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

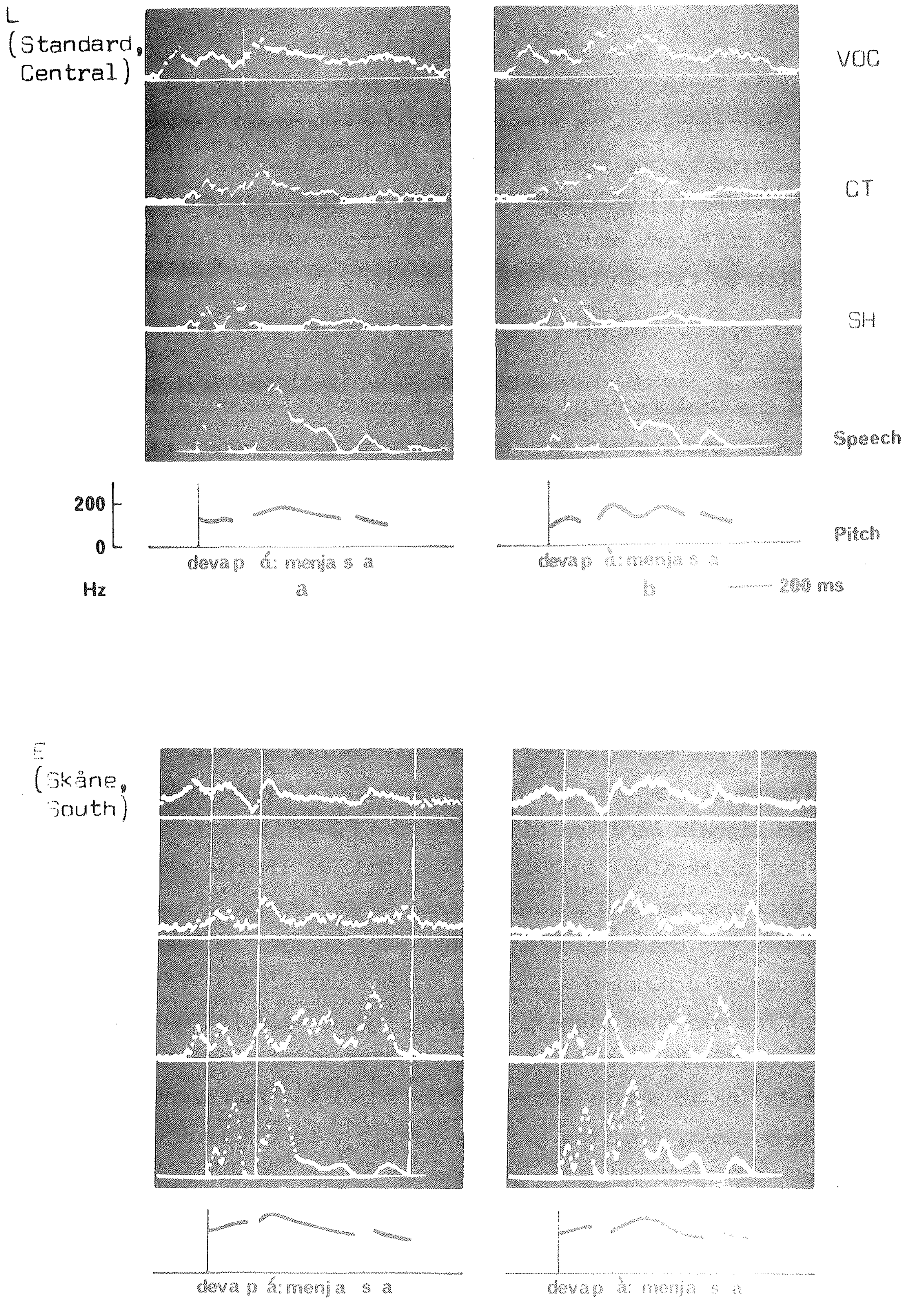


Figure 1. The averaged EMG signals of the vocalis, the crico-thyroid and the sternohyoid muscles, the speech signals and the pitch contour for the utterances a) de va ´pa:men ja sa b) de va `pa:men ja sa. Above Speaker L (Standard, Central) and below Speaker E (Skåne, South).

