(ARIPUC 15, 1981 p. 13-47)

INTONATION CONTOURS AND STRESS GROUP PATTERNS IN DECLARATIVE SENTENCES OF VARYING LENGTH IN ASC DANISH - SUPPLEMENTARY DATA

NINA THORSEN

Intonation contours in non-compound declarative sentences containing from one to eight stress groups were analysed with special regard to the manifestation of prosodic boundaries. The results are compared with a previous analysis of a very similar material. The neat and easy match between the prosodic and some of the major syntactic boundaries exhibited by the earlier corpus is not immediately replicated by the present utterances. Resetting of the intonation contour taken rigidly to manifest prosodic phrase group boundaries leads to counter-intuitive phrase contours, but on the other hand, the alternative or supplementary boundary criterion suggested in the paper cannot be applied automatically and unambiguously. This strengthens the claim that in syntactically unambiguous non-compound sentences, prosodic structure does not directly reflect or map syntactic structure. However, the four speakers' productions show a high degree of intra- as well as inter subject coherence, and the variation found in the intonation contours with varying sentence length is not random. Nevertheless, the principles governing the intonational properties, especially of the longer utterances, are not easily recovered, and it seems - maybe not surprisingly - that the semantic content of the (constituents of an) utterance must also be considered. The need for further research, also into the perception of prosodic structure, is evident.

I. INTRODUCTION

In the previous volume of ARIPUC (Thorsen 1980a) were published the results of an analysis of eight non-compound declarative sentences, containing from one to eight stress groups, recorded six times by each of four speakers. This is a report of an analysis of an only slightly dissimilar material, recorded by the same speakers, under similar conditions to the 1980 material, and the reader is referred to Thorsen (1980a) for an introduction to the subject with references to the relevant literature, as well as for accounts of recording procedures and measurements. However, I do wish to stress once more the fact that the two materials (forthwith identified as the 1980- and 1981 material) were not originally intended to investigate the relation between prosodic and syntactic boundaries but to test a hypothesis about the equidistant distribution on the frequency scale of the stressed vowels in declarative sentences. It was the refutation of this hypothesis which led to considerations and questions about the syntactic/prosodic interplay. I also wish to point out that temporal relations have not been investigated. Finally: everything that is said in the following pertains exclusively to syntactically unambiguous non-compound sentences.

II. MATERIAL

The material consists of eight non-compound statements, all variations on the same theme:

1. *Til / Tíflis.*
2. *Túkke $ skal til / Tíflis.*
3. *Búster $ skal med / bússen $ til / Tíflis.*
4. *Kísser $ skal med / bússen $ i / nát $ fra / Tíflis.*
5. *Lissi $ skal med / bússen $ klokken / ét = i / nát $ fra / Tíflis.*
6. *Pýtte $ skal med / bússen $ til / Thísted $ klokken / ét = i / nát $ fra / Tíflis.*
7. *Hútters $ skal med / bússen $ fra / kírken = i / Thísted $ klokken / ét = i / nát $ til / Tíflis.*
8. *Knúdsen $ skal med / bússen $ fra / pládsen = ved / kírken = i / Thísted $ klokken / ét = i / nát $ til / Tíflis.*

Sentence no. 8 translates: 'Knudsen is taking a bus from the square by the church in Thisted at one o'clock tonight for Tiflis.' The stressed vowels are indicated with an acute accent. "$" denotes the boundary between major syntactic

constituents (noun phrase, verb phrase, complements of place and time); "=" indicates complement internal boundaries; "/" denotes prosodic stress group boundaries (see further page 21). Note that syntactic and prosodic stress group boundaries never coincide. The syntactic boundaries occur after the first post-tonic syllable in the stress group, except that it occurs directly after *ét* and *nât*.

The stressed vowels are all short, but the sequence /ir/ in *kîrken* is realized as a diphthong [iɐ̯], which was measured at the midpoint in time of the voiced stretch; they are all high except [ɐ̯] in *ét* and [æ] in *nât*; they are surrounded by unvoiced obstruents except [n] in *nât*, [l] in *Lîssi*, and the stressed vowel in *klokken ét* ([-ŋ ˈeɐ̯]); [l] in *plâdsen* is devoiced by the preceding aspirated stop.

Sentence no. 1 occurred in a small dialogue, - all the others were naked but were mixed with a material recorded for a different purpose, being evenly distributed over two full pages of recording material, which appeared in three different randomizations, each being read twice on two separate occasions (only once by JR), giving a total of six (three) readings by each subject.

III. RESULTS

Stylized tracings of the eight sentences are depicted in figure 1-4, the grand mean in figure 5. (The last post-tonic vowel could not be measured in JR's sentence 7 and is therefore also lacking in the mean.) As with the 1980 material, the calculation of a mean over all subjects (mean of means) is justified by the rather good qualitative similarity between them. The stressed syllables are connected with full and broken lines, denoting the intonation contour proper. (Note that these lines are "imaginary" - they are not directly present in the course of fundamental frequency (Fo).) Broken lines occur wherever the connection between two stressed vowels is less steep than the preceding as well as succeeding ones; two slopes are considered different if they differ by an (arbitrary) amount of 0.5 semitones/second. (This criterion of slope identity or difference was adopted in order to obliterate slope differences that may occur merely as a consequence of slight differences in the timing of the stressed vowels. 0.5 semitones/second may be too narrow a step, but lacking data on difference limens for the perception of slopes such as these a rather strict criterion is preferable.)

The stressed vowels in the tracings have been corrected for differences in inherent Fo levels, see further Thorsen (1980a p. 6). No correction has been attempted for the unstressed vowels, partly because the difference in inherent Fo level in unstressed vowels and sonorant consonants is smaller than in stressed vowels (Reinholt Petersen 1978), and partly also be-

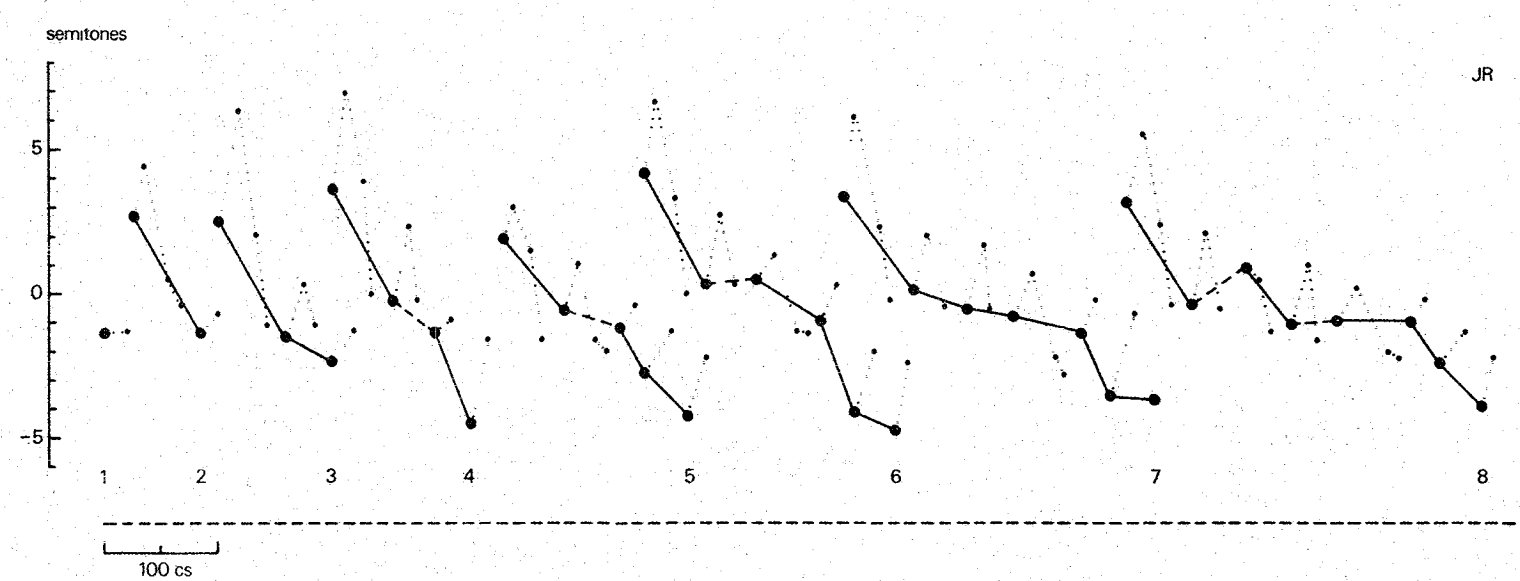

Figure 1

Intonation contours (full and broken lines) and stress group patterns (dotted lines) in declarative sentences containing from one to eight stress groups. Large dots represent stressed syllables, small dots unstressed syllables. Zero on the logarithmic frequency scale corresponds to 100 Hz. Subject: JR.

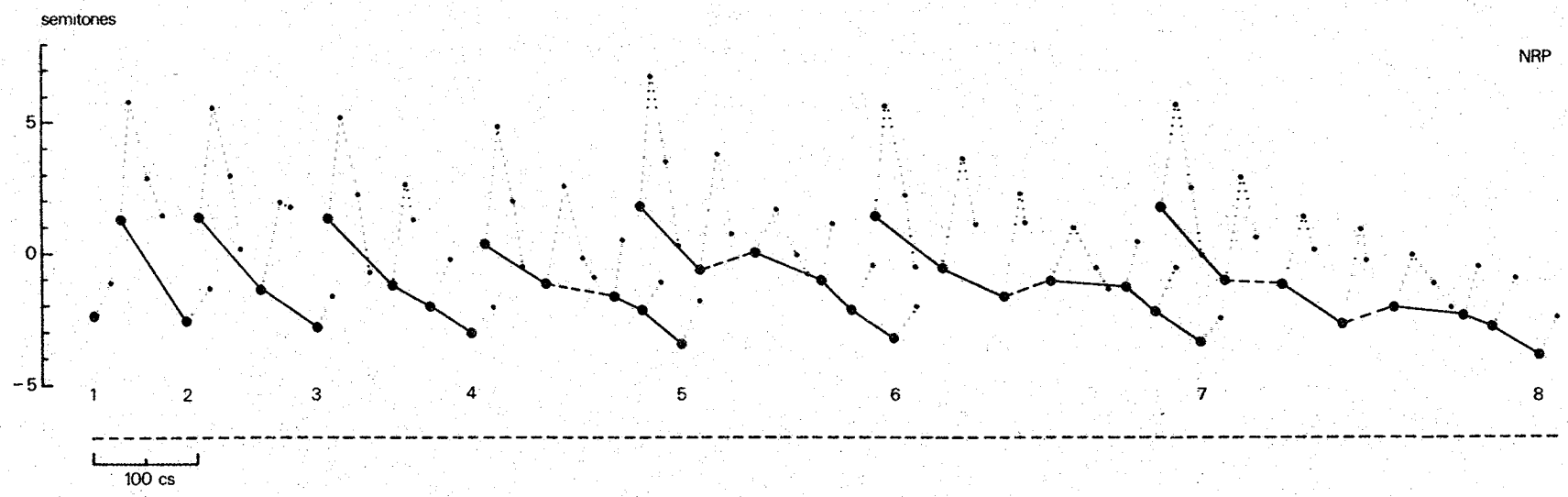

Figure 2

Subject: NRP. See further the legend to figure 1.

