Second Occurrence Focus is Prosodically Marked

Results of a production experiment

David Beaver, Brady Clark, Edward Flemming and Maria Wolters
Stanford University and Rhetorical Systems

Contact.   David Beaver (`dib@stanford.edu`)
Department of Linguistics
Stanford University
Stanford, CA 94305-2150
USA

Abstract.   A second occurrence focus is an expression which is in the scope of a focus sensitive operator, is the semantic focus of that operator, and which is a repeat of an earlier focused occurrence. Second occurrence foci are intonationally distinct from the original occurrence of the material. Indeed, second occurrence foci are often claimed to lack any intonational marking, e.g. pitch accent. This apparent dissociation of semantic and intonational focus is commonly used as an argument against certain theories of focus; e.g., alternative semantics (Rooth, 1985) and structured meaning semantics (Jacobs, 1983; Krifka, 1992; von Stechow, 1989). Here we report on a production experiment designed to test whether second occurrence foci are prosodically marked. We find that while there is no significant pitch accent on second occurrence foci, there are other prosodic effects. In particular, we observe that second occurrence focus is marked by increased duration and intensity. This result is of significance to semanticists interested in the interpretation of focus and to intonational phonologists interested in the acoustic realization of focus.

Keywords: focus, semantics, prosody

1. Introduction

In this paper we present experimental data concerning the acoustic realization of focus, data which is relevant for any model of the semantics/phonology interface. We concentrate on a phenomenon that has been studied in the semantic literature and which involves expressions which are *focus sensitive*. An expression is focus sensitive if its interpretation is dependent on the placement of focus, keeping in mind that the linguistic realization of focus (e.g., by prosody, syntactic position or morphology) varies cross-linguistically. In English, focus is typically marked by a pitch rise on a stressed syllable.[1]

The basic generalization about focus sensitive operators like "only" is that they are truth-conditionally sensitive to focus in certain contexts. Consider the examples in (1). In a situation where Jan gives Bill and Malachi a book each, and gives nobody anything else, (1a) is true while (1b) is false.

(1) a. Jan only gave Bill [a book]$_F$.

  b. Jan only gave [Bill]$_F$ a book.

While all theories of focus sensitivity agree that expressions like "only" and focus interact in linguistic contexts like that exemplified by (1), there is controversy about 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 is the interaction illustrated by (1) optional in certain contexts? This question provides the main distinction between different contemporary theories of focus. A number of examples have been cited in the literature involving an apparent dislocation between the associated semantic focus of focus sensitive operators like "only" and prosodic prominence. Most examples fall under the rubric of "second occurrence focus", where a repeated focused item apparently lacks pitch accent. Early references on second occurrence focus include Gussenhoven (1984), Hajičová (1973), Hajičová (1984), Koktová (1986), Koktová (1987) and Taglicht (1984). The example in (2), from Partee (1999:215), illustrates second occurrence focus. The two sentences are to be read as a dialogue between speakers A and B.

(2) A: Everyone already knew that Mary only eats [vegetables]$_F$.

  B: If even Paul knew that Mary only eats [vegetables]$_{SOF}$, then he should have suggested a different [restaurant]$_F$.

Partee (1999) summarizes the problem succinctly. If "only" is a focus sensitive operator (i.e., needs an intonationally prominent element in its

scope) then the two occurrences of "only eats vegetables" in (2) should have the same analysis. However, if there is no phonological reflex of focus in the second occurrence of "vegetables" then this leads to the notion of "phonologically invisible focus". The notion of inaudible foci "at best would force the recognition of a multiplicity of different notions of 'focus' and at worst might lead to a fundamentally incoherent notion of focus" (Partee 1999:215–216).

We will present the results of a production experiment in which we examine the acoustic correlates of second occurrence focus in the scope of focus sensitive operators "only" and "always". Our results confirm that there are acoustic correlates, small but significant effects involving duration and intensity of the focused item. These conclusions show that the phenomenon of second occurrence focus does not provide an argument that current theories of focus are "fundamentally incoherent", although the data does suggest some need for revision of current theories of the interface between sound and meaning.

The paper is organized as follows. First we consider the theoretical background in more detail, discussing the realization of focus in English, the main accounts of focus sensitivity and some of the conclusions reached in the literature about second occurrence focus. Then we introduce the experiment we carried out and its associated results, before concluding with a brief discussion of the significance of what we have shown.

2. The Marking of Focus

There is a substantial literature on the phonetic correlates of focus marking, although little of it pertains specifically to focus associated with focus sensitive particles (most studies concern corrections, contrast, or answers to questions). While it is not safe to assume that all varieties of focus are marked in exactly the same way (Bartels and Kingston, 1984), it is clear that all types of focus are normally marked by aligning a pitch accent with the primary stressed syllable of the focused item. A pitch accent is "a local feature of a pitch contour - usually but not invariably a pitch change, and often involving a local maximum or minimum" (Ladd 1996:45–46) — cf. Bolinger (1958) and Pierrehumbert (1980). In simple declarative sentences, as in (1a), the pitch accent will typically involve a pitch peak, but in yes-no questions, focus is often marked by low pitch.

More specifically, focus is usually marked by a nuclear pitch accent, i.e. the last pitch accent in a phonological phrase — see Cohan (2000) and Ladd (1996:225ff). The nuclear accent is perceived as more

prominent than preceding accents in the same phrase. For example, in an utterance of (1b) there could be a pitch accent on "Jan", but this first accent is perceived as less prominent than the nuclear accent that marks the focus, "Bill".