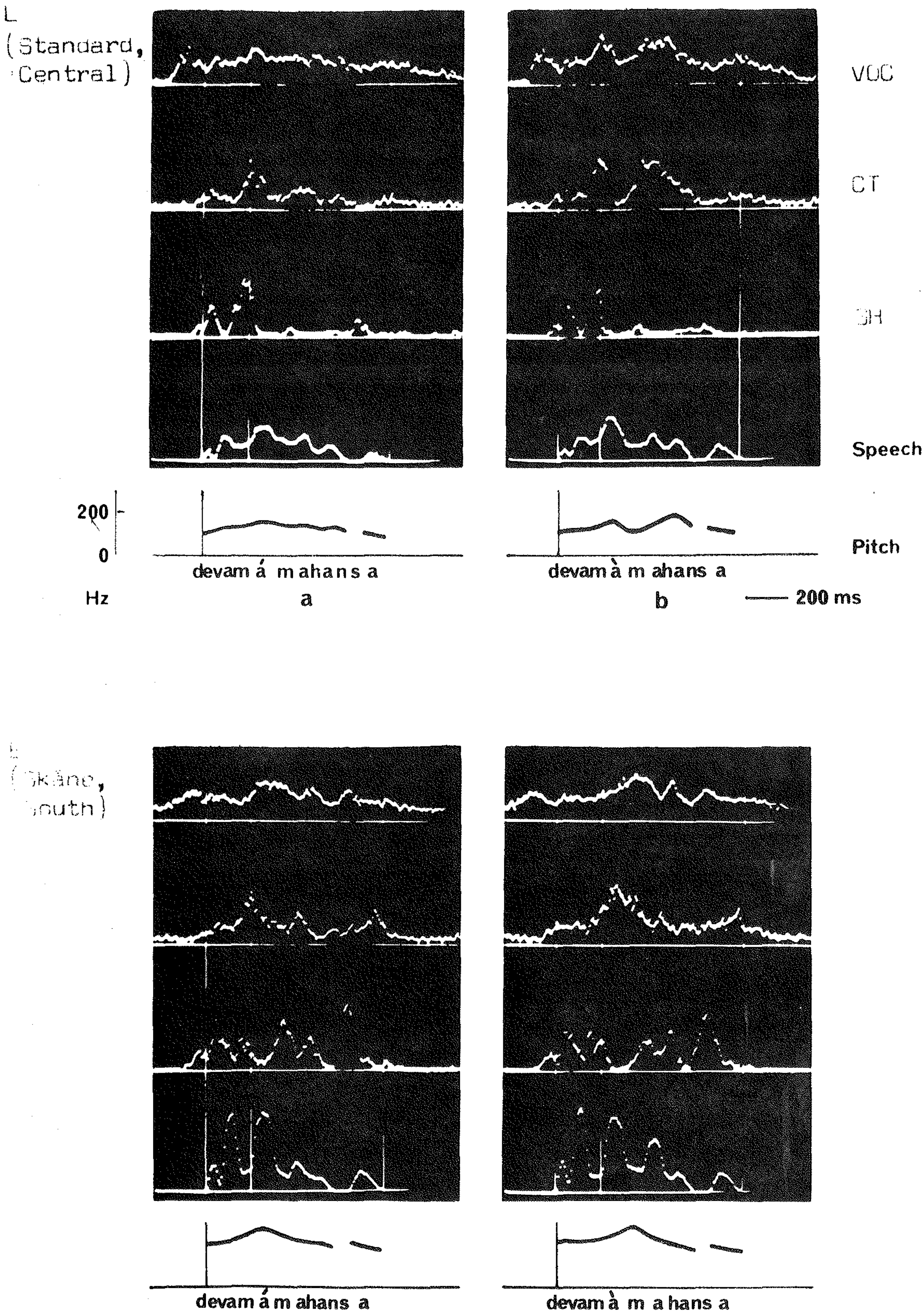


Figure 2. The same as in Figure 1 for the utterances
a) de va ´mamma han sa
b) de va `mamma han sa.

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.

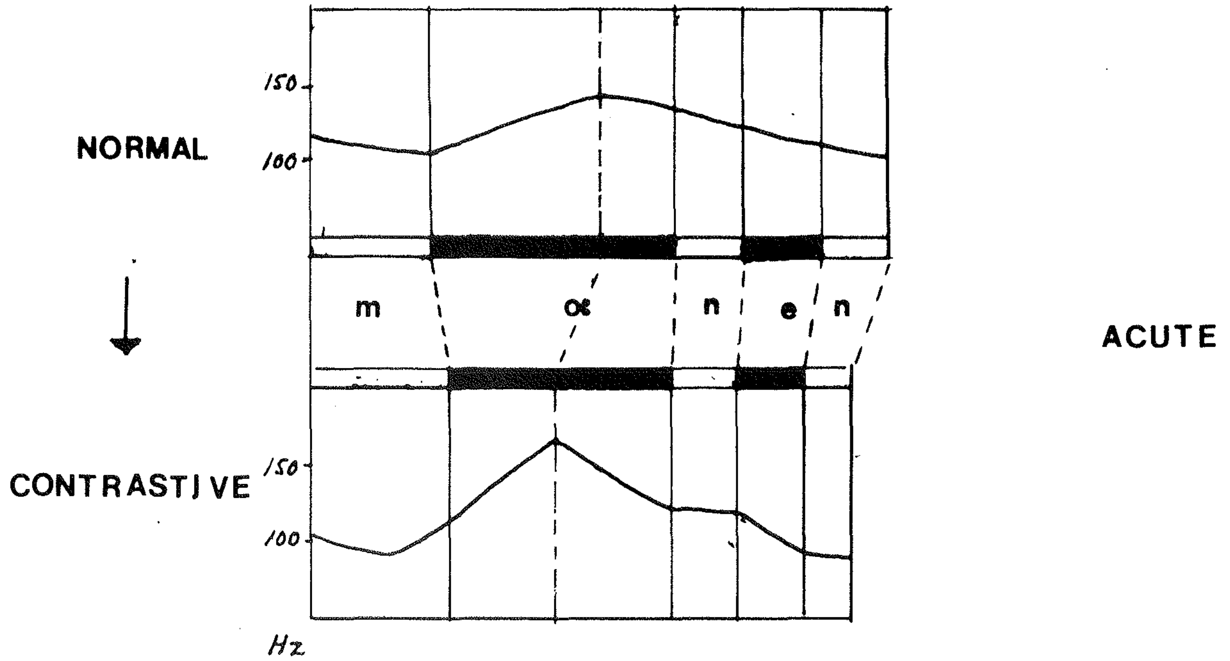
The fact that the second EMG peak of the grave accent (Speaker L) is higher than the first one can be related to the larger range of pitch rise after the transient pitch fall at the end of the first syllable.

