

*Anne Wichmann**

F0 troughs and prosodic phrasing: an investigation into the linguistic information contained in a speaker's baseline when reading

Abstract

This paper examines some aspects of the prosody of reading aloud. The pitch contour of a read text was examined, and the low troughs, or valleys, in the contour were identified. The pitch of these troughs, and their position in the text, was found to relate systematically to the structure of the text. This suggests that the speaker's pitch baseline may carry important linguistic information.

The framework within which a speaker exploits pitch for linguistic purposes is the "linguistic range" - a subsection of the physical range of the speaker's pitch (measured as fundamental frequency or f_0). Excursions outside that range are sometimes regarded to be of paralinguistic origin, such as the expression of extreme surprise or dismay (Laver 1994). Within the linguistic range, a speaker's pitch habitually spans a range which changes over the course of an utterance, from wide to narrow and from high to low. The two framing components of utterance range are known as the baseline and topline. These are notional lines drawn roughly through all the peaks of a contour (the topline) and the troughs (the baseline). It is generally assumed that both tend to decline over an utterance, (hence the term "declination") and gradually converge.

Much attention has been paid to the behaviour of the topline. The positions and relative height of peaks highlight the information structure of a sentence, indicating focal and non-focal elements, and can also indicate the topic structure of the text. Less work has been done on the behaviour of the baseline. The downward trend is thought to be "re-set"

* *Anne Wichmann*
Department of Cultural Studies
University of Central Lancashire
Preston PR1

slightly at clause boundaries (Bruce et al 1991) and a major reset generally occurs at the beginning of a new topic or paragraph.

However, the point at which the baseline of an utterance falls to the lowest boundary of a speaker's neutral linguistic range, in other words the point at which the declining trend lines converge and "hit bottom", is widely assumed to occur only at the end of complete utterances - sentences and paragraphs, giving a clear "I've finished" signal. (Menn and Boyce 1982) This fall to the bottom of the range is often referred to as a "low terminal" (Brazil et al 1980). It is an important element in intonation theory because it is seen as delimiting an intonational domain, a concept which is essential for intonation analysis. This domain is assumed to be usually coterminous with a sentence.

It seems then that the baseline of a speaker's pitch span carries important information: syntactic information regarding clause and sentence structure, semantic information regarding topic structure, and interactional information as in turn-taking. Preliminary studies reported in earlier papers (Shockey and Wichmann 1992; Wichmann 1993) suggested that in spontaneous monologue the low terminal could also occur inside sentences. The present study describes an instrumental analysis of the baseline behaviour of a formal read text and suggests that the results are similar to those for spontaneous speech.

Data

The data used in this analysis is a Radio 3 news summary from the Lancaster/IBM Spoken English Corpus (SEC). This was chosen firstly as an example of attitudinally neutral speech and secondly for the lack of complexity in the structure of the text. A total of twenty-five sentences were analysed, representing the first six of a total of nine news items, and excluding the initial metatextual sentence ("Here is the news"). The speaker is an experienced professional broadcaster.

Analysis

In a normal f₀ trace there are many minor fluctuations, some representing local segmental effects and others of physiological origin. In order to exclude these, and identify only marked f₀ minima or troughs, a window of 0.5 seconds was imposed. It was assumed that f₀ peaks of prominent syllables would not be more than 25cs apart, and since the

present study was concerned to find evidence of intonational phrasing above the level of the foot, this was doubled. Thus any f0 trough was recorded which was not followed by a lower trough within a time window of 50cs. A second pass identified each trough which was lower than those either side i.e. the troughs of the troughs. These are referred to in this paper as “low troughs”.

Relationship between low troughs and text

Low troughs typically mark the ends of sections of spoken text in much the same way as punctuation delimits sections of written text. When the low trough is the end-point of a falling contour, which is the usual case, then it comes right at the end of the section. However, in many cases in this data the trough was in fact the lowest point in a falling rising contour, in which case it comes immediately before the rise which completes the section.

Low endpoints are most commonly associated with finality and therefore with the ends of sections of text at least the size of a sentence. One might therefore expect the number of low troughs to be equal to the number of sentences in the text. (Although the sentence is a unit of writing, the term is used here because the spoken text was a reading aloud of a written text.) In fact there were 51 low troughs in the 25 sentences analysed. This means that each sentence contained an average of 2 low troughs, suggesting a final low end-point and one other. Except for two sentences this is indeed the case. The question is therefore: where do the “others” occur? The exact distribution is shown in Table 1.

