FPP	future participle suffix
PRS	present tense suffix
PRT	preterite tense suffix (past tense meaning)
1S	1st person singular suffix ('I', 'my')
2S	2nd person singular suffix ('you', 'your')
3S	3rd person singular suffix ('he'/' she'/' it', 'him'/' her', 'his'/' her'/' its')
1P	1st person plural suffix ('we', 'us', 'our')
PL	plural suffix
OBJ	object case suffix (used for definite objects)
DAT	dative case suffix ('to', 'for')
LOC	locative case suffix ('in', 'on')
ABL	ablative case suffix ('from', 'of')
COM	comitative suffix ('with')
QPC	question particle
NEG	negative suffix ('not')

An exploration of prosody and turn projection in English conversation*

Barbara A. Fox
University of Colorado at Boulder

The present study explores the possibility that pitch-accented syllables which project upcoming turn completion may be phonetically different from pitch-accented syllables which do not so project. This possibility was raised by Schegloff (1987b) with regard to "peak" accents.

This study reports on various phonetic details of the pitch peaks of 43 utterances, which were culled from naturally-occurring American English conversations. Each utterance selected had two pitch peaks in the predicate of its syntactic structure. The goal of the study was to determine if, in fact, there were phonetic differences between "non-last" accented syllables (that is, those which do not project upcoming turn completion) and "last" accented syllables (that is, those which do project upcoming turn completion). A variety of phonetic details were selected, most of which are implicated in the literature on prosody in the creation of prominence. It was hypothesized that "last" accents would be phonetically more prominent than "non-last" accents.

All of the findings except one fail to support this hypothesis. While several of the phonetic properties examined showed trends in the expected direction, only one, Duration, showed a statistically significant difference between "non-last" and "last" accents.

1. Introduction

Although the Sacks, Schegloff and Jefferson paper on turn-taking is now a quarter of a century old (Sacks, Schegloff and Jefferson 1974), research into the *linguistic* underpinnings of the proposed mechanisms for turn-taking is a fairly recent enterprise. For example, aside from Goodwin's early work (e.g. C. Goodwin 1979, 1981), Ford and Thompson (1996) is one of the first studies I know of to re-examine the SSJ turn-taking model from a more rigorously linguistic perspective. The virtual explosion of research within the last few

years on this topic (e.g. Ford, Fox and Thompson 1996a and b, to appear; Fox, Hayashi and Jasperson 1996; Hayashi, Mori and Takagi, to appear; Hayashi and Mori 1998; Lerner and Takagi 1999; Ford and Mori 1994; Tanaka 1996, 2000; Selting 2000; Couper-Kuhlen and Selting, eds. 1996; Schegloff 1996, 1998; Lerner 1991, 1996; Ochs, Schegloff and Thompson, eds. 1996) is thus testimony to the growing awareness among linguists that turn-taking is both organized by and organizing of linguistic practices and hence deserving of linguists' attention.

Nevertheless, in spite of this growing attention to turn-taking (and other facets of conversational organization), there remain large gaps in our understanding of exactly how turn-taking is achieved. The need for further linguistic research is especially evident in the area of prosody and turn-taking.

While quite a bit has been done on the prosody of turn-final syllables, or even final words (e.g. Duncan 1974; Duncan and Fiske 1977; Wells and Peppe 1996; Lehiste 1979; Local, Wells and Sebba 1985; Local, Kelly and Wells 1986; Berkovits 1984a, b), very little has been done on how prosody might project upcoming turn completion from earlier in the turn.

Perhaps not surprisingly, it is Schegloff (cf. Schegloff 1987a and b, 1996, 1998) who has made the only proposal I know of with regard to turn-taking and prosody *before* the last syllables of the turn. Schegloff (1987a) proposes the following relationship between prosody and turn-taking in discussing the last accented syllable in the utterance *he's about the only good regular out there*:

> The relevance of a pitch peak of this sort — but mark, not of all pitch peaks — is that it marks the enhanced likelihood that the next possible completion of the turn constructional unit will be an actually intended turn completion. That is, the developing grammatical structure of an utterance in the course of its production is potentially compatible with alternative points of possible completion. Pitch peaks, and their suppression, are one means by which speakers can indicate which syntactically possible completions are built to be completions on this occasion and which not. A pitch peak can thus project intended turn completion at the next grammatically possible completion point. (1987b, transcribed from Xerox Parc lecture)

When asked during the question and answer period after this public lecture why the pitch peak on *good* doesn't also project in this way, Schegloff said he didn't know. Clearly, if Schegloff is correct, some pitch peaks project upcoming completion and others do not. But we do not yet know what distinguishes the former from the latter.

Schegloff's proposal thus opens up a fascinating arena of research which

very much requires contributions from linguists. The larger arena asks: How are prosodic practices in a given language implicated in turn projection? The more narrow focus of Schegloff's proposal asks: What distinguishes pitch peaks (a term which will be discussed more fully below) which project upcoming turn completion from pitch peaks which do not so project?

It is the goal of the present paper to begin an exploration of the more narrowly focused question. This question could have been approached in a number of different ways. In the present paper I have adopted one particular approach to the question, one which naturally arises from the particular utterance for which Schegloff's proposal was offered. This approach compares multiple pitch peaks within a single utterance to determine if there are phonetic differences between what we can for now call the "non-projecting" pitch peak(s) and the "projecting" pitch peak.

Other approaches to the question would have yielded different insights. For example, it would have been possible to make a collection of utterances with a single pitch peak which appears to project upcoming turn completion at the next place of grammatical completion, and contrast those cases phonetically with cases in which the single pitch peak does not appear to project upcoming completion (in spite of being at or near the end of the turn). Or, it would have been possible to make a collection of utterances in which each exhibits a pitch peak that the recipient analyzes as projecting upcoming turn completion but where the speaker displays that analysis to be incorrect; a phonetic study of such pitch peaks could then be compared to pitch peaks for which the recipient displays an understanding of "no upcoming completion." Contrasting pitch peaks with other types of accented syllables could also be done.

The current study focuses exclusively on prosody in order to address a fundamental, and obviously naive, hypothesis, which is that the phonetic realization of a syllable indexes projection *by itself*, without regard to the syntactic, semantic, or sequential position in which that syllable is placed. That the study finds little evidence to support this hypothesis — and, indeed, may be compatible with quite a different hypothesis regarding the relationship of prosody to projection — is hardly surprising, given what we know about the complex constellation of practices involved in turn projection (see, for example, Ford, Fox and Thompson 1996a). Nonetheless, there were two motivations for exploring this hypothesis in detail: first, there is a great deal of literature in the field of intonation which holds that there are two kinds of accented syllables, nuclear and pre-nuclear accents — which as I will discuss below may be seen to bear a relationship to "projecting" and "non-projecting"

accents — and that these nuclear and pre-nuclear accents exhibit distinct phonetic realizations. This claim would seem to strongly support the possibility that speakers can index through prosody alone whether an accented syllable is designed to be heard as the last accented syllable of the turn or as not the last accented syllable of the turn. Second, it was necessary to explore the role of prosody by itself so that the next step can be taken, which is to explore the phonetic realization of syllables in different syntactic, semantic, and sequential locations. As we will see, the results of the study suggest that phonetic realizations of accents seem to show too much variation to index projection by themselves. Rather, they are very likely heard in and through the syntactic, semantic, and sequential positions in which they occur. For example, a very prominent syllable may not be heard as being a "last" accent (thereby projecting upcoming completion of the turn) if it occurs in a position which is not syntactically or semantically possibly complete. In such cases, the accent may not be "interrogated" (Schegloff, pc) for whether it is doing projection or not, which leaves that particular prosodic movement free to accomplish other work (showing contrast, doing repair, etc.).