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



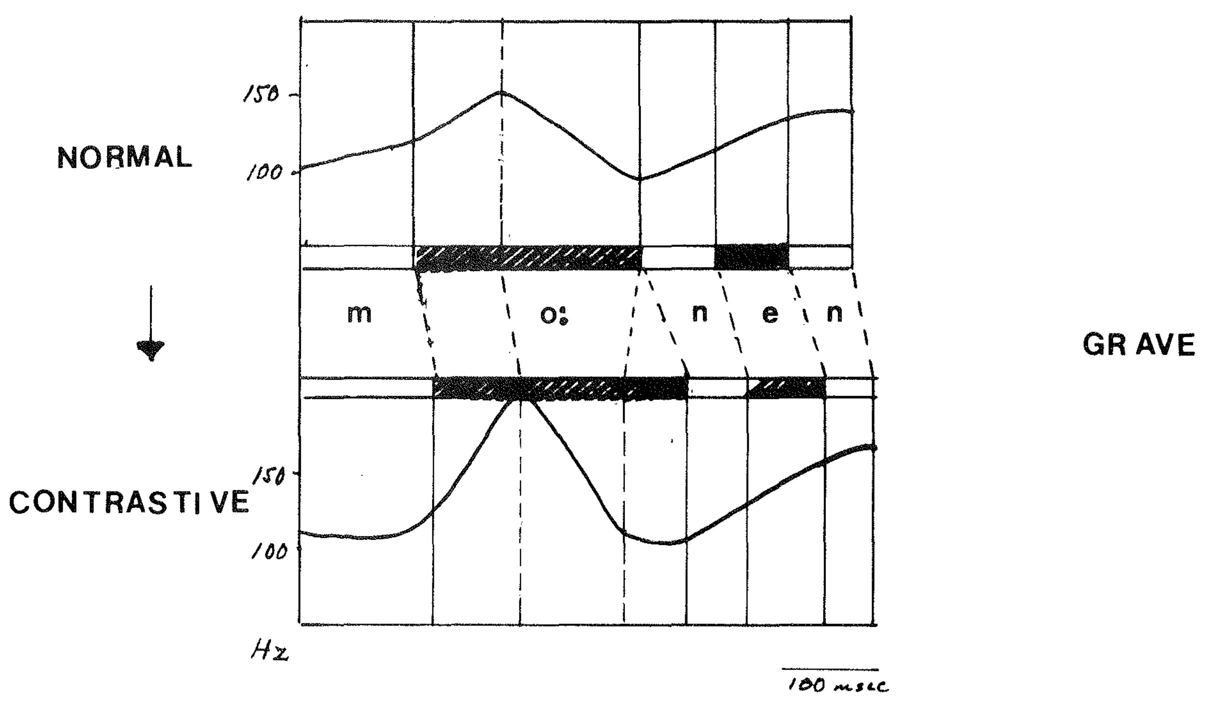


Figure 3. Accents under normal and contrastive stress. Averaged acoustic segments and pitch curves. Sequence /mo:nen/. Speaker L.

