PHONETICS LABORATORY
DEPARTMENT OF GENERAL LINGUISTICS
LUND UNIVERSITY

WORKING PAPERS

10 · 1975

On the occasion of the

60th anniversary of

Kerstin Hadding
7.2.1975

Congratulations from the Institute of Linguistics

Anja Bodén

Ulf Rylander

Majastina Lindgren

Kenneth Hyltenstam

Thore Pettersson

Faith Ann Johansson

Ragnvald Sanne

Christer Platzack

Britt Malmberg

Eva Larsson

Gösta Bruce

Eva Wallstedt

Robert Bannert

Eva Wigforss

Gunvor Hvenegård

Kurt Ellmann

Eva Gårding

[illegible]

Anders Löfqvist

Karl-G. Jönsson

Sidney Wood

Richard Kowinski

[illegible]

Previous numbers and contents:

WP 1 • 1969	K. HADDING M. HIRANO T. SMITH	Electromyographic study of lip activity in Swedish CV:C and CVC: syllables
	KURT JOHANSSON	Försök avseende vokaltransitions riktning och dess betydelse för "plats"-distinktionen bland tonande klusiler
WP 2 • 1970	MONA LINDAU	Prosodic problems in a generative phonology of Sweden
WP 3 • 1970	GÖSTA BRUCE	Diphthongization in the Malmö dialect
	EVA GÅRDING	Word tones and larynx muscles
	KURT JOHANSSON	Perceptual experiments with Swedish disyllabic accent-1 and accent-2 words
WP 4 • 1971	ROBERT BANNERT	Hat das Deutsche zugrundeliegende stimmhafte Spiranten?
	EVA GÅRDING	Laryngeal boundary signals
	KARIN KITZING	Contrastive acoustic analysis of vowel phonemes, pronounced by some North German and South Swedish high school pupils (A summary)
	SIDNEY WOOD	A spectrographic study of allophonic variation and vowel reduction in West Greenlandic Eskimo
WP 5 • 1971	MICHAEL STUDDERT-KENNEDY KERSTIN HADDING	Auditory and linguistic processes in the perception of intonation contours
	ANDERS LÖFQVIST	Some observations on supraglottal air pressure
WP 6 • 1972	ROBERT BANNERT	Zur Stimmhaftigkeit und Quantität in einem bairischen Dialekt
WP 7 • 1973	GÖSTA BRUCE	Tonal accent rules for compound stressed words in the Malmö dialect
	EVA GÅRDING PER LINDBLAD	Constancy and variation in Swedish word accent patterns
	KERSTIN HADDING MICHAEL STUDDERT-KENNEDY	Are you asking me, telling me or talking to yourself?
WP 8 • 1973	EVA GÅRDING	The Scandinavian word accents
WP 9 • 1973	PHONETIC SYMPOSIUM	Postgraduate students Stockholm • Lund
	SIDNEY WOOD	Speech tempo

Correspondence concerning this series should be addressed to:

Robert Bannert
Phonetics Laboratory
Lund University
Kävlingevägen 20
S-222 40 LUND
Sweden

PHONETICS LABORATORY
DEPARTMENT OF GENERAL LINGUISTICS
LUND UNIVERSITY

WORKING PAPERS

10 · 1975

TEMPORAL ORGANIZATION OF SWEDISH TONAL ACCENTS: THE EFFECT OF VOWEL DURATION

Robert Bannert and Anne-Christine Bredvad-Jensen

I. Introduction

Since Meyer's publications (Meyer 1937, 1954) the increasing interest of phoneticians in the Swedish word accents has resulted in many investigations, e.g. Malmberg (1953, 1955), Hadding-Koch (1961), Öhman (1965), Alstermark and Erikson (1971), Erikson and Alstermark (1972), Gårding and Lindblad (1973), Löfqvist (1973). A production model of the Swedish tonal accents was proposed by Öhman (1967) and criticized and amended by Gårding (1970). Both proposals are discussed by Bruce (1974) who provides new acoustical evidence concerning the tonal accents of the Stockholm dialect in stressed but non-focal sentence position. Gårding (1973) established a preliminary typology, based on the characteristics of the f_0 manifestations in bisyllabic words, not only for the Swedish accents but also for the other Scandinavian (Norwegian and Danish) word accents.

Investigating the effect of vowel duration on the f_0 contour of the stressed vowel, Erikson and Alstermark (1972) presented data for a Stockholm speaker in the stressed vowel of accent 2 in nonsense words. They discuss two ways in which the f_0 contour of accent 2, which in this dialect is mainly falling throughout the vowel segment, may be modified when vowel duration is decreased:

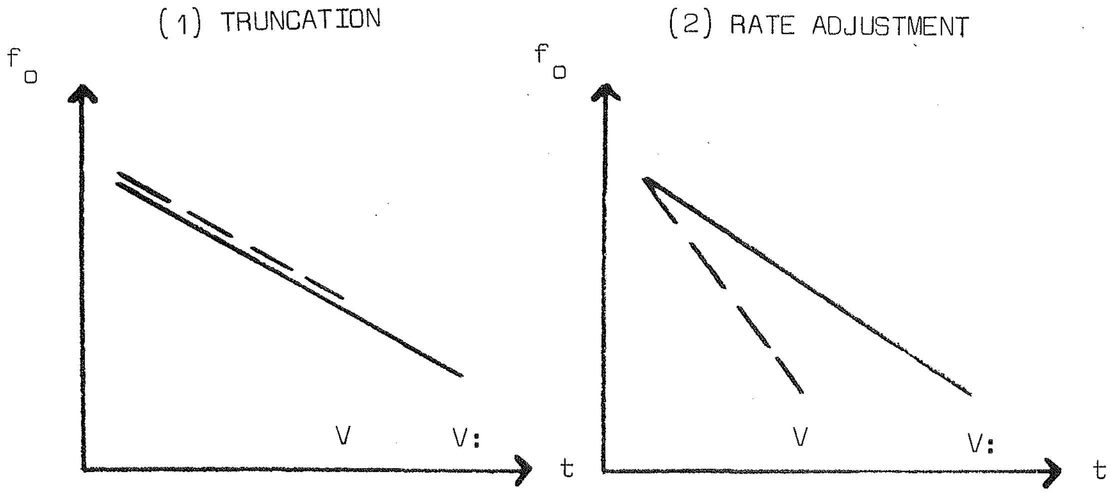
(1) TRUNCATION: the falling f_0 contour of the short vowel, starting at the same frequency level as that of the long one, merely ends earlier, which means that its last part is cut off or truncated.

(2) RATE ADJUSTMENT: the f_0 contour of the short vowel, starting from the same level as that of the long one, falls more rapidly towards the end, thus implying a reorganization of the tonal contour due to the shorter vowel duration.

The complete shape of the long vowel contour, the fall from the initial f_0 maximum to the final f_0 minimum, is preserved. The f_0 contour of the short vowel is compressed and the final f_0 minimum has to be reached in a shorter time. Therefore the curve has to fall more rapidly.

These two hypotheses are illustrated like this:

Figure 1. Erikson's and Alstermark's hypotheses:



The findings of Erikson and Alstermark support the truncation hypothesis in the case of accent 2 of their Stockholm informant.

The aim of this paper is to provide further information on the effect of vowel duration on the f_o contours of the stressed vowel segments of Swedish tonal accents. We preferred to work with meaningful words (in order to be able to use naive informants), and we studied both accents of several speakers representing different dialects.

As a consequence of the varying segmental features of the meaningful words in our material it is often not possible to correlate an observed change of the f_o contour with one phonetic feature only. Nevertheless we believe that our data will satisfy the purpose of an introductory investigation. Of course, further and more strictly controlled studies will be necessary.

II. Material

As we intended to work with informants who have no phonetical training, we chose four quadruplets of meaningful, bisyllabic, non-compound test-words from Elert's list of minimal pairs contrasting accent 1 and accent 2 (Elert 1972).

The most radical shortening of the vowel segment is achieved, as Erikson and Alstermark demonstrated, as a combination of phonological (by means of the long-short opposition) and phonetical conditioning (by means of the voiced-voiceless opposition). One pair of testwords (first line of each set below) contained a long vowel followed by a voiced stop. The corresponding pair (second line of each set below) contained a phonemically short vowel followed by a voiceless stop. The following four sets of testwords (a-d) were chosen:

	accent 1	accent 2
(a)	tag-en (def.pl. of "grip")	tagen (past.part. of taga = "to take")
	tack-en (def.pl. of "thanks")	tacken (solemn imperative of 2nd pers. pl. of tacka = "to thank")
(b)	stig-en (def.sing. of "path")	stigen (past.part. of stiga = "to rise")
	stuck-en (def.sing. of "stucco")	stucken (past.part. of sticka = "to stick")
(c)	steg-en (def.pl. of "step")	stege-n (def.sing. of "ladder")
	skott-en (def.pl. of "shot")	skotte-n (def.sing. of "Scotsman")
(d)	Hagen (e.g. family name)	hage-n (def.sing. of "grove")
	back-en (def.sing. of "crate")	backe-n (def.sing. of "hill")

The four sets of testwords contain the following structural (segmental) variations; firstly between the pair with long vowel and that with short vowel, secondly between the four sets themselves. The difference of voicedness of the medial consonant following long and short vowel respectively, however, is built into the material and common to the four sets. The feature of voicedness of the follwoing consonant is not expected to affect the f_o of the preceding vowel (Lehiste and Peterson 1961). In Löfqvist's data (1973), however, there seems to be such an effect on the f_o peaks but this raising or lowering effect is not consistent, not even for the two speakers (1 and 3) of the same dialect (Skåne).

In set (a) the pair with long vowels is segmentally and phonemically identical to that with short vowels, but there is an allophonic difference of vowel quality between the long and the short vowel, in IPA symbols [ɑ] and [a] respectively. This difference of vowel quality is also found in Erikson's and Alstermark's nonsense material.

The pairs in set (b) differ in the phonemic vowel quality, (/i/, vs /ʉ/), those in set (c) differ in the phonemic vowel quality (/e/ vs /o/) and in consonantal context. (The long vowel is preceded by a dental and followed by a velar stop, and vice versa for the short vowel.) The pairs of set (d), finally, show the same allophonic difference in vowel quality as set (a), both containing the vowel phoneme/a/, but the initial consonants differ. Vowel quality affects f_o (Lehiste and Peterson 1961), open vowels having the lowest f_o, close vowels the highest.

Between the four sets there is also a difference as to the number of initial consonants. Sets (a) and (d) are alike in having one single initial consonant, whereas in sets (b) and (c) the stressed vowel is preceded by an initial consonant cluster consisting of /s/ and the unaspirated voiceless stops /t/ and /k/ respectively. According to Rapp (1971) the number of prevocalic initial consonants does not seem to affect the f_o of the following vowel.

III. Informants

Our material is derived from seven speakers representing several Swedish dialects:

1. MV, Malmö	female,	20 years old
2. EW, Helsingborg	" ,	32 " "
3. EH, Stockholm	" ,	20 " "
4. LGP, Nybro,	male ,	23 " "
5. BD, Jämshög,	female,	19 " "
6. BS, Jönköping,	" ,	21 " "
7. JT, Ryd,	male ,	28 " "

The geographical distribution of the dialcets investigated is shown in figure 2. Each informant's speech is typical of the local area.

IV. Recordings

The testwords were embedded in the carrier sentence: Jag sa _____ där (I said _____ there). The test sentences were spoken with primary stress on the test words uttered as answers to the question: "What did you do there?". Each testword was repeated seven times with a falling sentence intonation and the items in each set were presented in the following order: long vowel accent 1, long vowel accent 2, short vowel accent 1, short vowel accent 2. The sentences were spoken at approximately the same rate and with a tempo most convenient to each speaker. The test sentences were recorded on a Studer A 62 tape recorder, speed 7.5 ips, in the sound-proof room of the Phonetics laboratory. The microphone was a Sennheiser MD 421 placed about 15 cms from the speaker.

As a buffer the informants produced the quadruplet Polen - pollen (accent 1) and pålen - pållen (accent 2), all four words differing in spelling. The buffer was not analysed.

V. Analysis

The productions of the four sets of testwords were analysed via a FONEMA analysis assembly and a Frøkjaer-Jensen intensity meter yielding a duplex oscillogram, a f_o curve and an intensity curve, all three written out on a Philips Oscillomink, paper speed 100 mm/sec. Duration was calibrated continuously by a time marker (100 msec). F_o was calibrated for each speaker from a test tape. Five repetitions of each testword were selected for measurement.

The beginning and the end of the stressed vowel segments were marked on the duplex oscillogram and on the f_o curve. The f_o curve during the vowel segment was drawn to connect the end points of the sweep oscillations. In a few cases the curve was smoothed. Consequently the f_o curves to be measured do not contain any articulatory ripple. For measurements of f_o we chose a number of points along the contour. The beginning (first vocal cord vibration) and the end (beginning of the occlusion of the stop), were taken for every stressed vowel segment. In order to catch the relevant features of the f_o contours between the beginning and the end thus defined we determined those points along the curve where it showed a marked change of direction. These points are:

(1) a f_o maximum or peak, defined as the turning point of the curve from rise to fall,

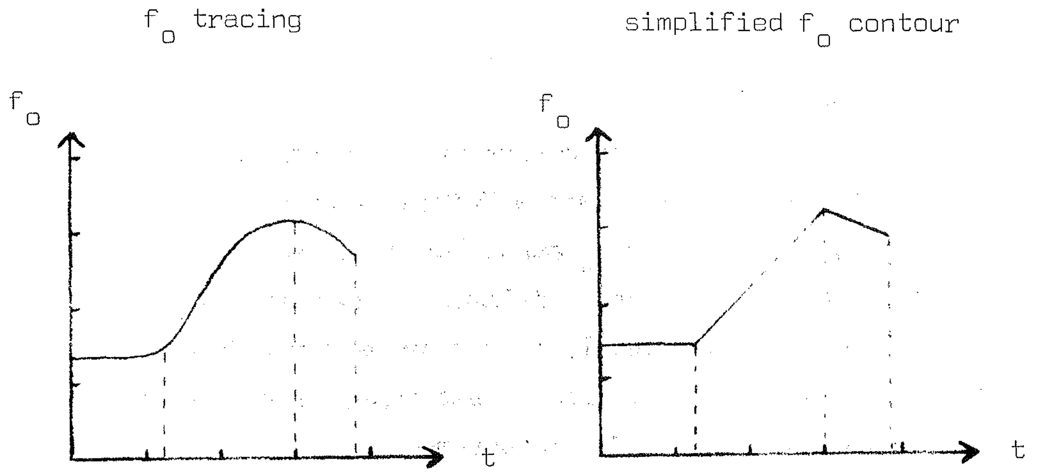
(2) a f_o minimum or dip, the turning point of the curve from fall to rise,

(3) a point where the contour shows a considerable change of rate,

(4) the beginning and the end of a level portion or plateau of the curve, turning points from fall or rise to level, and from level to fall or rise respectively.

The procedure for quantization is illustrated in figure 3. An example of an observed f_0 tracing is given to the left. Upon it the points chosen for measuring f_0 and their projection upon the time axis are indicated. To the right is shown the result, a simplified f_0 contour containing all essential information of the contour.

Figure 3. Procedure for quantization of the f_0 contours.



Given the simplified curves in our figures one can easily infer from the f_0 tracings of accent 1 and accent 2 of /tagen/ and /tacken/ for each informant (tables 2-8 below) the characteristics of the real f_0 contours in each testword.

The number of the intermediate points varies from 0 to 3. Therefore, the simplified f_0 contours in the diagrams are determined by 2 to 5 points of measurements. Their location along the time axis was measured as well as the segment durations.

Durations are measured to the nearest 10 msec and f_0 frequencies to the nearest 10 Hz. The arithmetic means are rounded to the half of the interval of measurement (5 msec and 5 Hz respectively). The range of measured f_0 frequencies and durations within each series of five repetitions was relatively small. Therefore the arithmetic means may be considered reliable. Differences of 5 msec or 5 Hz in the data should not be considered relevant.

VI. Results

Tables 2–8 give the arithmetic means of the f_0 frequencies measured at the beginning and at the end of the stressed vowel segments and at the intervening points, together with the mean durations of these points from the beginning of the segment. The f_0 contours of the four sets of testwords for each informant are diagrammed schematically in figures 4–8, 10, and 12. The diagrams of the left-hand column pertain to accent 1, those to the right to accent 2. The diagrams represent the following sets of testwords from top to bottom: (a) tagen, tacken, (b) stigen, stucken, (c) stegen, skotten, (d) hagen, backen. Vowel durations are normalized so that all the long vowel segments are given the same length in the figures. Accordingly the short vowel length indicates the duration of the short vowel relative to the corresponding long one. Thus it may be seen that the informants shorten the vowel segments to a different degree. The locations of the points representing f_0 measurements are proportional to their durations from the beginning of the segments.

Below each table (2–8) the optimal f_0 contours (one repetition of the long vowel pair ´tagen – `tagen and the short ´tacken – `tacken) for accent 1 and accent 2 of each informant are given as tracings from the oscillograms. With their aid correct f_0 contour can be inferred from each of the schematized.

Like Erikson and Alstermark, we use the contours of the long vowels as the reference when comparing the contours of the long and the short vowels. This does not imply any causal relationship. It would be equally reasonable to relate a comparison to the short vowels.

(a) The shortening of the stressed vowel

The degree of shortening of vowel duration from long vowel to short vowel varies among the speakers. The arithmetic means of the stressed vowel segment durations of the four testwords, the relative duration of the short vowel as a percentage of the corresponding long one (the vowel ratio V/V:), and the difference in % are given in table 1.

Table 1. Segment durations, vowel ratios and difference of vowel ratios between accent 1 and accent 2.

	Accent 1			Accent 2			Difference V/V: %
	MEAN DURATION msec		V/V: %	MEAN DURATION msec		V/V: %	
	LONG	SHORT		LONG	SHORT		
1. MALMÖ	176	114	65	189	118	63	2
2. HELSINGBORG	164	95	58	173	79	46	12
3. STOCKHOLM	229	88	38	239	79	33	5
4. NYBRO	174	98	56	176	93	53	3
5. JÄMSHÖG	205	98	48	205	83	40	8
6. JÖNKÖPING	276	110	40	261	98	38	2
7. RYD	221	120	54	208	105	50	4

It is found that the duration of the short vowel of accent 1 relative to the long one is always larger than that of the short vowel of accent 2 in spite of the differences in vowel quality. Thus it seems that the stressed vowel segment of accent 2 is shortened more than that of accent 1. The same relationship is found for speakers 1 from Skåne and 2 from Västergötland in Löfqvist (1973).

(b) F_o contours

1. MALMÖ (MV, Table 2, Figure 4)

Manifestation in the long vowels

The f_o curve of accent 1 falls during most of the long vowel segment, preceded by a short rise in three testwords but not in ´stegen. In ´stigen, the other testword beginning with the /st-/-cluster, this initial rise of only 5 Hz might also be looked upon as a plateau. It is considerably higher in ´tagen (20 Hz) and ´Hagen (40 Hz).

Accent 2 has a mainly rising contour with a short fall during the final quarter of the stressed vowel segment. In `stigen, however, the contour did not rise immediately at the beginning of the vowel but stayed at the initial low level for about 17 % (35 msec) of the segment. The appearance of this initial plateau seems to be a consequence of the larger duration (210 msec or about 25 %) of the /i/ segment compared to the duration of the long vowel /e/ (170 msec) in `stegen. Increased vowel duration due to a slower speaking rate makes a low level contour precede the rise of accent 2 in this dialect. The rise is delayed (Gösta Bruce, personal communication).

Vowel shortening

The short vowel segments display the most characteristic features of the long vowels, that is a falling f_0 in accent 1, a rising f_0 in accent 2. But the short f_0 curves differ from the long ones in one main respect. Minor variations apart, the final portions of the long vowel contours of both accents are missing in the short vowels. They are truncated.

2. HELSINGBORG (EW, Table 3, Figure 5)

Manifestation in the long vowels

This speaker belongs to the same dialect as the Malmö informant. The f_0 contour of the tonal accents of both speakers are very similar in shape. The curve of accent 1 falls throughout the long vowel segment from an initial high level for this speaker, independent of initial consonants.

The f_0 contour of accent 2 is characterized by a rise preceded by a rather short and slight initial fall and followed by a final fall. The f_0 in the long and short vowels of accent 2 started from approximately the same frequency, whereas this was the case for only two test pairs for accent 1, viz. ´stigen - ´stucken, and ´Hagen - ´backen.

Vowel shortening

When vowel duration is decreased radically the shape of the long vowel contour is retained in the short vowel segments. But again, several minor changes are observed.

The short vowel contours of accent 1 start with a short rise while for the long vowels they are level at the beginning. The peaks are either higher than the high level of the long vowel contour (´stucken), lower

(´skotten) or at the same level (´tacken, ´backen). After the peak the short vowel contours fall more rapidly than for the corresponding long vowels, representing instances of rate adjustment as is the case with the initial rise. Thus the short contours are totally reorganized (compressed).

Compression is also found in the short vowel contours of accent 2. The complete shape of the long vowel contours (slight initial fall or level, rise, fall or level) is preserved, except for the absence of the initial fall in `skotten which may be due to the different place of articulation of the preceding stop. The short vowel contours rise more rapidly than for the long vowels. The peaks of the short vowel contours are lower in `tacken and at about the same level in `stucken, `skotten, and `backen.

Comparing this informant's changes with those of the Malmö speaker, we find one main difference. The short vowel contours of the Malmö speaker for both accents represent mainly a truncation of the final part of the long vowel contours, whereas the complete shape of the long vowel contours of the Helsingborg informant is preserved but compressed in the short vowels.

Thus we find that two speakers of the same local dialect use different means for manifesting the tonal accents in short vowels followed by a voiceless consonant.

3. STOCKHOLM (EH, Table 4, Figure 6)

Manifestation in the long vowels

The f_0 curve of accent 1 of the long vowel is characterized by a rise which, except for ´Hagen, is rather slight while in ´stigen and ´stegen the contour ends in a plateau. The short rise of accent 2 seen in `tagen and `hagen is missing in `stigen and `stegen, which might be due to the initial consonant cluster. Otherwise the long vowel of accent 2 shows a falling f_0 throughout the vowel segment.

Vowel shortening

The characteristic pattern of the long vowel f_0 curves of the tonal accents is preserved in the short vowel segments. The short vowels of accent 1 are characterized by a mainly rising contour, those of accent 2 by a mainly falling one. The final parts of the long vowel contours of both accents are truncated. The rising-falling pattern of `tagen appears also in `tacken, but in `backen there is no rise, the $f_0$ contour falling throughout the short vowel segment as it is in `stucken and `skotten, which begin with a consonant cluster. Besides the truncation of the final parts the short vowel f_0 curves of both accents are shifted upwards along the frequency axis by about 20 Hz in most cases. These considerable frequency shifts of the short vowel contours do not seem to be a consequence of differences in segmental structure between the four sets of test words because they appear in all of them. The frequency-shifted short vowel contours may be due to the voicelessness of the following consonant or to the phonologically short vowel or both.

4. NYBRO (LGP, Table 5, Figure 7)

Manifestation in the long vowel

Accent 1 is manifested as an initial fall towards about one third of the long vowel segment followed by a rise during the remaining part of the vowel segment, thus displaying a f_0 minimum in the first half of the vowel.

Accent 2, on the contrary, has a f_0 maximum at the beginning of the vowel, in `stigen and `hagen manifested as a plateau, although the greater part of the segment is characterized by a fall.

Vowel shortening

When vowel durations are decreased by about 50 %, the remaining short vowel contours, in general, preserve the main pattern of the long ones.

It is notable, however, that for this informant the asymmetrical consonant context of the stegen - skotten pair does not change the shape of the contours.

The changes in the short vowel contours of this informant are rather difficult to interpret.

In ´backen the contour seems to be compressed although it does not rise more rapidly. The remaining three short vowel contours of accent 1 are truncated because the f_0 minimum appears at about the same point of time as in the long vowel contours. The long final rise of the long vowel contours is considerably shorter in the short ones indicating that the final part of the rise is truncated.

All the short vowel contours of accent 2 show some kind of reorganization. In ˋstucken and ˋskotten the curves of the short vowels are compressed. Both short [a] vowels have similar contours. They share one feature, namely the fall, with the other two contours. As it is steeper in ˋtacken and ˋbacken than in their corresponding long vowels, even these curves may be counted as instances of compression.

5. JÄMSHÖG (BD, Table 6, Figure 8)

Manifestation in the long vowels

The f_0 manifestations of both tonal accents in the long stressed vowels of this informant show a rather similar contour: a fall at the very beginning and a rise which is longer in accent 1 than in accent 2. But whereas accent 2 always has a rather long and considerable fall (about 30 Hz) from the f_0 peak to the segment boundary, this fall is only slight or nonexistent in accent 1.

Vowel shortening

When vowel duration is decreased the short vowel f_0 contours of both accents end with a f_0 maximum. In general, the short vowel segments of both accents are characterized by a rising contour, and in most cases the brief initial fall is preserved. The final falling parts of the long vowel contours of both accents never remain, they are truncated. Thus the f_0 manifestations of the tonal accents in the stressed vowels become rather similar to each other. There are, however, some differences between the remaining short vowel contours of accent 1 and 2. The short vowel contours of accent 1 end mostly at a higher frequency level than the corresponding ones of accent 2. Furthermore, the f_0 difference between the final f_0 and the f_0 minimum is larger in the short accent words than in the corresponding words of accent 2. The rising rate is higher in all short accent 1 words than in the corresponding ones of accent 2, except for ´backen.

The correct manifestations of the tonal accents in the minimal tonal pair ´backen and `backen by the informant were checked perceptually by the informant herself some weeks after the recordings had been made and also by a trained phonetician (GB) of our laboratory. It is true that the perceptual difference between accent 1 and accent 2 in this testpair and for this speaker is rather small.

The f_0 values of backen were then measured on narrow-band spectrograms made on a Voice Print PV 10. The f_0 was measured at the beginning, the minimum, and the final maximum in the 4th and 5th partials and calculated to the nearest 5 Hz, taking the mean of both if necessary.

In the case of backen we find it necessary to consider briefly the contour of the second (and final) syllable of the testwords as well. For the Jämshög speaker f_0 tracings of hagen and backen from the oscillograms are given in Figure 9. It will be seen that for this informant the high f_0 level of the second syllable of accent 1 is preserved in the testword with the short vowel while the rise of accent 2 is changed into a more level contour.

6. JÖNKÖPING (BS, Table 7, Figure 10)

Manifestation in the long vowels

The informant's accent 1 is manifested in the long vowels as a f_0 fall during at least the first half of the vowel, an f_0 minimum, and a final, slight rise towards the segment boundary. In ´tagen and ´Hagen, both with the vowel /a/, the fall is preceded by a short rise by 15 Hz.

The f_0 contour of accent 2, starting from about the same frequency level as accent 1, remains at this high level or rises slowly towards the middle of the vowel and falls towards the end of the vowel segment.

Vowel shortening

1. tacken, stucken, and skotten

When the vowel duration is decreased the fall of accent 1 and the high frequency level of accent 2 are preserved on the whole. The final part of the long vowel contours are truncated. But some slight variations may be noted. The short vowel curve of accent 1 in ´stucken is merely cut off at the segment boundary. In ´tacken the fall of f_0 dominates the vowel after the short rise at the very beginning. But the f_0 curve falls more rapidly than in the long vowel. Towards the segment boundary the rapid fall is slowed down as is the case in ´skotten. But here the f_0 falls at the same rate as in the corresponding long vowel.

2. backen

The contours of the short [a] vowel of backen of both accents deviate totally from the short vowel contours of the other three testwords which show a similar overall pattern. The correct performance of the tonal accents in ´backen and `backen was checked perceptually some months later. Narrow band spectrograms and measurements were made of all productions, as for the Jämshög informant.

As to segment durations, the short [a] vowels do not deviate from the short vowel segments of the other testwords (see Table 7). They are not even the shortest segments.

The most pertinent difference between the short vowel contours of accent 1 and 2 in backen lies in the steeper rise and in the final fall of the contour of accent 2.

In the case of backen we find it necessary to consider briefly the contour of the second (and final) syllable of the testwords as well.

Figure 11 gives the f_0 oscillogram tracings of one representative production each of hagen and backen for both accents. It can be seen that the f_0 contour of the second syllable of accent 1 has changed, too. The rise of the second syllable of ´Hagen has become a high level. The pattern of the f_0 contour of the second syllable of `hagen and `backen is not changed, however, remaining rising even after the short vowel followed by the voiceless medial stop.

7. RYD (JT, Table 8, Figure 12)

Manifestation in the long vowels

Accent 1 is manifested as a falling f_0 reaching its lowest level in the second half of the long vowel, in ´stigen as late as at the segment boundary.

The contour of accent 2, unlike accent 1, rises, after an initial short fall in `tagen and `stegen, towards a f_0 maximum at about the middle of the long vowel segment. It then falls during the second half, reaching the same low frequency level at the segment boundary as the corresponding curve of accent 1.

Vowel shortening

When vowel duration is decreased the final parts of the long vowel contours of both accents are missing in all cases but one (´stucken). Apart from the truncated long vowel fall of accent 2, the remaining short vowel contours, except for ˋstucken, are compressed resulting in a steeper rise.

(c) Conclusions

To sum up we will attempt in the following table to characterize the main change in the short vowel contours compared to the long ones for each informant and for each accent.

Table 9. Main changes in the short contours.

		Accent	
Figure	Informants and place	1	2
4	MV, Malmö	Truncation	Truncation
5	EW, Helsingborg	Compression	Compression
6	EH, Stockholm	Truncation	Truncation
7	LGP, Nybro	Truncation	Compression
8	BD, Jämshög	Truncation	Truncation
10	BS, Jönköping	Truncation	Truncation
12	JT, Ryd	Truncation	Truncation
			Compression

Table 9 suggests that the short contours of our seven informants are derived from the long ones by the principles of truncation in most cases and by compression in a few cases irrespective of the shape of the f_0 contours of the word accents.

Within the same dialect speakers can differ in the kind of change observed. Furthermore, a speaker may show truncation in one accent and compression in the other. Both principles may be combined within one accent.

VII. Discussion

The alternation of voiced stop versus voiceless stop in the testpairs was built into the material on the assumption that a following consonant does not affect the f_0 contour of a preceding vowel (Mohr 1971, Leandersson and Lindblom 1971). An investigation of the effect of the segmental context on a given f_0 contour of a Gothenburg male informant was carried out at the laboratory. It reveals very clearly that the whole f_0 curve of accent 2, a rising-falling pattern (the peak being located at a point about one third from the beginning of the vowel segment), is raised considerably, by about 20 Hz, before a voiceless stop compared with a nasal, and somewhat less compared with the corresponding voiced stop. The f_0 contour of accent 1, however, which falls throughout the vowel, appears not to be affected by the different features of the following consonants /m, b, p/. Partly contradictory evidence is reported in Löfqvist (1973). That is why we do not want to give attention to the frequency-shifted short vowel contours found with some of the informants. It is tempting to associate the f_0 raising in short vowels, before voiceless stops, and in fast speech (Gårding etal., 1975) with one common mechanism: a greater tension of the muscles of the vocal organs.

Apart from minor variations, it seems that for the whole material of our seven informants the short vowel contour of the first, stressed vowel are achieved by two main strategies or programmes:

(1) Certain parts of the f_0 contours are not manifested. They are truncated. The remaining parts of the contours of both accents may (a) still be different from each other or (b) they may resemble each other.

(2) The complete shape of the contour is compressed. Therefore the curve has to rise or fall more quickly (rate adjustment).

In general, we only find support for the two hypotheses discussed by Erikson and Alstermark (1972). We agree with them in calling the one effect of vowel duration on the f_0 of Swedish word accent for truncation, but we prefer to call the other effect for compression which is a more general term for covering that change than is rate adjustment. A compressed contour must be produced with an increased speed for the rise or the fall, although the absolute values of the f_0 maxima or f_0 minima of the long contours need not be reached.

As to these tonal accents several linguists (e.g. Elert 1970, 46 and references there) hold that there is only one phonemically relevant tonal accent, namely accent 2. It is manifested in the first stressed syllable and is followed by a contour similar to that of accent 1 (e.g. Gårding 1970, Gårding and Lindblad 1973), which is considered to be the manifestation of the sentence intonation. The data of Alstermark and Erikson (1971) support this assumption as does Bruce (1974). He points out that the first syllable of accent 2 in the Stockholm dialect may be a manifestation of the "pure" accent 2 and that the "pure" accent 1 is not realized in sentence stressed position where the f_0 contour is determined by the sentence intonation instead.

Thus it seems that shortening of the vowel of the first syllable in an accent 2 word merely affects the tonal accent while it changes the sentence intonation contour in accent 1.

A deviating contour is found in one of the short vowels of the Jönköping informant: the short vowel contour of ´backen shows a rise, unlike the falling f_0 curves in all the other testwords. As the f_0 curves of `backen are characterized by a rise, the short vowel contours of both accents in the stressed vowel have become very similar. And yet, the tonal contrast in these two word is preserved, the tonal difference being signalled in the second syllable.

The observed changes in all the testpairs of the Jämshög informant and the radical reorganization of the short vowel f_0 contour of the Jönköping informant result in diminishing the tonal contrast in the first syllable. A decrease of a given contrast or the neutralization of a given distinction are often found in segmental phonology.

But as the domain of the tonal accent 2 is at least a bisyllabic word, the second syllable (or another following syllable depending on the stress and accentuation rules of the dialect, see Bruce 1974) will be available for the accent manifestation if there is no neutralization due to the deletion of stress.

At least one instance of radical deviation and the resulting similarity of the short vowel contours in the first syllable, as well as other reasons, motivate a study of not only the f_0 curves of one syllable but also those of the whole word (consisting of at least two syllables) in further investigations.

References

Alstermark M. and Y. Erikson. 1971. Swedish word accent as a function of word length. STL-QPSR 1, 1-13

Bredvad-Jensen A.-C. 1973. An empty space on the map. Phonetic Laboratory, Lund University. Working Papers 9, 54-57

Bruce G. 1974. Tonaccenter och satsbetoning. Institutionen för lingvistik, Lunds universitet. Lingvistiska månadsmeddelanden. Februari

Elert C.-C. 1970. Ljud och ord i svenskan. Uppsala

Elert C.-C. 1972. Tonality in Swedish: Rules and a list of minimal pairs. Studies for Einar Haugen. Scherabon Firchow et al. (ed.), 151-173. The Hague

Erikson Y. and M. Alstermark. 1972. Fundamental frequency correlates of the grave word accent in Swedish: The effect of vowel duration. STL-QPSR 2-3, 53-60

Gårding E. 1970. Word tones and larynx muscles. Phonetics Laboratory, Lund University. Working Papers 3, 20-46

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

Gårding E., R. Bannert, A.-C. Bredvad-Jensen, G. Bruce, and K. Nauclér. 1974. Talar skåningarna svenska? Åttonde sammankomsten för svenskans beskrivning. Lund

Gårding E., O. Fujimura, H. Hirose and Z. Simada. 1975. Laryngeal control of Swedish word accents. Phonetics Laboratory, Lund University. Working Papers 10, 53-82

Hadding-Koch K. 1961. Acoustico-phonetic studies in the intonation of Southern Swedish. Traveaux de l'Institut de Phonétique de Lund III. B. Malmberg (ed.). Lund

Leanderson R. and B. Lindblom. 1972. Muscle activation for labial speech gestures. Acta Otolaryng. 73, 362-373

Lehiste I. and G. Peterson. 1961. Some basic considerations in the analysis of intonation. JASA 33, 419-425

Löfqvist A. 1973. On intrinsic and extrinsic F_o variations in Swedish tonal accents. Institutionen för lingvistik, Lunds universitet. Lingvistiska månadsmeddelanden. Oktober, 1-16

Malmberg B. 1953. Sydsvensk ordaccent, en experimentalfonetisk undersökning. Lunds Universitets Årsskrift. NF Av. 1, Bd. 49, Nr 2, 3-35

Malmberg B. 1955. Observations on the Swedish Word Accent. Haskins Laboratories, Report. Mimeographed

Meyer E.A. 1937. Die Intonation im Schwedischen, I: Die Sveamundarten. Studies Scand. Philol. Nr 10, Univ. Stockholm

Meyer E.A. 1954. Die Intonation im Schwedischen, II: Die norrländischen Mundarten. Studies Scand. Nr 11, Univ. Stockholm

Mohr B. 1971. Intrinsic variations in the speech signal. Phonetica 23, 65-93

Öhman S. 1965. On the coordination of articulatory and phonatory activity in the production of Swedish tonal accents. STL-QPSR 2, 14-19

Öhman S. 1967. Word and sentence intonation: A quantitative model. STL-QPSR 2-3, 20-54

Rapp K. 1971. A study of syllable timing. STL-QPSR 1, 14-19

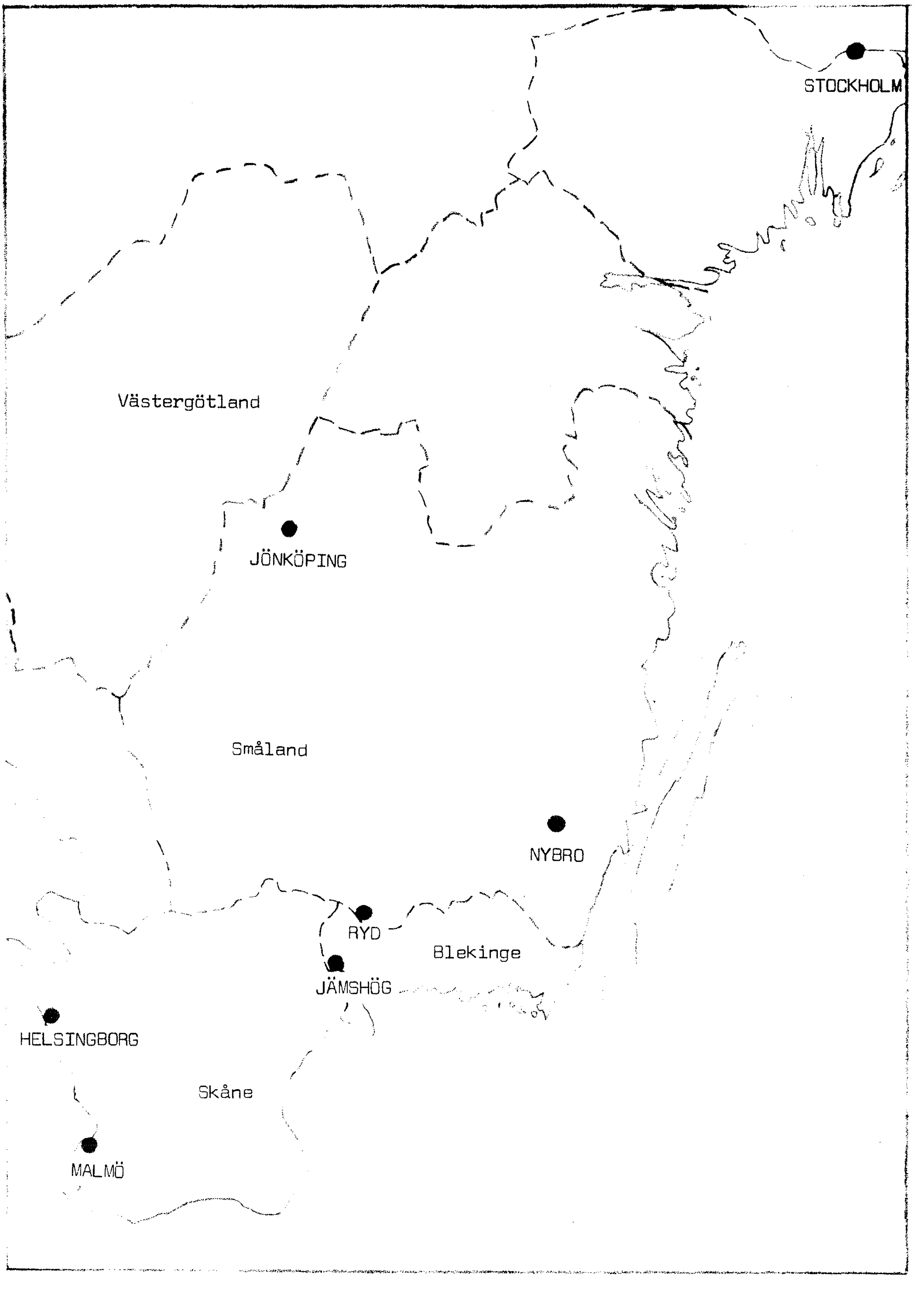


Figure 2. The geographical distribution of the informants.

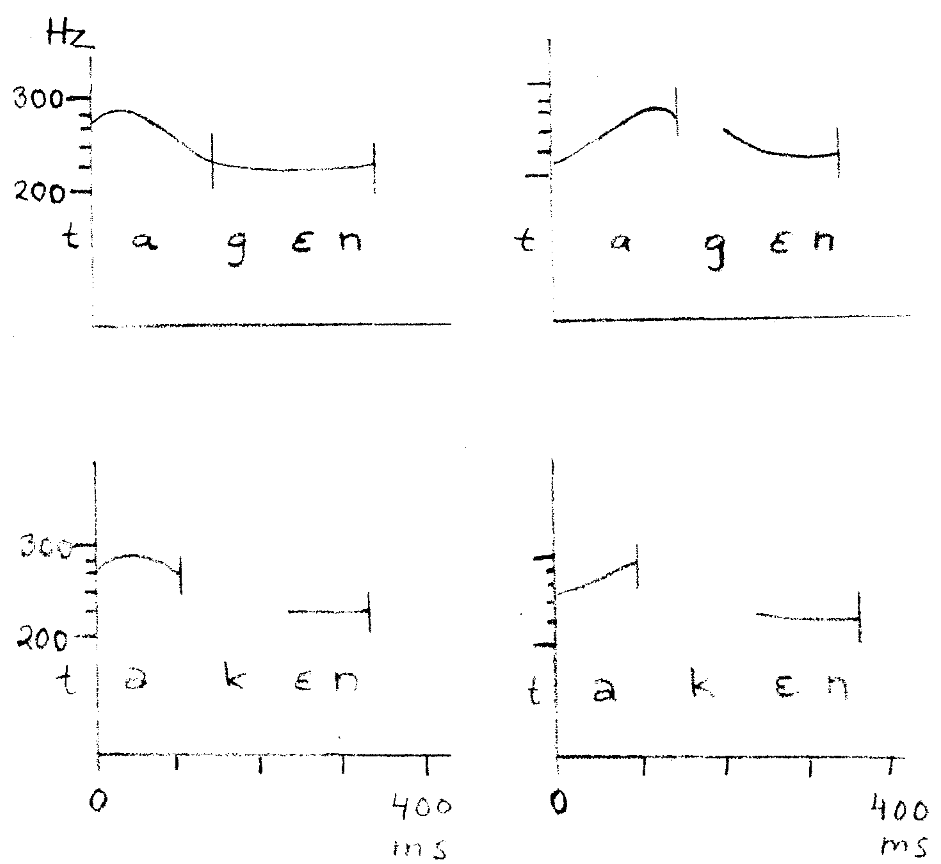
Table 2. MALMÖ

F_0 values (in Hz) and durations (in msec).

		Accent 1			Accent 2			
tagen	f_0:	270	290	220	230		300	280
	t :	0	25	145	0		140	175
tacken	f_0:	275	295	260	250			300
	t :	0	35	105	0			105
stigen	f_0:	295	300	205	230	230	290	260
	t :	0	30	185	0	35	160	210
stucken	f_0:	300	305	220	235			315
	t :	0	25	125	0			130
stegen	f_0:	285		200	240		275	265
	t :	0		170	0		125	170
skotten	f_0:	280		240	230			285
	t :	0		95	0			105
hagen	f_0:	235	275	195	210		280	260
	t :	0	50	205	0		165	200
backen	f_0:	245	270	230	205			285
	t :	0	55	130	0			130

F_0 tracings of tagen and tacken, accent 1 to the left, accent 2 to the right.

Malmö.



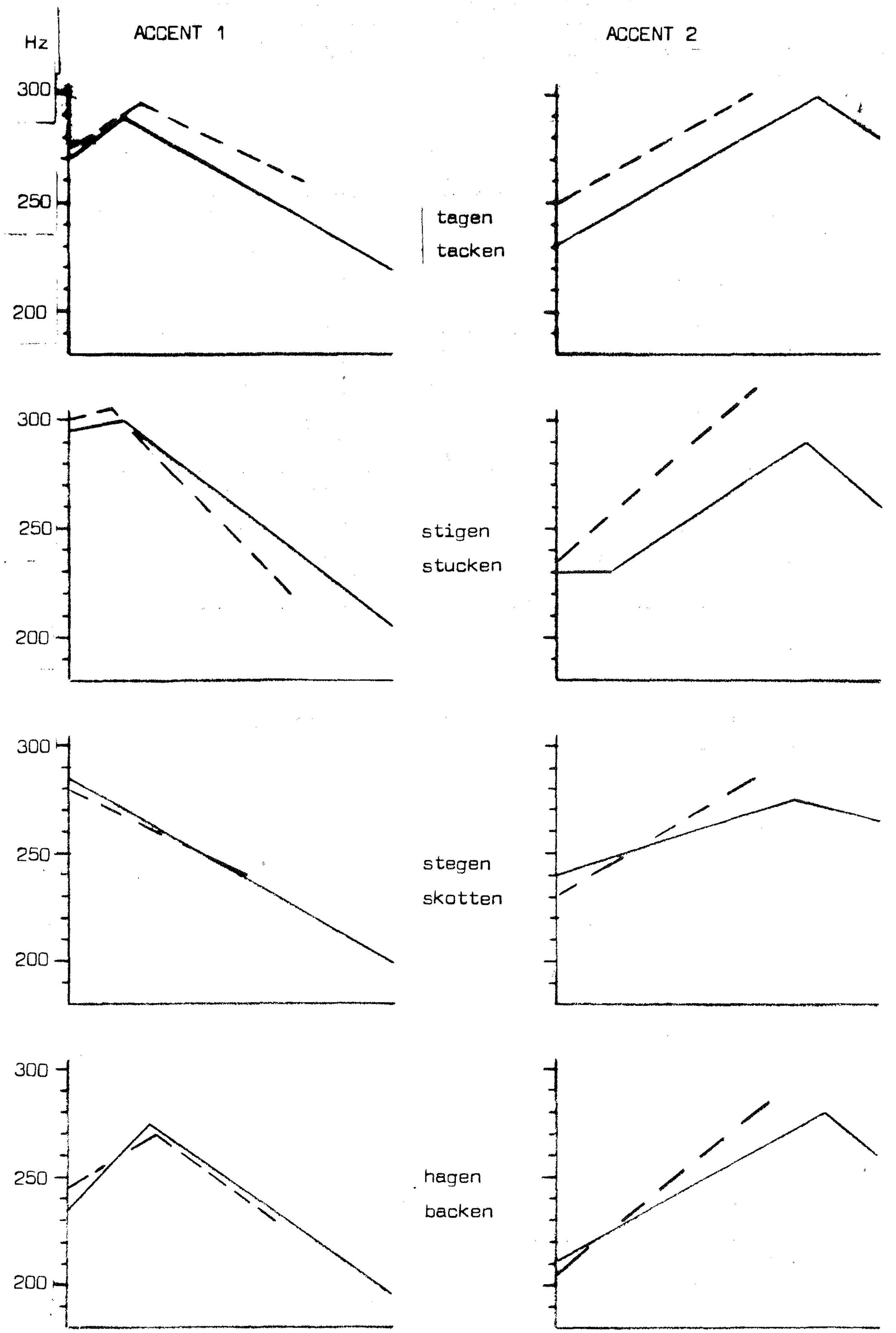


Figure 4. Simplified f_o contours of the first vowel segment. MALMÖ
The duration of the long vowels is normalized.
Dashed line: short vowel

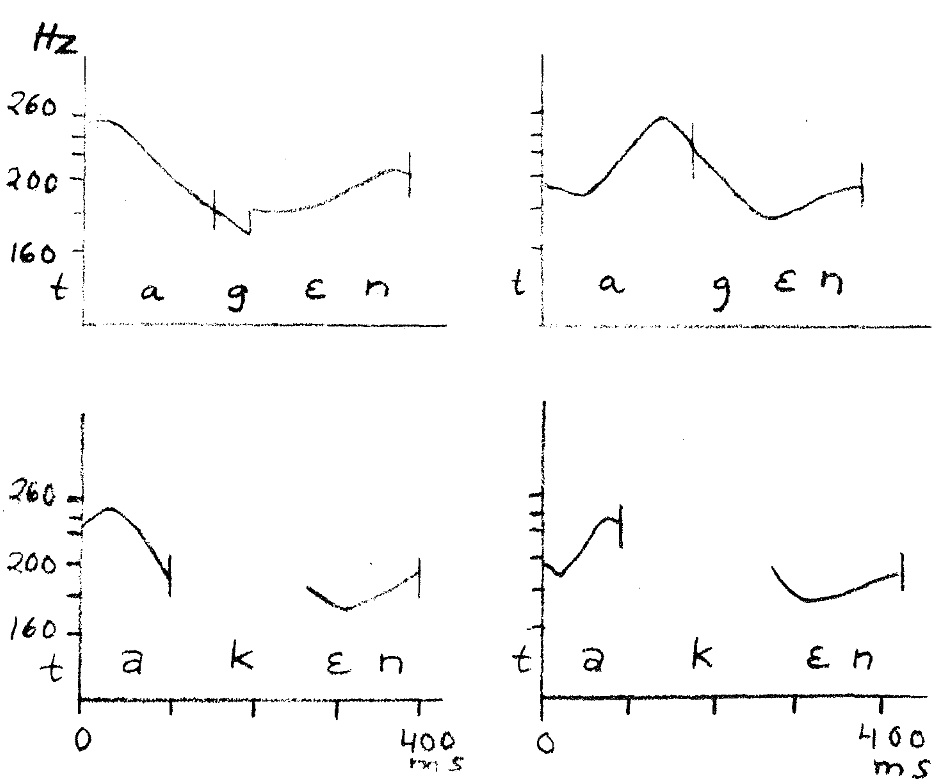
Table 3. HELSINGBORG

F_o values (in Hz) and durations (in msec).

		Accent 1			Accent 2			
tagen	f_o:	255	255	190	195	190	250	220
	t :	0	40	150	0	55	155	185
tacken	f_o:	240	255	195	195	190	235	230
	t :	0	30	100	0	35	65	85
stigen	f_o:	215	215	160	180	175	210	200
	t :	0	40	170	0	35	140	170
stucken	f_o:	220	225	175	185	185	210	210
	t :	0	15	75	0	15	50	65
stegen	f_o:	240	240	170	180	175	220	205
	t :	0	45	180	0	45	135	170
skotten	f_o:	210	225	160	180		220	210
	t :	0	25	90	0		40	60
hagen	f_o:	235	235	175	190	185	240	220
	t :	0	35	155	0	40	135	165
backen	f_o:	230	235	170	185	185	235	225
	t :	0	35	115	0	25	80	105

F_o tracings of tagen and tacken, accent 1 to the left, accent 2 to the right.

Helsingborg.



