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. 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 that (i) there is no mean pitch rise on SO foci, (ii) SO foci are marked by longer duration and greater energy, and (iii) listeners are able to detect the difference between SO foci and nonfoci. 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.*

1. Introduction. In studying information-structural notions like focus, topic, and contrast, semanticists depend on phoneticians and phonologists to categorize the types of marking available for interpretation. If only it were that simple! Despite enormous progress in the last few decades, the development of theories of prosody has been hindered by an absence of clear evidence as to what the underlying semantic distinctions are. As Ladd (1996:102) writes: ‘For the present, proposals about intonational meaning are not a reliable source of evidence on intonational phonology’.1 Our point of departure is that both the phonological study of prosody and the semantic study of information structure stand to benefit from collaboration between semanticists and laboratory phonologists.

That prosodic marking affects truth-conditional meaning was observed by Paul (1888:312ff.). Consider our examples in 1, where small caps 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, that is, 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 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.

* This work was partially funded by a Stanford University Office of Technology Licensing grant held by David Beaver, and by the Scottish Enterprise Stanford-Edinburgh Link grants ‘Sounds of Discourse’ and ‘Synthesis’. This article has benefited from commentary of Language referees 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 from numerous seminar audiences. Our greatest intellectual debt is to Mats Rooth, whose insightful observations motivated us to perform an experimental study.

1 The reader should set Ladd’s comments against a background of significant advances in the study of prosody in the previous twenty years. These advances are exemplified by the development of prosodic transcription standards (e.g. the ToBI system of Silverman et al. 1992) and the development of several corpora with prosodic transcriptions (e.g. the BRN corpus; Ostendorf et al. 1995). Note that the term intonation (and, similarly, intonational phonology) is often used with a broader meaning than the etymology would suggest, and includes suprasegmental features other than pitch, such as duration and energy. The term prosody reliably includes all suprasegmental features.
Many standard approaches to grammar separate phonology and semantics into components that cannot exchange information directly. In such approaches, the relationship between prosody and meaning is mediated by syntax. Therefore, many authors, from Halliday (1967) and Chomsky (1972) on, have concluded that there are syntactic features that 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, for example, Kayne 1998 and Tancredi 1990. Many other authors also postulate some syntactic effect of prosodic focus, though they might use features or derivations in a quite different way; see, for example, Krifka 1992, Rooth 1985, Steedman 2000, and von Stechow 1985/1989.

What all of 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 are pragmatic. This view strikes such a radically discordant note with the approaches listed earlier that it is worth exploring whether the claims of the pragmatists are justified. The current article deals with one particular claim that the pragmatists 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 phonological 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.

Pragmatic theories of focus could be seen as part of an enterprise of keeping the interfaces between phonology, morphosyntax, and semantics as narrowly circumscribed as possible. These pragmatic theories lean on extragrammatical mechanisms to make up the shortfall, much as in the model of Grice (1989). If the interpretive effects of focus could be explained pragmatically, then the phenomenon of focus would provide us with little insight into how suprasegmental information is represented at the morphosyntax/phonology and semantics/morphosyntax interfaces. But if there is a grammaticized connection between focus and meaning, then that places minimal constraints on the suprasegmental information that must be represented at the morphosyntax/phonology interface. It also places a lower limit on what information must be passed between syntax and semantics. What we show here is that the major argument that 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 suprasegmental information like that marking focus is treated differently from 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 here has been studied extensively in the semantic literature and involves expressions that 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 of focus (e.g. by prosody, syntactic position, or morphology) varies crosslinguistically. All theories of focus-sensitivity agree that focus interacts with expressions like only in linguistic contexts like that exemplified in 1. A special case of the controversy over whether the interpretation of focus is mediated via syntax, however, involves 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?

Examples are given in the literature that involve an apparent dislocation between the interpretation of only and the position of prosodic prominence. Most of these examples fall under the rubric of SECOND-OCURRENCE (SO) focus, where a repeated focused item apparently lacks a pitch accent. Example 2, adapted from Partee (1999:215), illustrates SO 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 we discuss shortly, is usually taken to indicate marking at a syntactic level, while our ‘SOF’ marking allows us to remain (temporarily) agnostic as to whether the item described as a SO focus is in fact focused at a syntactic level or any other.

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

Partee (1999) summarizes the problem succinctly. If only is a focus-sensitive operator (i.e. needs a prosodically prominent element in its scope), then the two occurrences of only eats vegetables in 2 should have the same analysis. But 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 1999:215-16).

Following up on observational evidence presented by Rooth (1996b), we present experimental data concerning precisely the cases where the pragmaticists have claimed interpretational effects without phonological marking. We show that in these cases the claim that there is no phonological marking is wrong. Specifically, we present the results of multisubject production and perception experiments in which we examine the acoustic correlates of SO focus in the scope of operators only and always, which are both assumed to be focus-sensitive in the bulk of prior literature. Our results confirm that there are acoustic correlates involving duration and energy of the focused item, and show that the acoustic correlates are perceptible. These conclusions make clear that the phenomenon of SO focus does not support an argument that current syntactic and semantic theories of focus are ‘fundamentally incoherent’.

For semantics and pragmatics, our results contribute to a long-running debate. But the phenomenon of SO 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 postnuclear domain, the part of an intonational phrase that follows the nuclear accent. As becomes clear, the results (i) imply an increased role for phrasal stress in marking focus, while showing that pitch accents are not required, and (ii) suggest that although postnuclear deaccenting restricts the appearance of pitch accents, phrasal stress distinctions must be permitted in the postnuclear domain.

---


3 But see discussion in §4: 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.
clear stretch. This latter point buttresses arguments Ladd (1996) has put forward for the existence of postnuclear prominence, although our conclusions differ in detail from his.

The article is organized around three hypotheses: first, that SO focus is prosodically distinguished from nonfocus; second, that SO focus is marked differently from other foci; and, third, that SO focus marking is perceptible. We thus first provide the technical background of the article as regards the syntax, semantics, pragmatics, and phonetics/phonology of focus. The production and perception experiments are then presented, followed by a more general discussion that connects the results to wider theoretical concerns and possibilities for future work.

2. BACKGROUND. The following discussion involves 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 has sometimes been referred to as PROSODIC OR INTONATIONAL FOCUS), F-MARKING (which is syntactic), and SEMANTIC FOCUS (an aspect of the representation of meaning).4 We must now explain why the absence or presence of certain acoustic properties has been claimed to be a decisive factor for theories of focus.

