PHONETICS LABORATORY
DEPARTMENT OF GENERAL LINGUISTICS
LUND UNIVERSITY

WORKING PAPERS

12 · 1975

Previous numbers and contents:

WP 10 · 1975	ROBERT BANNERT ANNE-CHRISTINE BREDVAD-JENSEN	Temporal organization of Swedish tonal accents: The effect of vowel duration
	GÖSTA BRUCE	Stockholm accents in focus
	EVA GÅRDING OSAMU FUJIMURA HAJIME HIROSE ZYUN'ICI SIMADA	Laryngeal control of Swedish word accents
	KURT JOHANSSON	Perceptual characteristics of vowels
	ANDERS LÖFQVIST	Some phonetic correlates of emphatic stress in Swedish
	BERTIL MALMBERG	Niveaux, choix et systemes approximatifs dans le langage
	KERSTIN NAUCLER	Some thoughts on reading and writing
	THORE PETTERSSON	In favour of the archiphoneme
	EVA WIGFORSS	Foreign accent and bilingualism
	SIDNEY WOOD	What is the difference between English and Swedish dental stops?
WP 11 · 1975	MONA LINDAU	Vowel features
		A phonetic explanation to reduced vowel harmony systems
	SIDNEY WOOD	The weakness of the tongue-arching model of vowel articulation
		Tense and lax vowels - degree of constriction or pharyngeal volume?

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

12 · 1975

Working Papers 12 · 1975 Lund University

Phonetics Laboratory General Linguistics

THE SIGNIFICANCE OF VOWEL FEATURES IN THE PERCEPTION OF COMPLEMENTARY LENGTH IN CENTRAL BAVARIAN

Robert Bannert

1. Introduction

Minimally contrasting pairs such as [fe:da] - [fet:a] (feather - male cousin) and [vi:sn] - [vis:n] (meadow - to know) present a special problem for the phonological analysis of the segmental phonemic units (segmental phonemes) of Central Bavarian. When applying the commutation test to them, it is not possible, for example, to exchange only the first (and stressed) vowel in these words while leaving all other segments unchanged. Only two kinds of stressed vowel-consonant sequences are permitted in Central Bavarian word structure, namely

(1) the long vowel followed by a short, weak or lenis consonant, and

(2) the short vowel followed by a long, strong or fortis consonant.

The interrelationship between these features of the vowel and the consonant is demostrated by the following:

	kind of sequence			
feature	V:	C	V	C:
length	+	–	–	+
strength (fortis)		–		+

As can be seen the consonant features of length and strength always have the same specification, that is they occur together. The consonant is either short and lenis or it is long and fortis. Vowel length and the consonant features, however, always have the opposite specification.

There is in fact a third phonetic feature of the consonant, namely voicing which has, however, not been recognized or discussed in the literature (cf Bannert 1972, 1974).

As these features (length of the vowel and length and strength of the consonant) are constrained with regard to each other, a complementary distribution of the above mentioned features of both segments exists. It should be obvious that the vowel-consonant sequences are characterized by a double redundancy of phonetic features, for only one of the three features needs to be known for the others to be predicted.

The phonological difficulty lies in the choice of the distinctive feature for this minimal contrast observed in the sequences: Which of the three existing phonetic features of the sequences is the distinctive one, that is to say which feature cannot be predicted and thus has to be assumed as being specified in the lexical representation of the words?

From a logical point of view this decision need not be difficult. But in order to satisfy the claim of explanatory adequacy of linguistic description the choice is crucial. In accord with predictive phonetics (cf Lindblom 1971) the linguistic description has to be made from phonetic evidence. As the vowel-consonant sequences of Central Bavarian are concerned, the description also has to include the phonemic vowel and consonant units of the dialect.

It is not very surprising to find that the descriptions in the relevant literature, which are based on purely auditive analysis ("Ohrenphonetik"), offer different solutions to this problem. This fact, however, is not unique for Central Bavarian. In a similar way the vowel-consonant sequences of the Nordic languages of Central Swedish, Norwegian, and Icelandic, which are characterized by complementary length, have not yet received a generally accepted phonological description, although acoustic and perceptual data are available in this case. There are of course several suggestions for treating this problem. For a presentation and discussion see Elert (1964).

Each phonological analysis of the vowel-consonant sequences in Central Bavarian explicitly rejects the choice of the length feature of the vowel as the distinctive feature. The reason given is that the length of the vowel is conditioned by the segmental context: It can of course be predicted as a consequence of one of the features of the consonant or the contact feature.

It is really quite obvious that this solution is not the only possible one, not only from a logical point of view, but even more so from a phonetic-phonological one. For the features of the consonant can be predicted in exactly the same way if the vowel is specified as to its length.

From the point of view of a modern phonetician the literature lacks basic and phonetically adequate descriptions, especially measured values, of the manifestation of the Central Bavarian vowel-consonant sequences. This is also true of the sounds and sound structures of the dialect as such. Until a language is described precisely (empirically, by using quantitative measures) on all three levels of the speech communication process (production, acoustics, and perception), it is not possible to achieve a satisfactory phonological description which also meets the claim of explanatory adequacy. In these investigations the significance of perception in the speech communication process has to come to the fore. It is imperative to investigate which feature of the sound sequences, originating in the production and being transmitted to the ear of the listener through the air, is used by the listener in order to perceive distinctive sound contrasts (differentiating between words).

In view of the lack of the necessary basic phonetic data on Central Bavarian I started an acoustic investigation of the sound structures in the dialect. The vowel-consonant sequences are in the focus of attention (see references).

Perceptual investigations were carried out parallel to these studies. I wanted to find out which phonetic features of the vowel or the consonant or both are needed by listeners in order to identify the sequences as the one kind or the other.

This papaer is a report on an investigation which was by way of an intermediate study. After having evaluated the results of the pilot study (Bannert 1975) it was shown to be necessary to carry out the present investigation due to the following reasons:

(i) <u>Redundancy of the phonetic features</u>

Even if only one feature in one segment is altered within the whole vowel-consonant sequence the responses of the listeners to this manipulation cannot be interpreted unambiguously because the other two features are retained in the sequence and may affect the identification. Therefore most of the redundancies of the vowel-consonant sequences were eliminated so that the effect of only one feature on perception could be revealed. For this purpose the consonant and the rest of the words were cut off, thus getting rid of the two consonant features, duration and strength (fortis-lenis).

(ii) The time dimension (quantity)

From the phonetic point of view it would seem to be likely that, if significant and stable differences between the durations of vowel and consonant are observed in the acoustic signal, the time dimension (that is, segment duration, temporal extension of an acoustic spectrum) may signal the kind of vowel-consonant sequence for the listeners or, at least, it may contribute to their correct identification.

Thus perception tests in different quantity languages have shown that listeners identify (categorize) the phonological categories (classes) long vs short according to their segment duration. A specific reduction of the duration of long vowels led the listeners to hear the corresponding short vowel (Southern Swedish: Hadding-Koch and Abramson 1964, Central Swedish: Jonasson and McAllister 1972, Standard German: Heike 1969, Finnish: Lehtonen 1970, Estonian: Lehiste 1971, Thai: Abramson 1962, different languages: Fliflet 1961).

Another kind of experiment suggests that listeners have some kind of "perceptual knowledge". Native speakers of quantity languages adjusted the segment duration of the "best" long and short vowels respectively by turning a knob on a speech synthesizer. These perceptually established vowel durations corresponded well with the measured vowel durations in the acoustic signal (Nooteboom 1973, Petersen 1974).

As far as Central Bavarian is concerned then, it is to be expected that listeners will use the time dimension (segment durations) - perhaps in addition to other features - to categorize the received acoustic stimulus either as being one or the other of the vowel-consonant sequences types. This is to be expected from the temporal regularities (time patterns) observed on the acoustic level (cf Bannert 1972, 1973a) and by way of extrapolation from other languages as mentioned above.

In order to test this hypothesis the segment durations of long and corresponding short vowels were varied. Vowel portions of minimal pairs differing in duration were presented to listeners while the postvocalic consonant containing the features of length and strength was eliminated.

(iii) <u>Contact features</u>

The contrast between the two kinds of vowel-consonant sequences is attributed to different features of contact between the vowel and the following consonant in many of the phonological analysis in the literature ("stark" and "schach geschnittener Akzent": Pfalz 1913 following Sievers 1881, "Abglitt": Gladiator 1971 following Pilch 1964).

These phenomena established on the auditive level certainly exist in the real world, but it should be remembered that their definition in the relevant literature is not very precise (cf for example Sievers 1881). It is definitely not based on acoustic or physiological data.

However, the correlates of these contact features are postulated. They are given as being expiratory pressure, articulatory energy, absorption of the flow of air from the lungs by the supraglottal cavities on the physiological level, and on the acoustic level as intensity. Noone has yet succeeded, however, in experimentally proving that the correlates of the contact feature are independent of other features such as time already utilized in speech production.

One aim of this investigation was to find out if any acoustic correlates whatsoever of the contact features, included in the vowel, are used by listeners in perception.

2. Hypothesies

Even if the features of the postvocalic consonant are eliminated, at least two features remain in the vowel: Besides the feature of duration which is the manifestation of the phonological dimension of quantity, there is also the phonetic feature of vowel quality (spectral pattern), which from a phonological point of view is the same e.g. in the phoneme /e/ of /fe:da/ - /fet:a/ but which certainly is not the same in the manifestation (e.g. in terms of formant frequencies) of these words.

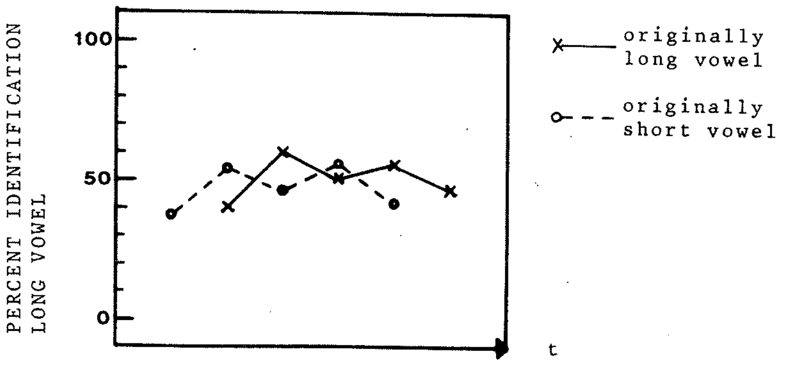
Languages with phonological long and short vowels (quantity languages) very often show some differences in formant structure between both categories of vowel length (cf Fant et al. 1969 for Swedish, Fischer-Jørgensen 1972 for Danish, Lehiste 1970 for Czech and Serbo-Croatian).

Although the qualitative difference between a long and the corresponding short vowel is very small on the auditive level (for comparison with other languages acoustic data on formant frequencies will appear in Bannert, forthcoming), it has to be assumed that this qualitative difference exists and that it may be utilized as a cue for perception by listeners (at least in a typical test situation).

Finally, the postulated acoustic correlate of intensity as the manifestation of the contact feature in the vowel has to be considered as well. According to the work of Fischer-Jørgensen and Jørgensen (1969) which, it is true, did not contain Central Bavarian material, it seems unlikely that such an acoustic correlate in the vowel exists. Indeed they could not find any evidence for such a correlate but concluded that the relevant acoustic correlate of the auditive contact phenomenon ought to be the duration of the vowel.

Therefore, as to the perception of the rather redundancy-free vowel fragments, the following general expectations can be set up. The hypothetical identification curves given as the response /V:C/ to each stimulus are shown as a function of the vowel duration (abscissa):

(i) <u>Redundancy</u>



Listeners cannot identify the VC-sequences (the words) without hearing the postvocalic consonant segment as well. In order to perceive correctly it is necessary for the listeners to have access to the whole VC-sequence with all the reduncancy of the features present. As the consonant is missing, the listeners will just guess when performing the test on the vowel portions only.

It is obvious that listeners, when hearing and processing (decoding) mutually completed VC-sequences, must have access to many redundancies being signalled by both segments of the sequences. Although the domain of the temporal contrast in Central Bavarian seems to be the whole VC-sequence on the acoustic level at least, and although perception has been shown not to be a segment-by-segment processing (Kozhevnikov and Chistovich 1965, Liberman et al. 1969), one could be inclined to suppose that listeners need not hear the whole sequence in order to identify the correct word.

This assumption, however, seems to be contrary to the view expressed by Lehiste (1970, 35-36):

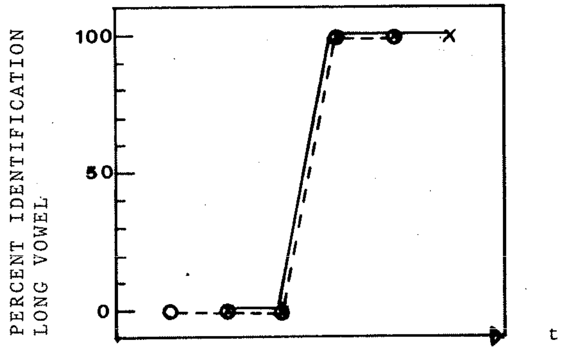
> "... suprasegmental features (including time, RB) can only be identified by comparison of items in sequence, and thus differ in a very essential way from features that may be identified by inspection of a segment (Jakobson, Fant, and Halle 1952; Lehiste 1967a)."

From this claim it follows that listeners would not be able to identify words contrasting in segment duration alone (that is, to identify e.g. the vowel categories long vs short), if they do not hear the following consonant or the whole word as well.

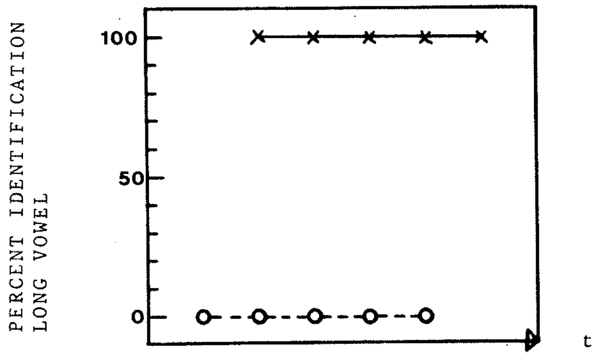
(ii) Segment features

Listeners are able to identify the VC-sequences from hearing only portions of the vowel segment with the following consonant segment missing. They do not guess, they use one or the other of the two phonetic features (dimensions) present in the vowel segment:

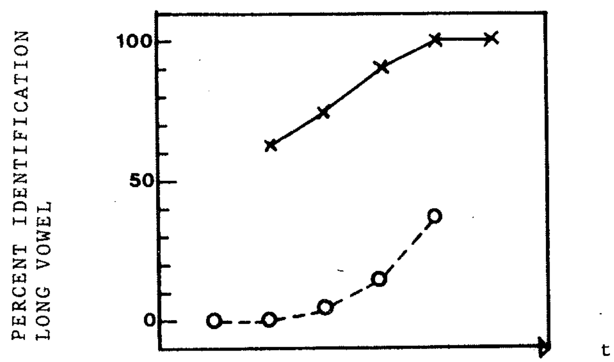
(a) Listeners identify the words exclusively from vowel duration.



(b) Vowel quality is the only cue used by the listeners when identifying the VC-sequences. Duration, although present, does not influence their perception.

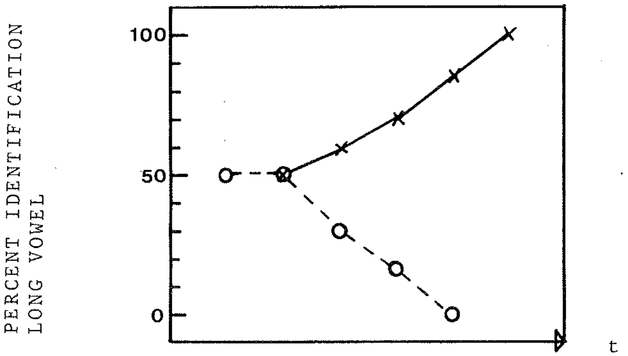


(c) Both features may of course counterbalance each other or interfere with each other, causing at least parts of the curves to deviate from their expected course towards the other alternative as indicated in the following figure:



(iii) "Contact"-features

Listeners can identify the VC-sequences from hearing only the vowel segment. They rely now on the manifestation of the postulated "contact"-feature manifested acoustically in the vowel. It should be present especially towards the end of the vowel segment, that is towards the segment boundary between the vowel and the following consonant. In this case the listeners, hearing the initial portions of the originally long and short vowels and categorizing them randomly, will be more likely to recognize the originally long vowel when longer portions of that vowel are presented to them. In accord to this, they should also identify the short vowel increasingly accurately with the larger portions they hear of this vowel segment.



3. Material

For this test meaningful, bisyllabic words were chosen from the material on the acoustic description of the vowel system of Central Bavarian (Bannert forthcoming). They were produced in isolation by informant A in Bannert (1972, 1973, 1974, and forthcoming), who also is one of the naive listeners (A 1) of the tests. The test words constitute three minimal pairs contrasting the two kinds of VC-sequences. After careful listening and inspection of the spectrograms the pairs of words with vowel qualities most like each other were chosen from the twelve renderings of each test word. The six test words are listed here:

	V:C		VC:		
/e/	[ke:gal]	-	[kek:al]	Gegal - Gekal	(name of a hill - roast chicken)
	[fe:da]	-	[fet:a]	Feda - Feta	(feather - male cousin)
/i/	[vi:sn]	-	[vis:n]	Wiesn - wissn	(meadow - to know)

The initial stop in the first word pair is voiceless and non-aspirated.

In addition to these pairs contrasting long and short vowels of the same phonemic quality category, a fourth minimal pair Nasn - nassn [nɔ:sn] - [nɑs:n] (nose - wet, inflected form) was included in the material for the purpose of checking the identification of vowel quality by the listeners. As well as differing in length the two vowels belong to two different phonemic vowel qualities, namely /ɔ/ and /ɑ/.

The target forman frequencies of F_1, F_2, and F_3 and the vowel durations of the eight test vowels are given in the following table:

Table 1. Phonetic values of the test vowels and the stimuli.

	Target formant frequencies (Hz)			Vowel duration (msec)	Duration of vowel stimuli (msec)				
Original	F_1	F_2	F_3		1	2	3	4	5
Gegal	300	2280	2840	180	80	100	120	140	160
Gekal	340	2280	2860	140	50	70	90	110	130
Feda	320	2250	2780	140	60	80	100	120	140
Feta	360	2200	2680	100	40	60	80	100	120
Wiesn	240	2260	2720	150	70	90	110	130	150
wissn	280	2100	2640	110	50	70	90	110	
Nasn	430	810	-	210	60	90	120	140	170
nassn	720	1200	2800	130	30	50	70	90	110

Broad-band and narrow-band spectrograms of the test words are shown in figure 1 a-d. It is to be noticed that the vowel quality /e/ appears in different segmental contexts: a symmetrical context of velar stops in Gegal vs Gekal, and an asymmetrical one in Feda vs Feta, where the vowel is surrounded by an initial labiodental voiceless fricative and a postvocalic dental stop. Due to this different consonantal context, the course of the formants through the vowel segments, especially F_2 and F_3, of the pair Feda vs Feta are totally different from the pair Gegal vs Gekal. The difference between the target formant frequencies of F_2 and F_3 in the long and the corresponding short vowels is greater in the former pair.

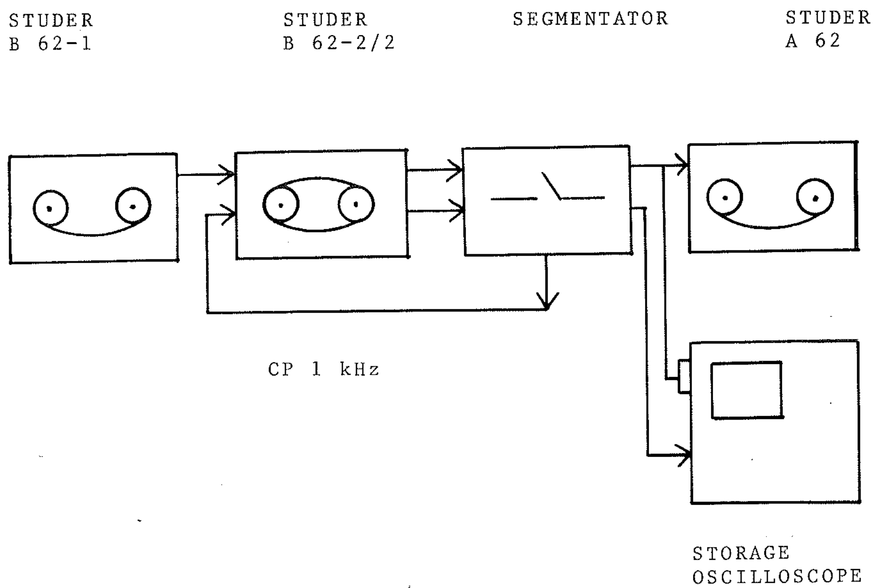
4. Preparation of the test

According to the test strategy outlined above the stimuli consist of the initial consonant followed by portions of the stressed vowel differing in duration, while the postvocalic consonant and the rest of the word are eliminated. The stimuli are of the structure /CV/. Starting from the pre-recorded natural test words, the test stimuli were prepared in the following way. Each vowel segment was divided into five portions using the electronic gate (segmentator) of the Phonetics Laboratory. The location of the cuts in the vowel segments and the size of the interval between them were determined perceptually starting from the end of the vowel and proceeding toward its beginning. After having listened to several different divisions of the vowel segments, including different interval durations, I found the steps of 20 msec used in this test the most suitable ones.

As a consequence of the applied method of cutting backward (from the vowel-consonant boundary) and the different durations of the originally long and short vowels, the initial parts of the vowel segments (stimuli nos. 1 of each word) had different durations. The initial portions of the long vowels had larger durations than those of the corresponding short ones. But this difference is of no importance for the test as the main purpose of this investigations was to vary segment durations as such.

The long vowel of Nasn (original duration 210 msec) is not included in its entity in the test because I considered it improbable that the listeners would hear anything else than just this word. The duration of the longest stimulus of Nasn is 170 msec, about the same as for Gegal.

The experimental set-up for preparing the stimuli is shown in the following diagram:



The beginning and the end of the signal ("window") which the segmentator is to cut out can be adjusted with an accuracy of 1 msec. The slope at the beginning and the end of the triggered signal may be varied stepwise from 0.47 to 100 msec independently from each other. The segmentated signal can be displayed on a storage oscilloscope (TEKTRONIX 5103 N).

The original word was copied from a STUDER tape recorder B62-2/2, operation speed 15 ips, together with a 1.000 Hz pilot signal (CP). Then the end of each vowel portion to be cut out was adjusted by means of magnifying the time scale on the storage oscilloscope, putting the triggering point in a zero crossing of the curve. The decay time was set to 10 msec resulting in stimuli, the ends of which were as short and, at the same time, as smooth as possible.

Each stimulus was then recorded twice on a STUDER A62, operation speed 15 ips.

Seven words yielded five vowel portions each, the vowel of wissn was divided into four portions only. Thus the present test consisted of 39 CV-stimuli. They were arranged in random order in five different series. Each stimulus was presented twice in each series with an interval of about three seconds. There was a pause of about four seconds between different stimuli.

The five stimuli, originating from each of the eight test words, are indicated in figure 1. Their durations are given in table 1.

5. Performance of the test

The test was presented from a NAGRA III tape recorder, operation speed 7.5 ips, via SENNHEISER HD 110 ear phones. Like the pilot study it was given in the homes of the listeners. The test series were preceded by a presentation of the longest stimulus (no. 5) of each of the eight test words. The purpose of this arrangement was to acquaint the listeners with their task. The listeners were told to imagine the following situation: A friend of yours is just going to utter one of the test words when he/she is sudenly interrupted and only manages to pronounce the very beginning (that is the initial consonant and the first vowel) of the words. The listeners were, then, asked when hearing only the initial fragments of the test words, eg.g [ke...], to identify the word as the one or other of a minimal pair. They had to underline the identified word on the answer sheets where the pairs were written. In cases of indecision they had to guess (decide on one word: forced choice). A short break was made after each series.

Nine listeners, the same as in Bannert (1972), participated in the test.

I have known all the listeners for many years. Except for C 1 they were born in the Central Bavarian area. But all of the listeners grew up there and are still living there. They do not talk Standard German except, perhaps, C 1. They are linguistically naive.

As all of the listeners had taken the pilot study earlier they were well acquainted with the test procedure.

This test was presented to the listeners twice, as was the pilot study, in order

(1) to enlarge the number of responses and thus the validity of the results and
(2) to check the reliability of the listeners' responses.

The time interval between the runs was three months. Each listener judged each of the forty stimuli ten times (five times in each run), so each stimulus received ninety responses (45 responses in each run); thus there were 3.600 responses in the whole test.

The differences between each listener's responses in the 1st and the 2nd run enable us to estimate the degree of certainty with which the listeners identified the stimuli. Before the results of the identification test are given and discussed, som remarks on these differences and thus the reliability of the identification scores are therefore necessary.

6. Reliability of identification

The reliability of identification or the certainty of judgment of the listeners can be viewed from three different aspects:

(a) With what degree of consistency did each listener judge each stimulus? If he responded differently in the 2nd run, to what degree did he deviate? How large is each listener's total deviation in relation to all the stimuli?

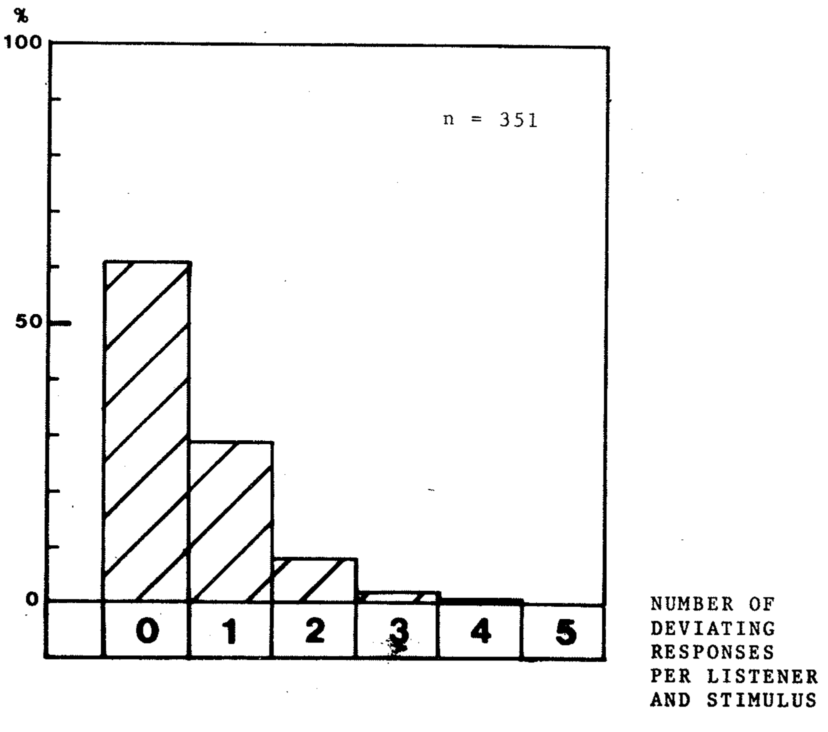
(b) How difficult was the identification of each stimulus by the whole group? Is each vowel identified to the same degree by each listener?

(c) To what degree did the listening group as a whole deviate in judging each stimulus?

These questions will be answered in the following.

(a) The listeners

The distribution of consistent and different responses between the two runs for the total number of possible cases of deviation for each listener and each stimulus (amounting to 351) is shown in the following diagram:



Out of these 351 cases 214 (61.0 %) showed no difference. This fact should be compared with the 59 % non-deviation found in the pilot study (Bannert 1975), which suggests that the listeners did equally well in the present test. In 101 cases out of the deviating 137 cases the identification of the test stimuli differ in 1 response only, which can be considered as being the result of non-linguistic factors such as insufficient concentration or fatigue. There is no case of a difference of 5 responses which would be the maximal difference. It can be considered as implying that the listener, when judging the stimulus the 2nd time, has changed his strategy, now picking a different cue from that in the 1st run when the difference in responses of a listener to a given stimulus between the runs is larger than 3 responses (more than half of the possible differences). There are only 8 cases (2.3 %) of such and apparent change of listening strategy. They are distributed irregurarly among the three test pairs with the same phonemic vowel and may therefore be attributed to individual factors rather than to features of the stimuli. These especially deviant cases are distributed amongst six listeners, two listeners showing 2 cases each.

As a measure of the degree of deviation (and thus the reliability or consistency) of each listener in identifying all the stimuli, the total deviation ratio for each listener was calculated. They are shown in figure 2. The degree of uncertainty for each listener's identification is determined by (1) the sum of stimuli which were identified differently by him and (2) the sum of differing responses (0 - 5 per stimulus) for all the 39 stimuli. It is expressed in percent relating each listener's real deviation to his optimal difference.

The nine listeners deviate to different degrees. Listener A 1 identifies most consistently, D 1 differs most. But compared to the possible degree of deviation, the listeners on the whole show a high degree of consistency and thus their judgments are definitely reliable.

(b) _The stimuli_

As a measure of reliability of stimulus identification, the group stimulus difference for each stimulus was calculated. It is obvious that the more listeners vary in their identification of a given stimulus the more difficult it is to identify. Therefore the group stimulus difference is defined as the product of (1) the sum of deviating responses of all the nine listeners per stimulus and (2) the number of listeners per stimulus who responded differently. The difference is expressed as a percentage of the optimal value of group stimulus differences. The group stimulus difference as a function of the vowel durations is shown in figure 3.

In general the degree of deviation is very small. Only four stimuli show a difference larger than 15 %, three of them pertaining to the pair _Gegal_ - _Gekal_.

Identification of the portions of the originally long vowel becomes more consistent with increasing duration as the stimulus product decreases. With the originally short vowels, however, identification becomes more uncertain with increasing vowel duration. This suggests at least that, as the formant transitions towards the following consonant and the intensity or any other alleged acoustic correlates of the contact features are included in the longest portions of the short vowel segments, the proposed contact feature does not facilitate the identification task for the listeners.

By calculating the means of the group stimulus differences, the degree of certainty, with which the whole group judged each vowel, can be expressed. The means for each vowel are shown in figure 4. It can be clearly seen that the pair _Nasn_ - _nassn_, which exhibits the largest difference in vowel quality, since the vowels represent two different phonemes, is identified with greatest certainty. The pair Gegal - Gekal, however, was judged with least consistency. It does not differ in vowel quality but only in duration.

As a rough but convenient measure of deviation (and thus the reliability of each listener's identification of each vowel in each of the runs), the vowel listener ratio was calculated. It may be argued that a listener's certainty or consistency of identification is reflected not only in the number of stimuli per vowel he identifies differently but also in the degree of deviation (that is, if he differs with only one response or five responses). Therefore, for each test word, the vowel listener ratio takes into account the number of stimuli per vowel to which each listener responded differently and the sum of deviating responses for each stimulus by each listener. It is expressed as a percentage of the maximal value of the difference thus defined.

The vowel listener ratios (as percentages) are shown graphically in figure 5. It is obvious that there is considerable variance between the originally long and short vowel. The vowel listener ratio is zero in only a few cases. It is never zero in Gegal and Gekal, which fits in the listener's impression that these words were the most difficult ones to identify.

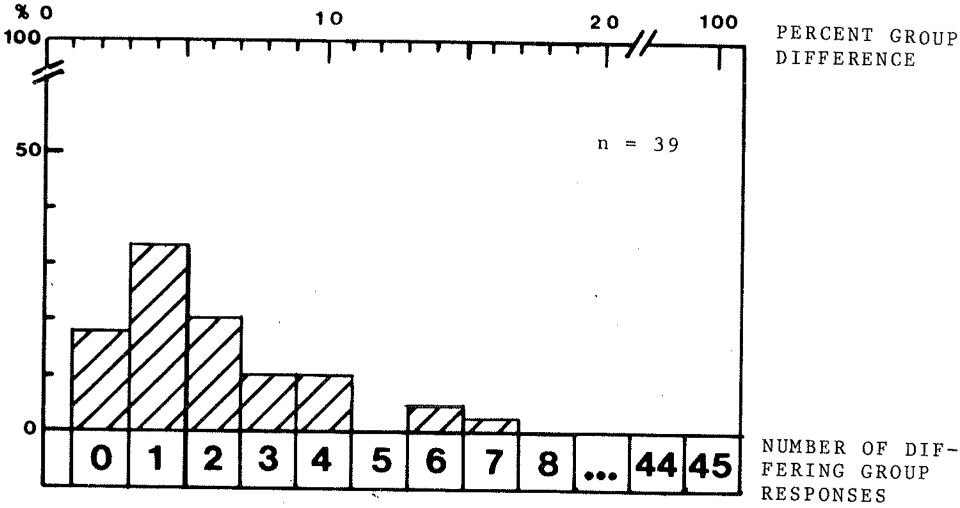
(c) The group

The group differences for each stimulus (9 listeners x 5 responses per stimulus, maximally 45) and each word (maximally 225) are shown in the following table:

Table 2. Group differences for each stimulus.