The present study asks the following question: Are "last" and "non-last" pitch peaks phonetically distinct (in particular, are "last" pitch peaks more prominent than "non-last" pitch peaks?)? That is, can a recipient tell from hearing the phonetic details of a particular pitch peak that that syllable will not be the last accented syllable in the turn? Or that it will be the last accented syllable in the turn? The motivation for asking the question in terms of "last" and "non-last" (versus, for example, "projecting" and "non-projecting") will be discussed in detail in Section 2.

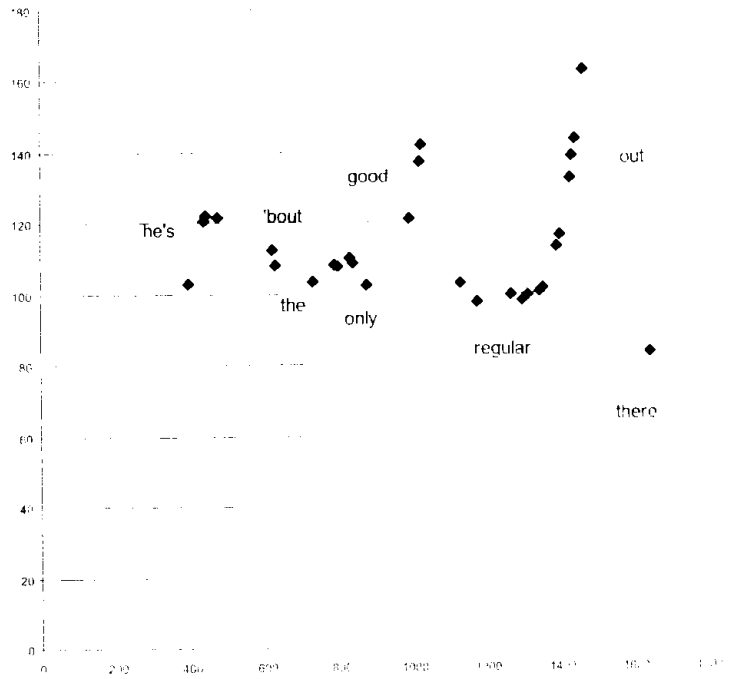
Let us now turn to developing the questions a bit more thoroughly.

1.1 Pitch peaks and turn projection

Schegloff's original proposal for the relationship between pitch peaks and turn projection was developed in connection with the analysis of a particular utterance. It may therefore be helpful here to consider that utterance in more detail (the utterance in question is given below in bold face):

(1)

If we listen to this utterance, we can notice several things. First, there are two accented syllables in it, *good* and *out* (and they are obviously "peaks"). These can be seen clearly in the pitch track given below:



Pitch Track 1

Second, one of these accented syllables occurs towards the end (although not exactly at the end) of the utterance and the other does not. Third, there are some interesting phonetic differences between the two accented syllables. Consider the table below, which displays certain aspects of the phonetic realization of each word.

From this table we can see that *out* is longer than *good*, shows an earlier peak alignment, has a steeper slope, a larger movement of pitch, and a larger step down from it to the next syllable. If we take this case as an example of an utterance which displays one pitch peak that projects upcoming completion of

Table 1 Phonetic Properties of *Good* and *Out*

	good	out
syl length	119 ms	155 ms
word length	119 ms	155 ms
peak alignment	0.91	0.84
slope	0.61	0.71
change in pitch	20 hz	50 hz
stepupto	20 hz	11 hz
stepdownfrom	38 hz	80 hz
amplitude	-14.96 db	-16.55 db

the turn (the one on *out*) and another that does not (the one on *good*),[1] then we might be encouraged to conclude that, indeed, there are clear phonetic differences between pitch peaks which project and those which do not. In particular, based on the figures given above, it would not be out of place to suggest that the difference is largely, though perhaps not entirely, one of prominence — the projecting pitch peak seems to be more prominent than the non-projecting pitch peak.[2]

We might be further encouraged in this hypothesis by the tradition followed by British intonationists, noted above, which distinguishes pre-nuclear and nuclear accents, where nuclear accents are said to come last in their tone group and to be more prominent than pre-nuclear accents (e.g. Cruttenden 1997; Crystal 1969). Although the exact relationship between pre-nuclear/nuclear tones and non-last/last accents remains to be worked out, it seems plausible that pre-nuclear tones project in general that another, nuclear, tone is still to come, and that nuclear tones project the upcoming completion of the utterance. If this is in fact the case, then we may be able to equate pre-nuclear tones with non-last accents and nuclear tones with last accents. While claims regarding the distinctions between pre-nuclear and nuclear tones are rejected by American intonationists (e.g. Pierrehumbert 1980; Silverman and Pierrehumbert 1990), we can make use of them for now as suggesting that the hypothesis tested by the current study is not completely unrealistic.

This is thus the main guiding question I have explored in the present paper. Are "last" pitch peaks more prominent phonetically than "non-last" pitch peaks in the same utterance?

Although the real question I am interested in has to do with real-time turn projection, in order to keep the project manageable, I chose an approach which is somewhat more static in its treatment of accents and turns. That is, the study

I will report on here actually compares the phonetic properties of non-last and last accents in a turn, on the important assumption that the terms non-last and last relate to turn projection in some precise way. For example, I assume that non-last accents project that the turn is designed to be heard as *not* complete at the next place of grammatical completion, while last accents project that the turn is designed to be heard as complete at the next place of grammatical completion.

There are a range of problems with these assumptions — for example, it is not easy to specify the exact relationship between the last accent and the end of the utterance; however, space considerations keep me from exploring these issues in detail here. In spite of the problems, however, I believe that the approach I have taken in the current study represents one reasonable way to begin an exploration of this very complex area.

