(ARIPUC 14, 1980)

INTONATION CONTOURS AND STRESS GROUP PATTERNS IN DECLARATIVE SENTENCES OF VARYING LENGTH IN ASC DANISH[1]

Nina Thorsen

Abstract: Four subjects recorded eight non-compound declarative sentences, containing from one to eight stress groups. Acoustic analysis reveals a tendency for fundamental frequency range to increase with increased utterance length, but in a non-linear and seemingly random fashion. The increase is brought about by higher starting points as well as lower ending points in the longer utterances. Concomitant with the range increase we find a decrease in overall downdrift in the longer utterances, but degree of downdrift is not simply inversely related to utterance length. With four and more stress groups the intonation contour is decomposed into prosodic phrase groups, i.e. the contour contains discontinuities in the shape of partial resettings. The prosodic phrase group boundaries are determined by but do not exactly coincide with major syntactic boundaries, and the data present an argument in favour of a hypothesis of prosodic categories as distinct entities with a non-isomorphous relation to syntactic structure.

1. Introduction

The relationship between stress and fundamental frequency (Fo) and the intonation contours of various types of short sentences in Advanced Standard Copenhagen (ASC) Danish have been described elsewhere (Thorsen 1978, 1979b). For the purpose of the present paper only a few points need be repeated: Stress in ASC Danish is signalled mainly by Fo. In neutral speech a stressed syllable will be (relatively) low and followed by a high-falling tail of unstressed syllables, i.e. the stressed syllable is one that is jumped or

---

1) Revised and expanded version of a paper published in ARIPUC 13 (1979), p. 1-7. Instead of referring extensively to vol. 13, I have chosen to repeat some passages from the earlier version. An abbreviated edition of this paper will appear in the proceedings of the "2nd Symposium on the Prosody of Nordic Languages, Trondheim 19-21 June 1980".

glided up from, depending on the segmental composition, cf. fig. 1 (full lines). The unit which carries this Fo pattern consists of the stressed syllable plus all succeeding unstressed ones, irrespective of intervening syntactic boundaries within the simple (i.e. non-compound) sentence. It is termed a stress group (SG). (A detailed account of the stress group and its tonal properties can be found in Thorsen 1980b - this volume, sections 2 and 7.)

The Fo patterns of SGs are predictable and recurrent entities (though allowing for certain context dependent modifications), wherefore the intonation contour may be defined solely in terms of the stressed syllables. (This does not necessarily mean that the course of the unstressed syllables is irrelevant, e.g. for the perception of intonation contours but it is, strictly speaking, redundant, cf. Thorsen 1980a.) This concept of intonation contour is different from the current 'topline' or 'baseline' concepts: To Bruce (1979) and Gårding (1979) (Swedish) the overall contour of an utterance is determined by a topline (connecting local Fo

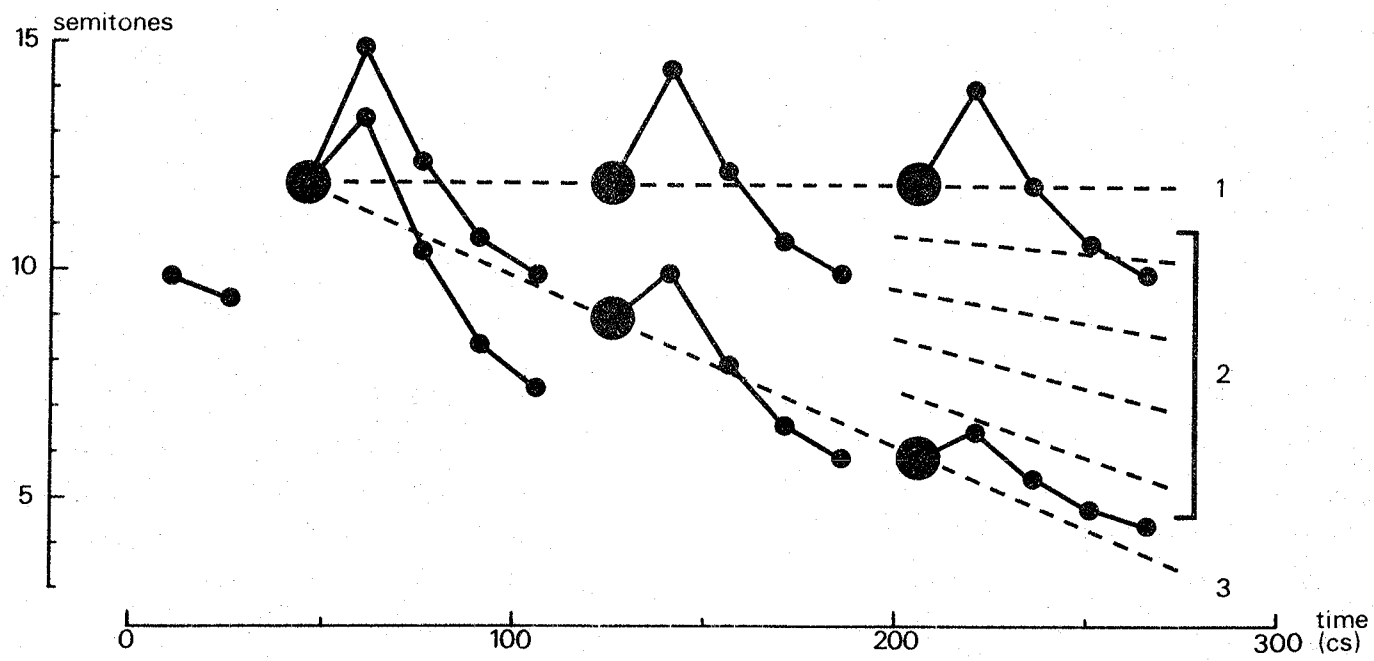

Figure 1

A model for the course of Fo in short sentences in ASC Danish. 1: syntactically unmarked questions, 2: interrogative sentences with word order inversion and/or interrogative particle and non-final periods (variable), 3: declarative sentences. The large dots represent stressed syllables, the small dots unstressed ones. The full lines represent the Fo pattern associated with stress groups, and the broken lines denote the intonation contours. Zero on the logarithmic frequency scale corresponds to 100 Hz.

maxima) as well as a baseline (connecting local Fo minima), with the topline declining more rapidly than the baseline. Maeda (1977) (American English) and Fujisaki et al. (1979) (Japanese) attribute the downdrift in declarative sentences to a baseline, but Fujisaki et al.'s baseline (termed the 'voicing component') is an abstraction in the sense that it need have no direct physical representation in the Fo course which is the combined result of the voicing and accent control mechanisms. 't Hart and Collier (1975) (Dutch) have the baseline (connecting Fo minima) as determinant of the declination in declarative sentences. Breckenridge and Liberman (1977), Pierrehumbert (1979), and Sorensen and Cooper (1980) (American English) let the topline (connecting local Fo maxima) determine the downdrift (but, apparently, the maxima always coincide with the stressed syllables of the utterance). Sternberg et al. (1980) (American English) define the downdrift in terms of the stressed syllables, but presumably this would be identical to a topline (their test material contained no unstressed syllables). Neither a topline (connecting local maxima) nor a baseline (connecting local minima) will serve as useful determinants of the overall contours in ASC Danish because they are both highly dependent upon the stress group composition, i.e. on the number of unstressed syllables (if any) in the SG, cf. fig. 1: The baseline may be but is not invariably coincident with a connection of the stressed syllables (if, e.g., the first SG contains more than two unstressed syllables); the topline may be but is not invariably coincident with a connection of the first post-tonic syllables (if, e.g., the second SG contains no post-tonic): in both instances the base- and toplines would steer a zig-zag course downwards. (Furthermore, there are individual differences in stress group patterns: compare fig. 3 (JR) to figs. 2, 4, and 5 and note that JR's unstressed syllables transgress the intonation contour in several instances.) A line connecting the stressed syllables, however, will always exhibit a smooth and gradual course. Furthermore, in an experiment where subjects had to identify utterances as interrogative, non-final, or declarative, solely on the basis of their fundamental frequency course, it turned out that the distribution of subjects' responses was more closely correlated with the stressed syllables in the stimuli (Thorsen, 1980a).

In simple sentences in ASC having no more than three SGs, the intonation contours were found to approach straight lines whose slopes varied according to sentence type. Declarative sentences have the most steeply falling contours at one extreme and syntactically unmarked questions have level contours at the other. In between are found various syntactically marked questions as well as non-final clauses, with a tendency for a trade-off relationship between syntax and intonation contour: The more syntactic information is contained in the sentence about its interrogative or non-final function, the more declarative-like is its intonation contour, and vice versa, cf. fig. 1 (broken lines).

