the function of the speech act, and is independent of which word occurs first in the linear order of the sentence.)

The fact that in Roberts' example the nuclear accent happens to occur right in the middle of the phrase that is the focus of *only* complicates the analysis, but Roberts' example still has the same basic property as the cases we have studied: a word that we would expect to receive accent occurs in post-nuclear position and in fact is not accented. Once again, we may ask whether in this case the word in question is marked as prominent in some other way, e.g. by lengthening and intensity effects.

Are there any cases in which a semantic focus lacks accent, but does not follow the nuclear accent? Dryer (1994) claims that there are, and presents a candidate example. In the final sentence of (22), the second occurrence focus expression *a book* precedes the nuclear accent on *many people*.

- (22) A: I hear that John only gave [a book]<sub>F</sub> to Mary.  
 B: True, but John only gave [a book]<sub>SOF</sub> to [many people]<sub>F</sub>.

Dryer is certainly right to suggest that the most prominent accent in B's utterance can be on *many people* rather than *a book*. However, it is not clear to us whether this sentence can be felicitously uttered with no pitch accent on *a book*.<sup>35</sup> If Dryer is right, then we obviously would have to give up on our hypothesis that pitch accentless realizations of foci are restricted to post-nuclear positions. However, even if B's utterance is felicitous with no pitch

<sup>35</sup> Perhaps what Dryer is observing is simply that *a book* in (22) need not carry a nuclear accent: it is naturally uttered without phrasal breaks between *a book* and the next accented expression, i.e. *many people*.

accent on *a book*, it remains open whether the information status of *a book* is marked by lengthening or greater intensity. If there is lengthening or greater intensity but no pitch accent, then the example shows that pitch accentless foci are not restricted to post-nuclear positions. However, it is only if there are no prosodic correlates of focus at all on *a book* that the example would be problematic for the strong theories of focus which we have set out to defend in this paper.<sup>36</sup>

## 7. Conclusion

The focus of a focus sensitive operator receives the strongest phrasal stress in the scope of the operator, usually realized as a nuclear accent. Second occurrence foci provide an extreme test case since they are typically deaccented. Yet our results indicate that second occurrence foci are still marked with the strongest stress in the scope of the operator.

This analysis has two implications for theories of focus-marking and accentuation. First, it implies that phrasal stress plays an important role in marking focus, and that pitch accent is not required, although it is usual. Second, post-nuclear deaccenting is purely a restriction on the appearance of *pitch accents* following a nuclear accent, while phrasal stress distinctions are permitted following a nuclear accent. This is compatible with analyses of phrasal stress and intonation such as Selkirk (1984:154f.) and Hayes (1995:396), but is inconsistent with analyses such as Bolinger (1958) which attribute all phrase-level prominence to pitch accents, implying that there are no prominence distinctions in the post-nuclear domain. Ladd (1996) has also proposed a number of arguments for the existence of post-nuclear prominence distinctions, but attributes these to post-nuclear pitch accents (e.g. pp. 215,

---

<sup>36</sup> We have not experimentally examined the example types due to Rooth, Roberts and Dryer discussed above, but we have considered further constructed data. For example, consider the following two variants of Rooth's (19):

- (i) In poor countries, people who sell vegetables normally only eat vegetables.
- (ii) In poor countries, people who have vegetables to sell normally only have vegetables to eat.

Whereas (i) is parallel to Rooth's original example, in (ii) the second occurrence focus *vegetables* precedes the contrastive focus *eat* rather than following it. Naive respondents who were asked to read (i) and (ii) at first generally did not place contrastive stress on *sell* and *eat*. When asked to emphasize this contrast, subjects' reactions to (i) and (ii) varied. One respondent commented, not atypically: "It's harder to put emphasis on *sell* and *eat* in [(ii) than in (i)]." This provides further suggestive evidence that position of the second occurrence focus relative to the nuclear accent could be important, although there are certainly other factors to be considered. For example, even if some productions of (ii) lacked any pitch accent, we would have to rule out these being cases where the entire VP *have vegetables to eat* was the focus of *only*.