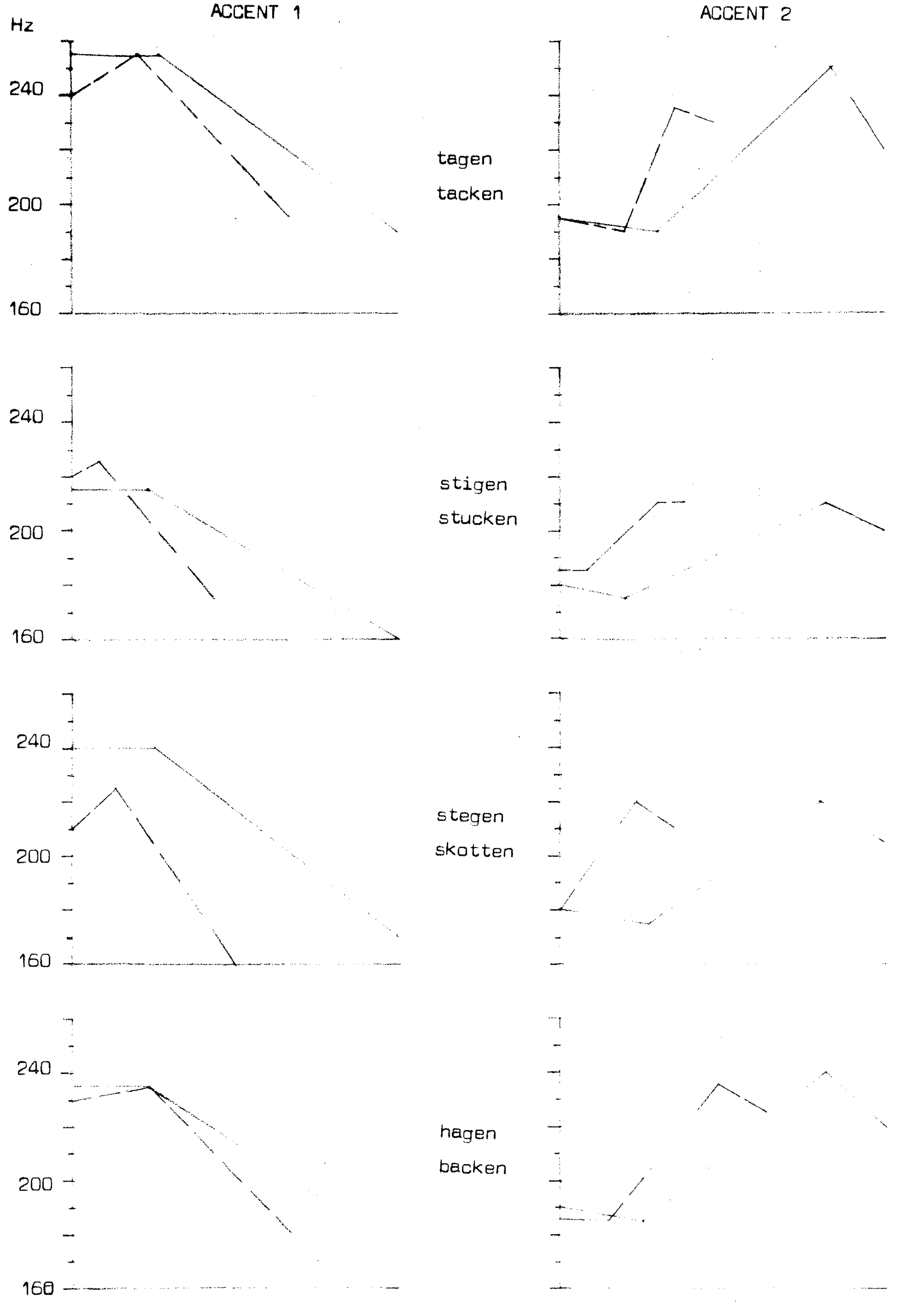


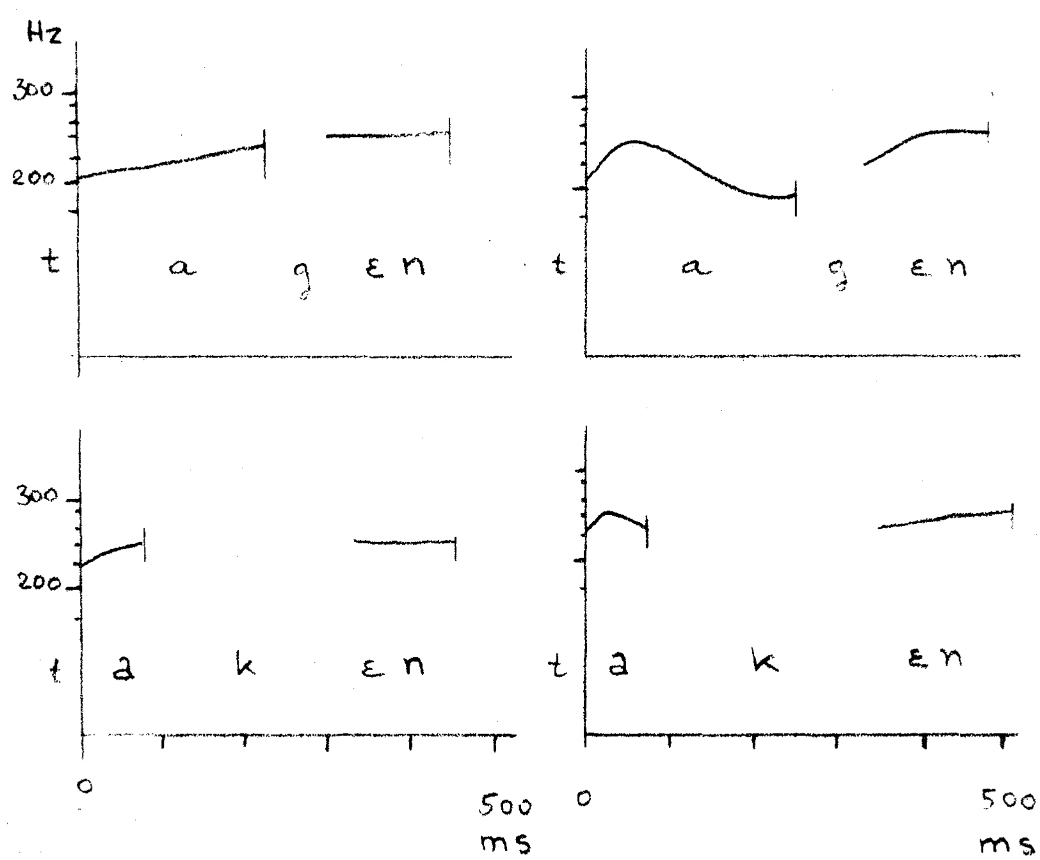
Figure 5. Simplified f_o contours of the first vowel segment. HELSINGBORG
The duration of the long vowels is normalized.
Dashed line: short vowel

Table 4. STOCKHOLM

F_o values (in Hz) and durations (in msec).

		Accent 1			Accent 2		
tagen	f_o:	220		240	240	260	195
	t :	0		240	0	50	260
tacken	f_o:	235		260	250	270	250
	t :	0		95	0	30	90
stigen	f_o:	245	255	255	270		215
	t :	0	120	190	0		215
stucken	f_o:	270		280	280		260
	t :	0		80	0		65
stegen	f_o:	235	260	260	260		200
	t :	0	200	255	0		240
skotten	f_o:	250		260	280		255
	t :	0		90	0		80
hagen	f_o:	230		285	240	250	200
	t :	0		230	0	50	240
backen	f_o:	235		270	255		245
	t :	0		95	250		80

F_o tracings of tagen and tacken, accent 1 to the left, accent 2 to the right.
Stockholm.



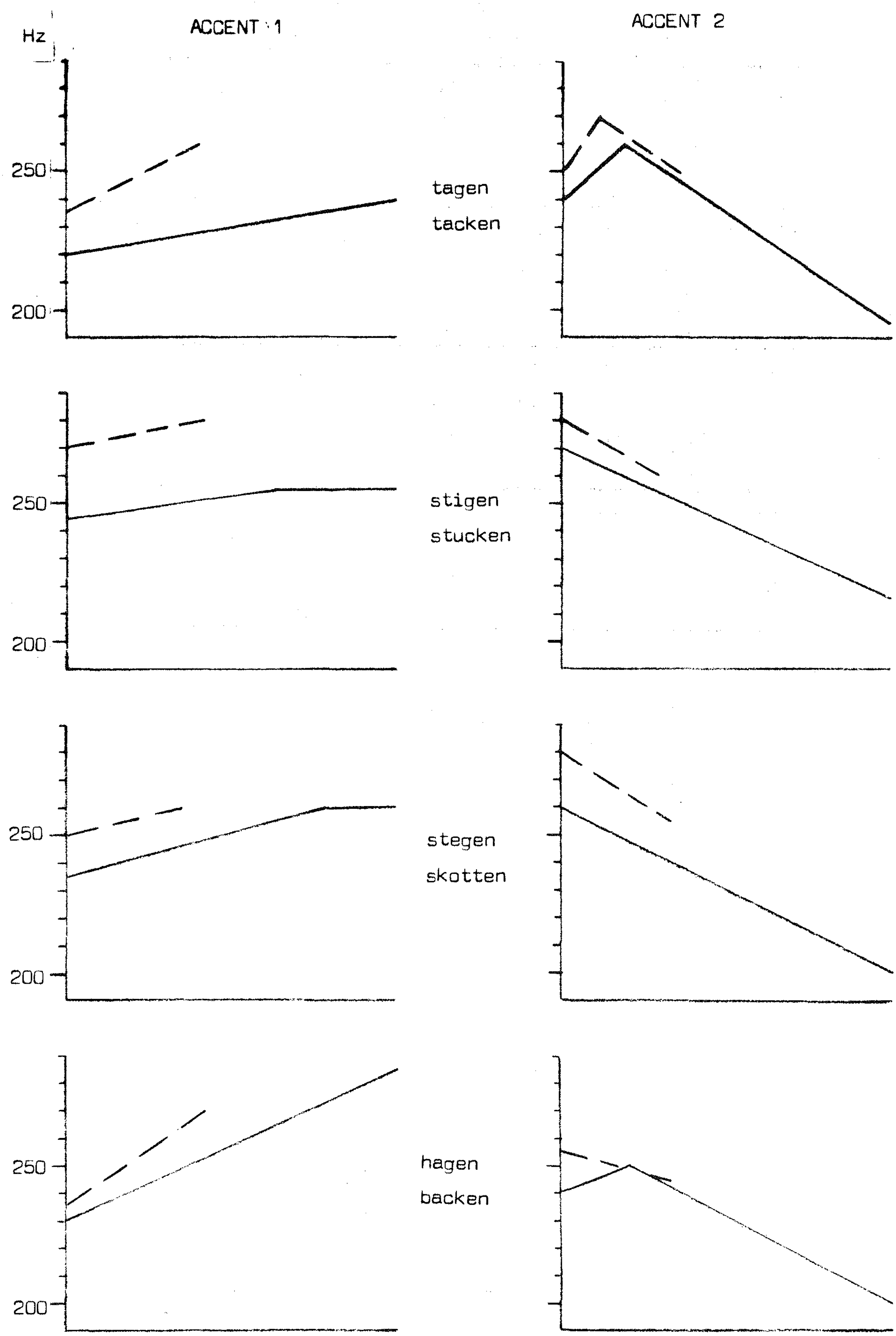


Figure 6. Simplified f_o contours of the first vowel segment. STOCKHOLM
The duration of the long vowels is normalized.
Dashed line: short vowel

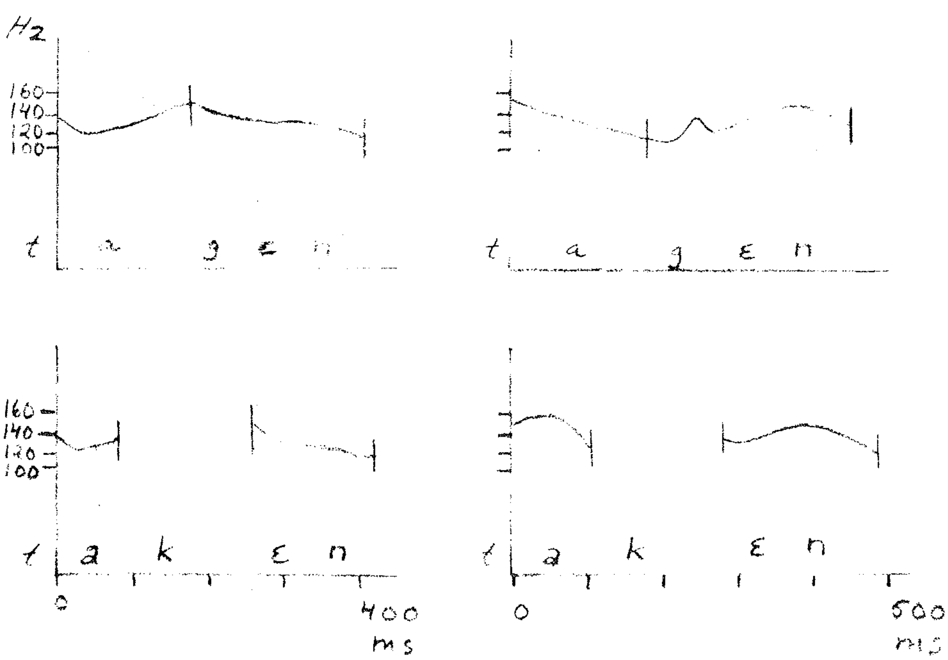
Table 5. NYBRO

F_o values (in Hz) and durations (in msec).

		Accent 1			Accent 2		
tagen	f_o:	135	115	155	150		115
	t :	0	45	180	0		180
tacken	f_o:	135	125	135	145	155	125
	t :	0	40	90	0	50	90
stigen	f_o:	130	120	140	140	140	110
	t :	0	50	150	0	55	160
stucken	f_o:	150	130	140	150	150	130
	t :	0	40	90	0	20	90
stegen	f_o:	135	130	140	145	115	115
	t :	0	55	160	0	145	175
skotten	f_o:	140	135	145	150	135	135
	t :	0	35	95	0	45	90
hagen	f_o:	125	110	140	155	155	120
	t :	0	75	205	0	50	190
backen	f_o:	140	120	150	140	145	130
	t :	0	25	115	0	65	100

F_o tracings of tagen and tacken, accent 1 to the left, accent 2 to the right.

Nybro.



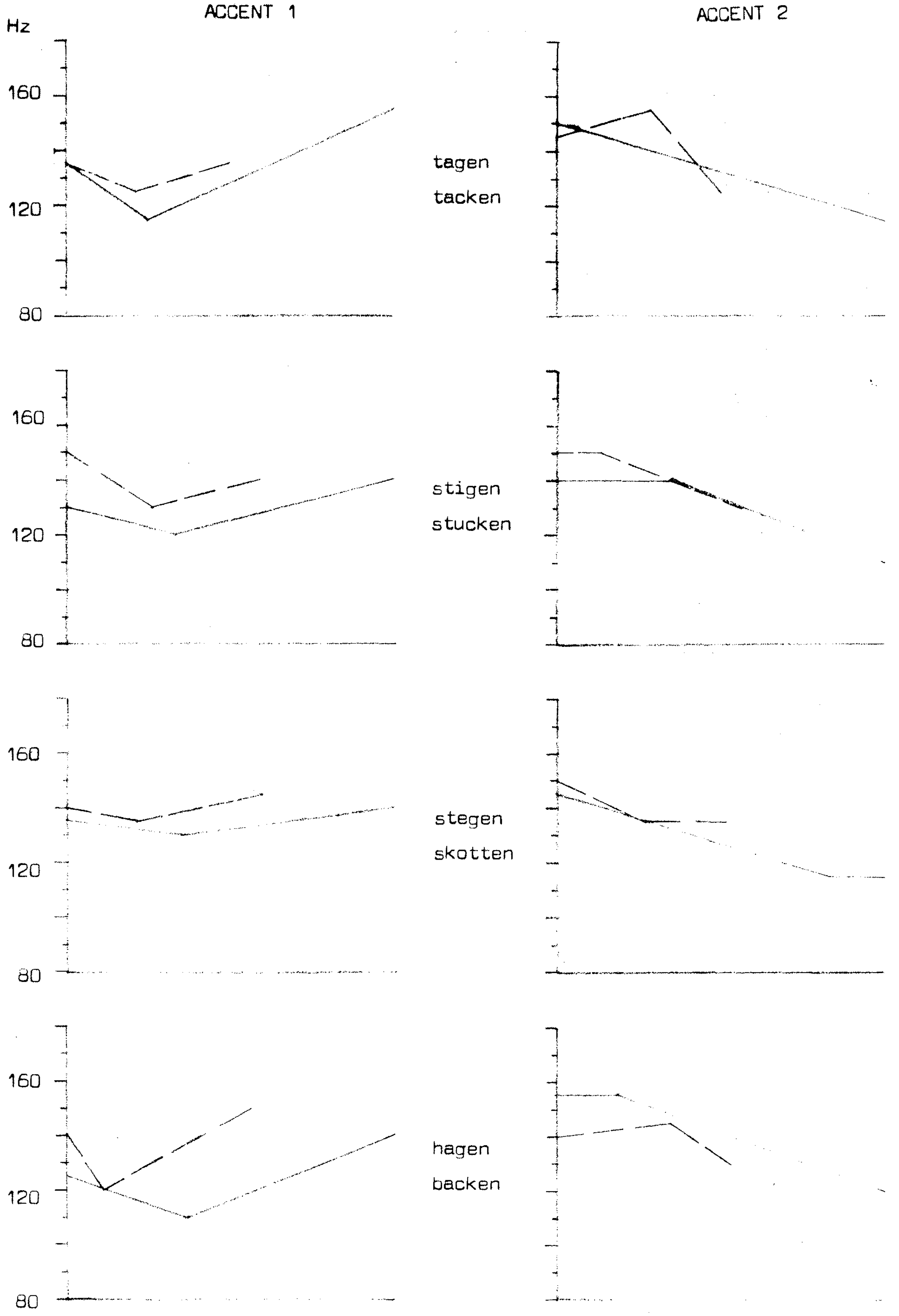


Figure 7. Simplified f_o contours of the first vowel segment. NYBRO
The duration of the long vowels is normalized.
Dashed line: short vowel

Table 6. JÄMSHÖG

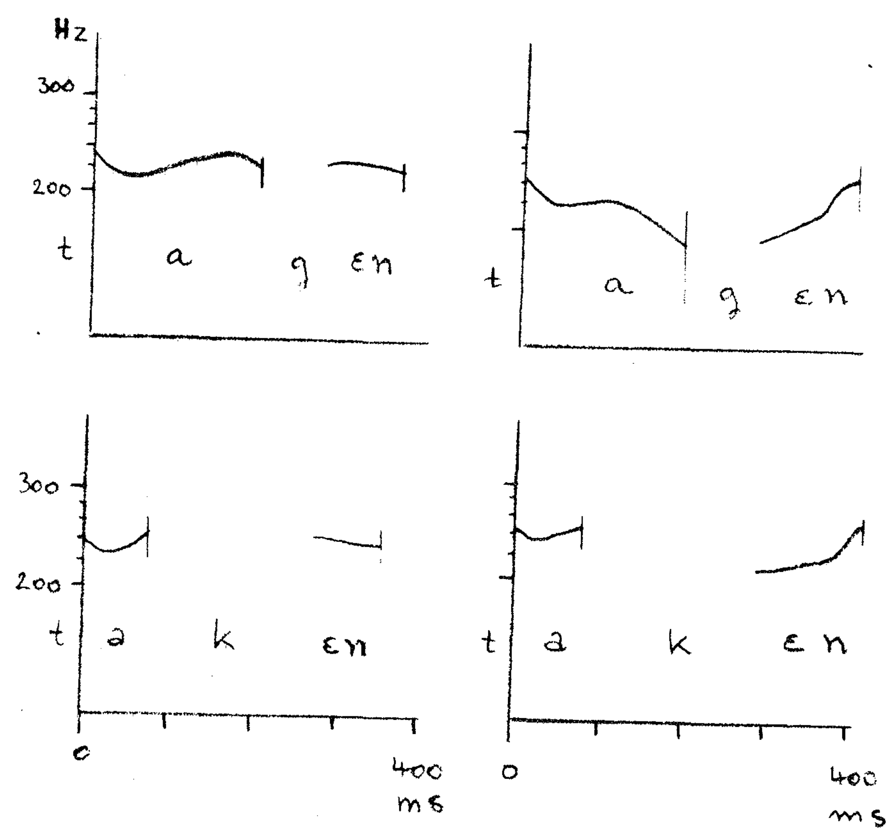
F_0 values (in Hz) and durations (in msec).

		Accent 1				Accent 2			
tagen	f_0:	250	210	240	220	250	230	240	190
	t :	0	50	170	210	0	40	100	200
tacken	f_0:	240	230		250	250	230		240
	t :	0	20		80	0	20		80
stigen	f_0:	235	210		240	240	215	235	210
	t :	0	35		200	0	30	115	215
stucken	f_0:	240	220		240	230			235
	t :	0	20		85	0			70
stegen	f_0:	220	210	240	235	230	220	240	220
	t :	0	35	150	210	0	20	115	205
skotten	f_0:	215			255	220	220		240
	t :	0			100	0	15		80
hagen	f_0:	230	210	245	235	225	210	235	200
	t :	0	30	165	200	0	20	110	200
backen	f_0[1]:	210	200		245	210	205		235
	t :	0	25		125	0	25		100

[1] measured on narrow-band sonagrams

F_0 tracings of tagen and tacken, accent 1 to the left, accent 2 to the right.

Jämshög.



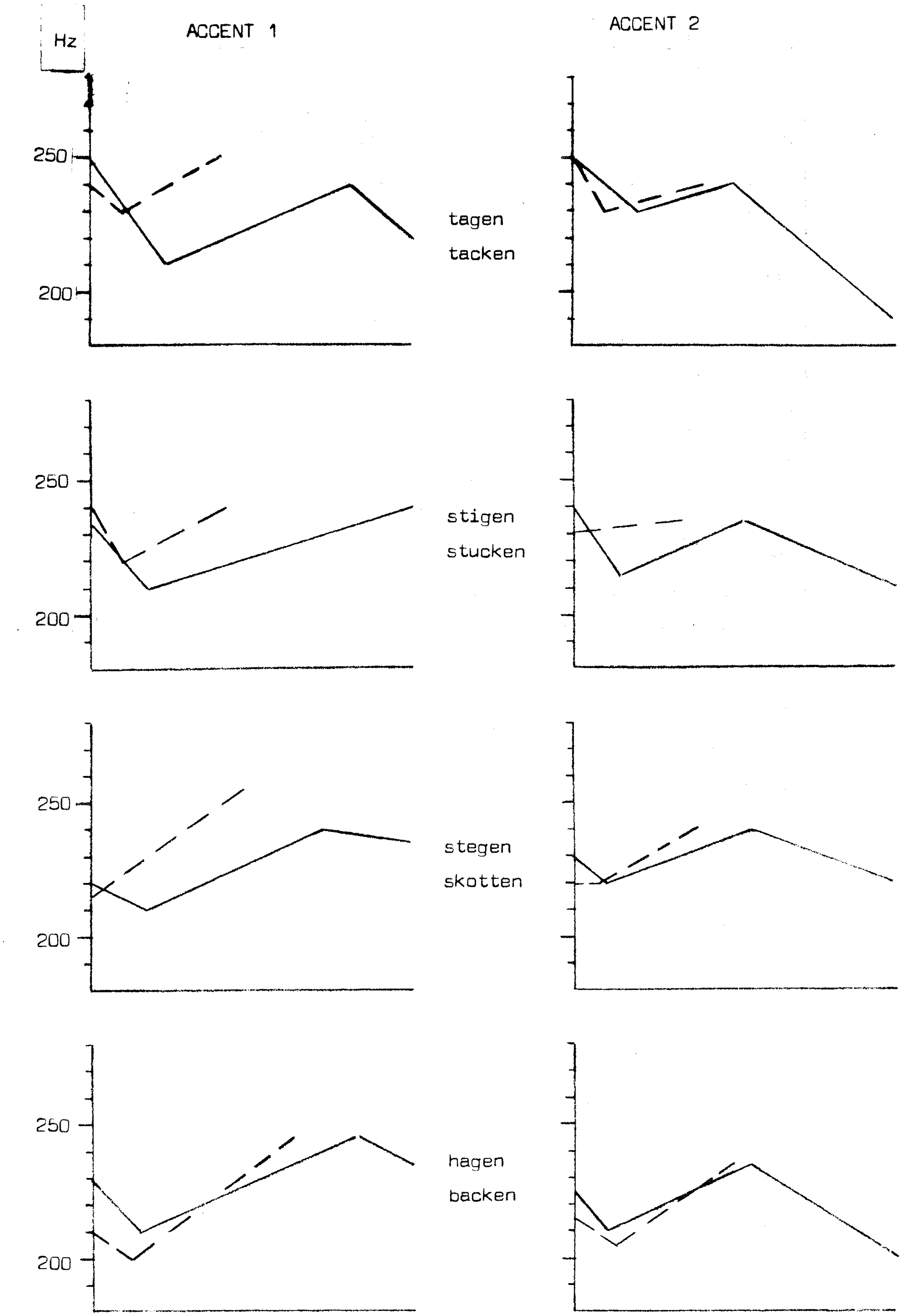


Figure 8. Simplified f_0 contours of the first vowel segment. JÄMSHÖG
The duration of the long vowels is normalized.
Dashed line: short vowel

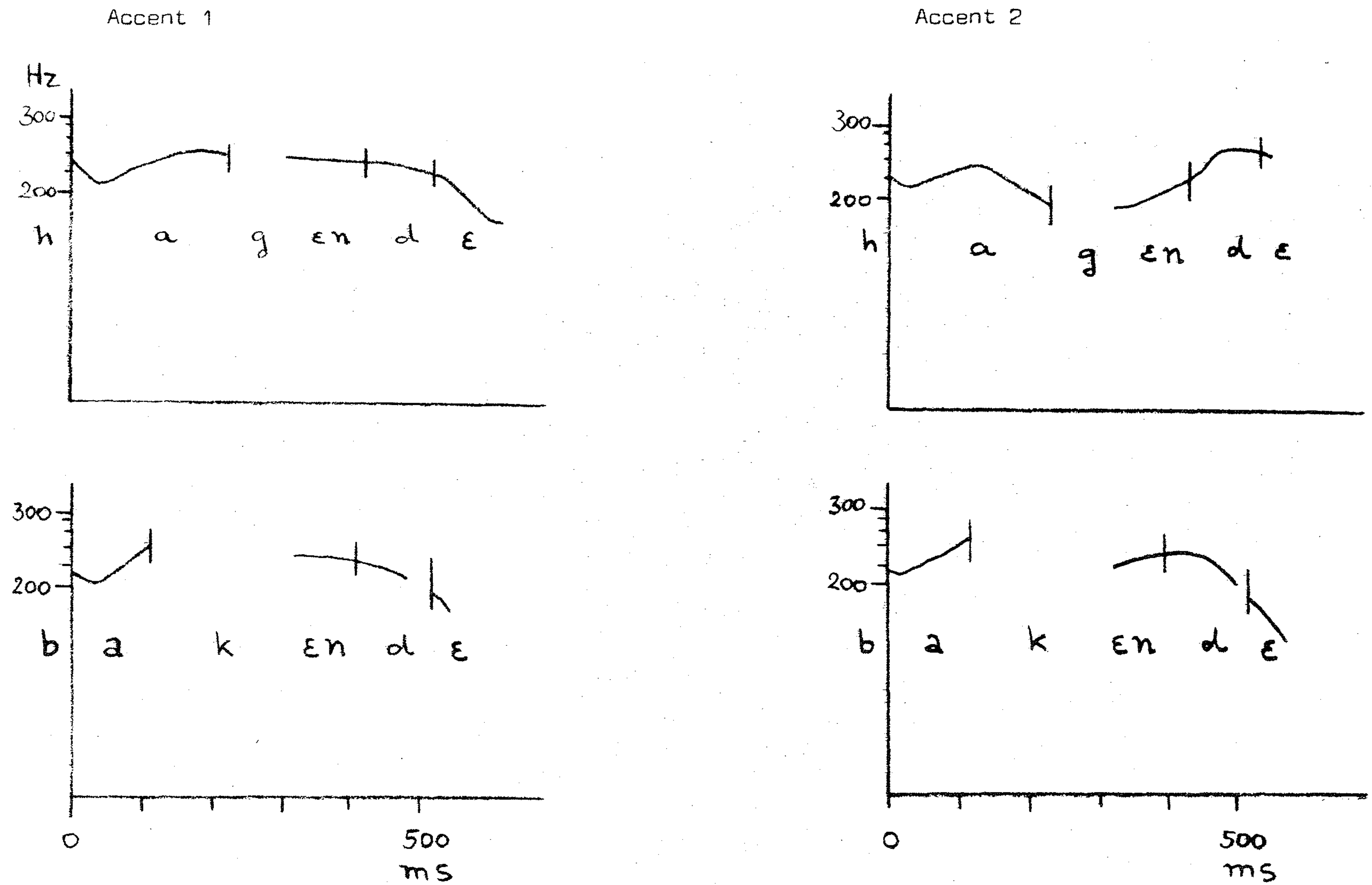


Figure 9. F_0 tracings of hagen (above) and backen (below), accent 1 to the left, accent 2 to the right. Jämshög informant

Table 7. JÖNKÖPING

F_o values (in Hz) and durations (in msec).

		Accent 1				Accent 2				
tagen	f_o:	230	245	170	190	230	240	230	240	200
	t :	0	20	210	290	0	20	60	140	290
tacken	f_o:	235	250	190	180	245				260
	t :	0	20	85	120	0				100
stigen	f_o:	245	180		195	240	240			195
	t :	0	140		235	0	130			210
stucken	f_o:	245			200	255				255
	t :	0			85	0				85
stegen	f_o:	260	185	180	195	255	270	255		210
	t :	0	120	200	280	0	105	165		270
skotten	f_o:	255	215		200	245	260			245
	t :	0	70		125	0	50			105
hagen	f_o:	230	245	165	185	230			250	190
	t :	0	30	210	290	0			120	275
backen	f_o[1]:	240	230		260	240	235	265		260
	t :	0	25		110	0	20	75		100

[1] measured on narrow-band sonagrams

F_o tracings of <u>tagen</u> and <u>tacken</u>, accent 1 to the left, accent 2 to the right.
Jönköping.

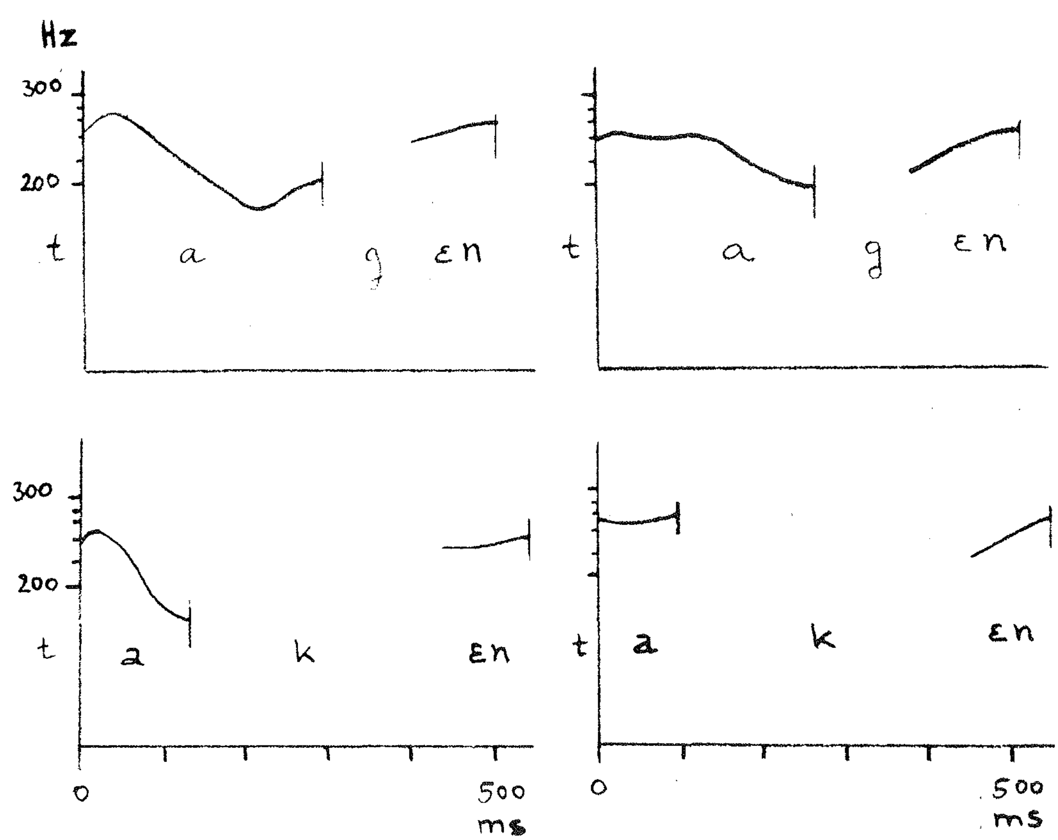


Table 6. JÄMSHÖG

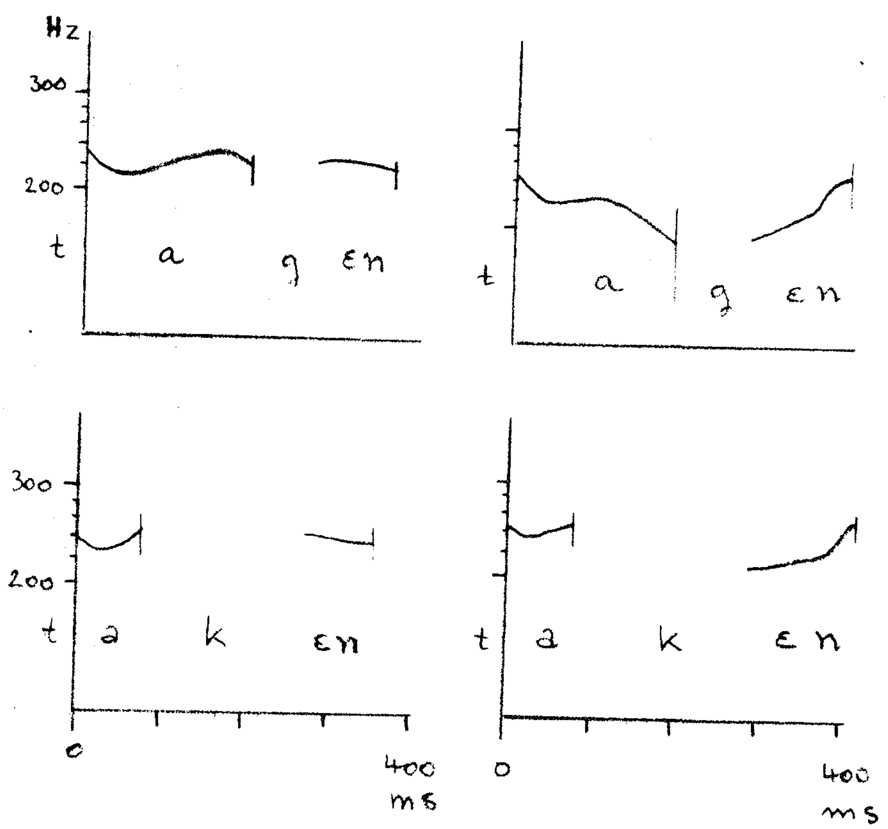
F_o values (in Hz) and durations (in msec).

		Accent 1				Accent 2			
tagen	f_o:	250	210	240	220	250	230	240	190
	t :	0	50	170	210	0	40	100	200
tacken	f_o:	240	230		250	250	230		240
	t :	0	20		80	0	20		80
stigen	f_o:	235	210		240	240	215	235	210
	t :	0	35		200	0	30	115	215
stucken	f_o:	240	220		240	230			235
	t :	0	20		85	0			70
stegen	f_o:	220	210	240	235	230	220	240	220
	t :	0	35	150	210	0	20	115	205
skotten	f_o:	215			255	220	220		240
	t :	0			100	0	15		80
hagen	f_o:	230	210	245	235	225	210	235	200
	t :	0	30	165	200	0	20	110	200
backen	f_o[1]:	210	200		245	210	205		235
	t :	0	25		125	0	25		100

1) measured on narrow-band sonagrams

F_o tracings of tagen and tacken, accent 1 to the left, accent 2 to the right.

Jämshög.



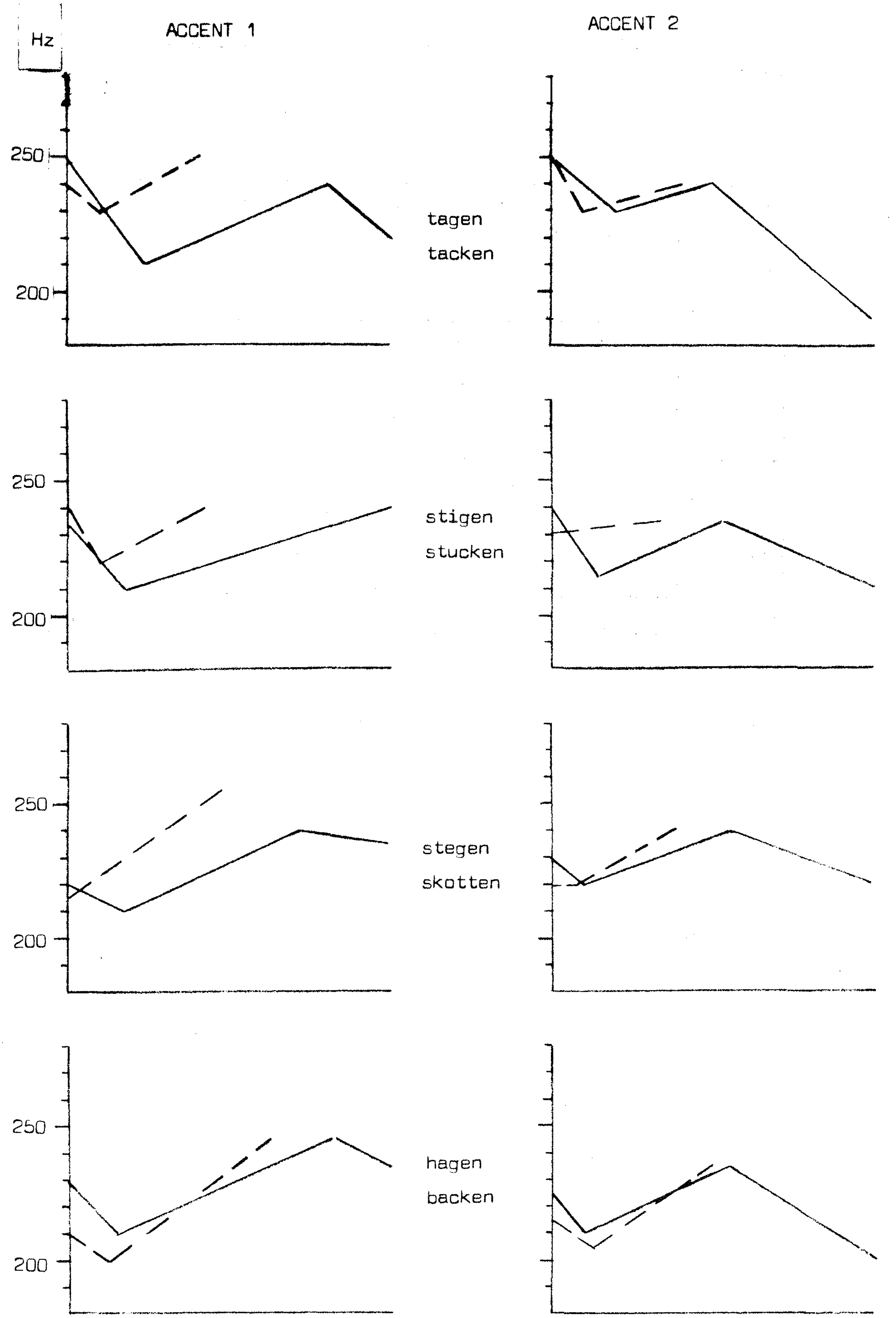


Figure 8. Simplified f_o contours of the first vowel segment. JÄMSHÖG
The duration of the long vowels is normalized.
Dashed line: short vowel

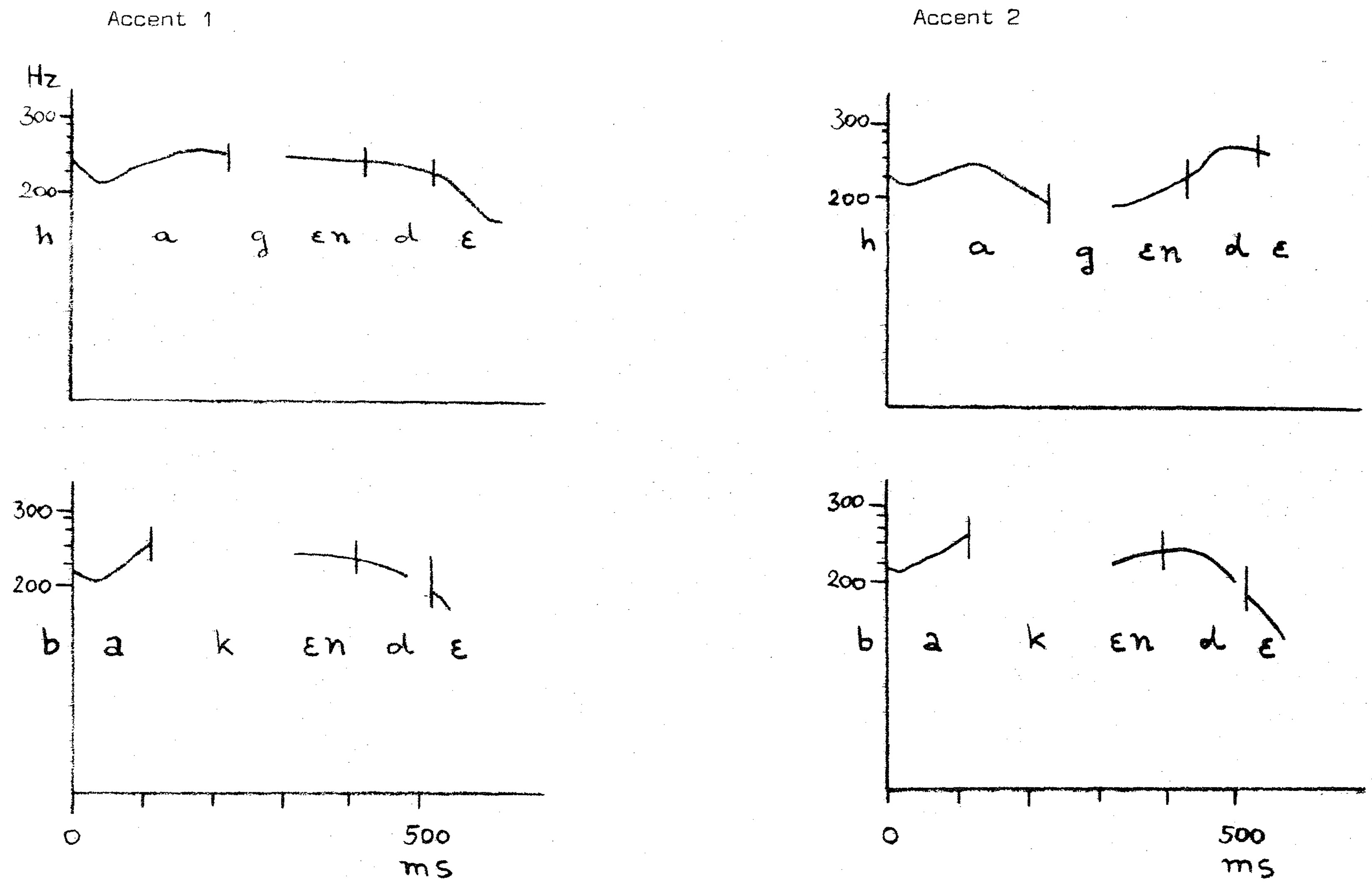


Figure 9. F_o tracings of hagen (above) and backen (below), accent 1 to the left, accent 2 to the right. Jämshög informant

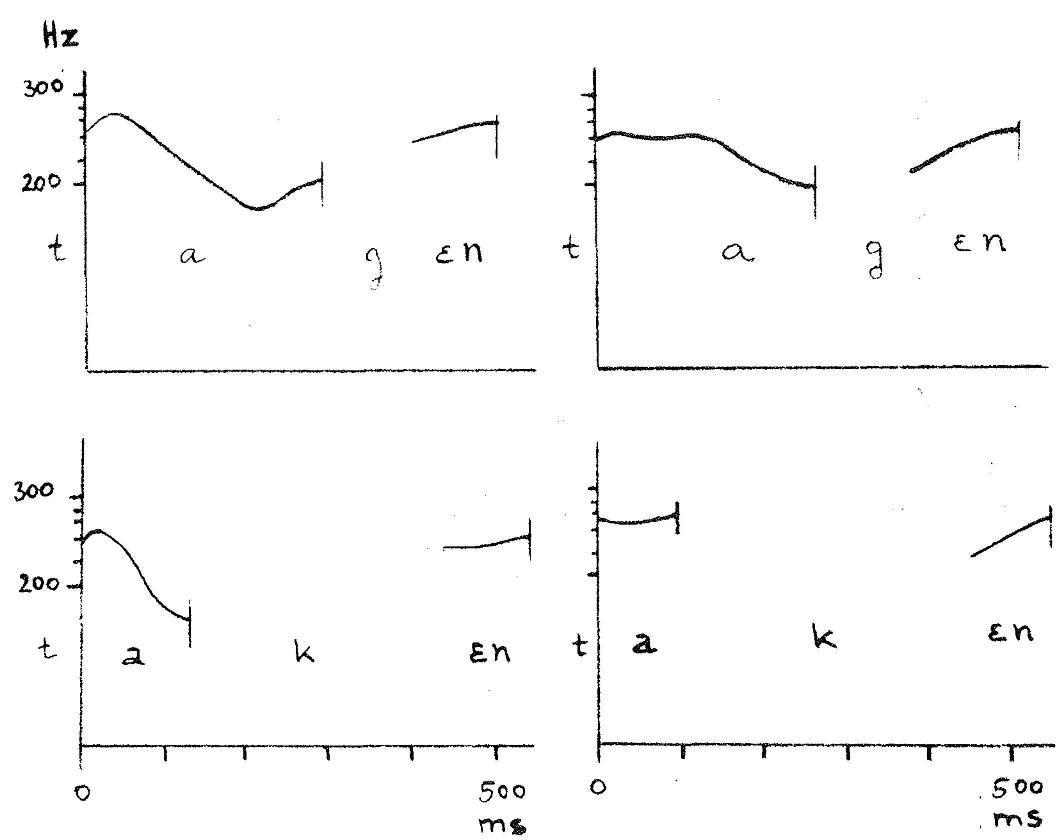
Table 7. JÖNKÖPING

F_o values (in Hz) and durations (in msec).

		Accent 1				Accent 2				
tagen	f_o:	230	245	170	190	230	240	230	240	200
	t :	0	20	210	290	0	20	60	140	290
tacken	f_o:	235	250	190	180	245				260
	t :	0	20	85	120	0				100
stigen	f_o:	245	180		195	240	240			195
	t :	0	140		235	0	130			210
stucken	f_o:	245			200	255				255
	t :	0			85	0				85
stegen	f_o:	260	185	180	195	255	270	255		210
	t :	0	120	200	280	0	105	165		270
skotten	f_o:	255	215		200	245	260			245
	t :	0	70		125	0	50			105
hagen	f_o:	230	245	165	185	230			250	190
	t :	0	30	210	290	0			120	275
backen	f_o[1]:	240	230		260	240	235	265		260
	t :	0	25		110	0	20	75		100

1) measured on narrow-band sonagrams

F_o tracings of tagen and tacken, accent 1 to the left, accent 2 to the right.
Jönköping.



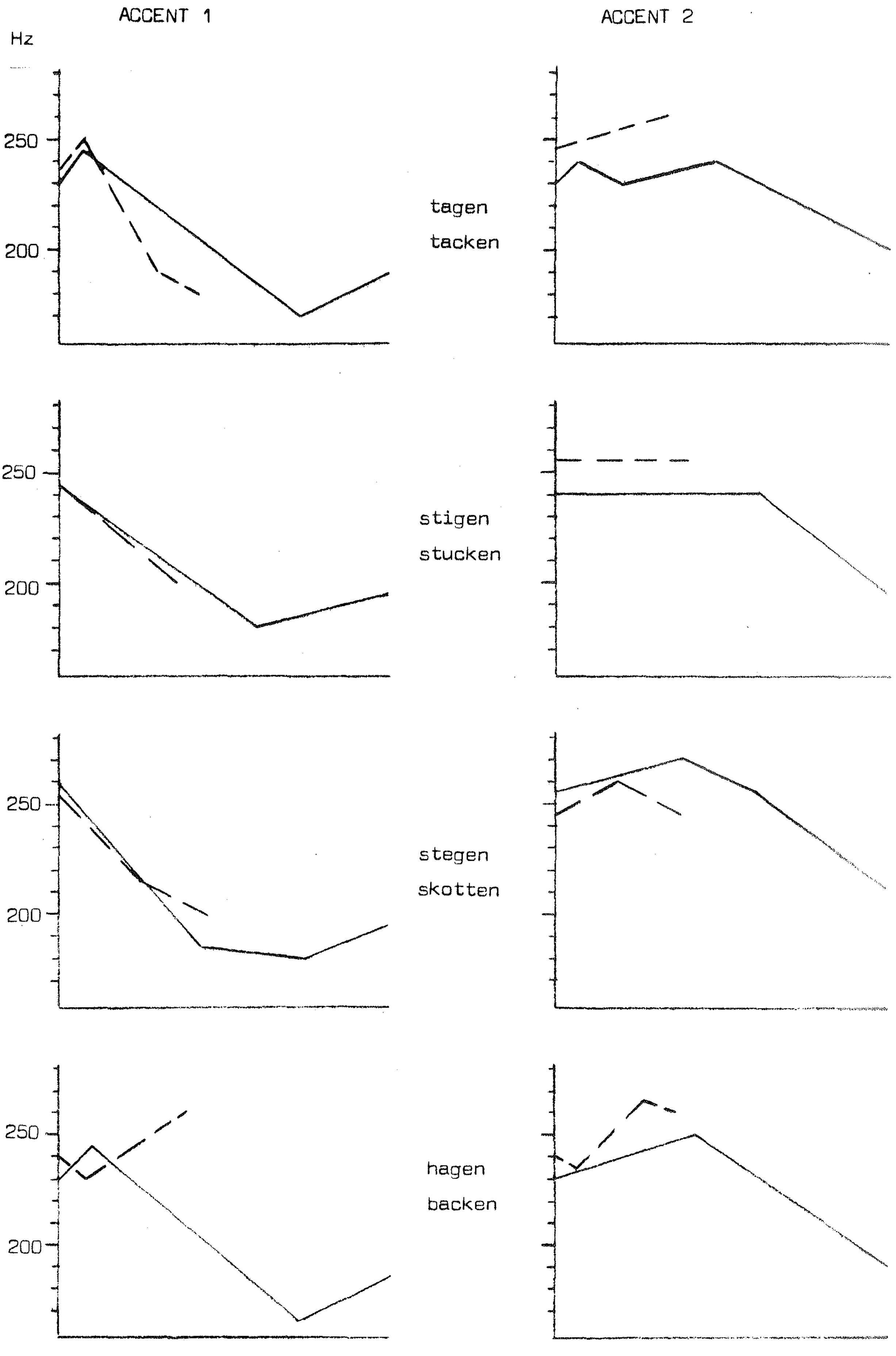


Figure 10. Simplified f_0 contours of the first vowel segment. Jönköping
The duration of the long vowels is normalized.

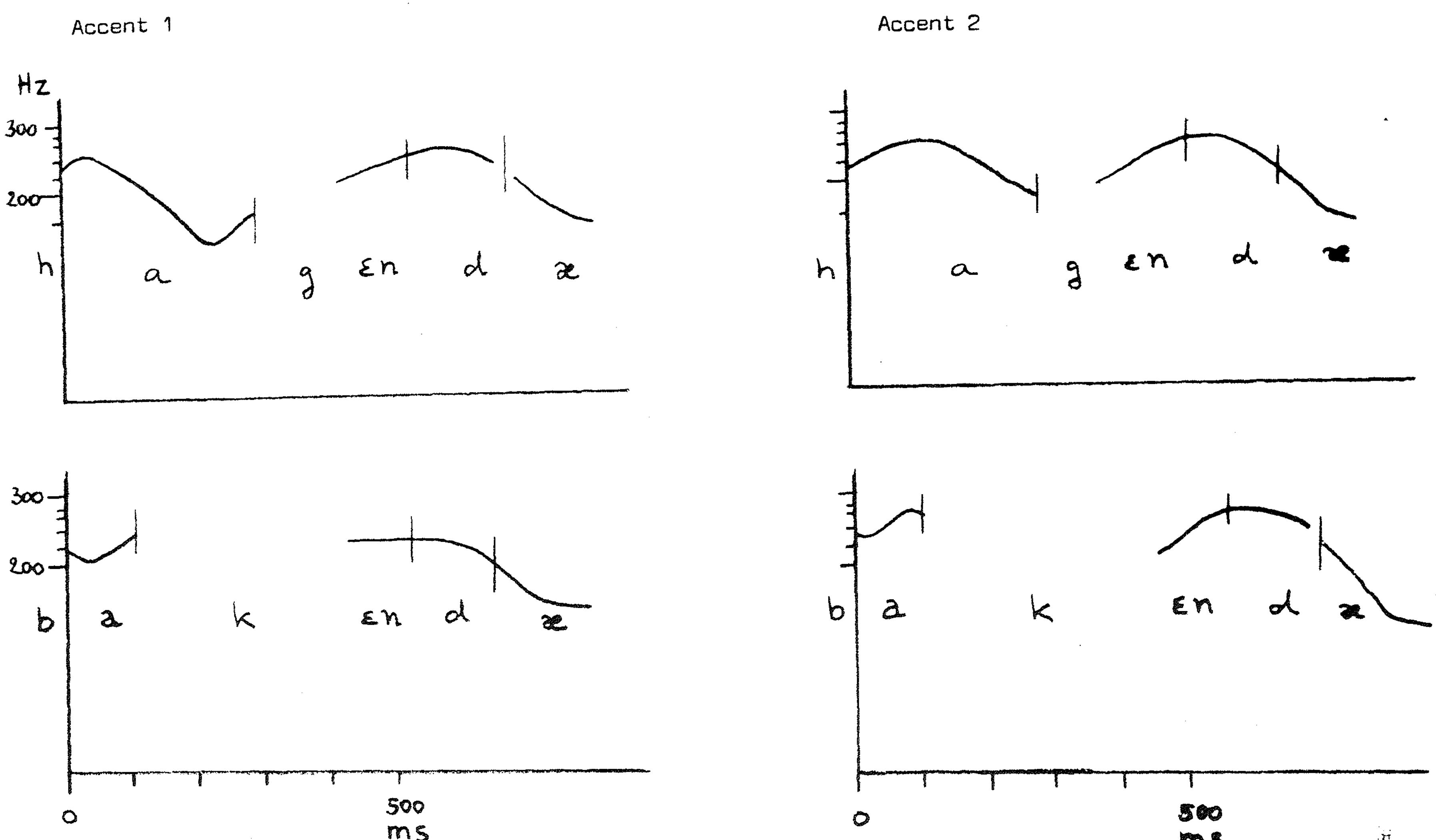


Figure 11. F_o tracings of <u>hagen</u> (above) and <u>backen</u> (below), accent 1 to the left, accent 2 to the right. Jönköping informant

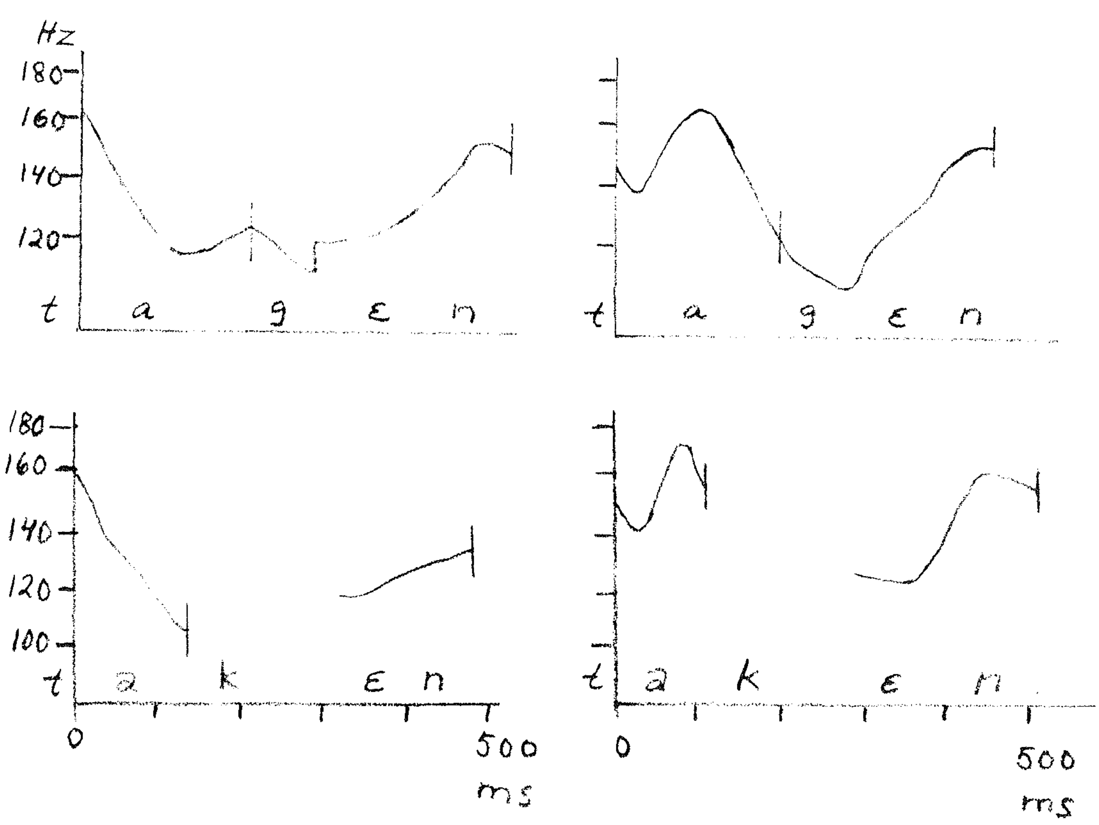
Table 8. RYD

F_0 values (in Hz) and durations (in msec).