	V:C						VC:						
	Vowel portions					word	vowel portions					word	
	1	2	3	4	5	(1-5)	1	2	3	4	5	(1-5)	
Gegal	1	0	2	6	1	10	1	7	1	4	0	13	Gekal
Feda	3	1	1	1	1	7	2	4	0	3	2	11	Feta
Wiesn	6	4	2	0	0	12	1	1	2	3	-	7	wissn
Nasn	2	1	1	0	1	5	2	4	2	0	3	11	nassn

The distribution of the 39 stimuli on the number of differing responses is given in the following diagram:



The group differences are very small indeed. For 36 of the 39 stimuli the group difference is lower than 10 % (4 differing responses). Seven stimuli do not differ at all. Therefore it may be inferred that the listeners as a whole responded very consistently to the stimuli. This can be interpreted as meaning that the listeners used the same strategy (picked the same features) in perceiving the test stimuli in both presentations.

In summary then, the comparatively small deviations in the indentifications scores of the two runs, although varying between the listeners and the stimuli, justifies adding the scores of both runs together and treating them as a whole, yielding the total group scores.

7. Identifications of the stimuli

The total group scores (identification of the stimuli in both runs as the word containing the sequence /V:C/) as a function of segment duration are shown in figure 6a-b.

When individual responses are added together for the group score, deviating tendencies are eliminated. Therefore each listener's total score is shown in figure 7a-d. It can be seen there that, apart from certain deviations, each listener's identification curves look very much the same as the total group curves. Thus it may be assumed that all the listener, by and large, perceived the stimuli in a similar way.

Among the identification curves of the group no instance supporting the guessing hypothesis can be found. It may therefore be concluded that listeners, when presented with different portions of just the vowel, are able to identify a given vowel-consonant sequence. They do not guess but rely on some feature of the vowel segment alone, as the following consonant is missing. Besides the segmental feature of quality, the vowels have the feature of duration. But it is obvious, especially in the pair Gegal - Gekal, that listeners identified the vowel portions picking duration as the cue without being able to compare the duration of the vowel portions with those of the following segment (or segments). Therefore it is assumed that listeners can establish two phonological categories of length (quantity), hearing the vowels as differing in segment duration. They must do it with reference to some absolute match (time value) and not a relative one.

The phoneme boundary between the two categories long and short is located at about 100 msec. The zone of ambiguity (transition of the identification curve) is not very sharp but rather extends over a period of about 40 msec.

All the identification curves observed correspond to the expected results of the 2nd hypothesis. The pair Gegal - Gekal, where the vowel quality of the long and the short vowel is nearly the same, is identified exclusively according to the duration of the vowel portions, the pair Nasn - nassn, on the contrary, only according to vowel quality (figure 6a). As two different vowel phonemes are concerned, their quality differs considerably, compared with the differences of vowel quality between the long and short members of the same quality phoneme.

If, as in Gegal- Gekal, vowel duration is the only cue in the presented vowel portions of the originally long and short vowels, since the vowel quality is the same, vowel portions with short duration are heard as short vowel, those with long duration as long vowels, irrespective of their origin.

Both features (duration and quality) are used in the identification of the pairs Feda - Feta and Wiesn - wissn (figure 6b). The originally long vowels of Feda and Wiesn are heard as the corresponding short ones with decreasing segment duration. Whereas the longest portion of the e of Gekal (130 msec) is identified as the long vowel in 87 % of the responses, the entire short e of Feta (120 msec) and the entire short i of wissn (stimulus no. 4, 110 msec) are both heard as the long vowel to an equally low degree, only 38 %, which is about half of that of the Gekal case. This difference must be due to differences in vowel quality between the long and short vowels, which is a consequence of the features of the segmental context (place of articulation of the surrounding consonants and its effect on the medial vowels in terms of formant transitions). In spite of the quality of the short vowels of Feta and wissn, duration affects the perception of their longest portions. Thus the feature of duration tends to override the feature of quality both with increasing and decreasing vowel length.

No instance pertaining to the 3rd hypothesis is found in the material. Consequently, the proposed feature of contact between the vowel and the following consonant does not affect the listeners' perception, although the complete formant transition towards the following long consonant k: is included in the longest stimuli (nos 4 and 5) of the short vowel of Gekal. The same is true of the longest stimuli (nos 4 and 5) of the short vowel of Feta. The identification scores become higher with increasing segment duration. Nor is there any effect of the contact feature towards the end of the short vowel of wissn. Stimulus no 4 is identified as being a long vowel to a considerably greater degree than is stimulus no 3, which is only 20 msec shorter.

The following attempt was made to determine the so-called phoneme boundary or the center of the zone of ambiguity of identification between the long and the short vowel category: The 50 % response line was drawn in the graphs of figure 6 from the shortest portion of the originally short vowel to the longest portion of the originally long vowel. The intersections of this 50 % line and the identification curves represent the category boundary. For the contrasting quality pair Nasn - nassn there is no crossover point, nor is there any for the short vowels of Feta and wissn, although their curves approach rather close to the 50 % level. It may be assumed, however, that the identification curves of these short vowels would rise beyond the 50 % line if the vowel segments were to be lengthened by means of electronic splicing (for this method see Bannert 1975).

The 50 % level of identification score intersects with the curves of _Gegal_ and _Gekal_ at about 100 msec, with those of _Feda_ and _Wiesn_ to a lower value, about 80 msec. It should be noted that the segment duration of [e:] in _Feda_ and of [i:] in _Wiesn_ were approximately the same (140 and 150 msec respectively), whereas the long [e:] in _Gegal_ had a longer duration, namely 180 msec (table 1). One can speculate that this difference reflects the perceptual knowledge of physiologically conditioned effects which listeners are supposed to have (cf Nooteboom 1973). In this case, vowel duration would be longer in symmetric syllables than in asymmetric, a hypothesis which has to be proven.

If the zone of ambiguity is defined as that part of the curve which lies between the 25 % and the 75 % level of identification, then the corresponding part on the time axis equals about 40 msec (the distance between 80 and 120 msec).

8. Conclusions

The results of the present test indicate that listeners can identify the members of minimal pairs of mutual complementation in Central Bavarian even when they hear portions of just the vowel segment which differ in duration. In performing the identification task listeners use the phonetic features of duration and spectral pattern, the first of which is the phonological feature of quantity of the language, manifested as segment duration in the vowel portions.

If the spectral structure of the originally short vowel differs largely from that of the originally long vowel (that is if the two spectra belong to two different phonemic vowel qualities), listeners make exclusive use of this cue, the dimension of duration having no influence whatsoever on perception. This is in close accord with the findings in Hadding-Koch and Abramson (1964) on Swedish material. The converse is also true: listeners judge according to duration alone if the short and the corresponding long vowel do not differ in vowel quality. No influence whatsoever of the alleged contact feature, manifested in the last portion of the vowel segment as intensity or formant transitions, is to be found in this material. If such a feature did exist the short vowel would be heard as short more often with increase of the final part of the vowel segment approaching the following consonant.

Duration and vowel quality compete with each other as cues for identification in minimal pairs showing some difference in vowel quality: The quality of the short vowel always dominates although the dimension of time becomes more important for perception with increasing duration. For long vowels, on the other hand, duration becomes the predominant cue for perception the shorter the vowel portion is made.

In the perception of the test vowels differences in vowel quality are used in preferance to differences in time. The listeners show greater consistency in identifying quality than length. Both observations are in agreement with the universal phonological fact that all languages use the phonetic mechanism of spectral structure (vowel quality) in producing sound contrasts in order to signal differences of meaning. But not all languages utilize segment duration (phonemic length or quantity) for the same purpose.

The findings of this investigation are therefore interpreted as supporting the view that the phonemic dimension of length is less effective than quality in differentiating meaning. They may thus be considered as phonetic evidence for attributing a lower rank to quantity than quality in a hierarchy of distinctive features.

References

Abramson A.S. 1962. The vowels and tones of Standard Thai: Acoustical measurements and experiments. International Journal of American Linguistics. Vol. 28, Number 2, Part III

Bannert R. 1972. Zur Stimmhaftigkeit und Quantität in einem bairischen Dialekt. Working Papers 6. Phonetics Laboratory, Lund University

Bannert R. 1973a. Mutual complementation of VC-sequences in Central Bavarian. Working Papers 9, 72-83. Phonetics Laboratory, Lund University

Bannert R. 1973b. A binary treatment of the distinctive prosodic feature of quantity. Working Papers 9, 84-97. Phonetics Laboratory, Lund University

Bannert R. 1974. The phonetic feature of voicing in the alveolar fricative of a Central Bavarian speaker. Mimeographed. Department of General Phonetics, Lund University

Bannert R. 1975. The perception of VC-sequences in Central Bavarian. Mimeographed. Department of General Phonetics, Lund University

Bannert R. Forthcoming. Quantität im Mittelbairischen. To appear in: Travaux de l'Institut de Linguistique du Lund, B. Malmberg and K. Hadding (eds). Lund

Bastian J. and A.S. Abramson. 1962. Identification and discrimination of phonemic vowel duration. Journal of the Acoustical Society of America 34, 743-744 (Abstract)

Durand M. 1946. Voyelles longues et voyelles brèves: Essai sur la nature de la quantité vocalique. Paris

Elert C.-C. 1964. Phonologic Studies of quantity in Swedish. Uppsala

Fant G., G. Henningsson, and U. Stålhammar. 1969. Formant frequencies of Swedish vowels. STL-QPSR 4, 26-31

Fischer-Jørgensen E. 1956. The commutation test and its application to phonemic analysis. For Roman Jakobson, 140-151. The Hague

Fischer-Jørgensen E. and H.P. Jørgensen. 1969. Close and loose contact ("Anschluss") with special reference to North German. Annual Report of the Institute of Phonetics 4, 43-80. University of Copenhagen

Fischer-Jørgensen E. 1972. Formant frequencies of long and short Danish vowels. Annual Report of the Institute of Phonetics 6, 49-57. University of Copenhagen

Fliflet A.L. 1961. Einige Beobachtungen über Anschluss und Silbe. Proceedings of the 4th International Congress of Phonetic Sciences, 610-615

Fry D.B. 1970. Prosodic phenomena. Manual of Phonetics, 365-410. B. Malmberg (ed.). Amsterdam

Gladiator K. 1971. Untersuchungen zur Struktur der mittelbairischen Mundart von Grossberghofen. Münchner Studien zur Mundartforschung, H. Kuhn und I. Reiffenstein (eds). Munich

Hadding-Koch K. and A.S. Abramson. 1964. Duration versus spectrum in Swedish vowels: some perceptual experiments. Studia Linguistica 18, 94-107

Heike G. 1969. Suprasegmentale Analyse. Marburger Beiträge zur Germanistik, Bd. 30, J. Kunz, E. Ruprecht und L.E. Schmitt (eds). Marburg

Jakobson R., G. Fant, and M. Halle. 1952. Preliminaries to speech analysis. The distinctive features and their correlates. Cambridge, Mass.

Jakobson R. and M. Halle. 1970. Phonology in relation to phonetics. Manual of Phonetics, 411-449. B. Malmberg (ed.). Amsterdam

Jespersen O. 1904. Lehrbuch der Phonetik. Leipzig

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

Keller R.E. 1961. German dialects. Manchester

Koekkoek B.J. 1955. Zur Phonologie der Wiener Mundart. Beiträge zur deutschen Philologie. W. Mitzka (ed.), Bd. 6, Giessen

Kozhevnikova V.A. and L.A. Chistovich. 1965. Speech, articulation, and perception. Joint Publications Research Service 30543. Washington, D.C.

Kufner H. 1956. The dialect of the Freutsmoos area (Bavaria). Dissertation. Cornell University

Kufner H. 1957. Zur Phonologie einer mittelbairischen Mundart. Zeitschrift für Mundartforschung 25, 175-184

Kufner H. 1961. Strukturelle Grammatik der Münchner Stadtmundart. Munich

Ladefoged P. 1971. Preliminaries to linguistic phonetics. Chicago

Lehiste I. 1967. Suprasegmental features, segmental features, and long components. Proceedings of the 10th International Congress of Linguistics. Bucharest

Lehiste I. 1970. Suprasegmentals. Cambridge, Mass.

Lehiste I. 1971. Experiments with synthetic speech concerning quantity in Estonian. Working Papers in Linguistics 9, 200-217. Ohio State University

Lehtonen J. 1970. Aspects of quantity in Standard Finnish. Studia Philologica, Jyväskyläensia VI

Liberman A.M., F.S. Cooper, D. Shankweiler, and M. Studdert-Kennedy. 1969. Perception of the speech code. Brain and Behaviour 4, 105-149. K.H. Pribram (ed.)

Lindblom B. 1971. Phonetics and the description of language. Proceedings of the 7th International Congress of Phonetic Sciences, 63-97. Montreal

Lisker L. 1974. On "explaining" vowel duration variations. Haskins Laboratories Status Report 37/38, 225-232

Nooteboom S.G. 1972. Some timing factors in the production and perception of vowels. Occasional Papers 13, 49-76. University of Essex

Nooteboom S.G. 1973. Perceptual reality of prosodic durations. Journal of Phonetics 1, 25-46

Petersen N.R. 1974. The influence of tongue height on the perception of vowel duration in Danish. Annual Report of the Institute of Phonetics 8, 1-10. University of Copenhagen

Pfalz A. 1911. Phonetische Beobachtungen an der Mundart des Marchfeldes in Nieder-Österreich. Zeitschrift für deutsche Mundarten, 244-260

Pfalz A. 1913. Die Mundart des Marchfeldes. Sitzungesberichte der Akademie der Wissenschaften in Wien. Phil.-hist. Klasse, 170. Bd., 6. Abhandlung

Pfalz A. 1935. Zur Phonologie der bairisch-österreichischen Mundart. Lebendiges Erbe, Reclam Festschrift. Leipzig

Pilch H. 1964. Phonemtheorie I. Teil. Bibliotheca Phonetica 1. Basel and New York

Sievers E. 1881. Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen. Leipzig

Slis I.H. 1972. The influence of articulatory effort on the timing of speech. Occasional Papers 13, 128-150. University of Essex

Studdert-Kennedy M. 1971. The perception of speech. Current trends in linguistics. Vol. XII. The Hague

Trubetzkoy N.S. 1958. Grundzüge der Phonologie. Göttingen

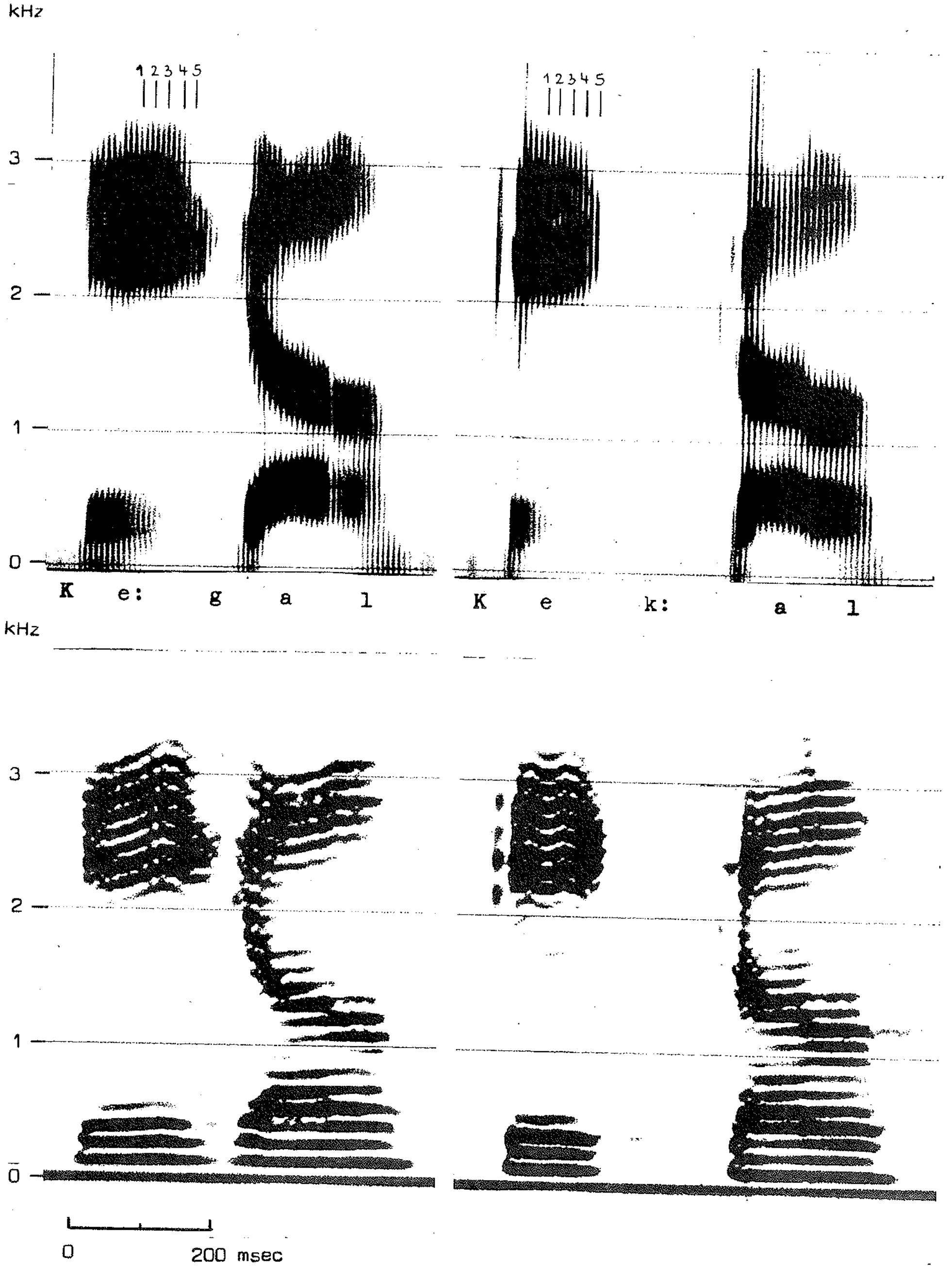


Figure 1a. Broad-band and narrow-band spectrograms of Gegal and Gekal. The vowel portions (stimuli 1-5) are indicated.

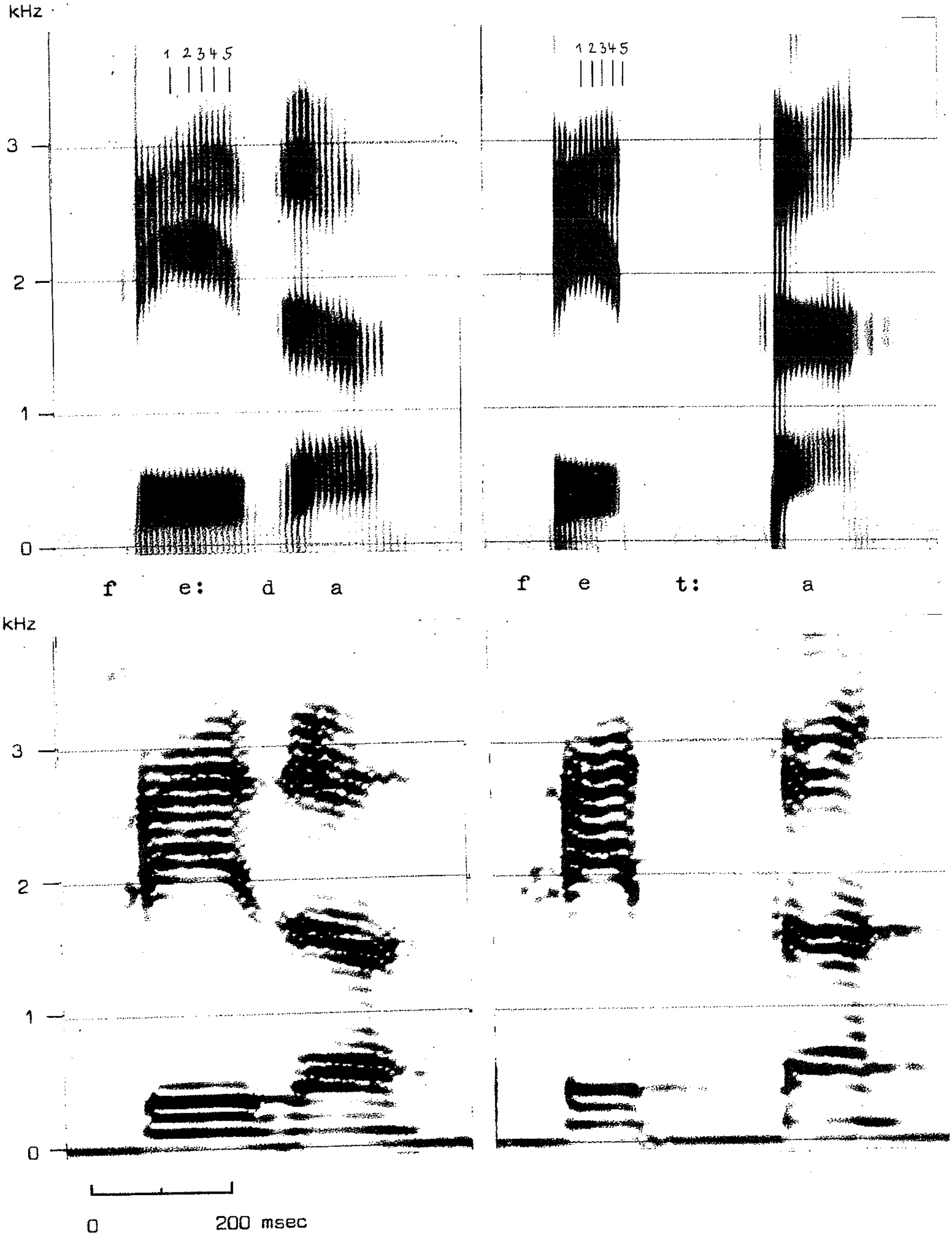


Figure 1b. Spectrograms of Feda and Feta. The vowel portions are indicated.

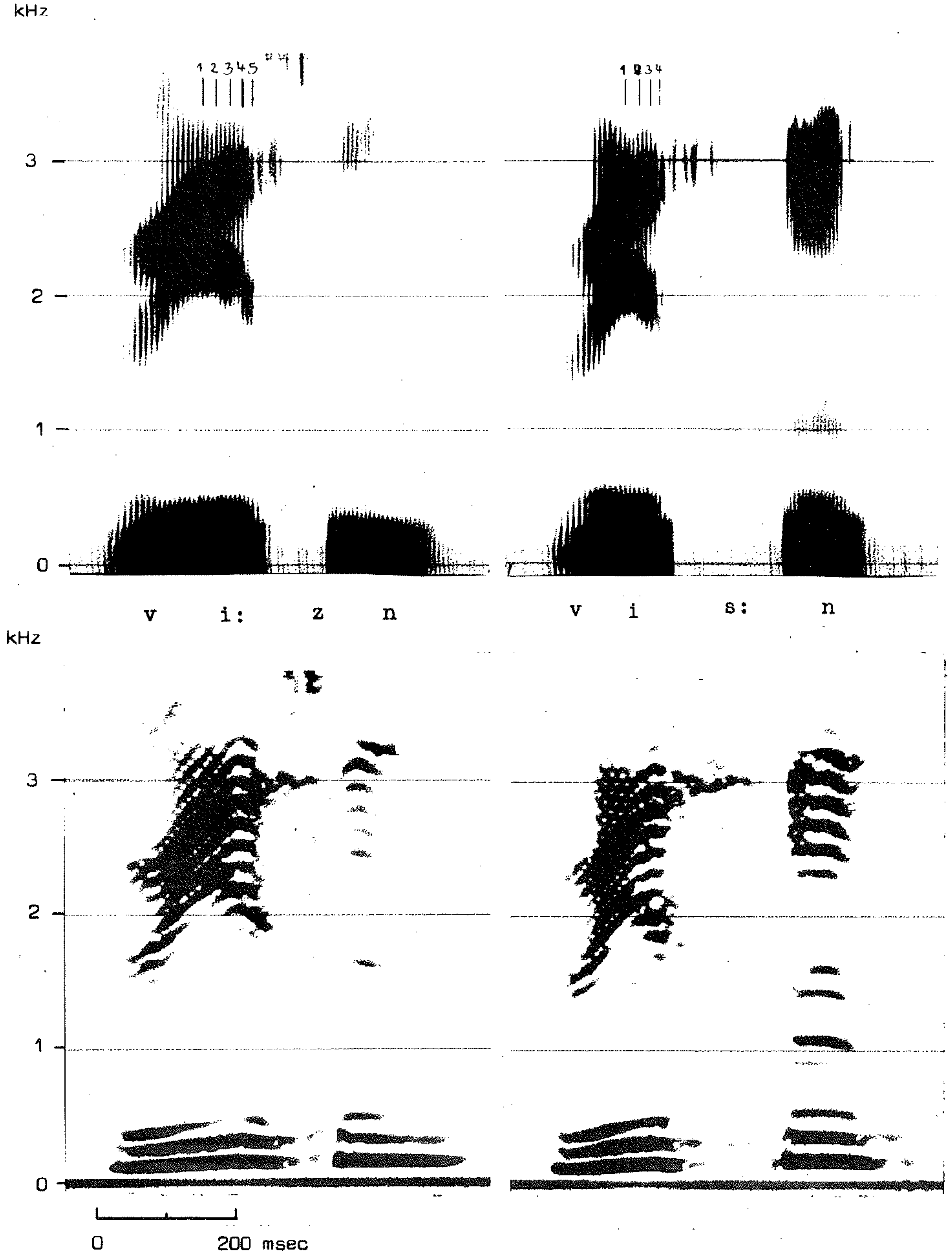


Figure 1c. Spectrograms of <u>Wiesn</u> and <u>wissn</u>. The vowel portions are indicated.

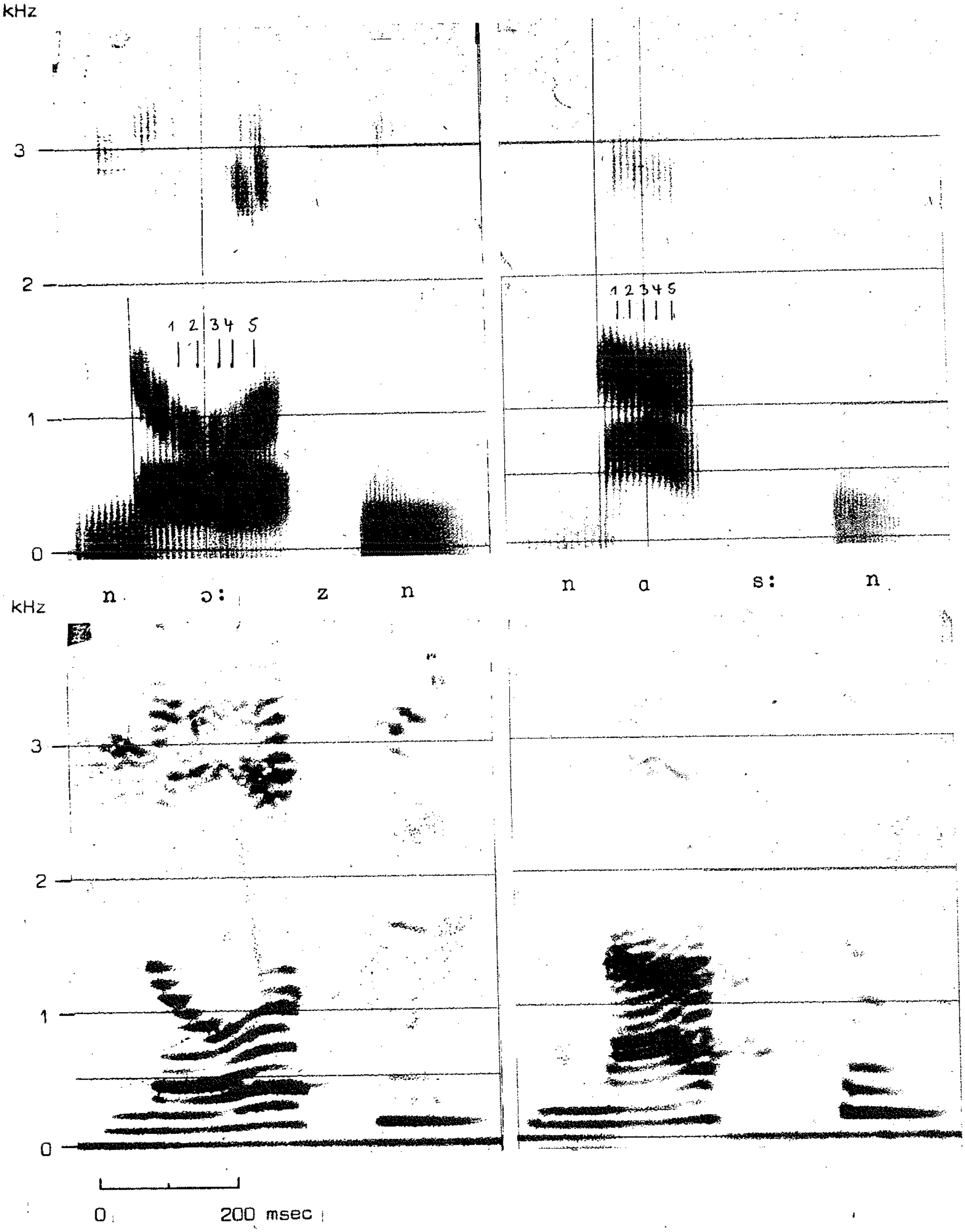


Figure 1d. Spectrograms of Nasn and nassn. The vowel portions are indicated.

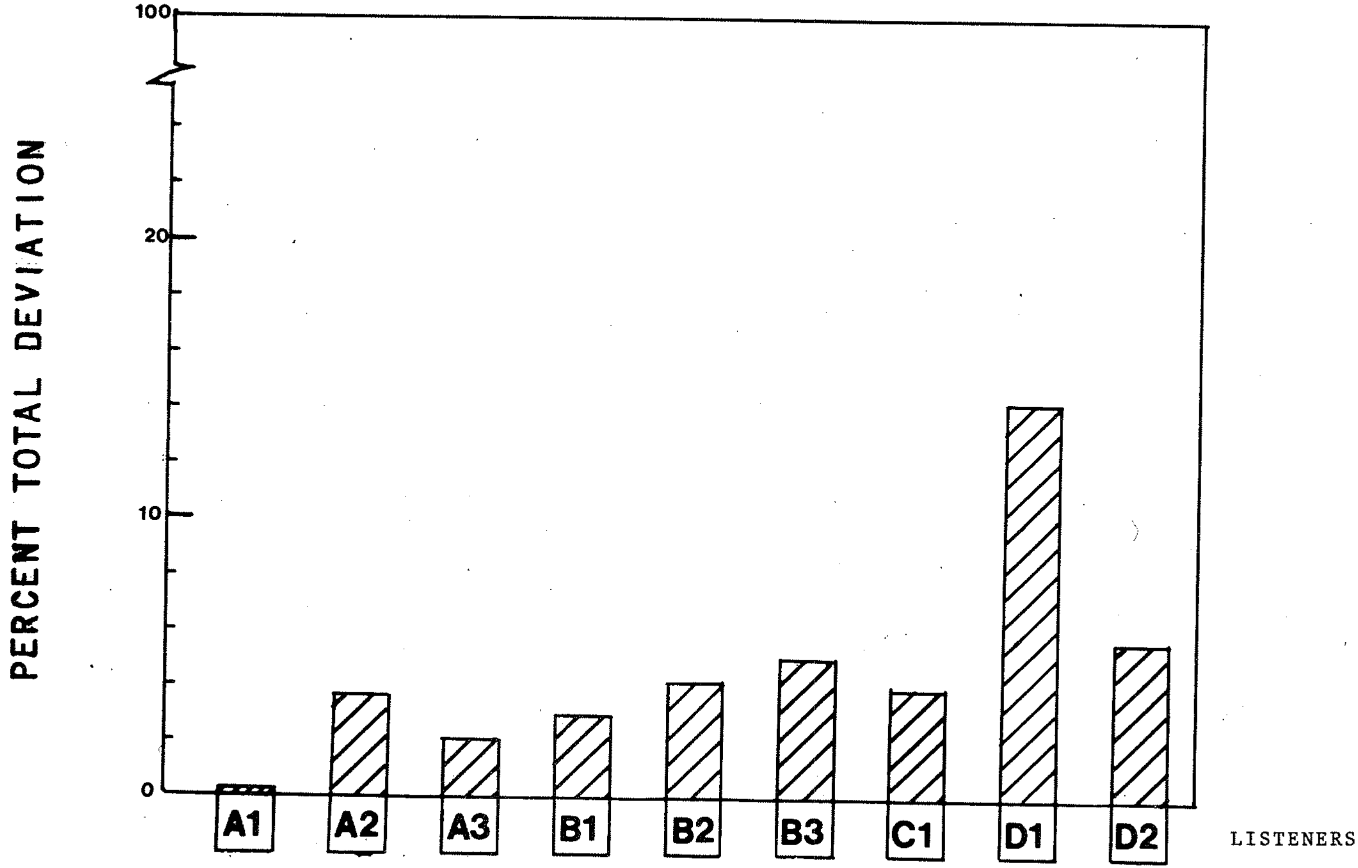


Figure 2. Total deviation of each listener as a measure of reliability of judgment.

