LUND UNIVERSITY
DEPARTMENT OF LINGUISTICS
General Linguistics
Phonetics

WORKING PAPERS

22 • 1982

LUND UNIVERSITY
DEPARTMENT OF LINGUISTICS
General Linguistics
Phonetics

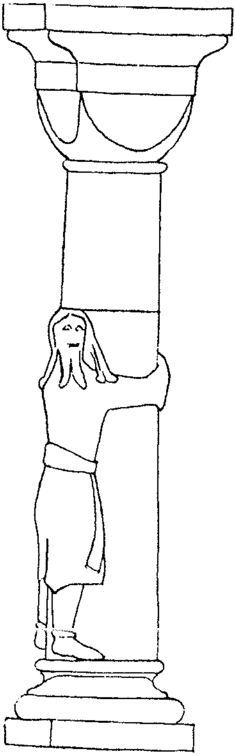
1666.

Eva Gårding

Mal
Rättelser & kommentarer

WORKING PAPERS
22 • 1982

ISSN - 0280 - 526X



Department of Linguistics and Phonetics
Helgonabacken 12
S-223 62 LUND

This issue has been edited by Eva Gårding
and Anders Löfqvist

ANNUAL REPORT

This is a summary of the main events since the publication of Working Papers 21, June 1981.

The new location of the department near the hospital has improved our opportunities for collaboration with the medical departments interested in speech research. To provide a common platform for exchange of knowledge and ideas, Sören Fex of the Institute of Logopedics and Phoniatrics and myself started a joint seminar around the themes Language * Speech * Sound * Hearing.

The program of this seminar, in English translation, has been as follows

21.9	Research applied to the communication problems of the hard of hearing	Arne Risberg, KTH, Stockholm
5.10	What we know and what we do not know about prosody	Gösta Bruce & Eva Gårding
19.10	Voice for speech and singing	Johan Sundberg, KTH, Stockholm
2.11	Speech physiology and linguistic units	Joseph Perkell, Research Laboratory of Electronics, MIT
16.11	Eye movements and speech disabilities	Nils Gunnar Henriksson, Ilmari Pykköö, Lucyna Schalén, Carsten Wennmo, Laboratory for Balance Research, University Hospital, Lund
14.12	Laryngeal kinetics	Bertil Sonesson, Department of Anatomy, University Hospital, Lund
1.2	Music as communication	Sören Nielzén, Department of Psychiatry, University Hospital, Lund
15.2	Voice quality: one speaker's phonology is another's pathology?	Karin Andersson, Gösta Bruce, Anders Löfqvist
1.3	Audiology I: Clinical and therapeutical aspects of loss of hearing	

8.3	Audiology II: Presentation of research and laboratory	Hans Eric Holst, Jarle Aursnäs, Ingvar Härlid, Department of Oto-Rhino-Laryngology, University Hospital, Lund
15.3	The need for pronunciation norms	Birgitta Heyman, The Drama School, Stockholm
19.4	Right-hemisphere language functions	Christina Hedqvist, Karen Johnson, Mc Gill, Montreal
3.5	Symbol, sign, language, speech, distinctive features	Bertil Malmberg
17.5	Communication disturbances due to trauma	Inger Vibeke Thomsen, State hospital, Copenhagen

The Friday seminar for discussions of research in progress has continued as before with the following program:

7.9	Research paradigms in phonetics	
14.9	Reports on dissertations	
28.9	Bee-phonetics	Gabriella Koch
	Techniques in discussion	Gabriella Koch Ulrika Nettelbladt Barbara Prohovnik
30.10	Speech production models	Sidney Wood
3.11	Method for measuring laryngeal resistance	Eva Holmberg, Inst f lingvistik, University of Stockholm
13.11	Experiments with the intonation model	Gösta Bruce
20.11	Direct and indirect measurement of subglottal pressure Intrasegmental timing	Anders Löfqvist
27.11	Acoustic and perceptual analysis of nasal consonants	David House
	Testing prosodic functions in aphasics	Christina Hedqvist
4.12	Physiological and perceptual analysis of stress in Greek	Antonis Botinis
11.12	Constancy and variation in Swedish interrogative intonation	Anne-Christine Bredvad-Jensen
18.12	Finland-Swedish intonation	Kerstin Tevajärvi
19.3	A review of Pierrehumbert's dissertation: The phonology and phonetics of English intonation	Gösta Bruce

26.3	Experiments with accelerated speech	Richard Schulman, Department of Linguistics, Stockholm
2.4	What is LPC?	Bengt Mandersson
16.4	A tonal grid as an expression of sentence intonation and its usefulness for the study of prosodic transfer	Paul Touati
30.4	The intonation workshop in Paris	Eva Gårding
7.5	On interrogative intonation Interspeaker variation	Anne-Christine Bredvad-Jensen
	Report on child-language research in Stanford	Ulrika Nettelbladt
14.5	Tests differentiating various forms of dyslexia	Kerstin Nauclér
4.6	Lund - Copenhagen Microsymposium on macroproblems Phonology - Phonetics	Nina Thorsen, Paul Touati, Gunnel Tottie, Thore Pettersson, Eva Gårding

Two graduate courses were given, one on speech physiology in the fall of 1981 by Anders Löfqvist, and one on vowel articulation in the spring of 1982 by Sidney Wood.

On the undergraduate level we have had the regular courses. A new group of logopedics students started their two-term phonetics training in January 1982. There are at present 33 first term students, 17 second, 11 third, 4 fourth, and 10 graduate students.

Bengt Mandersson defended his thesis 'Resolution of convolved signals in biomedical applications' at the Lund Institute of Technology in January. Sidney Wood's dissertation 'X-ray and model studies of vowel articulation' WP 23, is now in the press.

Gösta Bruce arranged the second all-Swedish symposium, 'From Text to Prosody', in September 1981. A collection of seminar papers by second-term students around the theme *Prosody* was published in the fall by Gösta Bruce as number 6 of our Swedish series, *Praktisk lingvistik*. Robert Bannert has substituted for Sven Öhman in Uppsala. He also led a seminar here, entitled *Interlanguage*.

Kurt Johansson has been appointed lektor (reader). Mona Lindau visited us in the fall and helped plan for a new project. Anders Löfqvist was a temporary consultant at the Royal Institute of Technology.

The *Röst- och talvård* section (Voice and speech therapy), housed in the same building as our department, has been referred to our department. We hope that this will be the beginning of joint activities in teaching as well as research, not just an administrative manoeuvre.

It has been recognized by the administrative board for the humanities that our laboratory needs special funds.

Finally, it is a special pleasure to report that Lund University will confer an honorary degree to ILSE LEHISTE on May 28th, in the cathedral used for a similar occasion in Ingmar Bergman's *Wild Strawberries*.

Lund, May 1982

Eva Gårding

CONTENTS

TEMPORAL AND TONAL CONTROL IN GERMAN

Robert Bannert

ABSTRACT

A fundamental issue for a model of prosody is whether or not the phonological features expressed by time (duration) and pitch (fundamental frequency) are independent variables. Vowel duration in Central Swedish has been said to be dependent on the fundamental frequency change associated with that segment. However, there is evidence that an Fo-dependent model for duration does not adequately account for data on Swedish and Danish. This report examines the relationship between fundamental frequency change and duration in German. Two statistical tests clearly indicate that time and pitch should be treated independently in German prosody. Finally, a model for German prosody is sketched based on these findings.

1. INTRODUCTION

In every model that generates speech utterances, be it articulatorily or acoustic-auditorily oriented, the two dimensions time and fundamental frequency (Fo) play an important role. To produce an utterance takes a certain length of time. Furthermore, every utterance has its own temporal structure and a tonal structure as well which, as it seems, exists in parallel.

Up to now a number of models for duration and intonation in several languages have been developed (see references). These models attempt to generate either the durational or the tonal structure of an utterance to be observed in the signal. The question that then arises is how to arrive at a more complete utterance containing both structures. Assuming the independent generation of the temporal and tonal features of an utterance, this goal can be reached by simply adding both structures together. This, however, leads to problems of coordination which have to be solved. It should be quite obvious that these two structures generated independently of each other will have to be changed or adjusted when put together. Tonal adjustments of the basic Fo-contours as a consequence of factors of time and sentence position were described by Bruce (1977, 74 ff.).

Recently Lyberg (1981) made an attempt to handle these two prosodic dimensions simultaneously. Taking up an idea of, among others, Öhman et al. (1979), he presented an Fo-dependent model for segment duration where segment durations are the automatic result of Fo-movements. The more complex such a movement, i.e. the larger the Fo-change or the Fo-difference over the segment, the larger the segment duration has to be. The alleged causal relationship between Fo-change and segment duration is expressed by the formula $D = k \cdot \Delta Fo$, where D = segment duration, k = constant, and ΔFo = Fo-difference. Comprehensive counter evidence against an Fo-dependent model for duration, however, has been presented by Bruce (1981) and Bannert (1982).

The aim of this paper[1] is to answer the crucial question from the outset as to which relationship exists between the temporal and tonal dimension in a model of prosody[2] for German. It is of course decisive to know whether segment duration must be treated independent of intonation or not. The answer to this question will have fundamental consequences for the design of any model of prosody.

2. MATERIAL AND ANALYSIS

In order to be able to answer the question formulated above, the following strategy was chosen. In a sentence with three accents the tonal structure was varied while the temporal structure (the segments) was kept constant. The temporal structure was also varied while the tonal structure was kept constant. This was done in the following way:

(1) <u>Tonal variation (segments constant)</u>

The statement

Der $\underset{1}{\underline{\text{Lahme}}}$ hat die $\underset{2}{\underline{\text{Maler}}}$ $\underset{3}{\underline{\text{gemahnt}}}$.

("The limping man reminded the painter.") containing three accent positions (underlined and numbered) was embedded in four different prosodic contexts. By this procedure, a tonal difference as large as possible between the four contexts (versions) of the test sentence was to be produced. The starting or neutral version (1) of the sentence contains three equivalent accents. In the other three versions, the sentence accent would be moved by focusing to the first (version 4), second (version 2), and third (version 3) accent position. The third accent position served only as a dummy, in order to be able to move the sentence accent away from positions 1 and 2. Only these two positions were investigated. The sentence presented above served also as a means of calibration throughout the recording session. As the calibration sentence is semantically quite natural, it served as a pattern for the rest of the material, and the validity of the measurements in the test material could be checked against this sentence.

Two versions were added where the test word appears in isolation: affirmative statement (version 5) and echo question (version 6). These isolated versions further increased the tonal and temporal variation within the test material.

(2) Segmental variation (intonation constant)

In each tonal version, two bisyllabic word pairs, containing the stressed vowel /a/, were used. In order to achieve temporal variation, different segmental features were introduced: consonant type (voiced nasal sonorant vs. voiceless obstruent), quantity (long vs. short vowel), and number of initial consonants. The test words and the distribution of the segments were as follows:

	consonant type								
	sonorant (nasal)				obstruent (voiceless)				
vowel quantity	C1	V1	C2	V2	C0	C1	V1	C2	V2
V:	M	ah	n	er	S	p	a	t	er
	[m	a:	n	ɐ]	[ʃ	p	a:	th	ɐ]
V	M	a	nn	er	S	p	a	tt	er
	[m	a	n	ɐ]	[ʃ	p	a	th	ɐ]

The test words, except Mahner ('admonisher'), are semantically void, although they pertain to a very common word structure (in English e.g. singer). Each test word appears both in the first and second sentence position, yielding a doubling of the words to be investigated. This arrangement was also necessary due to semantic reasons. Each sentence contained test words with either sonorants or obstruents respectively, but always a long and a short vowel. Thus the following five sentences were used:

> Der Lahme hat die Maler gemahnt. (calibration sentence)
> Der Mahner hat die Manner gemahnt. (sonorants)
> Der Manner hat die Mahner gemahnt. (sonorants)
> Der Spater hat die Spatter gemahnt. (obstruents)
> Der Spatter hat die Spater gemahnt. (obstruents)

As each sentence appeared in 6 versions, the material consisted of 30 sentences in all. They were arranged in random order and spoken by four speakers[4] from Northern Germany seven times in randomised series with a buffer sentence at the beginning and at the end of each series. In a dialogue-type situation, the author spoke the context sentences during the recording which was made in the studio of the Department of Phonetics at Kiel University. Two signals were recorded simultaneously on a Revox tape recorder, speed 7 ips, the audio signal from a microphone and the Fx-signal from a laryngograph (Fourcin and Abberton 1971). The speakers were well acquainted with the recording situation and produced the sentences with satisfactory naturalness and consistency.

The acoustic analysis rendered an oscillogramme and an Fo-contour, recorded on a Siemens Oscillomink with a paper speed of 100 mm/s. Segments could be defined easily. The following parameters were measured by hand:

(1) Duration: sentence (S), word (W), segments (C0 = ʃ ; C1 = m, p, l; V1 = a (:); C2 = n, t, m, l; V2 = ɐ,ə; the sequence V1C2.

(2) Fo (test words with sonorants only): Fo-value at each segment boundary, starting with the beginning of C1 and thus yielding five Fo-values. If there was an Fo-maximum or Fo-minimum within a segment, the number of Fo-measurements increased. Thus the course of the Fo-movement throughout the word is unequivocally determined.

These basic values were used to calculate the mean, standard deviation, the VC-ratio (section 4.2), and the Fo-change (ΔFo) (section 5.2.2).

The results will be presented below. First, as an orientation, the tonal structures of the six versions will be shown in section 3. Second, some central aspects of the temporal structure of the material under the varying prosodic conditions is treated in section 4. Third, the central question concerning a possible relationship between the duration of segments and the Fo change is investigated by statistical methods in section 5.

3. TONAL STRUCTURE

By varying the prosodic context in the different versions, the tonal contours in sentence positions 1 and 2 were expected to be different. This was the case with all four speakers. Representative tonal contours of the four sentence versions (1-4) and the two isolated versions (5, 6) for two speakers are shown in Figure 1. The curves for each accent position are superimposed with the beginning of the accentuated vowel as the line-up point. The third position is included just for the sake of completeness and will not be treated.

Comparing the tonal contours of the speakers, it can be seen that all the contours, except for version 1 (neutral), are very similar. The tonal differences in version 1 may be the result of not giving equal weight to the accents by each speaker. While speakers U (and M) manifest the accent in all three positions, speakers R (and C) de-accentuate the word in the third position.

In this paper I do not want to treat the tonal features of the Fo-contours. A more detailed report on the tonal aspects of the German accents in statements and two kinds of questions, including also syntactic variation, is being prepared (Bannert forthcoming 2).

4. TEMPORAL STRUCTURE

Primary data on the temporal structure of the test material will not be presented due to limitation of space. Nevertheless, to concentrate on a central issue, the relations within the accentuated syllable will be examined. While the durations of the vowel and the following consonant are presented in the temporal space, the relative temporal relations are shown in the VC-ratios.

4.1 The temporal space

Within accentuated syllables, temporal relations of a certain kind seem to exist. One can think of the relationship between two successive segments as being confined to a temporal field or space and expressing a characteristic pattern of rhythmic structure (Bannert 1976, 1979). A special instance of temporal relationships, the pattern of complementary length (Lehiste 1970), exists in, among other languages, Central Bavarian, a High German variety (Bannert 1976). This pattern is also found in

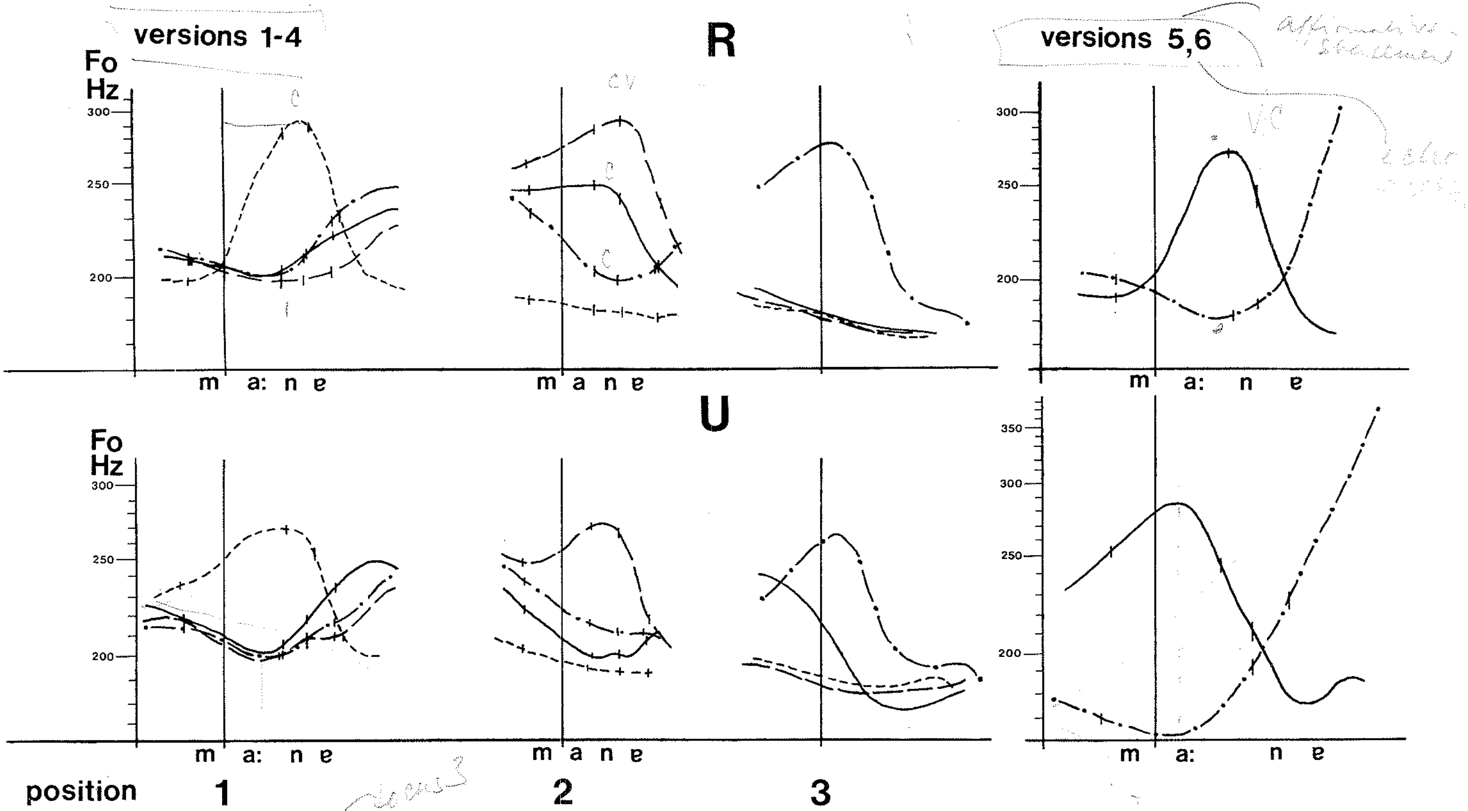


Figure 1. Representative contours of the six versions for two speakers (R, U). Line-up point is the onset of the stressed vowel. Sentence versions (1-4) to the left, isolated versions (5,6) to the right.

neutral
3 equivalent accents

Der Mahne hat die Manner gemahnt

Standard German, although there it has been reported to function in a different way, namely in the contrast between voiceless (fortis) and voiced (lenis) plosives (Kohler 1979 a, b).

The temporal relations between the accentuated vowel and the following consonant can be shown in the vowel-to-consonant plot. The data points expressing the temporal relations are grouped into two clearly separated fields or spaces. Figure 2 depicts the temporal space of the two VC-sequences. The mean values of vowel duration V1 are plotted against the mean values of the duration of the following consonant C2 in all the versions (1-6), both positions, and for all four speakers, making a total of 120 points. Apart from the length contrast of the consonant, Figure 3 shows the same temporal picture, as was found in Central Bavarian. There, however, the duration of the consonant following the short vowel is much larger. Both fields, which are shown by the points of the VC-sequence with long and short vowels respectively, are clearly separated and expand in different directions with increasing segment durations. If the durations are increased drastically, e.g. in version 6, the field of the long vowels extends in the direction of the long vowel, the field of the short vowels, however, extends in the direction of the consonant. Thus the temporal relations between the long and short vowel and the following consonant in stressed syllables in North German, too, can be accounted for by using the pattern of complementary length.

4.2 The VC-sequence

Apart from expressing temporal relations between the stressed vowel and the following consonant in absolute terms, the temporal structure of the segments can also be represented in relative terms. The VC-ratio, i.e. the temporal relationship between the parts (vowel, consonant) and the whole (vowel + consonant = the sequence) was introduced in Bannert (1976). It is assumed that the VC-sequences constitute some temporal basic unit in languages with quantity and stress (Bannert 1976, 1979). Changes of the

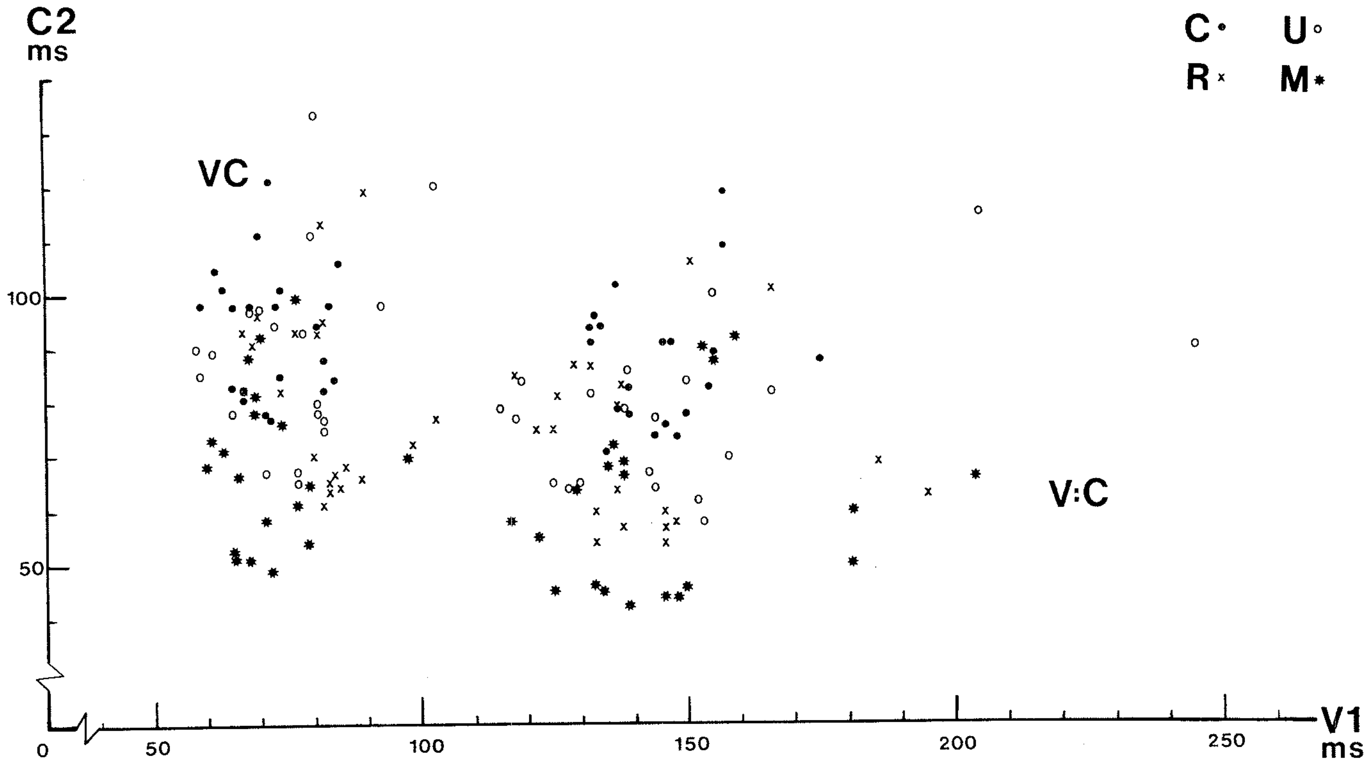


Figure 2. The temporal space. Mean durations of the accentuated vowels plotted against mean durations of the following consonants. All six versions, both positions, and all four speakers.

segment durations of vowels and consonants which are caused by varying segmental and prosodic conditions are brought about in such a way that, according to the principle of complementary length, the long and short vowel and the following consonant tend to make up a certain part of the whole sequence. This temporal relationship between the parts and the whole will remain more or less unchanged.

The VC-ratios, which show the portion of the vowel in relation to the duration of the whole VC-sequence, expressed in percent, are given for each speaker and consonant type in Figure 3. Certain patterns of VC-sequences found earlier in other material are also seen here very easily: the long vowel makes up a larger portion of the VC-sequence than does the short vowel. The VC-ratios are smaller when the consonant is a voiceless obstruent than when it is a voiced sonorant. The ratio varies only a little and, above all, not in a systematic way throughout the six different versions and the two positions. It can also be seen that the ratios keep within a relatively limited area for each length category of the vowels. Even if the ratio does not remain constant, the variation due to large durational differences appear to be minimal. Inherent features, like consonant type, do affect the VC-ratio (cf. Bannert 1976).

However, there is also in this material a relatively constant temporal feature. As shown earlier for Central Bavarian, Central Swedish, and Icelandic (Bannert 1976), this feature is not the VC-ratio itself, but rather the temporal distance (TD) between the ratios of the /V:C/-sequences and the /VC/-sequences. The temporal distance seems to be rather independent of inherent features. It is defined as follows:

$$TD = Q_{V:} - Q_V$$

where TD = temporal distance, V(:) = long or short vowel, Q = VC-ratio

Mean values and standard deviation for the temporal distance for each speaker and each consonant type are given in Table 3. The

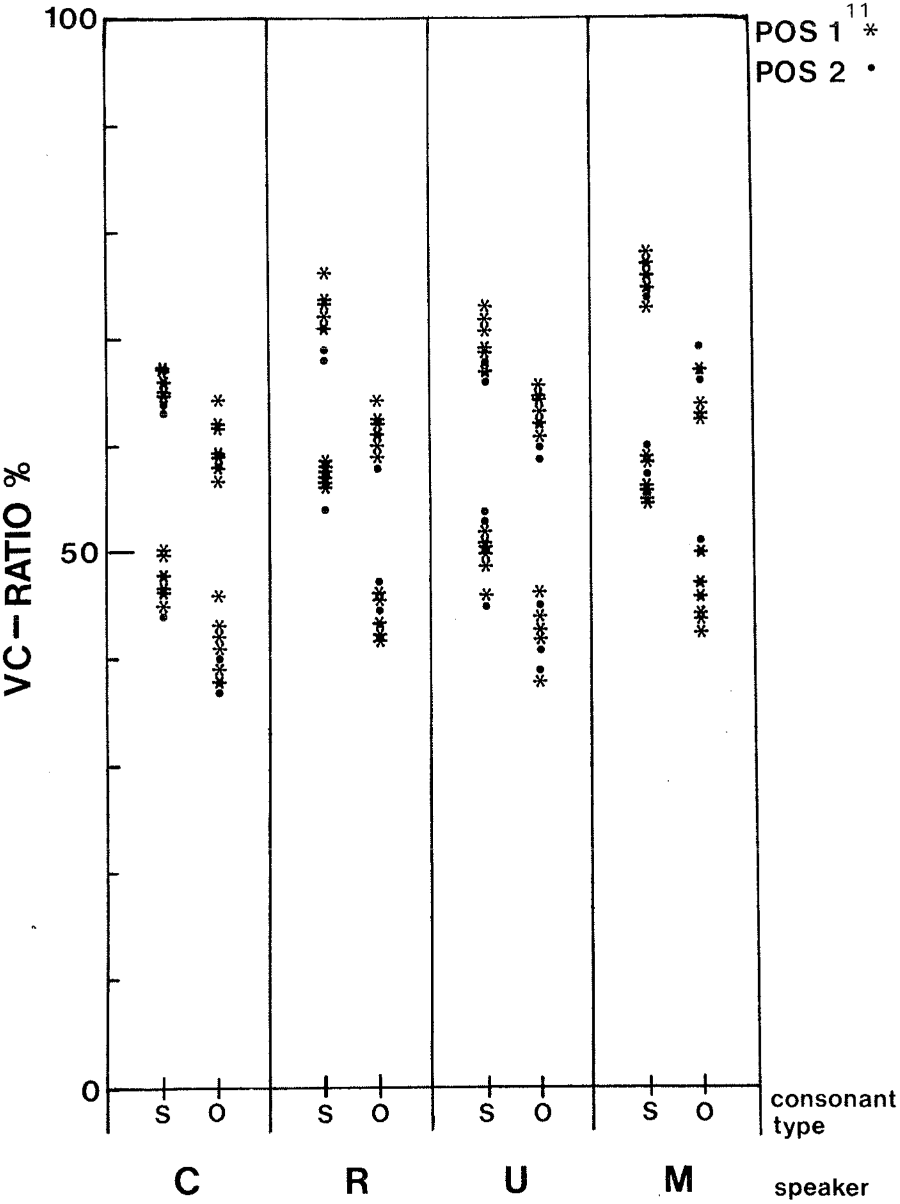


Figure 3. The vowel-to-sequence ratios for each speaker, consonant type, and position.

total mean is 18.25%.

Table 1. Mean values and standard deviation for the temporal distance (TD) between the VC-ratios for each speaker and consonant type in percent.

	speaker							
	C		R		U		M	
	SONORANT	OBSTR.	SONOR.	OBSTR.	SONOR.	OBSTR.	SONOR.	OBSTR.
$\bar{x}$	18.0	19.4	15.0	16.8	18.5	20.0	19.1	19.2
s	1.98	1.08	1.63	1.27	3.81	2.65	2.48	1.33

5. DURATION AND INTONATION

In an Fo-dependent model for segment duration the definitive duration of a given segment is the result of the Fo-change over this segment (Lyberg 1981). Lyberg showed that, in Central Swedish, there exists a strong, positive and linear relationship between the segment duration and the Fo-difference of the long vowel in sentence initial and final position with and without focus. In this North German material, however, no such relationship is to be found. And what is more, the tonal movement is determined by the feature of quantity as it is in Swedish (Bannert 1982).

5.1 Independent quantity

The phonological feature of quantity is manifested by a distinct difference in duration between the long and the short vowel. This is true of all the different tonal contours in all six versions. What is, then, the effect of quantity on the Fo-contour of the vowel, and also on larger units, e.g. the word?

Two alternatives can be proposed (cf. Erikson and Alstermark 1972, Bannert and Bredvad-Jensen 1975, 1977):

(1) TRUNCATION. The Fo-contour of a word containing a long vowel is identical with that of one containing a short vowel, i.e. in a word with a short vowel, the Fo-contour just ends earlier, due to the shorter duration of the short vowel, compared to the corresponding word with a long vowel. Line-up point is the beginning of the accentuated vowel.

(2) COMPRESSION. The Fo-contour of a word containing a long vowel is not identical with that of one containing a short vowel, i.e. the Fo-contours of both words are identical in their course (appearance), the contour in the word with the short vowel, however, is compressed in time. It shows the same tonal movements, but makes them faster.

In all cases with tonal movement, all four speakers clearly show compression. As a representative example, Figure 4 shows simplified Fo-contours of speaker M. They correspond to the mean values of Fo and duration. The Fo-contours in the words with long and short vowels are identical, by and large, at the beginning (in the initial consonant C1) in all versions. The consonant C2, following the stressed vowel, and starting earlier after the short vowel, counted from the beginning of the stressed vowel, affects the Fo-contour radically. The rise of Fo in statements is interrupted and the fall is initiated. The rise in questions is brought about earlier. In other words, certain points or levels of the tonal contour are tied to certain segments or parts of segments. Hence, quantity determines the course of the Fo-contour. As quantity in German appears not only in accentuated syllables, but also in non-accentuated ones, it has to be concluded that quantity is an independent feature, while the tonal contour, at least in certain respects, is dependent on quantity.

5.2 Statistical examination

Before the question about any causal relationship between the

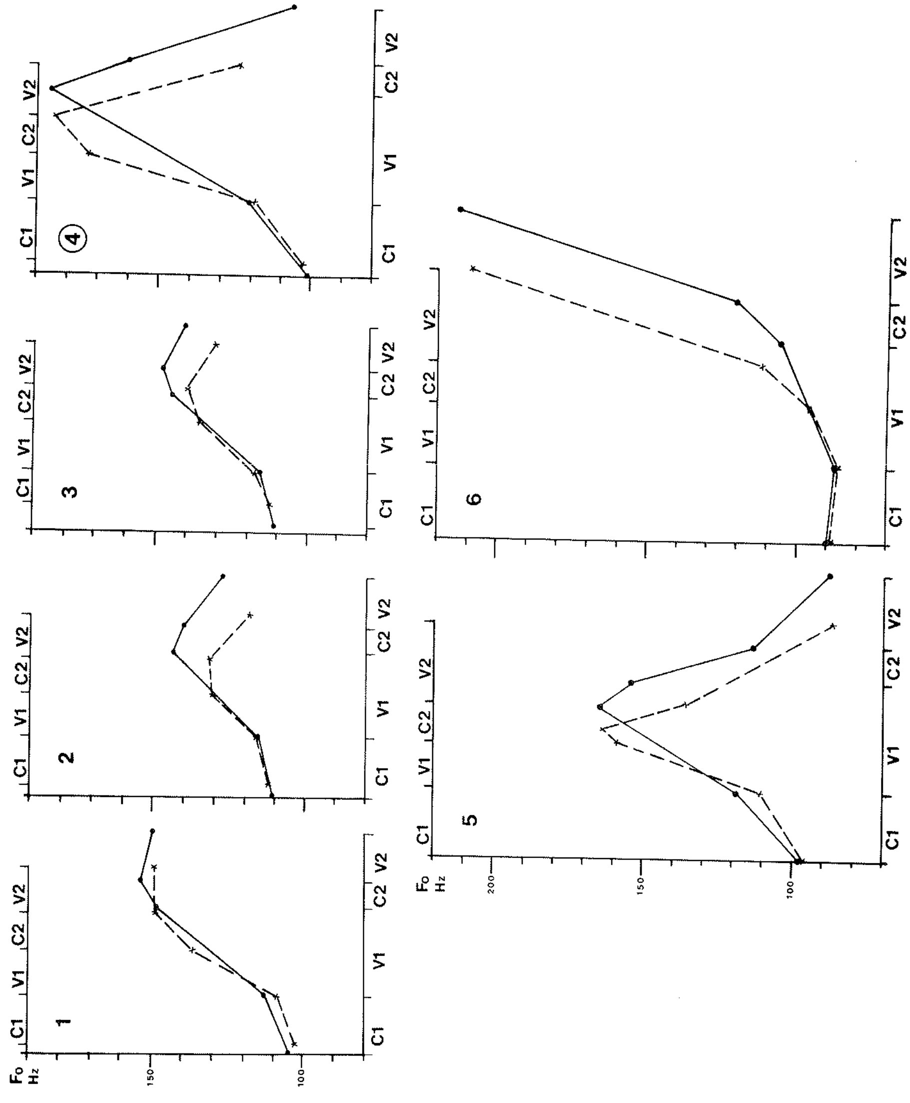


Figure 4. The effect of quantity on the tonal contours. Superimposed Fo-contours of words with long and short vowels. Line-up point is the beginning of the stressed vowel. Mean durations and Fo-values. Sentence versions 1-4 (above), position 1; isolated versions 5,6 (below). Speaker M.

segment duration and the Fo-change can be answered, it has to be settled whether or not there is any relationship at all between these two features. In order to determine step by step if any relationship exists, two methods[5] will be employed. First, it will be determined if the durations of the segments, the VC-sequence, and the word in all the four sentence versions and both positions, are significantly different from each other. If they are different, it could be assumed then that this difference is related to the tonal difference. Second, in any case, it will be tested if there exists any relationship between durations and Fo-changes, i.e. if both variables are correlated. Only if it can be established that they really are correlated, will it be warranted to ask which variable is determined or caused by the other.

