Prosodic variation for topic shift and other functions in local contrasts in conversation

Margaret Zellers

Department of Language & Linguistic Science, University of York, UK
(Current affiliation: Dept. of Speech, Music, & Hearing, KTH, Stockholm, Sweden)

Short title: Prosodic variation in local contrasts

Correspondence address:
Department of Speech, Music & Hearing
Lindstedtsvägen 24
100 44 Stockholm
Sweden
Phone: +46-8-790-7567
Fax: +46-8-790-7854
Email: zellers@kth.se
Abstract

Speakers and listeners have been shown to use phonetic cues to help them in tracking the ongoing structure of conversational interaction, but fragmentation between qualitative and quantitative research means that the forms and functions of these cues have been given varying characterizations. The current study explores prosodic variation in contrastive structures in conversational data, using a combined methodology adopting aspects from both qualitative (conversation analysis) and quantitative (experimental phonetics/phonology) approaches. Statistical and conversation-analytical methods used together reveal relationships between prosodic variation and interactional function, such as variations in pitch range across adjacent turns being linked to the presence of “stepwise” (Sacks 1972 [1992]) topic changes.

1 Introduction

In investigating the organization of language in use, it is important to consider two levels of such organization: a local, unit-by-unit framework that manages immediate transitions from one element to the next, and a global organization that holds a large body of talk together and gives it coherence (cf. Schegloff 2011). Conversationalists have been consistently shown to track in an ongoing manner the direction of talk in order to both understand what is currently happening, and to plan future contributions (Goodwin & Goodwin 1992). Phonetic cues such as variation in segmental production, pitch contours, and speech rate have been shown to be essential for this process (cf. e.g. Local, Kelly & Wells 1986; Ogden 2003, 2006; Uhmann 1992 inter alia).

The research presented here examines contrastive structures, which although they find their roots in (relatively) local lexical comparisons, have functions related to the global issue of topic structure. This is demonstrated in example (1), in which U accomplishes a topic change from her parents’ opposition to a proposed journey to Israel, to events in her New
York City neighborhood, using a contrastive structure. Recycling V’s structure you have more of a chance of getting killed (l. 4), U replaces one potential source of harm (the plane ride, ll. 4-5) with another one launching her new story: my own neighborhood (ll. 7-8).

(1)

1  U: the probability of me: getting hhh killed in a
2  terrorist hhh hhh hhh hhh hhh situation is highly
3  unlikely
4  V: you have more of a chance of getting killed in the
5  plane ride over there [than xxx]
6  U: [oh yeah and you know what] else
7  I have more n- chance of getting killed in my own
8  neighborhood
9  U: hhh [take the I took the dog out last night
10  V: [yea::h] exa::ctly
11  U: about eight fifteen
12  U: [.hhh]
13  V: [m hm]
14  U: I'm walking I got to the end of the block and I
15  hear four or five big gunshots

(CallFriend-engn:42; transcripts follow Gail Jefferson’s conventions, cf. e.g. Sacks, Schegloff & Jefferson 1974)

Similarly to other work on contrast in a Conversation Analytic paradigm (Ford 2000; Deppermann 2005), contrast is defined initially following Mann & Thompson (1992):

“...two items are said to be in contrast if they are:
a. comprehended to be the same in many respects,

b. comprehended as differing in a few respects, and

c. compared with respect to one or more of these differences.” (37)

We may apply this to example (1) in the following way: more [of a] chance of getting killed in is the same, while the object of the preposition is different: the plane ride over there (ll. 4-5) or my own neighborhood (ll. 7-8). The comparison between these sources of harm is apparent in the juxtaposition of apparent to real danger which was launched in line 2 with terrorist [] situation, stepping down to the seemingly less hazardous plane ride (l. 5), and then my own neighborhood (ll. 7-8), which might under other circumstances be considered a perfectly safe location.

In response to criticisms of monosystemic approaches by e.g. Local (1996) and Couper-Kuhlen & Selting (1996), the current study takes a multiple-methods approach to the investigation of contrastive structures in conversation. It is qualitatively informed by Conversation Analysis (CA) and especially its sub-discipline, the Phonetics of Talk-in-Interaction (PTI; cf. Local & Walker 2005; Ogden 2012b; Walker 2013), in that it prioritizes the situated use of language. All the data are taken from naturally-occurring conversations, and discussion of the use of the structures in question will involve consideration of contextual features. However, the study also adopts quantitative methods used primarily in laboratory studies of prosodic (and especially intonational) structure of language. In drawing together these different methodologies, this study also attempts to address fragmentation between areas of linguistic inquiry adopting a formal/structural approach on the one hand, and a functional/communicative approach on the other (cf. Scobbie & Stuart-Smith 2012; Zellers & Post 2012).

The current study explores ways in which meaningful prosodic variation can be identified in conversational data using statistical methods typical of more experimentally-
based research paradigms. Specifically, conversational contrasts are marked with pitch variation even if they do not bear canonical “prosodic contrast” (cf. Section 2.1.2). Global rhythmic features are found to be more closely linked with prosodic contrast marking in conversational sequences than global pitch features. Finally, variation in pitch range of the turn following a conversational contrast is linked to whether or not the contrastive structure was used to make a topic change.

2 Background

2.1 Defining contrast

Mann & Thompson’s characterization of contrast is flexible in that it is able to address both the semantics of contrast as well as its form. While the general definition of “contrast” given above provides a basis for understanding this term, in practice the term has been used to cover a wide variety of phenomena which may be intuitively easy to identify but are more difficult to characterize in strict terms. The essential intuition is something along the lines of “same-but-different”. Halliday & Hasan (1976) partially assume this intuitive understanding of contrast when they discuss the functions of substitution, in which one item is “removed” and something else “replaces” it. The replacement must attain some threshold of similarity to the original item, or it cannot be said meaningfully to be a substitution. Just where this threshold falls can be problematic to identify; Deppermann (2005) argues that in principle any two items can be placed in contrast if the appropriate conversational means are adopted. Namely, a context must be created in which the two items are shown to have some kind of link or co-class membership (cf. Sacks 1966 [1992]). Thus, a terrorist situation, the plane ride, and my own neighborhood may not in isolation be immediately obvious co-class members, but the conversational participants in (1) give them co-class membership by framing them as situations in which potential danger must be assessed.
2.1.1 Contrast in CA

CA insists that only usages which conversational participants themselves orient to can be said to be meaningful (cf. e.g. Levinson 1983; Drew 2004). In other words, identifying constructions that meet the formal qualifications for contrast is not sufficient; it also remains to be shown how they are used and understood in everyday language use.