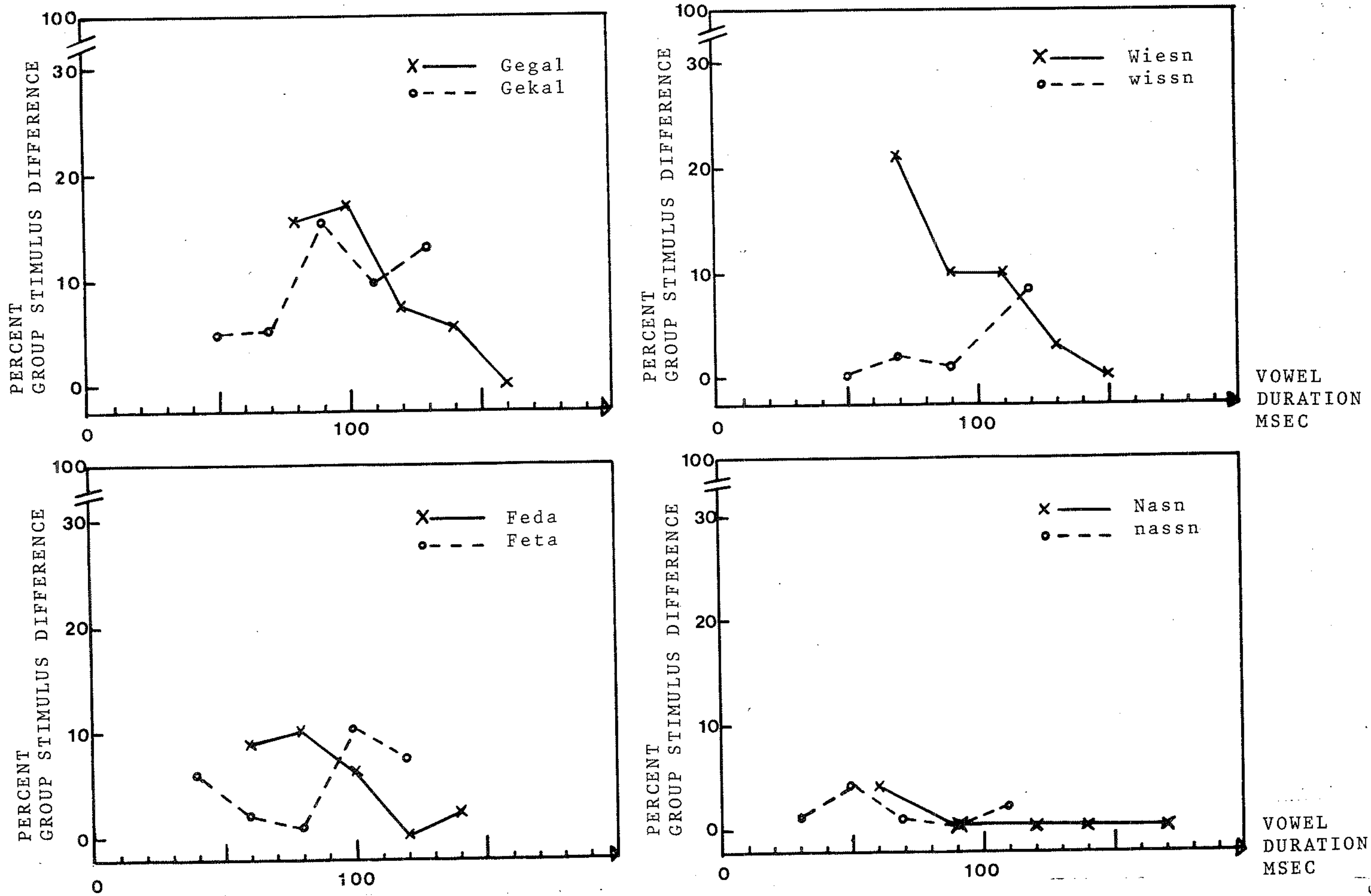


Figure 3. The group stimulus differences expressing the inconsistency with which each stimulus was indentified by all the listeners.

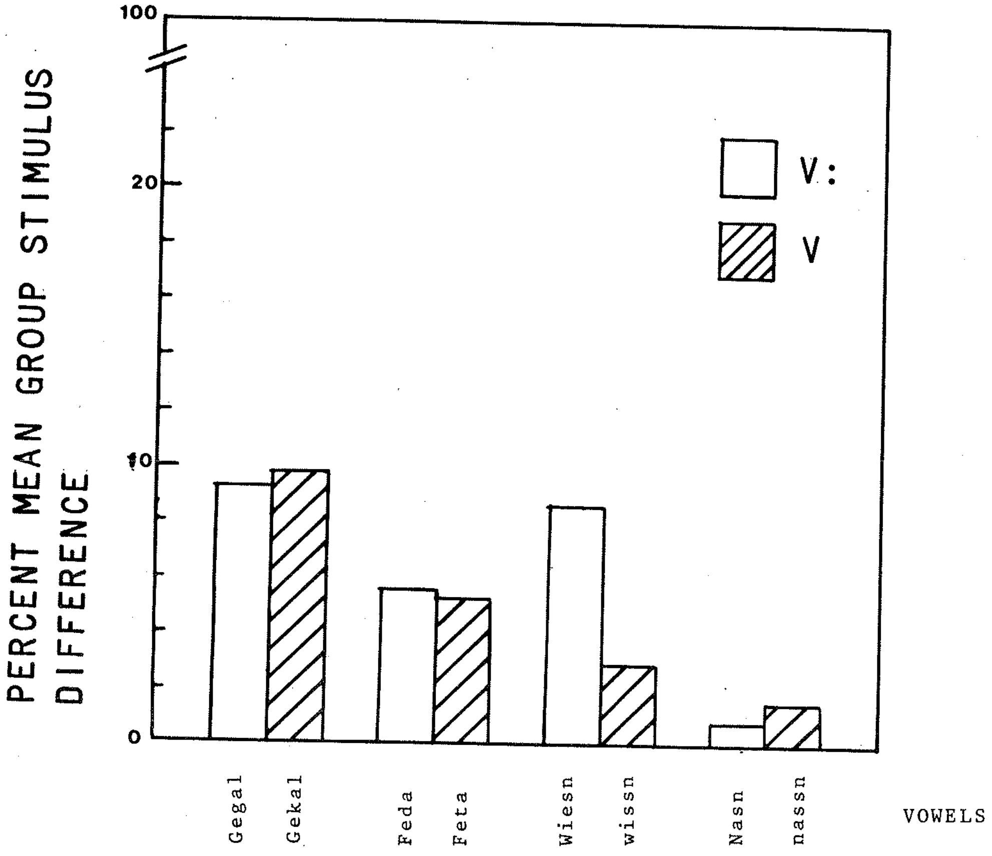


Figure 4. Means of group stimulus differences for each vowel.

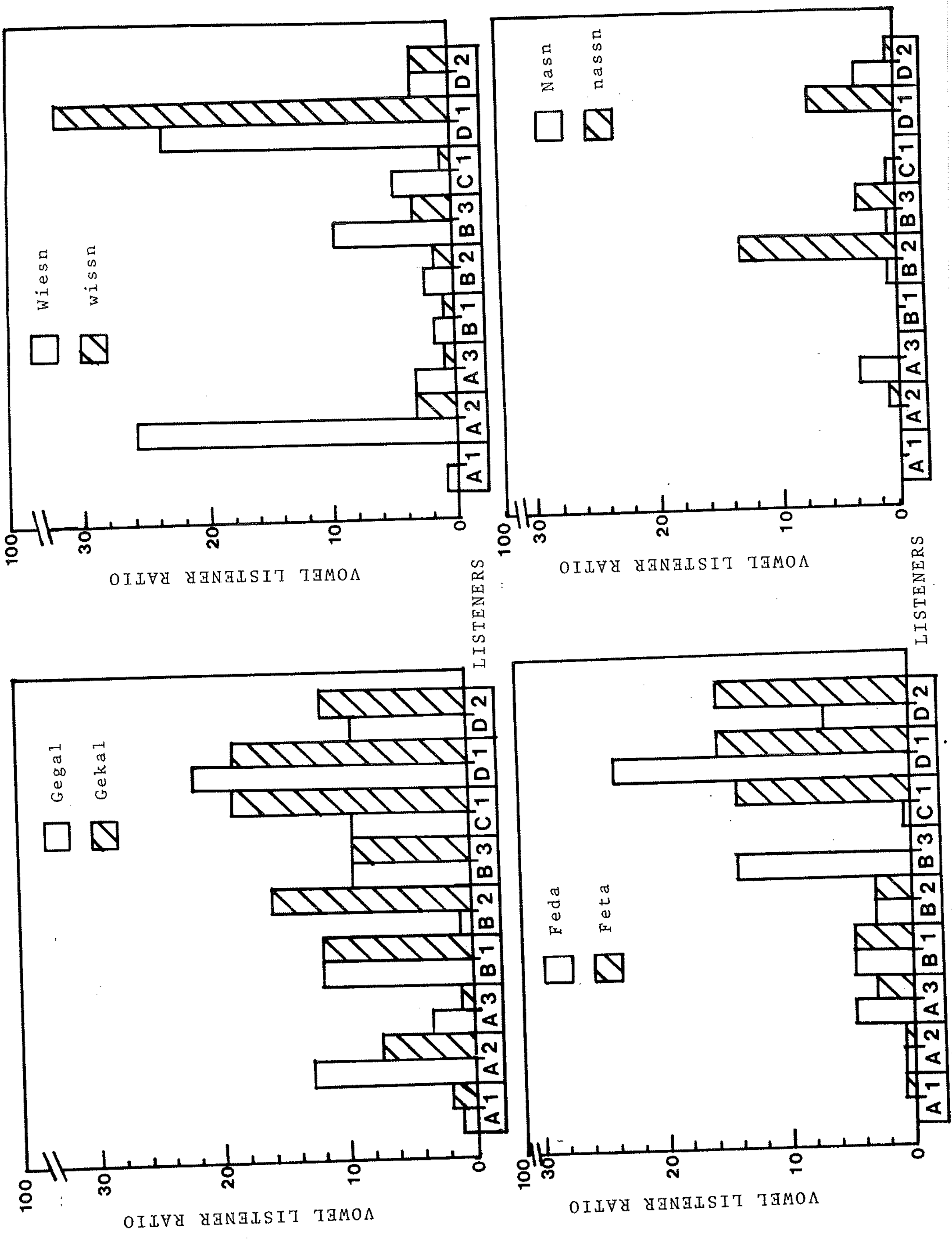


Figure 5. The vowel listener ratios as a measure of uncertainty of identification of each vowel by each listener.

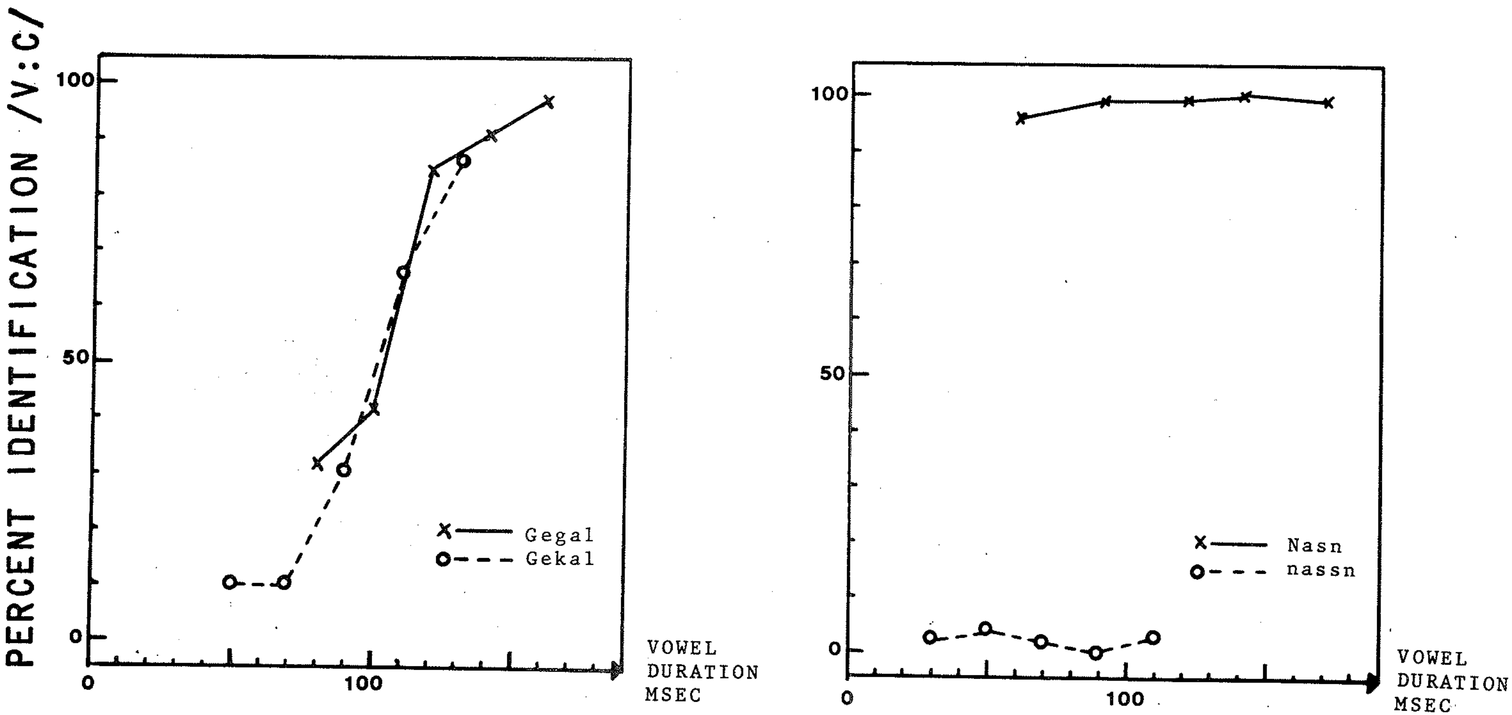


Figure 6a. Identification of the vowels of <u>Gegal</u>-<u>Gekal</u> (left) and <u>Nasn</u>-<u>nassn</u> (right) as /V:C/.

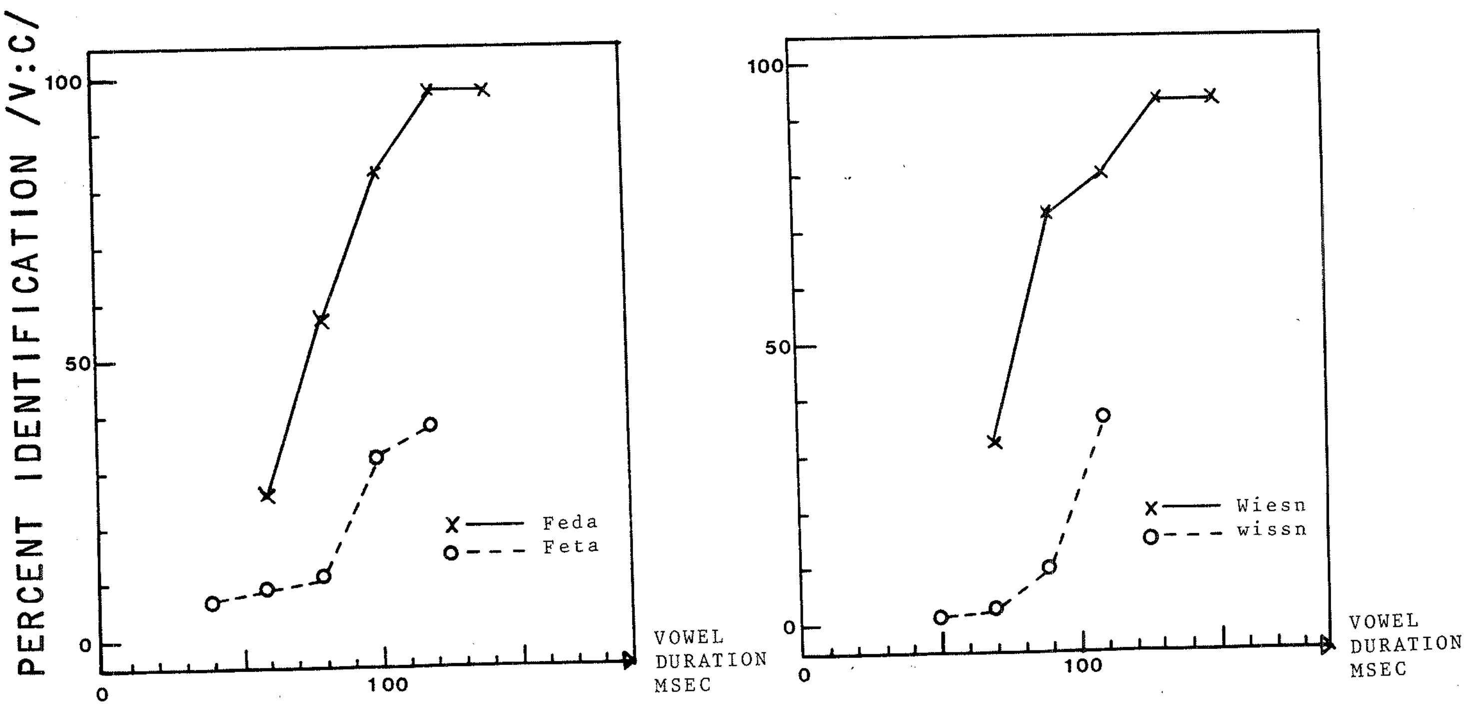


Figure 6b. Identification of the vowels of Feda-Feta (left) and Wiesn-wissn (right) as /V:C/.

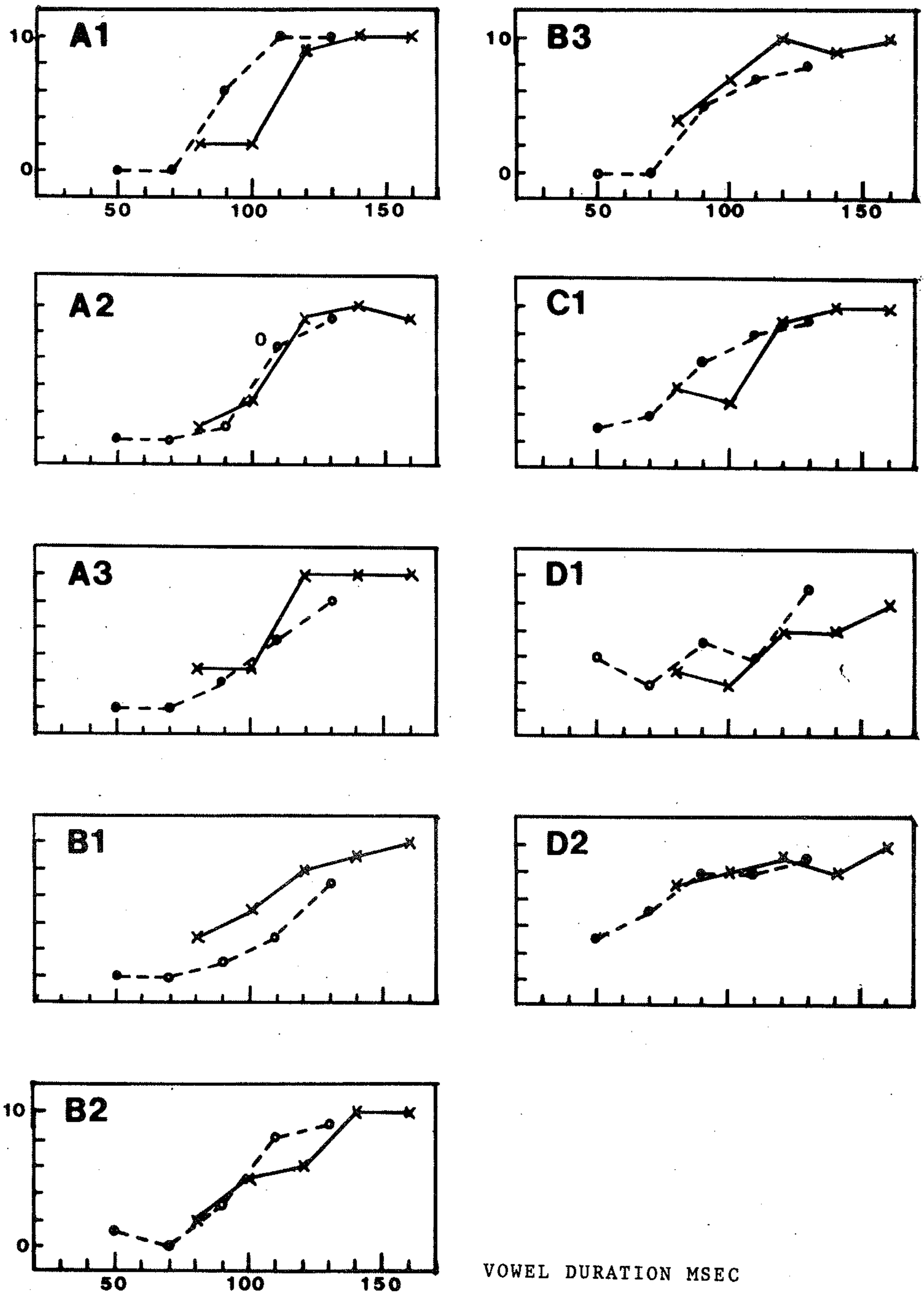


Figure 7a. Identification curves of Gegal and Gekal as /V:C/ for eac listener (both runs, totally 10 responses per stimulus). Dashed line: originally short vowel, solid line: original long vowel.

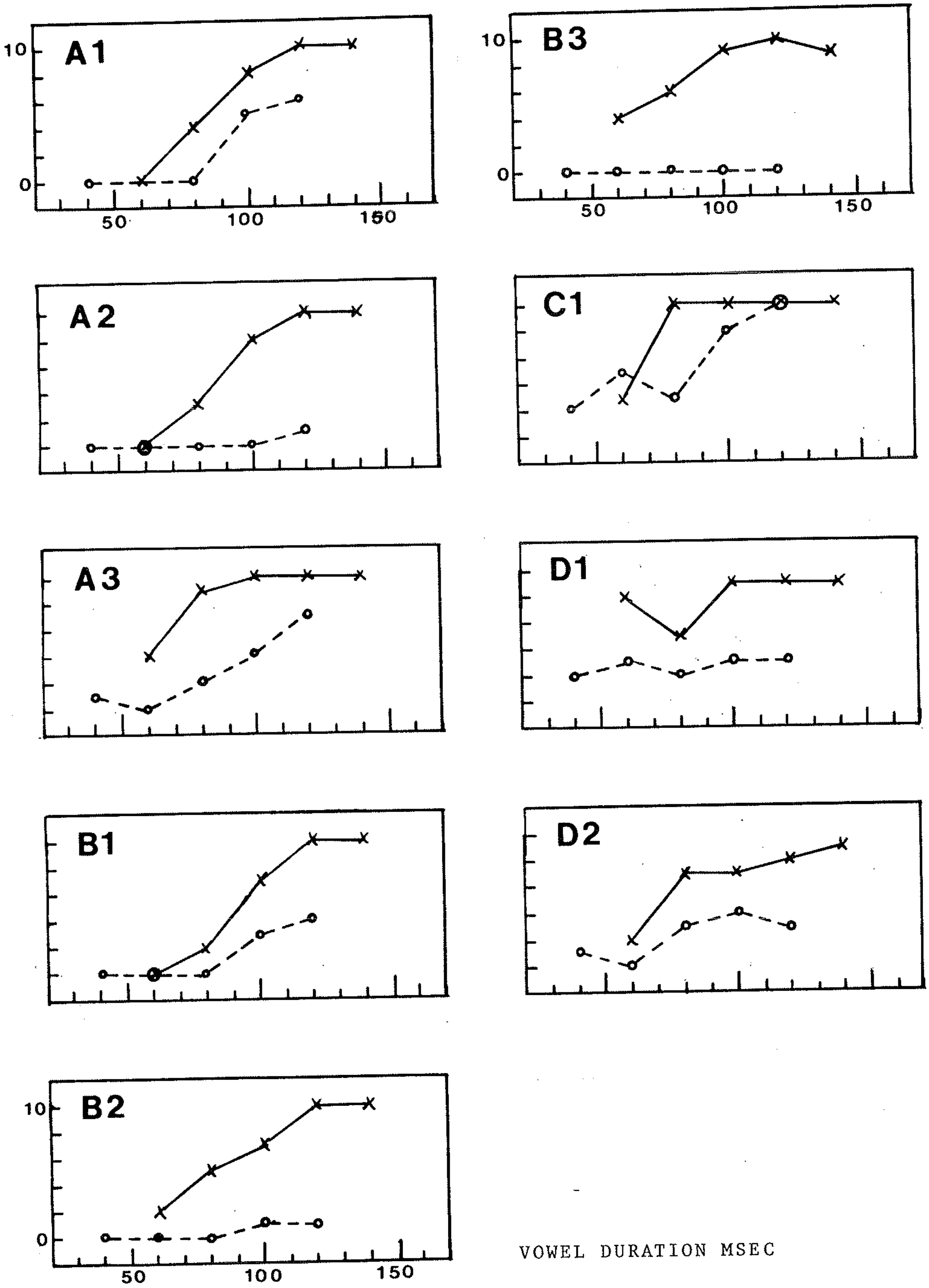


Figure 7b. Identification curves of Feda and Feta as /V:C/ for each listener.

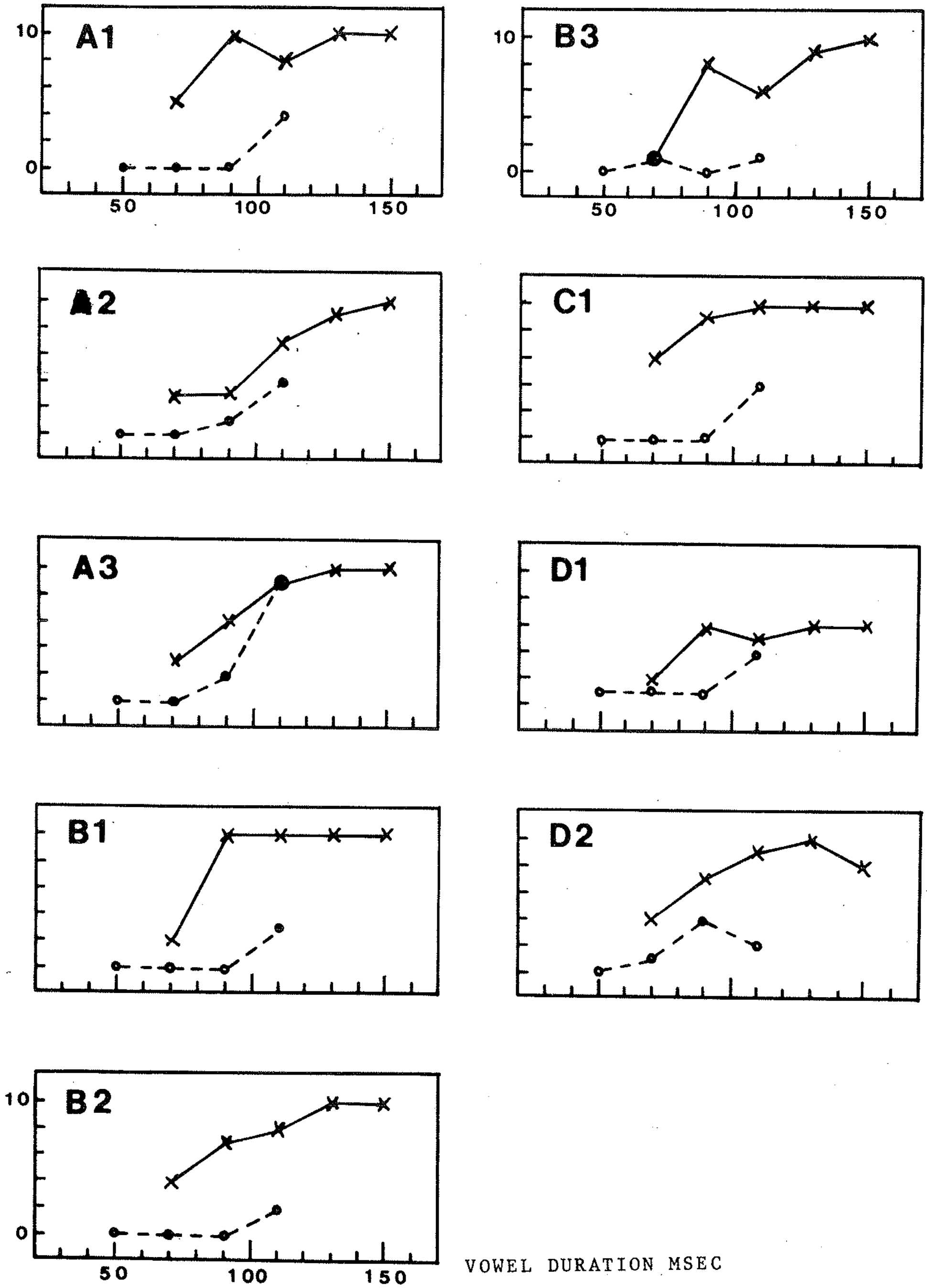


Figure 7c. Identification curves of <u>Wiesn</u> and <u>wissn</u> as /V:C/ for each listener.

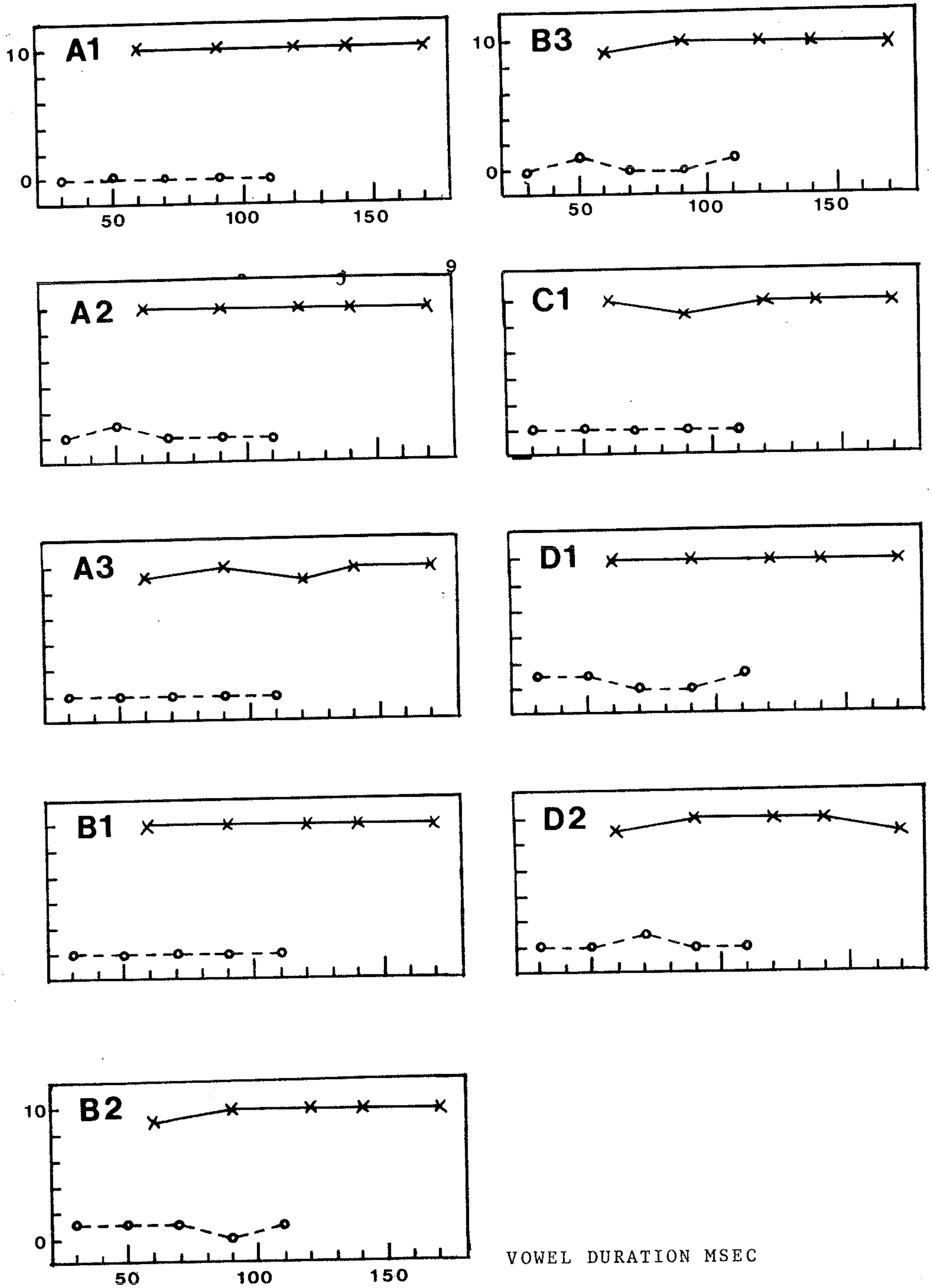


Figure 7d. Identification curves of <u>Nasn</u> and <u>nassn</u> as /V:C/ for each listener.

Working Papers 12 · 1975 Lund University
Phonetics Laboratory General Linguistics

TEMPORAL ORGANIZATION AND PERCEPTION OF VOWEL-CONSONANT SEQUENCES IN CENTRAL BAVARIAN*

Robert Bannert

The phonological description of Central Bavarian sound structure faces a central problem, namely the analysis of stressed vowel-consonant sequences, According to the literature these sequences are of two kinds: A long vowel is followed by a lenis, weak or short consonant, while a short vowel is followed by a fortis, strong or long consonant. The contrast between the two sequences is illustrated by a word pair like [ke:gal] (name of a hill) and [kekal] (roast chicken).

Although no phonetic data are provided in the literature, the following four features are postulated as the one distinctive feature of both segments:

(1) Weak of strong accent ("schwach bzw. stark geschnittener Akzent", Pfalz 1913 following Sievers 1881).

(2) Offglide ("Abglitt", Gladiator 1971 following Pilch 1964).