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 involve only within an NP (hence ‘NP-only’). The semantic interpretation of 3a to 3b is related in 4.

(3) a. Only felines are immortal.
   b. Every immortal is feline.

(4) ONLY(FLINE)(IMMORTAL) ⇒ EVERY(IMMORTAL)(FLINE)

In 4, ONLY represents the meaning of the word only. Let us call the first argument of only its restrictor, that is, FELINE, and the second its scope, that is, IMMORTAL. Then only is just like every except that the operator we use to represent its meaning has its arguments reversed.5 In the case of NP-only, configuration fully determines the restrictor and scope, but, as discussed below, the situation is more complex for VP-only.

The standard mechanisms for argument selection are (i) syntactic configuration and (ii) 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

---

4 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, focus in question-answer pairs, and focus in clefts or left dislocations. These foci are sometimes taken to share phonological features, but this is far from obvious (cf. Bartels & Kingston 1984, Hedberg 2003, and Hedberg & Sosa 2001).

5 For evidence that only and every are related in this way semantically, see, for example, 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). 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 adherents of these approaches. It is presumably because of the nonstandard 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.
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 (repeated from 1a,b with explicit F-marking) and their intended meanings 6a,b.

(5) a. Jan only gave Bill [money]F.

(6) a. \text{ONLY} (\text{MONEY}) (\forall x [\text{JAN GAVE BILL} x])
   \text{i.e.} everything Jan gave Bill was money.
   b. \text{ONLY} (\text{BILL})(\forall x [\text{JAN GAVE} x \text{ MONEY}])
   \text{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. 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 has 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, some sentence connectives (because, counterfactual if-then), negation, questions, and emotive predicates (regret, be glad that) (Hajicova et al. 1998, Rooth 1996a, Beaver & Clark 2007:Ch. 3).

Clearly not all focus-sensitive expressions are quantifiers, which makes it inappropriate to label the arguments of those expressions as restrictor and scope. All of the expressions can be thought of semantically as multi-argument operators, however, such that one argument position is filled by a focused expression. We term this argument the semantic focus of the operator. 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 phonological focus marking.

\footnote{It might be argued that \textsc{jan} should be kept out of the argument of \textsc{only} in 6a,b, in which case we could think of the meaning of 5a, for example, as underlingly consisting of separate NP and VP meanings combined conjunctively: \textsc{only} (\textsc{money}\ (\forall x [\text{subj gave bill} x]) \land \text{subj} = \textsc{jan}. The interpretation of the subject in sentences with VP focus-sensitive operators is not a concern of this article. See Beaver & Clark 2003 for an analysis.}

\footnote{It must be stressed that this is not a phenomenon peculiar to English: see, for example, the many Dutch particles discussed by Foolen (1993).}

\footnote{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 phonologically focus-marked. See the introduction to §2.}
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 have 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 the minimalist program (Kayne 1998), the focus-sensitive operator only can attract the F-marked Bill in 5b to the specifier of the phrase that only heads.9

In situ approaches became dominant in the semantics community after Rooth (1985) observed a range of problems with movement accounts.10 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. 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 (1985/1989) and Krifka (1992).11 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 nonfocused constituents.

2.3. Strong theories of focus-sensitivity. The in situ and movement models described above explain the interpretational effects that we term focus-sensitivity in terms of grammatical mechanisms conditioned on syntactic F-marking. Two questions arise with regard to 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 relevant to this article, 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

9 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 (2006) 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.
10 Rooth’s (1985) empirical arguments against movement-based accounts include the interaction of focus-sensitive expressions with multiple foci and syntactic islands. Krifka (2006) suggests 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.
11 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.
surprising: if there are some environments where the semantic focus is not what is F-marked, 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.12

Rooth 1992 introduced a more nuanced view, one that allows for F-marking to be dissociated from the semantic focus. Rooth discussed 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 about 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.13 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 that have been termed focus-sensitive have a free variable; the argument that we have been calling the semantic focus is introduced lexically as a free variable. 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]* 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 such a view, free variables involved in focus-sensitivity are resolved pragmatically in much the same way that 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 do not discuss the details of how pragmatic models might achieve this.14 What is pertinent to the current enterprise is that such models would predict that an F-marked constituent might not become the semantic

12 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 vs. an invite*).

13 Rooth also considers the possibility of what he terms an intermediate theory, one that allows for lexical stipulation, but only of a very limited sort. We too von Fintel’s (1994, 1997) account of quantificational adverbs to be an example of an intermediate theory.

focus, and conversely that the semantic focus could be a constituent that is not F-marked.

It should now be broadly clear how 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. 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 is grammaticized but optional. We do 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 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 here. But first, there is one further subtle distinction to elucidate: that between phonological focus marking and syntactic F-marking. We have described above how models may differ with regard 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 psychoacoustic 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 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 that is phonologically focus-marked.

Relating phonological focus to F-marking are theories of Focus Projection (e.g. 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. But for our purposes it is enough that in models of focus projection the principle given in 7 is usually adhered to.

(7) FOCUS PROJECTION
An F-marked constituent must contain at least one phonologically focus-marked subpart.16

15 Beaver and Clark (2003, 2007) provide syntactic and semantic criteria for obligatoriness of F-marking and show that F-marking is obligatory in the scope of some expressions that are usually taken to be focus-sensitive.

16 Note that what we call F-marking corresponds to what some 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 here.
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 also 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 of SO focus.

2.5. The argument from SO focus. We have already seen (an adaption of) an example of SO focus from Partee, repeated here.

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

Let us repeat the analysis of SO focus examples given in §1, but using some of the concepts now 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, that is, 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, contrary to what is expected on the movement and in situ accounts of focus sketched above.

It is this type of analysis that has allowed theorists to argue that SO-focus data counterexemplifies weak theories of focus interpretation. As mentioned previously, many authors hold that SO-focus data provides evidence against the major in situ focus accounts (alternative semantics and structured meaning semantics). These include Büring (1997), Dryer (1994), Kadmon (2001), Martí (2003), Roberts (1996), Rooth (1992), Schwarzschild (1997), Vallduví (1990), and Williams (1997), all of whom interpret SO-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 SO-focus data against grammaticized theories of focus-sensitivity—most arguments in the literature from SO focus have basically the same structure.17

(i) Weak, grammaticized theories of focus-sensitivity require the semantic focus to be F-marked.
(ii) By focus projection, these theories also predict that the semantic focus should contain a phonological focus.
(iii) a. SO foci are semantic foci,
    b. but SO foci contain no phonological focus-marking.
(iv) Therefore we should prefer a strong, pragmatic account of focus.