Taking as her starting point Mann & Thompson's (1992) characterization of contrast, Ford (2000) demonstrates how contrastive structures in conversational settings are often followed by explanations or solutions. She shows that these explanations or solutions are oriented to as the “normal” continuation of talk following such a contrast; that is, they are expected by conversational participants, and treated as absent when they do not occur, for example being requested by another speaker.

Deppermann (2005) points out that contrast is something that speakers create in the course of interaction. Following Edwards (1997), he shows that context can be, and is, used to place two lexical items in opposition that would not otherwise have a contrastive interpretation. Two particular functions that Deppermann identifies are that of correcting a prior (mis)interpretation, and that of justifying a “deviation categorization”, that is, explaining a specific local meaning for a word and presenting it as adequate in the context.

Instead of employing the term “contrast”, Barth-Weingarten (2009) prefers to describe the phenomena she investigates as “parallel-opposition constructions”. Her focus is on the form, particularly the prosodic form, of turns which project an upcoming contrast. She finds that in most cases, speakers use prosodic means including narrow focus (cf. Section 2.1.2) on items to be contrasted, “continuing” (i.e. holding at a mid-level) intonation, lack of final lengthening, filled pausing, and rush-through (i.e. local increase in speech rate at potential boundary sites) to project that a contrast will be made.
2.1.2 Contrast in experimental prosody

In quantitative experimental approaches to phonetics and prosody (hereafter EP), contrast is treated as a subset of the larger phenomenon of focus, the prosodic marking of certain portions of an utterance as prominent. The existence of broad, narrow and (often) contrastive focus as linguistic tools is treated as given in intonational phonology, while their forms are more open for debate, and may overlap, as can be seen in example (2).

(2)

(a) What happened?

John went to the store with MARY. (Broad focus)

(b) With whom did John go to the store?

John went to the store with MARY. (Narrow focus)

(c) Where did John go with Mary?

John went to the STORE with Mary. (Narrow focus)

(d) Did John go to the restaurant with Mary?

John went to the STORE with Mary. (Contrastive/corrective focus)

Although the prominence distribution may be the same (compare 2c and 2d), contrastive focus is generally treated as having a different phonetic (and possibly phonological) form from narrow focus. Calhoun (2010) argues that contrastive interpretations in English are based on the comparative prominence of a given item compared to other items in the context. Katz & Selkirk (2011) characterize this comparative prominence as involving longer duration, higher intensity, and larger fundamental frequency (F0) movements compared to non-focused elements than narrow-focus not leading to a contrastive interpretation. Across languages, whether phonological or phonetic characteristics are involved in the discrimination of contrastive from narrow focus, comparative prominence appears to be a common thread (cf. e.g. Braun 2006; de la Mota Gorriz 1997; Féry & Kugler 2008; Hanssen, Peters &
2.1.3 Comparing approaches to contrast

In EP, contrast is a kind of characteristic form that can be used, both in the syntax and in the prosody; however, the contexts in which it is studied are very constrained and treated as self-evident, generally being tied to the kind of constructions shown above in example (2). In CA, contrast can refer to the kind of form described in EP, referring to the same-but-different syntactic construction, and/or the presence of a pitch accent marking this construction, but contrast is considered to be an activity done by the speaker. It need not take a specific syntactic or prosodic form, although there are notable regularities; the point is that the interlocutors treat the tokens in question as having same-but-different characteristics. This difference in definition reflects the difference between “formal” (e.g. EP) and “functional” (e.g. CA/PTI) approaches that the current research attempts to address. Although there are obvious points of connection, the underlying assumptions are dramatically different.

2.2 Defining topic

As a first approximation to the definition of discourse topic (hereafter simply “topic”), we can use the idea of “aboutness”: the topic is what a discourse (or part of a discourse) is intuitively about, whether this is a referent, an idea, or something else. Attempts at formal definitions of topic have made reference to information that is active or accessible at a given point in a conversation (Grosz & Sidner 1986; Chafe 1994).

The different assumptions underlying linguistic analysis that have already been shown in the two approaches to contrast summarized above are equally apparent when it comes to the discussion of topic. CA approaches are more interested in the functional purpose of topic and topic shifts as behaviors in interaction, while EP approaches treat topic as a structural feature showing characteristic forms.
2.2.1 Topic in CA

Maynard (1980) characterizes both topic change and topic maintenance as being related to an overarching desire on the part of conversational participants to maintain ongoing talk. Thus, jumps from one topic to another unrelated one are used to keep talk going when participants have run out of things to say on a current topic.

Within the CA literature, two kinds of topic change have been identified. One is disjunctive, wherein one topic is brought to a close and then another is begun. However, there is also another kind of topic change, which has been described by Sacks (1972 [1992]) and Jefferson (1984a, 1993) as “stepwise”. These changes are gradual, achieved by the use of a conversational “pivot” which is related to what has gone before but also adds something new, which is then taken up as the topic. This creates a sort of “hidden” topic change, wherein it can be difficult to identify an exact location at which the topic changes. Jefferson (1984a) points out that this kind of topic change work may be done in order to protect social relations between conversational participants; that is, shifting topic may be a dispreferred activity. For example, it might threaten an interlocutor’s face by implying that the current topic was not desirable. Therefore, an “other-attentive” topic change, whereby the new topic is presented as closely tied to the old topic, is a preferable method for changing topic.

Disjunctive topic changes, like Button & Casey's (1984, 1985) itemised news enquiries and topic-initial elicitors, also involve attention to interlocutors’ needs, as can be seen in the prosodic variation associated with them. Couper-Kuhlen (2004) gives a detailed description of the prosody of English news announcements (cf. Button & Casey 1985) which launch something new: they tend to be preceded by pausing and inbreaths, and are louder and have higher pitch peaks than what came before. The features of recipient responses to news announcements are also relevant for the launch of new topics; a news-receipt “oh” that ends low in the speaker's pitch range, is lengthened, has creaky phonation quality, and follows a
pause is likely to end a sequence; it may do this by proposing that the recipient was
uninformed but is now informed, making further talk irrelevant (Local 1996). These types of
turns, which might be described as “disaffiliating”, are often prosodically very similar to one
another, while turns that “affiliate” with another speaker's talk tend to be prosodically
matched with the preceding turn (Goldberg 1978, 2004; Müller 1996; Gorisch, Wells &
Brown 2012). Thus prosody can function to signal both the structure of the interaction in
question, as well as the alignment or relationships between interlocutors in the course of the
interaction.

No studies, as far as the author is aware, have addressed the prosodic characteristics
of stepwise topic changes. This may be because, as “hidden” topic changes, they do not have
any obvious prosodic characteristics to be described. However, there is at present no data
available to confirm or challenge this suggestion, which will be addressed in the current
study.

