DEVELOPING THE SWEDISH INTONATION MODEL

Gösta Bruce

ABSTRACT

A series of experiments were conducted to test an earlier model for Swedish intonation where local specifications for accentuation were inserted in a global baseline-topline structure. The course of the Fo downdrift is found to be stepwise at the accents rather than gradual over the utterance. The speaker signals utterance length by raising the Fo peak and the Fo values of succeeding accents in proportion to the number of upcoming accents, rather than by varying the slope inversely with utterance length. The involvement of the speaker is expressed as an increase in the overall Fo range with a preservation of the Fo relations between the accents within the utterance. The Fo bottom of the speaker's range appears to be fixed, while the Fo maximum is variable. These findings indicate that a more adequate account of the relationship between accentuation and intonation should give a more explicit role to their interaction than is achieved by merely adding accentuation to intonation. It is suggested that the overall Fo course of an utterance is essentially formed by the relations between successive, local excursions for accentuation.

0 INTRODUCTION

In a model for Swedish intonation (Bruce & Gårding 1978, Gårding & Bruce 1981) we have shown how a phonetic representation of an utterance in terms of an Fo contour can be generated from a phonological representation using only a few linguistically relevant parameters.

The aim of the present study is to examine some of the predictions and implications of the intonation model, and also to try to extend the predictive power of the model to new kinds of linguistic variation. I will concentrate on the problem of the interaction of accentuation and intonation proper, and in particular how accents are assigned Fo values in different contexts. This will lead us into questions of speech planning and ultimately into a discussion of the consequences for the intonation model itself.

The first section of the paper gives a brief review of the intonation model for Swedish. Then a series of experiments, designed to test the model, is reported and discussed. The first of these experiments examines the nature of the Fo downdrift that characterizes a declarative type of intonation. The next two experiments are devoted to the effect of utterance length on Fo. Utterance length is varied by a syntactic expansion either only to the right (experiment two), or both to the right and to the left (experiment three). In experiment four the attitude of the speaker - two degrees of involvement - is also varied. In experiment five, finally, the effect of introducing a clause boundary in a sentence is examined. In the final section problems of Fo scaling are discussed and the results of the experiments are interpreted in terms of a revised model for Swedish intonation.

1 BRIEF REVIEW OF THE INTONATION MODEL

The point of departure for the present study is the Lund model for Swedish intonation. The model was developed to account for the intonation of different Swedish dialects, but it can be used to study the intonation of other languages as well. Recently it has been applied to contrastive intonation studies and studies of prosodic interference between a primary and a secondary language (Gårding 1982).

In the present paper I will concentrate on Standard Swedish, which is the dialect spoken in e.g. Stockholm. Figure 1 illustrates the model for this variety of Swedish.

The input to the model is a phrase or a sentence represented as a string of segments and their relative durations. To this representation is attached a prosodic transcription with information about stress location and word accent (accent 1 or accent 2), placement of sentence accent and sentence type (statement or question). An example of two input phrases - a minimal pair - is given in orthographic form in Figure 1.

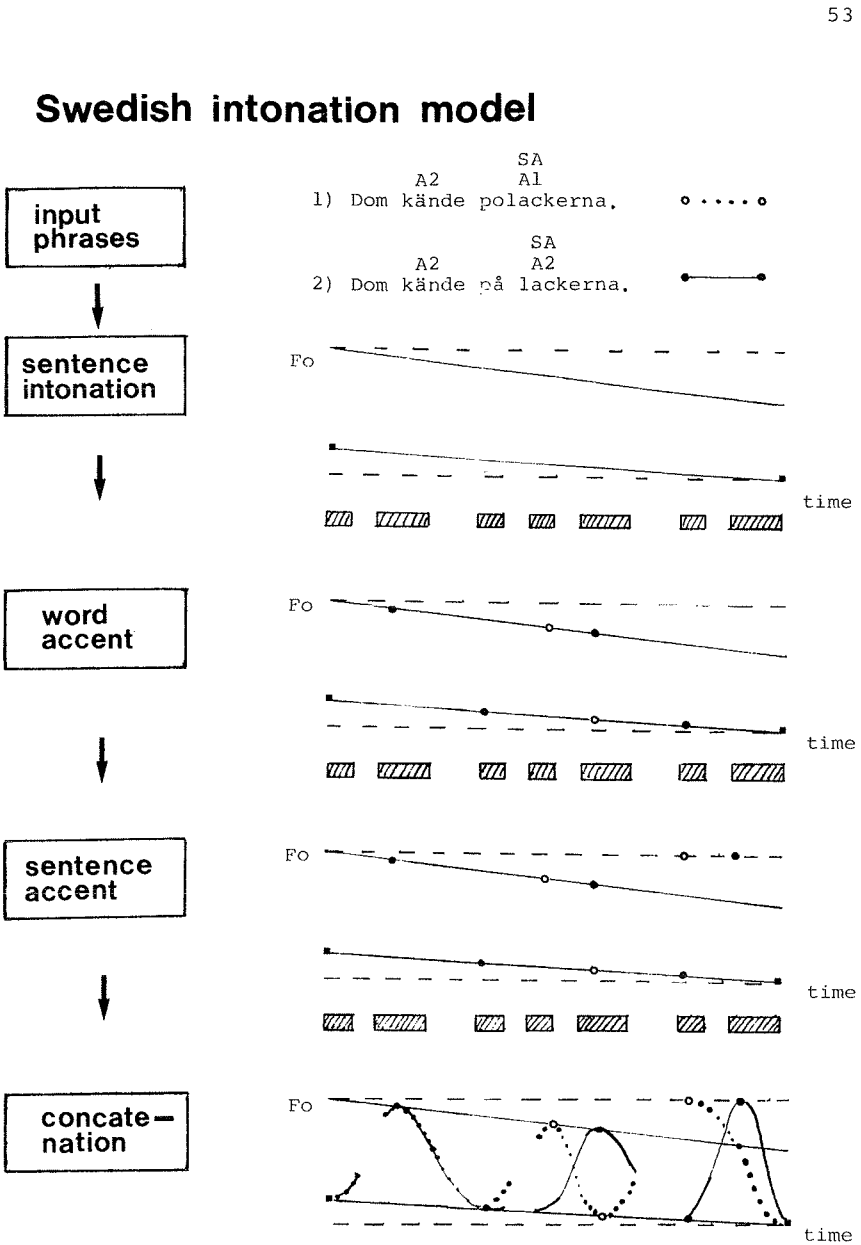


Figure 1. Swedish intonation model. Application of pitch rules for Standard Swedish (after Bruce & Gårding 1978).

The basic assumption underlying the model is that intonation proper can be separated from accentuation and that accentuation can be divided into a word accentuation part and a phrase or sentence accentuation part.

In model terms this is expressed in the following way. The first step is to generate sentence intonation, which takes the form of a baseline - topline structure consisting of four reference lines to give the frame for accentuation. The word accents are represented as combinations of high and low points with a different timing for the two word accents (A1/A2) as the distinctive feature. The highs and lows of the word accents are inserted on the interior topline and baseline respectively. Sentence accent (SA) is represented as a high following the high-low of the word accent in focus. It is inserted on the exterior topline. The exterior baseline is not relevant for this dialect.

For a declarative intonation type as illustrated in Figure 1 the interior lines fall and converge, while the exterior lines remain more or less level. This global fall or downdrift gives successively lower Fo values to the word accent highs and lows, while the Fo value of the sentence accent high is independent of the position in the utterance.

For question intonation (Gårding 1979), the downdrift is largely suspended, and the exterior, focal topline is lifted, consequently widening the total Fo range.

The final operation is to connect neighbouring points to form the output Fo contour in which the highs and lows of the accents become turning points. A simple truncation of the Fo contour has been applied to voiceless consonants, while other known segmental effects on the Fo contour have been ignored here. This concatenation procedure also involves adjustments due to time restrictions on the execution of tonal commands (see Bruce 1977, chapter 5) such as anticipation and undershooting of highs and lows, which have been excluded from Figure 1.

Various similar representations of the relationship between intonation proper and accentuation, where the accents are principally superposed as local humps on the global, successive downdrift of intonation, have been suggested by a number of authors (see e.g. Thorsen 1978 for an overview). This approach has been termed the "layered" theory of intonation, as the Fo contour is assumed to be built up by superposition of layers (Pierrehumbert 1980, p. 208).

In this paper the terms 'sentence accent' and 'phrase accent' will be used synonymously. We have used 'sentence accent' in the model reviewed here, which was based on test material consisting of relatively short sentences with only one occurrence of this kind of accent per sentence. In an extended material with longer sentences, where a division into prosodic phrases seems to play an important role, two or more instances of this kind of accent occur with one such accent per phrase. Replacing the term 'sentence accent' with 'phrase accent' in later sections of the paper should be seen in this light.

2 EXPERIMENT ONE: THE NATURE OF Fo DOWNDRIFT

The first experiment was designed to examine some of the predictions and implications of the model concerning the nature of the Fo downdrift in a declarative intonation type.

In particular, the following prediction and some of its implications were tested: the Fo downdrift has a gradually decreasing course. It follows from this prediction that the Fo value of the first accent of an utterance is lower, if it is preceded by unaccented syllables than if it is not, and also that Fo gradually falls on unaccented syllables between two accents. It also implies that the very bottom level of Fo is not reached until the last syllable of the utterance.

A further implication is that the downdrift before and after focus will have the same course, i.e. placement of focus has no effect on the course of the Fo downdrift. The model also predicts that word accents are executed as rise-falls both before and after focus. As the starting points and end points of the interior topline and baseline are taken to be constant for a given intonation, the slopes of these lines vary inversely with utterance length. According to the model this means that Fo will fall more abruptly in a shorter utterance compared to a longer one.

In the material that was designed to test these predictions the following parameters are systematically varied: number of stress groups (2, 4), placement of focus, and the number of unstressed syllables before (0, 2), between (1, 3) and after (1, 3) the stresses. A stress group contains a stressed syllable and all succeeding unstressed syllables (if any) within an utterance.

Test material - experiment one:

ˇUngen ˇnallar
The kid is stealing

ˇMamman ˇlämnar honom / (Ja) ˇmammorna hann ˇlämna honom
The mother leaves him / (Well) the mothers managed to leave him

Man hann ˇlämna honom ˇungen / (Ja) man hann ˇlämna nåra ˇnallar
One managed to leave him the kid / (Well) one managed to leave some teddy bears

ˇMamman ˇlämnar ˇungen ˇnallar
The mother leaves the kid teddy bears

(Ja) ˇmammorna hann ˇlämna våran ˇunge nåra ˇnallar
(Well) the mothers managed to leave our kid some teddy bears

The phonetic composition of the material is designed to minimize segmental effects on the Fo contour. Differences in intrinsic Fo of vowels and obstruent consonants have been avoided. Each stressed syllable carries accent 2. My own earlier investigations (Bruce 1977) have shown that the phonetic difference between accent 1 and accent 2 is basically one of Fo timing.

Restricting the present study to only one accent is therefore justified, as the word accent distinction is not relevant here and the results for one accent can be extrapolated to the other accent.

The test material contains meaningful Swedish sentences. Each sentence has the form of an answer to a question. Various focus assignments in the answer (response sentence) are elicited by changing the question (context sentence). A neutral version of the answer was not included in this experiment, as earlier investigations (Bruce 1977, section 3.3) have shown it is similar to the final focus version.

Two informants representing the Stockholm dialect were used. A male informant (TB) recorded the whole material seven times. A female informant (UN) recorded a subset of the material ten times - the sentence with four stress groups and three unstressed syllables between the stresses plus sentences with two stress groups of a slightly different composition. Data from the main informant (EH) from my thesis (Bruce 1977) - female, Stockholm dialect - has been used for comparison.

The recordings were processed by hardware pitch and intensity meters. Fo measurements were made by hand.

Figures 2 and 3 - with one and three unstressed syllables respectively between the stressed syllables - show repeated Fo contours for the male informant. Table I gives the mean values and standard deviations of successive Fo maxima and Fo minima of of the test sentences for both the male and the female informant. For a comparison with the corresponding Fo values from the main informant in my thesis see (Bruce 1977: chapts. 4, 5).

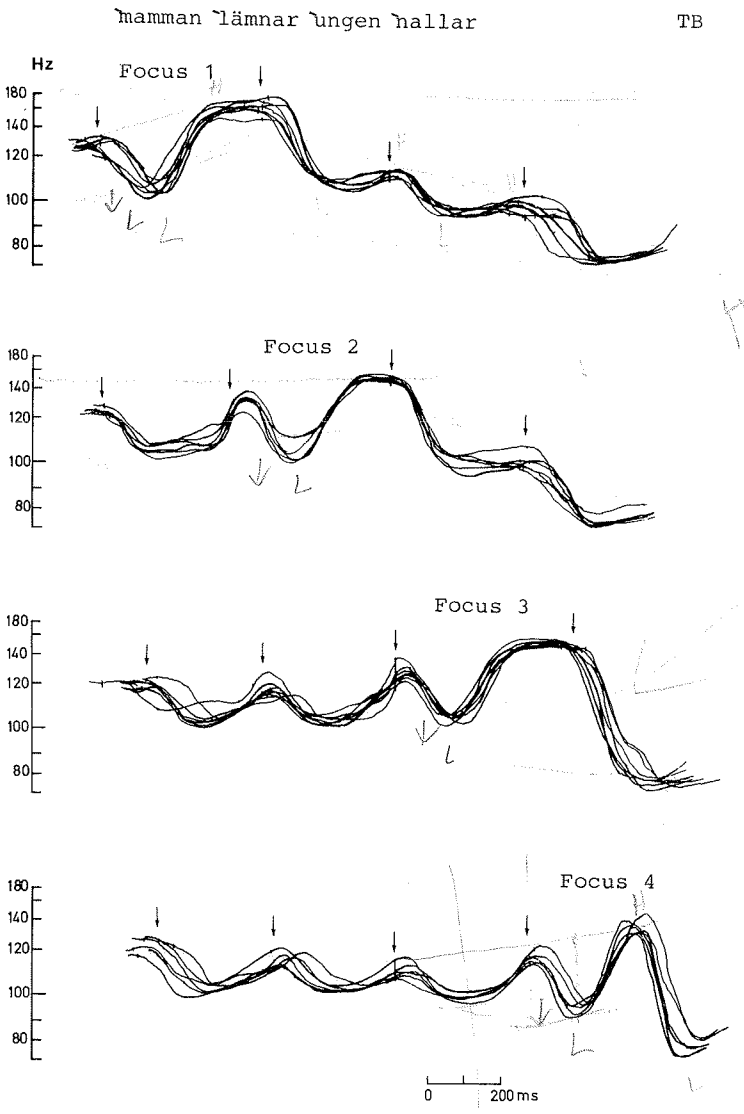


Figure 2. The nature of Fo downdrift. The effect of placement of focus in a sentence containing four stress groups with one unstressed syllable between the stressed syllables. Repeated Fo contours by a male informant (TB). The arrows indicate the stress group boundaries. The line-up point is at the third stress group boundary.