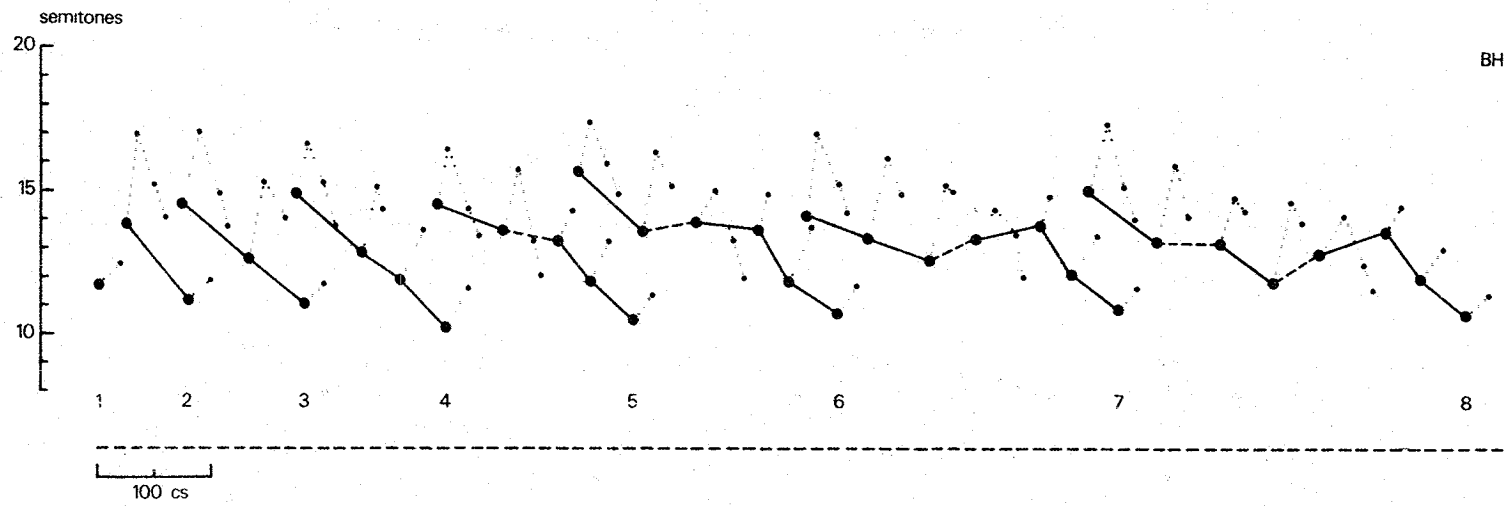

Figure 3

Subject: BH. See further the legend to figure 1.

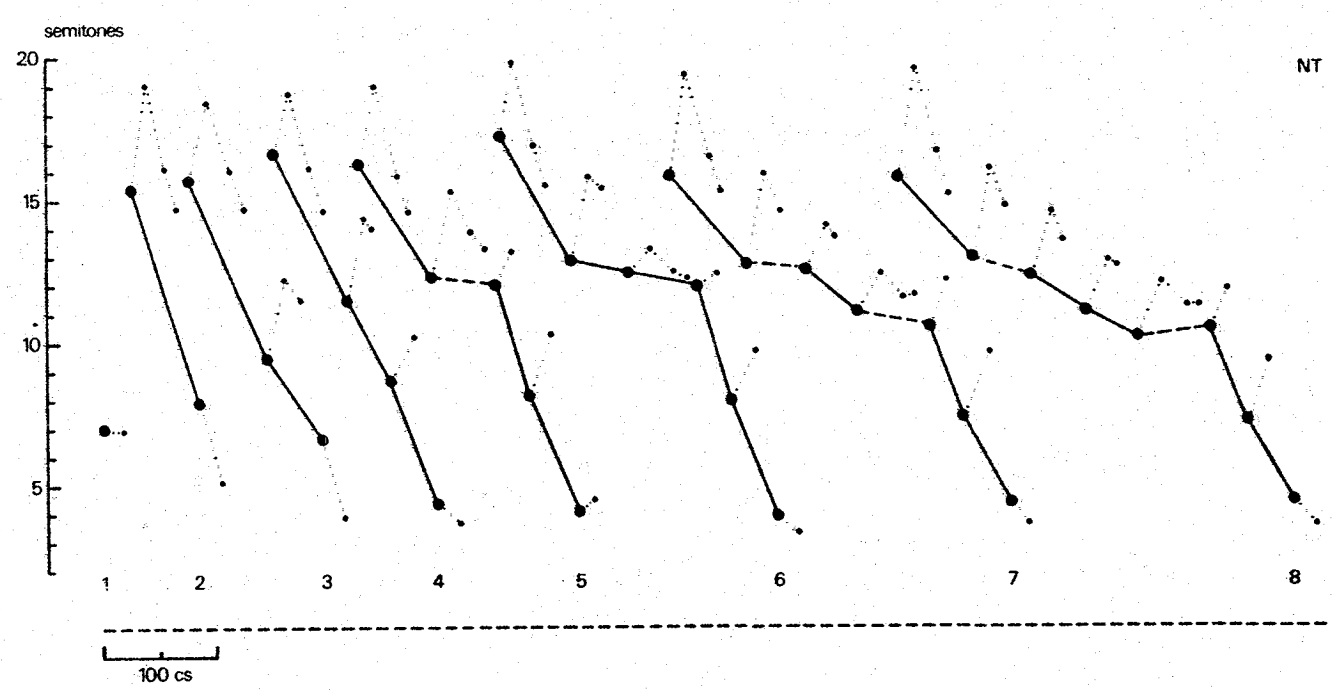

Figure 4

Subject: NT. See further the legend to figure 1.

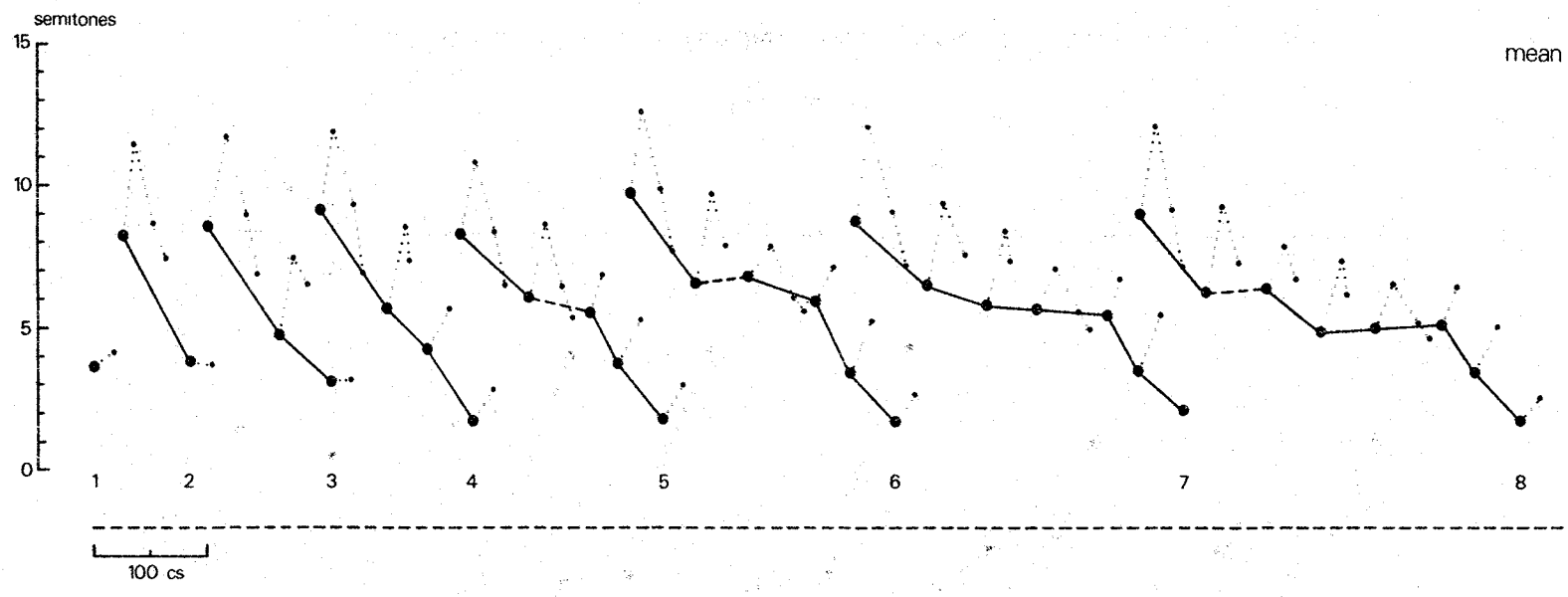

Figure 5

Average over four subjects (mean of means). See further the legend to figure 1.

cause reliable quantifications of these differences are lacking. However, Reinholt Petersen (1980) found that inherent Fo level differences in a stressed vowel are to some extent carried over into the first post-tonic vowel, i.e. a post-tonic vowel will have a higher fundamental frequency after a stressed vowel with a high than after a stressed vowel with a low tongue position, everything else being equal. It might therefore have been appropriate also to raise those post-tonic vowels that occur after non-high stressed vowels, but again the exact magnitude of these corrections is not well established and no such compensation has been attempted.

> It is worth noting, as with the 1980 material, that standard deviations on Fo and time measurements are small: in sentence no. 8, for instance, they range (with individual subjects) for the stressed vowels between 4.9% and 1.1% of the mean and for the unstressed vowels between 5.2% and 0.5% of the mean in the Fo measurements. The standard deviations on the total duration of sentence no. 8 range between 3.2% and 1.7% of the mean. The figures must therefore be fairly reliable indications of the subjects' behaviour. - Furthermore, there is no tendency for the first recording(s) of each item to be deviant from the following ones (which is not necessarily excluded by the small standard deviations, if only later recordings are sufficiently undispersed) - i.e. there is no apparent sign of subjects having to go through a learning procedure which then made for more or less automated readings of later repetitions of a given item. I note this expressly because sentences no. 7 and 8 are very long indeed and might have caused this kind of behaviour. The small standard deviations, i.e. the good "production stability" across a subject's six (three) renderings of each utterance, are interesting also in the light of the difficulty I have in recovering the rules that the subjects must have employed for their productions, see further below.