		Accent 1			Accent 2			
tagen	f_0:	160	115	120	145	140	160	120
	t :	0	130	230	0	15	100	210
tacken	f_0:	160		115	150	140	170	155
	t :	0		130	0	25	85	110
stigen	f_0:	150		110	145		165	115
	t :	0		185	0		80	195
stucken	f_0:	160		115	150		165	155
	t :	0		100	0		60	90
stegen	f_0:	150	110	115	150	145	155	115
	t :	0	180	225	0	25	95	195
skotten	f_0:	165	135	120	155	145	160	155
	t :	0	50	115	0	35	80	100
hagen	f_0:	155	115	115	135		165	115
	t :	0	120	245	0		110	230
backen	f_0:	155	110	115	130	135	165	155
	t :	0	95	135	0	50	95	120

F_0 tracings of tagen and tacken, accent 1 to the left, accent 2 to the right.

Ryd.



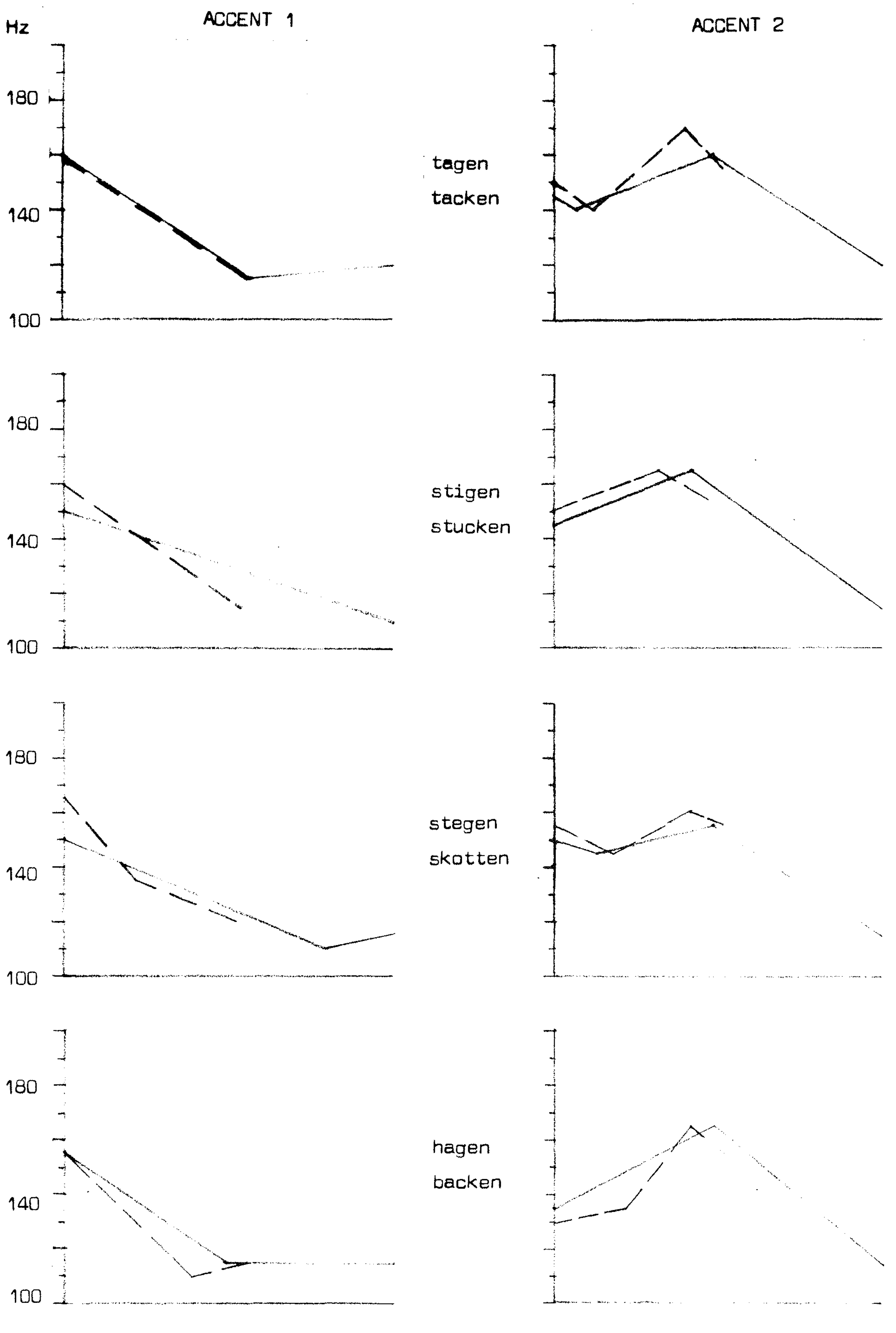


Figure 12. Simplified f_0 contours of the first vowel segment. RYD
The duration of the long vowels is normalized.
Dashed line: short vowel

STOCKHOLM ACCENTS IN FOCUS

Gösta Bruce

INTRODUCTION

The main purpose of this investigation is to determine how the fundamental frequency contours of Swedish sentences vary with accent (accent 1 or acute and accent 2 or grave) and focus in different dialects. In accordance with current terminology, the term focus is used to denote the new information in the sentence, i.e. "the information in the sentence that is assumed by the speaker not to be shared by him and the hearer" (Jackendoff 1972; 230). The old information, non-focus, is "the information in the sentence that is assumed by the speaker to be shared by him and the hearer" (op.cit).

The main result of my present report, which only concerns the Stockholm dialect, is briefly that the f_o patterns of the accents in focus are specific, and not to be found in non-focus position.

PROCEDURE

A typical test sentence in my speech sample has the form of an answer to a question. The question is formulated in three different ways in order to make the speaker choose one of three possible parts of the sentence as the focus and carrier of primary stress. In the following example "One can buy white blouses " focus is indicated by capital letters:

Question 1:	What white things can one buy?
Answer 1:	One can buy white BLOUSES.
Question 2:	What blouses can one buy?
Answer 2:	One can buy WHITE blouses.
Question 3:	What can one do with white blouses?
Answer 3:	One can BUY white blouses.

There are three sets of test sentences in my material (A, B and C).

$$\text{(A)}\quad \text{Man kan}\left\{\begin{matrix}\text{l\`{ä}mna}^{1}\\ \text{an\'{a}mma}\end{matrix}\right\}\ \text{några}\ \left\{\begin{matrix}\text{l\`{å}nga}^{2}\\ \text{l\'{ä}ngre}\end{matrix}\right\}\left\{\begin{matrix}\text{n\`{u}nnor}^{3}\\ \text{n\'{u}mmer}\end{matrix}\right\}.$$

$$\left(\text{One can}\left\{\begin{matrix}\text{leave}\\ \text{accept}\end{matrix}\right\}\ \text{some}\ \left\{\begin{matrix}\text{long}\\ \text{longer}\end{matrix}\right\}\left\{\begin{matrix}\text{nuns}\\ \text{numbers}\end{matrix}\right\}\right).$$

The numbers denote three possible focus locations. In each position there are words with accent 1 (´) or accent 2 (`). These words are disyllabic and stressed on the first syllable with the exception of the acute trisyllabic verb anámma, which is stressed on the second syllable. The first and second positions are separated by three unstressed syllables, the second and third by one.

(B) Man kan làmna[1] nåra {lànga[2] / lángre} {lámm[3] / laméller / bumeránger} .

(One can leave some {long / longer} {lambs / laminas / boomerangs} .)

The second set has the same basic pattern as the first. The accent 2 verb only is used in the first position, and in the third position the number of syllables of the accent 1 noun is varied.

(C) Man kan lämna[1] nåra {Lànga lamm[2] / Làngalameller / Lànganummer / Làngabumeranger} .

One can leave some {lambs / laminas / numbers / boomerangs} from Lång.

In the third set the noun phrase - adjective and noun - occupying the second and third positions of the first and second sets is replaced by a compound noun (accent 2) whose first component has accent 2 when standing alone and whose second component has accent 1 when standing alone. The compound word has the same segmental structure as the corresponding noun phrase in the second set. The first component of the compound is to be interpreted as a place name.

Words with sonorant consonants have been chosen as far as possible to provide continuous, undisturbed f_o curves of the utterances. Vowels with approximately the same degree of opening - non-high vowels - have been used to avoid differences in intrinsic f_o (cf. Lehiste and Peterson 1961), which might complicate the interpretation of the contours. Moreover all the vowels are phonologically short.

The speech material, which consists of 50 sentences, was recorded in the sound studio of the institute according to routine procedures. The speaker was a female student of phonetics from Stockholm. Each test sentence appeared three times in three different order arrangements.

The recording was analysed with the aid of a pitch extracting device. The figures 1–6 are examples of the extracted f_o curves. Intensity curves and oscillograms were used for segmentation.

RESULTS AND DISCUSSION

Accents in focus

Sentence final position. The observations made here are based on the f_o tracings presented in the figures. One such tracing is only one instance of a test sentence. The other repetitions of a test sentence look about same.

The upper part of figure 1 shows the f_o patterns used by my informant for disyllabic words in focus in sentence final position. These f_o patterns are in agreement with those shown by earlier investigations (Meyer 1937, Öhman 1965 and 1967, Gårding 1967, Alstermark and Eriksson 1971, Gårding and Lindblad 1973). The main characteristics are summarized below.

ACCENT 2: nunnor [ˋnʉn:ɔr] = $C_1V_1C_2V_2C_3$, "nuns".
Rise in C_1, f_o maximum at the beginning of V_1, fall, f_o minimum at V_1C_2, rise, f_o maximum at the beginning of V_2, fall, f_o minimum at the end of V_2

ACCENT 1: nummer [ˊnʉm:ɛr] = $C_1V_1C_2V_2C_3$, "numbers".
f_o minimum at the beginning of V_1, rise, f_o maximum in C_2, fall at C_2V_2, f_o minimum at the end of V_2.

Non-final position. The mid part of figure 1 shows the corresponding f_o patterns in a non-final position: làmna, anámma, lànga, and långre. The accent patterns of the two positions are similar except for the final part. In sentence final position for both accent 1 and accent 2 words there is an f_o fall (see above), which is missing in non-final position. Instead there is a fall in the stressed syllable of the following stressed word (fig. 1, lower part). This fall seems to correspond to the final fall of the accented word in sentence final position. When there are several unstressed syllables between a word in focus and the following stressed syllable, the peak between the focus rise and the fall of the following stressed syllable becomes a plateau (fig. 1). Moreover the length of the plateau is determined by the accent of the word in focus. An accent 1 word in focus will have a longer plateau, as can be seen in the lower part of figure 1.

Accents out of focus

As can be seen in figure 2 the f_0 pattern of the accent 2 word långa in non-focus position is partly similar to the corresponding focus pattern. We find an f_0 maximum at the beginning of the stressed vowel and a fall. In non-focus position, however, the f_0 minimum is not reached until C_2V_2 (compare compounds below). In addition the rise of the second syllable of the word in focus is totally absent out of focus. It seems as if the rise - in focus position - starts before the target f_0 minimum value is reached. In spite of greater f_0 expansion (higher starting point) in post-focal position compared to pre-focal, the same f_0 minimum value is reached.

The accent 1 word lángre (fig. 2) shows quite a different pattern in non-focus positions than in focus position. In both pre-focal and post-focal positions we find a fall in C_1 to the f_0 minimum, which is reached at the beginning of the stressed vowel. The fall is much larger in post-focal positions, having started from a much higher frequency. The low level is maintained throughout the word. The rise and the following peak, features which are regarded as typical of accent 1, are totally absent. We notice also that the f_0 minimum value here is lower than in focus positions. The non-focus value may - also for accent 1 - be interpreted as the target, which is not reached in the focus word, the rise preventing the target f_0 minimum from being reached.

Corresponding differences between focus and non-focus position are found also for the pairs l̀ämna/anámma and nùnnor/númmer (fig. 2). In the last pair the difference in frequency range in the two positions after focus is large, but the pattern remains the same, i.e. the timing of the fall and the reaching of the f_0 minimum in both the accent 1 and the accent 2 word is approximately the same in the two positions. Although the fall is about 75 Hz in the first position after focus and only 10-15 Hz in the second position after focus, the same f_0 minimum value is obtained in both cases.

The obvious difference between focus and non-focus position is the f_0 rise and the f_0 peak, which is present only in the focus words independently of accent. It is evident from the lower part of figure 2 that it is the same tonal phenomenon that is found in both the accent 1 and the accent 2 words. The accent 1 and the accent 2 words in focus can be decomposed - regarding their f_o manifestations - into one accent-dependent part which is different for the two accents, and one accent-independent part (focus part) which is the same for both accents. The timing of the focus part is different, however, for the two accents.

The f_o data presented here show that accent 1 - like accent 2 - has a stable f_o pattern of its own even out of focus. According to my analysis accent 1 is a "true" word accent on a par with accent 2. It is not merely stress and sentence intonation as has been proposed (see e.g. Haugen 1967, Elert 1970, Gårding 1970).

The accent distinction

In figure 3 the f_o manifestations of accent 1 and accent 2 in non-focus position are compared. It appears that the accent distinction - as far as the f_o patterns are concerned - is retained in non-focus position, both before and after focus (for another interpretation see Gårding 1967). The main difference is in the timing of the fall in connection with the stressed syllable. The fall of an accent 1 word comes in the prevocalic consonant of the stressed syllable, whereas in an accent 2 word the fall does not start until the stressed vowel has begun. The timing difference, which is about 100 msec, results in an f_o maximum for accent 2 and an f_o minimum for accent 1 at the beginning of the stressed vowel. The frequency difference between accent 1 and accent 2 is greatest in a position directly after focus and least in the second position after focus, where we find a very slight difference. For the positions before focus, the difference is intermediate. It can be observed that the fall is often steeper in the accent 1 words.

In figure 5 we can observe that the f_o peak value of the pre-tonic syllable of <u>l´ängre</u>, which belongs to the preceding word, is approximately the same as the corresponding value of the stressed syllable of <u>l`ånga</u>. It seems to be the same peak for both accents; there is only a difference in timing.

The difference between position before and after focus should be noted. In position before focus the f_o peak - in the stressed vowel of accent 2 and in the pretonic vowel of the stressed syllable of accent 1 - is normally reached by a directly preceding rise, while in position after focus the f_0 maximum is the final part of a tonal plateau and has no directly preceding rise.

Although there is a systematic f_0 difference between accent 1 and accent 2 in the non-focus positions it should be tested, whether the accent distinction is not only an acoustical but also a perceptual reality. According to phonological analyses of Swedish the accent distinction is neutralized in non-primary stress position.

Accents in compounds and in two word phrases

So far we have considered the influence of focus position on f_0 in simple disyllabic words. In this section compounds and two-word phrases will be treated. In the Stockholm dialect compounds as a rule have accent 2. Secondary stress is attributed to the stressed syllable of the last component of the compound. Like simple accent 2 words compounds in focus have two tonal peaks: the first peak occurs in the stressed syllable of the first element, and the second one is tied to the syllable carrying secondary stress. This means that in long compounds like Långalameller and Långabumeranger (fig. 6) the peaks are well separated.

In post-focal position, compounds, like simple accent 2 words, have no second peak: when the fall from the peak of the first component has been completed, f_0 remains low throughout the word (fig. 6).

As is evident from figure 3 one difference between LÅNGALAMM (compound noun in focus) and långa LAMM (noun phrase with noun in focus) is a peak in the second syllable of långa (in the phrase), which does not occur in the corresponding compound (compare in fig. 3 also LÅNGANUMMER - långa NUMMER). At first glance one might assume that this peak is part of accent 2. The peak is, however, not present in the second syllable of långa in the examples långa LAMELLER and långa BUMERANGER, but occurs later in the phrase (fig. 4). The examples of figures 3 and 4 show that the peak occurs in the syllable preceding the stressed syllable of the accent 1 word in focus. Therefore the peak can be assumed to be tied to the following accent 1 and not to the preceding accent 2.

Comparing phrase pairs like långa LÁMM – lángre LÁMM and långa NÚMMER – lángre NÚMMER (fig. 5) we find for both phrase types a peak in the pretonic syllable of the accent 1 word in focus. In a phrase like lángre NÚMMER this peak cannot possibly be tied to any accent 2. As mentioned above we find a peak in the pretonic syllable of an accent 1 word even in non-focus position (fig. 5). So it is clear that this peak is somehow tied to accent 1, although it can appear in a word preceding the accented one.

It is obvious from the data presented here that the f_o pattern of a compound like LÅNGANUMMER is not merely a combination of one accent 2 part and one accent 1 part, as has been suggested (Öhman 1965, Alstermark and Eriksson 1971). The two word phrase långa NÚMMER contains a peak, which is not present in the compound, in the syllable preceding the last stressed syllable (compare also duration differences between the stressed syllable of accent 1 and the secondary stress syllable, Lindblom and Rapp 1973).

That the f_o pattern of compounds is distinct from that of two-word phrases is also evident from the lower part of figure 4. In non-focus position the two-word phrase contains a tonal modification in connection with the last stressed syllable. The corresponding compound has no such tonal modification, as has been mentioned above.

An additional observation is that in the compounds of figure 6 the f_o minimum is reached at C_2V_2; i.e. later than in the simple disyllabic accent 2 words in focus (fig. 1), but at the same point as in the non-focus accent 2 words (fig. 2). Whether or not the target f_o minimum is reached evidently does not have anything to do with the compound/simple word distinction, but rather is dependent on the timing of the rise. In these compounds the rise towards the second peak does not start in conjunction with the first peak, so the fall is not interrupted but is allowed to reach its target minimum. As can also be seen in figure 6, there is no difference in f_o minimum after the fall from the peak of the stressed syllable between focus and non-focus compounds.

SUMMARY

We can summarize the main findings of this investigation of the Stockholm accents as follows:

The main result is briefly that the f_0 patterns ofthe accents in focus are specific, and not found in non-focus position.

Accents in focus. Accent 2 has two tonal peaks and accent 1 one peak. In sentence final position there is for both accents an f_0 fall in the last syllable, which is missing in non-final position. In non-final position there is, however, a corresponding f_0 fall in connection with the stressed syllable of the following stressed word.

Accents out of focus. Words out of focus have a different f_0 pattern from words in focus: the tonal peak that is found in focus - i.e. the second peak of accent 2 words and the only peak of accent 1 words - is missing in non-focus position.

The accent distinction. The difference in f_0 manifestation for the two accents is retained even out of focus: accent 1 is distinguished from accent 2 by a difference in timing of the f_0 fall tied to the stressed syllable; thus at the beginning of the stressed vowel accent 1 has an f_0 minimum and accent 2 an f_0 peak.

Accents in compounds and two-word phrases. Out of focus compounds, like simple accent 2 words, have no second peak. A compound like L̀ÅNGANUMMER does not have entirely the same f_0 pattern as a two-word phrase consisting of one grave and one acute word: l̀ånga NÚMMER. The two-word phrase contains a tonal modification - in both focus and non-focus position - which is not present in the compound.

BIBLIOGRAPHY

Alstermark M. and Y. Eriksson. 1971. Swedish word accent as a function of word length. STL-QPSR 1:1-13

Elert C.-C. 1970. Ljud och ord i svenskan. Stockholm: Almqvist & Wiksell

Gårding E. 1967. Prosodiska drag i spontant och uppläst tal; i Holm G. (utg.): Svenskt talspråk. Stockholm: Almqvist & Wiksell, 40-85

Gårding E. 1970. Word tones and larynx muscles. WP 3. Phonetics Laboratory, Lund University

Gårding E. and P. Lindblad. 1973. Constancy and variation in Swedish word accent patterns. WP 7. Phonetics Laboratory, Lund University

Haugen E. 1967. On the rules of Norwegian tonality. Language 43, 185-202

Jackendoff R. 1972. Semantic interpretation in generative grammar. Cambridge, Massachusetts, and Londong: The MIT Press

Lehiste I. and G. Peterson. 1961. Some basic considerations in the analysis of intonation. JASA Vol. 33. No. 4:419-25

Lindblom B. and K. Rapp. 1973. Some temporal regularities of spoken Swedish. PILUS No. 21

Meyer E.A. 1937. Die Intonation im Schwedischen I: die Sveamundarten. Studies Scand. Philol. No. 10. University of Stockholm

Öhman S. 1965. On the coordination of articulatory and phonatory activity in the production of Swedish tonal accents. STL-QPSR II:14-19

Öhman S. 1967. Word and sentence intonation: A quantitative model. STL-QPSR II-III:2054

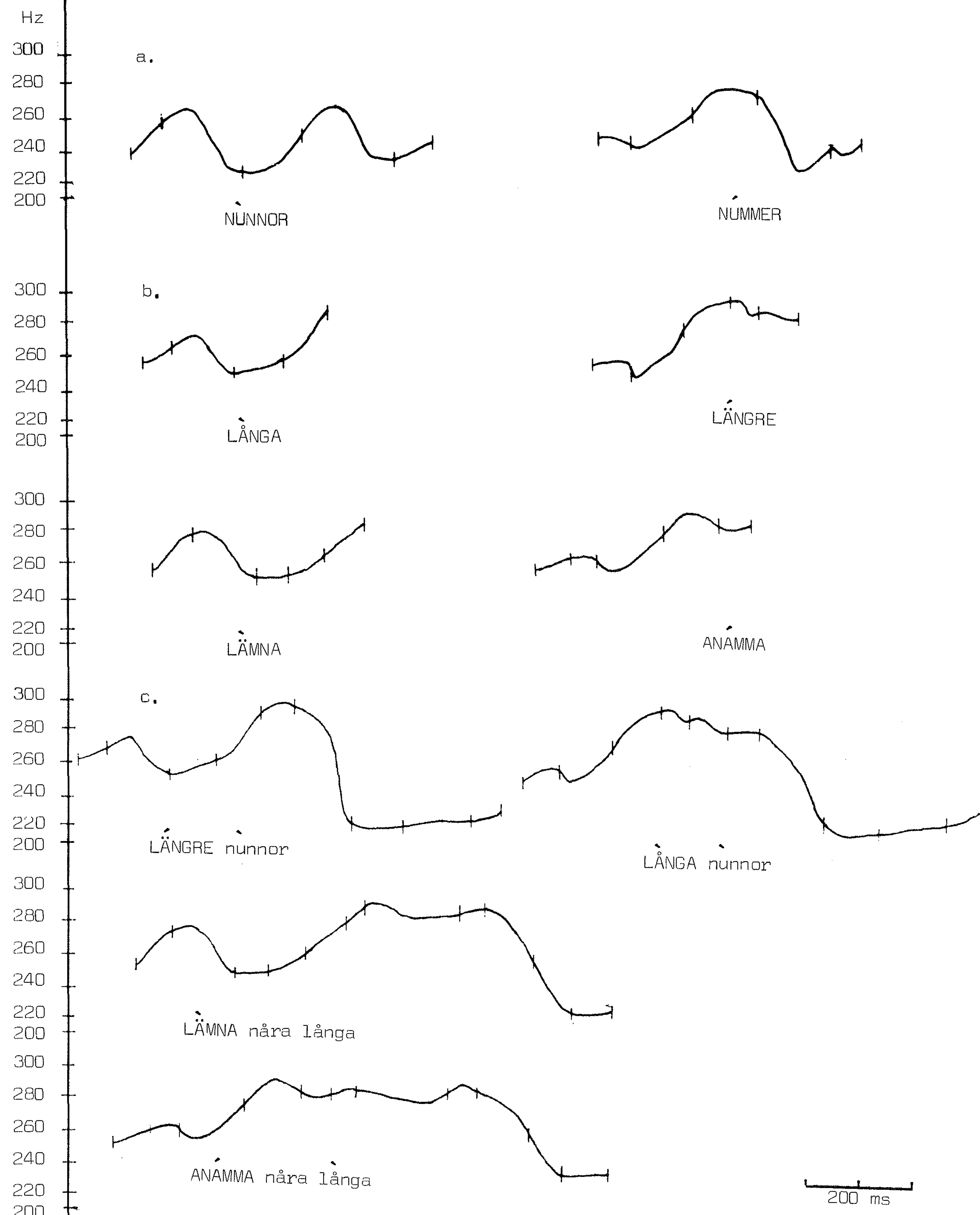


Fig. 1. Accents in focus. (a) Words in sentence final position.
(b) Words in non-final position.
(c) Phrases in non-final position. First word in focus.

Focus is indicated by capital letters.

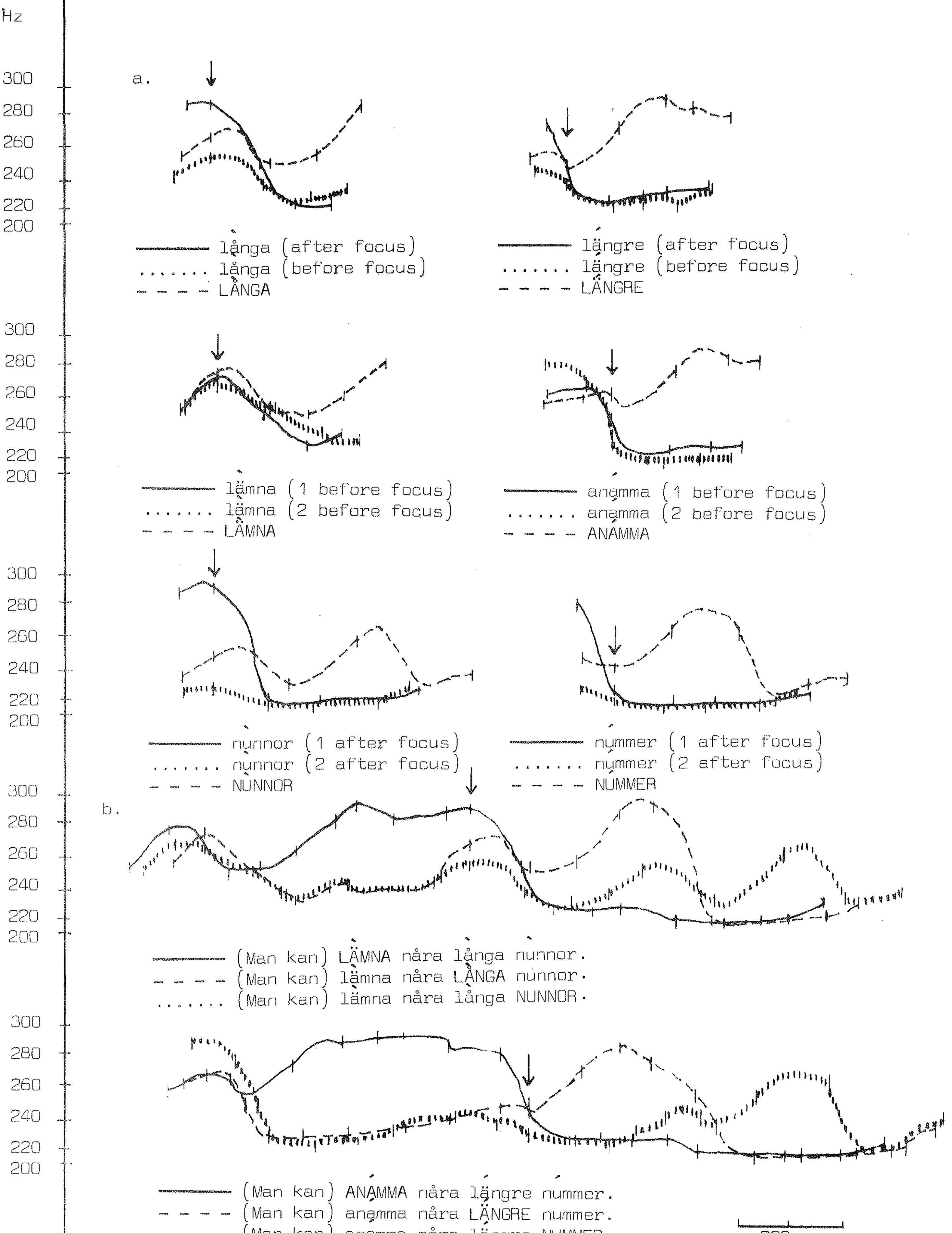


Fig 2. Accents out of focus with corresponding words in focus as a reference. (a) Words. (b) Sentences with different focus locations. ↓ = line-up point

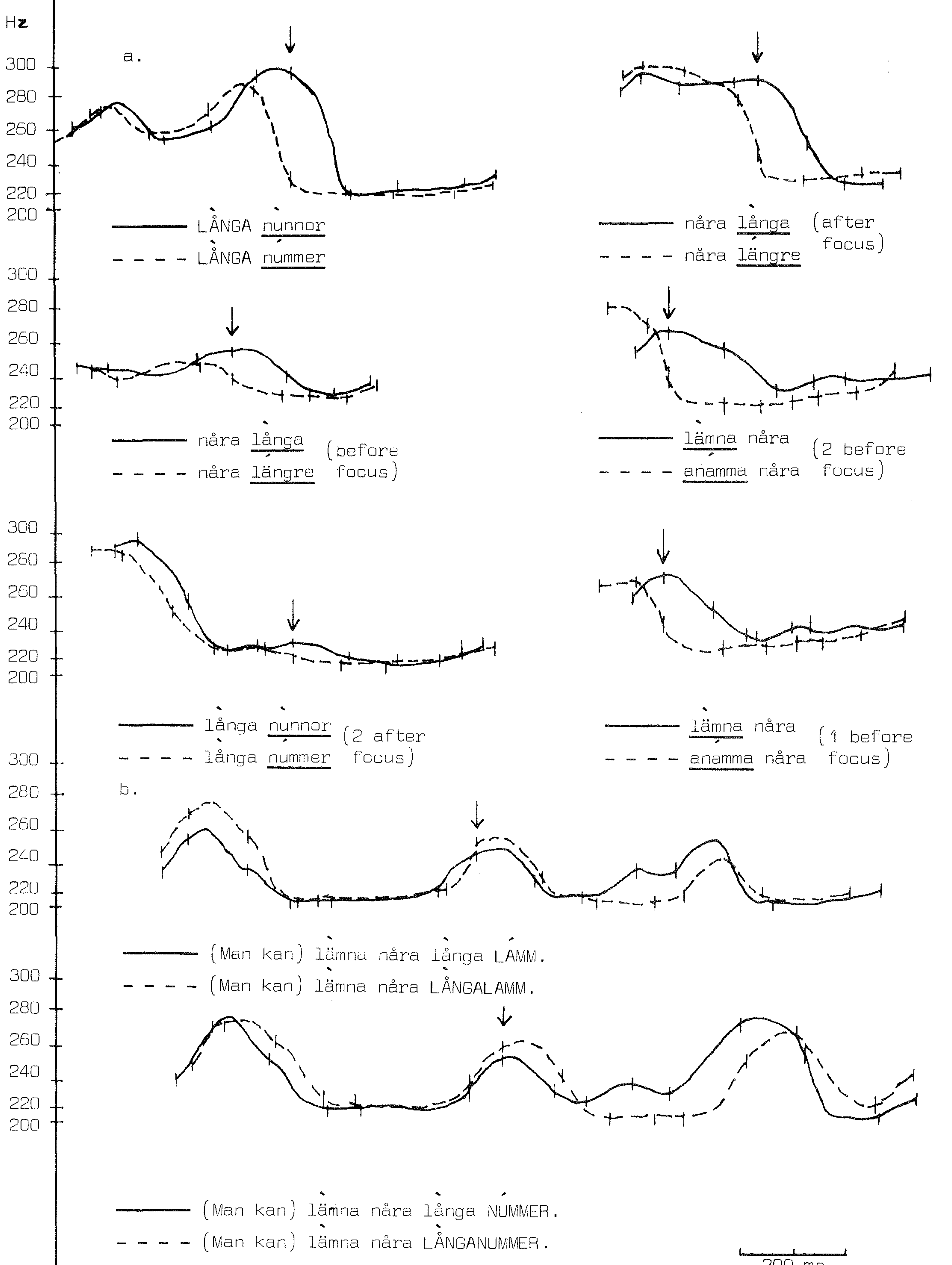


Fig 3. (a) The accent manifestations out of focus.

(b) Sentences with two-word phrases or compound words. Last word in focus.

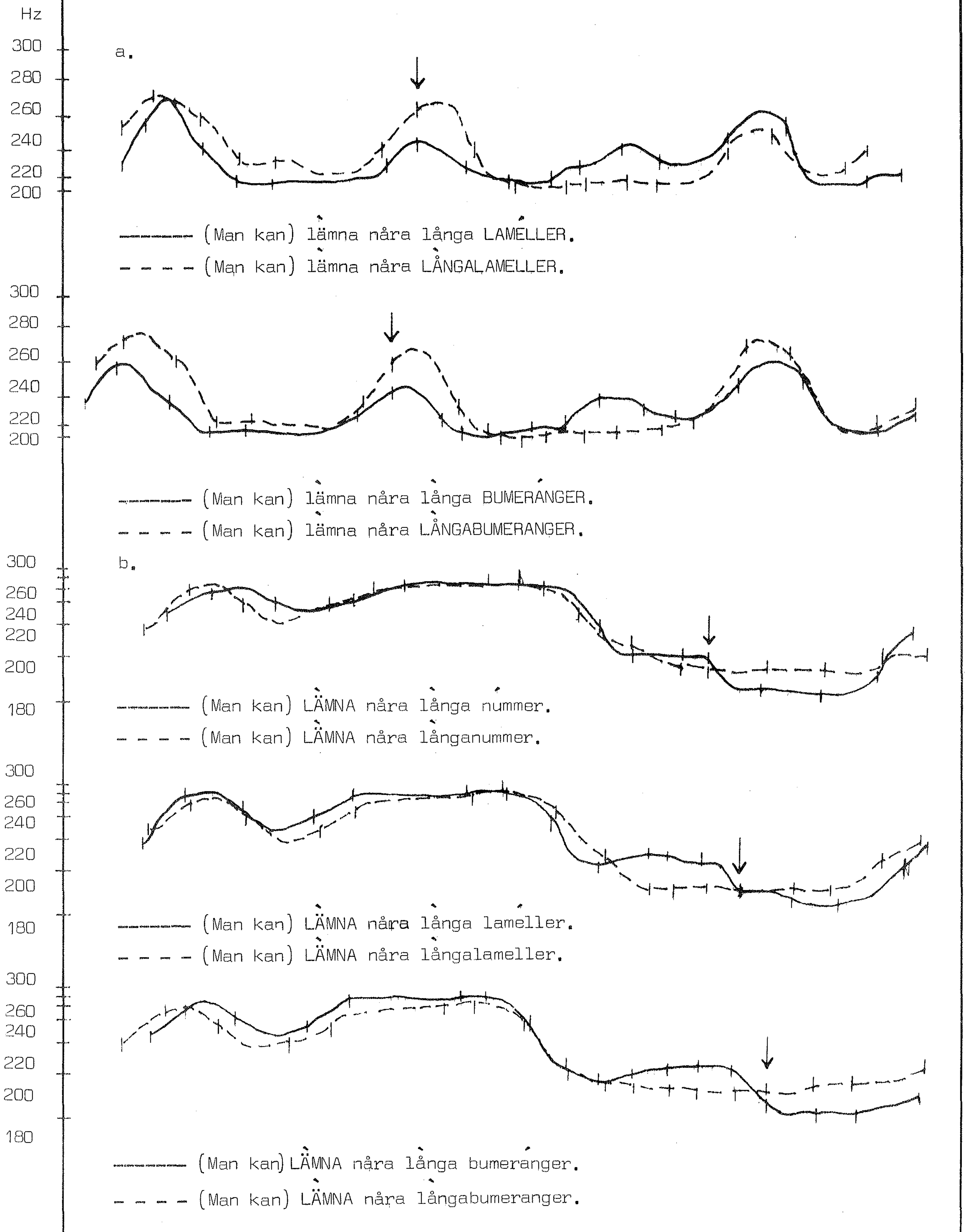


Fig. 4. Sentences with two-word phrases or compound words.
(a) Last word in focus. (b) Out of focus

200 ms

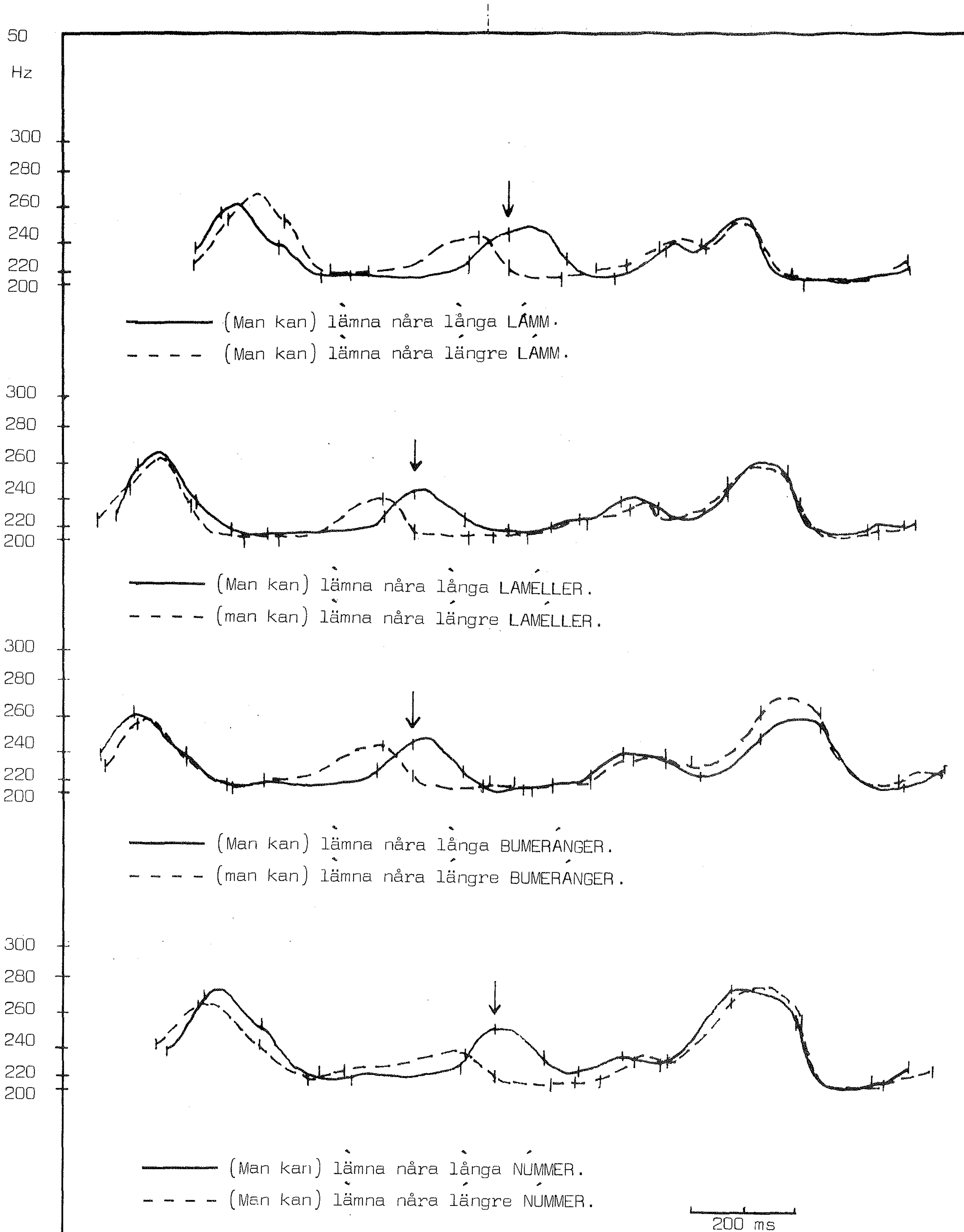


Fig 5. Sentences with two-word phrases: accent 1 + accent 1 or accent 2 + accent 1. Last word in focus.

LARYNGEAL CONTROL OF SWEDISH WORD ACCENTS

E. Gårding, O. Fujimura, H. Hirose and Z. Simada

INTRODUCTION

The Swedish accents have attracted the interest of a great number of linguists and phoneticians.

Kerstin Hadding, to whom we dedicate this paper, was one of the first to analyse them by means of the spectrograph. She was interested in contextual and dialectal variation and tried to find invariant dialect independent features in the accent manifestations. The results were published in her doctoral thesis, Acoustico-phonetic studies in the intonation of Southern Swedish (1961). Recently electromyography (EMG) brought about a renewed interest in the physiology of speech. A pilot study of laryngeal muscle activity in connection with the production of accents was carried out by Öhman and his collaborators (1967).

Along similar lines we made a number of EMG recordings of Swedish accents at the Research Institute of Logopedics and Phoniatrics of Tokyo University. A preliminary report was published a year later (1970 ref. 11). Since then the material has been used to discuss a model for intonation (ref. 9 and 25) and to elucidate the nature of boundary signals (ref. 10).

The aim of our experiment is to study the production of the accents as it is reflected in the speech wave and in the EMG signals from some selected laryngeal muscles. An EMG investigation of how the accents are related to laryngeal muscle activity is interesting not only as a complement to the acoustic picture. A study of the fast changes of pitch involved in the accents and the delicate control that is necessary to achieve this effect is likely to throw some light on the laryngeal mechanisms that regulate pitch in general.

EXPERIMENTAL PROCEDURES

Subjects and speech material

Our test material consisted of some 20 sentences each containing a test word that had one of the two accents associated with the phonetic condi-

Table 1. Test sequences and variables.

Test sequence	Syllabified notation	VARIABLES: Dialect Speaker	Accent	Contrastive Stress	Whisper	Speed
de va mo:nen han sa	$C_1V_1-C_2V_2C_3$	L E	+	+	+	
– pa:men ja sa	$C_1V_1-C_2V_2C_3$	L E	+			
– mamma –	$C_1V_1C_{21}-C_{22}V_2$	L E	+			+
– ma:ma –	$C_1V_1-C_2V_2$	L	+			+
– be:ben –	$C_1V_1-C_2V_2C_3$	L	+		+	+
– bebben –	$C_1V_1C_{21}-C_{22}V_2C_3$	L	+			+

The first line of the table means that the corresponding test sequence was spoken by both L and E with the two accents (acute [´mo:nen] and grave [`mo:nen]) under normal stress, and contrastive stress [˝mo:nen], [̏mo:nen], and that the accent contrast also was produced in whisper. The rest of the table reads accordingly.

tions listed in Table 1. The test words were embedded in semantically neutral carrier sentences in a rising-falling statement intonation and they were uttered by one female speaker (E) of a southern dialect (Skåne) and a male speaker (L) of standard central Swedish. The dialects are known to have different manifestations of word accents. Each test sentence was uttered fifteen times in repetition.

Electromyography

We selected the vocalis (VOC) and cricothyroid (CT) muscles as the targets of our EMG study since they are known to be active in laryngeal control of voicing and pitch changes (e.g. Hirose et al. 1970). In addition the sternohyoid (SH) was examined. This muscle has been reported to be active for pitch-lowering (Ohala, 1970).

The EMG data were obtained by means of double-ended hooked-wire electrodes which were inserted through the skin and other tissues of the neck. (For a full description of the technique see Hirose et al. 1970.)

The electromyographic signals were amplified by high gain DC preamplifiers. Three EMG signals from different muscles and the speech signal were simultaneously recorded by a four-channel FM magnetic tape recorder. The recorded signals were fed to an off line PDP-9 computer via an AD converter for processing. In this process the EMG signals were sampled every 250 microseconds and digitized into 6-bit levels. The absolute values were taken for the samples and these were integrated over a range of 10 msec by use of a running window. (For more detail see Simada and Hirose 1970.) The smoothed signals obtained for 10 selected utterances were summed at every corresponding time sample. The sampling times were determined in relation to a time moment (line-up point), representing a selected speech event, e.g. the explosion of [m], in the test word.

RESULTS

Each of the EMG curves in the Figures 1-2, 5-6, and 9-10 represents an average of 10 utterances. They have been obtained from the following muscles, from top to bottom in each figure: the VOC, the CT, and the SH. The fundamental frequency curve shown as the lowest trace in the figure is a hand made average of three of the test utterances. The line-up point on the time axis for the summation process was selected at the voice onset

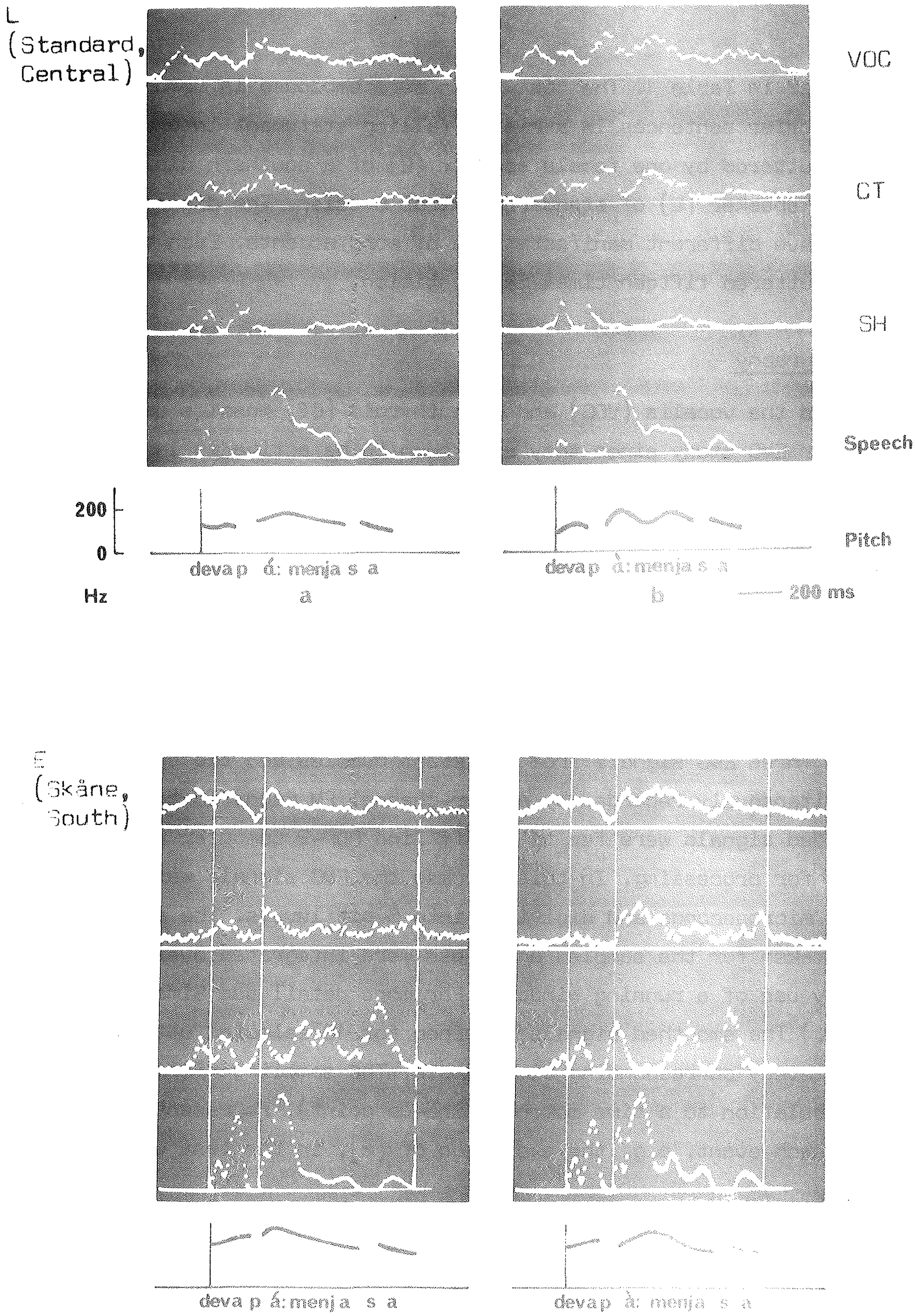


Figure 1. The averaged EMG signals of the vocalis, the crico-thyroid and the sternohyoid muscles, the speech signals and the pitch contour for the utterances a) de va ´pa:men ja sa b) de va `pa:men ja sa. Above Speaker L (Standard, Central) and below Speaker E (Skåne, South).

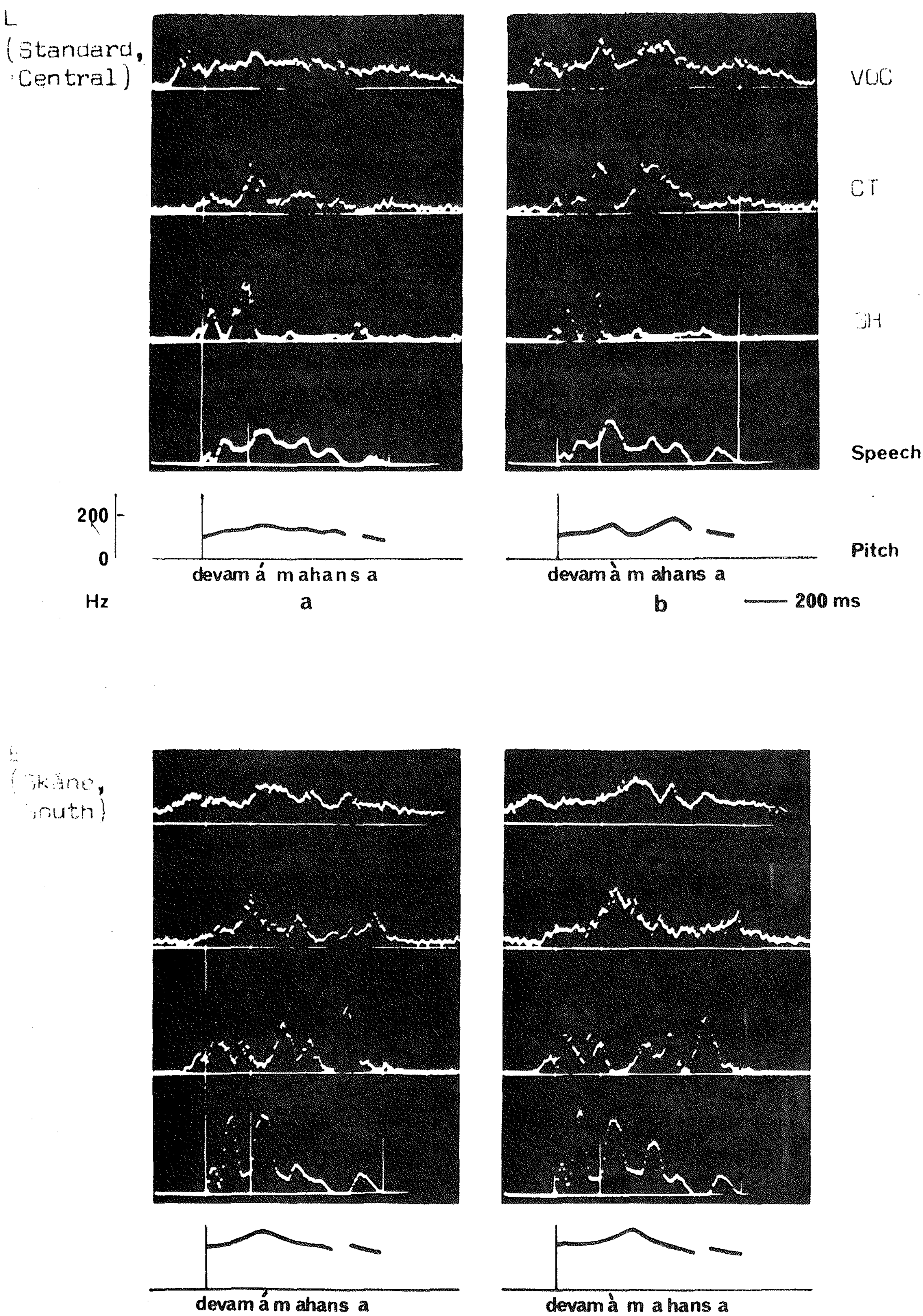


Figure 2. The same as in Figure 1 for the utterances
a) de va ´mamma han sa
b) de va `mamma han sa.

after initial /p/ or at the release of initial /m/. We shall discuss our findings in the order given by the variables in Table 1, i.e. accent, contrastive stress, whisper and rate of speech.

1. Acute versus grave under primary stress

Standard Central Swedish. Speaker L

Acoustic data. As can be expected in this dialect, the acute accent in the given prosodic circumstances (primary stress and statement intonation) is characterized by a late fundamental frequency maximum (pitch peak) in the stressed syllable. The grave accent has two peaks, an early peak in the stressed syllable followed by a late peak in the second syllable (Figures 1 and 2). The pitch value of the peaks does not vary much with accent and location.

