

SENTENCE INTONATION IN TEXTUAL CONTEXT -  
SUPPLEMENTARY DATA

NINA GRØNNUM THORSEN

*The experiments presented here follow up a previous investigation (Nina Grønnum Thorsen "Intonation and text in Standard Danish", J. Acoust. Soc. Am. 77, 1985, 1205-1216) and were designed to show whether a sequential fundamental frequency lowering of individual sentence components is present in a semantically, but not syntactically, coherent sequence (a text), when the number of sentences exceeds three. The results show that such a sequential lowering may appear, though it is not evenly distributed across the text. However, the textual intonation contour is sensitive, not only to the number of sentences that make up the text, but also to the length of individual sentence components.*

I. INTRODUCTION

In a previous investigation of intonation in texts in Standard Danish (Thorsen, 1985) it was established that each declarative sentence in a read text is associated with its own declining intonation contour, but together two or three such contours describe an overall slope. Individual sentence contours are steeper in a succession of declarative terminal sentences than in a corresponding string of coordinate main clauses. However, since the texts contained no more than three sentences, the experiment left unresolved the question whether the sequential lowering of individual sentences observed should be ascribed to a particular initial and final effect (text initial and final sentences being higher and lower, respectively, than medial items which under that hypothesis are not differentiated among themselves), or to a truly sequential lowering (evenly) distributed over all the sentences in a text. The experiments reported below were designed to address this issue,

and, furthermore, to see what effect, if any, changing the length of individual sentence components would have.

## II. METHODS

10 declarative sentences were made up; 4 sentences with three prosodic stress groups in each (stressed vowels are indicated orthographically with acute accents here):

1. Káren skal séjle ved Sámsø. (Karen will be sailing  
round Samsø.)
2. Sáanne skal på cåmping i Skótland. (Sanne is going  
camping in Scotland.)
3. Kálle skal våndre i Fránkrig. (Kalle is going to trek  
in France.)
4. Thómas spiller skák i Cånada. (Thomas plays chess in  
Canada.)

4 sentences with two prosodic stress groups each:

5. Káren skal til Sámsø. (Karen is going to Samsø.)
6. Sáanne skal til Skótland. (Sanne is going to Scotland.)
7. Kálle skal til Fránkrig. (Kalle is going to France.)
8. Thómas skal til Cånada. (Thomas is going to Canada.)

2 sentences with four prosodic stress groups:

9. Káren skal séjle med sin fár ved Sámsø. (Karen will be  
sailing with her father round Samsø.)
10. Sáanne skal på cåmping med Kálle i Skótland. (Sanne is  
going camping with Kalle in Scotland.)

The stressed vowels are all low and the initial consonant is either an aspirated stop or an unvoiced fricative, except that in 'Skótland' the stop is unaspirated, and in 'Fránkrig' a devoiced [x̥] intervenes. This will minimize differences in segmental effects on the stressed vowels (from which the lower lines derive in the figures that follow). Segmental effects on fundamental frequency ( $F_0$ ) are smaller in unstressed syllables (from which the upper lines derive), though they are not completely negligible, cf. Reinholt Petersen (1979, 1980).

The following combinations of sentences constituted the texts<sup>1</sup> to be recorded:

1	2	4	3
2	1	3	4
3	4	1	2
4	3	2	1
5	6	7	
5	6	7	8
9	10		

Note that each three-stress sentence occurred in all four positions in the four-sentence texts. Sentences no. 1, 3, 5, 7, 9, and 10 occurred in isolation as well. These 13 items were randomized 6 times to make 6 pages of reading material, each page beginning and ending with two filler sentences. Two subjects from the previous investigation (Thorsen, 1985: NRP (male) and NT (the author)) recorded the material in one session lasting about 20 minutes.

The recordings were made with a Revox A-77 tape recorder and a Sennheiser MD 21 microphone in a quiet room. The tapes were processed with hardware intensity and fundamental frequency meters (F-J Electronics) and registered on an Elema 800 Mingo-graph at a paper speed of 100 mm/s. Measuring accuracy was 1 Hz for the male and 2 Hz for the female speaker, and 0.5 cs in the time dimension.