(3) Force of articulation or intensity of the consonant (Koekkoek 1955, Keller 1961). This opposition is labelled Fortis-Lenis.

(4) Length of the consonant (Kufner 1956).

I shall propose here that also the Central Bavarian vowel-consonant sequences represent a case of temporal organization of a vowel and the following consonant, which is called complementary length or mutual complementation of

---

* Paper presented at the Eighth International Congress of Phonetic Sciences, Leeds, August 17-23 1975

vowel and consonant (Lehiste 1970) and which is found in other germanic languages, too (e.g. Standard Central Swedish).

In view of the lack of phonetic measurements, I commenced an investigation of Central Bavarian vowel-consonant sequences on two levels of the speech communication process. Firstly, an acoustical analysis and description of the vowel and the consonant of such sequences was made. Secondly, the perception of minimally contrasting pairs was studied using natural stimuli in which several acoustical features of the vowel and the consonant were altered.

In this report I shall first give the durations of both segments and some conclusions I have drawn regarding the temporal organization of the sequences in two word pairs. I shall then give an account of the perception tests where vowel and consonant durations were altered.

Temporal organizatinn of vowel-consonant sequences

Segment durations of vowels and consonants were measured in meaningful, bisyllabic words which consists of a /CVC/-root and the suffix /al/ denoting the diminutive. The two test pairs, the first of which shows the minimal contrast in the vowel-consonant sequences, are the following:

/CVCal/

V:C		VC	
[ke:gal]	(name of a hill)	[kekal]	(roast chicken)
[pi:bal]	(chicken)	[pupal]	(little doll)

The initial consonant is a voiceless, non-aspirated stop, either /k/ or /p/. The manifestation of the postvocalic stop varies. Following a short vowel it is voiceless and non-aspirated. Following a long vowel, however, it may be a voiced stop or even a voiced fricative. This variation lies behind the phonological dispute.

The absolute durations of the segments were forced to vary considerably in order to study the temporal organization of the vowel and the consonant under widely varying conditions. To achieve this, the test words were all placed in five different prosodic patterns. All were statements:

(1) The testword produced alone:	E	emphatic
	N	neutral with normal tempo
	F	neutral with fast tempo
(2) The test word placed in a sentence:	S_m	test word in medial position, answer to the question: "What have you seen?": /i hap s___ gseŋ/ (I have seen the ___).
	S_i	test word in initial position, carries contrastive stress: /s ___ tat i meŋ/ (It's the ___ I would like to have).

The data on segment durations reported here pertain to speaker 2. He is 45 years old, a carpenter by profession and lives in the village of Polling near Weilheim/Obb. (about 30 miles south of Munich). The measurements of the other two informants show a similar picture.

Figure 1 gives the segment durations of the vowels and the consonants of the four test words in the five prosodic patterns in msec. Not only the long and short vowels, but also the consonant following them show clear differences of duration. It is therefore justified to talk of long and short consonants in Central Bavarian.

If the absolute differences of duration between long and short vowels and consonants within the pairs are compared (Figure 2), it turns out that the difference of segment duration between the consonants in all the prosodic patterns is considerably larger than that between the long and short vowels. This may explain the findings on vowel ratios in Bavarian reported by Zwirner (1961) and therefore is likely to weaken one important point in the argumentation of Gladiator (1971).

The proportion of each segment in the total duration of the vowel-consonant sequence may be expressed by calculating the vowel-to-sequence ratio (V/(V+C)x100 %, Figure 3). The relative segment duration within the sequences does not remain constant throughout the prosodic patterns, although its values appear to fall into two classes: It is higher in those patterns in which the test word is stressed particularly (given special prominence), it is lower in the cases of more neutral stress. The variation of the vowel-to-sequence ratio is similar to the variation of vowel duration, especially in the sequence /VC:/ (cf. Figure 1). This is because the absolute duration of the consonants varies little in comparison to that of the vowels.

The same temporal pattern has been found for corresponding Central Swedish vowel-consonant sequences (Bannert 1972), calculated on data given by Elert (1964).

<u>Perception of vowel-consonant sequences</u>

As the starting stimuli for the listening tests, one rendering each of the following four minimal pairs was used. The words were produced in isolation by speaker 1:

	V:C		VC:	
C_m=stop	[fe:da]	(feather)	[fet:a]	(male cousin)
	[ke:gal]	(name of a hill)	[kek:al]	(roast chicken)
C_m=fricative	[o:fa]	(stove)	[of:a]	(open)
	[vi:sn]	(meadow)	[vis:n]	(to know)

The vowel-consonant sequences of the test words were manipulated manually (using tape cutting and splicing techniques) and electronically (by means of a segmentator). The main aim was to change the temporal relations between the two segments. The part of the tape containing the segment boundary between the vowel and the consonant was not touched. Ten listeners were asked to identify the stimuli either as the one or the other member of the four minimal pairs.

Here I shall give the identification scores of the listeners to three manipulations:

(1) Shortening of the initial long vowel in [o:fa] (Figure 4).
(2) Shortening of the long consonants (including the elimination of the explosion of both stops). The duration of stops was decreased by cutting out portions of the silent interval of occlusion (Figure 5).
(3) Lengthening of the short stops in [fe:da] and [ke:gal]. The periodicity of the occlusion had been gated out previously.

All these changes of the duration of only one segment at a time led to the identification of a given vowel-consonant sequence as its counterpart (cf. Fliflet 1961, Heike 1969, Fischer-Jørgensen and Jørgensen 1969). This also happened when the relatively strong explosion of the considerably shortened, originally long stops was eliminated. This explosion may be considered part of the manifestation of the feature fortis.

If the results of the identification test are considered together with the temporal regularities observed in the vowel and the consonant, two things become evident:

(1) The auditive Fortis-Lenis contrast of the postvocalic consonant is always combined with clearly different durations of this segment (and the feature of voicing, cf. Bannert 1972, 1974, and forthcoming). This contrast is likely to be a consequence of the segment durations (cf. Ladefoged 1971).
(2) Although the contact between the vowel and the consonant was retained unaltered, each kind of sequence could be turned into its respective counterpart. This total reversal of identification was obtained by varying segment duration alone and, as a consequence, the temporal relations between the vowel and the consonant.

Therefore, in consideration of these preliminary phonetic data it seems likely that the Central Bavarian vowel-consonant sequences in fact constitute a case of temporally distinctive organization of phonological segments. Hence the phonological analysis should take this fact in to account and handle the phenomenon of complementary length in Central Bavarian by applying the temporal distinctive feature of QUANTITY (cf. Lehiste 1970, Bannert 1973).

The phonological solution suggested here will lead to an analysis of the segmental phonological units (phonemes) of Central Bavarian, especially the stops, which differs from those hitherto found in the literature (Bannert forthcoming).

References

Bannert R. 1972. Zur Stimmhaftigkeit und Quantität in einem bairischen Dialekt. Working Paper 6. Phonetics Laboratory, Lund University

Bannert R. 1973. A binary treatment of the distinctive feature prosodic of quantity. Working Paper 9, 84-97. Phonetics Laboratory, Lund University

Bannert R. 1974. The phonetic feature of voicing in the alveolar fricative of a Central Bavarian speaker. Mimeographed. Department of General Phonetics, Lund University

Bannert R. Forthcoming. Quantität im Mittelbairischen. To appear in: Travaux de l'Institut de l'Linguistique du Lund. B. Malmberg and K. Hadding (eds). Lund

Elert C.-C. 1964. Phonologic studies of quantity in Swedish. Uppsala

Fliflet A.L. 1961. Einige Beobachtungen über Anschluss und Silbe. Proceedings of the 4th International Congress of Phonetic Sciences, 610-615

Fischer-Jørgensen E. and H.P. Jørgensen. 1969. Close and loose contact ("Anschluss") with special reference to North German. Annual Report of the Institute of Phonetics 4, 43-80. University of Copenhagen

Gladiator K. 1971. Untersuchungen zur Struktur der mittelbairischen Mundart von Grossberghofen. Münchner Studien zur Mundartforschung. H. Kuhn and I. Reiffenstein (eds). Munich

Heike G. 1969. Suprasegmentale Analyse. Marburger Beiträge zur Germanistik, Vol. 30. J. Kunz, E. Ruprecht and L.E. Schmitt (eds). Marburg

Keller R.E. 1961. German dialects. Manchester

Koekkoek B.J. 1955. Zur Phonologie der Wiener Mundart. Beiträge zur deutschen Philologie, Vol. 6. W. Mitzka (ed). Giessen

Kufner H. 1956. The dialect of the Freutsmoos area (Bavaria). Dissertation, Cornell

Ladefoged P. 1971. Preliminaries to linguistic phonetics. Chicago

Lehiste I. 1970. Suprasegmentals. Cambridge, Mass.

Pfalz A. 1913. Die Mundart des Marchfeldes. Sitzungsberichte der Akademie der Wissenschaften in Wien. Phil.-hist. Klasse, Vol. 170, Part 6

Pilch H. 1964. Phonemtheorie I. Teil. Bibliotheca Phonetica 1. Basel and New York

Sievers E. 1881. Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen. Leipzig

Zwirner E. 1961. Phonometrischer Beitrag zur Geographie der prosodischen Eigenschaften. Proceedings of the 4th International Congress of Phonetic Sciences, 500-518

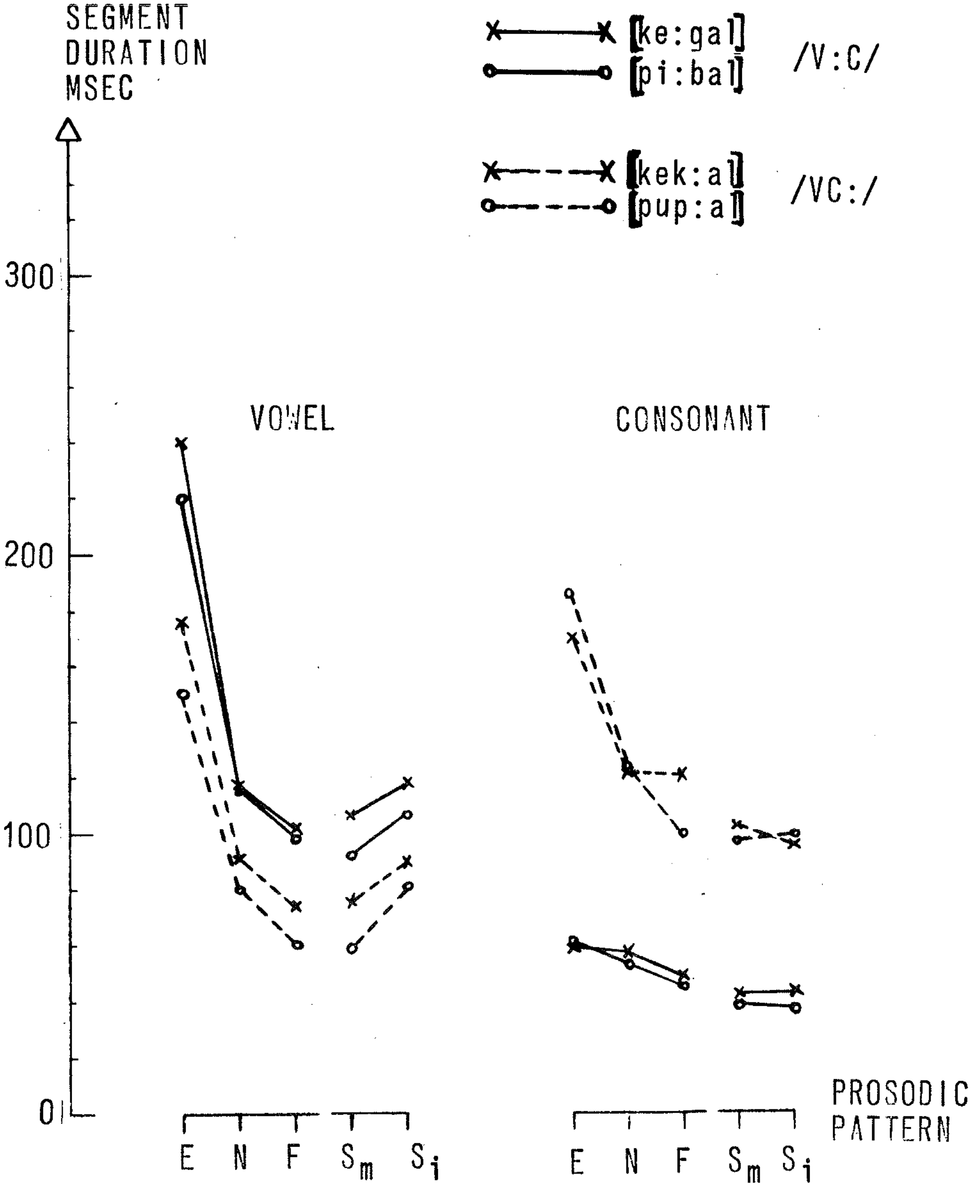


Figure 1. Durations of the vowels (left) and the consonants (right) in the four test words in msec.
Prosodic patterns see text.

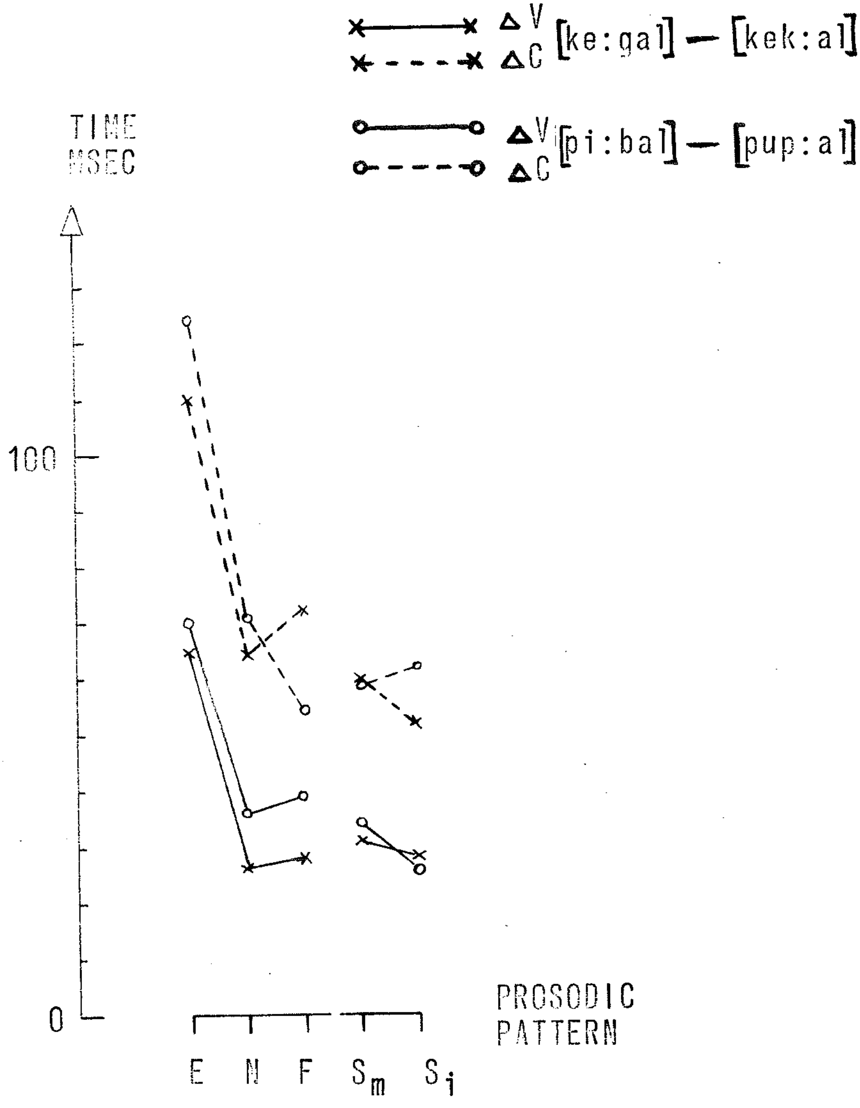


Figure 2. Difference of durations (msec) between the long and the corresponding short segments of the two word pairs.

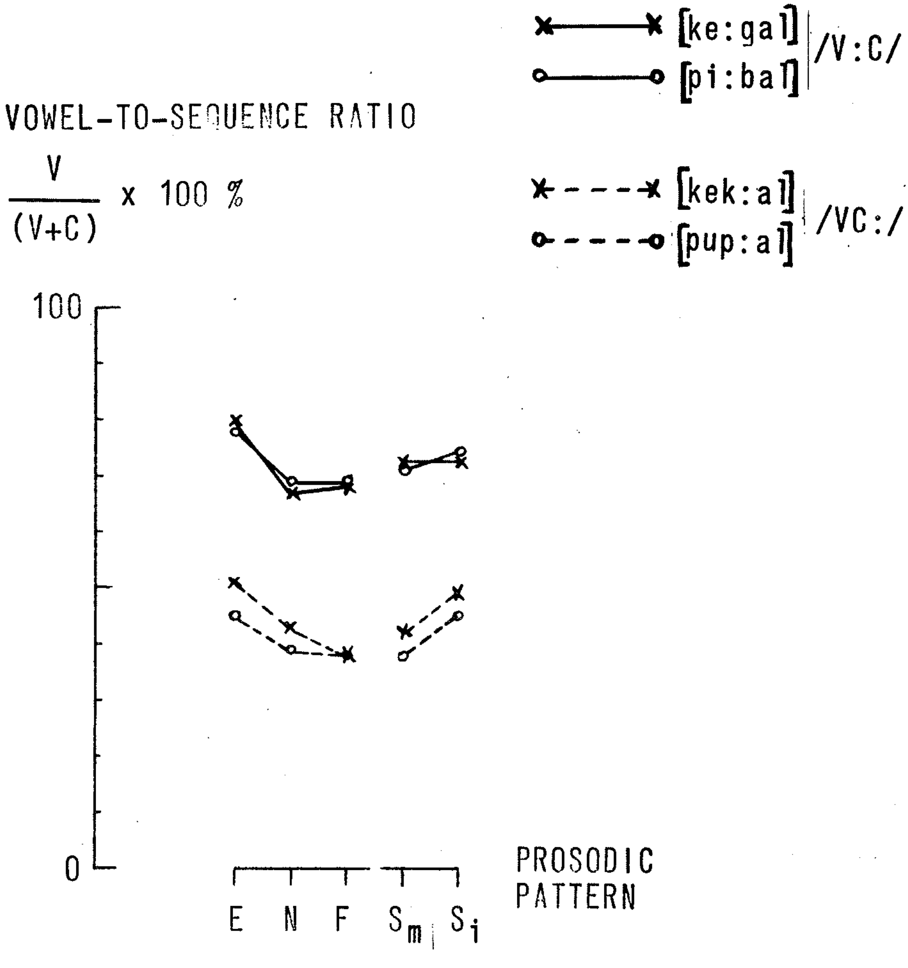


Figure 3. Vowel durations in relation to the duration of the entire vowel-consonant sequence, expressed by the vowel-to-sequence ratio in %.

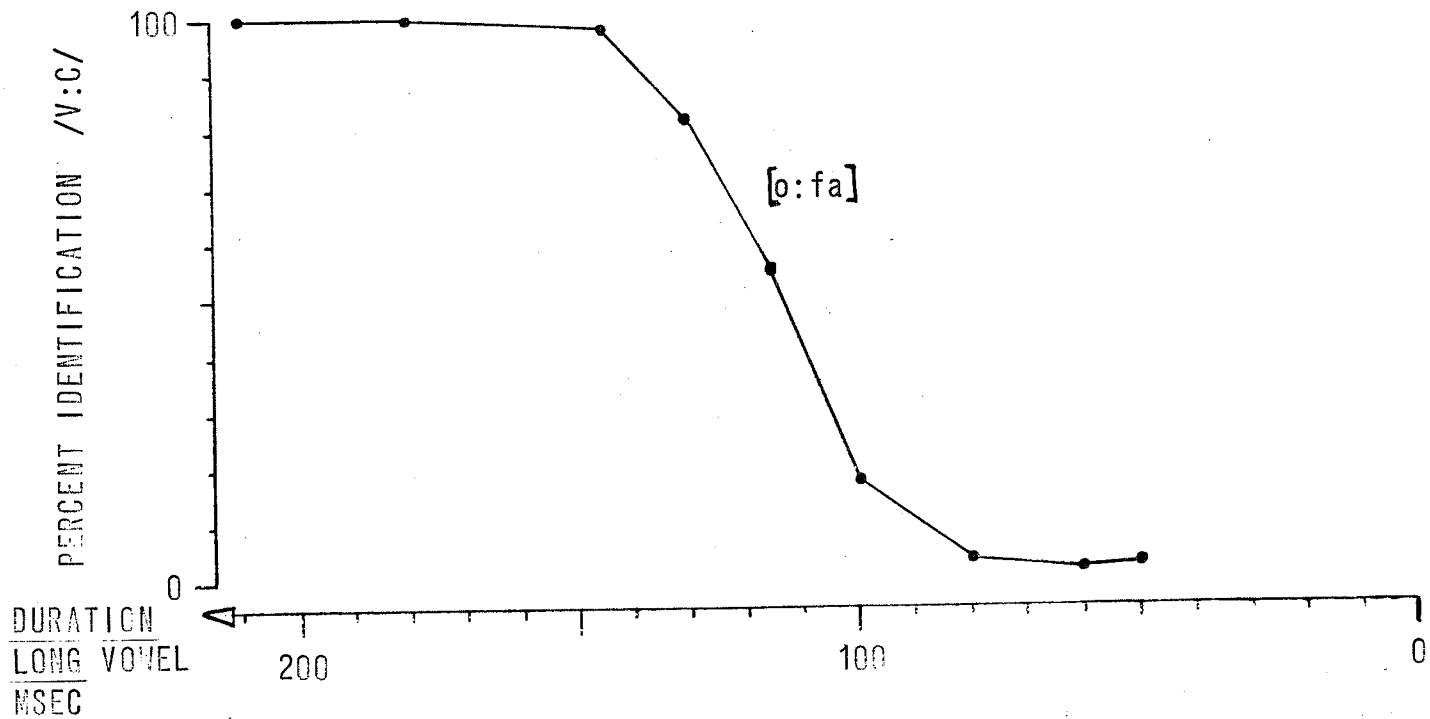


Figure 4. Identification of the word [o:fa] when the long initial vowel is shortened.
Ten listeners responded to each stimulus ten times each; a total of 100 responses.

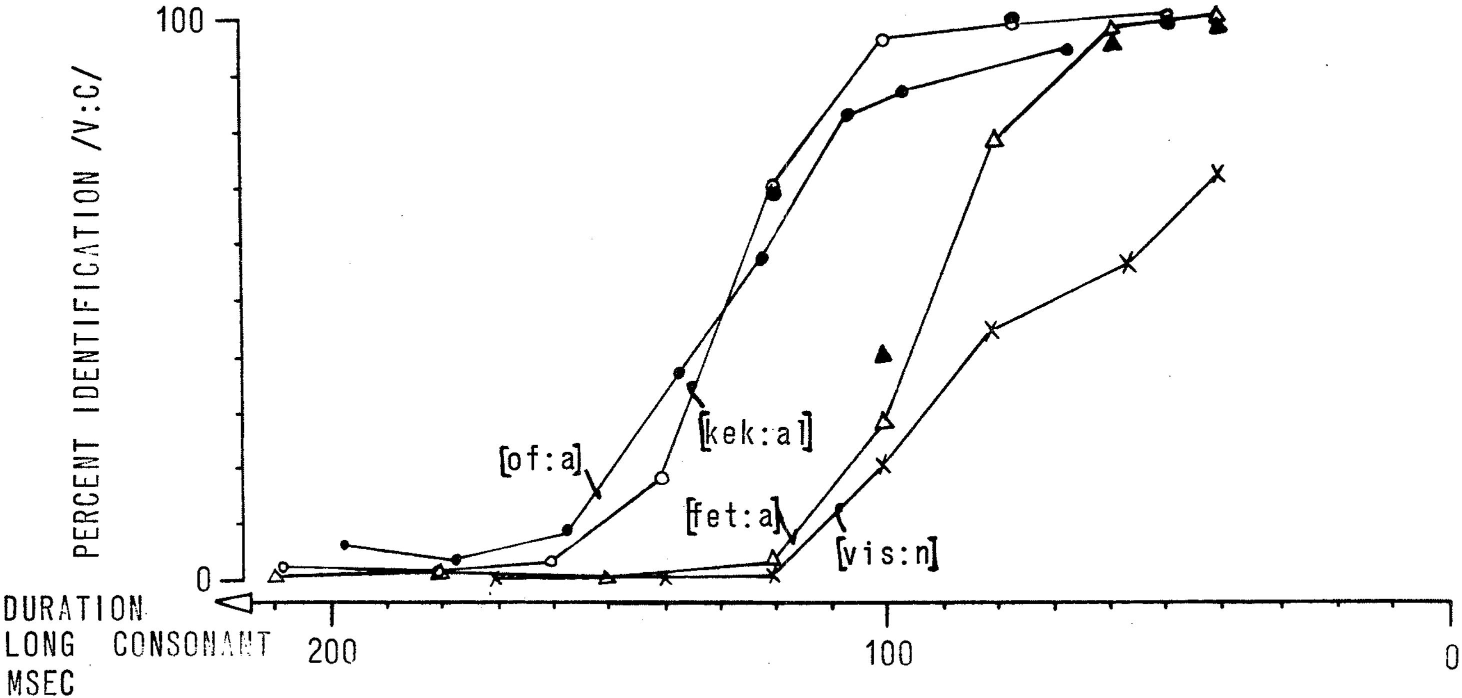


Figure 5. Identification of the four test words with long consonants when their durations are decreased.
Explosion eliminated in [fet:a] (▲) and [kek:a] (●).

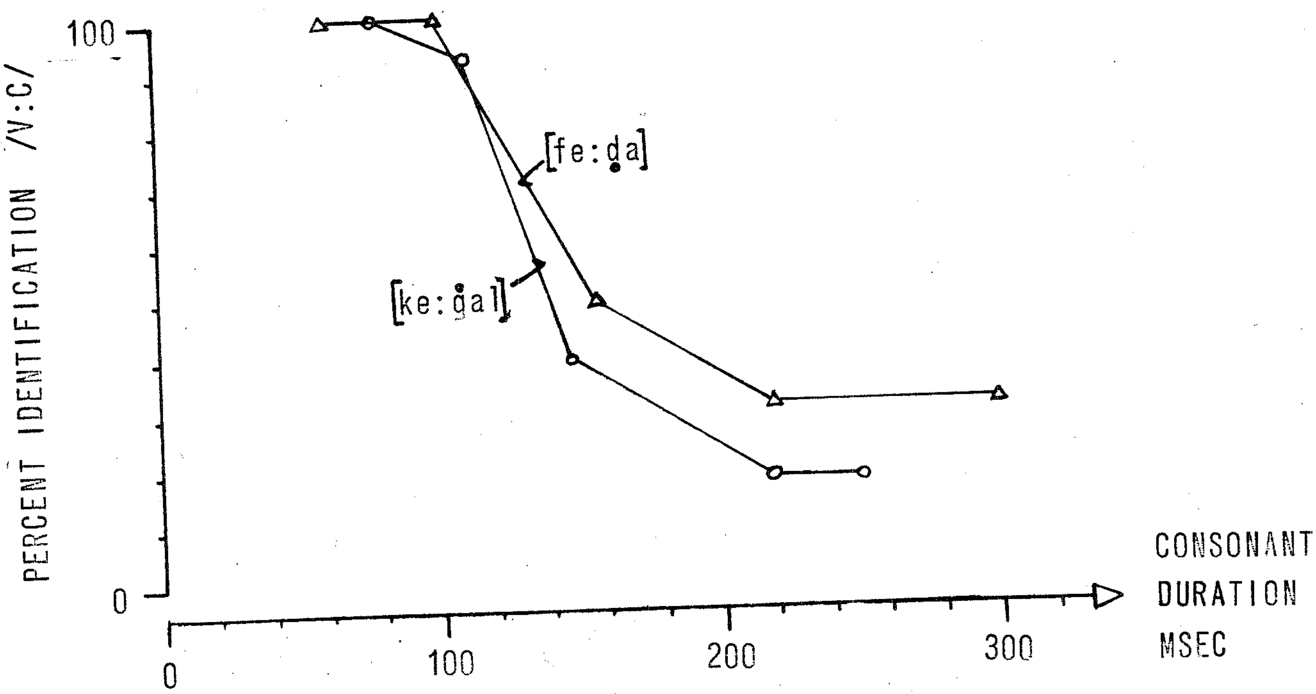


Figure 6. Identification of the words [fe:da] and [ke:ga] when the short stops are lengthened. The periodicity of the occlusion had been gated out previously.

Working Papers 12 · 1975 Phonetics Laboratory
Lund University General Linguistics

SWEDISH ACCENTS IN SENTENCE PERSPECTIVE *

Gösta Bruce

Introduction

The main purpose of this investigation is to determine how the fundamental frequency (f_o) contours of Swedish sentences vary with word accent (accent 1 or acute and accent 2 or grave) and focus. In accordance with current terminology, the term 'focus' is used to denote the new information in a sentence (cf. Jackendoff 1972). The old information in a sentence is said to be out of focus or is simply called non-focus.

The present report concerns two Swedish dialects - the Stockholm dialect (from Central Sweden) and the Malmö dialect (from the South). The main result is briefly that the f_o patterns of the accents in focus are specific and are only partly similar to the non-focus patterns.

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. The following set of test sentences has been used:

	1		2	3	
Man vill	{ `lämna / a´namma }	nåra	{ `långa / ´längre }	{ `nunnor / ´nummer }	.
(One will	{ leave / accept }	some	{ long / longer }	{ nuns / numbers }	.)

The numbers denote three possible focus locations. In each position there are words with accent 1 (´) or accent 2 (`).

---

* Paper presented at the Eight International Congress of Phonetic Sciences, Leeds, August 17-23, 1975

These words are disyllabic and stressed on the first syllable with the exception of the acute trisyllabic verb a´namma, which is stressed on the second syllable. 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-Peterson 1961), which might complicate the interpretation of the contours. All the vowels are phonologically short.

Results and discussion

The observations made here are based on the f_o tracings of Figs. 1-4. The f_o patterns of the accents in focus are in agreement with those shown by earlier investigations e.g. Meyer 1937, Malmberg 1959, Hadding-Koch 1961, Gårding & Lindblad 1973.

The Stockholm dialect

The f_o pattern of the accent 2 word `lämna in non-focus position is partly similar to the corresponding focus pattern (Fig. 1). We find an f_o maximum in the stressed vowel and then a fall. The rise of the second unstressed syllable of the word when in focus is totally absent when out of focus. Instead we find a low f_o. The accent 1 word a´namma shows a similar beginning for both non-focus and focus position (Fig. 1). We find an f_o peak in the pretonic syllable and a fall in the prevocalic consonant to a f_o minimum, which is reached in the beginning of the stressed vowel. The rise of the stressed syllable and the following peak - features which are regarded as typical of accent 1 - are found only in focus position. In non-focus position the low f_o is maintained throughout the word. For both accent 1 and accent 2 it seems as if the rise in focus position starts before the target f_o minimum is reached. Corresponding differences between focus and non-focus positions are also found for the pairs `långa/´längre and `nunnor/´nummer (Fig. 1).

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

In sentence final position for both accent 1 and accent 2 words in focus the focus peak is sharp as compared with non-final position (Fig. 2). In non-final position there is no immediate f_o fall following the f_o rise as in final position. The peak is flatter and when there are several unstressed syllables between the focus syllable and the following stressed syllable - as in the first focus position - the peak between the focus rise and the fall of the following stressed syllable becomes a plateau.

The Malmö dialect

For the accent 2 word ˋlämna there are obvious similarities between focus and non-focus position (Fig. 3). We find a rise in the stressed vowel, and the peak is reached in the postvocalic consonant. The difference lies in the unstressed post-tonic vowel. In focus position there is a steep fall, which is not present in non-focus position. In the accent 1 word aˊnamma we observe a rise in the prevocalic consonant and a peak in the stressed vowel (Fig. 3). In focus position this peak is immediately followed by a steep fall beginning in the middle of the stressed vowel and reaching its minimum in the post-tonic vowel. In non-focus position there is instead a gentle f_o fall in the post-tonic syllable. For the pair ˋlånga/ˊlängre (Fig. 3) there are corresponding differences between focus and non-focus position, but for the pair ˋnunnor/ˊnummer (Fig. 3) the difference is not evident. It appears that the f_o pattern of the accent 1 word ˊnummer in focus, when regarded in isolation, is almost identical to one of the non-focus patterns. This is due to a fall, which is present in sentence final position independently of focus location and therefore erases a potential difference between focus and non-focus. There is, however, a difference of transition from the preceding word (Fig. 4). When the preceding word is in focus the transition is a deep trough, as a consequence of the focus fall. When the following word is in focus, however, the transition is relatively shallow. This difference of transition is evident also from the f_0 patterns of the accent 2 word ˉnunnor (Fig. 4).

In contrast to the Stockholm dialect we find a fall as the f_0 manifestation of focus in the Malmö dialect. This fall is independent of accent, but the timing of the fall is later for accent 2. This pattern of focus manifestation is perfectly parallel to the Stockholm pattern. Thus in both dialects we have major f_0 change as the reflex of focus. The difference is in the direction of the change - a rise in Stockholm corresponds to a fall in Malmö (Fig. 5). In non-focus position there are no such changes. By shifting the focus to different positions of a sentence in the way described in this paper, it has been possible to separate the contribution of focus to the f_0 contour, thereby isolating the basic word accent patterns.

