Contrastive Focus vs. Discourse-New:
Evidence from Phonetic Prominence in English

Jonah Katz (Institut Jean Nicod, Paris)
Elisabeth Selkirk (University of Massachusetts, Amherst)

1. Introduction

This article presents new evidence from English that the theory of grammar makes a distinction between the *contrastive focus* and *discourse-new* status of constituents. The evidence comes from a phonetic investigation which compares the prosody of all-new sentences with the prosody of sentences combining contrastive focus and discourse-new constituents. We have found that while the sentences of these different types in our experimental materials are not distinguished in their patterns of distribution of pitch accents and phonological phrase organization, they do differ in patterns of *phonetic prominence*—duration, pitch and intensity, which vary according to the composition of the sentence in terms of contrastive and/or new constituents. The central new finding is that contrastive focus constituents are more phonetically prominent than discourse new constituents that are contained within the same sentence. These distinctions in phonetic prominence are plausibly the consequence of distinctions in the phonological representation of *phrasal prosodic prominence (stress)* for contrastive focus and discourse-new constituents in English.

1.1 Background

Certain basic generalizations about the patterns of pitch accenting found in sentences with different types of information structure in English are widely agreed upon. For example, in most standard varieties of English, in an all-new, out-of-the-blue utterance spoken in a context with no shared prior discourse, all non-pronominal noun phrases necessarily carry pitch accents, regardless of position in the sentence (see Gussenhoven 1983 et seq, Selkirk 1984 et seq, Rochement 1986 and others). Imagine a situation where family members have just sat down to an evening meal together and one asks whether

anything newsworthy has happened that day. Responses like those in (1) would contain constituents that are discourse-new[1] and the pitch accenting shown would be appropriate:

(1) Has anything newsworthy happened today?
   A. Elíza màiled the cáramels.
   B. I rèad that the Tímes is òffering néwspaper subscriptions to the póor.
   C. Wíttgenstein spòke to Ánscombe at the fáculty meeting.

(Obligatory pitch accents in the sentences are marked by acute accents, optional ones by grave accents.) It's been proposed that the locus of obligatory pitch accent in such cases is a position of phrasal stress prominence (Ladd 1980, 1996, 2008, Selkirk 1984, 1996, Truckenbrodt 1995, Calhoun 2006, 2010, Féry and Samek-Lodovici 2006, Kratzer and Selkirk 2007), though other scholars adhere to the view that there is no such thing as phrasal stress prominence and see pitch accent distribution within the sentence as defined independently (Bolinger 1961, 1972, Gussenhoven 1983, 1991).

The empirical picture in English is radically different when sentences like those in (1) are uttered in other types of discourse contexts. Consider the case where the sentence in (1A) is uttered as a correction to a previous statement, as in (2BDF).

(2) A. Sárah màiled the cáramels.
   B. No, Elíza *mailed the caramels.*

   C. Elíza àte up the cáramels.
   D. No, *Elìza* máiled *the caramels.*

   E. Elíza màiled the póster.
   F. No, *Elìza màiled* the cáramels.

[1] All-new sentences like these which contain no (contrastive) focus within them have been widely referred to as 'broad focus' sentences. For reasons to be explained in section 1.3, this term (due to Ladd 1980) is not used in this paper.

The correction sentences (2BDF) have semantic/pragmatic properties that are quite distinct from the all-new, 'out of the blue' sentence (1A). Each of them involves what could plausibly be called a "contrastive focus" constituent, indicated by underlining[2]. And the surrounding material in the sentence, which is italicized, is discourse-given, having been mentioned in the immediately preceding sentence. These information structure properties of contrastive focus and discourse-givenness correlate closely with the different patterns of pitch accenting seen in the correction sentences. The contrastive focus—the correction constituent-- necessarily carries a pitch accent, even if it's a verb. As for the discourse-given constituents, those that appear to the right of the focus constituent necessarily fail to carry any pitch accent at all, while those that precede the focus may optionally be pitch accented. In this paper will see that, when all the relevant data is taken into account, the conclusion must be drawn that the theory of grammar makes a three-way distinction between contrastive focus, discourse-new and discourse-given, as proposed already by Chafe 1976, for example.

In the Rooth 1992, 1996 theory of the semantics of (contrastive) focus, focus constituents introduce alternatives into the discourse. Focus on *mailed* in (2D) gives rise to an alternatives set for (2D) that would include propositions with different substitutions for *mailed*: {Emily mailed the caramels, Emily ate up the caramels, Emily hid the caramels, etc.}. The members of the alternatives set are the ordinary meanings corresponding to these sentences. For Rooth, such an alternatives set constitutes the focus meaning of the sentence (as opposed to its ordinary meaning). As shown in Rooth 1992, 1996 and in other work (e.g. Hamblin 1973, Hagstrom 1998, Kratzer and Shimoyama 2002; Horn 1972), alternatives sets like these are exploited in various ways by the semantics and the pragmatics. In the case of the correction sentence (2D) in the context (2C), for example, focus on *mailed* implicates that Emily did not eat the caramels, that Emily did not hide the caramels, etc. The Rooth theory of focus as crucially involving alternatives sets is assumed in this paper, though unlike Rooth we will use the term 'contrastive focus' instead of 'focus' alone to refer to the semantic property involved in the establishment of

[2] Using the terminology introduced by Ladd 1980, examples like these with an internal (contrastive) focus constituent are instances of 'narrow focus' sentences.

alternative sets. Using 'contrastive focus' to refer to alternatives-based focus has a mnemonic value, recalling the implicit contrast that is set up between different members of the focus alternatives set. The longer term can be cumbersome, though, so for short we will often instead use the capitalized term 'Focus', to refer to alternatives-based contrastive focus.

Consider the further examples of contrastive focus sentences in (3BCD), which appear in an invented conversation beginning with sentence (3A). Each of the subsequent sentences in this conversation introduces an alternatives set in semantic interpretation. Sentences (3BCD) each include a single constituent which is new to the discourse and is moreover a contrastive focus (underlined); and in each sentence the remaining material may count as given in this conversational context, hence the italics.

(3) A: Brattleboro elected odd people to the City Council this year.
B. Kéene *didn't elect odd people to the City Council.* [It's not a regional trend.]
C: *Bràttleboro elècted òdd people to the* Schóol Board, too.
D: Yeah, *they* only *elècted* church pástors *to the School Board.*

In (3B) an explicit comparison or contrast is being made between the neighboring towns *Keene* and *Brattleboro.* Focus-marking of *Keene* is responsible for the introduction of the alternatives set {Brattleboro elected odd people to the City Council this year, Keene elected odd people to the City Council this year, ….} that constitutes the focus meaning of (3B) . In the next contribution to the conversation, another contrastive parallel, one whose presence is explicitly signaled by the particle *too,* is established between (3C) and (3A); the alternatives set corresponding to the focus meaning of (3C) would be {Brattleboro elected odd people to the City Council, Brattleboro elected odd people to the School Board, …}. And in (3D), a focus-sensitive particle *only* associates with the alternatives set introduced by the focus constituent *church pastors*: {Brattleboro elected odd people to the School Board, Brattleboro elected church pastors to the School Board, …}. The substitutions within a shared frame that are seen in these alternatives sets have

been taken to be the hallmark of the semantic interpretation of (contrastive) focus since the early semantic analysis of focus by Jackendoff 1972.

As is clear from the pitch accenting patterns in the paradigms in (1)-(3), discourse-given constituents receive a different phonological treatment from contrastive focus constituents and from simple discourse-new constituents. For example, if a noun phrase in (1)-(3) is either a contrastive focus or discourse-new, it obligatorily bears a pitch accent, while a noun phrase that is discourse-given may lack a pitch accent entirely, or only optionally bear one, depending on position. The similarity in the pitch accenting properties of contrastive focus and discourse-new constituents, as compared to given ones, might appear to support a theory according to which contrastive focus and discourse-newness are just two sides of the same coin and are represented identically in the grammar. Indeed, many scholars (including Bolinger 1961, Jackendoff 1972, Ladd 1980, Gussenhoven 1983, 1992, 2004, Selkirk 1984, 1996, Rooth 1985, 1992, Kratzer 1991, Krifka 1992, 1993, Schwarzschild 1999; Kahnemuyipour 2004, 2009, Büring 2007, Féry & Krifka 2008) adopt the position that examples of discourse-new and (contrastive) focus constituents like those above are both instances of a same fundamental information structure category, a category which in both cases has been assigned the same F-marking notation in syntactic representation. But others (including Halliday 1967, Chafe 1976, Rochemont 1986, Pierrehumbert & Beckman 1988, Kiss 1998, Vallduví & Vilkuna 1998, Frota 2000, Face 2002, D'Imperio 2002, LeGac 2002, Selkirk 2002, 2007, 2008, Neeleman and Szendroi 2003, Kratzer 2004, Féry and Samek-Lodovici 2006, Kratzer and Selkirk 2007, 2010) would see the common capacity to bear obligatory pitch accents in English as obscuring a fundamental semantic (and syntactic) distinction between contrastive focus and discourse-new. It is the distinction in their phonetic properties, this paper shows, that supports the postulation of a grammatical distinction between contrastive focus and discourse-new. The fact that both may obligatorily bear pitch accents, as in the examples in (1-3), follows if both bear at least a minimum level of phrasal stress prominence, and if the presence of pitch accenting is phrase-stress-driven in English, as argued by Ladd 1996, 2008, Selkirk 2002, Féry and Samek-Lodovici 2006, Calhoun 2006, 2010, Kratzer and Selkirk 2007.

It turns out that the same experimental materials which provide the basis for our finding that contrastive focus and discourse-new constituents are phonetically and phonologically distinguished by the grammar of English also provide an argument that the two must be distinguished by the semantics[3]. The experimental materials include sets of minimal triplets, like those in (4), each with different combinations of post-verbal constituents that are (putatively) contrastive focus and discourse-new: Focus-new, new-Focus, new-new (see section 2.1 for further examples, and Appendix A for a full list). Because there is no antecedent in these mini-discourses for either of the verbal complements in the target sentence (marked in boldface), we assume that both count as discourse-new. The inclusion of focus-sensitive particles like *only* in the target sentences signals that one of the verbal complements is a contrastive focus (in addition to being discourse-new).

(4) A. Focus-new

Gary is a really bad art dealer He gets attached to the paintings he buys. He acquired a few Picassos and fell in love with them. The same thing happened with a Cezanne painting. **So he would *only* offer [that Modigliani] to [MoMA]**. I bet the Picassos would have fetched a much higher price.

B. new-Focus

Gary is an art dealer. Lately he's been very picky about which museum he deals with; he doesn't do business with the Metropolitan or the Guggenheim. **He would *only* offer [that Modigliani] to [MoMA]**. He says that's the only place with a good enough space to hang it in.

C. new-new

Gary was a successful art dealer, and could afford to be pretty demanding with his clients. He would never make a deal unless the price was right and he respected the buyer. H**e will probably offer [that Modigliani] to [MoMA]**. But only for a six figure sum.

[3] This argument was provided to us by Angelika Kratzer (p.s.), and appears in joint work by Kratzer and Selkirk (2010, in preparation).

It should be pointed out that definite descriptions like *that Modigliani* or proper names like *MoMA*, such as are seen in the target sentences in (4), may in principle qualify as either discourse-new or discourse-given, depending on the context. When definite expressions are used in an 'out of the blue' context, as in the utterance context for (1A), where the obligatory pitch accenting of both *Emily* and *the caramels* shows they count as discourse-new, there is of course the assumption that the referents of these definite expressions are entities in a model of discourse that is shared by speaker and hearer. Apparently, in order to qualify as discourse-given, and show the relevant deaccenting properties in English, it is not enough to be contained in the shared discourse model; to count as given an entity must be salient in the active preceding discourse (Prince 1981, Schwarzschild 1999). Since there is no salient antecedent for any of the target sentence verbal complements in the mini-discourses presented to the speakers in our experiment, we assume that both verbal complements count as discourse-new. The systematic presence of pitch accent on these complements, reported in section 3 of this paper, provides support for this assumption.

Consider now condition A, the (putative) Focus-new case. The intended reading is that only the Modigliani painting is being offered to MoMA; the implicature is that neither the Cezanne nor the Picassos are being offered to MoMA. At issue in the focus-related meaning of this sentence is the alternatives set {he would offer that Modigliani to MoMA, he would offer the Picassos to MoMA, he would offer the Cezanne to MoMA}. *Only* associates with this alternatives set, and is responsible for the interpretation that, of these alternatives, only the proposition *He would offer that Modigliani to MoMA* is true. In condition B, it is the alternatives set {he would offer that Modigliani to MoMA, he would offer that Modigliani to the Metropolitan, he would offer that Modigliani to the Guggenheim} that *only* associates with. The intended reading for B is that only MoMA has been offered the Modigliani, with the implicature that the painting has not been offered to the Metropolitan or the Guggenheim. The question now is just how these two distinct semantic interpretations are supplied by the grammar. In both cases, the focus-sensitive particle occupies a preverbal position from which it can take scope over the complements of the verb; it is not the presence of *only* or its position in the sentence that

distinguishes the cases. It must be the syntactic representation of the constituent with which *only* associates that provides the distinction.

For a theory that distinguishes between the representation of contrastive focus (Focus) and discourse-new, there is no problem. A Focus-marked constituent will contribute to the establishment of the alternatives set with which the particle *only* is associated in semantic interpretation (Rooth 1992); a new constituent which has no Focus-marking will not. The target sentences in conditions A and B of (4) would be supplied with syntactic representations that are distinguished by this Focus-marking, would be assigned distinct alternatives sets, and would thereby receive distinct readings.

(5) A. Focus-new

He would *only* offer that [Modigliani]$_{\text{Focus}}$ to [MoMA]$_{\text{"new"}}$.

B. new-Focus

He would *only* offer that [Modigliani]$_{\text{"new"}}$ to [MoMA]$_{\text{Focus}}$.

But what of the undifferentiated theory of focus which represents both contrastive focus and new constituents in identical fashion, as F-marked (Jackendoff 1972 et seq)? This theory would assign the identical F-marked syntactic representation seen in (6) to both verbal complements in the A and B target sentences:

(6) With F-marking, both conditions A and B have the same target sentence.

He would *only* offer that [Modigliani]$_{\text{F}}$ to [MoMA]$_{\text{F}}$.

But if both *Modigliani* and *MoMA* are simply F-marked, we would get the wrong alternatives set, assuming standard accounts like Rooth (1992), and hence the wrong meaning. Both F-marked objects are in the scope of *only*; both would contribute to the formation of the alternatives set. The members of this alternatives set would be propositions of the form "He would offer that *x* to *y*", with *x* and *y* being drawn from the cross-product of paintings and museums in the lists seen in (7):

(7) (i) Wrong alternatives set with simple F-marking for contrastive focus and new

He would offer that	_____	to	_____
	Modigliani		MoMA
	Cezanne	x	Metropolitan
	Picasso		Guggenheim
	etc.		etc.

(ii) Formal characterization of the incorrect alternatives set:

{p: ∃x ∃y [painter(x) & museum(y) & p = λw offer(x)(y)(Gary)(w)]}

A possible response to this argument against uniform F-marking on contrastive focus and discourse-new constituents would be to propose that one of the F-marked complements has been removed from the scope of *only* by a covert movement operation, with the result that only the F-marked complement that remains in situ within the scope of the focus-sensitive particle contributes to the formation of the alternatives set and so is interpreted as a contrastive focus[4]. But this proposal is not tenable[5]. Consider an alternative target sentence for the paradigm (4A-4B) where, rather than containing *only*, the sentence contains the negative particle *not* and in addition one of the verbal complements contains the negative polarity item *any*:

(8) So he wouldn't ever offer any of his Modiglianis to MoMA.

In a target sentence in condition A, it is *any of his Modiglianis* which would be interpreted as the contrastive focus in the sentence; in a target sentence in condition B, *any of his Modiglianis* would not be the contrastive focus, but only qualify as discourse-new. But since the proper interpretation of the negative polarity item *any* requires that it remain in the scope of negation under both interpretations, regardless of contrastive focus status, target sentences of the form in (8) show that the needed difference in contrastive focus interpretation in the two sentences cannot be obtained through a covert movement out of the scope of the negation by a discourse-new constituent. In sum, we see that

[4] This proposal was suggested by Michael Wagner (p.c.).
[5] The following counterargument is due to Kratzer 2010b.

consideration of the semantic interpretation of the target sentences we will examine in this paper already provides an argument that the theory of grammar must provide a morphosyntactic marking which distinguishes between contrastive focus and discourse-new constituents.

A final point needing clarification concerns the question whether it is necessary to posit a morphosyntactic feature [New] alongside the feature [Focus] which represents contrastive focus. Supposing that the grammar makes a three-way distinction between Focus, new and given, it is not necessary to represent both "new" and "given"; it would suffice to mark just one of the two. Since, logically speaking, new is simply not given and given is simply not new, the semantics does not at first blush provide an argument for which of the two requires representation[6]. Accounting for facts of sentence phonology, however, appears to require the positing of a [Given] feature. In English and German, for example, it is discourse-given constituents which display marked behavior like deaccenting; this departs from the default, "elsewhere", properties which are manifested by the phonology of new constituents (see, e.g. Féry and Samek-Lodovici 2006, Truckenbrodt 2006, Kratzer and Selkirk 2007). For this reason, we do not assume the existence of a feature [New] in this paper. Any use of the term "new" to annotate or refer to discourse-new, non-Focus constituents should not be taken to presuppose the existence of a [New] feature.

1.2 Phonetic investigation of Focus vs. discourse-new

Because the mere presence of a pitch accent does not reliably distinguish a Focus constituent from a discourse-new one in English, finding evidence for a sound-related distinction between the two hypothesized categories will necessarily involve other properties of their prosody. It's been suggested that in English contrastive focus is marked by a L+H* pitch accent, while a simple H* pitch accent is found with discourse-new elements (Pierrehumbert and Hirschberg 1990), though this proposal seems not to be tenable (see discussion in section 3.1). It has also been proposed that in English a Focus

[6] Though see the argument in Selkirk 2008 that an account of second occurrence focus requires that such constituents be marked for both [Focus] and [Given].

carries a higher degree or level of prosodic prominence than any other constituent (Jackendoff 1972, Ladd 1996, 2008, Truckenbrodt 1995, 2006, Neeleman and Reinhart 1998, Rooth 1996, Williams 1997, Selkirk 2002, Neeleman and Szendroi 2004, Wagner 2005, Féry and Samek-Lodivici 2006, Kratzer and Selkirk 2007, 2010, Calhoun 2006, 2010). Such proposals have not been accompanied by any attention to the relation between this hypothesized abstract prosodic prominence and the phonetic properties that might realize it[7]; in the course of this paper we construct an initial hypothesis about what that relation might be. In the experimental phonetics literature it has been argued that Focus and discourse-new constituents in English, and certain other languages, are in fact distinguished by the degree of phonetic prominence assigned to them, in particular by amount of pitch protrusion, duration and/or intensity (Cooper et al 1985, Eady and Cooper 1986, Xu and Xu 2005, Breen 2007, Breen et al 2010). These previous phonetic studies make no attempt to relate the findings concerning phonetic prominence to any hypothesized phonological representation nor to any theory of how phonetics might interface with information structure in the organization of a grammar. Nonetheless the results are suggestive, despite certain shortcomings in design discussed in section 1.3. In accordance with previous proposals regarding the phonological and phonetic prominence of Focus as compared to discourse-new, then, a further working hypothesis we adopt here is that, in English, a Focus constituent shows a greater degree of phonological prominence than a discourse-new constituent, and as a consequence shows a greater degree of phonetic prominence. The question is how to best test this hypothesis.