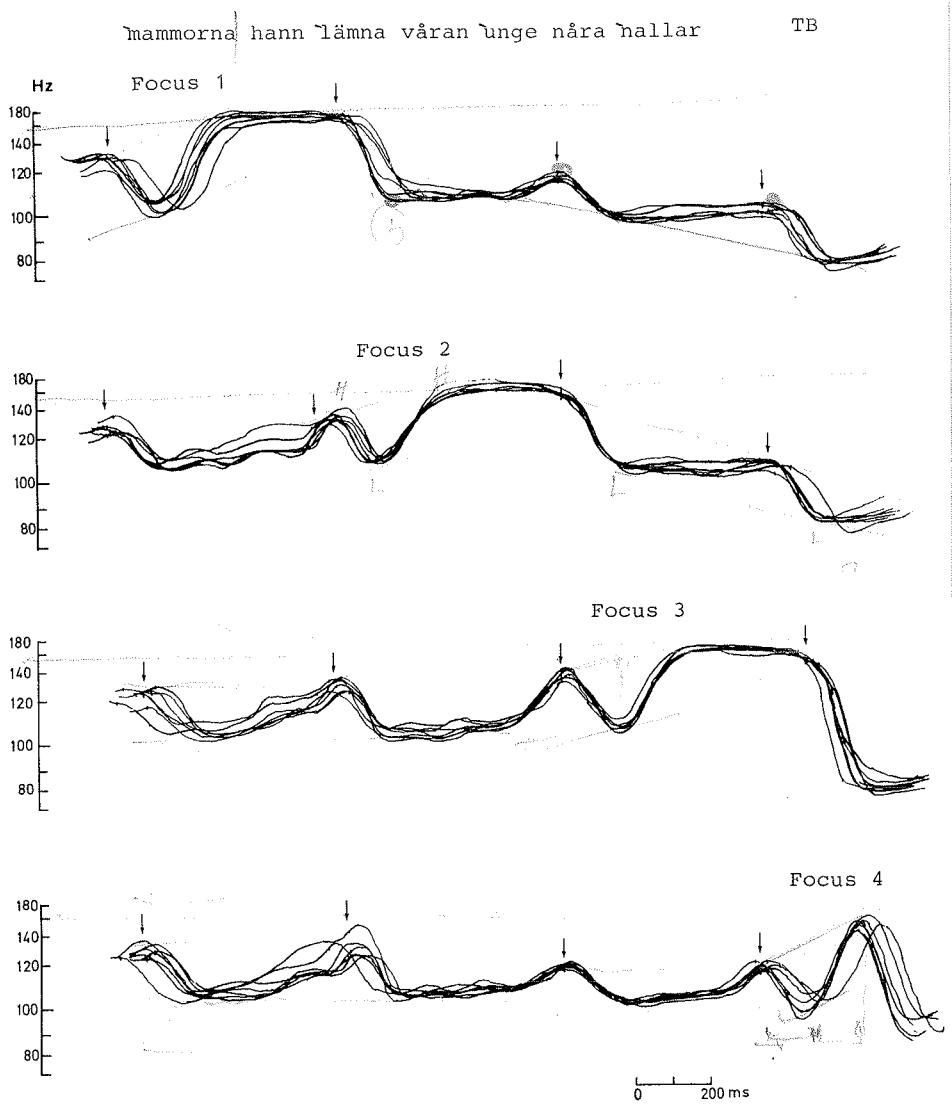


Figure 3. The nature of Fo downdrift. The effect of placement of focus in a sentence containing four stress groups with three unstressed syllables between the stressed syllables. Repeated Fo contours by a male informant (TB). The arrows indicate the stress group boundaries. The line-up point is at the third stress group boundary.

2.1 Results of experiment one

The main results of experiment one for informant TB are summarized below. The following findings agree with the predictions of the model:

1. The starting points and end points appear to be constant and independent of utterance length. When an utterance starts with an accented syllable, the Fo onset is naturally higher, though, than when the initial syllable is unaccented. There is a tendency for the end point value to be slightly higher in utterances with final focus, which is ascribed to undershooting.

2. The Fo values of succesive maxima and minima are similar in the two versions of the four stress group utterance.

3. Focal Fo maxima have the same frequency in all positions except the final one. Here it is lower, however, even if we take into consideration the possibility of undershooting in this position. When three instead of two syllables follow the accented syllable the focal maximum of the final position is raised but it is still lower than non-final focal maxima.

4. The very bottom Fo level appears to be reached only on the last syllable of an utterance. So the only true Fo downdrift in this material is found after the last accent, when several unaccented syllables follow.

The following findings were not predicted by the model:

5. The downdrift appears to have a stepwise rather than continuous course. The stepwise declination of Fo becomes more evident when there are unaccented syllables between the accents as in Figure 3, but it is less apparent from Figure 2 where the accents are close together.

6. The actual course of the Fo downdrift appears to be dependent on factors such as the location of phrase accent (focus) and of the word accents.

7. The pivot of the declination is the focus of an utterance. Before focus the declination is absent or gentle; after focus Fo decreases more rapidly. More specifically, when focus is non-final, pre-focal declination is practically absent. But when focus is final, there is a gentle declination. In contrast to the prediction of the model, the Fo values of pre-focal maxima and minima do not decrease faster in a two stress group utterance as compared to a four stress group utterance.

8. The step down occurs in connection with the word accents. In unaccented syllables between two accents there is no downward slope but instead an Fo plateau. This is true for both baseline and topline.

9. The Fo values of the first accent of an utterance appear to be the same regardless of its position in relation to the onset of the utterance. When unaccented syllables precede, there is no Fo downdrift up to the first accent.

10. The Fo range of the rise of a post focal accent is narrower than that of a pre-focal accent. So there is a tendency for the word accents before focus to be executed as rise-falls and after focus as simple falls. Accent maxima immediately preceding a phrase accent (focal) maximum tend to have higher values than other pre-focal maxima.

The prosodic behaviour of the second informant (UN) is very similar as far as the material is comparable. Points 1, 3-6, 8 and 9 are supported, while point 10 is partly supported; the tendency for post-focal accents to be simple falls is not typical of this informant. Point 7 is not fully comparable owing to her accentuation. When focus is non-initial, she manifests two phrase accents - one for the first stress group and one for the focus, except in one of the sentences containing two stress groups.

Table I. The nature of Fo downdrift. Mean Fo and standard deviations (Hz) of successive minima and maxima of the test material recorded in experiment one: (a) male informant TB, (b) female informant UN. The focal (phrase accent) maximum of each utterance is in italics.

(a) Speaker TB n = 7

		min	max	min	max	min	max	min	max	min	max	min
`ungen `nallar												
initial	$\bar{x}$		135	104	*162*	80						
focus	s		3,0	4,0	7,6	3,7						
final	$\bar{x}$		134	104	126	98	*144*	87				
focus	s		2,4	2,0	6,3	4,6	5,7	9,6				
`mamman `lämnar honom												
initial	$\bar{x}$		129	103	*165*	80						
focus	s		2,8	2,7	5,8	4,4						
final	$\bar{x}$		126	105	131	100	*155*	84				
focus	s		6,1	3,0	2,4	3,0	11,2	7,8				
man hann `lämna honom `ungen												
initial	$\bar{x}$	111	136	105	*169*	79						
focus	s	3,8	3,5	2,9	5,3	5,2						
final	$\bar{x}$	107	128	103	126	102	*149*	82				
focus	s	3,7	5,5	1,9	4,0	3,7	3,5	4,5				
`mamman `lämnar `ungen `nallar												
initial	$\bar{x}$		127	103	*164*	105	113	93	99	76		
focus	s		2,7	3,2	10,2	2,7	2,0	1,2	4,0	3,5		
2nd	$\bar{x}$		124	104	131	103	*165*	100	102	78		
focus	s		4,8	2,8	5,6	4,8	8,2	3,4	4,7	3,4		
3rd	$\bar{x}$		123	104	120	104	131	105	*165*	78		
focus	s		3,7	2,4	6,4	1,9	5,0	3,4	2,9	5,1		
final	$\bar{x}$		119	103	117	103	112	98	118	94	*143*	78
focus	s		6,4	3,8	6,4	2,0	4,4	2,3	6,4	2,8	10,3	4,9
`mammorna hann `lämna våran `unge nåra `nallar												
initial	$\bar{x}$		129	101	*179*	105	115	95	99	77		
focus	s		6,3	4,3	8,0	2,7	1,7	2,5	3,5	3,9		
2nd	$\bar{x}$		126	105	133	106	*167*	101	105	75		
focus	s		7,5	2,0	5,0	2,8	5,7	3,5	2,5	2,9		
3rd	$\bar{x}$		126	104	132	103	137	105	*169*	80		
focus	s		6,1	2,8	6,1	2,5	7,9	1,7	8,9	3,9		
final	$\bar{x}$		128	105	132	104	119	98	115	95	*144*	86
focus	s		5,9	2,5	8,7	2,0	2,4	1,9	2,0	4,1	7,5	6,4

(b) Speaker UN n = 10

		min	max	min	max	min	max	min	max	min	max	min
ja `mammorma hann lämna honom												
initial focus	$\bar{x}$	189	224	169	*297*	147						
	s	5,2	10,3	6,3	14,6	2,4						
final focus	$\bar{x}$	187	209	166	230	164	*275*	158				
	s	4,3	6,8	2,2	16,8	6,0	8,3	5,6				
ja man hann `lämna nåra `nallar												
initial focus	$\bar{x}$	191	215	170	*294*	146						
	s	6,2	12,7	5,5	18,6	3,7						
final focus	$\bar{x}$	189	202	163	189	153	*256*	167				
	s	4,9	10,3	5,1	7,4	5,7	13,9	7,6				
ja `mammorna hann `lämna våran `unge nåra `nallar												
initial focus	$\bar{x}$	191	236	175	*314*	169	180	153	169	143		
	s	6,9	12,9	6,0	8,4	4,6	11,8	3,5	5,8	2,6		
2nd focus	$\bar{x}$	184	215	169	257	178	*308*	156	167	144		
	s	5,7	9,1	3,9	10,3	6,8	8,9	3,7	7,9	2,4		
3rd focus	$\bar{x}$	189	219	170	267	177	220	171	*306*	144		
	s	5,2	7,8	7,2	7,1	8,2	15,6	5,5	9,9	3,2		
final focus	$\bar{x}$	186	218	171	269	191	205	168	192	158	*265*	149
	s	5,5	6,4	5,0	7,1	5,0	8,6	2,6	6,7	6,3	9,7	8,5

Apparently, it dependes on the composition of the sentence, whether she manifests one or two phrase accents. This is evident from a comparison of the two sentences containing two stress groups. If the first stress group is an NP, which is also the topic of the utterance, it is attributed a phrase accent. If, on the other hand, the first stress group belongs to the VP, which is also the comment of the utterance, it will have no phrase accent. When focus is initial, only the first stress group of an utterance will receive a phrase accent. Therefore the declination before focus is not comparable to that of the first informant. Point 2 is not testable owing to the composition of the test material recorded by the second informant.

The main informant in my thesis (Bruce 1977) also conforms very well with TB and UN. This is true for points 3, 5-8 and 10. Point 4 is not supported; in post-focal words in final position, when the accented syllable is an antepenult (or even earlier), the bottom Fo level may be reached before the utterance final syllable. Points 1, 2 and 9 are not testable owing to the composition of the test material.

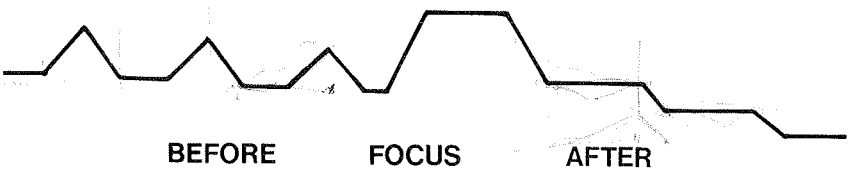


Figure 4. Downdrift in Swedish. Stylized Fo contour of a Standard Swedish utterance.

Figure 4 is an attempt to summarize in a schematic form the main features of the Fo downdrift in Swedish, as revealed by experiment one: the pivot position of the focus, the absent or gentle declination before focus as opposed to the stepwise declination after focus coinciding with the accented syllables, the rise-fall character of the accents before focus and the tendency towards simple falls after focus.

2.2 Discussion

Downdrift in Fo is a well attested phenomenon among the languages of the world, e.g. in languages as distinct as Germanic, West African and Japanese. A number of theories have been proposed to explain downdrift, mostly as a direct consequence of human production constraints. The most widely held opinion until recently has been to ascribe downdrift in Fo to the observed gradual fall in subglottal pressure over an utterance (Lieberman 1967, Atkinson 1973, Collier 1975). But other investigations (see references in Maeda 1976 and in Pierrehumbert 1979) show that the fall in subglottal pressure is hardly sufficient to account for the whole Fo drop over an utterance.

Perceptual experiments by Pierrehumbert (1979) indicate that listeners expect a downdrift in Fo.

Recently, however, it has been suggested that downdrift is not a direct consequence of human production or perception constraints but is learned behaviour, actively controlled and linguistically purposeful (Ohala 1978, Pierrehumbert 1979). It is useful in the textual organization of speech. According to this view downdrift in Fo can be seen as a phonetically motivated and natural process that may become integrated in the linguistic system of a language. It is therefore best regarded as only indirectly caused by production and perception constraints. For a similar reasoning on final lengthening see Lindblom (1978).

Downdrift will normally be exploited in languages, if language-specific demands do not interfere. The absence of downdrift in a three tone language like Yoruba, for example, is a case in

point (Hombert 1974). It will counteract the possibility of perceptual confusion of tones.

Downdrift was not ascribed a very significant role in my description of Standard Swedish (Bruce 1977). This is at least partly due to the fact that downdrift is not directly transparent, as it is not usually distributed evenly over an utterance. The placement of phrase accent to the sentence final position - even in a neutral version of a sentence - will tend to arrest downdrift up to this point, and most of the Fo drop of an utterance will be executed in the terminal Fo fall.

The results of experiment one support the view that downdrift is actively controlled and is part of the linguistic system of Standard Swedish. The absent or gentle declination up to the focus of an utterance and the stepwise declination after focus with the steps coinciding with the accented syllables are indicative of planned behaviour. But these results suggest another kind of planning of the overall Fo course than is predicted by the model. Rather than a total preplanning of the global, declining Fo course in relation to utterance length with local accent humps inserted on this gradual downdrift of intonation, it seems more likely that only certain Fo events are preplanned.

A more adequate account of the relationship between accentuation and intonation proper should give a more explicit role to their interaction than can be obtained from a mere addition of accentuation and intonation proper. This is also the view expressed in t'Hart and Collier (1979). Recently Pierrehumbert (1980) has presented an alternative to the "layered" theory of intonation, which will be discussed in a later section of this paper.

From a production point of view, planning only certain Fo events - the Fo movements associated with accentuation and their relations - appears to be relatively simple. It is easier to plan the step size of the accents if there is no Fo downdrift between successive accents. But a stepwise declination may also be optimal from a perception point of view.

3 EXPERIMENT TWO: UTTERANCE LENGTH - EXPANSION TO THE RIGHT

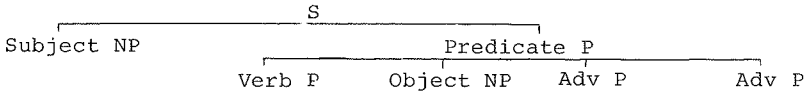
Experiment two was designed to test the model further, in particular to examine in more detail the relationship between Fo course and utterance length. The model predicts that the Fo range of either the baseline or the topline is constant over utterance length, while the slope varies inversely with utterance length due to the constancy of the starting points and end points. However, the results of experiment one indicate that the relationship between Fo downdrift and utterance length is not of a simple inverse type.

In order to test this prediction the following parameters were systematically varied in the test material.

(1) Number of stress groups (from one to five) and consequently utterance length
(2) Placement of focus, although not fully covered: neutral focus assignment or focus on the first or last stress group

As in experiment one a context sentence (question) was used to elicit a neutral focus assignment on either the first or the last stress group as the focus of the response sentence.

The increase in the number of stress groups from one to five in this experiment is achieved by a syntactic expansion to the right. This means in each case the addition of a new phrase with a new syntactic function. The longest utterance of the test material has the following syntactic composition:



Nån av ˈmammorna hann ˈlämna honom ˈungarna med ˈnallarna bland ˈlångorna

This gives a test material consisting of the following five sentences:

Test material - experiment two:

Nån av ˈmammorna.
One of the mothers.

Nån av ˈmammorna hann ˈlämna honom.
One of the mothers managed to leave him.

(Det var med ˈnallarna bland ˈlängorna.)
(It was with the teddy bears among the barns.)

Nån av ˈmammorna hann ˈlämna honom ˈungarna.
One of the mothers managed to leave him the kids.

Nån av ˈmammorna hann ˈlämna honom ˈungarna med ˈnallarna.
One of the mothers managed to leave him the kids with the teddy bears.

Nån av ˈmammorna hann ˈlämna honom ungarna med ˈnallarna bland ˈlängorna.
One of the mothers managed to leave him the kids with the teddy bears among the barns.

Each utterance has an onset of two unstressed syllables and a corresponding offset of two unstressed syllables. Utterance no. 2 is an exception. It has an offset of three unstressed syllables. Between each stressed syllable in an utterance there are three unstressed syllables. Each stressed syllable carries accent 2.

The same requirements on the phonetic composition of the test material as in experiment one were also complied with in this experiment.

The female informant UN (who had recorded a subset of the material in experiment one) recorded the whole material ten times.

A subset of the material was recorded by a male informant (OE, Uppsala dialect) seven times. This subset contained sentences with three and five stress groups plus the sentence in parentheses above with two stress groups, all with neutral focus assignment.