2.2.2 Topic in EP

Experimental phonetic or phonological approaches to topic structure have focused on
disjunctive topic changes, which have distinctive prosodic features. Increased pitch peak
height and expansion of the pitch range used by a speaker are widely identified as cues
occurring at the beginning of a new topic group (Cruttenden 1986). Most of the research done
on the prosody of discourse structure has focused on boundary cues such as pitch reset
for English; across other languages, Botinis 1989; Bruce 1982; Sluijter & Terken 1993; Tseng
2004 inter alia). Longer pauses have been shown to occur adjacent to stronger topic
boundaries (Swerts & Geluykens 1993); and low pitch register, compressed pitch range, and
falling pitch contours are more likely to occur or are more pronounced at topic ends (Swerts,
Bouwhuis & Collier 1994).
In EP studies, the existence of topic structure of some form (at the very least, topic boundaries) is generally taken for granted, just as the existence and structure of contrast is, and obvious disjunction is a key feature. However, the consistent quantitative results of many studies are encouraging when it comes to considering the validity of these results, and certainly a unified pattern across many types of language use emerges.

### 2.3 Convergence and dealing with fragmentation

The background assumptions and methodologies in qualitative (CA/PTI) and quantitative (EP) approaches are different enough that bringing them together can be somewhat problematic. However, this is not to say that there has been no approach at all to doing so. EP has always been interested in prosodic meaning and the pragmatics of intonation contours, drawing on features of the interactional context to try to explain these differences (cf. Cruttenden 1986; Ladd 1996). Furthermore, EP has become interested in interactional functions such as turn-taking, which are foundational for PTI (cf. e.g. Gravano & Hirschberg 2009). A number of EP studies relate variation in multiple phonetic cues to conversational functions, e.g. questioning/interrogativity (D’Imperio 2000; House 2003; Rialland 2007). PTI, on the other hand, is becoming increasingly interested in larger bodies of data and describing, often quantitatively, what is common to many interactions with similar functions (e.g. Kurtić, Brown, & Wells 2010; Gorisch et al. 2012); and PTI researchers have often been critical of monosystemic approaches to language research (cf. Local 1996; Couper-Kuhlen & Selting 1996).

An explicit attempt to bring together PTI and EP approaches was made by Zellers & Post (2012). They found that variation in speech rate at topic boundaries and within topic groups showed a complex pattern, with new topic boundaries showing overall slower speech rates than other categories, but faster speech rates at the initial edge; that is, new-topic utterances were spoken unusually quickly for the first few syllables, then unusually slowly
over the rest of the utterance. While slow overall speech rates in new-topic utterances were consistent with other literature in an EP-type tradition (Crystal 1969; Quirk, Greenbaum, Leech, & Svartvik 1985; Koopmans-van Beinum & van Donzel 1996), fast speech at topic boundaries was an unexpected finding. However, speeding up at boundaries was consistent with the phenomenon of rush-through discussed in CA literature (Schegloff 1987), wherein a speaker who is completing a turn-constructional unit but who wishes to hold the floor will increase speech rate, avoid pausing, and/or continue an intonational contour into the beginning of a new turn in order to extend his or her own turn. Local & Walker’s (2004) analysis of abrupt-joins also supports this conclusion. Abrupt-joins may occur in interactional contexts where a speaker is completing one action and beginning a new, different one, and they regularly display a set of prosodic characteristics including stepping-up in loudness and pitch at the boundary between the actions. An increase in speech rate is also observed at this point (normally before the boundary), where “finality” cues might otherwise lead to an expectation of decreased speech rate. The mystery of the two opposing speech rate patterns in Zellers & Post's (2012) data is therefore better explained with reference to both research traditions, rather than just one.

While Zellers & Post (2012) explicitly made reference to multiple methods in their investigation of speech rate, their methodology could still be better characterized as EP with a hint of CA, rather than a truly blended approach - in particular because they used read speech, and topic-structure theories consistent with those applied in an EP tradition rather than a CA one. The study reported here develops a new methodology which attempts to achieve a better balance between CA and EP approaches to language.

3 Study: Contrast and Topic

3.1 Methodology development
The methodology presented here represents an attempt to bring together the strong points of two different empirical traditions: quantitative (CA/PTI) and qualitative (EP). The strengths of these two traditions may be summarized (without intending to imply a monopoly by either tradition on the characteristics given) as follows:

- **PTI**: attention to multiple cues (phonetic and otherwise) as they relate to one another; focus on everyday conversational language use; awareness of individual characteristics of any given case of language use
- **EP**: quantification of various phonetic aspects; generalizability of findings due to close control of experimental input; replicability

The task, therefore, is to determine in which parts of the empirical process these strengths should be applied in order to make the most use of them. A proposed three-part process for the investigation of production data is as follows.

### 3.1.1 Selection of data set

A constant (and valid) concern in EP approaches is to ensure that tokens that are compared are similar enough for valid results to be obtained. This has often meant comparing instances of the same sentence produced in different contexts, in order to be sure that the items are comparable. However, when it comes to dealing with spontaneous conversational data, such an undertaking is all but impossible. If we wish to do principled statistical investigations of conversational data, therefore, there must be an alternative way of specifying “sameness”. CA provides this alternative by demonstrating that although the same lexical or sentence tokens may recur extremely infrequently, conversational actions such as assessing, requesting, and inviting (among many others) recur frequently, and with recognizable characteristics as to the types of turns that must appear and, often, the phonetic characteristics of these turns. Thus we have a method for constraining items for a data set without requiring them to have exactly the
same lexical and/or syntactic structure. The data used in this methodology are therefore turns at talk from spontaneous conversation, with the specific format of the turns used in this study discussed in Section 3.2.

### 3.1.2 Analysis of prosodic characteristics

The standards for data-set size and quantification are extremely variable within CA. In earlier studies, in particular, the size of a “collection” of data is often not given; more recent studies sometimes give an approximate number, even specifying that it is not clear whether some items belong in a collection or not. Quantification is often limited to counting of tokens, and phonetic analyses are often impressionistic. Applying quantitative methods to this data can address both of these problems. A statistical analysis could provide a more objective basis for deciding which items are canonical in their production and which are not, allowing for a clearer definition of the collection of data; and it can specify what the acoustic parameters are that lead to the impressionistic findings, leading to better accuracy in predicting further members of one set or another. Quantification can also lead to better testability in a perceptual setting, since it could identify a typical variation level that listeners are, presumably, attuned to through their everyday language use. Some CA-style studies using corpora are already beginning to address this issue (e.g. Kurtić et al. 2010; Gorisch et al. 2012). The current study uses linear mixed models, a statistical method typical of EP-style approaches, to investigate patterns of prosodic variation in the data.