For the acute accent the pitch curve reaches its peak at the end of V_1 regardless of the structure of the syllable. The rise starts towards the end of the initial consonant when this is a bilabial nasal (Figure 2). After the stressed syllable there is an overall slow and smooth decrease of pitch during the rest of the test word until the beginning of the first word of the frame when the curve makes a more rapid fall.

For the grave accent the peak occurs at the end of the first third of a long vowel and at the middle of a short one. This peak is reached by a rapid rise starting in the preceding consonant when the vowel is short. After the peak the curve falls abruptly to a minimum at the end of the syllable. With a long vowel the curve has a flat ending after the minimum has been reached. With a short vowel the fall is interrupted by the consonant which now contains the minimum. At the onset of the second syllable the pitch starts rising again. The second peak is reached towards the end of V_2 (some 150 msec from the beginning of the rise). The pitch remains at the same level during C_3 (always [n] in the test words) and falls at the beginning of the frame.

The relative distribution of acoustic energy on the two syllables of the test words is similar for the two accents. The durations of the acoustic segments are also about the same except that the intervocalic consonants are slightly longer in the grave words. This finding is in agreement with Elert's data for the Stockholm dialect (1964 p. 156).

Physiological data. As can be seen in the averaged EMG records in Figures 1 and 2 there is general cooperation between the VOC and CT muscles. Similar cooperation for pitch control has been found in other investigations (e.g. Hirano et al. 1969).

Like the pitch curve, the VOC - CT curves exhibit one peak for the acute accent and two for the grave accent.

The activity pattern related to the accents is very consistent (Figures 1 and 2). Apart from prosodic activity we also notice the influence of articulation in the VOC. Figure 1 for instance, shows how this muscle is suppressed for the obstruents. We notice also that the EMG peaks for the grave accent are somewhat higher than for the acute accent although the resulting pitch values are similar. (This will be discussed below.)

For the acute accent the VOC - CT peaks relate to the pitch in the following way.

The activity of the muscles starts to increase about 50 msec before the release of C_1. The peaks are reached about 90 msec later, that is soon after the onset of the vocalic segment. The resulting pitch peak comes 70-90 msec later, a time lag typical of an acoustic signal as compared to an EMG peak. For the grave accent the rise towards the first EMG peak starts some 30 msec earlier than the corresponding activity for the acute peak, a difference in timing which agrees well with the acoustic record. The duration and rate of the rise are about the same for all the peaks but the rate of fall differs. The first peak of the grave accent has a steeper fall (see e.g. Fig. 1) than the other peaks which are not followed by an additional peak.

The SH muscle does not seem to be involved in the accent distinction.

Southern Swedish. Speaker E

Acoustic data. As is well known from many earlier investigations, the acute accent is characterized by an early pitch peak in the stressed syllable, while the grave accent has a peak late in the same syllable. The initiation of the rise is correlated to the location of the peak. For the acute accent (early peak) the rise starts in C_1 (= m) whereas for the grave accent (late peak) the rise starts later, at the border or at the beginning of V_1. The rate of fall after the peak is about the same in both cases.

The first syllable has much stronger intensity than the second one regardless of accent. However, the second syllable of the grave accented words contributes more to the total energy than the corresponding syllable in the acute words. (The relations in [mamma] are 3:1 for acute and 3:2 for grave.) The accents are distinguishable also by the intensity envelope, the acute accent having an earlier and faster drop than the grave one.

The segmental durations are about the same for the two accents, except that intervocalic consonants are slightly longer in the grave words.

Physiological data. The muscular activity pattern connected with the accents is rather consistent all through the test words but the VOC - CT cooperation differs to some extent from Speaker L.

In the acute words the VOC starts rising at about 70 msec before the release of initial [m]. The activity of the CT begins a little earlier (10 msec) than the VOC. The VOC peak is reached some 140 msec from the beginning of the rise, while the CT peak comes a little later. The CT remains active for about 20 msec longer than the VOC. For the grave words there is no conspicuous difference from the acute ones in the timing of the initiation of the muscular activity but the EMG peaks are reached later than for acute corresponding to the later pitch peaks.

Typical of the grave accent is that the VOC and CT remain at a high level of activity longer than for the acute accent.

Discussion

Our data suggest that both a pitch rise (as in L's acute) and a pitch fall (as in his grave accent) can be controlled by the state of contraction of the VOC and CT muscles. Whether a syllable has rising or falling pitch may depend on how the contraction and relaxation of these muscles is timed in relation to the syllable. A fall as in the grave accent can be explained as a consequence of a relaxation of muscles that have been activated earlier.

The SH muscle which has been shown to have a pitch lowering effect in some cases of American English speakers (Ohala 1970) does not seem to be involved in the pitch fall present in the word accents. The SH curves are always similar regardless of accent. Hence the SH curve does not reflect the steep pitch fall in the stressed syllable which is characteristic of the grave accent in Central Swedish (Speaker L). Moreover there are in our data many instances of SH peaks without any corresponding pitch falls. These peaks can be related to jaw opening and consonant release. Another regular feature is that the SH muscle is suppressed when the CT is active.

As stated above, the pitch peaks of the two accents are about the same. Figures 1–2 (Speaker L) show that this is not the case for the EMG peaks. The grave accent has somewhat higher EMG peaks than the acute one.

A probable explanation is that the steep rise (and the subsequent steep fall) characteristic of the first peak of this accent demands higher EMG activity than a slow rise as in the acute accent. An additional (not alternative) possibility is that since pitch starts to rise earlier in the grave accent than in the acute one, the effect of voicing initiation is added to the effect of pitch raising for the grave accent.

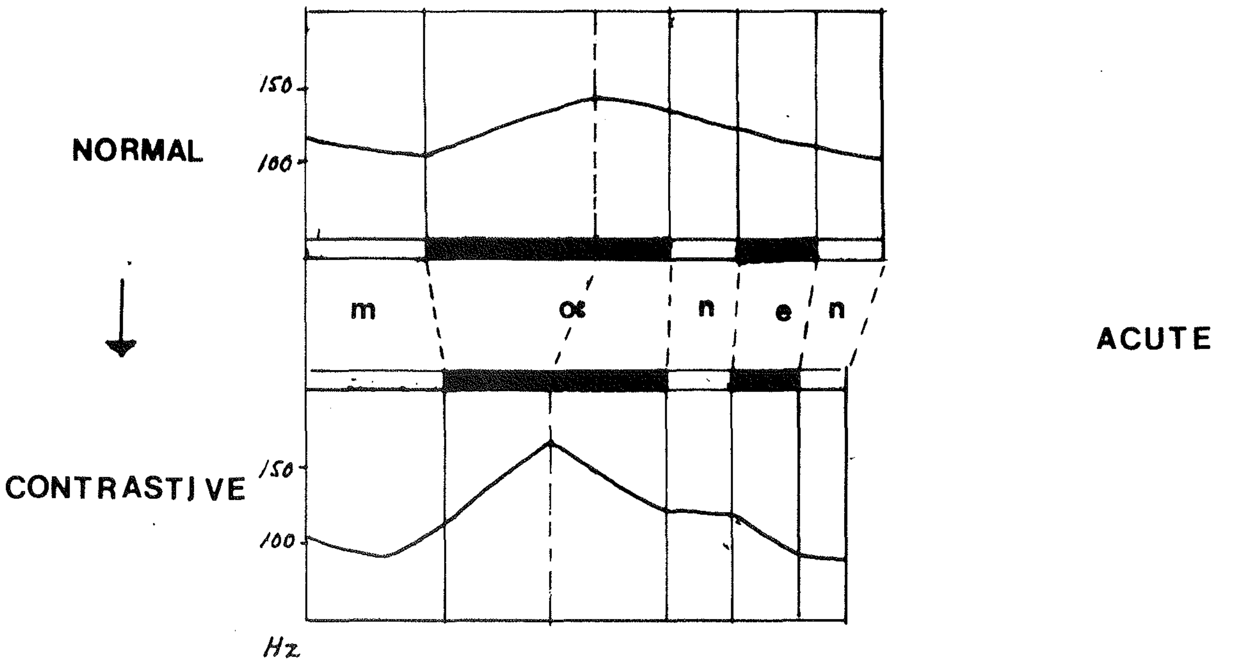
The fact that the second EMG peak of the grave accent (Speaker L) is higher than the first one can be related to the larger range of pitch rise after the transient pitch fall at the end of the first syllable.

2. Accents under neutral and contrastive stress

Acoustic data. Figures 3 and 4 (based on averages of four utterances of each test sentence) compare the difference in segmental durations and pitch curves between neutrally and contrastively stressed renderings of the test words [ˊmo:nen] and [ˋmo:nen]. The acoustic segment of the stressed initial consonant is lengthened and for the acute accent of both speakers there is a slight compensatory shortening of the segments that make up the second syllable. The part of the frame introducing the acute test word is likewise shortened for both speakers.

The following table lists the duration of the second syllable in percentage of the duration of the first one as a function of accent, stress and speaker (dialect).

	Acute		Grave	
	Sp L	Sp E	Sp L	Sp E
Neutral stress	57	90	56	82
Contrastive stress	46	37	46	70



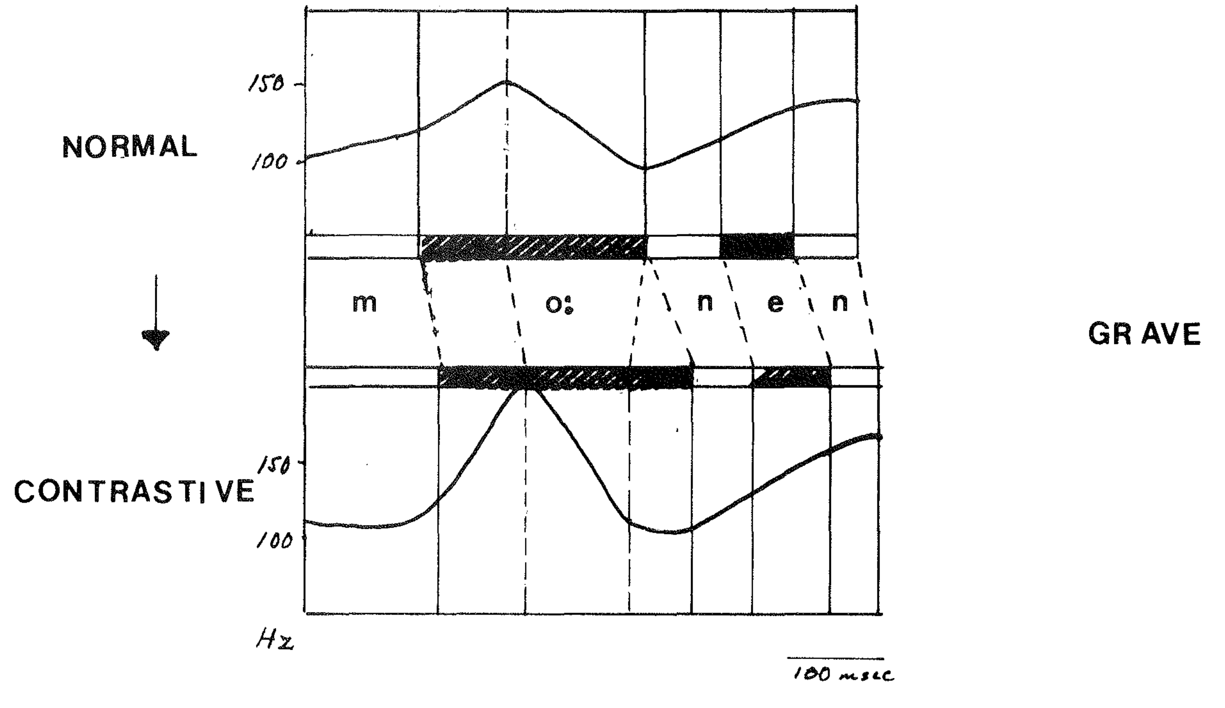


Figure 3. Accents under normal and contrastive stress.
Averaged acoustic segments and pitch curves.
Sequence /mo:nen/. Speaker L.

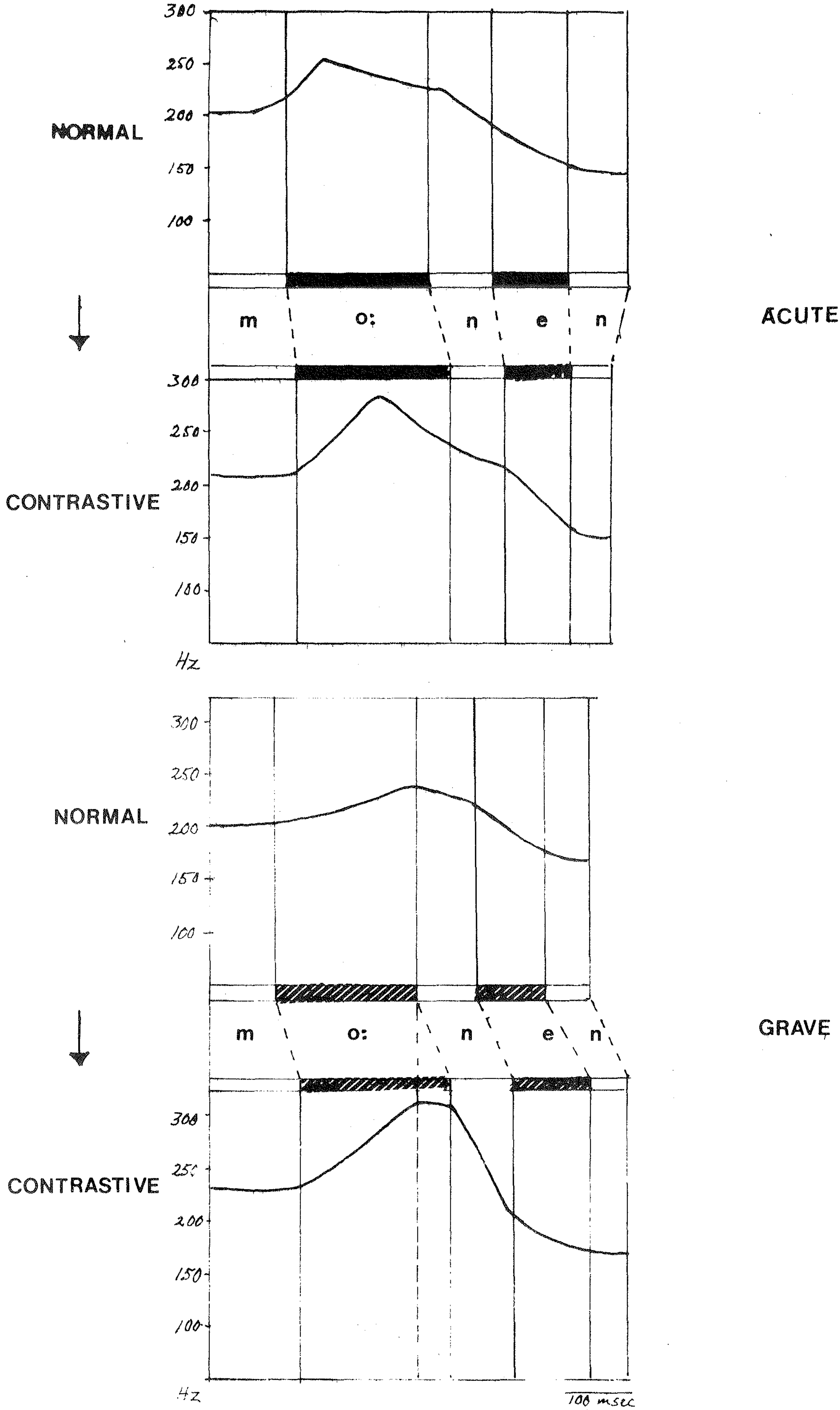


Figure 4. The same as in Figure 3. Speaker E.

We notice that under neutral primary stress Speaker L gives greater predominance to the first syllable than Speaker E who has a more equal relation between the two syllables.

The pitch curves have similar overall shapes but the contrastive accents have higher peaks.

For Central Swedish contrastive acute (late peak, Speaker L) the higher peak is achieved by a steeper rise. The rise starts earlier and its duration is somewhat shorter than under normal stress and therefore the peak location is now nearer to the middle of the vocalic segment. Consequently the rise, characteristic of the stressed syllable of the acute accent, has turned into a rise-fall under contrastive stress. The pitch fall of the second syllable remains unperturbed by the added degree of stress. The higher peaks of the same speaker's contrastive grave accent are also achieved by steeper rises. The first peak retains its early position relative to the vocalic segment. The steep rise is immediately followed by a fall with a similar rate of change but towards the end of the first syllable the fall rate is the same for both stress conditions. The second peak is reached by a steep rise from a minimum that has about the same value and the same timing as before.

The higher peak of Speaker E's contrastive acute accent is achieved mainly by a longer lasting pitch rise. The peak is followed by a steeper fall over the rest of the syllable. Because of the longer rise the typical early peak of the acute accent under normal stress is located nearer the middle of the syllable resulting in a rising-falling contour rather than a falling one. From the second syllable onwards the rate of fall is similar for the two stress conditions.

In the contrastive grave accent the higher peak is achieved by a steeper rise. The subsequent fall which is steeper than in normal stress here hits the beginning of the second syllable (intervocalic n) but in V_2 the falling rate is similar to that of neutral stress.

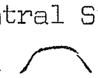
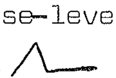
Common for both dialects is that contrastive stress modifies both syllables in the grave accent and only the first syllable in the acute one. The changes in the stressed syllable of the acute accent also have something in common in that the location of the pitch peak moves in the direction of the center of the vocalic segment resulting in a rising-falling pitch movement. As far as pitch is concerned the contrastive acute accents of the two speakers are more similar than the neutral ones.

Physiological data. As seen in Figure 5 for Speaker L, the high pitch of contrastive stress is strongly reflected in the VOC activity. The CT activity seems to be much less involved in the pitch elevation. The change to higher pitch is due either to some other intrinsic laryngeal muscles (e.g. the lateral cricoarytenoid) which are not included in the present experiment or to increased subglottal pressure. The size of the pitch change, ~ 50 Hz, is not inconsistent with the latter assumption. According to Flanagan (1971) the pitch of a closed vowel like /u/ may rise from 100 to 140 Hz when the subglottal pressure increases from 8 to 10 cm H_2O and pressure changes of this magnitude are reasonable in our situation (see e.g. Ohala 1970 p. 69 ff).

Speaker E's production of contrastive stress is different from L's (Figure 6). The higher pitch peaks of the test words uttered with contrastive stress correspond to higher CT peaks and steeper rises in the VOC.

Discussion. Comparison of dialects

Under contrastive stress L retains the same relations between the durations of the two syllables as he uses in neutral stress. Speaker E, on the other hand, changes the ratio for the acute accent in favour of the first syllable. Contrastive stress then makes the durational relations of the two syllables accent-dependent in her speech: a predominant first syllable for the acute accent versus equal durations of the two syllables for the grave accent.

Under neutral stress the pitch patterns accompanying the accents are as we have seen quite different. It has often been noted as a paradox that a Skåne grave accent is similar to a Central Swedish acute one, both having one peak rather late in the first syllable giving a rising-falling contour. The similarity is in fact great enough to create confusion in perceptual tests (Johansson 1970). This similarity is restricted to bisyllabic words only. A comparison of polysyllabic words shows how the added syllables change the contours in different ways. In Central Swedish the rising-falling contour becomes a plateau, rise-level-fall  , whereas in Skåne it turns into a rise-fall-level  (Gårding and Lindblad 1973). The present material makes it possible to show that also in the bisyllabic words the accents of the two dialects have many fea-

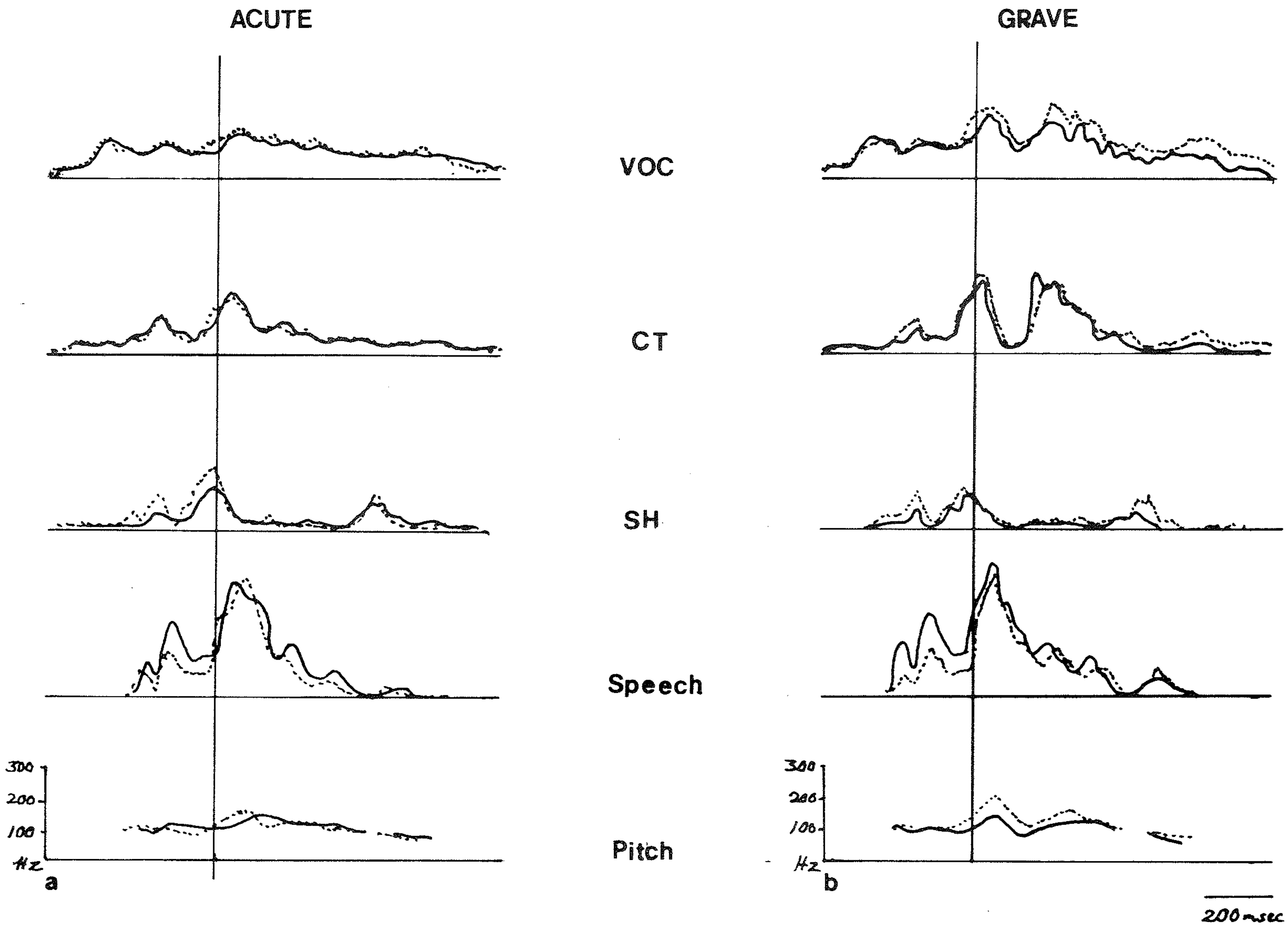


Figure 5. Accents under normal and contrastive stress. Averaged EMG signals, speech signals and pitch contours (as in Figure 1) for the utterances a) de va ´mo:nen han sa b) de va ˋmo:nen han sa. The solid curves refer to normal stress and the dotted ones to contrastive stress. The vertical line indicates the reference point for the averaging process (release of /m/). Speaker L.

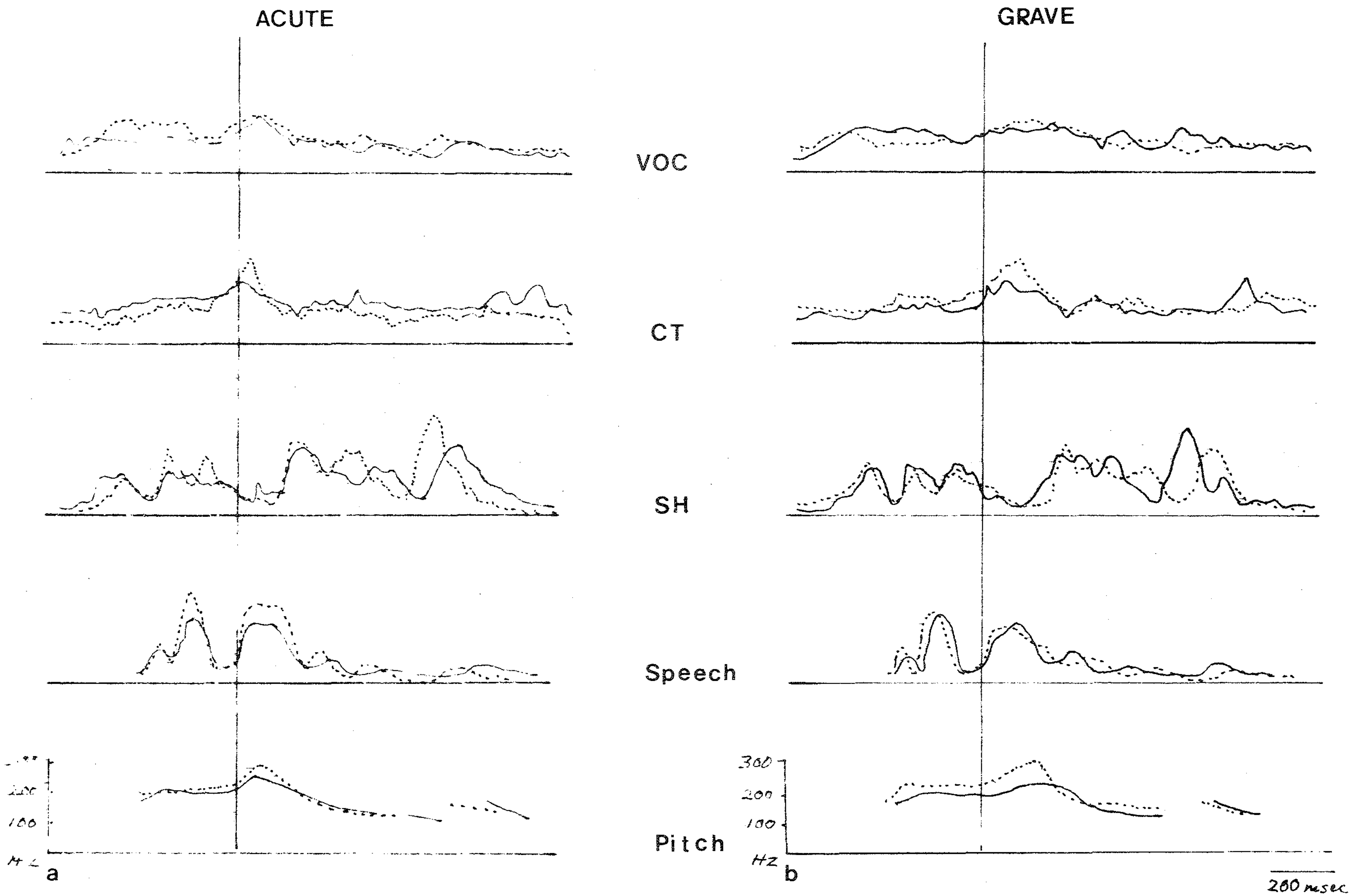


Figure 6. The same as in Figure 5. Speaker E.

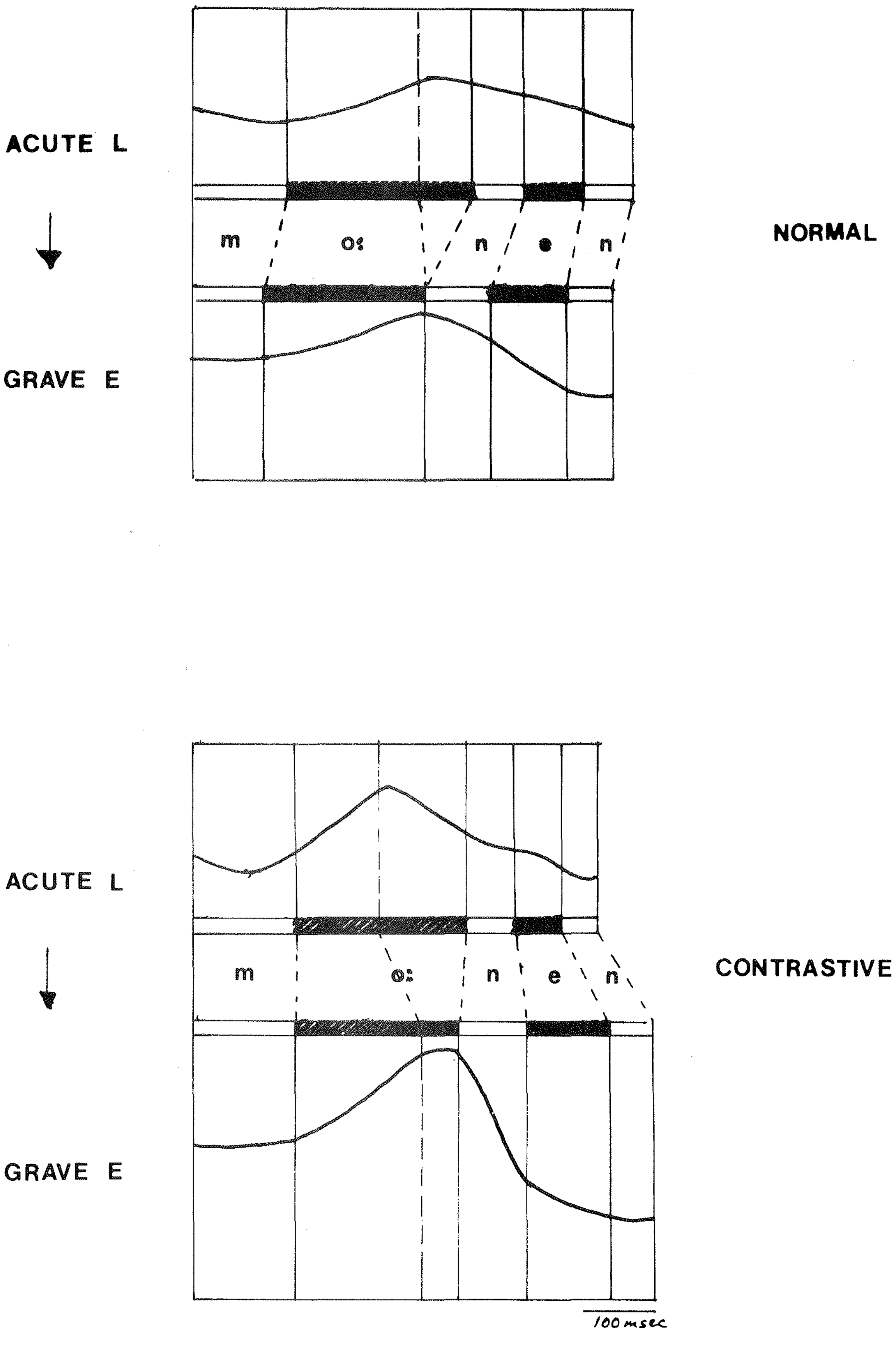


Figure 7. Dialectal variation. Averaged acoustic segments and pitch curves for L's acute and E's grave accent under two stress conditions.

tures that are not similar at all.

Figure 7 compares in some detail pitch curves from the grave (E. Skåne) and the acute (L. Central Sweden) accents under neutral and contrastive stress. We notice that E's peak occurs later relative to the stressed vowel than L's. A comparison of the peak location in open and closed syllables suggests that E's peak is tied to the syllable boundary whereas L's peak is related to the vocalic segment. The rate of fall after the peak is also different. For E, grave, the fall is slower at the beginning of the second syllable whereas L. has a rather smooth slope of pitch right through the remainder of the test word. (A similar smooth fall is characteristic of E's acute accent.)

The muscular behaviour is closely correlated with the difference in pitch (Figures 5 and 6). We notice particularly that the Skåne CT peak comes later than the Central Swedish peak and the muscle is active over a longer stretch of time. This activity should cover part of the second syllable and account for the slow rate of fall in the beginning of this syllable.

Practically all the differences that we have just noted are more marked in contrastive stress: the difference in peak location, the difference in the subsequent rate of fall and the corresponding difference in the CT activity.

All this suggests rather deep-going differences in the production and manifestation of accent contrasts in the two dialects (Skåne was formerly a Danish province and the Skåne dialect has many prosodic features in common with Danish (13)). One characteristic of contrastive stress is common to both speakers, the higher pitch values. Also in this respect, however, the similarity disappears at the physiological level. Speaker E's higher peaks are correlated to higher CT peaks whereas L does not show such a correlation. It is obvious that the higher pitch values for this speaker are achieved by a different mechanism.

The following timing characteristics of the pitch curves remain constant in the two stress conditions and in Speaker L's fast speech (see section 3). The dash should be read as "occurs in".

Speaker L

Acute. The rise – C_1

The peak – V_1 (end)

The low – V_2 (or later)

Grave. The first rise – C_1

The first peak – V_1 (beginning)

The low – the end of the first syllable

The second rise – the beginning of the second syllable

Speaker E

Acute. The rise – C_1

The peak – V_1 (beginning or middle)

The low – V_2

Grave. The rise – V_1 (beginning)

The peak – the end of the first syllable

The low – V_2

We notice that in spite of the different contours connected with the acute and grave accents in the two dialects the curves for a particular accent nevertheless have important timing similarities. If we just consider the timing of the turning points, i.e. peaks and lows, and if we accept the conventional syllabification (Table 1), we find that for the acute accent the turning points occur in V_1 and for the grave accent at the boundary between the first and second syllable. In other words, although different in form, the accent commands for a particular accent have similar targets. It is perhaps worth remembering that the acute accent developed in words that were monosyllabics in Old Norse and that the grave accent is the reflex of the accent of polysyllabics.

3. Accents in normal versus fast rate of speech

Figure 8 is based on averages of Speaker L's utterances of the test words [ˊma:ma], [ˋma:ma] and [ˊmamma], [ˋmamma]. The utterances are produced at two rates of speech, here called normal (also used in the other test sentences) and fast.

Acoustic data. As shown by the figure all the acoustic segments are shortened with an increased speaking rate. The shortening is not uniform however. Long segments, [a:] and [m:], are shortened more (20 %) than the

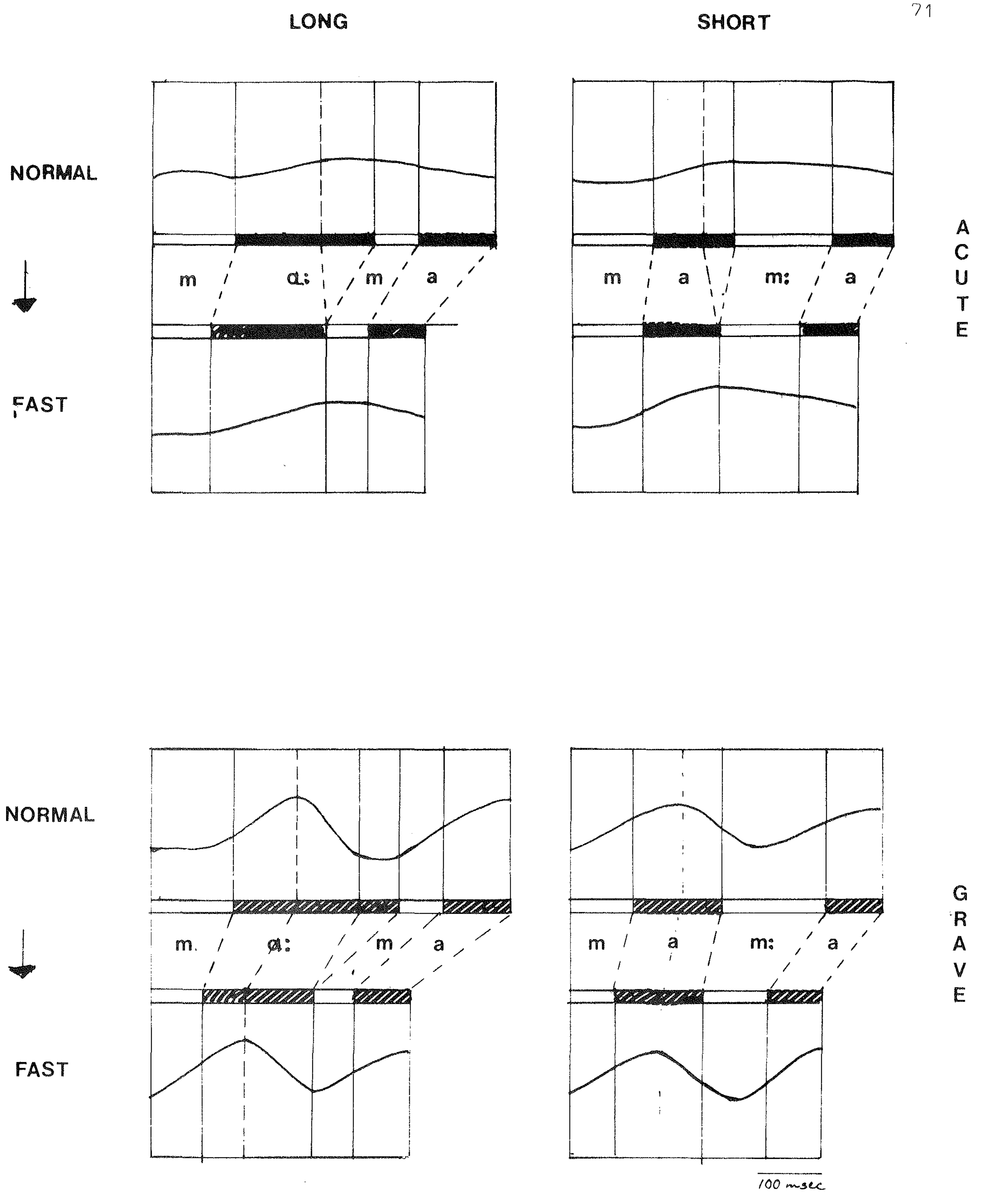


Figure 8. Accents, tempo and vowel length. Averaged acoustic segments and pitch curves for L's acute and grave accents in normal and fast renderings of the sequences /ma:ma/ and /mamma/.

corresponding short segments (10 %). An initial consonant - here always part of a stressed syllable - is shortened more (20-40 %) than an intervocalic one (10 %), which is part of an unstressed syllable. Long vowels with the grave accent are shortened more (35 %) than long vowels with the acute accent (20 %).

From the above follows that the ratio between the stressed and unstressed syllable of the test words is smaller at fast speaking rate. It also follows that there is a change in the vowel and consonant quotients (V/V: and C/C:) which become larger in faster speech.

The most conspicuous difference between the slow and fast speech curves is that at the faster rate the peaks come closer to each other and that the smooth lows are turned into rather sharp dips.

Otherwise the overall shapes of the pitch curves and the location of the peaks relative to the acoustic segments remain largely the same regardless of the speaking rate. There is a small displacement of the peaks in the grave accented words towards the middle of the syllable. The peak values are always higher in faster speech.

Differences in the location of the lows of the grave accent may be tied to the segmental and syllabic structure of the test words. For [ˋma:ma] the range gets smaller and the rise and fall rates are similar. The low has shifted upwards on the frequency scale. For [ˋmamma] the low remains unchanged and the range increases with the increased peak value. Consequently rises and falls are here somewhat steeper.

Speaking rate has similar effects on the intensity curves. In fast speech the peaks are higher and come closer to each other.

Physiological data. With increased speaking rate all the EMG peaks are slightly higher, corresponding to the higher peak values of the pitch curves. Like the two pitch peaks of the grave accented words, the peaks of the corresponding VOC and CT records come closer in the fast utterances - (Figure 9). We notice that at normal rate the first CT peak falls to the base line of activity in [ˋma:ma] whereas at fast rate the fall turns into a rise well before the base line. For [ˋmamma] the minimum has the same value at both speeds (just as in the pitch curves) but the rise from the minimum is steeper at fast rate.

The SH has two major activity peaks probably related to the jaw opening movement of va and ma. Also these peaks come closer to each other

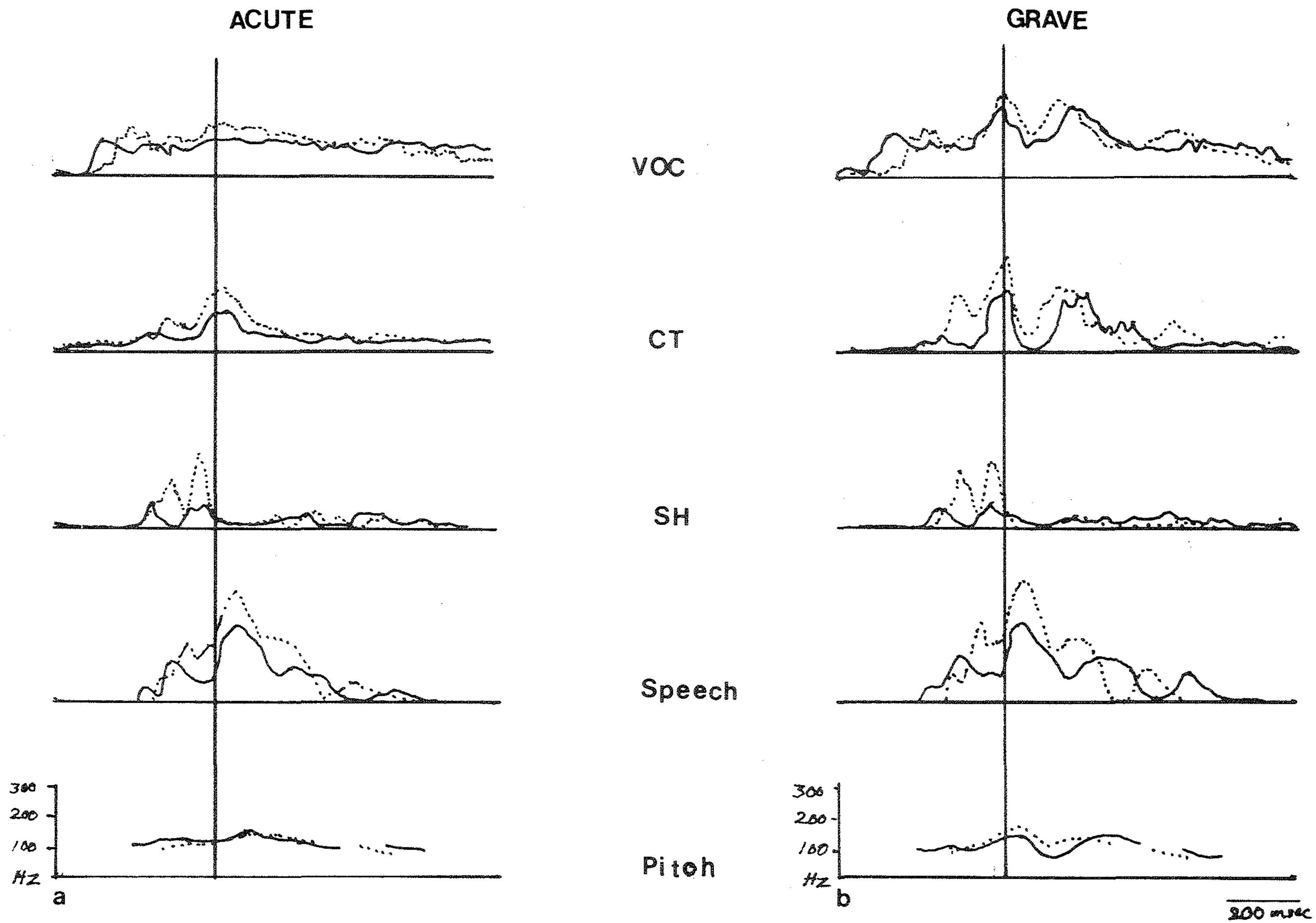


Figure 9. Accents and tempo. Averaged EMG signals, speech signals and pitch contours (as in Figure 1) for the utterances a) de va ´ma:ma ja sa b) de va `ma:ma ja sa. The solid curves refer to normal rate and the dotted ones to fast rate. The vertical line indicates the release of /m/. Speaker L.

with increased tempo.

Discussion

It is natural to assume that in fast speech, as under other strained conditions, we try to preserve the intelligibility of an utterance. This effort leads in our case to a non-uniform contraction of the acoustic segments. The segmental shortenings in fast speech affect the stressed and long segments more than the other ones.

The contraction of the pitch curve is not uniform either but certain features are preserved. The range and the timing of the peaks relative to the acoustic segments are largely intact. For this reason these features may be regarded as essential to the accent contrast.

The muscular activity is well correlated with pitch also in fast speech. A higher activity level in fast speech is probably needed to bring pitch through the same ranges in a shorter time. Why the increased activity is so large that it also results in higher pitch may perhaps be explained as an effect of overshoot.

Increased muscle activity in faster speech rate has also been observed in an EMG study of labial articulation (Gay & Hirose 1973). In this case the higher activity level is combined with observed faster movements of the articulators.

The smooth rises and falls in slow speech versus more abrupt changes in fast speech may indicate a difference in the activity pattern of the neuromuscular units (NMU) involved (slow versus instant recruitment of NMUs).

From the general increase of intensity in fast speech as shown by the oscillograms we can infer that there may be a corresponding increase in the activity level of the respiratory muscles. The increased intensity could of course also be due to increased activity in the other adductors of the larynx.

4. Accents in phonation and whisper

Figure 10, referring to Speaker E, compares data derived from phonated and whispered renderings of the test sentences with [ˊmo:nɛn] and [ˋmo:nɛn]. The whisper was rather strong.

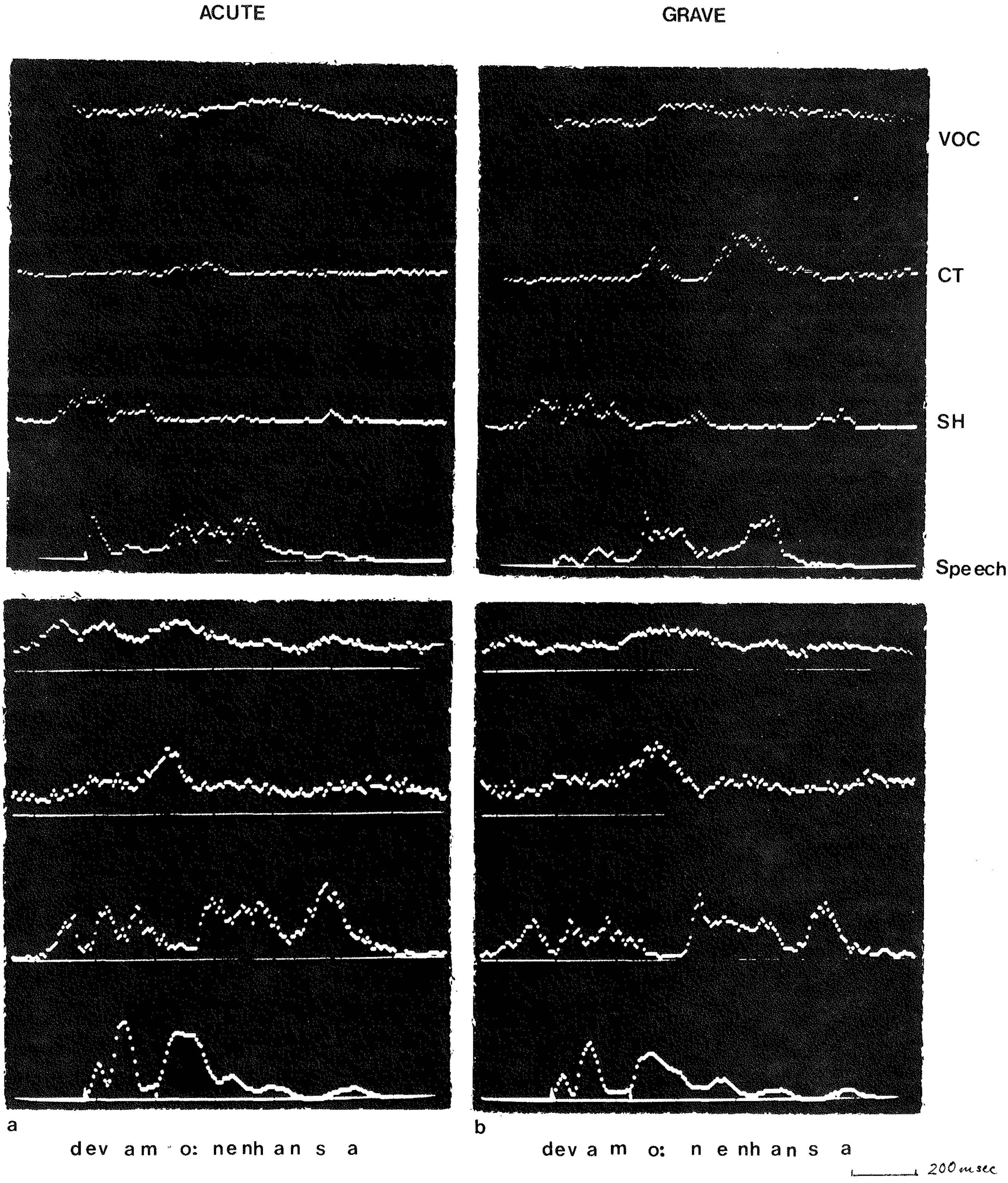


Figure 10. Accents in whispered and phonated speech. Averaged EMG and speech signals of the utterances a) de va ´mo:nen han sa b) de va `mo:nen han sa. Above in whisper and below under contrastive stress. Speaker E.

Acoustic data. The table below lists the duration of the second syllable in percentage of the duration of the first one, as a function of accent, stress and phonation.

	Acute	Grave
Normal stress	90	82
Contrastive stress	37	70
Whisper	40	75

The durations of the two syllables are about equal for both accents in normal stress for this speaker but they change drastically in favour of the first syllable under contrastive stress and acute accent. For the grave accent the equilibrium between the syllables is preserved. Whisper seems to have the same effect on the durations as contrastive stress. The similarity is striking. In the following as in Figure 10 the whispered utterances will be compared only with the contrastive ones.

The intensity of a whispered acute accent has a broad maximum over the center of the test word but the intensity envelope of a whispered grave accent has two peaks, one at the beginning and one at the end of the test word.

Physiological data. For the whispered acute accent there is a general overall suppression of the muscular activity. The small peaks of activity that do exist are displaced in comparison to the phonated utterances with contrastive stress. For VOC the highest peak of activity is not in the stressed vowel but between the end of the test word and the beginning of the frame, probably in connection with [h]. In CT the overall suppression is greater. A small peak is found towards the end of the stressed syllable.

For the whispered grave accent, the CT is considerably more active than for the acute one with peaks corresponding to those of the phonated utterances.

For both accents SH is active only before the test word.

Discussion

Since VOC and CT have been found to be active in voicing and pitch control it is not surprising that their activity should be suppressed in whisper (cf. Faaborg-Andersen 1965).

But the qualitative difference in the muscular behaviour for the two accents is astonishing. It can however be tied to other observations.

In an experiment Kloster-Jensen (1958) found that the recognition of whispered accents largely depended on the whisperer's technique. Whispered accents whose recognition scores were 100 % were performed as follows: The acute accent had one marked stress and the grave accent had an additional marked stress on the second syllable. Hadding (1961) also showed that successful grave whispers had a prominent second syllable.

From all this we may conclude that the speaker has made a special effort in connection with the grave accent and that this effort has triggered a program of signals to the CT muscle similar to that used in contrastive stress.

SUMMARY

Introduction. A Swedish word carries one of two accents, "acute" or "grave". They are commonly described as tonal. In this study of the accents we compare acoustic data and EMG data from the vocalis (VOC), cricothyroid (CT) and sternohyoid (SH) muscles. The data are obtained from two speakers representing dialects with different accent manifestations (L, Central Swedish and E, South Swedish). The accents were produced in disyllabics in sentences with statement intonation under primary stress and contrastive stress, in fast speech (Speaker L) and in whisper.

Primary stress. The fundamental frequency (pitch) curves conform to earlier observations. The acute accent for Speaker L has a late maximum (peak) in the stressed syllable giving rising pitch to this syllable and a rising-falling contour to the whole word. Speaker E has a corresponding early peak resulting in mainly falling pitch for both the stressed syllable and the word. Speaker L's grave accent is characterized by two peaks, one for each syllable (rise-fall-rise) whereas Speaker E's has one late peak in the stressed syllable producing a rising-falling contour similar to L's acute accent. As observed in other investigations, the CT activity correlates well with rising pitch. EMG peaks precede pitch peaks by about 80 msec. The VOC and CT are in ge-

neral active simultaneously particularly for Speaker L. When pitch is falling the two muscles show decreasing activity. Hence the pitch falls of the accents could be the result of a relaxation of the CT and VOC muscles from a contracted state. The SH is not involved in the pitch falls connected with the accents. The relation between the durations of the first and second syllable does not distinguish the two accents. Speaker L gives predominance to the first syllable for both accents and E has a more equal relation between the syllables.

Contrastive stress. For both speakers the absolute duration of the contrastive grave accent is larger than that of the contrastive acute. For L the relation between the first and second syllable is the same as under primary stress and similar for both accents. For E the contrastive acute is characterized by a predominant first syllable whereas the contrastive grave has retained the earlier equal relation. The overall impression of the pitch curves is the same as under primary stress but the location of the peaks is slightly displaced towards the middle of the vocalic segments. This results in a smaller accent contrast in terms of peak location for Speaker E. The peaks have higher frequency values for both speakers. There is a slight elevation of the lows and the net result is larger pitch ranges. The peaks and lows are reached by higher rise and fall rates which results in level parts in the pitch curves not present under primary stress. By and large the muscular activity follows the pitch curve. For Speaker E the higher pitch peaks are reflected in higher activity in the VOC and CT muscles. Speaker L has about the same activity for both stress conditions and his high peaks are probably due to increased subglottal pressure or activity in some other internal laryngeal muscle than those investigated.