1.2 The present study

The study I will report on here begins to answer the questions posed by examining a somewhat small corpus (43) of utterances, culled from phone conversations and face-to-face videotaped interactions, among native speakers of American English, of clauses with two or more pitch accented syllables, where each pitch obtrusion could be considered a "peak" (details and definitions will be provided later), and where one of those accented syllables was close to or at the end of the turn. For the rest of the report I will use the term "pitch" to mean f_0. I created pitch tracks for all of these utterances, and coded each "pitch peak" for a variety of phonetic properties. I then compared what appeared to be "last" accented syllables with "non-last" accented syllables for these phonetic properties to see if there were indeed phonetic differences between such syllables. I reasoned that if there were phonetic differences between "last" and "non-last" (or: those that project upcoming completion and those that do not), then Schegloff's proposal would be supported in a very obvious way, whereas if I found no phonetic differences, we might suppose at least three possible alternative explanations: (a) Schegloff's proposal is not correct; (b) Schegloff's proposal covers only a subset of utterances; or (c) the phonetic details of accent can only be understood within their emerging syntactic, semantic and sequential environments.

It should be immediately clear how many problematic pieces there are to such an endeavor. For example, it is very difficult to determine which syllables should count as accented. Moreover, it is not at all a trivial matter to decide what

the unit should be within which there are to be two (or more) accented syllables: an intonation unit, a clause, an utterance? In addition, it is well known that the phonetic details of a particular accented syllable may bear only an indirect relationship to the way in which that syllable is perceived in context (see Pierrehumbert 1979; 't Hart, Collier and Cohen 1990; Ladd 1996); it is thus not obvious that exploring the phonetic details of accents is a profitable line of inquiry. It is also not clear that it is wise to compare accents which may be doing very different kinds of semantic work (for example, indicating the "focus" of the utterance vs. indicating some other kind of discourse information; see Brown 1983; Cullicover and Rochemont 1983; Gussenhoven 1984; Lambrecht 1994). These are all important questions and I am not able to explore their answers here. For this reason I consider this study to be exploratory.

In spite of all of these problems, the project seems significant to me on several levels. First, for those of us interested in turn projection, that is, the practices used to display and interpret how and when a turn might come to possible completion intonation, is clearly crucial (see Ford, Fox and Thompson 1996a). Second, the details of prosody and turn projection have as yet remained largely unexplored. Third, although I am unable to pursue the question further in the current report, the findings of this study may have a bearing on the debate regarding differences between pre-nuclear and nuclear accents alluded to above.

The remainder of this report is organized as follows. Section 2 describes the major categories used in the study. Section 3 describes the database and how the collection was created. Section 4 describes the phonetic properties that were coded. Section 5 provides the results of the analyses, and Section 6 concludes with a discussion of possible explanations and implications for the findings.

Now let us turn to the definitions employed in the study.

2. Definitions

2.1 What is a pitch peak?

In Schegloff's proposal, the term "pitch peak" is used. While the term "peak" is used in a variety of works (e.g. Liberman and Pierrehumbert 1984), it is not defined precisely. I based my definition on the definition of H* given in Ladd (1996). The definition I used follows:

> a pitch peak is an accented syllable which is (a) stepped up to, or (b) if stepped down to, then a significant glide occurs such that the highest pitch on the accented syllable is higher than the highest pitch of the previous syllable.

Although this study confines itself to pitch peaks, as defined above, it is reasonable to ask if downward pitch obtrusions might not project in just the same way as upward pitch obtrusions. The question is an obvious one to pursue further; nonetheless, for the present study I have confined my exploration to upward pitch excursions, in part to limit the scope of the study.[3] Future work on this topic must expand the scope to include downward pitch obtrusions.

2.2 The notion of accent

Intonationists tend to agree that an accented syllable in English (at least of the nuclear sort) exhibits, in addition to some kind of pitch movement to or from (and perhaps on) that syllable, greater amplitude and duration than non-prominent syllables. But such phonetic properties are scalar: when do we consider a syllable to have reached a sufficient degree of these properties to achieve "accent" status? This is not at all a trivial concern when one uses conversational data, since a syllable may show, for example, a step up to it and a step down from it, but may sound only somewhat prominent. Should such a syllable be included in this study? How prominent does it have to be?

Given that upward pitch movements are not necessarily "peaks", I needed a way to distinguish certain upward pitch excursions from others, and that way has turned out to require some appeal to prominence created by pitch, duration and amplitude. I will thus use the term *accent* in the rest of this report. I use it to mean a pitch-accented syllable — that is, a syllable made prominent through pitch movement, as well as through increases in duration and amplitude.[4] I excluded from this study syllables which sounded prominent but which lacked pitch movement that could be identified as creating a "peak".

2.3 The units

To compare a non-last and last accent, we must determine within what domain those two accents must occur. There were three obvious choices for unit: Turn Constructional Unit (see Sacks, Schegloff and Jefferson 1974); Clause, and Intonation Unit (see Chafe 1987; also known as tone group, intonational phrase, breath group, among others). For the purposes of the current study I

chose to work with the unit "clause"; space considerations prevent me from discussing the motivations for this choice here.

3. Data

The utterances included in my corpus come from many different audiotaped phone conversations and videotaped face-to-face conversations. The native language of all participants from whom I selected examples was standard American English. A total of 17 conversations form the source of my corpus of instances, representing a range of ages, regional accents, education level, and social class. Included in those conversations are 7 conversations from the American English CallHome database.[5]

I listened to every utterance from all of these conversations using an audiotaped version of the recording, playing the audiotape through my computer speakers. When I came across an appropriate utterance, I digitized it using a speech analysis program,[6] and then listened to it repeatedly to make sure it fit the pattern I was looking for. When possible I made use of CSpeech's automatic pitch tracking function to create a pitch track just so that I could be sure the utterance really fit my collection. I then created a pitch track "by hand", that is I used the energy wave to measure distance between cycles.

The collection of utterances for the present study consists of 43 utterances that were not syntactically interrogatives or pragmatically questions. These instances show two or more pitch peaks (as defined above) within a single clause. I excluded accents on grammatical subjects from this study, so the two or more pitch peaks had to fall within the predicate (which I defined as everything that was after the subject). The motivation for this decision comes from the fact that subjects are very unlikely to be accented in spontaneous English conversation, and hence when they are, the accent is going to be doing something very special (see, for example, Lambrecht 1994). I allowed an infinitival complement clause to be considered part of the same clause as its matrix clause, if the matrix clause indicated mood or aspect of some sort (as in: *I'd really like to take my bike out this summer*). The motivation for this was the mergedness of the two clauses: it has been noted that infinitival clauses of this sort could well be analyzed as being closer to single clauses than complex clauses (Thomas-Ruzic 1998). No other complex clauses were included. Moreover, the end of the clause had to be a place of possible turn completion (as that was determined by me).