This observation corresponds to Jackendoff's (1972) generalization that focus is marked by the strongest stress in the sentence. Jackendoff refers primarily to phrasal stress in the sense of Chomsky and Halle (1968), but it is clear that these stresses correspond to nuclear accents, although the precise relationship between pitch accents and stress is a matter of continuing dispute — see Ladd (1996:45ff) for discussion.

As the name suggests, a pitch accent is primarily a pitch event, so its main correlates are to be found in the fundamental frequency (f0) contour — e.g. a local maximum in f0. However, it is well established that accented words can differ from unaccented words in other respects such as increased duration and intensity (Sluijter and van Heuven, 1996; Turk and White, 1999) and more differentiated quality of stressed vowels (de Jong, 1995; Harrington et al., 2000).

We now turn from the general issue of how focus is marked to the more specific issue of how second occurrence focus is marked. While it is generally accepted that second occurrence focus is not marked by pitch movement, it remains possible that second occurrence focus is marked as prosodically prominent by other factors, such as duration, intensity and vowel quality. For example, Ladd (1996:226f) argues that focus can be marked independently of pitch accent in this way. In the next section we discuss in more detail why this issue is of importance to semanticists studying focus sensitivity, before proceeding to the experimental investigation of the marking of second occurrence focus.

3. Theories of Focus Sensitivity

Following Rooth (1992), theories based on lexical stipulations we will term *weak*, and those that explain focus sensitivity without recourse to such stipulation we will term *strong*, while theories which allow for very limited lexical stipulation we term *intermediate*. We will now describe each of these approaches.

A weak theory of focus interpretation introduces semantic objects, focus semantic values, which are then manipulated by construction-specific rules (Rooth 1992:107). Both structured meaning semantics (Jacobs, 1983; Krifka, 1992; von Stechow, 1989) and alternative semantics (Rooth, 1985) are weak theories of focus interpretation. What unites weak theories of focus interpretation like structured meaning semantics and alternative semantics is their usage of specific rules for

individual focus sensitive constructions, like association with focus by "only". Both accounts introduce special semantic objects, focus semantic values, and it is the responsibility of construction-specific rules to make use of this information. Neither approach characterizes a notion of **possible focus sensitive construction**:

> An analysis which fails to address this requirement might be saying a lot about specific constructions, but it says nothing about focus in general. By omission, it maintains that there is no uniform semantic or pragmatic phenomenon of focus. (Rooth 1996a:278)

In a strong theory of focus interpretation there is no construction-specific reference to focus. According to a strong theory of focus interpretation, pragmatic factors optionally link the interpretation of "only" to its associated focus. Rooth (1992) and von Fintel (1994, 1995) explore similar pragmatic theories of focus interpretation. Other pragmatic approaches to focus sensitivity can be found in Geurts and van der Sandt (1997), Roberts (1995), Roberts (1996), and Schwarzschild (1997).

Let us now exemplify a relatively weak theory of focus cast, like Rooth's analyses, in terms of what we will call *focal alternatives*. If a phrase ϕ contains a focused expression F, then the focal alternatives are obtained by considering what the meaning of ϕ would have been had F been different. Thus the focal alternatives to "gave [Bill]$_F$ a book" are properties expressible as "gave u a book", for some u. We will notate the focal alternative that has u in place of the focus ϕ_u.

We can now state what the meaning of "only" might be in a weak theory, restricting ourselves for the moment to cases in which the focus is an NP referring to an individual. In (3), FOC is the meaning of the focused expression, SUB is the meaning of the subject NP, ϕ the meaning of the VP to which "only" is adjoined:

(3)   "NP only VP" $\mapsto \forall u[\phi_u(\text{SUB}) \rightarrow u = \text{FOC}]$

For the case of (1b), this will come down to the requirement that every individual to whom Jan gave a book is Bill. A theory of focus which included a semantic rule like this for "only" is weak because the rule makes direct reference to the semantic value of the focused expression: it is not clear how such a theory might be constrained to prevent semantic rules performing arbitrary manipulations on the semantic value of the focus.

Intermediate theories of focus, like weak theories, allow for lexical stipulation that effectively forces "only" to associate with focus. However, in an intermediate theory the type of semantic rule that can be stated is very limited. Rooth postulates an operator $\sim$ which manages

the interface between focus sensitive constructions and the focus feature $[\ ]_F$. The configuration $\phi \sim Y$ induces a constraint on a variable Y, which is assumed to lack any surface expression. This constraint forces Y to be a set of focal alternatives to ϕ: in the case at hand, Y will be a set of properties expressible as "gave u a book". Let us assume for simplicity that Y includes every such property. Lexical semantic rules are then not allowed to make reference to focus directly, although they can make reference to the output of the $\sim$ operator.