1.2.1 Comparing Focus and new within the same sentence

It is clear that phonetic/phonological prominence is not associated with any absolute phonetic values for duration, pitch or intensity, just as a High tone can vary in its quantitative realization in terms of pitch (F0). So a comparison of the phonetic values for pitch, duration, etc. from a pitch-accented Focus constituent in one utterance with the

[7] Calhoun's (2010) paper is a partial exception to this statement. Although she does not provide a theory of which phonetic aspects are relevant to phonological prominence in which ways, she provides an impressionistic discussion of how phonetic properties might indicate prominence and phrasing distinctions in specific utterances.

values for pitch, duration, etc. from a pitch accented nonFocus discourse-new constituent in another utterance is not the best testing ground for our hypothesis. Considerable overlap in these values is in principle possible in such cases. However, within a sentence-sized utterance, variation in the pitch values of sequences of pitch accented entities is under considerably more grammatical control (Liberman and Pierrehumbert 1984, Ladd 1988, van den Berg *et al* 1992, Truckenbrodt 2002). Given this, a better test of our working hypothesis is an investigation of the *relative phonetic prominence* of Focus and discourse-new constituents *within the same sentence*. So the paradigm of experimental stimuli used in our experiment includes combinations of contrastive focus and discourse-new constituents, as schematized in (9).

(9) a. Focus-new: [..… […]$_{\text{Focus}}$ […]]
b. new-Focus: [..… […] […]$_{\text{Focus}}$]
c. new-new: [..… […] […]]

Note that it is the very strong tendency for constituents of both the Focus and discourse-new types discussed here to bear pitch accents in English that allows us to investigate the hypothesis that Focus constituents have a greater degree of phonetic (and phonological) prominence than discourse-new constituents within the same sentence. (This tendency is documented in our own experimental materials in section 3.1.) A difference in the presence of pitch accent on the two constituents might lead to differences in phonetic prominence that are the consequence of the presence or absence of the pitch accent itself, and not necessarily correlated with a distinction between Focus and new.

It is perhaps not surprising that there has been no prior phonetic investigation of sentences combining Focus and discourse-new constituents. The very existence of sentences that combine Focus and new constituents has gone virtually unreported in most of the literature on information structure and prosody. Such sentences are taken up by Reinhart and Neeleman (1998), Selkirk 2002, Neeleman and Szendroi (2004) and Féry and Samek-Lodovoci (2006), however, in discussion of a general theory of stress

assignment in sentences with Focus[8]. Calhoun (2006) also discusses sentences of this type from the Switchboard corpus. These authors observe, on the basis of intuition, that greatest stress prominence in the sentence appears to be assigned to the constituent bearing the (putative) contrastive focus and that this Focus may be followed or preceded in the sentence by other (discourse-new) constituents which also bear phrasal prominence (and pitch accent). Here we provide evidence from a controlled production experiment supporting the claim that such prominence patterns for sentences combining Focus and new do in fact exist, and investigate their relevance for the grammar of prosody.

The paradigm of sentences for our experiment includes minimal triplets of the type schematized in (9) and illustrated above in (4). The members of the minimal triplet differ in the Focus vs. new status of two syntactic phrases in a sentence: Focus-new, new-Focus and new-new. As shown below in sections 4-6, there is a three-way distinction in patterns of phonetic prominence in these three sentence types, so these findings provide support for the hypothesis that the grammar does indeed make a distinction between contrastive focus and discourse-new.

1.2.2 Other experimental investigations of the Focus-new distinction in English

The current experiment is by no means the first to compare the prosodic properties of all-new sentences with sentences containing contrastive focus in English. But earlier studies differ from ours in employing sentences in which the material surrounding the (putative) contrastive focus is discourse-given, as in responses to alternative questions (Cooper et al

---

[8] In a key example sentence offered by Neeleman and Szendroi 2004, appearing as the final sentence of the dialogue repeated here, the noun phrase *Superman* is a Focus that is both preceded and followed by discourse-new phrases (*Johnny, some kid):*

Father: What happened?

Mother: You know how I think our children should read decent books. Well, when I came home, rather than doing his homework, [$_{IP}$ Johnny was [$_{VP}$reading [$_{DP}$SUPERMAN] to some kid]]

1985), correction statements (Breen 2007, Breen et al 2010) and responses to wh-questions (Eady and Cooper 1986, Xu and Xu 2005, Breen 2007, Breen et al 2010). The sentence types compared in this earlier work have been of the following schematic types (analogous to the all-new sentence (1A) vs. the contrastive focus sentences (2BDF) above):

(10)

	I	II	III
a.	new	new	new
b.	Focus	Given	Given
c.	Given	Focus	Given
d.	Given	Given	Focus

A second difference in the design of the investigation undertaken in this paper is that these earlier studies compare phonetic properties of a Focus constituent in one sentence with the phonetic properties of a discourse-new constituent in another sentence. The studies cited above report differences in duration, pitch and intensity between the Focus constituents found in contrastive focus sentence types (10bcd) and the discourse-new constituents found in the all-new sentence type (10a) in all sentence positions (I, II, III). Focus constituents had longer duration than the corresponding new constituents in all-new sentences. The Focus constituents also showed greater pitch protrusion and greater post-Focus pitch compression. But the design of these experiments requires us to interpret these results with some caution.

One shortcoming of the design of these earlier experiments is that fully controlled comparison of the prosody of (putative) Focus and new constituents is not possible, since a (putative) Focus constituent in contrastive focus sentence types and its corresponding new constituent in the all-new sentence type do not appear in prosodically identical environments within their respective sentences. For example, we know that in a sentence of type (10c) in English, the (putative) Focus in medial position would be followed by accentless material, whereas the medial new constituent in (10a) would be followed by accented material; comparable distinctions in surrounding prosodic context appear with the other Focus sentence types. Such a difference in prosodic context itself could

potentially be responsible for differing phonetic properties, rather than a putative syntactic difference in Focus vs. new status (and a consequent difference in phonological representation). This shortcoming in design is of particular importance in the study of pitch patterns, which are context-dependent. So, for example, the fact that the sentence-medial constituent in (10c) has significantly higher F0 than the corresponding medial constituent in (100a) could be the consequence of the putative contrast in information structure status (Focus vs. new) of that constituent. But it could also result from a difference in the preceding prosodic context. If the obligatorily pitch-accented (and phrase-stressed) new constituent in initial position in (10a) triggers a downstepping or lowering of pitch range on what follows, as in all-new sentences in other Germanic languages (van den Berg et al 1992, Truckenbrodt 2004), then the new constituent in medial position will accordingly be lowered in pitch, while still bearing pitch accent (and phrasal stress). In sentence type (10c), on the other hand, the medial Focus is preceded by a Given constituent, and because that initial Given constituent lacks the obligatory pitch accent (and phrasal prominence) of a new constituent, it may fail to trigger the same degree of downtrend on what follows. As a consequence the medial constituent in (10c)—the putative Focus--would show higher F0 than the corresponding medial new constituent in (10a). This means that in a comparison of the pitch values of the medial Focus and new constituents here, it's not certain one can attribute the difference in F0 to the Focus-new distinction. The Given-new distinction in the preceding context and its consequent phonological and phonetic effects could just as well be responsible. Which is to say that the prosodic differences displayed could be consistent with a common F-marking representation for the contrastive focus and discourse-new medial constituent.

For these reasons, we need to compare sentences containing sequences of (putative) Focus-new, new-Focus and new-new constituents in order to provide crucial evidence for the existence of a contrastive focus vs. discourse-new distinction in grammar. The experiment we report here constitutes a first attempt to investigate a paradigm of sentence types like (9) which controls for sentence-internal context, with a mind to determining whether a systematic greater prominence for Focus-marked constituents, as compared to merely discourse-new constituents in the same sentence, does in fact emerge.

At this point some terminological clarification is in order. In phonetic studies examining sentences belonging to the types schematized in (10), the term "narrow focus sentence" has typically been employed to refer to sentences types that contain an internal contrastive focus, as in (10bcd), while the term "broad focus sentence" has been used to refer to the all-new sentence type (10a). This terminology, due originally to Ladd 1980, does not adequately characterize the sentences types at issue. The term "narrow focus" has come to be associated with Focus constituents that appear in Given contexts within the sentence, as in (10bcd). This practice may indeed be partly responsible for a general failure to recognize that a Focus may also co-occur with discourse-new constituents in a sentence, or with some combination of Given and new, even though these latter sentence types contain "narrow focus" too. As for the term "broad focus", it has typically been used to refer to sentences which do not contain a "narrow" contrastive focus, and which moreover consist of constituents that are all discourse-new. Such sentences are assumed by Ladd 1980, following Jackendoff 1972, to be Focus constituents themselves. But the notion that every sentence must contain at least one Focus (if not "narrow" then "broad") is questionable on semantic/pragmatic grounds (Kratzer and Selkirk 2010, in preparation). And even if an entire sentence were itself a Focus constituent, this would imply nothing about its internal composition in terms of Given vs. new. In other words, the broad/new focus terminology is descriptively inaccurate. In this paper, the term "contrastive focus sentence" refers, depending on the circumstances, either to a sentence that contains a Focus constituent within it or to the particular class of contrastive focus sentences examined in our experiment, which combine sentence-internal Focus and discourse-new constituents. We use the term "all-new sentence" to refer to sentences lacking any (internal) Focus and which moreover contain no Given material.

1. 4 What's ahead

In the next section, we lay out details of the design of our experiment, including the sentence types examined. Section 3 reports data on the distribution of pitch rises and falls in the intonational contours of these sentences and motivates the assumptions that we will

make concerning the distribution of pitch accents and phonological phrase organization. Sections 4-6 report on the findings concerning phonetic prominence, specifically duration (section 4), pitch (section 5) and intensity (section 6). In these, reports of the phonetic findings are followed by discussions which develop our understanding of the implications of these findings for the semantics, syntax and phonology of the (putative) Focus vs. discourse-new distinction, leading to a summary in Section 7.

2. The design of the experiment

2.1 Experimental materials

The goal of this study is to compare the pronunciations of sentences which are essentially identical except for the Focus vs. new status of key constituents. Ensuring that these sentences are indeed produced with the intended information structural meanings in mind poses a methodological challenge. Just how can the experimenter induce a speaker to pronounce a sentence with the desired information structure in a way that approximates natural usage? Elicitation of sentences combining Focus and new constituents requires materials that are somewhat more complex than is typical in studies of standard 'narrow focus' sentences. These latter have usually been elicited in *wh*-question-and-answer dialogues or in dialogues with correction statements like (2), making up experimental paradigms like (10).

Another method for cuing the presence of a contrastive focus is to include a focus-sensitive particle or adverb like *only* or *even* in the sentence. For these particles to be interpretable semantically or pragmatically, they must associate with an alternatives set introduced by a Focus constituent lying within their scope (see Rooth 1992, 1996, Beaver and Clark 2008 on this issue). Our materials used preverbal focus-sensitive particles in the Focus-new and new-Focus sentences to indicate that one of the constituents following the particle is a contrastive focus. The all-new sentences that we used contain no focus-sensitive particles. This was illustrated above with the set of minimal triplet target sentences contained in the three mini-discourses in (4); (11) gives additional examples.

(11) Sample minimal triplet discourses from the experimental materials

1A. The Red Sox had an exhibition game for charity, and they gave the players various bright-colored uniforms. Bill Mueller and Nomar Garciaparra have really played well this year. **But they *only* gave [Manny] [the yellow one].** That's the one that's reserved for the most valuable player.

1B. The Red Sox had an exhibition game for charity, and they had special bright-colored uniforms made for the occasion. There were a lot of different colors; a couple of the jerseys were orange, one was purple. **But they *only* gave [Manny] [that yellow one].** That was a lousy color**.**

1C. The Red Sox had an exhibition game for charity, and they gave all the players crazy bright-colored uniforms to wear for the occasion. The whole thing was pretty funny to watch. **They gave [Manny] [the yellow one].** It was so ugly.

2A. Bill chooses the most awful companions. He was dating that horrible lawyer last year, and then there was Kate, who we all hated. **He *even* took [Minnie] to [a Mariners game]**. And she's insufferable.

2B. Bill is a sports freak. He'll go to any kind of sporting event, regardless of the team that's playing. **He *even* took [Minnie] to [a Mariners' game]**. And they haven't been in contention for years!

2C. Bill's had a pretty busy week. He had meetings all through the weekend, and then he went to Seattle for a conference. **He took [Minnie] to [a Mariners' game]**. I bet that was fun.

The semantic interpretation of the pre-verbal focus-sensitive particle (italicized) requires association of the particle with one of the constituents that follow it. But it does not pinpoint which constituent is the Focus: it could be the verb phrase, the verb, or one of the VP-internal constituents. Our experiment exploits this variation. We compare the case of Focus on the first post-verbal constituent to the case of Focus on the second, and compare these (putatively) Focus-containing sentences with Focus-less all-new sentences. (In what follows we refer to these post-verbal constituents as 'complements', though they do include instances of non-argument constituents in the verb phrase.)
To bias the speaker in favor of one or the other location of Focus in the contrastive sentences, we used a preceding and following discourse which would tend to reinforce

that interpretation. From the discourse the speaker could also infer the new status of constituents preceding and following the Focus in the contrastive sentences.

2.2 Recording the materials

Stimulus materials consisted of minimal triplet discourses, as described above. There were 18 such triplets, for a total of 54 discourses. A full list of materials is in Appendix A. Each subject participated in three sessions. Each session consisted of six discourses in each condition, for a total of 18 discourses. Two members of a minimal triplet never co-occurred in the same recording session. Sessions were scheduled at least two days apart. During each session, subjects completed the eighteen discourses first in one random order, then in another. The repetition ensured a full set of recordings from each subject in the case of errors in reading or recording, or difficulties in acoustic analysis. Subjects were seated in a soundproof recording booth and read the sentences into a condenser microphone positioned on a desk in front of them, slightly to one side.

Each stimulus consisted of a paragraph presented on a computer screen, generally three to five sentences in length, as shown in the examples in section 2.1. The first several sentences (the *context*) were presented in plain type, the last two sentences in bold type. The first bold sentence was the *target sentence*, and the last sentence in the paragraph was designed to reinforce the desired information structure for the target sentence. Subjects were instructed to silently read each paragraph several times, to get a feel for it. They were instructed to next push a button, which prompted a recording of the preceding context sentences read by one of the experimenters, a native speaker of American English. In the A and B conditions, the noun phrases from the preceding context which corresponded to the intended contrastive focus in the target sentence were pronounced with some emphasis, e.g. *Bill Mueller* and *Nomar Garciaparra* in discourse 1A in (11). The subjects were instructed to read aloud the two sentences in bold type after the context was played, completing the paragraph in 'as natural a manner as possible'.

One sentence (number 8 in Appendix A) was consistently produced with a pitch accent on the verb in condition A and no accent on the first complement. Because this item did not receive the intended information structure in condition A, we excluded all instances of the sentence in all conditions from our analysis. In addition, one token of another sentence (number 9 in Appendix A) was improperly recorded by the experimenters, and could not be analyzed. This left a grand total of 509 utterances from the five subjects.

Subjects were one female and four male students at the University of Massachusetts. Three were graduate students in the linguistics department, two were undergraduates. All subjects were naive to the purposes of the experiment. None reported having been diagnosed with speech, hearing, or reading disorders.

2.3 Measurements

The paramount consideration in constructing contexts and target sentences was the intended information structure for each item. Ideal stimuli, from the point of view of phonetic analysis, would contain complements of a uniform length and similar segmental makeup in all 18 sentence frames; in practice this was not possible. But given the minimal triplet design, and the statistical analysis of the results, the diversity in make-up of the sentence frames is not problematic. The complements of the minimal triplets corresponding to each of the 18 sentence frames were identical across the three conditions.

Generally, the boundaries selected for durational measurements of the complements were the closest easily-identifiable phonetic events, as described below. This method resulted in strings measured for duration that are consistent across conditions within each sentence frame, but differ between sentence frames. Because the experiment was a balanced design (i.e., the same measured strings were used the same number of times in each condition), between-item variation (across the 18 sentence frames) in the duration of the measured string will not affect differences between conditions. In addition, by-item random effects were included in the statistical model to account for variance in the

overall duration of the measured strings. The string measured for duration in the first complement was two or three syllables in length, including the pitch accented syllable in the lexical word; the measured string in the second complement was generally one or two syllables in length and also included the pitch accented syllable. The last syllable of the sentence (which was not the pitch accented one) was excluded from the durational measurement wherever possible to avoid utterance-final glottalization and 'trailing off'[9].

Concerning segmentation, for each item some phonetic property or cluster of properties was used as a boundary criterion for all instances of that item. If there were obstruents in close proximity to the pitch accent, these were used as criteria. Generally, characteristic movements in formants and waveform envelope were chosen as criteria when obstruent-sonorant boundaries were unavailable. For instance, in the minimal triplet involving (11-1AB), in *the yellow one*, with pitch accent on *yel-*, there are no obstruents to use as a right boundary. In this case, the criteria chosen for the right boundary of *yellow* were an F1 minimum and an amplitude minimum in between the ultimate and penultimate syllable, corresponding to the segment /w/. (See Appendix B for characteristics of duration measurement points for each of the minimal triplets.)

Measurements were normalized by subject using a Z-transform. The Z-transform expresses each data point in terms of its number of standard deviations above (positive values) or below (negative values) the subject's mean value. It helps reduce the amount of variance between subjects due to differences in speech rate or pitch register, for instance, and therefore results in simpler statistical models. A separate mean and standard deviation were calculated for each subject based on all measured pitch values for that subject.

[9] A reviewer suggests that a better method would be to measure strings of roughly equivalent length (such as the pitch accented syllable) for each sentence frame. While this would likely reduce between-item variance, it would also result in more difficulty in segmentation, which would plausibly introduce greater experimenter error. While between-item variance can be controlled for in statistical analysis, experimenter error cannot. Additionally, these materials are counterbalanced for which lexical items appear in each condition, meaning that between-item variance will not affect differences between conditions (Raaijmakers *et al*. 1999).