Utterances which meet these criteria, and which, furthermore, are acoustically clear enough to allow a pitch track to be created (which means no overlap, no special voice quality, no low volume, no overly noisy data, and so on)[7] are quite rare. From all those hours of interaction, I was only able to find 43 clear examples.[8]

4. What to code for? What might be possible phonetic differences?

In the literature on nuclear and pre-nuclear accents, it is often claimed the nuclear accents are more "prominent" than pre-nuclear accents. "Prominent" generally means longer and louder in this literature, but from examining Schegloff's "*he's about the only good regular out there*" example we can propose some additional hypotheses. Compare *good* and *out* in this example. We can notice the following differences (where the accent on *good* counts as the non-last accent, while the accent on *out* counts as the last accent):

i. There is a bigger step up to the accented syllable from the preceding syllable for the non-last accent.
ii. There is lower amplitude on the last accented syllable; since the values for amplitude are negative, a larger number is actually a lower amplitude.
iii. There is a bigger step down from the last accented syllable to the next syllable than for the non-last accent
iv. There is a greater change of pitch on the last accented syllable than on the non-last accent
v. There is an earlier peak for the last accented syllable
vi. There is longer duration on the last accented syllable

I thus used these properties as the starting point for my coding. Each accented syllable that was included within the study was coded for each of the properties above, as well as for:

vii. Whether the syllable was level, rising or falling, or rising and falling
viii. Whether the word was turn-final or not
ix. How many syllables there were to the end of the word, and to the end of the turn
x. What the highest peak in the syllable is
xi. Whether the last accent is where a declination line would predict it to be

These properties will be described below.

The most important variable was what I called "accent-type". I distinguished three accent-types: non-last, final and terminal. A non-last accent was a syllable which was not the last accent in its clause. I broke "last" accents into two subtypes: final and terminal. A final accent was the last accent in the clause but it did not occur on the terminal word of the clause; a terminal accent was the last accent in the clause and it fell on the terminal word of the clause. I decided to distinguish final and terminal accents under the assumption that terminal accents might show, in addition to prominence-lending pitch movement, some nonprominence-lending pitch movement marking the end of the turn (what Pierrehumbert calls a boundary tone). It was my assumption that final accents would not show such nonprominence-lending movements.[9]

5. Results

While I hypothesized that "last" accented syllables would be more prominent as indicated by the coded measures than "non-last" accented syllables, in fact the evidence does not strongly support that position. For some measures there appears to be a trend in the expected direction, but these differences are overwhelmed by large standard deviations and relatively small numbers of instances.[10] There is only one location of statistically significant interactions in comparing non-last and last accents: Duration. If we split the category "last" into final and terminal accent types, then we find one other site of significant difference: Peak Alignment. The details for each measure are given below.

5.1 Change of Pitch (in Hertz) on syllable

My first hypothesis was that more prominent accents would show greater pitch movement on them. I therefore calculated the number of Hertz (Hz) moved both up and down on each syllable. As can be seen in Table 2 below, there is a trend for last accented syllables to show a greater degree of change, but the differences are not statistically significant.

Table 2 Change in Pitch (in Hertz)

non-last	26	non-last	26
last	37	final	40
		terminal	36

5.2 Degree of Step Up To (in Hz)

While we might expect more prominent syllables to show a greater step up them, on the basis of example (1) above I hypothesized that in fact non-la syllables would show a greater step up to them than would last syllables. hypothesized in this manner for two reasons: first, because prior studies ha suggested that American English speakers often produce a "dip" just before th first accented syllable in an utterance ('t Hart et al 1990), which would produ a greater step-up-to the accented syllable for the first accented syllable than f a later accented syllable; and second, because previous studies have suggeste that the process of declination, by which the pitch of an utterance gradual drifts down over its course, creates a pattern in which the first accented syllab often (though not always, see 't Hart et al 1990) reaches a higher peak than lat accented syllables. I calculated Degree of Step Up To by finding the endir pitch (in Hertz) of the syllable before the accented syllable in question and l finding the beginning pitch (in Hertz) of the accented syllable and taking th difference. As can be seen from Table 3 below, there is a trend in this directio but it is not significant.[11]

It is important to note in this context that while the last accent may show smaller step-up-to, it may nonetheless be heard to be as high in pitch as, or least equal in pitch to, the non-last accent, precisely because of the expectatio created by declination (cf. Pierrehumbert 1979; Bolinger 1986; 't Hart et 1990). That is, because recipients expect the last accent to reach a lower pe than a non-last accent, they may compensate for that drift down by hearing lower peak as if it were comparable in height to an earlier peak that actual reaches a higher peak. It may therefore be the case that the last accents in m corpus were heard as just as prominent as the non-last accents, in spite showing a smaller step-up-to. This possibility is explored in greater depth Section 5.3 below.

Table 3 Degree of Step Up To (in Hertz)

non-last	33	non-last	33
last	23	final	22
		terminal	23

5.3 Highest Pitch Reached (in Hz)

We might expect the most prominent syllable in an utterance to also be the highest peak in the utterance. That is, we might expect that the most prominent syllable in an utterance would have as its highest pitch a level which is higher than the highest pitch of other prominent syllables in the utterance. On the other hand, given the existence of declination (mentioned above), it is possible that the last accent will tend to reach a peak that is actually lower than the peak of the non-last accent, given the general downdrift of pitch in an utterance, regardless of which syllable is perceived to be more prominent. My hypothesis was that the last accent would show a higher peak than the non-last accent, on the assumption that declination could be manipulated in favor of achieving prominence. Consider Table 4 below.

Table 4 Highest Pitch Reached (in Hz)

average difference between non-last and last: –22 Hz
(Standard Deviation = 67 Hz)

What these figures mean is that the peak of the non-last accent is, on average, 22 Hz higher than the peak of the last accent (with quite a bit of variation, evidenced by the large standard deviation). These data suggest that the last accent is not more prominent than the non-last accent by this measure.

Of course, as mentioned above, a last accent that is lower than the prior accent may nonetheless be heard as equal in pitch to, or even higher in pitch than, that prior accent, due to the expectations generated by declination. If the last accent reaches a peak that is predicted by declination, then it may indeed be heard as at least equal in pitch to the non-last accent, and therefore as at least equal in prominence to the non-last accent.

How can we determine if, in fact, the last accent is heard as lower in pitch than the non-last accent, because it is in fact lower in pitch, or if the last accent is heard as at least equal in pitch to the non-last accent, due to the expectations created by declination?

One way to check which interpretation of these data is more accurate is to determine what degree of declination would have been expected for the last accent, and how different from that expectation the actual peak on the last accent was. If the peak on the last accent is lower than what would be predicted by declination, then we can assume that the last accent might well be heard as lower in pitch than the non-last accent; if, however, the peak on the last accent is at the declination line, as predicted, or higher, then it might be reasonable to conclude that the last accent is heard as equal in pitch to, or even higher in pitch than, the non-last accent (in spite of being physically lower in pitch).[12]

I therefore compared the actual peaks reached by last accents and compared those with what would have been expected for those accents, given declination. I used 't Hart et al's (1990) formula for calculating declination, modified in order to allow for the fact that I was not calculating declination across an entire utterance but only a segment of an utterance (i.e., from the beginning of the non-last accent to the beginning of the last accent).[13] For each last accent, I calculated the predicted peak for that accent, and created a ratio of actual peak for that accent over the predicted peak for the accent. The resulting figures are given below in Table 5.[14]

Table 5 Actual vs. Predicted Values for the Last Accent (in Semitones)

Ratio of Actual Peak Value/Predicted Peak Value:	1.05
Standard Deviation	2.52 ST

The ratio is very close to 1.0, suggesting that the actual value and the predicted value are nearly identical. This finding would suggest that the peaks on the last accents are very close to what would have been expected given the process of declination, and therefore that the last accents, although lower in pitch, may be *perceived* as equal in pitch (and therefore, potentially, in prominence) to the non-last accents. However, the standard deviation is very large, and thus the results are inconclusive. What this means is that sometimes the last accent reaches a peak that is quite a bit lower than would have been predicted by declination, while at other times it is higher than would have been predicted. It is my understanding, then, that we cannot assume that the last accents are perceived as equal in pitch to the non-last accents — they may, in fact, be perceived as lower in pitch, and therefore potentially as less prominent than non-last accents.