Rooth would assume a representation of (1b) like that in (4):

(4)
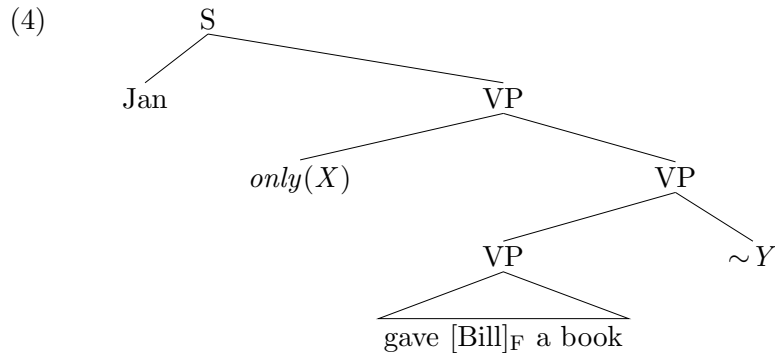
An intermediate theory of focus might allow a configurational stipulation in the lexical entry for "only" forcing the variables X and Y to be identical. In addition to this configurational requirement, we now need to give a variant of the semantics that makes reference to X, but no direct mention of the interpretation of the focus. This requires a modification of the semantics above:[2]

(5)   "NP only VP" $\mapsto \forall P \in X[P(\text{SUB}) \to P = \phi]$

In the intermediate theory we are considering, $X = Y =$ the set of properties expressible as "gave u a book". So (1b) is interpreted to mean: every property of the form "gave u a book" which holds of Jan is the property "gave Bill a book". Thus, once again, every individual to whom Jan gave a book is identical to Bill, which is the desired result.

Given that construction-specific rules are not available in the strong theory of focus interpretation, pragmatic factors alone link the interpretation of "only" to focus. Thus we might, for example, assume the same semantics for "only" but no configurational restriction. Then X and Y in (4) would only be identical if that happened to result from some process of pragmatic resolution. So in classic examples like (1b), association of the focus sensitive expression with focus is predicted to be optional in certain contexts. In such a theory a reading like that

for (1b) might potentially be available for (1a), although that would depend on mechanisms independent of focus making available the set of alternatives of the form "gave u a book", and on pragmatic factors causing a preference for resolution of X to this set rather than to the value produced by an application of the $\sim$ operator.

This takes us to an empirical argument which can potentially distinguish the three theories of focus interpretation, weak, strong and intermediate. While weak and intermediate theories of focus interpretation stipulate a focused constituent in the scope of an operator like "only", a heavily degrammaticized strong theory of focus interpretation like that suggested by Rooth (1992) predicts that we should find examples where context alone provides the set of alternative properties required by the semantics of VP "only". So a strong theory predicts that there should be cases where there is no focus-marked constituent in the syntactic scope of "only" (Partee 1999:217).

A number of theorists have argued that second occurrence focus examples are straightforward counterexamples to weak theories of focus interpretation like alternative semantics and structured meaning semantics, arguing in favor of a strong account of association with focus phenomena; e.g. Dryer (1994), Roberts (1996), Schwarzschild (1997), Vallduví (1990) and Williams (1997). The arguments in the literature from second occurrence focus have basically the same structure,[3]

(i)   Weak theories require focus marked elements in the scope of focus sensitive expressions.

(ii) In the case of second occurrence focus there is no such element.

(iii) Therefore focus sensitivity is purely pragmatic.

This paper hinges on step (ii) of this argument, which would be undermined if it turned out that second occurrence focus has prosodic correlates. Recently, suggestive evidence has been presented indicating just that. Rooth (1996b) recorded himself uttering SOF dialogues like (6).

(6)   a.   A: Do you want Sue to only [name]$_F$ Manny today?
            B: No. I only want [Eva]$_F$ to only [name]$_{SOF}$ Manny today.
      b.   A: Do you want Sue to only name [Manny]$_F$ today?
            B: No. I only want [Eva]$_F$ to only name [Manny]$_{SOF}$ today.

Rooth compared second occurrence focus NPs (e.g., "Manny" in (6)b) to controls where the NPs were not in focus at all (e.g., "Manny" in (6a)). Rooth found that the pitch track was flat in second occurrence focus position for examples like (6). However, focus with "only" in second occurrence focus position is "clearly" detectable. He shows second occurrence focus correlates with phonetic parameters other than pitch, specifically with vowel duration.

In related work, Bartels (1997) studied the possibility of non-pitch correlates of prominence (relative syllable lengthening and amplitude) on second occurrence expressions. While Bartels demonstrated that a systematic experimental approach to focus phenomena was possible, her conclusions concerned the different realization of second occurrence focus from regular focus. She did not add controls in which the test words were not in focus at all (i.e. neither focus or second occurence focus) (Bartels 1997:24). Thus her experiments do not determine whether second occurrence focus is marked prosodically, but merely establish that if it is marked, then it is marked differently from ordinary focus.

Are there, as Rooth suggested, prosodic correlates of the focus feature in second occurrence focus position? If the answer is affirmative, the argument from second occurrence focus unravels. In the remainder of this paper we will report on work that resolves the issue in a systematic study. We will not only confirm that duration effects mark second occurrence focus, but also provide evidence for another prosodic correlate.

4. Method

Our goals were to examine three hypotheses, that second occurrence foci are marked by pitch movement, that they are marked by increased duration, and that they are marked by increased intensity.

Prior literature discussed above, although not based on systematic experimentation, leads to a strong expectation that the first hypothesis would be false, a weak expectation that the second would be true, and no expectation about the third hypothesis. Note that if any of the three hypotheses are true, then the argument from second occurrence focus does not hold.