This article addresses step (iii)b of the above argument. This step could be attacked in two ways. First, one might argue, purely theoretically, that observations of a lack of acoustic prominence cannot possibly justify a claim of a lack of phonological focus marking, so that even if there were no acoustic prominence, step (iii)b would be unjusti-

17 See for example Rooth (1996b:206) for a similarly organized argument.
fied. Yet we seek to undermine the standard argument even more completely. We seek to show that it is irrelevant whether the absence of acoustic prominence could, in principle, evidence a lack of phonological focus. For what we test is whether there is an acoustic prominence in the first place. If there is, then (iii)b is false. To set this up, we now consider the nature of acoustic prominences that might potentially signal phonological focus marking.

2.6. The phonetics and phonology of focus in English. In order to characterize the phonetics and phonology of focus marking it is necessary to briefly review the basic elements of English prosody. We adopt here the model developed in Pierrehumbert 1980 and subsequent works (especially Beckman & Pierrehumbert 1986 and Ladd 1996), which also forms the basis for the ToBI system for the transcription of prosody (Beckman & Elam 1997). In this model, the prosody of an utterance involves three interrelated components: intonation, phonological phrasing, and stress. An utterance is divided into phonological phrases, each of which is characterized by a tonal melody and is usually marked by lengthening of the final syllable or syllables (Shattuck-Hufnagel & Turk 1996). Phonological phrasing is related to, but distinct from, syntactic phrasing (see Shattuck-Hufnagel & Turk 1996 for a review).

The intonational melody is represented as a sequence of high and low tones. These tones are of two types: pitch accents and boundary tones. 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) and is associated with a stressed syllable (cf. Bolinger 1958 and 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. For example, a sentence like 1 can be produced as a single phrase 8a, or the subject can be phrased separately from the verb phrase as in 8b. In the latter case, the sentence-medial phrase boundary is marked by a low boundary tone (L−).

(8) a. Jan only gave Bill money.
   \[H^{*} H^{*} L-L%\]

   b. Jan only gave Bill money.
   \[H^{*}L-\] \[H^{*} L-L%\]

The stress pattern of an utterance specifies the relative prominence of syllables. Phonetically this prominence is realized by ‘a complex of properties that can be related to greater force of articulation, including increased intensity and duration, and shallower spectral tilt’ (Ladd 1996:58). Pitch accents can only be associated with the syllables that receive the strongest stresses in a phrase (Pierrehumbert 1980, Selkirk 1995).

Focus marking typically involves the alignment of a pitch accent with the primary stressed syllable of the focused item. Although focus is most frequently marked by an H* accent (e.g. Hedberg 2003, Hedberg & Sosa 2001), there are many exceptions to this generalization. For example, in yes-no questions, focus is often marked by an L + H* accent (Hedberg & Sosa 2002). More specifically, focus is usually marked by a nuclear pitch accent, that is, 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 8a there is a pitch accent on

---

18 The ToBI system distinguishes two levels of phrasing: utterances are divided intointonational phrases, which in turn consist of one or more intermediate phrases. This distinction is not important here, and we make no claims as to the number of levels prosodic phrasing can take (cf. Wagner 2005).
Jan, but this first accent is perceived as less prominent than the nuclear accent that marks the focus, Bill. This observation corresponds to Jackendoff's (1972:230) generalization that focus is marked by 'the main stress . . . in the sentence', since the main sentence stress receives a nuclear accent. It is possible for a sentence to contain more than one nuclear accent if it contains multiple phrases, as in 8b, so it is more accurate to say that focus is marked by the strongest stress in a phonological phrase. It is also possible for a sentence to contain more than one focus (Jackendoff 1972:258ff.), although the presence of multiple nuclear accents does not necessarily imply the existence of multiple foci (Ladd 1996:248ff.). That is, not every nuclear accent marks focus.

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 (F0), which is a measurable property of the speech signal. So the main acoustic correlates of a pitch accent are to be found in the F0 contour. For example, an H* pitch accent is typically realized by a local maximum in F0. It is well established, however, 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 & van Heuven 1996a, Turk & 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; rather, they may be correlates of stress, since pitch-accented syllables also have the highest levels of stress in a phrase. Treating duration and intensity as correlates of stress rather than pitch accent is consistent with evidence that longer duration and higher intensity correlate with stress even in the absence of pitch accent (Beckman 1986, Sluijter & van Heuven 1996a,b).

Bolinger (1958) has 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 §6, we argue that the distinction between phrasal stress and pitch accent is helpful in analyzing the differences between the marking of SO 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 stress. Accordingly, duration and intensity can also play a role in marking focus. Ladd (1996:226ff.) has gone further, suggesting that certain foci (e.g. in the answer to a question) can be marked without a pitch accent, so that nonpitch measures may become the primary indicator of focus. It is thus natural to examine the role of nonpitch measures in SO focus, and, as discussed below, we are not the first to do so.

### 2.7. The phonetics and phonology of SO focus

We now turn from the general issue of how focus is marked to the more specific issue of how SO focus is marked. It is generally accepted that SO foci do not contain a pitch movement, although we know of only one systematic experimental study of pitch movement in SO focus—Bartels 1997, discussed below. But it remains possible that a phonological correlate of SO focus corresponds acoustically to other measures discussed above, such as duration, intensity, and vowel quality.

We presented in §2.5 an explicit argument from the phenomenon of SO focus, an argument that would be undermined if it could be demonstrated that SO focus has phonological correlates. Rooth (1996b) presented suggestive evidence indicating just that, and his observations provided the impetus for the more systematic experimental studies that we report here.
Rooth recorded himself uttering minimal pairs of dialogues differing with regard to which item was the SO focus. He found that the pitch track was flat in SO-focus position, but that he could auditorily detect the marking of SO focus in his own productions. Furthermore, he noted that this marking of SO 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 SO-focus expressions, these expressions have greater duration and absolute intensity than their nonfocused counterparts.