The results of range variation and of variation in starting and end points resemble the 1980 material to a point where they are hardly worth accounting for. The appropriate figures and tables and accompanying comments are given in an appendix.

A. STRESS GROUP PATTERNS

A prosodic stress group has been defined previously as a stressed syllable plus all succeeding unstressed syllables, irrespective of intervening syntactic boundaries within the same intonation contour. This definition still holds. The stress group patterns in figure 1-5 look the same, basically, and no trace of either word boundaries or stronger syntactic

Table 1

Least squares regression line slopes of the stressed vowels and of the first post-tonic vowels in each stress group depicted in figure 1-5 and their correlation coefficients.

		Sentence no.	JR Slope/corr.	NRP Slope/corr.	BH Slope/corr.	NT Slope/corr.	mean Slope/corr.
stressed	vowels	2	-7.1	-6.0	-4.8	-12.3	-7.5
		3	-4.9/-.97	-3.8/-.99	-3.3/-1.0	-7.8/-1.0	-5.1/-.99
		4	-6.4/-.99	-3.1/-.98	-3.5/-1.0	-8.4/-1.0	-5.4/-1.0
		5	-3.5/-.98	-1.8/-.97	-2.2/-.95	-5.8/-.95	-3.4/-.97
		6	-3.8/-.96	-1.8/-.93	-1.8/-.91	-4.6/-.93	-3.0/-.94
		7	-2.3/-.95	-1.2/-.89	-0.8/-.74	-3.3/-.94	-1.9/-.94
		8	-1.7/-.90	-1.2/-.89	-0.9/-.75	-2.7/-.94	-1.6/-.92

(table 1 continued)

unstressed	vowels	2	-7.8	-10.0	-7.7	-20.3	-11.3
		3	-7.6/-.97	-6.2/-1.0	-4.8/-.97	-11.9/-.99	-7.7/-1.0
		4	-7.0/-.98	-5.1/-.99	-3.5/-.98	-9.7/-.98	-6.4/-1.0
		5	-3.0/-1.0	-3.6/-1.0	-2.8/-.97	-6.9/-.96	-4.1/-.99
		6	-3.8/-.98	-3.2/-.99	-2.2/-.96	-5.6/-.95	-3.7/-.98
		7	-2.5/-.94*	-2.3/-.99	-1.7/-.94	-4.3/-.94	-2.3/-.98*
		8	-2.0/-.93	-1.9/-.97	-1.4/-.94	-3.6/-.94	-2.2/-.96

* The last post-tonic could not be measured with JR and is omitted from the calculations.

boundaries (which occur after the first post-tonic syllable in most instances) can be detected, i.e. the prosodic stress group cuts across the syntactic structure (still: in syntactically unambiguous non-compound sentences).

In the 1980 material a trend was found towards higher rises to the post-tonic vowel in a stress group immediately preceding a discontinuity (a resetting) in the intonation contour. A similar trend is not present here, and I am inclined to believe that I over interpreted the 1980 material on this point. If a consistent trend had been found to higher rises to the post-tonics (or any other systematically deviant stress group pattern) in certain places, then that of course might have been added to the list of criteria for prosodic boundary assignment, cf. below.

B. INTONATION CONTOURS

1. OVERALL DOWNDRIFT

a. Degree of downdrift In table 1 are given least squares regression line slopes for the stressed vowel data points and the first post-tonic vowel data points, respectively, in sentences 2 through 8, cf. figure 1-5. These regression lines may be taken as an expression of the degree of overall downdrift. For ease of reference I shall forthwith call the lines connecting the stressed vowel data points (i.e. the intonation contour proper) and the lines that would connect the first post-tonic vowel in each stress group "baselines" and "toplines", respectively, but note that this is not the way these terms are ordinarily understood, see further Thorsen (1980a p 2-3). As with the 1980 material the "topline" is steeper than the "baseline" slope, and "baseline" and "topline" slopes are highly correlated across the seven utterances (r = 0.93 (JR), 1.00 (NRP), 0.94 (BH), 0.97 (NT)). The Pearson product moment correlation coefficients come out with rather high values, i.e. straight lines are fairly good approximations to the data. Exceptions are the "baselines" of BH's sentence 7 and 8, with correlation coefficients of -.74 and -.75, respectively, but note that the jagged "baseline" is somewhat smoothed out in the "topline" which would be rather better fitted to a straight line (cf. the coefficients of -.94). A similar tendency (though slight - due to the generally high "baseline" coefficients) can be observed with the other subjects, but I do not think the smoother "toplines" warrant considerations about their being in some sense primary and "baselines" secondary, for a number of reasons:

The stressed vowel and the first post-tonic in each stress group are highly correlated across a given sentence, cf. table 2. The correlations have been calculated both from data where no compensation for differences in intrinsic Fo level between stressed vowels of different tongue height is performed as well as from the "compensated" data (as they appear

Table 2

Pearson product moment correlations between the stressed vowel and first post-tonic vowel Fo's in six sentences, containing from three to eight stress groups, i.e. three to eight data pairs. The correlations have been calculated on data where no compensation for differences in intrinsic Fo level between stressed vowels of different tongue height is performed ("uncomp.") as well as on data where such a compensation is carried through ("comp."). The second decimal is rounded off to the nearest integer. In JR's sentence no. 7 the correlation is calculated on the first six vowel pairs only, since the last post-tonic could not be measured.

	JR		NRP		BH		NT	
sentence no.	uncomp.	comp.	uncomp.	comp.	uncomp.	comp.	uncomp.	comp.
3	1.00	1.00	0.99	0.98	0.99	0.97	0.97	0.96
4	0.99	0.95	0.93	0.96	0.91	0.99	0.97	0.99
5	0.97	0.99	0.98	0.96	0.92	0.99	0.98	0.99
6	0.97	0.98	0.92	0.94	0.91	0.95	0.97	0.99
7	0.92	0.96	0.92	0.93	0.88	0.89	0.96	0.98
8	0.97	0.91	0.93	0.96	0.88	0.90	0.95	0.99

in the figures): Stressed and post-tonic vowels do not lead separate lives, but I would argue, as on a number of previous occasions, that the post-tonic vowels are the dependent variables in this relation. Firstly, a stress group obviously has to have a stressed vowel in it, but not necessarily an unstressed one, secondly the stressed vowels in an utterance seem to be stronger perceptual cues to the identification of sentence intonation than the post-tonics (see Thorsen 1980b). Finally, one might argue that if prosodic boundaries are indeed signalled via a non-smooth declination, then the slightly more irregular "baseline" is a better carrier of such information than the "topline", and I shall thus consider "baselines" only in the section on prosodic phrase group boundaries. These arguments do not necessarily deprive the "topline" of any perceptual relevance at all - on the contrary, the smooth "topline" may be seen both as a carrier of information about the degree of overall downdrift, and as a reference which sets off the more irregular "baseline".

A tendency appears in table 2 for the correlations on "compensated" data to be (albeit only slightly) higher than on uncompensated data. Given that the correlations are generally very high (excepting BH's sentence 7 and 8) and should be so if post-tonic vowels are predictable from the stressed ones, then the even better correlations that we obtain from data where intrinsic Fo level differences are compensated for is a point in favour of just this procedure. And one might speculate that if a partial compensation had also been performed in the first post-tonic vowel after non-high stressed vowels (cf. above) then correlation coefficients would have been still closer to unity.

That slope variation is not a linear function of utterance length is seen also in figure 6 and 7. Rather, the slope of the overall downdrift decreases asymptotically with utterance length (whether defined in terms of number of stress groups - figure 6 - or in terms of actual duration - figure 7) and reaches a mean saturation value of about -1.5 semitones/second (stressed vowels) and -2.0 semitones/second (post-tonic vowels), respectively. (In the 1980 material the corresponding values were -2.0 and -2.5 semitones/second.) For a brief discussion of this point, see Thorsen (1980a p. 20-21).

b. Shape of the downdrift In the 1980 material a tendency was found (by visual inspection of the tracings) with some of the subjects in some of the utterances towards greater "baseline" declination in the early part of the utterance, i.e. a tendency towards an asymptotic declination throughout the longer sentences. A similar tendency is not manifest in the present material. On the contrary, with two subjects, BH and NT, the final part of the intonation contour of the longer sentences is decidedly more steeply falling than the beginning. With JR

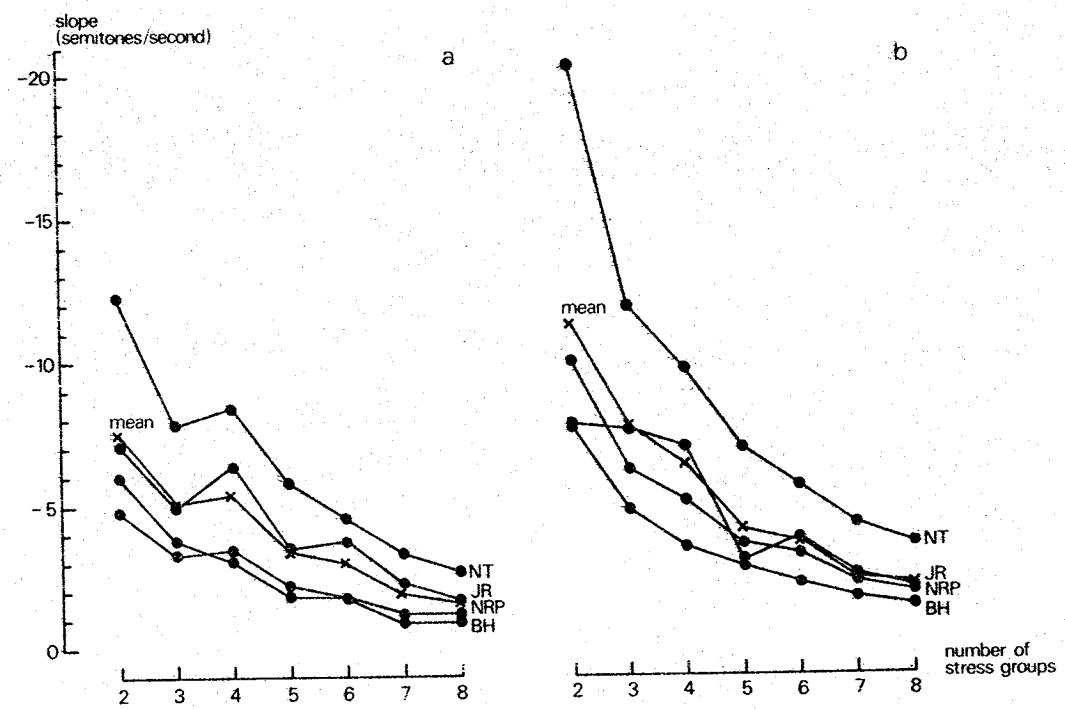

Figure 6

Slope of the overall downdrift of seven utterances, containing from two to eight stress groups, depicted as a function of the length of the utterance (in terms of number of stress groups). Four subjects and their grand mean (crosses). In (a) is depicted the downdrift as determined by the stressed vowels; in (b) downdrift is determined by the first post-tonic vowels. In (b), JR, sentence no. 7, the slope is calculated from the first six vowels only.