Fast speech (Speaker L). All the segments are compressed but in a non-uniform manner. The compression affects the stressed and long segments and the segments of the grave accent more than the other ones. The overall shape of the pitch curves and the location of the peaks relative to the acoustic segments remain unchanged but the peak values are higher. The higher peak values are correlated to a high activity level in the CT and VOC.

Whisper (Speaker E). Whisper has the same effect on the relation between the durations of the first and second syllable as contrastive stress. For the acute accent the first syllable becomes predominant in contradistinction to the grave accent for which the durations of the two syllables are more equal. There is a general suppression of the muscular activity of the VOC and CT for the whispered acute accent. For the whispered grave accent however the CT is active. Here the speaker seems to have made a special effort which may have triggered a program of muscular activity similar to that used in phonated contrastive stress.

Comparison of dialects. As an example of the dialectal variability in the accent manifestations it has often been noted that the Skåne grave accent is similar to the Central Swedish acute one. Our data show that these accents differ in the timing of the pitch peak and in the corresponding VOC and CT activity.

Under contrastive stress there are durational differences for our two speakers which may be dialect dependent. Speaker L achieves contrastive stress mainly by means of higher peak values and ranges whereas E also changes the relation between the durations of the two syllables. Since E's accent contrast (timing of peak within the same syllable) is more subtle than L's (number of peaks within the word) and since contrastive stress tends to displace the peaks and diminish the contrast, she may reinforce the difference between the accents by using durational means. The same durational contrast between the accents is also used in whisper.

In spite of the different manifestations there are important timing similarities in the accents of the two speakers. If we restrict our attention to the timing of the turning points of the curves (peaks and lows) and accept the conventional syllabification we notice that for the acute accent the turning points are related to V_1 and those of the grave accents are related to the boundary between the first and second syllable for both speakers. This indicates that both the vowel (acute) and the syllable (grave) may be targets for the pitch commands. It also recalls the fact that the acute accent is the accent of the Old Norse monosyllabics whereas the grave accent is carried by words that were polysyllabic during the same period.

ACKNOWLEDGEMENTS

We should like to thank Lars Gårding for his patient performance as Speaker L and his general interest in our work. We have also profited greatly from the advice and experience of M. Hirano and J. Ohala.

References

(1) Atkinson J.E. 1973. Aspects of Intonation in Speech. Dissertation. The University of Connecticut

(2) Elert C.-C. 1964. Phonologic Studies of Quantity in Swedish. Almqvist & Wiksell, Uppsala

(3) Elert C.-C. 1972. Tonality in Swedish: Rules and a list of minimal pairs. Studies for Einar Haugen. Eds. Firchow, Hasselmo & O'Neil. Mouton, The Hague, 151-173

(4) Faaborg-Andersen K. 1965. Electromyography of Laryngeal muscles in Humans, Technics and Results. Aktuelle Probleme der Phoniatrie und Logopädie 3. Ed. Trojan. S. Karger, Basel, New York

(5) Fintoft K. 1970. Acoustical analysis and perception of tonemes in some Norwegian dialects. Universitetsforlaget, Oslo

(6) Flanagan J.L. 1970. Speech Analysis. Synthesis and Perception. Springer, Berlin, Heidelberg, New York

(7) Hirose H. and T. Gay. 1972. The activity of the Intrinsic Laryngeal Muscles in Voicing control. Phonetica 25, 140-164

(8) Gay T. and H. Hirose. 1973. Effect of Speaking Rate on Labial Consonant Production. Phonetica 27, 44-56

(9) Gårding E. 1970. Word tones and larynx muscles. Working Papers 3, Phonetics Laboratory, Lund University, 20-46

(10) Gårding E. 1971. Laryngeal boundary signals. Working Papers 4, Phonetics Laboratory, Lund University, 23-47

(11) Gårding E., O. Fujimura, and H. Hirose. 1970. Laryngeal control of word tones. Annual Bulletin No. 4. Research Institute of Logopedics and Phoniatrics, University of Tokyo, 45-53

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

(13) Gårding E., R. Bannert, A.-C. Bredvad-Jensen, G. Bruce, and K. Nauclér. 1974. Talar skåningarna svenska? Paper read to the 8th conference conc. the description of Swedish. Lund

(14) Gjerdman O. 1952. Accent 1 och accent 2, akut och gravis. Nysvenska studier 32, 125-154

(15) Hadding-Koch K. 1961. Acoustico-phonetic Studies in the Intonation of Southern Swedish. Gleerups, Lund

(16) Hirano M., J. Ohala, and W. Vennard. 1969. The function of the laryngeal muscles in regulating fundamental frequency and intensity of phonation. J. Speech & Hear. Res. 12, 616-28

(17) Hirose H., Z. Simada, and O. Fujimura. 1970. An Electromyographic Study of the Activity of the Laryngeal Muscles during Speech Utterances, Annual Bulletin, No 4. Research Institute of Logopedics and Phoniatrics, University of Tokyo, 9-25

(18) Jensen M.K. 1958. Recognition of Word Tones in Whispered Speech. Word 14, 187-196

(19) Johansson K. 1970. Perceptual experiments with Swedish disyllabic accent 1 and accent 2 words. Working Papers 3, Phonetics Laboratory, Lund University, 47-74

(20) Malmberg B. 1959. Bemerkungen zum schwedischen Wortakzent. Z. Phonetik 12, 193-207

(21) Malmberg B. 1967. Structural linguistics and human communication. Springer, Berlin, Heidelberg, New York

(22) Meyer E.A. 1937. Die Intonation im Schwedischen, I: Die Sveamundarten. Studies Scand. Philol. No. 10, Univ. Stockholm

(23) Meyer E.A. 1954. Die Intonation im Schwedischen, II: Die norrländischen Mundarten. Studies Scand. Philol. No. 11, Univ. Stockholm

(24) Ohala J. 1970. Aspects of the control and production of speech. UCLA Working Papers in Phonetics No. 15

(25) Öhman S. 1967. Word and Sentence Intonation: A Quantitative Model. STL-QPSR 2-3, 20-54

(26) Öhman S., A. Mårtensson, R. Leandersson, and A. Persson. 1967. Cricothyroid and Vocalis Muscle Activity in the production of Swedish Tonal Accents: A Pilot Study. STL-QPSR 2-3, 55-57

(27) Segerbäck B. 1966. La réalisation d'une opposition de tonèmes dans des dissyllabes chuchotés. Gleerups, Lund

(28) Simada Z. and H. Hirose. 1970. The function of the Laryngeal Muscles in Respect to the Word Accent Distinction. Annual Bulletin No. 4. Research Institute of Logopedics and Phoniatrics, University of Tokyo, 27-40

PERCEPTUAL CHARACTERISTICS OF VOWELS

Kurt Johansson

My intention in this paper is to report some investigations concerning how adults and children perceive vowels, or rather how they describe their perception of vowels

a) when they have complete freedom for description

and

b) when the frame of description is limited in some way or other by the experimenter.

The investigations were carried out during two terms of seminar work with students at the Department of linguistics, Lund University. The tests were carried out by students of phonetics and presented in independent reports. This is a revision, based on the original data from these papers.

INTRODUCTION

While quite a lot has been said about the discrimination and identification of vowels, descriptions of perceptual characteristics seem to be rather scarce. Descriptions concerning similarity are given by Göte Hansson (e.g. 1967). Eli Fischer-Jørgensen (e.g. 1967) exemplifies another type of investigation, where the task had been to describe vowels by using certain previously fixed adjectives.

Generally speaking the perceptual dimension has been much less investigated than has the articulatory and the acoustic, in spite of the fact that its importance has often been stressed and that it has often been argued that the distinctive features should be given perceptual labels rather than labels pertaining to the other levels. Of course, the essential thing could hardly be to make a consistent choice of terminology from one level or another, but rather that available data on a (possible) psycholinguistic reality are reflected by the choice.

It is beyond the scope of this paper to give a detailed account of the extensive discussion that has taken place on the domain of distinctive features, particularly since the appearance of Jakobson-Fant-Halle's Preliminaries to Speech Analysis (1952). The purpose is instead to make a contribution to this discussion with the previously mentioned seminar investigations as a starting-point.

SCOPE OF INVESTIGATIONS

The investigations concerned <u>two main fields</u>:

A. Associations between vowels and adjectives

B. Associations between vowels and colours.

The same <u>material</u> was used for tests I–V viz. the long Swedish vowels [i:, e:, ɛ:, ɑ:, o:, u:, y:, ʉ:, ø:] plus the short [a]. The vowels were uttered as monophthongs by a male phonetician and recorded on tape. Formant frequencies of these vowels are given in Table 1.

	F1	F2	F3
i	200 Hz	2300 Hz	2900 Hz
e	350	2200	2650
ɛ	500	2000	2300
a	800	1250	–
ɑ	500	900	–
o	350	650	–
u	250	600	–
y	225	2200	2400
ʉ	250	1850	2400
ø	400	1600	2200

Table 1. Formant frequencies of the vowels used in tests I–V.

<u>Field A</u> comprised the following tests:

I. Judgements using 6 given adjectives that are frequently used in phonetics for describing vowels perceptually.

II. Free association between vowels and adjectives.

III. Judgements using 7 given adjectives. This investigation was based on the free associations in test II.

<u>Field B</u> was composed of the following tests:

IV. Free associations between vowels and colours.

V. Judgements using given colours. This investigation was based on the free associations in test IV.

VI. The same as IV and V but with French oral and nasal vowels and French listeners.

All tests started with some buffer stimuli. In each randomized test each stimulus appeared on three different occasions, each time presented three times in a succession with about one-second pauses.

FIELD A (ASSOCIATIONS BETWEEN VOWELS AND ADJECTIVES)

Test I

The starting-point for this investigation were 6 adjectives, often used for perceptual descriptions of vowels:

ljus - mörk ("light":"dark") for [i]:[u]

tunn - fyllig ("thin":"mellow") for [i/u]:[a/ɑ]

klar - dov ("bright":"dull") for [i]:[y]

The material was first presented to 14 phonetically naive Swedish university students who were asked to give one or more adjectives for each vowel, with second choices in brackets. However, some of the listeners did not indicate their primary choice at all. But if these people had been excluded from the analysis, the results would still have been very much the same.

A second test that on the whole could be regarded as a repetition of the first one was undertaken some time later. Here, only one adjective was to be listed for each vowel. The 10 listeners had also participated in the first part.

Though the circumstances are not quite comparable, the scores are treated as if they had been achieved on the same occasion. (Table 2.)

	i	e	ɛ	a	ɑ	o	u	y	ʉ	ø
light	32	28	7	18	5			5		
dark	5	2	15	8	17	18	20	14	16	21
thin	25	18	9	26	7	3	3	15	12	5
mellow	5	3	20	10	14	11	14	20	13	19
bright	13	27	21	20	33	5	3	12	7	9
dull	3	3	9	1	6	43	43	13	31	28

Table 2. Scores for the six adjectives used in test I. The responses from two similar tests have been pooled.

There are some clear tendencies. They are perhaps still more apparent from table 3 below, where I have given the three most common adjectives for each vowel, with the one most preferred first.

	1	2	3
i	light	thin	bright
e	light	bright	thin
ɛ	bright	mellow	dark
a	thin	bright	light
ɑ	<u>bright</u>	dark	mellow
o	<u>dull</u>	dark	mellow
u	<u>dull</u>	dark	mellow
y	mellow	thin	dark
ɯ	<u>dull</u>	dark	mellow
ø	dull	dark	mellow

Table 3. The three most common adjectives for each vowel in test I (cf. table 2). When underlined an adjective dominates completely.

"Light/thin/bright" are the adjectives principally found for [i] and [e]. For [e] the order is "light/bright/thin".

If we compare [i] and [y] which are rather akin from an articulatory as well as from an acoustic angle, we find for the rounded vowel that the lowering of F2 and (particularly) F3 have caused a shift towards "mellow", "dark", and "dull".

For [ɯ], where there has been a further lowering of F2, the listeners prefer "dull", but also "dark" and "mellow" appear. As a matter of fact [ɯ] belongs to the same perceptual group as [u], [o], and [ø]. In all cases the same adjectives are chosen, and in the same order: "dull/dark/mellow", with "dull" as the dominating choice.

"Mellow" and "dark" also appear for [ɛ] and [ɑ], in different order, however, and in both cases "bright" is the most common adjective. Particularly for [ɑ] it dominates completely. "Bright" is also found for [a], but here together with "thin" and "light".

On the whole the results are satisfactory though the different vowels are not kept totally apart. We get two clearly separated groups, [i/e] and [u/o/ɯ/ϕ], but also [ɛ] and [ɑ] that seem to be connected somehow. [a] and [y] seem to constitute intermediate forms.

The main tendency would have been the same if the two tests had been treated separately. Some changes would have appeared, however: "dull" instead of "dark" for [ɛ] and [y] (in one test). For [y] "thin" in one test was exchanged for "bright" in the other. In connection with [ɯ] "thin" is chosen in one case instead of "dark", showing a certain and expected kinship with [y] and [i]. There would also have been some changes of the order between the adjectives, compared to the order in table 3.

These were the main results of this test. In order to proceed further along this line, a semantic normalization has to be made. Now, some of the adjectives seem to be synonymous. Particularly "dark", "dull", and "mellow" seem to interchange freely.

Test II

For this test the subjects were instructed to give one adjective only, that best described the vowel, with no restrictions on their choice. That the listeners were not restricted to a certain kind of adjective means for instance that also colour adjectives appear, albeit sparsely. Further, not only are "purely" perceptual terms to be expected, but also terms referring to the articulatory level. If the motor theory of speech perception is correct, one might expect quite a few adjectives of that kind, though certainly not so many for vowels as for consonants, where the tactile feedback is more obvious. In fact, terms of that type did appear, but they are very rare and whenever they occur they do not always have a connection with the actual articulatory features of the vowel in question. See however table 4.

Two groups of 10–12 listeners each took part in this investigation, group A consisting of students aged about 20 and group B of school children aged 10 to 11. Group A was tested once and group B twice.

Results

Under these circumstances it is quite obvious that the distribution of choices must be very great. In all about 70 different adjectives were offered by group A and about 60 by group B. In most cases an adjective only appears occasionally, but as could be seen from table 4 there are some more frequent choices. In a few cases different adjectives have been gathered under the same heading. For instance some kantig ("angular") responses have been added to spetsig ("sharp"). This does not alter the facts, however.

	i	e	ɛ	a	ɑ	o	u	y	ɯ	ø
sharp	28	21			5		2	13	9	
light	8	4		7	12				5	2
bright				2	4					
cold				4	4					
hard		4		7		10	2	5		2
thin	2		2		3		3		2	
dark						6				15
gloomy		2	3	11	2		5	9	4	
dull				2		11	17		2	6
coarse		2		2		3	2			2
soft					4				3	
round		2					3		2	
bleating	3		20							
close								2	4	
half-open						7	10			

Table 4. Scores for free association between vowels and consonants. Groups A and B are pooled. The Swedish adjectives were spetsig, ljus, klar, kall, hård, tunn, mörk, dyster, dov, grov, mjuk, rund, bräkig, sluten, halvöppen respectively.

In the investigation reported above [i] was "light/thin/bright". "Light" is the only adjective of these that to some extent maintains its place, but here "light" must give way to "sharp" which is the adjective preferred by most listeners. "Angular", "piercing" (skarp), "keen" (vass) are other adjectives of a similar meaning, which also occur occasionally for this vowel.

[y] and [ɯ] have their highest scores for "sharp", too, indicating a kinship with [i]. "Gloomy" for [y] in this test might correspond to "mellow/dark" in test I.

[u] and [o] are still "dull", [o] is "hard" as well, [u] "half-open" (!), however.

[ø], which was also labelled "dull/hard/mellow" in test I, gets the highest score for "dark". An influence from the vowel letter or the vowel sound in the matching adjective offered by the listener cannot be excluded, here or elsewhere.

[ɛ] is labelled "bleating", probably for the same reason.

[ɑ] is primarily "light", while the dominating adjective in the earlier investigation was "bright". The earlier "thin/bright/light" for [a] has become "gloomy/light/hard".

No doubt some consistency could be found between tests I and II, also when groups A and B are treated separately. Considering the low scores in test II definite conclusions are impossible.

Test III

This investigation is a consequence of test II. The most frequent adjectives from that test have been made up into one set, i.e. "sharp", "light", "gloomy", and "hard". Further, the opposites of "light" and "hard" were included, i.e. "dark" and "soft". "Bright" was included as a potential opposite of "gloomy". No more opposites appear. Personally I miss "dull", which obviously belongs to the most frequent group, and which furthermore was the most frequent choice in test I. In order to correct, at least to some extent, a wrong choice of adjectives, there was also an open alternative, where the listeners might choose freely.

Three groups of listeners were tested:

Group I: a group of 11 students of phonetics aged 25 to 60

Group II: 17 school children aged 10 to 11

Group III: 5 phonetically naive adults aged 21 to 30.

Results

The variety of responses is large in all groups. There is slightly more consistency among the adults than among the children, but there are common tendencies within all groups. Below (table 5) the groups are treated as one group. As can be seen, all adjectives have more than 100 responses, none more than 185. In order to facilitate a comparison with the earlier tests the three most common adjectives for each vowel and each test are given in table 6, with the most frequent adjective first.

	i	e	ɛ	a	ɑ	o	u	y	ɯ	ø	total
sharp	32	10	14	19	7	2	1	23	8	9	125
light	28	11	11	14	8	1	8	22	6	2	111
bright	6	14	4	21	24	8	2	6	6	9	100
hard	22	31	19	36	18	11	13	19	8	8	185
dark	1	10	11	2	21	22	33	5	19	29	153
gloomy	2	14	25	4	11	17	14	15	24	35	161
soft	6	8	10	2	6	30	23	8	26	3	122
other choice	2	1	5	1	3	8	5	1	2	4	32

Table 5. Choice between given adjectives. The scores for the three groups in test III are pooled. "Other choice" also includes blank responses.

"Sharp" and "light" evidently give a good characterization of [i]. For the back rounded vowels there is "dull", not included in test III, together with "dark", "mellow", "soft", and "gloomy". These are also found for [ɯ], and, with the exception of "soft", for [ø]. Obviously as a result of the energy shift towards higher frequencies in comparison with [u] and [o], such adjectives as "sharp", "light", and "bright" appear.

With such adjectives as "sharp" and "mellow/gloomy", [y] has features in common with [i] as well as with the rest of the rounded vowels.

[e] is "hard", "sharp", but also "light" and "bright".

[ɛ] is difficult to account for. The responses lack uniformity. There are "bright", "hard", and "sharp", but also "bleating", "gloomy", and "mellow".

[ɑ] is "bright" and "light", but also "dark".

"Bright" and "light" are found for [a], too, but also "hard", "thin", and "gloomy".

		1	2	3
i	I	light	thin	bright
	II	sharp	light	–
	III	sharp	light	hard
e	I	light	bright	thin
	II	sharp	(light)	(hard)
	III	hard	bright	gloomy
ɛ	I	bright	mellow	dark
	II	bleating	(gloomy)	(thin)
	III	gloomy	hard	sharp
a	I	thin	bright	light
	II	gloomy	light	hard
	III	hard	bright	sharp
ɑ	I	bright	dark	mellow
	II	light	sharp	(bright/cold/soft)
	III	bright	dark	hard
o	I	dull	dark	mellow
	II	dull	hard	half-open
	III	soft	dark	gloomy
u	I	dull	dark	mellow
	II	dull	half-open	gloomy
	III	dark	soft	gloomy
y	I	mellow	thin	dark
	II	sharp	gloomy	hard
	III	sharp	light	hard
ɯ	I	dull	dark	mellow
	II	sharp	light	(gloomy/close)
	III	soft	gloomy	dark
ø	I	dull	dark	mellow
	II	dark	dull	(coarse/hard/light)
	III	gloomy	dark	bright/sharp

Table 6. The most common adjectives for each vowel in tests I, II, and III, respectively. When underlined an adjective dominates completely. Only exceptionally (from test II) have adjectives with scores lower than 5 been included (in brackets).

Discussion

With this material as a starting-point it does not seem impossible to reach a satisfactory perceptual description of the vowels. In spite of the fact that the problem has been attacked from somewhat different angles in the three tests, and that listeners of different age and phonetic knowledge have been used, there is good agreement on certain points.

[i] receives similar judgements in the three tests, and so do the rounded vowels, particularly the back vowels. "Light":"dark", or in the language of the listeners of these tests "sharp":"dull", seems to be the dimension most easily agreed upon. Considering the predominance on one hand of high, and on the other hand of low frequencies, this is hardly unexpected.

[ɑ] ("bright") seems to constitute another extreme, but it is hardly so well defined. "Mellow", which has often been used for description of low vowels, does not seem adequate.

The dimension "dull":"bright" primarily seems to concern [u/o]:[ɑ/a], but "bright" appears in effect for all unrounded vowels.

There are intermediary forms, e.g. the front rounded vowels [y] and [ʉ], which have "sharp" in common with [i], and "dull", "mellow", and "gloomy" with [u] and [o].

The data seem to some extent to suggest that the extremes, particularly [i] and [u], i.e. the acoustic extremes, are used as reference vowels, in other words as more absolute units than the other vowels. This would mean that when an [i] is judged as "sharp", it is not so much in comparison with the other vowels but rather in the same way as a whistle (or perhaps an [s] among the obstruents) may be perceived as "strident", "sharp", or "piercing". Similarly [u] would be perceived generally as "dull" by virtue of the very low frequencies characterizing it.

My intention is not to squeeze the data too much. What was, however, quite clear from test II was that the vowels were not to any great extent perceived in terms of colour or production. Much further elucidation is needed. I hope to be able to return to this problem later (as well as to the perception of consonants with investigations along the same line).

The great problem in investigations of this kind is of a semantic nature. The orthography, too, constitutes a "disturbance". How will it ever be possible to understand what a person means by a certain term and to relate this to what other persons mean? One way will perhaps be comparisons in pairs, where one vowel comes first the first time, the other vowel the second time. If the listeners are to describe the second vowel relative to the first one, it would at least be possible to get at what the different listeners consider to be terminological counterparts.

Eventually also the influence of different voice qualities (male, female, and child's voice) on the judgements of different types of listeners must be taken into consideration. The material must also be varied as to duration and certainly also as to consonantal context.

FIELD B (ASSOCIATIONS BETWEEN VOWELS AND COLOURS)

"... the structure of sound and colour systems shows marked agreements. Moreover, cases of pronounced coloured hearing, especially in children or retained from childhood, in which acoustic impressions and particularly speech sounds, "appear bound non-arbitrarily, regularly and consistently with the same colour experiences", show the close connection of the vowels o and u with the specifically dark colours, and of e and i, on the other hand, with the specifically light colours." (R. Jakobson, Child Language, Aphasia, and Phonological Universals, 1968, p. 82.)

Now and then statements of the above kind may be found in the literature. It is difficult to say how common a pronounced colour perception is, but the investigations reported above show that very few listeners, when facing the task to associate freely between adjectives and vowels, actually choose colour adjectives. Otherwise I can give an example that at least children may perceive vowels as colours. A six-year-old son of a friend of mine assigned colours with great confidence not only to isolated vowels but also to words containing two vowels. If the word contained one "yellow" and one "blue" vowel, it was reported as "green", and so on. Now, after a couple of years at school, at least his more spontaneous ability is said to have decreased.

Jakobson gives (op.cit., p. 83) examples of two systems of colour associations:

1.

ɑ red		œ rose
o blue	ø light blue	e yellow
u dark brown	y grey	i silver white

2.

	ɑ red	
o blue-red		e light green
u dark blue		i canary yellow

One of the conclusions in connection with test II above was then that colour perception can hardly be the primary type of perception. From the tests below we may in other words only find out, if we nevertheless may have the ability to associate vowels and colours in a fairly consistent way. It might also be possible to relate the ability of children in this respect to that of adults.

Test IV

This test was carried out with the same listeners as in test II, both group A (adults) and group B (children). Each group, consisting of 10-12 listeners, judged the material once, and their task was to associate freely to colours.

Results

The total scores for both groups are reported in table 7. In table 8 the three most common colours for each vowel are given.

As can be seen the choice has primarily been between two colours for each vowel. It may be suspected that the orthography may have played an important role in some cases, though the listeners had been requested to try to neglect it. This is probably not true to any greater extent for [i], though it is characterized as "white" (cf. test VI). Other possible orthographical and vowel sound influences are that [ø] is "green" (grön), [o] is "blue" (blå), [ʉ] is "yellow" (gul), and [a/ɑ] are "black" (svart).

[e/ɛ/y] are primarily "green", [y] with a tinge of "white", revealing kinship with [i].

	i	e	ɛ	a	ɑ	o	u	y	ɯ	ø
white	41	2	6	4	3	3	2	12	2	2
yellow	11	7	17	1	1	6	20	9	29	1
green	1	17	23		2	2	8	14	4	41
blue	5	10	2	9	6	30	16	5	1	5
black		2	3	21	26	2	3	2		
red	3	12		21	24	1	7	7	5	17
lilac	7	4				1	1	8	2	2
pink			14			1	1			
grey	1	9	1	9		18	4	2	7	1
brown		6	3	4	7	5	7	10	18	

Table 7. Scores for free associations between vowels and colours. Groups A and B are pooled.

	1	2	3
i	<u>white</u>	yellow	–
e	green	red	blue
ɛ	green	yellow	pink
a	<u>red/black</u>	–	–
ɑ	<u>black</u>	<u>red</u>	–
o	<u>blue</u>	grey	–
u	yellow	blue	–
y	green	white	brown
ɯ	<u>yellow</u>	brown	–
ø	<u>green</u>	red	–

Table 8. The three most common colours for each vowel (cf. table 7). When underlined the colours dominate completely. Only colours with scores higher than 10 are included.

Some differences could be seen between the groups. The children are somewhat less consistent and sometimes prefer other colours.

As for [i/e/ɛ] there are no real differences. Nor are there for [a/ɑ], though group A favours "red" and group B "black". [y/ɯ] are "brown" for group A, [y] "green/white" and [ɯ] "yellow" for group B. Group A chose "yellow" for [u] and "blue" for [o], while group B preferred "blue" and "grey/blue", respectively.

It is obvious that somehow the orthographic influence must be mastered. In tests V and VI we have tried to solve this problem by using colour plates in such a way that the listeners would not have to write down the names of the colours.

Test V

Here the listeners were given 6 colours to choose among, "white/yellow/green/blue/red/black" plus an open choice providing the possibility to choose other colours, if necessary. The above colours were, together with "brown" and "grey", the most frequent in test IV.

The material was presented to two groups, group A consisting of 19 adults aged 20 to 57, group B of 21 children aged about 11.

In order to avoid the orthographic influence a plate had been made containing the six colours mentioned above. Each colour was provided with a number to be written on the response sheet instead of the name of the colour.

Results

In table 9 below the responses for groups A and B are pooled.

The three most common colours for each vowel are more easily seen in table 10.

There are some differences, but also some similarities, in comparison with test I.

Here, too, [i] is associated with "white". "Red" reappears for [ɑ], but not for [a], and "black" is not particularly common for either of these vowels. [u] has changed from "yellow/blue" to "black/green". [o] and [ɛ] are still "blue" and "green", respectively, while [e] has changed from "green/red" to "blue/red". Also for the rounded front vowels there are some changes: for [y] from "green/white" to "yellow/red",

	i	e	ɛ	a	ɑ	o	u	y	ɯ	ø
white	50	14	16	25	20	3	12	13	5	5
yellow	33	19	7	26	21	11	15	33	19	12
green	1	18	28	8	16	22	22	14	21	31
blue	12	28	20	19	15	49	11	15	14	17
black	4	8	12	15	14	14	24	9	23	37
red	11	21	18	18	23	11	14	20	15	7
other choice	9	12	19	9	11	10	22	16	23	11

Table 9. Scores for associations between vowels and 6 given colours. Groups A and B are pooled.

	1	2	3
i	<u>white</u>	yellow	–
e	blue	red	–
ɛ	green	blue	–
a	yellow	white	–
ɑ	red	yellow	white
o	<u>blue</u>	green	–
u	black	green	–
y	yellow	red	–
ɯ	black	green	–
ø	black	green	–

Table 10. The three most common colours assigned to each vowel in test V (cf. table 9). When underlined the colour dominates completely. Only colours with scores higher than 20 are included.

for [ɯ] from "yellow/brown" and for [ø] from "green/red" to "black/ green".

Also in this test the adults are more consistent than the children.

As the material is the same in the two tests reported above I have taken the liberty to treat the adult groups as a single group and the children's groups as another, in order to get, if possible, the differences that may occur between groups of different age. See tables 11–13.

	i	e	ɛ	a	ɑ	o	u	y	ɯ	ø
white	44	6	6	19	17	4	6	13	1	3
yellow	22	12	15	10	8	12	22	24	15	6
green		15	22	2	6	12	16	7	17	41
blue	7	24	9	19	16	40	10	9	9	11
black	1	2	6	14	15	4	5	4	7	13
red	10	15	10	17	20	5	9	10	11	9
other choice	7	16	22	9	8	13	22	23	30	7

Table 11. Total score for the adult groups in tests IV and V concerning vowels in relation to colours.

	i	e	ɛ	a	ɑ	o	u	y	ɯ	ø
white	42	8	10	10	9	3	6	8	4	2
yellow	25	13	18	16	13	3	15	23	14	6
green	1	19	32	6	10	14	20	11	12	30
blue	11	20	13	12	9	43	11	12	7	14
black	3	6	6	13	19	12	19	7	16	24
red	7	16	8	25	31	6	11	10	10	16
other choice	7	14	9	14	5	15	14	25	33	4

Table 12. Total score for the children's groups in tests IV and V concerning vowels in relation to colours.

		1	2	3
i	A	<u>white</u>	yellow	–
	B	<u>white</u>	yellow	–
e	A	blue	(red/green)	–
	B	blue	(green)	(red)
ε	A	green	(yellow)	–
	B	<u>green</u>	(yellow)	–
a	A	(blue/white)	(red)	–
	B	red	(yellow)	–
ɑ	A	red	(white)	(blue)
	B	<u>red</u>	(black)	–
o	A	<u>blue</u>	–	–
	B	<u>blue</u>	–	–
u	A	yellow	(green)	–
	B	green	(black)	(yellow)
y	A	yellow	–	–
	B	yellow	–	–
ɯ	A	(green)	(yellow)	–
	B	(black)	–	–
ø	A	<u>green</u>	–	–
	B	green	black	(red)

Table 13. The three most common colours for each vowel. The two adult groups from tests IV and V are pooled (A), and so are the children's groups (B). When underlined a colour dominates completely. Colours with scores lower than 15 are not included. Colours in brackets have scores lower than 20.

As can be seen from table 13, there are certainly very great similarities between the two groups. Only on three points do there seem to be more apparent differences, viz. for [a], [u], and [ɯ]. For [a] "red" appears, however, for adults as well as for children, and it is the dominating colour for [ɑ]. "Yellow" and "green" are associated with [u] by both groups. [ɯ] looks more irregular, but if e.g. most of the "brown" responses of test IV had been replaced by "yellow", if the conditions had been quite the same as in test V, which does not seem too unlikely, "yellow" would have dominated for the adults as well as for the children. The differences are indeed very small between the two groups.

An arrangement of the same kind as the Jakobsonian might have the following appearance:

i	white	y	yellow	u	yellow/green
e	blue	ɯ	yellow	o	blue
ɛ	green	ø	green		
a	red			ɑ	red

In spite of the supposition that the listeners may have had difficulties in disregarding the orthography, the results are so unanimous that one does not hesitate to state that the listeners really possess the ability to associate fairly consistently between vowels and colours. This goes for both groups.

In a tentative arrangement like the one above there are several overlappings, which do not seem to be entirely unsystematic. No doubt the arrangement could serve as a starting-point for continued investigations.

To couple the results from the test concerning colours to those concerning adjectives in general does not appear particularly meaningful at this stage, but in both cases the data indicate that it is possible to reach a perceptual description that is not founded on the linguist's intuition alone.

Test VI

The following French vowels were included in this test [i, e, ɛ, a, ɑ, o, u, y, ø, œ, ɛ̃, œ̃, ɑ̃, ɔ̃].

For practical reasons the French student carrying out the test was not able to use prerecorded material, but had to pronounce the vowels himself at each test situation. The listeners were tested one at a time.

18 Frenchmen took part in this test, and their first task was to associate freely between vowels and colours. Three of the listeners did not consider it possible to associate with colours and are not included

in the tables below.

14 of the remaining listeners also took part in a second experiment. This differed from the first in that the listeners gave their responses by pointing to coloured squares - "white", "yellow", "red", "blue", "green", "brown", and "black". Each colour except "white" was divided into shades from lighter to darker.

Each listener was asked to give only one colour per vowel and experiment.

Results

In table 14 the responses to both experiments are reported together.

	i	e	ɛ	a	ɑ	o	u	y	ø	œ	ɛ̃	œ̃	ɑ̃	ɔ̃
white	12	1												
yellow	12	6		1		4	3	6	8	9	5	1	1	2
brown		2	7		1		3	4	8	8	4	11	4	9
green	2	6	15	1				1	2		7	4	3	3
blue	2	5	2	5	4	7	11	10	4	4	2	5	6	2
black		1	1	7	10	3	3	1	3	3	4	4	2	2
red		2	3	13	14	13	5	4	4	4	2	4	9	5
other choice	1	5	1	2		2	4	2		1	5		5	6

Table 14. Scores for associations between French vowels and colours. Two tests pooled.

Of course the responses are too few to permit any far-reaching conclusions, but certain similarities to the Swedish material seem to exist.

Here, too, [i] is "white" and "yellow", which is an indication that it is really possible to free oneself from the orthography (French blanc, jaune). [ɛ] is "green", [a/ɑ] "red" in both cases. Rounded vowels show a certain tendency towards "yellow/brown".

The nasal vowels have as a rule caused considerable difficulty. One of the listeners remarked, after having taken part in the test, that the nasals were no real sounds. Another listener called them "semi-sounds".

Where a comparison with the corresponding oral vowels is possible, the judgements seem to be of a similar nature, with somewhat lower scores for the nasals.

Summing up it could be said that the colour experiments suggest that a coupling of vowels to colour perception is possible. Evidently there is reason to expect results of a more universal bearing in this connection. The orthographical difficulties will certainly not be easy to handle, but the method used in test VI, with a colour plate containing squares with a number of variants of the same colour would seem to be a useful method for solving some of these problems. Presenting several shades of the same colour will certainly make it easier for listeners who may not be willing to accept a "pure" colour, and it facilitates a mapping of the direction to which the colour perception within a certain square points.

Bibliography

Chomsky, N., Halle, M., The Sound Pattern of English, 1968

Fant, G., The Nature of Distinctive Features, 1967, To Honor Roman Jakobson, pp. 634-642

Fant, G., Lindblom, B., Studies of Minimal Speech Sound Units, STL - QPSR - 2/1961, pp. 1-11

Fischer-Jørgensen, E., Perceptual Dimensions of Vowels, 1967, To Honor Roman Jakobson, pp. 667-671

Hansson, G., Dimensions in Speech Sound Perception, 1967

Jakobson, R., Child Language, Aphasia, and Phonological Universals, 1968

Jakobson, R., Fant G., Halle, M., Preliminaries to Speech Analysis, 1952

Jakobson, R., Halle, M., Fundamentals of Language, 1956

Ladefoged, P., Preliminaries to Linguistic Phonetics, 1971

Liberman, A.M., Cooper, F.S., Harris, K.S., MacNeilage, P.F., A Motor Theory of Speech Perception, Proceedings of the Speech Communication Seminar, Vol. 2, Stockholm: Royal Institute of Technology, 1962

Malmberg, B., Distinctive Features of Swedish Vowels: Some Instrumental and Structural Data, 1956, _For Roman Jakobson_, pp. 316-321

Malmberg, B., Réflexions sur les traits distinctifs et le classement des phonèmes, 1967, _To Honor Roman Jakobson,_ pp. 1247-1251

SOME PHONETIC CORRELATES OF EMPHATIC STRESS IN SWEDISH

Anders Löfqvist

This paper contributes some information on subglottal pressure, oral pressure, fundamental frequency and segment duration during a Swedish utterance in which emphatic stress is systematically shifted from one word to another. The term "emphatic" is used in the broad sense of "extra" stress in addition to the "normal" word and sentence intonation. It seems unclear to me whether "contrastive" might have been a more appropriate term and whether there are any differences between these two types of stress - emphatic and contrastive - as discussed by Bierwisch (1966). Since it is beyond the scope of this paper to discuss matters of terminology and different degrees and types of stress nothing more will be said about them here, cf. Lehiste (1970), Lieberman (1970), Atkinson (1973) for reviews.

Method

Subglottal pressure was recorded by a no 18 gauge spinal needle inserted through the crico-thyroid membrane; the needle was placed in a specially designed collar which prevented it from entering too far into the subglottic space and also from moving during the recording session. At the insertion the bevel of the needle was placed parallel to the airflow in order to avoid spurious pressure recordings, cf. Hardy (1965). A polyethylene tube, 14.5 in. long and with an inner diameter of 0.118 in., connected the needle to a differential pressure transducer, Statham PM 131 TC $\pm$ 2.5. Oral pressure was sensed by another polyethylene tube, 17.5 in. long and 0.118 in. inner diameter introduced into the pharynx through the nose; the tube was coupled to a second differential pressure transducer of the same type. The different lengths of the tubes were chosen to make the frequency response of the two recording systems as equal as possible since the length of the needle was added to that of the tube for subglottal pressure, cf. Edmonds et al. (1971). The output from the transducers was amplified, low pass filtered at 200 Hz and recorded together with the speech signal on a FM tape recorder, Consolidated Electrodynamics Corp. VR 3300, at a recording speed of 3 3/4 ips.

A separate recording of the speech signal was made through an Electrovoice 664 dynamic microphone on a Scully 280 tape recorder at a recording speed of 7.5 ips.

Throughout the recording session the pressure signals were constantly monitored since the tubes and especially the needle often clogged and had to be cleaned to give a correct pressure signal.

The pressure and F_o measurements were made on a LINC-8 computer with a program written by Robert Krones. Before the pressure signals were fed to the computer they were low pass filtered once more, oral pressure at 110 Hz and subglottal pressure at 180 Hz, to remove most of the voicing.

For the durational measurements the speech signal was recorded on a Mingograph at a paper speed of 100 mm/sec. and measured by hand.

Material and measurements

The speech material consisted of the utterance "Kvinnor får lättare snuva än män" [ˈkvIn:ʊɹfɔˈɭɛt:aɹəˈsnɯ:vaɛnˈmɛn:]
In version 1 it was pronounced with a "normal" intonation pattern for a declarative utterance without any emphasis. In version 2 emphatic stress was placed on the word "kvinnor", in version 3 on "lättare", in version 4 on "snuva" and in version 5 on "män". The different versions were read several times by a native male speaker of Southwestern Swedish (västgötska) and seven repetitions of each version could be used for the analysis.

Subglottal pressure was measured at the points given in Table I. These points are peak subglottal and oral pressure for the voiceless obstruents /k, f, t, s/; the middle of the vocalic segment of the vowels /I, U, ɛ, a, ɯ/; peak subglottal pressure for the final vowel /ə/ in "lättare" which in this case tended to occur towards the end of the vocalic segment; the middle of the nasal consonant /n/ in "kvinnor".

Fundamental frequency was measured at the same points as subglottal pressure during the vowels and the nasal consonant.

Results

Figures 1-5 show representative examples of the recorded parameters during the production of each version of the test sentence and the results of the measurements of subglottal pressure, fundamental frequency and

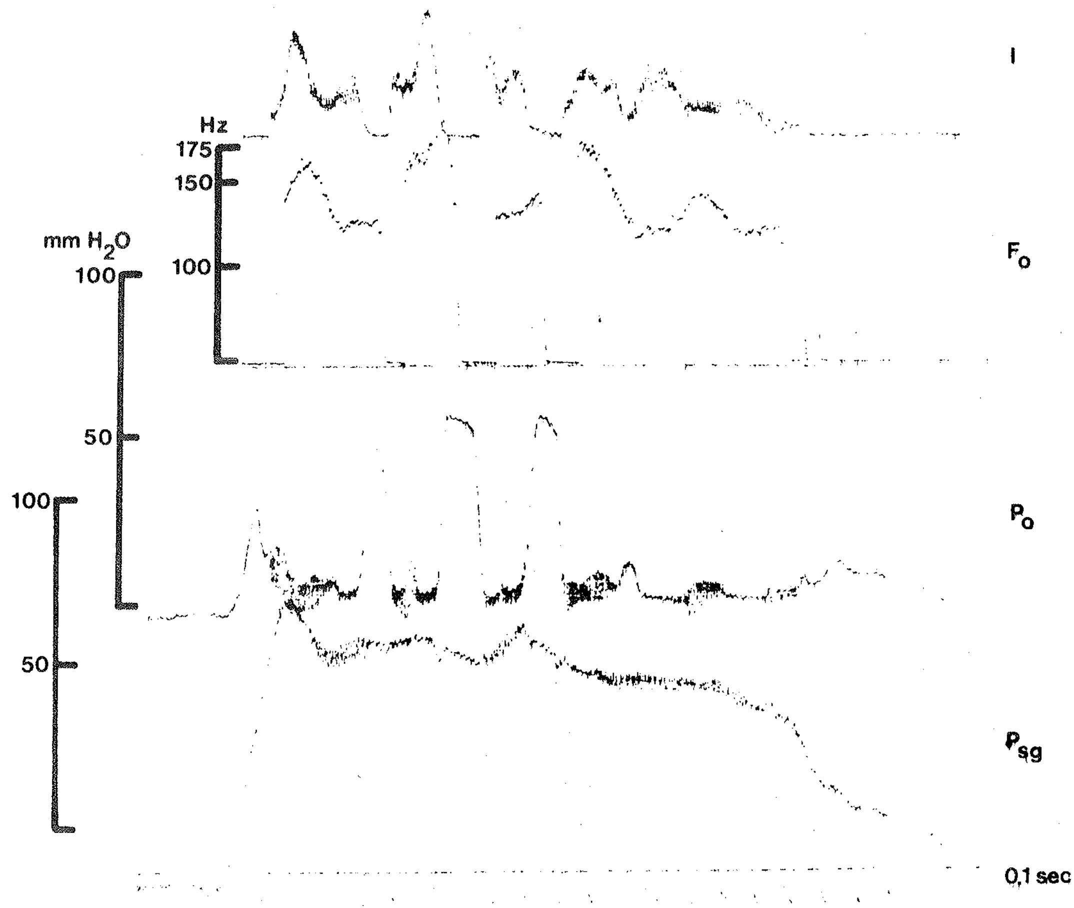


Figure 1. Version 1 of the test utterance - normal intonation. The curves represent from top to bottom: intensity, fundamental frequency, oral pressure, subglottal pressure and time, 0.1 sec.

segment duration are summarized in Tables I-III respectively.

Emphatic stress is associated with an increase in subglottal pressure during the emphasized word. The difference in P_{sg} between words with and without emphatic stress is, however, not the same throughout the utterance but appears to vary according to the position of the word within it; it is greater in the beginning and at the end and slightly less in between. With the exception of the final vowel /ə/ in "lättare" the difference is about the same for the stressed and unstressed vowels in the polysyllabic words although the absolute pressure is higher during the stressed vowels.

For the consonants the difference between the emphatic and non-emphatic versions is greater in prestress than in poststress position.

Another tendency which can be noted is that subglottal pressure during the non-emphasized words in an utterance where one word carries emphatic

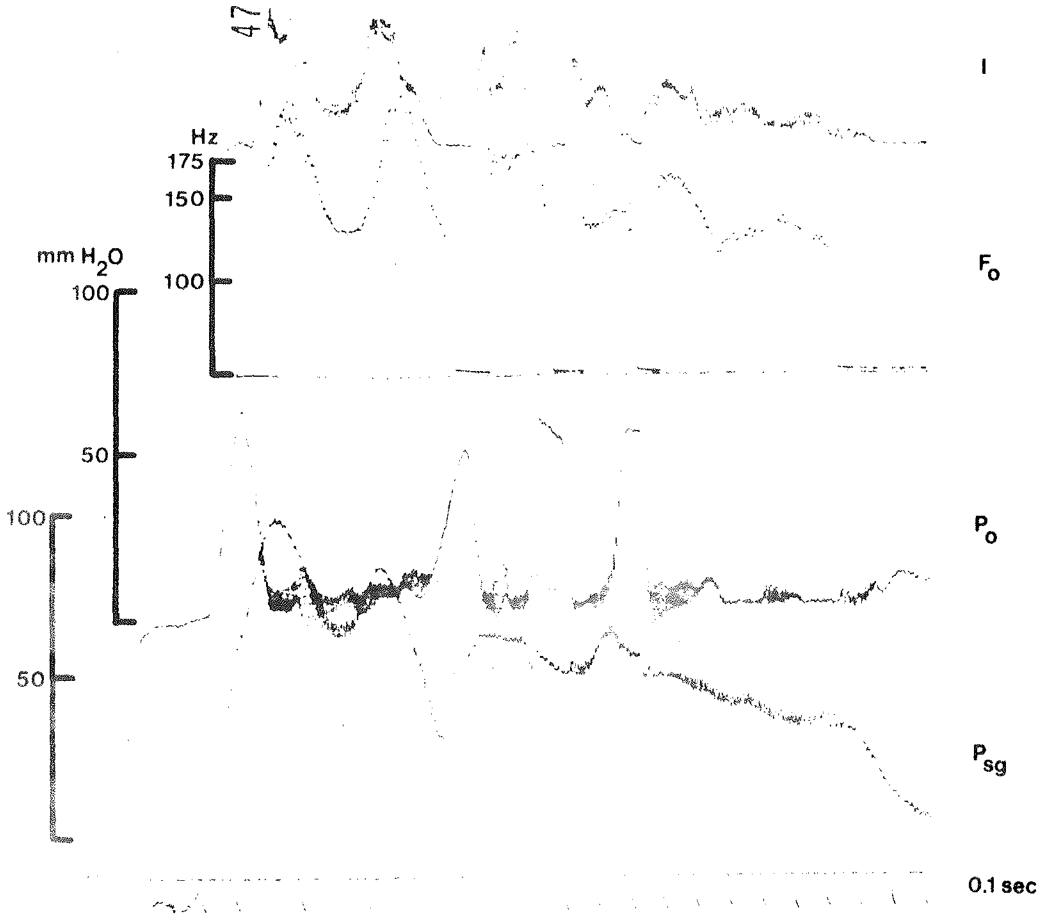


Figure 2. Version 2 – emphasis on "[illegible]".

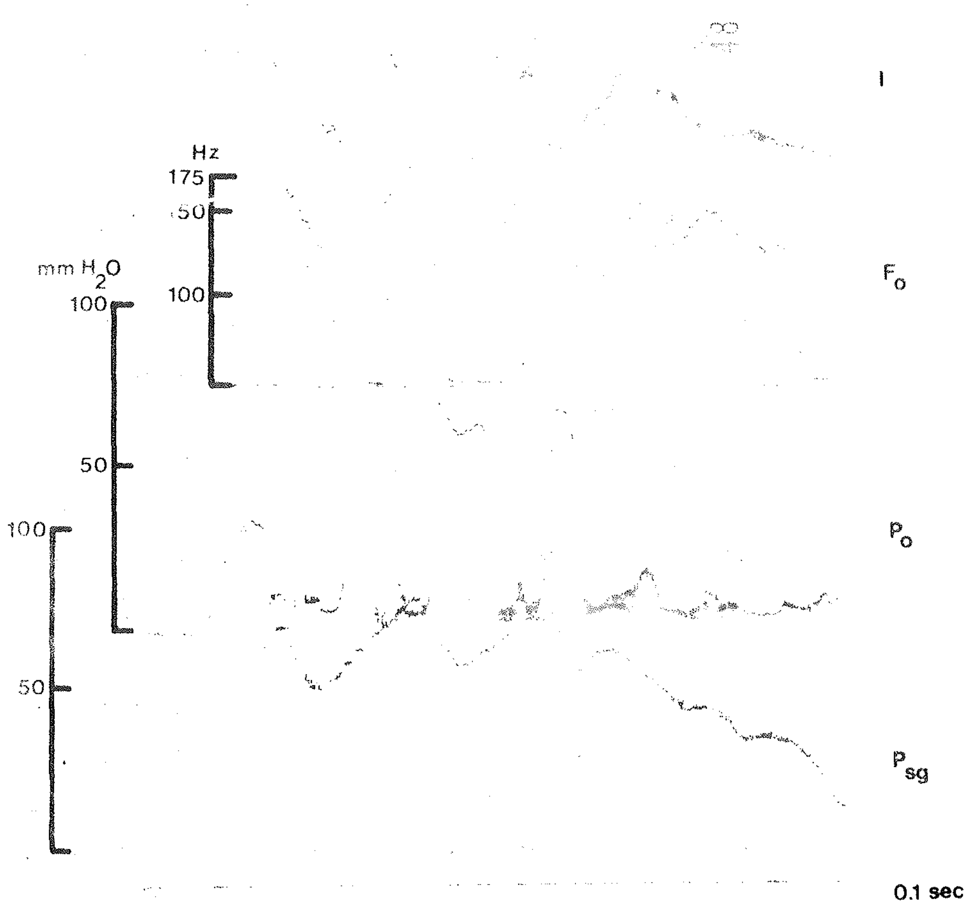


Figure 3. Version 3 – emphasis on "lättare".

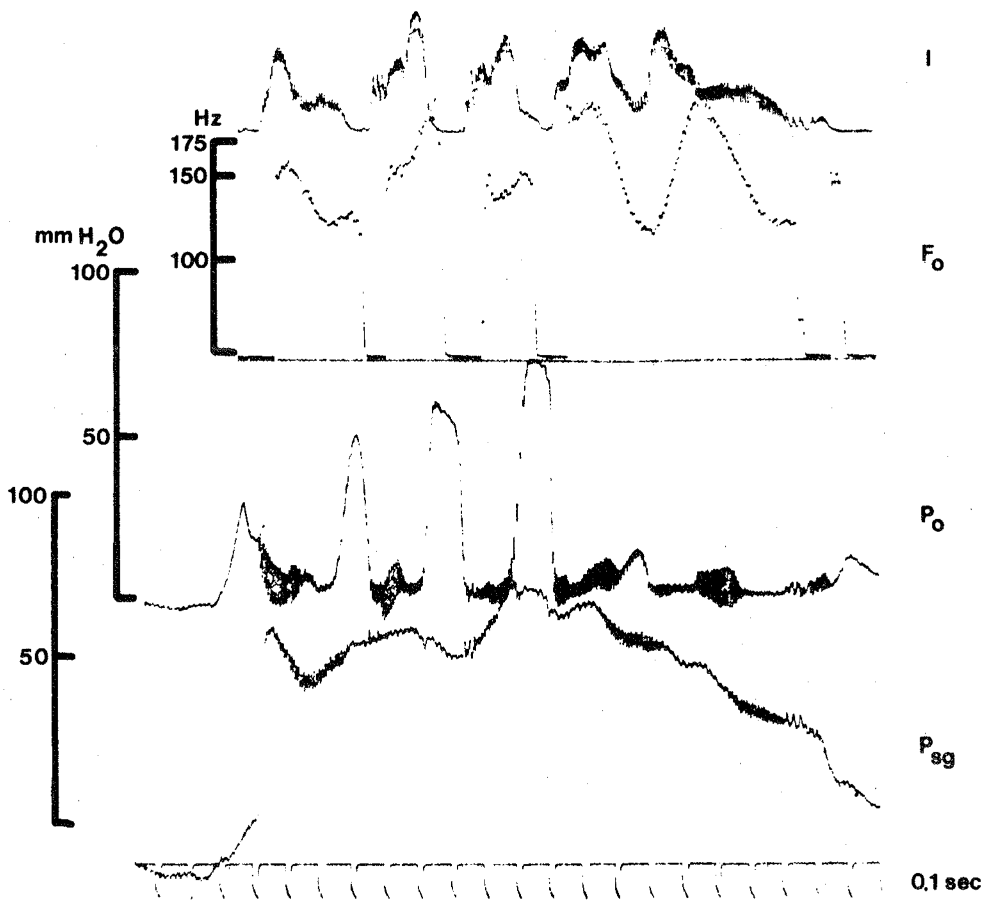


Figure 4. Version 4 - emphasis on "snuva".

49

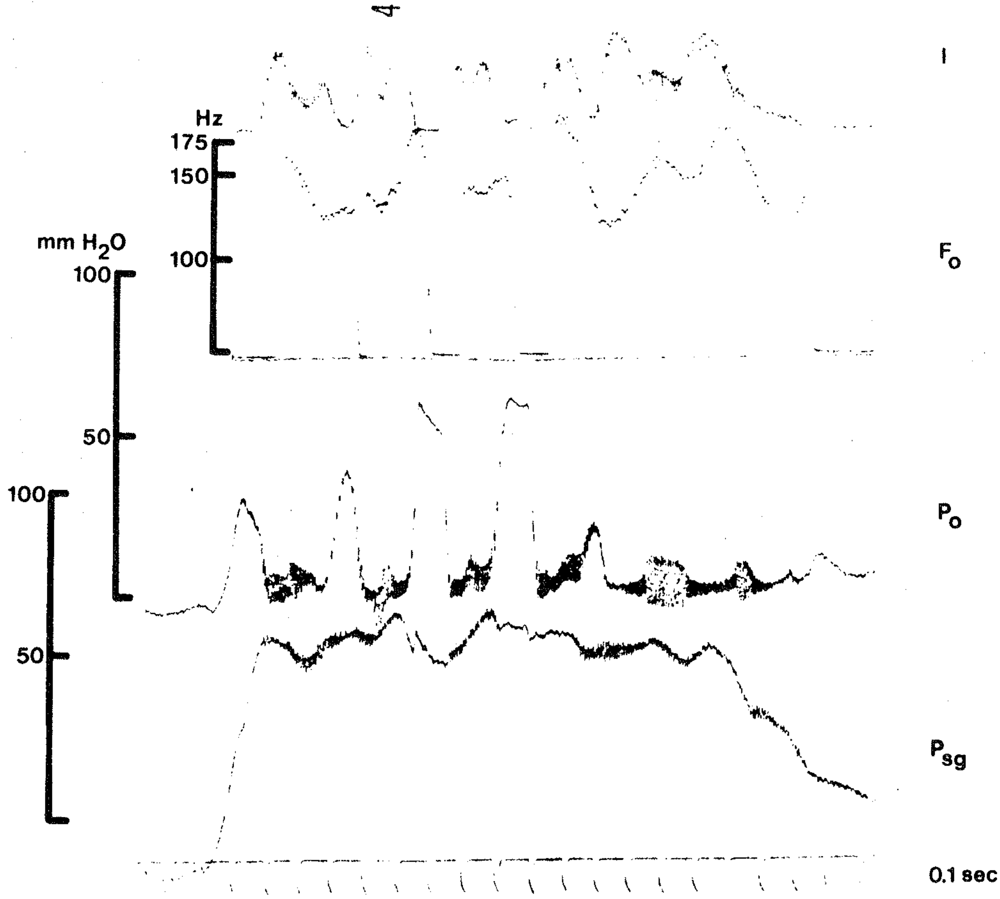


Figure 5. Version 5 - emphasis on "män".

stress is lower than when no emphatic stress occurs in the same utterance. For some reason subglottal pressure during the word "snuva" appears to be an exception to this.