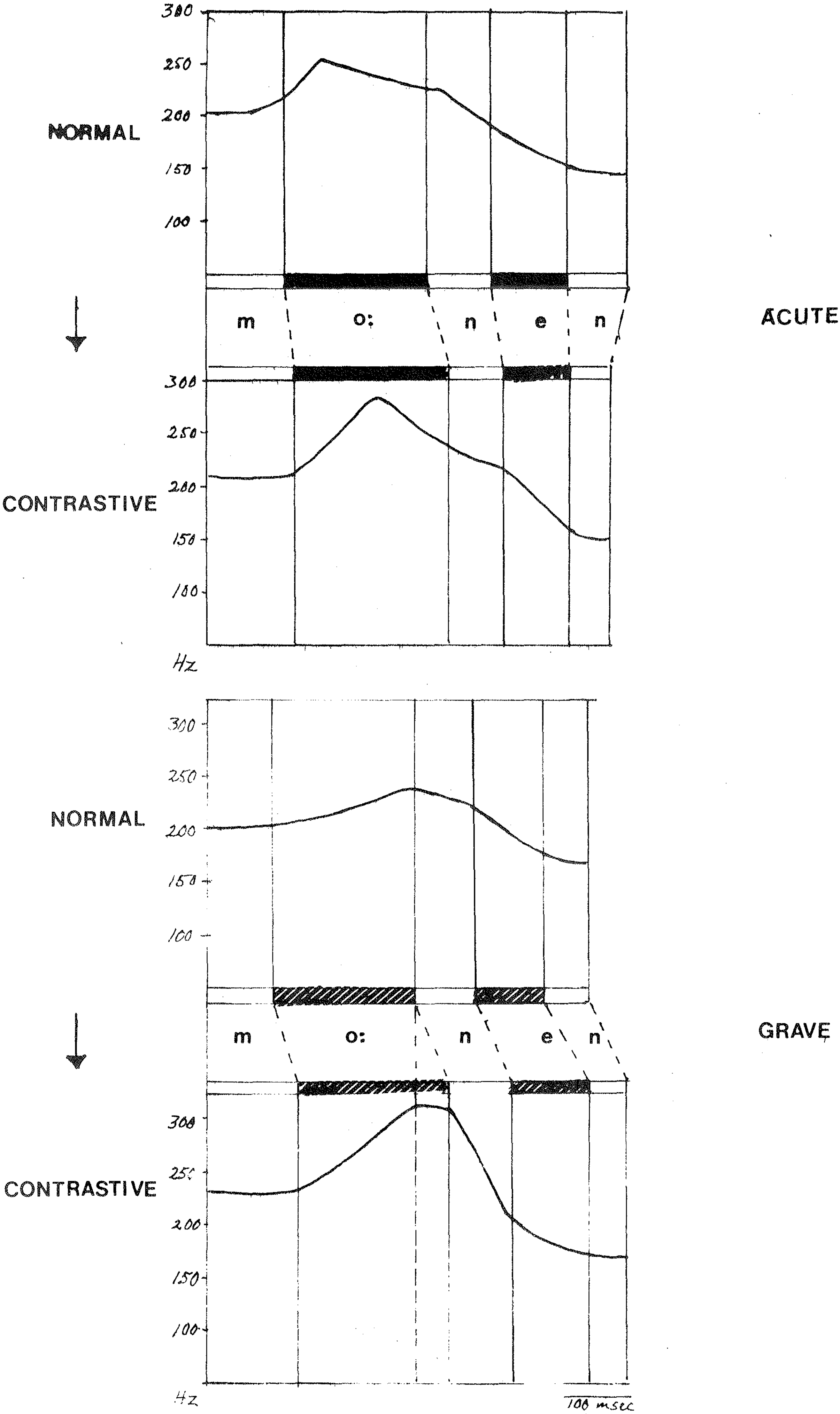


Figure 4. The same as in Figure 3. Speaker E.

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.

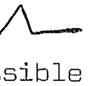
Common for both dialects is that contrastive stress modifies both syllables in the grave accent and only the first syllable in the acute one. The changes in the stressed syllable of the acute accent also have something in common in that the location of the pitch peak moves in the direction of the center 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-