### 3.1.3 Interpretation of results

The results of the statistical analysis are interpreted both with reference to communicative principles (CA) as well as traditional linguistic categories (EP). This means that conversational actions may be as important, or more important, than the formal structural
features of the turns investigated. In particular, variation within what might be considered to be linguistic categories (e.g. broad versus contrastive focus) can be further analyzed in light of what the structure is being used to accomplish (e.g. a demarcation or topic shift). One implication of this is that outlying data need not be (and should not be) treated as errors; cf. Ogden (2012a). Since each conversation occurs in and creates its own specific context, the variation that sets one particular instance apart from the general pattern formed by others is often the key to establishing what the conversational norms are. This, in turn, can lead to new insights into what kinds of actions or situations lead to variation in the data.

3.2 Data selection

The data to be discussed in the analysis below are drawn from a larger data set of 75 tokens of contrast found in conversational data. The conversational data is taken from TalkBank (www.talkbank.org; MacWhinney 2007), a freely-available database of spontaneous conversations, and were drawn from the following:

- 16 telephone conversations (CallFriend corpus; approx. 8 hours)
- 5 call-ins to a radio show (Gulf War corpus; approx. 15 minutes)
- 7 face-to-face conversations (Saarbrücken Corpus of Spoken English {SCoSE}, approx. 1 hour; Santa Barbara Corpus of Spoken American English {SBCSAE}, approx. 3 hours)

All conversations were between native speakers of American English (mostly northern, though a few of the CallFriend conversations were between speakers of southern American English). All conversations were between two parties only at the time the contrast tokens were produced (although some had other participants at other times); the reason for this was to control for possible variation due to next-speaker selection processes (cf. Sacks et al. 1974).
Some characteristics of contrast have been discussed in Section 2.1. However, for purposes of this study, it was important to have a specific set of formal criteria with which to select the tokens. For purposes of this study, contrastive tokens had to have similar syntactic structure with the contrast made as an overt replacement of one (or more than one) element with another, as demonstrated in example (3) (cf. also example 1).

(3)

S: .hhh they're tied down to these dogs
S: I'd rather they be tied down to babies you know
(CallFriend-engn:51)

The characteristics used to define similar syntactic structure were:

- Recycling, or repetition of a portion of an earlier turn; in (3), they [be] tied down to.

(Near-recycling, as in “go” for “walk”, was permitted if it was clear contextually that the two terms were essentially equivalent)

- Since contrasts were made overtly, contrasting elements had the same syntactic class; in (3) these dogs and babies are both noun phrases (though with different internal structure)

Contrasts used in this study were obligatorily not instances of repair. Although cases of contrast used for repair might have been of interest here, it was important for purposes of the study to make sure the tokens in question were as similar as possible. Repair being a conversational function that is well-known to show certain features (cf. Schegloff, Jefferson & Sacks 1977; Selting 1996), it was possible that repair tokens would have muddied the waters of the current investigation; therefore they were not included in the collection of data.

The contrasts investigated in the current study are a subset of this larger collection, consisting of 36 contrasts which occur in turns adjacent to one another in the conversation,
rather than at a distance. That is, they are in turns that either follow one another immediately, or there is an intervening continuer or acknowledgement token: “oh”, “okay”, “right”, “uh huh” or “yeah”. Such acknowledgement tokens or continuers are not necessarily oriented to as turns at talk (cf. Jefferson 1984b; Gardner 2001), so they were not considered as intervening turns. An important feature of the contrasts used in this study is that the contrastive turn follows a preceding turn that is hearable as syntactically complete. Although syntactic completeness is not considered sufficient for identifying the boundary of a turn-constructional unit (TCU), in many cases it is the strongest cue (cf. Selting 2000 for discussion of the problems inherent in defining and bounding TCUs). This requirement was imposed to help ensure that contrasts done by one speaker and contrasts done by two different speakers were comparable.

The small size of the data set was a result of the strict requirements used to ensure that the data were comparable; sequences that were ambiguous in any way with regard to these requirements were discarded, although a planned future direction for this research is to use statistical methods to address such ambiguous cases. The size of the data set was also limited by the quality of recordings and the conversational situations in the available corpora; background noise (especially in face-to-face conversations as in SBCSAE) often meant that turns with the appropriate structure had to be discarded due to inability to obtain accurate acoustic measurements. Since identifying and validating potential tokens was a time-consuming process, and since the study was intended to be exploratory in nature, it seemed most practical to proceed with a small dataset, with more time to be invested in gathering larger bodies of data following proof-of-method.

The contrast sets referred to here have, minimally, two parts: the contrastive turn, and the antecedent with which it is contrastive. (In 5 of the 36 items, there were two contrastive turns; that is, after the first contrastive turn, another following turn with similar
(4)

1     B: we'll probably get as much as they can give us
2     A: yeah
3     B: cause our priority score's so high
4     A: (0.3) yeah
5     B: but umm
6  Pr A: until Newt gets a hold of it
7  A  B: yeah right that's one thing actually
8     A: yeah
9  C  B: an then the other thing is that
10 Po B: .hhh because (0.2) it (.) wasn't voted on until after July first
11     B: .hhh now they're in the next fiscal year

(CallFriend-engn:22; Pr = pre-turn; A = antecedent; C = contrastive turn; Po = post-turn)

3.3 Phonetic analyses

Prior to taking any acoustic measurements, sequences were categorized according to whether or not the contrast was prosodically marked, as determined by impressionistic listening for increased relative prominence of one or more of the contrasting elements compared to the rest of the surrounding (contrastive) turn. Despite the difficulties inherent in the formal description of prosodic contrast (cf. Section 2.1.2), for the current data this categorization was very straightforward to achieve. A second listening by the original labeler several months after the
initial categorization resulted in the same categorization for all items except one (which was
categorized as having prosodic marking the first time but not the second); results reported
below which include this feature are based on the first categorization.

The acoustic analysis was conducted using Praat (Boersma & Weenink 2012). Most
of the recordings were in two channels, one for each speaker. When this was the case, the
channels were separated automatically in Praat and the relevant turns extracted. An overall
measurement of the mean F0 for each speaker was taken by Praat across the whole available
channel (20-30 minutes of recording), with ceiling and floor values set by gender to help
avoid measurement errors. When there was only one recording channel, or when the two
channels were not reliably separable (as in some of the SBCSAE recordings), mean F0
measurements for each speaker were taken over as many stretches of speech as possible, with
the appropriate ceiling and floor values in place, then averaged together to provide a value for
mean speaker F0. All measurements were taken in both Hertz and semitones (as calculated by
Praat), but semitones are used in the statistical analyses, as log scales have been shown to be
more representative of auditory perception in linguistic contexts than Hertz (cf. e.g. Patterson
& Ladd 1999; Nolan 2003), and also provide a means of normalizing across different
speakers.