To test the hypotheses we ran a production experiment in which subjects read preprepared written materials. The materials (apart from dummy discourses we also included) take the form of minimal pairs of three sentence discourses designed to reveal how second occurrence foci differ acoustically from non-focal material.

An example of a minimal pair of discourses used as stimuli is given in (7) and (8). These three-sentence discourses are designed to probe second occurrence focus effects in the scope of a focus sensitive operator operator in third sentence. In this case, the relevant operator is "only", while in other test pairs the operator was "always". Note that another focus sensitive operator "even" is also present. The presence of this operator in the subject NP of the third sentence leads the sentence to be uttered with a nuclear pitch accent in the initial NP (here on "state prosecutor"), while the fact that the post-nuclear material is repeated seems to lead to an unaccented realization (see section 6 for further discussion).

In these particular examples, the relevant potential foci are "Sid" and "court". For all the stimuli we used, the material following the second focus sensitive operator in the third sentence does not differ between the two members of the discourse pair. Thus the phonological context for the segment of text containing the the two potential foci should not differ between the two elements of the minimal pair. We can therefore attribute acoustic differences between the potential foci in the two pairs to second occurrence focus effects. In all the minimal pairs, it is variations in the content of the first two sentences of the discourse which are used to set up context and produce the effects to be measured.

(7) 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.

(8) 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 (7b), "Sid" is the focus of the focus sensitive particle "only", whereas in (8b), "court" is in focus, and "Sid" is non-focal. Sentences (7c) and (8)c contain a textually identical VP ("only named Sid in court today") to the (b) sentences, but this VP follows an earlier phrase "even the state prosecutor" containing what we take to be the nuclear focus of the sentence (presumably on "state prosecutor"). As a result of this earlier nuclear focus, the VP in the (c) sentences is expected to involve relatively little pitch movement. According to the research discussed in the previous section, we now have a classic case in which the second occurrence of the focus in the VP in each of the (c) sentences might be expected to receive no intonational prominence.

Crucially, the focus in each of the (b) sentences was different, so the textually identical VPs in the (c) sentences should differ only in so far as they contain different expressions which are second occurrence foci. In (7c), "Sid" is the second occurrence of a focus, and "court" is non-focal, whereas in (8c), "Sid" is non-focal, and "court" is the second occurrence of a focus. Thus the hypotheses above can be operationalized in terms of the acoustic differences between "Sid" in each of the two final sentences, and the acoustic differences between "court" in each of the two final sentences.

If an expression is the second occurrence of a focus in the final sentence of a discourse, we will say that the expression is in the *re-focused* condition, and otherwise we will say it is in the *non-focal* condition. Our general hypothesis can then be formulated as follows:

> Potential foci in the re-focused condition differ in acoustic realization from the same potential foci in the non-focal condition.

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

We examined four parameters in testing for differences in acoustic realization: peak fundamental frequency (f_0), f_0-range (the difference between the minimum and maximum f_0 values during the word), duration, and intensity. These parameters were selected based on previous research concerning the acoustic correlates of focus, reviewed in Section 2. Focus in declarative sentences of the kind used here would usually be marked by some kind of high pitch accent, which would result in higher

peak f_0, and might also result in a significant pitch change on the focus, which would be reflected by an increased f_0-range. Increased duration and intensity are well-established non-tonal correlates of focus.

We used a total of 14 discourse stimuli, made up of 7 minimal pairs like those above. In each case, the repeated focus sensitive particle was "only" or "always", while the focus sensitive particle "even" was used to induce the nuclear focus in the subject of the final sentence of the discourse. In all cases the potential second occurrence foci in the final sentence are not sentence final, this being contrived by the use of an additional adverb — "today" in the pair above. The reason for adding the 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.

Following standard procedure, the stimuli were arranged in a pseudo-random order and intermingled with unrelated filler stimuli, also discourses[4]. To control for recency effects which might occur when a subject encounters both members of a minimal pair of discourses, we ensured that paired discourses were always separated by at least four other discourses, either other real stimuli, or dummy items. We constructed four different sets of such randomized, interleaved stimuli, and each participant in the study received one stimulus set. There were a total of 44 discourses in each stimulus set, of which 28 were stimuli for the experiment. Thus the stimulus set included two repetitions of each discourse. The participants in the study were 21 undergraduates at Stanford, all US born native speakers of English, none with any training in linguistics. Since each stimulus discourse incorporated both conditions, the total number of probes of each condition was $28 \times 21 = 588$.

After recording all the speakers, the word boundaries were marked on the expressions taken to be potential foci in the second and third sentences, although only the marking on the third is relevant to the hypotheses being examined. The word boundaries were hand-coded by the authors, based primarily on examination of spectrograms. For each word, boundaries were marked at acoustic landmarks (such as stop closures) near the word onset and offset. The landmarks were selected (i) to be consistently identifiable across utterances of a word, and (ii) to include the vowel of the test word, since this is the expected locus of pitch and duration effects. All analyses are based on paired comparisons of the same words in non-focal and re-focused conditions, so consistency across conditions is of the greatest importance.

Having hand-coded all the relevant word-boundaries, the maximum f_0, f_0-range, R.M.S. intensity and duration were then extracted automatically. All f_0 values were converted to a logarithmic scale, and then normalized with respect to the mean and standard deviation for each

speaker. f_0-range was calculated as the difference between the maximum and minimum normalized f_0 values during the word. Discourses were excluded from analysis if they involved significant disfluency, or if the f_0 extraction algorithm failed, or yielded implausible values.