5.2.1 Differences in segment duration

Using the statistical method of the Mann-Whitney-Test, the question of whether the segment durations in the four sentence versions were significantly different from each other was tested. In most cases, the versions differed from each other in their tonal aspects (cf. Figure 1). In the test, every version was compared with all the others, i.e. version 1 with 2, 1 with 3, 1 with 4, 2 with 3 etc. Thus for each of the segments six comparisons were made for each position and each speaker.

If one can find significant differences in duration between the different versions, and if all other factors, e.g. speech tempo, can be considered constant, it has to be assumed that these durational differences have to be traced back to the tonal differences. Thus, looking at the high tonal peak of focus in position 1 (version 4) and comparing it to the other three rather similar and low contours, one could expect that the segment durations of version 4 in position 1 would be different from those of versions 1, 2, and 3. And, if there exists any relationship between the Fo change and the segment duration, it should show up in a systematic way, in some kind of pattern.

The result of the statistical testing is shown in Table 2. As can be seen, there is no regular picture to be found. No pattern

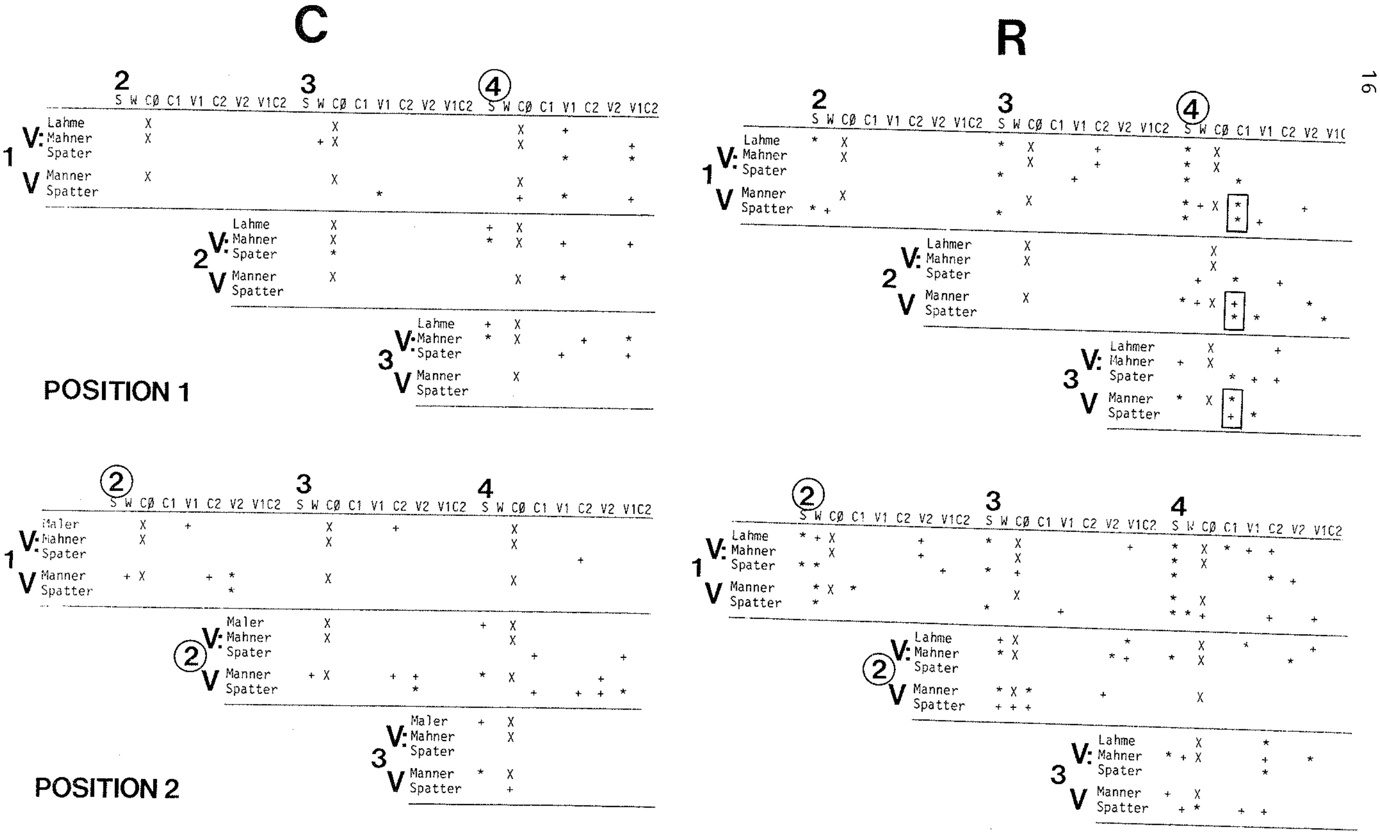


Table 2a. Significant differences of durations of sentence, word, segments, and VC-sequence for each test word between the four versions. Speaker C (left) and R (right). Position 1 above, position 2 below. + 0.01 < p < 0.05, * p < 0.01

U

POSITION 1

		2								3								④							
		S	W	C0	C1	V1	C2	V2	V1C2	S	W	C0	C1	V1	C2	V2	V1C2	S	W	C0	C1	V1	C2	V2	V1C2
1 V:	Lahme		+	X		*			+			X								X					
	Mahner	+		X								X								X			+		
	Spater																		+						
1 V	Manner			X								X								X					
	Spatter																								
2 V:	Lahmer											X							*	X	+	*			+
	Mahner											X								X		+	*		*
	Spater																								
2 V	Manner											X								X					
	Spatter																								
3 V:	Lahmer																			X					
	Mahner																			X			*		*
	Spater																		+			+			+
3 V	Manner																			X					
	Spatter																					+			

POSITION 2

		②								3								4							
		S	W	C0	C1	V1	C2	V2	V1C2	S	W	C0	C1	V1	C2	V2	V1C2	S	W	C0	C1	V1	C2	V2	V1C2
1 V:	Lahme		*	X		+		*	*			X								X					
	Mahner		+	X		+			+			X								X					
	Spater		*			*	+		*												+				
1 V	Manner	+		X								X								X		+	*		
	Spatter											*	*								*				
② V:	Lahme										*	X		*		*	*		*	X		*		*	*
	Mahner										+	X		*			*		+	X		+			+
	Spater										*			*	+		*		+			*			+
V	Manner										+	X				*				X		*	+		
	Spatter											*								+		+	+		
3 V:	Lahme																			X					
	Mahner																			X				+	
	Spater																								
3 V	Manner																			X	+		*		
	Spatter																								

M

POSITION 1

		2								3								④							
		S	W	C0	C1	V1	C2	V2	V1C2	S	W	C0	C1	V1	C2	V2	V1C2	S	W	C0	C1	V1	C2	V2	V1C2
1 V:	Lahme	+		X						+		X				*		*	*	X	+	*			*
	Mahner	+		X								X						*	*	X	*	*	*		*
	Spater																		*	*		*	*		*
1 V	Manner			X								X			*			+	*	X	*				
	Spatter		+					+											+				+		
2 V:	Lahme											X						+	*	X				*	
	Mahner									+		X							*	X	*	*	*		*
	Spater																		*	*		*	*		*
2 V	Manner											X		+				+	*	X	*			*	
	Spatter																	+	*	+			+	+	+
3 V:	Lahme																		*	X	+	*		*	*
	Mahner																	*	*	X	*	*		*	*
	Spater																	+	*	*	*	*	*		*
3 V	Manner																	+	*	X	*		*	*	
	Spatter																		*	+			+		

POSITION 2

		②								3								4							
		S	W	C0	C1	V1	C2	V2	V1C2	S	W	C0	C1	V1	C2	V2	V1C2	S	W	C0	C1	V1	C2	V2	V1C2
1 V:	Maler	+	*	X	+	+			*	+		X						*	*	X		*			*
	Mahner			X		+			*			X						+		X		+			
	Spater		*	*		+	*		*																
1 V	Manner	+	+	X	+							X						*		X					
	Spatter			+			+	+														+			
② V:	Maler										*	X	+		+		*	+		X	+				
	Mahner										+	X		+				+		X					
	Spater										*	+		*	*		*	+	*	*					*
V	Manner									+	*	X	*			*			*	X	+				
	Spatter														*		+							+	+
3 V:	Maler																		*	X		*	+		*
	Mahner																	+		X					
	Spater																								+
3 V	Manner																	*	*	X				+	
	Spatter																	+					+		

Table 2b. Significant differences of durations of sentence, word, segments, and VC-sequence for each test word between the four versions. Speaker U (left) and M (right). Position 1 above, position 2 below. + $0.01 < p < 0.05$, * < 0.01

of systematic variations of durational differences can be detected, not even for one speaker alone.

In some cases, however, fragments of a systematic behaviour can be discerned. They are, though, limited to a certain position (indicated in Table 2). The initial consonant C1 of speaker R has a different duration, but only in the words containing short stressed vowels in position 1. The word duration and almost all durations of the long vowels of speaker M in position 1 are significantly different. The third speaker U, however, shows some instances of significantly different durations in position 2 (version 2 against versions 1, 3 and 4). And last, the fourth speaker C does not show any kind of regularity.

Therefore it has to be concluded that there is no general relationship between the segment duration and the corresponding Fo change.

5.2.2 Segment durations and Fo-change

In contrast to Lyberg (1981, 44), here the Fo-change was defined more completely. In a complex tonal contour, consisting of two different movements and a maximum or a minimum in between, e.g. a rise-fall, the total tonal change is to be considered the tonal difference in both directions. In my opinion it is wrong to calculate the tonal fall only and disregard the rise or vice versa.

By way of simplification, though, the difference in direction of the tonal movements will not be treated. It is known that it takes somewhat more time to increase Fo than to decrease it (Sundberg 1979, Ohala & Evan 1973). However, it seems unlikely that a division of the Fo-changes according to whether they rise or fall would lead to different results.

The Fo-change of a simple, uninterrupted Fo-movement is defined as: $Fo = \left| Fo_{\text{initial}} - Fo_{\text{final}} \right|$.

The Fo-change of a complex Fo-movement is defined as:

$$Fo = \left| Fo_{\text{initial}} - Fo_{\text{max/min}} \right| + \left| Fo_{\text{max/min}} - Fo_{\text{final}} \right| .$$

Due to restrictions of space, neither the Fo-measurements nor the Fo-changes can be presented here. But examining Figures 1 and 5, one can get a fair impression of the size and range of the Fo-change.

The relationship between the segment durations and the Fo-changes was examined by means of linear regression. The calculation was carried out with the values of the sonorant words Lahme, Maler, Mahner, and Manner in several sub-categories according to the variables of test word, quantity, and position. As a representative summary, Table 3 shows the correlation coefficient (r) as a measure of the strength of the relationship according to segment, speaker, and quantity. The level of significance is indicated. For each speaker, the calculation is based on 1008 values of segment durations and the corresponding values of Fo-differences (6 segments x 7 renderings x 4 versions x 3 test words x 2 positions).

Table 3. The relationship between the duration of the segments C1, V1, C2, V2, the sequence /V1C2/, and the word /C1V1C2V2/ and the Fo-change over these segments in the four sentence versions, expressed as the correlation coefficient (r). The levels of significance are indicated. n for each segment and speaker: /V:/ = 118, /V/ = 56.

		SEGMENTS					
Speaker	Quantity	C1	V1	C2	V2	V1C2	Word
C	V:	0.053	0.132	0.083	0.138	0.046	-0.162*
	V	0.184	-0.015	0.019	0.338**	-0.036	0.037
R	V:	-0.075	0.181*	0.200*	0.253**	0.256**	0.210*
	V	0.137	0.037	0.019	0.362**	-0.115	0.421**
U	V:	0.090	0.473**	0.335**	-0.009	0.582**	0.361**
	V	-0.225*	0.121	0.239	0.436**	0.258*	0.353**
M	V:	0.392**	0.557**	0.307**	0.419**	0.582**	0.706**
	V	0.614**	0.323**	0.460**	0.504**	0.451**	0.789**

It can be seen that the values of (r) are quite small, perhaps except for the word (W) of speaker M. Therefore it has to be concluded once more that there is no statistically provable and systematic relationship between the segment duration and the Fo-change in this North German material.

6. CONSEQUENCES FOR A MODEL OF PROSODY

On the basis of the present findings, showing that there is no general and systematic relationship between Fo-change and segment duration, the question of the temporal and tonal control in German can be highlighted in a few essential respects. A comparison with the situation in the related languages Danish and Swedish will render a deeper dimension of understanding to these findings.

It could be shown that there is no relationship between the temporal and tonal dimension in this North German material. It does seem, however, that speakers are also free to use the temporal dimension, apart from the tonal, when focusing. The double signalling of sentence prominence appears to be more effective than the use of one dimension only. Similar variations between speakers in using the temporal and tonal dimension for focus is also to be found in my Central Bavarian material (Bannert 1976). Furthermore, the four North German speakers of this investigation behave partly like the Danish speakers in Thorsen (1980) and one of the Swedish speakers in Bruce (1981), partly like the Swedish speakers in Bannert (1979, forthcoming 1), and the other speaker in Bruce (1981). While segment durations of the Danish speakers and one of Bruce's speakers remain unchanged in focus, the duration of all segments in focus is increased with the other Swedish speakers. And this in spite of the large tonal difference between the Central and Southern variety of Swedish. Therefore, it has to be assumed that an increase of duration in focus need not necessarily be a result of the changed Fo-contour. On the contrary, as languages may differ in this respect, time and tone have to be treated basically as independent features.

Thus some consequences for a model of prosody in German emerge. Starting from an input[6], defined and specified by necessary linguistic information, segmental and prosodic, the model firstly will generate the basic temporal and tonal structure of the utterance separately and independently of each other. This can be done either in parallel or successively. Then both basic structures will be joined or added. Due to various conditions, certain adjustments will appear to be necessary, e.g. the adjustment of the basic tonal contours as a consequence of temporal limitations (varying number of unstressed syllables between the stressed ones). It seems clear that these adjustments will be less in German than in Swedish. There the tonal structure is much more complex due to the word accents. An important factor of the mutual temporal and tonal adjustments is speech tempo. The output contains the final prosodic form of the utterance.

A simple sketch of the model for generating the prosody of a German utterance is shown in Figure 5. Although the model contains the important components, it is just a beginning; developing the model further will require extensive work. However, the point of departure is rather favourable. Quite obviously, it is not necessary to start from scratch. As mentioned in the introduction, models for time and intonation for related languages do exist. Therefore it seems to me most advantageous to start out from suitable models developed for other languages and to examine their usefulness for German. What is the new approach here, is the treatment of the two main dimensions of prosody, time and intonation, at the same time and basically independent of each other. In order to generate the temporal basic structure, a model like Klatt's (1979) presents itself. In order to generate the tonal basic structure, the Danish model (Thorsen 1979) and the Swedish model (Bruce 1977, Bruce 1982, Gårding 1982, Gårding & Bruce 1981) seem convincing.

It can be expected that future work with this model will promote

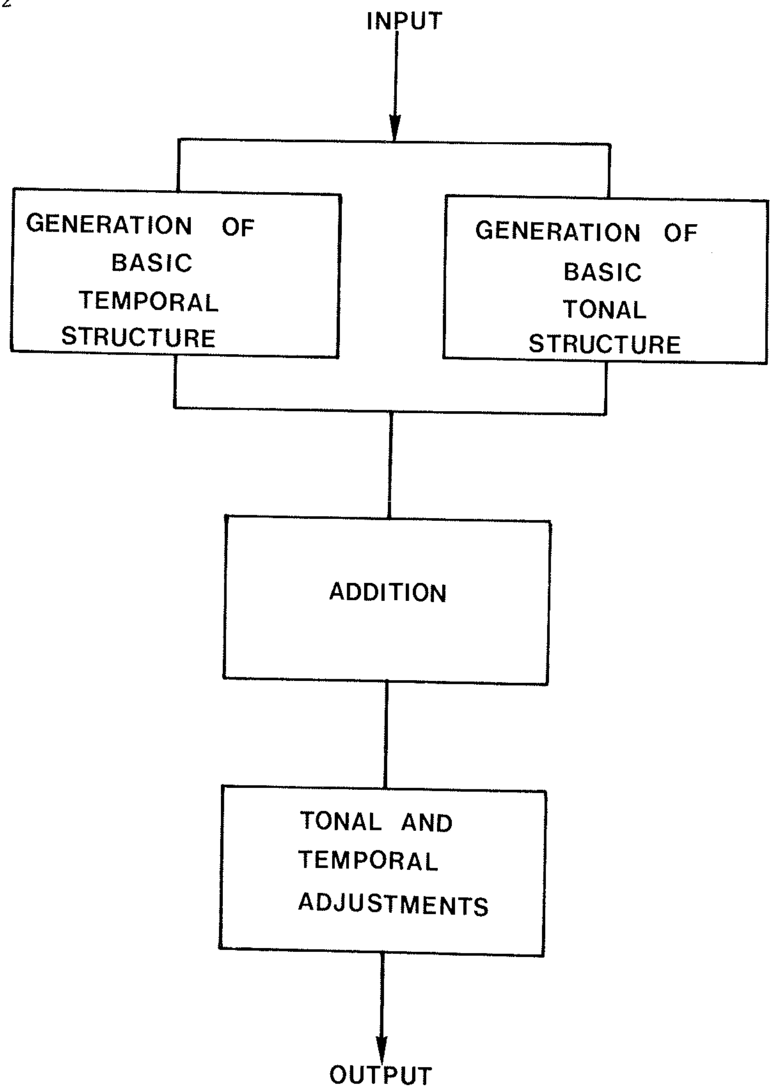


Figure 5. The basic components of a model for prosody in German.

phonetic knowledge and understanding of prosody in German. It will not only bring new insights into the nature of prosody, and thus lead German intonation studies from the level of merely describing auditory impressions to the level of acoustic measurements. Later on, when the level of perception has been reached, the development of speech synthesis will be pushed forward. Of course, applied phonetics, too, will profit by this, and the teaching of rhythm and intonation, important for intelligibility and acceptability of spoken German, will be improved.

FOOTNOTES

1. This investigation was begun at the same time as Bruce (1981) as a reaction against Lyberg's proposal of an Fo-dependent model for segment duration (Lyberg 1979). While Bruce studied Central Swedish, the same variety of Swedish as Lyberg, my work aimed at extending the issue to another language.

2. Prosody is used here in a restricted sense. It means the rhythm and melody of language, i.e. the temporal and tonal structure of utterances. Concentrating on the basic linguistic components of prosody, other features, e.g. vocal quality, are excluded for the sake of simplification.

3. For a detailed description of the method of moving focus within a sentence, see Bruce (1977).

4. Data concerning the speakers:

Abbreviation	Sex	Place of birth	Year of birth
C	female	Lübeck	1957
R	female	Fintel (Rothenburg/ Wimme)	1952
U	female	Bleckede (Lüneburg)	1952
M	male	Marne (Dithmarschen)	1952

5. The two statistical tests were run by using the SPSS-programme at the Computer Center of the Christian-Albrechts-University in Kiel.

6. In this paper, I do not intend to explore any details of the basic components of the model. Thus no further specification of the input will be presented here. A more detailed exposition of the model will be given in Bannert (forthcoming 1, 2).

ACKNOWLEDGEMENTS

This investigation was begun during a short stay at the Department of Phonetics at the Christian-Albrechts-University of Kiel in the fall of 1979 and continued during the winter term of 1980/1981. I would like to express my sincere gratitude to the Director of the Department, Prof.Dr. Klaus J. Kohler, for kindly giving me the opportunity of using the extensive facilities of the Phonetics Laboratory in Kiel.

I would also like to thank Bill Barry, Gösta Bruce, Wim van Dommelen, Olle Engstrand, Eva Gårding, Lennart Nordstrand, and Gert Zimmerman for fruitful discussions, statistical help, and valuable comments on a preliminary version of this paper, and Mr. Fuchs and Mr. Jansen for all their efficient technical assistance.

REFERENCES

Bannert, R. (1976) Mittelbairische Phonologie auf akustischer und perzeptorischer Grundlage. Travaux de l'Institut de Linguistique de Lund X, Gleerups: Lund; Fink: Munich

Bannert, R. (1979) The effect of sentence accent on quantity. Proceedings of the 9th International Congress of Phonetic Sciences, Vol. II:253-259. Copenhagen

Bannert, R. (1982) An Fo-dependent model for segment duration? Reports from the Institute of Linguistics, Uppsala University 8:59-80

Bannert, R. (Forthcoming, 1). Independence and interdependence of prosodic features

Bannert, R. (Forthcoming, 2). Towards a model for German prosody

Bannert, R. and Bredvad-Jensen, A.-C. (1975) Temporal organization of Swedish tonal accents: The effect of vowel duration. Working Papers, Phonetics Laboratory, Lund University 10:1-36

Bannert, R. and Bredvad-Jensen, A.-C. (1977) Temporal organization of Swedish tonal accents: The effect of vowel duration in the Gotland dialect. Working Papers, Phonetics Laboratory, Lund University 15:133-138

Bruce, G. (1977) Swedish word accents in sentence perspective. Travaux de l'Institut de Linguistique de Lund XII, Lund: Gleerups

Bruce, G. (1981) Tonal and temporal interplay. In T. Fretheim (Ed.) Nordic Prosody II, pp. 63-74. Trondheim: Tapir

Bruce, G. (1982) Developing the Swedish intonation model. Working Papers, Department of Linguistics and Phonetics, Lund University 22:(this issue)

Carlsson, R. and Granström, B. (1973) Word stress, emphatic stress, and syntax in a synthesis by rule scheme for Swedish. Quarterly Progress and Status Report of the Speech Transmission Laboratory, Royal Institute of Technology, Stockholm 2/3: 31-36

Erikson, Y. and Alstermark, M. (1972) Fundamental correlates of the grave word accent in Swedish: The effect of vowel duration. Quarterly Progress and Status Report of the Speech Transmission Laboratory, Royal Institute of Technology, Stockholm 2/3: 53-60

Fourcin, A.J. and Abberton, E. (1971) First applications of a new laryngograph. Medical and Biological Illustrations 21: 172-182

Gårding, E. (1981) Contrastive prosody: A model and its application. Studia Linguistica 35:146-165

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

t'Hart, J. and Cohen, A. (1973) Intonation by rule: a perceptual quest. Journal of Phonetics 1:309-327

t'Hart, J. and Collier, T. (1975) Interpreting different levels of intonation analysis. Journal of Phonetics 3:235-255

Klatt, D. (1979) Synthesis by rule of segmental durations in English sentences. Proceedings of the 9th International Congress of Phonetic Sciences, Vol. II:290-297. Copenhagen

Kohler, K.J. (1979a) Parameters in the production and the perception of plosives in German and French. Arbeitsbericht 12: 261-291. Institut für Phonetik, Universität Kiel

Kohler, K.J. (1979b) Dimensions in the perception of Fortis and Lenis plosives. Phonetica 36:332-343

Lehiste, I. (1970) Suprasegmentals, Cambridge, Mass.: M.I.T. Press

Lindblom, B., Lyberg, B. and Holmgren, K. (1981) Durational patterns of Swedish phonology: Do they reflect short-term memory processes? Indiana University, Linguistics Club

Lyberg, B. (1979) Final lengthening - partly a consequence of restrictions on the speed of fundamental frequency change? Journal of Phonetics 7:187-196

Lyberg, B. (1981) Temporal properties of spoken Swedish. Monographs from the Institute of Linguistics, University of Stockholm 6

Nooteboom, S.G. (1972) Production and perception of vowel duration. A study of durational properties of vowels in Dutch. Philips Research Reports Supplement 5, Eindhoven

Chala, J. and Ewan, W. (1973) Speed of pitch change. Journal of the Acoustical Society of America 53:345 (A)

Öhman, S., Zetterlund, S., Nordstrand, L. and Engstrand, O. (1979) Predicting segment durations in terms of a gesture theory of speech production. Proceedings of the 9th International Congress of Phonetic Sciences, Vol. II:305-311.Copenhagen

Sundberg, J. (1979) Maximum speed of pitch changes in singers and untrained subjects. Journal of Phonetics 7:71-79

Thorsen, N. (1979) Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish. Annual Report of the Institute of Phonetics, University of Copenhagen 13:59-85

Thorsen, N. (1980) Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish. Annual Report of the Institute of Phonetics, University of Copenhagen 14:121-205

STRESS IN MODERN GREEK: AN ACOUSTIC STUDY

Antonis Botinis

INTRODUCTION

This paper reports an experimental investigation of the contribution of fundamental frequency, duration, and intensity to the construction of word stress and sentence stress in Modern Greek. By word stress is meant the phonetic prominence over a syllable (of the type CONflict-conFLICT) to convey linguistic information. By sentence stress is meant that part of the sentence which carries the most important information; in the terminology of generative grammar, it is the *focus*, referring to the new information in contrast to *presupposition*, the information already shared by the speaker and the listener (Jackendoff 1972). Thus, sentence stress is semantically rather syntactically conditioned.

The paper addresses three basic questions: (1) What is the acoustic manifestation of word and sentence stress in Modern Greek? (2) Which is the most consistent acoustic parameter of word and sentence stress? (3) Are the acoustic parameters equally consistent across the utterance or does their contribution to the construction of stress depend on the position of sentence stress in the sentence?

EXPERIMENTAL PROCEDURE

Fundamental frequency (Fo), duration, and intensity data were recorded from two speakers who produced utterances with word stress and sentence stress. Oscillographic analyses were made of these recordings.

Speech material

Eight meaningful Greek sentences were constructed. All the sentences were declarative, with the same syntactic structure, i.e., subject-verb-adverb. The words composing the sentences were prosodic minimal pairs and the position of sentence stress varied according to the contextual situation. The first two sentences were pronounced *neutrally*, i.e., the speaker had no contextual information. The remaining six test sentences were elicited as answers to questions formulated in different ways, to make the speaker choose one of the elements of the sentence as the focus and the carrier of the information required by the questioner. The segmental structure of the sentences was composed exclusively of sonorants in order to have a continuous Fo-contour. Below is the list of the test sentences (capital letters indicate sentence stress).

Contextual frame	*Test sentences*
1. None	*i mána múmene móni* (Mother was alone.)
2. None	*i naná mu méni moní* (Nana is single.)
3. *pja súmene móni?* (Who was alone?)	*I MÁNA múmene móni* (MOTHER was alone.)
4. *pja su méni moní?* (Who is single?)	*I NANÁ mu méni moní* (NANA is single.)
5. *ti súkane i mána móni?* (What was mother doing alone?)	*i mána MÚMENE móni* (Mother WAS STAYING alone.)
6. *ti su káni i naná moní?* (What is Nana doing single?)	*i naná MU MÉNI moní* (Nana IS STAYING single.)
7. *pos súmene i mána?* (How was mother?)	*i mána múmene MÓNI* (Mother was ALONE.)
8. *pos su méni i naná?* (How is Nana?)	*i naná mu méni MONÍ* (Nana is SINGLE.)

The test sentences were read in a random order, five times by each speaker. The speakers read the sentences as in everyday speech.

Subjects.

The main subject is a male student, 24 years old, brought up and educated in Athens; he speaks what is considered to be standard Athenian. In addition the present investigator, male, 32 years old, was recorded. Both subjects speak about the same sociolect.

Dialect.

The dialect investigated is Modern Demotic Greek, a dialect spoken in southern Greece, especially in Athens and Peloponnesos.

Experimental equipment.

The recordings took place in a sound-treated room at the phonetics laboratory, Lund University. The frequency response of the tape recorder was flat within ± 2 dB from 30 Hz to 14 000 Hz and the signal to noise ratio was 63 dB. The microphone was flat within the frequencies 35 to 17 000 Hz. The Fo-contour was extracted by a F-J Electronics pitch extracting device along with a duplex oscillogram and recorded on a Siemens oscillominc, with a paper speed of 100 mm/sec. Intensity was also recorded using a Fonema analysis unit.

Acoustic analysis.

Oscillograms of the eight utterances were made from the tape recording. From these oscillograms the first and the last repetitions of the test sentences were rejected. From the remaining three, the one which combined the best execution (in the investigator's judgement) and the finest oscillogram display was selected for acoustic analysis.

Intensity was calibrated in 5 dB steps and fundamental frequency in 10 Hz steps. Both intensity and fundamental frequency scales were linear. For the segmentation of the utterances, a duplex oscillogram and the intensity contour were used. Intensity was measured at the peak of the intensity contour for each syllable; peak Fo values for the stressed syllables and Fo minima for the unstressed ones were also measured.

RESULTS

The acoustic parameters of Fo, duration, and intensity are shown in Figs. 1, 2, 3, and 4. Figure 1 represents the neutral con-

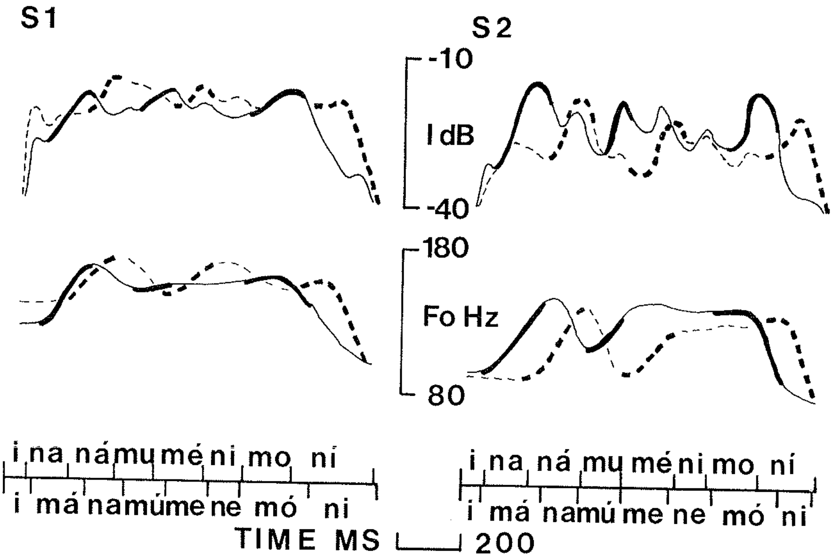


Figure 1. Neutral sentence. Intensity contour (upper), Fo-contour (middle), and duration (bottom) of the sentences *i mána múmene móni* (solid lines), and *i naná mu méni moní* (dashed lines). The thick lines represent the stressed syllables and the thin lines the unstressed syllables, subject 1 on the left and subject 2 on the right.

figuration of the two sentences; the other three figures represent the acoustic manifestation of the sentences with sentence stress located either at the first position, second position or third (final) position, in response to the contextual situation demanded by the question.

In Fig. 1 (the neutral sentence) the Fo-contour rises on each word stress although only moderately on the last stressed syllable (for absolute values see Table I). The Fo-contour of the unstressed syllables is either falling or level. The stressed syllables are longer than the unstressed ones; at the end of the utterances the syllables are longer no matter whether stressed or unstressed. The intensity contours have about the same configuration as the Fo-contours, with the addition of small peaks on the unstressed syllables.

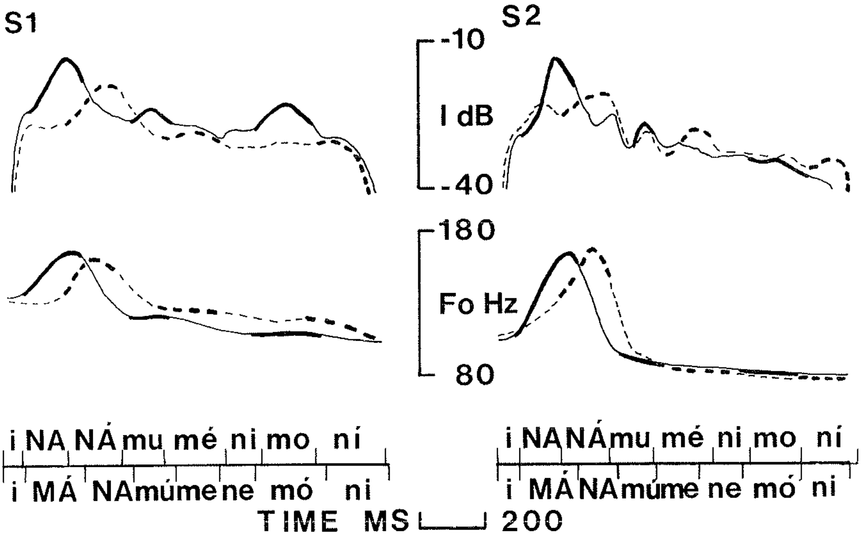


Figure 2. Focus 1 position. Intensity contour (upper), Fo-contour (middle), and duration (bottom) of the sentences *I MÁNA múmene móni* (solid lines), and *I NANÁ mu méni moní* (dashed lines). The thick lines represent the stressed syllables and the thin lines the unstressed syllables. Capital letters indicate focus, subject 1 on the left and subject 2 on the right.

In Fig. 2 (focus on the first poistion), the Fo-contour rises on the first stressed syllable (carrying the focus of the sentence) and then falls and flattens towards the end of the utterance. Duration preserves the same relations among the stressed and unstressed syllables as in the neutral sentences. The intensity is highest on the focus syllable, fearly high peaks on the other word stresses right to the end of the utterances, in contrast to the Fo-contour which remains low after the focus.

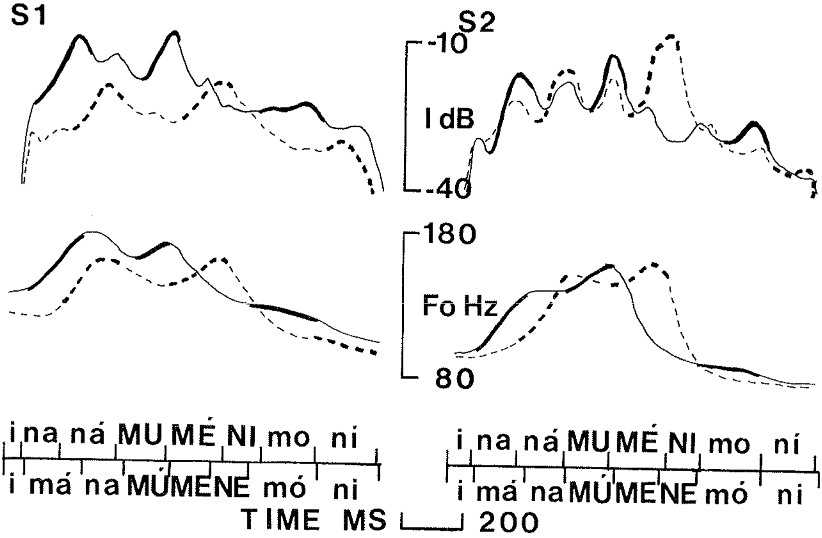


Figure 3. Focus 2 position. Intensity contour (upper), Fo-contour (middle), and duration (bottom) of the sentences *i mána MUMENE móni* (solid lines), and *i naná MU MÉNI moní* (dashed lines). The thick lines represent the stressed syllables and the thin lines the unstressed syllables. Capital letters indicate focus, subject 1 on the left and subject 2 on the right.

In Fig. 3 (focus on the second position), the Fo-contour rises on each stressed syllable until the focus syllable and then falls and flattens to the end of the utterances. Duration preserves the same structure as the neutral and the first-focus position sentences. The word stresses have higher intensity peaks than the unstressed syllables and preseve the same relationship to the end of the utterances, though the peak levels are lower after focus. The intensity contour is highest on the focus syllables.