Table 1

23	end of sentence
12	end of penultimate sentence element (11 = adverbial; 1 = appositional phrase)
11	end of first sentence element
2	end of co-ordinate clause (i.e. before “and “ or “but”) (10 = subject; 1 = adverbial)
1	before “of” -phrase
1	before sentence-final NP object
1	after subject element in 2nd position (after initial adverbial)
==	
51	total

Of these 51 low troughs, as can be seen in the table, 23 occur at the end of a sentence, 12 before a final adverbial or appositional phrase, and 2 at the end of a co-ordinate main clause. If we accept the fact that a sentence is at least potentially complete before a final adverbial, appositional phrase or at the end of a main clause, it can be claimed that these 37 low troughs (72.5%) co-occur with either the end or with the “beginning of the end” of a sentence, and could be considered to indicate finality.

However, there are 14 remaining troughs which cannot be accounted for in this way. 11 of these occur at the end of the first sentence element, one before an “of” phrase, one before a sentence-final NP object, and one after the second sentence element. Disregarding the last three examples it is clear that 11 low troughs (over 20%) are associated with the beginning of a sentence (or rather the “end of the beginning”) rather than with the end.

Baseline values in relation to strength of (syntactic) boundary

Although the low troughs in the baseline are clearly not only associated with the ends of utterances, they are nonetheless all at syntactic (phrase or clause) boundaries of some kind. It is perhaps reasonable to assume that the actual f_0 value of an individual low trough may bear some relationship to the strength of the boundary it accompanies. For example, the end of a sentence may be marked by a particularly low endpoint, at the bottom of the speaker’s linguistic range, while low troughs at lesser boundaries might be expected to be higher in the speaker’s range.

In order to establish whether the value of low troughs varies with relation to the strength of the syntactic boundary, the f_0 value of each trough was noted and an average value calculated for each of the three main positions observed, (end of sentence, before final adverbials, after sentence initial subject elements). The results are as follows:

	average f_0 (Hz)
end of sentence	82
before adverbials	83.5
initial subject element	85.4
other	91.4

It seems from this that in the data analysed here, the end-of-sentence troughs do indeed have an average f0 value which is lower than at other boundaries. However, all these average f0 values are within one standard deviation of all low f0 troughs. (Average f0 for all low troughs = 84.1 Hz; stdev = 7.1) The difference between them may therefore not be significant. The results nonetheless suggest a tentative ranking of f0 values, whereby the closer the trough is to the end of a sentence, the lower it is likely to be. This would reflect the generally assumed declination of the baseline over a sentence-length utterance.

However, if the troughs after a sentence-initial element (average f0 value 85.4 Hz) are broken down into two groups, (i) those which are not only sentence initial but also topic initial, and (ii) those which are sentence initial but topic internal, then this ranking no longer holds. The average f0 of those troughs occurring at the end of a topic initial element is 78.5 Hz, which is lower even than the end-of-sentence endpoints. The revised ranking of average f0 values in relation to the text is as follows:

	average f0 value (Hz)
topic-initial subject element (right-hand marker)	78.5
end of sentence (left hand marker)	82
sentence-final adverbial clause or phrase (left hand marker)	83.5

Discussion

It was to be expected that low f0 troughs should be associated with the end of a sentence or a co-ordinated main clause in this kind of read-aloud speech. What is more interesting is the association between made f0 troughs and less strong syntactic boundaries, i.e. before final sentence adverbials and after the initial NP subject.

(i) before final adverbials

The average f0 values before a sentence-final adverbial phrase or clause are close to the average value of all low troughs. Although the average is higher than that for the ends of sentences, this averaging procedure obscures the fact that in some sentences the f0 before the final adverbial is in fact lower than that at the end of the same sentence. This has been

referred to as “early closure” (G. Knowles, personal communication) and as a “false alarm followed by an afterthought” (Collier 1993). It could of course be symptomatic of the mental processing of the written text by the speaker, perhaps suggesting that the syntactically superfluous material (i.e. the adverbial) is processed separately, as an afterthought. This can however only be substantiated if other known prosodic signals of finality occur at this point, such as pause, lengthening, vocal creak and lowered amplitude. This remains to be investigated, but my impression is that it is not the case, and that the reader has not simply failed to look far enough ahead. However, assuming that low *f*₀ alone is some indication of finality, whether it reflects planning or lack of planning, its occurrence at a point at which the sentence could be complete is not counter-intuitive.