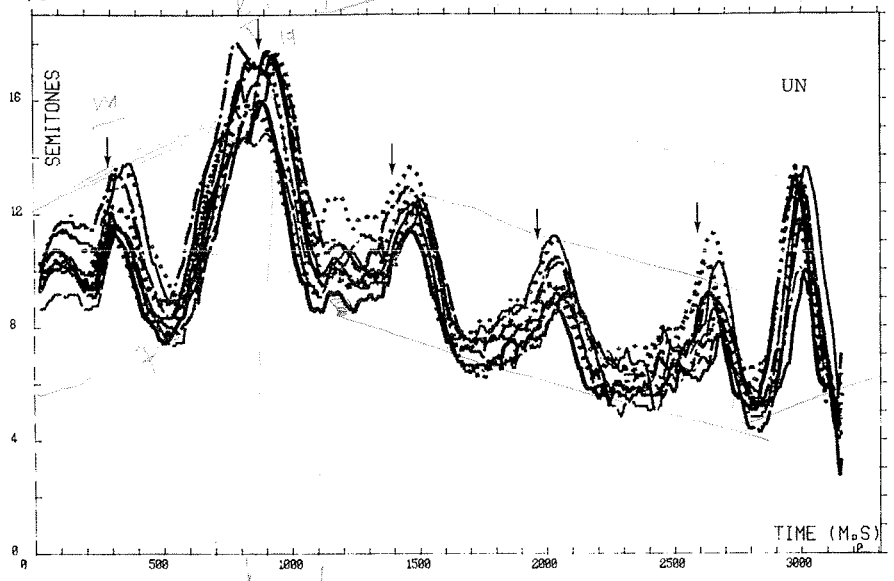
Fo analysis was carried out by the use of an autocorrelation LPC algorithm at the phonetics laboratory in Uppsala, Fo measurements being made interactively with the computer.

3.1 Results of experiment two

The main results of experiment two are summarized below. Figures 5 and 6 show repeated Fo contours of the longest utterance (5 stress groups, neutral version) for the female (10 repetitions) and the male informant (7 repetitions) respectively. The Fo contours are given in both a Hz and a semitone scale. Figures 7 and 8 show stylized baselines and toplines (mean values of successive Fo minima and maxima respectively) of utterances of varying length (2-5 stress groups) for both informants. Table II gives the means and standard deviations of successive Fo minima and maxima and ranges of successive Fo rises and Fo falls.

It should be pointed out that for both informants and in all three versions for the female informant (neutral, initial and final focus) the first stress group is assigned a phrase accent. In addition, the final focus and neutral versions of an utterance containing at least two stress groups have a second phrase accent on the last stress group.

1. The starting points and end points of an utterance appear to be constant and independent of utterance length (cf. experiment one). There is a tendency, however, for the shortest utterances (one or two stress groups) to have somewhat higher Fo values of the end points (see Table II).



nån av ˋmammorna hann ˋlämna honom ˋungarna med ˋhallarna bland ˋlängorna

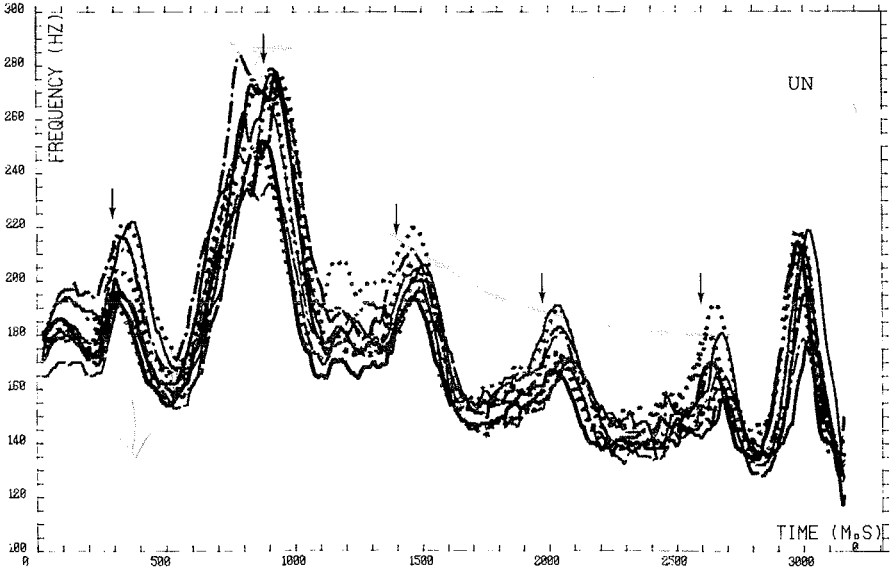
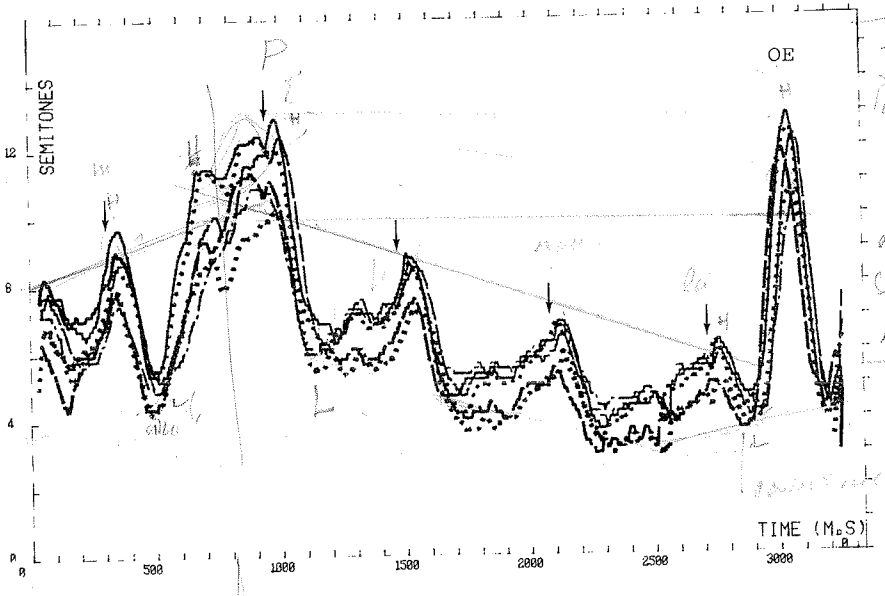


Figure 5. Fo course and utterance length (expansion to the right). Repeated Fo contours of the neutral version of a sentence containing five stress groups by a female informant (UN) with a semitone scale (above) and a Hz scale (below). The arrows indicate the stress group boundaries. The line-up point is the onset of the utterance.



nån av ˈmammorna hann ˈlämna honom ˈungarna med ˈhallarna bland ˈlängorna

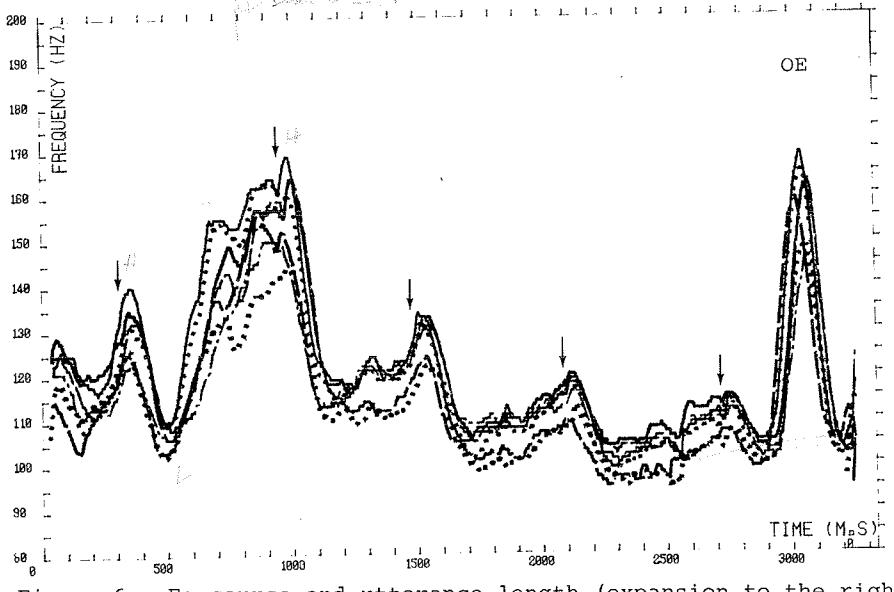


Figure 6. Fo course and utterance length (expansion to the right). Repeated Fo contours of a sentence containing five stress groups by a male informant (OE) with a semitone scale (above) and a Hz scale (below). The arrows indicate the stress group boundaries. The line-up point is the onset of the utterance.

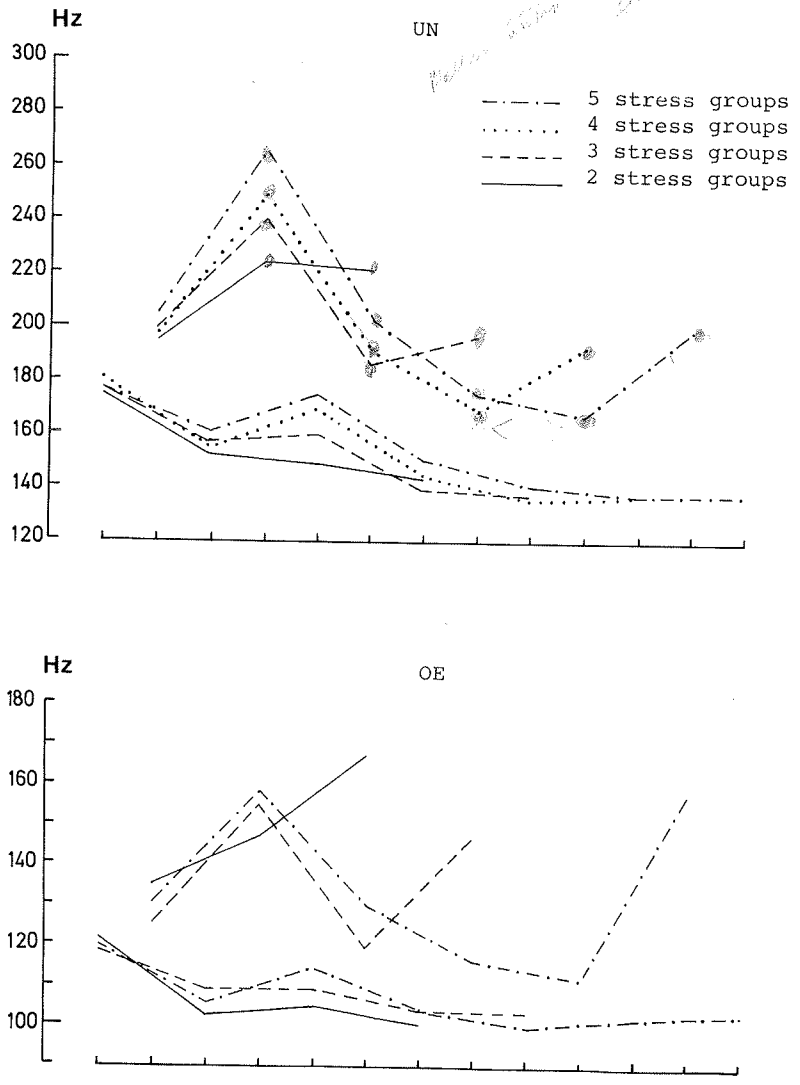


Figure 7. Fo course and utterance length (expansion to the right). Stylized baselines and toplines - means in Hz of successive Fo minima and Fo maxima - of utterances of varying length (2-5 stress groups). Neutral version for a female informant (UN) (above) and a male informant (OE) (below).

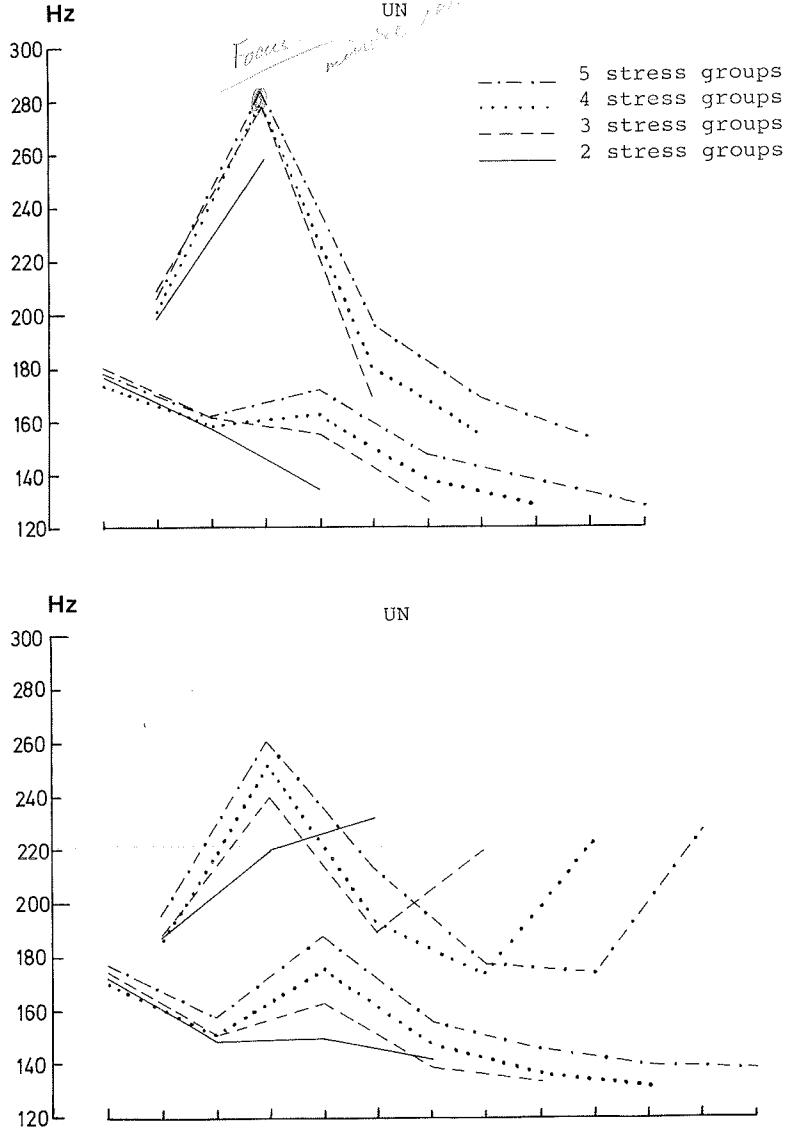


Figure 8. Fo course and utterance length (expansion to the right). Stylized baselines and toplines - means in Hz of successive Fo minima and Fo maxima - of utterances of varying length (2-5 stress groups). Initial focus version (above) and final focus version (below) for a female informant (UN).

2. The Fo maximum of the first phrase accent (in the first stress group) for all three versions (neutral, initial and final focus) varies with utterance length (number of stress groups): the longer the utterance, the higher the Fo maximum (see Table II and Figures 7 and 8).

3. Up to the Fo maximum of the first phrase accent, the Fo values are more or less constant. The single stress group utterance is an exception; the onset is the same, but after that point the Fo values are lower than for the longer utterances. The values are comparable to those of the final stress group in a longer utterance.

4. After the Fo maximum of the first phrase accent the Fo values (minima as well as maxima) are higher in a longer utterance except for the final phrase accent maximum, which does not show any systematic variation but appears to be relatively constant.

5. The step size of the first accent fall that follows the first phrase accent maximum - as well as that of subsequent ones - appears to be relatively constant (in Hz) across variations in utterance length. This relative constancy in step size is independent of the increase with utterance length of the Fo peak value of the first phrase accent and the preservation of higher Fo values for succeeding accents.

6. For each word accent after the first phrase accent maximum there is a successive lowering of the Fo values of both maxima and minima. The Fo declination appears to be exponential in nature. Consequently the step size of successive accent falls decreases similarly within an utterance.

7. While the range of successive accent falls decreases gradually within an utterance, the accent rises preceding the falls tend to have a constant Fo range independently of their position in the utterance.

8. The neutral and final focus versions display similar Fo patterns; but the Fo values for the final focus version tend to be somewhat lower initially and higher finally in the utterance. The first phrase accent maximum is the pivot point.