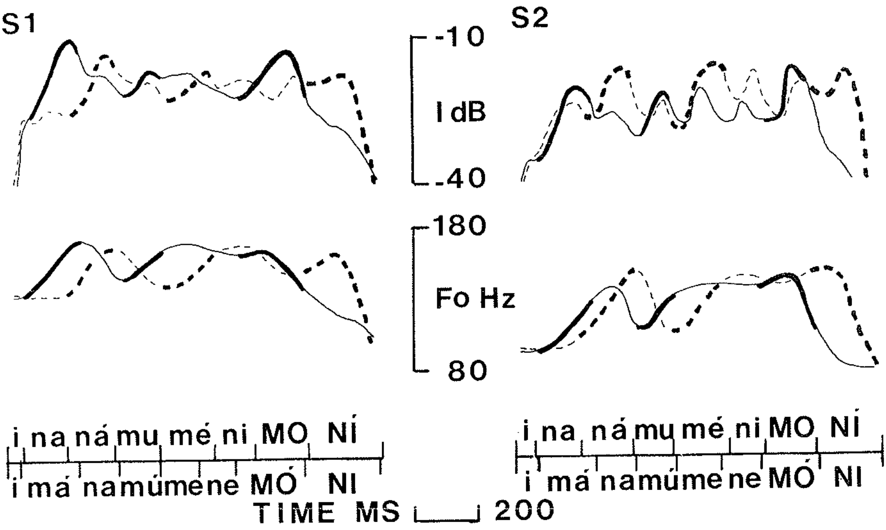


Figure 4. Focus 3 (final) position. Intensity contour (upper), Fo-contour (middle), and duration (bottom) of the sentences *i mána múmene MÓNI* (solid lines), and *i naná mu méni MONÍ* (dashed lines). The thick lines represent the stressed syllables and the thin lines the unstressed syllables. Capital letters indicate focus, subject 1 on the left and subject 2 on the right.

The overall structure of Fig. 4 (focus on the third position) is quite similar to that of Fig. 1 (the neutral sentence). The stressed syllables are marked for Fo, duration, and intensity relative the unstressed ones.

All eight investigated declarative sentences have shown some degree of declination, both for Fo and intensity, which is more evident after focus.

Subject 1

		Sentence (a)								Sentence (b)							
NEUTRAL		i	má	na	mú	me	ne	mó	ni	i	na	ná	mu	mé	ni	mo	ní
	Fo	115	160	140	140	145	145	155	100	130	130	165	140	160	160	140	150
	D	70	170	120	130	120	110	210	200	60	130	140	110	150	110	150	250
	I	18	25	22	26	23	21	25	10	22	21	27	24	26	24	26	24
FOCUS 1		I	MÁ	NA	mú	me	ne	mó	ni	I	NA	NÁ	mu	mé	ni	mo	ní
	Fo	130	160	112	115	110	105	108	100	120	120	150	120	120	110	108	110
	D	60	190	140	140	130	110	210	200	50	130	160	130	180	100	160	220
	I	21	28	17	20	17	17	21	15	18	17	23	15	16	13	13	14
FOCUS 2		i	má	na	MÚ	ME	NE	mó	ni	i	na	ná	MU	MÉ	NI	mo	ní
	Fo	128	180	150	170	130	120	115	100	112	112	150	130	150	110	105	100
	D	60	170	130	140	120	120	210	200	50	120	170	160	180	110	170	190
	I	21	31	27	32	24	20	22	17	16	16	23	19	24	11	11	14
FOCUS 3		i	má	na	mú	me	ne	MÓ	NI	i	na	ná	mu	mé	ni	MO	NÍ
	Fo	128	175	135	160	165	150	150	100	130	130	160	130	155	160	140	155
	D	40	170	120	130	120	110	210	230	50	130	150	130	170	120	160	230
	I	17	30	25	26	24	22	28	16	18	20	28	23	25	23	24	24

Subject 2

		Sentence (a)								Sentence (b)							
NEUTRAL		i	má	na	mú	me	ne	mó	ni	i	na	ná	mu	mé	ni	mo	ní
	Fo	95	135	105	125	135	128	132	80	95	95	135	95	120	120	125	135
	D	60	190	120	130	150	130	200	120	60	140	160	130	160	100	150	190
	I	13	27	22	24	23	20	26	6	11	18	25	17	22	20	17	22
FOCUS 1		I	MÁ	NA	mú	me	ne	mó	ni	I	NA	NÁ	mu	mé	ni	mo	ní
	Fo	100	150	95	90	85	85	83	80	100	120	160	90	85	85	85	85
	D	60	180	120	120	130	130	180	140	60	130	150	140	180	110	150	170
	I	15	28	19	18	13	10	10	4	15	20	22	15	17	13	10	10
FOCUS 2		i	má	na	MÚ	ME	NE	mó	ni	i	na	ná	MU	MÉ	NI	mo	ní
	Fo	95	130	130	150	95	90	90	80	95	100	142	135	152	85	80	80
	D	70	160	130	150	130	130	190	180	60	140	150	140	180	110	170	170
	I	15	26	24	28	21	18	18	6	15	22	26	25	31	18	13	8
FOCUS 3		i	má	na	mú	me	ne	MÓ	NI	i	na	ná	mu	mé	ni	MO	NÍ
	Fo	90	118	102	120	125	125	135	80	90	90	135	100	130	135	130	140
	D	60	170	120	120	120	120	190	190	50	140	160	120	170	110	160	190
	I	8	22	18	22	22	19	25	6	6	20	25	20	26	25	22	25

Table I. The acoustic parameters Fo, duration, and intensity of the sentences *i mána múmene móni* and *i naná mu méni moní*. The neutral sentences as well as the three focal positions for both subjects are shown.

DISCUSSION

Before interpreting the results given in Table I, the main points will be recapitulated. First, word stress is realized differently before and after focus. Before focus, as well as in the neutral statements, all the three parameters contribute to a stressed syllable; after focus only duration and intensity are present in word stress. Second, sentence stress is realized as a relative increase of Fo and intensity in comparison with an unfocused word stress in the same position of the sentence; the structure of the words bearing sentence stress is the same as those without focus as far as duration is concerned. Lastly, the effect of downdrift, the tendency of human languages to decline in pitch from the beginning to the end of an utterance (Breckenridge 1977, Bruce and Gårding 1978, Collier 1972, Thorsen 1978), is quite apparent in the majority of the utterances analysed.

It is convenient to discuss sentence stress first. The word *mána* in focal position (Fig. 2) will be compared with the same word without focus (Fig. 3 and 4) in order to isolate the sentence stress component. It is easy to be led astray by the absolute acoustic values (Table I). For instance, in sentence (a), subject 1, when the word *mána* is in focus, Fo rises to 160 Hz on the stressed syllable. When *mána* is not focused, Fo rises to 180 and 175 Hz (the neutral sentence not considered). If the stressed syllable is instead compared with the unstressed one, the Fo difference for *mána* in focus is found to be 48 Hz whereas the differences for the unfocused *mána* are 30 and 40 Hz. The other acoustic parameters can be examined in the same way. Differences in duration are found to be almost constant, no matter whether the word is in focus or not, while intensity is 11 dB higher for the stressed syllable in focus position, 4 and 5 dB higher in non-focal position. Thus, rising Fo and higher intensity are acoustic elements characterizing sentence stress for subject 1. The same holds true for subject 2. Duration does not seem to contribute to the acoustic structure of sentence stress.

When the word *mána* is compared with *naná* in pre-focal position, we see that the stressed syllables have noticeably higher values than the unstressed syllables, for all three acoustic parameters, for both subjects. Fo is always realized as a rise to denote word stress in Modern Greek (Botinis 1979). Hyman (1977) has observed that in other languages word stress can be realized both as rising and falling Fo. This has been reported for Swedish as well (Gårding et al. 1970, Hadding-Kock 1961). Thus, Fo, duration, and intensity are present on the word stress in pre-focal and focal position.

The acoustic pattern of the post-focus stressed syllables is quite different from the structure of the pre-focus stressed syllables. The most striking difference is that variations of Fo are reduced to a minimum, about 10 Hz. Bruce (1977) and O'Shaughnessy (1979) have reported similar data for Swedish and English respectively. Although these small variations are above the perceptual threshold (Lehiste 1970), it seems highly questionable whether they contribute to the perception of a post-focal word stress distinction.

Duration and intensity run parallel to each other in post-focus stressed syllables, both increase in the majority of the utterances. Durational differences are only at a minimum in one sentence (focus in second position) by subject 2, see Fig. 3. Although the durational difference in this case is 10 msec. (above the perceptual threshold, Lehiste 1970), and although there is a 10 Hz Fo difference, the dominant acoustic factor in this position seems to be relative intensity. The final syllable *-ni* (of both *móni* and *moní*, i.e., both stressed and unstressed) has lower intensity than *mo-*. This may be due to the general difference in intrinsic intensity between /i/ and /o/ and to the contribution of the final juncture. But the intensity difference is much larger in *móni* (12 dB) than in *moní* (5 dB). This is probably the decisive factor for the perception of this stressed syllable.

This modern Greek data suggests that it would be too simple to set a hierarchy of acoustic cues to describe stress (Fry 1955, 1958). Rather, it is my opinion that the different acoustic parameters contribute in different ways across the syntagm. The different combinations to denote the same phonetic concept, stress, show that they can be independent of each other; moreover, there seem to be no solid grounds to support a view of a one-to-one mapping between phonetic features and the acoustic signal (Bailly and Summerfield 1980).

CONCLUSIONS

Finally, I shall try to answer the questions posed at the beginning. It is difficult to generalise as the present study has been limited to a small number of speakers and utterances[1].

Firstly, while all three acoustic parameters (Fo, duration and intensity) contribute to the structure of word stress, only duration and intensity are present after focus; on the other hand, relative intensity and Fo both contribute to the sentence stress. Secondly, of the three acoustic parameters examined, relative intensity differences are always present for both word and sentence stress. Lastly, the three acoustic parameters do not make a constant contribution across the syntagm but their relative contribution depends on the position of sentence stress.

---

1. I am planning both perceptual and physiological experiments related to the relative importance of the physiological and acoustic parameters for the perception of stress.

BIBLIOGRAPHY

Bailey, P. & Summerfield, Q. (1980) Information processing in speech: Observations on the perception of s-stop clusters. Journal of Experimental Psychology: Human perception and Performance 6: 536-563

Botinis, A. (1979) Stress and stress degrees in Modern Greek. Unpublished paper, Phonetics Laboratory, Department of General Linguistics, Lund University

Bruce, G. (1977) Swedish word accents in sentence perspective. Travaux de l' Institut de Linguistique de Lund XII. Lund: Gleerup

Bruce, G. & Gårding E. (1978) A Prosodic TYpology for Swedish Dialects. In Gårding, E. Bruce, G. & Bannert, R. (Eds.) Nordic Prosody, 219-228. Travaux de l' Institut de Linguistique de Lund XIII, Department of Linguistics, Lund University

Breckenridge, J. (1977) The declination effect. Journal of the Acoustical Society of America 60: S90

Collier, R. (1972) From pitch to intonation. Ph.D. thesis, University of Louvain

Fry, D. (1955) Duration and intensity as physical correlates of linguistic stress. Journal of the Acoustical Society of America 27: 765-768

Fry, D. (1958) Experiments in the perception of stress. Language and Speech 1: 126-152

Gårding, E. Bruce, G. & Bannert, R. (Eds.) Nordic Prosody. Travaux de l' Institut de Linguistique de Lund XIII. Department of Linguistics, Lund University

Gårding, E., Fujimura, O. & Hirose, H. (1970) Laryngeal control of Swedish word tone: A preliminary report of an EMG study. Annual Bulletin of the Research Institute of Logopedics and Phoniatrics 4: 45-54, Tokyo

Hadding-Kock, K. (1961) Acoustico-phonetic studies in the intonation of Southern Swedish. Travaux de l' Institut de Linguistique de Lund III. Lund: Gleerup

Hyman, L. (Ed.) (1977) Studies in stress and accent. Southern California Occasional Papers in Linguistics 4

Hyman, L. (1977) On the Nature of Linguistic Stress. In Hyman, L. (Ed.) Studies in stress and accent. Southern California Occasional Papers in Linguistics 4

Jackendoff, R. (1972) Semantic Interpretation in Generative Grammar. Cambridge, Mass.: The MIT Press

Lehiste, I. (1970) Suprasegmentals. Cambridge, Mass.: The MIT Press

O' Shaughnessy, D. (1979) Linguistic features in fundamental frequency patterns. Journal of Phonetics 7: 119-145

Thorsen, N. (1978) Aspects of Danish intonation. In Gårding, E. Bruce, G. & Bannert, R. (Eds.) Nordic Prosody, 23-32. Travaux de l' Institut de Linguistique de Lund XIII, Department of Linguistics, Lund University

STRESS IN MODERN GREEK: A PHYSIOLOGICAL STUDY

Antonis Botinis

INTRODUCTION

This study deals with variations in subglottal pressure associated with word and sentence stress in modern Greek. We have shown previously (Botinis 1982), that the acoustic parameters of fundamental frequency (Fo), duration, and intensity contribute in different ways to the manifestation of word and sentence stress in Greek. All three of these acoustic parameters contribute to word stress, but only duration and intensity after focus. On the other hand, relative Fo and intensity both contribute to sentence stress.

This paper raises the following questions: (1) Which of the acoustic parameters co-vary with subglottal pressure (Ps)? Do variations in Ps mostly influence one or more of the acoustic parameters?

An indirect method for estimating Ps from records of oral pressure (Po) has been applied (Holmberg 1980; Smitheran and Hixon 1981). Its validity has been shown empirically (Löfqvist, Carlborg and Kitzing 1982). This method exploits the fact that during the production of voiceless stops, Po and Ps are identical. By constructing a suitable linguistic material where voiceless stop consonants and vowels alternative, it is possible to obtain indirect measurements of Ps. This is done by linearly interpolating the Po records of the stop consonants and thus estimating the Ps associated with the intervening vowels.

EXPERIMENTAL PROCEDURE

Speech material

Eight meaningful Greek sentences were constructed. All the sentences were declarative with the same syntactic structure, i.e., VP-NP-AP. The words to be tested were minimal pairs with stress on either the first or second syllable (*pápa* "pope" *papá* "priest"). The words were placed in the carrier sentence *γrápse ... padú* "write ... everywhere". The method of having minimal pairs only in middle position was adopted for both accuracy and simplicity. By shifting sentence stress between initial, medial, and final positions, the minimal pairs could be examined in post-focal, focal, and pre-focal position. The first two sentences were pronounced *neutrally*, i.e., the speaker had no contextual information. The remaining six sentences were elicited as answers to questions formulated in different ways to make the speaker choose one of the elements of the sentence as the focus and carrier of the information required by the question. The test sentences are given in the following table (capital letters indicate sentence stress).

Contextual frame	*Test sentences*
1. None	*γrápse pápa padú* (Write "pope" everywhere.)
2. None	*γrápse papá padú* (Write "priest" everywhere.)
3. *tí na káno ton pápa padú?* (What shall I do with pope everywhere?)	*ƔRÁPSE pápa padú* (WRITE "pope" everywhere.)
4. *tí na káno ton papá padú?* (What shall I do with priest everywhere?)	*ƔRÁPSE papá padú* (WRITE "priest" everywhere.)
5. *tí na γrápso padú?* (What shall I write everywhere?)	*γrápse PÁPA padú* (Write "POPE" everywhere.)
6. *tí na γrápso padú?* (What shall I write everywhere?)	*γrápse PAPÁ padú* (Write "PRIEST" everywhere.)
7. *pu na γrápso pápa?* (Where shall I write "pope"?)	*γrápse pápa PADÚ* (Write "pope" EVERYWHERE.)
8. *pu na γrápso papá?* (Where shall I write "priest"?)	*γrápse papá PADÚ* (Write "priest" EVERYWHERE.)

The test sentences were read in random order, five times by 3 speakers. The speakers produced the sentences at a normal tempo and loudness.

Subjects

The main subjects are two males, 26 years old, brought up and educated in Athens; they speak what is considered to be standard Athenian. In addition, the present investigator, male, 32 years old, was recorded. None of the speakers has any history of speech, hearing, neurological, or respiratory disorders. All three speakers have Greek as their native language and speak about the same sociolect. The dialect investigated is Modern Demotic Greek, a dialect spoken in southern Greece, especially in Athens and Peloponnesos.

Recordings

The recordings took place in a sound-treated room at the phonetics laboratory, Lund University. Oral pressure was sampled through a plastic tube, 10 cm long and with an inner diameter of 2 mm. The tube was held between the lips and the open end of the tube was positioned just behind the upper teeth. The tube was coupled to a differential pressure transducer. After suitable amplification, the pressure signal was recorded on one channel of a multichannel instrumentation tape recorder. During the experiment, the pressure signal was also monitored on an oscilloscope in order to detect clogging of the tube. Static calibration on the recording system was performed before and after the recording session using a water manometer.

Conventional acoustic recordings were made simultaneously. The voice signal was recorded in direct mode on the tape recorder.

Record analysis

Oscillograms of the eight utterances by all three speakers were made from the tape recording. From these oscillograms the first and the last of the five repetitions were rejected. For the segmentation of the utterances, a duplex oscillogram and the Po-contour were used. Po was measured in cm H_2O on a linear scale.

RESULTS

The pressure variations of the minimal pairs *pápa* and *papá* in their carrier sentences, in neutral configuration, are shown in Fig. 1; the acoustic parameters of duration, Fo, and intensity are shown as well. The utterances are produced by subject 1. In the first utterance, when stress is on the first syllable, the Ps rises to a high level at the very beginning of the syllable, i.e., during /p/ and then falls gradually at the beginning of the next syllable. In the second utterance, stress is on the second syllable and the Ps does not change but remains level during the vowel. In both sentences, the stressed syllables are longer than the unstressed ones and the Ps declines from the beginning to the end of the utterances.

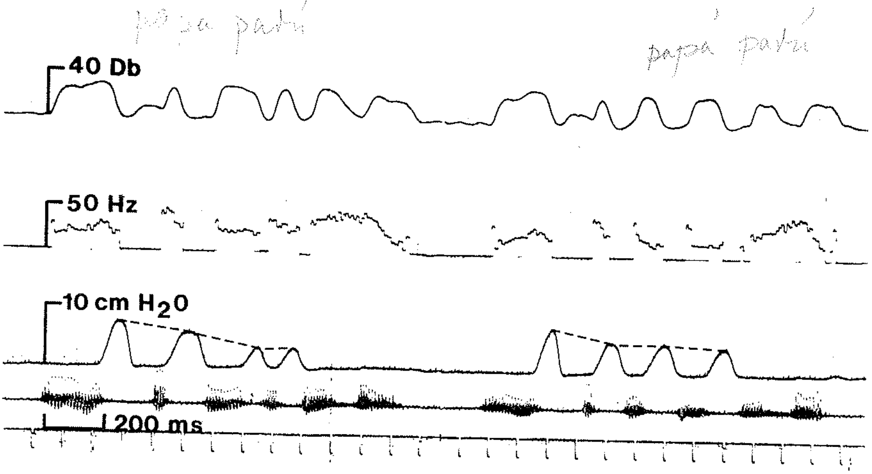


Figure 1. Record analysis of the utterances *γrápse pápa padú*, left, and *γrápse papá padú*, right. The curves represent from top to bottom, intensity, fundamental frequency, oral pressure, and duplex oscillogram. The interpolated broken line between the pressures associated with the voiceless stops is an estimation of subglottal pressure during the vowels.

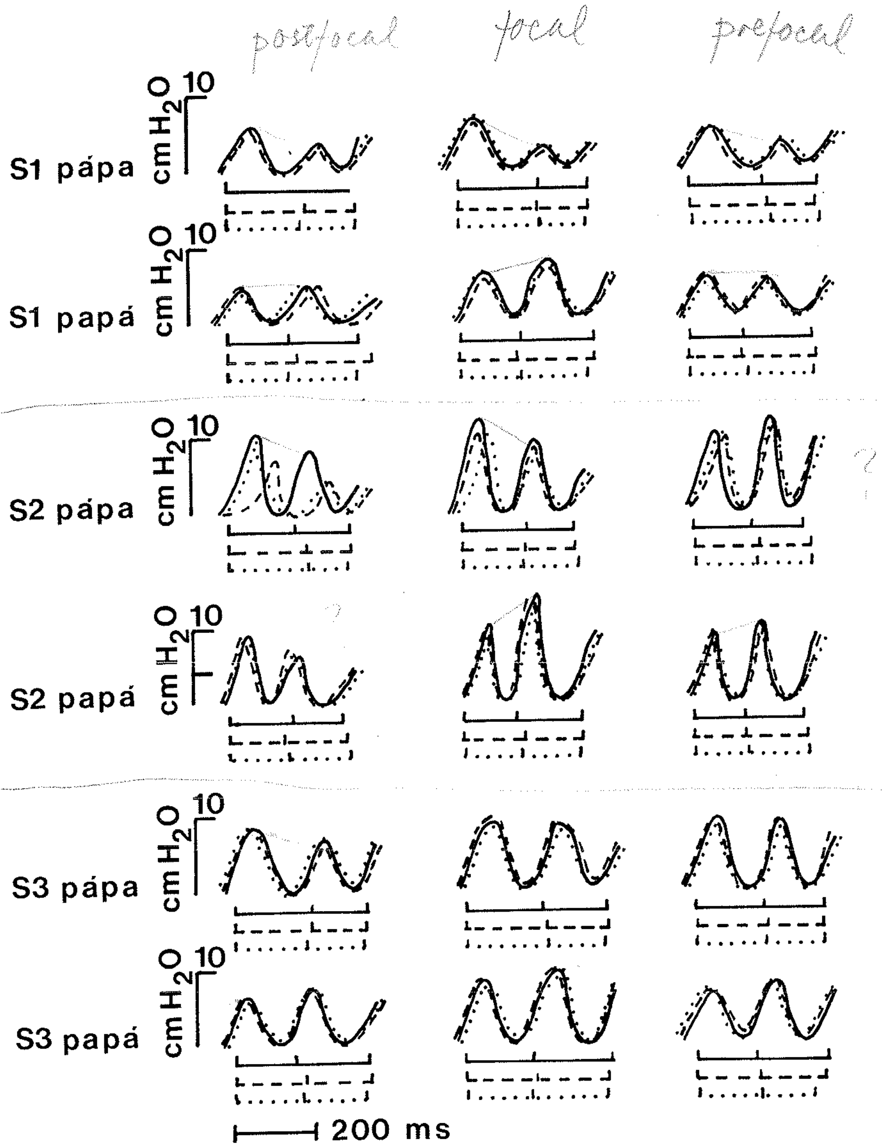


Figure 2. Oral pressure and duration of the minimal pairs *pápa* and *papá* in post-focal, focal, and pre-focal position for subject 1, subject 2, and subject 3. The repetitions are represented by solid lines, dashed lines, and dots.

In Fig. 2 the oral pressure curves of the three repetitions of the minimal pairs in post-focal, focal, and pre-focal position, for all three subjects is shown; let us consider the first speaker. When the minimal pairs occur in post-focal position, they have the same structure as in the neutral utterances, i.e., about 2.5 cm H_2O higher Ps on the unstressed syllable in /pápa/. In the word /papá/, both the stressed and the unstressed syllables have about the same Ps. The minimal pairs in focal position exhibit higher Ps on the stressed syllables, about 1.5 cm H_2O, in comparison with the minimal pairs in non-focal position. The minimal pairs in pre-focal position show the same kind of Ps variation as when in post-focal position. We could generalize our results by saying that on the one side we have the stressed syllables utilizing Ps against the unstressed syllables, on the other side we have the focused words utilizing Ps even more against the unfocused words. The same holds true for all three speakers; although the way and the degree they utilize Ps may vary, the structure of the Ps variations across the speakers is basically the same.

DISCUSSION

The analysed data of modern Greek show that every stressed syllable is associated with increase in Ps and that there is a declination of Ps from the beginning to the end of the utterances. These results agree with Ladefoged's (1967) earlier findings of an increased activity of the respiratory muscles on stressed syllables, whether emphatically stressed or not. However, van Katwijk (1974) and Ohala (1977) report increased expiratory activity only on emphatically, and not on normally stressed syllables.

If we combine the results reported in Botinis (1982) with the results obtained in this study, we see that every stressed syllable has higher relative intensity associated with higher Ps (Ladefoged et al. 1963; Isshiki, 1964).

The pre-focal stressed syllables have higher Fo, longer duration, and higher intensity; at this position it is difficult to examine the influence of the Ps on the acoustic parameters. If we consider the post-focal stressed syllables, we see that the higher Ps has an effect only on intensity, whereas Fo remains low and flat. This has important theoretical implications. First, it implies that the human speech mechanism does not produce a quasi-constant Ps, but a varied one that can produce linguistic effects. Second, variations on Ps do not necessarily cause an effect on Fo.

Studies reviewed in Ohala (1978) argue that Fo is exclusively regulated by the laryngeal muscles. Moreover, correlations found between Fo and intensity (which is causally correlated with Ps) show that they both are influenced by the larynx. Thus, according to Ohala, not only are observed Fo variations independent of Ps variations but, to some extent, Ps variations are caused by laryngeal activity.

The Greek data agree with Ohala's first point; the laryngeal muscles are alone responsible for the Fo-contour. In post-focal position, the Ps for a stressed syllable is higher; if the vocal cords maintained a constant level of tension, the Fo-contour should automatically rise higher. But this is not the case. The structure of the language requires that the Fo-contour be flat in post-focal position. Thus, the laryngeal muscles will have to keep Fo flat, although Ps variations could cause an effect on Fo. We disagree with Ohala's second point in the same post-focal position. Here, there is no laryngeal activity that could raise the Ps. Moreover, it is clear at this point that the Greek speakers do not produce a constant Ps but a varied one. Since focal elements are so common in speech, one can hardly expect good correlations between Ps and Fo (Collier 1975).

Lieberman (1967) raised the question whether the observed Fo variations were accomplished by the activity of the laryngeal muscles or the respiratory system. Lieberman had records only on Ps and assumed that the laryngeal muscles maintained a constant level of tension. This was thought to be the case in declarative sentences. In yes-no questions, where there is a terminal rise in Fo without any variation in Ps, Lieberman assumed the laryngeal muscles to operate.

The hypothesis "Ps causes Fo" does not agree with the Greek data reported here. In post-focal position, where stress is associated with Ps variations, flat Fo is the rule, although minor Fo variations, up to 10 Hz, may occasionally be found. Atkinson (1978) argues that different physiological mechanisms may be involved at different Fo levels. In particular, in declarative sentences and for low Fo (80-100 Hz), Fo is mostly governed by Ps. However, in the analysed Greek utterances where Fo is between 80-100 Hz at post-focal position, Ps still does not have the expected effect on Fo, i.e., 3-7 Hz/cm H_2O (Baer 1979). Answering the question whether it is the larynx or lungs that controls Fo (pitch), the answer is the *larynx* (Vanderslice 1967, Ohala 1977).

Of course variations in Ps can influence Fo, other things being equal, but Fo must take the configuration required by the structure of a particular language. In the case of Greek, Fo serves one purpose (language structure) and Ps another (stress distinction). It is possible that a stressed syllable may gather all the acoustic and physiological parameters, but it is far from true that *every* stressed syllable is associated with an Fo (pitch) change (Hyman 1977).

In Greek, there are two facts that suggest that the acoustic parameters are independent of each other and not produced by the same mechanism. First, hightened Ps in post-focal position does not produce any effect on the Fo-contour. Second, at post-focal position, intensity is combined with duration to give the impression of word stress, whereas for sentence stress intensity is combined with Fo.

In Swedish, the Fo variations are distinctive for both word and sentence stress, though the Fo spans a smaller range after focus (Bruce 1977). In a general study of the Scandinavian languages

(Gårding 1977), it was shown how the Fo manifestations are different for the "same" tonal phenomena. Even for the Swedish dialects the Fo-contour may exhibit a great variability depending on the dialect.

A recent comparison of the intonational patterns of Swedish, Greek, and French (Gårding et al. 1982), demonstrated how the three languages utilize Fo in different ways for the "same" linguistic entities, i.e., word stress, sentence stress, etc.

The above findings suggest that there must be a great variability in the ways speakers produce the same linguistic effect either in one language or across languages. This is a strong evidence against any correspondence between phonetic features, articulation, and the acoustic signal.

CONCLUSIONS

At the end of this study we have been able to give direct answers to questions that may be difficult to answer. This has been done by constructing a suitable linguistic material which isolated the complex components of the acoustic parameters in speech.

The first question, which of the acoustic parameters co-vary with Ps, the answer is intensity, occasionally Fo, depending on the structure of the sentence. The second question, do variations in Ps mostly influence one or more of the acoustic parameters, the answer is mostly intensity.

In this paper it has been shown how a language, and in principle any language, uses different acoustic elements to convey the same idea. A substantial question should be proper now: To what extent do speakers of different languages use different physiological mechanisms to produce the same acoustic effect? In other words how much is universal and how much is specific in the production of human speech across languages?

BIBLIOGRAPHY

Atkinson, J. E. (1978) Correlation analysis of the physiological factors controlling fundamental voice frequency. Journal of the Acoustical Society of America 63: 211-222

Baer, T. (1979) Reflex activation of laryngeal muscles by sudden subglottal pressure changes. Journal of the Acoustical Society of America 65: 1271-1275

Botinis, A. (1982) Stress in modern Greek: An acoustic study. Working Papers, Phonetics Laboratory, Department of General Linguistics, Lund University 22: (this issue)

Bruce, G. (1977) Swedish word accents in sentence perspective. Travaux de l' Institut de Linguistique de Lund XIII. Lund: Gleerup

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

Fromkin, V. (1978) Tone - a Linguistic Survey. New York: Academic Press

Gårding, E. (1977) The Scandinavian word accents. Travaux de l' Institut de Linguistique de Lund XI. Lund:Gleerup

Gårding, E., Botinis, A. & Touati, P. (1982) Acomparative study of Swedish, Greek, and French intonation. Working Papers, Phonetics Laboratory, Department of General Linguistics, Lund University 22: (this issue)

Holmberg, E. (1980) Laryngeal airway resistance as a function of phonation type. Journal of the Acoustical Society of America 68: S101 (A)

Hyman, L. (Ed.) (1977) Studies in stress and accent. Southern California Occasional Papers in Linguistics 4

Hyman, L. (1977) On the nature of linguistic stress. In Hyman, L. (Ed.), Studies in stress and accent. Southern California Occasional Papers in Linguistics 4: 37-83

Isshiki, N. (1964) Regulatory mechanism of voice intensity variation. Journal of Speech and Hearing Research 7: 17-29

van Katwijk, A. (1974) Accentuation in Dutch: An experimental study. Ph.D. dissertation, Utrecht University

Ladefoged, P. (1967) Three Areas of Experimental Phonetics. London: Oxford University Press

Ladefoged, P. & McKinney N., P. (1963) Loudness, sound pressure, and subglottal pressure in speech. Journal of the Acoustical Society of America 35: 454-461

Lieberman, P. (1967) Intonation, Perception, and Langyage. Cambridge, Mass.: MIT Press

Löfqvist, A., Carlborg, B., & Kitzing, P. (1982) Initial validation of an indirect measure of subglottal pressure during vowels. Working Papers, Phonetics Laboratory, Department of General Linguistics, Lund University 22: (this issue)

Ohala, J. J. (1977) The physiology of stress. In Hyman, L. (Ed.) Studies in stress and accent. Southern California Occasional Papers in Linguistics 4: 145-168

Ohala, J. J. (1978) Production of tone. In Fromkin, V. (Ed.) Tone - a Linguistic Survey. New York: Academic Press 5-41

Smitheran, J. R. & Hixon, T. J. (1981) A clinical method for estimating laryngeal airway resistance during vowel production. Journal of Speech and Hearing Disorders 46: 138-146

Vanderlice, R. (1967) Larynx vs lungs: Cricothyrometer data refuting some recent claims concerning intonation and archetypality. UCLA Working Papers in Phonetics 7: 69-79

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.

Swedish intonation model

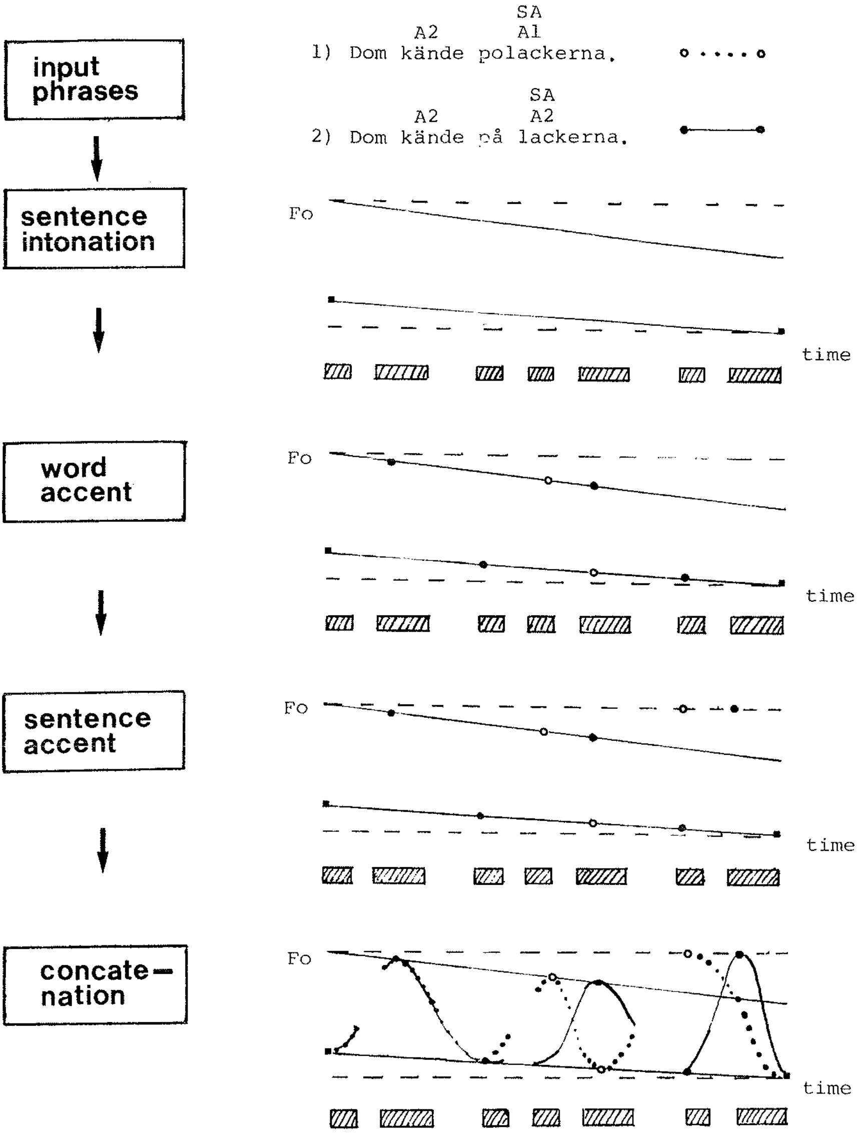


Figure 1. Swedish intonation model. Application of pitch rules for Standard Swedish (after Bruce & Gårding 1978).