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

3. A production experiment. We now report on a production experiment designed to test the hypothesis that SO focus is acoustically distinguished from nonfocus. This hypothesis competes with the null hypothesis implicitly assumed in the argument from SO focus (see §2.5) that SO focus is not prosodically marked. Earlier, we alluded to a possible explanation for the strong intuition of some researchers that SO focus is NOT focus-marked: SO focus is not realized acoustically in the same way as ordinary focus. A secondary goal of the experiment is to confirm the hypothesis that SO focus is marked differently from first-occurrence focus by singling out respects in which SO foci differ acoustically both from nonfocused expressions and from ordinary foci.

3.1. Description of the experiment. We ran a production experiment in which naive subjects read prepared written materials.

Subjects. Twenty US-born native speakers of English (ten female, ten male) were recruited, none with any training in linguistics.

Stimuli. An example of a minimal pair of discourses used as stimuli is given in 9 and 10. These three-sentence discourses are designed to probe SO-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 operator, even, is also present in 9c and 10c, typically causing the initial subject NP to bear nuclear accent, and leaving only and its associate (Sid or court) in a postnuclear position.

(9) 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.
(10) 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 9 and 10, the relevant potential foci are Sid and court. For all of the stimuli we used, the material following the second focus-sensitive operator in the third sentence (e.g. 9c and 10c) does not differ between the two members of the discourse pair. Thus, the phonological context for the segment of text containing 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 SO-focus effects.

Crucially, the focus in each of the (b) sentences in 9 and 10 was different, so the textually identical VPs in the (c) sentences should differ only in so far as they contain different expressions that are the SO focus. In 9c, Sid is the second occurrence of a focus, and court is nonfocal, whereas in 10c, Sid is nonfocal, 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 say that the expression is in the SO-FOCUS CONDITION, and otherwise we say it is in the NONFOCAL CONDITION.

Note that our stimuli, like those of Rooth (1996b), 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 SO foci Sid and court in the two discourses. Finally, note that we chose our stimuli to be maximally conservative with regard to the null hypothesis. In all stimuli, the SO focus follows another first-mentioned focus, which typically receives a nuclear accent (see §2.2). In other words, the SO-focus expressions investigated here not only are repeated but also appear in the postnuclear domain, thereby making accenting of the SO-focus expression very unlikely. We chose stimuli of this type to bias heavily against our hypothesis. We are interested in whether foci are acoustically marked even if they occur in prosodic environments that prevent marking by pitch accents.

We used a total of fourteen discourse stimuli, made up of seven minimal pairs like those above. In each case, the repeated focus-sensitive operator was only or always, while the focus-sensitive even was used to induce the nuclear accent in the subject of the final sentence of the discourse. In all cases the potential SO foci in the final sentence are not sentence-final, this being contrived by the use of an additional adverb, for example, today in 9c and 10c. 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 twenty-eight experimental stimuli were arranged in a pseudo-random order and intermingled with sixteen unrelated filler stimuli, also discourses. To control for recency effects that might occur when a subject encounters both members of a minimal pair of discourses,

19 A referee points out that in all of our stimuli, the choice of SO-focus placement is between two NPs, as opposed to an NP and some other constituent. The referee observes that conceivably the presence of the two NPs might induce the speaker to use a contrastive intonation. It is important to realize that such contrastive marking, in and of itself, would not be expected to yield a net effect on our measures, since we consider relative differences between the two NPs. Because of this design feature, however, it might be argued that our conclusions do not strictly cover SO-focus marking in general, but only SO-focus marking in the presence of potential contrast.

20 Fillers were drawn from a separate experiment, reported in Wolters & Beaver 2001, concerning accenting of pronouns.
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 twenty-eight experimental discourses consisted of seven minimal pairs repeated twice (for an example of a minimal pair, see 9 and 10 above). After recording all of 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), for example, Sid and court in 9c and 10c above. This procedure yielded 28 (discourses) \( \times \) 20 (subjects) \( \times \) 2 (probes per target sentence) = 1,120 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.

After all of the relevant word boundaries had been hand-labeled, 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), F0 range, and maximum, minimum, and mean F0 (all in Hz). A pseudo-energy measure was created by multiplying the standardized rms intensity value by the duration of the target. These measures were selected based on previous research on the acoustic correlates of focus, reviewed in §2.6.

For each phonetic measure taken on the target items we conducted both a by-subjects (henceforth Fl) and a by-items (henceforth F2) analysis of variance (ANOVA). Each ANOVA corresponded to a two-way crossed 2 \( \times \) 2 repeated measures design of SO Focus (whether the preceding context identified the target phrase as SO 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 (minF) was calculated from the Fl and F2 analyses. This maximally conservative approach was taken because the results reported below provide evidence against a long-standing opinion that SO focus is not phonetically marked. Although we report F1, F2, and minF values below,
our conclusions are based solely 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 SO focus.

A blind review process was performed prior to the statistical analysis. Overall, 0.7% of the data was excluded from the analysis of duration, intensity, and energy and 2.8% of the data was excluded from the pitch analysis (for a detailed description of the exclusion criteria, see Appendix B.I).\(^{22}\) The higher data loss for the pitch analysis is due to the well-known problems of pitch-tracking algorithms.\(^{23}\)

### 3.2. Results

The analyses revealed a significant main effect of SO focus on duration \((F(1,19) = 6.9, F(2,16) = 15.3, \text{min}F(1,24) = 4.8)\) and energy \((F(1,19) = 8.2, F(2,16) = 17, \text{min}F(1,24) = 5.5)\). SO-focus phrases were on average 6 ms longer than nonfocused phrases and were realized with 3.4% more energy. There was also a marginal effect of SO focus on standardized rms (root-mean-square) intensity \((F(1,19) = 7.1, F(2,16) = 8, \text{min}F(1,19) = 3.8)\), the standardized F0 range \((F(1,19) = 8.8, F(2,16) = 8, \text{min}F(1,17) = 4.2)\), and the standardized minimum F0 \((F(1,19) = 5.2, F(2,16) = 9.1, \text{min}F(1,23) = 3.3)\). SO-focus phrases were realized with 0.31 dB higher intensity, a 10.5% higher relative F0 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 SO focus was found for maximum F0 and mean F0. The means for all measures in both conditions of SO focus and the mean differences between the two conditions of SO focus are given in Table 1. Significant results are marked with an asterisk, marginal results with a plus sign.