9. For the initial focus version the focal maximum is higher, and the succeeding Fo maxima and minima decrease faster to lower (absolute) values than for the final focus and neutral versions. The latter difference may be related to the absence versus presence of a phrase accent on the last stress group.

These results suggest that each utterance containing at least two stress groups is divided into two relatively independent prosodic phrases: one consisting of the first stress group and the other containing the rest of the utterance (1-4 stress groups). Each phrase is assigned a phrase accent.

The division into two prosodic phrases is also apparent from a comparison of the Fo values of the two minima surrounding the first phrase accent maximum (cf. Figures 7 and 8). While the preceding Fo minimum is more or less constant across variations in utterance length, the Fo minimum following the phrase accent maximum increases with the number of following stress groups. This will introduce a change in the overall Fo course of an utterance, which becomes more apparent with increased utterance length.

Generally speaking, a prosodic phrase contains one or more stress groups. As was indicated in the preceding section, it is connected with major syntactic phrases such as Subject NP and Predicate P and also with topic comment structure. But the ultimate division into prosodic phrases is probably the speaker's own choice depending on his/her communicative intension.

The issue of division into phrases will be examined further in experiment three.

Table II. Fo course and utterance length (expansion to the right). Fo means and standard deviations in Hz of successice minima and maxima and Fo ranges of successive rises and falls in the test material recorded in experiment two; (a) female informant UN, (b) male informant OE. The phrase accent maxima are in italics.

(a) Speaker UN

		min	max	min	max	min	max	min	max	min	max	min	max	min	n
nån av `mammorna															
neutral	$\bar{x}$	171	175	140	*205*	149									10
focus	s	6,9	10,9	5,5	13,3	9,6									
	Δf		+5	-36	+66	-56									
nån av `mammorna hann `lämna honom															
neutral	$\bar{x}$	176	195	153	*225*	150	*221*	144							8
focus	s	5,5	10,6	3,3	15,2	3,8	12,5	4,0							
	Δf		+19	-42	+72	-75	+72	-78							
initial	$\bar{x}$	177	198	158	*258*	134									10
focus	s	5,6	5,7	4,2	6,6	4,1									
	Δf		+21	-41	+100	-124									
final	$\bar{x}$	172	187	148	*220*	150	*231*	141							9
focus	s	4,1	9,3	5,2	11,5	3,3	4,2	5,1							
	Δf		+15	-39	+72	-70	+81	-90							
nån av `mammorna hann `lämna honom `ungarna															
neutral	$\bar{x}$	178	199	157	*240*	160	186	140	*196*	137					11
focus	s	5,7	4,8	4,4	10,2	4,1	7,8	1,9	11,5	8,4					
	Δf		+21	-42	+83	-80	+26	-46	+56	-59					
initial	$\bar{x}$	180	208	160	*278*	155	169	130							10
focus	s	5,7	13,1	10,1	13,5	5,6	9,8	3,3							
	Δf		+28	-48	+117	-123	+14	-39							
final	$\bar{x}$	174	187	151	*239*	163	190	139	*219*	133					10
focus	s	6,7	13,2	7,7	5,8	5,2	4,3	6,4	15,7	6,8					
	Δf		+13	-37	+88	-76	+27	-51	+80	-86					

(a) Speaker UN

		min	max	min	max	min	max	min	max	min	max	min	max	min	n
nån av `mammorna hann `lämna honom `ungarna med `nallarna															
neutral	$\bar{x}$	181	197	155	*249*	169	191	145	169	136	*192*	137			9
focus	s	7,0	4,9	4,4	12,7	4,3	6,7	3,1	5,3	4,1	16,7	10,9			
	Δf		+17	-42	+95	-80	+22	-46	+25	-34	+55	-53			
initial	$\bar{x}$	174	201	159	*278*	162	180	138	155	128					10
focus	s	4,5	7,3	6,0	13,9	6,2	7,5	3,9	5,6	3,6					
	Δf		+27	-42	+119	-116	+18	-42	+17	-27					
final	$\bar{x}$	170	185	151	*251*	175	193	148	174	137	*222*	131			10
focus	s	5,8	11,4	7,6	9,1	3,9	6,1	2,9	4,4	4,5	8,6	4,0			
	Δf		+15	-34	+101	-76	+18	-45	+27	-38	+87	-91			
nån av `mammorna hann `lämna honom `ungarna med `nallarna bland `längorna															
neutral	$\bar{x}$	117	205	162	*265*	175	204	150	176	141	168	137	*199*	137	10
focus	s	5,3	12,1	5,9	17,3	7,4	7,8	4,7	9,6	5,3	10,7	5,4	15,7	8,2	
	Δf		+28	-44	+103	-90	+28	-54	+26	-35	+27	-31	+63	-62	
initial	$\bar{x}$	178	206	162	*283*	171	197	147	169	137	153	128			10
focus	s	5,5	4,4	2,8	11,0	7,8	8,3	4,0	7,9	4,3	7,3	3,7			
	Δf		+28	-44	+121	-112	+26	-50	+22	-32	+16	-25			
final	$\bar{x}$	177	195	157	*260*	187	212	156	177	145	174	139	*227*	138	10
focus	s	8,8	11,7	8,1	17,2	10,8	16,0	5,5	7,5	4,3	10,3	5,2	14,0	6,7	
	Δf		-18	-37	+103	-73	+25	-57	+22	-31	+28	-35	+88	-88	

(b) Speaker OE

det var med `nallarna bland `längorna									
neutral	$\bar{x}$	122	135	103	*147*	106	*168*	100	7
focus	s	6,2	6,1	2,8	7,6	2,5	9,3	2,6	
	Δf		+13	-32	+44	-41	+62	-68	
nån av `mammorna hann `lämna honom `ungarna									
neutral	$\bar{x}$	119	125	109	*154*	109	119	103	*147*	103	7
focus	s	3,7	6,6	9,5	4,3	1,8	4,1	4,9	14,0	8,2	
	Δf		+6	-16	+45	-45	+10	-16	+44	-44	
nån av `mammorna hann `lämna honom `ungarna med `nallarna bland `längorna											
neutral	$\bar{x}$	120	130	106	*158*	114	130	104	116	99	112	102	*157*	103	7
focus	s	4,3	5,9	3,0	9,2	3,4	4,7	3,8	3,9	3,4	3,3	2,1	7,9	1,5	
	Δf		+10	-24	+52	-44	+16	-26	+12	-17	+13	-10	+55	-54	

The results of experiment two also confirm the results of experiment one as far as they are comparable. The only qualification is the following: after a phrase accent maximum the word accents are usually manifested as rise-falls and not as simple falls as they were schematically characterized in the preceding section. To some extent this difference may be due to different prominence relations. In a version of an utterance where the phrase accent is also focal (experiment one, experiment two: initial focus version) the following (postfocal) accents may become downgraded in prominence. Also in the present experiment there is a tendency for the rise of at least the last accent of a declining Fo course to be less pronounced, if it is not followed by a phrase accent (initial focus version). There is no such tendency if it is followed by a phrase accent (neutral and final focus versions). Even so the characterization of postfocal accents as simple falls is an overgeneralization.

Moreover it should be added that "true" downdrift - a gradually decreasing Fo course - is found not only after the last accent, when several unaccented syllables follow, but also in the single stress group utterance, i.e. with only one accent.

3.2 Discussion

The results of experiment two are interesting from the point of view of speech planning. Although the starting points and end points appear to be constant and independent of utterance length, there is no simple variation in slope of either baseline or topline due to variations in utterance length as predicted by the model. Instead the speaker will signal a longer utterance primarily by increasing the peak value of the first phrase accent and by keeping the Fo values of succeeding accents higher as compared to a corresponding shorter utterance. It is possible that the listener may also benefit from this adjustment of Fo peak value as an expression of utterance length.

This variability of the first phrase accent peak is not expressed in our model, which in the present version predicts a constant phrase accent peak value for all contexts (see Section 1). The model also expresses the independence of the interior (slanting) and exterior (horizontal) auxiliary lines. This prediction is not supported by the results of experiment two. Instead, a higher peak value of the first phrase accent will give higher Fo values to the succeeding minima and maxima except for the final phrase accent maximum.

We have seen that, while the Fo peak of the first phrase accent varies with utterance length (number of stress groups) the very onset of an utterance (the Fo starting point) is constant and independent of variation in utterance length. These facts suggest that the Fo onset and the Fo peak (of the first phrase accent) may have different functions and express different things.

It seems justified to say that the peak value of the first phrase accent has the function of anticipating utterance length. As the Fo end point appears to be constant, the speaker anticipates a greater number of accentual downsteps by reaching a higher Fo point of departure for the downstepping, i.e. to ensure that the Fo bottom of the speaker's voice will not be reached until the end of an utterance (see further discussion in 5.2 and 7.1). The role of the Fo onset is less clear. There is some evidence from another investigation (Bruce 1981) that the Fo onset value of an utterance may be related to that of the immediately preceding utterance and is part of the organization of running discourse. It would therefore be natural that the Fo onset does not vary with utterance length.

The relationship between Fo downdrift and utterance length has been examined for Japanese (Fujisaki et al. 1979), Danish (Thorsen 1980, 1981) and American English (Sorenson & Cooper 1980, Pierrehumbert 1980) as well. A simple inverse relationship between slope and utterance length is not typical of any of these investigations.

Fujisaki's investigation of Japanese shows that there is an exponential decay of the Fo course of an utterance - a steeper fall in the beginning and then a leveling out to an asymptotic value. The shape of the downdrift is approximately the same independently of utterance length.

Thorsen also recognizes the exponential character of the Fo course of a longer utterance in Danish, although the tendency is less clear. But the characteristic feature is still as asymptotic decrease in the overall downdrift with increased utterance length - in her investigations ranging from two to eight stress groups. Together with a widening of the Fo range an overall downdrift will be preserved in a longer compared to a shorter utterance by a non-linear decrease in slope.

Sorenson & Cooper's study of declination in American English shows that the initial Fo peak of an utterance tends to be higher for a longer than for a shorter utterance, while the final Fo peak is nearly constant. A faster declination is also found in the early part of an utterance.

In her thesis Pierrehumbert (1980) presents a very interesting hypothesis concerning the implementation of the course of accentual downsteps in Amcerican English. To account for the overall Fo course of an utterance containing a number of accentual downsteps, she proposes a local implementation rule. The rule computes the value of a given pitch accent as a constant ratio of the value of the immediately preceding pitch accent. The overall Fo course of an utterance is the result of a recursive application from left to right of this local rule. No look-ahead is presupposed, but the rule still generates the exponential character of the Fo course, which is also asymptotic to the bottom of the speaker's voice.

This local approach is found to be superior to the non-local approach of the "layered" theory (see Section 1) where the Fo contour of an utterance is the result of superposing essentially independent layers: a globally specified sentence intonation on which the pitch accents are superimposed.

Applied to the Swedish material of experiment two the local approach also appears to correctly predict the constant ratio of successive downsteps and the consequent exponential character of the Fo course. This will be examined in more detail in Section 7.

A difference between Pierrehumbert's data on American English and the present data on Swedish is the sensitivity of the Fo peak to utterance length, which is not calculated with but is still recognized in a footnote by Pierrehumbert (1980). However, a local implementation is possible both with and without this variability of the Fo peak of an utterance.

The interesting implication for speech planning of the local approach is that the speaker does not have to plan the Fo course of the whole utterance in advance. This would have been necessary if the speaker divided the total Fo range by the number of upcoming accentual downsteps in the utterance. Instead, the speaker seems to plan for utterance length by adjusting the Fo peak value to the number of upcoming accentual downsteps in the utterance and by keeping the Fo values of subsequent accents higher than in an utterance with fewer accentual downsteps. Equal prominence of successive accents in a declining Fo course can be achieved by letting the Fo values of each accent be a constant ratio of the corresponding value of the preceding accent.

4 EXPERIMENT THREE: UTTERANCE LENGTH - EXPANSION TO THE LEFT

In experiment two utterance length was varied by syntactic expansion to the right. This implies that in an utterance containing several stress groups, one stress group belongs to the subject NP and the remainder to the Predicate Phrase. Recently Gårding et al. (1982) have argued that this bias in test material should be counterbalanced by also including syntactic expansion to the left, i.e. in the present case expansion of the subject NP. The argument is that the emphasis on syntactic expansion to the right and on Fo downdrift may conceal an Fo updrift in the early part of an utterance, which would be revealed when the

utterance is expanded to the left.

A test material was devised to test this hypothesis for Standard Swedish. The test sentences used in experiment three were composed by varying the theme from experiment two. The point of departure was the sentence with three stress groups, which was expanded both to the left and to the right. One sentence with three stress groups, one with four and two with five were included:

Test material - experiment three:

[*Nån av ˆmammorna*] [*hann ˆlämna honom ˆungarna*]
One of the mothers managed to leave him the kids

expansion to the left:

[*Nån av ˆmammorna med ˆnallarna*] [*hann ˆlämna honom ˆungarna*]
One of the mothers with the teddybears managed to leave him the kids

[*Nån av ˆmammorna med ˆnallarna bland ˆlångorna*.] [*hann ˆlämna honom ˆungarna*]
One of the mothers with the teddybears among the barns managed to leave him the kids

expansion to the right:

[*Nån av ˆmammorna*] [*hann ˆlämna honom ˆungarna med ˆnallarna bland ˆlångorna*]
One of the mothers managed to leave him the kids with the teddybears among the barns

The syntactic structure of the five stress group utterance with expansion to the left is the following (compare the corresponding syntactic structure of the five stress group utterance with expansion to the right in the preceding section):

Nån av ˆmammorna med ˆnallarna bland ˆlångorna | hann ˆlämna honom ˆungarna

As can be seen the same constituents as in the material of experiment two were used. The syntactic expansion to the left was achieved merely by transposing the constituents. Only neutral focus assignment was included in the present experiment. The same informants as in experiment two recorded the test material of experiment three seven times in a new recording session.

As in experiment one the recordings were processed by hardware pitch and intensity meters, and Fo was measured manually.

4.1 Results of experiment three

The most important results of experiment three are summarized below. Figures 9 and 10 show stylized Fo contours based on Fo means of successive minima and maxima in utterances where the direction of syntactic expansion was varied. Table III gives the means and standard deviations of Fo minima and maxima of the test material in experiment three.

1. The starting points and the end points appear to be constant and independent of direction of syntactic expansion (cf. the results of experiment one and two).
2. The position of the first phrase accent and its Fo maximum varies with expansion to the left. It marks the end of the first prosodic phrase (subject NP).
3. Up to the first phrase accent maximum there is no downdrift in Fo and the baseline and topline are virtually flat.
4. The Fo maximum of the first phrase accent varies with utterance length. This is true of expansion to the right (cf. experiment two), but there is a tendency also for expansion to the left to give a higher Fo maximum at the first phrase accent.
5. After the Fo maximum of the first phrase accent the Fo minima and maxima decrease in a similar way independently of the number of stress groups (one, two or three) contained in the subject NP (left expansion). The Fo course of the right expansion is also the same as in experiment two.

