F0 Troughs and Prosodic Phrasing

Anne Wichmann
Department of Languages
University of Central Lancashire, Preston PR1, UK

ABSTRACT

This paper describes prosodic boundary phenomena in naturally occurring speech. In two different speaking styles, read and spontaneous, a systematic low trough in the f0 contour is found to correlate with both syntax and topic structure. There is also evidence for systematic topline declination between such boundaries, but only in the read speech.

INTRODUCTION

There is considerable interest in the nature and function of prosodic boundaries. Knowles (1991) identifies a set of features (temporal, intonational and segmental) which co-occur with perceived tone group boundaries in the Lancaster/IBM Spoken English Corpus (SEC). This contains 50,000 words of prosodically transcribed speech. Bruce et al (1991), working on Swedish, investigate the temporal and intonational features of prosodic phrasing in association with syntactic boundaries.

The identification of prosodic boundaries forms the basis of earlier descriptions of larger prosodic domains, for example the "pitch sequence" (Brazil et al 1980) and "paratone" (Brown 1977). The phonetic cues to the boundaries of these domains are sought in pitch maxima on prominent syllables and in pitch movement. Brown (op cit) also claims an internal systematicity consisting of regularly descending pitch height on stressed syllables within each paratone.

There are therefore two issues to be addressed:
(i) What are the phonetic features which cause phrase boundaries to be perceived?
(ii) Do these boundaries define prosodic domains which also display systematic internal features?

The present paper describes the distribution of marked f0 minima (troughs) in relation to other prosodic features and in relation to the text itself. In its weakest form this feature is an f0 trough which is low in relation to preceding and following f0 minima; in its strongest form it is an f0 trough which is close to the speaker's base line and displays approximately the frequency in each case. The following describes the distribution of the strong form of this boundary feature in two different speaking styles, read speech (a news broadcast), and spontaneous speech. The aim of this research is to identify significant prosodic features which can be labelled in a speech data-base, providing information for a large-scale study of prosodic segmentation.

ANALYSIS I - NEWS BROADCAST

The section of text analysed here is the first news item (six sentences) in a broadcast news summary, taken from the Spoken English Corpus. For this speaker, the lowest in each series of f0 troughs was approximately 75Hz.

Syntactic correlates

In the section of text analysed there were 14 occurrences of the low f0 trough. Of these, all coincided with a syntactic boundary; 6 of these were sentence boundaries (one for each sentence); 5 were finite clause boundaries within a sentence, and 3 were sentence-internal

phrase boundaries. Of the phrase boundaries, one marked the end of the NP subject of the first sentence, and two marked the beginning of a sentence-final adverbial (prepositional) phrase. Of the 15 finite clauses in the text, only four are **not** marked prosodically with a low f0 trough.

Textual correlates

The placement of a strong prosodic boundary at minor syntactic boundaries does not appear to be random. The f0 trough which co-occurs with the end of the noun phrase subject of a sentence has a textual function. The end of the first noun phrase at the beginning of a news item is often treated as prosodically separate; the final prominent syllable is frequently given a falling contour instead of the non-final contour (rise or fall-rise) normally expected at a non-final syntactic boundary. This alone gives the impression of a degree of finality, which is reinforced if the pitch falls to the speaker's base-line at this point. This is the prosodic equivalent of printing the first phrase of a newspaper paragraph in block capitals or bold print. The prominence given in this way, either visually or aurally, serves to mark a new topic in the text.

The f0 trough at the end of sentences is a predictable indication of syntactic completeness. Its occurrence before a sentence-final adverbial is typical for professional news broadcasts. Readers tend to mark a degree of finality at the earliest point at which the sentence is **potentially** complete. This may be related to mental processing of the text ahead; it may also be that there is a maximum or minimum length for a prosodic phrase (if the boundaries in question can be assumed to demarcate units of some kind).

Co-occurrence with other prosodic features

The prosodic transcription of this section of text contains 6 major tone group boundaries. All of these co-occur with a low f0 trough. There are also 23 perceived minor boundaries, only 6 of which are signalled in this way. Thus 12 of the 14 f0 troughs occur at perceived boundaries. The two occasions where the feature does **not** co-occur with a tone group boundary are before adverbial phrases. Pausing in read speech tends to be at syntactic boundaries; it is therefore not surprising that there is a high degree of co-occurrence in this text between low f0 troughs and pauses. Of the 14 f0 troughs, 11 co-occur with pauses, 6 at the end of sentences, four at clause boundaries and one at a phrase boundary. The remaining f0 trough at a clause boundary co-occurs with final syllable lengthening.

Internal regularities

Across each section of text between two f0 troughs there was a systematic declination of f0 maxima. This suggests that, in this speaking style at least, there may be a prosodic unit - a phrase or declination domain - which is larger than a tone group, and can be defined both in terms of its boundary features and in terms of internal regularity.