227). Some of these observations might be attributed to post-nuclear phrasal stress rather than pitch accent *per se*, since they do not appear to involve significant pitch movements. The only previous experimental investigation of post-nuclear prominence of which we are aware is Huss (1978), but this study examines lexical stress, not phrasal stress, showing that stress minimal pairs such as *éxport* (noun) vs. *expórt* (verb) are distinguished in post-nuclear position (although the difference is subtle).

As discussed in section 6, studies of our production data indicate that pitch accenting is optional on second occurrence foci, although it occurs in only a minority of utterances. A more complicated question is whether the majority of second occurrence foci are just like ordinary foci, but with pitch movement suppressed. Our answer is a qualified *yes*: it is clear that pitch is usually suppressed in the relevant examples, while other parts of second occurrence focus marking are similar to ordinary focus marking, but we cannot conclude that pitch is the *only* suppressed feature.<sup>37</sup> If we hypothesize that pitch suppression forces (most) second occurrence foci to be marked without a pitch accent, so that they are realized instead with phrasal stress only, we arrive at a testable prediction for future work: *all* prosodic features that are shown to correlate with phrasal stress rather than pitch accents should be employed in second occurrence focus marking. Vowel quality, for example, may be such a prosodic feature.

As regards the semantics and pragmatics of focus, we have demonstrated that the argument from second occurrence focus is invalid. This argument implies a need for *strong*, pragmatic theories of focus rather than *weak* theories in which the interpretative effects of focus are mediated via the syntax-semantics interface. The argument took as its empirical basis a claim that certain semantic foci are not phonologically focus marked, and, under further assumptions, hence not F-marked. However, we have shown that there are no grounds for denying that these semantic foci are phonologically focus marked: second occurrence foci have acoustic properties which, in the absence of phonological focus marking, would be unexplained.

In section 6.3 we considered further data from the semantics literature, based on variants of the canonical second occurrence focus paradigm. We argued that in several cases the observed occurrence of semantic foci lacking pitch accents could be given a similar explanation to the canonical cases. In particular, these examples involve semantic foci occurring in a post-nuclear tail. The relevant cases are Rooth's *rice farmer* example (19) and Roberts' *dinner invitation* example (21). One type of example in the literature (22), due to Dryer, is potentially more problematic, since it is claimed to involve a semantic focus lacking phonological focus marking but in pre-nuclear po-

<sup>37</sup> As a matter of fact, Jaeger (2004, section 5) presents some evidence that intensity marking, in contrast to durational marking, is also subject to a relatively strong reduction on second occurrence focus expressions.

sition. Further study is required to establish whether there are any acoustic correlates of phonological focus marking in these cases, and, indeed, whether there is a pitch accent.

Apart from the issue of whether the semantic focus is in post-nuclear position, another significant question is whether exact repetition of material is critical in producing the phenomenon of de-accented second occurrence foci. Although there is prior discussion in the literature (cf. footnote 34), this question has never been investigated experimentally. Indeed, none of the example types discussed in 6.3 have been studied experimentally in any published work, so these example types remain as avenues for further research.

In the introduction to this paper we mentioned a pessimistic observation of Ladd's, i.e. "that proposals about intonational meaning are not a reliable source of evidence on intonational phonology." (Ladd, 1996, p.102) We hope to have shown that work on intonational meaning can provide a reliable source of evidence on intonational phonology, and conversely that intonational phonology can provide a reliable source of evidence on intonational meaning. Without collaborative work across linguistic sub-disciplines, there would be a danger of circularity. The methodological moral of this paper is that progress in the study of intonation is possible if interface phenomena, like second occurrence focus, are studied simultaneously from both semantic/pragmatic and phonetic/phonological perspectives.