F0 measurements were taken at loci of F0 maxima and minima. Some of these measurements could be affected by local perturbations of F0 due to segmental context. But because the experiment had a balanced design, segmental perturbations would not affect differences between conditions. Differences between individual lexical items in this regard are incorporated into the statistical model as a random effect. 'Peak' measurements were taken at the highest point of the pitch accented syllable in each complement (m3 and m6)[10]. F0 measurements were also taken near the right and left edges of the pitch accented word of each complement (respectively m2-m4 and m5-m7); these are positions where F0 was expected to be lower than the peak of the accented syllable. These measurements were done manually and in such a way as to record the largest pitch movement possible. In cases where F0 decreased monotonically all the way to the preceding or following boundary, we selected the point in each complement furthest toward the relevant boundary where pitch tracking was possible. In cases where F0 showed a plateau internal to the complement, we selected the point furthest from the boundary where an F0 minimum was reached. In cases where F0 showed a reversal (i.e., descending followed by or preceded by ascending) internal to the complement, we measured the minimum value. One measurement was taken preceding the first complement (m1) -- the highest point in the focus-sensitive adverb (in conditions A and B) or the verb (condition C). The location of these measurement points in one particular utterance is illustrated in Appendix C.

Glottalization sometimes made F0 difficult to recover from the low points in the pitch contour, especially utterance-finally. In cases where the F0 tracker appeared to be giving spurious results, first we attempted to extract an F0 value by examining the spacing

[10] A reviewer points out that confining the peak measurement to the pitch accented syllable may understate the height of the pitch peak in cases where the peak might be realized farther to the right than the accented syllable. This situation, however, did not arise in the analysis of the experimental materials; there was an F0 maximum internal to the accented syllable in all cases.

between harmonics in an FFT spectrum[11]. If this proved impossible, the measurement was simply excluded. A total of 44 measurements, about 1.2%, were excluded for this reason. All F0 measurements were converted to the ERB scale before further analysis, because this scale is meant to reflect the perception of pitch, unlike the Hertz scale. Because most linguists are more familiar with the Hertz scale, however, we sometimes report measurements in both scales in what follows.

For intensity measurements, a Praat script automatically extracted the highest intensity value from each complement, using the temporal window from durational measurements. These data were then examined for abnormally high or low values, which were investigated by hand.

2.4 Statistical analysis

Differences between conditions were analyzed with linear mixed-effects regression models, implemented with the lme4 package (Bates 2007) in the statistical environment R. This type of model offers several advantages over the repeated measures ANOVA models that are common in speech and language research (Quené & van den Bergh 2004, Baayen *et al*. 2008). Foremost among these are the greater power of these models to detect effects in a data set, their robustness to violations of the sphericity and homogeneity-of-variance assumptions underlying ANOVA models, and their ability to incorporate crossed random effects. In this case, the fixed effects of interest (condition and position) are crossed with the random effects of speaker identity and the individual lexical items used in each sentence frame; unlike repeated-measures ANOVA, mixed-effects regression allows us to incorporate variance between speakers and between lexical items into a single model.

[11] To do this we examined the first four visible peaks in the spectrum; if they appeared to be evenly spaced, the value of the fourth peak was divided by four. Although not as precise as a pitch-tracking algorithm, this method at least gives a rough measure of F0 in cases where the algorithm fails

For the analyses of duration, F0 and intensity, each model included fixed effects for condition ({A, B, C}), position ({1st complement, 2nd complement}), interactions between condition and position, and repetition ({1st, 2nd}). Recall that conditions A, B and C vary according to the (putative) Focus vs. new status of a complement and the Focus vs. new status of the complement that precedes or follows. Crossing condition and position creates six categories, each corresponding to a unique combination of focus status and context, as illustrated in Table 1.

<INSERT TABLE 1 ABOUT HERE>

For the F0 analysis, fixed effects were also included for the F0 of the first measured point in the sentence and for the type of pitch movement being considered ({rise preceding accent, fall after accent}).

The significance of fixed effects was assessed using Markov chain Monte Carlo (MCMC) sampling. Roughly speaking, this procedure generates hypothetical sets of parameters over and over again, then compares these parameters to the actual ones the model has fitted to the data, in order to assess the probability of obtaining such extreme parameters by chance. Baayen *et al*. (2008) give a more detailed description of this procedure. The significance of these effects is reported with the effect size β, which is the coefficient of the relevant effect, and a p-value. The significance of random effects was assessed with a likelihood-ratio test between models including and excluding the random effect of interest; these effects are reported with a chi-square statistic and p-value from the likelihood-ratio test.

3. Tone and Phonological Representation

The goal of this paper is to determine whether the Focus vs. new status of constituents in a sentence is associated with differences in phonological representation and/or phonetic interpretation, and in this way to determine if linguistic theory must distinguish in principle between the grammatical representations of the two in the interface with the

phonology/phonetics. Aspects of phonological representation potentially relevant to characterizing such differences in English include the representation of tone and of prosodic structure (consisting of prosodic constituency and prosodic prominence, see e.g. Selkirk 1996, Ladd 1996, 2008). This section examines the tonal make-up of the sentence types in the study, draws conclusions based on this evidence about the prosodic constituency of these sentences, and asks if these properties could be what speakers manipulate to encode a putative Focus vs. non-Focus distinction. We will see that from the point of view of tonal and prosodic phrasing properties the Focus and discourse-new constituents of the three sentence types are identical in the majority of cases. This finding is of central importance to the interpretation of the findings we report in sections 4-6, which document distinctive patterns of duration, pitch scaling and intensity in these sentence types. The reasoning is as follows: if the Focus and discourse-new items in our materials are distinguished by presence, absence, or type of pitch accent, or by different reflexes in prosodic phrase structure, then the observed phonetic differences in F0, duration, and intensity might in principle derive from these phonological distinctions. If, however, the Focus and discourse-new items do not differ in the distribution of pitch accents and phrase boundaries, then we must appeal to some other property to explain any phonetic differences amongst them.

In sections 4-6, we use mixed-effects regression models to ask about differences in phonetic detail between the various conditions. In this section, we are asking a rather different type of question: what are the general phonetic and phonological properties of the sentences in these conditions? For this kind of question, descriptive statistics are more useful; we are asking not about the *average* properties of some set of utterances, nor about whether phonetic patterns differ significantly on average between different conditions, but about the range and consistency of phonetic patterns across all conditions.

According to the Pierrehumbert 1980 and Beckman and Pierrehumbert 1986 theory of English intonation, assumed with some modifications in much previous work[12] and in this paper, the intonational pitch contours of English are phonologically represented by tonal entities of two different types—pitch accents, which align with positions of prosodic stress prominence, and peripheral tones, which align with the edges of prosodic constituents. These tonal entities constitute targets in the phonetic interpretation of the phonological representation of tone. A high tone (H) typically corresponds to the high end point of a pitch rise or fall, whether it coincides with a high pitch peak flanked by a rise and a fall, or merely marks a turning point at the end of a rise or the beginning of a fall. A low tone (L) typically corresponds to the low end point in a pitch rise or fall.

Our expectation was that a pitch accent would appear on each of the verbal complements in each sentence of our materials, given the status of these complements as contrastive focus and/or discourse-new. And given the surrounding discourse and the status of these sentences as declaratives, we expected these pitch accents to belong to the H* class. In a neutral declarative narrative (like the ones in our experiment), other pitch accent types would tend not to occur; these include the L* which is often the final pitch accent in English yes-no questions, and the L*+H, H+L* or H+H*, which appear in pragmatically special circumstances (Pierrehumbert and Hirschberg 1990, Grice 1995).

Moreover, we expected that the post-verbal complements in these utterances might occupy distinct phonological phrases in prosodic structure; prior work on English (and other languages) shows that phonological phrase breaks have a strong tendency to appear between two verbal complements in all-new utterances (see, e.g. Selkirk 1986, 2000, 2011). If so, a tonal reflex of that phrasing should appear-- in the form of a L tone coinciding with the right edge of each complement, creating a phrase-final pitch fall from the H* pitch accent of each complement. That phrase-edge L tone is analyzed as a peripheral L- phrase accent by Beckman and Pierrehumbert 1986 and in the English ToBI

[12] Selkirk (1984, 1995, 2000), Beckman and Pierrrehumbert (1986), Ladd (1996, 2008), Gussenhoven 1990 and in the ToBI transcription system for English (Beckman and Ayers Elam 1993, Veilleux et al 2008)

system (Beckman and Ayers Elam 1993, Veilleux et al 2008); as part of a phrase-final H*+L pitch accent in Gussenhoven 1990.

The intonational contours in our materials strongly suggest that these expectations regarding (i) the presence of a H* pitch accent in both complements and (ii) the presence of L tone at the right edge of both complements are borne out. This distribution of pitch accents and edge tones and the consequences for the prosodic structure representation of the various sentence types are discussed in sections 3.1 and 3.2.

<INSERT FIGURE 1 ABOUT HERE>

Figure 1 shows schematic F0 contours for each focus/new condition. The conditions represent differences in the focus/new status of the verbal complements: A = Foc-new; B = new-Foc; C = new-new. The m3 and m6 measurements correspond to F0 peaks on the final pitch accent within Complement 1 and Complement 2 respectively. The peak at m1 corresponds to the F0 peak found in the focus-sensitive particle (e.g. *only)* in the contrastive condition A and B sentences and to the peak within the verb in all-new sentences. The valley measurements m2, m4, m5, m7 were taken at the left- and rightmost edges of the word containing the pitch accent.

3.1 Pitch accents on Complements 1 and 2

The contours in Fig. 1 reflect the generalization that, in all but a tiny number of cases, there is some sort of F0 peak in both complements in all conditions. In particular, 97% of the complements show an identifiable preceding rise and 99% of complements show an identifiable following fall. The high F0 mark in each complement (m3, m6) is on average 0.55 ERB (22 Hz) higher than the values at the left edge of the accented word (m2, m5) and 0.86 ERB higher (34 Hz) than values at the right edge of the accented word (m4, m7). The histograms below, pooled across position and condition, show that the vast majority of all utterances feature a preceding rise and a following fall in F0 surrounding the pitch accented syllable.

<INSERT FIGURE 2 ABOUT HERE>

We take these facts to show that in the phonological representations of the utterances produced in our experiment, a H* pitch accent most often appears in association with the prominent syllable in both complements in all of the conditions, as exemplified in (12):[13]

(12) Complement pitch accents in all conditions (A, B and C)

H* H*

…….. [Manny] [the yellow one]

Comp 1 Comp 2

Of course, the histograms in figure 2 show a range of small and large values, including a few cases where the direction of pitch movement is in the opposite direction of what one would expect from a H* pitch accent. It is possible that some of the complements in our materials bear weak or no pitch accents. Crucially, however, this is only possible for a tiny minority of the data points shown above; if the presence vs. absence of pitch accent were being systematically manipulated to encode the Focus vs. non-Focus contrast, we would expect far more flat and negative values for pitch movement (about a third of the data points), as well as a bimodal distribution in the histograms. These data are thus inconsistent with systematic use of the presence vs. absence of a pitch accent as a Focus-marking strategy.

Note that the repertoire of pitch accents posited in Pierrehumbert 1980, Beckman and Pierrehumbert 1986 and the English ToBI transcription system includes a L+H* pitch accent, which could also potentially describe the pitch peaks appearing in our materials. We make no attempt to distinguish between H* and L+H* pitch accents here. Pierrehumbert & Hirschberg 1990 do claim that the L+H* pitch accent is a marker of

[13] A reviewer notes that declination might result in a fall following stressed syllables even if they did not bear pitch accents. We agree with this point, but note that the large values for many of the falls, in conjunction with the fact that 97% also feature a preceding rise, at least suggest that the vast majority of the complements bear pitch accents.

contrastive focus in English, but studies of corpora have failed to find a systematic correspondence between the Focus status of a constituent and the choice of H* vs. L+H* (Hedberg 2003, Calhoun 2006). Another important reason for not attempting to distinguish between the two is that previous research indicates that the H* vs. L+H* contrast is not reliably distinguished in perception or in production. For example, studies of consistency among trained ToBI transcribers show frequent inconsistency in assignment of H* and L+H* accents in English (Pitrelli, Beckman and Hirschberg 1994; Syrdal and McGory 2000). Moreover, Ladd and Schepman 2003 dispute the H* vs. L+H* distinction on the grounds that, like L+H* pitch accents, pitch accents typically analyzed as H* show the alignment of a low pitch valley at the onset of a stressed H*-bearing syllable. Given uncertainties like these in differentiating H* and L+H*, in our analysis of the data gathered in our experiment we did not undertake transcriptions of the pitch accents of the utterances by independent ToBI transcribers. Instead we have relied on phonetic measurements, like those summed up in Figures 1 and 2. Our data testifies to the consistent presence of F0 peaks on Complements 1 and 2 in all conditions but does not permit any conclusions concerning a putative H* vs. L+H* distinction. We have opted to represent them all with H*, though we could just as well have represented them all with L+H*, as suggested by Ladd and Schepman.

As we will see below, the F0 data make it highly doubtful that a categorical difference in pitch accent type could be based on the magnitude of the pitch rise, or on the steepness of slope that a greater pitch rise produces. In Section 5 we report that the absolute magnitude of the pitch rise and of the height of the pitch peak does not reliably distinguish Conditions A, B and C at either complement. What does reliably distinguish the three conditions is the size of the pitch rise at the first complement *relative to* the pitch rise at the second, as well as the size of the pitch fall at the first *relative to* the fall at the second. These findings provide yet further support for our ultimate decision to represent the pitch accents found in the two complements in terms of the same type of tonal entity, as shown in (12), rather than in terms of a H* vs. L+H* distinction. We will suggest below that it is not a contrast in the tonal representation or the phonetic interpretation of individual pitch accents that correlates with a difference in Focus/new

status, but rather a difference in the phonological prominence (stress) of the pitch-accented words, which is reflected phonetically in the within-sentence relation between the phonetic prominence of these words.

It should be pointed out here that the presence of pitch accents on both complements in Condition A sentences, in which a Focus constituent is followed by a new constituent, is important in establishing a basic fact about the phonology of English sentences containing contrastive focus, namely that it is simply false that material appearing to the right of a focus of contrast is necessarily "deaccented" in English (contra Büring 2008, Xu and Xu 2005, Calhoun 2010). We see that post-Focus discourse-new material does not undergo the necessary "deaccenting" of post-Focus material that is discourse-given (Féry and Samek-Lodivici 2006, Kratzer and Selkirk 2007, Selkirk 2008) [14].

3.2 Phrase-final L tones in Complements 1 and 2

The question we address in this section is whether—regardless of Focus vs. new status—both complements share the additional property of being final within a phonological phrase. We do this by examining the distribution of peripheral L tones in these sentences. If, in all conditions, a L tone is found both sentence-medially, at the right edge of Complement 1, and sentence-finally, at the right edge of Complement 2, this is evidence that the pitch accents in both complements are in phonological phrase-final position. The presence of such a L tone signals the presence of the right edge of a phonological phrase at the right edge of each complement, whether we assume the Beckman and Pierrehumbert 1986 or Gussenhoven 1990 theory of English intonational contours. In the present section we argue that the majority of our materials do indeed show a medial phrase-final L tone and hence show the prosodic constituency given in (13):

[14] While post-Focus deaccenting may be necessary for new material contained within the same phonological phrase as the Focus (Féry and Samek-Lodovici 2006), the current findings demonstrate clearly that "deaccenting" is not found if the new constituent that follows the Focus is located in a distinct phonological phrase.

(13) Complement pitch accents, edge tones, phonological phrases in Conditions A, B, C

H* L-	H* L-	tonal tier
( )$_{\varphi}$	( )$_{\varphi}$	phonological phrase
…….. [Manny]	[the yellow one]	
Comp 1	Comp 2	

At issue is the shape of the low 'valley' between the H* peaks of Complements 1 and 2. The general theory of pitch accents and peripheral tones adopted here allows us to distinguish two rough shapes of valley: a V-shaped valley and a U-shaped valley. A V-shaped valley would be created by a single L target between two H* targets; it would show roughly straight slopes from the point of the L at the valley floor up to the two H* on the flanking peaks. A U-shaped valley would be created by a sequence of two target Ls creating a low level stretch flanked on either edge by rises to the flanking H* peaks. The medial pitch valley between *Manny* and *yellow* in Figure 5 below is plausibly a U-shaped valley; its edges are defined by a right-phrase-edge L and a following pre-accentual L, as in a L+H*. Both the U-shaped valley in Figure 5 and the instances of seemingly V-shaped valley in Figures 3 and 4 suggest the presence of a final L tone at the right edge of in Complement 1. We ask now whether these valley shapes are representative of the data we have obtained.

<INSERT FIGURE 3 ABOUT HERE>
<INSERT FIGURE 4 ABOUT HERE>
<INSERT FIGURE 5 ABOUT HERE>

The schematic pitch contours in Figure 1, based on mean values at seven measurement points, show a sharp fall from the H* peaks of the two complements (m3, m6) down to the right edge of the complements (m4, m7, respectively) in all conditions. They also show that m4 is followed by another low point at m5, which coincides with the left edge of the pitch accented word in Complement 2. Examining the distribution of data points in

histogram form (Figure 6) allows us to ask about the generality of the relation between the low F0 values at m4 and m5.

<INSERT FIGURE 6 ABOUT HERE>

These data provide evidence for a L tone target at m4—the right edge of Complement 1--in a majority of cases in all conditions. As Figure 6 shows, data for pitch movement between the two low measurement points generally include a mix of positive and negative values, with modes close to zero. Although the median case in each condition shows a small amount of downtrend, there are many cases where the F0 value at m4, located at the right edge of Complement 1, is equal to or even lower than the following measurement: such cases constitute 35-43% of the data points in each condition. These data point to either a V-shaped valley with a single L target at m4, or a U-shaped valley in which the first of the L tones that define the U-valley floor appears at m4. As for the cases where m4 is higher than m5, most can also plausibly be analyzed as U-shaped valleys, showing varying degrees of downtrend between m4 and m5. Most cases of downtrend between the two low points are small in magnitude. Median values range from 0.06 to 0.18 standard deviations depending on the condition. These values are an order of magnitude smaller than, for instance, the pitch fall following complement 1, which ranges from 0.79 to 1.17 standard deviations across conditions. In sum, data from all three conditions suggest the presence of a sentence medial L tone that is final in Complement 1 in a majority of cases. This in turns supports the hypothesis of a phonological phrase break between Complement 1 and Complement 2 in a majority of cases, regardless of the Focus or new status of the complements, as indicated in (13).