5.4 Degree of Step Down From (in Hz)

We might also expect more prominent syllables to show a greater step down from them. I therefore hypothesized that last syllables would show a greater step down from them than would non-last syllables. I calculated Degree of Step Down From by taking the ending pitch (in Hertz) of the accented syllable in

question and the beginning pitch of the next syllable (whether it belonged to the same word or to a new word), and then taking the difference between them. As can be seen in Table 6 below, there is a trend in this direction, but it is not significant.[15]

Table 6 Degree of Step Down From (in Hertz)

non-last	33	non-last	33
last	44	final	44
		terminal	44

5.5 Peak Alignment

Silverman and Pierrehumbert (1990) report on work by Silverman (Silverman 1987) which found "nuclear peaks to be aligned much earlier in their syllables than prenuclear peaks" (p. 74). That is, they found that the highest pitch occurred earlier in the syllable for nuclear accents than for prenuclear accents. In their study, however, Silverman and Pierrehumbert found no such difference between nuclear and prenuclear accents if rhyme length and number of postnuclear syllables (to end of word) were figured into the calculations. Since all of this previous work was done on read speech in an experimental setting (and the Silverman and Pierrehumbert work used invented names as the data, rather than clauses or sentences), I wanted to see how peak alignment would behave in my conversational corpus. To calculate peak alignment, I measured the rhyme length (in milliseconds) of each accented syllable and measured how far into the rhyme the pitch attained its highest point; I then divided the peak measurement by the rhyme length, and plotted those figures by how many unaccented syllables followed the accented syllable within the same word. A smaller number indicates an earlier peak alignment, that is, that the peak occurs earlier in the rhyme. I hypothesized, contra Silverman and Pierrehumbert, that last accents would show an earlier Peak Alignment.

If we compare non-last and last, we can see that there are slight trends in the direction predicted by Silverman (1987) that is, that last (or nuclear) accents show an earlier peak alignment than non-last (or prenuclear) accents, but these differences are not significant.[16]

However, if we separate out final and terminal accent types, we can see that in fact there are some significant differences among the accent types.

Table 7 Peak Alignment

Number of unaccented syllables left in word	0	1	2
non-last	0.63	0.75	--
last	0.59	0.63	0.71

Table 8 Peak Alignment

Number of unaccented syllables left in word	0	1	2
non-last	0.63	0.75	--
final	0.75	0.76	--
terminal	0.45	0.57	0.71

Under the condition where the accented syllable is the last syllable of the word, non-last accents show a significantly later peak alignment than do terminal accents ($t = 2.25$, $p < 0.05$). This pattern almost certainly arises from the fact that terminal accents very often show a fall as well as a rise (33% of terminal accents show a fall and a rise), and if both a rise and a fall must take place within the same syllable, the speaker may position the peak quite early in the word, to accommodate the fall. A non-last accent, even if it is the last syllable in the word, does not have this constraint — even if a fall is to come, it is much less likely to come on that syllable (only 39% of non-last syllables show a fall, and only 20% show rise and fall), since it can be delayed to the following word.

It is interesting to note that final accents also show a significantly later peak alignment than do terminal accents ($t = 3.33$, $p < 0.01$), if the accented syllable is the last syllable of the word. The explanation for this pattern may be the same as the explanation offered above for the difference between non-last and terminal accents: terminal accents in this condition very often must accommodate a rise and a fall within that one syllable, while final accents can delay at least part of the fall to the next word.

The difference between non-last and final accents in this same condition is close to significant but does not quite reach it ($t = 2.0$, p between 0.1 and 0.05). If future research finds this difference to be a stable one, it will provide an interesting challenge to prior work, since here we see that non-last accents actually show an *earlier* peak alignment than final accents. I have no explanation for this possible finding.

If we move now to looking at cases where there is one unaccented syllable left in the word after the accent, we see a trend for the terminal accent to show an earlier peak alignment, but the differences are not significant.

5.6 Duration (in milliseconds)

It is well known that syllables perceived to be more prominent in English tend to be longer than syllables perceived to he less prominent. I therefore hypothesized that last accents would be longer (in milliseconds) than non-last accents. This is the one hypothesis that is strongly supported by the data.[17] As can be seen from the tables below, non-last accented syllables are significantly shorter than are last accented syllables, and words containing non-last accented syllables are significantly shorter than are words containing last accented syllables. Moreover, if we examine the second of the two tables, we can see that terminal syllables and words are the longest of all, showing a significant superiority in length over both non-last and final accents. These findings are in keeping with a great deal of prior research on lengthening of turn endings and on lengthening of highly prominent syllables (e.g. Local, Wells and Sebba 1985; Local, Kelly and Wells 1986; Lehiste 1979; Berkovits 1984a, b; Bolinger 1986; Ladd 1996).

Table 9 Duration (in milliseconds)

	Syllable Length	Word Length
non-last	195	249
last	280	436
non-last	195	249
final	235	343
terminal	304	486

5.7 Amplitude

It is also well known that more prominent syllables in English tend to be louder than less prominent syllables. I therefore at first hypothesized that last accents would be louder than non-last accents; but then I noticed in the data that amplitude tends to be lowered at the ends of turns. These two conflicting pressures may cancel each other out — there is no difference among the accent types.[18]

Table 10 Amplitude (in decibels)

non-last	−20.4
final	−20.6
terminal	−22.2

5.8 Movement Type: Level, rising, falling, rising-falling

Because I believed that prominent syllables may show greater pitch movement than non-prominent syllables, I hypothesized that non-last accents would tend to show just level or rising pitch, without a more complex pitch movement, while I hypothesized that last accents would tend to show more complex pitch movements. While there are some interesting trends here, none of the differences are significant. The closest to reaching significance is the distribution of level accents: there are no last accents which show level pitch movement throughout the syllable while 7% of the non-last accents show level pitch (χ^2 = 3.14, p between 0.1 and 0.05). This trend is consistent with the hypothesis that last accents would show more dynamic pitch movements on their syllables.

