

*Effects of Pitch Accent Position, Type, and Status on Focus Projection**

Pauline Welby

The Ohio State University

Key words

focus

focus projection

intonation

pitch accent

Abstract

This paper examines predictions made by two theories of the relationship between pitch accent and focus. The empirical evidence presented suggests that listeners are sensitive to a variety of factors that may affect the focus projection ability of pitch accents, that is the ability of a pitch accent on one word to mark focus on a larger constituent. The findings suggest that listeners' interpretation of focus structure is most sensitive to the presence or absence of a pitch accent on a focused constituent and the deaccenting of following unfocused material (pitch accent position). Preliminary evidence suggests that the status of a pitch accent as nuclear or prenuclear may also affect listeners' interpretations, though to a lesser extent than accent position. Finally, the results show that focus projection is affected only minimally, if at all, by the type of pitch accent (at least for the two accent types compared (H* vs. L + H*)).

1 Introduction

The same sentence uttered with different tunes or intonation patterns will be appropriate in very different contexts. For example, (1) is a possible exchange while (2) is not. Words in small capital letters are pitch accented.

- (1) Q: What newspaper do you read?
A: I read the [DISPATCH]_{Focus}
- (2) Q: What newspaper do you read?
A: # I READ the [Dispatch]_{Focus}

The two versions differ in that (1) has a pitch accent, a prominence-leading pitch movement, aligned to *Dispatch*, while (2) has a pitch accent aligned to *read*. The difference in acceptability follows from what each intonation pattern conveys about the

* *Acknowledgments:* I thank Mary Beckman, Allison Blodgett, Julie Boland, Paul Davis, Mariapaola D'Imperio, Svetlana Godjevac, Rebecca Herman, Martin Jansche, Keith Johnson, H el ene L evenbruck, Julie McGory, Laurie Maynell, Craig Roberts, Amy Schafer, Scott Schwenter, Shari Speer, Steve Winters, and the Pragmatics and Psychoids discussion groups at Ohio State for valuable comments and other assistance. I also thank Martine Grice, Carlos Gussenhoven, and Amy Schafer for their comments and suggestions.

Address for correspondence: The Ohio State University, Department of Linguistics, 222 Oxley Hall, 1712 Neil Avenue, Columbus, OH 43210, U.S.A.; e-mail: <welby@ling.ohio-state.edu>.

information structure of the discourse. In the examples, *Dispatch* provides information that updates the hearer's knowledge base, and the denotation of *Dispatch* is thus the focus of the sentence. In this example, it provides New information that answers the *what newspaper* of *What newspaper do you read*. The verb *read*, which is used in the immediately preceding question, conveys Given information. (See Chafe, 1983; Prince, 1981; & Vallduví, 1992 *inter alia* for further discussion of information structure.) In (1) pitch-accented *Dispatch* is prominent, and the answer is an appropriate response to the question. The answer in (2), where Given *read* is pitch accented, but not New *Dispatch*, is not an appropriate response.

Previous work on English has shown that constituents that are focused are usually assigned pitch accents while constituents that are not focused may lack pitch accents (e.g., Brown, 1983). A focused item is often, but not always, interpreted as New to the discourse. It can also be interpreted as Given but standing in contrast to another item, as in (3), where Given *read* is focused in the response. Note that (3) contains two foci. In the example, the parentheses indicate that the utterance has two prosodic phrases.

(3) Q: Do you read the Dispatch?

A: Yeah, (I [READ]_{Focus} the Dispatch) (but I don't [ENJOY]_{Focus} it).

In some cases, a pitch accent on a constituent seems to be able to "project" focus to a larger constituent. For example, a single pitch accent on the head of the noun phrase (NP) complement of a verb (the object NP) has been claimed to project focus to the entire verb phrase (VP). In (4), for example, the information conveyed by the VP *read the Dispatch* is New information—it has not been mentioned in the immediately preceding question, yet only the head noun of the object NP *the Dispatch* bears a pitch accent.

(4) Q: How do you keep up with the news?

A: I [read the DISPATCH]_{Focus}

Other research has found that other factors, such as the persistence of surface structure and the persistence of grammatical role also play an important role in accent assignment (e.g., Terken & Hirschberg, 1994), but these factors were not manipulated in the current experiments.

Other factors which may affect focus projection include type of pitch accent and intonational phrasing. These factors, which have received little attention in the literature, are examined here in a set of two experiments. Experiment 1 tests the hypothesis that only a pitch accent that is the last pitch accent in its phrase can mark focus. For example, the sentence in (5) can be spoken without a medial prosodic break (as in (5a)) or with a break (as in (5b)).

(5) Q: How do you keep up with the news?

a. A: I READ the DISPATCH.

b. A: (I READ) (the DISPATCH).

Experiment 2 tests the hypothesis that only a certain type of pitch accent can project focus. Imagine the answer in (6) spoken as a simple answer to a question, as in (6a), and as a contradiction, as in (6b). It is hypothesized that the contrastive pitch accent on *Dispatch* (6b) will not successfully project focus and will therefore not be an

appropriate response to a broad focus question (such as *How do you keep up with the news?*).¹

- (6) Q: How do you keep up with the news?
 a. A: I read the DISPATCH.
 b. A: I read the *Dispatch* (not the Times).

The experiments also examine predictions made by two well-known accounts of the relationship between accent and focus, Gussenhoven (1983, 1984, 1992, 1999) and Selkirk (1995) (see also Selkirk, 1984). Since focus in both these accounts is closely connected to the notions Given and New, it is important to define these terms. None of these papers give explicit definitions of Given and New, but simplifying somewhat, we can induce the following working definitions from the examples provided and the experimental materials used:

In an answer, the denotation of a linguistic expression is considered Given if it is coreferential with material from the preceding question. Everything not Given is New.

There are a number of differences between Selkirk's and Gussenhoven's theories. Selkirk approaches focus projection from the point of view of the listener who must map the phonology of an utterance onto the pragmatics. Gussenhoven offers a generation model in which the speaker maps the pragmatics of an utterance onto its phonological realization. However, he also approaches the question from the point of view of the listener in testing listeners' interpretation of intonation patterns in two experiments evaluating his model.

According to the Selkirk (1995) account, focus assignment proceeds through the assignment of a feature to a syntactic constituent (F-marking), and the interpretation of that feature. Accented words are F-marked and the F-marking of the head of a phrase or an internal argument of the head licenses the F-marking of the entire phrase (the Phrasal Focus Rule). A constituent that is not F-marked is interpreted as Given. An F-marked constituent that is not dominated by any other F-marked constituent is a Focus (FOC), and a Focus can be interpreted as Given or New. An F-marked constituent that is not a Focus is interpreted as New. Selkirk's theory does not include an explicit model of intonation, but generally follows the Pierrehumbert (1980) model of intonation. In Selkirk's model, all accents are assumed to have equal status with respect to their ability to project focus regardless of pitch accent type or prosodic phrasing. Selkirk (1995, p. 562), however, suggests that "[i]t may turn out that only certain of the pitch accents of English license the projection of focus."

Selkirk's rules predict, for example, that the sentence *I read the Dispatch*, uttered as a single intonation phrase with a single pitch accent on *the Dispatch*, is ambiguous. It is an appropriate answer to the question *What newspaper do you read?*, a question that calls for narrow focus on the object NP (as in (1)). In that case, the NP *the Dispatch* is the Focus of the sentence. The sentence uttered with the same accent pattern is also predicted to be a possible answer to the question *How do you keep up with the news?*, a question that calls for focus on the entire verb phrase (as in (4)). In that case, focus is

¹ *Dispatch* is italicized in the example following common typographical convention for conveying contrast. The accent type is described in the Experiment 2 section.

projected to the verb phrase via the Phrasal Focus Rule.

In Gussenhoven's (1999) theory, semantic constituents (defined as arguments, predicates, and modifiers), not syntactic constituents, are claimed to be marked for focus. In Gussenhoven's model of intonation, English has three pitch accents (fall (HL), fall-rise (HLH), rise (LH)). Nine additional pitch accents result from modifications of basic pitch accents. Gussenhoven's Sentence Accent Assignment Rule (SAAR) applies to all pitch accent types. According to the SAAR, every focused argument (subject and internal verbal complements), predicate (verbs, predicate nominals, predicate adjectives, etc.), and modifier (adverbials and other adjuncts) must be accented. The exception is that a predicate adjacent to an argument may be unaccented. A *focus domain* is defined as "any constituent whose focus can be marked with a single pitch accent, such as a focused argument, a focused modifier, or a focused argument-predicate combination" (p. 45).

Like Selkirk's account, Gussenhoven's account also predicts that the sentence *I read the Dispatch* uttered as a single intonation phrase with a single pitch accent on *the Dispatch*, is ambiguous between a structure with object-NP focus and VP focus.

The two theories make different predictions about the appropriateness of answer like the one in (7), where there are pitch accents on both the verb and the object NP. Selkirk's Basic Focus Rule predicts that assignment of a pitch accent to a constituent will F-mark that constituent.

(7) Q: What newspaper do you read?