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:

ˋUngen ˋnallar
The kid is stealing

ˋMamman ˋlämnar honom / (Ja) ˋmammorna hann ˋlämna honom
The mother leaves him / (Well) the mothers managed to leave him

Man hann ˋlämna honom ˋungen / (Ja) man hann ˋlämna nåra ˋnallar
One managed to leave him the kid / (Well) one managed to leave some teddy bears

ˋMamman ˋlämnar ˋungen ˋnallar
The mother leaves the kid teddy bears

(Ja) ˋmammorna hann ˋlämna våran ˋunge nåra ˋnallar
(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).

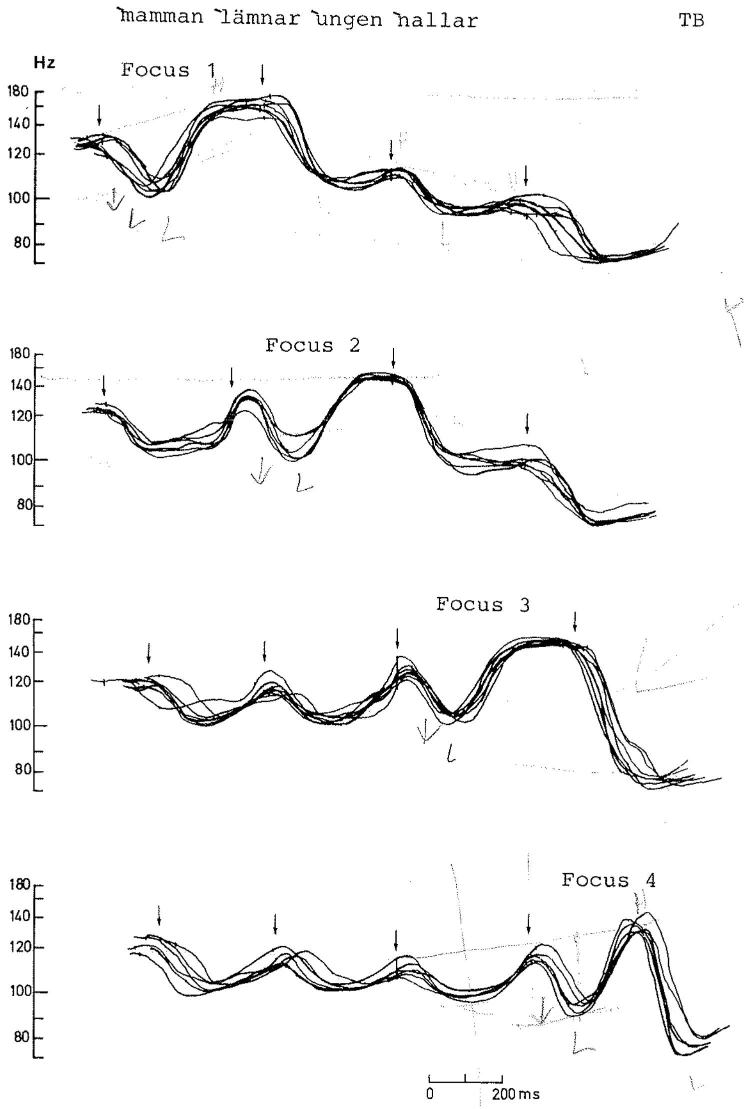


Figure 2. The nature of Fo downdrift. The effect of placement of focus in a sentence containing four stress groups with one unstressed syllable between the stressed syllables. Repeated Fo contours by a male informant (TB). The arrows indicate the stress group boundaries. The line-up point is at the third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

[illegible]

ˋmammorna hann ˋlämna våran ˋunge nåra ˋhallar TB

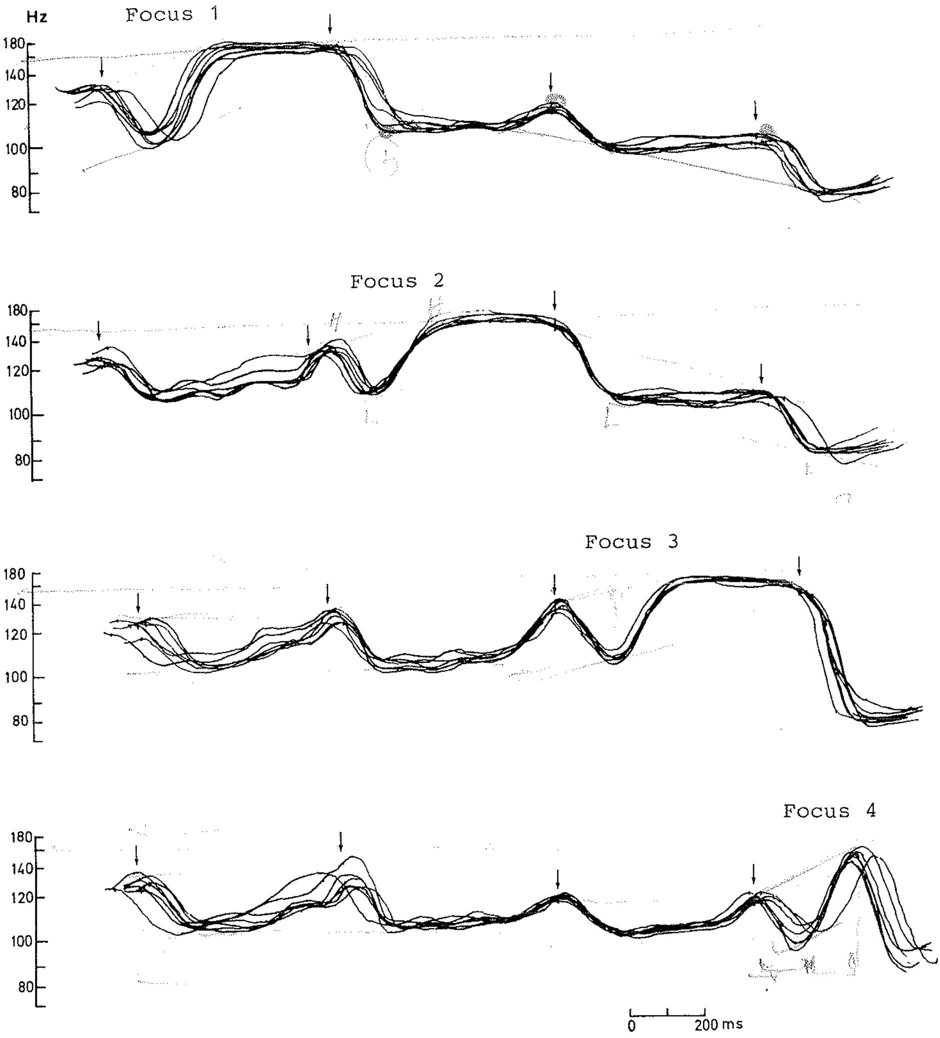


Figure 3. The nature of Fo downdrift. The effect of placement of focus in a sentence containing four stress groups with three unstressed syllables between the stressed syllables. Repeated Fo contours by a male informant (TB). The arrows indicate the stress group boundaries. The line-up point is at the 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
ˋungen ˋnallar												
initial	x̄		135	104	*162*	80						
focus	s		3,0	4,0	7,6	3,7						
final	x̄		134	104	126	98	*144*	87				
focus	s		2,4	2,0	6,3	4,6	5,7	9,6				
ˋmamman ˋlämnar honom												
initial	x̄		129	103	*165*	80						
focus	s		2,8	2,7	5,8	4,4						
final	x̄		126	105	131	100	*155*	84				
focus	s		6,1	3,0	2,4	3,0	11,2	7,8				
man hann ˋlämna honom ˋungen												
initial	x̄	111	136	105	*169*	79						
focus	s	3,8	3,5	2,9	5,3	5,2						
final	x̄	107	128	103	126	102	*149*	82				
focus	s	3,7	5,5	1,9	4,0	3,7	3,5	4,5				
ˋmamman ˋlämnar ˋungen ˋnallar												
initial	x̄		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	x̄		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	x̄		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	x̄		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
ˋmammorna hann ˋlämna våran ˋunge nåra ˋnallar												
initial	x̄		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	x̄		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	x̄		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	x̄		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 ˋmammorma hann lämna honom												
initial	$\bar{x}$	189	224	169	297	147						
focus	s	5,2	10,3	6,3	14,6	2,4						
final	$\bar{x}$	187	209	166	230	164	275	158				
focus	s	4,3	6,8	2,2	16,8	6,0	8,3	5,6				
ja man hann ˋlämna några ˋnallar												
initial	$\bar{x}$	191	215	170	294	146						
focus	s	6,2	12,7	5,5	18,6	3,7						
final	$\bar{x}$	189	202	163	189	153	256	167				
focus	s	4,9	10,3	5,1	7,4	5,7	13,9	7,6				
ja ˋmammorna hann ˋlämna våran ˋunge några ˋnallar												
initial	$\bar{x}$	191	236	175	314	169	180	153	169	143		
focus	s	6,9	12,9	6,0	8,4	4,6	11,8	3,5	5,8	2,6		
2nd	$\bar{x}$	184	215	169	257	178	308	156	167	144		
focus	s	5,7	9,1	3,9	10,3	6,8	8,9	3,7	7,9	2,4		
3rd	$\bar{x}$	189	219	170	267	177	220	171	306	144		
focus	s	5,2	7,8	7,2	7,1	8,2	15,6	5,5	9,9	3,2		
final	$\bar{x}$	186	218	171	269	191	205	168	192	158	265	149
focus	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.

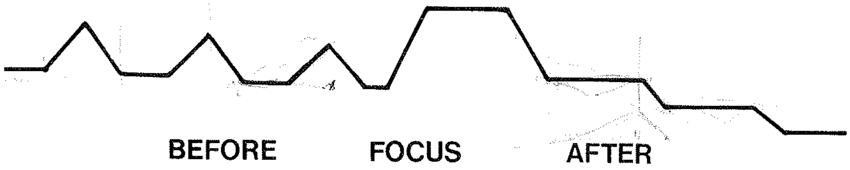


Figure 4. Downdrift in Swedish. Stylized Fo contour of a Standard Swedish utterance.

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.

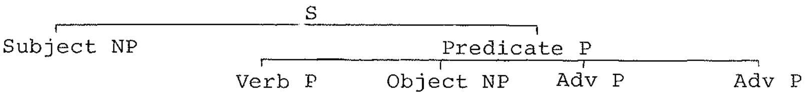
In order to test this prediction the following parameters were systematically varied in the test material.

(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 ˋmammorna hann ˋlämna honom ˋungarna med ˋnallarna bland ˋlängorna

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

Test material - experiment two:

Nån av ˜mammorna.
One of the mothers.

Nån av ˜mammorna hann ˜lämna honom.
One of the mothers managed to leave him.

(Det var med ˜nallarna bland ˜längorna.)
(It was with the teddy bears among the barns.)

Nån av ˜mammorna hann ˜lämna honom ˜ungarna.
One of the mothers managed to leave him the kids.

Nån av ˜mammorna hann ˜lämna honom ˜ungarna med ˜nallarna.
One of the mothers managed to leave him the kids with the teddy bears.

Nån av ˜mammorna hann ˜lämna honom ungarna med ˜nallarna bland ˜längorna.
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).

(k-pivot) high

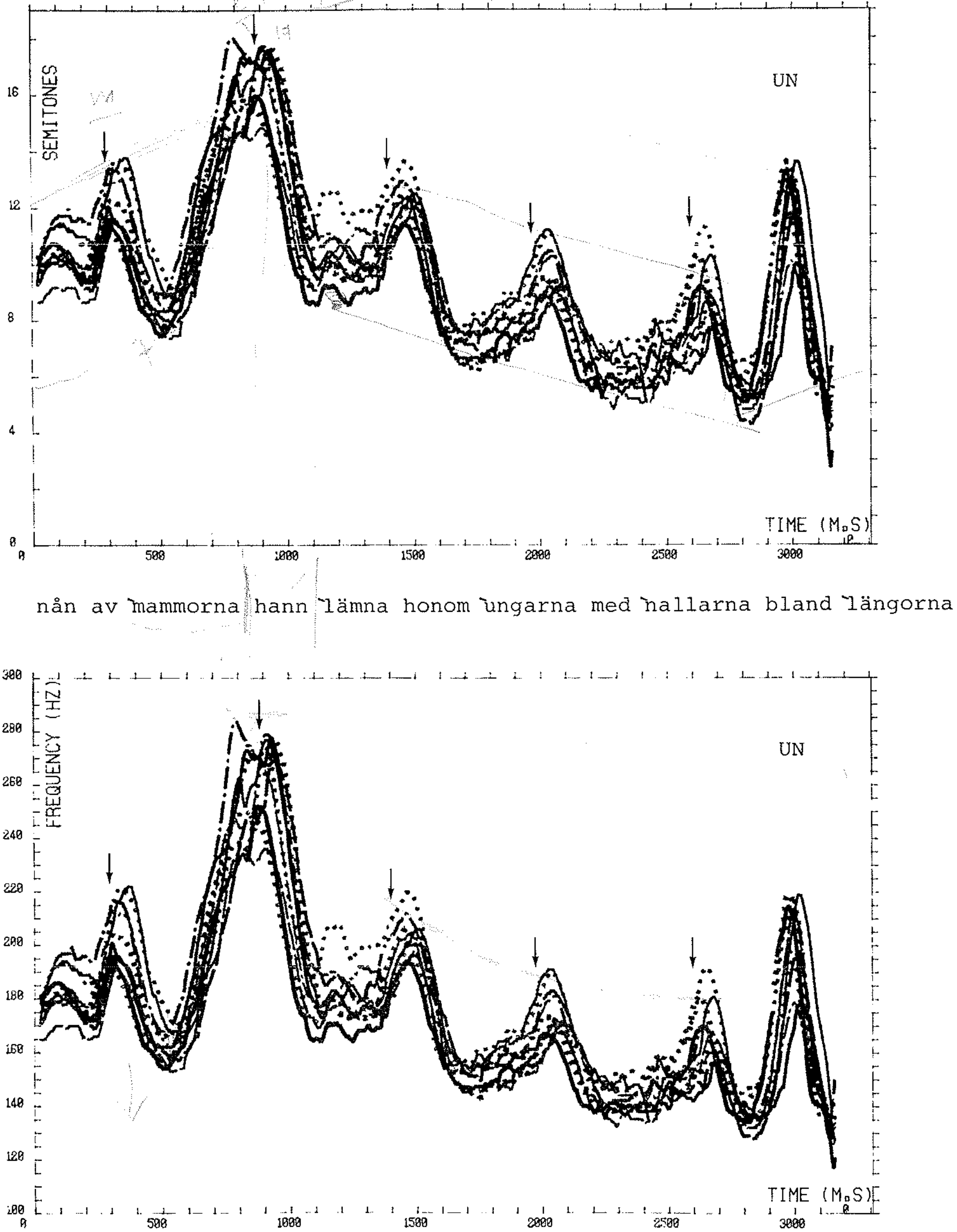
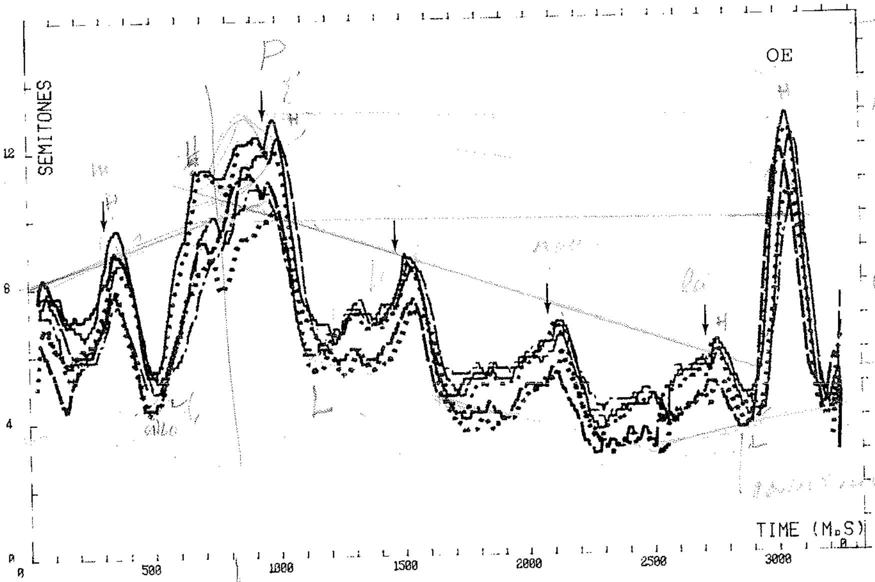


Figure 5. Fo course and utterance length (expansion to the right). Repeated Fo contours of the neutral version of a sentence containing five stress groups by a female informant (UN) with a semitone scale (above) and a Hz scale (below). The arrows indicate the stress group boundaries. The line-up point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



nån av ˋmammorna hann ˋlämna honom ˋungarna med ˋhallarna bland ˋlängorna

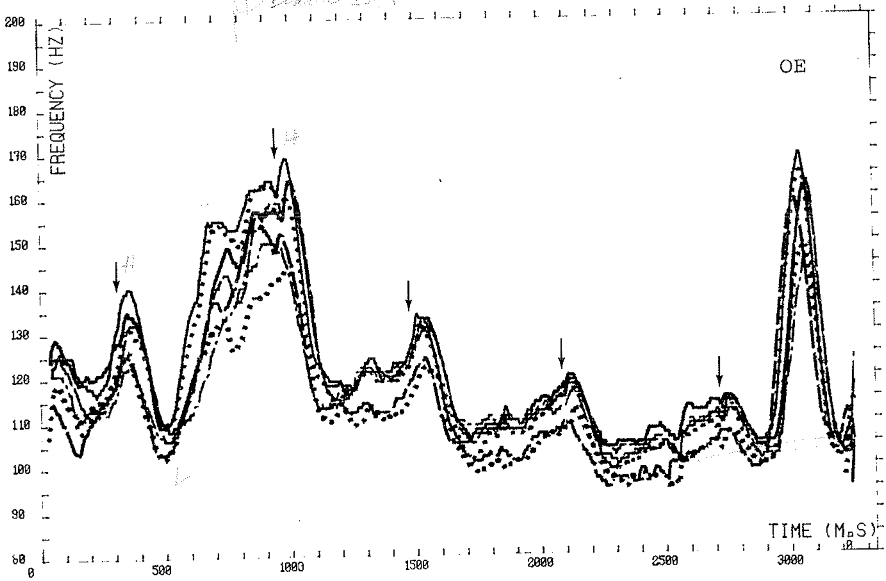


Figure 6. Fo course and utterance length (expansion to the right). Repeated Fo contours of a sentence containing five stress groups by a male informant (OE) with a semitone scale (above) and a Hz scale (below). The arrows indicate the stress group boundaries. The line-up point is the onset of the utterance.

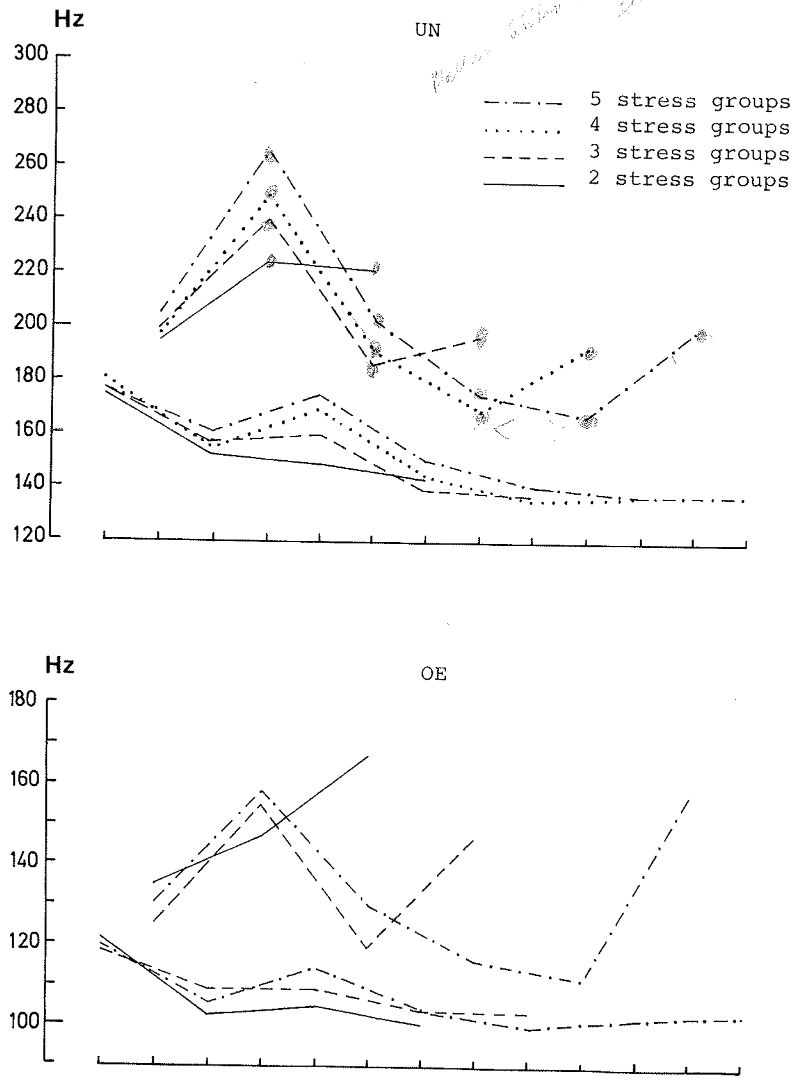


Figure 7. Fo course and utterance length (expansion to the right). Stylized baselines and toplines - means in Hz of successive Fo minima and Fo maxima - of utterances of varying length (2-5 stress groups). Neutral version for a female informant (UN) (above) and a male informant (OE) (below).

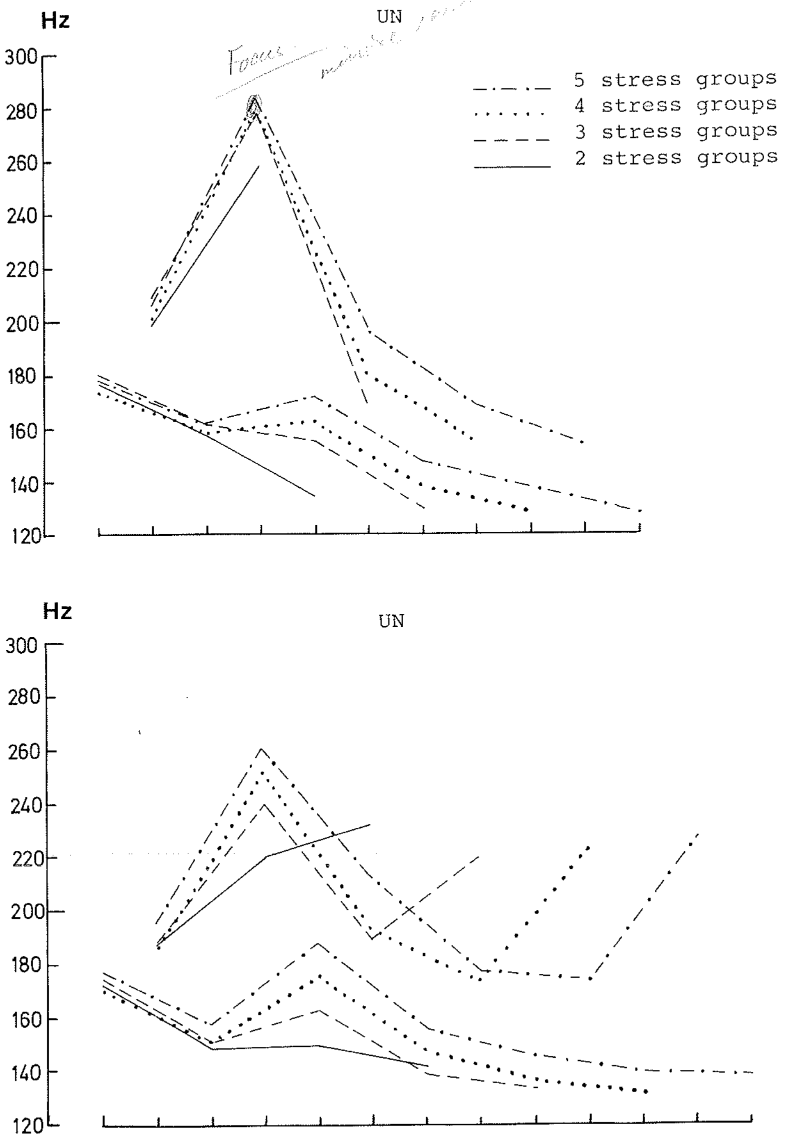


Figure 8. Fo course and utterance length (expansion to the right). Stylized baselines and toplines - means in Hz of successive Fo minima and Fo maxima - of utterances of varying length (2-5 stress groups). Initial focus version (above) and final focus version (below) for a female informant (UN).

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 ˋmammorna															
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 ˋmammorna hann ˋlämna 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 ˋmammorna hann ˋlämna honom ˋungarna															
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 ˋmammorna hann ˋlämna honom ˋungarna med ˋnallarna															
neutral focus	$\bar{x}$	181	197	155	*249*	169	191	145	169	136	*192*	137			9
	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 focus	$\bar{x}$	174	201	159	*278*	162	180	138	155	128					10
	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 focus	$\bar{x}$	170	185	151	*251*	175	193	148	174	137	*222*	131			10
	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 ˋmammorna hann ˋlämna honom ˋungarna med ˋnallarna bland ˋlängorna															
neutral focus	$\bar{x}$	117	205	162	*265*	175	204	150	176	141	168	137	*199*	137	10
	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 focus	$\bar{x}$	178	206	162	*283*	171	197	147	169	137	153	128			10
	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 focus	$\bar{x}$	177	195	157	*260*	187	212	156	177	145	174	139	*227*	138	10
	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 ˋnallarna bland ˋlängorna															
neutral focus	$\bar{x}$	122	135	103	*147*	106	*168*	100							7
	s	6,2	6,1	2,8	7,6	2,5	9,3	2,6							
	Δf		+13	-32	+44	-41	+62	-68							
nån av ˋmammorna hann ˋlämna honom ˋungarna															
neutral focus	$\bar{x}$	119	125	109	*154*	109	119	103	*147*	103					7
	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 ˋmammorna hann ˋlämna honom ˋungarna med ˋnallarna bland ˋlängorna															
neutral focus	$\bar{x}$	120	130	106	*158*	114	130	104	116	99	112	102	*157*	103	7
	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 lookahead 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 ˋmammorna] [hann ˋlämna honom ˋungarna]
One of the mothers managed to leave him the kids

expansion to the left:

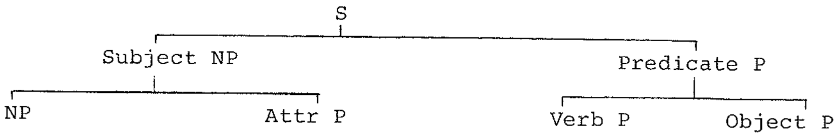
[Nån av ˋmammorna med ˋnallarna] [hann ˋlämna honom ˋungarna]
One of the mothers with the teddybears managed to leave him the kids

[Nån av ˋmammorna med ˋnallarna bland ˋlängorna] [hann ˋlämna honom ˋungarna]
One of the mothers with the teddybears among the barns managed to leave him the kids

expansion to the right:

[Nån av ˋmammorna] [hann ˋlämna honom ˋungarna med ˋnallarna bland ˋlängorna]
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 ˋmammorna med ˋnallarna bland ˋlängorna | hann ˋlämna honom ˋungarna

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.

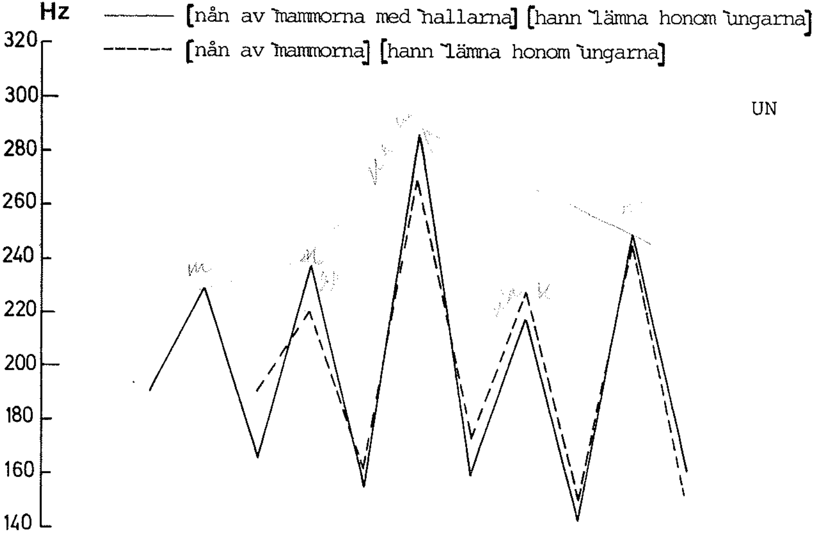
As in experiment one the recordings were processed by hardware pitch and intensity meters, and Fo was measured manually.

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.

PA



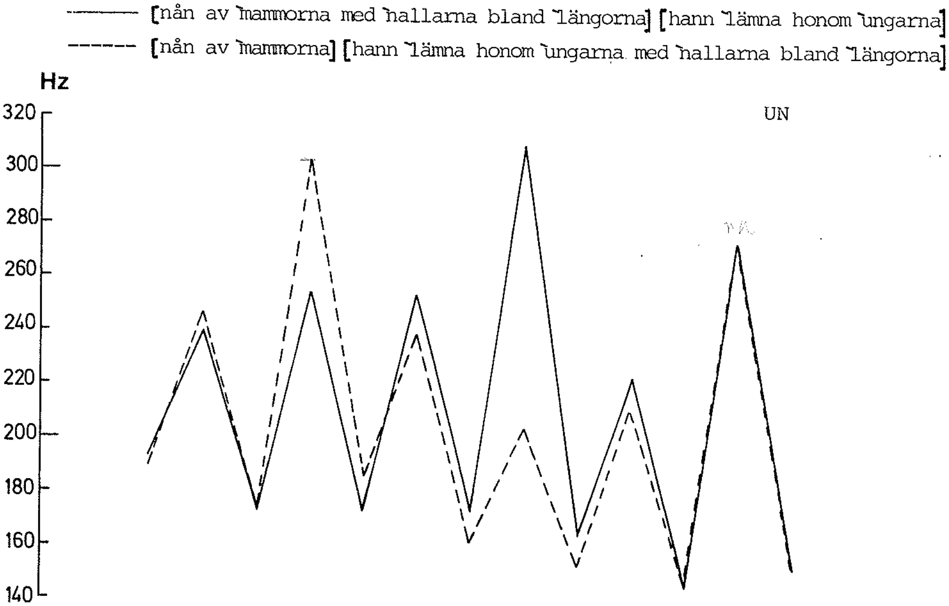


Figure 9. Fo course and utterance length (expansion to the left and to the right). Stylized Fo contours - means in Hz of successive Fo minima and Fo maxima - of a three stress group utterance compared to a four stress group utterance (above) and of two versions of a five stress group utterance (below)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

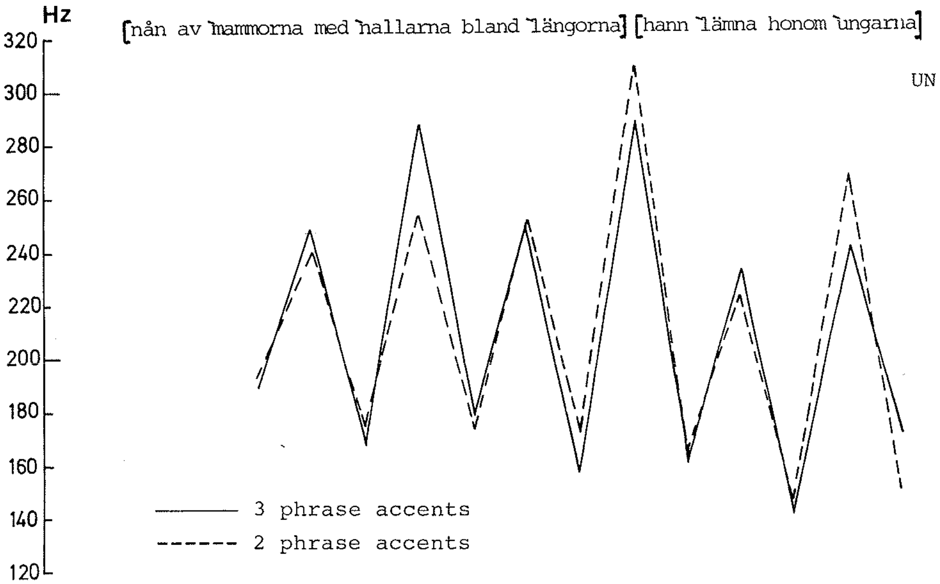


Figure 10. Fo course and utterance length (expansion to the left). Stylized Fo contours - means in Hz of Fo minima and Fo maxima - of two variants of the five stress group utterance with left expansion by a female informant (UN).

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 ˋmammorna hann ˋlämna honom ˋungarna															
neutral	x̄	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 ˋmammorna med ˋnallarna hann ˋlämna honom ˋungarna															
neutral	x̄	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 ˋmammorna hann ˋlämna honom ˋungarna med ˋnallarna bland ˋlängorna															
neutral	x̄	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 ˋmammorna med ˋnallarna bland ˋlängorna hann ˋlämna honom ˋungarna															
neutral	x̄	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	x̄	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

		min	max	min	max	min	max	min	max	min	max	min	max	min	n
nån av ˋmammorna hann ˋlämna honom ˋungarna															
neutral	x̄	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 ˋmammorna med ˋnallarna hann ˋlämna honom ˋungarna															
neutral	x̄	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 ˋmammorna hann ˋlämna honom ˋungarna med ˋnallarna bland ˋlängorna															
neutral	x̄	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 ˋmammorna med ˋnallarna bland ˋlängorna hann ˋlämna honom ˋungarna															
neutral	x̄	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.

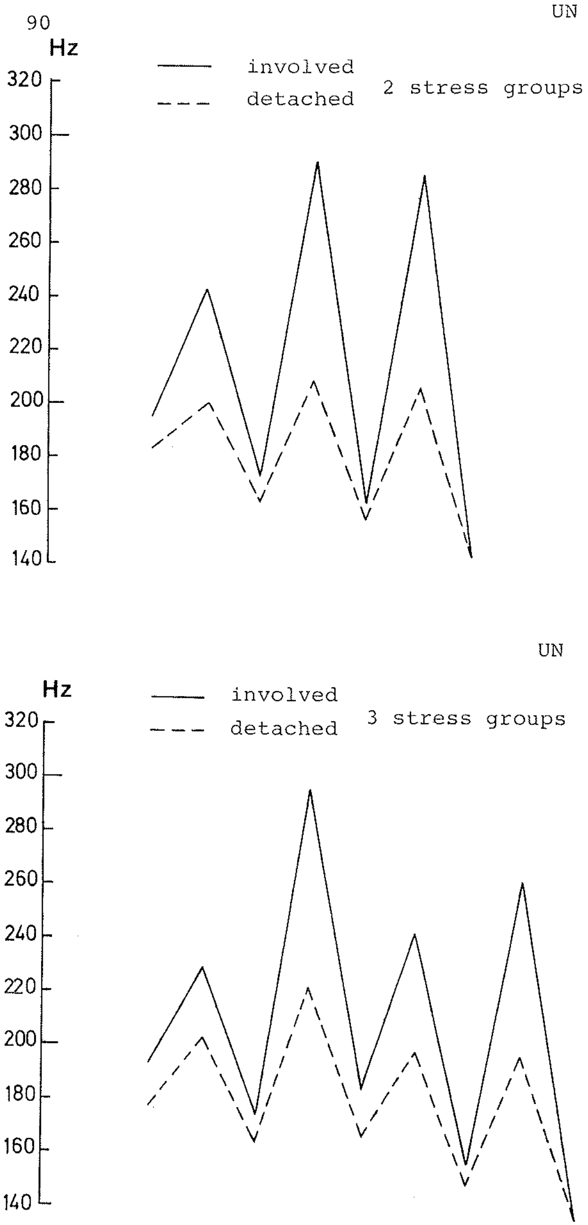


Figure 11. Fo course and attitude (involvement). Stylized Fo contours - means in Hz of successive Fo minima and Fo maxima - of an involved and a detached version of a two stress group utterance (above) and of a three stress group utterance (below) by a female informant (UN).