Measurements from remaining minimal pairs of discourses were then subjected to two-way paired t-tests for each acoustic parameter.

5. Results

As expected on the basis of prior literature, we found no significant differences in fundamental frequency measures between the non-focussed and re-focussed conditions. There is thus no reason to believe that pitch is a marker of second occurrence focus.

However, we obtained strong evidence that second occurrence focus is marked by other factors: there is a small but significant lengthening of second occurrence foci, i.e. words in the re-focussed condition, above that found for the same expression in the same sentential context but lacking focus. Across all trials, this lengthening averaged 6ms. There is also a statistically significant increase in average intensity in the re-focussed condition.

The main results are summarized in table 6. The top section of the table shows the results across all discourses and all speakers. Results are given for duration, maximum fundamental frequency ("Max f_0"), pitch range ("Range f_0") and average intensity of the potential focus relative to the average intensity across the sentence ("Rel. r.m.s. intensity").[5] For each of these, we present the average value for the potential focus when in non-focal condition ("μ Non-focal"), the average value for the potential focus when in re-focussed condition ("μ Re-focussed"), the average amount by which the measurement in non-focal condition is larger than the measurement in re-focussed condition ("μ Increase"), and the statistical significance as a probability derived from a two-tailed paired t-test ("p-value"). Cases in which the difference in a measurement is significant at the 5% level are marked using a box around the p-value, and cases where the increase of a measurement is significant at the 1% level are marked with a double box.[6]

At the time we initiated this study, and prepared the stimuli, we had no reason to believe that the focus sensitive particles "only" and "always" would yield different effects. As it happens, for four of the discourse pairs the focus sensitive particle used was "only", and for three the particle was "always". Notably, independent evidence is presented by Beaver and Clark (2002) that the focus sensitivity of "only" and "always" results from distinct mechanisms. It is therefore of interest to break down the results according to the focus sensitive particle involved, which we have done in the lower two sections of table 6.

What these results show is that the effects we observed across the full set of discourses primarily reflect marking of second occurrence focus in the scope of "only". For the four discourse pairs involving "only" we derived highly significant duration and intensity effects, but no significant maximal pitch or pitch range effects. However, for the three pairs of discourses involving "always", none of the dimensions

Table I. Summary of results

	μ Non-focal	μ Re-focussed	μ Increase	p-value
All pairs ($N = 456$):				
Duration *(s)*	0.265	0.271	0.006	0.0037
Max f_0 *(Hz)*	174	176	2	0.27
Range f_0 *(Hz)*	60.3	64.3	4.00	0.14
Rel. r.m.s. intensity *(dB)*	2.64	2.68	0.036	0.029
Pairs with "only" ($N = 262$):				
Duration *(s.)*	0.229	0.236	0.007	0.008
Max f_0 *(Hz)*	178	180	2	0.47
Range f_0 *(Hz)*	62.6	67.7	5.04	0.24
Rel. r.m.s. intensity *(dB)*	2.65	2.70	0.054	0.018
Pairs with "always" ($N = 193$):				
Duration *(s.)*	0.315	0.318	0.004	0.19
Max f_0 *(Hz)*	169	171	1.5	0.36
Range f_0 *(Hz)*	57.2	59.8	2.58	0.37
Rel. r.m.s. intensity *(dB)*	2.64	2.65	0.012	0.58

we investigated is a good predictor of whether a potential focus is in non-focal or re-focussed condition.

We cannot put this difference between "only" and "always" down to the reduced number of discourses involving "always": while the number of discourses involving "always" was only slightly lower than the number involving "only", the significance values for "only" are dramatically higher. Performing the same analysis on only half the discourses involving "only" gives a smaller number of discourses (and hence lower degrees of freedom in the *t*-test), yet still produces qualitatively similar increases and significance values as are found with the complete data set.

6. Discussion

The statistical significance of the duration and intensity results provide strong evidence that second-occurrence foci are focus marked at some level of linguistic representation, since we have been careful to ensure that there is no difference between the re-focussed and non-focussed condition except focus. This corroborates Rooth's informal findings regarding duration effects, and is sufficient to undermine any argument for a pragmatic/strong theory of focus that relies crucially on second occurrence foci not being formally marked. As far as second occurrence focus phenomena are concerned, semantic/weak theories of focus are quite defensible.

Accented words are known to differ from unaccented words in duration and intensity, as discussed in section 2. The results we have presented show that non-tonal features may be manipulated independently of pitch accent to mark prominence, and provide the first systematic experimental evidence, that under some circumstances, these features are sufficient to mark focus.

In the remainder of this section we will first consider the significance of the results we have obtained for perception, and second attempt to form some general hypotheses about the phenomenon of non-tonal prominence as a marker of focus.

6.1. Perceptual Significance of the Results

Although the differences in mean duration and intensity are statistically significant, they are small — differences of this magnitude are probably not perceptible even under ideal listening conditions. Of course we have only measured a few acoustic parameters, and second occurrence focus may be marked more clearly in terms of other parameters, e.g. duration in relation to surrounding words. Listening tests would be the most direct way of establishing the perceptibility of the focus marking identified here. It is also important to bear in mind that these are mean differences — it may be that subjects were inconsistent in marking second occurrence focus, so averaging over clearly marked and unmarked instances yields a small mean difference. For example, subjects might have imperfectly understood some of the discourses, resulting in failure to note the second occurrence focus. There are some patterns in the data that are consistent with this possibility: each subject read each discourse twice, and the duration difference between non-focal and re-focused words was significantly larger in the second repetition. This could indicate that subjects understood the discourses better as

the experiment progressed (or when encountering them for the second time).