The F_0 data in Table II go in the same direction as the pressure data in that emphatic stress is correlated with an increase in fundamental frequency. In general the increase in F_0 accompanying emphatic stress is much greater during the unstressed than during the stressed vowels. The only exception to this generalization is the vowel in "män" but this word is monosyllabic and also occurs in utterance final position. During the nasal consonant /n/ in "kvinnor" fundamental frequency is actually somewhat lower in the emphatic than in the non-emphatic version of this word. The F_0 variations during the non-emphasized words in an utterance where one word has emphatic stress seem to be the same as when no emphatic stress occurs in the same utterance.

Increased segment duration is another correlate of emphatic stress. This increase is especially large for the phonologically long segments of a word, i.e. the consonants /n, t, n/ in "kvinnor", "lättare" and "män" as well as the vowel /ʉ/ in "snuva".

Discussion

The function of emphatic stress seems to be to make the emphasized segment stand out against the rest of the utterance and the importance of fundamental frequency, duration and intensity for the perception of both normal and emphatic stress has been shown for several languages, Lehiste (1970); F_0 has often been claimed to be the most important factor whereas intensity would be of minor importance. One of the conclusions arrived at in the study of the intonation of Southern Swedish by Hadding-Koch (1961) is that "Tonal features are probably the chief indicators of contrastive stress, i.e. pointing to one of two or more possibilities, usually by means of a marked rise or rise-fall. In Accent-2 words this contrastive pointing may be realized by the unstressed syllable instead", (Hadding-Koch, 1961, p. 190).

Substituting the term "emphatic" for "contrastive" we see that the present results are in agreement with this statement. First, the fundamental frequency curve spans a much larger range of frequencies when a word has emphatic stress than otherwise. This is illustrated by the word

Table I. Subglottal pressure at selected points of the utterance, mm of water.

Sentence version	k	(v)	i	nn	o	(r)	f	(år l)	ä	tt	a	(r)	e	s	(n)	u	(v)	a	(än m)	ä	(n)
1	39		69	56	63		64		63	58	57		66	61		54		49		43	
2	53		86	56	78		50		59	51	51		60	56		52		43		37	
3	38		66	55	57		63		73	62	67		70	62		58		47		36	
4	39		62	50	55		60		61	57	54		66	70		64		56		39	
5	31		61	51	54		59		61	55	53		62	58		57		52		55	
with emphasis	53		86	56	78				73	62	67		70	70		64		56		55	
without emphasis	39		69	56	63				63	58	57		66	61		54		49		43	
difference abs	14		17	±0	15				10	4	10		4	9		10		7		12	
difference %	36		25	±0	24				16	7	18		6	15		19		14		28	

Table II. Fundamental frequency at selected points of the utterance, Hz.

Sentence version	(kv)	i	nn	o	(r får l)	ä	(tt)	a	(r)	e	(sn)	u	(v)	a	(än m)	ä	(n)
1.		167	130	127		185		129		143		171		122		127	
2		186	122	193		177		127		137		163		121		119	
3		159	130	121		191		152		193		187		125		118	
4		157	126	121		184		127		144		189		163		122	
5		159	130	124		176		135		147		178		125		169	
with emphasis		186	122	193		191		152		193		189		163		169	
without emphasis		167	130	127		185		129		143		171		122		127	
difference abs		19	–8	66		6		23		50		18		41		42	
difference %		11	–6	52		3		18		35		11		34		33	

Table III. Duration of selected segments of the utterance, msec.

Sentence version	(kv)	i	nn	o	(r får l)	ä	tt	a	r	e	s	n	u	v	a	(än m)	ä	n
1		98	78	93		88	129	47	18	78	106	42	136	56	78		145	141
2		121	176	158		93	124	54	16	81	109	44	126	58	84		129	132
3		76	70	89		106	193	86	19	86	119	39	145	59	88		138	141
4		85	73	91		95	120	58	22	75	134	41	201	76	118		150	155
5		74	71	84		91	119	51	18	84	116	40	132	61	91		155	224
with emphasis		121	176	158		106	193	86	19	86	134	41	201	76	118		155	224
without emphasis		98	78	93		88	129	47	18	78	106	42	136	56	78		145	141
difference abs		23	98	65		18	64	39	1	8	28	–1	55	20	30		10	83
difference %		23	126	70		20	50	83	6	11	26	–2	40	36	38		7	59

"kvinnor"; not only are the F_0 maxima higher during this word in version 2 of the utterance but the minimum is also slightly lower, cf. Table II and Figs. 1-2.

Second, the unstressed vowels in the words "kvinnor", "lättare" and "snuva" show a much greater increase in F_0 as a result of emphatic stress than do the stressed vowels in the same words. These words all have tonal accent 2 and the emphatic stress thus seems to be carried by the unstressed syllables in this case.

The patterns of P_{sg} variations found in the present study are in accordance with those reported in other investigations of the same parameter, cf. the references given below. During a declarative utterance with normal intonation subglottal pressure is highest at the beginning and then drops during the utterance to reach its minimal value towards its end. Overlaid on this pattern there are peaks occurring during the syllables carrying sentence stress. When emphatic stress is placed on a word higher peaks occur for it. We note on the other hand that when the emphasized word is at the end of the utterance there is an increase in P_{sg} but this is not necessarily the peak subglottal pressure of the utterance which it tends to be in other positions. The increase in subglottal pressure for emphatic stress is thus relative rather than absolute and the variations can probably in part be accounted for by reference to the pressure-volume relationship of the lung, Rahn et al. (1946).

The expiratory force of the relaxation pressure increases with increasing lung volume and it is thus greater at the beginning than at the end of an utterance since the volume of air in the lungs has decreased during the utterance. Assuming the net addition of expiratory effort for emphatic stress to be the same all over the utterance one would expect the difference between the emphasized and non-emphasized syllables to be the same irrespective of their position but that absolute pressure is higher in the beginning and successively lower the closer we get to the end of the utterance. Although the scarcity of the material makes it unwise to draw any far reaching conclusions it appears that the latter part of the assumption is true whereas the former is not since the increase in pressure for emphatic stress tends to be greater in the beginning and at the end of the utterance, cf. Table I. This can hardly be due to any mechanical factors but are independently controlled manifestations and the reduction of subglottal pressure during the non-emphasized words in an utterance where one word has emphatic stress would seem to be of the same kind.

From Figs. 1-5 it is evident that the F_o and P_{sg} curves follow each other rather closely, especially during the words with emphatic stress - perhaps with the exception of the vowel /a/ in "snuva". It is well known that, other factors being equal, an increase in subglottal pressure causes fundamental frequency to rise. The magnitude of this influence appears, however, rarely to exceed 5 Hz per cm of water, Öhman and Lindqvist (1965), Ohala (1970), Hixon et al. (1971). Increasing F_o is, on the other hand, usually accompanied by increasing glottal resistance, Broad (1968), Perkins and Yanagihara (1968), Shipp and McGlone (1971), which in itself would cause subglottal pressure to rise. Since the activity of the laryngeal muscles was not studied in the present investigation nothing can be said about the specific contribution of respiratory and laryngeal mechanisms to F_o control nor about the influence of glottal resistance on P_{sg}.

Acknowledgement

The recordings of subglottal pressure were made during a stay at the Phonology Laboratory, Department of Linguistics, University of California, Berkeley, and I am indebted to its director, William S-Y. Wang, for giving me the opportunity to use its facilities and to John Ohala and Robert Krones for technical assistance and discussions during the experiments; a special thank is due to James Schmitt who performed the tracheal puncture. This work is supported in part by a grant from the National Science Foundation, GS 2386 A1, to the Phonology Laboratory.

References

Atkinson, J.E. 1973. Aspects of intonation in speech: Implications from an experimental study of fundamental frequency. Dissertation, University of Connecticut

Benguerel, A-P. 1973. Corrélats physiologiques de l'accent en français. Phonetica 27:21-35

Bierwisch, M. 1966. Regeln für die Intonation deutscher Sätze. Studia Grammatica 7:99-201

Broad, D. 1968. Some physiological parameters for prosodic description. SCRL, Santa Barbara, Monograph no 3

Carlson, R., Erikson, Y., Granström, B., Lindblom, B. and Rapp, K. 1974. Neutral and emphatic stress patterns in Swedish. Papers from the Institute of Linguistics, University of Stockholm 23:27-41

Carton, F. and Marchal, A. 1972. La pression sous-glottique, corrélat de la mise en valeur dynamique ("accent d'insistance") en français contemporain. Proc. 7th Intl. Cong. Phonet. Sci. 871-879. The Hague: Mouton

Edmonds, T.D., Lilly, D.J. and Hardy, J.C. 1971. Dynamic characteristics of air pressure measuring systems used in phonetic research. J. Acoust. Soc. Amer. 50:1051-1071

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

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

Hixon, T., Klatt, D. and Mead, J. 1971. Influence of forced transglottal pressure changes on vocal fundamental frequency. J. Acoust. Soc. Amer. 49:105 (A)

Ladefoged, P. 1967. Three areas of experimental phonetics. Cambridge: Cambridge University Press

Ladefoged, P. 1968. Linguistic aspects of respiratory phenomena. Ann. N.Y. Acad. Sci. 155:141-151

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

Lieberman, P. 1967. Intonation, perception and language. Cambridge, Mass.: The M.I.T. Press

Lieberman, P. 1970. A study of prosodic features. Haskins Laboratories, Status Report on Speech Res. 23:179-208

Ohala, J. 1970. Aspects of the control and production of speech. UCLA Working Papers in Phonetics 15

Perkins, W.H. and Yanagihara, N. 1968. Parameters of voice production: I Some mechanisms for the regulation of pitch. J. Speech and Hear. Res. 11:246-267

Rahn, H., Otis, A.B., Chadwick, L.E. and Fenn, W.O. 1946. The pressure-volume diagram of the thorax and lung. Am. J. Physiol. 146:161-178

Shipp, T. and McGlone, R.E. 1971. Laryngeal dynamics associated with voice frequency change. J. Speech and Hear. Res. 14:761-768

Strenger, F. 1958. Mesure de la pression d'air sous-glottique, de la pression acoustique et de la durée de la prononciation des différents sons du langage suédois au cours de la phonation. Journal Français O R L 30:101-114

Öhman, S. and Lindqvist, J. 1965. Analysis-by-synthesis of prosodic pitch contours. STL-QPSR 4:1-6

NIVEAUX, CHOIX ET SYSTEMES APPROXIMATIFS DANS LE LANGAGE

Bertil Malmberg

Les termes qui se réfèrent aux paires conceptuelles *langue* – *parole*, *compétence* – *performance*, *modèle* – *manifestation* pullulent dans la littérature linguistique, surtout depuis les débuts du mouvement saussurien et du structuralisme. Une dichotomie analogue est reflétée aussi dans l'opposition glossématique entre le *système* et le *procès* (le texte) et dans le contraste plus généralement philosophique entre *invariance* et *variance*. Les problèmes derrière ces dichotomies ont eu une nouvelle actualité avec la naissance de la grammaire générative. Ils ont aussi, depuis longtemps, leur répercussion dans les tentatives d'analyse des styles et des niveaux de communication en linguistique. Plusieurs questions théoriques et descriptives se rattachent à ces distinctions. Il y a d'abord chez Saussure lui-même une définition du concept de *langue* qui admet plusieurs interprétations, depuis la forme purement abstraite jusqu'à l'usage établi (discutées par Hjelmslev[1]). Il y a aussi celle de la compétence et de la grammaticalité chez Chomsky, et ainsi de suite. On n'approfondira pas ces problèmes ici[2].

Déjà à l'époque des premiers phonologues, on discutait le niveau à choisir pour l'établissement des systèmes d'invariants (systèmes phonologiques), et dans ce contexte aussi la notion de la pertinence et de la distinctivité. Le point de vue synchronique avait rendu nécessaires une délimitation et une définition de l'objet de la description. L'introduction des principes de la théorie de l'information en linguistique avait actualisé la notion de redondance et son rôle linguistique. La notion de pertinence rendait indispensable une précision du niveau de la communication auquel la fonction distinctive appartenait. On en avait un point de départ dans les ouvrages de J. von Laziczius et de Karl Bühler. Ce qui était redondant au niveau cognitif (fonction de symbole dans le système de ce dernier), pouvait être distinctif au niveau émotionnel, etc. Une longue durée vocalique pouvait exprimer l'emphase dans une langue sans quantité phonologique (le dit " emphaticum " de Laziczius[3]). D'autres phénomènes linguistiques (phonétiques ou autres) peuvent caractériser l'individu ou le groupe et jouer ainsi un rôle dans les contacts humains[4]. Ce serait dire une banalité que de rappeler ici la différence qu'il y a entre un indice de qualités individuelles, indépendantes de la langue employée (âge, sexe, état de santé) et le symptome d'appartenance à tel ou tel groupe, interprété en vertu d'une expérience de variations dialectales ou sociales et contre le fond d'une norme ou d'un usage valables pour telle ou telle population. L'interprétation de ce dernier type de traits suppose donc une connaissance d'un ou de plusieurs codes.

Je résume brièvement les points de vue exposés dans le travail de jeunesse cité dans la note 2. La notion de distinctivité ou de pertinence n'est valable qu'à un niveau déterminé de la communication. Si nous disons que la durée vocalique est pertinente en finnois, non pertinente en espagnol, ceci veut dire que, dans la première langue, deux éléments lexicaux ou grammaticaux (deux signes) peuvent être différenciés à l'aide du choix d'un phonème long ou d'un phonème bref (fi. tulee ' il vient ' ~ tuulee ' il fait du vent ', ou tule ' viens ' [impératif] ~ tulee ' il vient '), tandis que, dans l'autre, cette possibilité est inexistante (esp. hermoso ayant le même sens lexical si le /o/ accentué est bref ou long), ce qui n'exclut pas que l'hispanophone puisse se servir d'une durée allongée pour exprimer l'emphase ou l'affectation (ce qu'il fait du reste assez couramment)[5]. La conclusion en sera que n'importe quel élément de l'énoncé linguistique, à condition d'être perceptible et susceptible de contraster avec un autre, peut assumer un rôle – linguistique ou extralinguistique – dans la communication humaine. Toute différence (de son, de morphème, de construction) peut donc être dans un sens ou dans un autre pertinente. C'est une raison pour préférer (avec Jakobson) le terme " distinctif " à ceux de " pertinent " (Martinet) et de " relevant " (Troubetzkoy, etc.). Toutefois, l'emploi même de ce terme préféré présuppose la remarque constante, sous-entendue ou explicite, que la distinctivité en cause ne joue qu'à un niveau choisi. C'est contre le fond des ressources disponibles à ce niveau que l'élément en question – que ce soit la valeur phonétique (timbre, intonation), le mot choisi (archaïsme, vulgarisme, dialectalisme) ou la construction préférée (un ordre des mots littéraire ou négligé) – remplit sa fonction linguistique.

Nous en sommes arrivé ainsi à mentionner quelques facteurs importants pour la détermination du niveau du langage qu'on appelle couramment style, concept dont nous réservons l'analyse pour une autre occasion[6]. Il importe pourtant de mentionner ici que les styles appartiennent aux niveaux linguistiques parmi lesquels le locuteur (l'écrivain) fait son choix. Ce choix peut être conscient et représenter une intention particulière ou une préférence personnelle. Mais il peut aussi être inconscient. Il représente alors le maximum dont un individu est capable. Le locuteur vise à la norme valable dans son entourage mais n'arrive pas à la réaliser. C'est le cas de l'enfant, du handicappé et de celui qui est socialement ou culturellement inférieur (enseignement défectueux, situation sociale pénible, etc.). On ne parle normalement de style que dans les cas où le choix de niveau est conscient en tant que possibilité parmi d'autres. Nous ne mentionnons que brièvement les questions de niveau et d'état de langue en attirant l'attention aux faits d'interférence qui peuvent se produire, pour passer ensuite aux problèmes descriptifs posés par la complexité de tout corpus linguistique de quelque étendue. Pour ce faire, nous partons de l'idée d'un modèle général contenant toutes les possibilités disponibles à l'intérieur d'une communauté linguistique et comprises comme un maximum de ressources communicatives.

Il m'a paru instructif de voir l'ensemble des ressources d'une langue et des règles qui en déterminent le fonctionnement comme un modèle, que nous appellerons ici modèle général - d'une complexité extrêmement grande. Nous faisons abstraction ici de cette complexité. Complexe ou simple, c'est ce modèle qui sert de dénominateur commun aux nombreuses compétences individuelles dont les faits concrets de parole sont à leur tour les réalisations. Le modèle général n'est jamais maîtrisé dans son ensemble par tous les membres d'une communité linguistique[7].

L'ensemble linguistique d'une communauté n'est pas homogène. Elle connaît des variations spatiales et sociales, horizontales et verticales. J'ai développé ailleurs[8] l'idée d'état de langue comme une couche arbitrairement délimitée, dans les deux sens, par le descriptiviste. Ce concept se réfère à un nombre d'énoncées, c'est-à-dire à un corpus dont le linguiste abstrait le système. Nous renonçons à traiter ici l'ordre des rapports entre différents systèmes ainsi établis pour ne considérer que l'acte de la création des énoncés comme la matérialisation d'un modèle. C'est la familiarité avec ce modèle qui permet au locuteur et à l'écrivain de s'exprimer, à l'auditeur ou au lecteur de comprendre.

Où le locuteur trouve-t-il son modèle? Deux alternatives s'offrent. Ou bien c'est tout simplement son expérience de sa langue telle qu'il l'entend dans son entourage depuis son enfance et dont il s'est formé inconsciemment une idée (une " théorie "). Ou bien c'est un modèle - littéraire ou stylistique - qu'il choisit consciemment et dans un but spécifique. S'il parle ou écrit une langue étrangère, le modèle est celui qu'on lui a appris à l'école et dont il a réussi, grâce à un travail méthodique d'apprentissage, à se rendre maître.

Le lecteur a dû voir déjà à travers ces quelques remarques l'idée que le modèle servant de guide à l'individu dans la production linguistique est une approximation du grand modèle général que représente le système des systèmes d'une langue.

En réalité, tout système - quel que soit son degré de perfection - est une approximation d'un tel modèle général, à moins que l'usage en question ne dépasse les limites admises par celui-ci, auquel cas il s'agit d'un système différent. Je prends comme premier exemple les déviations régionales et stylistiques, propres aux habitudes d'un locuteur. Un facteur socio-psychologique intervient ici. Deux cas se présentent. Celui qui parle un dialecte, ou un parler régional différent de la langue officielle, réalise consciemment - peut-être avec fierté - un modèle qui est distinct du modèle général et dont il sait qu'il s'en écarte. Celui qui, tout en voulant réaliser le modèle, se rend inconsciemment coupable de déviations sous l'influence d'un parler local, ne réalise qu'une approximation de celui-ci et manque son but. L'effet peut être le même. La genèse est différente. De même, celui qui parle un sociolecte vulgaire peut le faire à dessein en signe d'une attitude sociale consciente - par exemple pour marquer son indépendance par rapport à l'établissement - ou inconsciemment parce que, faute d'instruction suffisante, il n'arrive pas à appliquer un modèle qu'il regarde en principe comme supérieur. Si la déviation est intentionnelle, le locuteur est content, même fier, de sa langue. Si elle est due à une compétence défectueuse, il peut arriver que, par suite de la réaction de l'entourage, il se rende compte de son infériorité et qu'il en ait honte. On trouve des exemples de ces deux attitudes devant la langue maternelle chez les minorités et chez les immigrés. Ces différences d'attitude peuvent jouer un rôle décisif pour l'évolution d'une langue. La description diachronique a à en tenir compte.

Nous formulons donc la thèse que toute production linguistique, orale ou écrite, se réalise à base d'un modèle qui est une approximation d'un modèle plus général. Ces approximations représentent une échelle qui va de la plus grande perfection possible - où seuls certains faits de vocabulaire et de style le séparent du modèle général – à la plus extrême pauvreté. On retrouve cette pauvreté dans l'évolution linguistique de l'enfant, dans l'aphasie et chez tous ceux qui, à cause de faiblesse intellectuelle ou physique, ne sont pas capables, ou sont devenus incapables, de dépasser un minimum de compétence linguistique. Les analphabètes représentant un degré zéro de compétence écrite. Les illettrés, malgré certaines connaissances de lecture et d'écriture, représentent un degré très bas de compétence écrite mais possèdent souvent une grande facilité d'expression orale. Le degré d'approximation dont nous parlons ne concerne pourtant pas leur facilité de communication mais la distance qui sépare leur réalisation linguistique du modèle, quel qu'il soit.

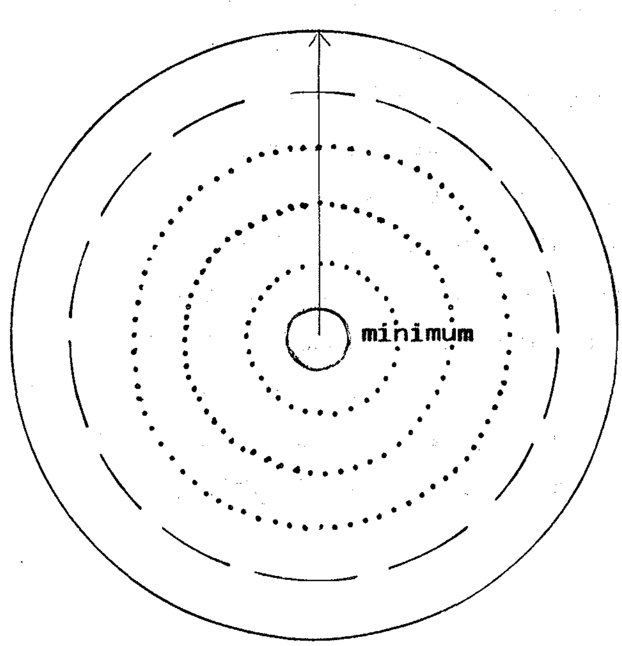
Supposons que tout degré d'approximation du modèle représente un système autonome. Ce n'est pas tout à fait exact, puisque toute couche linguistique d'une société est soumise à la double influence des plus évoluées et des plus pauvres, et puisque l'individu choisit, inconsciemment ou à dessein, différents niveaux selon le contexte dans lequel il parle (ou écrit) et selon le but de son message. Il y a constamment des faits d'interférence qui empêchent la réalisation " pure " du niveau choisi. Dans l'évolution de l'enfant aussi bien que dans l'apprentissage d'une deuxième langue, le niveau atteint à un stade donné sera le système intermédiaire qui correspond à la théorie que l'individu s'est formée de la langue cible. Un exemple en est la généralisation des formes régulières dans la morphologie chez l'enfant (et souvent chez l'élève). L'absence de formes irrégulières[9] est dans ce cas une mesure parmi d'autres du degré d'approximation du modèle. L'absence de /œ̃/ en français est un des facteurs qui mesurent le degré d'approximation du français parisien d'un modèle encore valable avec /œ̃/. Il va sans dire que je simplifie en laissant de côté ici la définition difficile de la norme. Je suppose donc, à tort ou à raison, que le Parisien cultivé tout en disant /ɛ̃/ a toujours le sentiment que /œ̃/ est plus correct que /ɛ̃/ et qu'il supprime un trait qui en principe est toujours présent comme ressource disponible mais non exploitée (en d'autres termes : disparu de la surface mais survivant dans le fond). Dans la mesure où ceci ne serait pas le cas, le /œ̃/ aurait disparu du modèle général, ou plus exactement le système plus complet contenant le /œ̃/ aurait cessé de rendre service comme modèle général. Quand, à l'époque actuelle, les locuteurs de Stockholm réalisent plus qu'autrefois la distinction /e:/ ~ /ɛ:/ (meta /mè:ta/ ' pêcher à la ligne ' ~

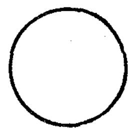
mäta /mɛ̃:ta/ ´ mesurer ´), ceci implique donc qu´ils reviennent à un modèle plus proche du modèle général (qui a toujours maintenu la distinction, autrement celle-ci n´aurait pas pu être réintroduite). Reste à voir si c´est l´enseignement scolaire ou l´influence croissante des provinciaux dans la capitale qui a été le facteur principal dans cette évolution.

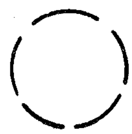
Ce dernier exemple nous rappelle l´impossibilité de mesurer en termes quantitatifs seuls la fidélité du locuteur à la norme. Des statistiques sur la fréquence de /œ̃/ dans le parisien d´aujourd´hui ne nous diront rien sur le nombre des locuteurs qui continuent à sentir comme supérieure la prononciation avec /œ̃/ dans des mots comme brun, parfum, chacun. Ce dernier chiffre sera linguistiquement plus intéressant que le nombre de ceux qui, à l´interieur d´un groupe choisi, ont trouvé inutile, dans les situations concrètes examinées, de réaliser une valeur vocalique dont la fonction leur est parfaitement consciente et qu´ils ne suppriment que par négligence ou par snobbisme. Les enquêtes de type faites-vous une différence entre ... et ... ? sont dans un certain sens plus appropriées que les enregistrements directs. Elles nous renseignent sur les attitudes prises par les locuteurs en face de leur langue.

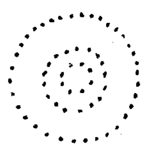
On a souvent critiqué ce genre d´interviews, en disant qu´elles ne renseignent que sur ce que les gens croient dire, non pas sur ce qu´ils disent. J´estime pourtant que, du point de vue du mécanisme linguistique valable dans un groupe, la première information est plus intéressante que la deuxième.

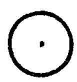
Nous proposons de faire représenter notre modèle par le dessin schématique suivant :



 = modèle général (nombre de locuteur = 0)

 = approximation atteinte par les " locuteurs cultivés "

 = différents degrés d'approximation du modèle (contenant la majorité de locuteurs)

 = minimum (enfants, handicappés ; nombre réduit de locuteurs) ; au milieux le point zéro

Le grand cercle renferme donc le maximum de possibilités de la langue à tous les niveaux (modèle général). Aucun membre de la communauté ne les maîtrise toutes. Le plus petit cercle au milieu représente le minimum auquel la langue est réduite chez les locuteurs qui ont perdu, ou qui ne sont pas arrivés à, une maîtrise acceptable d'un système. Chez l'enfant et chez l'élève ce minimum est une évolution par rapport au zéro marqué par le point central. Les cercles intérieurs représentent différents degrés d'approximation du maximum général. Le plus grand cercle intérieur (traits) est le niveau atteint, en moyenne, par les locuteurs adultes " cultivés " (avec toutes les réserves quant au sens de ce mot). Tel cercle peut signifier le maximum dont certains locuteurs sont capables et en même temps correspondre au niveau choisi à dessein par d'autres. La flèche allant de l'origo jusqu'à la limite extérieure du cercle indique donc d'un côté la direction du chemin à parcourir pour l'enfant, ou pour l'élève ; et de l'autre les différents niveaux choisis dans des buts spécifiques ou sous l'influence d'un contexte quelconque (social, régional, etc.) et par n'importe quel locuteur ou écrivain. Un tel choix peut être représenté par un point sur cette ligne. Il va sans dire que n'importe quel degré d'approximation ainsi symbolisé peut se référer soit à un trait de langue particulier (de phonologie, de grammaire, de vocabulaire), soit à un ensemble de traits ou à un état de langue entier.

Notre dessin schématique a donc aussi pour but de visualiser les mélanges de couches qui caractérisent tout état de langue et particulièrement toute langue de culture, parlée et écrite à l'intérieur d'un domaine politique et culturel, étant en même temps uni et diversifié, homogène et hétérogène, et avec une diversification qui augmente en allant du centre géographique à la périphérie et du haut en bas dans la hiérarchie sociale. Toutes ces couches se rencontrent, s'entre-croisent et interfèrent. N'importe quel point choisi comme représentant de tel ou tel niveau se trouvera donc en réalité entre deux cercles sur le schéma. Il y a dans toute société moderne un bilinguisme ou trilinguisme ou multilinguisme, dont il faut se rendre compte dans différents contextes, scientifiques et autres[10], mais qui posent aussi et surtout des problèmes pour la description linguistique. Le rapport qu'il y a entre l'individu ou le groupe et ces différentes possibilités linguistiques est un aspect important de la réalité sociologique d'abord, mais aussi un aspect de la réalité dont la description est une tâche essentielle du linguiste. Même si, dans toute situation de choix parmi les ressources linguistiques, il y a une opposition entre un idéal (système, norme, usage, mode) et une réalisation concrète (parfaite ou défectueuse), une simple dichotomie de type saussurien ou chomskyen ne suffit pas pour rendre compte de l'extrême complexité des relations entre l'homme et son langage.

Quelle que soit la façon dont nous comprenons le concept de bilinguisme[11], il reste vrai que celui-ci implique toujours une participation sociale et culturelle au moins double. C'est aussi le cas du bilinguisme de petit format connu de toute société un peu complexe. C'est en réalité ce genre de bilinguisme qui pose les plus grandes difficultés au descriptiviste et qui rend la collaboration du locuteur si importante pour une juste évaluation des attitudes de celui-ci, donc pour savoir si une déviation représente de la part du locuteur un choix conscient d'une valeur déterminée ou un niveau conditionné par une incapacité quelconque de la part de celui-ci ou par tel ou tel autre facteur, linguistique ou extralinguistique (un simple lapsus, une indisposition accidentelle, etc.).

Tous les linguistes depuis Saussure ont vu que le jeu entre ce que nous voulons dire et ce que nous disons doit se trouver au centre des préoccupations des linguistes. Tous n'ont pas vu la complexité de ces rapports.

N o t e s

1. " Langue et parole ", Cahiers Ferdinand de Saussure 2, 1943, pp. 29-44 (réimprimé dans Essais linguistiques 1971, pp. 77-89).

2. Je m'en suis occupé dans " Système et méthode ", Vetenskapssocieteten i Lund, Årsbok 1945, pp. 5-52 (Autour du problème langue - parole ; réimprimé dans " Linguistique générale et romane ", 1973, pp. 20-32).

3. Ungarische Jahrbücher XV, 1935, pp. 495-510, etc. Cf. mon livre " Die Quantität als phonetisch-phonologischer Begriff ", 1944, pp. 31-32 et les renvois y faits.

4. Développé plus en détail dans l'ouvrage cité dans la note 2.

5. Le fait que de tels allongements soient le plus souvent accompagnés de phénomènes musicaux et dynamiques n'enlève rien à la valeur démonstrative de notre exemple.

6. Le problème du style fera l'objet d'un chapitre de mon livre en préparation " Signes et symboles ; les bases du langage humain ".

7. Ceci est de toute façon entièrement vrai pour le vocabulaire, sûrement aussi pour certains faits de syntaxe et, surtout, de style. Il n'y a d'autre part guère de doute qu'une grosse majorité des locuteurs d'une communauté linguistique maîtrisent la phonologie de la langue et que les faiblesses de morphologie se limitent à certains faits exceptionnels de conjugaison et de déclinaison - pluriels d'emprunts, etc. - qui peuvent sans doute échapper à certains.

8. Voir mon rapport au X^e Congrès international des linguistes, Bucurest 1967 (impr. 1968) et mon article " Description synchronique et état de langue " (Phonétique et linguistique romanes. Mélanges offerts à M. Georges Straka, I, 1970, pp. 223-229 ; réimprimé dans " Linguistique générale et romane ", 1973, pp. 155-159).

9. On sait que l'enfant apprend souvent d'abord quelques formes irrégulières fréquentes pour les remplacer ensuite par des formes régulières incorrectes, une fois le système appris, et que ce n'est qu'après avoir été corrigées par l'entourage que ces formes irrégulières cèdent finalement la place aux formes correctes.

10. Je pense par exemple aux reflets des différences sociales dans l'enseignement, lequel favorise sans doute les élèves déjà familiarisés avec le niveau de langue propre à l'école au détriment de ceux qui se trouvent devant la tâche d'y accéder. Un cas extrême est celui des minorités linguistiques.

11. Il y en a qui ne regardent comme bilinguisme que le cas plutôt exceptionnel d'une maîtrise également parfaite de deux langues, d'autres qui regardent comme bilinguisme simplement la possibilité de se faire comprendre en plus d'une langue. Je ne discute pas ce problème ici.

SOME THOUGHTS ON READING AND WRITING

Kerstin Nauclér

Linguistic proficiency and linguistic activity are taken for granted by most of us. Certainly, every parent regards the first attempts of his child to talk as most remarkable and unique but all the same as completely natural. All children are expected to start talking and, as is well known, nothing can prevent a normal child from picking up the trick how to speak and understand the language of his linguistic community. To acquire other linguistic skills such as reading and writing, however, the child has to be sent to school to get instruction. In the hands of a trained teacher the child is expected to learn to read and write, albeit not as easily as he learned to talk.

Speaking and understanding are primary linguistic activities and they are basic to reading and writing proficiency. By the age of 4, the vocabulary of the child is large, his language is fluent and his memory span for heard words is quite adequate for comprehending simple sentences (Conrad 1972). Yet he is generally not sent to school until the age of 6 or 7. What more are we waiting for to happen? According to Mattingly (1972) the primary linguistic ability alone is not sufficient, since reading is highly related to the speaker-hearer's awareness of this linguistic ability. It is not enough to be able to listen and talk, one must be aware of this activity too. The pleasure some children show in playing with words as in punning and rhyming are obvious signs of this awareness, as is also the invention of "secret" languages. Maybe this awareness is reinforced by the reading skill so that it might be more appropriate to talk about a mutual influence.

Conrad (pp. 217 ff.) argues that what the child of 4 has not yet developed is a phonological short-term memory. Such a memory is necessary to enable him to remember the segment just processed while working on the next one. Without this memory no word would emerge from the letter-to-sound processing. Conrad draws this conclusion from an experiment with adults who showed a preferred use of what he calls the phonological short-term memory (STM) code when recalling visually presented verbal items with either similar or dissimilar sounds. Although the items in one of the two lists were visually quite different, this information was apparently ignored.

A similar paradigm was then used for young children with mental ages ranging from 3 to 11 years. Two sets of coloured pictures were used. The names of the pictures were all familiar to the subjects. One set consisted of "cat", "rat", "bat", "hat", "mat", "man", "tap", "bag" and the other of "fish", "girl", "bus", "spoon", "horse", "train", "clock" and "hand". The children had to match from memory a certain number of cards (which were turned face down after having been shown and named by the experimenter) with the cards from one of the two vocabulary sets. Up to the age of 5 there was no difference in the ability to recall either of the two sets. Beyond 5 years there was an increasing advantage for the non-rhyming set. Conrad concludes that before this crucial age it is irrelevant to STM performance whether or not the pictures have rhyming names but from the age of 5 or 6 years the subjects perform as if they were using phonological information as STM code.

Even though the child has some awareness of his linguistic activity and even if he has developed a phonological STM code, we cannot be sure that he will have no problems in mastering the language in its written form. Lots of apparently normal children fail to learn to read or learn only very slowly. One possible reason is the visual similarity of the letters. All the vowels except "i" are small and round, and yet they should be associated with very different sound qualities. Such a relation must be found illogical. During his short life the child has learnt that a picture of a cat or a dog is a cat or a dog no matter if the picture is turned on its head or to the right or the left. It must be confusing to discover that one letter turns into a new one if turned in different directions, as is the case with "b", "d", "p", and "q".

To a beginner the letters mean only form, not sound. This fact was illustrated very clearly once in a test given to the beginners in the schools of Great Britain. The instructions were to continue to place the letters of the alphabet on two rows in the following sequence:

```
A       E F   H
  B C D     G
```

This test became known because only beginners were able to solve the task in the intended way. Letters mean form only as long as they do not have an acoustic content. Once the letters have been given this property it is no longer possible to perceive them as pure graphic forms. Deaf mutes tend to confuse tachistoscopically presented letters that are similar in form, while people with good hearing and deaf with articulated speech confuse letters that represent similar sounds (Conrad pp. 227 ff.).

Lindell (1964) has shown that spelling problems do not mean deficient visual perception. Pupils with spelling problems have no difficulty in detecting differences between ordinary pictures but gain a lower score than pupils without spelling difficulties when tested for differences between word pictures.

Rozin et al. (1971) report very good results from an experiment with nine American children with severely retarded reading ability who were taught to read English by using Chinese characters. Two conclusions can be drawn from this experiment. Either the Chinese characters are too different in form to be confused or "phonetic mapping" is avoided when reading logograms.

In his neuropsychological investigation of reading disabilities, Doehring (1968) has shown that reading disability was most highly correlated with visual and verbal tasks that required sequential processing of the related material.

Isabelle Liberman (1973) is seeking the causes of reading errors in an unawareness of the phonemes of the spoken words. In her opinion, a child of 5-6 years cannot segment a word into phonemes, but has no problems when asked to count the syllables of the word. This argument is consistent with data from Japan. According to Makita (1968), fewer than 1 % of the Japanese pupils have reading difficulties. In the West the corresponding percentage is between 10 and 20. As the difference cannot be physiological it must be related to the different writing systems. The Japanese have two different types of characters, viz. "kanjis" (ideograms) and "kanas" (phonograms). The kanas represent 46 sound units, mostly the consonant-vowel type syllable. The kanji characters number about 1,800 and have inherent meaning as well as phonological form. In school, the children are first taught to recognize and reproduce the kanas, and then to represent the lexical items by the kanjis. It is a common observation that when a person forgets how to write the kanji for infrequent words (there are such a large number of characters to be mastered), he often represents such words with the kana equivalents (Sasanuma & Fujimura 1971).

Many skilled readers claim they go directly from the graphic picture of the word to its semantic representation without using any kind of "inner speech" or "latent articulation". There have been numerous discussions about such neuromotor activity during silent reading. Edfeldt (1960) in an investigation regarding EMG activity in laryngeal muscles during silent reading found less activity in good readers than in poor. He also found that an easy text resulted in less silent speech activity than did a difficult one. Smith (1973) asks how the maning of a sentence that is not available during unvocalized silent reading will suddenly appear when words are subvocalized. Conrad maintains the opinion "that articulation almost always occurs, that it is probably task-relevant, but that sound evidence that it is necessary is lacking" (p. 211). In some recent articles Klapp et al. (1973) and Pynte (1974) discuss a mechanism of readiness for pronunciation during the reading process.

In order to find out whether there was any kind of phonetic activity when reading non-alphabetical written words, Erickson et al. (1973) conducted an experiment on the visual, phonetic and semantic confusability of the Japanese kanji characters. The subjects, all native Japanese, reported not being aware of using any kind of silent speech or acoustic memory, which of course must be considered as redundant in connexion with non-alphabetical text. The subjects were shown a series of characters and then asked to name a subsequent character. Confusion turned out to be much greater when the words represented by the series of characters were phonetically similar. In spite of the nature of the kanji characters, the subjects used a "phonetic" code. In the discussion of their results, Erickson et al. bring up the investigations carried out among Japanese aphasic patients by Sasanuma & Fujimura (1971 and 1972). A group of aphasics with apraxia of speech made a significantly greater number of errors in kana processings than in kanji processings, especially in writing tasks. The aphasics without apraxia exhibited no such clear difference. It was assumed by Sasanuma & Fujimura that the kana transcriptions can by-pass the phonological processor and have direct access to the lexical items. The conclusion that the kana and the kanji transcriptions were to be processed in different modes is inconsistent with the way the normal Japanese acted in the experiment reported by Erickson et al. If the phonological processor was damaged then there should be a notable difference in perception too, they argue. Such a difference was not reported. They prefer to regard the behaviour of the aphasic patients as an indication of different levels of linguistic awareness.

Everything reported so far seems to ignore visual perception in reading and yet the graphic picture of the words is what is most important to the skilled reader. He does not code words into sounds before grasping their meaning. If that was his strategy, he would not hesitate when confronted with a sentence like: "The bouy and the none tolled hymn they had scene and herd a pear of bear feet in the haul" so long as the words are familiar to him (LaBerge 1972). Smith (pp. 73 ff.) emphasizes that meaning can be extracted from sequences of written words independently of their sounds, and that meaning must be comprehended if sounds are to be appropriately produced.

An active reading process like the analysis-by-synthesis procedure in speech perception is proposed by many investigators. Bower (1970) has shown that a fast reader picks up the structure of the sentence before recognizing the individual constituents of the sentences. This means that the appearance of a word is more important than its sound for reading (pp. 143 ff.). Kolers (1972) comes to a similar conclusion in a very interesting reading experiment where he rotated the lines of the text in various ways. Such texts preserve all the normal features but in a somewhat unfamiliar pattern. The subjects developed considerable skill at reading the rotated lines. It was discovered that in misreading a word the subjects usually substituted for it a word of approximately the same length. This seems to be the same phenomenon as Liberman (1972) refers to as "shortcuts" in the linguistic processing, indicating that familiar material does not receive complete linguistic processing. An active reading strategy must imply exactly this sort of guessing-game.

It is generally agreed that only a small part of a text is focused by the eye, which moves "in saccades" and makes only 3 or 4 fixations per second. Slow and fast readers make the same number of fixations. Malmquist (quoted by Platzack 1974, p. 26) reports that a skilled reader can take in 2.5 words per fixation at most, a slow reader not more than 0.5 words per fixation.

It can thus be assumed that becoming a skilled reader implies a lot more than just processing rapidly the letters put together into words and sentences. The reader's awareness of the phonotactical and grammatical structure of his language is a pre-requisite as is his knowledge of the phonology and the morphology of the words. This knowledge increases the redundancy of the visual information and a larger part of the text can be processed at the same time (Platzack, op.cit.). Spelling errors might reveal something about the development of this knowledge and about the reading process itself. I have therefore collected all the misspellings from about 200 compositions written by about 100 Swedish pupils during their fourth and sixth year in school. Since it is known from e.g. Lindell (op.cit.) that the difference between a good and a poor speller is the number, not the kind of misspellings, the errors collected from a 'normal' class can be regarded as representative.

The spelling errors were grouped conventionally into insertions, omissions, substitutions and transpositions. Only errors not due to ignorance of spelling rules are discussed here. As in the investigations by Lindell, the omissions and the substitutions turned out to be by far the most frequent errors. Lots of substitutions are purely graphic, defined as errors that would be discovered by the author if he was to read the text aloud himself. The "genuine" substitutions would probably not be detected by the writer since they are equivalent to his acoustic or articulatory idea of the word. About 3/4 of the omissions are consonants and almost 2/3 of the omitted consonants belong to clusters. A great deal of the substitutions and the omitted consonants can be explained by reference to the coarticulation and assimilation tendencies of the spoken language. Gårding (1974) has given rules for assimilation and reduction in spoken Swedish. One type of reduction rule concerns the assimilation of consonants. This is a very common phenomenon in the written compositions too. There seems to be an even stronger tendency in the misspellings than in spoken language for two adjacent consonants to share the same contact surface, since several omissions in the written material of consonants in clusters with the same or almost the same place of articulation are never omitted in speech. Gårding also mentions the fact that initial clusters are never reduced, only final ones. This seems to be true for the written material in general. Many omissions are found in final position, very few initially. Almost 100 % of the single consonants are omitted in final position. This goes even for words where such an omission is impossible in the spoken language. As is well known and as was mentioned also by Gårding, the information value is low at the end of a word. It would be an interesting task to compare the causes of the spelling errors with the origin of the reduction rules of the spoken language.

There was a certain tendency among the substitution errors for a target phoneme to be replaced by a phoneme with the same distinctive features except one. This observation has been reported in the literature many times. Wickelgren (1966) concluded that a consonant is coded in the short-term memory not as a unit but as a set of distinctive features. Sasanuma & Fujimura (1972) and Trost & Canter (1974) found that the substitution errors made by aphasic patients were approximations for the target phonemes. It has also been suggested that certain graphic errors found in compositions written by students at the age of 19 or 20 show a similar pattern (Hultman, personal communication).

If the tendency to confuse phonemes with the same or almost the same place of articulation and to omit consonants in clusters under similar conditions can be proved to occur more generally in misspellings, then we might assume that we use an articulatory memory during writing tasks. Maybe this hypothesis can also explain why the aphasics with apraxia had greater difficulty in writing the kanas than the kanjis. It might be that they could not evoke their articulatory memory just as they cannot evoke the memory of other movements. The articulatory memory would only be necessary for the sequential processing of the quasi-phonological kana symbols, but not for the ideograms. If this is true there must be another explanation for the phonological activity shown when the 'normal' Japanese confused the rhyming kanjis. This was a memory task and it is plausible, as is emphazised by LaBerge, that the phonological code is the strategy we use when the task is to store and retrieve directly from the short-term memory. This strategy cannot simply be generalized to comprehension tasks (p. 243).

Summary

The child who has not yet learnt to read and write has an undivided acoustic-articulatory image of a word. To be able to read and write, he must link this acoustic-articulatory unit with a sequence of letters.[1] This can only be achieved if the acoustic-articulatory image is first divided into sound segments corresponding to the letters, a task which involves segmentation of a continuous acoustic signal and overlapping articulatory gestures and relating the segments to neatly defined letters. Little by little the sound segments and the sequence of letters are integrated and perceived as a unit. There is no longer one acoustic and one visual picture of the word but a picture that is simultaneously and indivisibly acoustic-articulatory and visual. When this integration has been accomplished, the reading process can develop in an analysis-by-synthesis manner like the auditory perception. The strategy used by a skilled reader is neither a letter-by-letter nor a word-by-word processing. If the child has difficultes in managing the sequential processing when starting to read there will be no integrated word picture and the child will fail to learn to read, or learn only very slowly.

Learning to read and write logograms does not require awareness of the segments of the acoustic word image since the visual word picture is not a sequence of symbols. No sequential processing is thus necessary for reading logograms.

The completely integrated word picture cannot be dissolved. Alexia and agraphia are fictitious exceptions, since it is the performance and not the competence that is damaged (Weigl & Bierwisch 1970). With an integrated word picture it is no longer possible to perceive the letters as pure forms. Consequently, it is not possible to use only visual processing in silent reading or writing. If this was the case, then a poor reader would not have more problems in separating word pictures than other pictures. Some sort of silent speech is no doubt involved in silent reading and writing as well as in other mental activities. The task is to find out what kind, how and when.

---

1. This is the way reading is frequently taught. Only recently have attempts been made to teach reading by presenting the whole written word as a unit (e.g. Söderbergh, 1972).

References

Bower T. 1970. Reading by eye. In Levin & Williams (eds), Basic Studies on Reading. New York

Conrad R. 1972. Speech and Reading. In Kavanagh & Mattingly (eds), Language by ear and by eye. Cambridge, Mass.

Doehring D. 1968. Patterns of impairment in specific reading disability. Bloomington & London

Edfeldt Å. 1960. Silent speech and silent reading. Stockholm

Erickson D., I. Mattingly and M. Turvey. 1973. Phonetic activity in reading: An experiment with kanji. Haskins Laboratories. Status Report on Speech Research. SR-33

Gårding E. Forthcoming. Sandhiregler för svenska konsonanter. Åttonde sammankomsten för svenskans beskrivning. Lund 1974

Klapp S., W. Anderson and R. Berrian. 1974. Implicit speech in reading, reconsidered. Journal of Experimental Psychology 100, 2:368-374

Kolers P. 1972. Experiments on reading. Scientific American 227:84-91

LaBerge D. 1972. Beyond auditory coding. In Kavanagh & Mattingly (eds), Language by ear and by eye. Cambridge, Mass.

Liberman A. 1972. General discussion of papers by Conrad and LaBerge. In Kavanagh & Mattingly (eds), Language by ear and by eye. Cambridge, Mass.

Liberman I. 1973. Segmentation of the spoken word and the reading acquisition. Haskins Laboratories. Status Report on Speech Research. SR-33

Lindell E. 1964. Den svenska rättskrivningsmetodiken. Lund & Copenhagen

Makita K. 1968. Rarity of reading disability in Japanese children. American Journal of Orthopsychiatry 38, 4:599-614

Mattingly I. 1972. Reading, the linguistic process, and linguistic awareness. In Kavanagh & Mattingly (eds), Language by ear and by eye. Cambridge, Mass.

Platzack Ch. 1974. Språket och läsbarheten. Lund

Pynte J. 1974. Readiness for pronunciation during the reading process. Perception & Psychophysics 16, 1:110–112

Rozin P., S. Poritsky and R. Sotsky. 1973. American children with reading problems can easily learn to read English represented by Chinese characters. In Smith F. (ed.), Psycholinguistics and Reading. New York

Sasanuma S. and O. Fujimura. 1971. Selective impairment of phonetic and non-phonetic transcription of words in Japanese aphasics patients: Kana vs kanji in visual recognition and writing. Cortex 7:1–18

Sasanuma S. and O. Fujimura. 1972. An analysis of writing errors in Japanese aphasic patients: Kanji vs kana words. Cortex 8:265–282

Smith F. 1973. Decoding: The great fallacy. In Smith (ed.), Psycholinguistics and Reading. New York

Söderbergh R. 1972. Reading and stages of language acquisition. Paper presented at the IXth International Congress of Anthropological and Ethnological Sciences, Chicago, USA, 1973

Trost J. and G. Canter. 1974. Apraxia of speech in patients with Broca's aphasia: A study of phoneme production, accuracy and error patterns. Brain & Language 1, 1:63–79

Weigl E. and M. Bierwisch. 1970. Neuropsychology and linguistics: Topics of common research. Foundations of Language 6:1–18

Wickelgren W. 1966. Distinctive features and errors in short-term memory for English consonants. Journal of the Acoustical Society of America 39:388–398

IN FAVOUR OF THE ARCHIPHONEME

Thore Pettersson

Halle's famous argument against an autonomous phonemic level was originally constructed in order to decide between two possible conditions on a phonological description. Since, according to Halle - and I believe most linguists would agree on this statement - phonological segments are theoretical constructs, they must be appropriately related to observable data. The weakest form of condition for such a phonological description reads as follows in Halle's (1959, 21) setting:

(1) A phonological description must provide a method for inferring (deriving) from every phonological representation the utterance symbolized, without recourse to information not contained in the phonological representation.

There are, however, many cases where distinctively different utterances will be pronounced in the same way. The Russian expressions mok by 'were (he) getting wet' and mog by 'could (he)' would be pronounced identically: ['mogbɨ]. Similarly, there is no way to decide directly from the phonetic data that the nasal in Swedish imbecill 'imbecile' and inbilsk 'conceited' represents phonologically distinct segments. Obviously Halle is right when he states that "it must be possible to read phonological representations regardless of whether or not their meaning, grammatical structure, etc., is known to the reader". Given quite natural assimilation rules and phonemic representations of the form /mok bi/, /mog bi/, /imbesil:/ and /inbilsk/ condition (1) will be satisfied. But many, not to say most, traditional phonologists would not accept these representations, since they are impossible to arrive at from the utterance alone. The only kind of data, available to the adherent of pregenerative linguistics, was speech forms in the Bloomfieldian sense, i.e. actual physical utterances. Every conscious appelation to introspection was, accordingly, held to be unscientific. Therefore one would ask for a decisive method to derive the underlying form directly from the utterance itself. Halle formulates the condition necessary to satisfy this demand in the following way:

(2) A phonological description must include instructions for inferring (deriving) the proper phonological representation of any speech event, without recourse to information not contained in the physical signal.

The most obvious way to arrive at a proper phonological representation, given condition (2), is to use one symbol for every sound one is able to detect in the physical signal. But since, actually, the number of possible sounds is unlimited, one must restrict the number of symbols used so as to represent only distinctive sound types and leave "allophonic" variations to be accounted for by phonological rules. The fact that such a restricted application of condition (2) requires that we logically ought to represent English hang with the same symbol for both the initial and the final segment of the word, i.e. give the sequence in question one of the phonological representations /ŋæŋ/ or /hæh/, has mostly been rejected because it is intuitively absurd. However, phonological representations are still theoretical constructs which need no psychologically motivated support. Therefore I cannot understand why we should reject an abstract representation /hæh/ as psychologically disturbing, when we just have decided to prefer the psychologically suspect phonological representation /mog bi/ for mok by to the psychologically motivated representation /mok bi/. Accordingly, accepting the restricted version of condition (2) entails our giving up all psychological or intuitive motivations for constructing phonological representations, and it also forces us to accept the principle 'once a phoneme, always a phoneme'. If we do not, our conception of the phoneme will be quite empty and arbitrary.

Given an arrangement of grammar based on condition (2) – which must be the only reasonable one, if we want to avoid psychological motivations – Halle's argument against the phonemic level is quite devasting. It runs as follows: in Russian all obstruent phonemes occur in voiced-voiceless pairs which often alternate morphophonemically. Thus we have alternations such as /gorət/ 'city' – /gorədə/ (g.sg. of the same word), /mosk/ 'brain' – /mozgə/. The three obstruents /c/, /č/ and /x/, however, do not possess voiced cognates. Followed by a voiced obstruent they are nevertheless voiced. This means that we have to write the same obstruent assimilation rule twice in the grammar, first for obstruents which are paired with regard to voice and then for the unpaired. The only way to save condition (2) without increasing the complexity of the grammar is to show that the argument given by Halle is somehow invalid.