The literature on intonation contours in sentences of varying length is generally only concerned with declarative utterances. There seems to be consensus on an overall downdrift being characteristic of such utterances, i.e. downdrift is a global rather than local phenomenon, but descriptions vary with respect to the extent and shape of the declination. A majority of the authors cited above adhere to the simplest possible model where range is constant over utterances of different length and consequently the rate of the downdrift is inversely proportional to the length of the utterance it spans. This is true of Bruce (1979) and Gårding (1979), the numerous works of Cohen, Collier and 't Hart, explicited in 't Hart (1979), Weitzman (1970) (Japanese, cited from Ohala 1978), Hirose (1971) (Japanese, cited from Ohala 1978), Silverstein (1976) (Hausa, cited from Ohala 1978), Sternberg et al. (1980), and Maeda (1977); Pierrehumbert (1979) finds support for this model in perceptual experiments. McAllister (1971) and Sorensen and Cooper (1980) find that range increases with increased length: the longer utterances start higher than the shorter ones, whereas the lower limit is nearly constant. (An examination of Sorensen and Cooper's data reveals, however, that in addition to the range variation, there is also a slope variation: the longer utterances have less steep slopes than the shorter ones.)

McAllister (1971), Fujisaki et al. (1979), and Sorensen and Cooper (1980) deviate from most other writers on the subject who describe the downdrift in terms of straight lines. Common to their descriptions is a more rapid decline in the early part of the utterance.

For ASC Danish I have previously hypothesized (1979b) that range would be constant over utterances of different length, and that slope would vary inversely with length, and the stressed syllables between the first and last ones would be equidistantly spaced on the (logarithmic) frequency scale. The experiments reported below were designed to test this hypothesis.

2. Material, subjects, and procedures

2.1 Material

Since declarative sentences have the widest range (cf. fig. 1), differences in slope would be most easily detected in them. Accordingly, eight simple statements were made up, containing from one to eight stress groups, all variations on the same theme (´ denotes the stressed vowels and the vertical bar denotes the boundaries between noun phrase and verb phrase, between verb phrase and (compound) complement, and between the two complements):

1. Til Thísted.
2. Tűkke | skal til Thísted.
3. Bűster | skal med bűssen | til Thísted.
4. Kísser | skal med bűssen | til "Kílden" i Thísted.
5. Líssi | skal med bűssen | til fésten | på "Kílden" i Thísted.
6. Aníta | skal med bűssen | til fésten for Kísser | på "Kílden" i Thísted.
7. Hűtters | skal med bűssen | til fésten for Kísser og Líssi | på "Kílden" i Thísted.
8. Knűdsen og Bítten | skal med bűssen | til fésten for Kísser og Líssi | på "Kílden" i Thísted.

(Sentence no. 8 translates as follows: Knudsen and Bitten are taking a bus to the party for Kisser and Lissi at "Kilden" in Thisted.) The stressed vowels are all short, high (except [ɛ] in 'festen'), and surrounded by unvoiced obstruents (except [l] in 'Kilden' and 'Lissi', and [n] in 'Anita' and 'Knudsen') in order to facilitate the subsequent interpretation of the tracings

(cf. Thorsen 1979a). - Note that the syntactic boundaries all occur after the first post-tonic syllable in the stress groups.

The sentences 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), giving a total of six recordings of each sentence by each speaker.

2.2 Subjects

Four phoneticians recorded the material, three ASC speakers (NRP male, BH and NT female) and one with a slightly more conservative pronunciation (JR male).

2.3 Procedures

The recordings were made with semi-professional equipment (Revox A-77 tape recorder, Sennheiser MD21 microphone, larynx microphone) in a quasi-damped room at the Institute of Phonetics. The tapes were processed by hard-ware intensity and pitch meters (F-J Electronics) and registered on a Mingograph (Elema 800). The signal from the larynx microphone was processed in the hold mode. This, in combination with adjustment of the zero-line to the lower limit of the subject's voice range and full exploitation of the record space of the mingograph galvanometer, yields a good solution of the frequency scale, generally allowing for a measuring accuracy of 1 Hz for males and 2 Hz for females.

Fo of each of the vowels and syllabic consonants was measured at 2/3 of the distance from vowel/consonant onset (cf. Rossi 1971, 1978) which was an uncontroversial procedure since all the vowels/consonants had monotonically falling movements, excepting a few instances where the first post-tonic was rising-falling and was measured at its maximum. The distance in time of each of these points from the onset of the first stressed vowel was also measured. The average Fo measurements were converted to semitones (re 100 Hz) and a correction made for intrinsic Fo level differences between stressed [ˈu], [ˈɛ], and [ˈi], in accordance with Reinholt Petersen's (1978) results: [ˈɛ] is raised by 1.2 semitones; [ˈu] is lowered by 0.5 semitones with BH and NT, and by 0.25 semitones with NRP and JR. No correction was attempted for the unstressed vowels or syllabic consonants, cf. Reinholt Petersen (1979). - The standard devia-

tions on the mean Fo values are generally small. E.g. in the longest sentence (no. 8) they range between 2.5% and 3.5% of the mean for the stressed vowels and between 3% and 4% of the mean for the unstressed vowels/syllabic consonants. This means that production stability is rather great and the figures to follow must be fairly reliable indications of subjects' behaviour.

3. Results

Stylized Fo tracings of the eight sentences are depicted in figs. 2-5 for individual subjects, the grand mean in fig. 6. - It is immediately clear that range is not constant over the utterances, nor are the stressed syllables equidistantly spaced on the frequency scale. (Note that in some stress groups with NRP and BH, namely the last but one in sentences 4-8, a syllable is apparently missing. These were instances where the Fo maximum was only reached in the second post-tonic vowel, which is the one shown in figs. 2 and 4. The behaviour of syllables with assimilated schwa and syllabic consonants are the object of a separate investigation, see Thorsen, forthcoming.)

3.1 Range

Range may be defined in different manners, but no matter what definition we choose, the hypothesized constancy is lacking.

3.1.1 Range determined by the stressed vowels

Fig. 7a depicts the variation in range as defined by the interval between the first and the last stressed vowel. Although there is an overall trend for range to increase with increased number of SGs, range does not increase monotonically with length. Combining this information with the fact that the longer utterances seem to be composed of two and three gradients, respectively, with partial resettings between them (located at the broken lines in figs. 2-6, see further below), a relationship between range variation and the number of partial resettings of the intonation contour suggests itself: introducing a partial resetting might decrease the range. Sentences 2 and 3 contain no discontinuity, sentences 4-6 contain one, and sentences 7 and 8 contain two par-