The first and last local low  $F_0$  point in each sentence (which occurs at or near the offset of the first and last stressed vowel) and the first and last local high  $F_0$  (which occurs in the first post-tonic syllable after the first and last stressed syllable) were measured. The connection of these two highs and lows, respectively, make up the upper and lower lines in the graphs to follow. The distance in time of each of these four measuring points from the onset of the first stressed vowel in each sentence was likewise measured.  $F_0$  and time measurements were averaged over the six repetitions of each item by each speaker. Average  $F_0$  values were converted to semi-tones (re 100 Hz) and average tracings drawn. Furthermore, an average over the four texts with four three-stress sentences was calculated. For a more detailed account of the procedures, see Thorsen (1985).

The use of only two speakers was justified by the high qualitative agreement across the four speakers in Thorsen (1985). The simplified representation of the data in terms of straight upper and lower lines derived from initial and final points in the tracings is justified by the generally valid, straight line approximation to intonation contours in short Danish sentences like the present ones, cf. e.g. Thorsen (1978, 1985).

The qualitative agreement between the two speakers was high, and a common mean could be calculated. Only those means are presented here. Both speakers generally paused between the sentences in a text. Pause durations range between 20 and 80 cs, averaging around 40 cs, with a tendency to increase through the text. The figures do not depict pause durations.

### III. RESULTS

#### A. ISOLATED SENTENCES

Figure 1 presents upper and lower lines of isolated sentences containing two prosodic stress groups (dotted lines), three prosodic stress groups (dashed lines), and four prosodic stress

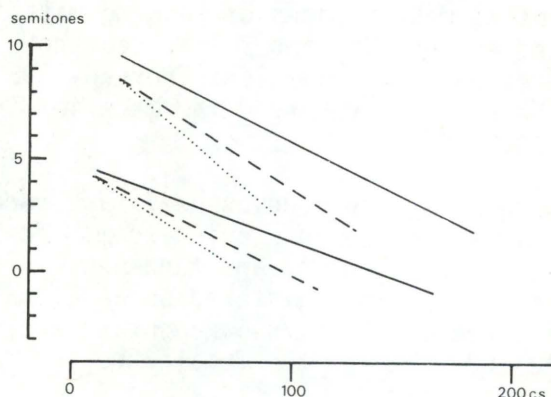


Figure 1

Upper and lower lines in three isolated terminal declarative sentences containing two (dotted lines), three (dashed lines), and four (full lines) prosodic stress groups. Each type represents the mean over four different sentences and two speakers, i.e.  $N=8$ . The onset and offset of upper and lower lines are derived from the first and last local high and low point in the  $F_0$  tracings. Zero on the logarithmic frequency scale corresponds to 100 Hz.

groups (full lines). These utterances only confirm what was found in previous investigations (Thorsen 1980, 1981, summarized in Thorsen 1983), that increasing the length of a terminal declarative sentence will affect beginning points (slightly), end points and slope. Beginning points rise slightly, end points lower, and slope gets less steep as the number of stress groups increase. (Increasing the length beyond four stress groups, however, will lead to a resetting of the intonation contour, cf. Thorsen 1983.) For a further discussion of these findings and references to the literature on other languages, see Thorsen (1980).

## B. MULTI SENTENCE TEXTS

Figures 2, 3, and 4 present upper and lower lines in the various texts.

### 1. Sentences with two prosodic stress groups - Figure 2:

In both texts, with three (dashed lines) and four (full lines) sentences, respectively, a clear overall textual slope can be discerned, which is in accord with the previous results (Thorsen 1985). Furthermore, the two medial components in the four-sentence text seem to be differentiated, i.e., the third sentence is lower than the second one, notably in the upper line.