References

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

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

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

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

Malmberg B. 1959. Bemerkungen zum schwedischen Wortakzent. Z. Phonetik 12

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

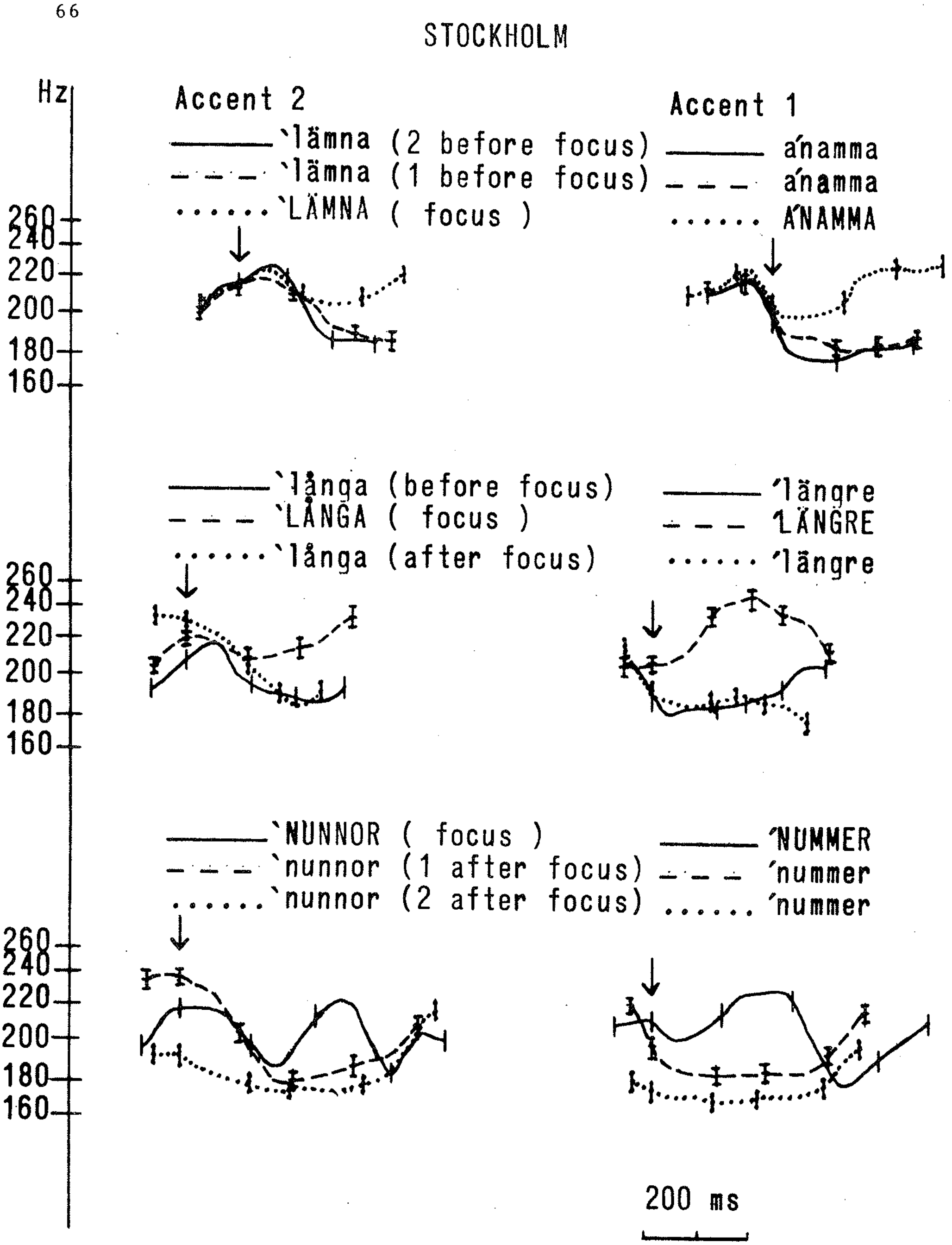


Fig 1. Accents in focus and out of focus. F_0 tracings of words derived from full sentences. Vertical bars indicate segment boundaries. Line-up poir (arrow) is at CV-boundary of stressed syllabl

STOCKHOLM

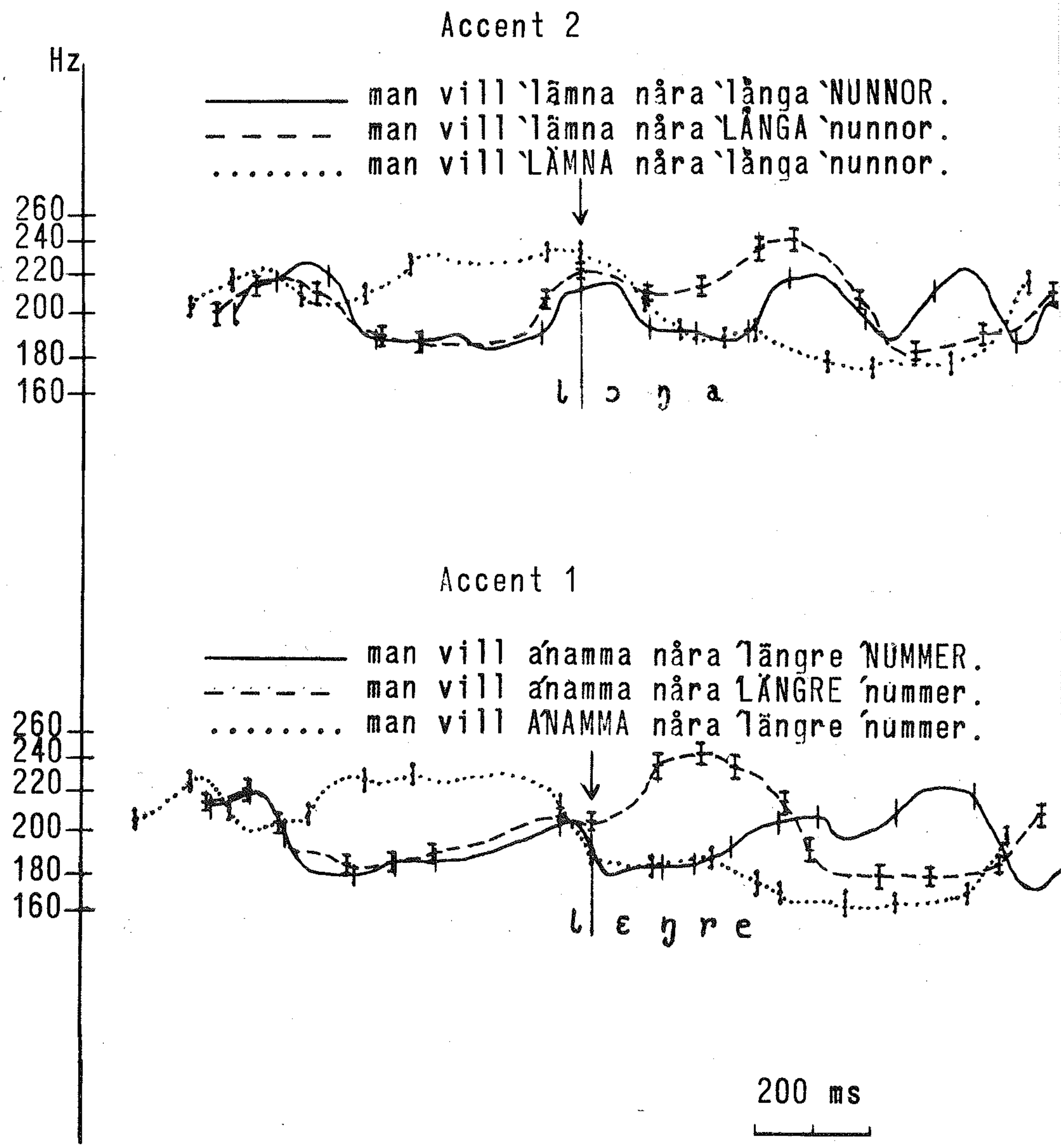


Fig 2. Accents in focus and out of focus. F_0 tracings of full sentences. Focus is indicated by capital letters.

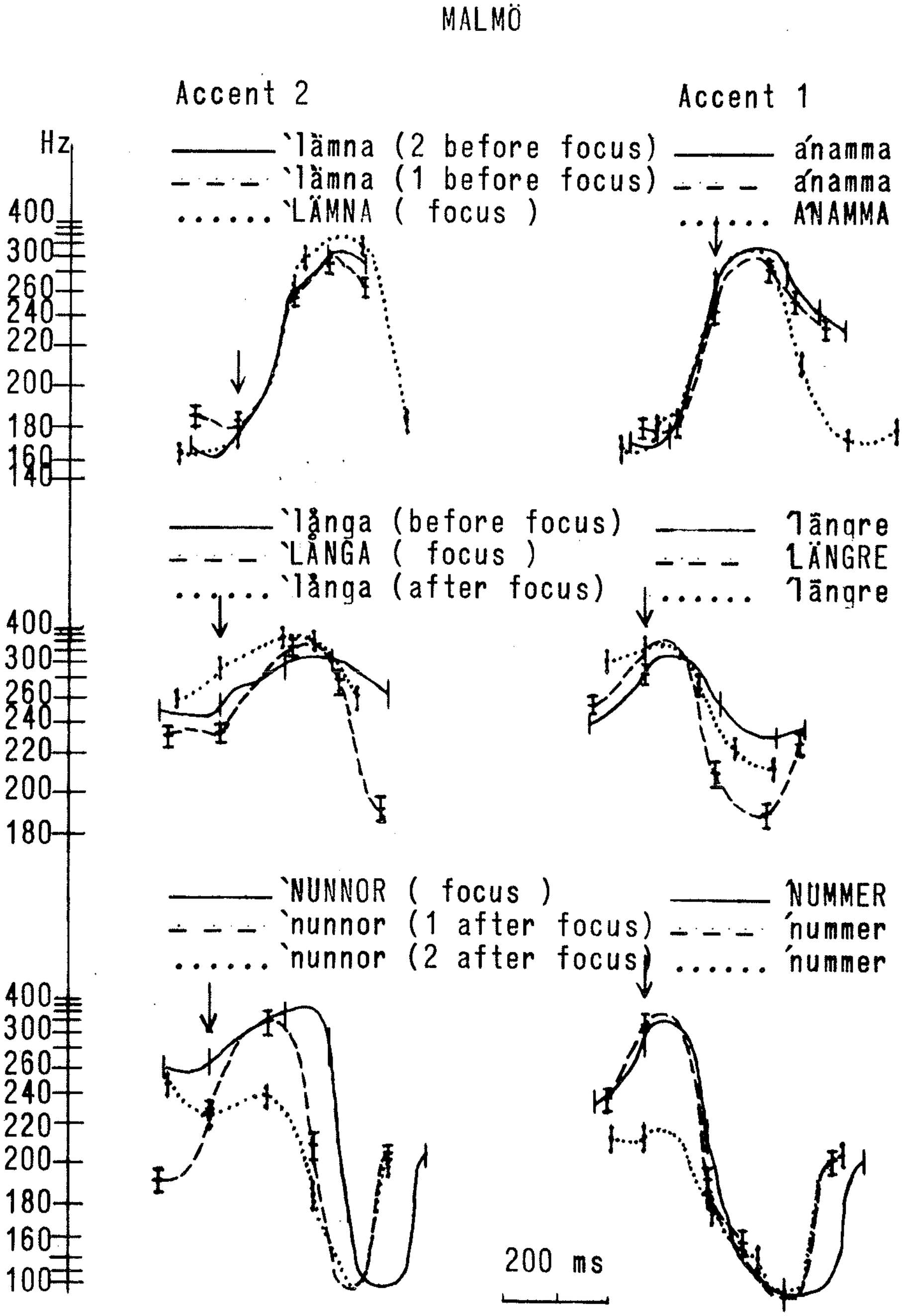


Fig 3. Accents in focus and out of focus. F_0 tracings of words derived from full sentences.

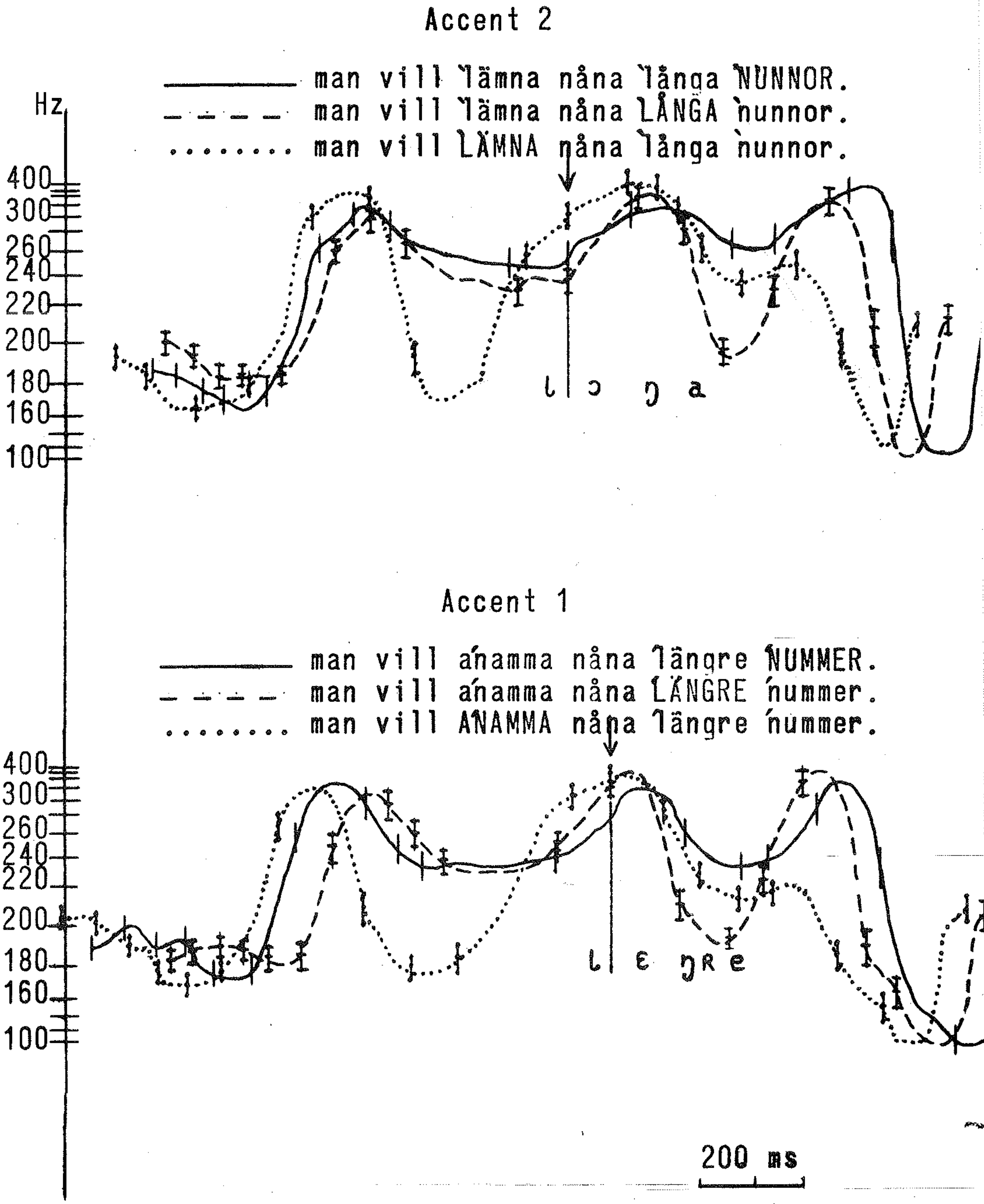


Fig 4. Accents in focus and out of focus. F_0 tracings of full sentences.

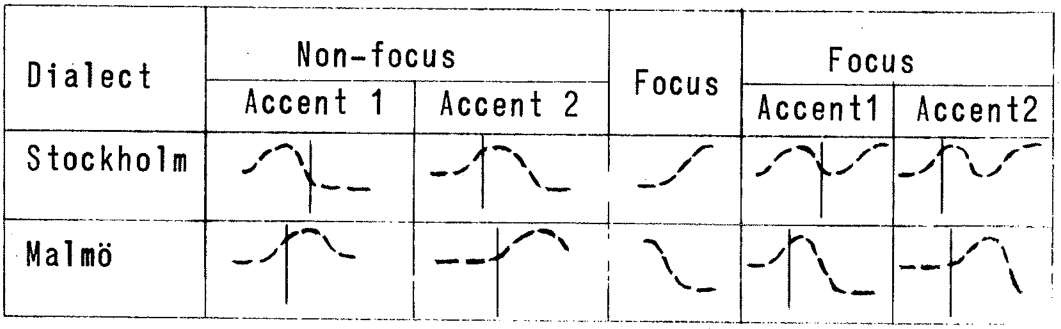


Fig 5. Constructive elements of the accents.

Working Papers 12 · 1975
Lund University
Phonetics Laboratory
General Linguistics

THE INFLUENCE OF TEMPO ON RHYTHMIC AND TONAL PATTERNS IN THREE SWEDISH DIALECTS*

Eva Gårding

Introduction

In The Scandinavian Word Accents (1973a) I proposed a typology for dialectal manifestations of accents. The Swedish part was based on Meyer's data (1937, 1954) and on more recently collected speech materials (1973b). Since the location of the fundamental frequency peaks plays a major role in the typology it is important to explore how it is affected by speech variables such as sentence stress (Bruce 1975) and tempo. Tempo was one of the variables in an EMG investigation dealing with accents (Gårding et al. 1975) and is the main concern of this paper.

My material consists of a number of noun phrases with two accented words. They have been uttered as statements in three kinds of tempo which I have labelled Slow, Normal and Fast. Let me give you one example.

Figure 1a shows average segment durations and fundamental frequency curves for the phrase en manli(g) nunna. The phrase contains two grave accents and means a masculine nun. It has been uttered by a Central Swedish speaker. My other speakers are from the South and from Finland (Figs 1b and 1c). It should be remembered that the Swedish spoken in Sweden is characterized by two accents with different tonal manifestations according to dialect. The Swedish spoken in Finland on the other hand has no accent contrast.

* Paper presented at the Eighth International Congress of Phonetic Sciences, Leeds, August 17-23 1975

In this report I shall study changes with a shift of tempo in segment durations (1), fundamental frequency (2) and in the coordination between fundamental frequency and segments (3). On the basis of my results I shall answer the following question: What change is needed in a generative program for rhythm and intonation to go from Normal to Fast and from Normal to Slow (4).

1. Segment durations

A first step in my analysis is to look at the relative changes in the duration of the segments. In Figure 2 the average utterances have been scaled to a common duration of one second. If the speaker had used a uniform stretching and shrinking of the utterance - or part of it - the diagram would have had parallel lines between the corresponding segments. As can be seen in the diagram there are some lines which are not parallel. They indicate deviations from uniformity and I shall now comment on some of them.

One such feature, which is common to all three speakers, is that the rhythm changes somewhat with tempo. In fast speech the last word occupies a larger part of the whole utterance than in slow and normal speech. This effect is strongest for the Central Swedish speaker. It is brought about by a relative shortening of the consonantal segments and the unstressed vowels of the early part of the utterance.

The larger effect of this rhythmical change may be caused by the more complex tonal pattern in the final grave accent of this dialect - two peaks versus one for South Swedish and East Swedish. This complex pattern may demand more time and hence longer segments.

There is a noticeable deviation from uniformity in the last segment for the Central and South Swedish speakers. In their speech the relative duration of the terminal vowel increases with tempo. This means in absolute terms that the duration of the final vowel is almost constant all through the tempi and therefore has a larger part of the fast-speech utterance. For the East Swedish speaker on the other hand there is a uniform tempo variation in this segment. One possible explanation is that the East Swedish speaker marks statement intonation by giving a gradual fall to the whole contour (Fig. 1c). The Swedish speakers also mark statement intonation by a fall but it is concentrated to the last syllable (Figs. 1a and 1b). It could be argued then that this syllable needs a certain duration to carry the fall.

2. Fundamental frequency

The second step in my analysis is to bring the fundamental frequency curves to a common time scale. There is no need for any change in the fundamental frequency since the pitch range seems to vary very little with tempo (cf. Gårding et al. 1975).

Figure 3 presents the normalized slow and fast-speech utterances superimposed on each other. There is one striking difference between the speakers. The East Swedish curves are almost the same. the others are not.

A reasonable guess is that this difference has to do with the accents which as I have said earlier are lacking in East Swedish. I shall come back to this later.

3. Coordination of segments and fundamental frequency

So far I have regarded segments and fundamental frequency separately. I shall now take the third step in my analysis and talk about their coordination.

For this discussion I shall introduce the term fix point meaning a point on the fundamental frequency curve which retains its relative time position in a segment regardless of tempo (Fig. 4).

It is clear that with a uniform change of segment durations and fundamental frequency curves there will be fix points all the way. This is in fact what happened in my East Swedish material. Here (Fig. 3) you can see that from the stressed syllable (which is marked by an arrow) onwards there is roughly speaking nothing but fix points. The location of the peaks over the stressed syllables in slow speech matches with the corresponding peaks in fast speech, etcetera. But what about the Central and South Swedish curves? They are not uniform but there are fix points in some crucial places. Figures 1a and 1b show that

(1) the peaks of the accented syllables are fix points and that

(2) the peak of the final syllable marking sentence intonation is also relatively fixed.

4. <u>What change is needed in a generative program for rhythm and fundamental frequency when tempo varies?</u>

I shall now try to answer my introductory question.

Let us first briefly consider a basic program that would generate the normal-rate tonal and rhythmic patterns for the East Swedish utterances. Here the fundamental frequency curve can be derived if we have certain information: phrase boundaries, type of sentence - in this case statement - location of stresses and duration of segments. To get fast and slow speech from this we just have to make a uniform change of the segment durations and of the fundamental frequency curves.

For the basic normal-rate patterns of the speakers of Central and South Swedish we need more information. We must know the accents, their shapes and fix points. The tempo as before determines the duration of the segments rather uniformly for South Swedish and the spacing of the accent shapes.

Figure 5 illustrates this situation for the South Swedish speaker. The accent shapes are here triangles. They represent the fundamental frequency manifestation of a laryngeal manoeuvre, the response to the accent command. The bases of the triangles in the schematic curves are constant which reflects the fact that the corresponding parts of the observed curves are fairly constant. This indicates that the same laryngeal manoeuvre is retained regardless of tempo. The main difference between the three tempi is that the triangles are spaced differently. The points of the triangles however, are fixed to the segments according to certain rules which are different for the two accents. In slow speech the triangles are separated and the spacing appears as a flat uneventful portion in the curve. In fast speech the triangles come closer to each other. The segments that in slow speech are stretched out in the flat portion of the curve have now been pushed up into the sides of the triangles.

This interpretation shows that the change of program for South Swedish when we go from one tempo to another is not much more complicated than for East Swedish. We make an almost uniform change of the segment durations but we keep the accent shapes with the peaks as fix points. Possible blanks and overlaps will then be filled out by interpolation rules. For Central Swedish we can use the same procedure. But here we must also have a rule that accounts for the deviation from uniformity in the rhythmic pattern.

I shall finally add some physiological comments to my analysis which so far has been rather descriptive. Prosody and articulation are performed by different sets of muscles innervated via different pathways of the nervous system. Hence it seems very improbable that centrally governed articulatory and phonatory events can follow each other in detail. Another factor of importance is that the laryngeal mechanism including the cricothyroid muscle is slower than at least some of the articulators.

In the light of these observations we can interpret our fix points as points of nervous coordination between articulation and intonation. It is then very natural that they should remain fairly independent of tempo. The behaviour of the accents in Central and South Swedish can therefore be regarded as a consequence of the greater inertia of the laryngeal mechanism as compared to the articulatory mechanism. The almost tempo independent covariation between fundamental frequency and articulation that we have observed in East Swedish may be due to the rather flat intonation where the peaks over the stressed syllables have no clear bases.

References

Bruce G. 1975. Swedish accents in sentence perspective. Paper presented at the Eighth International Congress of Phonetic Sciences, Leeds 1975

Gårding E. 1973a. The Scandinavian Word Accents. Working Papers 8, Phonetics Laboratory, Lund University, 119

Gårding E. & P. Lindblad. 1973b. Constancy and Variation in Swedish Word Accent Patterns. Working Papers 7, Phonetics Laboratory, Lund University, 36-110

Gårding E., O. Fujimura, H. Hirose and Z. Simada. 1975. Laryngeal control of Swedish word accents. Working Papers 10, For Kerstin Hadding, Phonetics Laboratory, Lund University, 65

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

Meyer E.A. 1954. Die Intonation im Schwedischen II: Die norrländischen Mundarten. Studies Scand. Philol. No. 11. University of Stockholm

AVERAGE DURATIONS AND FUNDAMENTAL FREQUENCY CURVES FOR THREE TEMPI

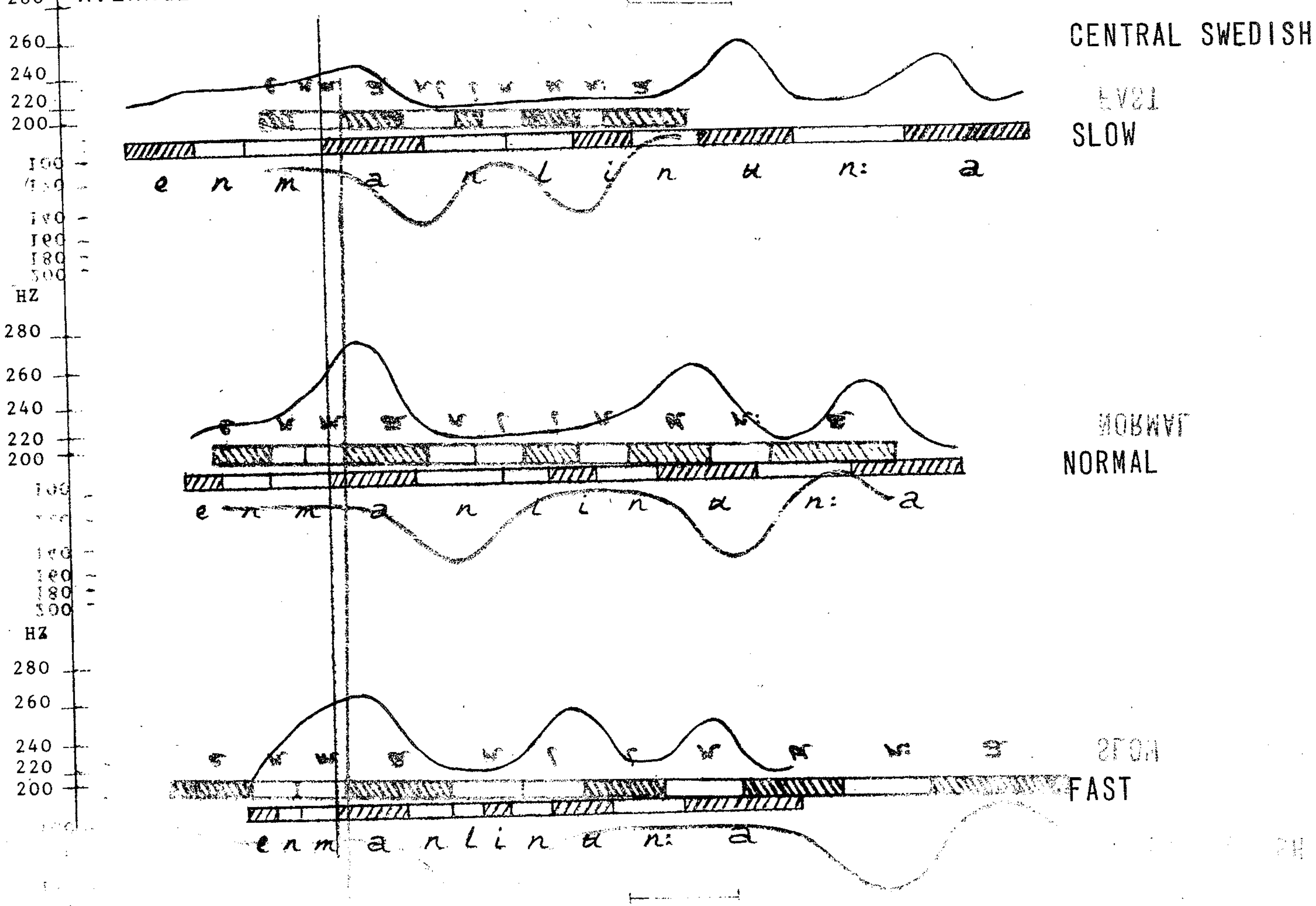


FIG 1 a

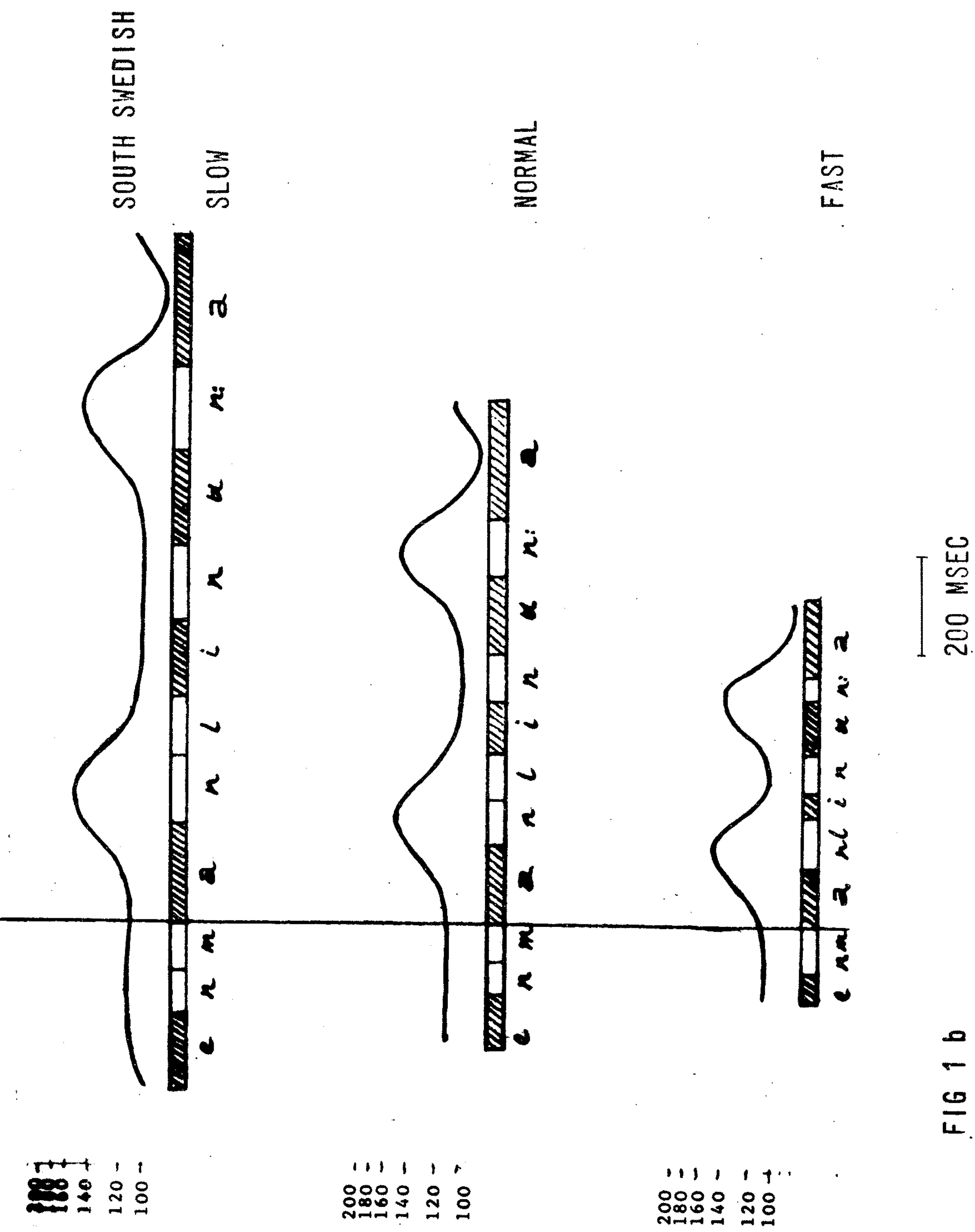

SOUTH SWEDISH
SLOW
NORMAL
FAST
200 MSEC
140
120
100
200
180
160
140
120
100
200
180
160
140
120
100

FIG 1 b

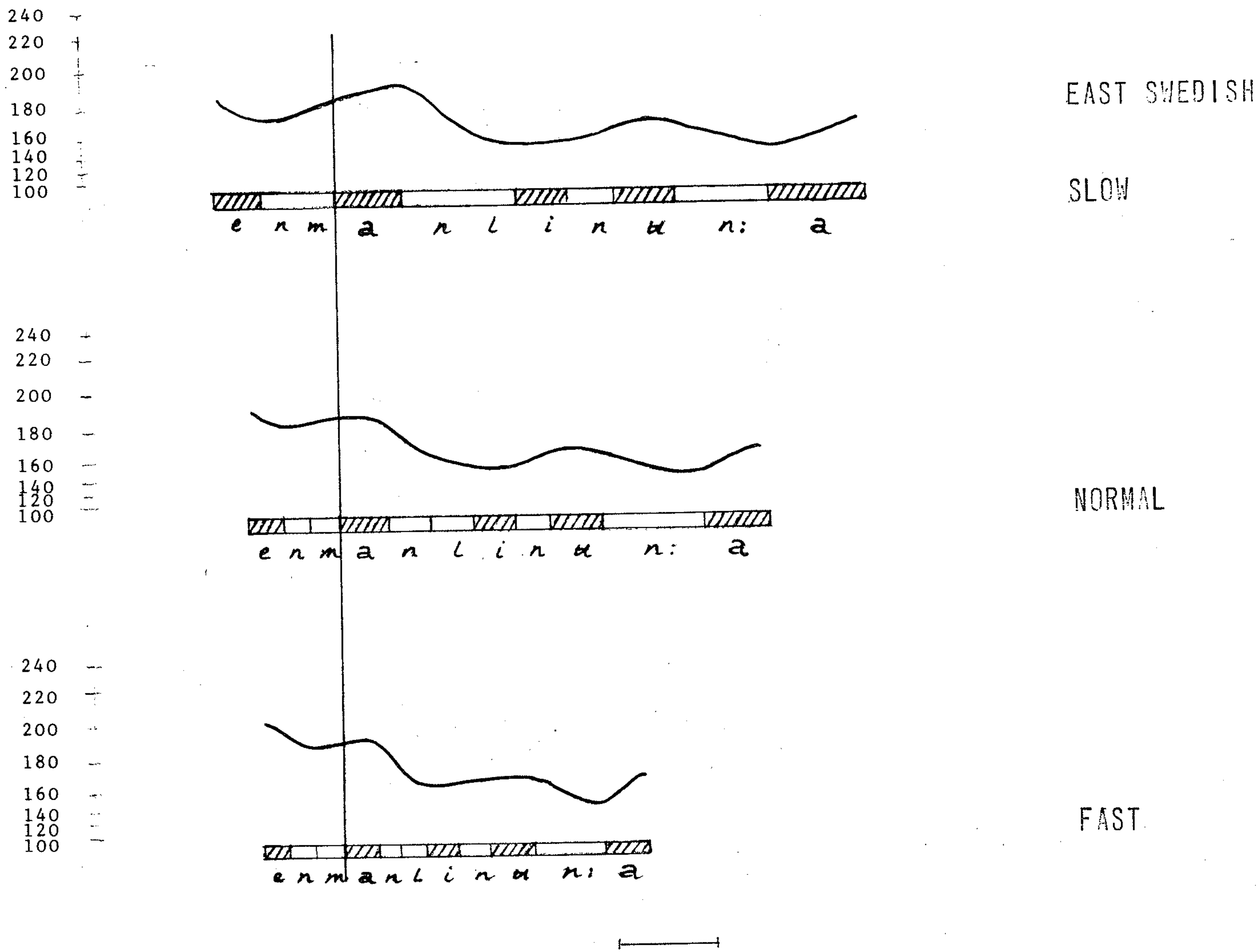

EAST SWEDISH
SLOW
NORMAL
FAST
240
220
200
180
160
140
120
100
e n m a n l i n u n: a
200 MSEC