The turns were labeled for word and syllable boundaries, syllable structure
(onset/rhyme), and a number of pitch points: onset and offset pitch, highest F0 peak (or end of
highest plateau, cf. Knight 2003), and lowest non-final valley. In order to avoid microprosodic
perturbations, only candidate locations within voiced continuant segments were considered,
although breathiness and creakiness could still lead to measurement problems, particularly for
final low F0 values. Within the voiced continuant space, Praat calculated the point of the
highest or lowest F0 value. Where relevant, syllable and/or segment boundaries were
identified following the segmentation principles put forth by Machač & Skarnitzl (2009).
Measurements were then taken automatically using a Praat script and a few measurement errors were corrected by hand (although in some cases, particularly for utterance-final pitch when creakiness was pervasive, it was impossible to get a valid measurement).

3.4 Statistical analyses

Statistical analyses were conducted using the free R software (www.r-project.org). The tests reported below used linear mixed models (R package \texttt{lme4}). Since this was an exploratory analysis and many statistical tests were conducted, the p-values reported below include a Bonferroni correction. Sample R syntax for the statistical analysis is given in the Appendix.

One important feature of all of the statistical analyses is that the data were treated as repeated measures, not only by speaker, but also by item. This is because in each case, the four turns (pre-turn, antecedent, contrast, and post-turn) were taken in order from a conversation. According to CA, an important feature of conversation is that it is sequential. That is, it does not simply occur that one thing happens, then another, then another; but rather than one thing happens which constrains the possibilities of what can follow, which in turn constrains the possibilities of what can come after that. Therefore, the four turns from each set are not independent of one another in how they are constructed, and should not be treated as such statistically.

The main question considered in all cases was how the prosodic features of the four turn types (i.e. pre-turn, antecedent, contrast, and post-turn) varied. The contrastive structures in the collection all conformed to the obligatory syntactic characteristics described in Section 3.2, but they varied on a number of other parameters. They could, but need not, have prosodic contrast marking, as identified by impressionistic listening. They could, but need not, lead to a shift in the topic of conversation. They could be done either by the same speaker or by two different speakers (i.e. with one speaker saying something and another speaker producing a
turn with contrastive form relative to that earlier turn). These characteristics were tested in the statistical analyses as binary factor variables alongside the conversational structure, but are only reported below in cases in which they contributed significantly to the statistical model.

3.4.1 Pitch

The mean pitch of turns (in semitones) did not show any consistent variation with regard to turn type. If, however, we instead take the difference between the mean pitch of the turn and the speaker's mean pitch, measured in semitones, there is observable regularity. Specifically, there was an interaction of turn type and prosodic contrast, where prosodic contrasts were characterized by further deviations from the speaker’s mean pitch. That is, prosodically contrastive items had mean pitch higher than the speaker’s mean in contrasts and post-turns. The mean pitch in pre-turns and antecedents in prosodically contrastive cases was below the speaker’s mean, suggesting a further prosodic differentiation across the turn boundary. Non-prosodically-contrastive turns did not vary significantly from one another in how far they differed from the speaker’s mean pitch (although as shown in Figure 1, there was a non-significant tendency for the pitch to be lower than the speaker’s mean in contrasts and post-turns). Statistics are given in Table 1.

[FIGURE 1 ABOUT HERE]

[TABLE 1 ABOUT HERE]

Another global pitch characteristic of the turns is the internal pitch range. Following Patterson & Ladd (1999), the range is defined as the distance between the highest pitch peak and the lowest non-final pitch valley, measured in semitones. Variation in the pitch ranges of different turn types did not show any significant results; however, variation in the size of the pitch range of turns compared to the pitch range of their preceding turn did. Specifically, there was an interaction between the turn type and whether or not the contrast led to a topic shift:
post-turns had a compressed pitch range compared to contrast turns when the contrast did not lead to a topic shift, but a slightly expanded pitch range compared to contrast turns if there was a topic shift, as shown in Figure 2. Statistics are given in Table 2.

[FIGURE 2 ABOUT HERE]

[TABLE 2 ABOUT HERE]

A commonly-recognized correlate of disjunctive topic shift is peak height. However, in this case, the height of each turn’s highest peak (in semitones) did not vary significantly according to any of the predictors, and peak height compared to the speaker's mean only varied on the basis of the turn type; contrastive turns, regardless of whether they were prosodically contrastive or not, and regardless of whether a topic shift was made in the contrastive turn, had significantly higher peak heights than other turn types. Figure 3 shows a trend towards contrastive turns (and post-turns) having higher pitch peaks when a prosodic contrast was made. However, as shown in Table 3, this trend does not achieve statistical significance.

[FIGURE 3 ABOUT HERE]

[TABLE 3 ABOUT HERE]

While the measurements reported thus far showed relatively robust statistical results, a number of other global pitch measurements did not show any statistically significant effects. In addition to “absolute” measures (i.e. those not taking into account either the speaker's characteristics or the characteristics of an adjacent turn), these included the utterance-initial and utterance-final pitch compared to the speaker's mean, and utterance-initial pitch compared to the final pitch of the previous utterance (either independently or relative to the speaker's mean). The lack of finding in this last comparative measure may be due to the fact that in many cases it was impossible to get a valid measurement for the final pitch due to creakiness,
so there were relatively few data points.

3.4.2 Speech rate and rhythm

Speech rate was measured in terms of average syllable duration (ASD). Syllables rather than segments were chosen as syllables were more reliably identifiable in the current data; phonetic reduction phenomena often led to difficulty in identifying plausible segment boundaries. However, ASD, either as an isolated measure or in comparison with a preceding turn, did not vary consistently with any of the factors tested (turn type, presence of prosodic contrast, speaker change, presence of topic shift). Turns were not consistent either in number of syllables or in the structure of the syllables; the diversity of syllable types across the different categories in the analysis is possible reason for this lack of result. A reviewer suggested that rhythm metrics used in typological studies might be appropriate, but neither %V (Ramus, Nespor & Mehler 1999) nor nPVI-V (Low, Grabe & Nolan 2000; Grabe & Low 2002), when tested, led to a statistically significant pattern; this is likely also due to the high variability in syllable structure in the data.

CA studies involving rhythm sometimes include in their metrics, in addition to ASD, the ratio of the number of stressed syllables to unstressed (or all) syllables in a turn. Uhmann (1992) reports that using both syllable rate and what she terms stressed-syllable rate makes it possible to make finer distinctions among activity types or speaker attitudes; for example, a combination of both of these rates can distinguish emphatic or parenthetical sections of talk. In the current study, the ratio compares the number of stressed syllables to the total number of syllables in each turn. Syllables counted as stressed in the analysis obligatorily had a non-reduced vowel, a local amplitude peak, and a pitch movement on, to, or from the nucleus.