Table 11 Movement Type

	Just Rise	Just Fall	Level	Rise-Fall	TOTAL
non-last	30 (53%)	11 (19%)	4 (7%)	12 (21%)	57
last	18 (42%)	13 (30%)	0	12 (28%)	43
non-last	30 (53%)	11 (19%)	4 (7%)	12 (21%)	57
final	7 (47%)	5 (33%)	0	3 (20%)	15
terminal	11 (39%)	8 (29%)	0	9 (32%)	28

6. Conclusions

On the basis of the very exploratory study conducted here, I would say that I found no strong evidence either to support or to reject the claim that last accents are more prominent than non-last accents in a way that could be made use of in real-time turn projection.

What are some possible explanations of these somewhat inconclusive findings? It seems to me that there are (at least) three possibilities:[19] (a) Schegloff's proposal concerning prosody and turn projection is not correct; (b) Schegloff's proposal covers only a subset of utterances; (c) the phonetic details of accent can only be understood within their emerging syntactic, semantic, and sequential environments.

For now I will reject the first explanation, since the data are not aligned

powerfully enough in one direction or another to allow us to eliminate a hypothesis which has other kinds of evidence in support of it (see, for example, Schegloff 1998). The second explanation, that Schegloff's original proposal covers only a subset of utterances, remains a possibility, given the findings of the current study. That is, while there are clear patterns for some instances of the dataset, these patterns may be "washed out" by the large variation across all the utterances taken together. Thus it may be that there is a group of cases that behave alike or perhaps several groups which have internal consistency but which differ from one another but there is a great deal of variation across groups. If this possibility turns out to be right, then we may find that there is in fact a group of cases which shows a clear difference between non-last and last accents, with each accent type showing a cluster of properties (e.g. greater duration, stepdownfrom and change in pitch for last accents). Instances outside that group or perhaps in other groups may show either a different pattern or perhaps no pattern at all.

The third explanation — that prosody and turn projection can only be understood within the syntactic, semantic, and sequential environment in which utterances actually occur — is, given what else we know about turn projection, the most plausible (see, for example, Ford, Fox and Thompson 1996a). For example, it is quite possible that speakers can put accents wherever and however they want, and the result is just a very large degree of variation across utterances, given that prosody can do other work besides projection (cf. Bolinger 1986). This possibility makes some sense, given that prosody is always working through other practices, such as syntax, semantics, gesture and pragmatics and so may bear a relatively small "work load" in terms of turn projection.

One particular version of this possibility deserves examination here. Schegloff (1998) has proposed, with regard to turn projection, that "grammar nominates and prosody seconds" a place of possible turn completion.[20] If this proposal is correct, then there ought to be a strong correlation between grammatical position and type of accent.

While my data at first appear to support this view of the relationship between grammar and prosody, in fact on closer reflection they may actually be silent on that topic. Consider the following data. A count of the non-last and last accents by part of speech yielded this distribution: 60% of all last accents were on nouns, while only 14% of *non*-last accents were on nouns. Moreover, while only 12% of last accents were on verbs, 46% of *non*-last accents were on verbs.

Although the accents in my corpus fall equally frequently on nouns and verbs in general (roughly 33% each of the total accents), there is thus a large

skewing as to where grammatically in the utterance each type of accent occurs. Non-last accents tend to occur on grammatical items that in English might project grammatically "more to come" (like verbs, for example), while last accents tend to occur on grammatical items that in English might project "upcoming completion" (like nouns, for example). At first it might seem reasonable to infer from this skewing that grammar is already doing much of the projection and that prosodic information is, at least under certain circumstances, free to engage in other projects, such as repair, or contrast.

Further inspection of the data, however, reveals that this apparently interesting skewing has more to do with the particular utterances selected for inclusion in the corpus than with prosody and turn projection in general. Recall that utterances were only included in this corpus if they exhibited two (or more) pitch peaks after the subject of the utterance; the utterances thus selected, if they had a first pitch peak on the verb thus also had to have another pitch peak, which in many cases would have fallen on a (lexical) noun — very often the direct object of the verb, or, in some cases, a locative or temporal noun. Thus since verbs quite often carry an accent, in this special corpus an additional accent would have been required, making the accent on the verb non-last. Of course, a verb could carry the second (and last) accent in an utterance, but in such instances a preceding lexical item would have had to carry a pitch peak, and this situation is somewhat less common than the first situation described, inasmuch as it requires two conditions to come together: (1) it would require pitch peaks on certain kinds of adverbs (like *never, still*) or on modals (like *can't*), and these items are just not as frequently accented as are lexical verbs and nouns (hut such cases did occur in my collection: for example, *(I) never met her*); and (2) it would require that no pitch peak occur on any item after the verb. Utterances that meet both conditions, especially given the relative infrequency of the first condition, are thus bound to be less common than utterances with a first accent on a verb and a second (and last) accent on a syntactically projected noun. The correlation between non-last accent and verb, and last accent and noun, tbus appears to be at least in large part an artifact of the collection. In fact, in my collection of utterances with a *single* pitch peak, fully 42% (15/36) of the accents fell on verbs, while only 19% (7/36) fell on nouns. In these instances, the "last" accent (which is also the only accent) correlates more strongly with verbs than with nouns, the opposite of the finding for utterances with two (or more) accents. I thus did not find clear evidence to support the claim that there is a strong correlation between grammatical position and accent-type.[21]

Thus, while I discovered no compelling evidence in favor of this third explanation, the data do not contradict it, and hence it remains a possible account of the findings. Clearly, further research is needed to explore the precise role of prosody in turn projection.

What are the implications of the findings for turn projection? Again, it seems to me that there are (at least) three possibilities.

The first possibility is that there are no implications of these findings for turn projection, because of the very large variation across instances. Even if some of the tendencies discussed here could be found to be statistically significant with a larger collection of cases, it would still remain to be shown how on any particular occasion a recipient could be able to hear an accented syllable as non-last or last, given that there is so much variability in the actual phonetic doing of accented syllables. Consider, for example, the measure of Duration. This is the one strong locus of difference between non-last and last accents in my study. But given that the minimum value for Duration on non-last accented syllables was 84 ms and the maximum was 455 ms, and the minimum value for last accented syllables was 127 ms and the maximum was 605 ms, an accented syllable with a duration of, let's say, 250 ms is well within the range of either accent type. So how would a recipient know on the basis of that duration if the accent was non-last or last? The same point holds even more clearly for the other measures, since they were not even statistically significant.

The second possibility, tied to the second explanation (above), is that there is a practice of turn projection such that a speaker can produce an accented syllable with a cluster of phonetic properties and thereby "do" upcoming completion (or: no upcoming completion). In this perspective recipients are attuned to not just one phonetic property but to clusters of properties which make an accented syllable sound more or less prominent. For example, the recipient may hear a syllable that shows a duration of 300 ms, a fairly large change in pitch, a small step up to the syllable, a peak 3/4 of the way through the syllable, and a fairly large step down, as a last accent. It may be such a cluster of properties that Schegloff heard in the example I started with; such a cluster may make an accented syllable sound more prominent, therefore projecting upcoming completion, while syllables which lack some of these properties may sound less like they project upcoming completion, even if in fact the turn does come to completion at the next place of possible completion.