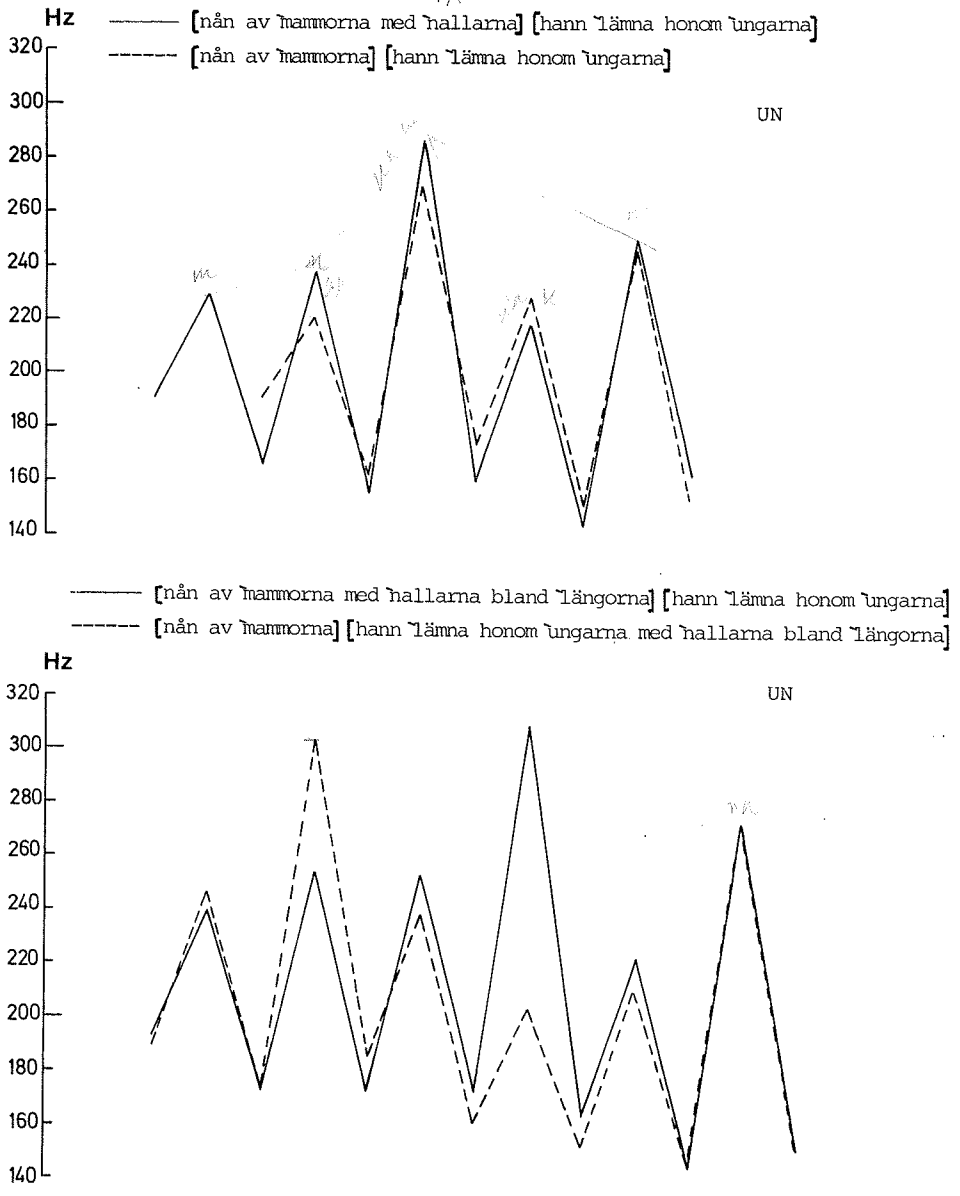


Figure 9. Fo course and utterance length (expansion to the left and to the right). Stylized Fo contours - means in Hz of successive Fo minima and Fo maxima - of a three stress group utterance compared to a four stress group utterance (above) and of two versions of a five stress group utterance (below) by a female informant (UN).

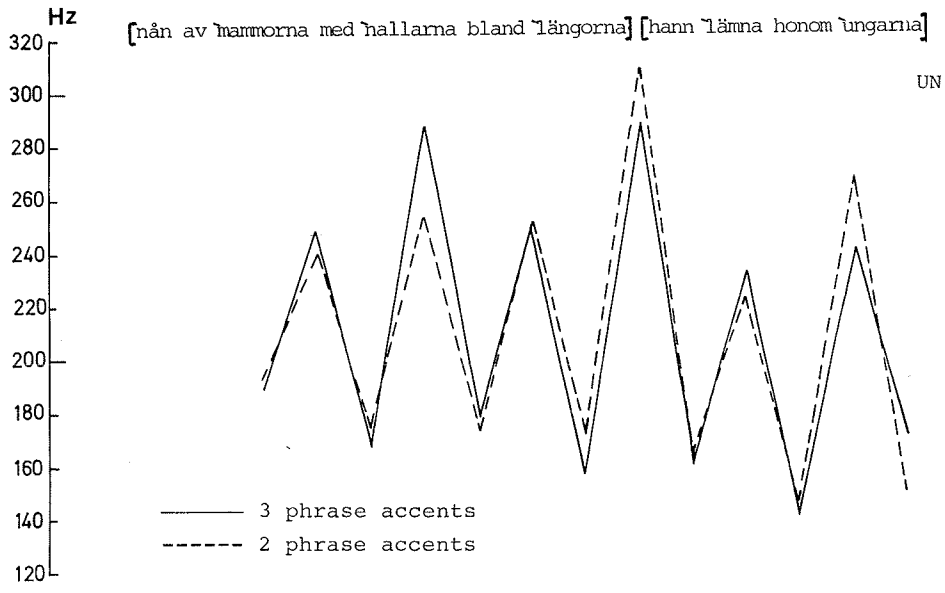


Figure 10. Fo course and utterance length (expansion to the left). Stylized Fo contours - means in Hz of Fo minima and Fo maxima - of two variants of the five stress group utterance with left expansion by a female informant (UN).

These results are representative of one informant (female) and should be considered preliminary. As to her manifestation of the five stress group utterance (left expanded version) there appears to be some variation, however. In four out of seven repetitions the Fo contour agrees with the above description - the first phrase accent marks the end of the subject NP, and there is no declination up to this point. In the remaining three repetitions she manifests as many as three phrase accents; one on the first stress group, another on the final stress group of the subject NP, and a third on the final stress group of the utterance. After the first phrase accent maximum the Fo course of accentual downsteps is comparable to that of the right expanded version of the five stress group utterance. It is arrested, however, by the insertion of a second phrase accent and is then started anew.

Table III. Fo course and utterance length (expansion to the left and to the right). Fo means and standard deviations in Hz of successive minima and maxima of the test material recorded in experiment three; (a) female informant UN, (b) male informant OE. The phrase accent maxima of each utterance are in italics.

(a) Speaker UN

		min	max	min	max	min	max	min	max	min	max	min	max	min	n
nån av `mammorna hann `lämna honom `ungarna															
neutral	$\bar{x}$	191	220	160	*269*	173	229	148	*247*	153					7
focus	s	4,5	13,2	6,5	8,0	8,6	9,8	3,9	15,5	11,1					
nån av `mammorna med `nallarna hann `lämna honom `ungarna															
neutral	$\bar{x}$	193	233	168	233	154	*293*	160	217	145	*253*	163			6
focus	s	13,7	20,4	8,2	19,4	3,8	13,3	6,3	13,3	3,2	16,6	9,4			
nån av `mammorna hann `lämna honom `ungarna med `nallarna bland `längorna															
neutral	$\bar{x}$	187	245	172	*304*	186	239	161	204	152	211	147	*273*	156	7
focus	s	4,9	23,3	12,2	10,6	6,1	14,6	6,1	6,7	3,9	4,5	2,7	13,2	19,0	
nån av `mammorna med `nallarna bland `längorna hann `lämna honom `ungarna															
neutral	$\bar{x}$	193	240	176	254	174	253	173	*310*	164	224	146	*270*	153	4
focus	s	6,5	0,0	6,3	21,0	8,5	6,5	11,9	4,1	2,5	13,2	2,5	14,1	11,9	
neutral	$\bar{x}$	189	248	168	*288*	179	251	159	*290*	161	234	144	*245*	174	3
focus	s	6,3	18,5	2,9	8,7	7,5	11,8	2,5	23,5	4,8	10,3	2,5	12,9	18,9	

(b) Speaker OE

nån av `mammorna hann `lämna honom `ungarna															
neutral	$\bar{x}$	125	130	108	*154*	107	121	106	*145*	102					7
focus	s	0,0	3,0	2,0	3,8	2,4	3,1	2,0	5,8	4,4					
nån av `mammorna med `nallarna hann `lämna honom `ungarna															
neutral	$\bar{x}$	126	131	111	*153*	110	129	104	118	105	*144*	98			6
focus	s	3,8	3,1	3,4	5,2	2,9	7,0	3,4	5,1	2,9	6,6	3,8			
nån av `mammorna hann `lämna honom `ungarna med `nallarna bland `längorna															
neutral	$\bar{x}$	127	133	111	*161*	113	129	103	121	100	115	107	*148*	100	6
focus	s	2,6	2,6	3,0	3,8	4,0	4,9	2,9	4,9	1,9	2,9	1,9	9,3	2,7	
nån av `mammorna med `nallarna bland `längorna hann `lämna honom `ungarna															
neutral	$\bar{x}$	125	132	110	*156*	111	128	103	121	101	116	105	*140*	98	8
focus	s	2,7	3,5	2,1	5,2	3,8	3,7	2,6	3,3	2,2	3,3	3,3	6,0	3,2	

For the second informant (male) the direction of syntactic expansion is not reflected in the Fo pattern. He exhibits practically the same Fo contour of the five stress group utterance, whether it is the left expanded or the right expanded version. The first stress group contains one phrase accent and the final stress group another one. Between these two maxima the Fo contour has an exponentially decreasing course.

The conclusion that can be drawn from the results of experiment three is that the direction of syntactic expansion, viz. expansion to the left, may be decisive for the Fo pattern, but this seems to be optional.

4.2 Discussion

The division of an utterance into (at least) two relatively independent, prosodic phrases is confirmed in experiment three (cf. experiment two). Each prosodic phrase contains a phrase accent - at least in the neutral case - to which the other accents of the phrase are subordinated.

It was hypothesized above based on findings in Gårding et al. (1982) that the extension of the test material by a syntactic expansion to the left would reveal an Fo updrift preceding the usual Fo downdrift of an utterance. What was found was instead no drift in Fo at all, i.e. a flat baseline and topline up to the first phrase accent maximum. The situation appears to be parallel to the one described in experiment one, where absence of declination was found up to the focus of an utterance.

The results of experiment three suggest that the direction of syntactic expansion may, but not necessarily, affect the Fo course of an utterance. This may be taken to support the view that there is a relatively weak correlation between syntax and prosody.

The increase in the Fo maximum of the first phrase accent (cf. experiment two) with increased utterance length is also found in the present experiment. This increase occurs both with a right and a left expansion.

Intuitively it makes more sense to conceive of an adjustment of the Fo peak to what remains to be done (number of stress groups to be executed) than to what has already been done. It is, however, interesting to compare with the durational patterns in Swedish described by Lindblom et al. (1976). The shortening of a stressed vowel was found to increase as a function of the number of following syllables but it was also to some extent dependent on the number of preceding syllables.

5 EXPERIMENT FOUR: ATTITUDE - INVOLVEMENT

Most of the phonetic research in prosody in recent years has been directed towards elucidating what is regarded as neutral prosody or symbol prosody, i.e. prosody without emotional or attitudinal colouring. Prosodic expressions of emotions and attitudes - symptom and signal prosody - have been considered to be of secondary importance or too difficult to handle and have therefore usually been ignored. There is reason to believe, however, that this neglect of symptom and signal prosody does not correspond to its relative importance. It has been suggested that symptom prosody may be linguistically as basic as symbol prosody and that research in this area could also give new insights into symbol prosody (see the discussion by Anward in Nordic Prosody, 1978, pp. 292-293).

In the experiments reported above the attitude of a speaker in a recording session was held more or less constant, although naturally some variation in attitude between speakers was observed. In the present experiment a first attempt has been made to include a systematic variation in attitude within the same speaker. There is strong reason to believe, however, that prosody mirrors emotions and attitudes incompletely, although some, relatively rough categories are kept distinct prosodically (see e.g. Hadding-Koch 1961, Brown et al. 1980). The aim was to elicit a simple, two-way distinction along the dimension involved - detached. It has been shown that a dichotomy of this kind has clear prosodic correlates (cf. Hadding-Koch 1961,

p. 122, Brown et al. 1980, p. 23). The prosodic correlate found to express most directly an involved - detached distinction is pitch range. It should be emphasized that this prosodic dimension (wide versus narrow pitch range) does not primarily express positive versus negative attitudes, but rather, as suggested here, involved - be it an expression of joy or anger - versus detached. Degree of involvement is also probably closely related to degree of overall emphasis (see Gårding & Lindblad 1973, Pierrehumbert 1980, p. 119).

The present experiment should make it possible to extend the intonation model for Swedish to include variation in attitude.

The test material that was used to vary the degree of involvement of the speaker is identical to a subset of the test material from experiment two. The variables are thus number of stress groups (2-5) with neutral focus assignment, and attitude (detached - involved). One informant (female, UN) recorded the test material six times.

As in experiment one and three the recordings were processed by hardware pitch and intensity meters, and Fo was measured by hand.

5.1 Results of experiment four

The main results of experiment four are summarized below. Figures 11 and 12 show stylized Fo contours - means of successive Fo minima and Fo maxima - of utterances, where the degree of involvement (detached - involved) was varied. Table IV gives means and standard deviations of Fo minima and maxima and Fo ranges in the test material of experiment four.

1. Differences in attitude (detached - involved) are clearly expressed as differences in Fo range; range varies with degree of involvement.

2. Variation in Fo range is achieved by a frequency expansion upwards; the lower limit of the range is fixed, while the upper limit is highly flexible.

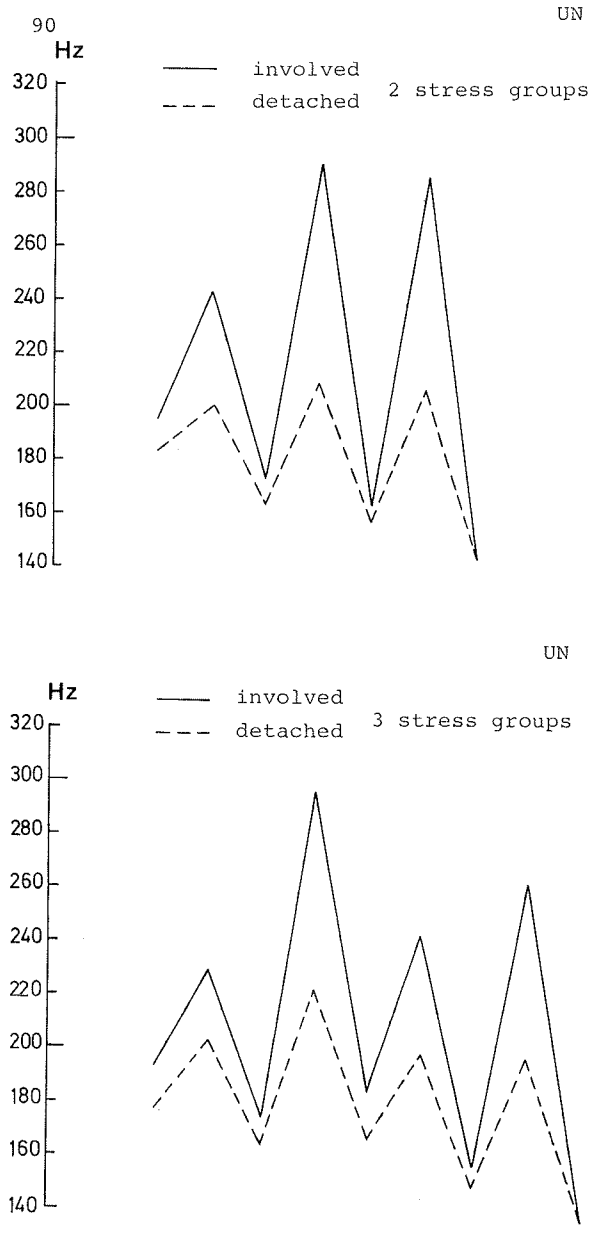


Figure 11. Fo course and attitude (involvement). Stylized Fo contours - means in Hz of successive Fo minima and Fo maxima - of an involved and a detached version of a two stress group utterance (above) and of a three stress group utterance (below) by a female informant (UN).