FIG 1 c

RELATIVE DURATIONS

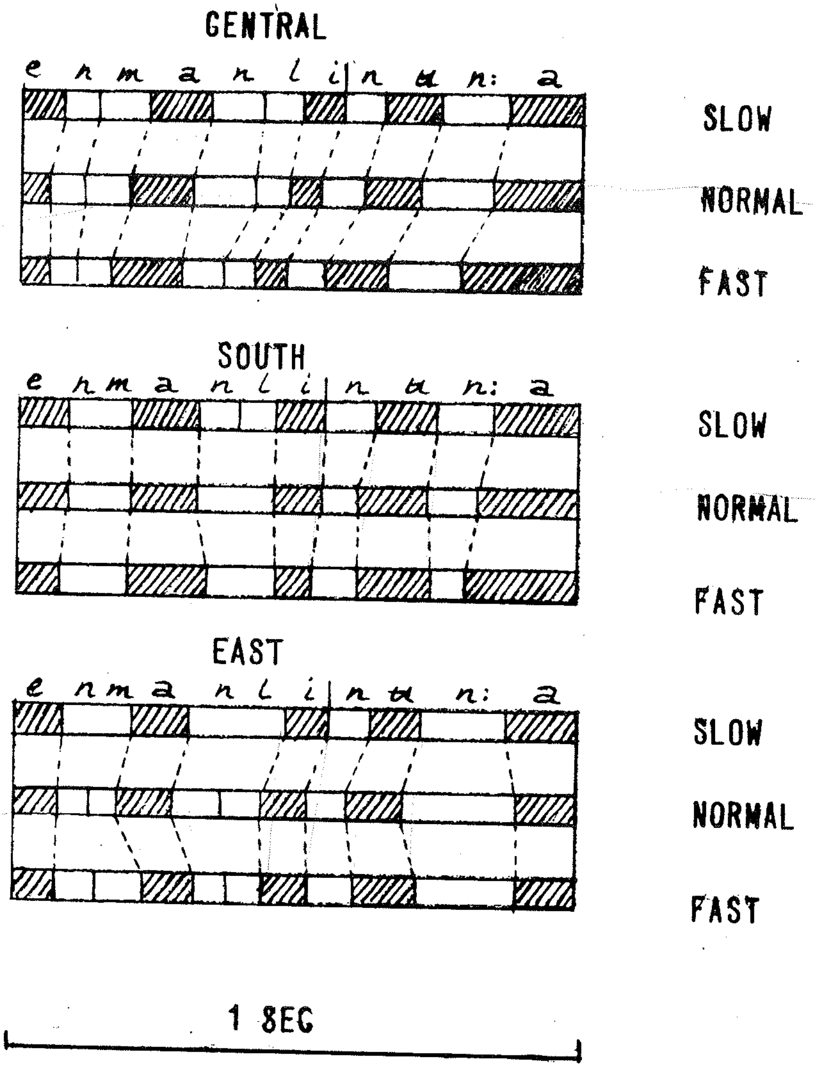


FIG 2·

RELATIVE FUNDAMENTAL FREQUENCY CURVES

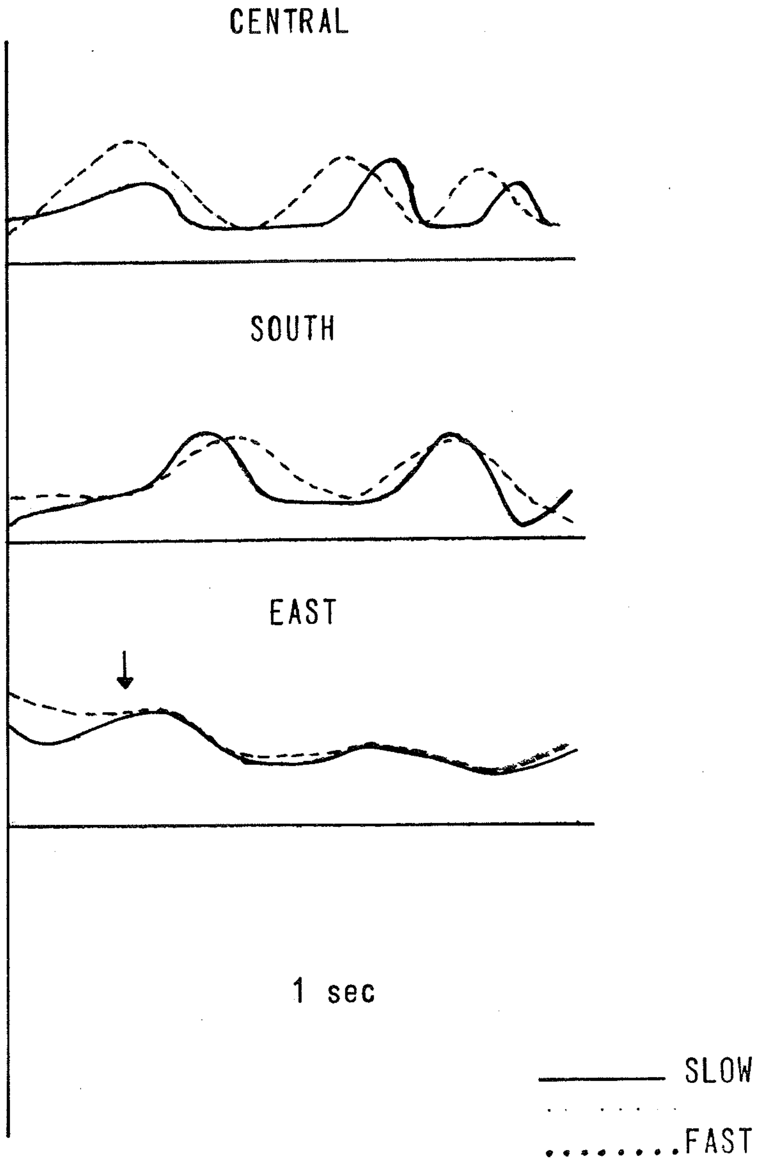


FIG.3

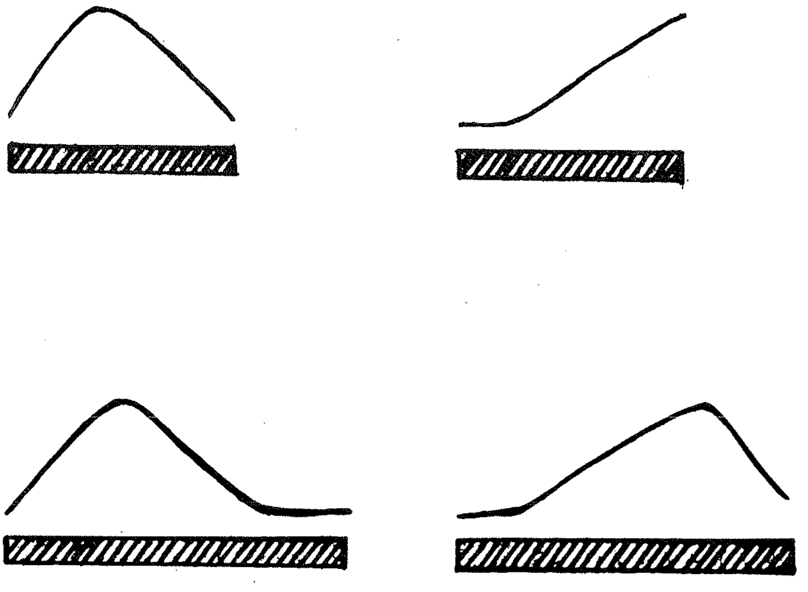


Fig. 4. ILLUSTRATION OF FIX POINTS

The peaks of the curves to the left are fix points but not the peaks of the curves to the right.

FUNDAMENTAL FREQUENCY VARIATION FOR THREE TEMPI

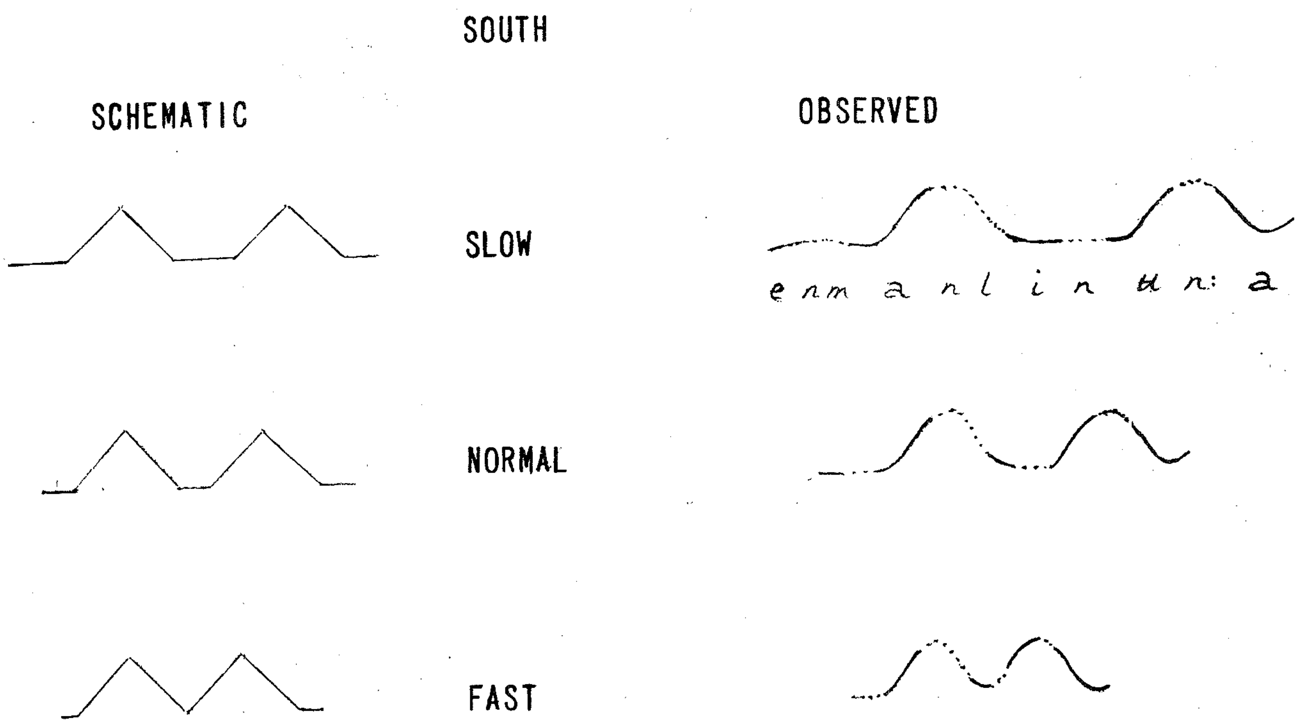


FIG 5

Working Papers 12 · 1975
Lund University

Phonetics Laboratory
General Linguistics

FUNDAMENTAL FREQUENCY OF THE VOICE IN CONTINUOUS SPEECH. – PRELIMINARY REPORT ON A DEVICE FOR DETERMINING MEAN AND DISTRIBUTION OF FREQUENCIES.

Peter Kitzing*, Hans-Erik Rundqvist** and Ewa Talo*

Abstract

There is no generally accepted and practicable method for quantifying average vocal pitch during continuous speech, even though this parameter is very often affected in different kinds of dysphonias and is the object of therapeutic measures in clinical work with voice disorders.

The present paper is a description of an apparatus for immediate presentation of the mean frequency and the distribution of frequencies of the fundamental of the voice during continuous speech. Some frequency distributions for normal and pathological voices illustrate the function of the apparatus.

There are various approaches to the study of the fundamental of the voice during continuous speech. Phoneticians are mostly interested in the linguistic implications of variations of pitch or intonation. On the other hand, the clinical interest of speech pathologists and phoniatricians is more often directed to vocal pitch as a correlate of laryngeal function. From this latter point of view it is of greater interest to obtain objective data about a person's total intonation

---

* Dept of Phoniatrics, ENT Clinic, University Hospital, Malmö.
** Dept of Electrical Measurements, Institute of Technology, Lund.
Address of corresponding author:
P. Kitzing, M.D.
Talvårdsavdelningen
Allmänna Sjukhuset
S-214 01 MALMÖ
Sweden

range and the mean and distribution of the fundamental frequencies of the voice during continuous speech rather than about language-dependent intonation patterns.

For phonetic research, many methods have been developed for extracting the fundamental of the voice from isolated sounds or continuous speech. Most of these methods consist of two steps. First, the vibratory frequency of the vocal folds or the corresponding fundamental frequency of the voice is recorded in some way. Then the recorded data are measured, which is usually a laborious and time-consuming procedure. This is the case for high-speed cinematography, for glottograms, for oscillograms and for sound spectrograms.

One very common method is to use a frequency meter to produce an intonation curve. Such meters have been described by Grützmacher and Lottermoser (1937), Risberg (1962) and Svend Smith (1968), all having in common a low-pass filter in their initial step. The bandwidth of this filter must be selected individually for different speakers, in order to pass the highest occurring fundamental frequencies while at the same time excluding the harmonics. This is not always feasible, so that measurements with this kind of frequency meter are liable to be somewhat uncertain (Fry, 1974). Moreover, the manual calculation of the mean fundamental frequency and the frequency distribution from an intonation curve is very time consuming.

Another possibility is to use a digital computer to extract the fundamental frequency of the voice. Several different computer programs have been reported recently (e.g. Maksym, 1973).

In the field of speech pathology and voice disorders it has not been possible to use either the laborious or the sophisticated methods mentioned above for routine clinical work. Instead, clinicians have checked the performance of their patients with aid of tuning forks or some musical instrument (e.g. Böhme and Hecker, 1970). While it seems quite possible to obtain a valuable estimate of the voice range in this way, the method is obviously far too subjective to offer a reasonably safe estimate of the average fundamental pitch of the voice in continuous speech. On the other hand, the average fundamental pitch in continuous speech is one of the most important aspects of vocal function, and voice therapy often aims at normalizing this parameter when it is pathologically disturbed.

The present paper is a description of a specially designed apparatus for on-line determination of the mean and distribution of fundamental frequencies of the voice. It is easy to use, and its function will be described and illustrated with some distributions for normal and disturbed voices.

Method

The principal components of the preliminary version of the system appear in the block diagram, Fig. 1. The signal from a microphone (Sennheiser MD 421 N) is amplified and filtered in a low-pass filter (Fonema Type 0031) set to cut off the frequencies above the fundamental. The resulting signal is converted by a Schmitt trigger circuit into rectangular wave form. The periods are measured in a digital period meter and the frequency calculated as the reciprocal of the period.

The rectangular wave is passed to the mean fundamental frequency meter (with digital display) and to the frequency distribution processor (the results of which are displayed as a histogram on an oscilloscope screen).

In the mean fundamental frequency meter the number of cycles is counted during a phonatory time of twenty seconds. As voiced and unvoiced speech sounds alternate in continuous speech, it is essential that the device operates only when a voiced input signal is present. At the end of the processing the mean fundamental frequency appears on a three digit display.

The frequency distribution processor consists mainly of a number of memory cells, each representing a given frequency interval. In the present preliminary version of the apparatus the number of memory cells has had to be limited to nine. The frequency interval of each cell was chosen to be 6/7 of a whole tone, which means that in its present form it can cover a range of about 16 semitones. This range can be adjusted in six steps as appropriate for different voices by means of a special frequency range selector. Each incoming cycle is assigned to the relevant cell, according to its frequency. The number of cycles counted in each cell is scanned at the end of the processing, and the distribution of the frequencies (strictly, of individual cycle periods) of the fundamental of the voice in the analysed speech sample is displayed as a histogram on the oscilloscope screen. Each histogram column represents the frequency interval of the corresponding memory cell while its height is proportional to the number of cycles counted in that interval. (For technical details cf. Rundqvist, 1974.)

In practice, the mean fundamental frequency can be read off directly from the digital display. But the histograms produced by the present version of the apparatus cannot be used without further processing. In order to compare histograms obtained on different occasions, the number of cycles counted in each cell must first be expressed as a percentage of the total since the absolute numbers can vary from one sample to another.

The speech sample analysed was a paragraph of neutral test ("Nordanvinden och solen", The North Wind and the Sun) comparable to the well known "Rainbow" passage (Fairbanks, 1960) read at the natural intensity of quiet conversation.

Preliminary results

The function of the new apparatus can be illustrated with the following examples.

In normal male (Fig. 2a) and female (Fig. 3) voices there is a fairly wide range of fundamental frequencies extending over six or seven columns in the histogram (5 or 6 tones), not counting columns with height lower than 5 %. The histogram shows either one or two peaks, neither normally exceeding a height of 30 %. Intonation can of course be restricted volontarily and this is illustrated by a sample of artificially monotonous speech with almost 80 per cent of all cycles occurring within the same interval (Fig. 4).

Restriction of the intonation range may be one feature of functional voice disorders as can be seen in Fig. 5, which is the record of a case of phonasthenia before and after voice therapy. Before therapy, the fundamental frequency distribution extended over four columns with two of them exceeding 30 %. After therapy, the average fundamental pitch was lowered about one semitone and the frequency distribution widened by one interval with no column in the histogram exceeding 30 %. Similar results have usually been found in other cases of functional voice disorder before and after therapy.

In organic voice disorders the same type of voice changes may occur. This is illustrated in Fig. 6, recorded from a 32 year old male with a left vocal fold paresis of unknown etiology. His speaking voice sounded quite normal although somewhat weak, but there were typical voice breaks when coughing and laughing. Being a talented amateur singer, he suffered from the loss of his singing voice. But in spontaneous speech he was able to control his voice unusually well and so did not experience great discomfort from his paresis during normal conversation. Therefore, there was no need for voice therapy but, of course, he was re-examined at regular intervals. One year after the appearance of the paresis, laryngolocial examination revealed his left vocal fold had regained normal respiratory and vibratory function, and the patient reported a normalization of his voice when speaking as well as singing or shouting. His mean voice pitch was now perceptibly lower and in his vivid intonation he covered a wide frequency range. The course of events described is objectively preserved in the fundamental frequency record of his speech, where his average fundamental pitch is lowered by two semitones and his frequency range widened from four intervals to six (7–10 semitones), not counting columns containing less than 5 % (Fig. 6).

In the cases of functional and organic dysphonias reported above, and in other frequency recordings not reported here, the amount and nature of voice change has been checked auditively by listening to tape recordings without prior knowledge of the fundamental frequency analysis. The results of this subjective evaluation, which of course could be expressed only in a general description, were in good accordance with the numerical data obtained by the objective frequency analysis.

Finally, the reproducibility of the method is illustrated by the two examples from a healthy male subject shown in Fig. 2, with a time interval of about 20 minutes between recordings.

Discussion

Our original goal was to construct a mean fundamental frequency meter for continuous speech which would easily yield objective data on-line for daily clinical examinations of patients with voice disorders. It became clear during preliminary discussions that a display of the distribution of frequencies would also be of great interest. This opinion was substantiated after searching the literature on the fundamental frequency distribution of speech. This literature is quite sparse, probably because of the hitherto very laborious data-gathering procedure. We are aware of only four papers about fundamental frequency distribution in continuous speech, namely Smith (1955), Risberg (1961), Saito et al. (1958) and Pisani (1971).

According to the highly experienced voice therapists of the State Institute for Speech Disorders in Copenhagen, Smith and Lauritzen (Smith 1955), intonation is always changed in a characteristic way after pedagogical voice therapy for functional voice disorders. As illustrated in Smith's paper, Lauritzen was able to demonstrate these changes in diagrams of the frequency distribution from four patients with functional voice disorders. He obtained his original data on fundamental frequency of the voice by measuring an oscillogram curve at sampling intervals of 20 ms. Two characteristic changes after therapy are reported. Firstly, there is most often an influence on the fundamental frequency of the voice, the directin of which seems to depend on the kind of voice disorder and the "placing" of the voice by the voice therapist. Secondly, there is usually an extension of the fundamental frequency range of the voice, displayed by a "widening of the peak" of the distribution or by "emergence of two peaks" (quotations from Smith).

In a pilot study on the range and rate of change of fundamental frequency in continuous speech, Risberg (1961) published frequency distribution charts from three male and two female speakers reading Swedish or English texts. Fundamental frequency data were sampled manually at successive 25 ms intervals from a Mingograph oscillogram produced by a modified Grützmacher analyzer. The speakers were described as being trained, semitrained or untrained. In the published diagrams there seems to be a tendency for widened frequency range to be positively correlated with the degree of the speaker's training.

Our own preliminary results are in good accordance with the two pilot studies just quoted.

Pisani (1971) reports another modification of the Grützmacher frequency analyser. By the use of a multichannel analyser the statistical distribution of voice fundamental frequency can also be detected. The paper includes uncommented records of fundamental frequency distribution from the same subject reciting two different pieces of poetry.

Another type of voice frequency distribution analysis has recently been reported by Müller (1973). The method is said to be based on the filter bandwidth of the ear and measures the time during which certain trigger levels are reached in 24 frequency bands reanging from 20 to 15 500 Hz. In this way the entire acoustic speech signal is analysed at various trigger levels for amplitude distribution in each of the 24 frequency bands. As the method yields an analysis of speech rather than voice, the resulting records arise from joint influence of at least three parameters, namely voice fundamental frequency, formant regions and filter bandwidth. This complex origin of the records complicates the interpretation of the resulting curves. The analysed signal to a certain degree stems not only from the voice source but also from the resonances of the vocal tract. This will render a correlation of these curves to the subtle auditive criteria of functional voice disorders very difficult.

In all methods of analysis discussed so far, as well as in our own preliminary investigations, the original signal analysed has been acoustical. As mentioned earlier in the introduction, the detection of the voice fundamental in an acoustic speech signal is subject to risks of error, mostly arising from unsufficient separation of harmonics from the fundamental. However, experience from the detection of laryngeal vibrations with the aid of an electroglottograph has convinced us that these difficulties can be overcome. Therefore we have recently substituted a glottograph for the microphone in our fundamental frequency analyser, so that the input signal corresponds to the laryngeal vibrations and is not an acoustic speech signal.

The method just described seems to be a practicable and economic way of obtaining objective data about normal and pathological voice pitch in spontaneous speech. The relevance of these data must of course be further substantiated by correlating the new fundamental frequency analyses to the auditive evaluation of the same voice samples by a panel of experienced listeners. Furthermore, the influence on intonation range of different types of text for speech samples, and the speaker's age, sex and dialect must be investigated. Once these aspects are fully understood, the method is expected to become a useful tool in clinical work with patients suffering from voice disorders.

References

Böhme G. and Hecker G. 1970. Gerontologische Untersuchungen über Stimmumfang und Sprechstimmlage. Folia phoniat. 22, 176-184

Fairbanks G. 1960. Voice and articulation drill book. Harper & Row, New York, Evanston, and London. 2:d ed.

Fry D.B. 1974. Prosodic phenomena. In Malmberg B., Manual of phonetics. North-Holland/American Elsevier, Amsterdam. 3rd printing

Grützmacher M. and Lottermoser W. 1937. Uber ein Verfahren zur trägheitsfreien Aufzeichnung von Melodiekurven. Akust Z. 2, 242-248

Maksym J.S. 1973. Real-time pitch extraction by adaptive prediction of the speech wave form. IEEE Transactions on Audio and Electroacoustics, AU 21:3, 149-154

Müller G. 1973. Statistische Langzeitanalyse von Sprechstimmen auf Frequensgruppenbasis. Arch. Ohr.-, Nas.-, Kehlk.-Heilk., 204 151-160

Pisani R.U. 1971. A melody detector. Acustica 25:3, 179-182

Risberg A. 1961. Statistical studies of fundamental frequency range and rate of change. Speech Transmission Laboratory, Quaterly Progress and Status Report 4, 7-8

Risberg A. 1962. Fundamental frequency tracking. Proc. 4th Int. Congr. Phonet. Sci., 227

Rundqvist H.-E. To be published. Device for registration of mean value and distribution of fundamental frequencies of the human voice in continuous speech

Saito S., Kato. K. and Teranishi N. 1958. Statistical properties of the fundamental frequencies in Japanese speech voices. J. Acoust. Soc. Japan 14, 111-116 (quoted from Risberg, 1961)

Smith S. 1955. Om tonehøjde og maaling af denne. Nord. Tidsskr. Tale og Stemme 15, 57-69

Smith S. 1968. A valve device for difficult cases of F_0 recording. Folika phoniat. 20, 202-206

Kitzing

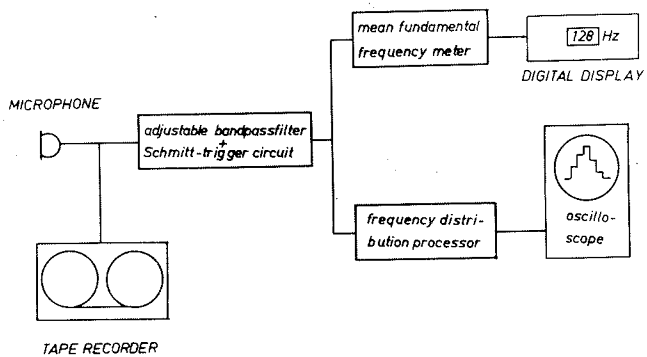


Fig. 1

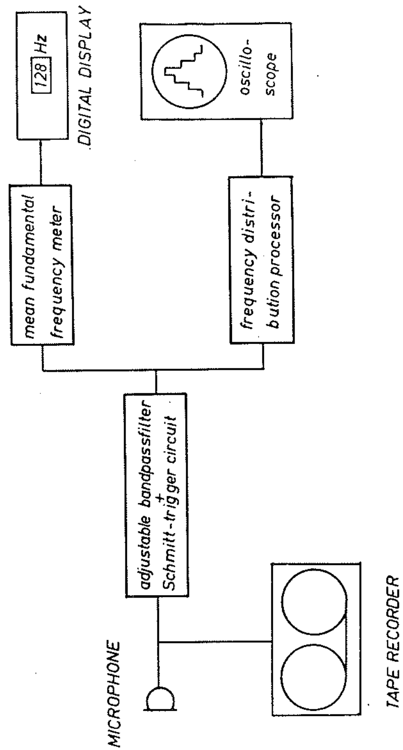


Fig. 1. After low pairs filtering to exclude harmonics, the voice signal is transformed by a schmitt trigger circuit into rectangular waveform with the same cycle period as the corresponding cycle of the voice fundamental. Further processing of the signal yields immediate digital display of the mean fundamental frequency (MFF) of the voice and a histogram of the distribution of frequencies on an oscilloscope.

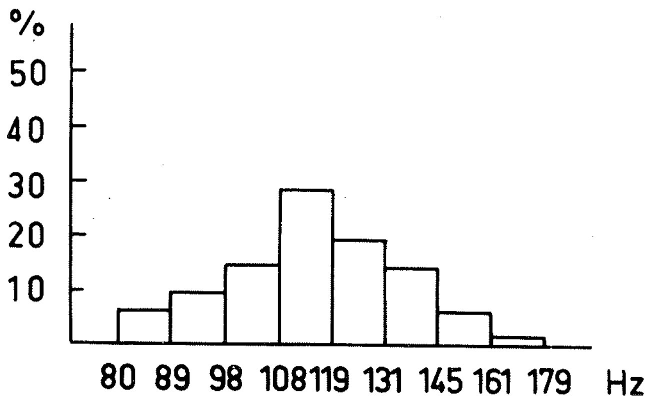


Fig. 2a

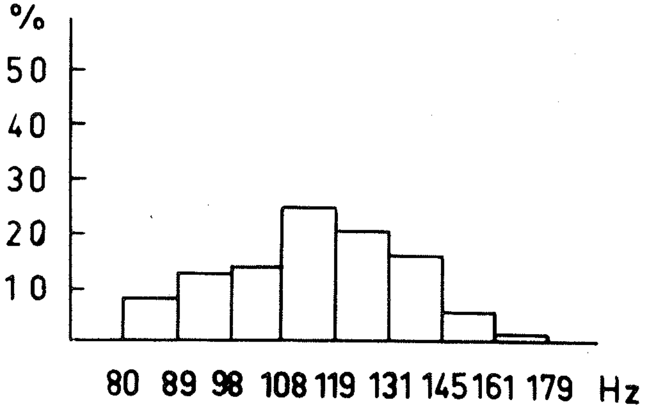


Fig. 2b

Fig. 2. a. Fundamental frequency distribution of a normal voice when reading a paragraph of neutral text in a natural way. The ordinate shows the number of cycles counted in each frequency interval expressed as a percentage of the total number of cycles counted in all cells. Male subject, aged 40. MFF: 112 Hz.

b. Distribution from the same subject 20 minutes later, demonstrating the reproducibility of the method. MFF: 110 Hz.

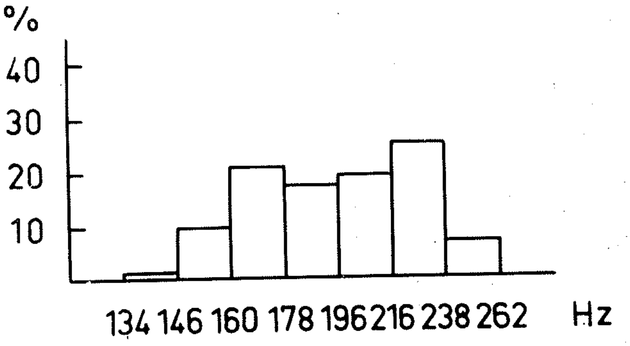


Fig. 3. Fundamental frequency distribution of a normal voice.
Female subject, aged 30. MFF: 192 Hz.

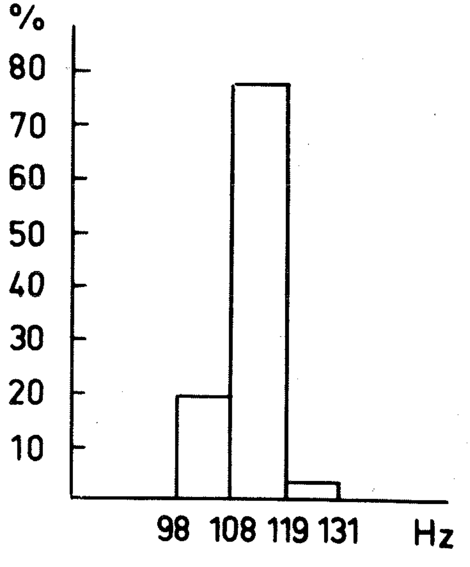


Fig. 4. Fundamental frequency distribution when reading in an artificially monotonous way. Male subject, aged 40. MFF: 109 Hz.