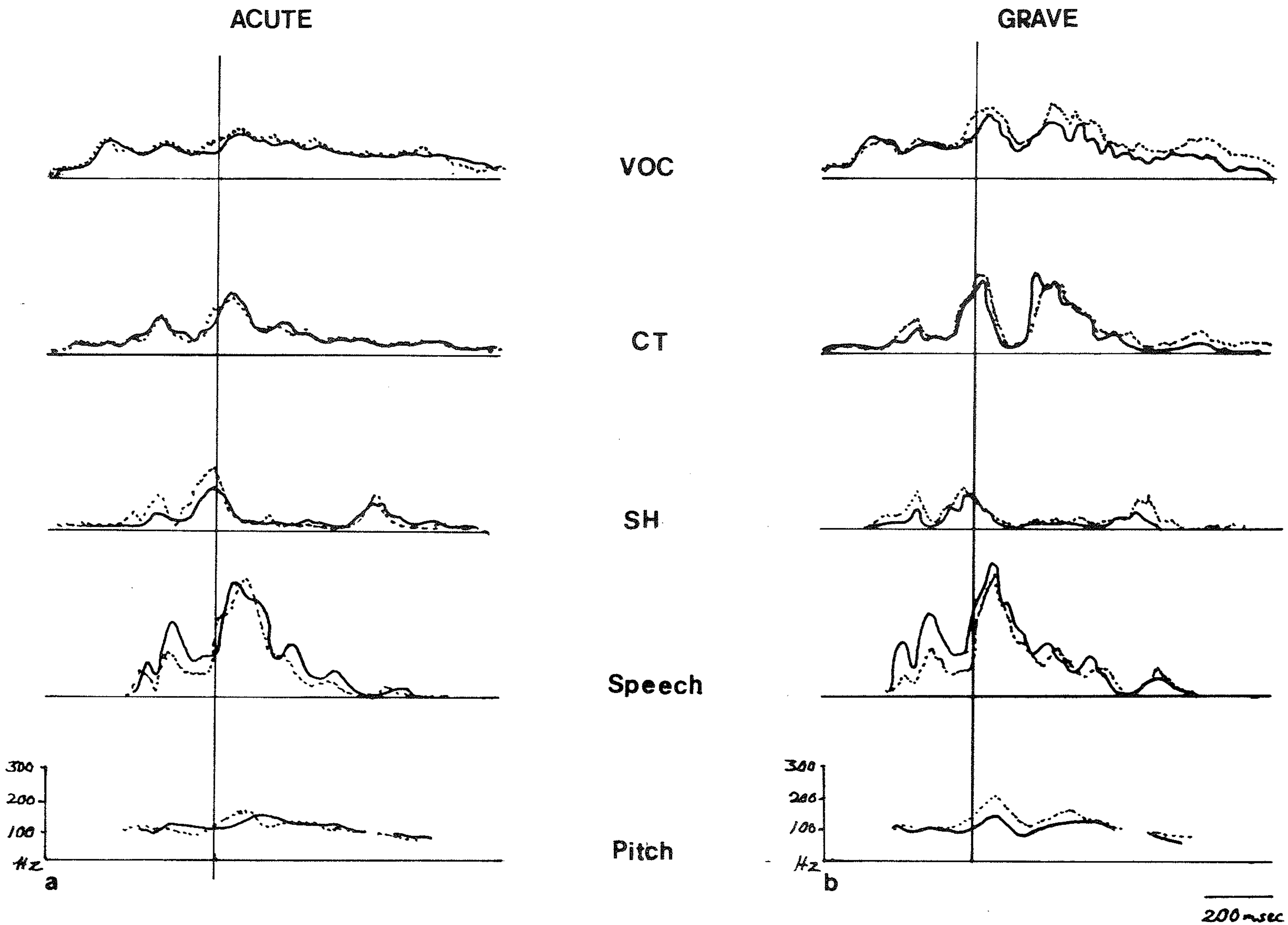


Figure 5. Accents under normal and contrastive stress. Averaged EMG signals, speech signals and pitch contours (as in Figure 1) for the utterances a) de va ´mo:nen han sa b) de va `mo:nen han sa. The solid curves refer to normal stress and the dotted ones to contrastive stress. The vertical line indicates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

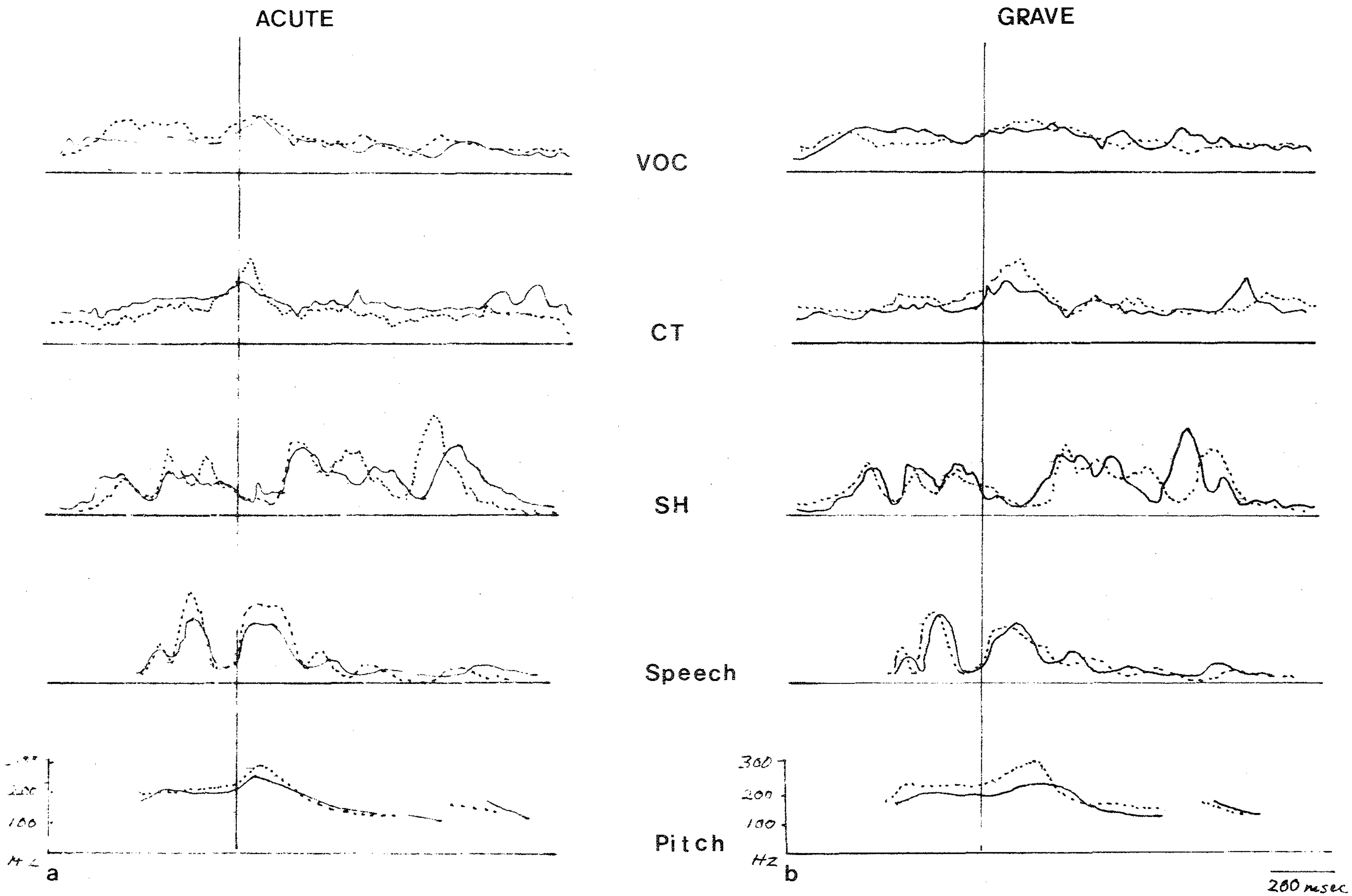


Figure 6. The same as in Figure 5. Speaker E.