ANALYSIS II - SPONTANEOUS SPEECH

The material analysed here was part of a spontaneous, unrehearsed monologue by a female speaker. As with the first text, the base line of the f0 contour was analysed. The same marked f0 troughs were observable and displayed a consistent f0 value. (159Hz)

Syntactic correlates

Each of the f0 troughs co-occurred with a syntactic (phrase or clause) boundary. However, they occurred overall less frequently than in the read speech and there were many syntactic boundaries which were not marked in this way. In the sequence analysed there were 35

finite clauses. Only 13 of a total of 22 f0 troughs occurred at such clause boundaries; a further 9 occurred at phrase boundaries (6 before final adverbials; 1 before a subject complement in final position; two after an NP subject).

Textual correlates

The topic structure in this text was identified independently, partly semantically and partly by means of other prosodic features: for example there was a marked increase in speech rate immediately preceding a new topic and the new topic itself co-occurred with higher than average f0 maxima. While the read text dealt only with one topic, i.e. one news item, the spontaneous text contained a range of topics. Although the relation between syntactic boundaries and f0 troughs was weaker than in the read text, the association with topic structure remained the same. The f0 troughs marked, not surprisingly, the end of a topic. They also marked the end of the noun phrase or clause announcing a **new** topic. (see Figure 1)

Co-occurrence with other prosodic features

In the same stretch of speech there were 46 pauses. Of these only 8 co-occurred with an f0 trough. 14 pauses were hesitation phenomena and associated with repetition or pause fillers ("er") in the running text. 9 Occurred immediately **after** clause boundaries, i.e. after either a conjunction or discourse marker *(and, anyway, for example)*. Low f0 troughs on the other hand, if they occurred at all, occurred **before** a conjunction or discourse marker, i.e. at the syntactic boundary. (see Figure 1)

Internal regularities

Unlike the read text, the systematic declination of f0 peaks did **not** occur in this text. The stretches between low f0 troughs were on average much longer than in the read speech, varying between one word and 68 words but averaging around 12 words. Unlike the read text, these stretches did not display a systematically declining topline, and therefore show no evidence of prosodic units which could be defined in this way.

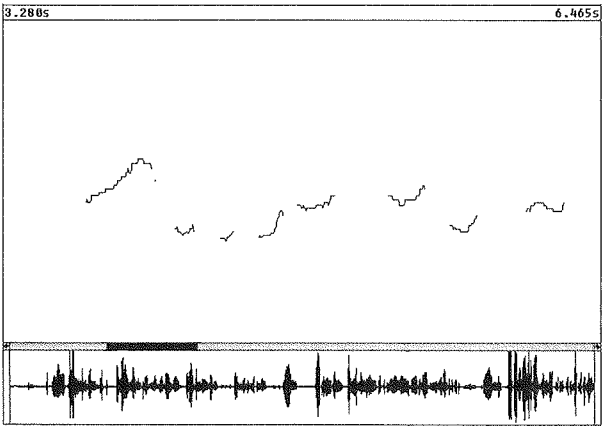

Figure 1. *Waveform and f0 contour of the text: "(She asks me things that) amaze me sometimes I mean for example ..". This is the start of a new topic. There is a low f0 trough on the first syllable of "sometimes" and a pause after "for example".*

DISCUSSION

These two speaking styles are markedly different both prosodically and syntactically. The spontaneous speech, for example, displays far fewer f0 boundary troughs than the read speech, and these do not co-occur to the same extent with other boundary features, particularly pauses. There are nonetheless similarities. The occurrence of low f0 troughs is in each case related in some way to syntax. Both texts also have in common that the low f0 troughs serve to indicate topic structure. This is particularly marked in the spontaneous speech. This suggests that the relationship to syntax is subordinate to, or a function of, the relationship to meaning. Further research is in progress to examine the distribution of f0 troughs in other speaking styles. The topic structure of both texts analysed so far is relatively simple. It remains to be seen whether the more complex topic structure of, for example, fiction is reflected in the nature of the prosodic boundaries.

Evidence for larger prosodic units is less convincing. Only the read speech displays any internal regularity between boundaries, suggesting that such regularity may only be a feature of impersonal read speech.

IMPLICATIONS

The results of this investigation support the view that the base line of the f0 contour holds important information. As Bruce (op cit) suggests, a lowered base line is a signal of phrasing. My research indicates that in naturally occurring speech these low points may signal phrase boundaries not only because they are low in relation to the context but also relative to the speaker's own range. Further research is necessary to establish whether this is common to all speakers and all speaking styles. If this is so, the occurrences of f0 minima could be of use in speech recognition, by giving not only syntactic information but also information about structure at the level of the text. Such insights might also serve to produce a more realistic sounding intonation contour in synthesised speech.

REFERENCES

Brazil, D., Coulthard, M., Johns, C. (1980) *Discourse Intonation and Language Teaching* (Longman, Essex).

Brown, G. (1977) *Listening to Spoken English* (Longman, London).

Bruce, G., Granström, B., Gustafson, K., House, D. (1991) *Prosodic Phrasing in Swedish* in Working Papers 38, 5-17 Lund University, Dept of Linguistics

Knowles, G. (1991) "Prosodic labelling: the problem of tone-group boundaries" in S. Johannson, A-B. Stenström (eds) *English Computer Corpora* (de Gruyter, Berlin)