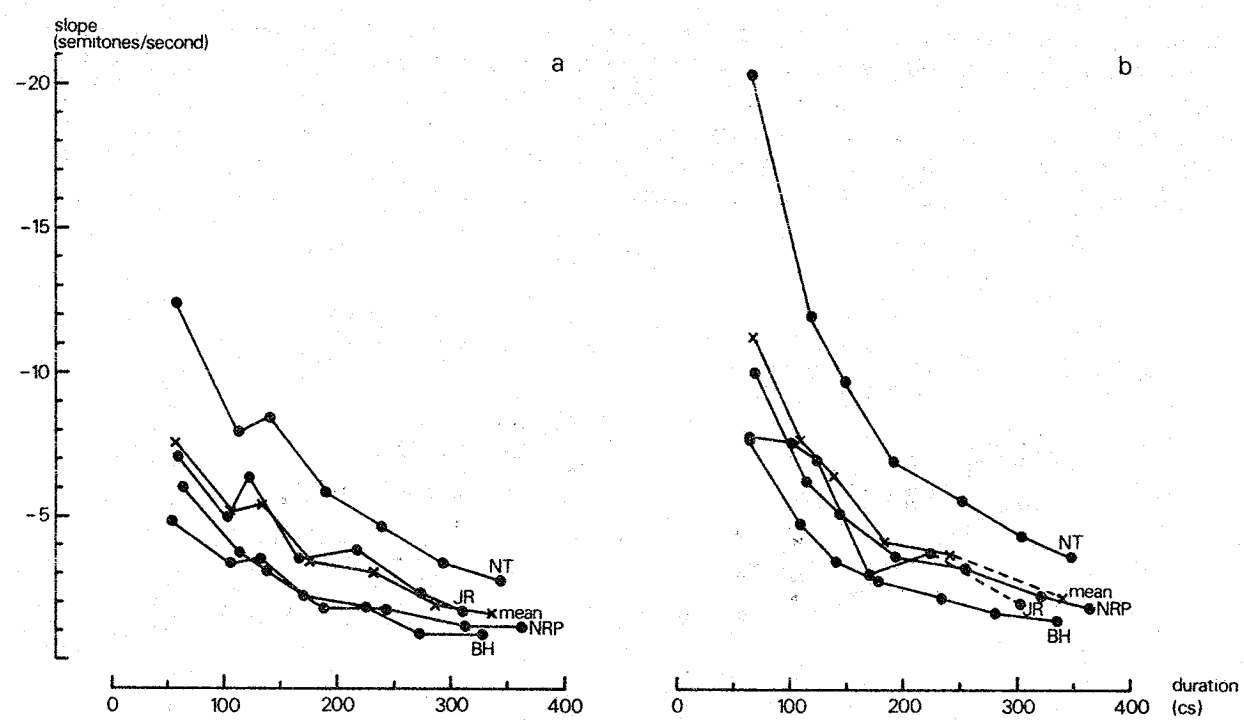

Figure 7

Slope of the overall downdrift of seven utterances depicted as a function of the duration of the utterance. In (b) data on sentence no. 7 is lacking with JR (and the mean). See further the legend to figure 6.

we find a steep beginning, a levelling out, and a steep fall at the end. The subparts of NRP's contours have approximately identical slopes. - The difference in the two materials should probably be ascribed to their different syntactic (and semantic) and prosodic make-up, and I think that this constitutes a good illustration of the precaution needed in the interpretation of the results of isolated analyses.

2. PROSODIC PHRASE GROUP BOUNDARIES

In Thorsen (1980a) a discontinuity in the intonation contour was said to occur at those places in the contour where the slope of the line connecting two stressed vowels is less steep than the preceding as well as succeeding ones (i.e. where a partial "resetting" of the contour takes place), and such discontinuities were taken to be manifestations of boundaries between prosodic phrase groups. The definition seemed a reasonable one, given the shape of the intonation contours and the neat and easy match between the prosodic and some of the major syntactic boundaries thus established.

Although the 1980 and 1981 materials resemble each other rather closely, they do differ in the number and distribution of major syntactic boundaries. Furthermore, the 1980 material had complements of purpose and place, the 1981 material has a complement of time and two complements of place in the longer sentences. The difference between the materials is greatest in sentence 6, 7 and 8 - which are also most evidently different in their intonation contours. The three pairs of sentences are listed here, indicating major syntactic boundaries with "$":

6: (1981) *Pýtte $ skal med bússen $ til Thísted $ klokken ét i nát $ fra Tíflis.*

(1980) *Aníta $ skal med bússen $ til fésten for Kísser $ på "Kílden" i Thísted.*

7: (1981) *Hútters $ skal med bússen $ fra kírken i Thísted $ klokken ét i nát $ til Tíflis.*

(1980) *Hútters $ skal med bússen $ til fésten for Kísser og Lissi $ på "Kílden" i Thísted.*

8: (1981) *Knúdsen $ skal med bússen $ fra plládsen ved kírken i Thísted $ klokken ét i nát $ til Tíflis.*

(1980) *Knúdsen og Bítten $ skal med bússen $ til fésten for Kísser og Lissi $ på "Kílden" i Thísted.*

(Sentence no. 8-1980 translates: 'Knudsen and Bitten are taking a bus to the party for Kisser and Lissi at "Kilden" in Thisted.'.) The major syntactic constituents contain one, two or three stressed syllables (with pertaining unstressed ones) as follows - the actual phrase grouping performed by subjects in the 1980 material is indicated in parentheses:

6:	(1981)	1 + 1 + 1 + 2 + 1	
	(1980)	1 + 1 + 2 + 2	(2+2+2: JR; 2+4: NRP, BH; 6 or 4+2: NT)
7:	(1981)	1 + 1 + 2 + 2 + 1	
	(1980)	1 + 1 + 3 + 2	(2+3+2: all subjects)
8:	(1981)	1 + 1 + 3 + 2 + 1	
	(1980)	2 + 1 + 3 + 2	(3+3+2: all subjects)

This is not precisely the way the contours appeared in Thorsen (1980a figures 1-5) due to the less tolerant criterion of slope identity adopted there, but it corresponds to the tracings in the lower part of figures 8-11 here. With a slightly wider step than 0.5 semitones/second the second discontinuity in JR's sentence 6-1980 would disappear and his grouping would be 2+4 as well.

On the basis of the grouping performed by subjects in 1980 I would venture the following hypotheses, although the risk of making too-far-reaching conclusions is considerable: (a) Four stressed syllables in one prosodic phrase group seems to be the maximum (cf. also the fact that sentence no. 5 was indeed divided prosodically into two groups). NT's exceptionally large range (about 15 semitones) will accommodate even larger prosodic phrases, but it is also possible to postulate a 4+2 grouping in no. 6-1980. (b) A syntactic constituent having only one stressed syllable in it ties up prosodically with a neighbouring constituent. (This point is one that may be disproven by a different material. If the neighbouring constituent(s) is (are) already maximally long (with four stressed syllables in it (them)) then a prosodic phrase with only one stressed syllable may be envisaged, unless a neighbouring constituent is to be cut up internally by a prosodic boundary. However, I do not think that a final syntactic constituent with only one stressed syllable in it will appear as an independent prosodic phrase and thus be preceded by a discontinuity, because such a final "rise" probably would violate the inherent feature of terminal declarative sentence intonation.) With these restrictions the groupings in sentence 6-1980 are all predictable, and the grouping in sentence 7-1980 is the only one possible; in sentence 8-1980 a 2+4+2 grouping is also possible, unless we add a further constraint that (c) prosodic phrase groups be of as nearly as possible equal size - then 3+3+2 is a better candidate than 2+4+2. - The 3+3+2 grouping in sentence 8-1980 might also be due to (d) a tendency for the boundary before the complement (at the second "+") to be stronger than the NP+VP boundary (at the first "+") - and this may be true of this particular utterance but is hardly a general phenomenon (cf. the discussion about the role that semantic content may have, in section 4. below).

If these restrictions on prosodic phrase grouping are carried over to the 1981 sentences, we get the following possibilities:

6-1981: 2 + 4 (hypothesis (d)) or 3 + 3 (hypothesis (c))

7-1981: 2 + 2 + 3 (hypothesis (d)) or 4 + 3

8-1981: 2 + 3 + 3 (the only one possible)

All of the above reasoning naturally rests upon the assumption that a major syntactic constituent will not be cut up internally (at least if it has no more than four stressed syllables in it), or in other words: a prosodic boundary will not occur independently of a major syntactic one (whereas a major syntactic boundary need not be accompanied by a prosodic one).

In figures 8-12 the two sets of intonation contours are shown (omitting the unstressed syllables) with syntactic boundaries indicated, and in the leftmost ("a") edition (if there is more than one) with broken lines at the discontinuities as defined above. Before proceeding any further, I wish to point out two things: Firstly, the "irregularity" of the contours cannot be due to faulty correction for intrinsic Fo level differences between vowels (cf. also A.1.b above). In figure 12 I have indicated in square brackets the lowest vowels in the sentences, and if one looks across the subjects at these particular stretches and compare them with the rest of the contours, it is clear that the raised non-high vowels cannot be made responsible for the breaks and turns in the contours. Secondly, it is also evident that a subject's Fo range will influence the demand for and degree of "resetting" of the contour, compare BH (figure 10) to NT (figure 11), for instance.

The prosodic boundary definition and assignment which worked so well with the 1980 sentences is less satisfactory with the present material, compare the actual "a" groupings to the predictions:

	predicted	JR	NRP	BH	NT
6-1981:	2+4 or 3+3	2+4	2+4	2+4	6
7-1981:	2+2+3 or 4+3	7	3+4	3+4	2+2+3
8-1981:	2+3+3	2+2+4	2+2+4	2+2+4	2+3+3

In sentence 6, three subjects fit one of the predictions, in sentence 7 and 8 this goes for only one subject. That could of course just be a hint that the assumptions on which the predictions are based are false - in other words prosodic boundaries can occur independently of major syntactic ones. - However, some of the phrase contours that arise in the figures ("a" editions) as a result of the "resetting" criterion applied seem distinctly counter-intuitive: (1) No boundary is assigned at the third "$" in JR's and BH's sentence no. 7 and 8, though the contour takes a sharp turn at that place. (2) The final phrase contour is rather sharply rising-falling in BH's sentence 7 and 8. (3) No boundary is assigned at the third "$" in NT's sentence 6, although the resemblance to the final phrase contour in sentence 7 and 8 is striking.