## Appendix

### A. Stimuli

1. a) *Twins Kate and Jane usually get lots of cards from their friends on their birthday. But Jim only sent Kate a card today.  
Even Jack only sent Kate a card today.*
- b) *Kate usually gets lots of nice presents on her birthday. But her brother only gave Kate a card today.  
Even her mother only gave Kate a card today.*
2. a) *Pete really needed an injection to ease his pain. But the nurse only gave Pete a pill today.  
Even the doctor only gave Pete a pill today.*
- b) *Both Pete and Edward are suffering from the flu. But the nurse only gave Pete a pill today.  
Even the doctor only gave Pete a pill today.*
3. a) *Both Sid and his accomplices should have been named in this morning's court session. But the defendant only named Sid in court today.  
Even the state prosecutor only named Sid in court today.*

- b) *Defense and Prosecution had agreed to implicate Sid both in court and on television. Still, the defense attorney only named Sid in court today. Even the state prosecutor only named Sid in court today.*
4. a) *The family cat either stays in the tent or caravan. But mom only let the cat in the tent today. Even the kids only let the cat in the tent today.*
- b) *The cat and the dog usually stay in the tent. But mom would only let the cat in the tent today. Even the kids would only let the cat in the tent today.*
5. a) *At the San Francisco zoo, the chimps love nuts and fruit. But tourists always throw nuts to the chimps there. Even the guides always throw nuts to the chimps there.*
- b) *At the Los Angeles zoo, both chimps and baboons love nuts. But tourists always throw nuts to the chimps there. Even the guides always throw nuts to the chimps there.*
6. a) *You might think that in the prestigious Clark company of architects, all drafts were done on the computer. But the intern always uses a pen for drafts there. Even the chief architect always uses a pen for drafts there.*
- b) *In some architecture companies, final versions of floor plans are drawn with pens, but this is different at Flemming Associated Architects. The intern always uses a pen for drafts there. Even the chief architect always uses a pen for drafts there.*
7. a) *You might think that Texas drugstores sell both small toys and sweets to kids. But they always sell sweets to kids there. Even Walgreens always sells sweets to kids there.*
- b) *You might think that Texas drugstores sell sweets to both adults and kids. But they always sell sweets to kids there. Even Walgreens always sells sweets to kids there.*

## **B. Further details on the production experiment**

This appendix contains further information on the work reported in section 4.

### **B.1. EXCLUSION CRITERIA**

As mentioned in section 4.3 not all elicited sentences could be used in the phonetic analysis. Some items were lost through experimental error, though this did not affect overall balance of the study, and some items were qualitatively not acceptable. In this section, we elaborate on our exclusion criteria.

A blind review of the data was performed and all experimental sentences that contained hesitations, stutter, or other significant reading effects were removed. Whenever a sentence was removed from the analysis, the matching

sentence (with the inverted second occurrence focus assignment) was also removed. This resulted in the removal of  $17 \times 2 = 34$  experimental items (corresponding to 6% of all items). A lost sound file and two targets on which no measures could be obtained resulted in another 0.9% data loss. Each item contained two targets (i.e. the direct and the indirect object NP), yielding 1048 targets for phonetic measures. Since subjects had read each stimulus twice, all phonetic measures were averages over those two instances (in some cases only one instance was available because the other instance had already been excluded). After averaging over the repeated items, 556 of the expected 560 items (7 dialogues eliciting 2 different focus structures each, read twice by 20 speakers) were available for the duration, energy, and intensity analyses (this corresponds to a data loss of 0.7%). The pitch analysis was performed on 544 items (2.8% data loss). This was due to further exclusion of items for which either (a) the pitch tracking failed to find a pitch value, (b) the observed pitch value was likely to be incorrect (i.e. the maximum pitch on the target deviated more than 2 standard deviations from the mean pitch of the whole utterance), (c) the pitch on the target had a range of 0, or (d) an otherwise usable item was removed due to problems (a) to (c) on its matching item.

## B.2. TESTS OF NORMALITY

The standard normality tests were conducted and in cases where a dependent variable was not sufficiently normally distributed the usual transformations were used to meet the normality conditions ( $\ln(x)$  for duration data;  $x^2$  for the energy measure;  $\sqrt{x}$  for rms intensity).