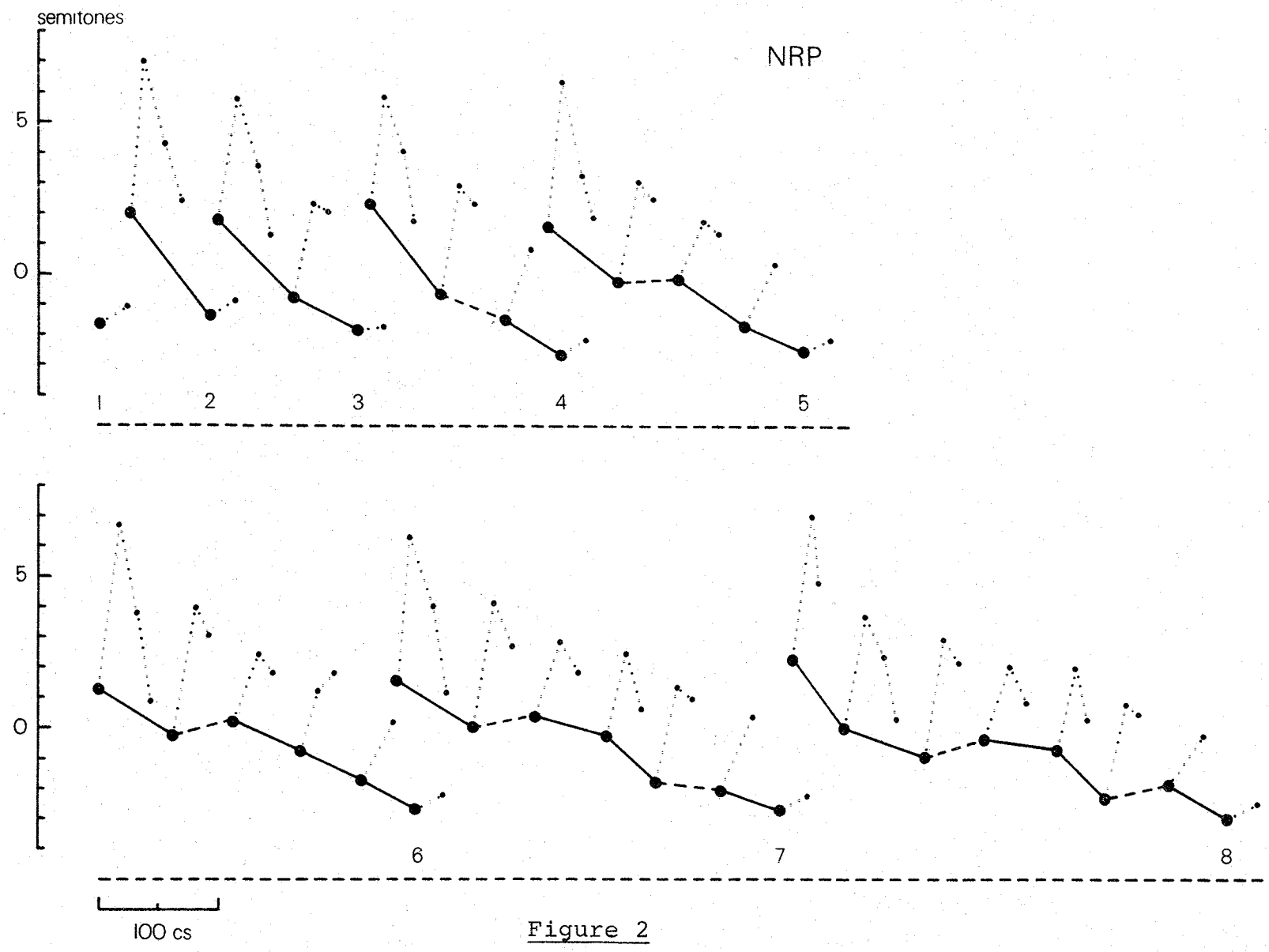

Figure 2

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: NRP.

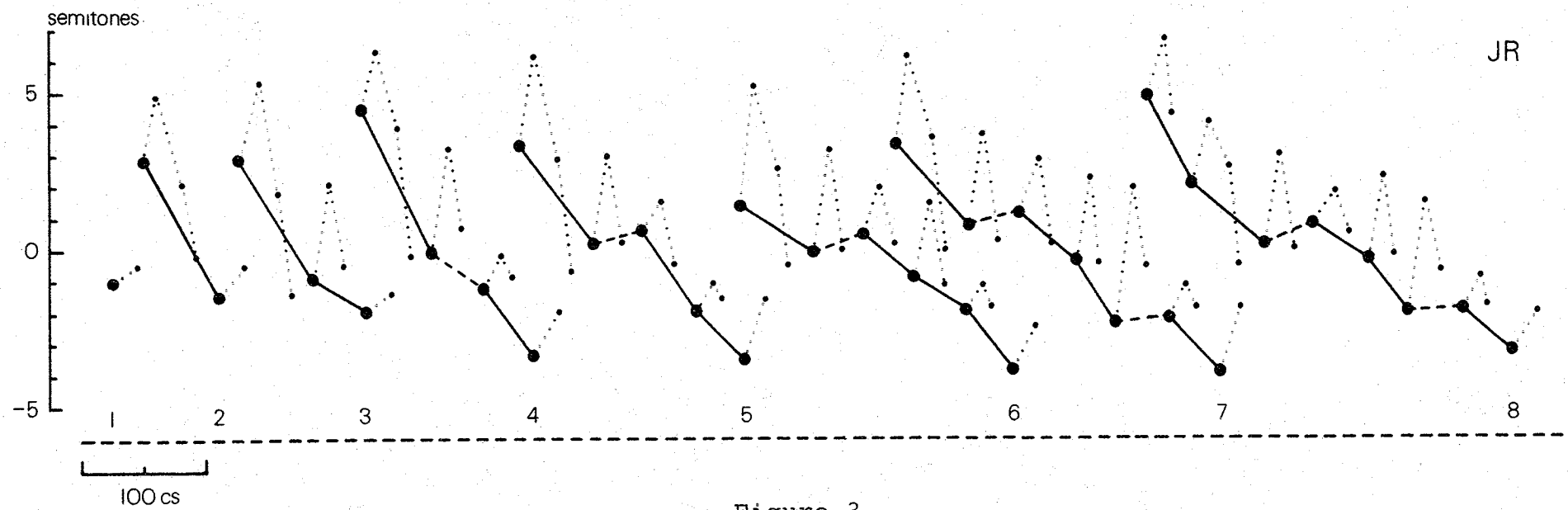

Figure 3

Subject: JR. See further the legend to figure 2.

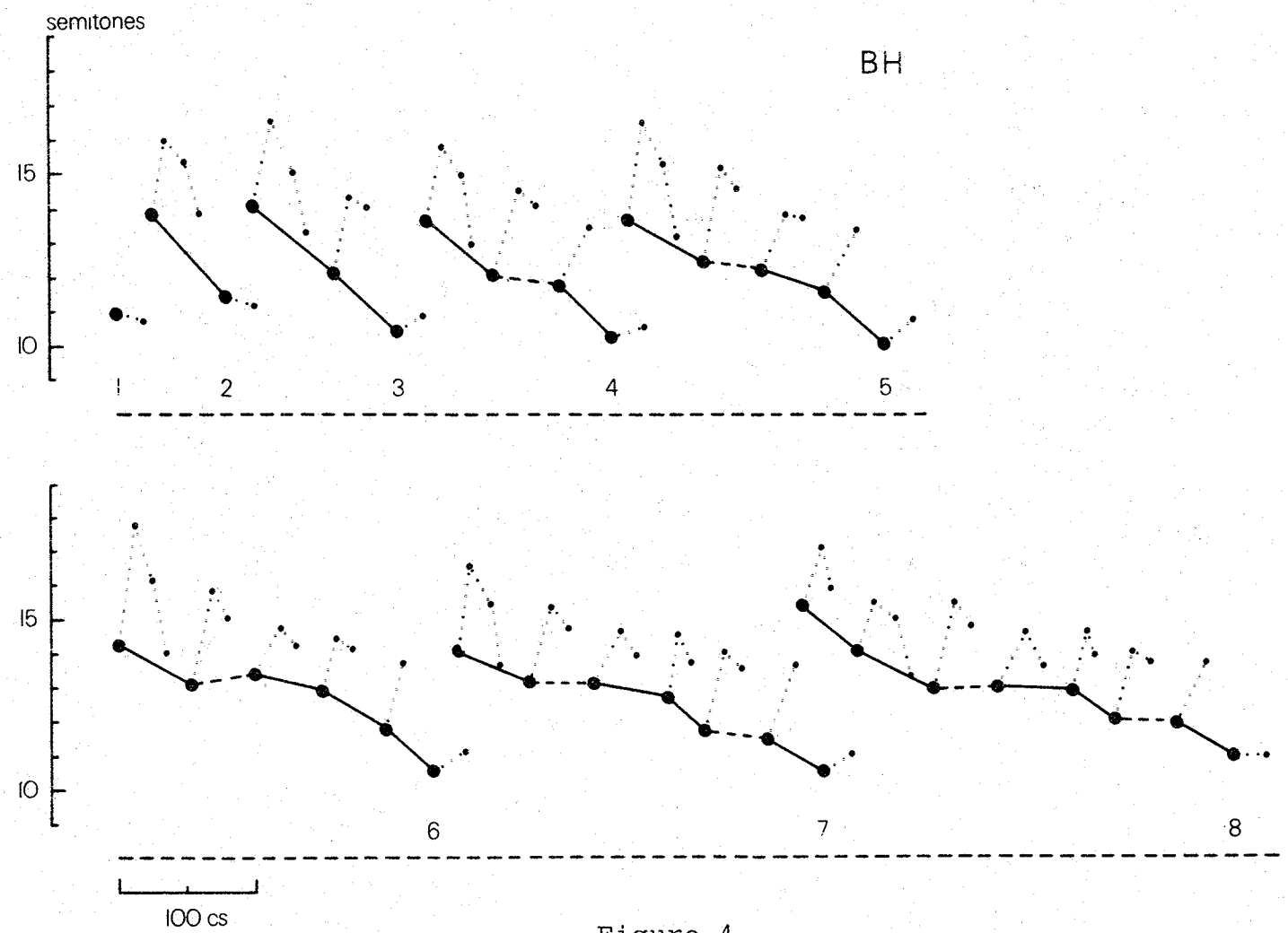

Figure 4

Subject: BH. See further the legend to figure 2.