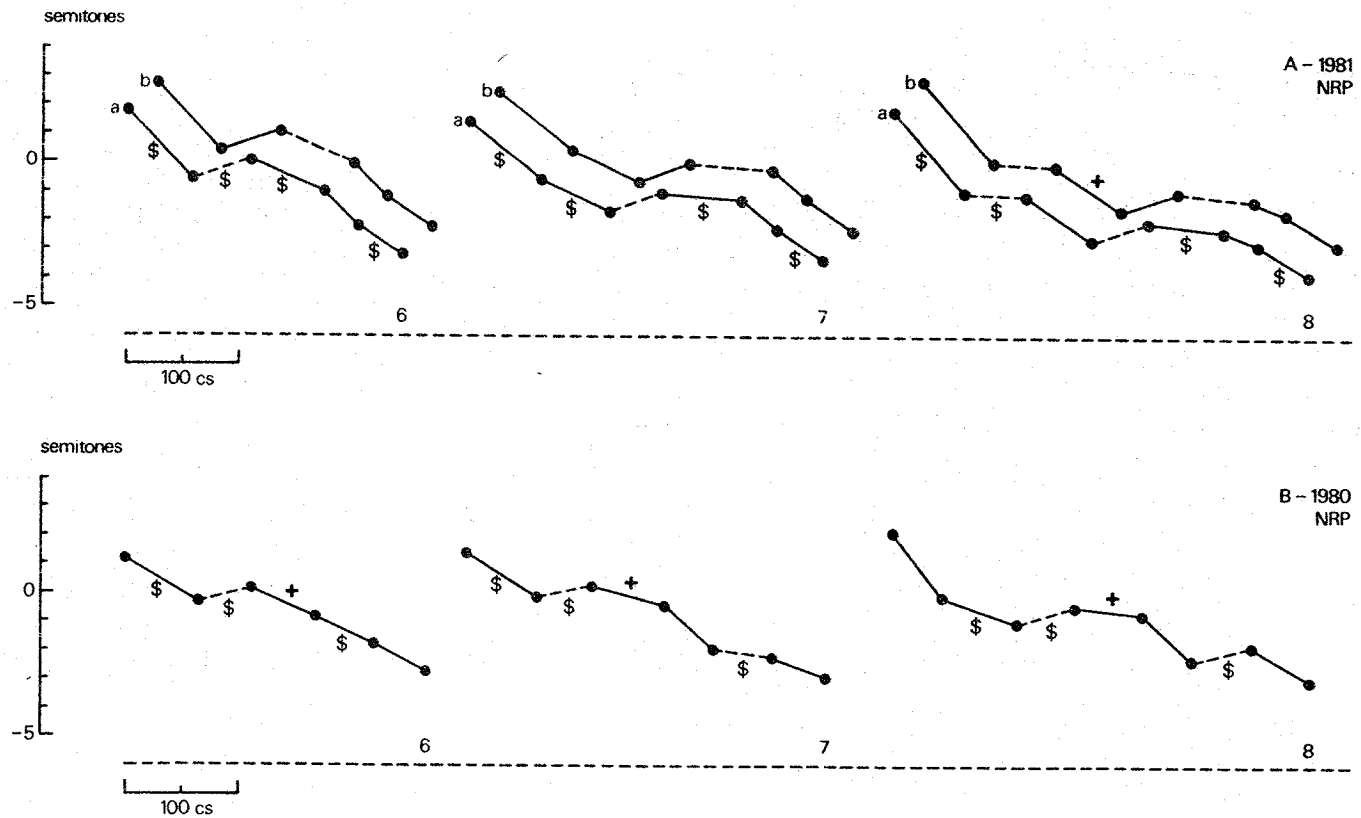

Figure 8

Intonation contours in declarative sentences containing from six to eight stress groups from two different materials: A-1981 at the top and B-1980 at the bottom. The unstressed syllables have been left out. The "a" and "b" editions are identical except for the placement of prosodic phrase group boundaries (indicated with broken lines). The "b" edition has been moved 0.5 semitones upwards and 25 cs to the right with respect to the "a" edition. Subject: NRP.

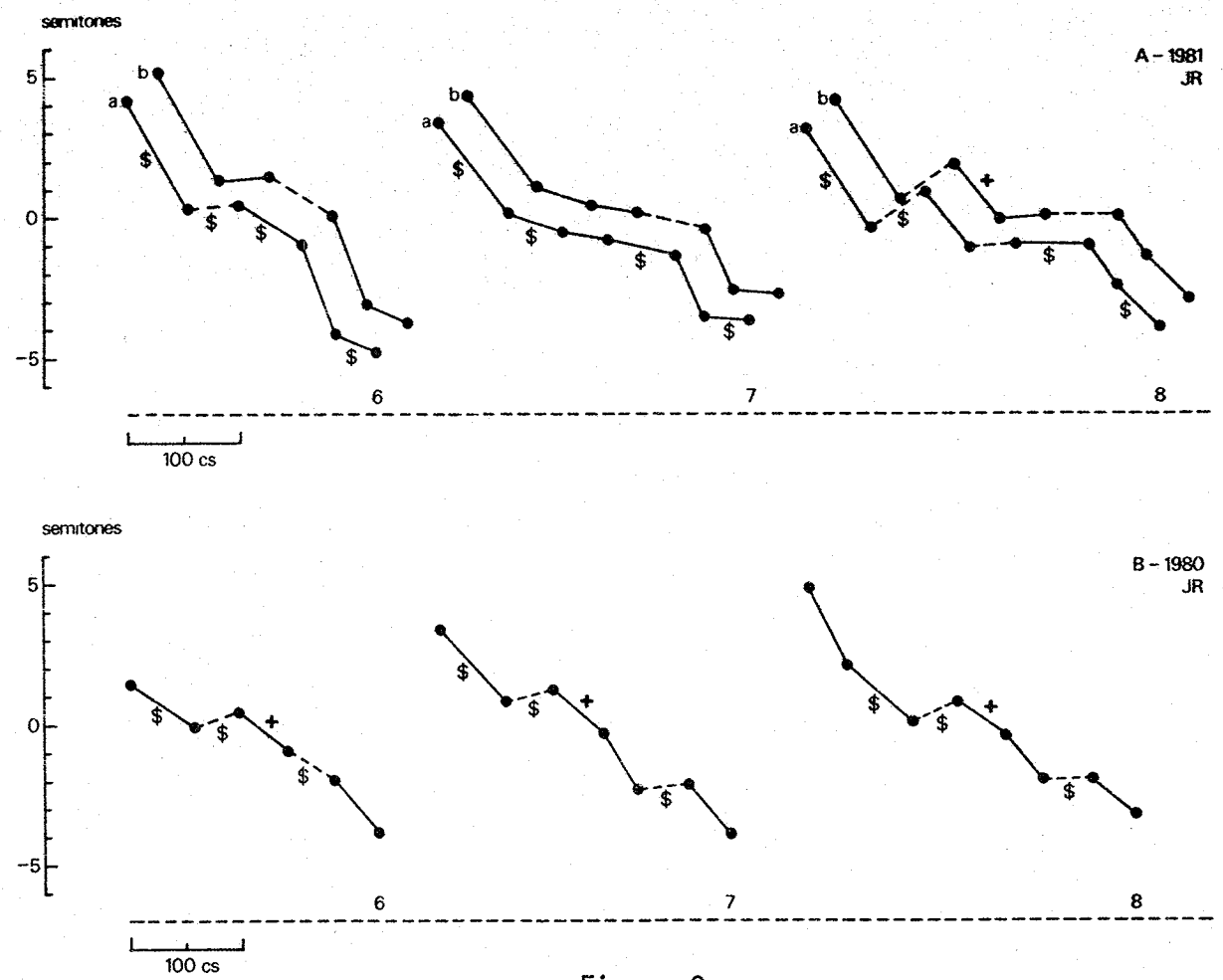

Figure 9

Subject: JR. See further the legend to figure 8.

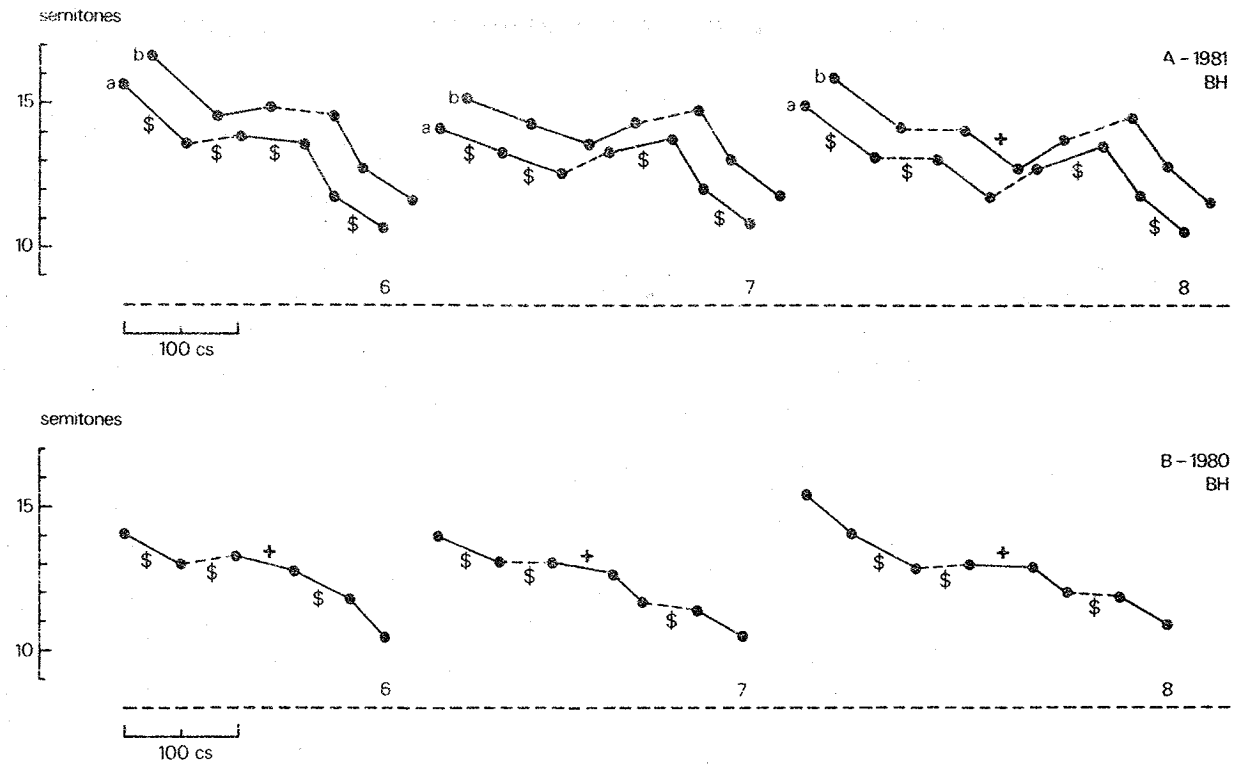

Figure 10

Subject: BH. See further the legend to figure 8.