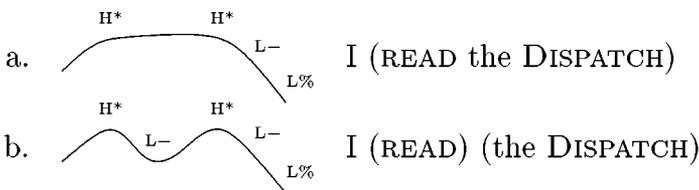
A: I READ the [DISPATCH]_{Focus}

In Selkirk's model, the pitch accent on the verb F-marks the verb, which must then be interpreted either as New (if F-marking projects to VP) or as a Focus (if F-marking does not project past V). Since the verb is neither New nor contrastively focused, Selkirk's model predicts that this answer should be inappropriate. Gussenhoven's model includes an optional rule which allows the unrestricted assignment of pitch accents to the left of the focal pitch accent within a focus domain, so his theory predicts that the answer should be appropriate.

Since neither Gussenhoven's nor Selkirk's theory makes reference to pitch accent type or to prosodic phrasing, both theories predict that these factors should not affect focus projection.

Figure 1

Hat pattern versus two peak accent pattern



The assumptions that both Selkirk and Gussenhoven make about the structure of intonational patterns like the ones in Figure 1 differ from the assumptions made here, where the Beckman and Pierrehumbert (1986) model of intonation is adopted (see also Pierrehumbert, 1980). In that model, there is a crucial difference between the (a) and (b) patterns in Figure 1. That is, although both patterns have pitch accents on the verb and on the object NP (indicated by H*), the (a) pattern (traditionally called the hat pattern due to its shape) contains a single *nuclear accent*, the second pitch accent. The pitch accent is nuclear because it is the last pitch accent in an intermediate phrase (marked by a L-phrase accent). The two-peak pattern in (b) contains two nuclear accents: each pitch accent is the only (and therefore last) pitch accent in its intermediate phrase. The nuclear accent or nuclear accents are claimed to have special prominence, and experimental evidence for this special status has been found (Ayers, 1996). Gussenhoven, however, does not recognize Beckman and Pierrehumbert's (1986) distinction between pitch accents and following phrase tones (phrase accents and boundary tones). He treats the H*L- and H*L-L% sequences in the two-peak pattern as undecomposable bitonal pitch accents (unmodified HL accents). Both the hat pattern and the two-peak pattern therefore contain sequences of pitch accents uninterrupted by a phrase boundary. Selkirk does not fully address the treatment of the medial L in the two-peak pattern (see Selkirk, 1995, p. 568, fn. 3). But since neither theory restricts focus marking to accents of a specific type (e.g., HL vs. LH or H* vs. L + H*) or to accents with a specific status as nuclear or prenuclear, it is clear that the predictions each theory makes for the hat pattern are identical to the predictions that theory makes for the two-peak pattern.

The two theories make predictions about the appropriateness of different combinations of focus structure and accent pattern. By manipulating the focus structure of questions and the intonation patterns of answers and obtaining judgments regarding their appropriateness in combination, we can test these predictions. A few studies have experimentally examined some of the predictions made by the two theories.

Gussenhoven (1983) performed a context-retrievability experiment in which listeners were asked to determine whether the question and answer of a question-answer pair came from the same or a different conversation. The critical comparison contained two types of questions, questions calling for focus on the entire verb phrase and questions calling for focus on a VP-internal argument of the verb (usually an object NP). An example is given in (8).

- (8) What do you do?
 What do you teach?
 I teach LINGUISTICS.

The answers in the critical comparison all contained a pitch accent on the object NP but no pitch accent on the verb. A control condition consisted of so-called nonmerging predicates, for which the verb and the object NP were not predicted to form a single focus domain. An example is given in (9).

- (9) Please tell me what happened last night.
 What do you remember from your last lesson?
 I remember NOTHING.

Gussenhoven found that, in contrast to the nonmerging predicate condition, listeners performed no better than chance in correctly judging whether question-answer pairs were correctly matched. This finding supports the prediction of the SAAR that when the entire VP is focused, only the VP-internal argument (the object NP), but not the verb, must be accented.

Using materials structured similarly to those used by Gussenhoven (1983), but with one condition with a pitch accent on the verb as well as the object NP, Birch and Clifton (1995) obtained appropriateness judgment ratings and response times (RTs) to a “make sense” judgment task. Response times (RTs) were faster when a New object NP was accented than when it was not accented. The experiment showed no difference in RTs for the VP-focus question when there was a pitch accent on both the object NP and the verb versus only the object NP. These findings are consistent with the predictions of both Selkirk and Gussenhoven’s theories. In the ratings task, however, Birch and Clifton found a small but significant preference (0.28 on a 1–5 scale) for pitch accents on both the verb and the NP object in responses to questions calling for VP focus. Birch and Clifton reconcile the results of these two experiments by suggesting that the small difference in appropriateness ratings found between the two conditions “does not necessarily mean that the accent is essential for efficient communication or that it is required for linguistic well-formedness. It could simply be seen by our subjects as typical or even ‘correct’ in some prescriptive sense” (p. 375).

2 Experiment 1

TABLE 1

Outline of tones in the ToBI transcription system

Pitch accents are prominence-lending tones associated to a particular syllable: H*, L*, bitonal accents such as L + H*.

Phrase accents fill up the space between the last pitch accent and the end of an intermediate phrase: H –, L –, !H –.

Boundary tones mark the edges of an intonational phrase: H%, L%.

Experiment 1 used the appropriateness ratings task of Birch and Clifton (1995) to investigate the influence of prosodic phrasing on focus interpretation. The materials were developed using the ToBI (Tones and Break Indices) transcription system, an explicit model for transcribing the intonation of certain varieties of English, including Standard American English (Beckman & Hirschberg, 1994). The system, outlined in Table 1, assumes a hierarchical structure in which pitch accents are contained in intermediate phrases, and one or more intermediate phrases are grouped together to form an intonational phrase. Following Beckman and Pierrehumbert (1986), the nuclear accent, defined as the last pitch accent in an intermediate phrase, is assumed to have particular prominence. Beckman (1996) claims that this difference in prominence has a specific effect on the relationship between intonation and focus structure:

[T]o produce an utterance with double foci explicitly on two different words (i.e., to make two equally stressed nuclear-accented syllables), it is necessary to make two intermediate phrases, with an intervening intermediate phrase break marked by a phrase accent.

(p. 35)

In other words, only nuclear pitch accents mark focus. By manipulating focus structure and intonation patterns and soliciting appropriateness judgments, the following hypothesis was tested:

HYPOTHESIS: Nuclear pitch accents, but not prenuclear accents, mark focus.

A sample set of question-answer pairs for Experiment 1 is given in (10). There are two possible question types, a question calling for focus on the VP and a question calling for focus on the object NP. The four possible answers contrast in the presence or absence of a H* pitch accent on the verb, the object NP or both, and in the presence or absence of a medial phrase boundary.

(10) Q (VP focus): What's that terrible smell coming from the neighbors' yard?

Q (object-NP focus): There's a terrible smell coming from the neighbors' yard.
What are they burning?

	H*		L-L%	
a. A:	They're BURNING	their garbage.		(verb H* pattern)
	H*	H*	L-L%	
b. A:	They're BURNING	their GARBAGE.		(hat pattern)
		H*	L-L%	
c. A:	They're burning	their GARBAGE.		(object-NP H* pattern)
	H*	L-	H*	L-L%
d. A:	They're BURNING	their GARBAGE.		(two-peak pattern)

Predictions: Both Gussenhoven and Selkirk predict that the verb-H* pattern (H* 0 L-L%, illustrated in (10a)) will be inappropriate in answers to both object-NP focus and VP focus questions. According to Gussenhoven, a focused argument must have a pitch accent. In Selkirk's model, an unaccented object NP is interpreted as Given. Both models also predict that the object-NP-H* pattern (0 H* L-L%, illustrated in (10c)) will be appropriate in answers to both types of questions. According to Gussenhoven's SAAR, a focused predicate (here, the verb) is de-accented when it is adjacent to a pitch-accented argument (here, the object NP). The object-NP-H* pattern is therefore appropriate in answers to both object-NP focus and VP focus questions. Selkirk also predicts that this pattern should be appropriate in answers to both question types. By the Phrasal Focus Rule, a pitch accent on the object NP licenses projection of focus to the VP.

Recall that Gussenhoven and Selkirk both predict that listeners should find the hat pattern (H* H* L-L%, (10b)) and the two-peak pattern (H* L-H* L-L%, (10d)) equally appropriate or inappropriate. Gussenhoven predicts that the hat pattern and the two-peak pattern will be appropriate in answers to both object-NP focus and VP focus

TABLE 2

Predictions of two theories for the Experiment 1 dialogs: ✓ predicted to be appropriate, # predicted to be inappropriate

<i>V</i>	<i>NP</i>	<i>Selkirk</i>		<i>Gussenhoven</i>	
		<i>NP obj. foc.</i>	<i>VP foc.</i>	<i>NP obj. foc.</i>	<i>VP foc.</i>
H*	0 L–L%	#	#	#	#
0	H* L–L%	✓	✓	✓	✓
H*	H* L–L%	#	✓	✓	✓
H* L–	H* L–L%	#	✓	✓	✓

questions. The pitch accent on the verb, while not required, is allowed by the optional prefocal pitch accent rule.