## References

- Anderson, S.: 1972, 'How to get even'. *Language* **48**, 893–906.
- Bartels, C.: 1997, 'Acoustic Correlates of 'Second Occurrence' Focus: Towards an Experimental Investigation'. In: H. Kamp and B. Partee (eds.): *Context-dependence in the Analysis of Linguistic Meaning (Proceedings of the workshops in Prague and Bad Teinach)*. University of Stuttgart, pp. 11–30.
- Bartels, C. and J. Kingston: 1984, 'Salient pitch cues in the perception of contrastive focus'. In: *Focus and Natural Language Processing, vol. 1: Intonation and Syntax*. Heidelberg, pp. 11–28.
- Beaver, D. and B. Clark: 2003, 'Always and Only: Why not all Focus Sensitive Operators are Alike'. *Natural Language Semantics* **11**, 323–362.
- Beckman, M. and G. A. Elam: 1997, 'Guidelines for ToBI Labelling (version 3.0)'. Unpublished ms., Ohio State University.
- Beckman, M. E.: 1986, *Stress and Non-Stress Accent*. Dordrecht: Foris.
- Bolinger, D.: 1958, 'A theory of pitch accent in English'. *Word* **14**, 109–149.
- Büring, D.: 1997, *The Meaning of Topic and Focus — The 59<sup>th</sup> Street Bridge Accent*. London: Routledge.



- Chomsky, N.: 1972, *Studies on Semantics in Generative Grammar*, Chapt. Deep structure, surface structure, and semantic interpretation, pp. 62–119. The Hague: Mouton.
- Chomsky, N.: 1976, 'Conditions on rules of grammar'. *Linguistic Analysis* 2, 303–351.
- Chomsky, N.: 2000, *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, Chapt. Minimalist Inquiries: The Framework, pp. 89–155. MIT Press.
- Clark, H. H.: 1973, 'The Language-as-Fixed-Effect Fallacy'. *Journal of Verbal Learning and Verbal Behavior* 12, 335–359.
- Cohan, J.: 2000, 'The Realization and Function of Focus in Spoken English'. Ph.D. thesis, University of Texas at Austin.
- Dryer, M. S.: 1994, 'The pragmatics of association with *only*'. Paper presented at the 1994 Winter Meeting of the L.S.A. Boston, Massachusetts.
- von Fintel, K.: 1994, 'Restrictions on Quantifier Domains'. Ph.D. thesis, UMass Amherst.
- von Fintel, K.: 1995, 'A Minimal Theory of Adverbial Quantification'. In: B. Partee and H. Kamp (eds.): *Context Dependence in the Analysis of Linguistic Meaning: Proceeding of the Workshops in Prague, February 1995, Bad Teinach*. pp. 153–193.
- Foolen, A.: 1993, 'De Betekenis van Partikels. Een Dokumentatie van de Stand van het Onderzoek met Bijzondere Aandacht voor 'Maar''. Ph.D. thesis, Katholieke Universiteit van Nijmegen.
- Geurts, B. and R. van der Sandt: 2004, 'Interpreting Focus'. *To appear: Theoretical Linguistics*.
- Grice, H. P.: 1989, *Studies in the Way of Words*. Cambridge, Massachusetts: Harvard University Press.
- Gussenhoven, C.: 1984, *On the Grammar and Semantics of Sentence Accents*. Foris, Dordrecht.
- Hajičová, E.: 1973, 'Negation and topic vs. Comment'. *Philologica Pragensia* 16, 81–93.
- Hajičová, E.: 1984, 'Presupposition and allegation revisited'. *Journal of Pragmatics* 8, 155–167.
- Hajičová, E., B. H. Partee, and P. Sgall: 1998, *Topic-Focus Articulation, Tripartite Structures, and Semantic Content*. Kluwer Academic Publishers.
- Halliday, M.: 1967, 'Notes on Transitivity and Theme in English (Part 2)'. *Journal of Linguistics* 3, 199–244.
- Harrington, J., J. Fletcher, and M. E. Beckman: 2000, 'Manner and place conflicts in the articulation of accent in Australian English'. In: *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge University Press, pp. 40–51.
- Hayes, B.: 1995, *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.
- Hedberg, N.: 2003, 'The Prosody of Contrastive Topic and Focus in Spoken English'. In: *Pre-proceedings of the Workshop on Information Structure in Context, Stuttgart University, November 15-16th*. pp. 141–152.
- Hedberg, N. and J. M. Sosa: 2001, 'The Prosodic Structure of Topic and Focus in Spontaneous English Dialogue'. In: *Topic and Focus: A Workshop on Intonation and Meaning*. University of California, Santa Barbara, Linguistic Society of America, Institute of Linguistics.
- Hedberg, N. and J. M. Sosa: 2002, 'The Prosody of Questions in Natural Discourse'. In: *Proceedings of Speech Prosody 2002 (the First International Conference on Speech Prosody)*. Aix-en-Provence, France, pp. 375–378.
- Horn, L.: 1996, 'Exclusive Company: *Only* and the Dynamics of Vertical Inference'. *Journal of Semantics* 13, 1–40.
- Huss, V.: 1978, 'English word stress in post-nuclear position'. *Phonetica* 35, 86–105.
- Jackendoff, R.: 1972, *Semantic Interpretation in Generative Grammar*. MIT Press.
- Jacobs, J.: 1983, *Fokus und Skalen*. Tübingen: Niemeyer.