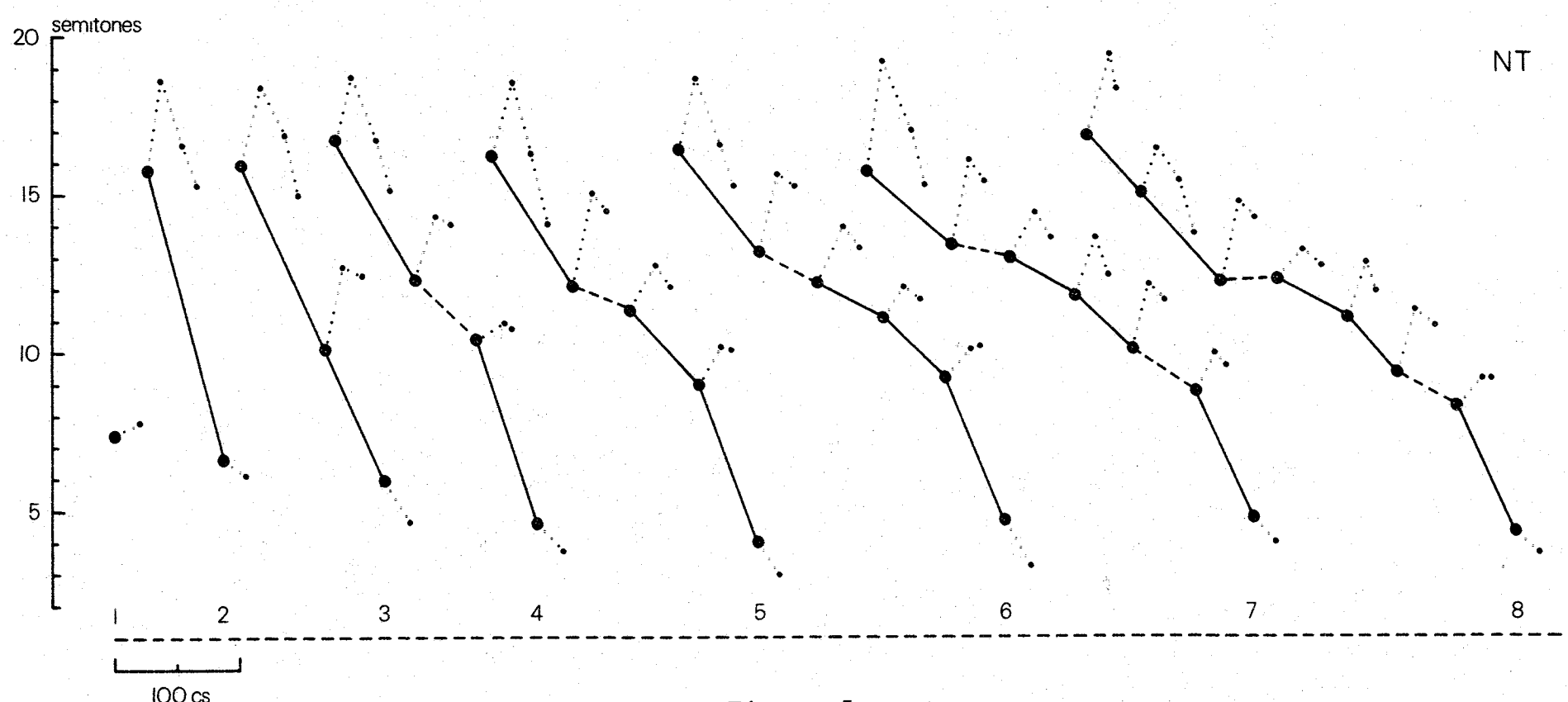

Figure 5

Subject: NT. See further the legend to figure 2.

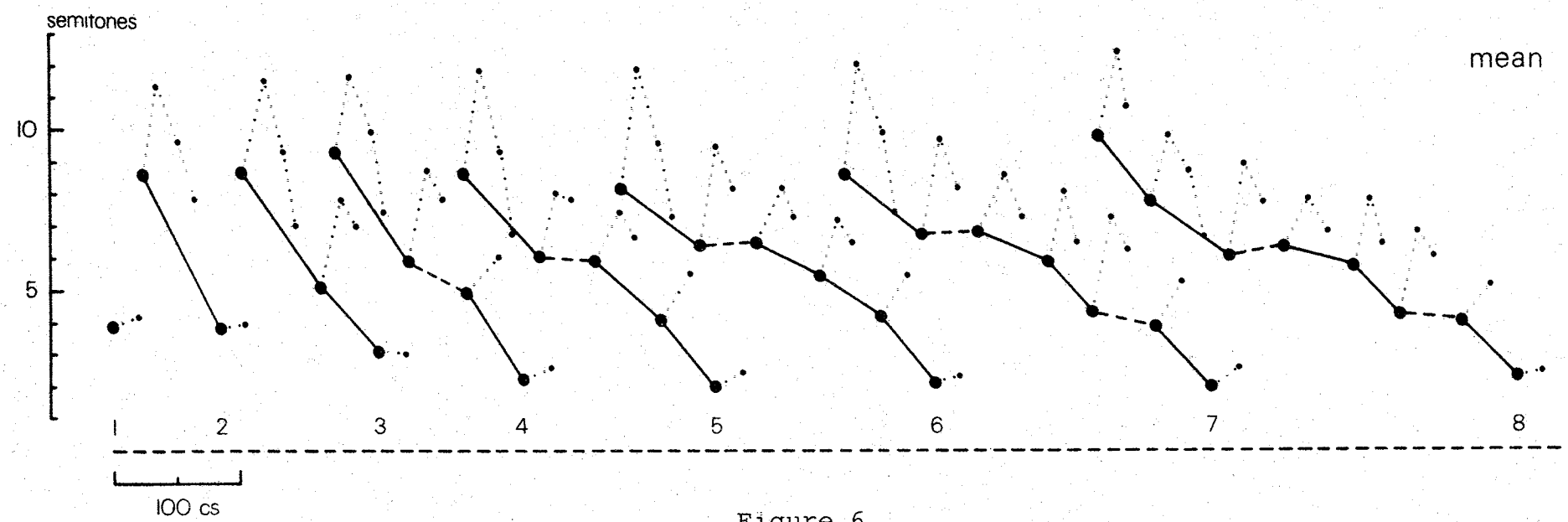

Figure 6

Average over four subjects. See further the legend to figure 2.

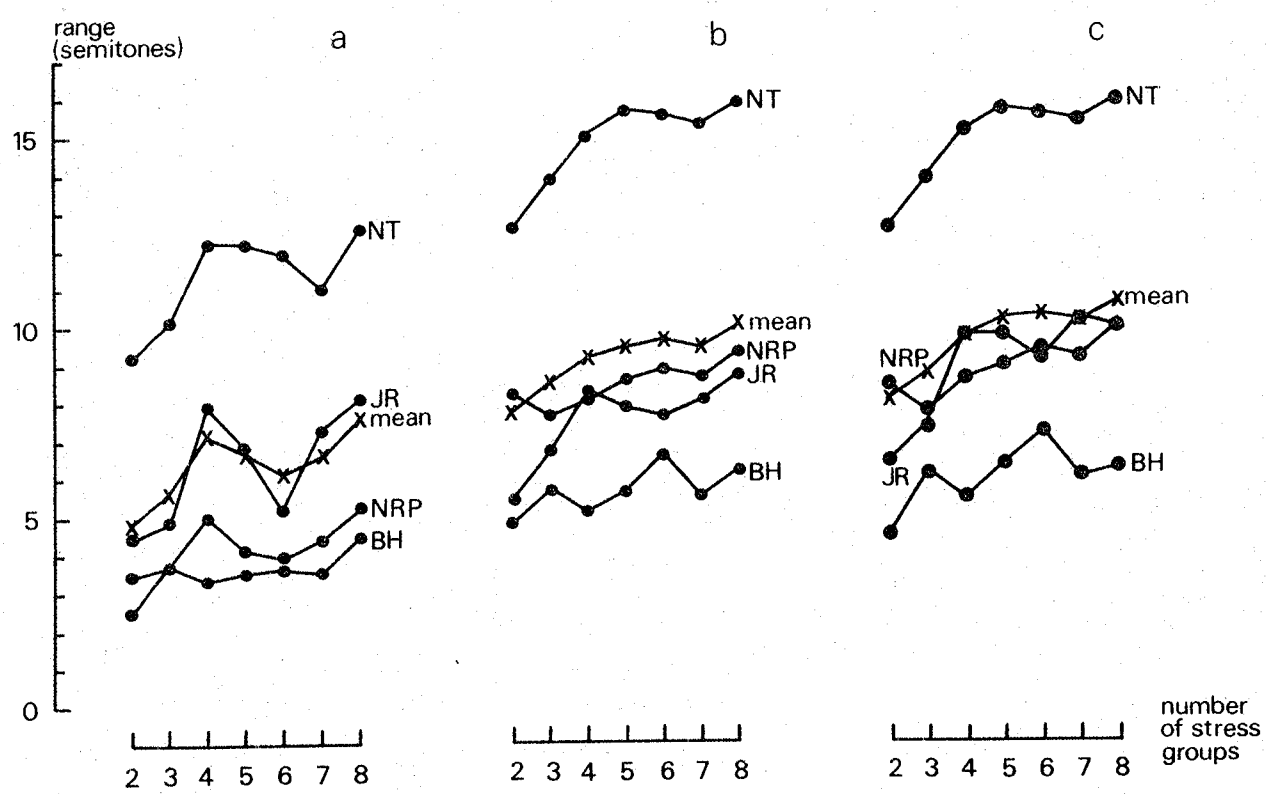

Figure 7

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 (i.e. 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 is identical to the one in (b).) Zero on the logarithmic frequency scale corresponds to 100 Hz.