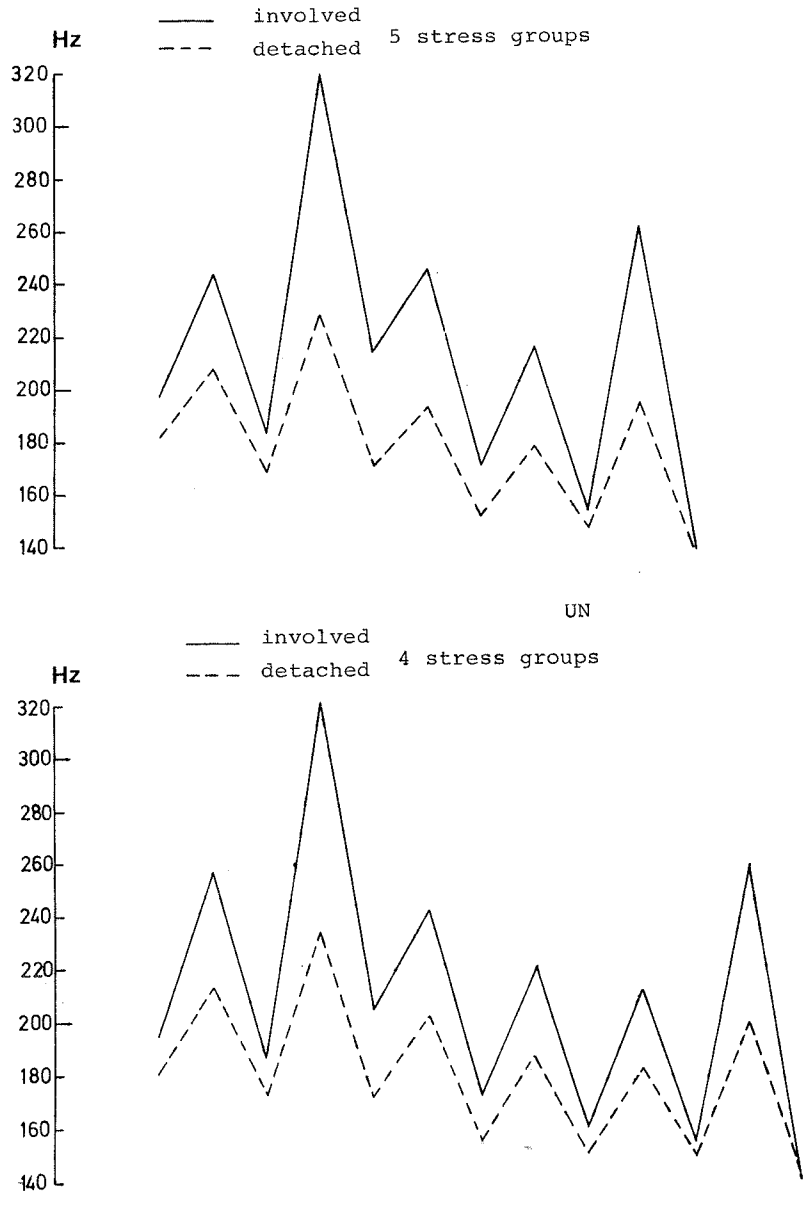


Figure 12. Fo course and attitude (involvement). Stylized Fo contours - means in Hz of successive Fo minima and Fo maxima - of an involved and a detached version of a four stress group utterance (above) and of a five stress group utterance (below) by a female informant (UN).

3. Fo at the endpoint appears to be constant and independent of attitude (detached - involved) and of utterance length (cf. experiments two and three).

4. The starting point varies with degree of involvement. But within the same degree of involvement the starting point is independent of utterance length (cf. experiments two and three).

5. All other Fo points (minima and maxima) are affected by variations in attitude (detached - involved). The Fo peaks appear to be more affected than the valleys.

6. The same Fo pattern is maintained across variations in attitude (detached - involved). This means that approximately the same relations seem to hold between successive Fo maxima and Fo minima within an utterance. The impression of difference between detached and involved is a matter of scaling.

The main trends from experiments two and three are repeated in the present experiment within one and the same attitude category (detached or involved). Utterance length affects the value of the Fo peak of the first phrase accent of an utterance and the following minima and maxima up to the final phrase accent maximum, while the starting points and end points are unaffected. The exponential decay of the Fo course of a longer utterance also occurs in the present experiment.

Concomitant with the increase in Fo range as a function of increased involvement is an increase in intensity, which is apparent from the records but which has not been measured. Differences in duration between the detached and the involved version of an utterance appear to be negligible.

5.2 Discussion

It can be considered highly expected and very natural that large Fo movements (a wide Fo range) are associated with a higher degree of involvement on the part of the speaker, while small Fo movements (a narrow Fo range) mark the detachment of the speaker. This pattern is naturally paralleled in the body language by large versus small gestures while speaking.

Table IV. Fo course and attitude (involvement). Fo means and standard deviations in Hz of successive minima and maxima and Fo ranges of successive rises and falls in the test material recorded in experiment four of an involved and a detached attitude by the female informant UN. The phrase accent maxima of each utterance are in italics.

Speaker UN

		min	max	min	max	min	max	min	max	min	max	min	max	min	n
nån av `mammorna hann `lämna honom															
detached	$\bar{x}$	184	200	163	*209*	158	*207*	144							6
	s	6,7	6,3	6,1	4,9	4,1	7,5	5,9							
	Δf		+16	-37	+46	-51	+49	-63							
involved	$\bar{x}$	196	243	174	*291*	164	*287*	143							6
	s	11,6	19,4	9,7	17,1	8,6	17,2	2,6							
	Δf		+47	-69	+117	-127	+123	-144							
nån av `mammorna hann `lämna honom `ungarna															
detached	$\bar{x}$	178	202	164	*222*	167	198	150	*198*	142					6
	s	5,2	7,5	3,8	12,5	4,1	4,2	3,2	8,8	2,6					
	Δf		+24	-38	+58	-55	+31	-48	+48	-56					
involved	$\bar{x}$	193	230	175	*296*	185	243	157	*262*	139					6
	s	15,1	22,1	16,7	22,2	15,2	20,4	6,1	20,9	2,0					
	Δf		+37	-55	+121	-111	+58	-86	+105	-123					
nån av `mammorna hann `lämna honom `ungarna med `nallarna															
detached	$\bar{x}$	182	207	168	*229*	171	193	152	178	148	*195*	139			6
	s	10,8	9,8	9,8	13,2	5,9	5,2	2,6	7,6	2,7	15,2	3,8			
	Δf		+25	-39	+61	-58	+22	-41	+26	-30	+47	-56			
involved	$\bar{x}$	198	243	183	*320*	214	246	172	214	154	*262*	143			6
	s	5,2	23,8	15,7	28,8	24,4	25,4	10,8	16,3	11,6	27,1	8,8			
	Δf		+45	-60	+137	-106	+32	-74	+42	-60	+108	-119			
nån av `mammorna hann `lämna honom `ungarna med `nallarna bland `längorna															
detached	$\bar{x}$	181	213	173	*234*	173	203	156	188	152	183	151	*201*	143	6
	s	4,9	8,2	5,2	17,4	7,5	6,1	4,9	4,2	2,6	5,2	3,8	7,4	6,1	
	Δf		+32	-40	+61	-61	+30	-47	+32	-36	+31	-32	+50	-58	
involved	$\bar{x}$	195	257	188	*322*	206	253	174	222	161	213	155	*259*	143	6
	s	5,5	23,2	12,1	17,2	7,4	5,2	4,9	6,1	4,9	11,7	4,5	20,8	2,7	
	Δf		+62	-69	+134	-116	+47	-79	+48	-61	+52	-58	+104	-116	

In the preceding experiments the Fo value of the end point was constant and insensitive to variation in utterance length, focus location and syntactic composition. The constancy of the Fo end point is true also of variation in attitude (detached - involved), as evidenced from the present experiment. Consequently both large and small Fo movements have the same offset, which is reached towards the end of an utterance. The end point seems to function as a reference equal to the bottom of the speaker's voice. To go lower would not be possible unless a change in voice quality is made, e.g. the use of a creaky voice, which is in fact exploited by many speakers at the very end of an utterance.

While there appears to be a definite floor for a speaker's voice, which is reached in almost every utterance, the corresponding Fo ceiling is out of range in normal speech (cf. Pierrehumbert 1980). Therefore it is particularly the top of a speaker's Fo range that is open to variation, although a wide range is signalled already from the very onset of an utterance.

We have seen in this experiment that there exists a certain relationship between successive Fo maxima and also between successive Fo minima. A high Fo peak of the first phrase accent as a signal of involvement means that the following Fo peaks of the same utterance will have higher values than in a corresponding utterance where the first phrase accent peak is relatively low as a signal of detachment.

In this connection it is interesting to draw attention to a perceptual experiment (Bruce 1977, chapter 7.2) that indicates that it is important also from a perceptual point of view to maintain certain proportions in Fo range between successive Fo peaks within an utterance. The results of the test show that an Fo peak of a fixed value is neglected by the listener, if the preceding initial Fo peak is relatively high, but it is taken into consideration if the preceding Fo peak has a relatively low value.

These results were interpreted in the following way: "The listener expects a speaker to use approximately the same fundamental frequency range during a whole utterance. It is assumed that the range in the beginning of an utterance sets the reference. Therefore, if the speaker begins with a relatively wide frequency range in connection with the first Fo peak, the listener will disregard minor changes in Fo like a small, medial Fo peak. But if the speaker starts with a relatively narrow range, the same absolute value of the medial Fo peak will be an integral part of the intonation pattern of that utterance, and the listener will pay attention to it" (Bruce 1977, p. 127).

6 EXPERIMENT FIVE: CLAUSE BOUNDARY

In the previous experiments we have studied the Fo pattern of utterances where focus location, utterance length (with both right and left expansion) and attitude (degree of involvement) have been varied. All test sentences so far contain a single clause, although they usually consist of two phrases. The present experiment will examine the effect on the Fo contour of introducing a clause boundary in a sentence.

The material consists of two different clauses containing two stress groups each. These can occur as constituent parts of a double clause sentence (coordinative with adversative coupling), they can be combined freely and they can also occur as single clause sentences. The two stress group clauses consist of one phrase each. The double clause sentences were elicited both with and without a physical pause at the clause boundary. The same requirements on the phonetic composition of the test material as in the preceding experiments were also imposed on this experiment.

Test material - experiment five:

Man ˇlämnar ˇmamman
One leaves the mother

Man ˇnallar ˇungen
One steals the kid

Man ˇlämnar ˇmamman (,) men man ˇnallar ˇungen
One leaves the mother but one steals the kid

Man ˇnallar ˇungen (,) men man ˇlämnar ˇmamman
One steals the kid but one leaves the mother

A neutral context for a double clause sentence of the present type in Standard Swedish will elicit one phrase accent on the first stress group of the first clause and another phrase accent on the final stress group of the second clause. In order to be able to compare the single clause sentence with each of the constituent clauses of a double clause sentence placement of focus (= phrase accent) has therefore been varied in the single clause sentences by varying the context sentence.

The female informant UN recorded the material eight times. The recording was processed by hardware pitch and intensity meters, and Fo was measured by hand.

6.1 Results of experiment five

This is a pilot experiment, and the results should be considered preliminary. Figure 13 compares typical Fo contours of the first clause of a double clause sentence with a physical pause at the clause boundary and of the corresponding single clause sentence. Table V gives the means and standard deviations of successive Fo minima and Fo maxima of the test sentences. The most important results are the following:

1. The Fo maxima and minima are higher for the first clause of a double clause sentence (both versions with and without a pause) than for a single clause sentence. This includes the Fo starting point, which is lower for a single clause sentence than for a double clause sentence. The starting point value of the single clause sentence is comparable to

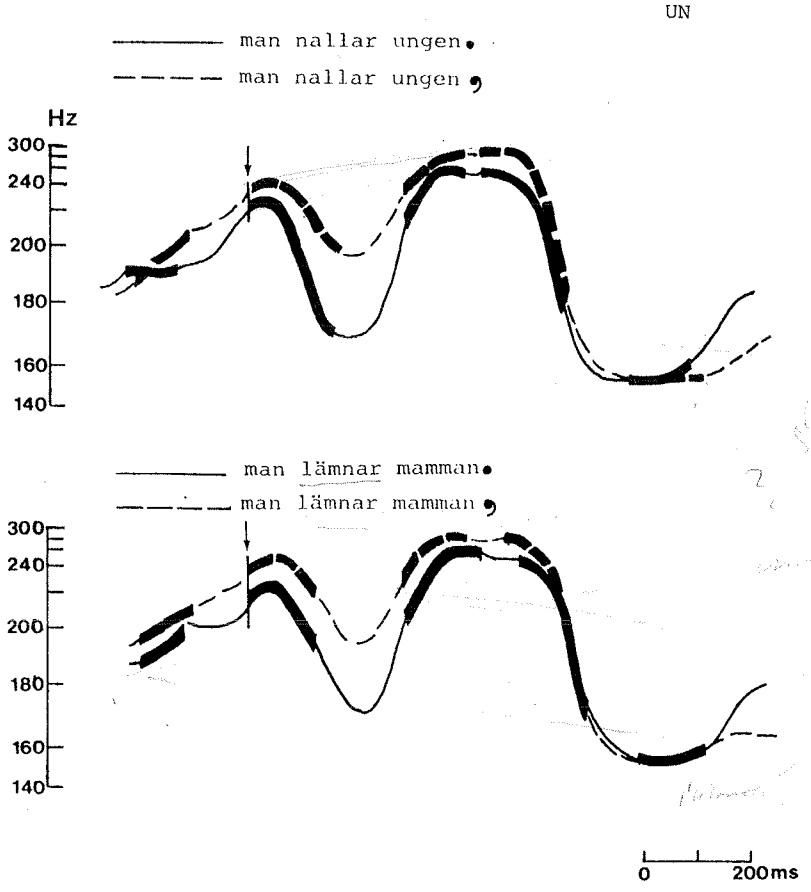


Figure 13. The effect of clause boundary. Typical Fo contours of the first clause of a double clause sentence with a physical pause at the clause boundary /,/ and of the corresponding single clause sentence /./ by a female informant (UN). The line-up point is at the CV-boundary of the first stress group.

that of the second clause of a double clause sentence.

2. The Fo values of the second clause of a double clause sentence (both versions) are all comparable with those of the single clause sentence.

3. The Fo end point of a single clause sentence has approximately the same value as that of a double clause sentence. Interestingly, the end point value of the first clause of of a double clause sentence (with an intruding physical pause) appears to be as low as in the single clause sentence (see Figure 13). The corresponding end point value of the first clause of a double clause sentence - without a physical pause - is noticeably higher.

The difference between a single clause utterance and the first clause of a double clause utterance with a pause is in the overall Fo course, which apparently involves both clauses of a double clause utterance. This will give higher Fo values to the first clause of the double clause utterance than to the single clause utterance, where the domain of the Fo course is this very clause. In both cases there is an Fo drop to the same end point value, which introduces a local break into the overall course of the double clause utterance before the internal pause. But immediately after the pause Fo is reset to the course of the downdrift. Also in the non-pause version of the double clause utterance there is a local Fo drop at the clause boundary before Fo is reset to the original course. This local Fo drop does not, however, reach the same low value as in the pause version but is enough to function as a prosodic boundary signal.

6.2 Discussion

It might be assumed that the results of the present experiment are comparable to those of the above experiments, where utterance length was varied. Disregarding the local disturbance introduced by the clause boundary, the Fo values of a single clause utterance of the actual type compared to a double clause utterance would then be proportional to those of a shorter versus

Table V. The effect of clause boundary. Fo means and standard deviations in Hz of successive minima and maxima of the test material recorded in experiment five by the female informant UN. Clause boundaries within an utterance are marked by a comma if there is a physical pause. Phrase accent maxima of each utterance are in italics.

Speaker UN

		min	max	min	max	min	min	max	min	max	min	max	min	n
Man `lämnar `mamman , men man `nallar `ungen														
neutral	$\bar{x}$	193	243	186	*281*	150	183	207	164	214	165	*242*	153	8
focus	s	7,6	16,7	7,3	13,6	2,7	4,6	9,2	3,5	11,9	6,5	11,0	2,7	
Man `lämnar `mamman men man `nallar `ungen														
neutral	$\bar{x}$	198	249	193	*291*	166	181	204	156	206	164	*239*	156	8
focus	s	8,0	11,0	5,3	15,1	3,2	6,4	5,6	10,8	8,2	5,2	11,2	5,8	
Man `lämnar `mamman														
initial	$\bar{x}$	186	230	169	*253*	150								8
focus	s	6,6	14,1	3,8	14,0	3,2								
Man `lämnar `mamman														
final	$\bar{x}$						176	194	165	198	161	*226*	154	8
focus	s						6,9	10,2	4,6	7,0	3,2	9,4	4,4	
Man `nallar `ungen , men man `lämnar `mamman														
neutral	$\bar{x}$	195	240	193	*286*	152	184	209	166	208	168	*236*	158	8
focus	s	5,3	9,6	8,9	8,2	3,7	3,2	7,4	3,2	9,2	3,8	8,8	2,6	
Man `nallar `ungen men man `lämnar `mamman														
neutral	$\bar{x}$	198	245	195	*299*	167	188	206	161	201	163	*236*	157	8
focus	s	7,0	7,1	8,5	11,6	5,3	7,0	5,6	5,2	5,2	4,6	4,2	4,6	
Man `nallar `ungen														
initial	$\bar{x}$	182	219	169	*245*	149								8
focus	s	8,1	6,7	5,3	10,0	2,4								
Man `nallar `ungen														
final	$\bar{x}$						181	199	166	204	161	*233*	153	8
focus	s						6,9	11,2	5,0	12,4	4,2	9,6	2,6	

longer single clause utterance. Judging from the present, preliminary experiment, this turns out not to be the case. It should be remembered, though, that in the above experiments the number of phrases was usually held constant when varying the number of stress groups, while in this experiment the number of phrases is varied at the same time. The results here are most readibly comparable with those of a single stress group utterance with one phrase versus a two stress group utterance containing two phrases (cf. experiment two) except for the starting point value.