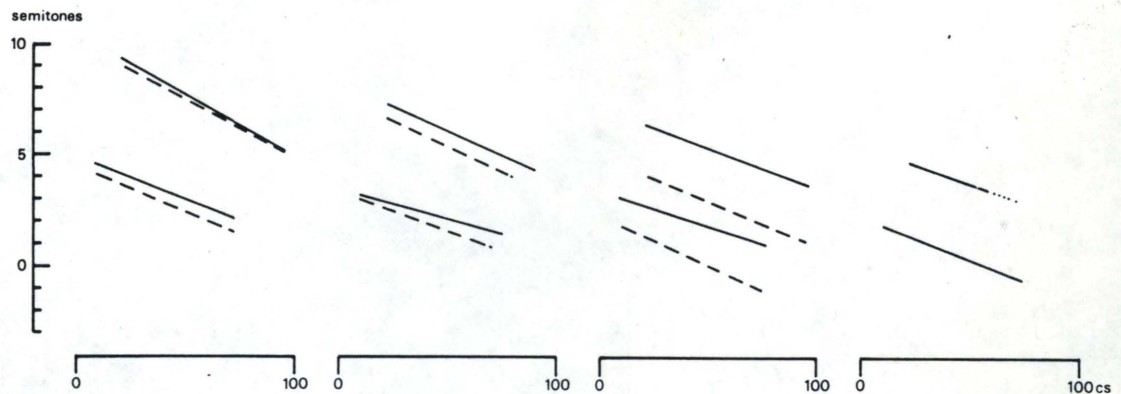


Figure 2

*Upper and lower lines in two texts, containing three two-stress sentences (dashed lines) and four two-stress sentences (full lines). Average over two speakers. The upper line offset in the last sentence of four could not be measured with the male speaker. See further the legend to figure 1.*

Since the two speakers speak in different ranges, standard deviations of their common means are very great, and statistical significance between such means is not to be expected.<sup>2</sup> However, with the female speaker, the difference between upper line onset and offset values, in the second and third sentences, turn out to be statistically significant (student's t-test) at the 0.05 level of confidence, whereas it does not reach statistical significance with the male speaker. - Simultaneously we should note, though, that the two medial sentences seem to cluster, i.e. the lowering of sentences is not evenly distributed across the whole text.

2. Sentences with 3 prosodic stress groups - Figure 3, full lines:

Here, too, an overall downward slope is discerned, but the slope levels out over the two medial components, which appear to be almost exactly identical as far as upper and lower line onsets and offsets are concerned, in contradistinction to the two-stress sentences (dashed lines).

3. Sentences with 4 prosodic stress groups - Figure 4 - will be dealt with in the discussion section below.

#### IV. DISCUSSION

Figure 5 summarizes Figures 1-4 and facilitates a comparison of the various texts. The isolated sentences, which are in a sense both text initial and text final, combine the high (upper and lower line) onsets of initial sentences with the low (upper and lower line) offsets of final sentences in multi-sentence

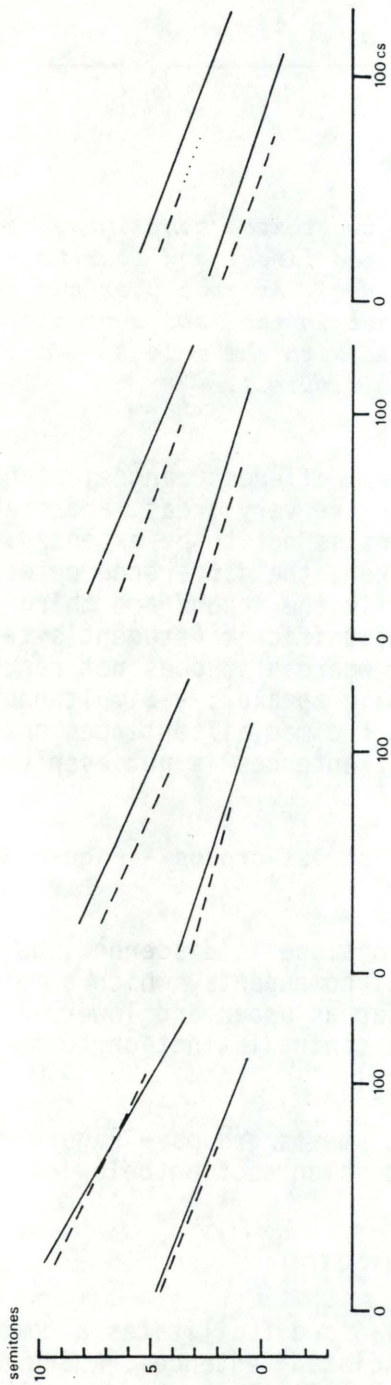


Figure 3

Upper and lower lines in a text with four three-stress sentences (full lines), averaged over four texts and two speakers, i.e.  $N=8$ . The dashed lines are a repetition of the full line tracings in figure 2. See further the legend to figure 1.