tial resettings. Range increases, as expected, from 2 to 3; it does not decrease from 3 to 4 (except with BH); it decreases through 4 to 6 (except with BH); it increases from 6 to 7 with NRP, JR, and the mean and decreases (as hypothesized) with BH and NT; it increases, as expected, from 7 to 8.

Concludingly, it seems that range increases with increased number of SGs, but in a non-linear and apparently random fashion, except possibly with BH whose variation is so slight, however, as to be nearly constant (she also has the smallest range of all subjects). - The average dispersion in range is 2.8 semitones, i.e. the largest average range is 2.8 semitones (corresponding to 58%) larger than the smallest one.

3.1.2 Range determined by the post-tonic vowels

Fig. 7b depicts the variation in range as defined by the interval between the first post-tonic vowel in the first and last SGs, respectively. The mean is closer to a straight line than in fig. 7a, which is mainly due to the smoother curve of JR. BH again (and now rather pronouncedly) follows exactly the pattern outlined above with increases from 2 to 3, from 4 through 6, and from 7 to 8, and decreases between 3 and 4, and 6 and 7. Note that the range spanned by the post-tonics is greater than for the stressed vowels, which is a reflection of the fact that the rise from stressed to post-tonic decreases from the earlier to the later parts of the utterance, cf. fig. 1, a phenomenon similar to the commonly noted faster decrease of toplines than baselines, cf. Gårding (1979), Breckenridge and Liberman (1977), and Sorensen and Cooper (1980). Interestingly, this faster decrease is carried by the unstressed syllables in ASC Danish, but by the stressed syllables (the topline) in American English. The average dispersion in range is 2.3 semitones, corresponding to 30% of the smallest one.

3.1.3 Range determined by the absolute Fo maximum and minimum

Fig. 7c depicts the variation in range as defined by the interval between the maximum, i.e. the first post-tonic vowel in the first SG, and the minimum, i.e. the last stressed vowel (except

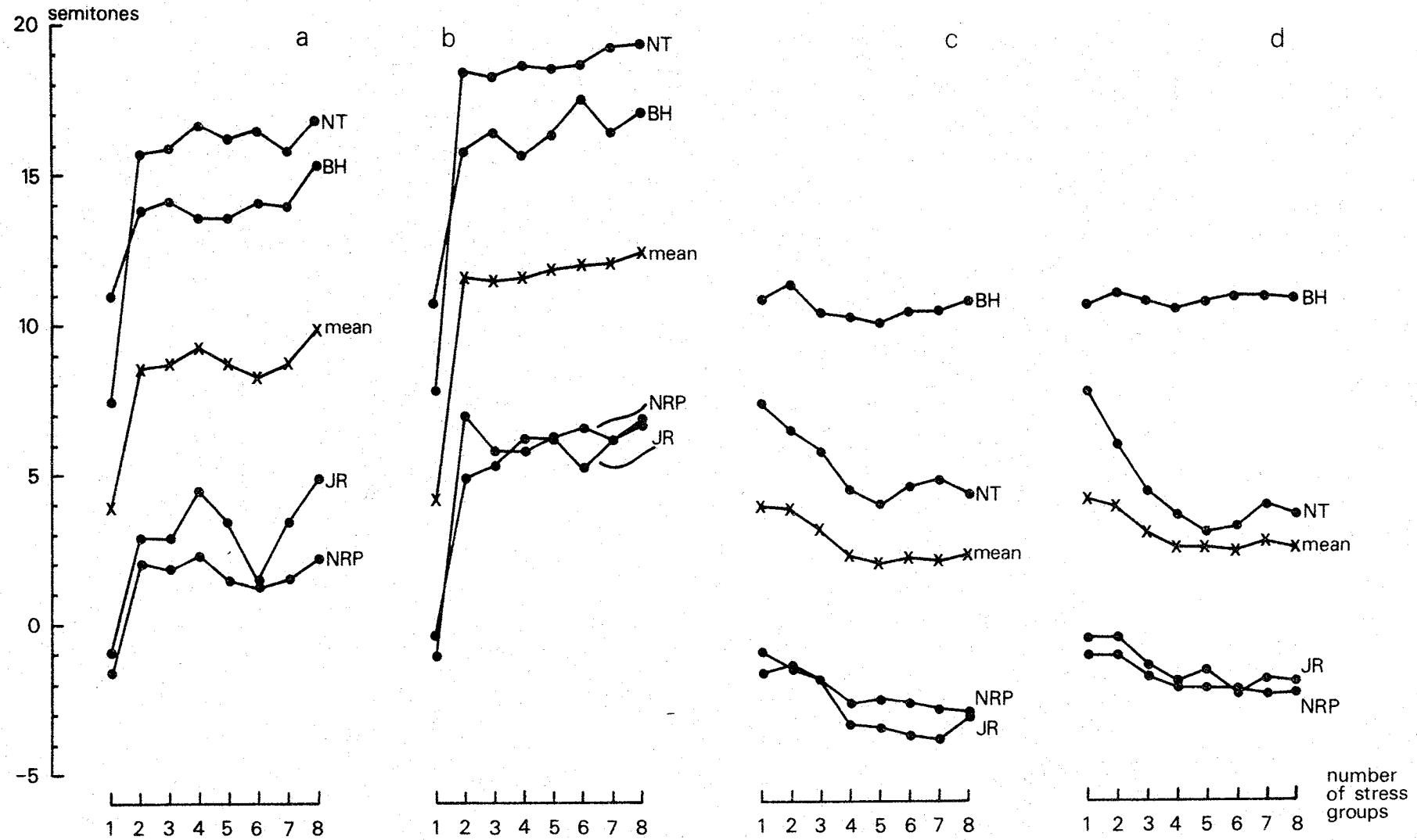

Figure 8

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 (i.e. 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. Zero on the logarithmic frequency scale corresponds to 100 Hz.

that NT's minimum is the last post-tonic vowel, cf. fig. 5, and her curve is identical to the one in fig. 7b). The pattern rather resembles that of fig. 7b, cf. above. - The average dispersion in range is 2.5 semitones, corresponding to 31% of the smallest range.

3.1.4 Starting and end points

The variation in range may be due to variation in starting and/or end points of the contours. Fig. 8a depicts the frequency of the first stressed vowel as a function of number of SGs in the utterance, and fig. 8b shows how the first post-tonic vowel in the first stress group varies with length. If we disregard sentence no. 1 which obviously groups itself with the endpoints (cf. fig. 8c and d) we note a slight trend towards higher starting points with the longer utterances, but the curves are not monotonically rising with higher sentence number and the fluctuation seems to be random, as was the case for range variation (except, possibly, again for BH). Concomitant with higher starting points, we find a (stronger) trend towards lower end points in the longer utterances (fig. 8c and d - stressed and post-tonic vowels, respectively) but there is a tendency for endpoints to stay constant from sentence no. 4/5 and upwards, which is probably a reflection of a physiological constraint.

3.1.5 Range - conclusion

The hypothesized constancy of fundamental frequency range over utterances of varying length is refuted by the data. The largest average range is nearly 60% greater than the smallest one, if range is defined in terms of the stressed vowels, and about 30% if range is determined by the post-tonic vowels or by the absolute Fo maxima and minima.

The increase in range with longer utterances is brought about by a combination of higher starting and lower end points and, contrary to the results of McAllister (1971) and Sorensen and Cooper (1980), up to five stress groups the end points decrease more than the starting points increase.

Table 1

Least squares regression line slopes of the stressed vowels and of the post-tonic vowels depicted in figs. 2-6, and their correlation coefficients.