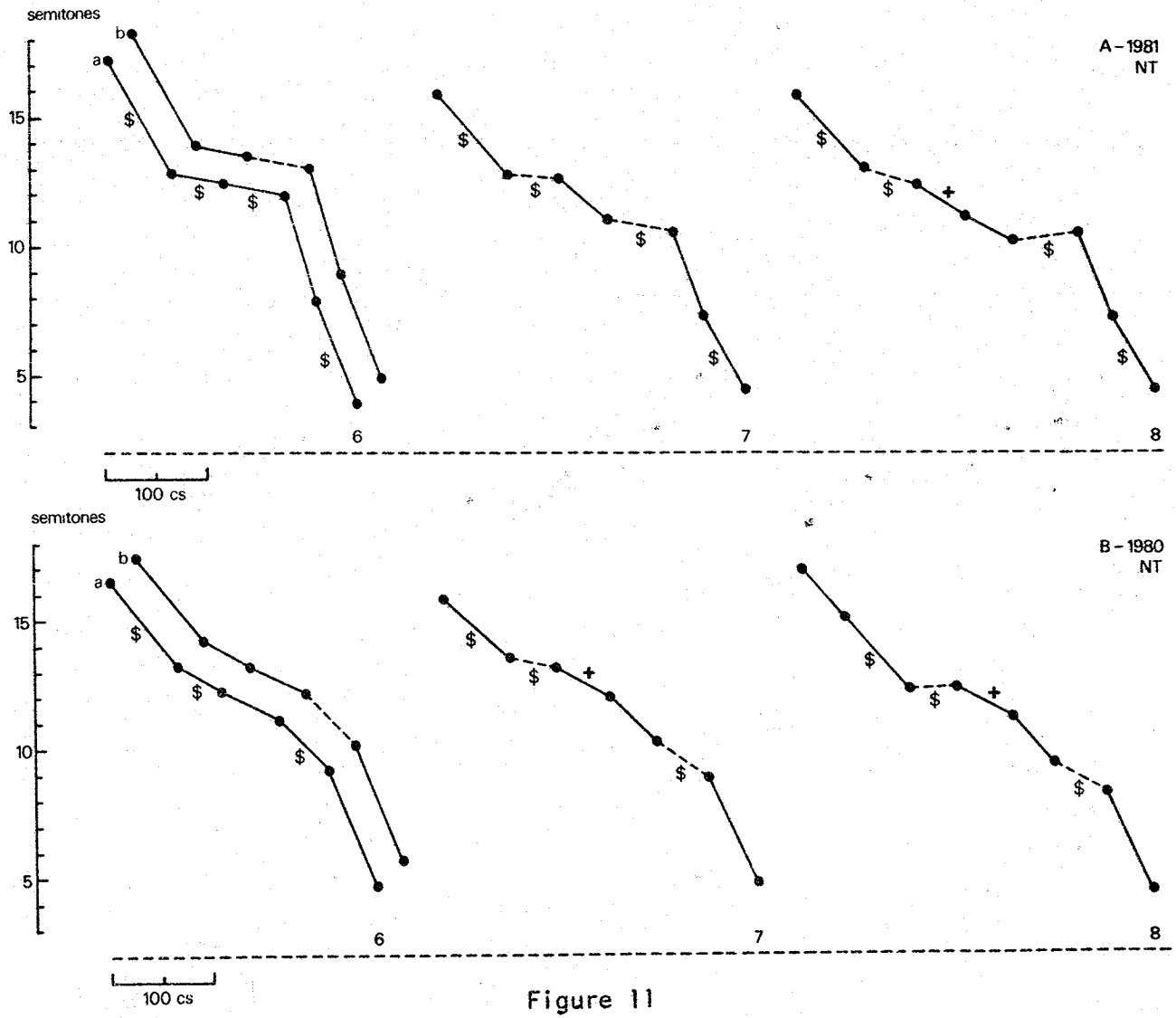

Figure 11

Subject: NT. See further the legend to figure 8.

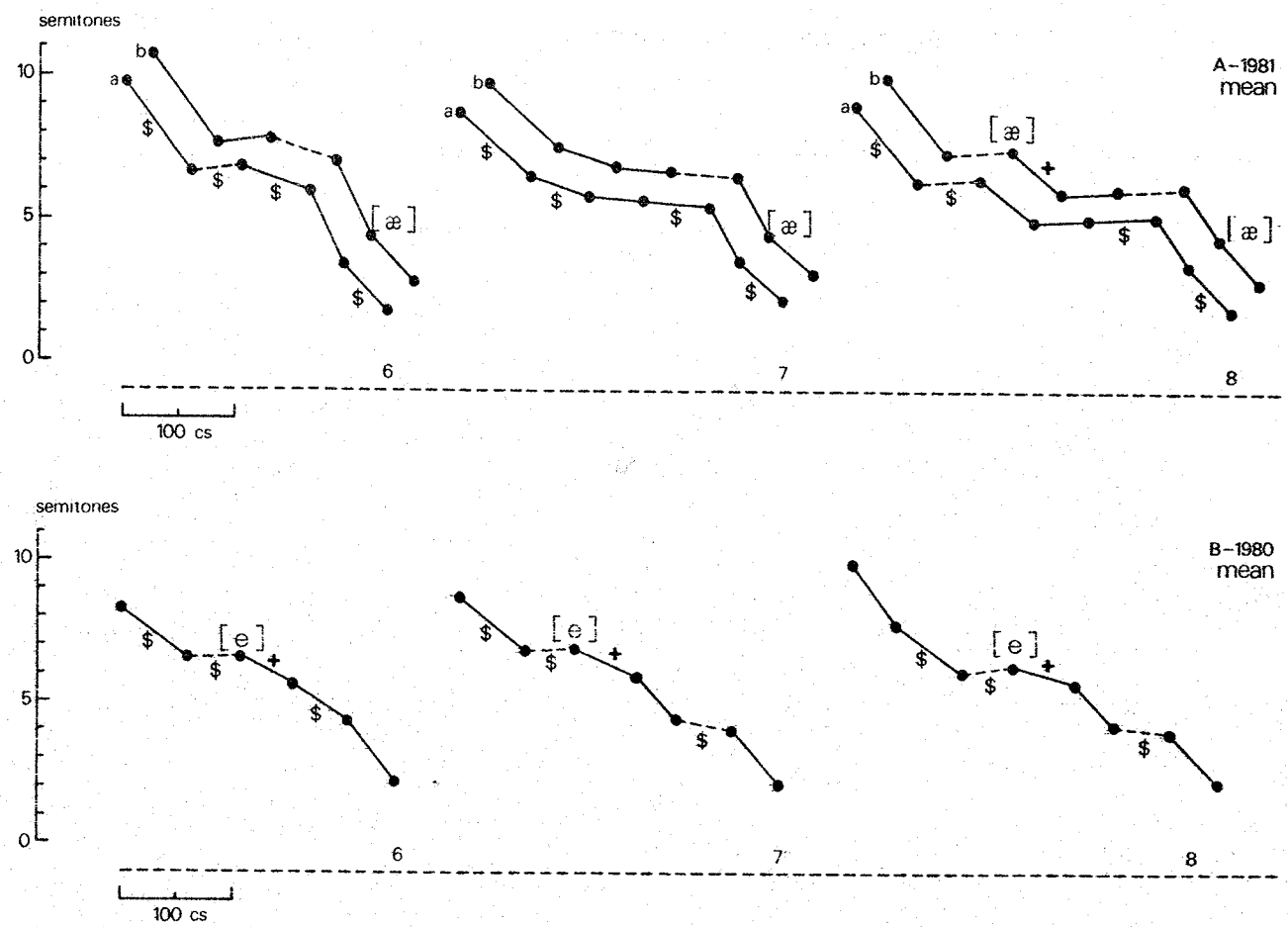

Figure 12

Average over four subjects (mean of means). See further the legend to figure 8.

In the "b" editions of the contours I have suggested an alternative boundary location: Before the antepenultimate stressed syllable of sentence 7 and 8 with JR and BH the contour takes a rather sharp turn, and a boundary is inserted. Since the corresponding final part of sentence 6 bears a striking resemblance to those of 7 and 8, the boundary is also moved (JR and BH) or inserted (NT) here. A similar boundary location can be defended for NRP as well, though his contours are on the whole less jagged-looking. In this way a certain coherence is established across subjects and sentences. Note that the 1980 criterion is not at crosscurrents with the present boundary assignment where the 1980 sentences are concerned, except that a boundary is suggested in NT's sentence 6 at the last "$". - Now the predictions and the "actual" groupings correspond very well indeed:

	predicted	JR	NRP	BH	NT
6-1981:	2+4 or 3+3	3+3	3+3	3+3	3+3
7-1981:	2+2+3 or 4+3	4+3	4+3	4+3	2+2+3
8-1981:	2+3+3	2+3+3	2+3+3	2+3+3	2+3+3

Unfortunately, it is no longer possible to assign prosodic boundaries automatically and unambiguously from the intonation contours alone: with the boundaries suggested in the "b" editions we get phrase contours that describe asymptotic declinations or even rises at the end (see e.g. BH), but when phrase final rises are admitted (and I do not see why, a priori, they should not be) then a number of alternative boundary locations present themselves (which only violate the hypothesis that a prosodic boundary cannot occur unaccompanied by a major syntactic one - a hypothesis that would therefore have to be discarded) as indicated by the crosses in the figures. A further difficulty would arise with an "isolated" sentence, i.e. one that could not be compared to any others in order that a 'coherence across sentences' criterion might be applied.

3. PROSODIC BOUNDARY PERCEPTION

At this point the need to know something about prosodic boundary perception becomes pressing, because if the discontinuities and resettings of these intonation contours have no perceptual significance per se, i.e. as prosodic boundaries properly speaking, then the search for definitions of them and criteria for retrieving them from the intonation contours (without regard to syntactic structure) becomes futile.