Alternatively, speakers might not provide clear marking of second occurrence focus because the location of focus is identifiable from context: if a linguist can tell from context which expressions are second-occurrence foci, then presumably hearers can too. These considerations suggest that the best test of the perceptual significance of the type of focus marking observed here would be to take utterances with clear duration differences on the re-focused item (natural or re-synthesized) and see if listeners can use this difference to identify the location of focus. That is, increased duration might be a cue to focus (in the absence of pitch accent), even if speakers do not consistently produce this cue in the contexts examined here.

6.2. The marking of focus with non-tonal prominence

Our results demonstrate that in one specific type of environment focus is marked using duration and intensity. It is natural to ask what features of the environment we considered are responsible for the observed effects, and, more generally, under what circumstances we might expect such phenomena to recur. In the remainder of this section we will consider a number of relevant parameters which should be investigated in further work.

The class of second occurrence focus sentences we looked at are those in which the second occurrence focus expression is the semantic focus of either "only" or "always", the nuclear focus of the sentence occurs to the left of the second occurrence focus expression, and the second occurrence focus expression is textually identical to the first occurrence. This leads us to isolate three features of relevance to the realization of a focus:

1. Which focus sensitive particle is present,

2. Phonological context of the focus,

3. Discourse status of the focus.

Regarding the focus sensitive particle, it was mentioned above that Beaver and Clark (2002) provide evidence that focus sensitivity of "only" results from a different mechanism than focus sensitivity of "always". They argue that focus sensitivity of "only" results from a weak/semantic mechanism, while focus sensitivity of "always" results from a strong/pragmatic mechanism. On the basis of their arguments, it is predicted that there should be a prosodic reflex of second occurrence

focus in the scope of "only". On the other hand, their account of "always" leads us to expect no requirement on the speaker to prosodically mark an element in its scope as being focal. This is fully in accord with the results obtained in the current experiment: speakers do mark second occurrence foci in the scope of "only", but there is no evidence that they mark second occurrence foci in the scope of "always".

Regarding phonological context, the main parameter to be considered is the position of the nuclear accent relative to the focus expression under consideration, and, in particular, the question of which comes first.

As mentioned in section 2, a nuclear accent is the last accent in a phonological phrase, so placing a nuclear accent early in a phrase, as in (9), implies that all following words must be unaccented. This is referred to as deaccenting because "Sandy" would be pitch accented in a more neutral (or *broad focus*) reading of the sentence.

(9)   Pat gave a [book]$_F$ to Sandy.

In the experimental materials, the second occurrence focus is preceded by a nuclear pitch accent, marking the focus associated with "even", so the absence of a pitch accent could be regarded as an instance of post-nuclear deaccenting, as noted by Bartels (1997:12) and Rooth (1992). However, it is possible to divide a sentence of this kind into two phonological phrases, each containing a nuclear pitch accent (10). A similar pattern of phrasing and accentuation seems acceptable in (11), where a variant of (2), Partee's 1999 example of second occurrence focus, is placed in a different context.

(10)   [Pat gave a [book]$_F$][to [Sandy]$_F$] (and a tie to Sam).

(11)   Why did they book a steak house for Mary's birthday dinner? [Even [Paul]$_F$ should have known][that Mary only eats [vegetables]$_F$].

In the analysis of the experiment reported here, we have not explicitly factored in the possible effects of phrasing. Informally, however, we note that in pitch tracks from the experiment, we see little evidence of phrase breaking. Thus the second occurrence focus does occur in the post-nuclear region of a single intonational phrase, and post-nuclear deaccenting is a possible explanation why there was no significant pitch movement on the second occurrence focus expression.

In looking through the literature on apparent dissociation between focus and pitch accenting, we observe that the great majority of examples involve an expression which semantically one would expect to be accented, but which occurs in post-nuclear position.

Consider the following classic example:

(12)   People who [grow]$_F$ rice generally only [eat]$_F$ rice. (Rooth 1992:109)

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

A more complex type of example, due to Roberts (1996), is given in (14). Roberts imagines a conversation where we are discussing a number of things we would rather Mary did not do, listed in (13):

(13)   a.   inviting Lyn for dinner

        b.   inviting Bill for dinner

        c.   staining the tablecloth at lunch

        d.   smoking before dinner

In this context, Roberts' considers the following dialogue:

(14)   A: Mary wasn't so bad after all. Of all the things we were afraid she might do, she only [invited Bill for dinner]$_F$.

        B: You got the person wrong. She only invited [Lyn]$_F$ for dinner. But it's true that she did only one of those terrible things she could have done.

On the desired reading of the penultimate sentence of Bs reply, the domain of "only" is the set of properties in (13). However, given that there is focal prominence on "Lyn", standard theories of focus sensitivity, like Rooth's (1985) weak theory, would lead us to expect that all the alternatives under consideration have the form "invited X for dinner".