		Sentence no.	NRP Slope/corr.	JR Slope/corr.	BH Slope/corr.	NT Slope/corr.	mean Slope/corr.
stressed	vowels	2	-5.10	-7.24	-4.50	-15.11	-7.93
		3	-3.31/ -.98	-4.82/ -.98	-3.57/-1.00	-8.96/-1.00	-5.21/-1.00
		4	-2.95/ -.97	-5.61/ -.98	-2.21/ -.96	-7.14/ -.98	-4.53/ -.99
		5	-1.81/ -.97	-3.66/ -.97	-1.66/ -.96	-5.26/ -.97	-3.13/ -.98
		6	-1.31/ -.95	-2.14/ -.94	-1.37/ -.93	-4.00/ -.96	-2.23/ -.96
		7	-1.26/ -.96	-2.66/ -.97	-1.20/ -.97	-3.23/ -.96	-2.10/ -.97
		8	-1.09/ -.89	-2.42/ -.95	-1.16/ -.95	-3.20/ -.97	-1.95/ -.96
unstressed	vowels	2	-10.52	-7.83	-7.27	-18.38	-11.00
		3	-6.43/-1.00	-6.37/-1.00	-5.21/ -.99	-11.76/ -.99	-7.56/ -.99
		4	-4.57/ -.99	-6.03/ -.99	-3.14/ -.95	-8.57/ -.97	-5.64/ -.99
		5	-3.52/ -.99	-4.39/ -.99	-2.55/ -.96	-6.64/ -.96	-4.12/ -.98
		6	-2.95/ -.99	-3.36/ -.99	-2.36/ -.96	-5.25/ -.96	-3.48/ -.98
		7	-2.30/ -.98	-2.91/ -.97	-1.61/ -.93	-4.28/ -.96	-2.82/ -.97
		8	-1.98/ -.95	-2.54/ -.96	-1.49/ -.91	-3.79/ -.96	-2.42/ -.96

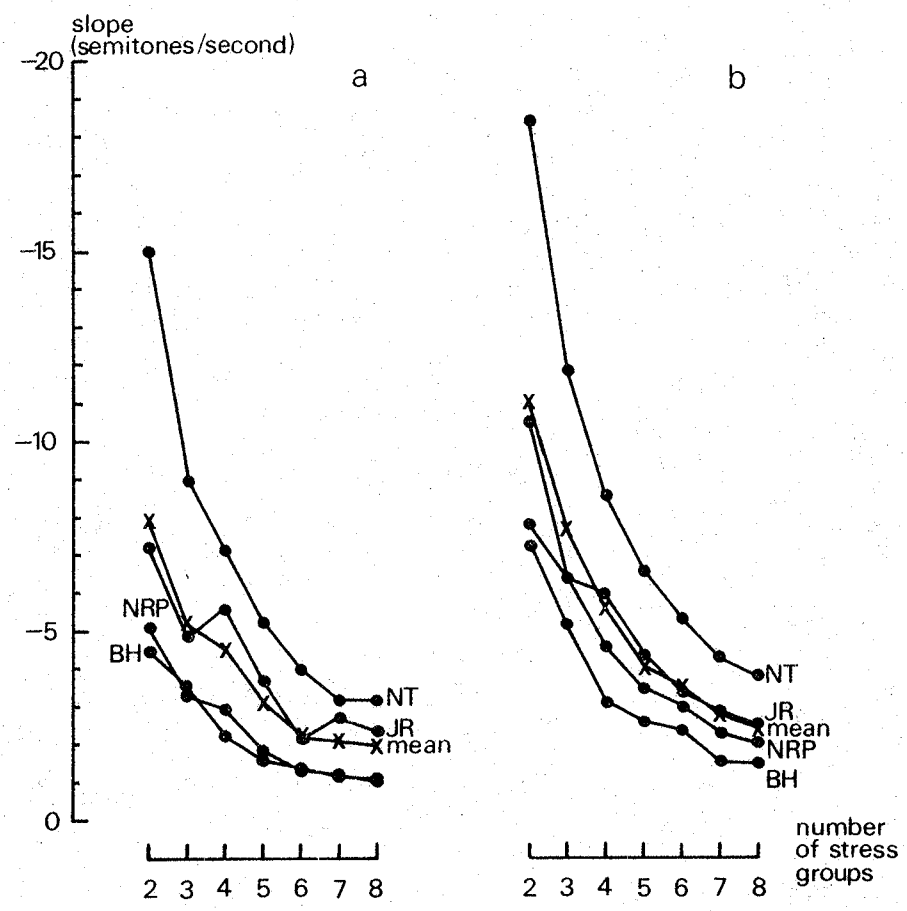

Figure 9

Slope of the overall downdrift of seven utterances, containing from two to eight stress groups, depicted as a function of length of the utterance (i.e. 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 post-tonic vowels (the "topline" - see further the text).

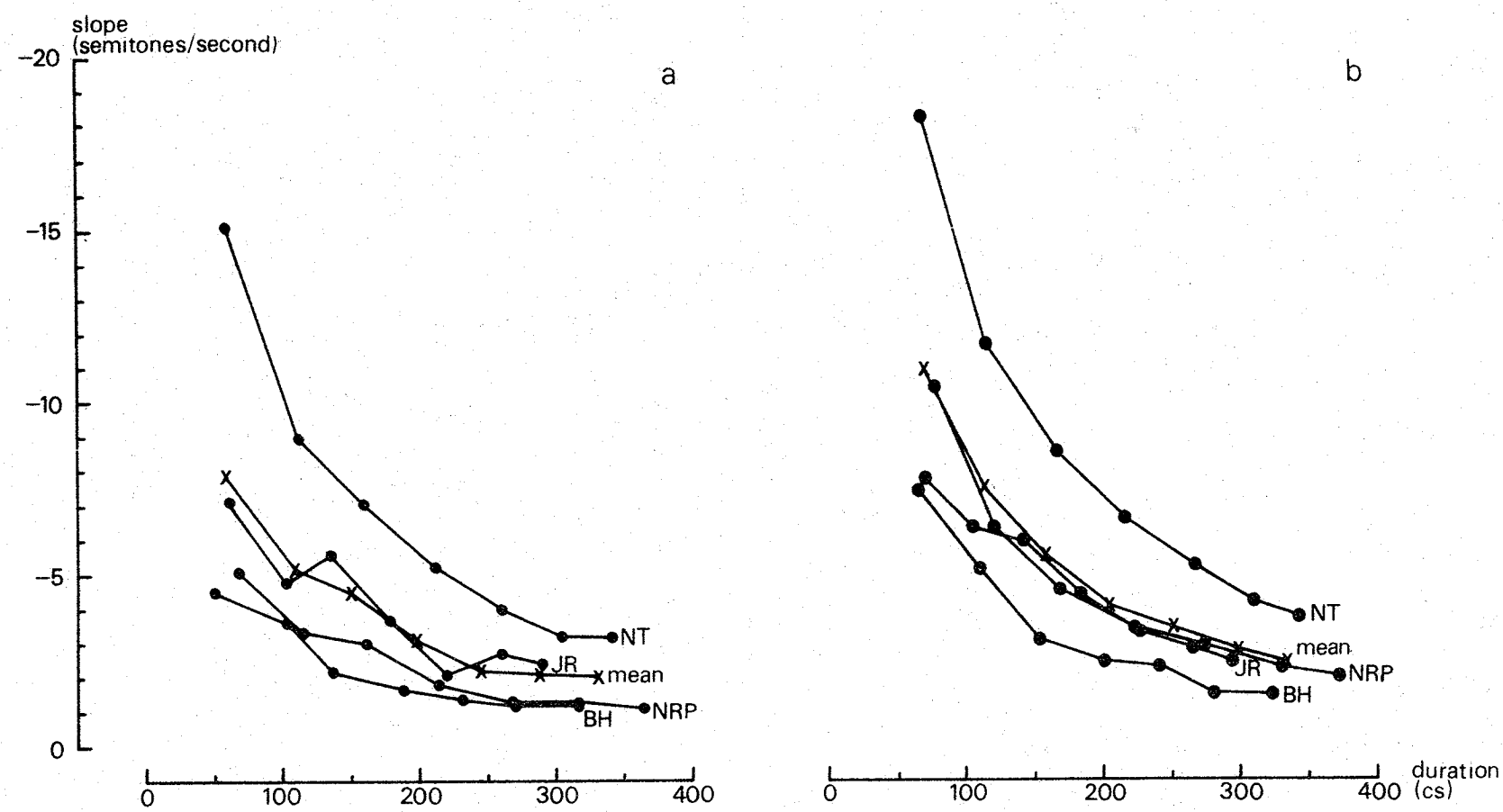

Figure 10

Slope of the overall downdrift of seven utterances, containing from two to eight stress groups, depicted as a function of the duration of the utterance (i.e. the time interval between first and last vowel measurement). 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 post-tonic vowels (the "topline" - see further the text).