Consider an alternative hypothesis concerning the pitch valley between the complements, one which posits not a L tone at the right edge of Complement 1, but instead a sequence of a H* pitch accent in Complement 1 followed by a L+H* pitch accent on Complement 2. This hypothesis is consistent with the averaged data presented in the schematic pitch contours in Figure 1, though only if the low pitch values at position m4 preceding the m5 minimum can be ascribed to an asymptotic phonetic interpolation between a H* tone and

a following L+ tone, one which would create the radically concave fall shape seen in the m3-m5 sequence in Figure 1. But Figure 6 shows that a H* plus L+H* sequence can't account for the full range of data; the H* plus L+H* proposal (even with asymptotic interpolation) is inconsistent with cases where m4 is lower than m5. Thus, although the materials are consistent with there being some variation with regard to these phrasing properties, we posit a medial phonological phrase break in the majority of cases. And crucially, the three conditions don't vary much with regard to the scaling of low F0 points. Figure 6 shows that all three feature a range of positive and negative values for downtrend, with modes close to zero. There are a few more cases of large downtrend values in condition C than the other conditions, but this difference amounts to 7 or 8 cases out of 170. The data therefore support similar phonological phrase analyses for all sentence types. In other words, although there may be variation in the presence of medial phonological phrase boundaries in our materials, speakers are not systematically exploiting this property to mark the Focus vs. non-Focus distinction.

3.3 The tone(s) preceding Complements 1 and 2

Sentences from conditions A, B, C do differ in the part of the intonational contour that precedes the complements. In Condition C, a subject pronoun and a verb precede the first and second verbal complements; the initial F0 measurement of the sentence (m1) was taken at the main-stressed syllable of the verb, which would bear any pitch accent that might appear. Inspection of pitch tracks in these cases indicates that the verb may or may not carry a pitch accent. In the utterance in Figure 3 the verb carries a pitch accent.

The sentences in Condition A and B included a focus-sensitive particle like *only,* situated between the subject pronoun and the verb. The initial high pitch peak of the sentences in Conditions A and B typically fell on that particle; in these conditions, the initial F0 measurement of the sentence (m1) was taken on the stressed syllable of the particle. In Figures 4 and 5 we have transcribed the pitch peak on the particle as H*. That initial H* peak is followed by a gradual (nonasymptotic) descent to a low target immediately preceding the H* pitch accent of Complement 1; this contour speaks to the absence of

any additional L tone immediately following the initial H* on the particle. Note that, as Figure 1 shows, the m1 pitch values on the focus particles in Conditions A and B were, on average, quite a bit higher than the m1 pitch value on the verb in Condition C sentences. The authors' impression in listening to the recordings was that speakers were choosing values for the pitch of the H tone on the focus particles that were especially high in their pitch range.

3.4 Summary

On the basis of the pitch contours of sentences in the three different Focus/new conditions in our study, we have established that the Focus and new complements share two essential phonological properties. The vast majority show a pitch peak, represented here as a H* pitch accent. And this H* immediately precedes a phrase-final L tone in a majority of cases, indicating that the complement phrases in our materials, whether sentence-medial or sentence-final, tend to be final in a phonological phrase. The fact that (putative) Focus and new complements in all conditions fail to differ systematically with regard to these properties means that any differing patterns of quantitative phonetic prominence exhibited by the Focus-new, new-Focus and new-new sequences cannot be explained in terms of presence vs. absence (or type) of pitch accent or in terms of medial vs. final position within a phonological phrase. Rather, they must be due to the Focus vs. new status of the complements themselves, or to some other aspect of phonological representation which would mediate between this Focus/new representation and its phonetic realization.

4.0 Focus Duration

We now turn to the question of whether a Focus constituent is distinguished from a new constituent by a greater degree of phonological or phonetic prominence. In this section we ask whether a Focus constituent displays greater duration when a comparison is made of Focus-new (A), new-Focus (B) and new-new (C) sentences. Our materials also allow us to ask whether the presence or location of Focus in a sentence has any effect on the

duration of surrounding constituents in the same sentence. We are able to make a controlled investigation of these questions because the constituents preceding and following the putative Focus in our materials are discourse-new and so have the same information structure status as the corresponding context constituents in the all-new sentences (cf. discussion in section 1), and because these context constituents also have the same prosodic structure across all conditions (as shown in section 3).

Section 4.1 reports on the duration findings themselves: a three-way distinction in phonetic duration patterns across the three sentence types. These findings support the claim that the theory of grammar (i) makes a distinction between contrastive Focus and discourse-new and (ii) allows for that distinction to be reflected in phonetic interpretation. In section 4.2 it is suggested that differences in the phonological representation of phrasal prominence (stress) in the three different sentence types are responsible for the three-way durational distinctions found in the phonetics.

4.1 Duration Findings

Shown below in Table 2 is the duration of the measured stretch of the two complements in each condition.

<INSERT TABLE 2 ABOUT HERE>

Recall that the measured stretch of each complement consisted of a string containing the pitch accented syllable and one or two adjacent syllables. For the sentences of any minimal triplet, the measured string was identical. (For details see section 2 and Appendix B.) In the statistical analysis that follows Complement 1 and Condition C are the baseline levels for position and condition that other levels of those variables are compared to. If Focus status is reflected in increased duration, we would expect that the relative duration of the two complements should differ amongst the three conditions, because these conditions differ in their Focus properties. Figure 7 shows the difference in duration between the first and second measured strings in each condition.

<INSERT FIGURE 7 ABOUT HERE>

Complement 1 is on average 35-40 ms. longer in condition A, where it is a Focus, than in the other two conditions, where it is new. This amounts to an 11% lengthening effect of Focus over new. The difference in Complement 1 duration between A and the control condition C is significant: $\beta = 0.51$, $p < 0.001$. Note that the difference in duration of the discourse-new Complement 1 in B and C, 2% of a standard deviation, is not significant. The relative pattern between conditions changes at Complement 2. Conditions A and C, which were significantly different at Complement 1, are roughly equal at complement 2. Conditions B and C, which were roughly equal at Complement 1, are different at Complement 2. Complement 2 is 25-30 ms. longer in condition B, where it is a Focus, than in either of the other two conditions, where it is new. This amounts to a Focus lengthening effect of about 10%. These differences in relative duration by complement and condition are reflected in significant interactions between position and condition: for B vs. C, $\beta = 0.40$, $p < 0.001$; for A vs. C, $\beta = -0.55$, $p < 0.001$.

A central finding is the greater duration of the Focus constituent relative to new constituents in the corresponding positions in other conditions. This finding supports the claim that the grammar treats Focus constituents differently from new constituents. The magnitude and direction of these effects do not differ significantly by subject or item, showing that they are statistically-robust generalizations.

As we noted earlier in section 1.2, the hypothesis that a Focus element receives greater phonetic prominence than a merely discourse-new element entails that the within-sentence relationship between the 1st and 2nd complement differs across the three conditions. This is exactly what the significant interactions between position and condition show. The difference in duration between the first and second complements is largest in condition A, intermediate in condition C and smallest in condition B.

Overall, the durations at Complement 2 are smaller than those at Complement 1: $\beta = 0.84$, $p < 0.001$. This simply reflects the fact that the strings measured in Complement 2 were shorter than those measured in Complement 1, because the utterance-final syllable was generally excluded from the measured sequence (cf. section 2). The exact size of this difference differs between speakers and between items; incorporating those differences into the model significantly improves fit: $\chi^2 (2) > 25$, $p < 0.001$ for both effects.

To summarize, our experiment has revealed a three-way distinction in duration patterning that correlates with a three-way distinction in Focus/new status, as seen in Figure 7. It is the significantly greater duration found with Focus constituents that has the result that the three sentence types Focus-new, new-Focus and new-new are distinguished in this aspect of their prosody. The duration of corresponding new constituents in the different sentence types is virtually constant, not affected by the presence or absence of a Focus constituent in the same sentence[15].

4.2 Implications of Focus duration findings for the grammar

4.2.1 Syntactic representation of contrastive focus vs. discourse-new

If the theory of grammar made no distinction between contrastive Focus and discourse-new and instead represented both of the complements in our sentence types in the same fashion, e.g. as F-marked, there would be no possible way to derive the observed three way distinction in duration patterning. Theories of focus and focus prosody which posit a common F-marking for both contrastive focus and discourse-new (e.g. Selkirk 1984,

[15] A strictly local effect of Focus lengthening is also reported by Cooper et al (1985) comparing narrow focus sentences with broad. Eady et al (1986) report a small but significant difference in duration between context-providing Given constituents in narrow focus sentences and context-providing New constituents in broad focus sentences, but only when these Given/New constituents are located in sentence-final position. This latter finding need not imply a syntagmatic effect of Focus lengthening, given that no such syntagmatic effects are observed in our contrastive vs. all-new comparisons. Rather, it may simply reflect a local paradigmatic difference in the durational properties of New vs. Given constituents. Clearly more work on these matters remains to be done.

1995, Schwarzschild 1999, Büring 2007) are committed to an undifferentiated representation like (14) for all three conditions under discussion here:

(14) Undifferentiated representation for contrastive focus and discourse-new:

Conditions A, B, C: [….. […]$_F$ […]$_F$]

Since (14) would provide the same representation to the phonological interface for sentences in all three conditions, their phonological and phonetic properties would wrongly be predicted to be identical.

Rather, these duration findings from English provide support for two theoretical proposals: (I) linguistic theory provides a grammatical representation for contrastive focus constituents that is distinct from that for discourse-new constituents; and (II) the grammar of a language may give a prosodic treatment to contrastive focus constituents that is distinct from that of discourse-new constituents. With respect to (I), following most previous work in generative grammar on the semantics and syntax of focus, it will be assumed that the syntactic representation of a sentence, which mediates between sound and meaning, contains a representation of the focus status of syntactic constituents. A three-way difference in the syntactic representation of the sentences in our Conditions A, B and C as in (15) is accordingly hypothesized:

(15) Differentiating Focus and discourse-new in syntactic representation:

Condition A: [….. […]$_{Focus}$ […]]

Condition B: [….. […] […]$_{Focus}$]

Condition C: [….. […] […]]

(Discourse-new constituents have been left syntactically unmarked in (15), as suggested in section 1.1.) The grammar of English then gives a specific characterization to the relation between the distinct syntactic representations in (15) and the phonetic durational patterns that correspond to them (cf. Figure 7), as envisaged in general terms by (II).

4.2.2 A phonological distinction between Focus and discourse-new

Consider now the assumption made widely in generative grammar that phonetic interpretation is "syntax-blind" and accesses only phonological representation (Chomsky & Halle 1968). This assumption narrows considerably the set of hypotheses that can be entertained about the nature of the possible relation(s) between syntactic representations like (15) and their phonetic interpretation. It divides the theory of that relation into two parts: (IIa) a theory of the relation between syntactic and phonological representations, one that allows differences in phonological form to be assigned depending on the Focus-marked status of a syntactic constituent, for instance; and (IIb) a theory of phonetic interpretation that is sensitive to those (and other) differences in phonological representation. In what follows we pursue the consequences of adopting the specific assumption that phonetics is syntax-blind and that a mediating phonological representation does enter into the definition of the Focus-phonetics relation. The question, then, is what the character of that mediating phonological representation is.

In section 3, we showed that pitch accents are located on both complements in the vast majority of cases in all three conditions, and that in all three conditions, there is a strong tendency for the pitch accented word in both the complements to be the last one in a phonological phrase. So neither presence of pitch accent or phrase-final position can be the aspects of phonological representation that are responsible for the duration difference between Focus and new constituents. But positing differing patterns of prosodic prominence (stress) in the phonological representation could predict the durational differences amongst the pitch accented sequences in our sentence types. This explanation seems plausible, given that stress at lower levels in English is reflected in duration (see Fry 1955, Huss 1977, Beckman 1986, Okobi 2006, inter alia, for the effects of stress on

syllable duration in English). Moreover, there is ample precedent for assuming a phonological representation of phrasal prominence patterns in English and in particular for assigning a Focus constituent greater prosodic prominence (stress) than any non-Focus within the same domain (see references in first paragraph of section 1.2). We need to posit three distinct prominence representations in the phonology that would mirror the different syntactic distributions of Focus in (15), in accordance with IIa. The phonetic component of English is hypothesized to assign greater duration to elements that bear greater abstract phonological prominence (stress)[16], in accordance with IIb. In this way, the quantitative phonetic patterns of Focus-related duration in English that are seen in Figure 7 would be accounted for.

Assuming, then, that a phonological representation of prosodic prominence is the basis for the durational patterns in Figure 7, the specific prosodic prominence representations in (16abc) could be posited for the three distinct cases. The Focus constituent in first and second complement position in (16a) and (16b), respectively, carries greater prominence than a new constituent in the same sentence. In (16c)—Condition C-- the two new constituents in sequence bear the same, lower, level of prominence.

(16) Prosodic Prominence Representations in Conditions A,B, C: A three-way contrast

a. Condition A: Foc-new

(x)	Intonational Phrase
(x) (x)	Phonological Phrase
(x) (x) (x) (x)	Prosodic Word
[He even took [Minnie]$_{Foc}$ to a [Mariners game]]	
H* H* L- H* L-	

[16]Works addressing the relation between phonological and phonetic prominence include van Heuven and Sluijter 1996, Keating and Shattuck-Hufnagel 2002, Keating 2003, Choi et al 2005.

b. Condition B: new-Foc

(x)	Intonational Phrase
(x) (x)	Phonological Phrase
(x) (x) (x) (x)	Prosodic Word
[He even took [Minnie] to a [Mariners game]$_{Foc}$]	
H* H* L- H* L-	

c. Condition C: new-new

( )	Intonational Phrase
(x) (x)	Phonological Phrase
( ) (x) (x)	Prosodic Word
[He took [Minnie] to a [Mariners game]]	
H* L- H* L-	

In these representations the height of the columns of stacked x-marks indicates the level of prosodic prominence (stress)[17]. In prosodic structure theory, a tight connection is made between prosodic constituent structure and prosodic prominence. The representations in (16) contain two levels of prosodic phrasal constituency-- the intonational phrase and the phonological phrase, a standard assumption (Selkirk 1986, Nespor & Vogel 1986, Hayes 1989). Each prosodic constituent immediately dominates at most one prominent lower constituent, its prosodic head (Selkirk 2007); the prosodic head, marked with an x, is the locus of what's typically referred to as the main stress of the containing constituent. Since nearly all phrasal complements in all conditions in our materials are pitch accented and since there is a strong tendency for both complements to be phonological phrase-final (cf. sec. 3), we hypothesize that the majority carry main, head, prominence at the phonological phrase level . Assuming this baseline phonological phrase-level prominence for all complements, we ascribe the even greater prosodic prominence of the Focus complement in Condition A (Focus-new) and Condition B

[17] This representation of prosodic constituency and prominence has been called a bracketed grid representation (Halle and Vergnaud 1987, Hayes 1995).

(new-Focus) to the status of the Focus as the prosodic head of the higher level intonational phrase constituent, hence the hypothesized representations in (16ab)[18].

In the phonological representations in (16abc), the Focus status of constituents is relevant only to prosodic prominence at the intonational phrase level. Most other aspects of these representations are determined by grammatical principles that make no appeal to information structure status. The pitch accenting that is common to Focus and discourse-new constituents is arguably a phonological fact: phonological phrase-level stress is associated with pitch accent in English. As for presence of phonological phrase-level stress, it's a matter of phonology, too, a result of the high ranking of the phonological constraint that calls for a phonological phrase to be prosodically headed. Finally, the presence of that phonological phrasing in the first place is determined by general principles of the interface between syntactic and prosodic constituency (Selkirk 2011). In other words, on this account, the phonological representation of discourse-new constituents is a matter of default phonology, not information structure.

Note that no head prominence of intonational phrase is posited in (16c) for the all-new Condition C sentence, contrary to common thinking on the prosody of all-new sentences in English, which has ascribed to all sentences a nuclear (most prominent) stress (Chomsky and Halle 1968, Selkirk 1984, 1995, Cinque 1993, Ladd 1996, 2008, Kahnemuyipour 2009). (16c) is posited here for Condition C sentences because the complements in Condition C in our experiment, both new, turn out to have virtually identical duration to the new complements in the corresponding syntactic positions in the Focus sentences in Conditions A and B (see Table 2). This fact could not be explained if either of the new complements in Condition C were instead to bear intonational phrase-level prominence (stress), as posited here for Focus constituents, in addition to the baseline phonological phrase-level prominence.

[18] The requirement that a Focus prominence carry maximal prominence within its domain (Truckenbrodt 1995) predicts that Focus will necessarily bear intonational phrase-level prominence only if there are other phonological phrase-level prominences within the same domain.

If this prosodic prominence-based account of the phonetic implementation of duration in our materials is correct, it has implications both for the general theory of prosodic prominence and for the analysis of the prosody of all-new sentences in English as compared to sentences containing Focus. Each of the phonological phrases in (16c) carries a "nuclear" main stress at the phonological phrase level, but the intonational phrase corresponding to the sentence as a whole does not contain its own, distinct, IP-level nuclear stress. Positing the representation in (16c) amounts to the claim that it is not a theoretically necessary property of prosodic structure representations that every prosodic constituent contain a single, most prominent, main-stressed, nuclear constituent-- its prosodic head. A general investigation of this claim, involving all levels of the prosodic hierarchy, is clearly needed; unfortunately limitations of space do not allow us to pursue this issue here. As for the specific claim—that sentence-level main stress may fail to be defined in English all-new sentences (like 16c))— it goes counter to the intuitions of generations of linguists who have agreed with Chomsky and Halle 1968 that the final constituent of a 'neutral' (= all-new), typically declarative, sentence of English carries the highest level of stress in that sentence. But there are other potential explanations for this intuition. One is that the reliable presence of phonological phrase-level stress within any sentence is the source of the intuition of a greater righthand stress prominence in the sentence. Another is that the intuition of "greatest prominence " or "highest stress" is due to the greater size and perceptibility of the pitch movements that appear in the sentence-final position that is standardly associated with main sentence stress. For example, the effect of intonational phrase-final pitch lowering found in English (Liberman and Pierrehumbert 1984, Choi et al 2005) can result in large and perceptible final pitch falls in declarative sentences, and could thereby make the final constituent more salient. Systematic research on the phonetic basis for perceptions of "main sentence stress" in English is clearly needed.

4.2.3 Summary

The purpose of the current section has been to bring standard theoretical assumptions from generative grammar to bear on possible accounts of the prosody of the Focus-new

distinction. Assuming that the basic grammatical architecture includes a syntax-blind phonetics, we have looked to the phonological representations which correspond to the syntactic representation of Focus and/or new constituents as the input to the phonetic interpretation of duration. Among the components of prosodic representation—tone, prosodic constituency, prosodic prominence (stress)—only the latter can provide the distinct representations that are required for the phonetic interpretation, given that the sentence types examined in each of the three conditions do not appear to differ in their tonal or phonological phrasing properties. The hypothesis we are entertaining, then, is that contrasts in prosodic (stress) prominence representation—specifically those in (16abc)—mediate between the Focus/new articulation of the sentence and the phonetic implementation of duration. This hypothesis allows an account of the data at hand, and opens up some very interesting questions for further research. In sections 5 and 6, we will see data concerning the pitch and intensity of Complements 1 and 2 that also show a three-way distinction in phonetic patterning depending on Focus/new status. This is consistent with the three-way distinction in phonological representation posited in (16).