Selkirk makes different predictions for the appropriateness of these two patterns in answers to the object-NP and the VP focus questions. She predicts that the hat pattern and the two-peak pattern should be judged appropriate in answers to VP focus questions: in that condition, both the *V* and its *NP* complement are *New* and can have pitch accents. As discussed earlier, Selkirk predicts that the hat pattern and the two-peak pattern should be judged inappropriate in answers to object-NP focus questions. The pitch-accented verb must therefore be interpreted as *New* or *Given* but focused (contrastively focused). Since the verb is neither *New* nor contrastively focused, the model predicts that this answer should be inappropriate. These predictions are summarized in Table 2.

2.1

Methods

Materials: The materials consisted of 48 question-answer dialogs like the example set in (10). For each dialog, there were two versions of the question and four versions of the answer. One version of the question called for object-NP focus and the other version for VP focus. Each answer consisted of a subject pronoun coindexed with the subject of the question sentence, a verb, and the object NP. The answer sentences differed according to which word or words (verb or head noun of object NP) bore a pitch accent and whether the pitch accent was nuclear or prenuclear. All tokens within each of the four answer versions conformed to the same intonation pattern. An example set with stylized pitch curves is given in Figure 2.

Combining the two question types with the four answer types resulted in eight experimental conditions. The question-answer pairs were recorded in separate sessions in a sound-attenuated chamber by two female native speakers of American English. The speaker of the questions was presented with a printed version of the complete dialogs (questions and answers) and instructed to speak as naturally as possible. The intonation of the questions was not manipulated. The speaker of the answers, an expert in the ToBI transcription system, was given ToBI-labeled written versions of each sentence and

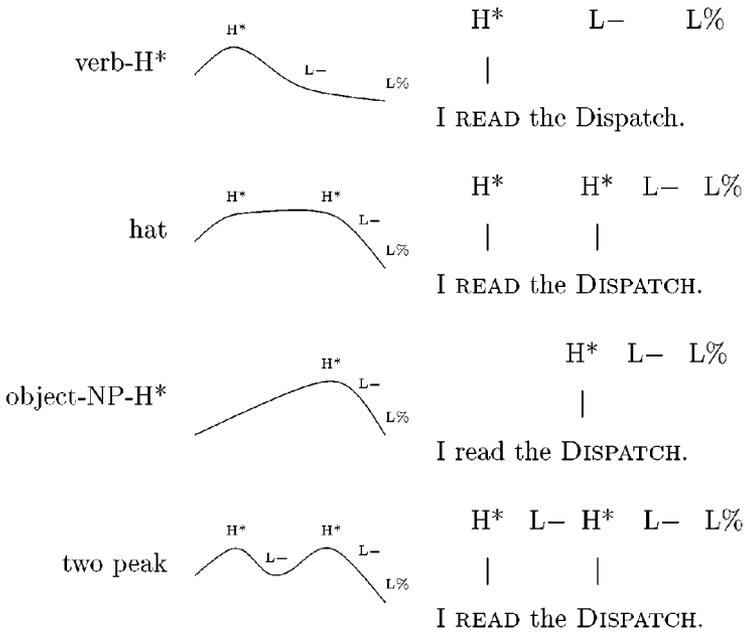
Figure 2

Structure of Experiment 1 questions and answers

Two versions of each question

- Object-NP focus: What newspaper do you read?
- VP focus: How do you keep up with the news?

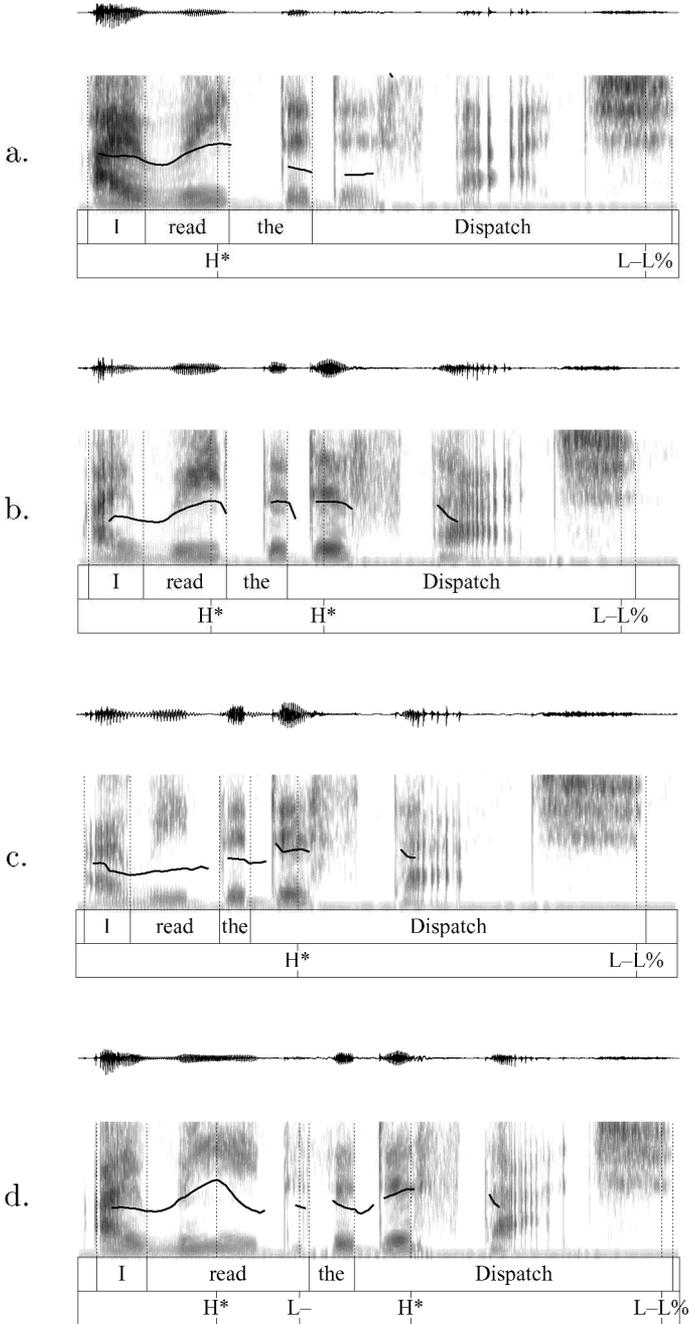
Four versions of each answer



asked to read the sentence with the ToBI pattern indicated. The speech was digitized at 8 kHz and each answer was saved as a separate file. Fundamental frequency (F0) was calculated and the author listened to the sound files and inspected the F0 tracks to make sure that the items had been produced as intended. The hat pattern (H* H* L-L%) and the two-peak pattern (H* L-H* L-L%) were distinguished by the presence of a medial intermediate phrase boundary in the latter pattern. A fall to a L target and a sense of disjuncture signaled the presence of the phrase boundary. Figure 3 shows example F0 tracks for all four answer types. Eight tapes were made, with each tape containing six dialogs of each question type/accnt pattern. Each tape contained exactly one version of each of the 48 question/answer pairs. There was a fixed order of dialogs and the order of the eight conditions was pseudorandomized.

Figure 3

Example F_0 tracks for the four conditions of Experiment 1



Subjects: Eighty undergraduate students at the Ohio State University (10 per tape) participated in the experiment. All were native speakers of American English with no reported hearing problems. English was the only language spoken during childhood.

Procedure: Subjects listened to taped dialogs in a quiet room and completed an appropriateness rating task. The instructions were modeled after those used in Birch and Clifton (1995). The instructions given were:

You will hear a series of 48 short question-answer dialogs. The response in each dialog is an appropriate and direct answer to its question in terms of the words that are used. However, some of the responses may not be appropriate in terms of the way that the words are said. In particular, some words may be emphasized or not emphasized in ways that make the response sentence less relevant to answering that particular question.

Listen carefully to each dialogue. PAY ATTENTION NOT ONLY TO WHAT IS SAID BUT TO HOW IT IS SAID. Then rate on a scale from 1 to 5 whether you think that the response sounds like an appropriate answer to the question. Circle 1 if the response is "definitely appropriate." Circle 5 if the response is "definitely NOT appropriate." Use the intermediate values to indicate that the answer is somewhere in between "definitely appropriate" and "definitely NOT appropriate."

There was a short practice session. In the practice session, the same dialog was repeated three times, each time with a different accent pattern. No feedback was given on whether any of the intonation patterns presented were appropriate or inappropriate.

2.2 Results

Table 3 compares the experimental results with the predictions made by each of the two theories for the two focus structures examined.

TABLE 3

Experiment 1 results compared with theoretical predictions. ✓ judged to be appropriate # judged to be inappropriate. ✓# The two peak pattern (H* L-H* L-L%) was given appropriateness ratings in between those of the other two patterns

<i>V</i>	<i>NP</i>	<i>Selkirk</i>		<i>Gussenhoven</i>		<i>listeners</i>
		<i>NP obj. foc.</i>	<i>VP foc.</i>	<i>NP obj. foc.</i>	<i>VP foc.</i>	
H*	0 L-L%	#	#	#	#	#
0	H* L-L%	✓	✓	✓	✓	✓
H*	H* L-L%	#	✓	✓	✓	✓
H* L-	H* L-L%	#	✓	✓	✓	✓#

Mean ratings by subjects and by items were obtained for each condition and submitted to ANOVA. The results of ANOVA showed a highly significant main effect of accent pattern by subjects and by items, $F_1(1, 79) = 134.137, p < .001$; $F_2(1, 47) = 163.255, p < .001$. There was no main effect of question type (object-NP focus/

VP focus), $F_1(1, 79) = .018, p = .893$; $F_2(1, 47) = .018, p = .892$, and no interaction between question type and accent pattern, $F_1(1, 79) = 1.802, p = .146$; $F_2(1, 47) = 2.194, p = .088$. Listener ratings are summarized in Figure 4.