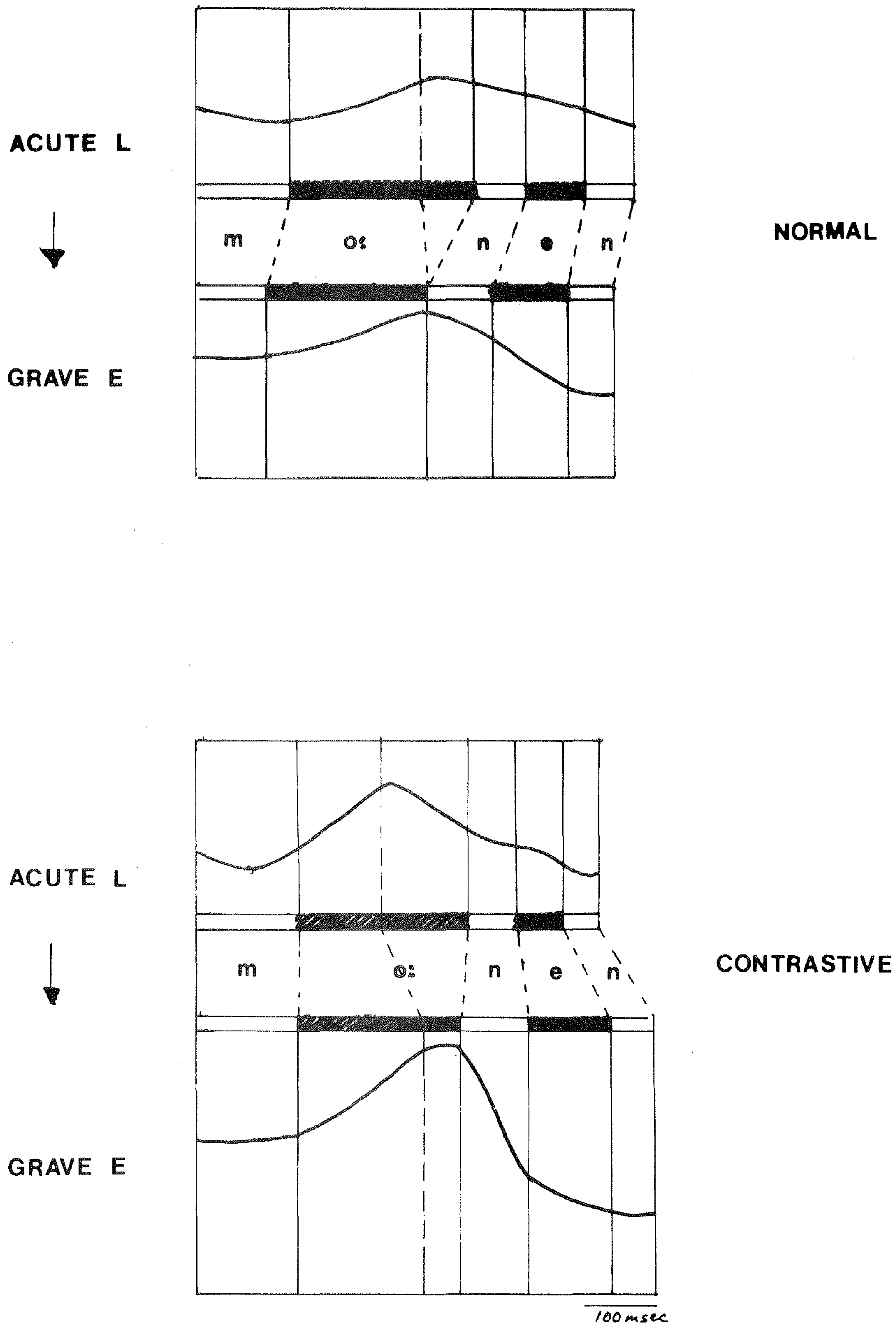


Figure 7. Dialectal variation. Averaged acoustic segments and pitch curves for L's acute and E's grave accent under two stress conditions.