Once again, the stand-alone ratio measure did not show statistically significant variation with regards to turn type. However, when the ratio of a turn was compared to the
ratio of the preceding turn, there was a statistically significant effect of prosodic contrast, as well as a marginally significant interaction between turn type and prosodic contrast. Specifically, if the contrast was not marked prosodically, the ratio of stressed syllables tended to decrease from previous utterances in all turn types, while if it was prosodically marked turns tended to have approximately equal ratios of stressed syllables. The marginal interaction shows post-turns following prosodic contrast having a decreased ratio of stressed syllables compared to their preceding contrastive turns. This is shown in Table 4 and Figure 4.

[FIGURE 4 ABOUT HERE]
[TABLE 4 ABOUT HERE]

3.5 Discussion

The findings reported here range from the relatively predictable to the frankly unexpected. Therefore, the last step of the analysis, re-applying CA interpretative principles to the data, is essential in order to demonstrate the validity of the findings. The discussion will thus be organized around conversational actions rather than around the individual prosodic characteristics.

Before entering into specific discussion, one general point may be raised. The current methodology has demonstrated the importance of considering sequential and contextual features. The features that varied significantly were relative features: comparisons between the features of the turn and the turn that preceded it. Analyses of situated language use must take these features into account; otherwise they risk missing important sources of variation.

3.5.1 Contrast

One broad characterization that we can make about contrast in this study is that pitch is only one of the contributing factors to prosodic contrast, and probably not the strongest one. While pitch peak height appears visually to be higher in prosodically contrastive turns and the post-
turns following them, cf. Figure 3, this difference was not statistically significant, except that the mean peak height for all contrastive turns (prosodically contrastive or not) was higher than that for other turn types. Pitch level raising contributes to prosodic contrast, with contrastive turns and their post-turns having a higher mean pitch than the pre-turns and antecedents preceding them when a prosodic contrast is made. However, this suggests more of a boundary phenomenon than simply contrast, since the post-turns in contrast cases also had raised pitch. The lower mean pitch values in pre-turns and antecedents preceding prosodically-marked contrasts also contribute to the perceptibility of this boundary, though given that there is overlap between the mean pitch comparisons in pre-turns and antecedents in prosodically contrastive and non-contrastive cases, this cannot necessarily be treated as anticipating the prosodic marking of the contrast.

It is interesting to note that the global pitch features of post-turns following prosodic contrast were similar to the contrastive turns themselves. Higher pitch may be associated with disjunction and the presence of something different or new, and the post-turns may thus have similar pitch features to show that they are linked with the contrastive turn. We can observe that contrast, even when it does not make a topic change, can serve a demarcative function. Returning to example (4) we see that the antecedent turn *yeah right that's one thing actually* (l. 7) serves to summarize or close down something that has gone before, while the contrastive turn *an then the other thing is that* (l. 9) opens something new. It would be difficult to justify treating this as a topic change, as the talk continues on the grant that B's research group is to be awarded; nonetheless there is a sense of one thing ending and another beginning. Some of the uses identified by other authors for contrast, such as setting up an opposition and then aligning with one side (cf. Ford 2000), would fit well with this kind of demarcative function; in order to strengthen the alignment with one side or another, a speaker needs to be clear about what the two sides of an issue are.
The rhythm data reported here were inconclusive, with the ratio of stressed syllables holding approximately steady in sequences where contrast was (impressionistically) prosodically marked, and decreasing across sequences where contrast was not prosodically marked. The marginally significant finding that post-turns in sequences with prosodic contrast marking had a reduced ratio of stressed syllables compared to their preceding turns raises the possibility that rhythmic characteristics may play a stronger role than pitch characteristics in marking prosodic contrast. This is not as surprising as it may at first seem; Kochanski, Grabe, Coleman & Rosner (2005) found that prominence in English is not primarily marked by pitch, but rather by variations in lengthening and amplitude. However, further research is necessary to determine whether this is a real effect or simply an artifact of the current data.

While the global pitch characteristics of these turns did not differ enough from other turn types to be able to make clear a distinction, the prosodically contrastive turns were nonetheless identifiable auditorily. It is possible that more specific comparisons (e.g. of adjacent pitch peaks, or of trends of pitch peak heights) would identify more F0-based features of prosodic contrast that make it auditorily distinct from non-prosodically marked contrasts. One possibility that has not been explored here is that it is not the strength of prominence, but its location, which differs between the prosodically contrastive and non-prosodically contrastive turns: pitch accent placement might vary, for example. This seems likely, as contrast may be used to select a particular set of items for comparison, but it is beyond the scope of the current work.

3.5.2 Topic shift

One of the findings of this research that was most surprising was that of the pitch range variation associated with topic change – specifically, that it did not occur in the contrastive turn leading to topic change, but in the post-turn following. As far as the author is aware, no similar findings have been reported. Other findings on variation in pitch span (Swerts et al.
1994 on list items; Zellers 2011) found increased pitch span with new unit beginnings. Zellers (2011) also found evidence of at least one intermediate level within topic structure. Falling pitch contours in new topic beginnings tended to have a pitch span of around 5 semitones, falling contours in pure continuations tended to have a span of around 3 semitones, and items which were intermediate between the two (i.e. adding information on the same topic, or elaborating on a previous utterance) had falling contours with a span of around 4 semitones.

The current data are not directly comparable with Zellers' (2011) findings, since she measured the size of the first (prenuclear) falling pitch movement, while the current study measured the overall pitch range. However, we may note a certain similarity: post-turns which did not follow a topic shift showed a compressed pitch range, having a range around 5 semitones smaller than the contrastive turn preceding them. Conversely, post-turns which did follow a topic shift showed an expanded pitch range, having a range around 1.5 semitones larger than the contrastive turn preceding them. Compression following nuclear prominence is a well-known phenomenon in EP (e.g. Eady & Cooper 1986, Xu & Ching 2005 for English; de la Mota Gorriz 1997 for Peninsular Spanish; Jun & Lee 1998 for Korean; Chen, Wang & Xu 2009 for Mandarin); and in fact this compression in post-contrast turns may also contribute to why we do not see a significant difference in pitch between prosodically contrastive and non-prosodically contrastive turns (cf. section 3.4.1). That is, the prosodic prominence may be increased by following compression, rather than by expansion in the contrastive turn itself. However, if that is the case, then it would appear that an increase in pitch range (and thus, potentially, an increase in prominence) in the post-turn is done on purpose or marked, and therefore potentially meaningful.