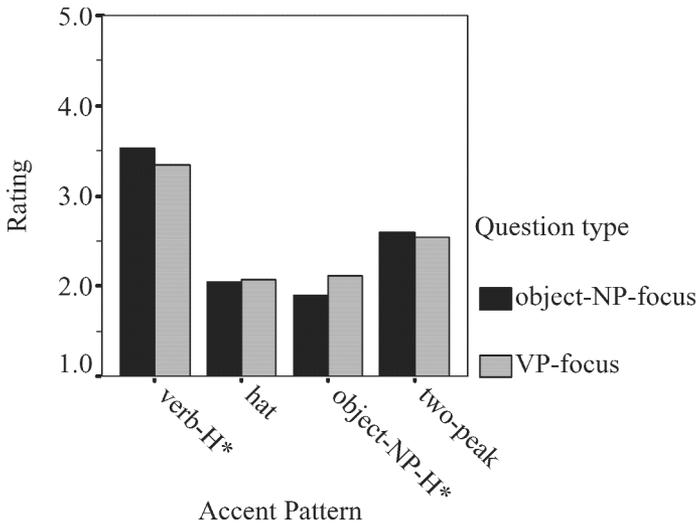


Figure 4

Appropriateness judgments for the two question types and four accent patterns of Experiment 1. (1 is best; 5 is worst)

The results of a planned Scheffe pairwise analysis showed that the rating for the verb-H* pattern ($H^* 0 L-L\%$), the pattern with the highest (worst) rating (3.379 for VP-focus condition, 3.530 for NP-focus condition), was significantly different from the ratings of all three other patterns ($p < .001$ for all comparisons). This result is consistent with Gussenhoven's and Selkirk's predictions.

The object-NP-H* pattern ($0 H^* L-L\%$) was found to be appropriate, as predicted by both Gussenhoven and Selkirk. The mean appropriateness rating for the object-NP-H* pattern (2.117 for the VP-focus condition, 1.898 for the NP-focus condition) was significantly different from the ratings for the verb-H* pattern ($H^* 0 L-L\%$) and the two-peak pattern ($H^* L-H^* L-L\%$) (2.544 for the VP-focus condition and 2.589 for the NP-focus condition). The difference between the object-NP-H* pattern and the hat pattern, however, did not even approach significance ($p = .949$).

The predictions of both Gussenhoven and Selkirk that listeners would give the same judgments to the hat pattern and the two-peak pattern were not confirmed. The hat pattern ($H^* H^* L-L\%$) was judged to be appropriate. The rating for the hat pattern is significantly different from those of both the two-peak pattern ($H^* L-H^* L-L\%$) and the verb-H* pattern ($H^* 0 L-L\%$). The ratings for the two-peak pattern ($H^* L-H^* L-L\%$) were significantly higher than those for the verb-H* pattern ($H^* 0 L-L\%$) but significantly lower than those for the hat pattern ($H^* H^* L-L\%$) and the object-NP-H* pattern ($0 H^* L-L\%$).

2.3

Discussion: Experiment 1

The appropriateness rating of the hat pattern is consistent with Gussenhoven's prediction and Selkirk's prediction for VP focus questions and with Gussenhoven's prediction for object-NP focus questions. The finding is inconsistent with Selkirk's prediction for object-NP focus questions because the pitch accent on the verb should create a focus structure that is incompatible with the Given status of the verb.

The predictions that the hat pattern ($H^* H^* L-L\%$) and the two-peak pattern ($H^* L-H^* L-L\%$) would be judged to be equally appropriate or inappropriate were not supported. The two-peak pattern is significantly less acceptable in a response to either an object-NP focus or a VP focus question. Both of these question types call for an answer with a single focus interpretation. It seems likely that the two-peak pattern is inconsistent with a single focus interpretation, but rather favors a double focus interpretation. In the two-peak pattern ($H^* L-H^* L-L\%$) answers, in which both pitch accents are nuclear accents in a Beckman and Pierrehumbert analysis, both accents may mark focus. This focus marking does not project to the VP, so the response is not interpreted as a VP-focus answer (i.e., it is not judged to be an appropriate response to a VP-focus question). Instead, nuclear accenting of both the verb and the object NP creates a double focus structure that is inappropriate to both the VP and object-NP focus questions in the experiment.

The two-peak pattern is predicted to be appropriate in responses that call for a double focus structure where the verb and object NP are separately focused. For example, in (11), *read* and *Dispatch* contrast with *write* and *Times*.

(11) Q: How do you keep up with the news?

Do you write for the Times or something?

$H^* L- \quad \quad H^* \quad L-L\%$

A: (No,) I [READ]_{Focus} [the DISPATCH]_{Focus}

If focus is marked with a nuclear pitch accent, then we would expect to find a double nuclear accent pattern in cases where there is clear double focus. Example (12), from a corpus of ToBI transcribed radio news speech, is a clear example of the double nuclear accent pattern being used to convey contrastive double focus (Linguistic Data Consortium, 1997, speaker m1b).

$H^* \quad \quad L + H^* L-L\%$

(12) It looks like a watch —

$H^* \quad L- \quad \quad H^* \quad L^* \quad L-H\%$

It [FUNCTIONS]_{Focus} [like an electronic PROBATION OFFICER]_{Focus}

The story from which (12) is taken was read by six different newreaders. The speakers differed in choice of pitch accents, phrase accents, and boundary tones. But five of the six speakers produced this sentence with the hypothesized double focus structure, placing a nuclear pitch accent on *functions* and a nuclear pitch accent on at least one of the words in *electronic probation officer*, separated by a phrase accent or a phrase accent-boundary tone sequence.²(overleaf)

The example from the corpus and the results of Experiment 1 are consistent with the hypothesis that it is necessary for a pitch accent to be nuclear to be interpreted as a focus marker. However, further study to test this hypothesis is clearly needed. One such test might ask listeners to rate the appropriateness of a double nuclear pattern versus a hat pattern in responses with a double focus interpretation, as in (13). A reading task or a dialog game task of the type used in Krahmer and Swerts (1998) might be used to see what accent patterns speakers produce in contexts favoring double focus.

(13) Q: How do you keep up with the news?

Do you write for the Times or something?

- a. A: (No,) I [READ]_{FOCUS} [the DISPATCH]_{FOCUS}
 $H^* L- \quad H^* \quad L-L\%$
- b. A: (No,) I [READ]_{FOCUS} [the DISPATCH]_{FOCUS}
 $H^* \quad H^* \quad L-L\%$

There is a problematic conflict between listener judgments for the two-peak pattern ($H^* L-H^* L-L\%$) in Experiment 1 and listener judgments found in Birch and Clifton (1995). Birch and Clifton found that subjects judged answers like those in (14) acceptable in response to VP-focus questions.

(14) Q: What did Tina do while her neighbors were away?

- A: She [WALKED] [THE DOG]_{FOCUS}
 $L+H^* \quad L- \quad L+H^* \quad L-L\%$

The answers have an accent pattern with $L+H^*$ pitch accents on both the verb and the object NP, separated by a L phrase accent. The items differ from the two-peak pattern in Experiment 1 only in the type of pitch accent (H^* vs. $L+H^*$). Since both of these pitch accents are nuclear, we would expect that both should mark focus and that projection of focus to the VP should be blocked as it is in the Experiment 1 items. Either the choice of pitch accent type or the particular combination of pitch accents and phrase accents in Birch and Clifton's materials may account for listeners' interpretation of the pattern as appropriate in response to VP-focus questions. It is possible that the intervening low phrase accent was obscured by the leading low tone of the second $L+H^*$ pitch accent and that the intended two-intermediate-phrase structure $L+H^* L-L+H^* L-L\%$ pattern was misparsed as the one-intermediate phrase $L+H^* L+H^* L-L\%$. Another possibility is that Birch and Clifton's materials sounded more emphatic than those used here, since the materials in the earlier study used $L+H^*$ as the critical pitch accent in all conditions. As discussed later, the $L+H^*$ accent may be more salient to the listener than the H^* accent (see introduction to Experiment 2). Different rules may also guide the focus interpretation of emphatic utterances (see (18) and discussion). Experiments which include both an accent type condition and a phrasing

² The one speaker who does not have a nuclear pitch accent on *functions* has a very dense distribution of pitch accents. He produces the sentence as one intonational phrase with pitch accents on all content words in the sentence and on *like*: (*It FUNCTIONS LIKE AN ELECTRONIC PROBATION OFFICER*).

condition (e.g., with patterns: $H^*L-H^*L-L\%$, $H^*H^*L-L\%$, $L+H^*L-L+H^*L-L\%$, and $L+H^*L+H^*L-L\%$) would clarify the individual contributions of accent type and phrasing and the contribution of their interaction.

In contrast to the two-peak pattern, the hat pattern ($H^*H^*L-L\%$) was found to be appropriate. That is, the presence of the prenuclear pitch accent in the hat pattern did not render the pattern inappropriate as a response to either a VP-focus or an object-NP-focus question. In fact, there is no difference in appropriateness between the hat pattern ($H^*H^*L-L\%$) and the object-NP- H^* pattern ($0H^*L-L\%$). The appropriateness of the hat pattern suggests that the presence of a prenuclear pitch accent does not affect focus structure interpretation. Specifically, the prenuclear H^* pitch accent on the verb does not mark focus and create a double focus interpretation. The contrast in appropriateness between the hat pattern and the two-peak pattern is consistent with an account in which only nuclear accents mark focus.