The results concerning the end point values are interesting to discuss in the light of a hypothesis proposed recently by Pierrehumbert (1980). This discussion will be postponed to the next section.

7 REVISION OF THE SWEDISH INTONATION MODEL

Rather than present a complete model for Swedish intonation to replace the one described in Section 1, I will only try to show in what direction the current model needs revising in order to accomodate the results of the experiments of this study. But first there are some problems of Fo scaling to be resolved.

7.1 Problems of Fo scaling

The experiments of this study all indicate that the Fo end point is the least variable Fo point of an utterance across variations in focus location, utterance length, syntactic composition and attitude (degree of involvement).

This lowest Fo value of an utterance is of course speaker-dependent. As was pointed out above (cf. 5.2) it can be considered a reference for other Fo values. It has been called the Fo floor or Fo bottom of a speaker's voice range (cf. Gårding 1979, Pierrehumbert 1980). In normal speech the voice does not reach up to a corresponding Fo ceiling or Fo top. The Fo peak

value of an utterance is highly variable; it varies with utterance length and especially with attitude (involved - detached).

These facts - that there is a constant floor and a variable ceiling - should be expressed in an intonation model. The traditional Fo scales - the linear Hz scale or the logarithmic semitone scale - do not express this assymetry directly. As pointed out by Pierrehumbert (1980, p. 175), the values of a semitone scale are asymptomic to 0 Hz and not to the Fo bottom of a speaker's voice range.

To capture these facts - constant floor, variable ceiling - Pierrehumbert (1980) devised a scale where Fo values are expressed as baseline units above the baseline. The baseline is equal to the bottom of the speaker's voice range. Each value is scaled as $(Fo - Fo_{bottom}) / Fo_{bottom}$. This formula thus encompasses both a difference between the actual Fo value and the bottom value and a division of this difference by the bottom value. The combination of these two operations is shown to be superior to choosing just one of them. The formulation of the rule for accentual downsteps in American English becomes simple and intuitively correct in this scale compared to a semitone scale.

In Pierrehumbert's conception the baseline declines slightly over an utterance. This means that the lowest Fo value the speaker is disposed to reach would be higher earlier in an utterance. Her proposal is based on an experiment on an utterance containing two phrases with a prosodic boundary in between but without a physical pause (Pierrehumbert 1980, chapter 3). The lowest Fo value at the utterance internal boundary was higher than at the utterance final boundary, which is followed by a pause. This is interpreted as support for a declining baseline.

This is worth discussing in the light of the results of experiment five where two versions of a sentence, consisting of two

clauses, were recorded: one with a physical pause at the clause boundary and another without a pause but yet with a clear boundary signal. In the non-pause version the results are very similar to those of Pierrehumbert's experiment; the lowest Fo value at the utterance internal boundary is higher than at the utterance final boundary. But when a physical pause is introduced between the two clauses of the sentence, there is no difference in the lowest Fo value at the two boundaries. My interpretation of this experiment is that the speaker is indeed disposed to reach the bottom of the voice range not only at the end of an utterance but also earlier, e.g. before a pause within an utterance.

In the light of this experiment the evidence for a declining baseline as suggested by Pierrehumbert does not seem compelling. I will instead make the simpler assumption that the baseline is flat; i.e. a speaker is disposed to reach the Fo bottom at any point of an utterance. Note that this conception of the baseline as potential rather than actual is distinct from the one given earlier (see Section 1).

There may be some problems in defining the exact value of a speaker's Fo bottom. But the important thing is to establish that the Fo bottom is the reference value for the Fo events. This is not taken into consideration in a Hz or a semitone scale.

To express Fo values in baseline units above the baseline is a way of normalizing both different Fo ranges within one and the same speaker and voices of different pitches, e.g. male and female voices. To illustrate this stylized Fo contours for the male and female speakers of experiment two are given in Figure 14a. These show successive minima and maxima of the five stress group utterance expressed as baseline units above the baseline (cf. Figures 5 and 6). A corresponding comparison between involved and detached attitude within the same speaker is made in Figure 14b. The suitability of this scale for intra and inter-individual comparisons needs to be examined more thoroughly.

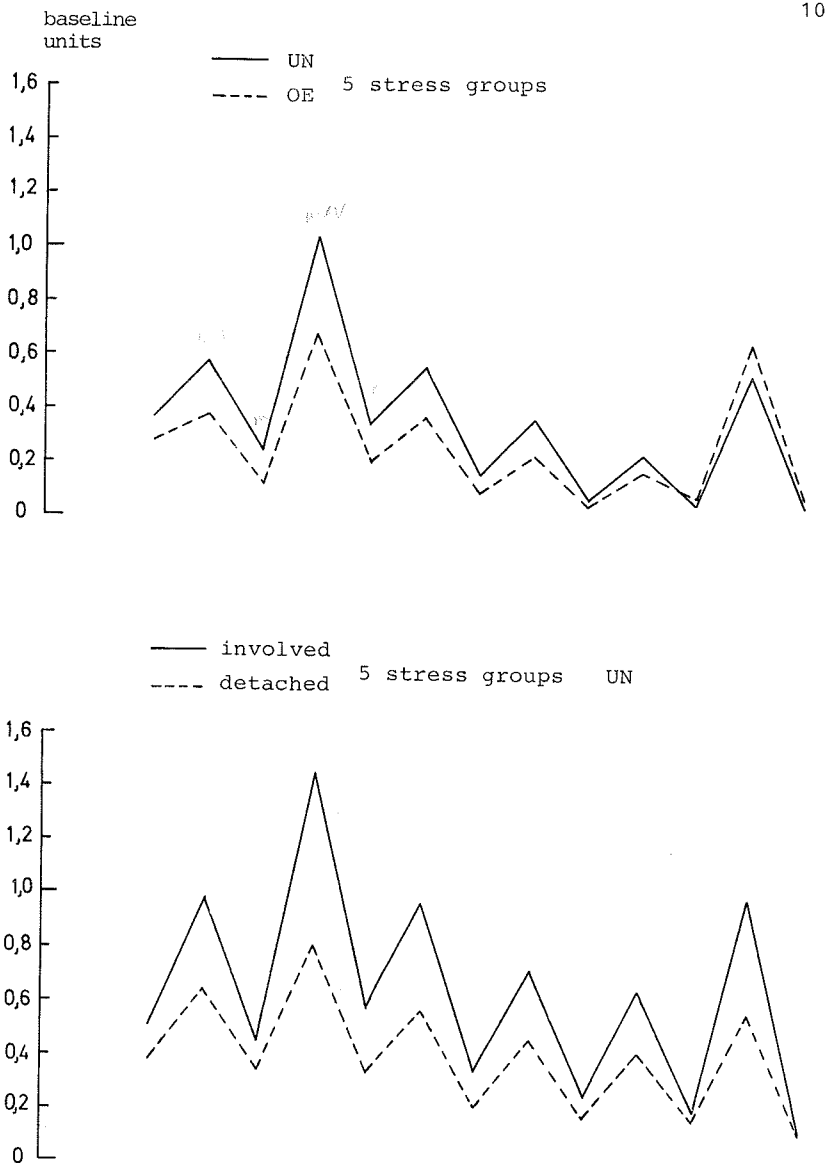


Figure 14. Normalization of Fo. Stylized Fo contours - means of successive Fo minima and Fo maxima expressed in baseline units above the baseline: (a) five stress group utterance by the female (UN) and male (OE) informants, (b) five stress group utterance of an involved and a detached version by the same informant (UN).

7.2 Outline of a revised model

In the current model for Swedish intonation (Bruce & Gårding 1978, Gårding & Bruce 1981) pitch relations are not expressed according to a specific type of scale. Only qualitative comparisons could therefore be made. In the present outline of a revised model I will try to make quantitative comparisons of pitch relations within and between speakers with reference to baseline units above the baseline.

In the revised version of the model for Swedish intonation a prosodic transcription of the same kind as in the current version (cf. Section 1) is presupposed as the input. A basic component of the revised version is the division of an utterance into prosodic phrases, each containing a number of word accents and a terminating phrase accent. The step by step generation of pitch layers, i.e. of sentence intonation followed by accentuation as in the current version, will be replaced by left-to-right implementation of the phonetic rules converting the prosodic input notation to an Fo contour. This underlines the interaction between accentuation and intonation proper more effectively than the layered implementation does. Word accents will still be represented as combinations of high and low points and phrase accents as low plus high points, where the low point is usually shared by word accent and phrase accent. These points will be converted directly to actual Fo values by the phonetic rules sketched below and connected to form the actual Fo contour. Although represented as high-lows and low-highs, pitch relations between successive accents expressed as pitch changes will play a larger role in the revised version than in earlier versions (cf. Bruce 1977, Bruce & Gårding 1978, Gårding & Bruce 1981). The old controversy of pitch levels versus configurations should be reexamined to see if it is still a real issue or merely a point of view (cf. e.g. Hadding-Koch 1961, pp. 44-46).

The following regularities are to be expressed in a revised version of the Swedish intonation model.

Utterance length. One variable that has been under examination in the present study is utterance length. Relevant expressions of utterance length were found to be number of stress groups and division into prosodic phrases. Unless stated otherwise, the regularities below concern an utterance consisting of two prosodic phrases.

The point of departure is a two stress group utterance, where the two phrases contain one stress group each. In this case the two minima surrounding the first phrase accent peak have nearly the same Fo value. A syntactic expansion to the right, i.e. an increase in the number of stress groups in the second phrase of the utterance reveals the following relationship. The minimum preceding the phrase accent maximum stays nearly constant, while the minimum after the maximum increases with the number of stress groups. It is therefore possible to assume that the accent minimum following the phrase accent peak has a basic value that is copied from the accent minimum preceding the peak and that is typical of the two stress group utterance, where the two phrases contain one stress group each. To this basic value is then added an extra value that increases with the number of stress groups contained in the phrase.

Also the phrase accent peak, which is identical to the phrase-initial word accent maximum, increases with the number of stress groups in the second phrase in a similar way to the following, phrase initial word accent minimum. The phrase accent peak is assumed to have a basic value that is found in the two stress group utterance and that is added to the preceding minimum. The same extra value as for the following accent minimum is then added to the basic value depending on the number of stress groups in the second phrase. This means that the step size from the first phrase accent peak to the following valley is more or less constant.

An equivalent way of expressing these facts to give more emphasis to the interdependence between the phrase initial accent maximum (=phrase accent peak) and the corresponding accent

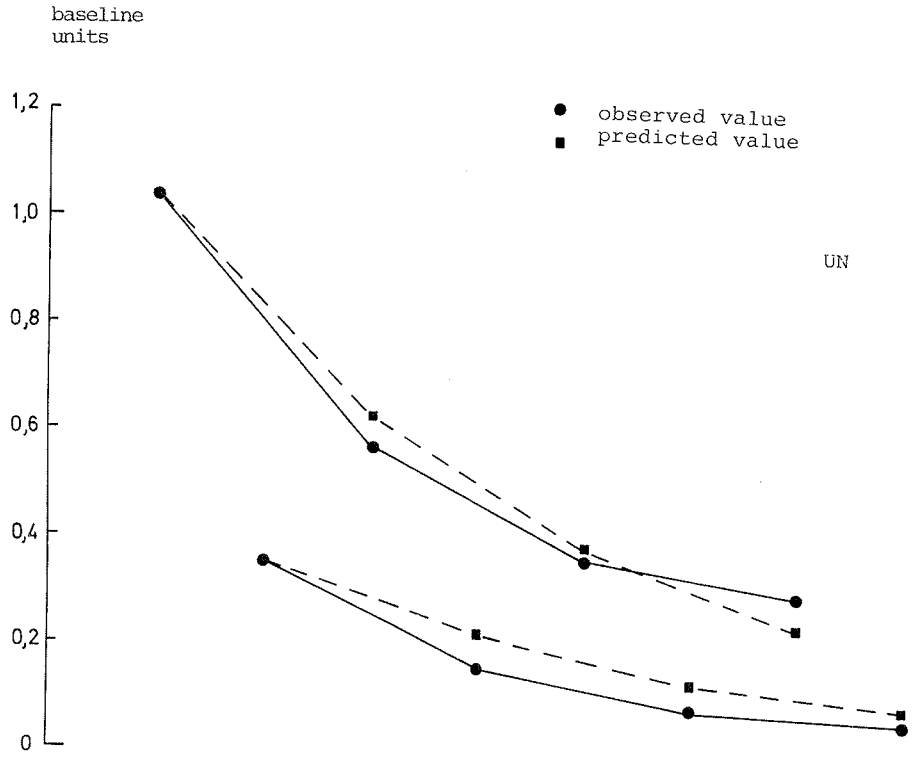


Figure 15. Prediction of Fo course. Observed and predicted toplines and baselines of the neutral version of a five stress group utterance - the second phrase except for the second phrase accent - by a female informant (UN). The initial maximum (topline) and initial minimum (baseline) values are given. Each Fo value is 0.6 x the preceding Fo value (for both maxima and minima).

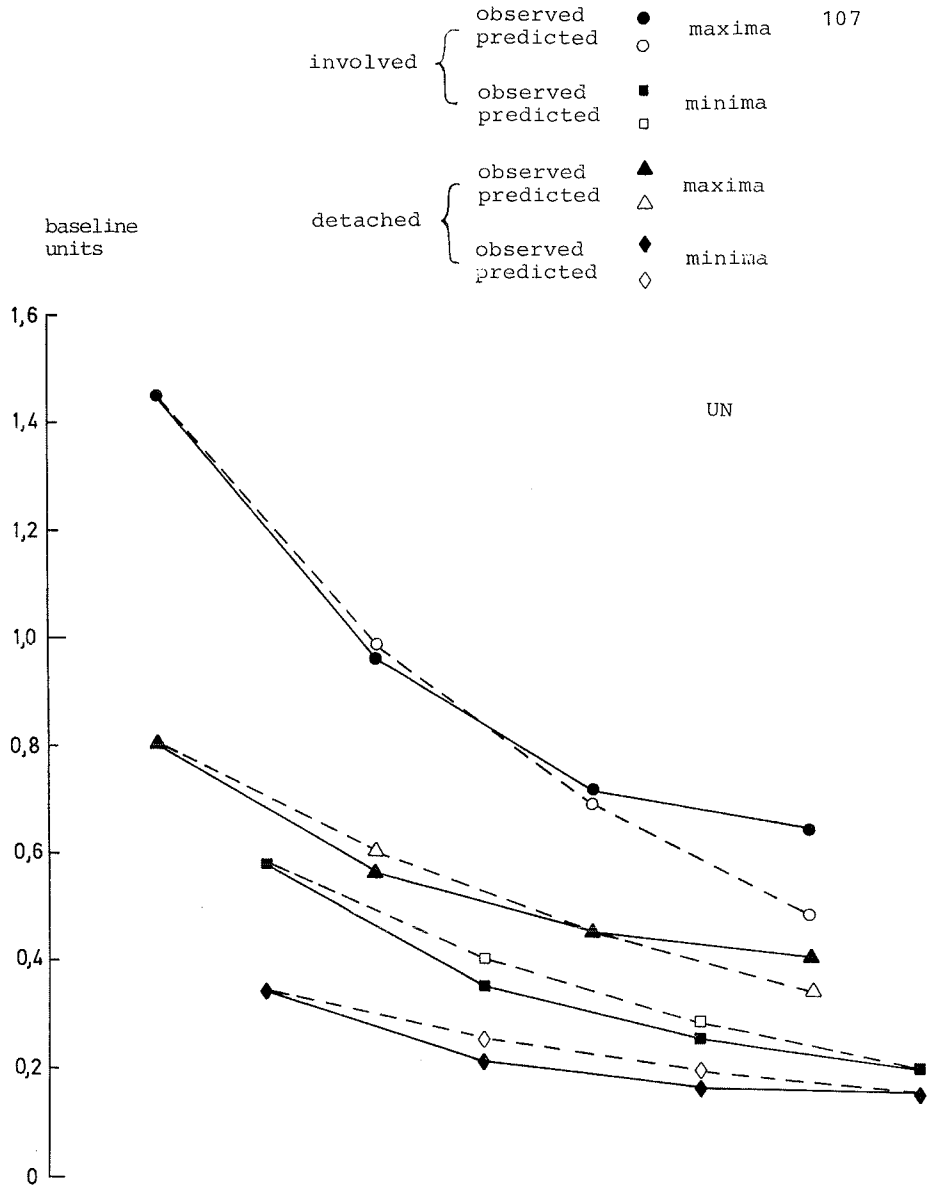


Figure 16. Prediction of Fo course. Observed and predicted toplines and baselines of the involved and detached version of a five stress group utterance by a female informant (UN). The initial maximum (topline) and minimum (baseline) are given for both degrees of involvement. Each Fo value is 0.7 x the preceding value for the involved version, and 0.75 x the preceding value for the detached version (for both maxima and minima).

minimum would be the following. The phrase initial accent maximum has a basic value to which is added an extra value that increases with the number of following accents (see above). The following accent minimum value will then result from a subtraction of the basic value from the preceding maximum value.