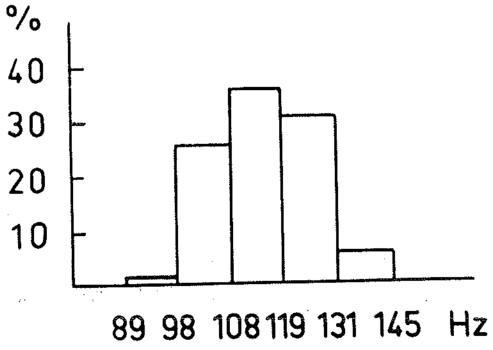


Fig. 5a

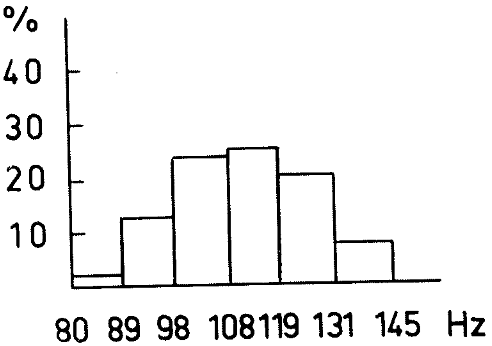


Fig. 5b

Fig. 5. Fundamental frequency distribution during continuous speech for a case of functional voice disorder (phonasthenia)
(a) before therapy. and
(b) after voice therapy and normalization of the voice function. Male subject, aged 32. MFF before therapy: 127 Hz; after therapy: 118 Hz.

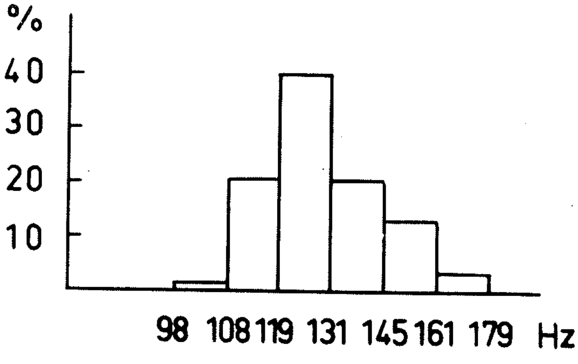


Fig. 6a

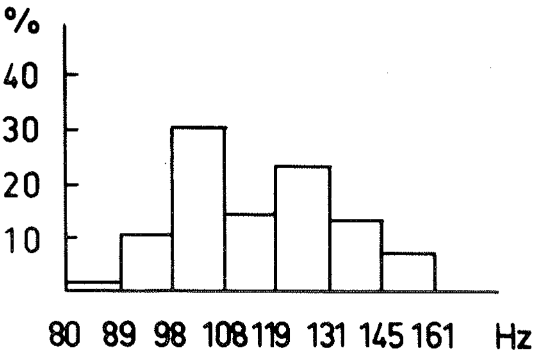


Fig. 6b

Fig. 6. Fundamental frequency distribution during continuous speech for a case of organic voice disorder (vocal fold paresis on one side)
(a) during paresis, MFF: 134 Hz, and
(b) after spontaneous recovery. MFF: 121 Hz. Male subject, aged 32

Working Papers 12 · 1975
Lund University

Phonetics Laboratory
General Linguistics

ON THE CONTROL OF ASPIRATION IN SWEDISH*

Anders Löfqvist

Swedish voiceless stops are aspirated in prestress and unaspirated in poststress position. (Aspiration will in the following be taken as the interval from stop release to the onset of glottal vibrations for a following segment.) This difference within the set of unvoiced stops is, however, not phonemic in Swedish but when it occurs aspiration serves as one of the cues for the distinction between voiced and voiceless stops since the former are always unaspirated.

Studies of subglottal pressure during the production of Swedish stops, Löfqvist (1975), reveal no difference in this parameter related to aspiration and the results are in agreement with a model of respiratory activity in speech which assumes that the respiratory system, ceteris paribus, generates a constant subglottal pressure irrespective of the presence or absence of aspiration after the stop release.

The control of aspiration would thus seem to depend on the coordination of glottal and supraglottal articulations. In this connection we can note that the difference in closure duration between Swedish voiced and voiceless stops is rather small except in those positions where the latter are unaspirated, Karlsson and Nord (1970), Löfqvist (1973); this is mainly due to an increase in closure duration for the voiceless set in these positions. From this we might hypothesize that the presence or absence of aspiration for the voiceless stops in Swedish is related to the duration of the oral closure.

This hypothesis is strengthened by the data given in Figure

---

*Paper presented at the 8th International Congress of Phonetic Sciences, Leeds, 17-23 August, 1975.

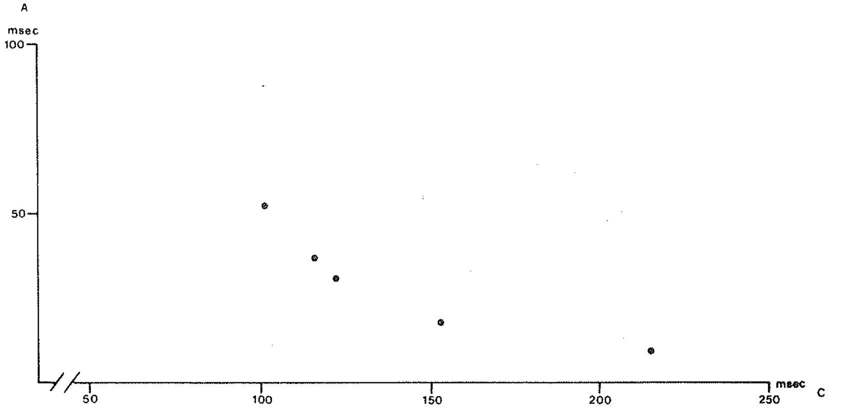


Figure 1. Aspiration (A) plotted against closure duration (C) for voiceless stops in various positions and under different stress conditions in bisyllabic words.

1, which shows the duration of the period of aspiration (A) plotted against the duration of the occlusion (C) for voiceless stops in various positions and under different stress conditions in bisyllabic words; each point represents the mean of 35 tokens and a negative correlation is apparent between aspiration and closure duration ($r = -0.9$). A simple model for the control of aspiration in Swedish voiceless stops based on these results would assume that the glottal gestures are invariant and that the duration of the oral closure controls the amount of aspiration. If the release occurs while the glottis is still open the stop will be aspirated, if it occurs when the glottis is already closed the stop is unaspirated and the degree of aspiration depends on when the release occurs in relation to the glottal gesture.

To get a picture of the glottal activity photoglottographic recordings were made during the production of the same test material, an example of which is shown in Figure 2. Let us for the moment concentrate on the timing of the glottal movements in relation to the supraglottal events and return later to the size of the glottal opening. The transillumination technique does not give an adequate measure of the degree of glottal opening since the amplitude of the signal depends, inter alia, on the relative positions of light source and light sensor and these positions change during the recording session; the temporal relationships in the curve appear, however, to remain stable irrespective of such changes. For the facts to be compatible with the model outlined above we would expect among other things that the interval from implosion to peak glottal opening remains stable across different positions. A plot of aspiration (A) versus the interval from implosion to peak glottal opening (T), Figure 3, shows no correlation ($r = 0.04$) and thus seems to be in agreement with the model as far as timing in bisyllabic words are concerned; a fair agreement can also be found with the material presented in Lindqvist (1972).

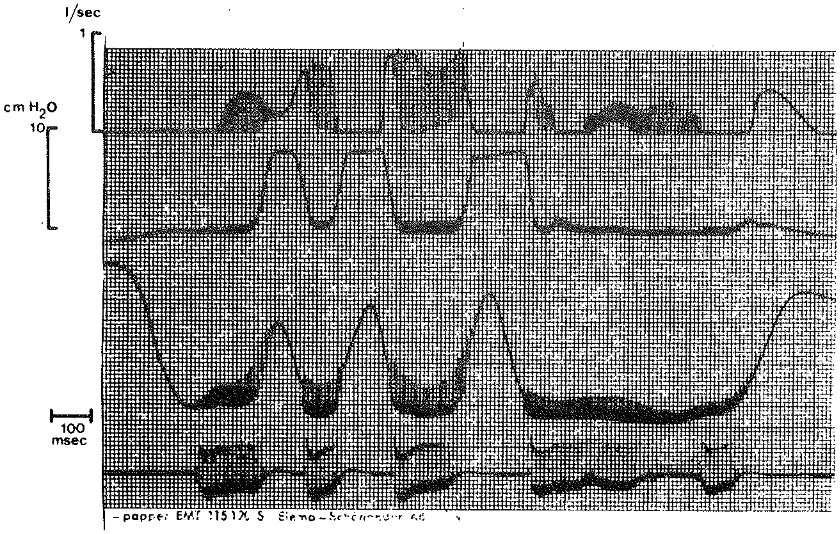


Figure 2. Record of the utterance "Men se 'teten igen". The curves represent from top to bottom, oral air flow, oral air pressure, photoglottogram and signal from larynx microphone.

If we try to extend the model to stops in words with more syllables it turns out, however, that the situation is not that neat. A plot of aspiration versus closure duration for stops in words with 2, 3 and 4 syllables, Figure 4, shows the same negative correlation ($r = -0.5$) but several cases can be seen where this relation does not hold. Furthermore, aspiration appears not to decrease beyond a certain limit and when this limit is reached a further increase in closure duration has no effect. At the same time the plot of aspiration against the interval from implosion to peak glottal opening for the same material, Figure 5, reveals a positive correlation ($r = 0.7$) indicating that peak glottal opening tends to occur earlier during the occlusion for the unaspirated and less aspirated stops.

The results obtained thus far suggest that several strategies can be used in the production of voiceless stops for the control of aspiration. One is to change the length of the closure period; in the unaspirated case an increase means that the glottis gets time to return to a position suitable for voicing to occur prior to the release of the oral closure. Another is to vary the moment at which glottal abduction and adduction occur and perhaps also the speed of the glottal movements. If peak glottal opening occurs late during the occlusion this ensures that the glottis is open at the release in the aspirated case and if it occurs early the glottis will be closed at the release in the unaspirated case; for the unaspirated stops the abduction may start during the preceding segment and thus cause the stop to be preaspirated. These two strategies seem to be combined in the production of Swedish voiceless unaspirated stops. A third strategy involves variations in the magnitude of the glottal opening; a reduction of peak glottal opening tends to occur in those positions where the stress pattern of the word makes the closure duration of the aspirated stops quite short.

As was mentioned above the transillumination technique can

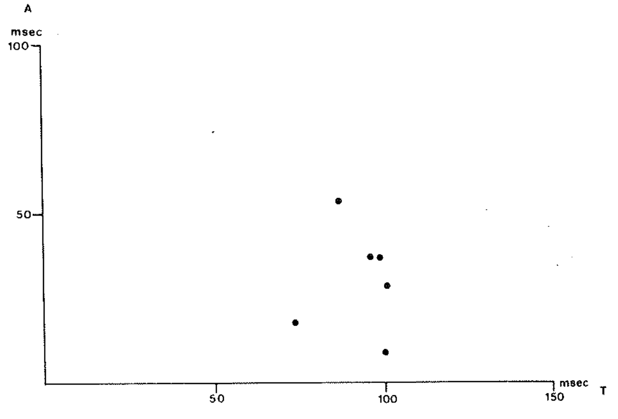


Figure 3. Aspiration (A) plotted against the interval from implosion to peak glottal opening (T) for voiceless stops in various positions and under different stress conditions in bisyllabic words.

not give a quantitative picture of the degree of glottal opening. Inspection of the records and comparisons of glottal opening for stops within the same test word do, however, suggest two general tendencies. Peak glottal opening is related to closure duration and increases with it. Glottal opening at the moment of release, suggested by Kim (1970) to be the determining factor for aspiration, is related to aspiration and tends to be larger the longer the period of aspiration. At the same time it should be noted that counterexamples to both these generalizations can be found.

The material discussed above suggests a framework for studying the relations between closure duration, voicing and aspiration in stop production and various strategies used for the control of aspiration in various languages.

References

Karlsson I. and Nord L. 1970. A new method of recording occlusion applied to the study of Swedish stops. STL-QPSR 2/3:8-18

Kim C.-W. 1970. A theory of aspiration. Phonetica 21:107-116

Lindqvist J. 1972. Laryngeal articulation studied on Swedish subjects. STL-QPSR 2/3:10-27

Löfqvist A. 1973. More on closure duration and aspiration for Swedish stops. Monthly Internal Memorandum, Phonology Laboratory, U C Berkeley, February 1973:27-32

Löfqvist A. 1975. A study of subglottal pressure during the production of Swedish stops. J. Phonetics 3:175-189

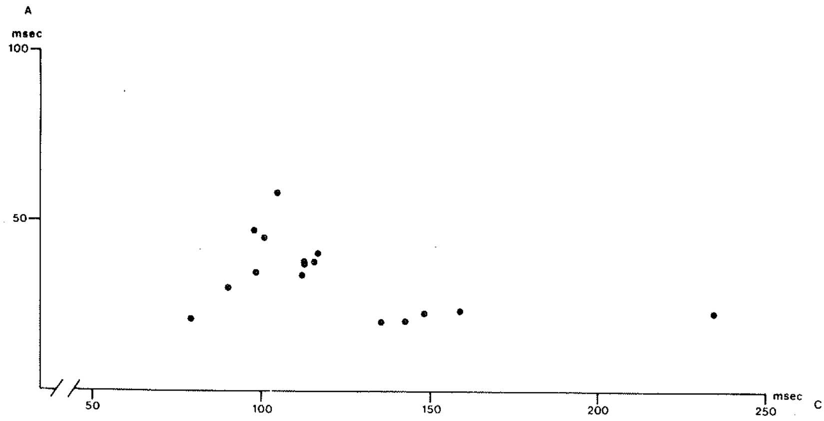


Figure 4. Aspiration (A) plotted against closure duration (C) for voiceless stops in various positions and under different stress conditions in words with 2, 3 and 4 syllables.

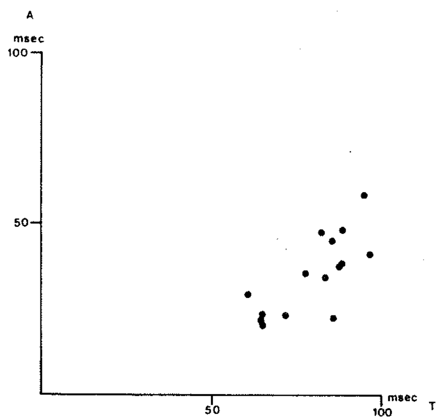


Figure 5. Aspiration (A) plotted against the interval from implosion to peak glottal opening (T) for voiceless stops in various positions and under different stress conditions in words with 2, 3 and 4 syllables.

Working Papers 11 · 1975
Lund University

Phonetics Laboratory
General Linguistics

Étude glottographique de quelques consonnes islandaises.

par

Magnús Pétursson

Introduction.

Les caractéristiques acoustiques et articulatoires des occlusives islandaises ont été étudiées par nous (Pétursson 1974a, passim et p. 191-205; 1974b p. 289-381) dans une série de recherches depuis 1967. Les résultats obtenus nous mettent aujourd'hui en position de décrire les consonnes occlusives islandaises avec une précision beaucoup plus grande qu'il y a quelques années seulement. En outre nos recherches ont permis de préciser de nombreux points obscurs dans les descriptions traditionnelles et elles ont apporté de solides arguments en faveur d'autres points de vue souvent opposés à ce qu'on a affirmé depuis longtemps. Résumons brièvement les faits tels qu'ils apparaissent à partir de nos recherches.

L'islandais moderne possède huit consonnes occlusives, une série de consonnes aspirées [p^h t^h c^h k^h] (orthographe p t k) et une série de consonnes non-aspirées [p t c k] (orthographe b d g ou p t k après s ou d'autres constrictives). Les deux séries de consonnes sont essentiellement sourdes, fait qui avait déjà été prouvé par Einarsson (1927, p. 38-49), bien qu'en certaines positions les non-aspirées puissent devenir des sonores. D'après nos recherches on peut espérer que dans la position intervocalique (dans le mot et dans la phrase) et dans la position après nasale et liquide sonore de 5 à 7% des occlusives non-aspirées deviennent des sonores (Pétursson 1974b, p. 326-328, 334-335).

On a supposé que les marques distinctives des deux séries de consonnes sont la force et l'aspiration. Les consonnes aspirées sont supposées être des fortes. L'aspiration est considérée automatiquement comme un indice de force (Jakobson, Fant, Halle 1972, p. 38). Dans nos recherches nous n'avons pas confirmé ce point de vue. Nous avons constaté que toutes choses égales par ailleurs les occlusives non-aspirées ont un contact organique plus large et un angle des maxillaires plus réduit que les occlusives aspirées, deux faits qui sont à mettre en relation avec une énergie articulatoire plus grande des non-aspirées

(Straka 1963, p. 35-43). L'occlusion des occlusives aspirées est plus brève que celle des occlusives non-aspirées, alors que la durée totale (occlusion+aspiration) des aspirées est plus grande que celle des non-aspirées. Ces faits sont résumés dans le tableau 1.

Tableau 1. Durée moyenne (en cs.) des occlusives islandaises en position initiale de mot.

Occlusive	occlusion	aspiration	occlusive	occlusion	explosion
p^h	9,1	4,99	p	9,16	0,9
t^h	8,36	3,96	t	8,45	1,54
c^h	7,88	4,96	c	8,42	2,32
k^h	7,32	5,59	k	8,0	1,72
Moyenne	8,16	4,75		8,50	1,62

Dans la présente étude nous abordons le problème du comportement glottal pour les occlusives islandaises, un domaine qui n'a pas été examiné pour les occlusives islandaises. Dans cette première contribution nous nous limiterons aux consonnes occlusives en position initiale de mot et après s à l'initiale dans les groupes [sp, st, sc]. A titre de comparaison nous avons également inclus les constrictives [f þ h] en position initiale. Nous espérons pouvoir examiner plus tard d'autres problèmes du consonantisme islandais à l'aide de cette méthode.

Matériaux linguistiques.

Les matériaux linguistiques que nous avons utilisés pour cette recherche consistent en 46 mots islandais isolés, contenant les consonnes et les groupes de consonnes que nous voulions étudier. Tous les mots devaient contenir une voyelle antérieure avant et après la consonne ou le groupe consonantique qui intéressait. Ceci était une condition essentielle afin de limiter le plus possible les mouvements de la racine de la langue. Pour les voyelles postérieures il existe le risque que la racine de la langue puisse changer la position du tube en polyéthèlene dans lequel est insérée la cellule photoélectrique qu'on fait passer par le nez dans le pharynx. Par conséquent nos mots contiennent les voyelles antérieures [i I e Y ö] ainsi que la diphtongue [ei]. Un mot contient la voyelle accentuée [a]. Les 46 mots ont été lus 21 fois dans un ordre dispersé, chaque mot étant précédé par le mot-clé segðu "dis" (impératif de segja).

La méthode glottographique.

Les enregistrements ont été réalisés à l'Institut de Phonétique de l'Université de Lund le 18 et le 19 février 1975.[1]. La fig. 1 résume schématiquement la disposition des appareils pour nos enregistrements. Sur nos tracés nous avons quatre lignes:

1. L'oscillogramme: le son pris au niveau du larynx avec un microphone collé au thyroïde (ligne LM).
2. Le photoglottogramme: enregistrement du rayon de lumière transmis à travers la glotte (ligne GL).
3. La pression intrabuccale enregistrée avec un tube en plastique de 1,9 mm de diamètre passé par le nez dans le pharynx (ligne PA). L'interférence avec l'articulation est minimale et cette disposition du tube permet de mesurer la pression intrabuccale pour les consonnes vélaires (Löfqvist 1971, p. 27).
4. Le débit d'air buccal enregistré avec un aéromètre électrique (ligne DA).

La lumière est transmise dans la trachée à la jonction du thyroïde et du cricoïde. Il s'agit d'un rayon de lumière froide d'une puissance de 7000 lux. Au-dessus de la glotte suspendue dans un tube en polyéthèlène qu'on fait passer par le nez et qu'on avale dans l'oesophage pour le stabiliser il y a une cellule photoélectrique enregistrant la lumière transmise à travers la glotte lorsque celle-ci s'ouvre. Le glottogramme a toujours des contours arrondis (figs. 2-9). Ceci s'explique parce qu'une certaine quantité de lumière rayonne à travers la glotte fermée (Köster, Smith 1970, p. 96; van den Berg 1968, p. 287-288), mais ce facteur n'a pas d'importance dans notre cas, car il s'agit de mesurer la forme glottale de consonnes sourdes pour lesquelles la glotte est de toute façon ouverte. Plus délicat est le problème de la localisation de la cellule photoélectrique dans la cavité pharyngale. On ne peut en effet pas savoir exactement où la cellule photoélectrique se trouve par rapport à la glotte.

---

1. Nous remercions les collègues de Lund pour leur généreuse aide dans la réalisation des enregistrements. Tout particulièrement nous remercions Mme le professeur Kerstin Hadding, directrice de l'Institut et Mme Eva Gårding qui nous ont donné la possibilité de faire ces enregistrements. Nous remercions également M. Anders Löfqvist, chargé de cours à l'Institut de Phonétique de Copenhague et collaborateur de l'Institut de Phonétique de Lund, pour le grand dévouement avec lequel il nous a aidé dans la réalisation technique des enregistrements. Nous exprimons notre gratitude à MM. Sidney Wood, Robert Bannert et Per Lindblad de l'Institut de Phonétique de Lund pour les discussions que nous avons eues avec eux à propos de ces problèmes.

Nous avons placé la cellule là où elle donnait de clairs signaux. Les endroits possibles de localisation peuvent pourtant être nombreux. Il s'est en effet avéré impossible de calibrer le glottographe avec le larynx vivant (van den Berg 1968, p. 288). L'amplitude glottale enregistrée ne révèle pas l'aperture glottale réelle, mais une aperture relative que nous avons essayée de transformer en valeurs numériques. Nous avons suivi dans ce cas la méthode proposée par Frøkjær-Jensen et al. (1971, p. 128; Mansell 1974, p. 143). Comme l'aperture glottale n'est pas tout à fait constante à partir du thyroïde le long des cordes vocales jusqu'aux arytaenoïdes, il est évident que les différences d'aperture glottale sur le plan horizontal peuvent influencer l'amplitude glottale enregistrée. Ceci peut dépendre de la localisation de la cellule par rapport à la glotte (Frøkjær-Jensen et al. 1971, Mansell 1974, p. 166). Cependant comme il s'agit pour nous d'établir des valeurs relatives ce problème reste pour nous marginal. De même la direction du rayon de lumière peut changer légèrement avec le moindre changement de position du sujet. Ceci est aussi un facteur pouvant influencer l'amplitude glottale sur le tracé, de sorte qu'on peut s'attendre à quelques irrégularités dans les tracés glottographiques, fait que nous constatons effectivement. Néanmoins étant donné que la cellule photoélectrique est relativement fixe dans le pharynx on peut avoir une certaine confiance que ces irrégularités ont une importance réduite. Bien qu'elles ne soient pas à négliger, nous pensons que leur importance n'est pas telle à influencer l'interprétation des tracés d'une manière décisive.

Les mesures.

Nous avons pris les mesures suivantes:

1. Durée de l'occlusion. Nous avons mesuré la durée de l'occlusion à partir du moment où l'amplitude de la voyelle précédente (ligne LM) décroît sensiblement. En général ce moment coïncide avec une chute rapide de la ligne du débit d'air. La rapide montée de cette ligne (DA) coïncide en général avec la rupture de l'occlusion qu'on peut également voir sur l'oscillogramme comme une brève période de bruits irréguliers.
2. Durée de l'explosion pour les non-aspirées.
3. Durée de l'aspiration pour les aspirées.
4. Le moment de l'ouverture glottale maximale par rapport au début de l'occlusion.

5. L´amplitude glottale mesurée selon une échelle relative. Pour le [pʰ] nous avons fixé cette valeur à 1.00 et nous transformons les valeurs relatives relevées pour les autres consonnes pour pouvoir les comparer avec le [pʰ].
6. La pression intrabuccale mesurée en cm H_2O.
7. Le moment de la pression intrabuccale maximale par rapport au début de l´occlusion.

Nous avons mesuré la durée en millisecondes. Pour les constrictives [f þ h] nous supposons que la durée du bruit fricatif équivaut à la durée de la consonne. En général nous pouvons établir une corrélation avec la ligne du débit d´air. Les quelques cas où nous constatons un retard dans le débit d´air pour [f þ] peuvent s´expliquer de deux manières:

1. La constriction a été très serrée au début de la consonne de sorte que le courant d´air diminue un moment. Le résultat est un infléchissement de la ligne du débit d´air au début de la consonne. Ceci est physiologiquement explicable, car on sait que sous l´effet du renforcement articulatoire la consonne se ferme (Straka 1963, p. 35-43). Une constrictive énergique a donc un canal de constriction plus fermé qu´une constrictive non-énergique. L´impédance accrue offerte à la sortie de l´air pour une constrictive énergique se traduit dans le tracé par un retard dans le flux d´air [2].
2. Il est possible que l´embouchure n´ait pas toujours été assez bien appuyée contre la bouche de sorte qu´une certaine quantité d´air ait pu échapper. Bien que nous ayons pris tout le soin possible, une telle éventualité n´est pas tout à fait à exclure.

Ces cas sont pourtant peu nombreux. En général il y a une bonne correspondance entre l´oscillogramme et la ligne du débit d´air.

2. Nous n´avons pas comparé systématiquement s´il existe une corrélation entre l´ouverture glottale et le débit d´air. D´après Mansell (1974, p. 163) une telle corrélation n´existe pas. L´examen préliminaire de nos tracés semble le confirmer.

Résultats.

Le tableau 2 résume les résultats des mesures pour les consonnes occlusives en position initiale de mot, la seule position où il y a un contraste linguistique entre une occlusive aspirée et une occlusive non-aspirée en islandais.

Tableau 2. Consonnes occlusives en position initiale absolue. Les chiffres 1 à 7 désignent les mesures.

Cons.	Nombre d´ex.	1	2	3	4	5	6	7
[pʰ]	21	128,5		36,0	80,5	1.00	7,9	128,5
[tʰ]	21	120,5		36,0	78,5	1.02	10,0	120,5
[cʰ]	20	109,0		50,5	78,5	1.79	8,9	109,0
[kʰ]	20	126,5		50,0	98,5	1.81	10,8	126,5
[p]	21	123,5	10,0		47,0	0.67	9,0	123,5
[t]	21	158,0	10,0		54,0	0.62	11,0	158,0
[c]	21	107,5	16,5		45,5	0.87	9,6	107,5
[k]	21	120,5	12,3		43,0	1.41	9,3	120,5

Selon le tableau 2 nous remarquons les faits suivants:

1. Durée de l´occlusion.

Dans nos recherches déjà achevées nous avions constaté qu´en moyenne la durée de l´occlusion des aspirées était plus brève que la durée de l´occlusion des non-aspirées. Dans le matériel actuel ceci est seulement confirmé pour la paire [tʰ/t]. Il est difficile d´expliquer à quoi cet écart par rapport à nos premiers résultats est dû. Il faut toutefois noter que la différence est toujours très réduite (elle n´atteint qu´une petite fraction de centiseconde) et que le fait de lire dans une embouchure peut avoir un effet sur la durée de l´occlusion. En fait on ne peut pas dire que les mots lus dans ces conditions aient une prononciation normale et naturelle, mais entre eux ils sont comparables.

2.-3. Durée de l´explosion et de l´aspiration.

Il se confirme le fait déjà constaté que les palatales et les vélaires ont la tendance à avoir l´explosion et l´aspiration la plus longue (Pétursson 1974b, p. 321, tableau 74).

4. Le moment de l´ouverture glottale maximale par rapport au début de l´occlusion.

La fig. 10 illustre le moment de l´ouverture maximale par rapport au début de l´occlusion. Pour les consonnes aspirées l´ouverture glottale maximale intervient toujours dans la seconde moitié de l´occlusion. Dans la plupart des cas l´ouverture glottale maximale intervient dans le début de la seconde moitié de l´occlusion. Dans un tiers des cas le maximum d´ouverture glottale intervient 10 ou 20 msec. avant la rupture de l´occlusion. Dans notre matériel nous n´avons aucun cas où l´ouverture glottale maximale tombe après la rupture de l´occlusion, ce qui est le plus fréquent en danois (Frøkjær-Jensen 1971, p. 132; Mansell 1974, p. 149 tableau 5).

Pour les consonnes non-aspirées le moment de l´ouverture glottale maximale intervient presque toujours dans la première moitié de l´occlusion. Dans de nombreux cas ce moment intervient 30-40 msec. après le début de l´occlusion.

La coordination temporelle (timing) entre l´activité glottale et les organes supraglottaux est par conséquent fondamentalement différente pour les consonnes aspirées et les non-aspirées, ce qui confirme une fois de plus que la différence essentielle entre les deux séries d´occlusives en islandais est avant tout une différence d´aspiration (Pétursson 1974a, p. 203-205; 1974b, p. 360-365).

5. L´ouverture glottale.

Pour la consonne [pʰ] nous avons fixé la valeur numérique 1.00 pour l´ouverture glottale et à partir de ce chiffre nous avons calculé une valeur relative pour les autres consonnes. Nous suivons sur ce point la méthode proposée par Frøkjær-Jensen et al. (1971, p. 128) qui analysent cependant seulement les consonnes labiales et le [h]. Ces chiffres n´indiquent pas l´ouverture glottale réelle, mais ils donnent une idée de différences relatives. Même avec cette restriction les chiffres font apparaître des facteurs intéressants (fig. 11). D´après les chiffres le lieu d´articulation aurait une certaine influence sur l´aperture glottale. L´aperture glottale serait d´autant plus large plus le lieu d´articulation est reculé dans la cavité buccale. Si nous admettons d´après Kim (1970, p. 109-111) que la durée de l´aspiration dépend de l´ouverture glottale au moment de la rupture de l´occlusion ceci expliquerait pourquoi les palatales et les vélaires ont toujours l´aspiration et l´explosion les plus longues. Il serait aussi possible de penser que le volume de la cavité derrière le lieu d´articulation pourrait avoir une influence sur la pression intrabuccale et par là sur l´ouverture glottale (Lindqvist 1972b, p. 15). On pourrait espérer que la pression intrabuccale serait d´autant plus élevée plus la cavité est réduite, supposant bien entendu que toutes les consonnes dépensent, toutes choses égales par ailleurs, la même quantité d'air. Cette hypothèse n'a cependant pas été confirmée.

6. La pression intrabuccale.

Si nous calculons la moyenne générale de la pression intrabuccale pour les aspirées (9,40 cm H_2O) et pour les non-aspirées (9,72 cm H_2O) nous constatons que la pression intrabuccale est légèrement plus élevée pour les non-aspirées. Pourtant la différence est si réduite qu'il est difficile de penser qu'elle puisse avoir une signification quelconque. On peut seulement dire qu'il n'y a pratiquement pas de différence entre les occlusives aspirées et les non-aspirées islandaises en ce qui concerne la pression intrabuccale.

7. Le moment de la pression intrabuccale maximale.

En général la pression intrabuccale maximale est atteinte à la fin de l'occlusion juste avant la rupture. Cela est le cas pour les aspirées et pour les non-aspirées. Pourtant il y a quelques exemples où le maximum de pression est atteint 10 ou 20 msec. après le début de l'occlusion. Pour [p^h] nous avons 4 cas où le maximum de pression est atteint 10 msec. et un cas où il est atteint 20 msec. après le début de l'occlusion. Pour [t^h] ces cas sont 4 et pour [p] nous avons deux exemples de ce type (comp. fig. 3). Lorsque la pression intrabuccale atteint son maximum juste après le début de l'occlusion, elle tombe généralement de 1 ou 2 cm H_2O un moment après et reste à ce niveau jusqu'au moment de la rupture de l'occlusion.

Les constrictives.

Le tableau 3 résume les mesures sur les constrictives dans notre matériel.

Tableau 3. Les constrictives [f þ h].

Cons.	Nombre d'ex.	1	4	5	6
[f]	21	155,5	68,5	1.18	8,1
[þ]	21	162,5	70,0	1.56	10,9
[h]	21	122,0	46,0	0.50	?

L'aperture glottale est la plus grande pour le [þ] ce qui correspond à la même tendance que pour les occlusives, à savoir que l'aperture augmente à mesure que le lieu d'articulation recule dans la cavité buccale.

D'après Lindqvist (1972b, p. 17) l'aperture glottale est plus grande pour les constrictives que pour les occlusives. Nos données semblent le confirmer en général. Pour [h] l'aperture glottale est la plus réduite. Ce même fait est également constaté par Lindqvist (1972b, p. 17). On peut éventuellement expliquer ceci par le fait que le [h] est prononcé avec un canal buccal ouvert sans constriction dans la cavité buccale. Son articulation ressemble en effet à l'expiration pure et simple. Or, on sait que la glotte est normalement plus fermée à l'expiration qu'à l'inspiration (Ladefoged, 1971, p. 9), ce qui pourrait expliquer le fait observé.

Un autre fait intéressant est le mouvement symétrique de la glotte pour les constrictives. L'aperture maximale est atteinte vers le milieu de la durée de la consonne (figs. 4-6, 14) et les mouvements d'ouverture et de fermeture ont à peu près la même durée.

La pression intrabuccale n'est en général pas plus réduite pour les constrictives que pour les occlusives (fig. 13), mais pour les constrictives la pression ne reste qu'un petit moment à un niveau (figs. 4-6), alors que pour les occlusives ce moment est beaucoup plus long (figs. 2 et 3).