What blocks the projection of focus to the VP? Since the hat pattern ($H^*H^*L-L\%$) and the two-peak pattern ($H^*L-H^*L-L\%$) are minimal pairs, differing only in that the latter pattern contains a low phrase accent, it is reasonable to believe that the presence of this phrase accent somehow blocks focus projection.

Schafer (1997) explicitly investigates the possibility that intonational phrasing affects the projection of focus. Her account assumes Selkirk's (1995) focus projection and interpretation rules and adds the hypothesis in (15).

- (15) Bounded Projection Hypothesis: F-marking cannot project beyond the prosodic phrase which contains the accented material. (Schafer, 1997)

The Bounded Projection Hypothesis makes different predictions for focus projection to the VP in (16a) and (16b). In the examples, the intended focus structure is indicated by square brackets and the *Focus* label. Underlining indicates the maximum projection of focus allowed by the hypothesis. In the (16a) answer, the F-marking on the pitch-accented noun *potatoes* can project to the NP *some potatoes*, but is blocked from projecting beyond the phrase boundary (L-) to the verb. Since F-marking cannot project to the head, further projection to the VP is not possible. The hypothesis predicts, however, that F-marking can project to the verb and then to the VP in examples like (16b), where there is no prosodic phrase boundary after the verb that would block projection.

- (16) Predictions of the Bounded Projection Hypothesis (Schafer, 1997)

- (a) What did the farmer do?

$L^* \quad \quad \quad L- \quad \quad \quad H^* \quad L-L\%$
The FARMER [delivered some POTATOES]_{Focus}

- (b) What did the farmer do?

$L^* \quad L- \quad \quad \quad H^* \quad L-L\%$
The FARMER [delivered some POTATOES]_{Focus}

- (examples from Schafer, 1997)

Schafer 1997 provides experimental evidence that shows that, as predicted, listeners

find the phrasing in (16b), but not the phrasing in (16a) appropriate in answers to VP-focus questions. This important result demonstrates that not only the presence or absence of pitch accents, but also intonational phrasing plays a role in the interpretation of focus structure.

In Schafer's account, F-marking of a pitch-accented verb can project to the VP.³ The VP is then an F-marked constituent undominated by another F-marked constituent and is therefore interpreted as the Focus. The Bounded Projection Hypothesis, therefore, does not account for the inappropriateness of the two-peak pattern in responses to VP-focus questions like (17).

(17) Predictions of the Bounded Projection Hypothesis (Schafer, 1997)

Q: How do you keep up with the news?

H* L- H* L-L%

A: ✓ I [READ the DISPATCH]_{Focus} (appropriate)

Findings of Experiment 1

Q: How do you keep up with the news?

H* L- H* L-L%

A: # I [READ the DISPATCH]_{Focus} (inappropriate)

2.3.1

Refining models of focus projection

Although Schafer (1997) does not make the correct predictions for the two-peak pattern answers in Experiment 1, the account does show that an account of focus projection must make reference not only to pitch accent distribution but also to prosodic phrasing. The results of Experiment 1 confirm the importance of phrasing and also suggest that a pitch accent's status as nuclear or prenuclear within a phrase is also relevant to focus projection. Future models of focus projection should take these facts into account.

Any account of focus projection must also be able to account for the correct focus interpretation of utterances in emphatic styles of speech. Emphatic speech may be characterized by a high density of intermediate or intonational phrases; every word may be contained in its own phrase.⁴ For example, the emphatic response in (18) may be felicitously uttered with nuclear accents on both the verb and the object NP, even though it has VP-focus and not contrastive focus on the verb and object NP.

(18) Q: How do you keep up with the news?

H* L- H* L-L%

A: I [READ]_{Focus} [the DISPATCH!]_{Focus}

(And that's the 10th time I've told you that today!)

³ Although the VP node syntactically dominates two intermediate phrases, as a projection of the V node it is associated with only one, the intermediate phrase that contains the verb. This association allows the F-marking of the VP. (Amy Schafer, p.c.)

⁴ A single word may even contain more than one phrase: (UN)(BELIEVABLE)

An account of the influence of prosodic phrasing on focus projection must be able to predict the inappropriateness of the double nuclear pattern in the experimental items such as (17) and also the appropriateness of the double nuclear pattern in examples of emphatic speech such as (18). This difference can be accounted for by specifying that in parsing emphatic speech, the listener does not interpret a prosodic boundary as a block to focus projection. If the listener recognizes before encountering the prosodic phrase boundary that the speech is emphatic, focus projection will not be blocked. Expansion of pitch range, high density of prosodic phrasing in previous utterances, and segmental cues (lengthened syllables, failure to flap, for example) *inter alia* might mark an utterance as emphatic. In addition, focus projection may be initially blocked but then allowed if the listener does not immediately recognize that the speech is emphatic.

2.3.2

Accounting for gradience

Recall that listeners gave the experimental items with the two-peak pattern ratings in between those of the worst pattern and those of the best patterns. The gradience of the responses suggests that some factors have stronger effects on focus projection than others. Listeners gave the worst rating to responses with the verb-H* pattern (H* 0 L-L%), indicating that they required New NP objects to be pitch accented. Listeners gave the best ratings to two patterns in which the New NP object was pitch accented, the hat pattern (H* H* L-L%) and the object-NP-H* pattern (0 H* L-L%). The presence of a nuclear pitch accent on the verb in the two-peak pattern (H* L-H* L-L%) rendered the pattern significantly less appropriate than the hat pattern (H* H* L-L%) and the object-NP-H* pattern (0 H* L-L%), but the presence of the required nuclear pitch accent on the object NP also seems to have made the pattern significantly more appropriate than the verb-H* pattern (H* 0 L-L%). This gradience may indicate that while both factors play a role in determining focus structure, the influence of the presence or absence of a pitch accent is stronger than that of an accent's status as nuclear or prenuclear.

Another factor that may have affected the acceptability of the two-peak pattern (H* L-H* L-L%) is the number of unaccented syllables intervening between the accented syllables of the verb and the object NP. It may be more natural to insert a phrase break between accents that are separated by several unaccented syllables, as in (19), where there are three intervening syllables, than in (20), where there are none. In that case, we should see better appropriateness ratings for answers with more intervening syllables. However, no such trend was found for the experimental items, where the number of interaccent syllables ranged from 0 to 4. It seems likely that the number of intervening syllables (or the temporal distance between accents) may play a role in phrasing, but only in cases where there are stretches of unaccented syllables longer than those found in the experimental items.

(19) Q: What kind of car did you rent?

H* L- H* L-L%

A: We RENTED [a MERCEDES]_{Focus}

(20) Q: What phone company do you use?

H* L- H* L-L%

A: I USE [SPRINT]_{Focus}

There is another possible explanation for the gradient pattern of the ratings. It is likely that in completing the appropriateness ratings task, listeners attempted to imagine a larger context in which the given response could be appropriate. As discussed above, the double nuclear pattern may be appropriate in emphatic speech (see (18)), which may be rich in intermediate/intonational phrasing. The experimental questions in combination with the hat pattern or the object-NP-H* pattern responses can be sensibly uttered out-of-the-blue, but the two-peak pattern requires additional context to be judged an appropriate response. The low ratings for the verb-H* pattern suggest that it was impossible for listeners to imagine an additional context that would make that pattern an appropriate response to the given questions.

3 Experiment 2

In Experiment 1, pitch accent type was held constant in order to examine the effects on focus projection of pitch accent position and status as nuclear or prenuclear. Experiment 2 investigated the effect of pitch accent type on focus projection. Although Selkirk (1995) suggests that we may find differences among types of pitch accents, no theory of focus projection predicts that a certain type of pitch accent should be better or less able to project focus. Utterances like (21a), where there is a L + H* accent on the object NP, are therefore predicted to be as appropriate as utterances like (21b), where there is a H* accent on the object NP, as responses to questions calling for focus on the VP.

(21) Q: How do you keep up with the news?

L + H* L-L%

a. A: I [read the DISPATCH]_{Focus}

H* L-L%

b. A: I [read the DISPATCH]_{Focus}

The H* and L + H* pitch accents differ in that the H* is realized with a peak usually within the stressed syllable while the L + H* has a leading L tone near the bottom of the speaker's pitch range and an F0 peak usually occurring late within or after the stressed syllable. These two accents have been discussed under a variety of names in the literature. The L + H* accent has been described as either having a contrastive meaning or evoking a salient scale. For example, Pierrehumbert and Hirschberg (1990) treat each pitch accent as an intonational morpheme and argue that by using a L + H* accent, a speaker "evoke[s] a salient scale ... to convey that the accented item—and not some alternative related item—should be mutually believed" (p. 296). Steedman (1990) also draws a distinction between the pragmatic functions of the L + H* versus H* pitch accents, relating them to concepts such as theme/rheme.