4. Intonation contours

4.1 Overall downdrift

It is apparent from figs. 2-6 that the intonation contours are not straight lines, i.e. the medial stressed vowels do not occur with equal semitone intervals, and the irregularity generally sets in with four and more SGs. Nevertheless, when the least squares regression line slopes (which may be taken as an expression of the overall downdrift) are calculated for the stressed vowel data points of figs. 2-6 they come out with correlation coefficients generally at or above .95, cf. table 1, and any further statistical treatment will hardly disclose the regularities that can be observed in the intonation contours any more succinctly than mere visual inspection will do (see further below). Slopes and correlation coefficients have also been calculated for the data points constituted by the first post-tonic vowel in each stress group (the "topline"). These latter slopes turn out (as expected) to be steeper than the corresponding stressed vowel slopes, cf. section 3.1.2.

The information in table 1 is displayed graphically in figs. 9 and 10. In fig. 9 the overall downdrift - stressed vowels (a) as well as post-tonic vowels (b) - is depicted as a function of number of SGs in the utterance, and in fig. 10 as a function of the actual duration, i.e. the time interval between the first and the last vowel measurement. In both instances, but most evidently in fig. 10, overall downdrift is seen to decrease asymptotically with increased length (minor fluctuations occur with JR), approaching a mean value of about -2 and -2.5 semitones per second (stressed and post-tonic vowels, respectively). Presumably, this non-linear decrease in the steepness of the slope of the overall declination is a pertinent feature of longer utterances. A linear decrease would result in zero declination, which is incompatible with declarative sentences. - The preservation of an overall downdrift is secured through a widening of the range. However, there are physiologically determined upper limits to a subjects total range, and theoretically we might envisage declarative utterances that are long enough to jeopardize the downdrift. But I think this problem is purely academic: spontaneous speech rarely contains utterances as long as, say, sentences 7 and 8 that do

not have internal clause boundaries or breath group pauses, which presumably both could lead to a decomposition into what might be termed intonational phrases, each with its own intonation contour (downdrift). Even for read speech, sentences 7 and 8 are rather long and it is conceivable that if subjects were compelled to expand them even further, they would indeed introduce pauses at convenient places.

4.2 Prosodic phrase group boundaries

If we term discontinuities those places in the contours where the slope of the line connecting two stressed vowels is less steep than the preceding as well as succeeding slopes, then sentences 2 and 3 contain no discontinuity, sentences 4, 5, and 6 contain one, and sentences 7 and 8 contain two, for individual subjects as well as the mean, denoted by the broken lines in figs. 2-6. (Whether sentence 6 with NT actually has the discontinuity in the indicated place is debatable - according to the definition just given, it occurs rather between the third and fourth stressed vowels.) Thus, the longer utterances seem to be composed of two and three, respectively, prosodic phrase groups, the boundaries between which coincide with major syntactical boundaries, viz. before the (compound) complement and between the purpose and place complements. Note that these prosodic phrase group boundaries are not accompanied by pauses but seem to be caused by the syntactic structure per se, as pointed out also by e.g. Cooper and Sorensen (1977) and Fujisaki et al. (1979). The degree of resetting of the intonation contour at these boundaries is slight with BH and NT, somewhat larger with NRP and JR where it takes the form of actual rises across the boundary in most instances. (Incidentally, the cross-boundary rises exhibited by NRP and JR are not indicative of a sentence accent or the like: all the utterances were perfectly neutral and contained no perceptual trace of extra prominence anywhere.)

Due to the particular behaviour of the post-tonic syllables (cf. section 5 below) the prosodic phrase group boundaries cannot be detected in the "topline" in the same manner as in the intonation contour proper.

4.3 The shape of the downdrift

Even though straight lines are good approximations to the overall downdrift of the stressed syllables (cf. the high correlation coefficients in table 1), it is also apparent from figs. 2-6 that with some of the utterances by NRP, JR, and BH (but not NT) the trend is towards greater declination in the early part of the utterances: sentences 3-8 by NRP, 3-5 and 8 by JR, and sentence 8 by BH (and sentences 3-5 and 8 of the grand mean). This is in line with the results of McAllister (1971) and Sorensen and Cooper (1980).

The "topline" does not show the same trend, partly because pre-boundary post-tonics exhibit higher rises (cf. section 5 below), partly because the rise in the final stress group is considerably smaller than in preceding stress groups with NRP, BH, and NT (where it actually falls) and thus the steepest "topline" declination is located towards the end of the utterances.

In order to compare the Danish to Fujisaki et al.'s Japanese data, we should look at individual gradients ('voicing components'): the second prosodic phrase group of sentences 5-7 and the first and second one of sentence 8 contain more than two stressed syllables and could thus exhibit an exponential decay. NRP shows a tendency for such a decay in two instances (sentence 5, and the first gradient of no. 8) but quite the reverse in the second prosodic phrase group of sentences 7 and 8. JR has a tendency for asymptotic declination in sentence 5 and the first gradient of sentence 8, but a reversal in 6 and 7, as well as in the second gradient of no. 8. BH exhibits faster decays at the end of the prosodic phrase groups in sentences 5-7 as well as in the second gradient of no. 8, a pattern repeated by NT. Exponentially decaying "toplines" are equally scarce.

4.4 Intonation contours - conclusion

The hypothesized simple inverse relationship between intonation contour declination and utterance length is not supported by the data. At and above 4 stress groups, the intonation contour, as defined by the stressed syllables, is decomposed into prosodic phrase groups, with partial resettings of the contour between them. However, an overall downdrift <u>is</u> preserved, which <u>does</u> become less

steep with increased length (although the relationship is not a linear one), and which, further, exhibits a tendency towards exponential decay, i.e. greater declination in the beginning of the utterance.

5. Stress group patterns

On the basis of the 1978 analysis of ASC Danish, the stress group was defined as a stressed syllable plus all succeeding unstressed syllables, irrespective of intervening syntactic boundaries within the simple (i.e. non-compound) sentence. Thorsen (1980c) corroborated this definition: word boundaries (which were simultaneously noun phrase-verb phrase boundaries) do indeed seem to be immaterial for the Fo patterns of stress groups in ASC Danish. (JR was also a subject for that investigation and a reservation had to be made for more conservative variants of Danish since word boundaries left a clear trace in his utterances.) However, the sentences for the 1978 and 1980c materials were comparatively short, containing three and four SGs, and they exhibited no intonation contour discontinuities. It is conceivable that syntactic boundaries, when they co-occur with prosodic phrase group boundaries, as in sentences 4 through 8 in the present investigation, will break up the regular Fo pattern.

The syntactic boundaries always occur after the first post-tonic syllable in the utterances; accordingly, we might expect the relationship between the first and second post-tonic in SGs before and after phrase boundaries to be different. For instance, the fall from first to second post-tonic could be smaller before the boundary, and thus the second post-tonic could be comparatively higher, in anticipation of the "rise" performed by the succeeding stressed syllable, to which it is affiliated syntactically. Or the fall from first to second post-tonic could be (substantially) larger, bringing the second post-tonic below the level of the succeeding stressed syllable, in imitation of the way sentence initial unstressed syllables behave (cf. fig. 1). - No such differences appear: the fall from first to second post-tonic seems completely unaffected by the discontinuities in the intonation contour and the syntactic boundary per se cannot be made responsible

for any changes in stress group patterns. This is not to say that the phrase group boundary does not affect the stress group pattern: it does - only the variation is not located at the syntactic boundary. Inspection of figs. 2-6 reveal a consistent trend for the rise from stressed to post-tonic (which belong to the same word in all instances) to be greater in pre- than post boundary position. In sentences 5 and 6, and at the first boundary in 7 and 8, the rises can be compared directly, since the post-tonic syllable is carried by a syllabic [n̩] in all cases ('... bússen til fésten ...' - [ˈb̥usn̩ d̥ˢe ˈfɛsd̥n̩]). The average rise is 2.8 semitones before the boundary, as compared to a rise of 1.9 semitones after the boundary, i.e. a difference of nearly 1 semitone. - Now, greater pre- than post boundary rises do not in themselves prove a boundary effect, because progressively decreasing magnitude of the rise to the post-tonic is a feature also of statements without any phrase group boundaries, cf. fig. 1. - Due to differences in the segmental composition of the post-tonic syllable in the stress groups (which entails possible differences in intrinsic Fo levels) part of the following argumentation is qualitative only. First of all, it does seem that a difference in the magnitude of the pre- and post boundary rises of one semitone is rather more than one would expect between neighbouring stress groups in a long statement without any prosodic phrase group boundaries, cf. fig. 1. Secondly, the pre-boundary rises seem to be rather high also in comparison to the <u>preceding</u> rise: In sentence 8, the first prosodic phrase group's second and third stress groups both have syllabic nasals for post-tonics, and the pre-boundary rise is larger than the preceding rise with all subjects (the difference is 0.8 semitones in the grand mean). In the second prosodic phrase group in sentences 7 and 8, the two stress groups under scrutiny both have vowels in their post-tonics, [i] in the pre-boundary stress group, [ʌ] in the preceding one, and although an unstressed [i] may have an intrinsically higher Fo than an unstressed [ʌ], this alone does not seem sufficient to explain the increase we get in the pre-boundary rise (in sentence 8 the pre-boundary rise is 0.5 semitones, in sentence 7 0.8 semitones higher than the preceding one - grand mean). In the first prosodic phrase group in sentences 6 and 7 the two rises are of very nearly the same magnitude. The first post-tonic is carried by a vowel, [a] and [ʌ], respectively, the

second one by a syllabic nasal, and if we assume that an unstressed syllabic nasal has a lower intrinsic Fo than unstressed vowels, then the pre-boundary rise is even more "excessive".