Downstepping. It is hypothesized that for equally prominent, successive accents, accent maxima following the first phrase accent peak and accent minima following the first minimum of the second phrase are a constant ratio of the immediately preceding values. In Figures 15 and 16 an attempt has been made to fit curves to the declining toplines and baselines of the second phrase of a five stress group utterance, excluding the last phrase accent (which is insensitive to utterance length) according to the formula: each Fo value of a maximum or a minimum is a constant ratio ($0 < k < 1$) of the corresponding, immediately preceding Fo value. The first value of each topline and each baseline - the Fo maximum value of the first phrase accent and the following accent minimum - is given. Figure 15 shows observed and predicted values of the neutral version of the five stress group utterance from experiment two (informant UN), while the data points of Figure 16 are from experiment four (involved - detached). The difference between observed and predicted values is fairly small on the whole. The k value chosen is the same for both maxima and minima within the same degree of involvement and approximately the same for both degrees of involvement. Therefore the relation might be expressed as amount of Fo fall between successive accentual downsteps.

The most obvious difference between observed and predicted values in the Figures concerns the last accent maximum - the one preceding the last phrase accent peak. In both Figures the observed Fo value is noticeably higher than the predicted value.

In this connection it is interesting to draw attention to the fact that accent rises appear to be nearly constant in interval whether they occur in the upper or lower part of a speaker's

frequency range, while accent falls decrease in interval with position in the overall range (cf. Tables II and IV). Therefore if we assume as before that each Fo minimum is a constant value of the preceding minimum but instead that each Fo maximum is a constant interval above the preceding minimum, the predicted Fo maximum values in Figures 15 and 16 will in fact be closer to the corresponding observed values.

Another possibility would be that successive downsteps are a constant interval expressed in a semitone scale. It is clear, however, that the step size in semitones decreases with position in the frequency range too and that this scale does not reveal any constancy here (cf. Figures 5 and 6).

The formula discussed here can also be applied to Fo values of the first phrase. Apparently, the value of the accent minimum following the starting point in the first phrase can be expressed as a constant ($0 < k < 1$) of the starting point value. If there is more than one stress group in the first phrase of an utterance, the k value for minima seems to approach one; i.e. baselines and consequently also toplines can be considered nearly horizontal.

Upstepping. The relation between an accent peak and a succeeding phrase accent peak seems to be easiest to express as a ratio of amount of Fo rise. The range of the rise is wider for the phrase accent peak, often twice as wide or more, but the exact ratio seems to be a free choice within certain limits.

The relationship between the first and the last phrase accent peak should also be expressed as a ratio of amount of Fo rise. The last phrase accent peak appears to be independent of utterance length. The basic relationship between the two peaks is therefore to be found in the two stress group utterance. The actual Fo maximum value of the last phrase accent is higher in the two stress group utterance than in e.g. the five stress group utterance. But if we take into consideration that the immediately preceding Fo minimum is also higher in the two

stress group utterance, the Fo rise of the phrase can be shown to be nearly constant in range. This means that phrase accents tend to have equal height basically.

Involvement. Another variable that has been examined in this study and that has already been considered in this section is degree of involvement.

What has been said above about utterance length, downstepping and upstepping, is true of both an involved and a detached attitude within the same speaker.

Within one and the same attitude (degree of involvement) starting points and end points are constant and independent of the number of stress groups contained in the utterance or in the phrase. Starting points vary with attitude, though, as do all other Fo points except end points. The most obvious expression of an involved attitude is therefore an increase in the overall Fo range. The relationship between all Fo points of two different degrees of involvement might possibly be expressed as a mere scaling factor. This would be true within one and the same speaker and perhaps also between speakers.

Figure 17 shows a preliminary test of the hypothesis that the same Fo pattern is maintained with a change in degree of involvement. The difference is assumed to be one of a mere scaling factor. The similarity between observed and predicted values in Figure 17 is less striking than in Figures 15 and 16, but there is still a fair amount of agreement. The most apparent difference is that predicted values undershoot observed values at the Fo maximum of the two phrase accents. This discrepancy might be due to a worse fit in the upper region of the Fo range. On the other hand, the agreement for the Fo maximum preceding the first phrase accent maximum is good, although the absolute Fo value of this point is higher than that of the last phrase accent maximum. It therefore seems plausible that the discrepancy has to do with a somewhat different implementation of phrase accents in the two attitudes.

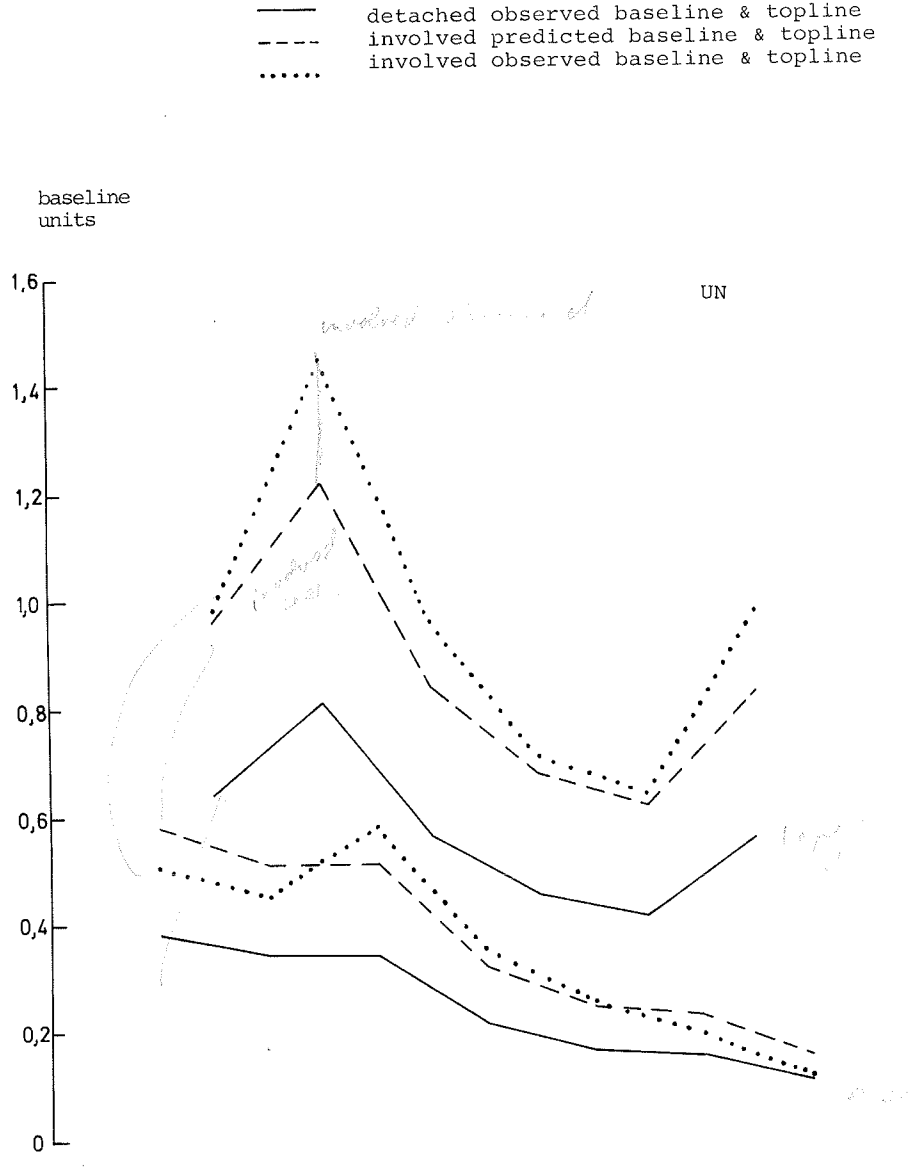


Figure 17. Prediction of Fo range. Observed topline and baseline for the detached version and observed and predicted toplines and baselines for the involved version of the five stress group utterance by a female informant (UN). Each predicted involved Fo value is 1.5 x the Fo value of the detached version.

Another possibility, which has yet to be tested, is that the Fo range is regulated by the Fo rise, which appears to be constant in range within the same degree of involvement. An increase in the overall range could therefore be expressed as an increase by a certain factor in the Fo rise only.

There is also some evidence for a dependence between the range of the Fo rise and the Fo starting point value. It may be hypothesized that the range factor is sensitive to the amount of increase in the starting point value, although this has to be further substantiated. There would then be a constant increase in the range of the Fo rise for each addition of a certain value to the Fo starting point value. The phrase accent rise is also sensitive to the range factor. This is true of both the basic value of a phrase accent peak as well as the added extra value, which depends on the number of upcoming accents.

A conclusion to be drawn from the present study is that the Fo course of an utterance is formed by the relations between successive local Fo excursions for the accents within and between the prosodic phrases that make up the utterance. In these local Fo excursions the rises and the falls can be assumed to have different functions. The Fo rise, the range of which tends to be constant for successive accents within an utterance, is seen as an expression of degree of involvement from the speaker. This is true at least of a neutral utterance, where no specific focussing is presupposed. The range of the Fo fall, on the other hand, tends to decrease for successive accents within the phrase for a declarative type of intonation. The relation between successive Fo falls is assumed to have the function of expressing statement versus question or perhaps rather definiteness versus indefiniteness.

7.3 Conclusion

The following main points, which are to be incorporated in a revised intonation model, can be established from the present study.

- The Fo end point is the least variable point of an utterance. It is the bottom of a speaker's Fo range and is a reference for all other Fo values of an utterance. This is expressed in Pierrehumbert's baseline units above the baseline.
- The overall Fo course of an utterance is expressed as the relations between successive, local Fo excursions for accentuation, with Fo plateaus in between.
- The Fo rise of such a local excursion for accentuation tends to have the same range independently of whether it occurs in the upper or lower part of a speaker's Fo range.
- There are two kinds of Fo rise, the above mentioned accent rise and the phrase accent rise, which has a wider range than the accent rise. The phrase accent rise terminates a prosodic phrase.
- The division of an utterance into prosodic phrases is important for Fo implementation.
- The range of the Fo fall of a local excursion for accentuation appears to decrease for successive accents of a declarative intonation type. This Fo downdrift can be described by a local rule, where each Fo minimum value is a constant ratio ($0 < k < 1$) of the immediately preceding Fo value. The rule operates within the phrase.

In the transition between two prosodic phrases the following regularities occur:

- The Fo peak after the phrase accent rise, which is identical with the first accent maximum of the following phrase, has a basic value, to which may be added an extra value that increases with the number of following stress groups in the phrase.

- The first accent minimum of the second phrase has a basic value that is the same as that of the preceding accent minimum of the first phrase, plus an extra value that increases in the same way as the preceding maximum.

The added extra value for the phrase initial maximum and minimum is an anticipation of utterance length expressed as a function of the number of following stress groups within the phrase. These higher Fo values in a longer compared to a shorter utterance will persist for the following accent maxima and minima in the phrase.

- The degree of involvement of the speaker is expressed as a variation in the Fo range of an utterance. The same Fo pattern is maintained across variations in involvement. The difference between detached and involved seems to be a matter of scaling. There is also a dependence between the Fo range and the value of the Fo starting point, which tends to increase with involvement but is otherwise constant and higher than the Fo end point. The exact nature of this relationship has yet to be established.

8 REFERENCES

Atkinson, J. (1973) Aspects of intonation in speech: implications from an experimental study of fundamental frequency. Unpublished doctoral thesis. University of Connecticut

Brown, G., Currie, K. and Kenworthy, J. (1980) Questions of intonation. London: Croom Helm

Bruce, G. (1977) Swedish word accents in sentence perspective. Lund: Gleerup

Bruce, G. (1981) Om prosodisk koppling i svenska. In Fries, S. and Elert, C.-C. (Eds.) Svenskans beskrivning 12, pp. 64-78. Umeå university

Bruce, G. and Gårding, E. (1981) A prosodic typology for Swedish dialects. In Gårding, E., Bruce, G. and Bannert, R. (Eds.) Nordic prosody, pp. 219-228. Department of linguistics, University of Lund

Collier, R. (1975) Physiological correlates of intonation patterns. Journal of the acoustical society of America 58: 249-255

Fujisaki, M., Hirose, K. and Ohta, K. (1979) Acoustic features of the fundamental frequency contours of declarative sentences. Annual bulletin of the research institute of logopedics and phoniatrics, University of Tokyo 3: 163-172

Gårding, E. (1979) Sentence intonation in Swedish. Phonetica 36: 207-215

Gårding, E. (1982) Contrastive prosody: a model and its application. Studia linguistica XXXVI: 146-165

Gårding, E. and Lindblad, P. (1973) Constancy and variation in Swedish word accent patterns. Working Papers, Phonetics laboratory, Lund university 7: 36-110

Gårding, E. and Bruce, G. (1981) A presentation of the Lund model for Swedish intonation. In Fretheim, T. (Ed.) Nordic prosody II, pp. 33-39. Trondheim: Tapir

Gårding, E., Botinis, A. and Touati, P. (1982) A comparative study of Swedish, Greek and French intonation. This issue

Hadding-Koch, K. (1961) Acoustico-phonetic studies in the intonation of Southern Swedish. Lund: Gleerup

t'Hart, J. and Collier, R. (1979) On the interaction of accentuation and intonation in Dutch. Proceedings of the ninth international congress of phonetic sciences. Volume II: 395-402. Institute of phonetics, University of Copenhagen

Hombert, J.-M. (1974) Universals of downdrift: their phonetic basis and significance for a theory of tone. In Leben, W. (Ed.) Studies in African linguistics, supplement 5: 169-183

Lieberman, P. (1967) Intonation, perception and language. Cambridge: MIT press

Lindblom, B. (1978) Final lengthening in speech and music. In Gårding, E., Bruce, G. and Bannert, R. (Eds.) Nordic prosody, pp. 85-101. Department of linguistics, University of Lund

Lindblom, B., Lyberg, B. and Holmgren, K. (1976) Durational patterns of Swedish phonology: Do they reflect short term motor memory processes. Mimeographed. Institute of linguistics, University of Stockholm

Maeda, S. (1976) A characterization of American English intonation. Unpublished doctoral thesis. Massachusetts institute of technology

Ohala, J. (1978) Production of tone. In Fromkin, V. (Ed.) Tone: A linguistic survey, pp. 5-39. New York: Academic press

Pierrehumbert, J. (1979) The perception of fundamental frequency declination. Journal of the acoustical society of America 66 (2): 363-369

Pierrehumbert, J. (1980) The phonology and phonetics of English intonation. Unpublished doctoral thesis. Massachusetts institute of technology

Sorensen, J. and Cooper, W. (1980) Syntactic coding of fundamental frequency in speech production. In Cole, R.A. (Ed.) Perception and production of fluent speech, pp. 399-440. Lawrence Erlbaum Associates, Hillsdale, New Jersey

Thorsen, N. (1978) Aspects of Danish intonation. In Gårding, E., Bruce, G. and Bannert, R. (Eds.) Nordic prosody, pp. 23-32. Department of linguistics, University of Lund

Thorsen, N. (1980) Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish. Annual report of the institute of phonetics, University of Copenhagen 14: 1-29

Thorsen, N. (1981) Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish: supplementary data. Annual report of the institute of phonetics, University of Copenhagen 15: 13-47