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

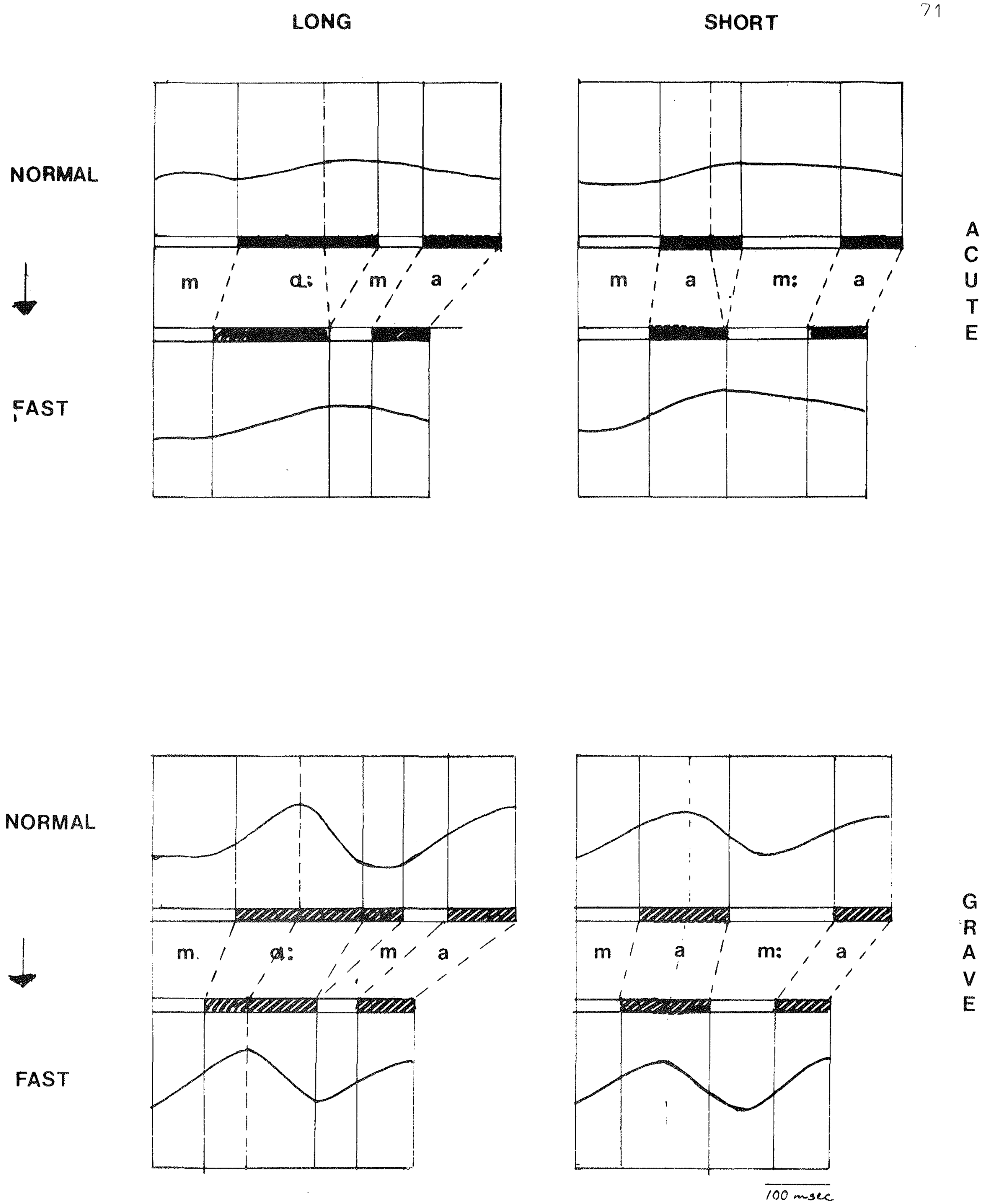


Figure 8. Accents, tempo and vowel length. Averaged acoustic segments and pitch curves for L's acute and grave accents in normal and fast renderings of the sequences /ma:ma/ and /mamma/.