The third possible implication is that the prosody of a particular utterance is always working through the specific environment of that utterance. If we consider Duration, for example, we could say that, although the duration of a syllable alone cannot tell us if an accented syllable is projecting upcoming completion or not, that duration *in a particular syntactic, semantic and sequential environment* may serve as part of a display that the turn is coming to possible completion (or not). Although finding evidence for such a hypothesis is difficult, it nonetheless deserves to be pursued further.

This study represents a first attempt at exploring the phonetic details of prosody and turn projection from early in the turn. Although the findings of the study are inconclusive, they do suggest further avenues of research. For example, a next study on this topic would need to explore cluster analyses of the data to see if, in fact, participants are orienting to clusters of phonetic properties in listening for turn projection displays. Or, alternatively, a next study could pursue in greater depth the relationships between syntax and prosody — are there patterns in where certain accent-types occur syntactically in the developing organization of an utterance? It will also almost certainly be necessary to conduct careful interactional analyses of small groups of utterances (including cases in which an accent is treated by the recipient as projecting upcoming turn-completion, and those in which an accent is treated explicitly as *not* indicating upcoming completion), to see how phonetic details of the sort examined in the current study are oriented to on particular occasions by participants. However future studies go, I hope the present study will have made a useful contribution to our understanding of prosody and turn projection.

Notes

* I am grateful to Cecilia Ford, Charles and Candy Goodwin, Robert Jasperson, and Emanuel Schegloff for the use of their data. I am also grateful to Alan Bell, Emanuel Schegloff, Elizabeth Shriberg, Sandra Thompson, and the editors of this volume for insightful and helpful comments on an earlier draft of this paper. Many thanks also to Robert Jasperson for giving me all manner of technical assistance in doing the analyses. Special thanks also to Elizabeth Couper-Kuhlen for helping with the details of the declination analyses.

1. There is one important complicating factor in this utterance, however: the accent on *good* is accomplishing repair, which may cause it to have the particular shape it does (see Jasperson 1998).

2. Note, however, that the non-projecting pitch peak displays a larger step up to it from the preceding syllable. This fact would seem to make the non-projecting pitch peak more prominent than the projecting pitch peak. I return to this question in section 5.

Note, further, that the recipient would not be able to hear that the projecting pitch peak is more prominent than the non-projecting pitch peak until s/he has heard the second (projecting) one. I return to this crucial issue in section 6.

3. I excluded all utterances with even a single downward pitch obtrusion.

4. I follow Bolinger (1986) in using the term *accent* rather than *stress* for a syllable which is actually produced in real discourse with prominence. Bolinger restricts the term *stress* to that syllable in a word — e.g. in a dictionary entry — which is capable of bearing accent.

5. CallHome is a digitized, and in some instances tagged, database consisting of phone conversations in American English.

6. I used CSpeech, which was created by Paul Milenkovic and Charles Reed at the University of Wisconsin, Madison.

7. Excluding such phenomena could obviously have caused a skewing in my collection.

8. The entire corpus of utterances I created pitch tracks for and database entries for included two other types of cases: (a) single-TCU turns with only a single accent; and (b) a somewhat random collection of other types of utterances. The entire corpus consists of approximately 100 utterances total.

9. This assumption may, in fact, not be correct. Final accents may indeed show nonprominence-lending movement.

10. It was suggested to me by Patricia Clancy that separating the data by gender might factor out much of the large variation. I did do this, and while there are clearly statistically significant differences between males and females on many of the measures, the variation within each gender is still so great that the results are once again not significant.

11. Only instances for which there was a step up to the accented syllable were included in this count. I treated a syllable as being stepped up to if there was at least a difference of 5 Hertz between the last measurement of the preceding syllable and the first measurement of the accented syllable.

12. I am very grateful to Elizabeth Couper-Kuhlen for suggesting this approach to the data, and for providing a great deal of support in doing the actual calculations.

13. Using 't Hart et al's (1990) formula for declination was problematic for my data, given that their formula was designed to predict downdrift for unaccented syllables (valleys) and not for accented syllables (peaks). 't Hart et al explicitly address the problems associated with computing declination based on accented syllables (the so-called topline), given that "it is reasonable to expect that possible differences in the strength of the accents are reflected in different peak heights (...). This would make the slope of the topline dependent on the strengths of the accents relative to each other" (p. 125). Nonetheless, it has been assumed by other scholars that the topline and the baseline (the line connecting the valleys to one another) are, in fact, parallel to one another (see, for example, Schuetze-Coburn, Shapley and Weber 1991). I therefore adopted the formula, with the awareness that it might be merely suggestive and not entirely accurate.

14. In order to use the formula given in 't Hart et al (1990), it was necessary to convert Hz values into semitone values. The figures are thus in semitones.

15. Only instances showing a step down from the accented syllable were included in this count. An instance was treated as showing a step down from if there was at least a 5 Hertz difference between the last measurement of the accented syllable and the first measurement of the following syllable.

16. Peak Alignment averages were calculated for instances whose Peak measurement was greater than zero.

17. It should be acknowledged that finding the exact length of a syllable or word in my data was extremely difficult. Because the recordings are quite noisy, and my tools are somewhat primitive, any given measurement could be off by 30 milliseconds or so.

18. I measured amplitude on the rhyme of the accented syllable, using the automatic amplitude function within CSpeech.

19. There is actually a fourth explanation, which is that the definitions used for making the collection are faulty and therefore the results are not meaningful.

20. But see Ford, Fox and Thompson (1996a, b) for another perspective on the relationships between grammar and prosody.

21. And a comparison of the pitch peaks on verbs when there was only one accent in the utterance with pitch peaks on verbs when there were two or more accents in the utterance revealed no statistically significant differences on the measures examined here, with the exception of Step Down From, which did show a significant difference. So, for example, if we compare the following two utterances,

(a) I understand that part [multi-accent utterance]
(b) I love you [single accent utterance]

we find only one feature on which they are different. This fact makes sense, given that the step down from a verb in a single-accent utterance would have been to the end of the utterance, a point at which lower pitch is expected, while the step down from a non-last verb in a multi-accent utterance would have exactly NOT been the end of the utterance, hence no special lowering of pitch would be expected. Step Down From is, however, not something that can be heard during the verb itself, so it could not be used by the recipient as a clue to whether the accent should be heard as the last one in the utterance or not. It appears, then, that there are no phonetic properties of the accented syllables themselves in these cases which would allow a recipient to hear if the verb is going to exhibit the last accented syllable in the utterance or not.

References

Berkovits, R.
1984a "Duration and fundamental frequency in sentence-final intonation". *Journal of Phonetics* 137: 255–265.
1984b "A perceptual study of sentence-final intonation". *Language and Speech* 27: 291–307.