There have been many attempts to invalidate Halle's argument. I fail to see, however, that anyone of those counterarguments I know of are really successful. Johns (1969, 375) tries to get rid of the argument in this way:

"The fact is that the voicing assimilation rule is always phonetic, since voicing is never distinctive in obstruents in the position before another obstruent. In other words, if we have a way of representing the concept of neutralization, which is surely not incompatible with the notion of a phonemic level, Halle's problem becomes totally spurious."

As a matter of fact, Halle demonstrates how to represent neutralization in such a way, viz. the concept of archiphoneme. But given condition (2) there is obviously no way of making the concept of archiphoneme compatible with the notion of an autonomous phonemic level. What one cannot detect in the physical signal has, according to this view, no relevance for a linguistic description, and, certainly, one does not pick up archiphonemes without reference to data not contained in the physical signal. The embarrassing fact is that we all know, whatever our attitude to the question of the existence of independent phonemics as opposed to morphophonemics may be, that the voicing assimilation in Russian actually _is_ a phonetic phenomenon. But as adherents of condition (2) we are not supposed to know this. Thus when Johns wants a phonological description powerful enough to account for neutralization he actually argues in favour of condition (1), i.e. the phonemic level which Johns claims to be compatible with the concept of neutralization is, in fact, a different phonemic level from the one Halle and his generativist followers have rejected. Johns' phonemic level is identical to Halle's morphophonemic level.

Derwing (1973, 186) following Householder (1967) rejects Halle's argument by the same means. But unlike Johns, he explicitly states: "The proper _morphophonemic_ rule involved in such a neutralization is therefore as follows: 'all obstruents are _unspecified_ for voicing before other obstruents' and the proper _phonetic_ rule is precisely the generalization which Chomsky and Halle suggest: 'all obstruents are voiced before voiced obstruents' (or, more generally, 'all obstruents assimilate in voicing to the final obstruent of an obstruent cluster')." Thus Derwing agrees with Halle's original argument against the phonemic level. His proposal could have been a counterargument against Halle, if there had been empirical evidence that absolutely _all_ obstruents are unspecified for voicing before other obstruents. But Derwings morphophonemic rule is _ad hoc_ constructed in order to save condition (2). That is to say that Derwing commits himself to the same kind of axiomization that he accuses Chomsky and Halle of. For consider what his proposal means: Russian _mog_ in the sequence _mog by_ should be morphophonemically represented as /MOK/, a representation that, certainly, no Russian native speaker could have incorporated in his grammar. There is no doubt about the fact that Halle's own proposal to separate /c/, /č/ and /x/ from other obstruents and assign their voicelessness by way of a special rule is the sounder solution. Observe that this does not force us to specify other obstruents for voice in position before an obstruent. The difference between Derwing and Halle is this: while Derwing insists that all obstruent morphophonemes in the given contexts are unspecified as to voice, Halle says that the three morphophonemes /c/, /č/ and /x/ are always unspecified whereas the other obstruents need not be specified in neutralizing contexts, i.e. their specification for voice is redundant (cf. Halle, 1959, 61, 63f.).

Linell (1974, 105) recapitulates the arguments of Johns and Derwing, but he goes on to postulate that there is "no clear empirical support" for the principle 'once a phoneme, always a phoneme'. Since it is exactly this principle that is at stake here, there would be no empirical evidence for Halle's argument either. But given condition (2) I cannot see that Linell's assertion holds. Obviously it can be empirically proven that the segment [g] in ['mogbɨ] is, in a certain sense, the same sound as [g] in [got], which is the phonetic output of Russian god 'year' and stands in phonemic contrast to kot 'male cat'.

Moreover, Linell cannot be convinced of the relevance of Halle's argument until he finds a case, where in spite of a neutralization rule morphophonemes could appear, which do not participate in the rule. Such a situation would, according to Linell, turn up in a case where X and Z are bundles of features and have the same feature specification except for one feature and the distinction X – Z is phonemic and, furthermore, we have two kinds of morphemes: (i) morphemes which are X in a context C__D as opposed to A__B and are subject to the neutralization rule in the context A__B, and (ii) morphemes that do not participate in the alternation and are X in the context A__B. As a matter of fact, Halle's original argument constitutes such a case, if we define the archisegment as a not fully specified segment in the underlying form. But since Linell seems to be sceptical about the psychological relevance of this concept, I shall present a case where such a rule really seems to exist.

In the indigenous system of Czech there is no phoneme /g/, old [g] having developed into [ɦ]. Czech has the same obstruent assimilation and final devoicing rules as Russian. Consequently words like gde 'where', kdo 'who' and k Brnu 'to Brno' ought to be phonemically represented as /kde/, /kdo/ and /kbrnu/. (Underived prepositions in Czech form one phonological word together with the following word.) Now, words beginning with a vowel are optionally preceded by a glottal stop [ʔ]. Czech Amerika can thus be pronounced [ʔamɛrika] or [amɛrika]. If the voiced preposition v 'in' stands in front of a word with initial vowel, the preposition will be devoiced to [f] as a consequence of the obstruent assimilation rule; but, N.B.!, this will happen irrespective of whether or not there is a glottal stop between the preposition and the vowel; v Americe 'in America' is therefore pronounced either [fʔamɛricɛ] or [famɛricɛ]. That the unvoiced preposition k 'to, till' remains unvoiced in the same position is not surprising: k otci 'to father' is pronounced [kʔɔci] or [kɔci]. These two variants are both commonplace in Standard Czech (Prague) pronounciation. In certain other Czech dialects the preposition v is not devoiced before a vowel; the pronounciation [vamɛricɛ] is thus also possible. But in such dialects where v does not devoice, the preposition k tends to become voiced; for k otci one would here say [gɔci], as if otci were immediately preceded by a voiced obstruent. Observe that the segment [g] never occurs before a vowel, liquid or nasal in any dialect other than in this sole case. As a consequence we find in Czech just that situation Linell asks for: though there is no phoneme /g/ and [g] regularly replaces [k] before voiced obstruents, we may nevertheless establish minimal pairs such as [kɔlɛji] - [gɔlɛji], corresponding to koleji (dat.sg. of kolej 'boarding-house') and k oleji 'to the oil' (formed from olej 'oil') respectively.

Now, if I were a classical phonologist and thus a true adherent of condition (2), I would say either that there must be some sort of boundary phoneme that could account for both the devoicing of v and the voicing of k, or I could quite straightforwardly give the Czech /g/ the status of an independent phoneme, since one can in any case state the existence of minimal pairs such as [kɔlɛji] - [gɔlɛji] and, furthermore, /g/ in Czech clearly has phonemic status amongst foreign words, giving rise to minimal pairs such as gáže 'salary' - káže 'preaches' and grog (phonetically [grɔk]) 'grog' - krok 'step'. Furthermore, as an argument against an autonomous phonemic level my example suffers from the fact that I must suppose an underlying segment - a voiced glottal stop - which it is impossible to realize at surface level. Such animals simply do not exist.

Now that I have given the fortress up, I may as well continue to play the role of the devil's advocate. The opponents of Halle's proposal referred to so far have tried to escape its consequences by invalidating condition (2), an approach which is doomed to failure. Whatever the value of Chomsky's (1964) analysis may be, it has at least clearly shown that the only possible position that can be taken by the strictly data-oriented phonologist is to faithfully stick to the principle of invariant biuniqueness. And there is no possibility to do this and simultaneously adhere to the principle of complementary distribution. In English /k/ is in complementary distribution with lax /a/. This means that both socked and Scot should be phonemically represented as /skkt/. The only solution is to give up the principle of complementary distribution and add as a prerequisite of a proper data-oriented description of the phonemic system of a given language, that it gives an exhaustive account of which natural phonological classes the phonemes of the language in question can be divided into. English has at least the two main classes of vowels and consonants. Since /a/ and /k/ can be shown to belong to separate classes, defineable both articulatorily and acoustically, the proper phonemic representations of socked and Scot will be /sakt/ and /skat/ respectively. For the same reason hang shall be represented as /hæŋ/, since /h/ belongs to the class of glides and /ŋ/ belongs to the class of nasals.

Halle's argument hinges on the fact that unless one does not give up the notion of an autonomous phonemic level, one is forced to divide the natural class of obstruents into two independent classes. However, Halle has not shown that the phonemes /c/, /č/ and /x/ really have the phonemic status within the Russian system that he ascribes to them. I shall now show good grounds for questioning his assumption.

It is true that the segment [ɣ] in certain contexts has a distribution which allows us to determine it as being an allophone of /x/. But this is not the whole truth. Firstly, [ɣ] is in these contexts a free variant of /x/. Russian v gorách by 'were it in the mountains' can be pronounced both [vgʌ'raɣbɨ] and [vgʌ'raxbɨ], whereas mok by obligatorily is pronounced ['mogbɨ]. Furthermore, it is not clear that /x/ shall be classified as an obstruent at all. The voiced counterpart is in any event to be characterized as a glide. Henning Mørk has brought to my attention the fact that modern Greek shows up with a similar distribution: here the non-palatal and palatal obstruents /x/ and /x,/ have as voiced cognates the glides /ɣ/ and /j/. Even on articulatory grounds it is reasonable to place Greek /x/

and /x,/ in the class of glides. They are, actually, not clearcut obstruents but rather approximants. This also holds for Russian /x/. It is a constant complaint of authors of handbooks on Russian phonetics that the sound corresponding to the cyrillic letter X is symbolized [x] and not [h].

There is another peculiarity about [ɣ] which Halle leaves totally out of account. The sound in question is generally a very rare one, but when it does occur it is far more often a variant of /g/ than of /x/. In elderly style a word such as kogda 'when' is pronounced [kʌˈɣda]. The exclamation gospodi! 'Lord!' is almost invariably pronounced [ˈɣospəd,ɪ]. Oblique forms of bog 'god' are pronounced either [ˈboɣə], [ˈboɣu] or [ˈbogə], [ˈbogu]. To my knowledge it has not been observed that these pronunciations reflect a semantic difference. Pronounced with the fricative the forms refer to the Christian god, whereas the stop implies a reference to a pagan god or the Christian god equalled with pagan gods. Observe that the plural forms bógi, bogóv etc. are always pronounced with the stop. Finally, the word buchgalter 'book-keeper' is, according to the literary norm as given by Avanesov & Ožegov (1960), always pronounced [buˈɣalt,ɪ̆r].

Thus, faithful to condition (2) we can state both that it is doubtful whether /x/ might be classified as an obstruent at all and whether /x/ has no voiced cognate /ɣ/ and whether [ɣ] should properly be regarded as an allophone of /x/ and not /g/.

What regards the affricates /c/ and /č/, these sounds are clear obstruents. But Halle's argument turns out to be completely vacuous, once it is considered that these sounds could equally well be analyzed as the phoneme sequences /ts/ and /tš/. When the dental stop and a sibilant come close but a morpheme boundary separates the two sounds as in détstvo 'childhood' (formed from the stem det and the suffix stvo) the combination is pronounced in exactly the same way as the affricate [c]. The reason for treating the affricates as single phonemes is that within morphology they alternate regularly with single stops: pekú 'I bake' alternates with pečëš' 'you bake'. A similar kind of morphological alternation occurs in Swedish, where the verb skära 'cut', phonemically /šæ:ra/, has the preterite skar. Nobody would suppose the sequence /sk/ to be a single phoneme, simply because it alternates with /š/. The standard argument against the biphonemic analysis of the Russian affricates, however, rests on their distribution. In certain cases a so-called unstable vowel is inserted before the last consonant of a word-final consonant cluster. When such a form ends in an affricate the unstable vowel is inserted before the affricate, not between the two segments constituting the sound; cf. otéc 'father' (*ottes) and its genitive otcá. The reason for this peculiarity is again the fact that the affricates go back to old single segments (Old Slavonic otĭkŭ). Something of the same principle also works for the combination stop – sibilant but the other way round. From the verb mstit' 'revenge' you form the deverbal noun mest' with the same type of unstable vowel before the sibilant. The sequence /st/ of this verb – and there is a host of equivalent cases to be found in modern Russian – corresponds to the simple phoneme /š,/: mšču, /mš,u/, 'I revenge' contrasts with mstiš' 'you revenge'. Thus, if you take /ts/ and /tš/ to be single phonemes, you must do the same thing with at least /st/ and /sk/. And that would be absurd. Consequently, the only sound phonemic analysis of the Russian affricates is the biphonemic one.

If my analysis of the Russian affricates and any of my assumptions on the voiced velar [ ] are accepted, then the main argument against the taxonomic phonemic level is definitively invalidated. It is true that Halle's argument has been repeatedly reformulated on material from other languages, for example by Bach (1964, 128) on German data, and Bach's argument has subsequently been reformulated on Norwegian data by Hovdhaugen (1971). But in both cases it is obstruents and the /r/-phoneme that are compared. Since /r/ in both German and Norwegian must be classified as a liquid, and liquids per se belong to a separate class, Bach's and Hovdhaugen's argument is not valid.

Another argument against an autonomous phonemic level was given by Kiparsky (1965, 4f.). It concerns the Old High German umlaut of a, which resulted in a fronted and non-raised vowel if the umlauted vowel was long but a fronted and raised vowel if the umlauted vowel was short. According to Kiparsky it is possible to formulate this phenomenon in one rule:

$$\begin{bmatrix} -\text{consonantal} \\ \langle -\text{long} \rangle \end{bmatrix} \rightarrow \begin{bmatrix} -\text{grave} \\ \langle -\text{compact} \rangle \end{bmatrix} / \_\_\_\_ C_n i$$

Thus there is no longer any reason to believe that umlaut first created fronted allophones of /a/ and raising afterwards incorporated the fronted allophones in the phoneme /e/. Vennemann (1972), however, has convincingly demonstrated that it is probably incorrect to handle the raising as part of the umlaut rule. Rather there was one phonological rule proper, i.e. the rule fronting back vowels before [i]. In addition to this rule there was a general constraint – as in Modern Standard German – with the effect of prohibiting short low front vowels at all levels of the grammar. Thus the old opinion that /a/ and /e/ (the result of raised [æ]) were distinct phonemes because they contrasted in environments other than before syllables containing [i] can be maintained.

The arguments against an autonomous phonemic level can doubtlessly be multiplied, but I strongly suspect that it will always be possible for a phonologist who strictly holds to condition (2) to refute them. It must be so, because he always has sturdy and substantial data to refer to. To take such a position is therefore a safe way of going about linguistics. Nothing can dislodge you and you are always rescued from the plague of thinking. Now I deliberately confess that I am greatly impressed by some recent highly critical studies of the metatheoretical basis for generative phonology. I have in mind books such as those by Botha (1971), Derwing (1973), Linell (1974) and Karlsson (1974). I also admit that I have very little to object to in the criticism that has been directed against generative grammar in general as in the books by Itkonen (1974) and Hiorth (1974), though the conclusions I would like to draw myself might diverge from those made by the authors mentioned. Finally, I declare that never have I been fully convinced of the viability of the phonology practised by Chomsky and by Halle from the early sixties, a thing which developed into the remarkable kind of morphological phonology exhibited in Lightner (1966, 1967) and above all in Chomsky & Halle (1968). But even the very fact that orthodox generative phonology has wallowed in the mire of exorbitant abstractions far from rhyme and reason and that a strict and calm data-oriented linguistics can obviously tell us much more about reality than the excessive speculations of generative phonologists ever could, cannot convince me that the neotaxonomic orientation of Derwing and Linell is the right way to set about it all.

As a matter of fact, there is very little of substance that divides us from one another, but the points in which I disagree are important in principle. In taking the standpoint of Derwing and Linell you do actually adhere faithfully to Halle's condition (2). But in spite of all that can be said against condition (1) it will remain my position. It must be possible to supply the phonological description you have constructed on the basis of physical data with relevant information provided through the use of introspective evidence.

Derwing (1973, 188ff.) discusses Spanish e-epenthesis and Russian palatalization. In Spanish there is a general phonotactic constraint prohibiting a word from beginning with s + stop. Words, inherited or borrowed from Latin, originally beginning with such a cluster, are in Spanish rendered: España, estudiante etc. But this also happens to new borrowings: English scar takes the form eskar, stop will be rendered estap. We could formulate this phenomenon as a phonological rule:

$\emptyset \longrightarrow e\ /\ \_\_\#\ sC$
where C is a voiceless stop

Now, Derwing asks, does the formulation of this phonological rule make it plausible that the underlying form of España is /span,a/? Of course not. The Spanish epenthesis is still a general phonotactic constraint, something similar to Vennemann's [æ]-raising rule of Old High German. Derwing's position prevents him from realizing the difference between phonotactic constraints and phonological rules proper. Most assimilation rules are of the latter type. Therefore it is not possible to place Spanish España on a par with Russian sledovat' 'follow', where the stem sled throughout the conjugation has the form [s,l,ɛd]. Why then do we not join Derwing in proposing the underlying form /s,l,edovat,/? The Russian palatalization is a phonological rule with the effect of making all consonants of a consonantal cluster – with some well defineable exceptions – conform in palatality with the last consonant of the cluster. Thus from the noun mesto, [m,ɛstə], 'place' we get the locative v meste [v,m,es,t,ĭ]. Accordingly, in other environments palatalization is of that typical phonetic character which we would have liked to have called allophonic, had palatalized consonants not been established in phonemic contrast with non-palatalized. But now, observe a peculiarity about the Russian palatalizing vowels. Whenever a palatalizing vowel in Russian is not preceded by a consonant, it is realized as [jɛ], [ja], [ju], [jo]. The only exception is the high front vowel [i]. As an adherent of condition (1) I can without exaggeration suppose that there really are no palatalized phonemes in Russian but a palatalizing segment, which I shall define as a high front syllabic or non-syllabic vocoid. In agreement with old taxonomic principles of economy my solution will reduce the number of consonant phonemes in Russian by half. Chomsky and Halle have been accused of making a phonology of English with the spelling as a model, correctly I think. If Derwing cannot accept my postulated underlying palatalizing segment I am inclined to accuse Derwing of writing a phonology for Russian with the spelling in mind. Accordingly, I propose that the underlying form of sledovat' is /sljedovatj/.

Both Derwing and Linell are remarkably anxious to admit that neutralization could be of allophonic character. Thus I fail to see why German Bund could not have an underlying form /bund/, why its phonological structuring necessarily must be /bunt/. As Karlsson (1974, 53, note 74) points out final devoicing is a rule which is natural, productive, and exceptionless. Karlsson also points to the fact that this is a rule that carries over when learning a foreign language. In Derwing's theory the last point, of course, is no argument, since his phonological component is build up of articulatory habits and does not contain rules. And it is, presumably no argument for Linell either, since it is not possible for Bund to occur on the articulatory plan with a voiced final stop. But is it really quite certain that this is also true of the perceptual structuring of the word? I shall not deny that Linell may be right that the identification of [t] in Bund with [d] in Bunde could be more a matter of morphology than of actual perceptual structuring. Nevertheless it is strange that German, Russian, Polish, Czech and a host of other languages, which have this rule in common, spell the devoiced obstruent with the symbol for the voiced sound. In Serbo-Croatian, which has an almost phonemic spelling, one writes the word for 'sparrow' vrabac and its genitive vrapca. This spelling, which obviously is in accordance with Linell's perceptual structurings, causes the Serbo-Croatian school children great difficulties. They spell intuitively vrabca and must be trained to achieve the phonemic and phonetic spelling vrapca. If they really hear nothing but an unvoiced stop in this word, and, certainly, they cannot produce anything but [vrapca], why should it then be so disturbing to write down the internalized phonological structuring directly without checking the morphological paradigm? My assumption is - and this assumption is supported by evidence from native speakers of Slavonic languages - that the children actually hear [b], though they can reasonably only perceive [p], exactly in the same way as I hear [n] in inbilsk, though I should be able to perceive nothing but [m].

In his famous article on Russian conjugation (1948) Jakobson develops a procedure of how to arrive at base forms (to be differentiated from phonemic forms) which are very much the same as the archiphonemic representations I shall propose her. In Russian all unstressed vowels except /u/ are neutralized into [ʌ], [ɪ̆] or [ə]. These three reduced vowels are in complementary distribution and could, accordingly, in classic taxonomic phonology be represented by the same vowel symbol, say /ə/. The noun nogá 'leg, foot' is pronounced [nʌ'ga]. But since the first syllable receives stress in the accusative: nógu and in the plural nominative: nógi, Jakobson would establish the base form of the stem as being /nog-/. In the loan-word kovbój 'cowboy', however, the first syllable is never stressed, and here Jakobson would have chosen the representation /kavbój/, whereas Troubetzkoy would have preferred to classify the unstressed vowel as the archiphoneme /A/, which gives the representation /kAvboj/. I shall in cases of uncertainty stick to this type of mixed morphoarchiphonemic-phonemic representation. In cases where I feel I have good reason to believe that a neutralization really only reflects the operation of a phonological rule, I shall take the morphophonemic representation to be identical with the phonemic. Thus I shall suppose that the underlying form of Serbo-Croatian vrapca is /vrabca/ and of German Bund is /bund/. This is an assumption that goes against Troubetzkoy, who would have preferred the representation /vraPca/ and /bunT/ respectively. Observe that I say feel and believe, which means that I have no other data for my decision than what I believe I have understood correctly from native Serbo-Croatian and German speakers' statements about their intuitions.

What then with a case such as Russian nogá? Would it not be a sound decision to take its underlying (base) form to be /nogá/ as Jakobson would? There are actually certain facts that are against such a solution. Morphophonemic change per se cannot be allowed to decide the question for me. As a matter of fact, Russian school children have great difficulties in learning the spelling of the Russian unstressed vowels. A Russian child who is not fully trained in spelling would therefore write noga as well as naga. In Byelo-Russian orthography the full consequence of this difficulty has been taken into account and the a-writing has been generalized for all non-stressed non-front non-high vowels. It looks, furthermore, as if the a-spelling is more natural, whereas non-correct o-spelling seems rather to be a case of hypercorrection. It could thus be so that what we have here is a kind of paradigmatic vowel alternation, very much the same as the German morphologisized umlaut. Since I, nevertheless, cannot be sure that this is the case, I prefer, until convincing evidence in either direction has come to light, to assume that the vowel in question is a not fully specified segment, i.e. the archisegment /A/, not specified with respect to the features back and low. Notice, however, that I do not exclude the possibility that either one of the phonological representations /nogá/ or /nagá/ could be equally correct. Observe that the representation /nagá/ implies that I take the a-o-alternation to be morphological as opposed to morphophonemic. The really embarrassing failure of orthodox generative phonology was that this distinction was never made.

Thus the Russian a-kan'e phenomenon can in a certain sense be equated with the problem of the proper specification of a stop following an [s] in English and Swedish. Observe that none of the variants of unstressed Russian o or a can be phonetically identified with their stressed counterparts. The unstressed variants all share certain features with both [o] and [a]. But Russian native speakers tend to identify the unstressed allophones with [a] rather than with [o]. In the same way, an English native speaker is said to identify the second segment of spin with [p] rather than [b]. Stampe (1972, 34ff.) argues in favour of the [p] solution because "in general stops after /s/ are phonologically voiceless". As evidence for this statement we have orthographical tradition, slips of the tongue like [hwipsr̥] for whisper, babytalk [phæk] for spank, the fact that intensivizing [s] occurs in pairs such as mash/smash and trample/strample but never occurs before a voiced stop. Stampe also refers to pre-school children's spontaneous spelling, such as SCICHTAP for Scotch tape and SKEEIG for skiing. (I have myself on the other hand registered spellings such as SGA for ska 'shall, will' and SBARA for spara 'spare, collect'.) All these facts and Stampe's own theory of the so-called "natural phonological processess" - for a printed summary of his theory see Stampe (1969) - tell us that every language possesses an inventory of phonemes "which is at least indirectly accessible to it's speakers' consciousness". Stampe means that this assumption is critical to an understanding of how alphabets are used and devised. I can follow Stampe this far, and as yet his theory is not incompatible with my own ideas of the impact of the concept of archiphoneme. But Stampe, furthermore, takes the fact that archisegmental theories of underlying representations are unable to identify the inventory of phonemes as a criterion of the fallibility of such theories. As far as I can tell, this cannot be a valid argument at all. There is not the slightest reason why alphabets should be able to represent archiphonemes, i.e. segments which in underlying form and/or at surface level are not fully specified. What is really at stake here is the simple fact that the voicelessness of the second segment of spin is of higher value than the features the segment in question shares with [b], i.e. its being unaspirated and lenis. For this reason all those peculiarities occur which could indicate an identification of [ph] and [p]. But this is nothing that per se could invalidate the concept of archiphoneme. For it is so that our identification of sounds is one thing and the factual specification of them quite another. Therefore I can argue that the last segment of Bund is /d/, whereas the second segment of spin is the archisegment /P/, i.e. a segment unspecified with respect to the feature of voice.

It would suffice to take the argument the other way round. In that variety of Swedish that I speak I make a systematic difference, phonemic and phonetic, between the vowels in fem 'five' and hem 'home'. I pronounce these words [fɛm] and [hem] respectively. However, until the age of about 30, when I first started studying linguistics, I had not the faintest idea that I made such a difference. As a Sturm und Drang poet of twenty I would consider these words to be perfect rhymes. Nevertheless it is selfevident that these words in my grammar are represented as /fɛm/ and /hem/. Therefore I mean that it is a very unwise thing to determine phonological representations from phonetic or introspective data only. A reliable phonological representation cannot be achieved without recourse to both physical and psychological information.

References

Avanesov, R.I. & Ožegov, S.I. (eds), Russkoe literaturnoe proiznošenie i udarenie, Gos. izd. inostrannych i nacional'nych slovarej, Moskva 1960

Bach, E., An Introduction to Transformational Grammars, Holt, Rinehart & Winston, New York etc. 1964

Botha, R.P., Methodological Aspects of Transformational Generative Phonology, Mouton, The Hague & Paris 1971

Chomsky, N., "Current Issues in Linguistic Theory", Fodor, J.A. & Katz, J.J. (eds), The Structure of Language. Readings in the Philosophy of Language, Prentice-Hall, Englewood Cliffs, New Jersey, 1964

Chomsky, N. & Halle, M., The Sound Pattern of English, Harper & Row, New York etc. 1968

Derwing, B.L., Transformational Grammar as a Theory of Language Acquisition, Cambridge University Press, 1973

Halle, M., The Sound Pattern of Russian, Mouton, 's-Gravenhage 1959

Hiorth, F., Noam Chomsky. Linguistics and Philosophy, Universitetsforlaget, Oslo etc. 1974

Householder, F.W., "Distinctive Features and Phonetic Features", To Honor Roman Jakobson, Volume II, Mouton, The Hague & Paris 1967, pp. 941-944

Hovdhaugen, E., Transformasjonell generativ grammatikk, 2. ed., Universitetsforlaget, Oslo 1971

Itkonen, E., Linguistics and Metascience (= Studia Philosophica Turkuensis II), Kokemäki 1974

Jakobson, R., "Russian Conjugation", Word 4 (1948), pp. 155-167. Reprinted in Jakobson, R., Selected Writings, II, Mouton, The Hague & Paris 1971, pp. 119-129

Johns, D.A., "Phonemics and Generative Phonology", Papers from the Fifth Regional Meeting of the Chicago Linguistic Society, Department of Linguistics, University of Chicago, Chicago, Illinois, 1969, pp. 374-381

Karlsson, F., Phonology, Morphology, and Morphophonemics (= Gothenburg Papers in Theoretical Linguistics 23), Göteborg 1974

Kiparsky, P., Phonological Change, M.I.T. Dissertation 1965. Reproduced by the Indiana University Linguistic Club 1971

Lightner, T.M., "Ob alternacii e~o v sovremennom russkom literaturnom jazyke", Voprosy jazykoznanija XV:5, 1966, pp. 64-80

Lightner, T.M., "On the Phonology of Russian Conjugation", Linguistics 35 (1967), pp. 35-55

Linell, P., Problems of Psychological Reality in Generative Phonology, RUUL (= Reports from Uppsala University, Department of Linguistics), no. 4, 1974

Stampe, D., "The Acquisition of Phonetic Representation", Papers from the Fifth Regional Meeting of the Chicago Linguistic Society, Department of Linguistics, University of Chicago, Chicago, Illinois, 1969, pp. 443-454

Stampe, D., How I Spent my Summer Vacation, Unpublished University of Chicago Dissertation 1972

Vennemann, Th., "Phonetic Detail in Assimilation: Problems in Germanic Phonology", Language 48:4 (1972), pp. 863-892

FOREIGN ACCENT AND BILINGUALISM*

Eva Wigforss

"Bilingualism...may be all degrees of accomplishment, but it is understood here to begin at the point where the speaker of one language can produce complete, meaningful utterances in the other language."
(E. Haugen. 1953: The Norwegian Language in America I, p. 6.)

INTRODUCTION

During the past decade there has been a steep rise in the number of immigrants coming to Sweden. The total number of immigrants in Sweden was estimated in 1974 to be 600.000. Of these a large porportion are children aged between 0 and 10. All children in Sweden are confronted with the compulsory school system at the age of seven and most of them will also have to attend compulsory nursery school in the future.

Since the school system is monolingual it causes a great language problem for those children who are not native Swedish speakers. In an attempt to ease the difficulties of those children who have another mother tongue than Swedish, the new supplement to the National Curriculum (Lg 69) provides that every immigrant child has the right to at least two hours a week of instruction in his mother tongue. The National Board of Education supports a project at the Department of Education (University of Lund) that aims find better instructional models for bilingual teaching and development.

This project: "Models for the bilingual instruction of immigrant children" has as a principal goal to ensure that the instruction of immigrant children in Sweden should aim at functional bilingualism for the pupils.

The main aims of the project can be summarized briefly under the following headings·

1. To investigate bilingual learning with reference to the functions and mutual relationship of the languages.

* This research was supported by the National Board of Education in Sweden, through the project "Models for bilingual instruction of immigrant children" at the Department of Education, University of Lund.

2. To design organizational models and to compile methodological guidelines for the bilingual instruction of immigrant children.
3. To arrange practical trials of these models in collaboration and consultation with the Malmö local authority.
4. To carry out a step-by-step evaluation of the practically tested models.

(School Research 1973:8)

Four groups of Finnish immigrant children in Malmö (the biggest city in southern Sweden) are being studied. The instructional models in the form so far evolved can be described as follows:

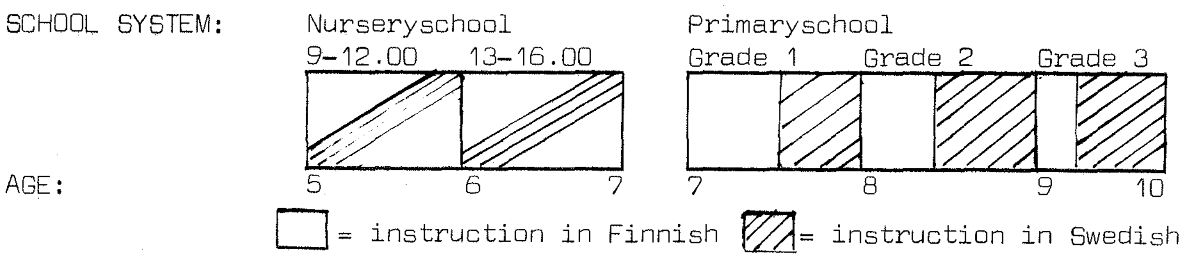


Figure 1. From 5 to 7 the child attends a monolingual Finnish nursery school for three hours a day. During this period a Swedish language teacher is present one hour a day to train the children individually or in small groups in the Swedish language. In grade 1 they will first have to learn to read and write in Finnish while Swedish reading and writing are not introduced until the end of the first grade. In the second and third years Finnish is successively diminished.

Since the project has a longitudinal scientific approach, bilingual development cannot be evaluated in one year and four groups of children at different ages are therefore being followed through the models. These groups are as follows: group A (born 1966), group B (born 1967), group C (born 1968) and group D (born 1969). Thus we follow the children from the ages of 5 to 10. The longitudinal research program is outlined below:

School-system:	age	Group of children: A	B	C	D
Nursery school	(5)		1972/73	1973/74	1974/75
Nursery school	(6)	1972/73	1973/74	1974/75	1975/76
Grade 1	(7)	1973/74	1974/75	1975/76	1976/77
Grade 2	(8)	1974/75	1975/76	1976/77	1977/78
Grade 3	(9)	1975/76	1976/77	1977/78	1978/79

Figure 2. Different terms are tabled above. It should be noticed that group A's first nursery school year started at 6 and that the group of children this first year was scattered between three Swedish nursery schools, but had regular training in Finnish every day.

The evaluation of these models is interdisciplinary, since so many factors are involved in a pupils bilingual and bicultural development. The evaluation is also characterized by a longitudinal approach (fig. 2) and is based on case studies, since the number of children in each group is very small (about 15 children).

BILINGUALISM

Linguists have traditionally viewed bilingualism as "languages in contact", that is, as the interaction between two language systems that otherwise exist in a "pure" state and that have been brought into unnatural contact with each other. Thus the main research focus has been laid on interference between the language systems involved and not enough emphasis laid on the investigation of the actual competence in the two languages. It is also unfortunate that little empirical research has been made into the linguistic development of Swedish and Finnish children (Hadding K. 1974).

Attempts to describe the linguistic competence of bilinguals have so far been made through measurements of their vocabulary, speed of translation between the two languages and pronunciation; but very little has been done in the area of syntactical and communicative competence in bilinguals. There will be a need in the future for empirical psycholinguistic and sociolinguistic research to find out how a bilingual uses his ability in a bilingual speech community. Results from interference studies are often based on error analysis, either in first language (L1, mother tongue) or second language (L2) performance. In the following, very preliminary analysis, I have used this methodology for a contrastive analysis of the phonetic structures of Finnish and Swedish.

I believe the main problem here is not so much to describe the character of interference as to interprete the errors. Some tentative attempts at such an interpretation have been made at the end of the article.

PRONUNCIATION AND FOREIGN ACCENT

Pronunciation should be viewed as a complicated interaction between the articulatory and perceptive system in man. Is it possible for someone to produce a speech sound that he cannot perceive as a speech unit, or vice versa? What role does imitation play in the act of learning the correct pronunciation of a language? Are there critical ages for articulation (and perception) after which it is more difficult to change a pronunciation pattern? There are many such questions that can be raised around the extent to which pronunciation can be affected by external or internal factors. So far we can only pose the questions!

Usually we consider a child of 4 - 5 years old as fully competent in the pronunciation of his mother tongue. There could of course be exceptions for such phonemes as Swedish /s/, /l/ and /r/ that children may not learn to master completely until the age of 6. But generally the prosodic patterns are well established before this. However, recent findings by Irwing (1974) have shown that the accurate articulation of speech sounds improves progressively as the child increases in age up to 7 and it would appear that the maturational process has possibly culminated before the age of 8. The auditory perception of speech sounds also progresses with age, with the most significant improvement between 5 and 7, for the identification of speech sounds in a word spoken by himself or someone else. This finding indicates that the perceptive and productive system in a child is not fully developed untill the age of 8. After this age there seems to be no further progress. This has great relevance for bilingual development.

Foreing accent should here merely be treated as a mispronunciation manifested in an acoustic deviation from the idiomatic norm of a language or a dialect. Though in some cases it is easy to state that a person has a foreign accent, it can be very difficult to state the exact nature of this accent from an auditive analysis alone. But this paper only aims to give a very general survey and the material will be subjected to acoustic analysis at a later date. It should be emphasized that prosodic factors (as for example intonation) are carriers of both syntactic and semantic information in the linguistic message and indications of extralinguistic factors such as attitudes, speech mode etc. One of the conclusions from "Acoustico-phonetic studies in the intonation of Southern Swedish" (Hadding-Koch 1961) was that "intonation is an important instrument for expressing the syntactical relation between utterances and parts of utterances." It is also important to bear in mind that a foreign accent cause misunderstanding at <u>all</u> linguistic levels as well as at psycholinguistic and sociolinguistic levels.

A CONTRASTIVE ANALYSIS OF THE PHONETIC STRUCTURE OF FINNISH AND THE MALMÖ DIALECT

There are many differences between the south Swedish dialect and Finnish. We shall here examine some vowel, consonant, prosodic and phonotactical differences. No further analysis of different Finnish dialects will be offered here since most of the children in the study have a dialect that does not differ in any essential way from standard Finnish as it is spoken in the Helsinki area. All phonetic notation is from The Principles of the International Phonetic Association (IPA 1949), but some symbols are supplemented by the use of diacritics.

It should also be emphasized that generally the Swedish sound system is more elaborate than the Finnish; that is, it contains more speech units.

<u>Vowel phonemes</u>:

The main differences between the vowel systems are:

(i) that the Malmö dialect has two more vowels than standard Finnish; namely /ɯː/ as in /hɯːs/ (Eng. "house") and /ŭ/ as in /hŭnd/ (Eng. "dog"). The first vowel is long and the second short.

(ii) that when vowel length in the Malmö dialect is manifested qualitatively it is quantitative in Finnish. Very often the lengthening of a vowel in Malmö is accompanied by diphthongization (Bruce 1970).

(iii) that there are differences in duration between long and short vowels in Finnish and in Swedish (Reuter 1973).

Vowel systems of Finnish and Swedish:

	FRONT				CENTRAL	BACK	
Finnish:	unrounded		rounded		rounded	rounded	
	long	short	long	short	short	long	short
Close	/i:/	/i/	/y:/	/y/		/u:/	/u/
Half-close	/e:/	/e/	/ø:/	/ø/		/o:/	/o/
Half-open	/æ:/	/æ/					
Open						/ɑ:/	/ɑ/
Swedish:							
Close	/i:/	/i/	/y:/	/y/		/u:/	/ʊ/
Half-close	/e:/	/e/	/ø:/ /ʉ:/		/ŭ/	/o:/	
Half-open	/ɛ:/	/ɛ/		/œ/			/ɔ/
Open		/a/				/ɑ:/	

Figure 3. It should be noted that this only gives a very superficial view of the two systems, and it is difficult the grasp the qualitative differences between the vowels. The vowel /ʉ:/ in the Swedish diagram has stronger liprounding than the vowel /ø:/.

Diphthongs:

There are 18 phonemic diphthongs in standard Finnish. The diphthongs of the Malmö dialect may occur both as phonemes and as allophones of the long vowels (Bruce 1970). It should be noted that the Finnish diphthongs may vary considerably from dialect to dialect.

The differences between the diphthong systems of the two languages will be explored further in coming reports, though a list of them is presented below.

The Finnish diphthongs are:
/ie/, /uo/, /yø/, /ai/, /oi/, /ui/, /ei/, /æi/, /øi/, /yi/, /au/, /ou/, /eu/, /iu/, /øy/, /æy/, /ey/, /iy/.

The Malmö dialect diphthongs are:

/i:/ — /e͡i/	/y:/ — /ø͡y/	/u:/ — /e͡u/
/e:/ — /ɛ͡e/	/ʉ:/[1] — /ø͡ʉ/	/o:/ — /ɛ͡o/
/ɛ:/ — /æ͡ɛ/	/ø:/ — /œ͡ø/	/ɑ:/ — /œ͡ɑ/

(This transcription according to Bruce 1970; p. 9.)

Consonant structure:

The main differences between the consonant structures are:

(i) in Finnish there are phonemic contrasts between long and short consonants, this difference is phonotactically related to the shortening and change in quality of the preceding vowel in Swedish. That is all Finnish consonants except /h/, /j/ and /v/ have a phonemic quantity, unlike Swedish consonants.

(ii) the following Swedish consonants are not present in the Finnish structure:

Voiced stops: /b/ and /g/ (they may occur in loan words spoken by educated people)
/d/ is mostly realized as voiceless in dialects.
Voiceless fricatives: /f/, /ç/, /ʃ/ are not present at all.
Liquid: /ʀ/ (that is the Malmö dialect uvular /r/)

Voiced stops: /b/ and /g/ (they may occur in loan words spoken by educated people),
/d/ is mostly realized as voiceless in dialects.
Voiceless fricatives: /f/, /ç/, /ʃ/ are not present at all.
Liquid: /ʀ/ (that is the Malmö dialect uvular /r/).

(iii) the voiceless stops in Finnish; /p/, /t/ and /k/, have no aspiration as they have in some positions in Swedish.

1 /ʉ/ is equal to /ɯ:/ in the vowel system above.

The following diagram shows the articulatory differences between the two consonant systems:

FINNISH:

PLACE OF ARTICULATION		MANNER OF ARTICULATION						
		STOP		FRICATIVE		NASAL	LIQUIDS	RHOTIC
		voiced	voiceless	voiced	voiceless			
Labial:	long		/p:/			/m:/		
	short		/p/			/m/		
Labio-dental:				/v/				
Dental:	long		/t:/					
	short		/t/					
Alveolar:	long				/s:/	/n:/	/l:/	/r:/
	short				/s/	/n/	/l/	/r/
Palatal:				/j/				
Velar:	long		/k:/			/ŋ:/		
	short		/k/			/ŋ/		
Glottal:					/h/			

SWEDISH:

Labial:	/b/	/p/			/m/		
Labio-dental:			/v/	/f/			
Dental:	/d/	/t/			/n/	/l/	
Alveolar:				/s/			
Palatal:			/j/	/ɕ/			
Velar:	/g/	/k/		/ɧ/	/ŋ/		/R/
Glottal:				/h/			

Figure 4. A classification of Finnish and south Swedish consonants by place and manner of articulation.

Phonotactic information:

The main phonotactic differences between the languages are:

(i) Finnish has no initial or final consonant clusters, whereas in Swedish consonant clusters of two and three are permitted both initially and finally in words. For example the Swedish words strand (Eng. "beach") is ranta in Finnish the Swedish consonant cluster /str/ having been reduced to /r/.

(ii) in native Finnish words, vowels occur in a harmony system according to following rules:

A. Back vowels /a(:)/, /o(:)/ and /u(:)/ occur together

B. Front vowels /y(:)/ /ø(:)/ and /æ(:)/ occur together

C. Front vowels /i(:)/ and /e(:)/ can occur with both A and B.

As an example of vowel harmony we can see how it works with the locative endings for "in" as in talo (Eng. "house") - talo-ssa ("in the house") and kylä (Eng. "village") - kylässä ("in the village"). There is no such harmony in Swedish.

(iii) /h/ may occur only initially in Swedish and not as a part of a consonant cluster. In Finnish it may occur before other consonants as in Lahti (town in Finland).

Prosodic information:

There are three main differences in this area:

(i) Intonation: there are two different kinds of tonal accent in Swedish, usually denoted "acute" and "grave" (Hadding 1961) whereas in Finnish there is only one, acute (that is similar to the Swedish acute).

(ii) Stress: normally the first syllable is stressed in Finnish words. Swedish stress rules are more complicated.

(iii) Quantity: according to the Swedish quantity rule a stressed syllable should either contain a long vowel, or a short vowel + a long consonant as for example in /m'ɑ:t/ (Eng. "feed") and /m'ɑ̆tɑ/ (Eng. "carpet"). In Finnish a stressed syllable could well contain a short vowel and a short consonant as in /tu'li/ (Eng. "fire").

There are different temporal relations between the vowels and consonants in Finnish and Swedish words. The exact nature of these relations has to be determined through instrumental analysis. Some research has been done in this area by Jonasson and McAllister (1972) who have investigated the temporal relations between vowels and consonants in Swedish and English.

Much more must of course be added for an exhaustive contrastive phonetic analysis of the two languages, but for the present this is sufficient.

METHOD

Material:

Several tests have been designed in order to obtain the widest possible grasp of what constitutes a foreign accent, that is, to find out how and under what circumstances it appears. Equal tests for both languages were designed.

Tests I and II: Two articulation tests were devised, based on the most frequent phonetic and phonotactical units in Finnish and Swedish. Each test consists of 61 items. Drawings and paintings of objects are shown to the child who is asked to name the objects in the language being tested. Both test I and II consist of the same pictures, though given in different presentations and order. I tried to find words that are known to most 3-4 year old Swedish and Finnish children. More difficult words are mixed with easier words (that is, more frequent in the children's languages). The examples below illustrate this:

Test I (Swedish version): Item 4. apa ("monkey") 5. vingar ("wings") 6. fjäder ("feather") 7. ekorre ("squirrel")

Test II (Finnish version): Item 7. saksit ("scissor") 8. suihko ("shower") 9. ruisko ("squirt") 10. joulukuusi ("Christmas tree")

The tests will be described and treated in more detail in forthcoming reports.

Test III and IV: In order to collect spontaneous fluent speech under compatible forms four pictures were drawn.

Test IIIa and IIIb with Swedish instructions; (a) a view of a Swedish garden with children playing, drawn to resemble their own neighbourhood; and (b) a picnic in the countryside. Picture IIIa has the following instructions: 1. "I have drawn a picture that I tried to make as like your garden as possible. Is it right?" 2. "What are they doing in the picture?"

Picture IIIb has the following instructions: 1. "What are they doing in the picture?" 2. "Have you ever been on a picnic? Tell me about it!" Test IVa and b ; with Finnish instructions; (a) a living room and (b) a child's bedroom.

Picture IVa has the following instructions: 1. "I have tried to draw a picture that resembles your living room. Is it right?" 2. "What are they doing in the picture?" Picture IVb has the follwoing instructions: 1. "What are they doing in the picture?" 2. "What is your bedroom like? Tell me about it!"

The different subjects of the four pictures were chosen for their bilingual environment. In their home the children mostly speak Finnish and in the garden they have to use the Swedish language to communicate with other children. This choice of subject is intended to create the best possible setting for each language.

Some objects in the pictures are the same as in the articulatory tests; for example "dog, cat, ball". This is to see whether pronunciation differs in spontaneous speech.

The testing interval for each picture is five minutes and there is wide variety as to how these minutes are used by the individual child. Some children can speak fluently for five minutes about the picture while others use only single words. The test situation depends on some of these effects.

These four tests are given regularly to each child in Oct-Nov and April-May.

Subjects:

All four groups of children were included in the four tests. All of the children have Finnish as their mother tongue and their competence in Swedish differs over a wide range from those who speak Swedish fluently to those who only know a few words.

The experimental group consists of 14 five-year-old children, 15 six-year olds, 6 seven-year olds and 15 eight-year olds. This makes a total of 50 children (23 are girls and 27 boys). The subjects and their background are partly described in Aronsson and Wigforss (1974; The bilingual nurseryschool in Malmö 1973-74).

Procedure:

The Finnish tests were given by one of the teachers in the nursery school who was quite well known to all children. The Swedish tests were given by the Swedish teacher in the nursery school and by myself in the primary school. Test sessions were tape-recorded on a Nagra and I made simultaneous notes during all test sessions except those I lead myself. The child's behaviour was also noted during test sessions. The tapes are of a high quality and will be subjected to instrumental phonetic analyse at a later date.

RESULTS

The results very briefly presented here are sampled from the first auditory analysis of the tape recordings from the four tests. I have tried to collect as many phonetic errors in the Swedish performances as I could find in the material, but I do not wish to claim that they are the only errors or that they have always been correctly percieved. The ear is a subjective tool.

This analysis, superficial though it is, will however lead to some tentative hypothesies, after which I will continue with further investigation of the material.

1. The first very remarkable thing that appears in the material is that some of the linguistic performances in the classroom are amazingly free from foreign accent. We can call this performance classroom-phrases. Most of these phrases have a very heavy south Swedish dialect accent of the following type:

   Sw. [gɛ̂e mæ͡eɛj gʋrːi:t] (Eng. "give me the rubber erasor") or Sw. [væ͡ɑᵊ tyᵊst] (Eng. "be quiet").

   Even if I have not yet found evidence in my material that isolated words from those phrases will also be correctly pronounced in other contexts, I strongly suspect that this will not be the case. I believe these phrases are due to imitation and some of the single phonetic units cannot yet be used in a productive manner in the child's phonological system.

2. Frequent errors in the pronunciation of vowels are misarticulations of:

   (i) the vowel /ʉ:/ in for example Sw. /rʉ:t / (Eng. "window"). In

Swedish we have the following contrasting pairs: /ri:tɑ/ (Eng. "to draw"), /ry:tɑ/ (Eng. "to roar") /re:tɑ/ (Eng. "to tease"), /rø:tɑ/ (Eng. "decay"), /rʉ:tɑ/ (Eng. "window") and /ru:tɑ/ (Eng. "to dig up"). All of these words are quite frequent so it is important to give the vowel its right quality. In the Finnish vowel system there is no such vowel as /ʉ:/ and a comparison will give us the following pattern:

Swedish has phonemically	/y:/	/ʉ:/	/u:/
Finnish has phonemically	/y:/		/u:/

This aspect causes considerable trouble for Finnish speakers not only regarding the pronunciation of the missing /ʉ:/ but also of the other two vowels. My impression of the children's mispronunciation of /ʉ:/ is that is seems too retracted and insufficiently rounded. A spectrographic analysis will probably reveal the more exact relation between these three vowels.

(ii) All long vowels seem to have much longer duration than the idiomatic long vowels of the Malmö dialect. Nor do they have the right quality apart from /i:/, /e:/ and /ɛ:/.

(iii) The vowel /a/ as in /tand/ (Eng. "tooth") is very often produced as a shortening of the long /ɑ:/ vowel (i.e. too dark).

3. Frequent errors in the pronunciation of the consonants are:

(i) Failure to aspirate voiceless stops.

(ii) No distinction between voiced and voiceless stops; all stops are voiceless. For example [pɔl] for [bɔl] (Eng. "ball"); [inti:ɑn] for [indi:ɑn] (Eng. "indian") and [klɑs] for [glɑs] (Eng. "glass").

(iii) The following voiceless fricatives are mispronounced:
/ɧ/ ⟶ [s] or [f] as in /ɧe:d/ (Eng. "spoon")
/ç/ ⟶ [s] or [k] as in /çyrkɑ/ (Eng. "church")

(iv) The lateral /l/ is often pronounced retroflexed, which gives it it a very strange character for the Malmö dialect.

(v) The liquid, uvular [R] is mostly realized as a Finnish alveolar [r].

(vi) The durations of the individual consonants are often longer than the surrounding segments. This could be due to stress too, and an instrumental analysis will show whether this is the case or not.

4. Frequent phonotactic errors are:

(i) the deletion of one or more consonants in initial consonant clusters for example /sku:/ --- [ku:] (Eng. "shoe") where /s/ is deleted.

(ii) use of vowel harmony, most frequently in the addition of Finnish syntactical endings to Swedish words as [ru:se:ngo:rdista] ("in Rosengård", a part of Malmö).

5. Frequent prosodic errors are:

(i) a very heavy stress on the first syllable in words that sometimes makes comprehension impossible.

(ii) use of only one tonal accent, the acute. The exact nature of how this affects the acoustic output has to be investigated instrumentally.

(iii) overall temporal relations that are neither Finnish nor Swedish, such as the relation between consonants and vowels which seems to vary unsystematically whereas in idiomatic Swedish there is a fairly constant relation between the vowels and consonants in a word.

CONCLUSIONS

The results are a few examples taken from the entire material. No case study based analysis has yet been done which would indicate whether or not errors are consistent from one individual to another. This bilingual group will probably show a great variety of errors, some being integrated parts of their performance and others due to occasional mistakes.

It should once more be emphasized that this is a very preliminary analysis. But nevertheless I think it is already possible to find three kinds of children, grouped around the kind of foreign accent they have. A very tentative grouping could look like this:

Group I Children who speak Swedish fluently and have a very typical south Swedish pronunciation. Most of these children have met the Swedish language at a very early age, through sisters and brothers or Swedish playmates. Some of them have probably started to communicate in Finnish and Swedish at the same age.

Group II These children have a very slight accent that is mostly characterized by misarticulations of /ŭ/ as in /hŭnd/ (Eng. "dog") and /ɧ/ as in /ɧe:d/ (Eng. "spoon") and /ç/ as in /çyrka/ (Eng. "church"). Most of them have probably met and learned Swedish before the age of 5, i.e. before they began to attend the Finnish nursery school; they were otherwise fluent in Swedish.

Group III These children have a very strong foreign accent, manifested through errors in both phonemic and prosodic patterns. Most of them often make all the errors described in the results. It is very easy to hear that they have a Finnish accent, since they place a heavy stress on the first syllable in words. Sometimes it is very difficult even to understand the exact nature of a word.

This group of children has not met Swedish as a communication language (except through TV). They have difficulties in understanding sentences and even words in Swedish.

If it is possible to group bilingual children in this fashion, this rises several interesting questions for teaching. Will these group differences remain through our instructional models? Will the third group still retain their foreign accent as adults?

Can we talk about critical ages for children, after which they cannot acquire a bilingual pronunciation that is free from accent? How important is it for a future bilingual to have met the intonational pattern of his second language before the age of 5?

The longitudinal approach based on case studies of the project may give some of the answers to these questions.

REFERENCES

Aronsson K. and Wigforss E. 1974. Den tvåspråkiga lekskolan i Malmö 1973-74. Report from the Institute of Education, University of Lund. No 53

Bruce G. 1970. Diphthongization in the Malmö dialect. Working Papers 3: 1-20. Phonetics Laboratory, University of Lund

Dahlstedt K. 1967. Svårigheter i svenskans uttal: En handledning vid undervisningen av finska, tyska och jugoslaviska invandrare. Modersmålslärarnas förening, Lund

Fishman J. 1970. Sociolinguistic perspective of the study of bilingualism. Linguistica 39:21-49

Hadding-Koch K. 1961. Acoustico-phonetic studies in the intonation of southern Swedish. Travaux de l'Institut de phonétique de Lund, III

Hadding K. 1974. Barnets språktillägnande. Mimeo. Phonetics Laboratory, Lund

Hadding-Koch K. and Studdert Kennedy M. 1963. A study of semantic and psychophysical test responses to controlled variations in fundamental frequency. Studia Linguistica XVII p. 65-76

Hansegård K.-E. 1968. Tvåspråkighet eller halvspråkighet. Stockholm: Aldus/Bonniers

Haugen E. 1953. The Norwegian Language in America I

Haugen E. 1965. Bilingualism in the America: A bibliography and research guide. University of Alabama press

Irwin R. 1974. Evaluating the perception and articulation of phonemes of children, ages 5 to 8. Journal of Communication Disorders 7: 45-63

Johanesson I. 1974. Aims and goals for bilingual - bicultural education. Unpublished paper. Department of Education, Lund

Johansson F.A. 1973. Immigrant Swedish phonology. Travaux de l'Institut de linguistique de Lund, IX

Jonasson I. and McAllister R.H. 1972. Foreign accent and timing: an instrumental phonetic study. Papers from the Institute of Linguistics, University of Stockholm (PILUS) 14: 11-40

Kangas T. 1973. Val av undervisningsspråk - forskning eller politik? Nordisk psykologi 25 (2): 105-112

Karlsson F. 1969. Suomen yleiskielen segmentaalifoneemien paradigma. Virittäjä 4. Helsingfors

Mackey Wm.F. 1965. Bilingual interference: its analysis and measurement. Journal of Communication 15: 239-249

Reuter M. 1971. Vokalerna i finlandssvenskan: en instrumentell analys och ett försök till systematisering enligt särdrag. Festskrift till Olav Ahlbäck 28.3.1971

Reuter M. 1973. Kvantitetsförhållanden i Helsingforssvenskan. Folkmålsstudier XXIII: 214-224

School Research 1973:8. Newsletters, current project. Models for the bilingual instruction of immigrant children

Weinreich U. 1963. Languages in contact: Findings and problems. Second printing. The Hague, Mouton

Wiik K. 1965. Finnish and English vowels. Åbo

WHAT IS THE DIFFERENCE BETWEEN ENGLISH AND SWEDISH DENTAL STOPS?