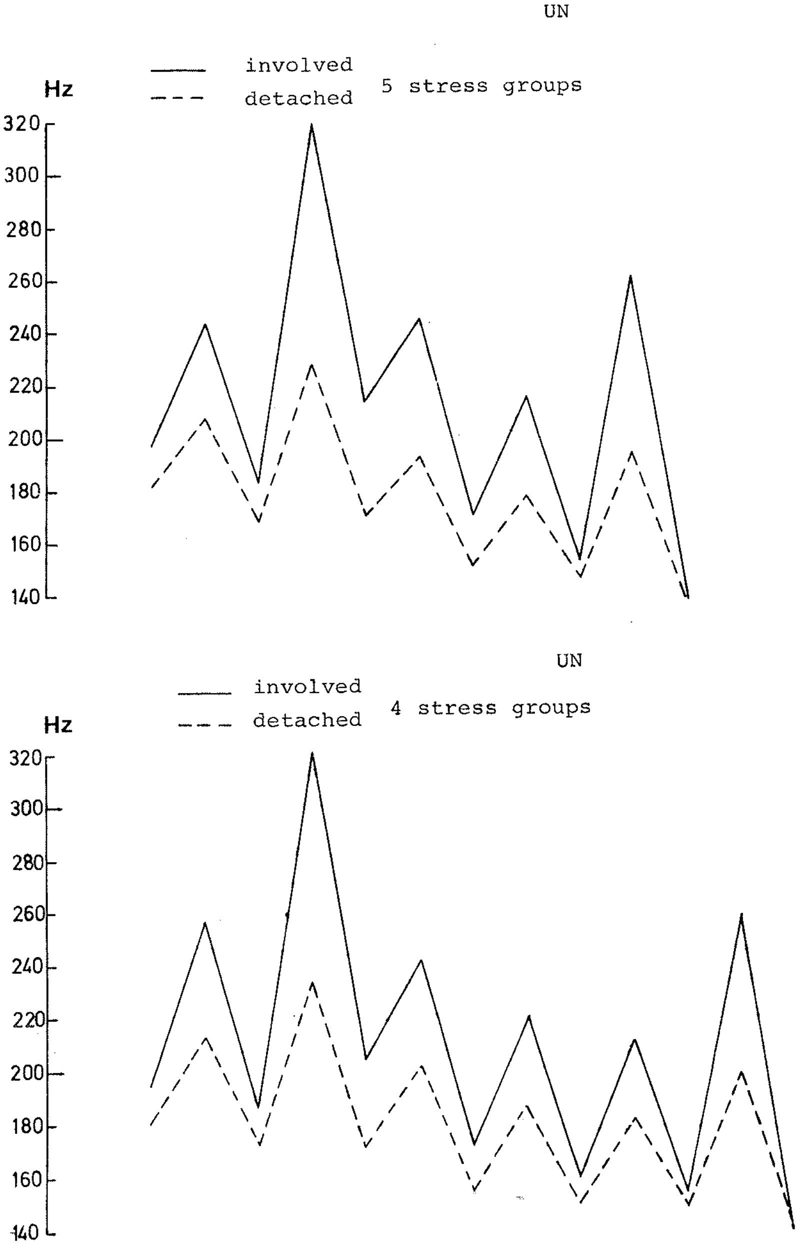


Figure 12. Fo course and attitude (involvement). Stylized Fo contours - means in Hz of successive Fo minima and Fo maxima - of an involved and a detached version of a four stress group utterance (above) and of a five stress group utterance (below) by a female informant (UN).

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	ˋmammorna hann ˋlämna honom													
detached	x̄	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	x̄	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	ˋmammorna hann ˋlämna honom ˋungarna													
detached	x̄	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	x̄	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	ˋmammorna hann ˋlämna honom ˋungarna med ˋnallarna													
detached	x̄	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	x̄	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	ˋmammorna hann ˋlämna honom ˋungarna med ˋnallarna bland ˋlängorna													
detached	x̄	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	x̄	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 ˋlämnar ˋmamman
One leaves the mother

Man ˋnallar ˋungen
One steals the kid

Man ˋlämnar ˋmamman (,) men man ˋnallar ˋungen
One leaves the mother but one steals the kid

Man ˋnallar ˋungen (,) men man ˋlämnar ˋmamman
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

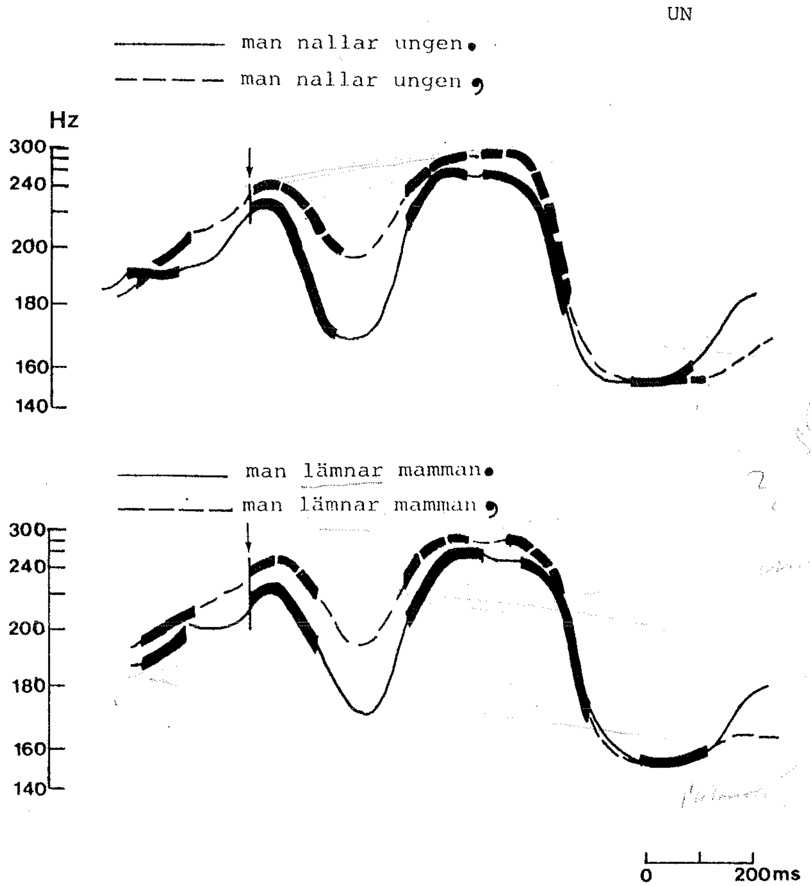


Figure 13. The effect of clause boundary. Typical Fo contours of the first clause of a double clause sentence with a physical pause at the clause boundary /,/ and of the corresponding single clause sentence /./ by a female informant (UN). The line-up point is at the CV-boundary of the first stress group.

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 ˋlämnar ˋmamman , men man ˋnallar ˋungen														
neutral	x̄	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 ˋlämnar ˋmamman men man ˋnallar ˋungen														
neutral	x̄	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 ˋlämnar ˋmamman														
initial	x̄	186	230	169	*253*	150								8
focus	s	6,6	14,1	3,8	14,0	3,2								
Man ˋlämnar ˋmamman														
final	x̄						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 ˋnallar ˋungen , men man ˋlämnar ˋmamman														
neutral	x̄	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 ˋnallar ˋungen men man ˋlämnar ˋmamman														
neutral	x̄	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 ˋnallar ˋungen														
initial	x̄	182	219	169	*245*	149								8
focus	s	8,1	6,7	5,3	10,0	2,4								
Man ˋnallar ˋungen														
final	x̄						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.

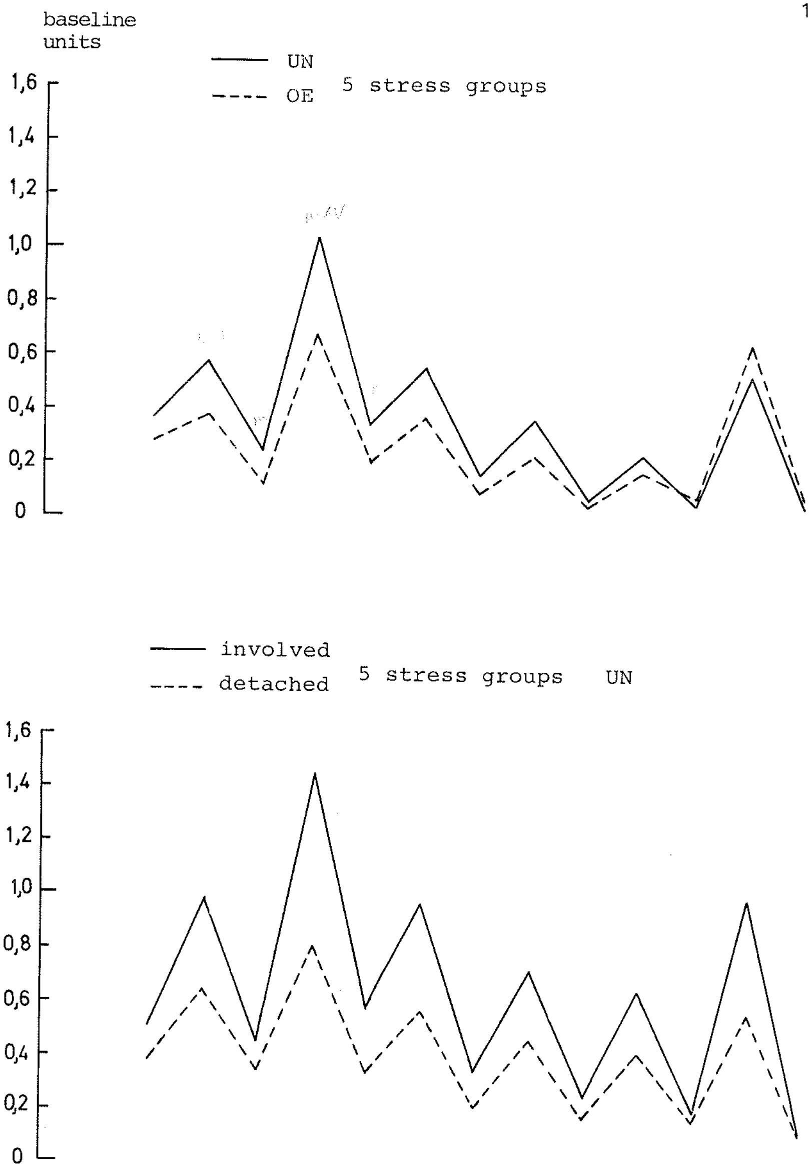


Figure 14. Normalization of Fo. Stylized Fo contours - means of successive Fo minima and Fo maxima expressed in baseline units above the baseline: (a) five stress group utterance by the female (UN) and male (OE) informants, (b) five stress group utterance of an involved and a detached version by the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

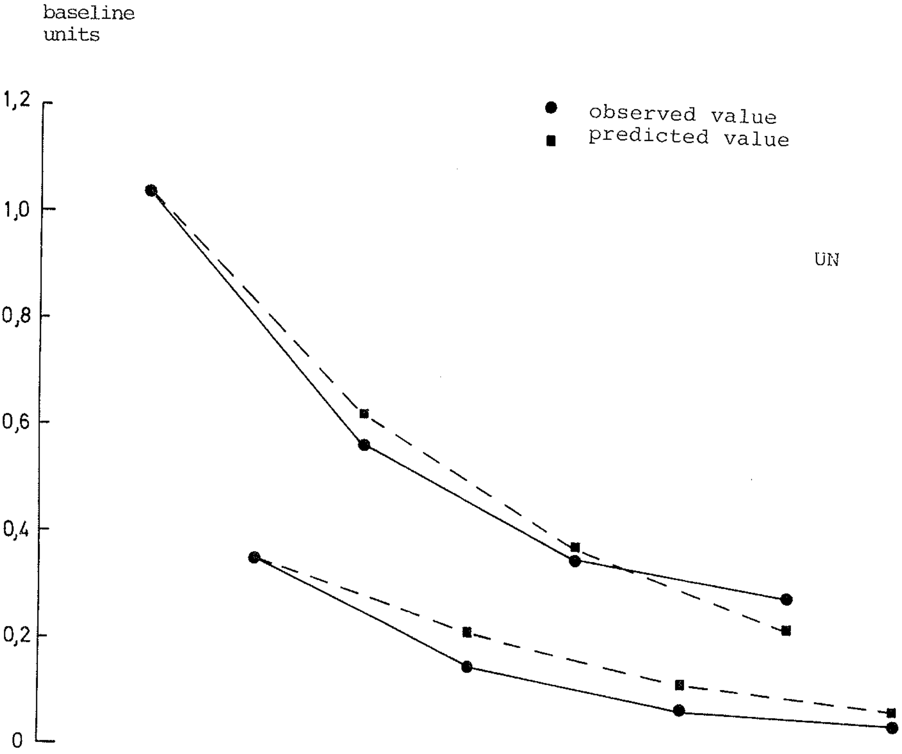


Figure 15. Prediction of Fo course. Observed and predicted toplines and baselines of the neutral version of a five stress group utterance - the second phrase except for the second phrase accent - by a female informant (UN). The initial maximum (topline) and initial minimum (baseline) values are given.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

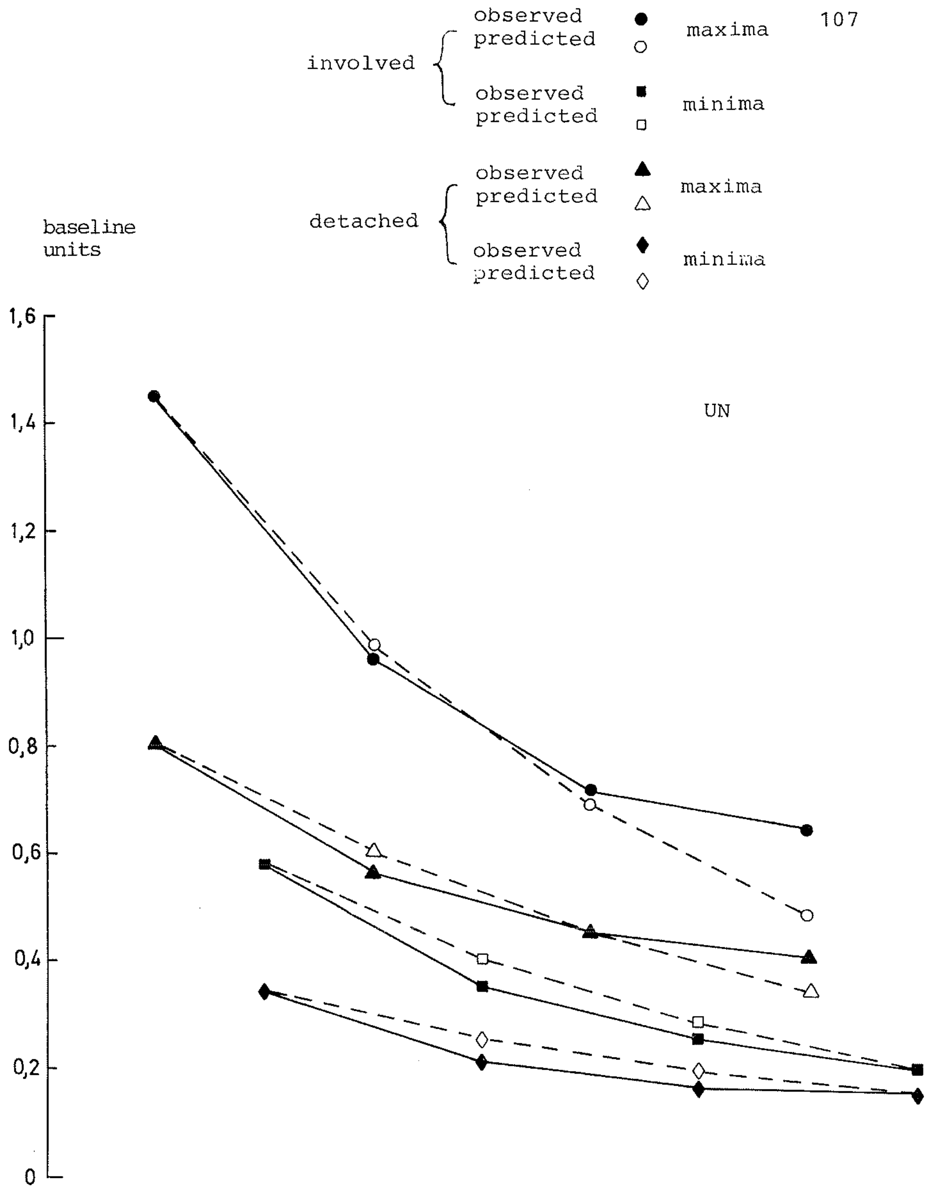


Figure 16. Prediction of Fo course. Observed and predicted toplines and baselines of the involved and detached version of a five stress group utterance by a female informant (UN). The initial maximum (topline) and minimum (baseline) are given for both degrees of involvement. Each Fo value is 0.7 x the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

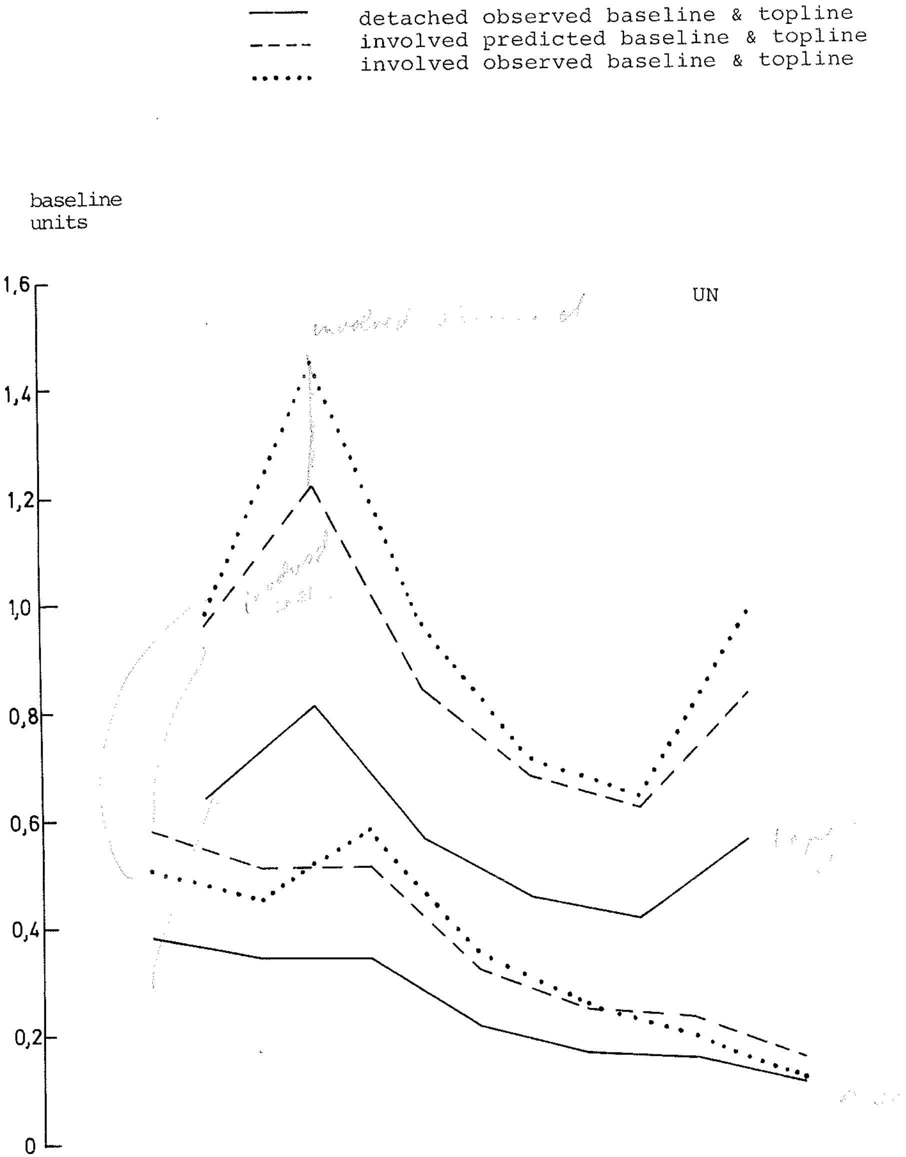


Figure 17. Prediction of Fo range. Observed topline and baseline for the detached version and observed and predicted toplines and baselines for the involved version of the five stress group utterance by a female informant (UN). Each predicted involved Fo value is 1.5 x the Fo value of the detached version.

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

PROSODIC EXPRESSIONS AND PRAGMATIC CATEGORIES*

Eva Gårding

Our earlier work in prosody has been mainly concerned with isolated sentences in a fixed situational frame (e.g. Gårding and Lindblad 1973, Bruce and Gårding 1978 and 1981). Our present goal is to show how prosody makes isolated sentences a coherent text and how it makes the coherent text part of situation-bound speech act. Such functions, here called pragmaprosodic, are now an object of current interest.

In this paper I shall present a preliminary frame for pragmaprosodic categories and report on some experiments designed to validate their existence.

THE FRAME

Prosody has at least four functions with a bearing on text and speech act. This is illustrated by Table 1, which is based on traditional analysis. The *hierarchical* function refers to prosodic means of ordering topics and comments which may be spread over several sentences. The *demarcative-connective* function refers to the absence or presence of boundary signals. These two functions are chiefly text-oriented. They divide the text into smaller units and arrange these units hierarchically. The two other functions are listener-speaker oriented. The *modal* function expresses orientation towards the listener, choice of utterance type, statement or question intonation, for instance. The *expressive* function, finally, conveys attitudes and emotions of the speaker.

Let me try to motivate the frame by looking more closely at the corresponding classes of prosodic expressions.

The *hierarchical* class comprises accent (stress) and intonation. Both can appear in one of three grades, neutral, upgraded and downgraded. Upgraded accent has a prominence-lending focal effect, downgraded accent may be anaphoric. Its physical domain is local, which means that it hits a minor part of the speech chain, a

---

* Paper presented to a conference on text linguistics in Lund, March 1981. Published in "Textstrategier i tal och skrift", Almqvist & Wiksell International, 1982.

Functions of pragmatic prosody

	TEXT ORIENTED		LISTENER–SPEAKER ORIENTED	
	hierarchical	demarcative-connective	modal	expressive
PROSODIC FEATURES		intonation g		accent g
	accent l	volume g	intonation g	intonation g
		tempo g		volume g
		voice quality g		tempo g
		tonal juncture l		lengthening l,g
		pause l		
	intonation g	accent l	tonal juncture l	pause l
		lengthening l		voice quality g

l= local domain, e.g. syllable, morpheme, word
g= global domain, e.g. phrase, sentence, text unit

Table 1

syllable or a word. The phonetic correlates of accent are varying combinations of fundamental frequency, intensity and duration. The actual combination of these correlates depends on the function (focal, non-focal) and the position in the intonation contour.

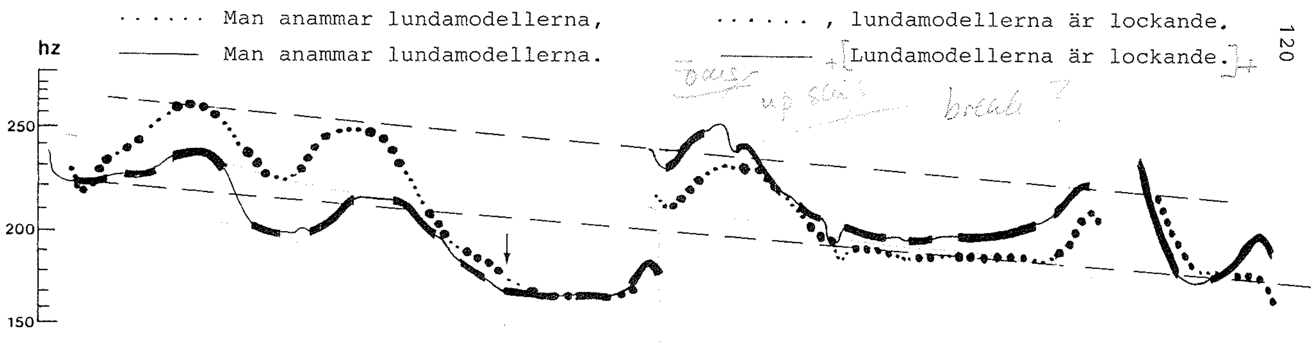
While the hierarchical function of accent is taken for granted, the corresponding function of intonation has been more or less neglected, at least experimentally. Since the domain of intonation is global, its hierarchical function is manifest in a major part of the utterance, a phrase or a sentence. The main acoustic correlates seem to be a continuation of the general course of a contour, a repetition of a specific contour, shifts up and down of a neutral contour often combined with expansion or compression of the frequency range.

/coordination
[illegible]
/subordination
[illegible]

Intonation and accent also appear in the *demarcative-connective* class. It is shown in Fig. 1, which is taken from an analysis by Bruce (1981), how a declarative intonation contour stretches over two sentences. The phonetic expression is a gradual fall of the baseline and the topline which connect the local maxima and minima of the fundamental frequency curve. In this way intonation reflects a larger constituent than the sentence, a text unit, which is set off from the rest of the text. This use of intonation at the same time has a hierarchical function.

Syntactic units may be delimited by tonal junctures, a local fall as in the declarative sentence of Fig. 1 or a local rise as in the interrogative contour of Fig. 2. There it is illustrated how a question contour connects an interrogative speech act spread over two sentences.

Accent may also be a demarcative signal. In many Swedish dialects (e.g. in Stockholm) it falls on the last accentable syllable of the phrase, if it has not been used to mark an earlier word in the same sequence for focus. As Bruce demonstrated so convincingly (1977), it is this demarcative function that gives rise to the double-peaked Accents 2. In other Swedish dialects, e.g. Finland Swedish, a different strategy is used to achieve the

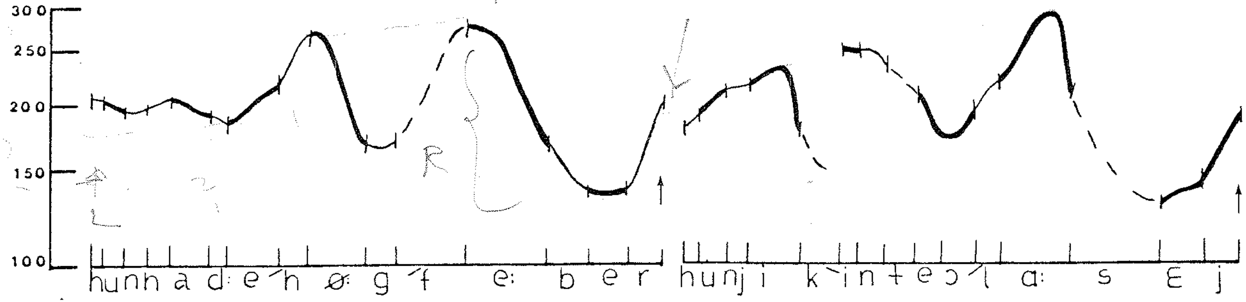


Pause up slide break ?

Fig. 1. F_o manifestation of declarative intonation stretching over two connected sentences (dotted line). Pause between sentences has been omitted. Solid line shows corresponding sentences, each with a declarative contour of its own. Hatched lines connect maxima and minima in the two-sentence text. Arrow points to tonal juncture fall. From Bruce 1980

+ plus up slide

same contour but slips down ? SA



Pause [illegible] [illegible]

Fig. 2. F_o manifestation of interrogative intonation in two connected sentences. Arrow points to tonal juncture rise. here coherence is created as the [illegible] down the same main contour

same goal. Here the initial part of the intonation contour is at an extra high level which in this way becomes a tonal juncture, signalling the beginning of the sentence. one use of?

The demarcative function of accent was implied in the concept of sentence accent (Bruce and Gårding 1978). However, the present investigation of connected sentences suggests that this accent is not related to the sentence but rather to the text. In any of the two test sentences that form a coherent text in the material used for this study (cf. Table 2), our alleged sentence accent only appears in the second sentence. Correspondingly, the initial tonal juncture of Finland Swedish belongs to the text unit.

The *modal* class includes intonation and tonal junctures. As has been mentioned already, a statement is characterized by gradually falling intonation stretching over the whole text unit, which may consist of one or several sentences. It is, then, the gradually falling fundamental frequency which marks the statement. Interrogative intonation is non-falling (Gårding 1979, Bredvad-Jensen 1980). In addition, there is a non-falling intonation at a low level, considerably lower than the question, which is very common in spontaneous narrative style (Gårding 1967). I regard this as the unmarked case. The speaker is too much involved in the textual part of what he/she is saying to care about the modal function of prosody. Tonal junctures, i.e. local falls and rises, strengthen the purport of the speech act. They are not obligatory.

The *expressive* class comprises a whole set of prosodic features, global as well as local. As a rule, they can be analysed as superimposed on the other classes. What makes them expressive is exactly their deviation from an expected pattern.

SOME EXPERIMENTS

The aim of the experiments is to explore the hierarchical function of accentuation and intonation, in particular how speakers

The hierarchical function of accent in texts consisting of two sentences

Text	Sentence 1 Ref.A Ref B	Sentence 2 Ref A Ref B	Manifested class of accents in referents (Ref) 1A/2A	 1B/2B	Relation of referents* 1A/2A	 1B/2B
1	Uno är åtta år 'Uno is eight'	Uno går i skolan 'Uno goes to school'	neutral/ downgraded neutral/ neutral	neutral/ neutral	Repetition	Equivalence
2	Janne är fyra år 'Janne is four'	Uno går i skolan	neutral/ neutral neutral/ upgraded	neutral/ neutral	Hyponymity Antonymity	Hyponymity Antonymity
3	Janne går i skolan 'Janne goes to school'	Uno går i skolan	neutral/ neutral upgraded/ upgraded	neutral/ neutral downgraded/ downgraded	Hyponymity	Repetition
4	Janne går på lekis 'Janne goes to the nursery school'	Uno går i skolan	neutral/ neutral upgraded/ upgraded	neutral/ neutral neutral/ upgraded	Hyponymity Antonymity	Equivalence Contrast
5	Janne går inte i skolan 'Janne does not go to school'	Uno går i skolan	neutral/ upgraded	downgraded/ downgraded	Antonymity	Repetition
6	Janne slutar klockan tre 'Janne finishes at three'	Uno går i skolan	neutral/ upgraded	neutral/ upgraded	Antonymity	Equivalence

*The term denotes second element as compared with first

Table 2

and listeners use these features to create coherence in a text. The first experiment concerns accentuation, the second intonation. In both experiments I use text units consisting of two sentences in which topics and comments have been varied systematically. These texts (Table 2 and 5) were read by several speakers. Recordings were made and analysed (Figs. 3 through 6). A closer analysis of these contours in terms of the earlier mentioned intonation model (Bruce and Gårding 1978) will be presented in a forthcoming paper. Some of these sentences were later presented in tests in which subjects were asked to match them in such a way that they formed natural text units (Tables 3 and 4). Similar experiments have been reported by Fónagy (1981).

The speakers, who were all phoneticians, represented four different dialects of Swedish, southern (EG), eastern (Stockholm, UN), far east (Helsingfors, KT), and received pronunciation, as taught in the drama school of the National Theatre in Stockholm (GS). Typical productions by Speakers UN and EG are presented in Figs. 3 and 4 for accentuation and in Figs. 5 and 6 for intonation.

The speakers produced the text units in three different orders. They were asked to ponder the content of the text and express it prosodically as clearly as possible. To secure natural productions, the relevant text units were interspersed with other text units, constructed after a different design.

1. Accentuation

The material used in this experiment is presented in Table 2. The sentences in the leftmost part of this table occur in pairs. Every sentence has a topic (Referent A) and a comment (Referent B) which are differently related to the topic and comment of the neighboring sentence. All the sentences have the same syntactic structure, a noun phrase followed by a verb phrase. The content has been chosen in such a way that we can expect the speaker to strengthen and weaken the accentuation of related topics and comments according to a predetermined plan (cf. middle part of Table 2). All sentences except one (no. 6, Table 2) have declarative intonation. For orientation let us look at the first example. There are three relations to be considered, the relations between the two referents A and B in the paired sentences and the relation between the two sentences 1 and 2. The relation between the referents is expressed in text-linguistic terms in the right-most part of the table. The middle part shows how accentuation has been expected to be used to bring out these relations. It may be neutral, downgraded or upgraded. We notice that the prosodic categories are very simple as compared to the text-linguistic ones.

In text 1, referent A, *Uno*, is repeated in sentence 2. This may happen in a communicative style used when grown-ups talk with children, as if pronouns were too weak to serve as referents. The repeated referent may be expressed by a downgraded accent. We notice how in both dialects (Figs. 2 and 4) the downgrading affects the prefocal part (*Uno går i*) by a narrower range of pitch. Referents B, on the other hand, are two equivalent pieces of information. Still, referent B of sentence 2 is upgraded. This accent is interpreted as having a demarcative function. Apart from the relation of the referents there is also the relation between the two sentences that together form a text unit. In text 1 they are given equal weight. This equivalence seems to be achieved by the use of the same contour in both sentences with a slight downshift of the intonation of sentence 2.

13

The referents of the sentences in text 2 can be related in different ways. We can give referents A, *Janne* and *Uno*, an equal neutral degree of accent which makes them equivalent pieces of communication (Fig. 3). We can also give *Uno* an upgraded accent and create a constrastive effect in meaning, equivalent to the expression *Uno, however*. A third possibility was used by two of the speakers: The whole of sentences 2 was contrasted with sentence 1 by the repetition of an expanded version of the same intonation pattern (Fig. 4).

Text 3 is deviant. For the recording it was supplied with the paranthesis *Everybody is at school*. In this way sentences 1 and 2 actually form part of a three-sentence unit. The treatment of the A referents varied. Either the referents were given equal weight, or else *Uno* of the second sentence seemed to be favoured. The intonation contour appropriate for the sentences of this text unit is not a simple declarative intonation as in earlier cases but one used in repetition of equal elements as in counting. The speakers have achieved this effect by using the same intonation contour for both sentences but with a downshift of sentence 2 in a compressed range. /e

Text 4 called for equal treatment of the referents by means of neutral accents in a repeated intonation contour. There is a slight downshift of the total pattern of sentence 2.

In text 5, on the other hand, an upgraded accent seems to be obligatory in the A referent *Uno* to express contrastiveness. At the same time there is an upshift of the pattern together with a compression of the range in one speaker (EG), and an upshift with expansion in the other (UN).

In text 6, finally, there is a choice between neutral accents (UN) or an upgraded accent in the referent of the second sentence (EG). Furthermore the whole of sentence 2, which is interrogative, seems to be contrastive with the whole of sentence 1, which is declarative. This contrastive effect may be due to the change of intonation contour. To summarize the comments on the productions: Accentuation was used as expected in three classes to express hierarchical relations. These accents, however, were often accompanied by variations in the basic contour, upshift, downshift, compression and expansion.