<table>
<thead>
<tr>
<th>DEPENDENT VARIABLE</th>
<th>NONFOCAL MEAN ((\mu_{\text{nf}}))</th>
<th>SO FOCUS MEAN ((\mu_{\text{sf}}))</th>
<th>MEAN DIFFERENCE ((\Delta(\mu_{\text{nf}} - \mu_{\text{sf}})))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration (s)</td>
<td>521 0.267</td>
<td>521 0.273</td>
<td>*0.006</td>
</tr>
<tr>
<td>Std. rms intensity (dB)</td>
<td>521 -27.28</td>
<td>521 -26.97</td>
<td>+0.31</td>
</tr>
<tr>
<td>Rel. energy</td>
<td>521 0.0106</td>
<td>521 0.0117</td>
<td>*0.0011</td>
</tr>
<tr>
<td>Max F0 (Hz)</td>
<td>483 162.2</td>
<td>483 163.0</td>
<td>0.8</td>
</tr>
<tr>
<td>Mean F0 (Hz)</td>
<td>440 143.1</td>
<td>440 142.5</td>
<td>*0.6</td>
</tr>
<tr>
<td>Minimum F0 (Hz)</td>
<td>473 127.5</td>
<td>473 124.1</td>
<td>*-3.4</td>
</tr>
<tr>
<td>Range F0 (Hz)</td>
<td>457 34.6</td>
<td>457 38.7</td>
<td>*4.1</td>
</tr>
</tbody>
</table>

Table 1. Summary of means and mean differences for nonfocal and SO-focus expressions.

\(^{22}\) Most exclusions were due to problems tracking F0 in creaky voice. This could be problematic in an experimental study of prosody, since creaky voice is presumably correlated with prosodic factors. For example, creaky voice is less likely to obscure the F0 of a nuclear accent than to obscure the F0 of material in the postnuclear tail. Since F0-tracking problems affected less than 5% of the data, we did not attempt to build creaky voice in as a factor in our statistical analysis.

\(^{23}\) We also checked for possible effects of the prescriptive dogma that expressions such as exclusives should not be separated from what they modify, and thus in ‘good English’ are not focus-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, see Nevalainen 1991. It is conceivable that prescriptive 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. We cannot see how this would produce any consistent effect, however. Furthermore, the use of three-sentence discourses in which only the third sentence provides our main data set means that we have a baseline measure of how subjects produce sentences with focus-sensitive expressions. 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. On the basis of informal sampling, however, 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 here. It seems that prescriptive rules did not interfere with the subjects’ ability to produce foci naturalistically.
Unsurprisingly, word position had a main effect on rms intensity, energy, and all standardized F0 measures. All of these main effects were in the Fl analyses and only the two main effects of standardized mean F0 ($F(1,19) = 11$, $F(2,16) = 5.7$, $minF(1,13) = 3.8$) and standardized minimum F0 ($F(1,19) = 17.1$, $F(2,16) = 11.4$, $minF(1,15) = 6.8$) were significant in the minF analyses.

For reasons addressed in §3.1 (under the heading 'Presentation and procedure'), we also tested for interaction effects between SO focus and word position. We found a marginal interaction effect of SO focus and word position on the standardized mean F0 in both the Fl and the F2 analyses. Whereas the relative mean F0 was higher for SO focus than for nonfocused 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 ($F(1,19) = 4.4$, $F(2,16) = 5.3$, $minF(1,20) = 2.4$). We also found a marginal interaction effect in the Fl analysis for standardized rms intensity ($F(1,19) = 3.5$, $F(2,16) = n.s.$, $minF = n.s.$) and in the F2 analysis for the standardized minimum F0 ($F(1,19) = n.s.$, $F(2,16) = 6.7$, $minF = n.s.$). Whereas the standardized minimum F0 was the same for SO-focused and for nonfocused indirect objects, it was lower for SO-focused than for nonfocused direct objects. The inverse was observed for intensity. Relative rms intensity on direct objects was the same in both SO-focus conditions, but was higher for SO-focused than nonfocused indirect objects. This means that the main effect of SO focus on intensity is only due to intensity differences on the indirect object. The main effect on minimal F0 is only due to F0 differences on the direct object. The choice of which list (out of the four available with different stimulus orders) was read by the subject was not significant in any of the analyses.

3.3. Discussion. As expected on the basis of prior literature, we found no evidence for systematic pitch marking of SO focus. Instead, we found a small but significant lengthening of the SO focus in comparison with the same expression in a sentential context that differs only in lacking focus. Across all trials, this lengthening averaged 6 ms. There is also a statistically significant increase in energy, as well as marginally significant increases in intensity and the F0-range in the SO-focus condition. Finally, there was a marginally significant drop in minimum F0 in the SO-focus condition. The nature of interaction effects between word position and SO focus for minimum F0 and intensity suggests a more complex relation between these measures and SO-focus marking.

The statistical significance of the duration and energy results provides strong evidence that SO foci are focus-marked at some level of linguistic representation, since we have been careful to ensure that there is no difference between the SO-focus and nonfocused 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 SO focus not being formally marked. As far as SO-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 §2.6. The results we have presented show that SO focus is marked by the same cues. By contrast, fundamental frequency plays a central role in the realization of pitch accent, but appears to play only a minor role in the marking of SO focus: the only identifiable effect on F0 measures is a marginally significant lowering of minimum F0 in SO-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 SO focus is marked by a clear pitch accent, the accent is high (H*), which would lead us to expect an increase in F0 measures (at least the maximum and minimum F0 measures) on the SO-focus expression (see §5.1 for further discussion). In any case, the perception experiment reported in §4 provides evidence that this F0 difference does not play a significant role in the perception of the marking of SO focus. So we conclude that greater duration and energy are the primary correlates of SO focus.

While the existence of a difference between SO foci and their nonfocal counterparts in production indicates that speakers are making a distinction between the two, the mean differences between SO foci and their nonfocal 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.

4. A perception study. We conducted a forced-choice perception experiment to test the hypothesis that hearers can distinguish between the prosodic realizations of SO foci and nonfoci.

4.1. Description of the experiment.

Subjects. Fourteen native speakers of English (ten female, four male; age range: 23–28) were recruited. Ten were linguistically naive (eight female, two male; age range: 19–26) and lacked prior experience in linguistic perception studies, while the remaining four subjects (two female, two male; age range: 24–28) had prior training in linguistics.

Stimuli. Forty minimal discourse pairs with the focus-sensitive operator only were quasi-randomly selected from sentences produced in the experiment described in the previous section. Three (out of the four available) minimal discourse pairs with the focus-sensitive operator only were used (the pairs stimuli consisted of sixteen tokens each of the pairs in Appendix A, 2 and 3, and eight tokens of the pair in 4).