- Jaeger, T. F.: 2004, 'Only always associates audibly. Even if only is repeated The prosodic properties of second occurrence focus in English'. Manuscript.
- de Jong, K. J.: 1995, 'The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation'. *Journal of the Acoustical Society of America* **97**, 491–504.
- Kadmon, N.: 2001, *Formal Pragmatics: Semantics, Pragmatics, Presupposition and Focus*. Oxford: Blackwell.
- Kayne, R.: 1998, 'Overt vs. Covert Movement'. *Syntax* **1**(2), 128–191.
- Krifka, M.: 1992, 'A Compositional Semantics For Multiple Focus Constructions'. *Informationsstruktur und Grammatik* **4**.
- Krifka, M.: 1997, 'Focus and/or Context: A second look at second occurrence expressions'. In: H. Kamp and B. Partee (eds.): *Context-dependence in the Analysis of Linguistic Meaning (Proceedings of the workshops in Prague and Bad Teinach)*. University of Stuttgart, pp. 253–276.
- Krifka, M.: 2004, 'Association with Focus Phrases'. Unpublished ms., Humbolt University, Berlin.
- Ladd, D. R.: 1980, *The Structure of Intonational Meaning: Evidence from English*. Indiana University Press.
- Ladd, D. R.: 1996, *Intonational Phonology*. Cambridge University Press.
- Ladefoged, P.: 2003, *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell.
- Levinson, S.: 2000, *Presumptive Meanings*. Cambridge: MIT Press.
- Martí, L.: 2003, 'Contextual Variables'. Ph.D. thesis, University of Connecticut, Storrs, CT and MIT Working Papers in Linguistics.
- de Mey, S.: 1991, "'Only' as a Determiner and as a Generalized Quantifier". *Journal of Semantics* **8**(1-2), 91–106.
- Montague, R.: 1974, 'The Proper Treatment of Quantification in Ordinary English'. In: R. Thomason (ed.): *Formal Philosophy: Selected Papers of Richard Montague*. New Haven: Yale University Press.
- Nevalainen, T.: 1991, *But, Only, Just: Focusing Adverbial Change in Modern English 1500-1900*. Helsinki: Société Néophilologique.
- Ostendorf, M., P. J. Price, and S. Shattuck-Hufnagel: 1995, 'The Boston University radio news corpus'. Technical report.
- Pallier, C.: 1997, 'Shuffle: A program to perform random sampling and to shuffle lists specifying constraints'. <http://www.ehess.fr/centres/lscp/persons/pallier/ressources/index.html>.
- Partee, B. H.: 1999, 'Focus, Quantification, and Semantics-Pragmatics Issues'. In: P. Bosch and R. van der Sandt (eds.): *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge University Press, pp. 213–231.
- Paul, H.: 1888, *Principles of the History of Language*. London: Swan Sonnenschein, Lowrey, & Company. translation of *Principien der Sprachgeschichte*.
- Pierrehumbert, J. B.: 1980, 'The Phonology and Phonetics of English Intonation'. Ph.D. thesis, MIT.
- Pierrehumbert, J. B. and J. Hirschberg: 1990, 'The Meaning of Intonational Contours in the Interpretation of Discourse'. In: Cohen, Morgan, and Pollack (eds.): *Intentions in Communication*. MIT Press, pp. 271–311.
- Prince, E.: 1981, 'Toward a Taxonomy of Given-New Information'. In: P. Cole (ed.): *Radical Pragmatics*. New York/London: Academic Press, pp. 223–255.
- Raaijmakers, J. G. W., J. M. C. Schrijnemakers, and F. Gremmen: 1999, 'How to deal with "The Language-as-Fixed-Effect Fallacy": Common Misconceptions and Alternative Solutions'. *Journal of Memory and Language* **41**, 416–426.