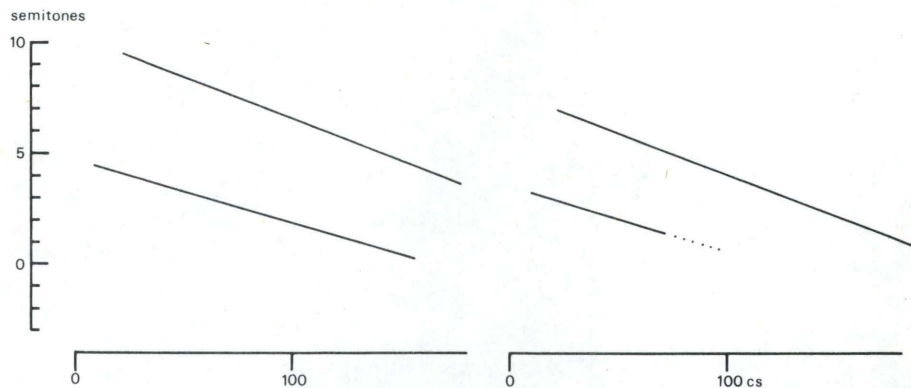


Figure 4

*Upper and lower lines in a text containing two four-stress sentences. Average over two speakers. See further the legend to Figure 1.*

texts; except that isolated sentences with two stress groups have offset values that do not quite reach the lower limit set by text final sentences. This is in complete agreement with the previous results (Thorsen 1985).

The most salient features of multi-sentence texts are: Onset values (whether of upper or lower lines) are fairly constant across texts of different length, both in terms of number of sentences and number of stress groups in each sentence. These onset values vary within a range of one semitone in the upper lines, slightly less in the lower lines. Considering the range of sentence onset values across a text, which amounts to about 5 semitones in the upper lines and 2.5 semitones in the lower lines, it appears justified to talk about a near-constancy in  $F_0$  onset across texts of different length. Likewise, offset values in text final sentences appear nearly constant. In other words, the total range spanned by any given text is independent of its length, at least within the range of text lengths investigated in this material (and thus the overall slope is inversely proportional to text length). But within this upper and lower limit, set by the first and last prosodic stress group in a text, sentences do distribute themselves differently according to their number and length. A sequence of three sentences (with two stress groups in each sentence) exhibit a smooth and linear fall in onset as well as offset values of both upper and lower lines. This compares well with the previous findings (Thorsen 1985). A text with two sentences with four stress groups in each has its second - and final - component beginning at a considerably higher onset value than any other text final component in this material, presumably in order to preserve a sentence intonation slope across its four stress groups which is steep enough to suit a text final terminal declarative sentence (without falling through the floor of the speaker's  $F_0$  range).