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 (SO) focus assignment.

The list of stimulus pairs was balanced so that two pairs were derived from each of the twenty speakers in the production experiment. For each of the three minimal dis-

---

24 Katy Carlson (p.c.) suggests that some of our recordings should be transcribed with a low pitch accent (L*) on the SO focus. If correct, this might explain the observed F0 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 F0 is expected to be low in any case. In support of the unaccented interpretation, we argue below (§5.2) that analyzing the usual realization of SO focus as unaccented enables us to relate the occurrence of this otherwise anomalous pattern of focus marking to the phenomenon of postnuclear deaccenting. By contrast, current generalizations about the use of L* accents (e.g. Pierrehumbert & Hirschberg 1990) do not provide any basis for expecting an L* accent in SO-focus environments, and in fact the few cases in our data where SO 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 SO-focus marking.
course pairs used, there were an equal number of instances in which the first NP was an SO focus and in which the second NP was an SO focus.\footnote{Recall also that subjects in the production study read every stimulus twice (see §3). So the stimuli used in the perception experiment controlled for repetition: if one stimulus in a pair was a token of a repeated production of the same discourse, then the other stimulus in that pair was also a repeat.}

To avoid systematic effects from our presentation of the stimuli, a number of order constraints were observed when the list of stimuli was constructed: (i) the same internal order of a pair (the first sentence has an SO-focus-marked first NP or an SO-focus-marked second NP) was not allowed to occur more than three times in a row; (ii) the same pair was not allowed to occur twice in a row. Furthermore, two versions of the experiment with inverse order were produced, and half of the subjects were given each version.

**Presentation and Procedure.** The acoustic stimuli were incorporated into an HTML page. Subjects first read through the instructions and 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 is given in bold face; see Figure 1) 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, as exemplified in Fig. 1.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Excerpt of HTML page displayed to subjects during the experiment.

Ten pairs of stimuli were displayed at a time. After every tenth pair of sentences, subjects were asked to take a break of thirty seconds to two minutes, which they could time themselves, before they continued on to the next page. Subjects were asked to proceed from pair to pair and always to listen to the first stimulus in a pair first. After that, they were allowed to listen to each of the stimuli at will. Subjects required fifteen to twenty-five minutes for the experiment and approximately two 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.

4.2. **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 twenty correct answers out of forty reveals that subjects on average perform significantly above chance \( t(13) = 7.7, p < 0.001 \) and that subjects performed ‘alike’ (i.e. the effect was not just driven by a few subjects; \( \chi^2(13) = 9.6, p > 0.7 \)). The average effect size is small to medium \( (\omega = 0.28) \). All subjects showed clear understanding of the task in the postexperiment interview.

None of the between-subject factors (i.e. gender, which of the two stimulus lists was used, or linguistic education) had a significant effect. Nor did ordering (i.e. when in
the list a stimulus was presented) have a significant effect, and this indicates that 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 ($\chi^2(1) = 1.2, p > 0.25$). Furthermore, all effects that hold for all subjects together also hold for linguistically naive subjects alone. There was also an item effect ($\chi^2(39) = 163.9, p < 0.001$). 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 our hypothesis that SO foci are perceptible. 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 on average led to better performance than the other two pairs. It is important for the current purpose that subjects performed significantly above chance for all three discourse pairs.26

4.3. Discussion. The results confirm that SO foci are perceptually distinguishable from nonfoci. Even though the phonetic correlates of SO focus observed in the production experiment are small, subjects were able to identify SO-focus-marked expressions correctly in 63% of all cases in the perception experiment. This shows that subjects are able to identify SO-focus marking without contextual clues, 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, overall performance decreased due to four items that were almost always judged contrary to our hypothesis of acoustically prominent SO foci (judged correctly by only three or fewer subjects of the fourteen total) and four other items that performed below chance (correctly judged by four to six subjects). The existence of items judged below chance is neither what is expected according to our hypothesis nor what would be expected in the absence of prosodic cues to SO focus. Since subjects should perform at chance in the absence of phonetic cues, these answers should have been judged correctly by 50% of all subjects.

Prior to the perception experiment, we had already examined the prosody on the second sentences in the read discourses, which contained one ordinary focus expression and no SO-focus expression. 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 contained a strong pitch accent on what was intended to be 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 SO-focus expression. The decision not to exclude those supposedly ill-formed stimuli a priori was made because this may have seriously confounded the study. It is clear, however, and, indeed, unsurprising, that the existence of occasional unnatural productions in the original reading task, perhaps due to the artificial experimental scenario, is one of the factors limiting performance on the perception task.

Finally, it is significant that the perception task that subjects were asked to perform referred to intended prominence rather than directly to SO-focus marking. The main purpose of the perception experiment was to establish whether the small but reliable

26 Stimuli from pairs 2 and 3 in Appendix A elicited 60.7% and 59.4% correct answers, respectively, while stimuli from pair 4 yielded 76.8% correct answers.
SO-focus marking we had previously observed is above the threshold of what can be humanly perceived. In that sense, participants were used as (very intelligent) minimal-difference detectors. That they performed significantly above chance means that the difference between the phonetic realization of SO 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 SO focus: it is logically compatible with our findings that hearers disregard the phonetic differences we have observed and rely entirely on context to identify SO foci.

With regard to the latter issue, it should be borne in mind that the point of this article 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 provide evidence that, contrary to claims in the literature, phonetic cues are available to infer the focus assignments of a sentence even in the case of SO focus. Our production results demonstrated that cues are reliably present, and the perception results show that such cues are in principle accessible to hearers. But since hearers did not identify SO foci as prominent with anything close to 100% accuracy under our experimental conditions, it must remain open to what extent SO-focus marking is used in online speech processing. Note here that relatively little is known about the processing of even first-mention focus marking, but see, for example, Watson et al. 2006 and Welby 2003.

5. General discussion. In this section we combine data from the production and perception experiments to study in more detail the important differences between SO-focus and ordinary-focus marking. At that point, we are in a position to consider why SO focus is realized differently from regular focus, and in what other environments similar effects should be expected.