It should be said that greater phonetic prominence of Focus as compared to discourse-new is not a universal; some languages show no prosodic prominence of Focus at all (e.g. Wolof (Rialland and Robert 2001), Yucatec Maya (Gussenhoven 2006, Gussenhoven and Teeuw 2008, Kügler and Skopeteas 2006, 2007), Chichewa, Chitimbuka and Durban Zulu (Downing 2008), Salish (Koch 2008)). The difference between English and these other languages can be understood as a difference in whether, or just how, particular languages give phonological expression to the morphosyntactic Focus feature that marks contrastive focus constituents in syntactic representation. Grammatical principles like that which governs the Focus-phonology interface in English, which specifies that a Focus-marked constituent in the syntax must carry the maximal prosodic (stress) prominence within some relevant domain (e.g. Truckenbrodt 1995), must be language-particular, instances of the language-particular spell-out of morphosyntactic features.

5.0 Focus Pitch Scaling

In this section we examine the F0 contours of Focus-new and new-Focus sentences as compared to new-new sentences. The specific question we are asking is whether or not the effects referred to as "Focus pitch boost" and "post-Focus pitch compression" that have been claimed to exist on the basis of earlier comparison of narrow focus (Focus/Given) and broad focus (all-new) sentences in English (e.g. Xu and Xu 2005) are supported by comparisons of the pitch patterns that we find in different conditions in our experimental materials. Section 5.1 reports the findings. Section 5.2 suggests that the distinct F0 patterns observed in the different sentence types receives an insightful account if we assume that a phonological representation of prosodic stress prominence like that in (16) is responsible for the Focus-related pitch boost and post-compression effects.

5.1 Pitch findings

The Focus pitch boost hypothesis holds that a Focus element should have heightened pitch, while the post-Focus pitch compression hypothesis holds that the stretch of the utterance following a Focus should have lowered pitch. If these effects hold outside of the Focus plus discourse-given contexts in which they have been reported (Cooper et al 1985, Eady and Cooper 1986, Xu and Xu 2005, Breen 2007, Breen et al 2010), the most straightforward prediction for our study is that the *relations* between the pitch profiles at Complements 1 and 2 *within the same sentence* should be distinct for sentences in Conditions A, B and C, where the composition in terms of Focus and new is varied. Because our materials are sentences in which both target constituents within a sentence are pitch accented and tend strongly to be prosodic phrase-final in all conditions, we are able to test this prediction. The measures we used to test this prediction are in (23):

(23) Testing for within-sentence differences in peak profiles as a function of Focus/new

A. The differences in magnitude between (a) the F0 rise at the left edge of first complement and the rise at the edge of the second complement and (b) the F0 fall

at the right edge of the first complement and the fall at the edge of the second complement.

B. The difference between the F0 at the peak of the first complement (m3) and the F0 at the peak of the second (m6).

Two mixed effects regression models were constructed to test differences amongst the conditions for these two dependent variables. If the predictions about relative pitch profiles for complements within the same sentence are correct, we would expect both condition A and condition B to show significant interactions with position relative to condition C. In other words, differences between the first and second complements should vary by condition. Broadly, we found that the pitch-excursion model reliably distinguishes all three conditions. The peak-difference model, on the other hand, reliably distinguishes condition B from the baseline condition C, but not A from C.

Section 5.1.1 examines the findings concerning pitch excursions (rises, falls) in Complements 1 and 2; section 5.1.2 examines findings concerning the height of pitch peaks and valleys in the complements. Section 5.1.3 looks at the effect of the pre-complement pitch peak, and section 5.1.4 reports on assorted remaining effects.

5.1.1 Patterns involving complement pitch excursion—relative and absolute

Figure 8 shows condition-dependent differences in the relation between the pitch excursion sizes at the two complements. The graph displays ratios between the excursion magnitudes at Complement 1 and Complement 2 for each condition, averaged across subjects. We observe a three-way distinction in relative excursion magnitude, following the pattern A > C > B. Precisely this pattern of difference in the three conditions is what we expect given the assumption that Focus items show greater pitch excursion relative to surrounding material than merely new items do, along with the assumption that there is greater pitch compression following Focus. In Condition A, the effect of Focus pitch boost is to increase pitch excursion at Complement 1 (Focus) and the effect of post-Focus

pitch compression is to diminish pitch excursion at Complement 2 (new), with the result that the excursions at Complement 1 are considerably larger than at Complement 2. In Condition B, there would be Focus pitch boost at Complement 2, leading to the result that the Complement 2 (Focus) pitch excursion would be larger than at Complement 1 (new), reversing the pattern seen in Condition A. In Condition C, neither Focus pitch boost nor post-Focus pitch compression are at play; in this default, baseline, condition, pitch excursions at Complement 1 (new) are somewhat larger than at Complement 2 (new).

<INSERT FIGURE 8 ABOUT HERE>

Excursion magnitude at Complement 1 in condition A is larger than at Complement 2 in the same condition: 0.22 ERB (10 Hz) larger on average for rises and 0.33 ERB (15 Hz) for falls. In condition B the *second* complement excursions are slightly larger than the first: 0.10 ERB (3 Hz) larger for rises and 0.20 ERB (6 Hz) for falls. For condition C the first rise is 0.13 ERB (8 Hz) larger than the second and the first fall is 0.11 ERB (8 Hz) larger than the second. For both rise and fall measures, the relationship between the three different conditions for difference between first and second complement excursion is the same: A > C > B. Note that this pattern holds for every subject.

Confirming the relational observations in the preceding paragraphs, the pitch-excursion model finds significant interactions between condition and position for both Condition A and Condition B relative to the baseline Condition C. These results concern a difference between differences: the difference between excursion magnitude at the first complement and excursion magnitude at the second complement is different in the three conditions. Condition B shows significantly smaller excursion magnitude at the first complement than the baseline condition C: $\beta = -0.19$, $p < 0.001$. At the second complement, this pattern reverses, resulting in a significant interaction between condition and position: $\beta = 0.32$, $p < 0.001$. Condition A shows a somewhat greater excursion magnitude at the first complement than baseline condition C, about 4% of a standard deviation, but the difference does not reach statistical significance. This pattern also reverses at the second complement, resulting in a significant interaction between condition and position: $\beta = -$

0.17, $p < 0.001$. This means that both condition A and condition B differ from condition C (in opposite directions) in the relationship between excursion magnitude at the first complement and the second complement. These results confirm the first of the predictions outlined above.

Note that we found that comparison of the absolute magnitude of the rises and and falls at complements in corresponding syntactic contexts does not reliably distinguish all three conditions at either complement, even though the means are consistent with the generalization that rises up to, and falls down from, a Focus complement are larger, in absolute terms, than those associated with a non-Focus complement. It's conceivable that contrasts in absolute rise or fall excursion size across the different conditions would come out significant if we had results from more subjects. But closer inspection of the data suggests that absolute magnitude is a less reliable indicator of Focus vs. new status than the relations between pitch profiles within the same sentence. For each of the rise and fall measurements, only a subset of subjects show the inter-sentence effect *Foc* > *new*. This is illustrated for just one measurement in Figure 9:

<INSERT FIGURE 9 ABOUT HERE>

This between-subject variability in absolute measures of F0 excursion, exemplified in Figure 9, contrasts with the findings concerning within-sentence pitch excursion relations, which were robust across all subjects.

To sum up, the data concerning pitch excursions validate the general prediction made by Focus pitch boost and post-Focus pitch compression for the sentence types in our study, namely that the *relations within the same sentence* between the pitch profiles at Complements 1 and 2 should be distinct for sentences in Conditions A, B and C, which vary in the Focus vs. discourse-new status of the complements.

The magnitude of virtually every fixed effect of interest (i.e., condition, position, and interactions), but crucially not the direction of these effects, varied substantially between

both subjects and sentence frames; in the mixed-effects models used here, this variance results in a large number of significant *random slopes* in both models. This means that, broadly speaking, models that allow a separate value (or slope) for each subject/frame sentence for these effects fit the data better than models that assume all subjects/frame sentences show the same size effect. Recall that the fixed effects of interest, however, are significant; this means that, even after taking into consideration the considerable variability subjects and frame sentences display in this regard, a model with those fixed effects included is significantly better than one without them.

5.1.2 Patterns involving complement pitch peak height—relative and absolute

We turn now to data on the height of pitch peaks and valleys. Our materials provide visible support for the existence of a Focus-related pitch-lowering effect. Post-Focus pitch compression is seen clearly in values for the low targets (m4 and m5) located between the two peaks in Complements 1 and 2 in Condition A, where Complement 1 is Focus (cf. Figure 1). The mean F0 values at m4 and m5, which define the valley following the first complement peak, are significantly lower in condition A than in condition B (m4: $\beta = 0.19$, $p < 0.001$; m5: $\beta = 0.10$, $p < 0.01$).

We turn now to consideration of the absolute and relative heights of the pitch peaks in Complements 1 and 2. The height of the Focus-marked Complement 1 peak in condition A is significantly lower than the new Complement 1 peak in the baseline condition C: $\beta = -0.23$, $p < 0.001$. As for within-sentence peak-peak relations, the amount of downtrend between the peaks at Complements 1 and 2 is, as expected, larger in condition A than in condition C, about 9% of a standard deviation, but this effect does not reach statistical significance. In condition B, the absolute height of the new Complement 1 peak is significantly lower than that of the new Complement 1 in condition C ($\beta = -0.35$, $p < 0.001$). As for within-sentence peak-peak relations, the amount of downtrend is smaller in the B condition (where Complement 2 is Focus) than in the C condition (where it is new). This results in a significant interaction of position and condition: $\beta = 0.37$, $p < 0.001$.

The significant difference between the (within-sentence) peak-peak downtrend relation of condition B vs. that of condition C suggests that relative pitch peak height, like relative pitch excursion, distinguishes the different Focus/new conditions. It is puzzling, then, that there is no significant difference between condition A and condition C in this regard. Moreover, the data on absolute pitch peak values of Complement 1 show two unexpected patterns: (i) the Focus constituent in Complement 1 in condition A is significantly lower than the new constituent in condition C (contrary to what Focus boost would seem to predict) and (ii) the new constituent at Complement 1 in condition B is significantly lower than the new Complement 1 constituent in condition C (while we expect them to be the same). It is possible that the solution to these puzzles is to be found in data not yet considered, namely the value of the F0 peak at m1, the pitch peak which immediately precedes the pitch peak at Complement 1 (cf. Figure 1). We turn to this data now.

5.1.3 Patterns involving the sentence-initial pitch peak

In conditions A and B the first pitch peak in the sentence falls on the focus particle; in condition C it falls on the verb. As Figure 1 shows, the mean F0 values of the m1 peak on the focus particle in conditions A and B (5.66-5.71 ERB; 202-205 Hz.) are much higher than on the verb in condition C (5.18 ERB;178 Hz.). It's plausible that this greater F0 protrusion on the focus particle is responsible for the lower-than-expected values of the pitch peaks at Complement 1 in Conditions A and B.

The presence of pitch compression/lowering following the focus-sensitive particles in Conditions A and B is confirmed by looking at static F0 measurements at various points in the sentence. The gross generalization is that all measurement points following the sentence-initial focus-sensitive particle in conditions A and B are lower than the corresponding points in condition C. (The lone exception is the peak of the focused second complement in condition B, which is nearly identical to condition C.)

What we are seeing in condition A and B sentences is arguably a pitch scaling effect independent of the lowering seen in post-Focus pitch compression. This may be the effect

documented by Gussenhoven et al (1997), who propose that the abstract reference line according to which F0 protrusion is calculated is made steeper when the initial F0 values in an intonational phrase are higher. We'll dub this the Initial High Effect. A steeper decline in the reference line after the initial super-high H tone associated with the focus particle in the A and B conditions could explain the overall lower pitch values throughout the sentence in these conditions. In particular, it solves the puzzle of why the Complement 1 pitch peak is higher in Condition C than in Conditions A and B. As for the remaining puzzle-- that the downtrend between the pitch peaks of the complements in Condition A is not significantly greater than in Condition C— this will have to remain a question for further research.

5.1.4 Assorted other patterns observed

We find a large positive correlation with the F0 of the first measured point m1 for both dependent variables: $\beta > 0.1$, $p < 0.001$ for both models. In other words, utterances that begin at a relatively high F0 have a tendency to stay at a relatively high F0. This indicates that m1 is serving as a control for the overall pitch register of individual utterances, helping factor out variance due to between-utterance differences in order to capture generalizations across utterances.

The excursion model included a fixed effect, Measurement, encoding the distinction between rise measurements and fall measurements. The effect of Measurement was significant: rises tend to be of smaller magnitude across the board than falls: $\beta = -0.35$, $p < 0.001$. This effect did not interact significantly with condition or position.

The magnitude of excursion measurements is somewhat larger on the second repetition within each experimental session: $\beta = 0.10$, $p < 0.01$. We have no explanation for this.

5.1.5 Summary

The findings reported here confirm the predictions of Focus pitch boost and post-Focus pitch compression for pitch patterning in our three different sentence types:

(17) Predictions of Focus pitch boost together with post-Focus pitch compression:
Relations between pitch profiles at Complements 1 and 2 *within the same sentence* are distinct for Conditions A (Foc-new), B (new-Foc), C (new-new).

In particular, differences in excursion magnitude are largest for condition A (with Focus-marked, IP-prominent, Complement 1), smallest for condition B (with Focus-marked, IP-prominent, Complement 2), and intermediate for condition C (with neither Focus-marking nor IP-prominence on either complement). The results concerning the relation between the heights of the pitch peaks on Focus and/or new within the same sentence are less clear. The amount of sentence-internal downtrend between the pitch peaks on Complements 1 and 2 is significantly different between conditions B (new-Foc) and C (new-new), but not between conditions A (Foc-new) and C (new-new), although the A vs. C comparison does go in the right direction. These distinctive pitch scaling effects emerge despite the confounding role of the pre-complement pitch peak, whose high F0 values in Conditions A and B correlate with greater compression of following F0 values. Taken together, our findings for pitch excursion and pitch peak height provide support for the generalization in (17).

5.2 Implications of pitch findings for the grammar

As with the duration findings, the F0 findings provide support for the hypothesis that the theory of grammar makes a distinction between the contrastive Focus and discourse-new status of constituents in a sentence. In 4.2 it was proposed that this distinction is represented through the presence of Focus-marking on contrastive focus constituents in syntactic representation and the absence of any such marking on discourse-new, as in (15abc). It was also hypothesized in 4.2 that phonetic interpretation does not directly access this grammatical Focus-marking, but rather that the Focus-phonetics relation is indirect, mediated in English by the phonological representation of prosodic prominence.

By raising the hypothesis of prosodic prominence-sensitive pitch scaling in this paper, we hope to encourage research that will help to better understand whether or not phrasal prosodic prominence (stress) should be considered, along with tone and prosodic constituency, in a theory of the phonetic interpretation of the phonological representation of sentences in a language like English. At this time, relatively little research on the phonetic effects of (putative) phrasal prosodic prominence has been done (but see van Heuven and Sluijter 1995 et seq, Turk and White 1999, Keating 2003 and Choi et al 2005 for reports of increase in duration with presence of accent[19]). The claim that phonetic pitch scaling (boost and compression) is directly Focus-sensitive and the claim that it is prosodic prominence-sensitive (and only indirectly sensitive to Focus) have very different theoretical and empirical implications. Most notably, the hypothesis that phonetic pitch scaling is prosodic prominence-sensitive predicts the within-sentence relational findings we have documented in section 5.1. The phonological representation of prominence itself encodes differences in the level of the abstract prominence assigned to distinct constituents of a sentence (as in (16abc)). And the most straightforward formulation of any phonetic principles of prosodic prominence-sensitive pitch boost and post-prosodic prominence pitch compression calls for the phonetic interpretation to boost and compress pitch in quantitative degrees that would reflect the respective levels of prosodic prominence, within some relevant domain. Such a phonology/phonetics interface homomorphism for prosodic *prominence* has a straightforward analogy in a documented homomorphism involving prosodic *boundaries* at different hierarchical levels. The gradient phonetic properties at work in marking prosodic constituent boundaries (edges) and the gradient likelihood of observing or blocking allophonic phonetic variation across these boundaries, such as fortition/lenition effects, mirror the strength or level of these boundaries (Wightman *et al*. 1992, Dilley *et al*. (1996), Byrd & Saltzman (2003), Féry (2003), Keating (2003), Choi *et al*. (2005), Yoon *et al*. 2007). By contrast, no such homomorphism would be definable between the syntactic or semantic representation of

[19] If the presence of pitch accent does indeed correspond to the presence of prosodic phrasal prominence in English, then accentual lengthening may simply be stress-driven lengthening.

Focus and degree of phonetic prominence. In other words, any directly Focus-sensitive pitch scaling would simply have to stipulate the presence of greater phonetic prominence on the Focus than on any other element in the relevant domain.

Another important point is that a theory of prosodic prominence-sensitive pitch scaling predicts that instantiations of prosodic prominence that are *not* Focus-related could, in principle, have phonetic effects similar to the boost and post-compression effects found with the intonational phrase-level prominence that corresponds to Focus constituents in our materials. We are not in a position to undertake here an investigation of the predictions for English, or other languages, of generalized operations of prosodic prominence pitch boost and post-prosodic prominence pitch compression. Though it is worth considering the possibility that the representation of phonological phrase-level prominence that was posited in (16c) for the all-new Condition C sentences might explain the significant pitch downtrend observed in this sentence type in our materials. If the phonological phrase-level prominence posited for Complement 1 in Condition C induces prosodic prominence-sensitive pitch boost and post-prosodic prominence pitch compression, we predict such pitch downtrend. Truckenbrodt 2004 reports that in German all-new sentences, the F0 of a phrase-final pitch accented word is raised and what follows is downstepped. He proposes that this F0 raising and downstepping are two sides of the same coin, both rendering the accented word more phonetically prominent. We would simply add here that the phonetic prominence Truckenbrodt observes is arguably a reflex of phonological phrase-level prominence.

6 Focus Intensity

Intensity is the final dimension of phonetic prominence that we undertook to investigate in relation to the variations in Focus/new status of constituents in the sentence types examined in our experiment. The findings mirror those for pitch.

6.1 Intensity findings

Shown below are the intensity values measured at each of the two complements in our sentence types, averaged over subjects:

<INSERT FIGURE 10 ABOUT HERE>

Just as with F0, both condition A and B differ significantly from baseline condition C (in opposite directions) with regard to the relative intensity of the first and second complement, but absolute intensity is not sufficient to distinguish between Focus and non-Focus materials in either position. At the first complement, condition A is not significantly different from control condition C; at the second complement, condition A is much lower in intensity than condition C, leading to a significant interaction between condition and position: $\beta = -0.37$, $p < 0.01$. Similarly, condition B is significantly lower in intensity than C at the first complement, but not at the second, leading to a significant interaction between condition and position: $\beta = 0.37$, $p < 0.01$.