The L + H* pitch accent may also be inherently more salient than the H* pitch accent. Since a H* accent has a gradual slope to its peak, adding another H* accent before a

nuclear H* does not greatly change the basic pattern. On the other hand, if a prenuclear H* is added before a nuclear L + H*, the leading low allows the L + H* to stand out. An earlier exploration of the relative prominence of L + H*, H*, and !H* nuclear accents, however, found a significant difference in prominence between the downstepped nuclear accent (!H*) and the two nondownstepped nuclear accents (L + H* and H*), but no difference between the prominence of L + H* and H* nuclear accents (Ayers, 1996).

Some scholars, including those working within the British school, do not recognize a categorical distinction between the H* and the L + H* pitch accents. Gussenhoven, for example, treats both the H* L–L% pattern and the L + H* L–L% pattern as HL pitch accents. Studies of intertranscriber consistency in ToBI analysis have also shown that the H*/L + H* distinction is one of the distinctions with the most intertranscriber disagreement (Herman & McGory, 2002; McGory, Herman, & Syrdal, 1999). Ladd and Schepman (2003) argue that not only the L + H*, but also the H*, has distinct L and H targets. They suggest collapsing the two categories into one pitch accent (L + H)*. Bartels and Kingston (1994) provide evidence that height of the F0 peak, rather than the presence of a L tone may be the most reliable cue to distinguishing noncontrastive H* from contrastive L + H*.

Experiment 2 was partly motivated by my own intuitions and those of other native speakers that the H* and the L + H* accents might differ in their ability to project focus, specifically that answers with a L + H* accent on the object NP and no other pitch accents were less appropriate as responses to questions calling for VP focus than the same answers uttered with H* accents. I suspected that an answer with a L + H* accent on the object NP was likely to be interpreted as an answer to a question calling for narrow focus on the object-NP focus, as in (22), and would therefore be dispreferred as an answer to a question calling for focus on the VP, as in (23).

(22) Q: Do you read the Lantern?

L + H* L–L%

A: I read [the DISPATCH]_{Focus}

(23) Q: How do you keep up with the news?

L + H* L–L%

A: ?I read [the DISPATCH]_{Focus}

One argument for the inaccessibility of the VP focus reading may follow from Pierrehumbert and Hirschberg's account of the meaning of the L + H* pitch account in which the L + H* invokes a scale of alternatives. It may be easier for the listener to interpret an utterance with a L + H* pitch accent on the object NP as narrowly focusing the object NP and invoking a scale of other noun-referent alternatives than to interpret the utterance as focusing the entire VP and invoking other complex events as alternatives (Mary Beckman, p.c.).

To test a possible difference in focus projection abilities of the accents, the following hypothesis was formulated and tested:

PITCH ACCENT TYPE FOCUS PROJECTION HYPOTHESIS: The L + H* pitch accent is not as effective as the H* pitch accent in projecting focus.

As a control condition, Experiment 2 also included cases in which the New NP-subject of a transitive verb was or was not pitch accented as in (24).⁵ Including the pitch accent position comparison allowed me to verify that subjects were attending to focus structure.

(24) Q: We're arranging carpools to the party. Who needs a ride?

a. A: [LAURIE]_{New} needs a ride.

b. A: # [Laurie]_{New} needs a RIDE.

3.1

Methods

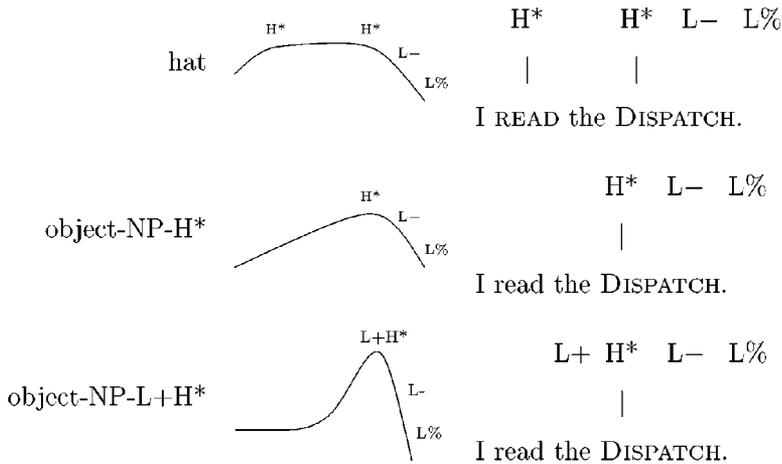
Figure 5

Structure of Experiment 2 accent type comparison questions and answers

One version of question

- VP focus: How do you keep up with the news?

Three versions of each answer

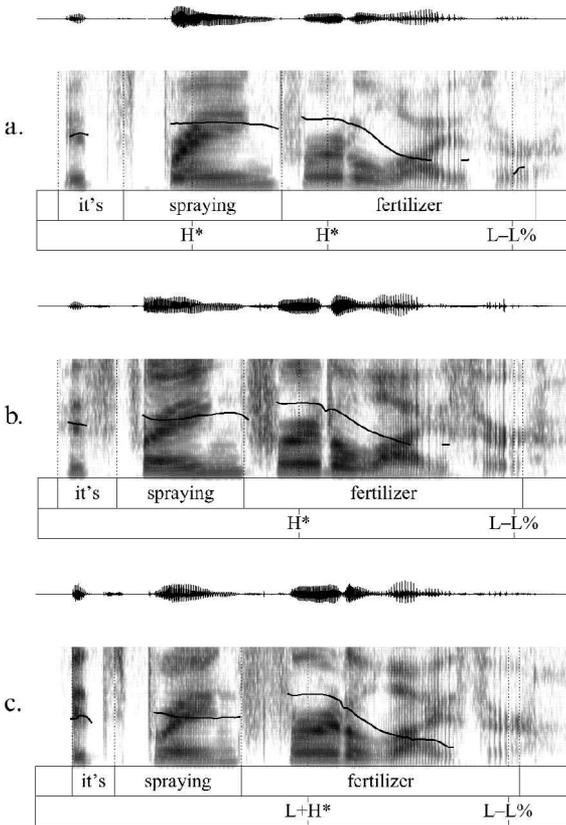


⁵ Utterances of the type included in the experiments require a pitch-accented subject. These utterances were of the form *subject-NP verb object-NP*. It is well known that unaccented New subject NPs are possible in some types of utterances with intransitive verbs (e.g., *Kennedy DIED*.)

3.1.1

Materials

The materials for Experiment 2 were designed to make two comparisons: a critical comparison of pitch accent type and a control comparison of pitch accent position. The materials relevant for the critical comparison consisted of 27 experimental dialogs. All contained questions calling for focus on the verb phrase. The answer in each dialog was recorded with three different intonation patterns, corresponding to the three experimental conditions: the hat pattern ($H^* H^* L-L\%$), the object-NP- H^* pattern ($0 H^* L-L\%$), and the object-NP- $L+H^*$ pattern ($0 L+H^* L-L\%$). An example set with

**Figure 6**

Example F_0 tracks for the three conditions of Experiment 2

stylized pitch curves is given in Figure 5. Figure 6 shows F_0 tracks for each of the three conditions.

The materials for the control comparison of pitch accent position consisted of a total of 18 dialogs, each of which contained a question calling for focus on the subject NP. The answer to each of these dialogs was recorded with two different intonation patterns, the subject- H^* pattern (H^* on the subject NP, $H^* 0 L-L\%$) and the object-NP- H^* pattern (H^* on the object NP, $0 H^* L-L\%$). Examples are given in Figure 7.

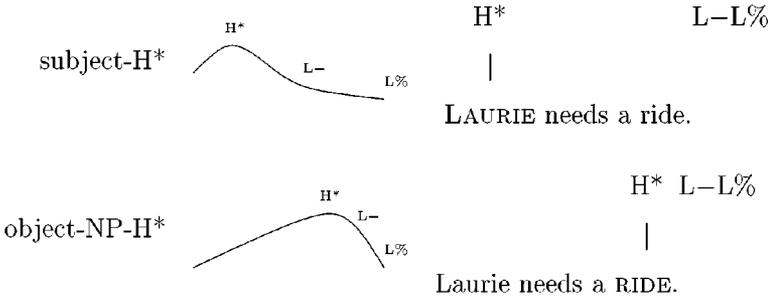
Figure 7

Structure of Experiment 2 accent position comparison questions and answers

One version of each question

- subject-NP focus: We're arranging carpools to the party. Who needs a ride?

Two versions of each answer



The questions and answers were recorded separately in a sound-attenuated chamber by two female native speakers of American English.⁶ As in Experiment 1, the speaker of the questions was instructed to speak as naturally as possible and the intonation of the questions was not manipulated. The speaker of the answers, an expert in the ToBI transcription system, was asked to read ToBI-labeled written versions of each sentence. The speech was digitized at 22.050kHz, each repetition of an answer was saved as a separate file, and F0 was calculated. The answers were then ToBI-transcribed by an expert labeler. The transcriber was the speaker of the answers, so steps were taken to minimize the chances that she would be influenced by her knowledge of her intentions during the recording. The file names were a series of numbers that did not indicate the intended ToBI pattern of the answer, and the labeler worked through the files in a randomized order. Items for which the ToBI transcription matched the intended intonation pattern were selected. When there was a choice between more than one repetition of an answer, an effort was made to choose answers with accents whose F0 peaks were fairly similar within a set. This was not always possible, and there was often variation in the height of the peaks across the three conditions. Three tapes were made, with each tape containing nine dialogs of each answer type for the accent type comparison (for a total of 27) and

⁶ These two speakers were different from the speakers in Experiment 1.

six dialogs of each answer type for the accent position comparison (for a total of 12). Each tape contained exactly one version of each dialog. The materials also included a total of 15 filler items: five contained questions with object-NP focus questions and answers with the object-NP-H* pattern (0 H* L-L%); five contained questions calling for contrastive focus on the object NP and answers with the object-L + H* pattern (0 L + H* L-L%); five contained questions calling for verb focus and answers with the object-NP-H* pattern. There was a fixed order of dialogs and the order of the conditions was pseudorandomized.