The productions of Speaker UN were used in a listening test with seven subjects. A subject was asked to match one of the six productions of sentence 2 with a given sentence 1, called the anchor. The matching was done with the aid of a card reader, a device permitting comparison between different stimuli and the

Accentuation

Results of matching task with card reader (7 subjects).
Productions U.N.

Sentence 2

Uno går i skolan
'Uno goes to school'

Sentence 1	1	2	3	4	5	6
1. Uno är åtta år 'Uno is eight'	4			3		
4. Janne går på lekis 'Janne goes to the nursery'	1			5	1	
5. Janne går inte i skolan 'Janne doesn't go to school'					7	

Table 3

Intonation

Results of matching task with card reader (7 subjects).
Productions U.N.

Sentence 2

Hon gick inte o(ch) la sej
'She didn't go to bed'

Sentence 1	1	2	3	4	5	6
1	7					
2		5				2
3			4	3		

Table 4

anchor, any number of times and at a leisurely pace. Three anchors were used from texts 1, 4 and 5, representing pairs of sentences with neutral, upgraded and downgraded accents on the referents.

The task turned out to be quite easy. It did not take long for the listeners (all students of phonetics) to find the correct pairmate. Informal tests indicate that the tendency is the same with sentence 2 as the anchor.

The results of the matching test are given in Table 3. Broadly speaking, the listeners preferred the original mate of the anchor.

All listeners agree that sentence 5.2 is the correct mate for 5.1. Here the anchor 5.1 seems to require an upgraded accent in referent A, a downgraded accent in referent B and a terminal declarative intonation contour. Only one sentence fulfills these requirements, sentence 5.2! No. 3.2 has the required accent pattern but a non-final intonation contour which does not fit the given context.

With anchor 1.1 the sentences 3.2 and 6.2 have wrong intonations, 3.2 and 5.2 a wrong downgraded accent on referent B. Sentence 2.2 seems acceptable in intonation and accent but its tempo is too slow for the anchor. The votes are about equally divided among the two remaining ones, 1.2 and 4.2. The upgraded accent of *Uno* in 4.2 sounds plausible but intriguing. It gives an element of surprise to the message and obviously puts the example in the expressive class.

With anchor 4.1 similar arguments pertaining to intonation, accentuation and tempo show 4.2 to be the most natural match. It received 5 out of 7 votes.

To produce the two sentences as a semantically coherent unit the speaker uses not only accentuation, as planned by the experimenter for some of the text units, but also intonation and tempo and perhaps other features as well, overall volume and

The hierarchical function of intonation in texts consisting of two sentences

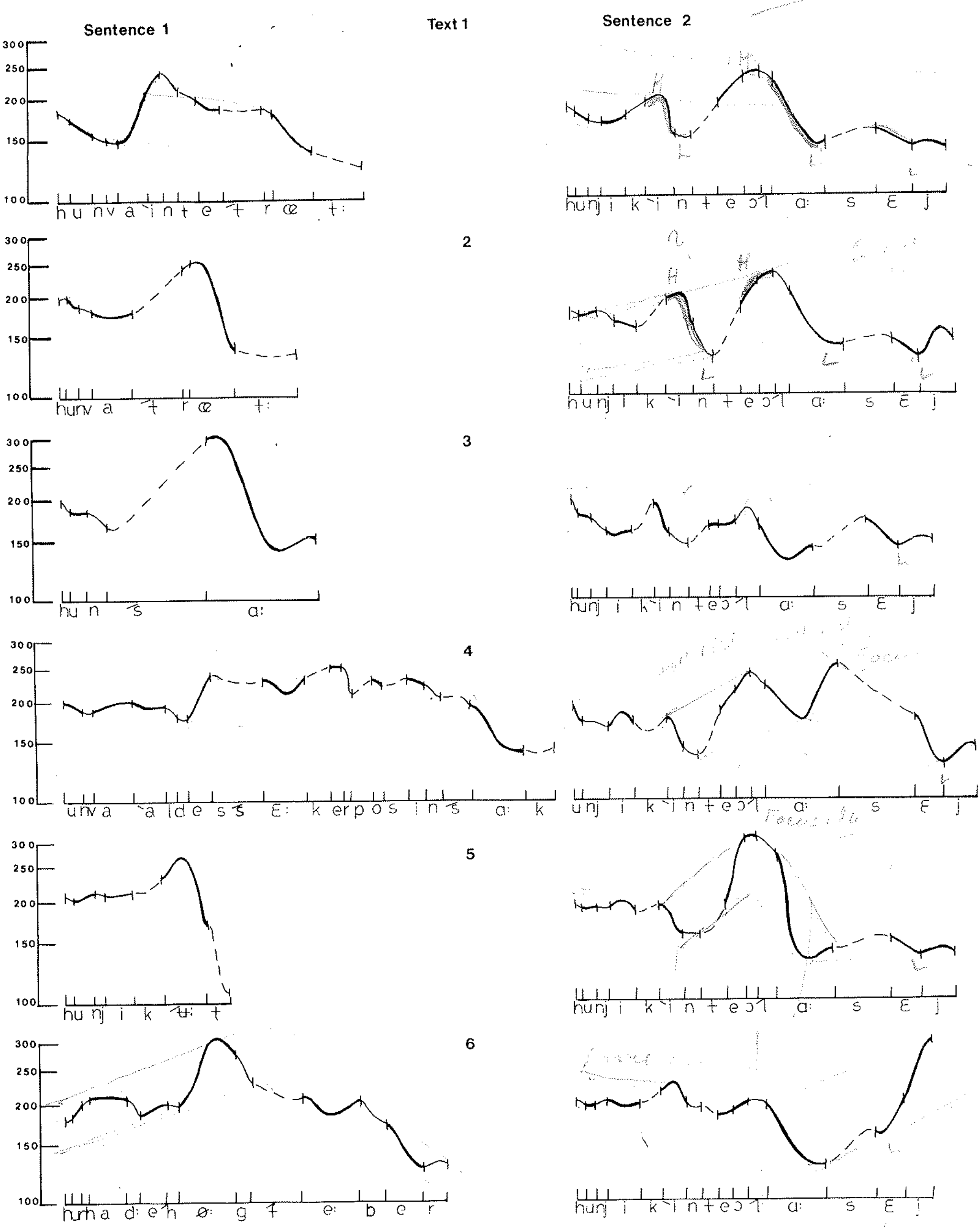
Text unit			Manifested class of intonation		Relation of intonations
Text	Sentence 1	Sentence 2	Sentence 1	Sentence 2	Sentence 1/Sentence 2*
1	Hon var inte trött 'She wasn't tired'	Hon gick inte o(ch) la sej 'She didn't go to bed'	Neutral	Neutral	Additive
2	Hon var trött 'She was tired'	Hon gick inte o(ch) la sej	Neutral	Upgraded	Adversative
3	Hon SA 'She SAID'	Hon gick inte o(ch) la sej	Upgraded	Downgraded	Out of place Couplement [illegible]
4	Hon var alldeles säker på sin sak 'She was quite sure'	Hon gick inte o(ch) la sej	Neutral	Upgraded	Concessive ?
5	Hon gick ut 'She went out'	Hon gick inte o(ch) la sej	Neutral	Upgraded	Contrastive
6	Hon hade hög feber? 'She had a high temperature?'	Hon gick inte o(ch) la sej?	Neutral	Neutral	Additive

/m

* The term denotes second element as compared with first

Table 5

appropriate - obviously [illegible] for
an altogether [illegible]
intonation

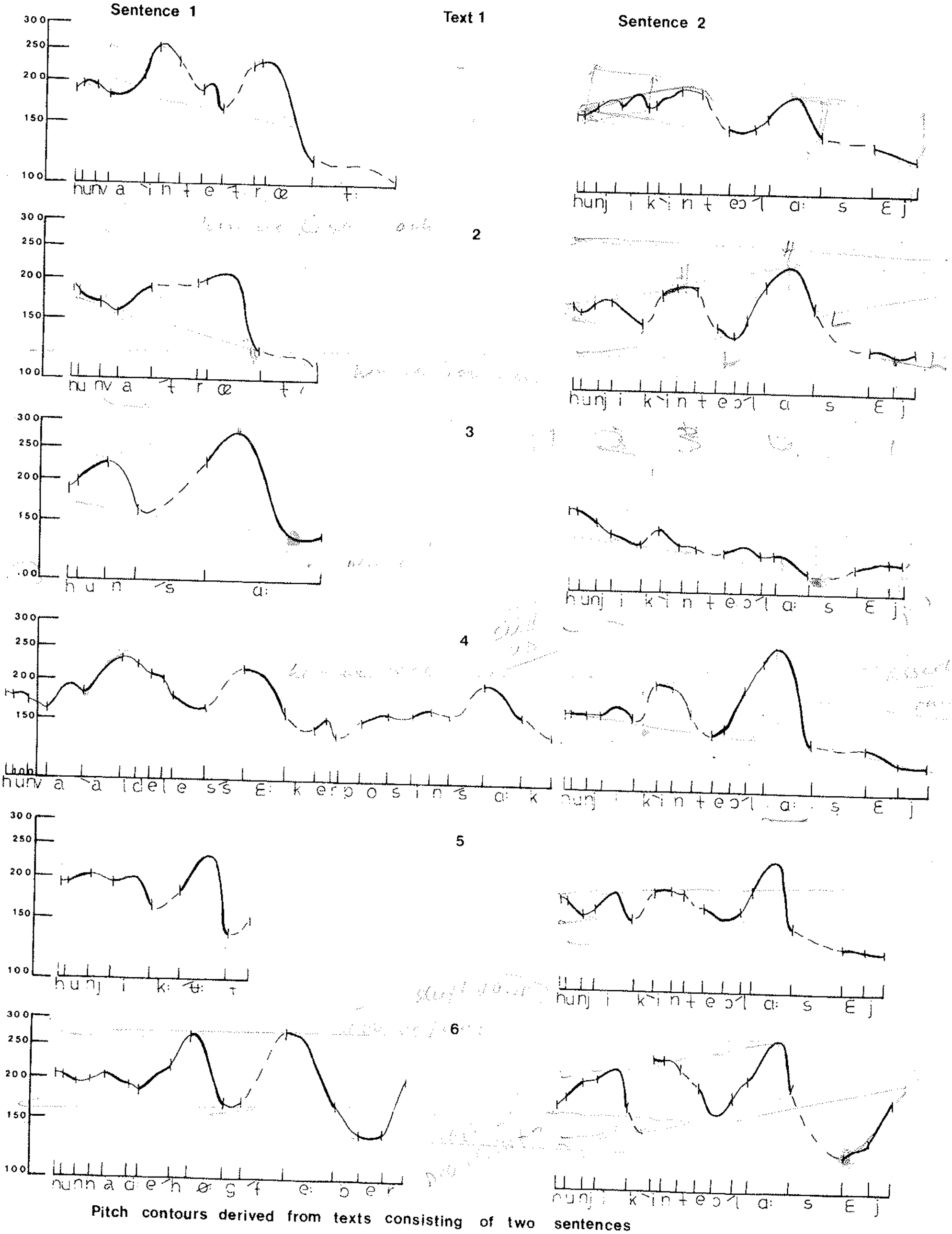


Pitch contours derived from texts consisting of two sentences

Speaker UN Stockholm

Fig. 5

notice fixed
position of rise, sound



Pitch contours derived from texts consisting of two sentences

Speaker EG Skåne

Fig.6

[illegible]

voice quality, for instance. To a large extent the use of such features seems to follow general principles. This is evidenced by the fact that listeners are able to recognize the missing part of the text by prosodic features only. The test also indicates that intonation and tempo are perhaps more important cues than accentuation.

2. Intonation

Table 5 shows the sentences which have neem constructed to test the power of intonation to express hierarchy of sentences. The text of sentence 2 is always the same, *Hon gick inte och la sej* 'She didn't go to bed', but it had been pronounced at different intonation levels which I analysed as neutral, shifted up and shifted down, depending on the situation described by sentence 1. The different relations between the sentences in the text have been described in text-linguistic terms in the rightmost column of the table. In the phonetic class only three terms are needed, neutral, upgraded and downgraded. The text-linguistic terminology is richer. The following comments on the production of these sentences refer mainly to Speaker EG, Fig. 6.

H been

Text 1 was pronounced with a neutral declarative sentence intonation in both sentences. These sentences are connected by a global fall stretching over both sentences (compare also Fig. 1 from Bruce 1980). This relation is textually additive in the sense that the information of sentence 2 is added to sentence 1. An equivalent syntactic arrangement would be to conjoin the two sentences by AND.

rising

In text 2 (Fig. 6) sentence 2 has been pronounced with shifted up intonation which has been interpreted by the experimenter as the prosodic expression of contrast. A common text-linguistic term for this relation between the comments is adversative. The shifted-up intonation plays the same role as a conjunction like BUT or an adverb like ON THE OTHER HAND.

rising

In text 3 shifted-down intonation in sentence 2 is natural after focus in the complement clause.

[illegible]

[illegible]

Shifted-up intonation in sentence 2 of text 4 expresses assertion and plays the same role as an adverb like CERTAINLY.

In text 5 shifted-up intonation has a contrastive effect and in the last example, no. 6, a [illegible] (with a terminal end ~~the same~~ contour has been used ~~in~~ ~~the interrogative contour.~~ for interrogative mode.

There is some interspeaker variation in the accentuation of these sentences. Speaker UN has a preference for a strong accent in the negation *inte* which in most cases is followed by an equally strong accent in the following predicate verb *la*. But again, the agreement is great enough to show that speakers use neutral, upgraded and downgraded intonation in a systematic way to interrelate phrases and sentences. These three prosodic categories may or may not coincide with syntactic coordination, superordination and subordination.

Test 2 (Table 4) was designed to test listeners' ability to recognize prosodic hierarchy and use it to achieve semantic coherence. Also in this test the instruction was to find the best match for three anchors (1, 2 and 3), one at a time. Another group of phonetics students served as subjects. The results are shown in Table 4.

Again the original pair-mate was found easily. In the cases where another candidate has been suggested, this is also a possible choice. A question intonation *Hon gick inte och la sej?* after *Hon var trött* as in no. 6 is a quite natural continuation and 4.2 after HON SA makes the sentence a direct quotation rather than an indirect one.

[illegible]

CONCLUSION

Our experiments seem to show that speakers use different levels of accentuation and intonation in a consistent way to express hierarchy between referents. Similarly, listeners rely on these cues to establish meaningful relations between sentences. The hierarchical function, then, is effected by three classes of phonetic expressions, neutral, downgraded and upgraded, which is called accent (stress), if the domain is local and intonation if it is global. It is semantic weight and semantic relations that determine the class.

ACKNOWLEDGEMENT

Kerstin Tevajärvi, an assistant working on the project *Swedish Prosody*, gave me efficient technical help.

REFERENCES

Bredvad-Jensen, A.C. (1980) Tonal Interaction between Attitude and Grammar. In Fretheim, T. (Ed.) Nordic Prosody II, pp. 51-62. Trondheim: Tapir

Bruce, G. (1977) Swedish Word Accents in Sentence Perspective. Travaux de l'Institut de Linguistique de Lund XII. Lund: Gleerups

Bruce, G. (1980) Om prosodisk koppling i svenska. In Fries, S. and Elert, C.-C. (Eds.) Svenskans beskrivning 12, pp. 64-77. Umeå

Bruce, G. and Gårding, E. (1978) A prosodic typology for Swedish dialects. In Gårding, E., Bruce, G., and Bannert, R. (Eds.) Nordic Prosody. Travaux de l'Institut de Linguistique de Lund XIII, pp. 219-228. Lund

Enkvist, N.E. and Nordström, H. (1978) On textual aspects of intonation in Finland-Swedish newscasts. Studia Linguistica XXXII:I-II: 63-79

Fónagy, I. (1981) Fonction prédicative de l'intonation. In Léon, P. and Rossi, M. (Eds.) Problèmes de prosodie, pp. 113-120. Paris: Didier

Gårding, E. (1967) Prosodiska drag i spontant och uppläst tal. In Holm, G. (Ed.) Svenskt talspråk, pp. 40-85. Stockholm

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

Gårding, E. and Bruce, G. (1980) A presentation of the Lund model for Swedish intonation. In Fretheim, T. (Ed.) Nordic Prosody II, pp. 33-39. Trondheim: Tapir

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

Lehiste, I. (1975) The phonetic structure of paragraphs. In Cohen, A. and Nooteboom, S.G. (Eds.) Structure and Process in Speech Perception, pp. 195-203. Berlin: Springer

+ was a mandolin from Mario?

A COMPARATIVE STUDY OF SWEDISH, GREEK AND FRENCH INTONATION

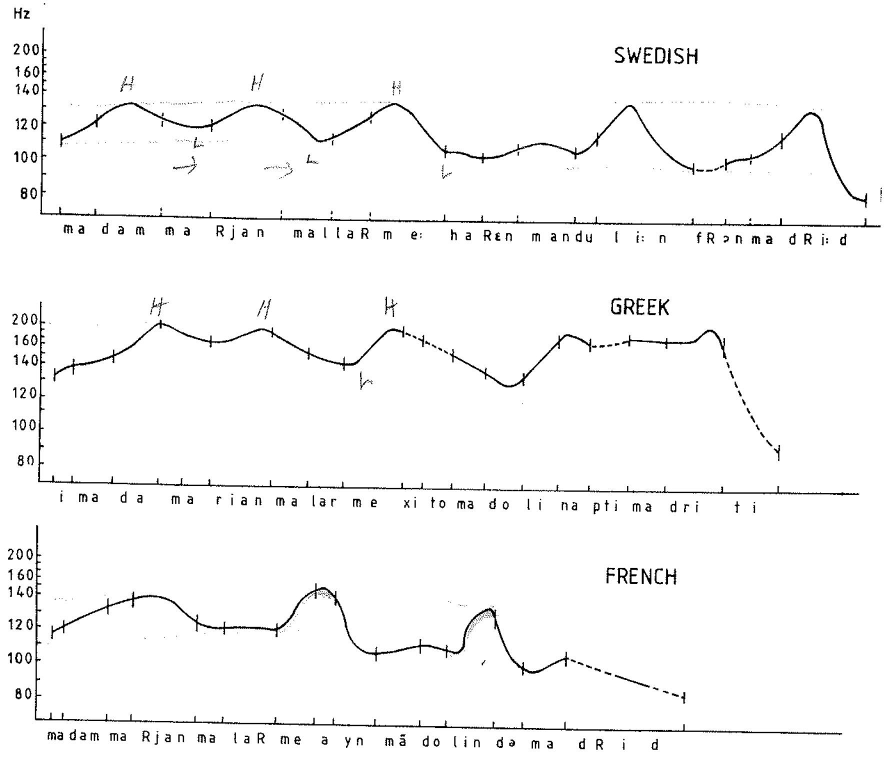
Eva Gårding, Antonis Botinis and Paul Touati

AN INTONATION MODEL

Most intonation contours look very complex at first sight. Experience shows, however, that we can get simple descriptions by indicating the important turning points, their pitch values and their positions in the sentence. The contour is then obtained by connecting the turning points with a smooth curve over the voiced segments.

What are then these important turning points? With an appropriate material, consisting of sonorant segments that do not disturb the local pressure conditions at the glottis, it can be shown that the turning points cooccur with crucial communicative events, such as lexical accents, focal accents at phrase and sentence level, morphological and phrasal boundaries. All of these linguistic variables are embadded in one and the same curve.

A model based on these general observations has been developed for Swedish dialects, (e.g. Gårding & Lindblad 1973, Gårding 1977, 1979, Bruce 1977, Bruce & Gårding 1978, 1981) and in this paper we shall demonstrate how it can be used to describe the intonation of other languages as well. For the demonstration we shall use one particular sentence for which the intonation will be generated in Swedish, Greek and French. The sentence is 'Madame Marianne Mallarmé' which has been uttered in a narrative style with equal semantic weight given to the words *Mallarmé*, *mandolin* and *Madrid* as if it had been the first sentence of a short story. Typical observed sentences are shown in Figure 1. The corresponding inputs to the model are given in the lower part of the figure. There are markings for word accents (A), for Swedish only one of the two existing accents is used (A1), phrase accents (PA) and sentence accent (SA).



PA PA SA
A1 A1 A1 A1 A1
Swedish St[madam: marian: malarme://har en manduli:n/frɔn madri:d]St

A A PA A PA SA
Greek St[i madam marian malarme//exi to mandolino/apo tin madriti]St

PA PA SA
French St[madam marian/malarme//a yn mãdɔlin/də madrid]St

word boundaries: spacing
phrase boundaries: / //
sentence boundaries: [ ]

A : word accent
A1: acute word accent
PA: phrase accent
SA: sentence accent
St: statement

Figure 1. Statement intonation in observed sentences above. Inputs to the model below. From Gårding 1981.

it is not an extra boundary.

Brackets and slashes show three kinds of syntactic boundaries, for the sentence, the major phrases and the phrase internal constituents. In this example we have tried to find a complete correspondence between syntactic and prosodic boundaries, which of course is very unusual in spontaneous speech. Nevertheless, the French speaker has put in an extra boundary before *Mallarmé*. The sentence boundary is labelled *St* for statement. Figure 2 presents the stages of the model (for a fuller account see Gårding 1981).

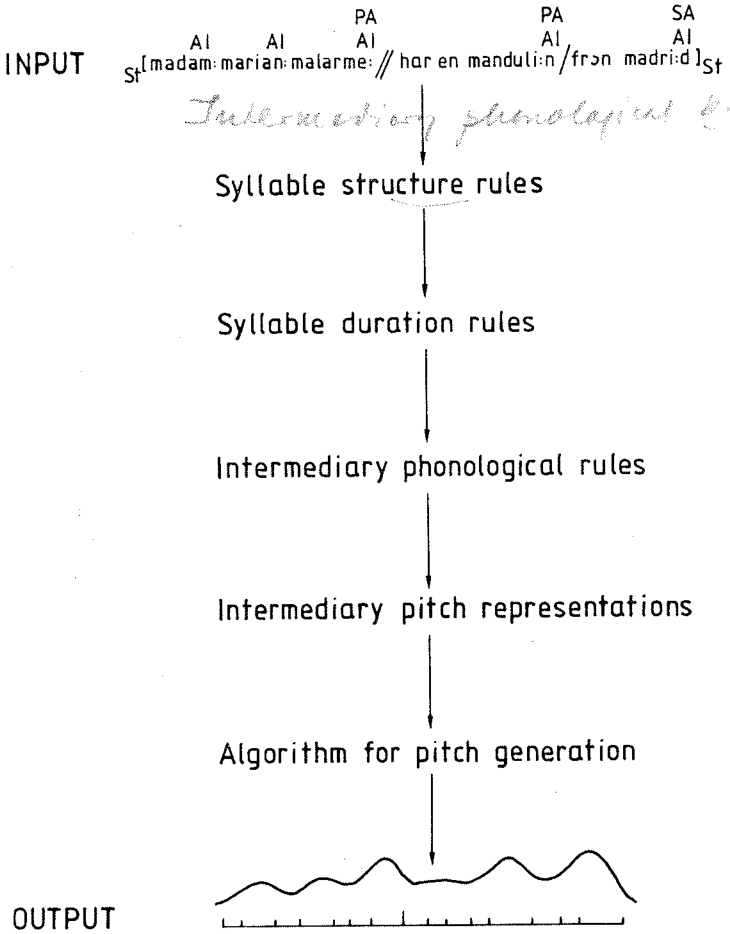


Figure 2. Model for prosody. From Gårding 1981.

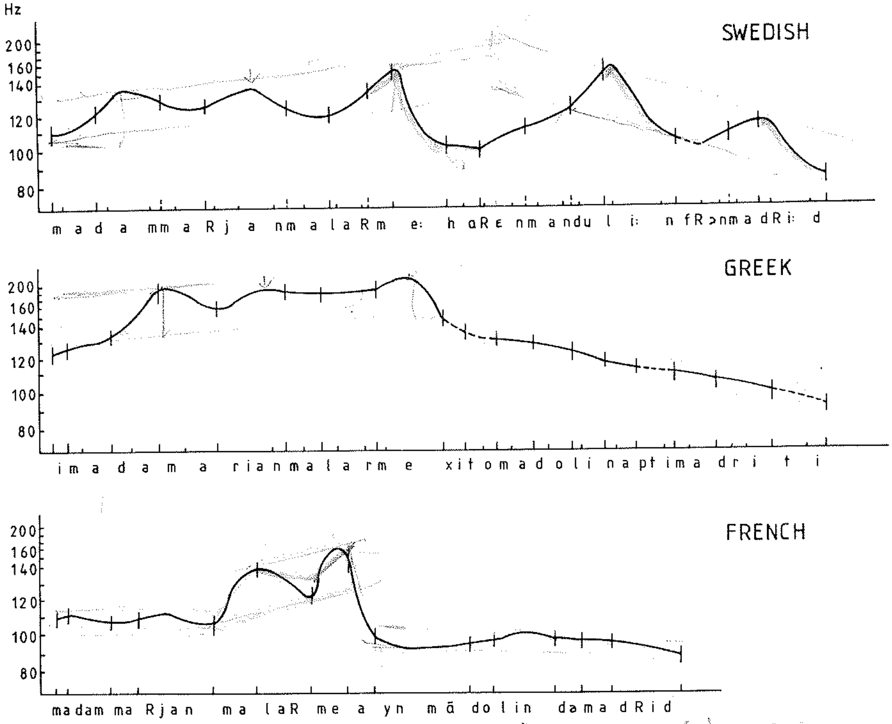
Here we shall assume that the input phrase has passed through the first two stages of the model, the syllable structure rules and the syllable duration rules, the effects of which have been demonstrated earlier (Gårding 1981). These rules give syllable durations and total durations consistent with the observed data. The French sentence is short mainly due to the lack of long accented syllables, the Swedish sentence is long because of the presence of such syllables.

Provided with the correct durations, the demonstration phrase is ready for pitch assignment. At the third stage, the intermediary phonological rules adjust context dependent symbols. One phenomenon which is fairly general across languages is that there are no pitch obtrusions in connection with accents after focus. Hence the word accent symbol which has already produced the duration appropriate for an accented syllable will here be deleted by an intermediary phonological rule. It is in this rule among others that our three compared languages differ. The pitch obtrusions characteristic of accents are not present after focus in Greek and French. They occur in Swedish, however (see examples in Figure 3).

The rules of the fourth stage, the intermediary pitch representations, transform the abstract symbols into more concrete ones. Global features, i.e. features connected with sentence and phrase, are expressed as a Rise, a Fall or a Level. Local features, those that pertain to syllables or words, are expressed as Highs and Lows which are going to become turning points in the final pitch contour.

The algorithm for pitch generation (Figure 4) works on this new notation. It has the following rules:

RULE 1 produces a tonal grid which represents the overall plan for sentence intonation. For a declarative sentence of the given structure it looks like the grid of Figure 5. The interior solid lines represent the bounds of not weakened accents within a phrase. The exterior hatched lines



in connection with

Figure 3. Different treatment of pitch ~~after~~ focus. From Gårding 1981.

[illegible]

syntactic & [illegible]

are used for pragmatic effects, e.g. for strengthened accents as in combination with focus, here called sentence accent (SA). To construct this grid we use knowledge that we have gathered from studying the same intonation contour in sentences with different phonetic and grammatical structure, uttered by different speakers. Results of interest for this grid will be found in Bruce's work (1982).

RULE 2 inserts Highs and Lows for sentence and phrase boundaries.

RULE 3 inserts the Highs and Lows of the sentence and phrase accents.

In French.

RULE 1. SENTENCE AND PHRASE INTONATION

DRAW THE TONAL GRID USING SENTENCE INTONATION AND MAJOR PHRASE BOUNDARIES

RULE 2. SENTENCE AND PHRASE BOUNDARIES

INSERT HIGHS AND LOWS ON THE GRID ACCORDING TO LANGUAGE AND DIALECT

RULE 3. SENTENCE AND PHRASE ACCENT

INSERT HIGHS AND LOWS ON THE GRID ACCORDING TO LANGUALGE AND DIALECT

RULE 4. WORD ACCENT

INSERT HIGHS AND LOWS ON THE GRID ACCORDING TO LANGUAGE AND DIALECT

RULE 5. CONTRASTIVE WORD ACCENT

ADJUST HIGHS AND LOWS ACCORDING TO LANGUAGE AND DIALECT

RULE 6. CONTEXT RULES

ADJUST HIGHS AND LOWS ACCORDING TO CONTEXT

RULE 7. CONCATENATION

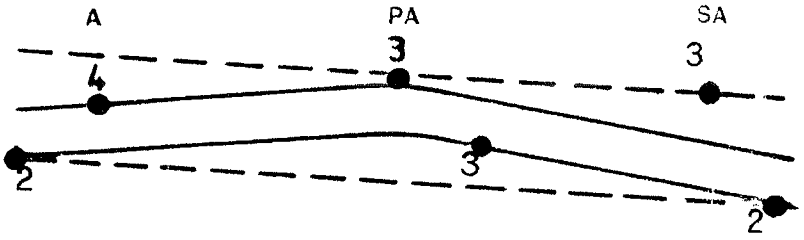
CONNECT NEIGHBORING GENERATED HIGHS AND LOWS

Figure 4. Algorithm for generating pitch contours.

Table 1. Rules of pitch algorithm generating declarative sentences consisting of two major phrases in narrative style.

Rule		Swedish (South)	Greek (Athens)	French (Standard)
1 SI PI	Sentence and phrase intonation	Rising-falling grid	do	do
2 SB PB	Sentence and phrase boundary	SB: Low....Low PB: Low....Low	SB: Low....Low PB: Low....High	SB: Low.... PB: Low....
3 SA PA	Sentence and phrase accent	SA: High-Low PA: High-Low	SA: High-Low PA: High	SA: Low PA: High
4 WA	Word accent	High	High	Void
5 CA	Contrastive accent	High-Low with large range	High-Low with large range	Low-High for first syllable of word
6	Context modifications A. Preparatory rule B. Assimilation rule etc.	A. Preparatory rule B. Assimilation rule	do do	do do
7	Concatenation	Connect neighboring points by a smooth line over the voiced segments	do	do

Figure 5. Typical grid with insertion of *Highs* and *Lows*



1. Grid: Sentence intonation and major Phrase Boundary
2. Sentence and Phrase Boundaries
3. Sentence and Phrase Accents
4. Word Accents

RULE 4 inserts the word accent Highs and Lows.

RULE 5 inserts contrastive accents.

RULE 6 takes care of context phenomena. Two examples are given: 6A is called *the preparatory rule.* It prepares for a following High or Low by inserting a point of the opposite value. 6B is called *the assimilation rule.* It pushes the inserted pitch value in the direction of a following value when a full movement is not warranted by the context.

RULE 7 connects all the generated Highs and Lows by a smooth line.

APPLICATION TO SWEDISH, GREEK AND FRENCH

Table 1 shows how the rules of the pitch algorithm are applied to Swedish, Greek and French. These rules are here expressed in general terms. Specifications are given elsewhere (e.g. Bruce 1977 and Bruce & Gårding 1978 for Swedish, and Touati, forthcoming, for French). The table has some redundancies. Rules 2 and 3 overlap when a phrase or sentence accent is final. In the case demonstrated here there is no conflict. The rules give the same result. In other cases, however, there is a need for priority rules. For instance, when in Swedish or Greek a phrase or sentence accent is initial and there is a clash between the word accent High and the phrase boundary Low, the accent High takes priority. Note that a final word accent in a phrase or a sentence is analysed as a phrase or sentence accent respectively.

Some of the rules in the algorithm are common to all three languages, others differ. A detailed comparison will be given in the next section. Here we shall just mention the principal features.

Phrase endings are Low for South Swedish but High for Greek. For French the phrase ending is also High, due to the phrase accent. The word accent rule is void for French. Of the two Swedish word accents only Accent 1 is used in the demonstration phrase. (It differs from Accent 2 in that its High occurs earlier. The manifestations of the two accents are dialect dependent. For details see Bruce & Gårding 1978).

The contrastive accent is a pitch expanded word accent in the South Swedish dialect and in Greek. In French contrastiveness is expressed by a Low-High in the first syllable instead of a Low in the contrasted word(s).

Figure 6 illustrates how the points given by the various rules are inserted into the grid which has been generated by Rule 1. The numbers 2, 3 etc. refer to the corresponding rules of the algorithm. Note that the lengths of the grids have been adjusted to the observed patterns. Table 1 and Figure 6 will serve as a basis for our comparison.

COMPARISON

Since our input phrases have been constructed in such a way that sentence accents and phrase accents occur in the same position, the comparison concerns mainly the manifestation of these accents and the intonation contour. The comparison is based on observations in a larger material than the one presented here, particularly for Swedish and French. The French material will be described in more detail in a forthcoming dissertation by Paul Touati. We shall follow the order given by the pitch generating rules.

Rules 1 and 2: Intonation and boundaries

The grids that show the general outline of the declarative sentence intonation are remarkably similar apart from differences

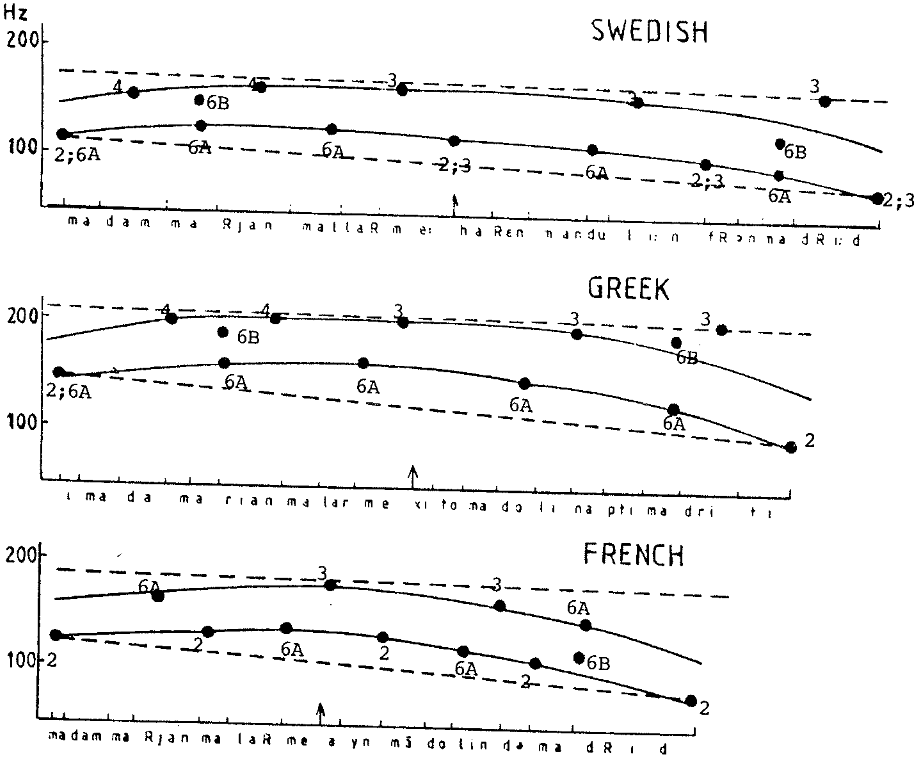


Figure 6. Generated Highs and Lows inserted on tonal statement grids for Swedish, Greek and French.

in overall length. They consist of a slightly rising part covering the subject and a falling part covering the predicate. The pivot marks the boundary between these two constituents. This boundary occurs on different lines, however. For Greek and French it is on the uppermost line, for Skåne Swedish it is on one of the lowest. It should be noted that, although the experimenters made precautions to keep the pragmatic contect fixed, there is still room for variation, particularly in the treatment of phrase boundaries. A slightly 'stronger' boundary in Swedish will result in a continuation of the fall to the lowest level, a similar situation in Greek will create a fall after the rise, whereas in French the duration of the rise and hence the whole syllable on which it occurs will be lengthened. It is clear that the function of the lines of the grid differs in the three languages. In Swedish the lowest lines mark the beginning as well as the end of phrases and sentences. In Greek the low line is reserved for the end of sentences and the high is used for the end of phrases. In French the low line is used for the end of a sentence and for the beginning of a phrase or sentence. The end of a phrase uses one of the upper lines with the highest level marking the boundary between the deepest constituents. The two accent languages, Swedish and Greek, use the highest line also in combination with phrase and sentence accents.