(ii) After NP subject

The occurrence of ‘finality’ signals close to the beginning of a clause or sentence requires more explanation. It is one thing to state that an NP subject at the beginning of a sentence can be followed by a low trough. The question is: what is the probability, given such a sentence, that the subject actually is marked off in this way? In this text, NP subjects occurred at the beginning of sentences and at the beginning of main and subordinate clauses. Of those which occur at the beginning of a sentence which is also the beginning of a new topic (in this text “topic” = news item) all (6 out of 6) were associated with a low trough. As already described above, these were also on average lower than any other low troughs in the text, including those at the ends of sentences.

Of those which occurred at the beginning of a sentence which was not topic-initial, only 3 out of 11 were associated with a low trough. There were 8 cases in the text where an NP subject stood at the beginning of a reported phrase. Of these, 3 were followed by a low trough. Four NP subjects began a co-ordinate or subordinate clause. Only one of these, that occurring before “and”, was followed by a low trough.

These results suggest that when an NP subject is followed by a low trough which is also low in the speaker’s range, this is a prosodic version of so-called “block language” in newspaper texts, where the first few words of a new paragraph also act as a kind of headline. The prosodic version could therefore be a rhetorical device, or possibly a journalistic mannerism. Of the 6 occurrences here, two have a fall-rise con-

tour, where the trough represents the lowest point of the fall, and therefore occurs not at the end of the phrase but within it, as explained earlier. The other 4 have falling contours, and the f0 trough is the endpoint of the fall. It is this combination - of fall and a low endpoint - which gives the strongest auditory impression of a headline or block language.

It is important of course to consider in the light of the above whether the pattern (a low f0 trough) occurs at the end of any initial sentence element, whether this contains the subject of the sentence or not. There are not enough examples in this text to generalise from. Two sentences begin with an adverbial clause or phrase, one is a cleft sentence, and one begins with a dummy subject “There”. Of these only the cleft sentence has a low trough after it, i.e. before the “that” clause.

Conclusions

It has been shown that the baseline behaviour of this speaker is systematically related to the text and appears to carry important information. Points at which the speaker’s voice reaches the baseline appear to be closely related to phrase and clause structure, i.e. to syntax. More detailed analysis suggests however that this in turn reflects the meaning of the text, in particular its topic structure. This speaker reaches his lowest pitch not at the end, or near the end, of sentences, but instead at the end of the announcement of a new topic. This is also consistent with the results of earlier analyses of a sample of spontaneous speech, referred to above.

The conventional association of low terminal, or baseline reset, with the signalling and perception of finality cannot be denied. However, to equate finality with the ends of sentences is too simple. It has been observed by others that speakers can delay or suppress a low terminal in order to keep a turn in conversation (Laver 1994). It also appears that it can be anticipated in order to highlight the topic structure of a text. This shows speakers’ ability to exploit prosodic signals for strategic reasons, motivated both by the text itself or by the interactive situation.

Implications

The patterns described here could have important consequences for speech technology. Firstly, the intonation in speech synthesis systems

could sound more natural if the realisation of the intonation component paid more attention to baseline behaviour. Secondly, it is possible that patterns in the baseline could provide important syntactic and possibly also discursal information in speech recognition

There are implications too for the annotation of speech corpora. The present study is on a small scale and the results are not necessarily representative of read speech as a whole, but any annotation of the speech signal in speech corpora could usefully include baseline information derived, if possible automatically, from the fundamental frequency values. In this way it would be possible to discover to what extent the results described above are generalisable.

References

- Brazil, D., Coulthard, M., Johns, C. (1980): *Discourse Intonation and Language Teaching*. Essex: Longman.
- Bruce, G., Granström, B., Gustafson, K., House, D. (1991): Prosodic Phrasing in Swedish. In: *Working Papers* 38, Department of Linguistics, Lund University, Sweden.
- Collier, R. (1993): 'On the communicative function of prosody; some experiments'. In: *IPO Annual Progress Report* no.28. Eindhoven: Holland.
- Laver, J. (1994): *Principles of Phonetics*. Cambridge: CUP.
- Menn, L., Boyce, S. (1982): 'Fundamental frequency in discourse structure'. In: *Language and Speech* 26(4), 435-451.
- Shockey, L., Wichmann, A. (1992): 'Discovering units in unscripted monologue' Paper given at the Cambridge conference of the British Association of Academic Phoneticians, April 2.
- Wichmann, A. (1993): 'F0 troughs and prosodic phrasing' Proceedings of ESCA workshop on prosody. In: *Working Papers* 41, Department of Linguistics, Lund University, Sweden.