3.1.2

Subjects

Thirty undergraduate students at the Ohio State University (10 per tape) participated in Experiment 2. All subjects met the requirements described in the Methods for Experiment 1.

3.1.3

Procedure

Subjects performed the same appropriateness rating task as in Experiment 1. As in Experiment 1, subjects were explicitly instructed to attend to intonation (or “how it is said”). The instructions given differed only minimally from the instructions given in Experiment 1 (described in the Methods for Experiment 1). In Experiment 2, subjects were instructed to use the entire scale, although the range of responses for the two experiments was comparable.

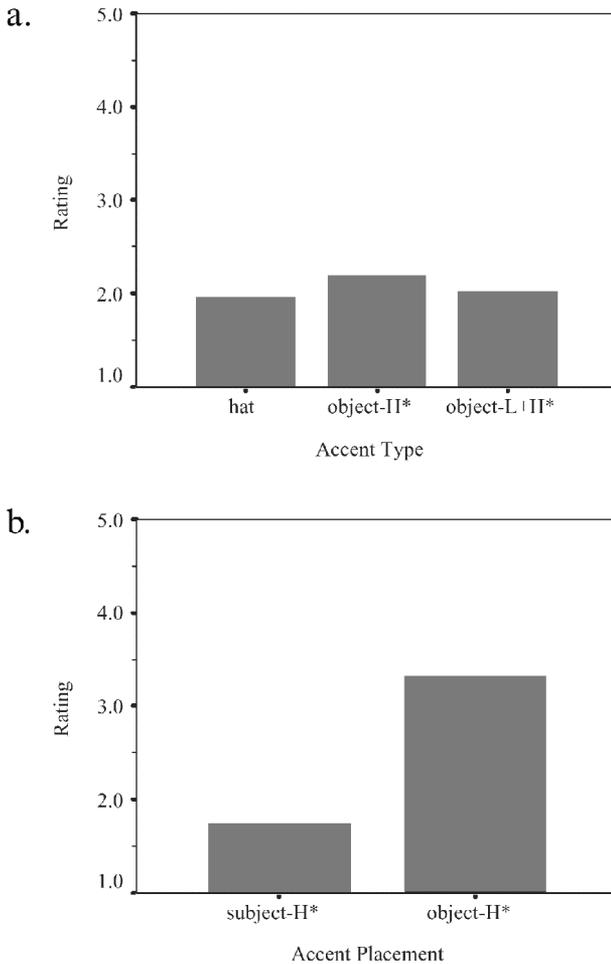
There was a short practice session in which a total of four dialogs, all with questions calling for focus on the object NP, were heard. The first two dialogs differed only in accent position of the answer. The last two dialogs differed in accent type (H* vs. L + H*). As with Experiment 1, no feedback was given on whether any of the intonation patterns presented were appropriate or inappropriate.

3.1.4

Predictions

Since neither theory restricts its claims to one accent type, both Selkirk and Gussenhoven predict that all three patterns, the object-NP-L + H* pattern (0 L + H* L-L%), the object-NP-H* pattern (0 H* L-L%), and the hat pattern (H* H* L-L%) should be equally able to project focus in the VP focus question condition. We therefore should expect no difference in appropriateness judgments for any of the three patterns. According to the Pitch Accent Type Focus Projection Hypothesis, however, answers with the object-NP-L + H* pattern (0 L + H* L-L%) are predicted to be inappropriate as responses to questions calling for focus on the VP.

For the control accent position condition, we expected that subjects would find answers with the object-NP-H* pattern (0 H* L-L%) inappropriate as responses to questions calling for focus on the subject NP.

**Figure 8**

Appropriateness judgments for the accent type and accent position of Experiment 2. (1 is best; 5 is worst)

3.2 Results

The results of Experiment 2 are charted in Figure 8. Recall that all the items in the control accent position comparison contained a question calling for focus on the subject NP. Subjects gave a mean rating of 1.744 to answers in the NP-subject-H* condition (H* 0 L–L%) and a mean rating of 3.317 to answers in the object-NP-H* (0 H* L–L%) condition, indicating a strong preference for New subject NPs to be pitch accented (a numerically lower rating on the 1–5 scale indicates that a pattern was found to be more appropriate). The results of ANOVA indicated that this result was significant by subjects, *F*(1, 29), $p < .001$.⁷ As expected, subjects preferred that a New subject NP be pitch accented. The positive results of this comparison show that subjects were attending to focus structure.

⁷ The 18 accent position dialogs were rotated among the three subject groups, with each group hearing 12 accent position dialogs. Each group therefore heard slightly different dialogs, so an items analysis for the control condition was not possible.

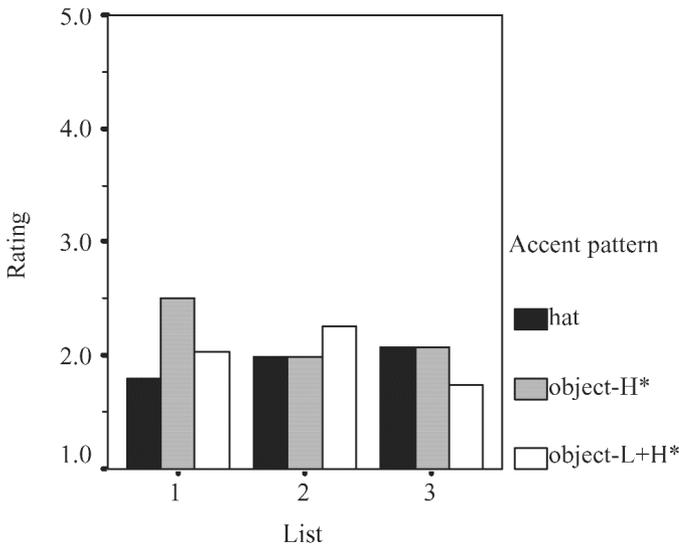


Figure 9
Interaction
between accent
and list in VP
focus cases,
Experiment 2

The Pitch Accent Type Focus Projection Hypothesis, however, which predicted that subjects would disprefer the object-NP-L + H* pattern (0 L + H* L – L%) as a response to a question calling for focus on the entire VP, was not supported. In general, subjects gave numerically low ratings to all of the critical accent patterns, indicating that all patterns were considered appropriate. Subjects gave a mean rating of 1.948 to the hat pattern (H* H* L – L%) and a mean rating of 2.011 to the object-NP-L + H* pattern (0 L + H* L – L%). Somewhat surprisingly, the object-NP-H* pattern (0 H* L – L%) received the highest (worst) rating, 2.185. Results of ANOVA did show a main effect of accent pattern by items, $F_2(1, 27) = 3.843, p < .05$, and a marginal effect of accent pattern by subjects $F_1(1, 29) = 3.080, p = .058$. Results of planned Scheffe pairwise comparisons showed that only the ratings for the hat pattern (H* H* L – L%) and the object-NP-H* pattern (0 H* L – L%) differed significantly ($p < .05$). The ANOVA, however, also showed a significant interaction between accent pattern and list ($p < .05$) by subjects.

3.3

Discussion: Experiment 2

The results of the accent position comparison confirm a strong preference for accented subject NPs in answers with subject focus, but do not provide support for the hypothesis that the L + H* accent does not project focus. It is possible that a normal peak height difference between the L + H* and the H* accent was suppressed in the selection of the materials. Bartels and Kingston (1994) argue that peak height is the most reliable cue to a contrastive pitch accent. In attempting to match peak height as closely as possible within a set, I may have excluded L + H* accents with F0 peaks high enough to clearly signal contrastive focus.

The small, but significant difference (0.237, 1–5 scale) between the object-NP-H* pattern (0 H* L – L%) and the hat pattern (H* H* L – L%) must be interpreted with caution in light of the interaction found between accent pattern and list (see Fig. 9). An examination of the subject means for each list showed that a preference for the hat pattern

over the object-NP-H* pattern was limited to subjects who heard List 1. List 1 subjects gave a mean rating of 1.79 to the hat pattern (H* H* L-L%) items and a higher mean rating of 2.5 to object-NP-H* (0 H* L-L%) items (indicating a preference for the hat pattern). They gave a mean rating of 2.03 to the object-NP-L + H* (0 L + H* L-L%) items. List 2 and List 3 subjects, however, did not make a distinction between ratings for the hat pattern and ratings for the object-NP-H* pattern.