- Roberts, C.: 1995, 'Domain Restriction in Dynamic Semantics'. In: E. Bach, E. Jelinek, A. Kratzer, and B. Partee (eds.): *Quantification in Natural Languages*. Kluwer Academic Publishers, pp. 661–700.
- 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.: 1985, 'Association with Focus'. Ph.D. thesis, UMass. Amherst: Graduate Linguistics Students Association.
- Rooth, M.: 1992, 'A Theory of Focus Interpretation'. *Natural Language Semantics* **1**, 75–116.
- Rooth, M.: 1996a, 'Focus'. In: S. Lappin (ed.): *The Handbook of Contemporary Semantic Theory*. London: Basil Blackwell, pp. 271–297.
- Rooth, M.: 1996b, 'On the Interface Principles for Intonational Focus'. In: T. Galloway and J. Spence (eds.): *SALT VI*. Ithaca, NY, pp. 202–226.
- Schwarzschild, R.: 1997, 'Why Some Foci Must Associate'. Unpublished ms., Rutgers University.
- Schwarzschild, R.: 1999, 'GIVENness, AvoidF and other Constraints on the Placement of Accent'. *Natural Language Semantics* **7**(2), 141–177.
- Selkirk, E.: 1984, *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, E.: 1995, 'Sentence Prosody: Intonation, Stress, and Phrasing'. In: J. A. Goldsmith (ed.): *The Handbook of Phonological Theory*. London: Basil Blackwell, pp. 550–569.
- Silverman, K.: 1987, 'The Structure and Processing of Fundamental Frequency Contours'. Ph.D. thesis, Cambridge University.
- Sluijter, A. M. C. and V. J. van Heuven: 1996a, 'Acoustic correlates of linguistic stress and accent in Dutch and American English'. In: *Proceedings of the Fourth International Conference on Spoken Language Processing, Volume 2*. pp. 630–633.
- Sluijter, A. M. C. and V. J. van Heuven: 1996b, 'Spectral balance as acoustic correlate of linguistic stress'. *Journal of the Acoustic Society of America* **100**(4).
- von Stechow, A.: 1985/1989, 'Focusing and Backgrounding Operators'. Technical report, Universität Konstanz.
- Steedman, M.: 2000, 'Information Structure and the Syntax-Phonology Interface'. *Linguistic Inquiry* **31**(4), 649–689.
- Taglicht, J.: 1984, *Message and Emphasis*. London and New York: Longman.
- Tancredi, C. D.: 1990, 'Not only *even*, but even *only*'. MIT, Cambridge, Mass.
- Tanenhaus, M. K., M. J. Spivey-Knowlton, K. M. Eberhard, and J. E. Sedivy: 1995, 'Integration of visual and linguistic information in spoken language comprehension'. *Science* **268**, 1632–1634.
- Terken, J. and J. Hirschberg: 1994, 'Deaccentuation of words representing given information: effects of persistence of grammatical function and surface position'. *Language and Speech* **37**, 125–45.
- Truckenbrodt, H.: 1995, 'Phonological phrases: Their relation to syntax, focus, and prominence'. Ph.D. thesis, MIT.
- Turk, A. E. and L. White: 1999, 'Structural influences on accentual lengthening'. *Journal of Phonetics* **27**, 171–206.
- Vallduví, E.: 1990, 'The Information Component'. Ph.D. thesis, University of Pennsylvania.
- Williams, E.: 1997, 'Blocking and Anaphora'. *Linguistic Inquiry* **28**(4), 577–628.
- Wolters, M. and D. Beaver: 2001, 'What does *he* mean?'. In: J. Johanna Moore and K. Stenning (eds.): *Proceedings of the Twenty-Third Annual Meeting of the Cognitive Science Society*. New Jersey, pp. 1176–1180.