Just like the F0 results, the relationship between the Focus conditions A and B here seems to be relatively clear and consistent, while the relationship of these conditions to the control condition C is inconsistent across subjects and statistically inconclusive. All five subjects had greater intensity in condition A than B for Complement 1, and greater intensity in condition B than A for Complement 2. The realization of both complements in condition C varied substantially across subjects, although the tendency was toward high intensity values in general in condition C relative to the other two conditions. As noted above, the results are much clearer when the relative intensity of the two complements within an utterance is considered, just as it was when relative pitch within the sentence was considered. This suggests that the intensity values may simply be tracking F0 values here. But because the influence of phrasing and accent on intensity has received considerably less attention in the literature than their influence on F0, we will not attempt to draw any further conclusions about intensity scaling.

6.2 Implications of the set of phonetic prominence findings for the grammar

What emerges in section 6.1 is that the intensity data, like the data from pitch and duration reported above, support the contention that the grammar must make a distinction between constituents that are contrastive Focus and those that are merely discourse-new in the syntactic representation which feeds into phonology and phonetics. The distinct patterns observed with intensity correlate with the distinct syntactic representations of Focus and/or new in (15abc).

Duration, pitch and intensity are key dimensions of phonetic prominence in English; they have been correlated with various levels of stress prominence in English (Fry 1958, Huss 1977, Beckman 1986, Choi et al 2005, Okobi 2006)[20]. The fact that we have found higher values on these three phonetic dimensions for Focus-marked constituents in English supports the two-part hypothesis laid out above: (IIa) that Focus-marked entities carries a higher level of phrasal prosodic (stress) prominence than new constituents in the sentences of Conditions A and B (as posited in (16abc)), and (IIb) that a general principle of phonetic interpretation calls for a systematic relation between prosodic (stress) prominence on the one hand and phonetic prominence on the other: the greater the one, the greater the other[21].

In defining the notion of phonetic prominence in general, one should keep in mind that languages could potentially differ in whether one or another dimension of phonetic prominence is actually called on (or to what degree or in what manner) in the realization of phonological prominence, and hence in the manner in which Focus-related phonetic prominence is realized. Japanese is a language in which Focus words show increase in F0[22], but no increase in duration (Maekawa 1997). Importantly, Maekawa shows durational reduction in the phrases surrounding the Focus in a sentence in Japanese, as well as effects on vowel formants under Focus. The difference between the durational

[20] Research on Dutch has shown additional phonetic reflexes of stress prominence, including spectral tilt (van Heuven and Sluijter 1995 et seq). Moreover, van Heuven et al 1995 have shown that supralaryngeal resonance and glottal pulse shape reflect word-internal stress in English.

[22] See Pierrehumbert and Beckman 1988 and Sugahara 2003 on F0 of Focus in Japanese.

reflexes in English and Japanese of what we hypothesize to be Focus-related prosodic prominence in phonological representation must have its source in the particulars of phonetic implementation in the two languages.

7.0 Summary, conclusion and prospectus

The primary goal of the experiment reported here was to determine whether there is a difference in the prosodic properties of contrastive focus and discourse-new constituents in English that would support the conclusion that linguistic theory must provide a grammatical representation of contrastive focus that is distinct from that of discourse-new. Our findings concerning the duration, pitch and intensity of contrastive focus constituents as compared to discourse-new ones provide phonetic support for the conclusion that contrastive focus does indeed have a distinctive Focus-marking.

The finding that the target Focus and new constituents within the sentences of the minimal triplets examined here have very similar pitch accenting and phonological phrasing properties made it possible, for the first time in phonetic research on the Focus-new contrast, to compare--within the same sentence-- the relation between the phonetic prominence of Focus-new, new-Focus and new-new constituents in sequence. Not surprisingly, a Focus constituent emerges as more phonetically prominent relative to a discourse-new constituent within the same sentence when it comes to duration, pitch profile and intensity. Indeed, this phonetic prominence relation must be what provides hearers with an essential cue to the presence and location of Focus in English sentences. As for the sentences containing just new-new sequences, it turns out that they are distinct in prosodic prominence patterning from both the Focus-new and new-Focus sentences in the minimal triplets. Put in other terms, we observed no neutralization of all-new "broad focus" sentences with "narrow focus" sentences that combine Focus and new.

In our discussion of the findings concerning the differences in patterns of phonetic prominence in the Focus-new, new-Focus and new-new sentence types, we have raised the question of how the relation between Focus-marking and phonetic prominence might

be accounted for in a theory of grammar. Is the relation direct or is it indirect, mediated by some aspect of phonological representation? The notion that, in English, Focus-marking in syntactic representation is reflected in the phonological representation of prosodic (stress) prominence forms part of an indirect theory. This theory must include, in addition, a theory of the relation between the phonological representation of (stress) prominence and its quantitative interpretation in terms of phonetic prominence. Our data are consistent with an indirect theory like this. The fact that the target Focus and new constituents in the three sentence types prove to be for all intents and purposes the same in relevant pitch accent properties and in prosodic constituent structure points to the third crucial aspect of the phonological representation of prosody-- prosodic prominence—as the only explanation for the observed differences in phonetic prominence. Moreover, since phonological representations of prosodic prominence provide an encoding of relations of prominence, as seen in the relative heights of the columns of x-marking in (16abc), our experimental findings showing the significance of within-sentence relations of phonetic prominence in characterizing the phonetic differences between Focus-new, new-Focus and new-new sentences can be easily explained by a simple principle of stress-sensitive phonetic interpretation. This principle specifies a straightforward relation between phonological and phonetic prominence, within some relevant domain: The greater the level of prosodic (stress) prominence, the greater the quantitative degree of phonetic prominence. Applying this principle to sentential representations like (16abc), the within-sentence relations in phonetic prominence that were observed are predicted.

We began this paper with a research question relevant to the theory of semantics and pragmatics: Does data from the phonetics of English provide support for making a grammatical distinction between the contrastive focus vs. discourse-new status of constituents of a sentence? In examining the evidence that it does, we have charted out a further research hypothesis: That the patterns of phonetic prominence that have been found to accompany the distribution of Focus and new in English (and other languages) are to be explained in terms of an interface relation between phonetic prominence and patterns of prosodic (stress) prominence in phonological representation.

Acknowledgments

The research for this paper began in the context of an NSF research grant to E. Selkirk (The phonological reflexes of Focus, #BCS-0004038), and has continued over a period of years, with improvements in our thinking that are due in part to the interchanges we've been able to have with colleagues. We'd like to especially thank Taka Shinya, Mariko Sugahara, Hubert Truckenbrodt, John Kingston, Angelika Kratzer, Caroline Féry, Michael Wagner, Irene Heim, Adam Albright, Edward Flemming, Mats Rooth, Roger Schwarzschild, Daniel Büring, Manfred Krifka, Lyn Frazier, Chuck Clifton and Carlos Gussenhoven for useful conversations and commentary. We also would like to thank audiences where we have presented this work, in various versions, including at presentations at our home institutions UMass Amherst, MIT as well as colloquia at Rutgers (2006) and the Institut Jean Nicod (2011), and conference presentations at McGill, Potsdam, the University of Delaware in 2010. We are also grateful for constructive input from Greg Carlson, Jennifer Cole, and two anonymous reviewers from *Language.*

References

Baayen, R., D. Davidson & D. Bates. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**, 390-412.

Bates, D. (2007). lme4: An R package for fitting and analyzing linear, nonlinear and generalized linear mixed models. Software application.

Beaver, David, and Clark, Brady. 2008. *Sense and sensitivity: How focus determined meaning*. Malden MA: Wiley-Blackwell.

Beckman, Mary E. 1986. *Stress and Non-Stress Accent.* Dordrecht: Foris.

Beckman, Mary E. and Pierrehumbert, Janet. 1986 Intonational structure in Japanese. *Phonology Yearbook* 3, 225-309.

Beckman, Mary, and Ayers-Elam, Gayle. 1993. *Guidelines for ToBI labelling.* URL: www.ling.ohio-state.edu/~tobi

Berg, Rob van den, Gussenhoven, Carlos, and Rietveld, Toni. 1992. Downstep in

Dutch:Implications for a model. In *Papers in laboratory phonology II: Gesture, segment and prosody*, eds. D. Robert Ladd and Gerald Docherty, 335-358. Cambridge: Cambridge University Press.

Bolinger, Dwight. 1961. Contrastive accent and contrastive stress. *Language* 37, 83-96.

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

Breen, Mara. 2007. *The identification and function of English prosodic features*, Ph.D. thesis, Brain and Cognitive Science, MIT.

Breen, Mara, Federenko, Evelina, Wagner, Michael and Gibson, Edward. 2010. Accoustic correlates of information structure. *Language and Cognitive Processes* Volume 25, (Issue 7&9), 1044-1098.

Büring, Daniel. 2007. Intonation, semantics and information structure. In *The Oxford Handbook of Linguistic Interfaces*, eds. Gillian Ramchand and Charles Reiss. Oxford: Oxford University Press.

Büring, Daniel. 2008. Been there, Marked that: A theory of second occurrence focus. In *Proceedings of NII Workshop on Focus-2005*, eds. Makoto Kanazawa and Christopher Tancredi. Tokyo. To appear in published version of proceedings.

Byrd, Dani and Saltzman, Elliot. 2003. The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31, 149-180.

Calhoun, Sasha. 2006. *Information structure and the prosodic structure of English: a probabilistic relationship,* Ph.D. dissertation, University of Edinburgh.

Calhoun, Sasha. 2010. The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language* 86(1), 1-42.

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In *Subject and Topic*, ed. Charles Li, 25-56. New York: Academic Press.

Choi, Jeung-Yoon, Hasegawa-Johnson, Mark, and Cole, Jennifer. 2005. Finding intonational boundaries using acoustic cues related to the voice source. *Journal of the Acoustical Society of America* 118(4): 2579–2587.

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

Cinque, Gugliemo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24:239-297.

Cooper, William E., Eady, Stephen, and Mueller, P. 1985. Acoustical aspects of

contrastive stress in question-answer contects. *Journal of the Acoustical Society of America* 77:2142-2156.

Dilley, Laura, Shattuck-Hufnagel, Stephanie, and Ostendorf, Mari. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24, 423-444.

D'Imperio, Mariapaola. 2002. Italian intonation: An overview and some questions. *Probus* 14(1), 37-69.

Downing, Laura J. 2008. Focus and prominence in Chichewa, Chitumbuka and Durban Zulu. In Marzena Zygis & Susanne Fuchs (eds.), *ZAS Papers in Linguistics 49. Papers in Phonetics and Phonology*, 47-65.

Eady, Stephen, and Cooper, William E. 1986. Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America* 80:402-415.

Face, Timothy L. 2002. Local intonational marking of Spanish contrastive focus. *Probus* 14(1), 71–92.

Féry, C. (2003). Gradient prosodic correlates of phrasing in French. In Meisenburg, Trudel & Selig (eds.), *Nouveaux Départs en Phonologie*. Niemeyer.

Féry, Caroline & Manfred Krifka. 2008. Information structure Notional distinctions, ways of expression. In Piet van Sterkenburg (ed.), *Unity and diversity of languages*, 123-136. Amsterdam: John Benjamins.

Féry, Caroline, and Samek-Lodovici. 2006. Focus projection and prosodic prominence in nested foci. *Language* 82:131-150.

Frota, Sonia. 2000. *Prosody and focus in European Portuguese: Phonological phrasing and intonation*: Outstanding dissertations in linguistics. New York, London: *f*Garland Publishing Co.

Fry, D.B. 1958. Experiments in the perception of stress. *Language and Speech* 1, 126-152.

Grice, Martine. 1995. Leading tones and downstep in English. *Phonology* 12: 183-233.

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19:377-417.

Gussenhoven, Carlos. 1990. Tonal association domains and the prosodic hierarchy in English. In *Studies in the pronuniciation of English—A commemorative volume in honour of A.C. Gimson*, Susan M. Ramsaran, ed., 27-37. London: Routledge.

Gussenhoven, Carlos. 1991. The English Rhythm Rule as an accent deletion rule. *Phonology* 8, 1-35.

Gussenhoven, Carlos 1992 Sentence accents and argument structure. In Iggy Roca (ed.) *Thematic structure: Its role in grammar.* Berlin, New York: Foris Publications, 91-106.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos. 2006. Yucatec Maya tone in sentence perspective. Poster, Laboratory Phonology 10, Paris.

Gussenhoven, Carlos, Repp, Bruno, Rietveld, Toni, Rump, H.H., and Terken, Jacques. 1997. The perceptual prominence of fundamental frequency peaks. *Journal of the Acoustical Society of America* 102:3009-3022.

Gussenhoven, Carlos, and Teeuw, Renske. 2008. A moraic and a syllabic H tone in Yucatec Maya. In *Fonología instrumental: patrones fónicos y variación,* Herrera, Esther and Martín Butragueño, Pedro, eds. El Collegio de Mexico

Hagstrom, Paul. 1998. *Decomposing questions.* Ph.D. dissertation, MIT.

Halle, Morris, and Vergnaud, Jean-Roger. 1987. *An essay on stress*. Cambridge MA: MIT Press.

Halliday, M.A.K. 1967. Notes on transitivity and theme in English (part II). *Journal of Linguistics* 3:199-244.

Hamblin, C.L. 1993. Questions in Montague English. *Foundations of Language* 10,41-53.

Hayes, Bruce. 1989. The prosodic hierarchy in meter. In. *Rhythm and Meter*, Paul Kiparsky and Gilbert Youmans, eds., 201-260. Orlando, FL: Academic Press.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: Chicago University Press.

Hedberg, Nancy. 2003. The prosody of contrastive topic and focus in spoken English. In *Proceedings of teh workshop on information structure in context*, 14-52. Stuttgart: Institut für Maschinelle Sprachverarbeitung.

Heuven, V.J., van & Sluijter, A.M.C. (1996). Acoustic correlates of stress and accent in Dutch and American English. In *Proceedings of the International Congress of Spoken Language Processing*, Philadelphia.

Heuven, V.J., van & Sluijter, A.M.C. & Shattuck-Hufnagel, S. & Stevens, K.N. (1995). Supralaryngeal resonance and glottal pulse shape as correlates of prosodic stress

and accent in American English. In *Proceedings of ICPhS 13*, pp. 630-633.

Heuven, V.J., van & Sluijter, A.M.C. (1996). Notes on the phonetics of word prosody. In Goedemans, R., Hulst, H., van der, Visch, E. (Eds.), *Stress patterns of the world, part 1: background*, pp. 233-269. The Hague: Holland Academic Graphics.

Horn, Laurence 1972 *On the Semantic Properties of Logical Operators in English*, Ph.D. dissertation, UCLA.

Huss, Volker. 1977. English word stress in post-nuclear position. *Phonetica* 35, 86-105.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge MA: MIT PRess.

Kahnemuyipour, Arsalan. 2004. *The syntax of sentential stress,* Ph.D. dissertation, Linguistics, University of Toronto.

Kahnemuyipour, Arsalan. 2009. *The syntax of sentential stress*. Oxford: Oxford University Press.

Keating, P. (2003). Phonetic encoding of prosodic structure. *Proceedings of the 6th International Seminar on Speech Production*, Sydney, December 7 to 10, 2003.

Keating, P. & Shattuck-Hufnagel, S. (2002). A Prosodic View of Word Form Encoding for Speech Production. *UCLA Working Papers in Phonetics* **101**, 112-156.

Kiss, Katlin É. 1998. Identificational focus vs. information focus. *Language* 74:245-273.

Koch, Karsten. 2008 *Intonation and Focus in Nle?kepmxcin (Thompson River Salish).* Ph.D. dissertation, University of British Columbia.

Kratzer, Angelika. 1991. The representation of focus. In *Handbuch Semantik/Handbook Semantics,* von Stechow, Arnim and Wunderlich, Dieter, eds, 825-834. Berlin, New York; de Gruyter.

Kratzer, Angelika. 2004. Interpreting focus: presupposed or expressive meanings? *Theoretical Linguistics* (30(1), special issue on *Interpreting Focus*, 123-136.

Kratzer, Angelika 2010 Distinguishing contrastive, new and given information. [Report of joint work with E. Selkirk] Keynote lecture, Georgetown Linguistics Society meeting, February 2010.

Kratzer, Angelika, and Selkirk, Elisabeth. 2007. Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24, 93-135.

Kratzer, Angelika, and Selkirk, Elisabeth. 2010. Distinguishing contrastive, new and given information. Talk delivered at the *International Conference on Information Structure,* sponsored by SFB 632: Informationsstruktur. Potsdam, July 2010.

Kratzer, Angelika and Selkirk, Elisabeth. (in preparation) Distinguishing contrastive, new and given information. Manuscript, UMass Amherst.

Kratzer, Angelika and Shimoyama, Junko. 2002. Indeterminate pronouns: The view from Japanese. In *Proceedings of the 3rd Tokyo Conference on Psycholinguistics,* Otsu, Yukio, ed., 1-25. Tokyo: Hituzi Syobo

Krifka, Manfred. 1992. A Compositional Semantics for Multiple Focus Constructions. In *Informationsstruktur und Grammatik*, Sonderheft der *Linguistischen Berichte*, Joachim Jacobs, ed.. Also in *Proceedings from Semantics and Linguistic Theory* 1, Cornell Working Papers in Linguistics 10, 1991

Krifka, Manfred. 1993. Focus, Presupposition and Dynamic Interpretation: The Case of Focus-Sensitive Particles. *Journal of Semantics 10*. 269-30

Kügler, Frank & Skopeteas, Stavros. 2006. Interaction of Lexical Tone and Information Structure in Yucatec Maya. *TAL 2006 International Symposium (Tonal Aspects of Language)*, 83-88. La Rochelle, France, April 27-29, 2006.

Kügler, Frank & Skopeteas, Stavros. 2007. On the universality of prosodic reflexes of contrast: The case of Yucatec Maya. In *Proceedings of the XVI. International Congress of Phonetic Sciences,* Trouvain, Jürgen & Barry, William J. (Eds.). Saarbrücken, Germany, 4-10 August 2007.

Ladd, D. Robert. 1980. *The structure on intonational meaning: evidence from English*. Bloomington IN: Indiana University Press.

Ladd, D. Robert. 1988. Declination 'Reset' and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84:538-544.

Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge Univ. Press.

Ladd, D. Robert. 2008. *Intonational phonology*. 2nd ed. Cambridge: Cambridge University Press.

Ladd, D. Robert and Schepman, Astrid. 2003. "Sagging transitions" between high pitch accents in English: experimental evidence. *Journal of Phonetics* 31(1), 81-212.

LeGac, David. 2002. Tonal alternations and prosodic structure in Somali. *Speech Prosody 2002,* Aix-en-Provence.

Liberman, Mark, and Pierrehumbert, Janet. 1984. Intonational invariance under changes in pitch range and length. In *Language sound structure*, eds. Mark Aronoff and Richard Oehrle, 157-223. Cambridge MA: MIT Press.