Clearly, something about the particular items that were encountered in the object-NP-H* pattern (0 H* L-L%) in List 1 made subjects disprefer that pattern. There are a number of possible explanations for this tendency. It may be that listeners imposed focus structures other than the intended focus structures on the experimental items. Consider, for example, the List 1 item *Q: I heard that guy down the block got arrested. What did he do? A: He committed murder.* A listener may prefer a response where there is contrastive focus on *murder* either to highlight the magnitude of the crime or because the listener imagines a context where the individual in question has been arrested before for some other offense—*this time he didn't just commit burglary, he committed murder.* In addition, some experimental participants reported that they found the object-NP-H* pattern to be unemotional and callous. In reviewing the materials in List 1, there are a number of items which address serious topics where this interpretation would be particularly inappropriate (e.g., the item discussed above). It is possible that for some of the items, listeners may have parsed the object-NP-H* pattern as L* H* L-L%, with a L* on the verb. If so, they might have interpreted the pattern as a contour discussed in Sag and Liberman (1975) as the surprise-redundancy contour. Since this contour conveys the idea that what is expressed ought to be known or obvious to the addressee, it would likely be interpreted as callous in responses to the informational questions in the experimental items.

Whatever the explanation for the interaction, any effect of pitch accent type on focus projection (at least for the H* vs. L + H* comparison) is sensitive to textual differences, in contrast to the robust effect found for pitch accent position and status as nuclear or prenuclear.

4 Conclusion

Taken together, the results of the two experiments suggest that in determining the focal structure of an utterance, listeners are most sensitive to the presence or absence of pitch accents in expected positions. The results of Experiments 1 and 2 showed that a pitch accent on the object NP can effectively project focus to the entire VP, as predicted by two theories of focus projection. The results of Experiment 1, however, also showed that listeners may be sensitive to the status of a pitch accent as nuclear or prenuclear. The results are consistent with a model in which only a nuclear pitch accent, a pitch accent that is the last accent in its intermediate phrase, marks focus. The presence of a phrase boundary between an accented verb and its accented NP complement may block projection of focus to the VP, although such projection would likely not be blocked in an emphatic speaking style. Follow-up studies are needed to examine this interpretation. Finally, the influence of pitch accent type, if present at all, is slight and relatively variable at least as measured in an off-line preference task. These results support the findings of Ayers (1996) that listeners judge nuclear accents to be more prominent than pre-

nuclear accents, but apparently do not distinguish between H* and L + H* nuclear accent types. The results also support the Beckman (1996) claim that only nuclear accents mark focus, and they add to Schafer's (1997) empirical evidence that prosodic phrasing influences focus projection.

There are a number of directions for future research on focus projection. Further work is needed to determine exactly which aspect or aspects of prosody (tonal, phrasal, or pausal information, for example) account for the effects found in Experiment 1 and in Schafer, 1997.

While the results are consistent with the hypothesis that nuclear pitch accents, but not prenuclear pitch accents, mark focus, follow-up studies are needed to confirm this interpretation, as noted in the discussion of Experiment 1. If further evidence supports this claim, its consequences need to be investigated. For example, both Gussenhoven's theory and Selkirk's theory predict that the two-peak pattern and the hat pattern should be equally effective in marking a double focus structure. The results of Experiment 1, however, suggest that the prenuclear pitch accent on *read* in (25a) will not effectively mark focus, while the nuclear accent in (25b) will.

(25) Q: Do you read the Dispatch?

a. A: # I [READ]_{Focus} [the DISPATCH]_{Focus} but I [don't ENJOY]_{Focus} it.

b. A: I [READ]_{Focus} [the DISPATCH]_{Focus} but I [don't ENJOY]_{Focus} it.

Most experimental work on focus projection has examined focus projection within the verb phrase (e.g., Birch & Clifton, 1995; Schafer, 1997, and this study), but our knowledge of the phenomenon would be extended by looking at other structures. For example, it has been observed that the nonfocused NP subject of a sentence is often pitch accented and even set off in its own prosodic phrase (e.g., Nakatani, 1997). For example, in (26), the (a) response, where the Given subject *John* is pitch accented, seems more natural than the (b) response, where the Given subject is not pitch accented.

(26) Q: What can you tell me about John?

a. A: Oh, JOHN's a GREAT GUY.

b. A: Oh, John's a GREAT GUY.

Speakers and listeners may treat subjects differently with respect to pitch accent assignment and focus interpretation, or the apparently different treatment may fall out of more general facts, such as a tendency to accent sentence-initial material. But for the moment, the question is open.

Experimental studies using constructed stimuli have looked at focus projection, but there is also a need to develop and examine spoken language corpora annotated for intonational and discourse structure. Such studies would allow us to evaluate the extent to which the accent patterns found in natural language support the theoretical predictions. These studies should address not only pitch accent position but also pitch accent type and status in the intonational hierarchy.

Finally, findings on the naturalness of intonation patterns have practical applications for language technology. For example, the finding that listeners judge patterns with a prenuclear pitch accent on a Given verb to be appropriate might motivate an addition to the intonation module of a text-to-speech synthesis system, a specification that assigns pitch accents to verbs before some accented noun phrases. This might introduce more variety and thus more naturalness into the system.

*Received: August 28, 2001; revised manuscript received: September 22, 2002;
accepted: December 03, 2002*

References

- AYERS, G. (1996). *Nuclear accent types and prominence: Some psycholinguistics experiments*. Unpublished doctoral dissertation. The Ohio State University, Columbus.
- BARTELS, C., & KINGSTON, J. (1994). Salient pitch cues in the perception of contrastive focus. In P. Bosch & R. van der Sandt (Eds.), *Focus and natural language processing*, Vol. 1: *Intonation and syntax* (pp. 1–10). IBM Deutschland Informationssysteme GmbH Scientific Center, Institute for Logic and Linguistics. (= Working Papers of the Institute for Logic and Linguistics, 6).
- BECKMAN, M. (1996). The parsing of prosody. *Language and Cognitive Processes*, **11**, 17–67.
- BECKMAN, M., & HIRSCHBERG, J. (1994). The ToBI annotation conventions. Unpublished manuscript. The Ohio State University and AT&T Bell Telephone Laboratories.
- BECKMAN, M., & PIERREHUMBERT, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, **3**, 255–310.
- BIRCH, S., & CLIFTON, C., Jr. (1995). Focus, accent and argument structure: Effects on language comprehension. *Language and Speech*, **38**, 365–391.
- BROWN, G. (1983). Prosodic structures and the Given/New distinction. In D. R. Ladd & A. Cutler (Eds.), *Prosody: Models and measurements* (pp. 67–77). Berlin: Springer.
- CHAFE, W. (1983). *Meaning and the structure of language*. Chicago: University of Chicago Press.
- GUSSENHOVEN, C. (1983). Testing the reality of focus domains. *Language and Speech*, **26**, 61–80.
- GUSSENHOVEN, C. (1984). *On the grammar and semantics of sentence accent*. Dordrecht / Cinnaminson: Foris.
- GUSSENHOVEN, C. (1992). Sentence accents and argument structure. In I. Roca (Ed.), *Thematic structure: Its role in grammar* (pp. 79–106). Berlin/New York: Foris.
- GUSSENHOVEN, C. (1999). On the limits of focus projection in English. In P. Bosch & R. van der Sandt (Eds.), *Focus: Linguistic, cognitive, and computational perspectives* (pp. 43–55). Cambridge, U.K.: Cambridge University Press.
- HERMAN, R., & MCGORY, J. T. (2002). The conceptual similarity of intonational tones and its effects on intertranscriber reliability. *Language and Speech*, **45**, 1–36.
- KRAHMER, E., & SWERTS, M. (1998). Reconciling two competing views on contrastiveness. In *Proceedings of the International Conference on Spoken Language Processing* (pp. 1299–1302), Sydney.
- LADD, D. R., & SCHEPMAN, A. (2003). “Sagging transitions” between high pitch accents in English: Experimental evidence. *Journal of Phonetics*, **31**, 81–112.
- MCGORY, J. T., HERMAN, R., & SYRDAL, A. (1999). Using tone similarity judgments in tests of intertranscriber reliability. *Journal of the Acoustical Society of America*, **106**, 2242 [abstract].
- NAKATANI, C. (1997). *The computational processing of intonational prominence: A functional prosody account*. Unpublished doctoral dissertation. Harvard University, Cambridge, MA.

- OSTENDORF, M., PRICE, P., & SHATTUCK-HUFNAGEL, S. (1997). *Boston University radio speech corpus*. (CD-ROM.) Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- PIERREHUMBERT, J., & HIRSCHBERG, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT.
- PIERREHUMBERT, J. (1980). *The phonology and phonetics of English intonation*. Doctoral dissertation. Massachusetts Institute of Technology. Distributed (1987) by the Indiana University Linguistics Club.
- PRINCE, E. (1981). Toward a taxonomy of Given-New information. In P. Cole (Ed.), *Radical pragmatics* (pp. 223–255). New York: Academic Press.
- SAG, I., & LIBERMAN, M. (1975). The intonational disambiguation of indirect speech acts. *Proceedings of the Chicago Linguistics Society*, **11**, 487–497.
- SCHAFER, A. (1997). *Prosodic parsing: The role of prosody in sentence comprehension*. Unpublished doctoral dissertation. University of Massachusetts, Amherst.
- SELKIRK, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT.
- SELKIRK, E. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 550–569). Cambridge, MA: Blackwell.
- STEEDMAN, M. (1990). Syntax and intonational structure in a combinatory grammar. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 457–482). Cambridge, MA: MIT.
- TERKEN, J., & HIRSCHBERG, J. (1994). Deaccentuation of words representing “given” information: Effects of persistence of grammatical function and surface position. *Language and Speech*, **37**, 125–145.
- VALLDUVÍ, E. (1992). *The informational component*. New York: Garland.
-