Our data indicates that SO focus is made prominent by greater duration and energy (and perhaps other prosodic correlates not considered in this study) but is usually not pitch-accented. Below, we suggest that SO-focus marking is an example of phrasal prominence without pitch accent. This realization can be understood as resulting from the fact that items in SO focus occur after the nuclear accent of their phrase, the environment of postnuclear deaccenting. We show that this interpretation has implications for the general theory of phrasal prominence (stress and accent) and the marking of focus. We then relate this analysis to prior discussion in the semantics literature of example types closely related to the SO-focus paradigm (§5.3).

5.1. What is nonstandard about SO-focus marking? We conducted a comparison between SO-focus marking and the marking of ordinary, nonrepeated focus, based on two questions. First, does standard SO-focus marking lack a pitch accent, and, second, can SO focus optionally be marked using (ordinary) focus marking, complete with pitch accent?

As discussed in §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 striking feature of SO-focus marking is the minor role played by fundamental frequency. The production study revealed no significant main effect for F0 differences between words in SO focus and unfocused words. Auditorily based transcription of a subset of the data indicates that SO-focus expressions

27 See Jaeger 2004:§4 for further details.
are occasionally marked with high pitch accents (H*), but in most cases there is no perceptual indication of a pitch accent on the SO-focus expression. For cases without a pitch accent on the focus expression, two representative pitch tracks along with their transcriptions are shown in Figures 2 and 3. The focus expression associated with even is marked by a clear F0 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.28

Figure 4 is a representative example of a SO-focus expression marked by an H* pitch accent (though a comparatively subtle one).29

So the results of the production study show not only that SO focus is consistently marked by greater duration and energy, but also that it is occasionally marked by F0. To put it in the phonological terms we used to set up this issue: SO 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 SO focus from lack of focus is based on a few stimuli in which SO focus is marked by a pitch accent. To show that this is not the case, we now turn to a brief analysis of the acoustic properties of the stimuli used in the perception experiment.30

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, that is, those stimuli for which all subjects gave the same response. We would expect these stimuli to exhibit a particularly clear distinction between SO 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 F0 could have been used to distinguish the prominence of the words.

Based on both the logistic regression analysis and the sample of best stimuli, we conclude that the primary correlates of SO focus are duration and energy in the sense that these are the measures that most reliably distinguish SO focus from lack of focus in production, and that they are the properties that listeners attended to in the perceptual

28 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 for example Silverman 1987 or Beckman & Elam 1997. Note also that the pitch tracks in Figs. 3 and 4 contain instances where the pitch-tracking algorithm erroneously halved the fundamental frequency for a short period (cf. Ladefoged 2003:83–84). The actual F0 is more or less level.

29 In the utterance pictured in Fig. 4, doctor, the focus of even, is highly prominent auditorily. 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 SO focus Pete, but we have not studied this possibility further.

30 More details are given in Jaeger 2004:§4.4–4.5.
Figure 2. First sample with an SO focus lacking any pitch accent.
Figure 3. Second sample with an SO focus lacking any pitch accent.
Figure 4. Sample with a pitch-accented SO focus.
task. Fundamental frequency is only optionally used in marking SO focus and is not required for perception of the prominence of SO focus. Our conclusions are supported by new evidence from a study on SO foci in German. Féry and Ishihara (2005) found that in German, too, postnuclear SO foci are marked by longer durations but not by pitch accents. Of course, if a speaker chose to mark SO focus with a pitch accent, it is quite possible that the associated F0-related cues would be more significant than duration and energy in giving rise to a percept of prominence, but this is not the typical realization of SO focus, nor is a pitch accent necessary for listeners to detect the prominence associated with SO focus.

5.2. Why are SO foci marked without a pitch accent? 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 SO focus is marked by phrasal stress only. As discussed in §2.6, theories that distinguish stress and pitch accent generally identify greater duration and intensity as primary correlates of stress, while pitch accents are realized in terms of the fundamental-frequency contour. So the greater duration and energy of SO focus are attributed to phrasal stress, and the absence of F0 differences between SO-focus and nonfocused material is attributed to the absence of pitch accents in both conditions (with a few exceptions in the case of SO focus, as noted above).

Positing that SO focus is marked by stress but not accent helps explain why SO focus is realized differently from regular focus, by relating SO focus to the well-established phenomenon of postnuclear deaccenting. As mentioned in §2.6, a nuclear accent is the last pitch accent in a phrase, so placing a nuclear accent early in a phrase, as in 11, implies that all following words in that phrase must be unaccented. This is referred to as deaccenting, because in an example like 11 the word Sandy would be accented in a more neutral (or broad-focus) reading of the sentence, but is not if the preceding book receives a nuclear accent (e.g. because it is narrowly focused).

(11) Pat gave a [book\] to Sandy.

In all of the examples used in our experimental study, the SO focus is preceded by an expression that 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 SO focus could be regarded as a case of postnuclear deaccenting (Bartels 1997:12, Rooth 1992). That is, the SO 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. The predominance of pitch-accentless realizations should be related to the fact that the material following the subject (e.g. only named Sid in court today in Fig. 3) 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 two factors favor unaccented realizations (Ladd 1996:66, Terken & Hirschberg 1994), so speakers may be disinclined to place this material in its own phrase where it must receive a nuclear accent.

31 See Jaeger 2004:§5.3 for discussion.

32 Note, however, that given material can in principle be accented (Hedberg 2003). It is really the postnuclear occurrence of given material that leads to deaccenting and prosodic subordination (phrasing of the material following the nuclear accent into the same phrase as the nuclear accent (Wagner 2005)). Hence, the account suggested here predicts that pitch accent should more frequently be a valid option for prenuclear SO foci (i.e. repeated foci in a prenuclear position) than for postnuclear SO foci. This is what Féry and Ishihara (2005) found for German.
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 postnuclear SO-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 focused constituent that occurs earlier in the sentence. This conflict is generally resolved in favor of deaccenting, but the SO focus can still be marked by the strongest stress in the scope of the item with which it is associated.

5.3. **What environments yield pitch-accentless foci?** Our results demonstrate that in one specific type of environment focus is marked using duration and energy, 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 SO-focus paradigm. In this section we consider these variants in the light of what we have learned.33

In the experimental materials, the SO 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 that semantically one would expect to be accentted, but that occurs in postnuclear position.

Consider the classic example in 12.

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

Like SO-focus examples, this case involves an apparent mismatch between the prosodic focus of a VP following a focus-sensitive operator (*only*) and the semantic focus of the operator. The prosodic 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 SO 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 deaccented SO foci. We leave examination of this question for further work.34

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 sentence B of 13, the SO-focus expression *a book* precedes the nuclear accent on *many people*.