Maekawa, Kikuo. 1997. Effects of focus on duration and vowel formant frequency in Japanese. In Y. Sagisaka and Nick Campbell, eds., *Computing Prosody: Computational Models for Processing Spontaneous Speech.* Springer Verlag.

Neeleman, Ad and Reinhart, Tanya. 1998. Scrambling and the PF interface. In Miriam *The Projection of Arguments*, eds. Butt, Miriam and Gueder, W., 309-353. Stanford: CSLI Publications.

Neeleman, Ad, and Szendroi, Kriszta. 2004. Superman sentences. *Linguistic Inquiry* 35:149-159.

Nespor, Marina, and Vogel, Irene. 1986. *Prosodic phonology*. Dordrecht: Foris.

Okobi, Anthony O. 2006. *Acoustic correlates of word stress in American English.* Ph. D. dissertation, Harvard-MIT Division of Health Sciences and Technology.

Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*, Ph.D. dissertation, Linguistics, MIT.

Pierrehumbert and Beckman 1988. *Japanese Tone Structure.* Cambridge MA: MIT Press.

Pierrehumbert, Janet, and Hirschberg, Julia. 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication*, eds. P.R. Cohen, James Morgan and Martha Pollack, 271-311. Cambridge MA: MIT Press.

Pitrelli, John, Beckman, Mary and Hirschberg, Julia. 1994. Evaluation of prosodic transcription labeling reliability in the ToBI framework. *Proceeding of the 3rd International Conference on Spoken Language Processing (ICSLP 94),* Yokohama, Japan.

Prince, Ellen. 1981. Toward a taxonomy of given-new information. In *Radical Pragmatics,* ed. Peter Cole, 223-255. New York: Academic Press.

Quené, H. & H. van den Bergh. (2004). On Multi-Level Modeling of data from repeated measures designs: A tutorial. *Speech Communication*, **43**(1-2), 103-121.

Raaijmakers, J., J. Schrijnemakers & F. Gremmen. (1999). How to Deal with "The Language-as-Fixed-Effect Fallacy": Common Misperceptions and Alternative Solutions. *Journal of Memory & Language* **41**(3), 416-426.

Rialland, Annie and Stéphane Robert. 2001. The intonational system of Wolof. *Linguistics* 39(5), 873-939.

Rochemont, Michael. 1986. *Focus in generative grammar.* Amsterdam and Philadelphia: J. Benjamins Publishing Co.

Rooth, Mats. 1985. *Association with focus.* Ph.D. dissertation, UMass Amherst.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75-116.

Rooth, Mats. 1996. Focus. In *The handbook of contremporary semantic theory*, ed.

Shalom Lappin, 271-297. London: Blackwell.

Schwarzschild, Roger. 1999. GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7:141-177.

Selkirk, Elisabeth. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge MA: MIT Press.

Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3:371-405.

Selkirk, Elisabeth. 1996. Sentence prosody: Intonation, stress and phrasing. In *The handbook of phonological theory*, ed. John A. Goldsmith, 550-569. Cambridge MA and Oxford UK: Blackwell.

Selkirk, Elisabeth. 2000. The interaction of constraints on prosodic phrasing. In *Prosody: Theory and experiments*, ed. Merle Horne, 231-262. Dordrecht: Kluwer.

Selkirk, Elisabeth. 2002. Contrastive focus vs. presentational focus: Prosodic evidence from right node raising in English. In *Speech Prosody 2002: Proceedings of the 1st International Conference on Speech Prosody*, ed. Bernard Bel and Isabelle Marlien, 643-646. Aix-en-Provence, France.

Selkirk, Elisabeth. 2007. Focus prominence drives focus phrasing: Bengali intonation revisited. In *Topic and focus: A cross-linguistic perspective.*, eds. Chung-min Lee, Matthew Gordon and Daniel Büring, 217-246. Dordrecht: Kluwer.

Selkirk, Elisabeth. 2008. Contrastive focus, givenness and the unmarked status of "discourse-new". *Acta Linguistica Hungarica* 55:331-346.

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In *The Handbook of Phonological Theory*, eds. John Goldsmith, Jason Riggle and Alan Yu. Oxford: Blackwell.

Sugahara, Mariko. 2003. *Downtrends and post-FOCUS intonation in Tokyo Japanese.* Ph.D. dissertation, University of Massachusetts Amherst.

Syrdal, Ann K. and McGory, Julia. 2000. Inter- transcriber reliability of ToBI prosodic labeling. *Proceedings of ICSLP-2000* 3, 235- 238.

Truckenbrodt, Hubert. 1995. *Phonological phrases: Their relation to syntax, focus and prominence*, Ph.D. dissertation, Linguistics, MIT.

Truckenbrodt, Hubert. 2002. Upstep and embedded register levels. *Phonology* 19:77-120.

Truckenbrodt, Hubert. 2004. Final lowering in non-final position. *Journal of Phonetics* 32:313-348. Truckenbrodt, Hubert. 2006. Phrasal Stress. In Keith Brown (ed.), *The Encyclopedia of Languages and Linguistics, 2nd edition, Vol. 9*, 572-579.

Amsterdam: Elsevier

Truckenbrodt, Hubert. 2006. Phrasal Stress. In Keith Brown (ed.), *The Encyclopedia of Languages and Linguistics, 2nd edition, Vol. 9*, 572-579. Amsterdam: Elsevier

Turk, Alice E. and White, Laurence. 1999. Structural influences on accentual lengthening in English. *Journal of Phonetics* 27(2), 171-206.

Vallduví, Enric, and Vilkuna. 1998. On rheme and kontrast. In *Syntax and semantics*, 79-108. New York: Academic Press.

Veilleux, Nanette, Brugos, Alejna, and Shattuck-Hufnagel, Stephanie. 2008. *The new English ToBI home page*. URL: www.tobihome.org

Wagner, Michael. 2005. *Prosody and recursion*, Ph.D. dissertation, Linguistics, MIT.

Wightman, Colin W., Shattuck-Hufnagel, Stefanie, Ostendorf, Mari, and Price, Patti J. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91(3), 1707-1717.

Williams, Edwin. 1997. Blocking and anaphora. *Linguistic Inquiry* 28(4), 577-628.

Xu, Y., and Xu, C.X. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33:159-197.

Yoon, Tae-Jin, Cole, Jennifer and Hasegawa-Johnson, Mark. 2007. On the edge: Acoustic cues to layered prosodic domains. Proceedings of International Congress of Phonetic Sciences, Saarbrucken, 2007.

Appendix A: Minimal triplet stimuli

1a. African languages have a number of very interesting speech sounds. The Bushman languages make use of a variety of ejectives and clicks. Zulu has a complex system of high and low tones. **But they only speak Wolof in Mali. And it has pretty standard speech sounds.**

1b. John's planning a long trip to Africa, to do some trekking and practice his language skills. For hiking, he'll go to Kenya or Tanzania, which are in East Africa. **But they only speak Wolof in Mali. So he'll be in West Africa for the language part.**

1c. I've been trying to learn about the languages and cultures of West Africa. The situation is complex. **They speak Wolof in Mali. It's in a completely different language family from the other two languages spoken there.**

2a. The ancient Greeks were very skilled blacksmiths, and were able to work with a number of different metals. They had great success in refining and using copper and nickel. **But they almost always used iron for arrowheads. It was strong and easier to sharpen.**

2b. Archeologists have been able to reconstruct the lives of the ancient Greeks from the set of artifacts they left behind. There were carefully crafted daggers, and relics of enormous shields. **They almost always used iron for arrowheads. Actually, they've found no trace of the bows.**

2c. The ancient Greeks were a more advanced society than many people seem to think. They were scientifically very advanced. Their craftsmanship was excellent. **They used iron for arrowheads. Other cultures took several more centuries to catch up.**

3a. Gary is a really bad art dealer. He gets attached to the paintings he buys. He acquired a few Picassos and fell in love with them. The same thing happened with a Cezanne painting. **So he would only offer that Modigliani to MoMA. I bet the Picassos would have fetched a much higher price.**

3b. Gary is an art dealer. Lately he's been very picky about which museums he deals with; he doesn't do business with the Metropolitan or the Guggenheim. **He would only offer that Modigliani to MoMA. He says that's the only museum with a space good enough to hang it in.**

3c. Gary was a successful art dealer, and could afford to be pretty demanding with his clients. He would never make a deal unless the price was right and he respected the buyer. **He will probably offer that Modigliani to MoMA. But only for a six-figure sum.**

4a. The Red Sox had an exhibition game for charity, and they gave the players various bright-colored uniforms. Bill Mueller and Nomar Garciaparra have really played well this year. **But they only gave Manny the yellow one. That's the one that's reserved for the most valuable player.**

4b. The Red Sox had an exhibition game for charity, and they had special bright-colored uniforms made for the occasion. There were a lot of different colors; a couple of the jerseys were orange, one was purple. **But they only gave Manny that yellow one. That was a lousy color.**

4c. The Red Sox had an exhibition game for charity, and they gave all the players crazy bright-colored uniforms to wear for the occasion. The whole thing was pretty funny to watch. **They gave Manny the yellow one. It was so ugly.**

5a. There are so many competing football teams in California. There are three or four popular big college teams, and then there are the Raiders and the Chargers in the NFL. **It seems they only watch the 49'ers in Salinas. This might be the most popular team in the state.**

5b. California is a big football state, with millions of fans. Each town has its own special fan population. In Los Angeles, the fans all care about how they look on TV. They're organized in competing clubs in San Diego. **People only watch the 49'ers in Salinas. It's an unpopular team everywhere else.**

5c. Central Coast California is a pretty interesting place for sports. There's a lot of amateur sports, but also a lot of professional team activity. **They watch the 49'ers in Salinas. There's also a minor league ballpark there.**

6a. Nora is a compulsive shopper. She's always running off to Target. She makes frequent passes through T.J. Maxx. **And she usually goes to Wal-Mart on Monday. Her housemates complain about all the stuff she brings home.**

6b. Nora doesn't manage her schedule very well. On Tuesday, she always has a lot of free time. And on Thursday, too. **But she usually goes to Wal-Mart on Monday. It would be easier if she picked a better day.**

6c. Nora's been looking for a special kind of espresso machine that she saw in Gretchen's house. She can't seem to find it anywhere. She tried everywhere in Northampton. **And she went to Wal-Mart on Monday. That didn't seem to work out, so she'll keep looking.**

7a. The U.S. Department of Defense just released their annual weapons inventory. They've got computer-guided precision bombs, and robot missiles. But they still make use of more conventional arms. **For instance, they mostly store mines in Idaho. They really don't keep any sophisticated weapons there.**

7b. The U.S. Department of Defense just released their annual report on the military infrastructure and its geographical distribution. The vast majority of military installations are in the South. There are facilities scattered in other places, though. **For instance, they mostly store mines in Idaho. It's pretty sparsely populated there.**

7c. The U.S. Department of Defense just released their annual report on the storage of arms and hazardous materials. I read the section on the installations in the West, which was pretty interesting. **For instance, they store mines in Idaho. Probably because there's not many people there.**

8a. Jane's a really controlling chef. She yells at Michael and is constantly interfering with Larry. **But she usually lets Lena prepare the mayonnaise. She thinks anyone else would overbeat it.**

8b. Jane's a controlling chef. She tries to do all the work herself. She slices the vegetables, browns the meat, and prepares the stock for sauces. **Oddly enough, she always lets Lena prepare the mayonnaise. Mayonnaise is really tricky.**

8c. Jane works as a chef. She is in charge of a number of cooks, and one of the hardest parts of the job is knowing their abilities and supervising the kitchen activities. **She lets Lena prepare the mayonnaise. She's an efficient worker, and frees up time for Jane.**

9a. Our soccer team went to compete in European tournaments last year. It was pretty intimidating, because the level of play was so high. But we surprised everybody with a number of second- and third-place finishes. **We even wound up winning in Mannheim. That was an incredible experience.**

9b. Our soccer team went to compete in European tournaments last year. In Madrid we played really well, and in Manchester, too. **We even wound up winning in Mannheim. And that's the toughest tournament in Europe.**

9c. Our soccer team went to compete in European tournaments last year. It went better than we expected, and we had a good time. **We wound up winning in Mannheim. They gave us a big trophy, and we keep it in the clubhouse now.**

10a. DeBeers is the largest ore corporation in the world. They employ thousands of people, including engineers, detonation teams, and earth-moving experts. **They primarily use miners for extracting aluminum. It's an amazing range of jobs.**

10b. DeBeers is the world's largest ore company. They've made billions from open pit copper operations. They're flushing the rivers of the Andes for gold. **And they primarily use miners for extracting aluminum. Aluminum is the most precious commodity.**

10c. DeBeers is one of the largest companies in the world. They extract everything, and they have digging operations in over thirty countries. **They use miners for extracting aluminum. This is the least profitable operation.**

11a. We used a lot of different carpenters when the house was being built. John David was pretty reliable, and that guy Gus did a lot of good work. **But we only hired Manny to work on the annex. We feel a lot of loyalty towards him.**

11b. The house really needed a ton of work. The front porch was rotting, and the driveway needed to be repaved. **But we only hired Manny to work on the annex. We figured that was all he could handle.**

11c. We didn't want to do any substantial repair work this year. There just wasn't enough money to do everything that needed to get done. **Even so, we hired Manny to work on the annex. And that worked out well.**

12a. Bill chooses the most awful companions. He was dating that horrible lawyer last year, and then there was Kate, who we all hated. **He even took Minnie to a Mariners game. And she's insufferable.**

12b. Bill is a sports freak. He'll go to any kind of sporting event, regardless of the team that's playing. **He even took Minnie to a Mariners game. And they haven't been in contention for years.**

12c. Bill's had a pretty busy week. He had meetings all through the weekend, and then he went to Seattle for a conference. **He took Minnie to a Mariners game. I bet that was fun.**

13a. The new corporate Vice President has some interesting ideas for improving the quality of life at the bank. He wants to create a high degree of responsibility and accountability for tellers. **But he only wants the managers to process loans. They're the ones who wind up with most of the responsibility.**

13b. The corporate Vice President has really missed the mark with his new policy. On any given day, the bank deals with dozens of customer complaints, new accounts, and business issues. **But he only wants the managers to process loans. All that other business they don't even see.**

13c. The corporate Vice President is making some minor changes to the division of labor in the bank. He supports a less centralized model, where each of the branches makes its own policies. **He wants the managers to process loans. And branch security will be outsourced to another company.**

14a. John's supposed to handle relations between the students and the Jordanian visitors who will be involved in the seminar. But he's neglected to set up meetings for Mike. And Sarah's out of town. **So he's only introduced Annie to Abdullah. He hasn't introduced anyone else to him.**

14b. John's been handling the Jordanian exchange students. Hussein and Khalil are pretty intimidated by the atmosphere in the department here, and he's giving them time to

get adjusted. **So he's only introduced Annie to Abdullah. Abdullah's more outgoing and doesn't mind being interviewed.**

14c. John's been handling the Jordanian exchange students. He organized an orientation seminar for them, where they talked about some of the differences that they might notice at our school. **He introduced Annie to Abdullah. They seem to be getting along very well.**

15a. Almost every player on the football field needs to be able to block. It's a fundamental skill for backs, receivers, and even kickers. But these players are all subject to restrictions. **They only let linemen use their hands. All the other players are at a disadvantage.**

15b. There are a lot of ways to lay down a good block in football. Checks with the hip are very effective at impeding forward progress. And you can really knock someone down with a shoulder block. **They only let linemen use their hands. That's a lot less effective than blocking with other body parts.**

15c. Football must have more rules than any other major sport. The rules about blocking alone could fill up an entire book. There are different rules for tackles upfield and tackles downfield. **They let linemen use their hands. Rules are determined by where the line of scrimmage is in relation to the play.**

16a. I had lunch with the boss the other day, and he's not overly impressed with the employees. He refuses to even talk about Richards and Tanner. **But he always says that O'Malley is a winner. I really like that guy, too.**

16b. The boss told me that the latest round of layoffs were performance-related. He wanted to get rid of the deadwood, the cheaters, the loafers and losers. **But he always says that O'Malley is a winner. He must be, since they didn't fire him.**

16c. The boss is pretty happy about the way the office is running. Production is up, and morale is good. The new hire is working out. **He says that O'Malley is a winner. I'm inclined to agree.**

17a. I tend to eat a lot of fast food when I'm on campus. There's the Blue Wall. Or the People's Market for a quick bagel and cheese. **I always go to the Newman Center on Monday. I have two classes near there.**

17b. I am really obsessive about my schedule. I concentrate my classes on Tuesday and Thursday. On Wednesday, I work all day. **I always go to the Newman Center on Monday. That's when they have the best specials.**

17c. I didn't do much this week. I studied for a midterm in my Poli-Sci class, and watched a really bad movie with my roommate. **I went to the Newman Center on Monday. I try to make it to mass once a week.**

18a. The Persian empire was famous for its textiles. Weavers produced silk, and fine cotton cloth. **They only produced linen in Ninevah. And it was the most valued of them all.**

18b. Trade was very difficult in ancient Persia. It might take weeks to travel between the capital and a trade center like Ephesus, or to the port in Persepolis. **And they only produced linen in Ninevah. That was a month's travel away.**

18c. The Persian empire had a complex supply web. Goods were transmitted between major cities primarily by cart and camel. **They produced linen in Ninevah. It would later be carried through the desert to Bactri.**

Appendix B: Strings measured for duration

Complement 1

Item	String	Onset	Offset
1	speak [wolo]f in	onset of /w/	onset of /f/
2	used [iron] for	glottal stop	onset of /f/
3	modig[liani] to	burst or F2 min	abrupt drop in or cessation of energy above F1
4	gave [manny] the/that	onset of /m/	onset of th
5	forty [niner]s	onset of /n/	onset of /z/
6	to [walmart] on	F1 min; ampl. min	tap or burst transient
7	store [mine]s	ampl. min; F3 rise	onset of /z/
9	up [winning] in	after burst	offset of ng
10	use [miner]s	offset of /z/	onset of /z/
11	hired [manny] to	onset of /m/	abrupt drop in or cessation of energy above F1
12	took [minny] to	onset of /m/	abrupt drop in or cessation of energy above F1
13	the [mana]gers	onset of /m/	before burst transient
14	introduced [annie] to	after burst	before burst transient
15	let [line]men	after burst	formant shift or midpoint of nasal sequence
16	o'[malley] is	onset of /m/	F2 max
17	the [newman] center	onset of /n/	onset of /s/
18	produced [linen] in	offset of /t/	offset of /n/

Complement 2

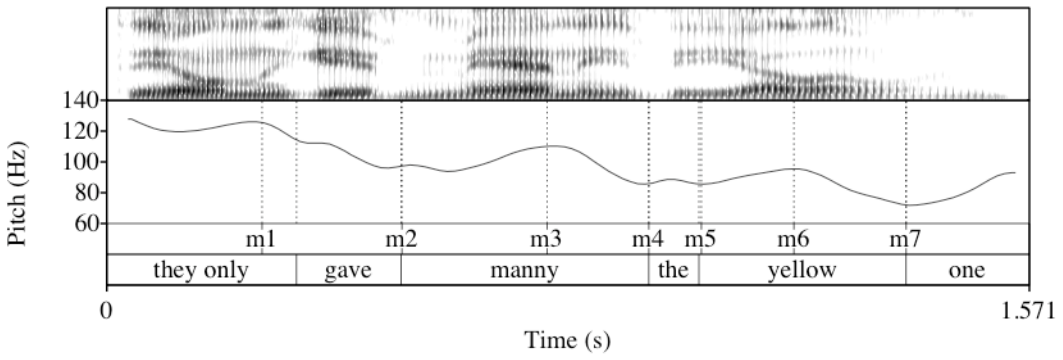
Item	String	Onset	Offset

			drop in energy above F2; low
1	wolof [in ma]li	offset of /f/	plateau in F1 & F2
2	for [arrow]heads	glottal stop	onset of /h/
3	to [mo]ma	onset of 1st /m/	onset of 2nd /m/
4	the/that [yellow] one	F2 max; ampl. min	F1 min; ampl. min
5	in s[alin]as	offset of /s/	offset of /n/
6	on [mon]day	formant shift or midpoint of nasal sequence	before burst transient
7	in [ida]ho	glottal stop or V onset	onset of /h/
9	in [mann]heim	formant shift or midpoint of nasal sequence	onset of /h/
10	extracting [alumi]num	offset of ng	onset of /n/
11	the [anne]x	glottal stop or F1 rise	before burst transient
12	the [marin]ers	onset of /m/	offset of /n/
13	process [loan]s	offset of /s/	onset of /z/ or /dz/
14	a[bdul]lah	closure	F1 & F2 rise; ampl. min
15	their [han]ds	onset of /h/	burst or onset of /z/
16	a [winn]er	F1 & F2 min	offset of /n/
17	on [mon]day	formant shift or midpoint of nasal sequence	before burst transient
18	i[n nine]veh	onset of /n/	onset of /v/

Notes:

- onset/offset of fricatives determined by appearance of high-frequency noise in the spectrogram.
- onset/offset of nasals determined by presence of periodic voicing and/or formant structure when adjacent to obstruents; shift in formants and decrease in waveform envelope when adjacent to vowels.