Bolinger, D.
1986 *Intonation and its Parts.* Stanford, CA: Stanford University Press.
Brown, G.
1983 "Prosodic structure and the given/new distinction". In *Prosody: Models and Measurements*, A. Cutler and D. R. Ladd (eds), 67–77. Heidelberg: Springer.
Chafe, W.
1987 "Cognitive constraints on information flow". In *Coherence and Grounding in Discourse*, R. Tomlin (ed.), 21–51. Amsterdam: John Benjamins.
Couper-Kuhlen, E. and Selting, M., eds.
1996 *Prosody in Conversation.* Cambridge: Cambridge University Press.
Cruttenden, A.
1997 *Intonation* [second edition]. Cambridge: Cambridge University Press.
Crystal, D.
1969 *Prosodic Systems and Intonation in English.* Cambridge: Cambridge University Press.
Cullicover, P. and Rochemont, M.
1983 "Stress and focus in English". *Language* 59: 123–165.
Duncan, S.
1974 "Some signals and rules for taking speaker turns in conversations". In *Nonverbal Communication*, S. Weitz (ed.), 298–311. New York: Oxford University Press.
Duncan, S. and Fiske, D. W.
1977 *Face to Face Interaction.* Hillsdale, NJ: Lawrence Erlbaum Associates.
Ford, C. and Mori, J.
1994 "Causal markers in Japanese and English conversations". *Pragmatics* 4: 31–61.
Ford, C. and Thompson, S.
1996 "Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns. In *Interaction and Grammar*, E. Ochs, E. Schegloff and S. Thompson (eds.), 134–184. Cambridge: Cambridge University Press.
Ford, C., Fox, B. and Thompson, S.
1996a "Practices in the construction of turns: the 'TCU' revisited". *Pragmatics* 6: 427–454.
1996b "Simultaneous practices of turn projection". Paper delivered at the annual meeting of the American Anthropological Association, November, 1996.
To appear "Social interaction and grammar". In *The New Psychology of Language*, Vol. II, M. Tomasello (ed.), Mahwah, NJ: Lawrence Erlbaum.
Fox, B., Hayashi, M. and Jasperson, R.
1996 "Resources and repair: a cross-linguistic study of the syntactic organization of repair". In *Interaction and Grammar*, E. Ochs, E. Schegloff and S. Thompson (eds.), 185–237. Cambridge: Cambridge University Press.

Goodwin, C.
1979 "The interactive construction of a sentence in natural conversation". In *Everyday Language: Studies in Ethnomethodology*, G. Psathas (ed.), 97–121. New York: Irvington.
1981 *Conversational Organization: Interaction Between Speakers and Hearers.* New York: Academic Press.
Gussenhoven, C.
1984 *On The Grammar and Semantics of Sentence Accent.* Dordrecht: Foris.
Hayashi, M. and Mori, J.
1998 "Co-construction in Japanese revisited". *Japanese/Korean Linguistics* 7: 77–93.
Hayashi, M., Mori, J. and Takagi, T.
To appear "Contingent achievement of co-tellership in a Japanese conversation". In *The Language of Turn and Sequence*, C. Ford, B. Fox and S. Thompson (eds.) Oxford: Oxford University Press.
Jasperson, R.
1998 *Repair after cut-off.* PhD Dissertation, University of Colorado.
Ladd, D. R.
1996 *Intonational Phonology.* Cambridge: Cambridge University Press.
Lambrecht, K.
1994 *Information Structure and Sentence Form.* Cambridge: Cambridge University Press.
Lehiste, I.
1979 "Perception of sentence and paragraph boundaries". In *Frontiers of Speech Communication Research*, B. Lindblom and S. Ohman (eds.), 191–201. New York: Academic Press.
Lerner, G.
1991 "On the syntax of sentences-in-progress". *Language and Society* 20: 441–458.
1996 "On the 'semi-permeable' character of grammatical units in conversation". In *Interaction and Grammar*, E. Ochs, E. Schegloff and S. Thompson (eds.), 238–276. Cambridge: Cambridge University Press.
Lerner, G. and Takagi, T.
1999 "On the place of linguistic resources in the organization of talk-in-interaction". *Journal of Pragmatics* 31: 49–75.
Liberman, M. and Pierrehumbert, J.
1984 "Intonational invariance under changes in pitch range and length". In *Language Sound Structure*, M. Aronoff and R. Oehrle (eds.), 157–233. Cambridge, Mass: The MIT Press.
Local, J., Wells, W. H. G. and Sebba, M.
1985 "Phonology for conversation: phonetic aspects of turn delimitation in London Jamaican". *Journal of Pragmatics* 9: 309–330.
Local, J., Kelly, J. and Wells, W. H. G.
1986 "Towards a phonology of conversation: turn-taking in Tyneside English". *Linguistics* 22: 411–437.

Ochs, E., Schegloff, E. and Thompson, S.
1996 *Interaction and Grammar*. Cambridge: Cambridge University Press.
Pierrehumbert, J.
1978 "The perception of fundamental frequency declination". *Journal of the Acoustical Society of America* 66: 363–369.
1980 *The Phonology and Phonetics of English Intonation*. PhD Dissertation, MIT.
Sacks, H., Schegloff, E. and Jefferson, G.
1974 "A simplest systematics for the organization of turn-taking for conversation". *Language* 50: 696–735.
Schegloff, E.
1987a "Analyzing single episodes of interaction: an exercise in conversation analysis". *Social Psychology Quarterly* 50: 101–114.
1987b Xerox Parc Lecture.
1996 "Turn organization". In *Interaction and Grammar*, E. Ochs, E. Schegloff and S. Thompson (eds.), 52–133. Cambridge: Cambridge University Press.
1998 "Reflections on studying prosody in talk-in-interaction". *Language and Speech* 41: 235–263.
Schuetze-Coburn, S., Shapley, M. and Weber, E.
1991 "Units of intonation in discourse: a comparison of acoustic and auditory analyses". *Language and Speech* 34: 207–234.
Selting, M.
2000 "The construction of units in conversational talk". *Language in Society* 29: 477–517.
Silverman, K.
1987 *The Structure and Processing of Fundamental Frequency Contours*. PhD Dissertation, Cambridge University.
Silverman, K. and Pierrehumbert, J..
1990 "The timing of prenuclear high accents in English". In *Papers in Laboratory Phonology I*, J. Kingston and M. Beckman (eds.), 72–106. Cambridge: Cambridge University Press.
't Hart, J., Collier, R. and Cohen, A.
1990 *A Perceptual Study of Intonation*. Cambridge: Cambridge University Press.
Tanaka, H.
1996 *Language, culture and social interaction: A comparison of turn taking in Japanese and Anglo/American English*. PhD Dissertation, University of Oxford.
2000 "Turn-projection in Japanese talk-in-interaction". *Research on Language and Social Interaction* 33(1): 1–38.
Thomas-Ruzic, M.
1998 *Language and activity: 'we', 'you' and 'I' in instructional talk*. PhD Dissertation, University of Colorado.

Wells, B. and Peppe, S.
1996 "Ending up in Ulster: Prosody and turn-taking in English dialects". In *Prosody in Conversation*, E. Couper-Kuhlen and M. Selting (eds.), 101–130. Cambridge: Cambridge University Press.