Rules 3 and 4: Accentuation

The manifestation of sentence accent is a steep fall in Swedish and Greek as compared to a very insignificant fall in French. We have regarded this as a support for our phonological analysis of a High-Low sentence accent for Swedish and Greek and a Low sentence accent for French.

The phrase accent is a fall of similar steepness in Swedish but a rise in Greek, corresponding to the phonological representations High-Low and High respectively. For French we have analysed the final accent of a phrase as a phrase accent High to make it parallel with the sentence accent which also has one point, Low. Since all accents in French are phrase or sentence final it would also be possible to analyse them as pitch boundaries. For the Swedish and Greek dialects under investigation, High seems to be the obvious choice for the word accents. The early position of the pitch peak typical of Accent 1 in South Swedish makes this accent mainly falling, whereas in Greek (as in Accent 2 of South Swedish) the peak has a late position in the syllable which gives it a mainly rising pitch.

Rule 5: Contrastive accent

Figure 3 shows some interesting phenomena not present in the demonstration sentence. With contrastive (focal, insisting) accent in Greek and French the pitch goes down to low level and remains there until the end of the phrase. In Swedish, on the other hand, the accents still have their pitch obtrusions. This may be due to the different status of the word accents in the phonological system of Swedish.

In Finland-Swedish, at least in the dialects without lexical accent distinction, the accents behave like in Greek or French (Kerstin Tevajärvi, forthcoming). In French we notice that under contrastive accent the Low of the initial boundary has been replaced by a rise (Figure 3).

Rule 6: Context rules

The context rules reflect perceptual and articulatory constraints. As could be expected they show great similarities across the languages. For the demonstration phrase we shall only need two. (For a more careful analysis of tonal context phenomena in Swedish see Bruce 1977 and for tone languages see Schuh 1978).

Rule 6A: The preparatory rule

This rule says that the generated Highs and Lows are preceded by inserted points of the opposite kind. Our data suggest that the purpose is to give prominence to the following High or Low. The Greek curve in Figure 1 can be used as a demonstration. In the subject phrase (*Madame Marianne Mallarmé*) there are several Lows preceding accent Highs. Similar phenomena can be seen in the Swedish sentence in the same figure.

In French the first Low of *Mallarmé* is prepared for by a High which, according to our interpretation, is introduced to emphasize the Low. If it had not been there, the initial boundary mark of the phrase would not have been present.

Rule 6B: The assimilation rule

The assimilation rule in the present material expresses undershoot phenomena. Before a steep SA fall the preceding Low is undershot in Swedish and Greek. This is so far unexplained. One possible reason is that the steep fall is a strong manifestation of the accent and therefore does not need a preceding full preparation.

The previously mentioned Greek curve in Figure 1 gives other examples of undershoot. In the subject phrase (*Madame Marianne Mallarmé*) there are three Lows preceding three Highs. The assimilation rule has raised the middle High. The reason may be a lack of time for the full preparation before a weakened accent.

This is in agreement with a universal tendency of weakening the second in a group of three 'equal' accents. A similar tendency can be seen in the Swedish sentence in the same figure.

These interpretations are built on observations in acoustic records of real speech that has been varied in a systematic way. We shall use synthesis to check our findings.

Rule 7: Concatenation

This is a mechanical rule and the same for all three languages. In Greek it accounts for tonal enclitization phenomena, i.e. the recognized fact that a group of unaccented syllables form a prosodic unit with the preceding accented one. In the observed sentence (Figure 1) concatenation prescribes the slow fall to the Low of the following accent, prescribed by the preparatory rule.

Table 2 shows the pitch representations proposed for the linguistic variables. It gives an overview of the tonal part of the prosodic structure of the three languages. At the same time it summarizes our comparison: Phrase and sentence intonation in declarative sentences are similar, accentuation and boundaries differ.

The analysis displayed in Table 2 is consistent with impressionistic statements about the languages. The South Swedish dialect

Table 2

PITCH REPRESENTATIONS FOR DECLARATIVE SENTENCES

	SB	PB	SA	PA	WA	CA
SWEDISH SOUTH	L...L	L...L	↑ HL ↓	HL	1. H(L) → 2. H(L)	↑ HL ↓
GREEK	L...L	L...H	↑ HL ↓	(L)H	(L)H	↑ (L)H ↓
FRENCH	L...	L...	(H)L	(L)H	-	(L)H

LETTERS IN PARENTHESES ARE CONSEQUENCES OF THE PREPARATION RULE

VERTICAL ARROWS INDICATE WIDENED RANGE

HORIZONTAL ARROW INDICATE DELAYED TIMING is often described as slow, monotonously chopped up,and down to earth. The falling phrase accents may be responsible for these impressions. The table also shows that Greek uses rising pitch much more than the other languages. This together with its preference for open syllables may explain why Greek has a hammering, insisting effect on a Swedish listener. The most conspicuous pitch characteristic of French to a Swede is the recurring phrase final rising pitch.

The importance of the analysis summarized in Table 2 goes beyond the impressionistic remarks just given. It makes it possible to compare the prosodic inventory and structure of various languages. In this way, it will also be an eff~~e~~icient aid in the analysis of foreign accent and the teaching of pronunciation.

REFERENCES

Botinis, A. (1982) Stress in modern Greek. Working Papers, Department of Linguistics, Lund University 22: (this issue)

Bruce, G. (1977) Swedish Word Accents in Sentence Perspective. Travaux de l'Institut de Linguistique de Lund XII. Lund: Gleerup

Bruce, G. (1982) Developing the Swedish intonation model. Working Papers, Department of Linguistics, Lund University 22: (this issue)

Bruce, G. & Gårding, E. (1978) A prosodic typology for Swedish dialects. In Gårding, E., Bruce, G. & Bannert, R. (Eds.) Nordic Prosody. Travaux de l'Institut de Linguistique de Lund XIII, pp. 219-228. Lund: Gleerup

Gårding, E. (1977) The importance of turning-points for the pitch patterns of Swedish accents. In Hyman, L. (Ed.) Studies in Stress and Accent, pp. 27-35. Los Angeles: University of Southern California

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

Gårding, E. (1981) Contrastive Prosody: A model and its application. In Studia Linguistica, Vol. 35, No. 1-2

Gårding, E. & Bruce, G. (1981) A presentation of the Lund model for Swedish intonation. In Fretheim, T. (Ed.) Nordic Prosody II, pp.

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

Schuh, R. (1978) Tonal rules. In Fromkin, V. (Ed.) Tone. A Linguistic Survey, pp. 221-256. New York: Academic Press

Tevajärvi, K. (forthcoming) Intonation in Finland-Swedish: Word and sentence stress in the Helsinki dialect

Touati, P. (1980) Etude comparative des variations de la frequence fondamentale en Suédois et en Français. Working Papers, Department of Linguistics, Lund University 19: 60-64

Touati, P. (forthcoming) Etudes contrastives de la prosodie du Français et du Suédois

Vaissière, J. (1980) The search for language-dependent features (preprint)

IDENTIFICATION OF NASALS: AN ACOUSTICAL AND PERCEPTUAL ANALYSIS OF NASAL CONSONANTS[1)]

David House

INTRODUCTION

In general terms, the spectral cues governing perception and identification of nasal consonants are described as dichotomous. Information concerning the presence of a nasal[2)] is thought to reside in the nasal segment proper (the nasal murmur) while the formant transitions of the adjoining vowel segments are usually regarded as comprising the major spectral cues necessary for specific nasal phoneme identification. In an attempt to investigate the extent to which vowel formant transition information alone is sufficient for nasal phoneme identification in a VC-syllable in the context of the vowels /i/, /a/, and /u/, three experiments were carried out.

The first experiment took the form of a speech perception identification test using OVE III speech synthesis based on the Haskins tests of the 1950's where vowel formant transitions were progressively altered while keeping the nasal segment constant. The failure of this test to elicit labial responses after the /u/ vowel and a general non-conformity of /u/ vowel context results led to a spectrographic analysis of nasals following /u/ in natural speech. The appearance of a characteristic spectral peak at about 1100 Hz in the /m/ segment led to the third experiment which once again took the form of a synthetic speech identification test. This time, however, natural speech syllables were replicated as closely as possible using an OVE III synthesizer in an attempt to elicit labial responses after /u/ on the basis of the /m/ spectral peak keeping other information at a minimum. A parallel identification test using natural recorded speech syllables was also carried out in order to determine actual nasal phoneme identification difficulty in the /u/ vowel context in isolated syllables.

PROCEDURE

The first experiment comprised 60 stimuli in which a single nasal formant at 200 Hz followed Swedish vowel formants for /i/, /a/, and /u/. F_2 transitions in the vowel segments were altered in steps of 500 Hz ranging from a final value of 500 Hz up to 2500 Hz. The test was presented for four Swedish listeners, the task being to identify each stimulus as one of the following nine syllables: /im, in, iŋ, am, an, aŋ, um, un, uŋ/.

In the second experiment, five sets of the nine syllables were read in randomized sequences by five subjects. All but one of the subjects were native speakers of Swedish, the investigator representing American English. Spectrograms of the middle three sets of the recorded syllables were analysed regarding F_2 transitions and possible distinguishing spectral characteristics in the nasal segments. Particular attention was paid to the segments containing the vowel /u/. Sections from the middle of the nasal segments were also analysed.

The third experiment, designed to investigate perception of syllables in which the nasal follows the /u/ vowel, comprised listener identification tests using both natural and synthetic speech stimuli. In the natural speech portion of the test, the syllables /um/, /un/, and /uŋ/ were randomized and read by one of the Swedish speakers. Each syllable appeared 20 times. In the speech synthesis portion, an attempt was made to replicate the three syllables of natural speech using an OVE III synthesizer in such a way as to enable the testing of the following criteria of identification: a) an N2 at 1100 Hz in the nasal segment of /um/ with no formant transitions, b) an F_2 transition to 1800 Hz in /un/ with no N2 in the nasal segment, and c) the absence of both formant transitions and N2 in /uŋ/ (Fig. 1). The second nasal formant of /m/ was produced by extending the formants of the vowel branch from the vowel segment into the nasal segment, thereby using F2 to emphasize the nasal segment spectrum at 1100 Hz. This method was not completely satisfactory, however, as a substantial (-18 dB re: the vowel) amplitude decrease of the vowel formants in the nasal segment was necessary to maintain satisfactory nasal acceptability due to the effect of using vowel formants in a nasal segment, despite increased bandwidths. Nonetheless, given the single nasal formant nature of OVE III synthesis, the stimulus was considered to be the best possible means of testing the identification criteria given the equipment available.

The three synthetic speech stimuli were randomized and recorded, each stimulus appearing 20 times. Finally, the same stimuli were presented in the form of a pair test where the /uŋ/ stimulus was matched against /um/ stimuli having varying vowel branch amplitudes in the nasal segment in order to test the possibility of an optimum F2 (N2) - N1 amplitude relationship. The test was presented twice, the first time with instructions to identify the labial member of the pair /um/ and the second time with instructions to identify the velar /uŋ/. 13 listeners participated in the third experiment.

RESULTS

For vowels /i/ and /a/ in the first experiment, results concerning F_2 transition patterns and nasal phoneme identification conform very well with the Haskins results. Low F_2 transitions elicited labial responses, transitions ending at 1500 Hz elicited dental responses, and higher F_2 transitions elicited velar responses. When the vowel /u/ preceded the nasal segment, however, the results changed drastically (Fig. 2). Low F_2 transitions elicited velar responses and high transitions dental responses while there was only a single labial preference.

In the second experiment, spectrogram analysis of F_2 transitions in the vowels /i/ and /a/ preceding the three nasal phonemes conforms well with the results obtained in the first experiment. Further, two of the five speakers, representing English and Swedish, produced /m/ segments having an extremely well defined second formant at 1100 Hz which was lacking in the other nasal segments. Although this /m/ resonance at 1100 Hz was not as

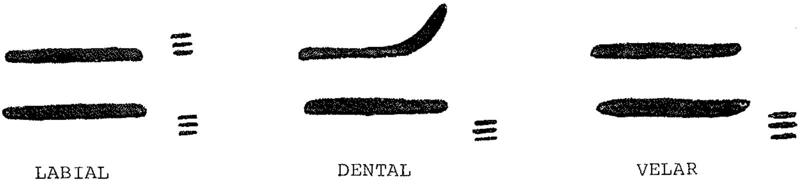


Figure 1. Schematic synthetic stimuli used for synthesis.

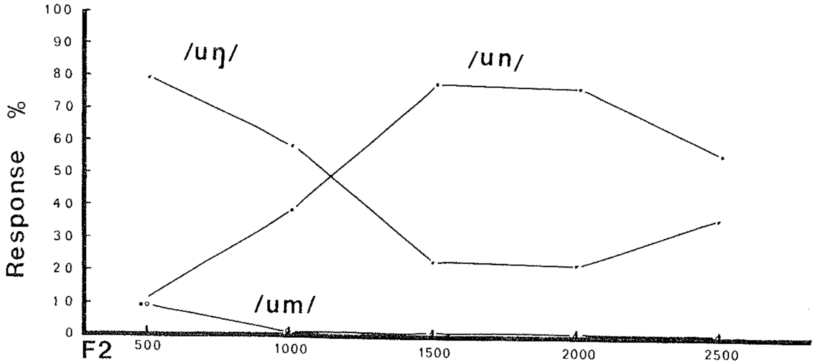


Figure 2. Frequency (Hz) of F2, final value along the abscissa, percent response for each stimulus along the ordinate.

sharply defined in the spectrograms of the other subjects, a spectral peak at this frequency could easily be discerned. Spectrograms of the /u/ vowel syllables differed from the other vowels in that no F_2 transitions were present in the labial or velar context (Fig. 3).

Identification of the natural speech syllables /um/, /un/, and /uŋ/ in the third experiment proved to be a very easy task. Correct identification of the three syllables was nearly 100%, slight confusion (less than 10%) occuring, as could be expected, between /um/ and /uŋ/ (Fig. 4). Identification of the synthetic stimuli where dental and velar responses were expected (Fig. 5) corresponded acceptably with the natural speech test (greater than 75%). Response where labials were expected did not correspond well with the natural speech results. Responses for this stimulus were distributed fairly evenly among the three possible phonemes in the identification test. Labial identification was obtained, however, in more than one-third of the responses (34.3%), a result that compares favorably with that of the first experiment (Fig. 2) where practically no labials were elicited. Results of the pair test where the /uŋ/ stimulus was paired against different versions of /um/ did not differ substantially from the results of the identification test. While a clear preference for the /uŋ/ stimulus was obtained when velar identification was requested (over 80%), no clear preference for any of the /um/ stimuli was obtained when labial identification was requested. Thus there was no optimum stimulus for eliciting a labial response.

DISCUSSION

Where the identification of /un/ and /uŋ/ are concerned, the results of this investigation demonstrate that a perceptually significant spectral cue for the identification of dental nasals following the vowel /u/ can be said to lie in the second formant transition of the vowel, while a spectral cue for identification of velars in the same context lies in a more general downward shift of energy from the vowel segment to the nasal segment.

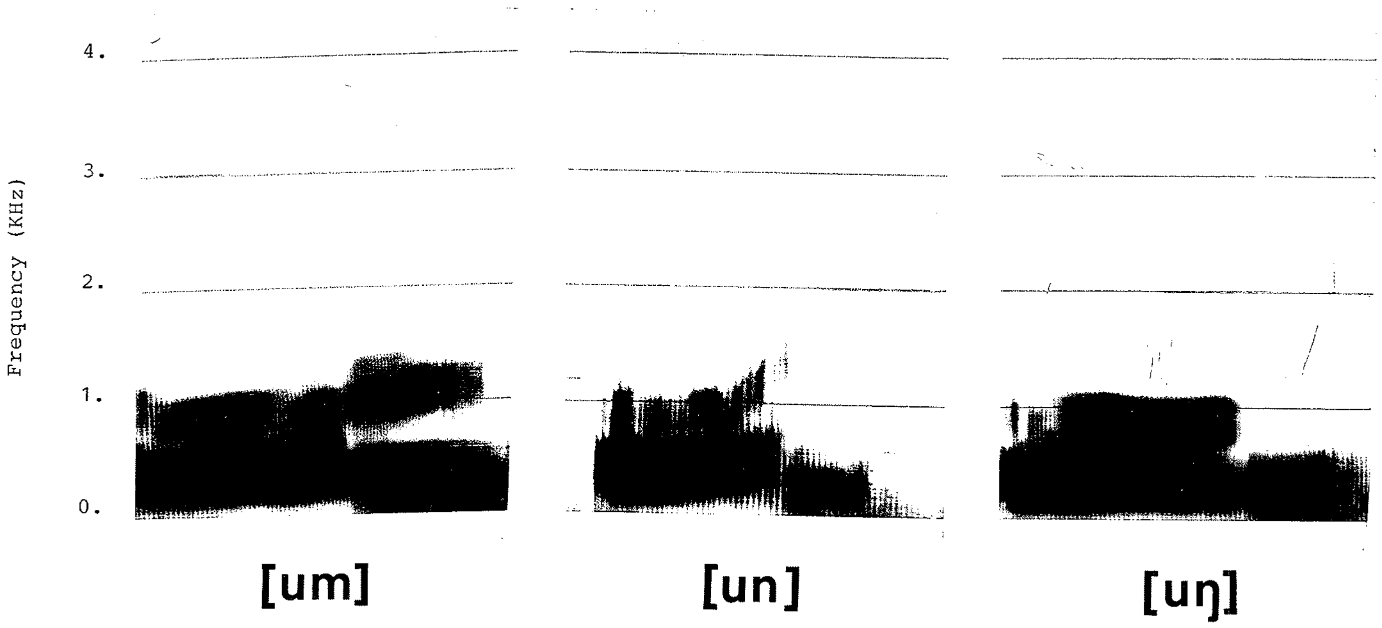


Figure 3. Representative spectrograms, Swedish Subject.

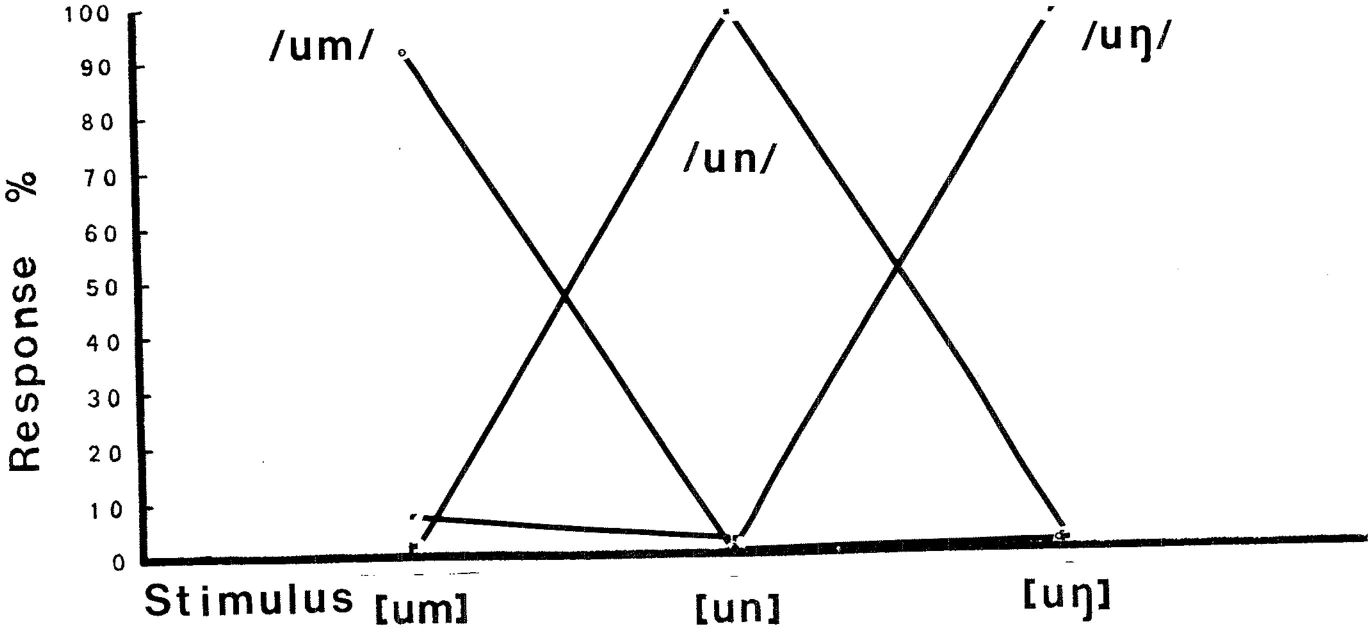


Figure 4. NATURAL SPEECH. Stimulus syllables along the abscissa, percent response for each stimulus along the ordinate.

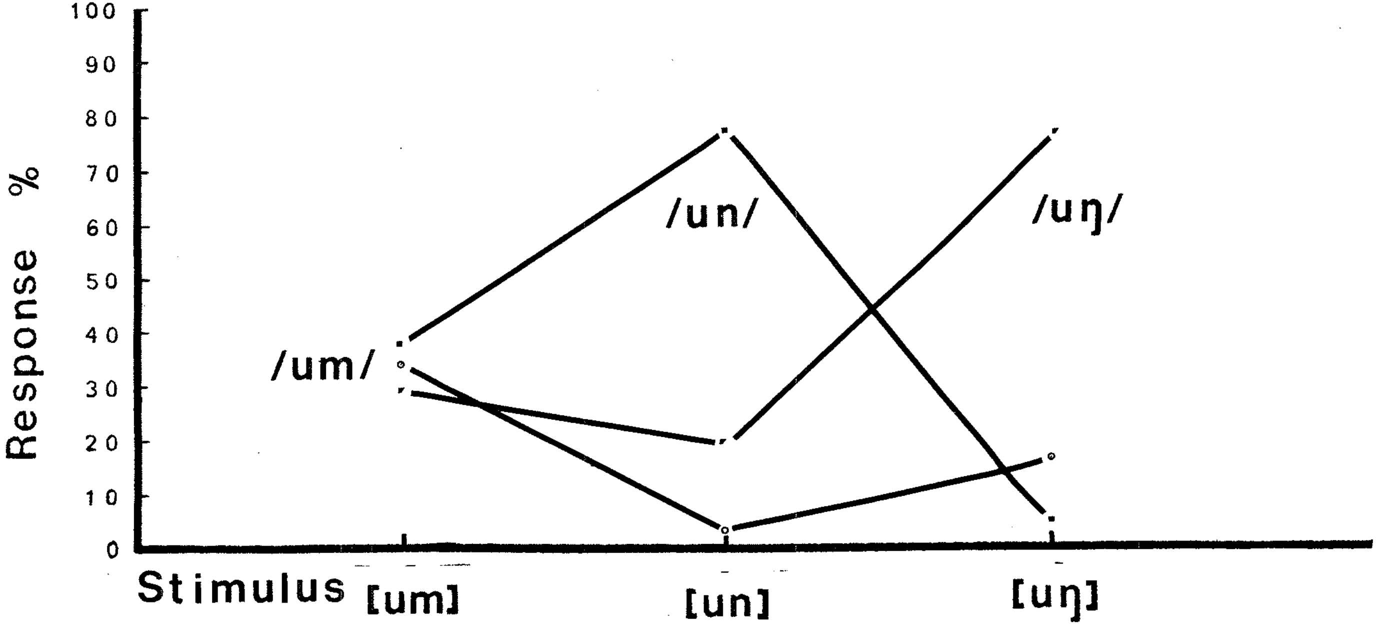


Figure 5. SYNTHETIC SPEECH. Stimuli along the abscissa, percent response for each stimulus along the ordinate.

Regarding identification of labials, results of this test (although possibly attributable to guessing and far from conclusive) do indicate a possibility that identification could be based on a spectral cue (perhaps a pole-zero pair emphasizing the envelope at 1100 Hz) located within the nasal segment. The OVE III synthesis available for this investigation did not enable this point to be sufficiently tested. A number of labial responses were obtained when expected, however, and very few where not expected showing that it is possible to elicit labials by virtue of a more complex nasal segment spectrum. As labial responses were not obtained in the first experiment, the addition of the second /m/ resonance can at least be seen as a step in the right direction.

This investigation further demonstrates the complicated nature of the perceptual mechanism and once again questions the concept of perceptually invariant spectral cues. Evidence points toward a double mechanism of spectral cues having no dependence upon phonemic units. Thus, the same phoneme /ŋ/ in different contexts relies upon different cue mechanisms. Following /i/, the formant transitions can provide the information. Following /u/, a general downward shift of energy from the vowel to the nasal segment can provide the information. This also indicates a possible difference in the point in time at which the perceptual mechanism determines phonemic identity.

SUMMARY

The purpose of this investigation was to determine the extent to which nasal phonemic identification is governed by formant transitions in the vowel, following the transition patterns for place of articulation as described in the Haskins perception experiments. While these patterns seem to hold true for nasal consonants following the vowels /i/ and /a/, a different type of cue is needed for non-dental nasals following /u/. The identification of /uŋ/ can be made on the basis of a downward energy shift from the vowel segment into the nasal segment. It is hypothesized that identification of /um/ is related to a pole-zero pair creating a spectral

peak at about 1100 Hz.

ACKNOWLEDGEMENTS

I am greatly endebted to Sidney Wood for his extensive help with the speech synthesis used in this investigation. Thanks also to Anne-Christine Bredvad-Jensen, Gösta Bruce, and Lennart Nord for their interest and encouragement.

REFERENCES

Fant, G. (1960) Acoustic Theory of Speech Production. The Hague: Mouton & Co.

Fujimura, O. (1962) Analysis of Nasal Consonants. Journal of the Acoustical Society of America 34: 1865-1875

House, A.S. (1975) Analog studies of nasal consonants. Journal of Speech and Hearing Disorders 22: 190-204

Liberman, A.M. et al. (1952) Perception of Synthetic Speech Sounds. Journal of the Acoustical Society of America 24: 597-606

Malécot, A. (1956) Acoustic Cues for Nasal Consonants. Language 32: 274-284

Nord, L. (1976) Perceptual Experiments with Nasals. Quarterly Progress and Status Report of the Speech Transmission Laboratory, Royal Institute of Technology, Stockholm 2/3: 5-8

Nord, L. (1976) Experiments with Nasal Synthesis. Quarterly Progress and Status Report of the Speech Transmission Laboratory, Royal Institute of Technology, Stockholm 2/3: 14-19

FOOTNOTES

1) This is a condensed version of a longer paper, copies of which are available from the author.

2) The term nasal will be used here to pertain exclusively to nasal consonants.

INITIAL VALIDATION OF AN INDIRECT MEASURE OF SUBGLOTTAL PRESSURE DURING VOWELS+

Anders Löfqvist, Björn Carlborg*, and Peter Kitzing**

ABSTRACT

Some methods for direct measurement of subglottal pressure during speech are invasive and thus cannot be used on a routine basis. The development of noninvasive techniques is thus desirable and a simple indirect method for measuring subglottal pressure from records of oral pressure during consonants has recently been proposed and applied to studies of glottal resistance during phonation. In order to be useful, indirect measurement procedures should be validated by comparisons with direct measurements, and the present experiment was designed for such a comparison. Miniature pressure transducers were used to obtain records of pressure below and above the glottis. Results showed nonsignificant differences and a high correlation between the direct and indirect measurements. This indirect method for measuring subglottal pressure thus appears to provide valid results.

INTRODUCTION

Measurements of subglottal pressure during speech may be obtained by four different methods: 1) A needle-tap is passed through the pretracheal wall and connected to a pressure transducer; 2) A small transducer is passed through the glottis and placed directly in the trachea (in a variant of this approach, a small catheter is passed through the glottis with the open end of the catheter in the trachea and the other end of the catheter coupled to a transducer outsied the subject);

+ To appear in Journal of the Acoustical Society of America

* Department of Otolaryngology and Experimental Research, Malmö General Hospital, Malmö

** Department of Phoniatrics, Malmö General Hospital, Malmö

3) The subject swallows a small inflatable balloon into the oesophagus with the balloon connected by a catheter to a transducer; 4) The subject is placed in a body plethysmograph and alveolar pressure is determined indirectly.

The first three of these procedures are invasive, cause a certain degree of discomfort for the subject, and thus cannot be used on a routine basis. The use of an oesophageal balloon requires, in addition, simultaneous recordings of lung volume, since the balloon itself will record intrathoracic pressure (Bouhuys, Proctor and Mead, 1966). If a body plethysmograph is used, the subject has to participate in rather elaborate calibration procedures (Hixon, 1972). Due to the practical problems associated with these methods for making measurements of subglottal pressure, a simple noninvasive technique has recently been proposed as a routine clinical procedure (Holmberg, 1980; Smitheran and Hixon, 1981).

This method makes use of the fact that pressure below and above the glottis is equalized during the closure period of voiceless stops (e.g., Shipp, 1973). During this period, the glottis is open and the oral cavity is closed. Measurements of oral pressure during this phrase of voiceless stop production thus provide measures of subglottal pressure as well. By constructing a suitable speech material, where voiceless stop consonants and vowels alternate in a regular manner, successive measurements of oral pressure can be made during the stop closures. A linear interpolation between the oral pressure registrations can then be made in order to get indirect measurements of subglottal pressure during the vowels occurring between the voiceless consonants.

Applications of this procedure to studies of glottal resistance during phonation under both normal and clinical conditions have recently been described by Holmberg (1980) and by Smitheran and Hixon (1981).

Glottal resistance is defined here as the ratio between the pressure drop across the glottis and the transglottal airflow. Its calculation thus requires measurements of subglottal pressure, and the simplified procedure outlined above has been used for that purpose. The reliability and usefulness of such studies thus depend on the validity of the indirect technique for estimating subglottal pressure. The measurement of transglottal airflow is not a major problem in this case.

In support of the validity of the method, two types of argument have been used. Smitheran and Hixon (1981) argue that values of glottal resistance during phonation obtained with the indirect technique are more or less identical to values of the same parameter obtained with direct measurements of subglottal pressure. Holmberg (1980) shows that calculated values of glottal resistance will discriminate between different phonation types where differences in glottal resistance can be assumed to occur on a priori grounds.

The technique appears to be a promising one, and arguments such as these are obviously useful in demonstrating its validity. A more direct evaluation of the procedure would, however, compare simultaneous direct and indirect measurements of subglottal pressure. The present experiment was designed for such a comparison.

METHOD

Subglottal pressure was recorded using a technique described in detail elsewhere (Kitzing and Löfqvist, 1975; Kitzing, Carlborg and Löfqvist, in press). Briefly, a small miniature transducer (Millar Instruments PC 350 5F) with adequate high frequency response was introduced through the nose and passed through the glottis under topical anesthesia of the pharyngeal and laryngeal mucosa.

The cable connecting the transducer to power supply and amplifier was positioned in the posterior commissure with the transducer itself about 2 cm below the glottis. A second transducer of the same type was used for recording oral pressure. This transucer was also introduced through the nose and placed in the pharynx about 4 cm above the glottis. The transducer output signals were recorded on FM tape together with a microphone signal for later processing.

The experimental procedure required anesthetization of the laryngeal mucosa and the introduction of catheters through the velopharyngeal port. Auditory and acoustic analysis of the voice during the experiment did not reveal any excessive breathiness or hoarseness. A complete velopharyngeal closure could also be made with the catheters in place. This was verified by recording egressive airflow during preliminary trials.

The speech material consisted of CVCVCV syllables. In order to make the material as similar as possible to that used in studies applying the indirect measurement procedure, the labial voiceless stop /p/ and the low back vowel /ɑ/ were used. A male speaker with a normal larynx and without any known history of voice disorders, one of the authors, served as subject. Twenty repetitions of the speech material were obtained.

For processing, the signals were played back on an ink-writer. During play-back, the pressure signals were low-pass filtered at 70 Hz to remove most of the rapid pressure fluctuations due to voicing and thus facilitate measurements.

The method of linear interpolation between the oral pressure records is shown in Fig. 1. The interpolation was made from the point at which oral pressure started to fall rapidly at the stop release, to the point at which a stable elevated pressure occurred shortly after stop implosion. This is illustrated by the broken line in Fig. 1.

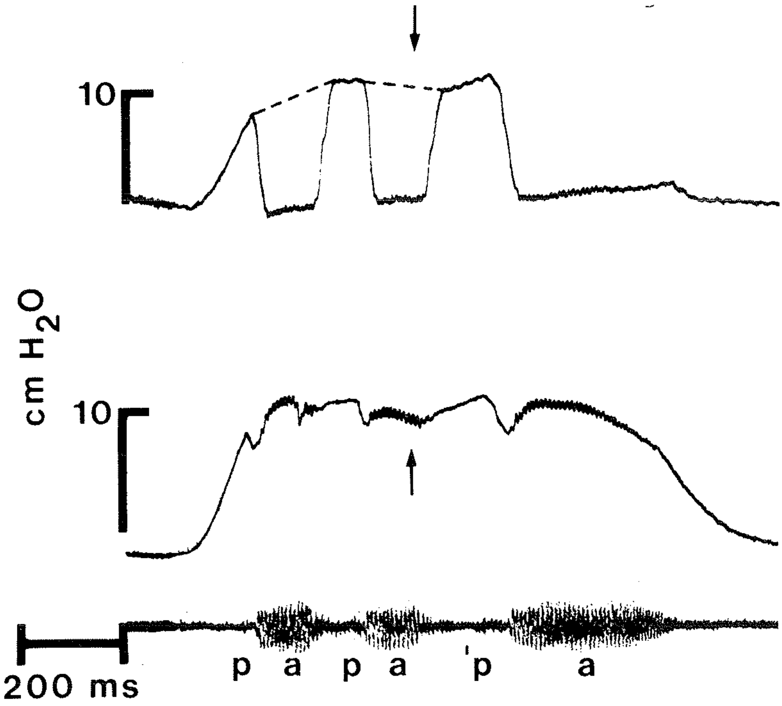


Figure 1. Simultaneous records of oral pressure (top), subglottal pressure (middle) and audio signal (bottom) during production of the nonsense syllables papa'pa. The pressure signals have been low pass filtered at 70 Hz. The broken line in the oral pressure curve illustrates the method of linear interpolation between the pressures associated with the voiceless stops in order to get an estimate of subglottal pressure during the vowels between the voiceless consonants. The arrows indicate the point during the utterance at which the measurement was made of direct and indirect subglottal pressure.

Pressure was measured in the middle of the second vowel of the utterance; this point is shown by the arrows in Fig. 1. For each utterance, two measurements of subglottal pressure were thus obtained. One represented a direct measure, and was made from the output of the transducer in the trachea. The other represented an indirect measure, and was made from the interpolated oral pressure curve.

RESULTS

The mean difference between the two sets of measurements was 0,85 mm of water, with a standard deviation of 3,73. A paired t-test showed this difference to be insignificant (t_{19} = 1.019). The two sets of measurements are plotted against each other in Fig. 2. A Pearson product moment correlation coefficient of 0,92 was obtained for the two sets. As can be seen from Fig. 2, the data points are about evenly distributed around the diagonal line drawn in the figure.