On the whole, it seems safe to conclude that the prosodic phrase group boundary results in a comparatively higher rise to the post-tonic in the preceding stress group. The cause of the greater pre-boundary rise may be sought in either of two processes (or in a combination of them): it is a signal for the prosodic phrase group boundary, and thus controlled by the speaker, or it is an automatic consequence of the higher position of the succeeding stressed syllable, i.e. the higher the following stressed vowel, the less of a fall must be executed by the preceding Fo pattern and, consequently, the higher the rise may be from the preceding stressed vowel. The latter explanation would be identical to, and could be taken as further support for, the explanation offered previously (Thorsen 1979b, 1980a) for the fact that the magnitude of the rise from stressed to post-tonic vowel varies with intonation contour, i.e. we get higher rises on less steep contours, cf. fig. 1.

Note that in this investigation JR does not deviate from the three ASC speakers as far as syntactic boundaries and fundamental frequency are concerned. His stress group pattern deviates from those of the other subjects by having more steeply falling unstressed syllables, but there is no trace of a syntactic boundary signalling in his traces. - I am inclined to think, now, that the word boundary signalling he performed in the previous (1979c) experiments does not constitute an example of a difference between ASC and more conservative variants of Danish. Rather, it demonstrates that it is possible for a speaker to signal word boundaries, also with fundamental frequency, if he so desires, a possibility which presumably is also open to ASC speakers.

In summary, if the results of the present investigation can be extended to cover simple sentences in general, they present an argument in favour of a theory expounded in Selkirk (forthcoming) that prosodic categories (in casu: stress groups and prosodic phrase groups) are distinct entities in the phonology that do not have an isomorphous relation to syntactic structure. Rischel (1972) argues in a similar fashion: Danish stress is best represented in a hierarchy (a tree structure) which is not necessarily congruent with the syntactic structure. - The autonomy of prosodic struc-

ture does not, of course, deprive it of a relation to syntax (cf. the questions posed below), on the contrary, prosodic categories can be seen as reconciling the syntactic structure to the phonetic output (in casu: the course of fundamental frequency).

6. Discussion

The purpose of the investigation was not to investigate the interplay between syntax and prosodic structure as such. Nevertheless, the tendencies that emerge raise some interesting questions concerning the hierarchy and domain of syntactic boundaries vs. the inherent features of declarative intonation. In this material, the syntactic boundary before the (compound) complement seems to be more manifest than the noun phrase/verb phrase boundary. The constituent which varied most in number of stress groups was the complement: What would the contours have looked like if instead the noun phrase and/or the verb phrase had varied? E.g. is the tendency towards a faster declination early in the utterance an inherent feature of declarative intonation or is it an artefact of the material that would disappear if the noun phrase or verb phrase were longer? With a short complement but a long noun phrase, would a prosodic phrase group boundary be introduced after the noun phrase, and would the verb phrase and complement merge into one prosodic phrase group? If the second of the two complements had consisted of only one stress group, would it have had to merge prosodically with the preceding complement in order to preserve a final declination? And where would the first complement be cut up then, if four stress groups are the maximum in a prosodic phrase group (compare sentence 6 and 7)? Or is the final declination dispensable as long as there is an overall downdrift in the utterance? How do clause boundaries manifest themselves? When unaccompanied by pauses, will they exhibit greater amounts of resetting than do clause internal boundaries? And will they affect stress group patterns?

Acknowledgement

Sincere thanks are due to Jeanette Holtse for her patient and competent graphics work.

References

Breckenridge, J. and M.Y. Liberman 1977: "The declination effect in perception" (unpublished manuscript)

Bruce, G. 1979: "Word prosody and sentence prosody in Swedish", Proc.Phon. 9, p. 388-394

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

Fujisaki, H., K. Hirose, and K. Ohta 1979: "Acoustic features of the fundamental frequency contours of declarative sentences in Japanese", Annual Bulletin, Research Institute of Logopedics and Phoniatrics 13, p. 163-173

Gårding, E. 1979: "Sentence intonation in Swedish", Phonetica 36, p. 207-215

't Hart, J. 1979: "Exploration in automatic stylization of Fo curves", IPO APR 14, p. 61-65

't Hart, J. and R. Collier 1975: "Integrating different levels of intonation analysis", JPh 3, p. 235-255

Hirose, H. 1971 "On Japanese accent: An acoustic analysis", Monthly Internal Memorandum, Phonology Laboratory, Berkeley, May, p. 16-40 (cited from Ohala 1978)

Maeda, S. 1977: "Sur les corrélatifs physiologiques de la fréquence du fondamental", Actes des 8emes Journées du Groupe de la 'Communication Parlée', Aix-en-Provence, 25-27 Mai, p. 43-50

McAllister, R. 1971: "Predicting physical aspects of English stress", STL-QPSR 1, p. 20-29

Ohala, J. 1978: "Production of tone", Tone: A Linguistic Survey (ed.: V. Fromkin), Academic Press, New York, p. 5-39

Pierrehumbert, J. 1979: "The perception of fundamental frequency declination", JASA 66, p. 363-369

Reinholt Petersen, N. 1978: "Intrinsic fundamental frequency of Danish vowels", JPh 6, p. 177-189

Reinholt Petersen, N. 1979: "Variation in inherent Fo level differences between vowels as a function of position in the utterance and in the stress group", ARIPUC 13, p. 27-57

Rischel, J. 1972: "Compound stress in Danish without a cycle", ARIPUC 6, p. 211-229

Rossi, M. 1971: "Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole", Phonetica 23, p. 1-33

Rossi, M. 1978: "La perception des glissandos descendants dans les contours prosodiques", Phonetica 35, p. 11-40

Selkirk, E.O., (forthcoming): "On the nature of phonological representation", The Cognitive Representation of Speech (eds.: J. Anderson, J. Laver, and T. Myers) North Holland Publishing Company

Silverstein, R.O. 1976: "A strategy for utterance production in Hausa", Studies in African Linguistics, Suppl. 6, p. 233-241 (cited from Ohala 1978)

Sorensen, J.M. and W.E. Cooper 1980: "Syntactic coding of fundamental frequency in speech production", Perception and Production of Fluent Speech (ed.: R.A. Cole), Lawrence Erlbaum Associates, Hillsdale, New Jersey, p. 399-440

Sternberg, S., C.E. Wright, R.L. Knoll, and S. Monsell 1980: "Motor programs in rapid speech: Additional evidence", Perception and Production of Fluent Speech (ed.: R.A. Cole), Lawrence Erlbaum Associates, Hillsdale, New Jersey, p. 507-534

Thorsen, N. 1978: "An acoustical investigation of Danish intonation", JPh 6, p. 151-175

Thorsen, N. 1979a: "Interpreting raw fundamental frequency tracings of Danish", Phonetica 36, p. 57-78

Thorsen, N. 1979b: "Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish", Proc.Phon. 9, p. 417-423

Thorsen, N. 1980a: "A study of the perception of intonation contours - Evidence from Danish", JASA 67, p. 1014-1030

Thorsen, N. 1980b: "Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish", ARIPUC 14 (this volume)

Thorsen, N. 1980c: "Word boundaries and Fo patterns in Advanced Standard Copenhagen Danish", Phonetica 37, p. 121-130 (also in ARIPUC 13, 1979, p. 121-134)

Weitzman, R.S. 1970: "Word accent in Japanese", Studies in the Phonology of Asian Languages, University of Southern California, 11 (cited from Ohala 1978).