   B: True, but John only gave [a book]SOF to [many people]F.

33 In the discussion in this section we concentrate on the importance of the relative positioning of SO (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 (1997:270–71). Krifka considers variants of SO-focus examples in which the SO focus is pronominalized, and argues that such examples require pitch accents on the pronominal foci. If so, then textual identity of the SO 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 §5.2 above.

34 Roberts (1996) introduced a more complex type of example, in which a VP (*invited Lyn for dinner*) is the semantic focus, one element (*Lyn*) is contrastive, and the rest of the VP is old. The crucial claim is that the semantic focus is not marked, since this would tend to produce a pitch accent on *dinner*, yet there is not (or need not be) any such accent. Once again, we may ask whether in this case the word in question (*dinner*) is marked as prominent in some other way, for example, by lengthening and energy effects.
Dryer is certainly right to suggest that the most prominent accent in B’s utterance can be on *many people* rather than a *book*. But it is not clear to us whether this sentence can be felicitously uttered with no pitch accent on a *book*. As a matter of fact, Féry and Ishihara (2005) present evidence from similar examples in German suggesting that prenuclear SO foci (*a book* in 13) are marked by pitch accents. These pitch accents are on average ‘weaker’ (e.g. lower maximum F0 for an H* accent) than pitch accents on first-mention focus (e.g. *many people* in 13). Although the new data is not from English, it is suggestive of a reason why prenuclear SO foci might appear to be unaccented, that is, because the accenting is relatively weak. While we are unable to draw a firm conclusion about Dryer’s examples at this stage, the Féry and Ishihara results support our methodological view: questions like those Dryer raises are best studied in the laboratory.

6. 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. SO focus provides an extreme test case for this generalization since it is typically deaccented. Yet our results indicate that SO focus is still marked with the strongest stress in the scope of the operator.35

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, postnuclear 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 those of Selkirk (1984: 154ff.) and Hayes (1995:396), but is inconsistent with analyses such as that of Bolinger (1958) which attribute all phrase-level prominence to pitch accents, implying that there are no prominence distinctions in the postnuclear domain. Ladd (1996) has also proposed a number of arguments for the existence of postnuclear prominence distinctions, and also argues that focus is marked primarily by metrical stress. The only previous experimental investigation of postnuclear 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. *exploír* (verb) are distinguished in postnuclear position (although the difference is subtle).

As discussed in §5, studies of our production data indicate that pitch-accenting is optional on SO foci, although it occurs in only a minority of utterances. A more complicated question is whether the majority of SO 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 SO-focus marking are similar to ordinary focus marking, but we cannot conclude that pitch is the only suppressed feature.36 If we hypothesize that pitch suppression forces (most) SO 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 SO-focus marking. Vowel quality, for example, may be such a prosodic feature.

35 For a recent theoretical elaboration using our results, see Büring 2006.

36 Jaeger (2004:§5) presents evidence that intensity marking, unlike durational marking, is subject to a relatively strong reduction on SO-focus expressions.
As regards the semantics and pragmatics of focus, we have demonstrated that an argument from SO focus found in the literature is invalid. The original argument implied 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. We have shown, however, that there are no grounds for denying that these semantic foci are phonologically focus-marked: SO foci have acoustic properties that, in the absence of phonological focus marking, would be unexplained. So the original argument fails.

Does it follow that pragmatic theories of focus are incorrect? Not at all. Indeed, our data shows that though SO focus is consistently marked and is typically above the threshold of what is perceptible, it may be that hearers can retrieve that marking with far less than perfect accuracy. Pragmatic theories could then describe an important source of information available in the processing of SO focus (as suggested to us by Jaye Padgett, p.c.). But such a processing claim would be independent of whether, as required by syntactic and semantic theories, association with focus is grammaticized, a claim about competence, not performance. It is now clear that standard arguments from SO focus cannot settle this competence issue.

We began this article with Ladd's pessimistic observation 'that proposals about intonational meaning are not a reliable source of evidence on intonational phonology'. We have shown that work on intonational (or, perhaps better, PROSODIC) meaning can provide a useful source of evidence for suprasegmental phonology, but we have also sought to demonstrate that study of suprasegmental phonology is a necessary component of work on intonational meaning. In the process, we have revealed SO focus to be a phenomenon of considerable phonetic and phonological subtlety.

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.

Appendix B: Further details on the production experiment

B.1. Exclusion criteria. As mentioned in §3.2, 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, stuttering, or other significant reading effects were removed. Whenever a sentence was removed from the analysis, the matching sentence (with the inverted SO-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. We then averaged over the two repetitions of each stimulus (or used only one if the other had been excluded), yielding 556 targets of a possible 560 (as calculated by taking two targets from each of seven dialogue items, each under two different focus structures, multiplied by twenty speakers, thus $2 \times 7 \times 2 \times 20 = 560$). These 556 targets were available for the duration, energy, and intensity analyses, corresponding to a data loss of 0.7%, as cited in the main text. The pitch analysis was performed on 544 targets (2.8% data loss). This was due to further exclusion of items for which either (i) the pitch tracking failed to find a pitch value, (ii) the observed pitch value was likely to be incorrect (i.e. the maximum pitch on the target deviated more than two standard deviations from the mean pitch of the whole utterance), (iii) the pitch on the target had a range of 0, or (iv) an otherwise usable item was removed due to problems (i) to (iii) 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 root-mean-square intensity).

REFERENCES


Beaver, David, and Brady Clark. 2003. Always and only: Why not all focus sensitive operators are alike. Natural Language Semantics 11.4.323–62.


OSTENDORF, MARI; PATTI J. PRICE; and STEFANIE SHATTUCK-HUFNAGEL. 1995. The Boston University radio news corpus. (Boston University technical report ECS-95-001.) Boston: Boston University.


SILVERMAN, KIM; MARY BECKMAN; JOHN PITRELLI; MARI OSTENDORF; COLIN WIGHTMAN; PATRI PRICE; JANET PIERREHUMBERT; and JULIA HIRSCHBERG. 1992. ToBI: A standard for labelling English prosody. Proceedings of the International Conference on Spoken Language Processing, vol. 2, 867–70. Banff: ICSLP.


TANCREDI, CHRISTOPHER D. 1990. Not only even, but even only. Cambridge, MA: MIT, MS.