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

be marked by placing a pitch accent on the final argument NP, here "dinner".

Having identified the accent pattern we would normally expect for such a case of VP focus, we can now say in exactly what sense the penultimate sentence in example (14) shows an unexpected pattern: "dinner" does not receive a pitch accent. Having observed this, it becomes clear that the lack of accent may be correlated with the fact that "dinner" occurs in post-nuclear position, where the nuclear accent is the contrastively stressed "Lyn". (Presumably the fact that "Lyn" rather than "dinner" is chosen by the speaker to be the nuclear accent relates to the function of the speech act, and is independent of which word occurs first in the linear order of the sentence.)

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

Are there any cases in which a semantic focus lacks accent, but does not follow the nuclear accent? Dryer (1994) presents just such an example. In the final sentence of (15), the second occurrence focus expression "a book" precedes the nuclear accent on "many people".

(15)   A: I hear that John only gave [a book]$_F$ to Mary.

        B: True, but John only gave [a book]$_{SOF}$ to [many people]$_F$.

On the basis of the experiment we have reported, and our own informal observations, we postulate that in (15) the second occurrence focus would normally be intonationally marked, either with pitch movement, or with lengthening, or both. However, as yet we do not have any experimental data in direct support of this claim.

What makes this example particularly important is that the lack of pitch movement on the semantic focus of "only" could not be a result of post-nuclear deaccenting. Thus if the lack of pitch movement is substantiated experimentally in examples with this basic form, then we would have to look at other explanations why the focus is not pitch accented.

We now turn to a consideration of discourse status, and a feature that is common to all the data we have considered so far, including Dryer's example: repetition. (Ladd, 1980) showed that there is a strong tendency not to accent repeated material in English, as exemplified in

(16). If (16b) (Ladd, 1996) were uttered in isolation, the word "German" would receive a pitch accent and "read" would not ("I don't read [German]$_F$"), but where "German" has already been mentioned, it is not accented.

(16)   a.   I've found you a book in German.

      b.   But I don't [read]$_F$ German.

      b'. ? But I don't read [German]$_F$.

Is it the case that non-tonal marking of focus only occurs when the focus is repeated material? We know of one type of example in the literature which bears on this question. In these examples, due to Krifka (1997:270–271), expressions in a reprise are realized with a different form to the original, through pronominalization.

Krifka argues that if the second occurrence expression is an exact copy of the corresponding first occurrence expression, then it should be completely destressed; see (17a) and (17b). However, if in the reprise the stressed part of the first occurrence is replaced by an unstressed pronoun, we get a less acceptable sentence; see (17c). Krifka observes that the sentence is fully acceptable with a secondary stress on the pronoun; see (17d).

(17)   A: Eva only gave xerox copies to the [POOR]$_F$ students.

      B: (a) No, [PETR]$_F$ only gave xerox copies to the [poor]$_{SOF}$ students.

          (b) ? No, [PETR]$_F$ only gave xerox copies to the [POOR]$_{SOF}$ students.

          (c) ?? No, [PETR]$_F$ only gave xerox copies to those students.

          (d) No, [PETR]$_F$ only gave xerox copies to THOSE students.

Krifka (1997:271) also presents examples which suggest that when some element of the first occurrence is changed, complete destressing of the second occurrence is dispreferred.

(18)   A: Eva only gave xerox copies to the [POOR]$_F$ students.

      B: (a) ?? No, [PETR]$_F$ only gave them to the poor students.

          (b) No, [PETR]$_F$ only gave them to the POOR students.

Again, we leave examination of these variants on the second occurrence focus regime to further work.

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.

References

Bartels, C.: 1997, 'Acoustic Correlates of 'Second Occurrence' Focus: Towards an Experimental Investigation'. In: H. Kamp and B. Partee (eds.): *Context-dependence in the Analysis of Linguistic Meaning (Proceedings of the workshops in Prague and Bad Teinach)*. University of Stuttgart, pp. 11–30.

Bartels, C. and J. Kingston: 1984, 'Salient pitch cues in the perception of contrastive focus'. In: *Focus and Natural Language Processing, vol. 1: Intonation and Syntax*. Heidelberg, pp. 11–28.

Beaver, D. and B. Clark: 2002, '"Always" and "Only": Why not all Focus Sensitive Operators are Alike'. MS., Stanford University.

Bolinger, D.: 1958, 'A theory of pitch accent in English'. *Word* **14**, 109–149.

Chomsky, N. and M. Halle: 1968, *The Sound Pattern of English*. New York: Harper and Row.

Cohan, J.: 2000, 'The Realization and Function of Focus in Spoken English'. Ph.D. thesis, University of Texas at Austin.

de Jong, K. J.: 1995, 'The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation'. *Journal of the Acoustical Society of America* **97**, 491–504.

Dryer, M. S.: 1994, 'The pragmatics of association with *only*'. Paper presented at the 1994 Winter Meeting of the L.S.A. Boston, Massachusetts.

Geurts, B. and R. van der Sandt: 1997, 'Presuppositions and backgrounds'. In: P. Dekker and M. Stokhof (eds.): *Proceedings of the Eleventh Amsterdam Colloquium*. pp. 37–42.

Gussenhoven, C.: 1984, *On the Grammar and Semantics of Sentence Accents*. Foris, Dordrecht.