There are at least two possible explanations for this expansion following the contrastive turn introducing a new topic. The first option is that it may simply be more appropriate to say that the post-turn is the first turn of the new topic. This would provide a
convenient explanation for the pitch expansion, but is difficult to reconcile with the fact that it
is the contrastive turn which introduces the new item(s) that make up the new topic. The
second explanation relies on a closer investigation of the form of the topic shifts. The turns in
question are very similar in form to the types of topic shifts that Sacks (1972 [1992]) and
Jefferson (1984a; 1993) have described as “stepwise” changes. Sacks defines a stepwise move
as involving “...linking up whatever is being introduced to what has just been talked about,
such that, as far as anybody knows, a new topic has not been started, though we're far from
wherever we began.” (1972 [1992]: 566) Jefferson (1984a: 202-204) further specifies the
form of such stepwise topic changes away from troubles-telling as regularly showing a series
of features: (1) a summary of the main problem; (2) turning to a matter which is ancillary to
the trouble; (3) topical stabilization of the ancillary matter; (4) production of a “pivotal”
utterance, which is on-topic but has “independent topical potential”; (5) the new topic is now
treated as the current topic.

It is not difficult to draw parallels between this structure, particularly steps 3-5, and
the structure of the contrastive topic shifts in the current data set. In each, the antecedent
represents something still linked to the prior subject of talk. Returning to example (1), you
have more of a chance of getting killed in the plane ride over there (ll. 4-5) continues the prior
discussion (of reasons why U’s proposed trip is not as dangerous as her parents think). The
contrastive turn serves as a “pivot”, which can be treated as on-topic in that it shows a clear
link to the antecedent, but which has “independent topical potential” in that it provides a new
perspective or a new referent. In example (1), U’s turn I have more n- chance of getting killed
in my own neighborhood (ll. 7-8) links to the previous talk about chances of death, but
introduces the new perspective that this could even happen at home. The post-turn then treats
the contrastive turn has having already achieved the topic shift, and continues with what is
now on-topic talk; in example (1), U tells her newly introduced story about hearing gunshots
while walking her dog (ll. 9-15).

If we accept these parallels and treat topic shifts initiated by contrast as stepwise, rather than disjunctive shifts (as indeed would probably already be the best analysis simply because there is no obvious topic boundary to identify), then the prosodic variation in the post-turns observed above takes on a new interpretation. Namely, we can propose that despite the stepwise nature of the topic shift, conversational participants are still orienting to the fact that a topic shift has occurred. The relatively small increase (1.5 st) in pitch range size is commensurate with the fact that stepwise shifts are more hidden than overt, but it is consistent with the idea that something new is still happening, rather than a simple continuation. In other words, Sacks' interpretation that “as far as anybody knows” nothing has happened is not quite correct. Rather, conversational participants are collaborating to both acknowledge that topic shift has occurred while treating the topic shift as non-disjunctive (or perhaps non-disruptive). This is an important distinction because it allows participants to reap the cognitive benefits of a topic shift (e.g. clearing short-term memory for previous referents, cf. Grosz & Sidner 1986) while also reaping the interactional benefits of non-disruption (conversation is ongoing, cf. Maynard 1980). Furthermore, it emphasizes that topic shift is an action that conversationalists accomplish and orient to in the course of language use.

4 Conclusion

The methodology reported here draws upon two very different theoretical backgrounds: CA/PTI, qualitative approaches, on the one hand; and EP, a set of quantitative approaches, on the other. Combining insights PTI and EP makes it possible to take advantage of strengths in both of the theoretical and methodological backgrounds, which opens up new possibilities for investigation. By prioritizing (or even simply including) qualitative sequential features of speech in a quantitative analysis, a high level of experimental control may be maintained.
The set of findings reported here are different in form and scale from what might be expected of either of these approaches, yet are interpretable from both points of view. The value of this research is not only in the novel finding that stepwise topic shifts may have a consistent prosodic structure that differs noticeably from disjunctive topic shifts, but also in that the reasoning which leads to this conclusion depends upon both quantitative and qualitative principles. In dealing with the finding that rhythmic variation may be more closely related to prosodic contrast marking than pitch characteristics, at least in terms of global features, we have recourse to evidence from EP accounts of prosodic perception (e.g. Kochanski et al. 2005), while at the same time being able to allow for different features of prosodic variation being relevant on the basis of different interactional needs in individual situations.

The current study is exploratory and as such, opens up a wide range of possibilities for future research. Already in progress is a study which will use a different statistical method, cluster analysis, to gain a better understanding of how bundles of phonetic features may be constructed to help accomplish different conversational activities, or to accomplish the same activity in different contexts. Future research can be more sophisticated in its statistical methods; given the high degree of variability in conversational speech, larger data sets and more powerful statistical tests will be essential modifications. Similarly, applying linguistic theory (for example, intonational categories) to well-established CA activity categories such as agreement/disagreement, or conversational openings and closings, could lead to a better understanding of how these structural linguistic means may be deployed in everyday language use. Bringing together qualitative and quantitative approaches thus opens up possibilities for an ever-more complex and nuanced picture of language use.

Acknowledgements
I am grateful to Richard Ogden for mentoring on this project and for detailed comments on an earlier version of this paper; to Becky Taylor and Traci Walker for suggestions guiding the direction of the research; to Dan Brenner, Antonio Origlia, Jessamyn Schertz, and Elaine Schmidt for discussion about rhythm; and to the associate editor, Francis Nolan, and two anonymous reviewers for their helpful comments. This research was supported by the ESRC Postdoctoral Fellowship ‘Multiple-methods approaches to discourse topic structure’ (ES/J005711/1).
Works Cited


Chen, S.; Wang, B.; Xu, Y.: Closely related languages, different ways of realizing focus.


de la Mota Gorriz, C.: Prosody of sentences with contrastive new information in Spanish.


Gravano, A.; Hirschberg, J.: Turn-yielding cues in task-oriented dialogue. Proc. SIGDIAL,


Knight, R.-A.: Peaks and plateaux: the production and perception of high intonational targets


Appendix: R Syntax

Sample R syntax for linear mixed models of variation in pitch mean differences is given below. Both models take turn type, topic shift and prosodic contrast as predictors; model 1 includes two random factors, item and speaker, while model 2 includes only item as a random factor. The ANOVA compares the two models (i.e. to see whether the random factor for speaker improves the model or whether it can be left out). Viewing the summary of model 2 shows the internal structure of the model, which constitutes the data reported in Tables 1-4.

```r
model1 <- lmer(pitchmeandiff ~ turntype * topshift * proscont + (1 | item) + (1 | speaker))
model2 <- lmer(pitchmeandiff ~ turntype * topshift * proscont + (1 | item))
anova(model1, model2)
summary(model2)
```
List of Figures

Figure 1: Mean pitch difference from speaker mean, in semitones, for each turn type (pre-turn, antecedent, contrast, post-turn). The height of the bar in relation to the 0 line represents the difference between the mean pitch of the turn, and the speaker's mean pitch; the different colors show whether or not there was a prosodic contrast in the contrastive turn.