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

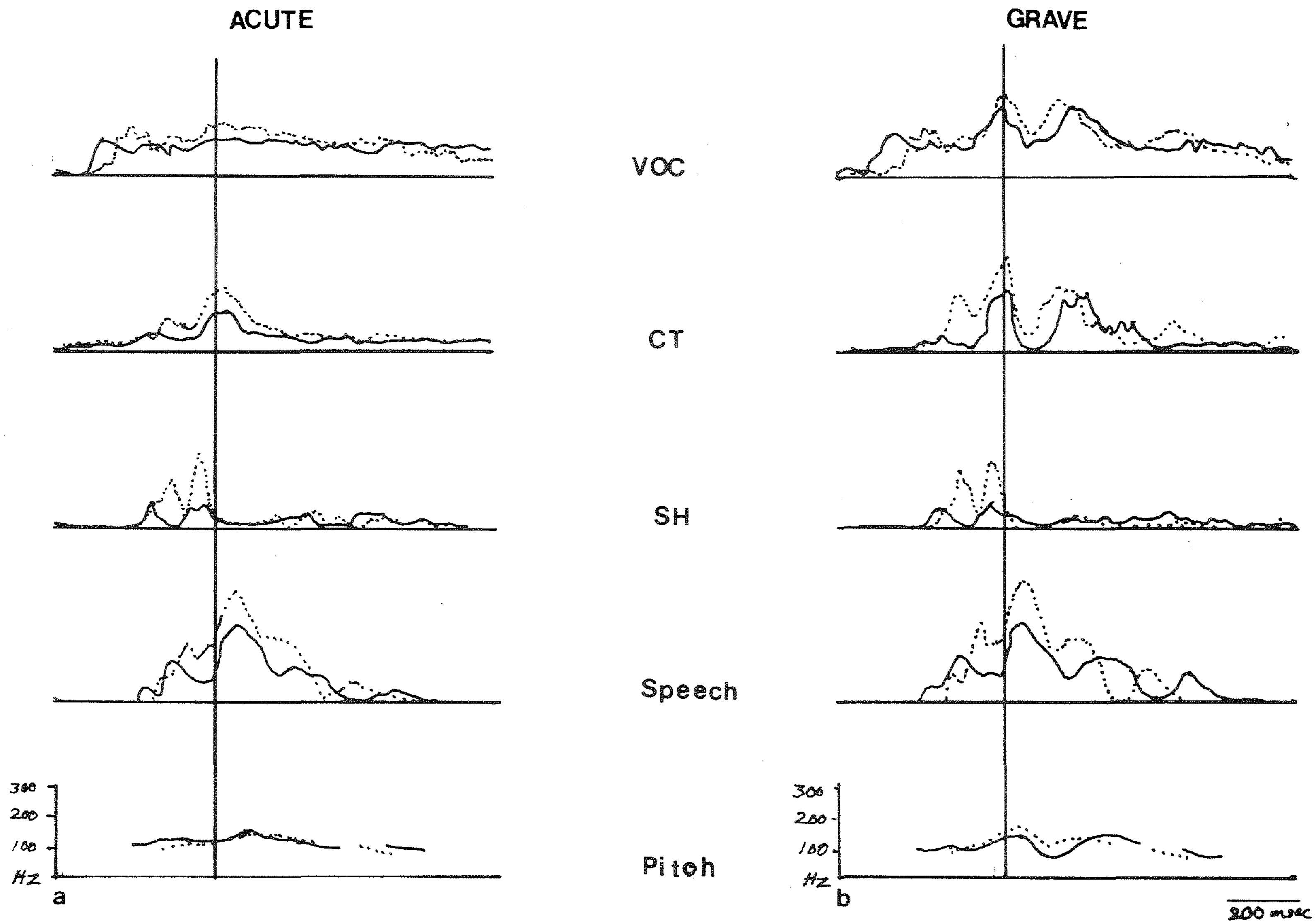


Figure 9. Accents and tempo. Averaged EMG signals, speech signals and pitch contours (as in Figure 1) for the utterances a) de va ´ma:ma ja sa b) de va `ma:ma ja sa. The solid curves refer to normal rate and the dotted ones to fast rate. The vertical line indicates the release of /m/. Speaker L.

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.

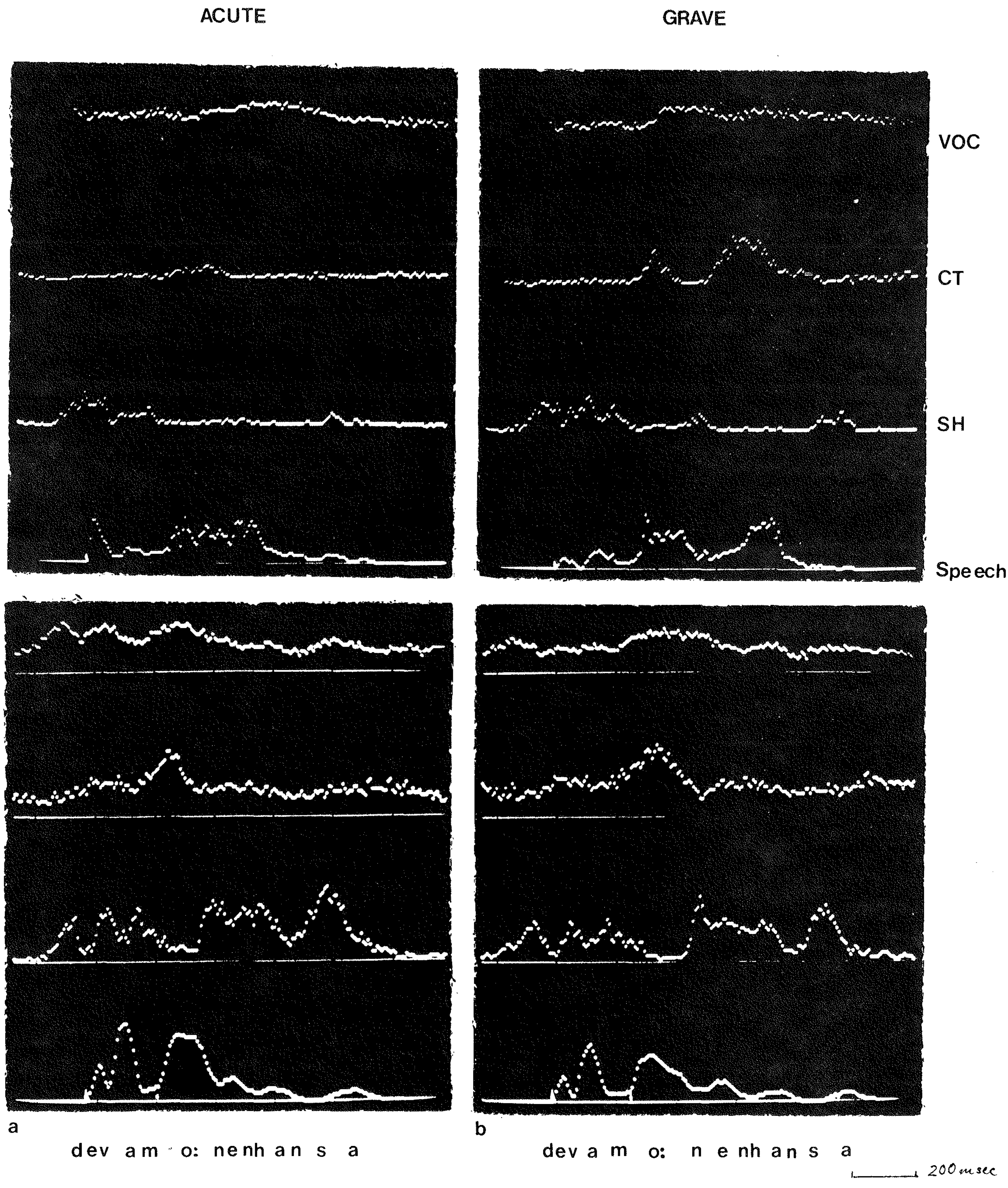


Figure 10. Accents in whispered and phonated speech. Averaged EMG and speech signals of the utterances a) de va ´mo:nen han sa b) de va `mo:nen han sa. Above in whisper and below under contrastive stress. Speaker E.

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 general 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