A number of studies are concerned with the prosody/syntax relation: Uyeno et al. (1980) find that Fo will disambiguate syntactically ambiguous sentence structures in Japanese (clause structures); Streeter (1978) reports that Fo will disambiguate ambiguous algebraic expressions in American English, whereas duration is a less stable cue; Lehiste et al. (1976) tell us that duration will effectively disambiguate syntactically ambiguous utterances (American English), whereas Lehiste

(1980) mentions an unpublished experiment by Olive and Lehiste where Fo was unsuccessful in disambiguating those same utterances. Although the evidence is slightly conflicting, I will assume - for the sake of the point I want to make below - that Fo can indeed disambiguate otherwise ambiguous utterances. - Cooper and Sorensen (1977) analyse Fo at syntactic (clause) boundaries and find significant differences as a function of syntactic coding but are not concerned with the perception of these differences. - Umeda et al. (1975) and Harris et al. (1981) investigate listeners' perception of (syntactic) boundaries in fluent speech, and Umeda and Quinn (1981) follow up the perceptual experiments with an analysis of word duration, which is shown to be positively correlated with the listeners' perception of boundaries; they also note that some listeners seem to be more sensitive to elongation than others, some are more sensitive to pitch contours than others, and some to a strong initial allophone, but they make no claim that prosody actually cues syntactic boundary perception in general.

Where disambiguating otherwise ambiguous utterances is concerned I think it reasonable to assume that we are actually dealing with perception of prosodic boundaries, which then lead to an interpretation of syntactic structure. But in syntactically unambiguous, non-compound sentences such as the ones under investigation here, I do not know and rather doubt whether prosodic boundaries will be perceived independently of (and thus be able to cue the perception of) syntactic boundaries. Given the complete speech signal a native speaker's prosodic boundary perception may just be a rationalization of his linguistic (and semantic) interpretation of the utterance. - I have listened to the recordings a number of times and tried to listen exclusively for prosodic boundaries (which may be an illusory attempt): I can really detect no boundaries at all with NRP and NT. With BH I can detect a boundary at the third "$" in most instances. With JR I can hear a boundary in sentence 8 at the third "$" and occasionally also at the second "$" in sentence 8. This would be a (weak) support for the boundary assignment in the "b" editions. However, with repeated listening I can induce myself to hear prosodic boundaries practically all over the utterances, and in an erratic fashion. - Evidently, reliable perceptual results should be obtained from low-pass filtered speech or synthetically produced signals.

If it turned out from perception tests with stimuli produced on the basis of intonation contours such as these, retrieved from syntactically unambiguous non-compound sentences, that listeners would not identify and locate consistently prosodic boundaries then it would hardly be appropriate to speak of them as boundaries any longer. It would also mean that we should content ourselves with a description of the intonation contours that arise as a result of syntactic structure (and semantic content, cf. below) - but we would still have to state that certain syntactic boundaries leave

no trace in the intonation contour, i.e. certain syntactic constituents will be tied together in one prosodic phrase - and we would also have to describe the conditions under which syntactic boundaries give rise to turns and breaks in the intonation contour and the possible phrase contour shapes that arise as a result. - Clearly, both of the materials investigated are too limited to deal satisfactorily with the last problem; further investigations are called for.

4. THE ROLE OF SEMANTICS?

Let us for the moment assume that the assignment of "boundaries" in the "b" editions in figure 8-12 is an adequate representation of the division into subparts, prosodic phrases, of the intonation contours and let me try and explain the one most striking difference between the intonation contours of the two materials. In 1981 there are a number of instances where the phrase contour rises before the final phrase contour (viz. JR 6 and 8 (only slightly), NRP and BH 6, 7 and 8). If we had had only BH's sentence 7 and 8 to judge from, a likely explanation could have been found in the fact that the final phrase contour must perform a rather steep fall, and since it contains three stress groups, it will have to start rather high up in the frequency scale, so the rise in the preceding contour could be an anticipation of this high start of the final contour. However, NRP and JR contradict this hypothesis. In fact, there is nothing in the intonation contours per se which can explain the phrase contour rises.

The final phrase group in all three sentences (6, 7 and 8) consists in the 1981 material of two complements, one of time followed by one of place (... *klokken ét i nát $ til Tíflis.*). The corresponding final phrase group in the 1980 material was one complement of place (... *på "Kílden" i Thísted.*). I do not think that two versus one complement, or a three-stress versus a two-stress phrase group have anything to do with the difference in intonation contour. I would venture a more semantically oriented explanation: The time complement is less intimately related to the preceding part of the utterance, it is more of a unit apart; that is, the boundary before it is stronger (at least in these particular utterances) and to signal this we get what might appropriately be termed "continuation rises" in the preceding contour. Note that the third "$" would also have been a very likely place for a pause to occur (not that it did). With this analysis we can distinguish two kinds of "rises" in an intonation contour: a phrase final continuation rise, which is a "local" deflection, and a resetting of the intonation contour which separates two phrases (see e.g. JR, sentence 8) and whose purpose is to keep the intonation contour within the speaker's Fo range.

It is tempting to illustrate and support the role of semantics in prosodic structuring by the "boundary" placement in sentence 6 in the two materials (JR, NRP and BH), even though the

intonation contours as a whole do not look widely different. In *Pýtte $ skal med bússen $ til Thísted /$/ klokken ét i nát $ fra Tíflis.* (1981) the break ("//") occurs after the first complement (and is preceded by a continuation rise) but in *Aníta $ skal med bússen /$/ til fésten for Kísser $ på "Kílden" i Thísted.* (1980) it occurs before the first complement. In other words, the prosodic phrasing is not governed by surface syntactic structure exclusively, the semantic content of the syntactic constituents is also taken into account, in this case inducing a "boundary" before the "heavier" time complement in the 1981 utterance, which makes the first place complement tie up with the preceding NP and VP. This argument is invalid, however, if *bússen til Thísted* is one syntactic constituent, rather than two (equivalent to *Thísted-bussen* 'the Thisted-bus'). In that case the explanation for the boundary placement in the 1981 utterance is not to be sought in a heavy succeeding time complement but in a weak boundary between *bússen* and *til Thísted.* Note, though, that this syntactic closeness is not accompanied by a stress reduction on the first element which is otherwise characteristic of close-knit syntactic relations in Standard Danish (like *køre bíl* 'to drive a car' and many others).

III. CONCLUSION

In Thorsen (1980a) I concluded that the results presented an argument in favour of a theory put forward in Selkirk (1980) that prosodic categories are distinct entities in the phonology that do not have an isomorphous relation to syntactic structure. This claim is certainly not weakened by the material analysed here. In syntactically unambiguous non-compound sentences the prosodic stress group will cut across any syntactic boundary, and when - in longer utterances - a division of the intonation contour into prosodic phrase groups is necessitated, this phrasing bears no simple relation to surface syntactic structure. Furthermore, I suggest that the matter is rendered even more complicated by the role that semantics may have to play in prosodic structuring.

How and to what extent the results would be applicable to free speech I cannot say. One might speculate that prosody plays a more important role in the production and perception of free speech - which is rarely so syntactically well-formed as the schematized material presented for reading in this investigation. That is: prosodic boundaries may be more evident (also when unaccompanied by pauses) in free speech and may of course take more and different shapes than encountered here.

ACKNOWLEDGEMENT

My sincere thanks are due to Jeanette Holtse for her efficient graphics work.

REFERENCES

Cooper, W.E. and Sorensen, J.M. 1977: "Fundamental frequency contours at syntactic boundaries", *J.Acoust.Soc.Am. 62*, p. 683-692

Harris, M.O., Umeda, N. and Bourne, J. 1981: "Boundary perception in fluent speech", *J.Phonetics 9*, p. 1-18

Lehiste, I. 1980: "Phonetic manifestation of syntactic structure in English", *Ann.Bull.Res.Inst.Log.Phon. 14*, p. 1-27

Lehiste, I., Olive, J.P. and Streeter, L.A. 1976: "Role of duration in disambiguating syntactically ambiguous sentences", *J.Acoust.Soc.Am. 60*, p. 1199-1202

Selkirk, E.O. 1980: "On prosodic structure and its relation to syntactic structure", *Indiana University Linguistics Club*, (Bloomington, Indiana)

Streeter, L.A. 1978: "Acoustic determinants of phrase boundary perception", *J.Acoust.Soc.Am. 64*, p. 1582-1592

Thorsen, N. 1980a: "Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish", *Ann.Rep.Inst.Phon., Univ.Cph. 14*, p. 1-29

Thorsen, N. 1980b: "A study of the perception of sentence intonation - Evidence from Danish", *J.Acoust.Soc.Am. 67*, p. 1014-1030

Umeda, N., Harris, M. and Forrest, K. 1975: "The placement of auditory boundaries in fluent speech", *J.Phonetics 3*, p. 191-196

Umeda, N. and Quinn, A.M.S. 1981: "Word duration as an acoustic measure of boundary perception", *J.Phonetics 9*, p. 19-28.

Uyeno, T., Hayashibe, H., Imai, K., Imagawa, H. and Kiritani, S. 1980: "Comprehension of relative clause construction and pitch contours in Japanese", *Ann.Bull.Res.Inst. Log.Phon. 14*, p. 225-236.

APPENDIX

1. RANGE

There is an overall tendency for range to increase with increased utterance length, cf. figure 13, but range does not increase monotonically, no matter whether range is determined by the interval between first and last stressed vowel (13a), first and last post-tonic vowel (13b), or absolute Fo maximum (first post-tonic vowel) and absolute Fo minimum (last stressed vowel, except with NT where the minimum is constituted by the last post-tonic vowel in sentences 1-4 and 6-8) (13c).

The range dispersion on the grand mean in figure 13a (stressed vowel interval) is 3.6 semitones (corresponding to 80% of the smallest range), lying between 5.9 (NT) and 1.8 semitones (NRP); in figure 13b (post-tonic vowel interval) it is 2.2 semitones (corresponding to 29% of the smallest range), lying between 3.9 (JR) and 0.9 semitones (BH), and in figure 13c (interval between absolute Fo maximum and minimum) it is 2.6 semitones (corresponding to 31% of the smallest value), lying between 5.7 (JR) and 0.9 semitones (BH).

2. STARTING AND END POINTS

Figure 14 depicts the level of starting points (14a: first stressed vowel, 14b: first post-tonic vowel) and end points (14c: last stressed vowel, 14d: last post-tonic vowel). If we disregard sentence no. 1, which obviously groups itself with the end points, there is only a slight and irregular tendency towards higher starting points with the longer utterances, cf. the slopes of the least squares regression lines on the data points of figure 14a-d and their correlation coefficients in table 3. End points decrease more, at least through sentence 1 to 4 (and with most subjects they also decrease more regularly than starting points increase, the correlation coefficients generally being numerically greater on the end point regressions). Only BH shows a deviant pattern: in the present material starting and end points increase and decrease, respectively, to approximately the same extent, and in the 1980 material (table 3B) starting points increase more than end points decrease. With JR in the 1980 material first and last post-tonic in- and decreased equally. - The near-constancy of the end points from sentence no. 5 and higher was also observed in 1980 and can probably be ascribed to a physiological constraint: the speaker has a lower limit to his Fo range, which he is bound to hit with utterances exceeding a certain length (in terms of number of stress groups).