Sidney Wood

During the past few months Gösta Bruce and I have cooperated in the production of several X-ray motion films of speech movements for investigations of vowel articulations and cavity formation. Each of us has his own informants and his own specific problem area, but the practical details have been arranged jointly. While we have hardly had time to analyse the films in depth it is already clear that they reveal many details of interest for the study of speech production over and above the vowel movements they were designed to portray. One such detail is the laminal articulation of dental stops by an English subject in contrast to the apical articulation of a Swedish subject.

Swedish and English dental stops

The English dental stops are usually described as having an alveolar place of articulation. There is a phonemic contrast of voice, /t – d/ (tin – din), but there is no contrast with other dental stops. The nearest contrasting stops are the affricates /č, ǰ/ (chin, gin), usually described as palato-alveolar. The Swedish dental stops are usually said to be **truly** dental or post-dental. Here too, the only contrast is one of voice, /t – d/ (tal – dal), not of place. In central and northern dialects (where /r/ involves the tongue tip) all dental consonants are retracted after /r/ with which they coalesce. In this respect Swedish (like Norwegian) resembles a number of the languages of India. The retracted set is referred to as supradental or retroflexed. The sandhi character of this process is obvious across word and morpheme boundaries as in har du ([haɖɯ]) but within morphemes there are minimal pairs such as bod – bord ([bu:d] – [bu:ɖ]). It is a matter of linguistic creed whether the minimal pairs are used as evidence for a phonemic contrast between retracted and non-retracted dental stops (/būd/ – /būɖ/) or are generated by the same sandhi rules as are necessary at morpheme boundaries (/būrd/ ⟶ [bu:ɖ], bord).

However, the Swedish subject described in this paper does not represent this type of dialect. He has instead a southern dialect with a dorso-uvular [ʁ] for /r/ and therefore lacks the retracted set of dentals, the sandhi rules not being applicable in this case. Nevertheless, it is still necessary to bear the retracted supradental set in mind during a general discussion of Swedish dentals, and especially for a comparison with English, since many Swedes unfortunately seem to prefer the retracted set when speaking English.

The traditional view expressed in pronunciation handbooks is that the English stops are articulated further back along the tooth ridge. This well established belief ought to be founded on reliable evidence since there is excellent tactile sensation in the tongue tip and it is relatively easy to make palatograms and linguograms.

Swedes are very conscious of an audible difference between English and Swedish dental consonants and look upon the quality of the English stops as a typical ingredient of an English accent in Swedish. Being so keenly aware of such an audible difference, they tend to be concerned not to let it tarnish their English pronunciation. My own personal view is that a Swedish [t] or [d] is more acceptable in English than the English consonants would seem to be in Swedish, but I cannot say whether other Englishmen are equally tolerant. Some Swedish language teachers instruct their pupils to aim at the retracted supradental reflex of /...rt.../ or /...rd.../ as an approximation to the English sounds. I find this gives a distinctly foreign quality to their English speech. Johansson (1967) has taken still radiographs of Swedish dental and supradental consonants, and found the latter to be completely retroflexed, with occlusion made by the underside of the tongue tip. Gimson (1962: p. 159) advises foreign students of English against the use of retroflex stops "since these sound over-retracted to the English ear".

Although both the X-rayed subjects have the same alveolar place of articulation for the dental stops, the Swedish subject still finds the English subject's [t] typically foreign in Swedish. The only apparent articulatory difference is between laminal and apical occlusions. Unfortunately, it is impossible to know how typical these two types are for the respective languages, there being no statistical evidence available to confirm any language-specific preference. It is tempting to speculate whether or not the perceived difference is in some way related to the articulatory difference between apical and laminal occlusions. Acoustically, there may be differences in the formant frequency transitions in adjacent vowels, in the spectral character of the plosive burst and in the duration of the burst and the timing of the voice onset of the following vowel.

Apical and laminal stops in the literature

Jones (1956: § 515) described the laminal articulation as a less common variant of the purely apico-dental, providing "a very unnatural effect" in the English speech of "many foreign people, eg French, Italians, Hungarians and some Germans". He describes the English stops as apico-alveolar. Jespersen (1899: §191) records an observation by Passy that French dental stops tended to be laminal. Jespersen recognized that the laminal articulation occurred frequently as an individual variant but he doubted whether the audible difference could be so great that a given language would prefer the one or the other. Malmberg (1967, pp. 95-96) describes the laminal articulation as an individual variant or as dialectal (frequent in central and northern dialects of Swedish and in French). He considered the acoustical differences to be insignificant. Catford (1968: p. 327) lists lamino-dental to lamino-post-alveolar stops as "variants of apical stops occurring in English and elsewhere". The traditional view is thus against a language-specific preference for either articulation apart from the apical being normal and the laminal a variant and against any audible difference between the two.

On the other hand, Ladefoged (1954: p. 19) has found phonemic contrasts between apical and laminal stops in the West African languages Twi, Ewe, Temne and Isoko, which indicates that it must be possible to obtain a useful acoustic difference in this manner and that some languages do categorize the two articulations. Chomsky and Halle (1968: §7.4.4) generalized from this and similar evidence and proposed a contrast between sounds with long constrictions (denoted [+distributed]) and sounds with short constrictions (denoted [-distributed]). According to their generalization, two contrasting places of articulation within the denti-alveolar region are always associated with a difference in the length of the constriction. A new case pointing in the same direction has recently been reported by Nihalani (1974) who has made a palatographic and radiographic examination of the stops of Sindhi. He disagreed with the standard view that the retracted dentals of Sindhi are retroflexed

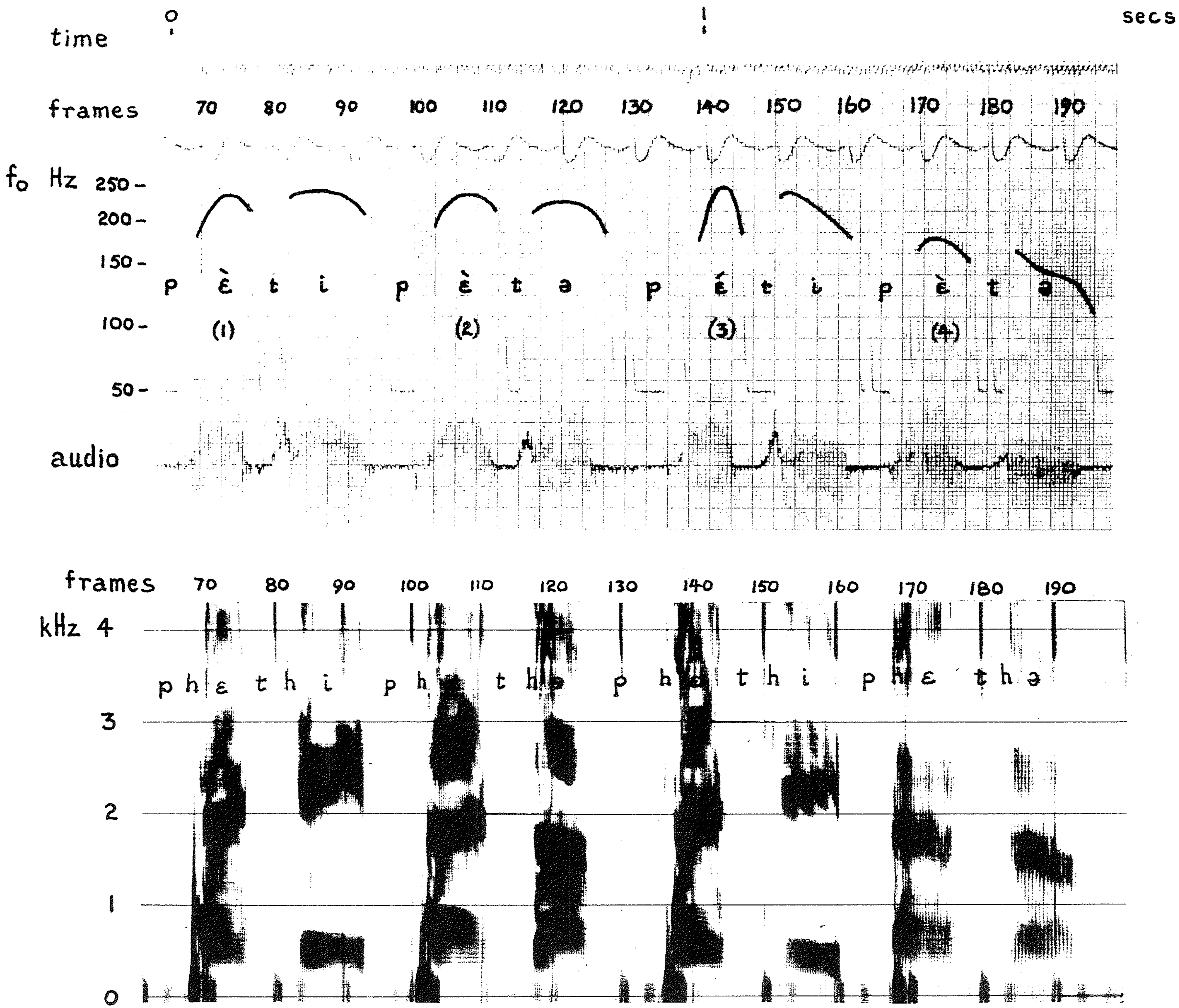


Fig. 1. (a) Oscillograms and (b) spectrograms of one of the English sentences, illustrating the synchronization pulse that appears at every tenth frame on the film.

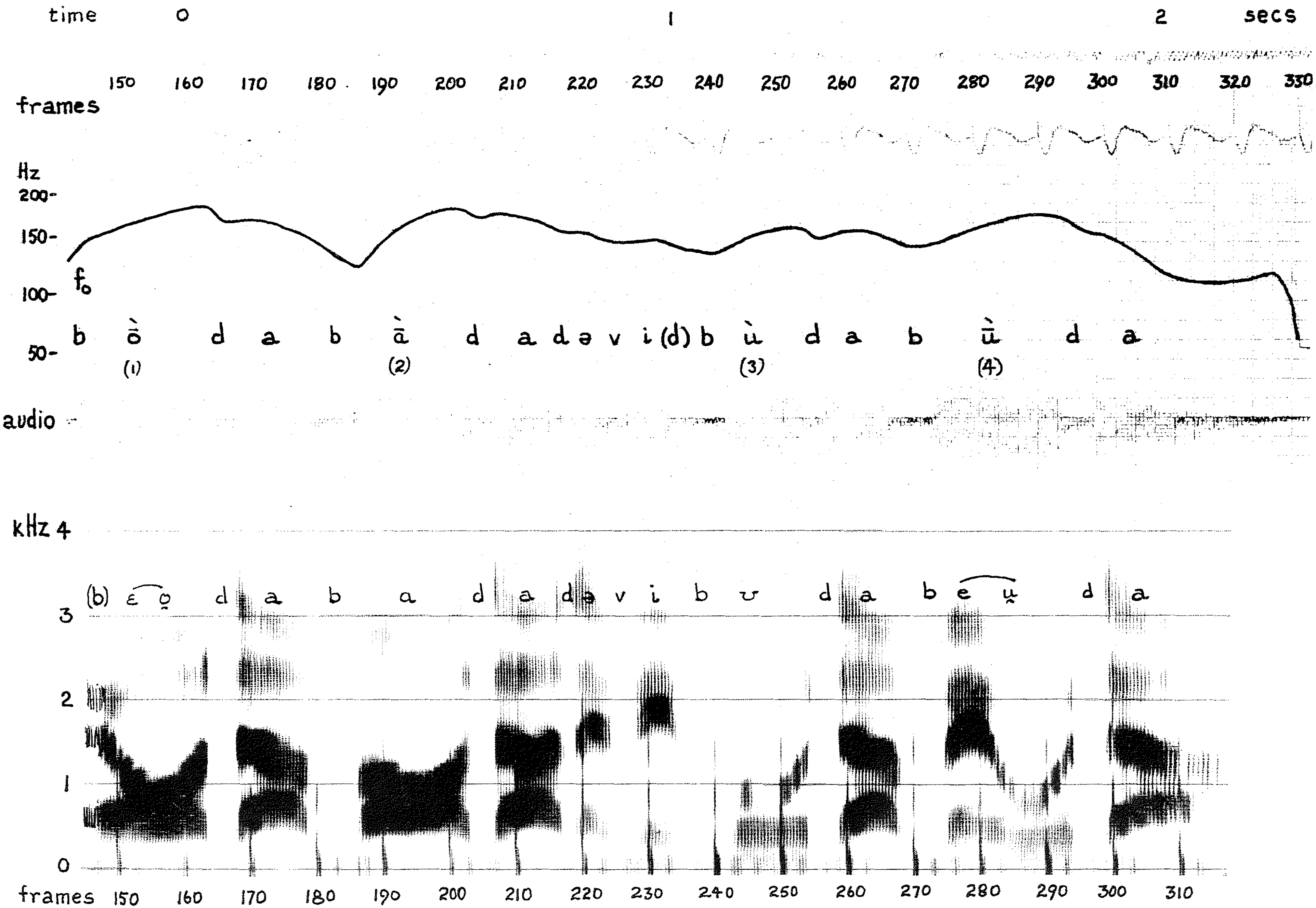


Fig. 2. (a) Oscillograms and (b) spectrograms of one of the Swedish sentences, illustrating the synchronization pulse that appears at every tenth frame of the film.

as in other Indian languages. He found that the non-retracted stops [t, th, d, dh] were laminal with a denti-alveolar location, while the retracted set were apical and post-alveolar but not retroflexed.

Cineradiography

The films were made at the County Hospital, Lund, with the consent of Professor Olof Norman and the assistance of members of the Roentgen Technology Section (Dr Thure Holm, Radiophysicists Gunnila Holje and Gudmund Swahn and Technician Rolf Schöner). The customary high voltage technique was used (120 kV) in order to take advantage of the greater transparency of bone to higher frequency X-radiation, thus producing adequate pictures of both bone and soft tissue.

The laboratory used is primarily intended for angiographic examinations with equipment specialized for observing events in soft tissue such as blood vessels. An added practical advantage is that the camera there provides a synchronizing pulse which appears on every tenth frame and which for clinical purposes can be recorded alongside an elctrocardiogram. We recorded this latter signal simultaneously with the microphone signal, separately on a parallel track. The microphone and synchronizing signals are subsequently processed separately for oscillographic records (Figs. 1a, 2a) but are mixed for spectrograms (Figs. 1b, 2b). We have no opportunity to screen the subject's speech from the relatively high ambient noise level in the laboratory (the high pitch whine of the rotating anode and the rapid intermittent tapping of the camera mechanism). To reduce the recorded level of the background noise relative to the speech we place a highly directional microphone close to the subject's mouth. The noise is audible on the speech recording but barely appears on spectrograms. In order to facilitate enlargement of the radiographs to life size, a long stiff coppar wire with blobs of solder at 1 cm intervals is taped to the subject's forehead and nose and bent so as not to interfere with the movement of the lips and mandible. The camera speed is 75 frames/second (one frame every 13.33 ms) and the duration of radiation is 3 ms per frame. We have 30 to 40 seconds effective time available per reel of film. As a safety precaution each subject is limited to one reel.

The speech utterances that were X-rayed were primarily designed for an investigation of vowel articulation. The present paper is based on articulatory data provided by the consonants framing the vowels. Dental stops (English [t] and Swedish [d]) were repeated four times in each sentence, always intervocalic and following a strong vowel. The utterances conformed to the grammatical structure of a model sentence in which vowels were substituted systematically. Each sentence contained four strong positions in which vowels were substituted systematically. The English sentences were "petty pet a petty-petter, potty pot a potty-potter, ... etc.". For Swedish sets of three vowels were rotated through four strong positions, thus: "Båda badade vid Bodda boda, Boda bådade vid Badda bada, Bada bodade vid Bådda båda, ... etc.". The English sentences were uttered at about 4.5 syllables/second, the Swedish sentences at about 4 syllables/second. A further set of English sentences, uttered at about 6.5 syllables/second, has not yet been analysed and could not be included in the present comparison of dental stops.

This paper is devoted to a comparison of two subjects only - myself (Southern British English, NE Kent) and Gösta Bruce (Helsingborg, NW Skåne).

The dental stops of the X-rayed subjects

A striking difference between the two subjects was in the formation of dental stops. Fig. 3 shows how the English speaker made his dental stop with a flat area of the tongue behind the tip (unbroken line, depicting the moment of occlusion). Fig. 4 shows how the Swedish subject made his dental stop with the very tip of the tongue (unbroken line). These tracings distinguish between mandibular and lingual gestures.

The actual shape of the tongue body at the instant of occlusion depends on the preceding vowel. The broken lines at Figs. 3 and 4 show the preceding vowel positions. These represent the turning point (lasting from 1 to 3 frames or about 10 - 50 ms) between the end of the CV movements and the beginning of the [t] or [d] movements. At these speaking rates (4 - 5 syllables/second, a little more leisurely than everyday speech) the gestures of the dental stops (mandible elevation, tongue tip raising) were not initiated until the vowel configurations had been completely formed. The cases illustrated are for a preceding palatal vowel [ɛ] and two pharyngeal vowels [o] and [ɑ]. For the Swedish subject, [o] is the final portion of a diphthong [ɛ͡o̯] for /ō/ (cf.

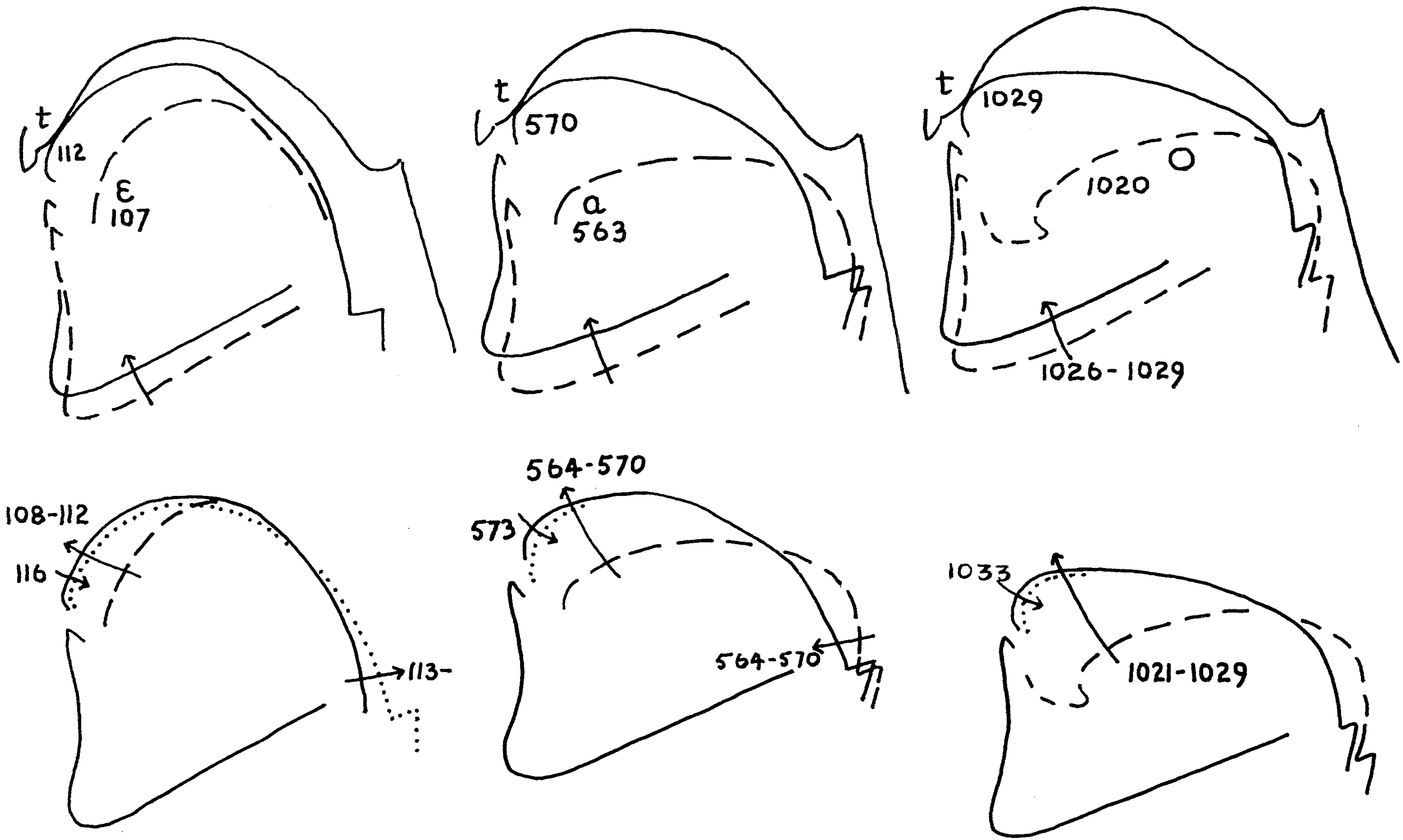


Fig. 3. Tracings of three laminal [t] occlusions after [ε, ɑ, o] by the English subject. Each frame number represents a 13.3 ms interval. Arrows indicate direction of movement.

--- preceding vowel —— occlusion ····· burst

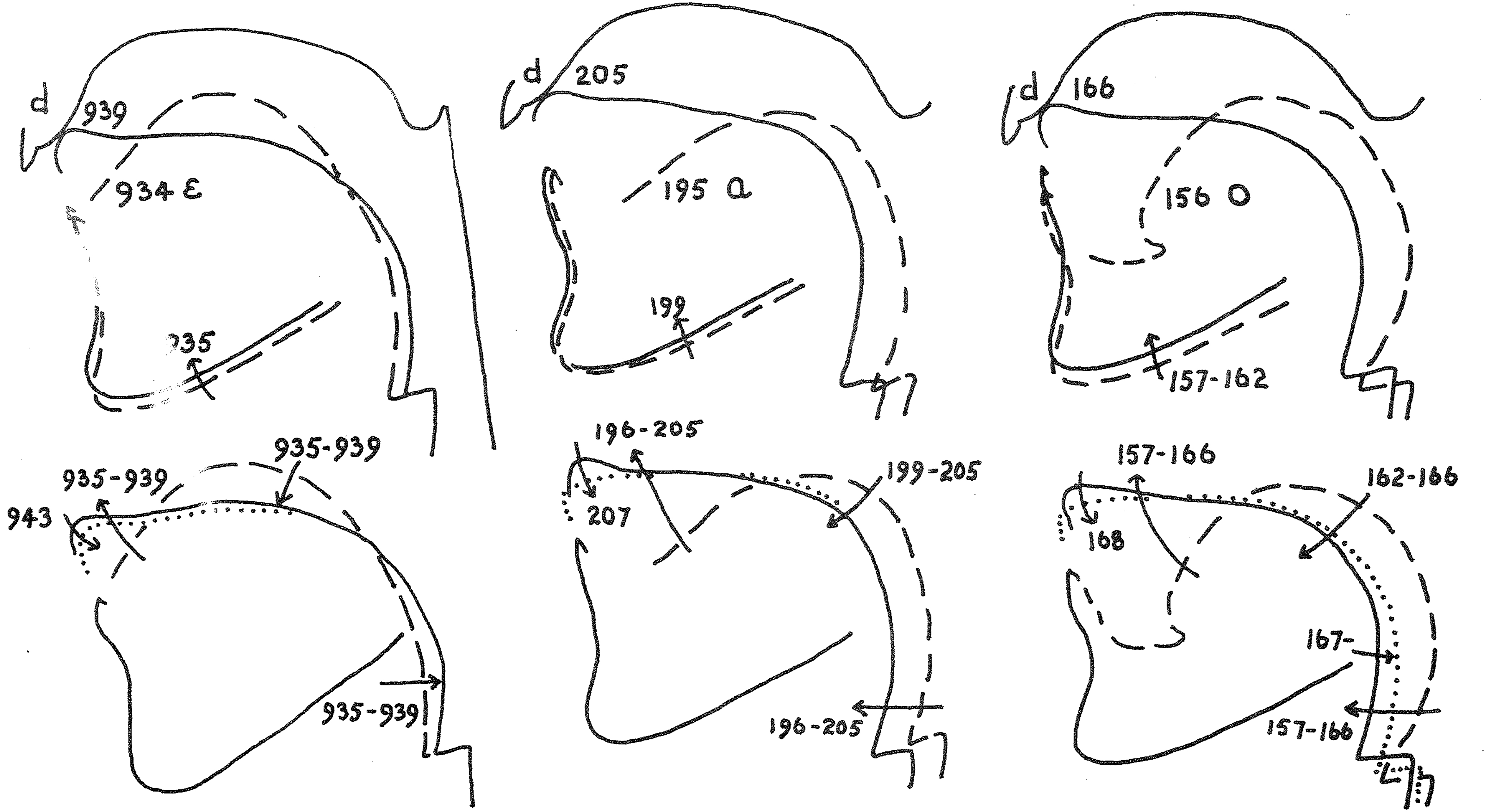


Fig. 4. Tracings of three apical [d] occlusions after [ε, ɑ, o] by the Swedish subject. Each frame number represents a 13.3 ms interval. Arrows indicate direction of movement.

--- preceding vowel — occlusion ····· burst

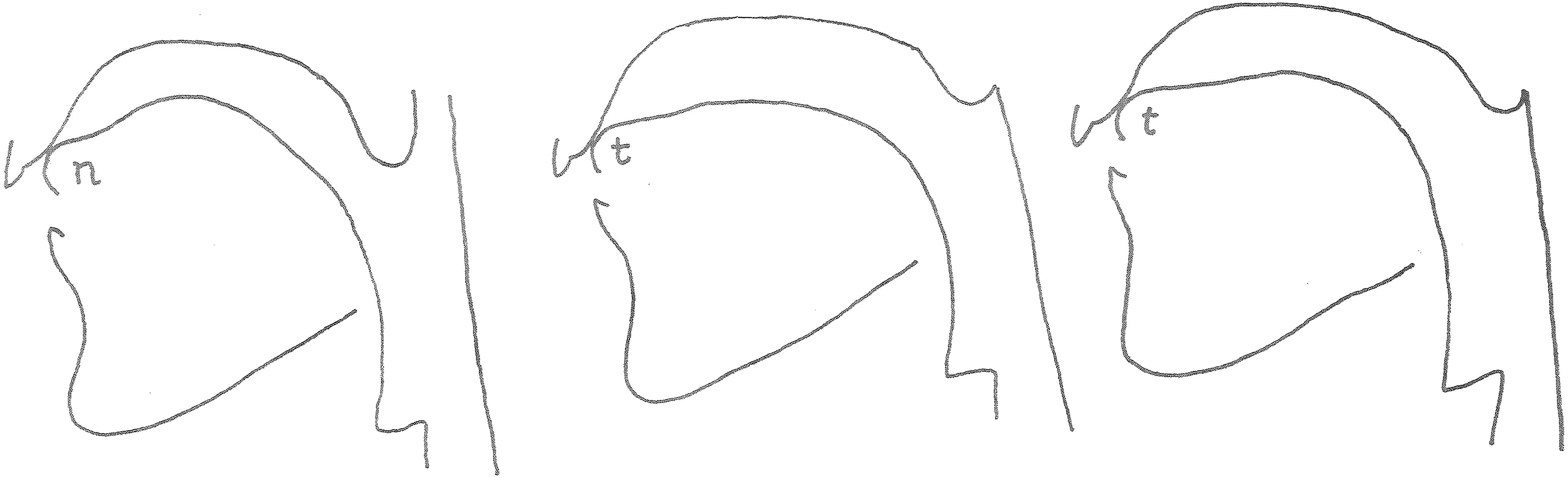


Fig. 5. Tracings of alveolar occlusions by the Swedish subject: [n] in nitton, [t] in nitton and [t] in sjuttio.

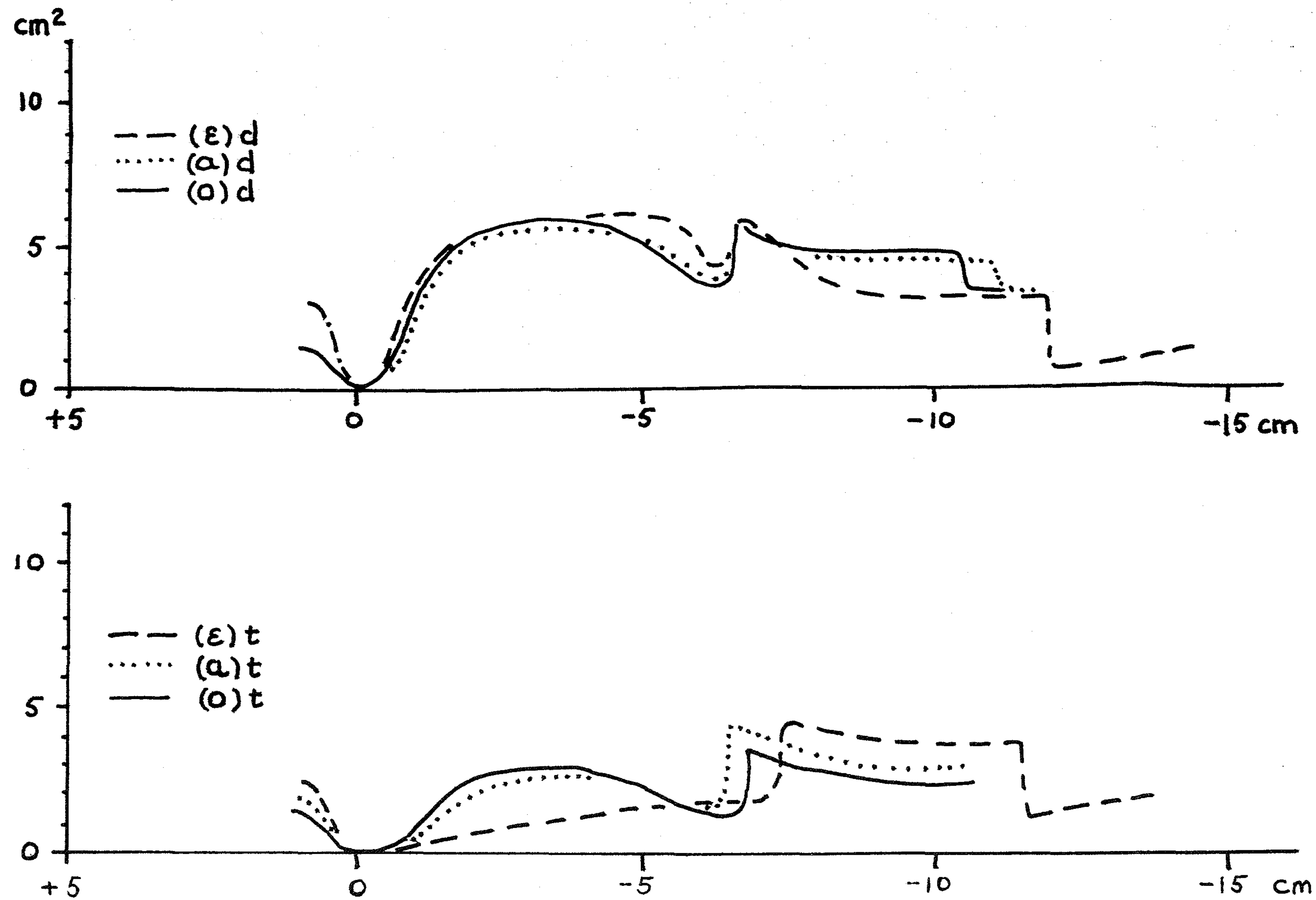


Fig. 6. Area functions of the Swedish subject's apical occlusions at Fig. 4 (above) and of the English subject's laminal occlusions at Fig. 3 (below).

Bruce 1970 and see the spectrograms at Figs. 2b and 7b). Similarly, the shape of the tongue body at the moment of release (dotted lines) depends on the configuration of the oncoming vowel towards which the tongue is moving. To make the examples as comparable as the material will permit, the cases illustrated are [ə] following the English [t] and weak [a] following the Swedish [d].

Comparison of Figs. 3 and 4 indicates practically the same place of occlusion for both subjects. This need not necessarily contradict the general belief in different places of articulation for English and Swedish stops since the Swedish subject is from Helsingborg and Malmberg (1967: pp. 94-95) has reported that the northern part of the province of Skåne is one of several Swedish areas where the dialects have stops that are alveolar rather than dental.

There is also a difference of voicing between the utterances of the subjects ([t] for the English, [d] for the Swedish). Malmberg (p. 95) also stated that there are frequent instances of different places of articulation for voiceless and voiced dental stops, [d] often being more retracted than [t]. He specifically mentioned the south of Sweden generally, reporting that [t] is dental but [d] post-dental. The Swedish subject included his name and the date in his film, and has thus provided examples of other dental consonants. Fig. 5 shows the initial [n] in nitton, the [t] in nitton and the [t] in sjuttio. These are all clearly alveolar and indicate that this subject does not have different places of articulation for [t] and [d].

There is an inherent difference of tongue body shape between the laminal and apical types of articulation. The crown of the tongue arch is clearly higher during the period of occlusion for the laminal articulation than for the apical. This is because the tongue curves upwards behind the laminal occlusion whereas the tongue tip approaches the tooth ridge from below for the apical occlusion. Consequently the vocal tract widens abruptly behind the apical occlusion but gradually behind the laminal occlusion (Fig. 6). The different vocal tract resonance conditions resulting from this will be discussed further on.

In order to compare the acoustical characteristics of [t] segments for both subjects, two separate series of teswords have been specially recorded: ['kvVta] (Swedish) and ['kVtə] (English), where [V] represents a range of vowels for commutation. Most resulting testwords were nonsense

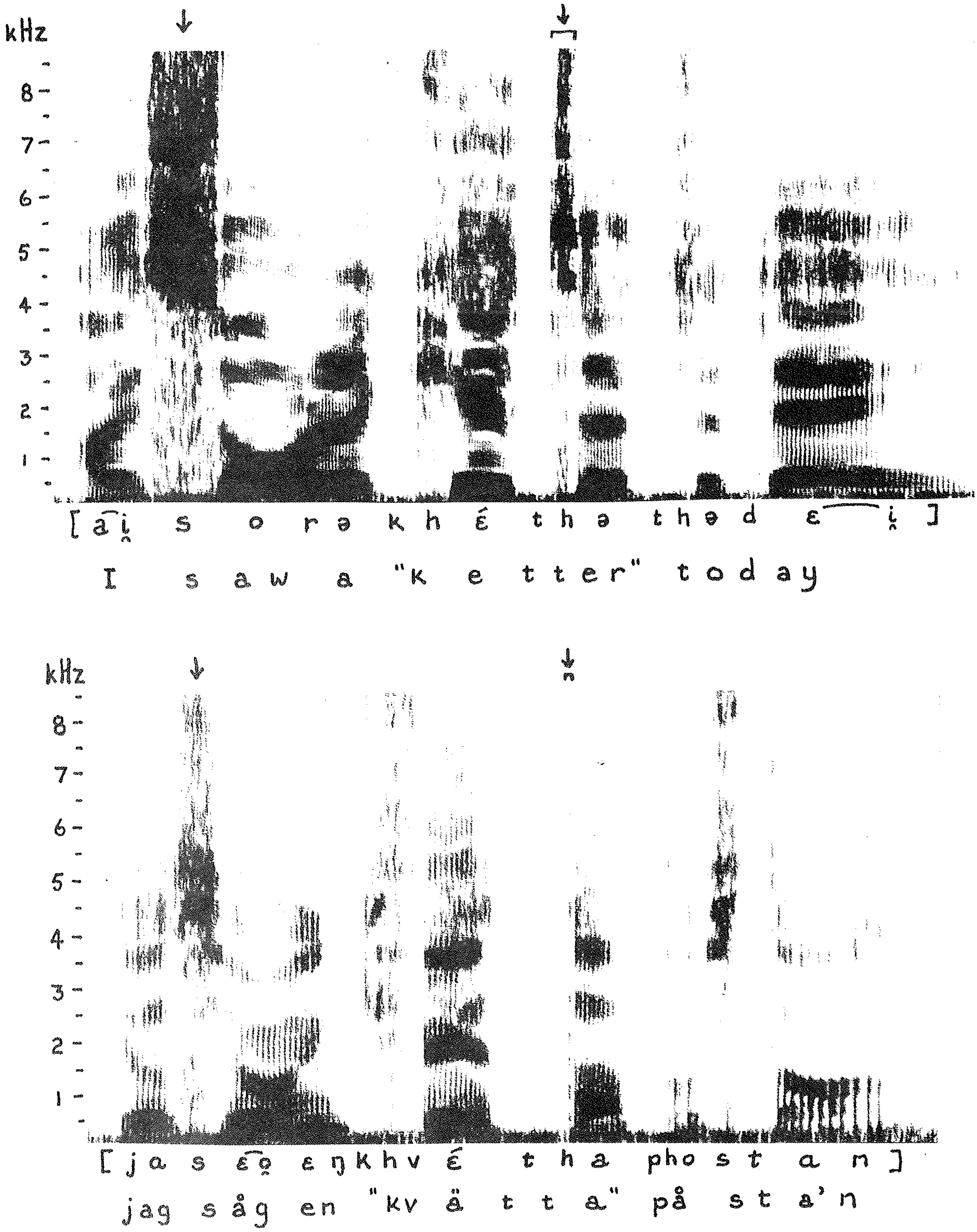


Fig. 7. Spectrograms of two examples of the sentences recorded for a comparison of [t] segments in (a) English (above) and (b) Swedish (below). The arrows indicate the locations of spectral sections (Fig. 8).

but a few were common real words - Swedish kvitta ("settle a debt") and English cutter, carter, cotter (and with a little imagination even kitter, catter, cooter). The velar stop was introduced in order to have a different lingual consonant alternate with the [t]. The fricative [v] was included in the Swedish testword in order to block spirantization of the palato-velar [k] to [ç] before palatal vowels. The testwords were placed in a frame sentence jag såg en ... på sta'n or I saw a ... today, where the [s] segments were also used for comparison.

Spectrograms of one pair of renderings are given at Fig. 7. One very striking difference is the duration of the [t] occlusions in the test-words - the Swedish occlusion was twice as long as the English although the total sentence duration is virtually the same for both speakers. This is related to the characteristically longer postvocalic consonant following short Swedish vowels.

Equally striking are the durations of the [t] bursts - about 40 ms in the English rendering against 10 or 20 ms in the Swedish. There is a corresponding difference in the voice onset. The spectral character of the English [t] burst is similar to that of the [s] - high levels above 4000 Hz with several peaks from 4300 to 7000 Hz. The Swedish burst appears to be weaker and more evenly spread up to about 7000 Hz with negligible levels above while the [s] has two peaks at 4000 - 5000 Hz, somewhat lower levels above and even less below.

There is also a difference in the transitional phases of the vowels towards the [t] occlusion - F2 in the English [ɛ] at Fig. 7 points to a locus at about 1800 or 1900 Hz while it points down more sharply in the Swedish [ɛ] to a locus at about 1600 Hz.

The spectral differences in the [t] bursts can be studied in greater detail at Fig. 8 which shows the spectra sampled at approximately 4 ms intervals (the time scale is very rough). The [s] spectra are included for comparison. This confirms the [s] like character of the English subject's [t] burst, with high levels from 4000 Hz to beyond 8000 Hz. The Swedish subject's [t] burst had a much lower overall intensity, mostly at 4500 Hz and below. The Swedish [t] burst does not resemble the same speaker's [s]. At the same time, the Swedish [s] is different from the English subject's - the Swedish [s] is almost symmetrical around two strong peaks at 4500 - 5000 Hz.

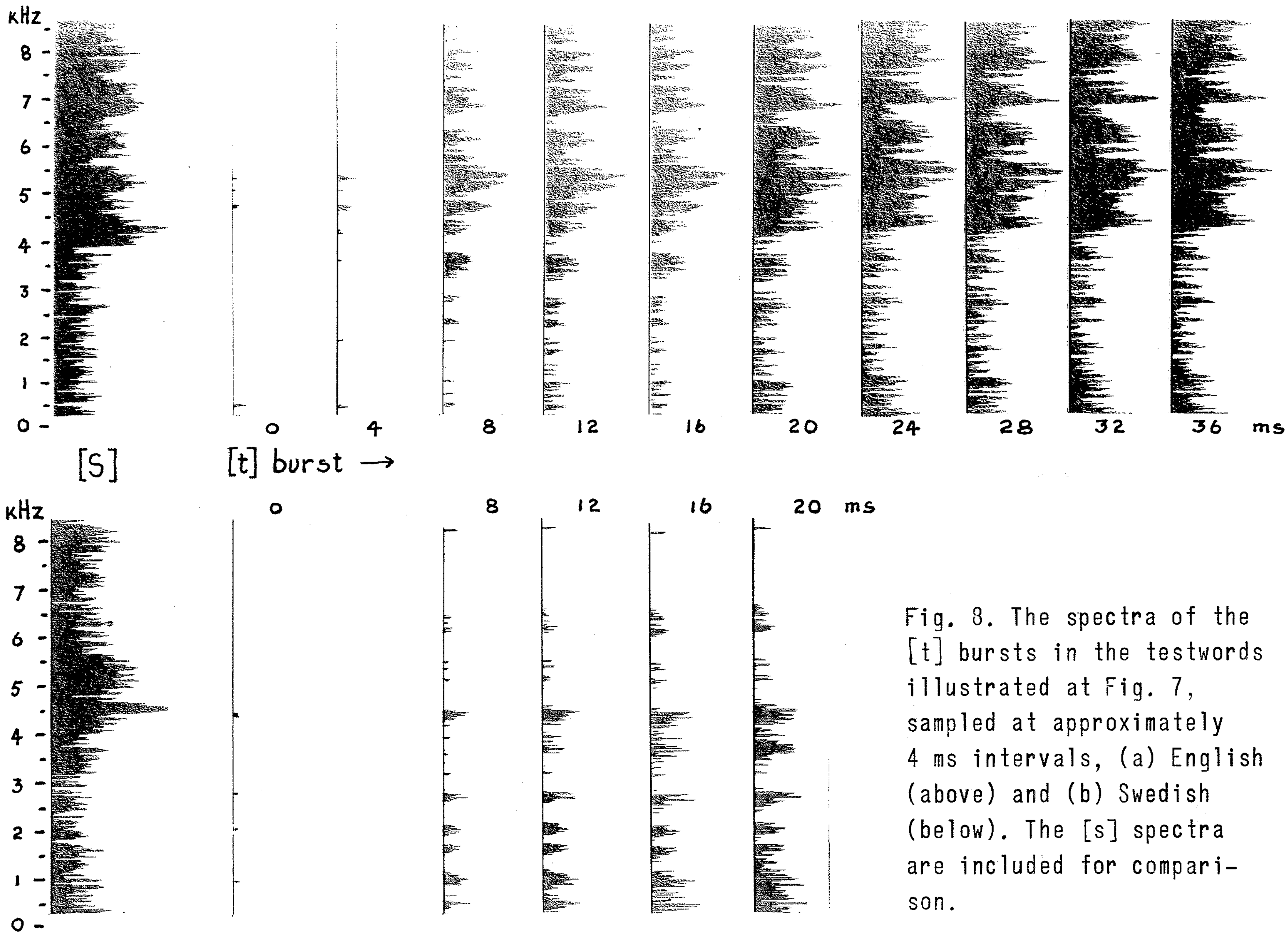


Fig. 8. The spectra of the [t] bursts in the testwords illustrated at Fig. 7, sampled at approximately 4 ms intervals, (a) English (above) and (b) Swedish (below). The [s] spectra are included for comparison.

Discussion

The difference traditionally described is that the English dental stops are more retracted than the Swedish, i.e. there is a difference in the place of articulation. Yet both of the subjects had alveolar stops. Paradoxically, the Swedish speaker had a lower locus, which traditionally would be ascribed to a more retracted place of articulation.

The most obvious difference seen on the X-ray films is that between laminal and apical occlusions, i.e. a difference of tongue shape rather than a difference between the points of aim of lingual gestures. The different vocal tract configurations associated with the two types of occlusion are illustrated at Fig. 6 while the consequent resonance conditions will be discussed below, where it will be argued that the observed loci and spectral differences are compatible with these two types.

There is finally the longer duration and greater overall intensity of the English subject's [t] burst. A necessary articulatory correlate of this is an open glottis to permit an adequate airflow, and a necessary consequence is the delay in voicing onset for the vowel. Glottal adjustment alone would account for a longer aspiration phase and delayed voice. But if the [s]-like quality of the burst is an essential component of the English [t], then the glottal adjustment must also be coordinated with a slow release of the occlusion to maintain a turbulent air stream through the tongue constriction and past the incisors.

Any cross-language comparison based on only two subjects is beset by the problem of how to generalize the results to embrace entire communities, especially when those results are novel and seemingly in contradiction to a long established and widely accepted view. The articulations X-rayed here are in accordance with the belief that English and NW Skåne dental stops have an alveolar place of articulation. Johansson's central Swedish subjects had purely dental occlusions (or more precisely denti-gingival). This is in accordance with the established belief that English and Central Swedish dental stops have different places of articulation. The novelty of the present results is that the dental stops of the two subjects are acoustically different, notwithstanding the same place of articulation. This suggests that the essential difference between English and Swedish dental stops must be something else and not the place of articulation. Gimson's advice to foreign learners of English (pp. 153, 154, 159) points in the same direction - he recommends practising affrication of English /t, d/ as [t^s, d^z]. This advice coincides with the difference observed above between the two subjects - the duration and spectral character of the burst (Fig. 8).

Are the observed acoustic differences related to the difference between apical and laminal occlusions? The traditional view has been that there is hardly any audible difference between the two types of articulation. But the contrasts reported by Ladefoged for certain West African languages indicate that there is a consistent difference that is large enough to be useful for phonemic oppositions. Stevens (1973) has recognized that the character of the turbulence of dental consonants is more dependent on the type of constriction (apical or laminal) than on the location of the constriction. For constrictions in the denti-alveolar region, turbulence excites mainly the front cavity resonances. But his calculations also indicate that a more gradual widening or tapering of the vocal tract behind the constriction results in greater coupling of the turbulence source near the constriction to the back cavity resonances. Consequently, he argues, for laminal articulations several resonances are excited over a range of frequencies above a certain critical frequency determined by the amount of tapering.

Chomsky and Halle believed that their contrast of constriction length would also cover the difference between hard and soft consonants in Slavic languages. They quoted evidence from Polish. Fant has published radiographs and area functions of Russian palatalized (soft) and non-palatalized (hard) consonants, including dental stops (1960: Figs. 2.6-8 and 2.6-9). The articulatory difference made by his subject is precisely that difference between the laminal and apical occlusions described above, and between Chomsky and Halle's long and short constrictions - for the soft dental stops the occlusion was laminal and long with the tongue body high against the hard palate, but for the hard dental stops the occlusion was apical and short with the tongue body depressed.

Stevens's argument is completely in accordance with the data presented by Fant for the Russian soft (laminal) voiceless dental stop relative to the hard (apical) voiceless dental stop (Fant's Fig. 2.6-10). The release burst of the soft laminal voiceless stop was characterized by low spectral levels below 3000 Hz, with a weak F2 at about 1900 Hz, and a high level in the region 6000 - 9000 Hz, while that of the hard apical voiceless dental stop was characterized by a relatively high spectral level below 500 Hz, a zero at 850 Hz, an F2 peak at 1600 Hz and a generally lower level at higher frequencies, declining after smaller peaks in the 3500 - 4500 Hz region.

The similarity between the two Russian [t] types and the [t] bursts of the present subjects (Fig. 8) is very striking. Like the Russian soft [t] burst, the English subject's [t] burst is relatively weak below about 4000 Hz but contains numerous high level peaks throughout the 4000 - 8500 Hz region. This corresponds to the greater excitation of the back cavity resonances when the vocal tract widens gradually behind the constriction, as pointed out by Stevens. Similarly, like the Russian hard [t] burst, the Swedish subject's [t] burst is relatively weak above 5000 Hz. They have peaks at similar frequencies - about 500, 1100, 1600, 3500, 4500 Hz. Unlike the Russian hard [t], the Swedish [t] did not have its maximum level at F2, and it had additional small peaks at 2100 and 2600 Hz. This type of spectrum corresponds to the excitation of the front cavity resonances only and the exclusion of back cavity resonances when the vocal tract widens abruptly behind the constriction, as pointed out by Stevens. The vowel formant transition loci of the English and Swedish [t] occlusions correspond to the F2 frequencies found by Fant in the Russian soft and hard [t] bursts, 1900 and 1600 Hz. The different spectral characters of the [t] bursts of the English and Swedish subjects, and the difference in locus frequency, are thus related to the apical and laminal manners of forming dental occlusions.

Stevens also notes a tendency for the longer laminal constriction to be released more slowly. In the case of voiceless stops the burst is prolonged and the voice onset for the following vowel delayed (he quotes data showing 20 ms or more following the laminal occlusion compared with a few ms after the apical occlusion). This was also observed by Ladefoged in the West African languages, where the laminal stops tended to be affricated. The same difference can also be seen on the spectrograms published by Fant for the Russian consonants (his Fig. A. 13-16) which show a much longer burst after the laminal soft [t], than after the apical hard [t]. The [t] burst of the English subject was nearly three times as long as the Swedish subject's. This difference is thus also in accordance with the laminal and apical types of occlusion.

Conclusions

One firm conclusion at least can be drawn from the present results - that the observed vowel formant transitions, the spectral characters of the bursts and the relative durations of the bursts can be related to the apical and laminal types of occlusion. Would it be presumptuous to assume that Swedish dental stops are generally apical and English dental stops generally laminal? This would contradict the widely held belief that dental stops are apical for most speakers of both languages, laminal for some in both languages.

If the widely held belief is correct, that dental stops in both languages are generally apical, then the difference between laminal and apical articulations observed in this pair of subjects will not be a typical difference between English and Swedish. Nor will the associated spectral differences be typical either and the spectra illustrated at Fig. 8 will then in part be individual speaker-dependent characteristics rathen than essential language-dependent features. What would this leave us with? Firstly, that the English [t] burst is [s]-like whereas the Swedish burst is not. This difference can be achieved with either type of articulation. Secondly, the English burst is longer and voice onset delayed for the next vowel. While this is a typical attribute of the laminal articulation, the latter is not necessary condition for affrication. Gimson's advice to learners, to make the burst [s]-like and affricated, favours both of these differences between English and Swedish. The longer burst requires the vocal folds to remain open for a longer period for English [t], the more intense [s]-like quality of the burst requires greater pressure behind the constriction, while the maintenance of the [s]-like quality thoroughout the whole of the longer burst requires the occlusion to be released more slowly.

How important is the often-quoted difference in place of articulation? For most Swedish dialects the dental stops are said to be truly dental while the English stops are alveolar. The fact that some Swedish dialects have alveolar stops without the "foreignness" of the English stops suggests that the difference in location is unimportant. Stevens has demonstrated that there are regions of the vocal tract where variations of constriction location yield minimal acoustical differences. The denti-alveolar part of the tract is one such region. There are other regions where differences of constriction location yield large acoustical differences, for example from the alveolar to the palatal. Consequently the acoustical difference between alveolar and purely dental consonants is very small, insufficient for a phonemic contrast. In the languages that have contrasting sets of dental stops with different locations, there is always the difference of constriction length (usually apical vs laminal) with the extra redundancy of affrication vs non-affrication. I am sure in my own mind that it is unwise for language teachers to emphasize the retracted English articulation. The danger is that this prompts the pupils to adopt something like the Swedish supradental (retroflexed) consonants. The first priority (eg for French learners) should be to learn aspiration. The next priority (eg for Swedish learners) to make the aspiration [s]-like and prolonged.

References

Bruce G (1970). Diphthongization in the Malmö dialect. Working Papers, Phonetics Laboratory, Lund University, 3:1-19

Catford J C (1968). The articulatory possibilities of man. Malmberg B (ed), Manual of phonetics, pp. 309-333. The Hague

Chomsky N and Halle M (1968). The sound pattern of English. New York

Fant C G M (1960). The acoustic theory of speech production. The Hague

Gimson A (1962). An introduction to the pronunciation of English. London

Jespersen O (1899). Fonetik, en systematisk fremstilling af laeren om sproglyd. Copenhagen

Johansson K (1967). Några perceptoriska försök med svenska dentaler och supradentaler. Unpublished licentiate thesis, Department of Linguistics, Lund

Jones D (1956). An outline of English phonetics. Cambridge. 8th edn

Ladefoged P (1964). A phonetic study of West African languages. Cambridge

Malmberg B (1967). Lärobok i fonetik. Lund

Nihalani P (1974). Lingual articulation of stops in Sindhi. Phonetica 30:197-212

Stevens K N (1973). Further theoretical and experimental bases for quantal places of articulation for consonants. MIT Research Laboratory of Electronics, QPR 108:247-252