Appendix C: Illustration of F0 measurement points



Special Matter

	Condition A	*Condition B*	*Condition C*
1st complement	Focus / __ new	new / __ Focus	new / __ new
2nd complement	new / Focus __	Focus / new __	new / new __

TABLE 1. Materials elicited in the experiment

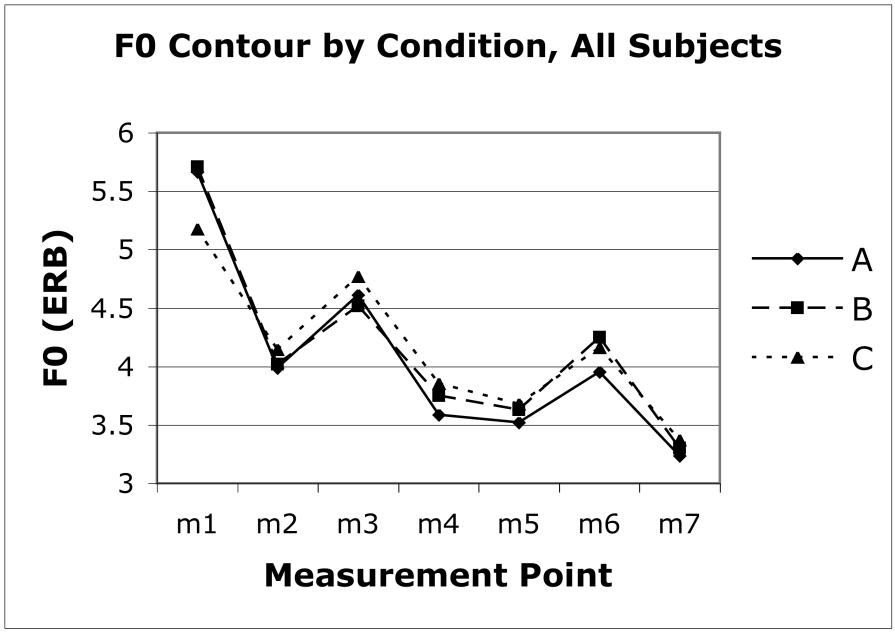


FIGURE 1. Schematic F0 contours for each condition, measured at seven points in the sentence and averaged across subjects

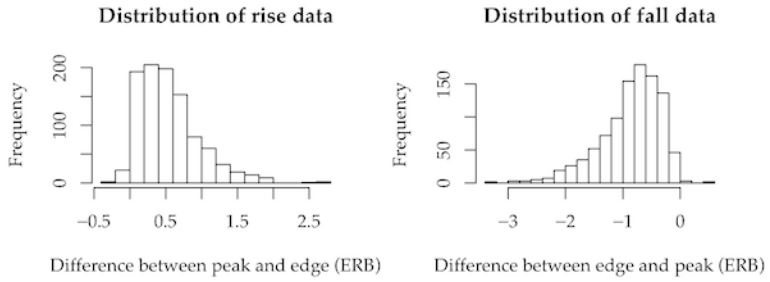


FIGURE 2. Histograms showing rises preceding (left) and falls following (right) accented syllables, pooled across subject, conditions, and position in the sentence. Y axis shows the number of data points falling within each bin (range of values).

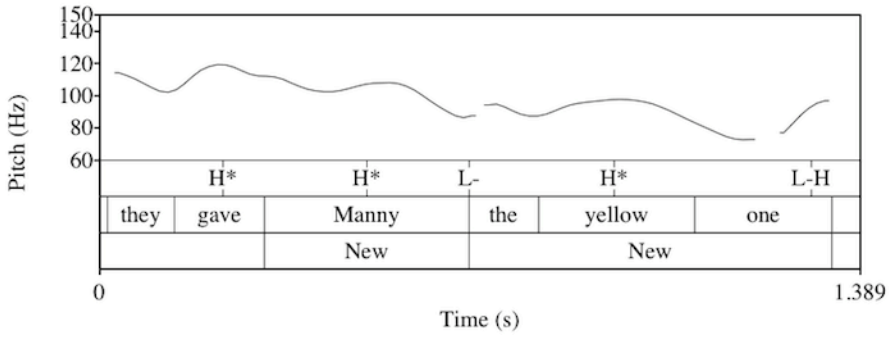


FIGURE 3. Smoothed pitch track, Condition C: new-new complements [speaker RGL]

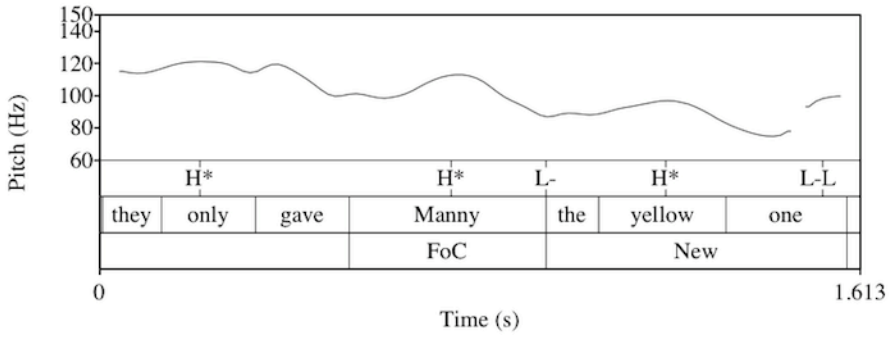


FIGURE 4. Pitch track for Condition A sentence: Foc-new complements [Speaker RGL]

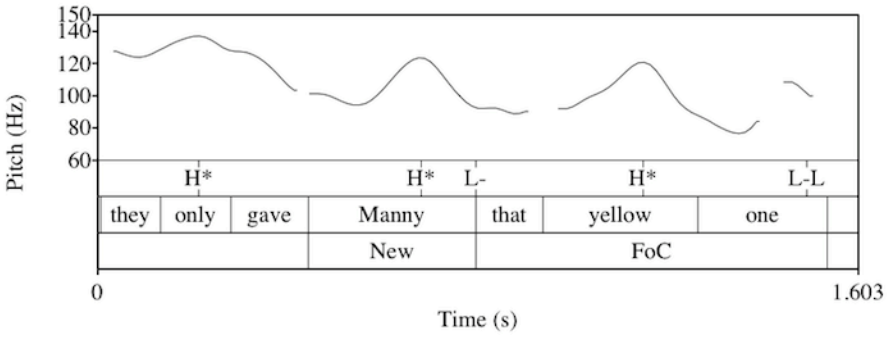


FIGURE 5. Pitch track for Condition B sentence: new-Foc complements [speaker RGL]

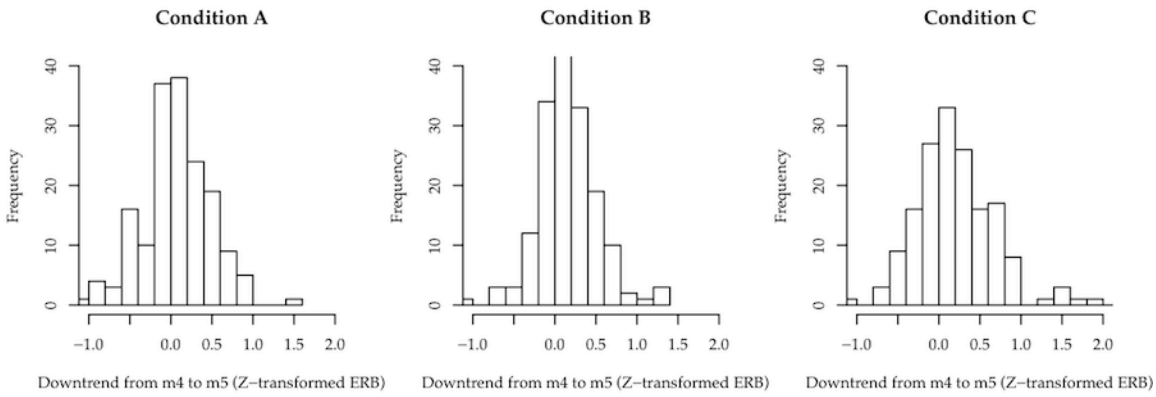


FIGURE 6. Histograms for each condition showing the amount of downtrend from m4 to m5. Negative values indicate uptrend. Y axis shows the number of data points falling within each bin (range of values).

CONDITION	DURATION (ms)	
	Complement 1	Complement 2
A	360	266
B	323	294
C	326	268

TABLE 2. Average duration of first and second verbal complements in three information structure conditions: A = Foc-new; B = new-Foc; C = new-new.

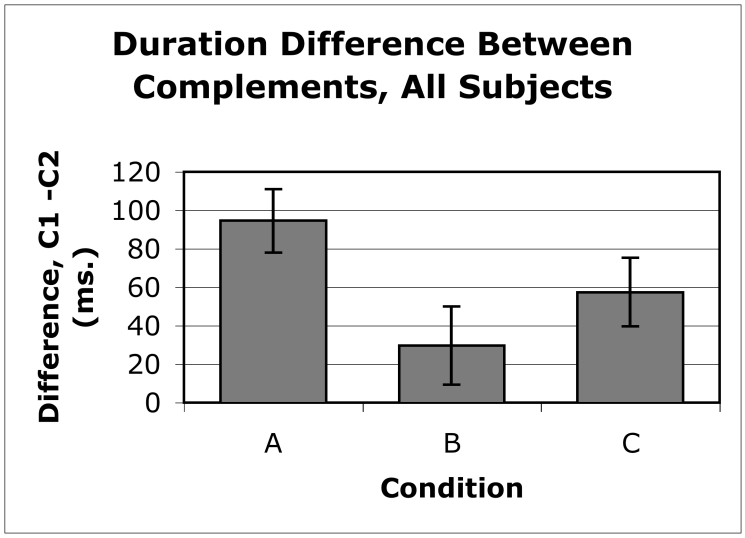


FIGURE 7. Average difference between measured stretches of first and second verbal complements in three information structure conditions: A = Foc-new; B = new-Foc; C = new-new. Whiskers represent 95% confidence intervals.

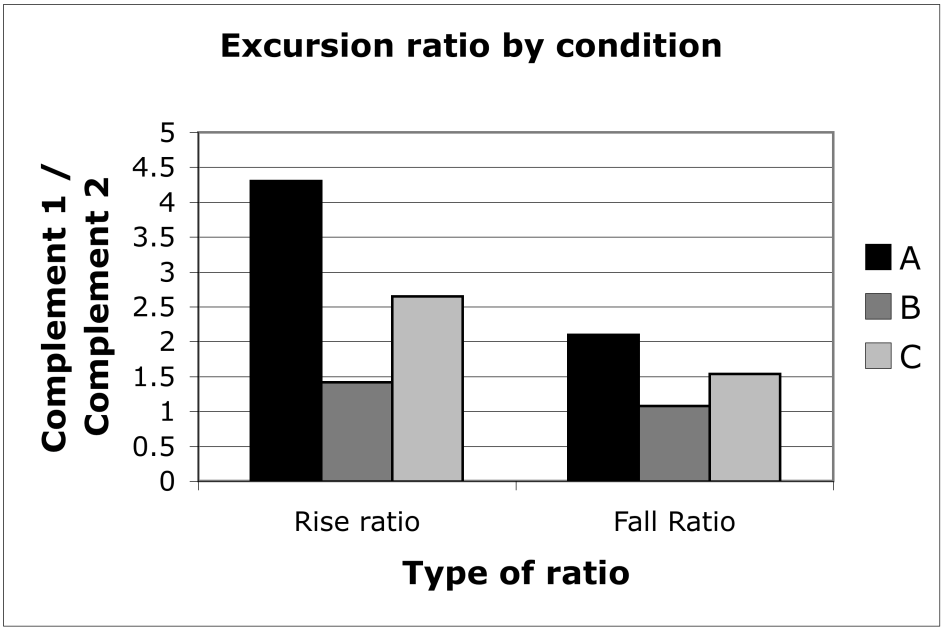


FIGURE 8. Ratios indicating the relation between pitch rises and falls at Complements 1 and 2, by condition

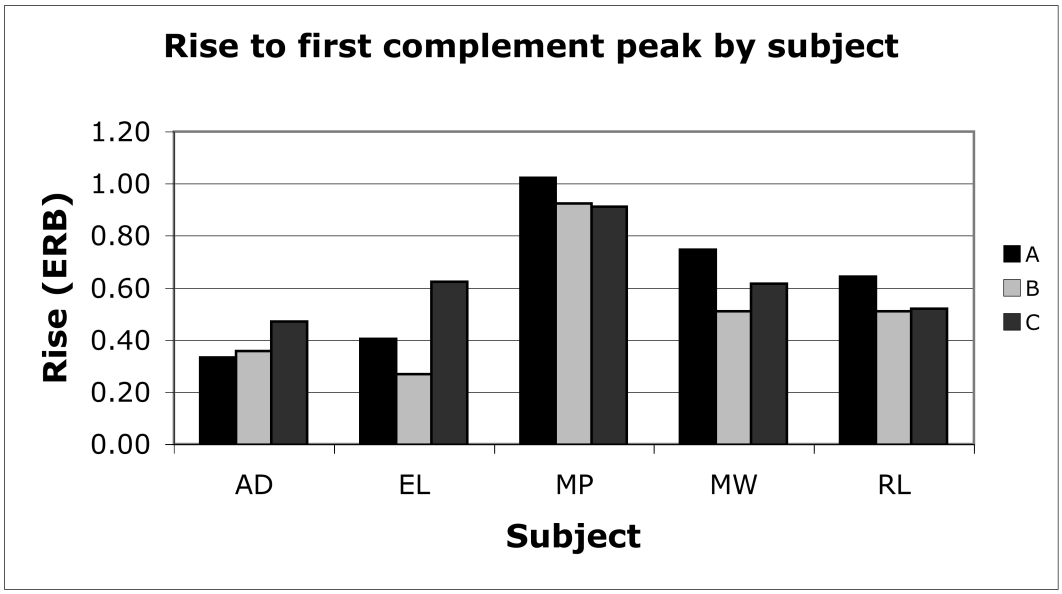


FIGURE 9. Between-subject variability in amount of rise to the first complement

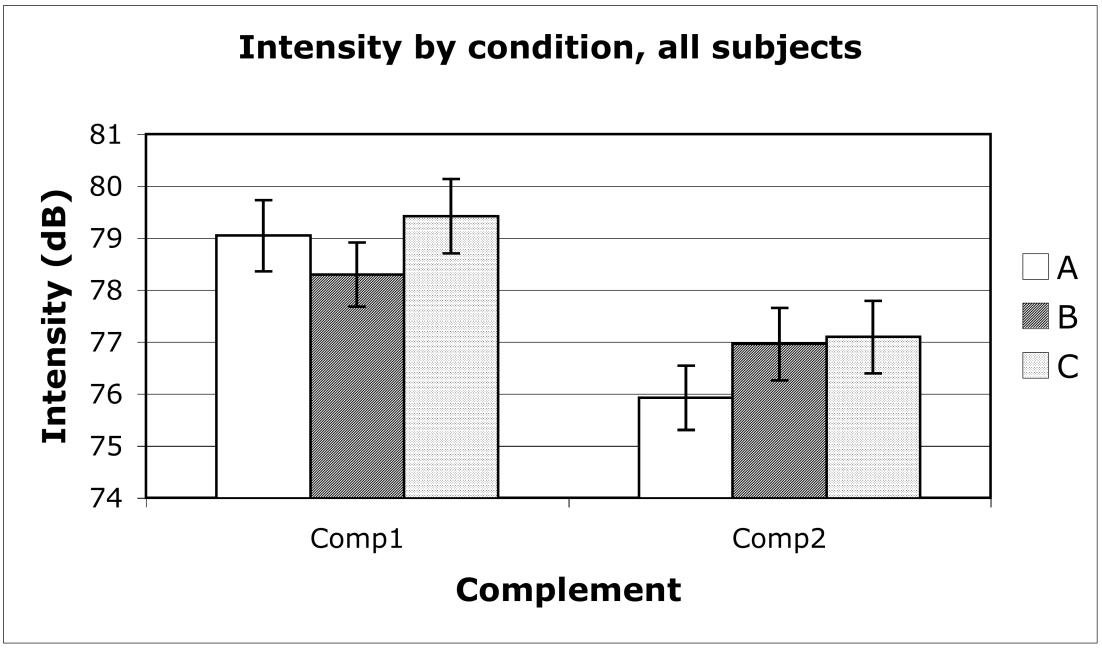


FIGURE 10. Average intensity peak within measured stretch including stressed syllable in both complements, for all subjects; A = Foc-new; B = new-Foc; C = new-new. Whiskers indicate 95% confidence intervals.