There are individual differences in the various curves in figure 13 and 14, and no clear-cut pattern in the association between range variation and starting versus end point variation can be found (i.e. even though end points decrease

more than starting points increase):

a. Stressed vowel range versus starting and end points — The correlation between stressed vowel range and first stressed vowel level in this material is as follows: JR 0.78, NRP 0.81, BH 0.95, NT 0.85 as opposed to a correlation between range and last stressed vowel level of: JR -0.92, NRP -0.67, BH -0.79, NT -0.97. Thus, the correlation is numerically stronger between range and starting point with NRP and BH. In the 1980 sentences the same calculations yield for range and starting point correlations: JR 0.79. NRP 0.49, BH 0.75, NT 0.86; range and end points: JR -0.67, NRP -0.80, BH -0.35, NT -0.96, i.e. JR and BH have stronger numerical correlations between range and starting points.

b. Post-tonic vowel range versus starting and end points — The correlations between the post-tonic vowel range and level of the first post-tonic in this material are: JR 0.92, NRP 0.85, BH 0.85, NT 0.78; between range and last post-tonic level: JR -0.43, NRP -0.62, BH -0.19, NT -0.88. In the 1980 material the same correlations yield: JR 0.88, NRP 0.62, BH 0.97, NT 0.60 and JR -0.85, NRP -0.56, BH 0.12, NT -0.96. Thus, in both materials range and first post-tonic correlate numerically more strongly than range and last post-tonic with JR, NRP and BH, vice versa with NT.

Concludingly we can say - as for the 1980 material - that fundamental frequency range is not constant over utterances of different length, neither is it a linear function of utterance length. The range variation is brought about by a combination of variation in starting and end points, and at least with some subjects end points lower more than starting points increase, until "saturation" is reached (with utterances of four to five and more stress groups).

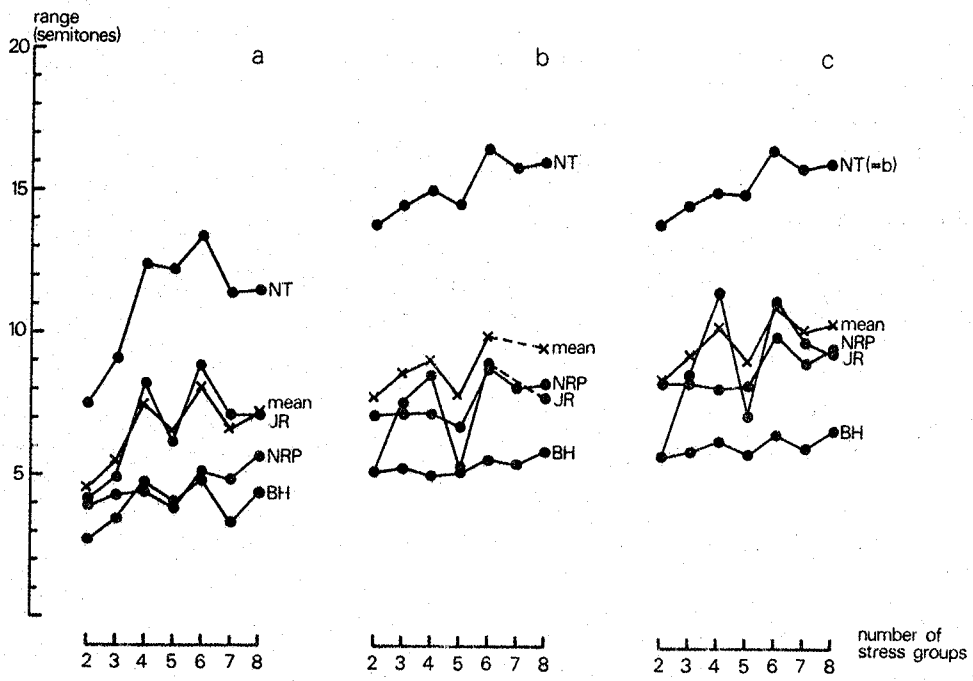

Figure 13

Range of fundamental frequency in seven declarative utterances, containing from two to eight stress groups, depicted as a function of the length of the utterance (in terms of number of stress groups). Four subjects and their grand mean (crosses). In (a) range is defined as the interval between first and last stressed vowel measurement in each utterance; in (b) range is defined as the interval between the first post-tonic vowel in the first and last stress group in each utterance; in (c) range is defined as the interval between the absolute Fo maximum (i.e. the first post-tonic vowel in the first stress group) and the absolute Fo minimum (i.e. the last stressed vowel) in each utterance. (NT's Fo minimum is constituted by the last post-tonic vowel and the curve in (c) is identical to the one in (b). In (b) data on sentence no. 7 is lacking with JR (and the mean).

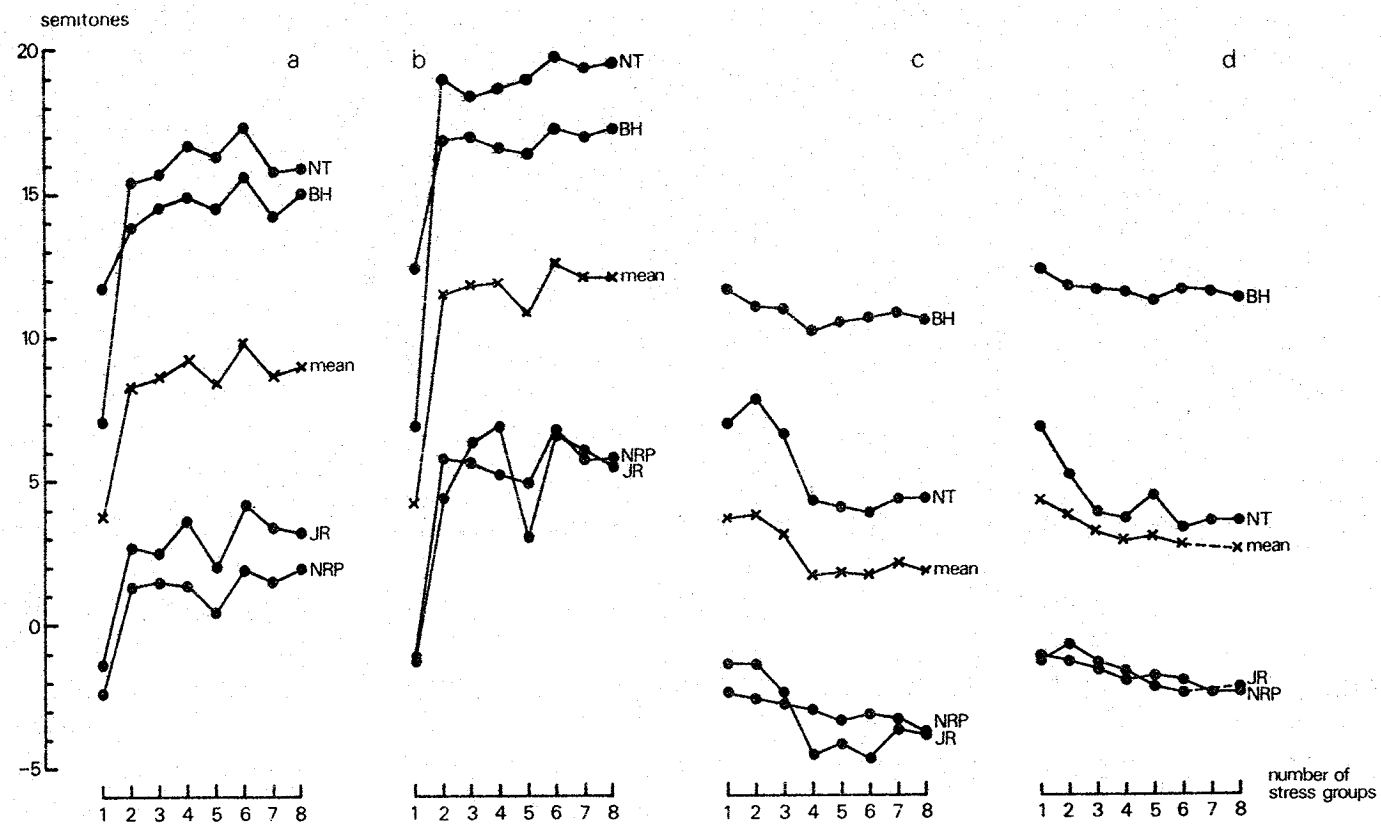

Figure 14

Fundamental frequency starting and end points in eight declarative utterances, containing from one to eight stress groups, depicted as a function of the length of the utterance (in terms of number of stress groups). Four subjects and their grand mean (crosses). In (a) is depicted the frequency of the first stressed vowel in each utterance; in (b) the frequency of the first post-tonic vowel in each utterance. In (c) is depicted the frequency of the last stressed vowel, and in (d) the last post-tonic vowel in each utterance. In (d) data on sentence no. 7 is lacking with JR (and the mean). Zero on the logarithmic frequency scale corresponds to 100 Hz.

Table 3

A: Least squares regression line slopes/Pearson product moment correlation coefficients on the data points in figure 14a-d. Sentence no. 1 is excluded in the calculations on the data in figure 14a and 14b.

	JR	NRP	BH	NT	mean
Starting points, stressed vowels (figure 14a)	0.14/0.18	0.08/0.35	0.13/0.49	0.08/0.27	0.10/0.12
Starting points, post-tonics (figure 14b)	0.09/0.14	0.06/0.23	0.07/0.44	0.11/0.76	0.11/0.42
End points, stressed vowels (figure 14c)	-0.42/-0.76	-0.18/-0.95	-0.11/-0.60	-0.52/-0.81	-0.31/-0.82
End points, post-tonics (figure 14d)	-0.22/-0.84	-0.19/-0.96	-0.10/-0.73	-0.38/-0.78	-0.23/-0.90

(table 3 continued)

B: Same as A on the 1980 material (Thorsen 1980a, figure 8, p. 15)

	JR	NRP	BH	NT	mean
Starting points, stressed vowels (figure 8a)	0.14/0.26	-0.04/-0.21	0.18/0.64	0.11/0.53	0.10/0.42
Starting points, post-tonics (figure 8b)	0.22/0.69	0.05/0.21	0.20/0.65	0.17/0.88	0.15/0.94
End points, stressed vowels (figure 8c)	-0.39/-0.86	-0.23/-0.91	-0.05/-0.31	-0.41/-0.83	-0.27/-0.84
End points, post-tonics (figure 8d)	-0.23/-0.82	-0.20/-0.90	0.02/0.29	-0.52/-0.78	-0.23/-0.82