Figure 2: Variation in pitch range of different turn types, in semitones. The 0 line represents the pitch range of the preceding turn; the different colors indicate whether or not a topic shift occurred in the contrastive turn of the sequence.

Figure 3: Peak height in different turn types, in semitones. The different colors represent whether the contrastive turn was prosodically contrastive (as identified by impressionistic listening). Although this graph shows F0 peaks to be higher in contrastive turns (and post-turns) when a prosodic contrast is made, this difference is not statistically significant when individual variation is taken into account in the statistical model.

Figure 4: Ratio of stressed syllables compared to ratio in previous turn; y-axis units are the difference in ratios between adjacent turns. The 0 line represents the ratio of stressed syllables in the preceding turn; the different colors show whether there was a prosodic contrast made in the contrastive turn.
List of Tables

Table 1: Phrase pitch mean difference from speaker mean: linear mixed model. * = p < .05.

Table 2: Pitch range: linear mixed model. * = p < .05. · = p ≈ .05.

Table 3: Peak height: linear mixed model. * = p < .05

Table 4: Ratio of stressed syllables compared to previous turn: linear mixed model. * = p < .05. · = p ≈ .05.
Table 1: Phrase pitch mean difference from speaker mean: linear mixed model. * = p < .05.

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>BIC</th>
<th>logLik</th>
<th>deviance</th>
<th>REMLdev</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>556.4</td>
<td>583.7</td>
<td>-268.2</td>
<td>547.4</td>
<td>536.4</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>item</td>
<td>(Intercept)</td>
<td>2.3080</td>
<td>1.5192</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>6.0788</td>
<td>2.4655</td>
</tr>
</tbody>
</table>

Number of obs: 114, groups: item, 29

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.24607</td>
<td>1.18226</td>
<td>-0.208</td>
</tr>
<tr>
<td>antecedents</td>
<td>0.27842</td>
<td>1.42347</td>
<td>0.196</td>
</tr>
<tr>
<td>contrasts</td>
<td>-1.42959</td>
<td>1.42347</td>
<td>-1.004</td>
</tr>
<tr>
<td>post-turns</td>
<td>-1.29105</td>
<td>1.42347</td>
<td>-0.907</td>
</tr>
<tr>
<td>prosodic-contrast</td>
<td>-1.30784</td>
<td>1.33285</td>
<td>-0.981</td>
</tr>
<tr>
<td>antecedents:prosodic-contrast</td>
<td>0.03355</td>
<td>1.60281</td>
<td>0.021</td>
</tr>
<tr>
<td>contrasts:prosodic-contrast</td>
<td>3.37835</td>
<td>1.60281</td>
<td>2.108 *</td>
</tr>
<tr>
<td>post-turns:prosodic-contrast</td>
<td>3.40450</td>
<td>1.60729</td>
<td>2.118 *</td>
</tr>
</tbody>
</table>
Table 2: Pitch range: linear mixed model. * = p < .05. · = p ≈ .05.

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>logLik</th>
<th>deviance</th>
<th>REMLdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>416.7</td>
<td>434.6</td>
<td>-200.4</td>
<td>416.9</td>
<td>400.7</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>item</td>
<td>(Intercept)</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>26.988</td>
<td>5.195</td>
</tr>
</tbody>
</table>

Number of obs: 69, groups: item, 29

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.923</td>
<td>1.964</td>
<td>1.998 ·</td>
</tr>
<tr>
<td>contrasts</td>
<td>-2.294</td>
<td>2.560</td>
<td>-0.896</td>
</tr>
<tr>
<td>post-turns</td>
<td>-8.418</td>
<td>2.777</td>
<td>-3.031 *</td>
</tr>
<tr>
<td>topic-shift</td>
<td>-3.550</td>
<td>2.378</td>
<td>-1.493</td>
</tr>
<tr>
<td>contrasts:topic-shift</td>
<td>1.265</td>
<td>3.169</td>
<td>0.399</td>
</tr>
<tr>
<td>post-turns:topic-shift</td>
<td>9.196</td>
<td>3.382</td>
<td>2.719 *</td>
</tr>
</tbody>
</table>
Table 3: Peak height: linear mixed model. * = p < .05

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>BIC</th>
<th>logLik</th>
<th>deviance</th>
<th>REMLdev</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>465.5</td>
<td>484.7</td>
<td>-224.8</td>
<td>462.5</td>
<td>449.5</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>item</td>
<td>(Intercept)</td>
<td>1.4625e-09</td>
<td>3.8242e-05</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>1.9091e+01</td>
<td>4.3693e+00</td>
</tr>
</tbody>
</table>

Number of obs: 81, groups: item, 29

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.949</td>
<td>1.261</td>
<td>0.752</td>
</tr>
<tr>
<td>contrasts</td>
<td>4.466</td>
<td>1.784</td>
<td>2.504 *</td>
</tr>
<tr>
<td>post-turns</td>
<td>3.228</td>
<td>1.871</td>
<td>1.726</td>
</tr>
<tr>
<td>topic-shift</td>
<td>2.630</td>
<td>1.669</td>
<td>1.576</td>
</tr>
<tr>
<td>contrasts:topic-shift</td>
<td>-3.766</td>
<td>2.345</td>
<td>-1.606</td>
</tr>
<tr>
<td>post-turns:topic-shift</td>
<td>-1.831</td>
<td>2.461</td>
<td>-0.744</td>
</tr>
</tbody>
</table>
Table 4: Ratio of stressed syllables compared to previous turn: linear mixed model. * = p < .05. · = p ≈ .05.

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>logLik</th>
<th>deviance</th>
<th>REMLdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>-43.97</td>
<td>-23.97</td>
<td>29.99</td>
<td>-86.29</td>
<td>-59.97</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>item</td>
<td>(Intercept)</td>
<td>0.00000</td>
<td>0.00000</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.02405</td>
<td>0.15508</td>
</tr>
</tbody>
</table>

Number of obs: 90, groups: item, 31

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.13181</td>
<td>0.06331</td>
<td>-2.082 *</td>
</tr>
<tr>
<td>contrasts</td>
<td>0.08422</td>
<td>0.08954</td>
<td>0.941</td>
</tr>
<tr>
<td>post-turns</td>
<td>0.16438</td>
<td>0.09391</td>
<td>1.750</td>
</tr>
<tr>
<td>prosodic-contrast</td>
<td>0.14548</td>
<td>0.07078</td>
<td>2.055 *</td>
</tr>
<tr>
<td>contrasts:prosodic-contrast</td>
<td>-0.08013</td>
<td>0.09990</td>
<td>-0.802</td>
</tr>
<tr>
<td>post-turns:prosodic-contrast</td>
<td>-0.20794</td>
<td>0.10403</td>
<td>-1.999 ·</td>
</tr>
</tbody>
</table>