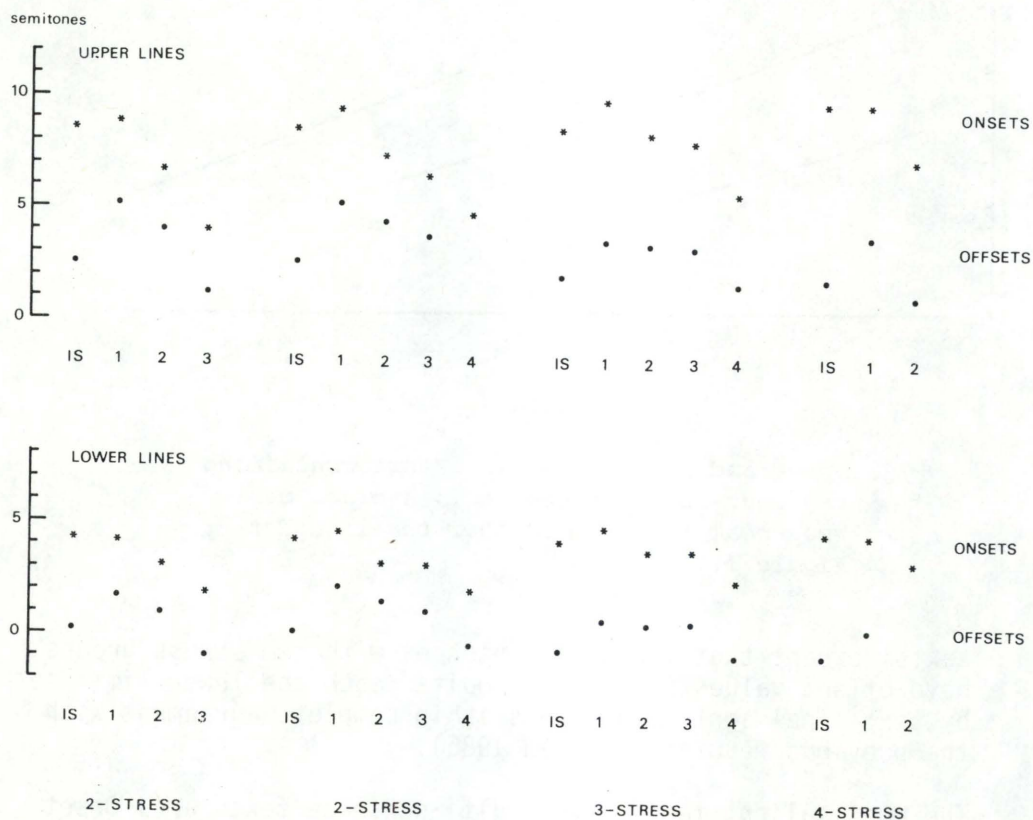


Figure 5

Frequency values of onsets (stars) and offsets (dots) of the upper lines (above) and lower lines (below) in Figures 1-4. The two leftmost sets of data pertain to two-stress sentences in isolation ("IS", derived from Figure 1) and in each of the successive positions in a text with three and four components, respectively (derived from Figure 2), as indicated below each graph. Similarly, the two rightmost sets of data pertain to sentences with three and four stress groups, in isolation (derived from Figure 1) and in a text with four sentences (derived from Figure 3) and with two sentences (derived from Figure 4), respectively.



Another instance of plasticity in the production of sentence and text intonation is offered by the somewhat different treatment of texts with four sentences, where the sentences have two and three prosodic stress groups, respectively. Onset values are higher, and offset values are lower (in upper as well as lower lines) in non-final sentences with three stress groups. This may, again, be seen as a consequence of sentence length, simply: the longer sentences will have higher onsets and lower offsets to preserve a sentence intonation slope suitable for terminal declarative sentences - just as is the case for isolated sentences of varying length, cf. Figure 1. The demand for a certain sentence intonation slope may also explain why a succession of two medial three-stress sentences are produced without any "downstep", in contradistinction to a succession of medial two-stress sentences, cf. Figure 3. If the longer sentences are to preserve a suitable slope, there is no room for a sequential lowering of medial components.

In Figure 2 there is a difference (notably in the upper lines) in the second component in the two texts: it is lower when only one sentence succeeds (dashed line) than when two sentences follow (full line). This difference persists, though diminished, if the two texts are aligned on the frequency scale to have identical upper line onsets. However, it turns out that with one of the speakers there is virtually no distinction between upper (or lower) lines in these sentences, and with the other subject the difference does not quite reach statistical significance. Thus, the evidence is not conclusive but seen together with the remainder of the present data and that published in Thorsen (1983, 1985), it does lend some support to the notion of look-ahead and preplanning in the production of intonation. For a further discussion of such mechanisms and their implications for intonation theory, see Thorsen (1985).

## V. NOTES

1. I employ "text" here to mean a sequence of semantically coherent but not necessarily syntactically coordinated sentences.
2. Standard deviations could have been considerably reduced if the data had been normalized, e.g. by a zero-line offset adjustment. However, there are theoretical and practical problems in such a normalization that still await a solution and since, furthermore, it is not really essential to the points I want to make, no normalization has been attempted.

## REFERENCES

- Reinholt Petersen, N. 1979: "Variation in inherent  $F_0$  level differences between vowels as a function of position in the utterance and in the stress group", *Ann. Rep. Inst. Phon., Univ. Copenhagen* 13, p. 27-57
- Reinholt Petersen, N. 1980: "Coarticulation of inherent fundamental frequency levels between syllables", *Ann. Rep. Inst. Phon. Univ. Copenhagen* 14, p. 317-354
- Thorsen, N. 1978: "An acoustical analysis of Danish intonation", *J. Phonetics* 6, p. 151-175
- Thorsen, N. 1980: "Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish", *Ann. Rep. Inst. Phon. Univ. Copenhagen* 14, p. 1-29
- Thorsen, N. 1981: "Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish - Supplementary data", *Ann. Rep. Inst. Phon. Univ. Copenhagen* 15, p. 13-47
- Thorsen, N. 1983: "Standard Danish sentence intonation - Phonetic data and their representation", *Folia Linguistica* 17, p. 187-220
- Thorsen, N. Grønnum 1985: "Intonation and text in Standard Danish", *J. Acoust. Soc. Am.* 77, p. 1205-1216.