Hajičová, E.: 1973, 'Negation and topic vs. Comment'. *Philologica Pragensia* **16**, 81–93.

Hajičová, E.: 1984, 'Presupposition and allegation revisited'. *Journal of Pragmatics* **8**, 155–167.

Hajiĉová, E., B. H. Partee, and P. Sgall: 1998, *Topic-Focus Articulation, Tripartite Structures, and Semantic Content*. Kluwer Academic Publishers.

Harrington, J., J. Fletcher, and M. E. Beckman: 2000, 'Manner and place conflicts in the articulation of accent in Australian English'. In: *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge University Press, pp. 40–51.

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

Jacobs, J.: 1983, *Fokus und Skalen*. Tübingen: Niemeyer.

Koktová, E.: 1986, *Sentence Adverbials*. Amsterdam: John Benjamins.

Koktová, E.: 1987, 'On the scoping properties of negation, focusing particles and sentence adverbials'. *Theoretical Linguistics* **14**, 173–226.

Krifka, M.: 1992, 'A Compositional Semantics For Multiple Focus Constructions'. *Informationsstruktur und Grammatik* **4**.

Krifka, M.: 1997, 'Focus and/or Context: A second look at second occurrence expressions'. In: H. Kamp and B. Partee (eds.): *Context-dependence in the Analysis of Linguistic Meaning (Proceedings of the workshops in Prague and Bad Teinach)*. University of Stuttgart, pp. 253–276.

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

Ladd, D. R.: 1996, *Intonational Phonology*. Cambridge University Press.

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

Pierrehumbert, J. B.: 1980, 'The Phonology and Phonetics of English Intonation'. Ph.D. thesis, MIT.

Roberts, C.: 1995, 'Domain Restriction in Dynamic Semantics'. In: E. Bach, E. Jelinek, A. Kratzer, and B. Partee (eds.): *Quantification in Natural Languages*. Kluwer Academic Publishers, pp. 661–700.

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

Rooth, M.: 1985, *Association with Focus*. Graduate Linguistics Students Association.

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

Rooth, M.: 1996a, 'Focus'. In: S. Lappin (ed.): *The Handbook of Contemporary Semantic Theory*. London: Basil Blackwell, pp. 271–297.

Rooth, M.: 1996b, 'On the Interface Principles for Intonational Focus'. In: T. Galloway and J. Spence (eds.): *SALT VI*. Ithaca, NY, pp. 202–226.

Schwarzschild, R.: 1997, 'Why Some Foci Must Associate'. unpublished.

Selkirk, E.: 1995, 'Sentence Prosody: Intonation, Stress, and Phrasing'. In: J. A. Goldsmith (ed.): *The Handbook of Phonological Theory*. London: Basil Blackwell, pp. 550–569.

Sluijter, A. and V. van Heuven: 1996, 'Acoustic correlates of linguistic stress and accent in Dutch and American English'. In: *ICSLP 96*. pp. 630–633.

Taglicht, J.: 1984, *Message and Emphasis*. London and New York: Longman.

Turk, A. E. and L. White: 1999, 'Structural influences on accentual lengthening'. *Journal of Phonetics* **27**, 171–206.

Vallduví, E.: 1990, 'The Information Component'. Ph.D. thesis, University of Pennsylvania.

von Fintel, K.: 1994, 'Restrictions on Quantifier Domains'. Ph.D. thesis, UMass Amherst.

von Fintel, K.: 1995, 'A Minimal Theory of Adverbial Quantification'. In: B. Partee and H. Kamp (eds.): *Context Dependence in the Analysis of Linguistic Meaning: Proceeding of the Workshops in Prague, February 1995, Bad Teinach*. pp. 153–193.

von Stechow, A.: 1985/1989, 'Focusing and Backgrounding Operators'. Technical report, Universität Konstanz.

Williams, E.: 1997, 'Blocking and Anaphora'. *Linguistic Inquiry* **28**(4), 577–628.

Notes

[1] A number of English focus sensitive expressions have been identified in the literature: questions, focusing sensitive operators like "only" and "even", counterfactual conditionals, adverbs of quantification, frequency adverbs, generics, emotive factives and attitude verbs, superlatives and modals (Rooth, 1996a; Hajičová et al., 1998).

[2] The semantics in (5) involves higher order quantification, unlike (3) which involves only first order quantification. For simplicity we have taken the semantics to be extensional in both cases. However, it is natural to consider X to be a set of intensional properties, in which case we would use Montague's $^\vee$ operator to extract the extension, replacing "$P(\text{SUB})$" by "$^\vee P(\text{SUB})$" in (5), and take ϕ also to be intensional.

[3] See Rooth (1996b:206) for a similar argument.

[4] In fact the filler stimuli were drawn from a separate experiment run simultaneously.

[5] Note that while table 6 shows absolute values of range and maximum pitch in Hz., the significance test was based on the z-score of the logarithm of the frequency, as noted in the previous section. The p-values for range and maximum pitch shown in the table thus relate not to the absolute values given in the same row, but to normalized values which we have not included here.

[6] In addition to the measures given in 6, we also examined mean fundamental frequency, finding again no significant correlation with condition. We also computed statistics by individual speaker and by individual potential focus expression. The number of degrees of freedom was so low for these statistics that results were in the majority of these cases insignificant for all measures, particularly given that when computing so many individual statistics we should set a high bound on what we count as significant.