DISCUSSION

The results of the present experiment indicate that indirect measurements of subglottal pressure interpolated between measures of oral pressure during voiceless stops provide a valid estimate of the "true" subglottal pressure during the vowel of the interpolated interval. In order to avoid possible influences from extraneous factors, measurements should be made in the middle of the vowel between the voiceless consonants, since airflow during the period of aspiration after stop release may perturb the subglottal pressure (cf., Löfqvist, 1975). With this caveat, the simplified procedure appears to be valid.

We should add a final note on its clinical application. Here, one may be interested in following changes in glottal resistance as a result of therapy or phonosurgery.

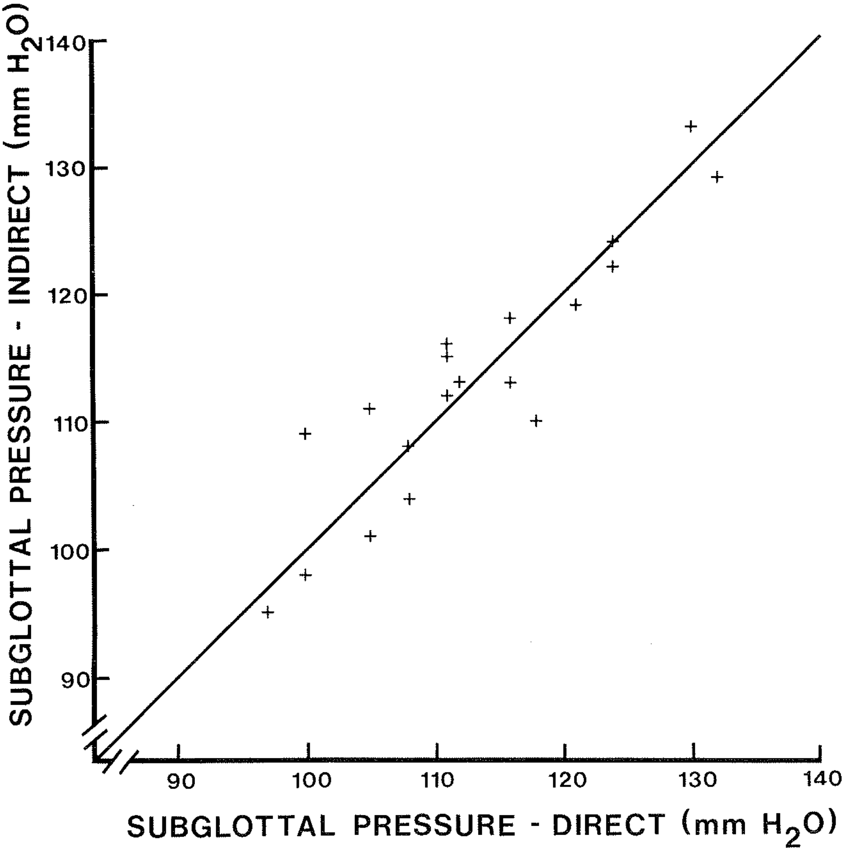


Figure 2. Plot of direct and indirect measurements of subglottal pressure (n = 20).

Comparisons of repeated measurements over a period of time are thus of more interest than the absolute values of glottal resistance on single occasions. In this case, minor deviations of the estimated subglottal pressure from the "actual" subglottal pressure would not seem to be a serious concern, since the error can be assumed to be constant across measurements.

ACKNOWLEDGMENT

This work was supported in part by grants from the Wallenberg Foundation and from the Segerfalk Foundation.

REFERENCES

Bouhuys, A., Proctor, D. and Mead, J. (1966) Kinetic aspects of singing. Journal of Applied Physiology 21: 483-496.

Hixon, T. (1972) Some new techniques for measuring the biomechanical events of speech production: One laboratory´s experience. ASHA Reports No. 7: 68-103.

Holmberg, E. (1980) Laryngeal airway resistance as a function of phonation type. Journal of the Acoustical Society of America 68: S101 (A).

Kitzing, P. and Löfqvist, A. (1975) Subglottal and oral air pressures during phonation. Preliminary investigation using a miniature transducer system. Medical and Biological Engineering 13: 644-648.

Kitzing, P., Carlborg, B. and Löfqvist, A. (in press) Aerodynamic and glottographic studies of the laryngeal vibratory cycle. To appear in Folia Phoniatrica.

Löfqvist, A. (1975) A study of subglottal pressure during the production of Swedish stops. Journal of Phonetics 3: 175-189.

Shipp, T. (1973) Intraoral air pressure and lip occlusion in midvocalic stop consonant production. Journal of Phonetics 1: 167-170.

Smitheran, J.R. and Hixon, T. (1981) A clinical method for estimating laryngeal airway resistance during vowel production. Journal of Speech and Hearing Disorders 46: 138-146.

RESOLUTION OF CONVOLVED SIGNALS IN BIOMEDICAL APPLICATIONS*

Bengt Mandersson

This thesis deals with the design of pulse shaping filters with finite impulse response (FIR) applied to the resolution of convolved signals. The thesis consists of three parts: I. Design of FIT Filters for Deconvolution; II. Least Squares Filters Using Bandpassing Subfilters; III. Digital Filtering of Ultrasonic Echo Signals.

The application of the pulse shaping technique to deconvolution problems yield filters with "inverse" properties relative to the input signals. This means e.g. that the magnitudes of the spectra of the filter impulse responses must be high where the magnitudes of the input signal spectra are low. This filtering technique is consequently sensitive to noise and to variations of the input signal waveforms. When designing such filters, the trade-off between noise sensitivity and pulse shaping performance must be taken into account. It is then of great importance to have methods for determining filters that are robust in the sense that they are as insensitive as possible to noise and to variations of the input waveforms.

Pulse shaping filters applied to deconvolution problems are often determined by minimizing the mean squared difference between the filter response to a specified input and some desired pulse shape (least squares filters). However, the solution can be sensitive to the detailed structure of the desired pulse. In many applications it is not always obvious what desired pulse shape to select in order to obtain the most suitable filter. Part I defines

---

* Summary of a doctoral dissertation in Telecommunication Theory, Department of Electrical Engineering, Lund Institute of Technology

"soft" versions of pulse shaping filters. These are also based on minimizing squared differences between the filter responses and desired pulses but with the modification that regions are inserted where the differences do not affect the solution. Both amplitude and temporal regions are examined.

The number of adjustable parameters in the design of FIR filters is quite high. Iterative solution schemes using a steepest descent algorithm are most easy to implement, since they do not need so much computer storage space and the amount of computation in each iteration is low. But the inverse filter technique leads to ill-conditioned minimization problems. Part II describes how the rate of convergence for the steepest descent algorithm can be increased by configuring the FIR filters as sets of bandpass subfilters in parallel. The computations are done in the frequency domain and an FFT routine is used to determine the DFT of the signals. This not only reduces the amount of computation but also suitably relates the temporal properties of the filter to its frequency characteristics.

In the ultrasonic pulse echo method axial resolution is strongly dependent on the durations of the impulse responses of the ultrasonic transducers. These impulse responses normally have narrow bandpass characteristics. Owing to this it is difficult to derive inverse filters suitable for ultrasonic signals. In Part III, the weighted least squares pulse shaping filter defined in Part I is applied to the ultrasonic signals in order to decrease the duration of the received ultrasonic pulses. The weighted least squares filter is found to be suitable for ultrasonic signals and improved resolution is achieved using this filter. This filter is used both to filter received echoes and to preshape emitted ultrasonic pulses.

INTONATION IN FINLAND-SWEDISH: Word and Sentence Stress in the Helsinki Dialect

Kerstin Tevajärvi

The aim of this paper is to present an acoustic study of the intonation of a less documented variant of Swedish i.e., Finland-Swedish.

Word and sentence stress in Swedish have been thoroughly investigated by Bruce, Gårding and others, but there are only a few reports on the prosody of Finland-Swedish. At the same time, there are many speculations and theories about Finland-Swedish which lack an empirical foundation. The aim of the present experiment is to contribute to such a foundation.

Finland-Swedish is a minority language spoken by about 300.000 people living in the costal areas and archipelago of Finland. Due to geographical and political reasons, the Swedish dialects of Finland differ from those in Sweden in many respects. The prosody is simpler, there is no word accent opposition, except for the dialect of Västra Nyland, and there are also differences on the segmental and syntactic levels. These differences are usually explained by interference from Finnish or by the archaic character of the dialect.

Prosodically the most striking difference is the lack of a word accent opposition and its effect on intonation. To study this problem, I chose a material consisting of short statements where the following parameters were varied; (1) the placement of focus (sentence accent), (2) the placement of stress, and (3) the number of syllables in the words in focus. This material was first

* This is a summary of a longer report. Copies of the full report are available from the author.

used by Gösta Bruce in his thesis *Swedish word accents in sentence perspective* (Bruce 1977) and later in intonation studies of several Swedish dialects. His material was very suitable since it was designed for studies of this problem; it had also been tested on other dialects and therefore my results could easily be compared. Furthermore, the Lund-model for Swedish intonation is based on this material (Bruce & Gårding 1978, Gårding & Bruce 1981).

Three informants, one male and two females with the same social dialect, were recorded in Lund. Measurements were made from mingograms.

Results: In words with short vowels the word accent is signalled by an Fo-rise at the beginning of the accentuated syllable, and an Fo-maximum in the middle of the accentuated vowel or at the VC-boundary. The Fo-maximum is followed by an immediate Fo-fall to the posttonic vowel (Fig. 1).

In Swedish dialects with a word accent opposition, the timing of Fo-changes is decisive for the perceptual distinction of the accents. Due to the lack of an opposition in Finland-Swedish it was hypothesized that there should be a less exact timing of the Fo-movements since a confusion of word accents was not possible. The instability of the accent has also been suggested before by Selenius (Selenius 1978).

Still in the present material, the timing of the Fo-rise and, in particular, the Fo-fall, was very precise both within and across speakers. The accent is stable in timing even though there is no opposition.

The sentence accent is also signalled by an Fo-rise with a maximum value in the accentuated vowel followed by an Fo-fall in the posttonic syllable, and also by a larger Fo-interval. The sentence accent Fo-contour is dominated by the steep fall in the posttonic syllable, the posttonic fall seems to be the most stable indication of focus. The difference between word and sentence accent is mainly a question of domain. The placement of focus affects the utterance as a whole, while the influence of the

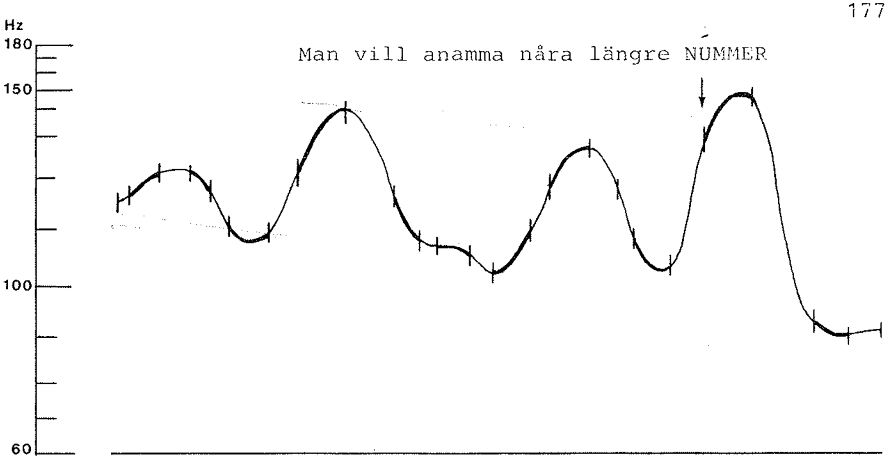


Figure 1. The arrow indicates the boundary between N and U. Focus on NUMMER. The thick lines are vowels, the thin lines consonants.

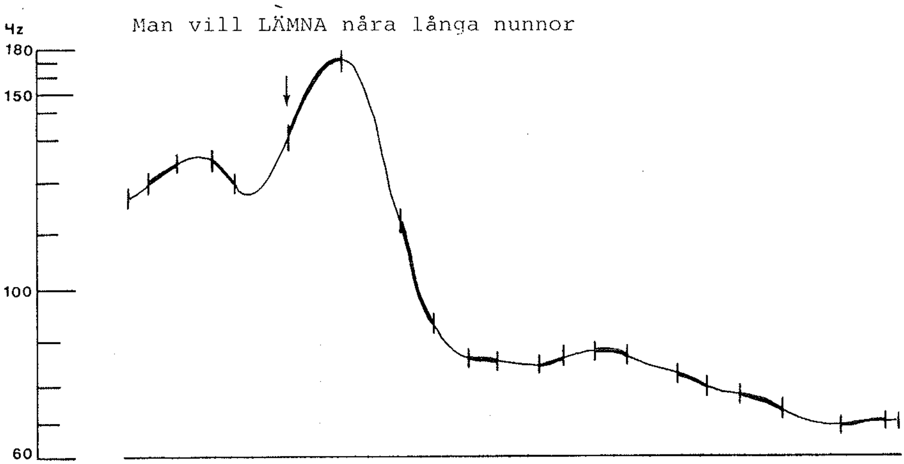


Figure 2. The arrow indicates the boundary between L and Ä. Focus on LÄMNA. No tonal manifestation of post-focal word accents.

word accent is limited to the surrounding syllables (Fig. 1 and Fig. 2).

The placement of focus is important for the planning of the utterance. Word and sentence accent co-occur and up to focus the tonal manifestation of the word accents is preserved. After focus the word accents are very deaccentuated and there are very low peaks or no tonal peaks at all manifesting word stress (Fig. 2). Deaccentuation occurs after focus as in other languages with one-peaked accents (Greek, French), even in the one-peaked Swedish dialects (1A, 1B) where word and sentence accent co-occur. It is never found in two-peaked variants where word and sentence accent are manifested by separated tonal gestures. Deaccentuation occurs for all informants regardless of Fo-range. Deaccentuation before focus is rare but can be seen in very monotonous informants.

By testing the Lund-model for Swedish intonation on Finland-Swedish a comparison can be made with other variants of Swedish. The model is a schematic description of the prosody in short Swedish statements, also accounting for dialectal variation (Gårding & Bruce 1981).

The model is based on Meyer's observations of word accents in different Swedish dialects. Gårding and Lindblad (Gårding and Lindblad 1973) made a further categorisation to four main types due to tonal contour, its timing, and geographical location. These four types are called South 1A (Malmö), Central 1B (Dalarna), East 2A (Stockholm), and West 2B (Gothenburg). Finland-Swedish is called Far East 0 (Helsinki) because of lack of word accent opposition.

The model is input-output oriented, where the input is information about word prosody, sentence prosody and the actual dialect. The output consists of sentences having correct intonation.

A comparison of Finland-Swedish with other variants of Swedish shows that the intonation in general resembles South 1A but without accent opposition (Fig. 3). The timing of the Fo-contour

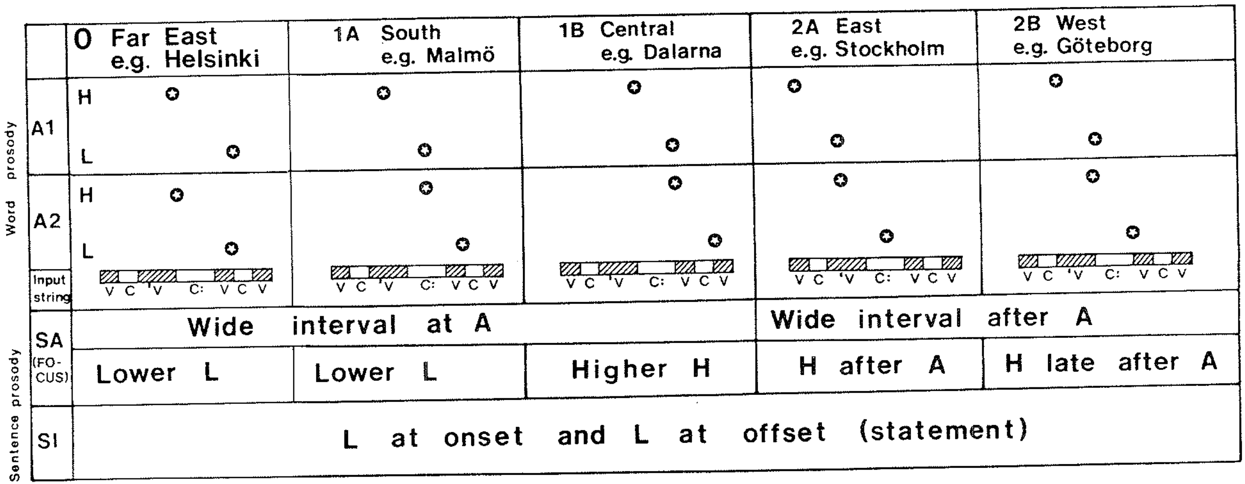


Figure 3. The linguistic part of the model, with a comparison of different Swedish dialects.

resembles the first peak of accent 2 in West 2B. The shape of this accent contour is said to be the consequence of a generalised two-peaked contour and a subsequent loss of the second peak.

REFERENCES

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

Bruce, G. and Gårding, E. (1978) A prosodic typology for Swedish dialects. In Gårding, E. (Ed.) Nordic Prosody. Travaux de l'Institut de Linguistique de Lund XIII

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

Selenius, E. (1978) Studies in the development of the 2-accent system in Finland-Swedish. In Gårding, E. (Ed.) Nordic Prosody. Travaux de l'Institut de Linguistique de Lund XIII

WORKING PAPERS
edited by Kerstin Hadding

RECENT NUMBERS AND CONTENTS

WP 16 1978		STUDIES IN GENERAL LINGUISTICS dedicated to BERTIL MALMBERG by students and colleagues on the occasion of his 65th birthday 22nd April 1978
	MILAN BÍLÝ	Some thoughts about functional sentence perspective, empathy, and reflexives
	EVA GÅRDING KRISTINA LINDELL	Tones in Northern Kammu: a phonetic investigation
	D.J. HACKMAN	The communicative competence of foreigners in Swedish: listener attitude and contextual appropriacy
	KERSTIN HADDING KERSTIN NAUCLÉR	Permissible and not permissible variations in pitch contours
	OLAV HAMMER	Historical Linguistics and generative phonology
	KENNETH HYLTENSTAM	A framework for the study of interlanguage
	EVA LARSSON	Effet communicatif de la dislocation d'un NP en français
	THORE PETTERSSON	Reading Chomsky
	CHRISTER PLATZACK	The Swedish prepositions <u>på</u> and <u>i</u> followed by measure phrases
	SVEN PLATZACK	The sign and its substance. A coloured view
	BARBARA PROHOVNIK	Linguistic aspects of bilingual aphasia
	CHRISTOPHER STROUD	The concept of semilingualism
	SIDNEY WOOD	Tongue retraction is not so useful after all

WP 17 1978		The PROSODY OF NORDIC LANGUAGES Symposium 14-16 June 1978 Abstracts
WP 18 1978	KENNETH HYLTENSTAM	Variation in interlanguage syntax
WP 19 1980		PAPERS GIVEN AT THE SEVENTH SWEDISH-DANISH SYMPOSIUM LUND APRIL 25-26 1980
WP 20 1981	KENNETH HYLTENSTAM	Understanding varieties of the target language
	THORE PETTERSSON	Fria adverbial
	BENGT SIGURD	COMMENTATOR A computer system simulating verbal behaviour
WP 21 1981	KERSTIN HADDING	Phonetics and I 1952-1979
	ROBERT BANNERT & KENNETH HYLTENSTAM	Swedish immigrants' communication: Problems of understanding and being understood
	ANTONIS BOTINIS	The Greek vowels
	ANNE-CHRISTINE BREDVAD-JENSEN	Tonal interaction between attitude and grammar
	GÖSTA BRUCE	Tonal and temporal interplay
	EVA GÅRDING	Optimal teaching of Swedish pronunciation
	EVA GÅRDING & GÖSTA BRUCE	A presentation of the Lund model for Swedish intonation
	EVA GÅRDING, GÖSTA BRUCE & URSULA WILLSTEDT	Transitional forms and their position in a prosodic typology for Swedish dialects
	LEA-LIISA LAHTI	On Finnish plosives
	KERSTIN NAUCLÉR & EVA SÖDERPALM	Slips of the tongue, Slips of the pen - Slips of the brain?
	BARBARA PROHOVNIK	Prosodic features in conduction aphasia

RECENT PUBLICATIONS ELSEWHERE BY STAFF AND RESEARCH STUDENTS

1979

Bílý, M. Anaphoric relations from the viewpoint of the referential and attributive use of language sign. Slavica Lundensia 7: 133-174

1980

Bílý, M. Pronominal vs. zero subjects in Czech and Russian. Slavica Lundensia 8: 29-50

1981

Bílý, M. Intrasentential Pronominalization and Functional Sentence Perspective (in Czech, Russian, and English). Lund Slavonic Monographs 1. (PhD Thesis)

Gårding, E. Contrastive Prosody: A model and its application. Studia Linguistica Vol. 35, no. 1-2: 146-166

Gårding, E. Den efterhängsna brytningen. In Lundaforskare föreläser 13: 5-11. Lunds Universitet: Gleerup

Gårding, E., Botinis, A., & Willstedt, U. Contrastive Phonetic Studies at the Linguistics Department of Lund University. In Hyldgaard-Jensen, K., & Skov-Larsen, J. (Eds.) Contrastive Linguistics in Scandinavia: 179-185. Nordiska Forskarkurser. Köpenhamn

Hyltenstam, K. Invandrarinriktad språkundervisning och interimspråksforskning. In Hyltenstam, K. (Ed.) Språkmöte, pp. 9-20. Lund: Liber Läromedel

Hyltenstam, K. (Ed.) Språkmöte. Svenska som främmande språk, hemspråk, tolkning. Lund: Liber Läromedel

Hyltenstam, K. (Ed.) Främmandespråksundervisning och översättarutbildning. Lund: Liber Läromedel

Hyltenstam, K. Från metoddiskussioner till elevdiagnostik. In Hyltenstam, K. (Ed.) Främmandespråksundervisning och översättarutbildning, pp. 9-18. Lund: Liber Läromedel

Löfqvist, A., Baer, T., & Yoshioka, H. Scaling of glottal opening. Phonetica 38: 265-276

Löfqvist, A., & Yoshioka, H. Interarticulator programming in obstruent production. Phonetica 38: 21-34

Löfqvist, A., & Yoshioka, H. Laryngeal activity in icelandic obstruent production. Nordic Journal of Linguistics 4: 1-18

Mandersson, B. Resolution of Convolved Signals in Biomedical Applications. Telecommunication Theory, Electrical Engineering Department, Institute of Technology, Lund (PhD Thesis)

Mc Garr, N.S., & Löfqvist, A. Obstruent production by hearing-impaired speakers: Interarticulator timing and acoustics. Haskins Laboratories, Status Report on Speech Research, SR-67/68: 83-106

Orrling, A., and Sigurd, B. Högskolan som verbal reform. Forskning och Framsteg 1981:8, 27-31

Pettersson, T. De mindre språken i gymnasiet. In Hyltenstam, K. (Ed.) Främmandespråksundervisning och översättarutbildning, pp. 39-54

Pettersson, T. Existential expressions in Russian and Modern Hebrew. Scando-Slavica 27: 213-224

Pettersson, T. On voice in the Russian verb. In Jacobsen, P. & Krag, H.L. (Eds.) The Slavic Verb. An anthology presented to Hans Christian Sørensen 16th December 1981, pp. 137-146. Copenhagen: Rosenkilde & Bagger

Sigurd, B. Intensivkurser. In Hyltenstam, K. (Ed.) Främmandespråksundervisning och översättarutbildning, pp. 82-94

Stroud, C. Tolkning, tolkutbildningen och lingvistik. In Hyltenstam, K. (Ed.) Språkmöte, pp. 178-215

Touati, P. Språkundervisning i studiecirklar. In Hyltenstam, K. (Ed.) Främmandespråksundervisning och översättarutbildning, pp. 68-81

Yoshioka, H., & Löfqvist, A. Laryngeal involvement in stuttering. A glottographic observation using a reaction time paradigm. Folia Phoniatrica 33: 348-357

Yoshioka, H., Löfqvist, A., & Hirose, H. Laryngeal adjustments in the production of consonant clusters and geminates in American English. Journal of the Acoustical Society of America 70: 1615-1623

1982

Bannert, R. An Fo-dependent model for duration? RUUL 8: 59-80 (Reports from the Institute of Linguistics, Uppsala University)

Bannert, R., Engstrand, O., Eriksson, H., & Nordstrand, L. Phonetics and identity. RUUL 8: 22-58 (Reports from the Institute of Linguistics, Uppsala Universtity)

Gårding, E. Prosodic Expressions and Pragmatic Categories. In Acta Universitatis Lundensis Band 38, Textstrategier i tal och skrift. Almqvist & Wiksell International

LECTURES AND CONGRESS PAPERS BY STAFF AND RESEARCH STUDENTS

1981

The 101st Meeting of the Acoustical Society of America
Ottawa 18-22 May 1981

McGarr, N. S. & Löfqvist, A. Laryngeal-oral coordination in the production of obstruents by hearing-impaired speakers. Journal of the Acoustical Society of America 70: S68 (A)

University of Stockholm
5 May 1981

Bannert, R. Faculty opponent at B. Lyberg's defence of his PhD dissertation: Temporal properties of spoken Swedish

The 6th Scandinavian Conference of Linguistics
Røros 19-21 June 1981

Bílý, M. Is there any "backward pronominalization?" In Fretheim, T. (Ed.) Papers from the Sixth Scandinavian Conference of Linguistics, pp. 26-33. Trondheim: Tapir

Symposium on Interdisciplinary Perspectives of Language Acquisition Research
Passau 14-15 July 1981

Bannert, R. Phonological variation and foreign language teaching

The 6th International Congress of Applied Linguistcs, AILA 81
Lund 9-15 August 1981

Gårding, E. Contrastive prosody: A model and its application. Special Lecture. Studia Linguistica 35: 146-165

Gårding, E., Botinis, A. & Touati, P. A comparative study of Swedish, French and Greek intonation. In Sigurd, B. & Svartvik, J. (Eds.) AILA 81 Proceedings I: Sections and Workshops, pp. 132-133.

Hyltenstam, K. & Magnusson, E. The acquisition of the voice contrast in stops by first and second language learners of Swedish. In Sigurd, B. & Svartvik, J. (Eds.) AILA 81 Proceedings I: Sections and Workshops, pp. 92-93.

Kitzing, K. The good language learner. In Sigurd, B. & Svartvik, J. (Eds.) AILA 81 Proceedings I: Sections and Workshops, pp. 96

Malmberg, B. Applied linguistcs, past, present and future. Keynote address at the opening ceremony. Studia Linguistica 35: 7-14

Sigurd, B. & Svartvik, J. (Eds.) AILA 81 Proceedings I: Sections and Workshops. Lund

Svantesson, J.-O. A dictionary on a desktop computer. In Sigurd, B. & Svartvik, J. (Eds.) AILA 81 Proceedings I: Sections and Workshops, pp. 356-357

Stroud, C. Text strategies in the speech of thought disordered schizphrenias. In Sigurd, B. & Svartvik, J. (Eds.) AILA 81 Proceedings I: Sections and Workshops, pp. 422-423

Från text till prosodi
Lund 18-19 september 1981

Bruce, G. Summary of earlier project meeting

Bannert, R. On temporal and tonal relations

Gårding, E. A model for comparing intonation

Sigurd, B. Prosody in connection with a text production program

The Leeds Experimental Phonetics Symposium
28-30 September 1981

Löfqvist, A. Intrasegmental timing: laryngeal-oral coordination in voiceless consonant production

De Nordiske Datalingvistikkdagene 1981
Dragvoll 22-23 October 1981

Mílý, B. Experience with "Commentator", a computer system simulating verbal behavior. In Lien, E. (Ed.) De Nordiske datalingvistikkdagene 1981. Foredrag fra en konferanse på Dragvoll 22-23 oktober 1981, pp. 39-46

102nd Meeting of the Acoustical Society of America
Miami Beach 30 November-4 December 1981

Baer, T., Löfqvist, A. & McGarr, N. S. Glottal area variations: a comparison between transillumination and high speech filming. Journal of the Acoustical Society of America 70: S12 (A)

Löfqvist, A., Kitzing, P. & Carlborg, B. Transglottal pressure during the laryngeal vibratory cycle. Journal of the Acoustical Society of America 70: S13 (A)

1982

Symposium on Child Prosody at Child Language Research Forum
Stanford 27-28 March 1982

Nettelbladt, U. Phonotactic and prosodic development in Swedish children.

Workshop on Prosody, Association Européenne de Psycholinguistique
Paris, Maison des sciences de l'homme 21-23 April 1982

Gårding, E. A comparative study of intonation

103rd Meeting of the Acoustical Society of America
Chicago 26-30 April 1982

Mc Garr, N.S., Baer, T., Löfqvist, A., Whitehead, R. & Metz, D. High-speed filming and acoustic analysis of deaf glottal pulses. Journal of the Acoustical Society of America 71: S55 (A)

COOLING 82. 9th International Conference on Computational Linguistics
Prague Czeckoslovakia 5-10 July 1982

Bílý, M. & Sigurd, B. Developing Commentator, a computer system simulating verbal production

XIIIth International Congress of Linguistics
Tokyo 29 August-4 September 1982

Bílý, M. Some remarks on the government-binding theory and anaphora

Bruce, G. Experiments with the Swedish intonation model

Gårding, E. A model for comparison of intonation

The 15th Annual Meeting of the Societas Linguistica Europae
Athens 8-11 September 1982

Bílý, M. & Pettersson, T. A typology of case

CURRENT RESEARCH PROJECTS

The immigrant and communication technology - understanding and being understood (The Bank of Sweden Tercentenary Fund)

Robert Bannert, Kenneth Hyltenstam, Gisela Håkansson, Kerstin Tevajärvi

Verbal reforms (The Bank of Sweden Tercentenary Fund)

Bengt Sigurd, Christer Platzack, Anna Orrling

Dictionaries for immigrants (The National Board of Education)

Bengt Sigurd, Agneta Svalberg

Commentator, a computer system simulating verbal behaviour (The Council for Research in the Humanities and Social Sciences)

Bengt Sigurd, Milan Bílý

Swedish Prosody (The Council for Research in the Humanities and Social Sciences)

Eva Gårding, Kerstin Tevajärvi

From text to prosody, a pilot project (The Council for Planning and Coordination of Research)

Gösta Bruce

The aerodynamics of phonation (The Wallenberg Trust)

Anders Löfqvist (with Peter Kitzing and Björn Carlborg, Malmö General Hospital)

PRAKTISK LINGVISTIK

edited by Eva Gårding

Nr 1-1979	Eva Gårding & Robert Bannert Optimering av svenskt uttal 10:- (incl. cassette recording)
Nr 2-1979	Kenneth Hyltenstam (Ed) Invandrarundervisning - aspekter på alfabetisering och den första nivån 10:-
Nr 3-1979	Robert Bannert Ordprosodi i invandrarundervisningen 20:- (incl. cassette recording)
Nr 4-1979	Eva Wigforss Svensk pluralbildning hos finska invandrarbarn - en psykolingvistisk utvecklingsstudie 20:-
Nr 5-1981	Robert Bannert Svårigheter med svenskt uttal: inventering och prioritering 30:- (incl. cassette recording)
Nr 6-1981	Gösta Bruce (Ed) Tema prosodi - seminarieuppsatser i fonetik 20:-

Inquiries to:

Praktisk lingvistik
Institutionen för lingvistik
Lunds Universitet
Helgonabacken 12
S-223 62 LUND
Sweden

(Prices in Swedish Crowns)

Swedish Post Office giro account 439 34 92 - 6

STUDIA LINGUISTICA

Eds Bengt Sigurd and Jan Svartvik

One volume of Studia Linguistica is published each year in two parts. The subscription price is 60 SEK (approximately $13). Orders may be sent to any international bookseller or the publishers, CWK Gleerup, Box 1205, S-221 05 Lund, Sweden.

TRAVAUX DE L'INSTITUT DE LINGUISTIQUE DE LUND
FONDÉS PAR BERTIL MALMBERG
PUBLIÉS PAR EVA GÅRDING ET BENGT SIGURD

I	*Carl-Gustav Söderberg.* A Typological Study on the Phonetic Structure of English Words with an Instrumental-Phonetic Excursus on English Stress. 1959.
II	*Peter S. Green.* Consonant-Vowel Transitions. A Spectrographic Study. 1959.
III	*Kerstin Hadding-Koch.* Acoustico-Phonetic Studies in the Intonation of Southern Swedish. 1969.
IV	*Börje Segerbäck.* La réalisation d'une opposition de tonèmes dans des dissyllables chuchotés. Etude de phonétique expérimentale. 1966.
V	*Velta Rūke-Dravina.* Mehrsprachigkeit im Vorschulalter. 1967.
VI	*Eva Gårding.* Internal Juncture in Swedish. 1967.
VII	*Folke Strenger.* Les voyelles nasales françaises. 1969.
VIII	*Edward Carney.* Hiss Transitions and their Perception. 1970.
IX	*Faith Ann Johansson.* Immigrant Swedish Phonology. 1973.
X	*Robert Bannert.* Mittelbairische Phonologie auf akustischer under perzeptorischer Grundlage. 1976.
XI	*Eva Gårding.* The Scandinavian Word Accents. 1977.
XII	*Gösta Bruce.* Swedish Word Accent in Sentence Perspective. 1977.
XIII	*Eva Gårding, Gösta Bruce, Robert Bannert* (Eds). Nordic Prosody. 1978.
XIV	*Ewa Söderpalm.* Speech Errors in Normal and Pathological Speech. 1979.
XV	*Kerstin Nauclér.* Perspectives on Misspellings. 1980.
XVI	*Per Lindblad.* Svenskans sje- och tje-ljud (Some Swedish Sibilants). 1980.

* Out of print.

Numbers V, VII-XII, XIV-XVI are available from the publisher, LiberLäromedel, Box 1205, S-221 05 LUND.

Numbers III, VI, XIII are available from the Department of Linguistics and Phonetics, Helgonabacken 12, S-223 62 LUND.

Department of Linguistics and Phonetics
LUND UNIVERSITY
Helgonabacken 12

PHONETICS

Eva Gårding	Fil Dr, Professor
Kerstin Hadding	Fil Dr, Professor Emeritus

Laboratory and Office

Karl-Gustav Jönsson	Engineer
Eva Körner	Secretary

Research and Teaching

Robert Bannert	Fil Dr, Docent
Anne-Christine Bredvad-Jensen	Fil Kand
Gösta Bruce	Fil Dr, Docent
Christina Hedqvist	Logoped (Department of Logopedics, Lund)
Kurt Johansson	Fil Lic, Lektor
Mona Lindau-Webb	Fil Dr (Phonetics Department, UCLA)
Anders Löfqvist	Fil Dr, Docent
Bengt Mandersson	Tekn Dr
Kerstin Nauclér	Fil Dr (Malmö General Hospital)
Ulrika Nettelbladt	Logoped
Kerstin Tevajärvi	Hum Kand
Sidney Wood	BA

Research Students

Catharina Andersson, Antonis Botinis, Anne-Christine Bredvad-Jensen, Christina Hedqvist, David House, Boris Larnert, Elisabeth Lindström-Sederholm, Eva Magnusson, Ulrika Nettelbladt, Ann-Christine Ohlsson, Kerstin Tevajärvi, Paul Touati, Sidney Wood.

AV-centralen i Lund 1982