Les groupes de s+occlusive.

En islandais les occlusives p t k ne peuvent pas être aspirées après s ou une autre constrictive faisant partie du même mot que l'occlusive. L'islandais partage ce trait phonétique avec les autres langues germaniques et autant que nous le sachions ce trait est commun à toutes les autres langues et dialectes européens. Le tableau 4 résume les données des groupes de s+occlusive.

Tableau 4. Groupes de s+occlusive à l'initiale de mot.

Groupe	Nombre d'ex.	1		2	4	5	6
		1e cons.	2e cons.				
[sp]	21	135,0	87,0	9,5	57,0	1.53	9,0
[st]	21	130,5	93,1	10,0	59,5	2.12	11,0
[sc]	21	122,5	103,0	18,8	69,0	2.09	9,0

Il ressort du tableau 4 qu'au niveau glottal les groupes de s+occlusive sont traités comme une unité. La glotte fait un seul mouvement d'ouverture et de fermeture pour les deux consonnes (figs. 7-9) comme s'il s'agissait d'une seule consonne occlusive ou constrictive.

Le maximum d'ouverture glottale tombe à l'intérieur du s (tableau 4, colonne 4) bien avant le début de la consonne occlusive (fig. 12). L'ouverture glottale est en général très grande (fig. 11) et la pression intrabuccale est élevée (fig. 13) avec pourtant une tendance assez prononcée à conserver deux niveaux correspondant à la consonne s et à l'occlusive respectivement (figs. 7-9). Dans nos tracés nous n'avons pas observé la tendance à deux sommets dont parlent Frøkjær-Jensen et al. (1971, p. 138 et fig. 5). pour les glottogrammes de s+occlusive en danois. Il faut pourtant noter que l'exemple de Frøkjær-Jensen et al. est un exemple où les deux consonnes appartiennent à des syllabes différentes.

Conclusions.

L'activité glottale est fondamentalement différente pour les occlusives aspirées et les occlusives non-aspirées islandaises. Cette différence se situe à deux niveaux:

1. Par rapport à l'activité supraglottique.
2. Au niveau de l'ouverture glottale.

Par rapport à l'activité supraglottique le sommet de l'ouverture glottale intervient très tôt après le début de l'occlusion, si la consonne est une non-aspirée. L'ouverture glottale est réduite et la glotte est déjà pratiquement fermée au moment de la rupture de l'occlusion. L'explosion est donc brève et la voyelle commence immédiatement après l'explosion.

Pour les consonnes aspirées l'ouverture glottale maximale tombe ou bien dans la deuxième moitié de l'occlusion ou bien 10 à 30 msec. avant la rupture de l'occlusion. Comme la glotte est très ouverte pour les aspirées, il faut un temps plus long, pour qu'elle puisse arriver à une position fermée. Cet intervalle constitue l'aspiration. Comme l'ouverture glottale semble augmenter avec le recul du lieu d'articulation (fig. 11), cela pourrait expliquer pourquoi l'aspiration est la plus longue pour les vélaires et les palatales. Il faut cependant aussi penser que les mouvements du dos de la langue sont plus lents que ceux de l'apex, ce qui peut partiellement expliquer que l'aspiration des vélaires et des palatales est plus longue que celle des labiales et des alvéodentales.

La différence de synchronisation entre l'activité glottale et l'activité supraglottique est le trait essentiel distinguant les aspirées et les non-aspirées islandaises. Ceci revient à dire que la différence essentielle entre les deux séries de consonnes est une différence d'aspiration (Pétursson 1974b, p. 357-365). Nos enregistrements apportent tous les arguements en faveur de la théorie de Kim (1970) que la durée de l'aspiration dépend essentiellement du degré d'ouverture glottale au moment de la rupture de l'occlusion. Par contre la durée de l'aspiration ne semble pas dépendre de l'ouverture glottale absolue (Lindqvist 1972b, p. 20). Nos tracés ne nous autorisent pas à décider si l'ouverture glottale pour les non-aspirées est une activité passive ou active (Frøkjær-Jensen et al. 1971, p. 134). Nous pouvons seulement constater la différence d'ouverture glottale entre les aspirées et les non-aspirées, mais nous n'avons pas de données nous permettant de dire que l'une est passive et l'autre est active. Jusqu'à la preuve du contraire nous supposons que l'ouverture glottale est toujours un mouvement actif et nous y voyons la preuve dans le fait que l'ouverture glottale n'intervient pas à n'importe quel moment par rapport à l'occlusion, mais elle est au contraire soigneusement coordonnée avec l'activité supra-glottique.

Pour les constrictives l'ouverture glottale maximale intervient à peu près au milieu de la durée et les phases d'ouverture et de fermeture glottales sont presque symétriques.

On a supposé que la pression intrabuccale serait un paramètre important assurant la distinction entre une occlusive aspirée et une non-aspirée. Guðfinnsson (1946, p. 39) suppose que la pression intrabuccale rompt l'occlusion des aspirées, alors que la rupture de l'occlusion des non-aspirées serait produite par l'activité musculaire. Nous apportons les premières données sur la pression intrabuccale des occlusives islandaises. Ces données montrent que les faits ne se présentent pas comme Guðfinnsson le supposait. La pression intrabuccale moyenne est plus grande pour les non-aspirées que pour les aspirées, mais la différence est si réduite qu'elle ne peut pas être significative (fig. 13). Il est permis d'en conclure, nous semble-t-il, que la pression intrabuccale ne peut pas différencier les deux séries de consonnes: elles se différencient essentiellement par l'aspiration.

Ce qui surprend un peu, c'est que la pression intrabuccale des constrictives est aussi grande que celle des occlusives. Ce même fait a été constaté par Slis (1970, p. 200). Ceci semble indiquer que c'est en premier lieu l'activité glottale qui détermine la pression intrabuccale et non pas le mode articulatoire de la consonne. La pression intrabuccale serait essentiellement indépendante de l'activité supraglottique.

Les groupes de s+occlusive sont traités au niveau glottal comme une unité, la glotte ne s'ouvrant qu'une seule fois pour les deux consonnes. Il s'agit d'un mode de coordination particulier entre l'activité glottale et supraglottique qui devrait expliquer pourquoi la consonne occlusive n'est pas aspirée après s. Pour obtenir une économie maximale de mouvement glottal la glotte ne s'ouvre qu'une seule fois. La difficulté de coordonner le mouvement glottal de deux manières différentes avec le même type d'activité supraglottique, à savoir s+occlusive, est probablement la raison pour laquelle les langues ont en général renoncé à créer un contraste d'aspiration après les groupes de s+occlusive. C'est ce principe que Rothenberg désigne comme le mouvement glottal unidirectionnel (Rothenberg 1968, p. 88-89). Kim (1970, p. 113-114) explique l'absence d'aspiration des occlusives après s comme un fait de coarticulation. D'après lui on posséderait déjà pendant le s l'information sur l'occlusive subséquente. Il ne faut par conséquent pas retarder un mouvement articulatoire pour la consonne subséquente si ce mouvement n'est pas incompatible avec la réalisation de la consonne précédente. Si l'ouverture glottale commence pendant le s, ceci signifie que la fermeture de la glotte commencera un peu avant ou pendant le début de l'occlusive subséquente. L'économie des mouvements, c'est-à-dire la réalisation d'un seul cycle balistique dont la durée moyenne oscille entre 150 et 125 msec. pour le débit normal de la parole (Rothenberg 1968, p. 82; d'après Lindqvist 1972b, p. 12 la durée de ce cycle est de 200 msec.), implique que l'occlusive sera non-aspirée après le s. Le pouvoir explicatif de cette hypothèse est si grand que nous avons cru pouvoir affirmer dans une discussion qu'il s'agissait d'un fait universel (Pétursson 1975, p. 653). Cette même hypothèse semble expliquer parfaitement pourquoi une occlusive aspirée ne peut pas exister après une liquide et nasale sourde en islandais du Sud, alors qu'elle peut exister après nasale est liquide sonore en islandais du Nord (Guðfinnsson 1964, p. 16-43; Pétursson 1973, p. 121-124).

Pourtant il ne s'agit pas d'une caractéristique universelle. Dans l'évolution des langues indiennes les groupes ṣṭh, sth, sph ont existé (Elizarenkova 1974, p. 172). En espagnol colombien dans les parlers de Carmen de Carupa (province de Cundinamarca), Corrales et de Monguí (province de Boyacá) il y a l'aspiration dans les groupes consonantiques [n, r, l + p, t, k] , [s + p, t, k] et dans les groupes de deux occlusives. Les consonnes intervocaliques peuvent aussi être aspirées dans ces dialectes. L'aspiration apparaît aussi en phonétique syntactique et semble complètement indépendante des limites du mot. Les langues européennes distinguent par contre soigneusement entre la phonétique du mot et la phonétique syntactique. Dans les langues germaniques l'aspiration ne peut pas exister après s dans le même mot, mais elle peut très bien exister après s en phonétique syntactique:

Allemand: Was tut sie? [va:s t^{h}u:t zi:] "que fait-elle?"

Islandais: Hás karl [hau:s k^{h}atl̥] "un homme enroué".

L'exemple de l'espagnol colombien (Rodríguez de Montes 1972)[3] est donc d'un intérêt typologique extraordinaire. L'aspiration de l'espagnol colombien dont tous les traits ne sont pas encore connus est peut-être un de ces effets de substrat qu'Amado Alonso (1967, p. 321) souhaitait qu'on puisse un jour découvrir en Amérique latine. Ainsi on explique les consonnes glottalisées de l'espagnol de Yucatán ("letras heridas" Zamora Vicente 1970, p. 390) comme étant dues à un substrat maya.

Bien que le type de combinaison de s+occlusive non-aspirée soit ainsi le type le plus répandu, ce n'est pas le seul type enregistré dans les langues. Théoriquement on comprend aisément qu'une occlusive aspirée soit possible après s. Une aspiration après s peut être obtenue de deux manières:

1. Par un retard du maximum de l'ouverture glottale jusque dans la deuxième moitié de l'occlusion de l'occlusive qui suit le s.
2. Par deux sommets d'ouverture glottale, un sommet correspondant à chaque consonne [4].

---

3. En espagnol le terme "aspiración" a généralement une signification différente de celle de ce terme dans les langues germaniques. En espagnol ce mot désigne habituellement une prononciation spéciale du h- initial provenant du f- latin, ou une prononciation spéciale de la jota [x] soit le souffle d'un -s implosif qui s'amuït (Zamora Vicente 1970, p. 55-73; 319-321; 396-397 et passim).
4. Il est à supposer que c'est ce modèle qui prévaut en phonétique syntactique dans les groupes /-s #occlusive aspirée/. Comme nous n'avons pas d'enregistrements de ces groupes ceci reste pour le moment seulement une hypothèse.

De ces deux possibilités la première nous semble la plus probable, parce qu'elle suppose une plus grande économie des mouvements phonatoires. Mais ceci ne peut être résolu qu'expérimentalement. Il faut aussi remarquer que l'économie, quoiqu'importante, n'est pas le seul facteur de l'évolution linguistique. En islandais du Nord les occlusives intervocaliques se maintiennent depuis des siècles bien qu'il n'y ait pas d'opposition aspirée/non-aspirée en position intervocalique. L'aspiration remplit linguistiquement le même rôle en danois et en islandais. Pourtant l'aspiration est plus longue en danois qu'en islandais, fait qui avait déjà été observé par Sveinbjörnsson (1894, p. 97). Les recherches expérimentales font apparaître que ceci est dû au fait que le maximum de l'ouverture glottale intervient plus tardivement en danois qu'en islandais (Frøkjær-Jensen 1971, p. 132; Mansell 1974, p. 149).

Dans l'interprétation de faits linguistiques il faut toujours rester ouvert aux possibilités d'explication qui peuvent se présenter. Il ne faut pas se laisser immobiliser par une théorie ou des théories déterminées, mais il faut se laisser guider par les faits dans l'interprétation. Rappelons ici une phrase oubliée de Rousselot qui renferme aujourd'hui encore une très grande vérité: " Les faits s'arrangent toujours; quant aux théories on les arrange au besoin" (Rousselot 1911, p. 113). En travaillant dans cet esprit en linguistique et en phonétique on sera toujours sûr que la recherche se fera sur des bases solides.

Phonetisches Institut
der Universität Hamburg.

Bibliographie.

Alonso, Amado (1967): Examen de la teoría indigenista de Rodolfo Lenz, dans Estudios lingüísticos, temas hispanoamericanos, pp. 268-321 , 3e éd. (Gredos, Madrid).

Berg, Jw. van den (1968): Mechanism of the larynx and the laryngeal vibrations, dans Malmberg: Manual of Phonetics, pp. 278-308 (North Holland, Amsterdam).

Einarsson, S. (1927): Beiträge zur Phonetik der isländischen Sprache (Bröggers Boktrykkeri, Oslo).

Elizarenkova, T.J. (1974): Oppozicija pridyxatel´nosti v istorii razvitija indoarijskix jazykov, dans Issledovanija po diaxroničeskoj fonologii indoarijskix jazykov, pp. 166-183 (Izdatel´stvo Nauka, Moscou).

Frøkjær-Jensen, B., C. Ludvigsen et J. Rischel (1971): A glottographic study of some Danish consonants, dans Form and Substance, phonetic and linguistic papers presented to Eli Fischer-Jørgensen, pp. 123-140 (Akademisk Forlag, København).

Guðfinnsson, B. (1946): Mállýzkur I (Ísafoldarprentsmiðja H.F. Reykjavík).

Guðfinnsson, B. (1964): Mállýzkur II (Heimspekideild Háskóla Íslands og Bókaútgáfa Menningarsjóðs, Reykjavík).

Jakobson, R., C. Gunnar M. Fant et Morris Halle (1972): Preliminaries to speech analysis. The distinctive features and their correlates, 10e éd. (MIT Press, Cambridge (Mass.)).

Kim, Chin-Wu (1970): A theory of aspiration, Phonetica 21, 107-116.

Köster, J.-P. et S. Smith (1970): Zur Interpretation elektrischer und photoelektrischer Glottogramme, Folia Phoniatrica 22, 92-99.

Ladefoged, P. (1971): Prelimaries to linguistic phonetics, 2e éd. (University of Chicago Press, Chicago).

Lindqvist, J. (1972a): A descriptive model of laryngeal articulation in speech, STL-QPSR 2-3/1972, 1-9.

Lindqvist, J. (1972b): Laryngeal articulation studied on Swedish subjects, STL-QPSR 2-3/1972, 10-27.

Löfqvist, A. (1971): Some observations on supraglottal air pressure, Working Papers 5, Phonetics Laboratory Lund University, pp. 27-60.

Mansell, Ph. (1974): Some airflow and glottogramm data on Danish whisper, ARIPUC 8, 133-171.

Pétursson, M. (1973): Phonologie des consonnes nasales en islandais moderne, La Linguistique 9, 115-138.

Pétursson, M. (1974a): Les articulations de l'islandais à la lumière de la radiocinématographie (Klincksieck, Paris).

Pétursson, M. (1974b): Aspects acoustiques et articulatoires du phonétisme islandais (Serv. Reproduction des Thèses, Lille).

Pétursson, M. (1975): Examen expérimental de la réalisation phonétique des nasales sourdes en islandais moderne, The Nordic Languages and Modern Linguistics 2, pp. 634-653 (Almqvist & Wiksell, Stockholm).

Rodríguez de Montes, María Luisa (1972): Oclusivas aspiradas sordas en el español colombiano, Boletín del Instituto Caro y Cuervo 27, pp. 583-586 (Bogotá 1972).

Rothenberg, M. (1968): The breath-stream dynamics of simple-released-plosive production (S. Karger, Basel).

Rousselot, P.-J. (1911): La phonétique expérimentale jugée par M. Jespersen, Revue de Phonétique 1, 105-113.

Slis, I. H. (1970): Articulatory measurements on voiced, voiceless and nasal consonants, Phonetica 21, 193-210.

Straka, G. (1963): La division des sons du langage en voyelles et consonnes peut-elle être justifiée? Travaux de Linguistique et de Littérature 1, 17-99.

Sveinbjörnsson, S. (1894): Spesimɛn islãdɛ, Le Maître Phonétique 9, 97-100.

Zamora Vicente, A. (1970): Dialectología española, 2^e^ éd. (Gredos, Madrid).

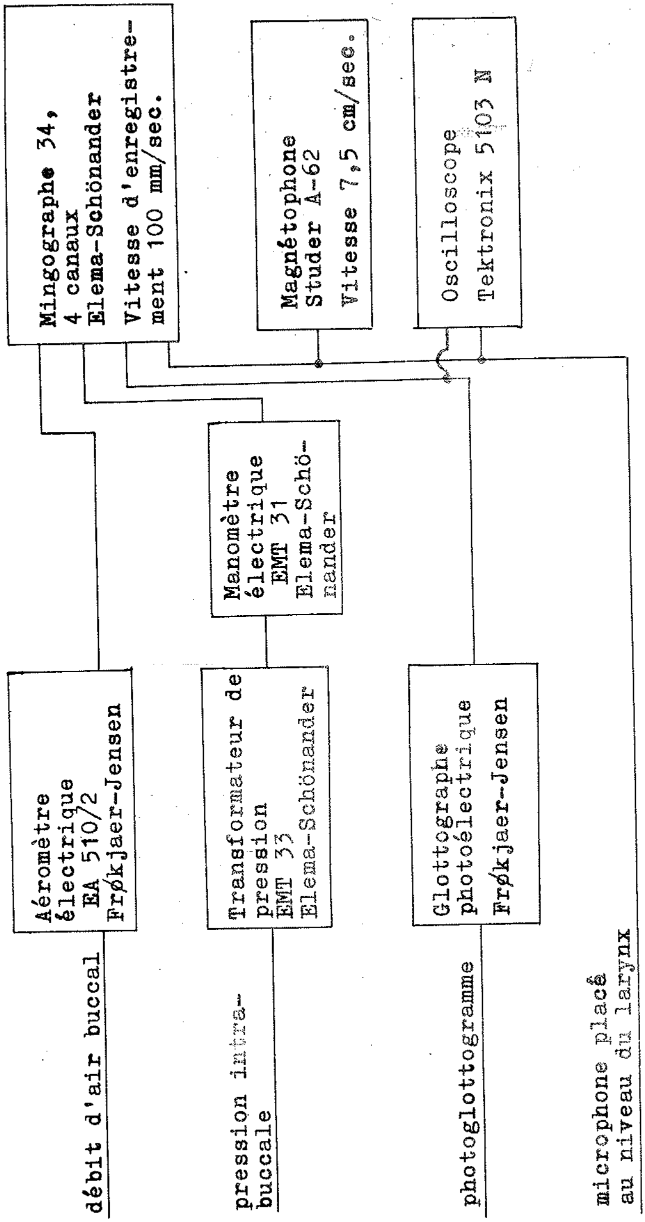


Fig. 1. Représentation schématique des appareils utilisés dans l´enregistrement.

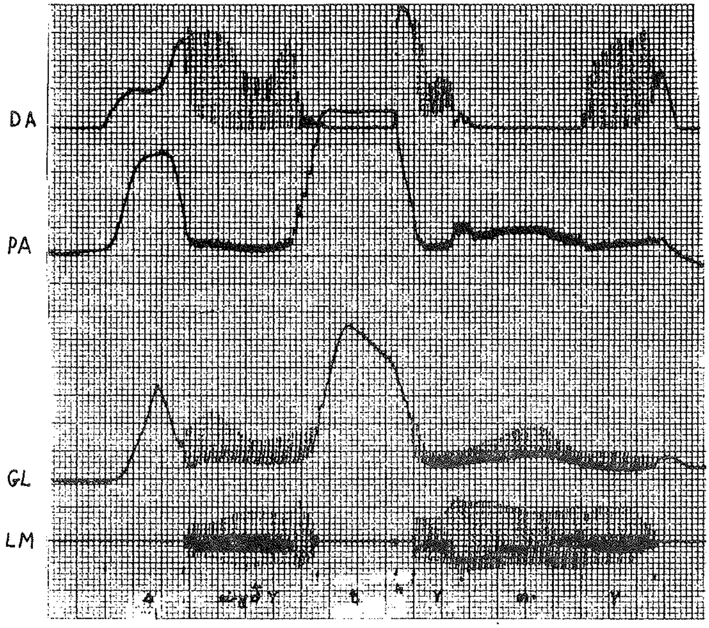


Fig. 2. Segðu tunnu "dis: fût".

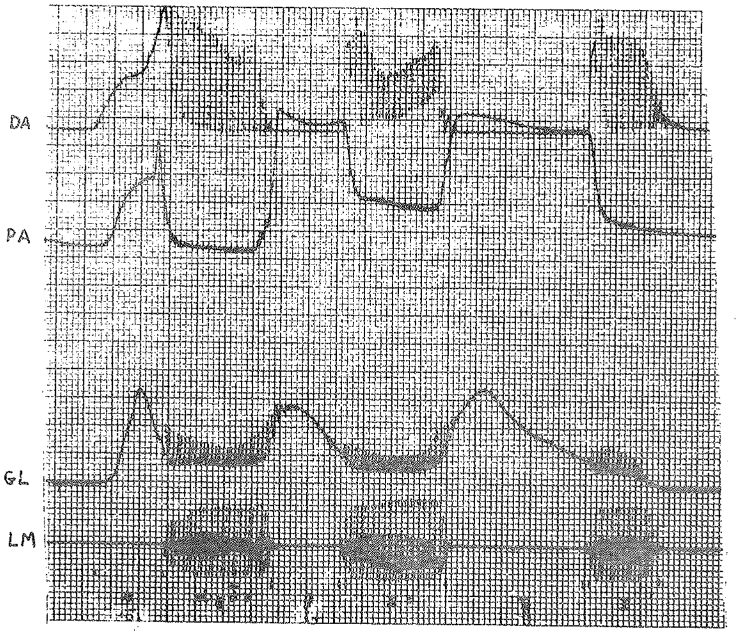


Fig. 3. Segðu biti "dis: morceau".

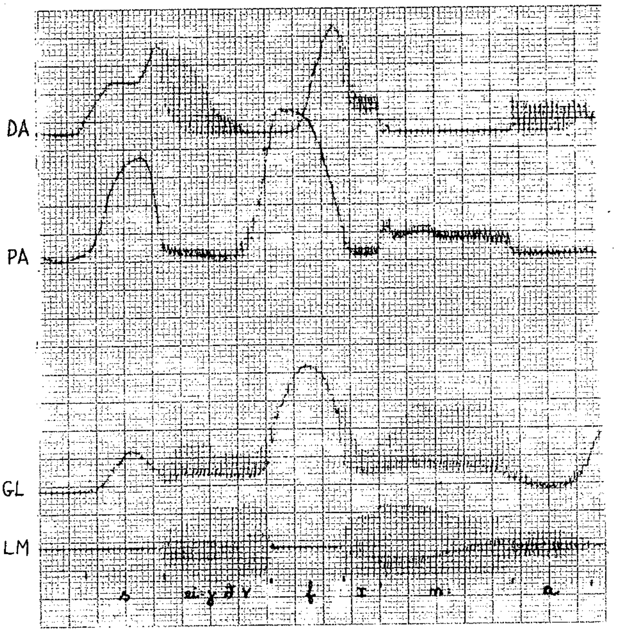


Fig. 4. Segðu finna "dis: trouver".

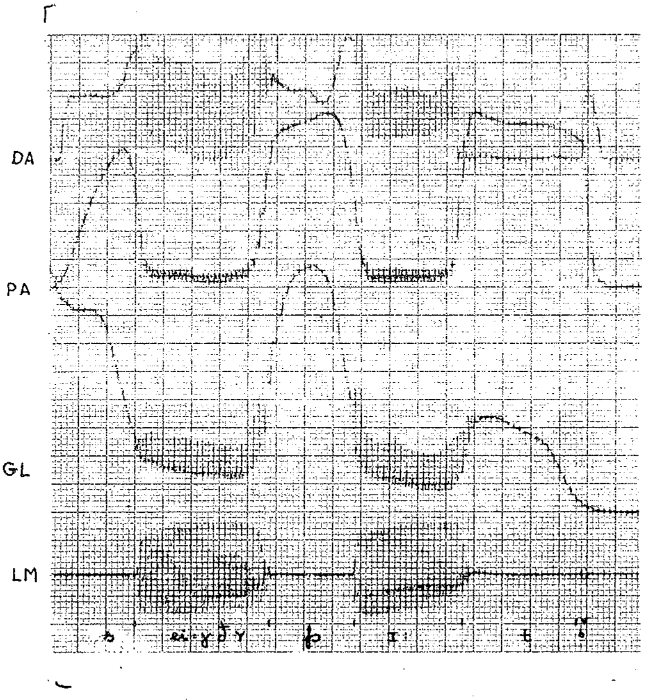


Fig. 5. Segðu þyt "dis: bruit de friction".

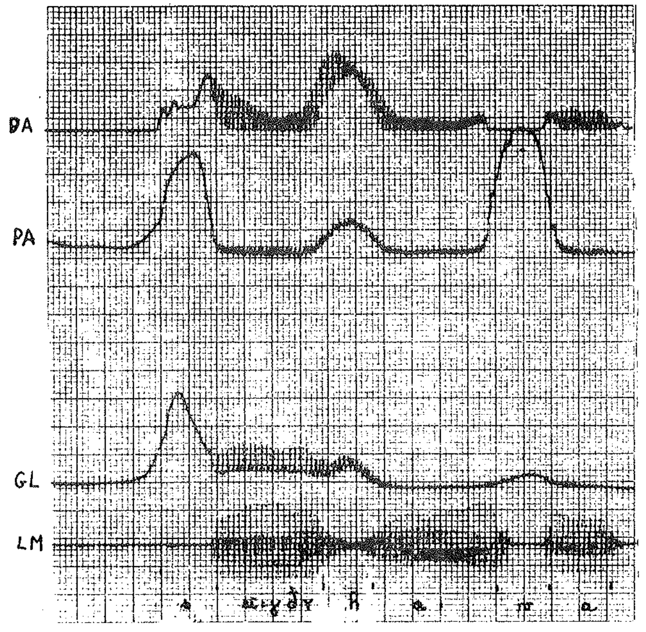


Fig. 6. Segðu hafa "dis: avoir"

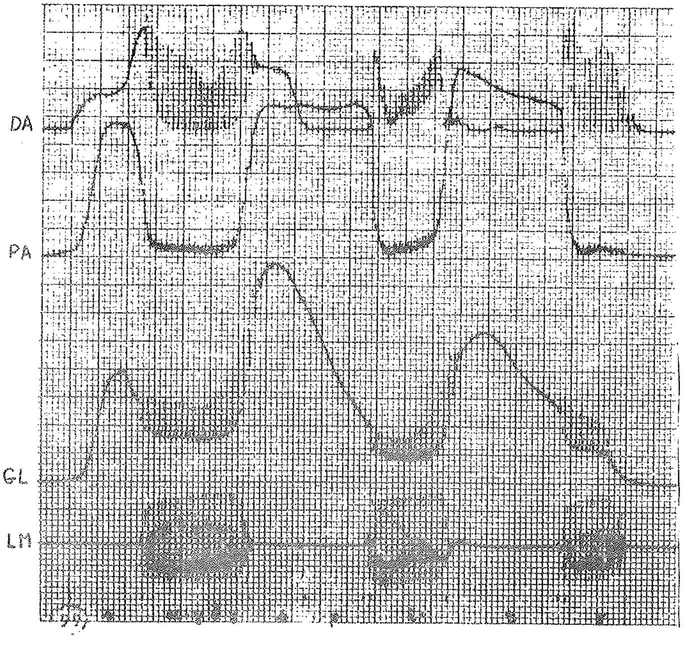


Fig. 7. Segðu spýtu "dis: pièce de bois"

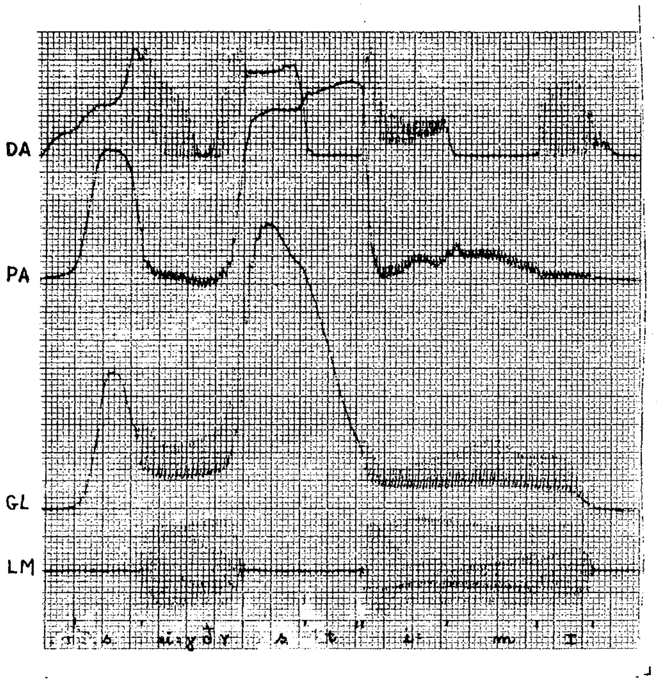


Fig. 8. Segðu <u>st</u>ími "dis: j´avance".

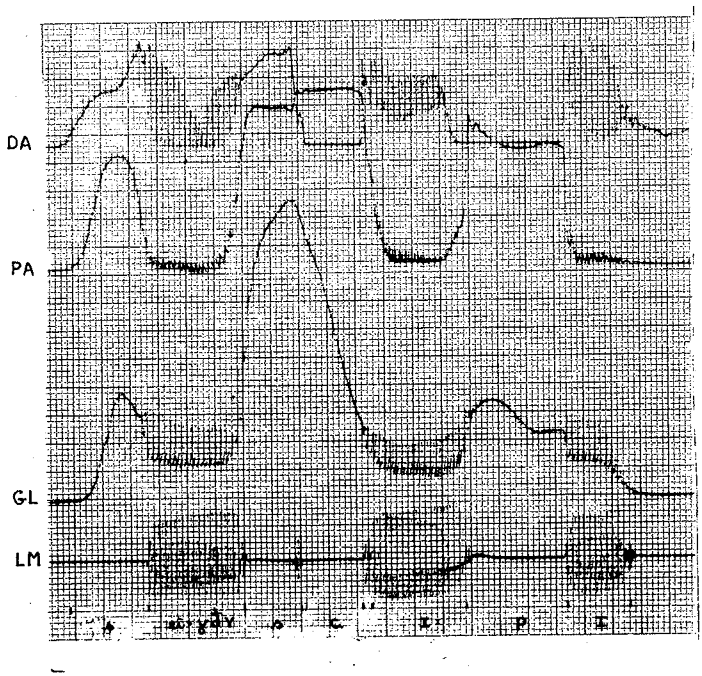


Fig. 9. Segðu <u>sk</u>ipi "dis: bateau".

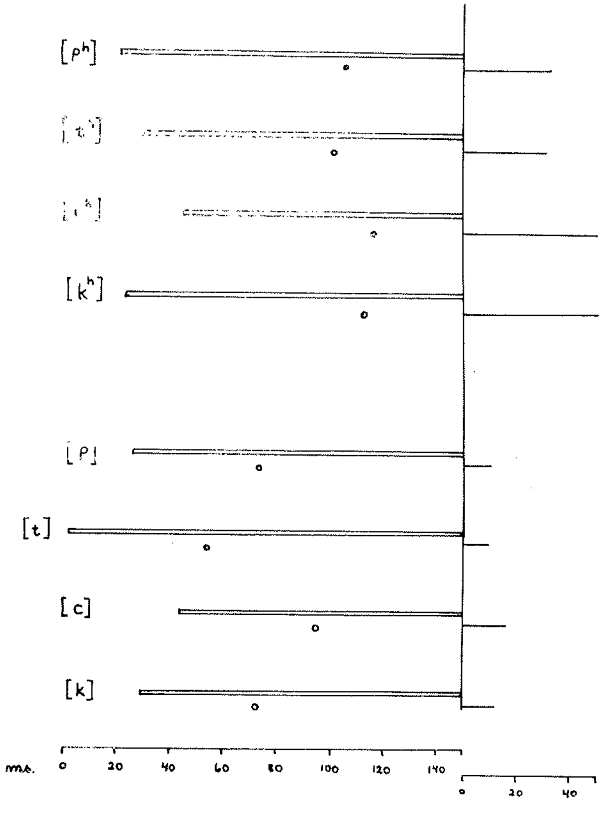


Fig. 10. Le moment de l'ouverture glottale maximale (indiqué par °) par rapport au début de l'occlusion (═══). La durée de l'occlusion a été normalisée par rapport à la rupture indiquée par le trait vertical.

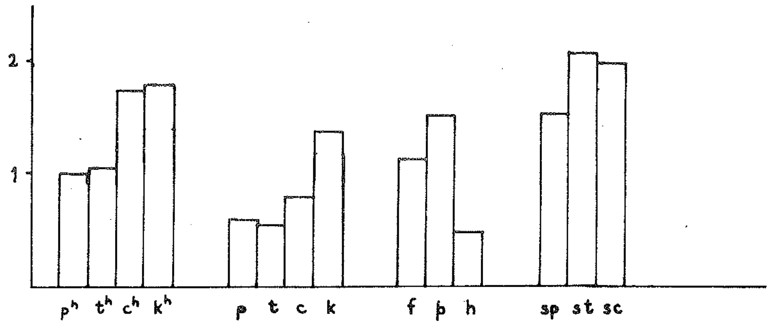


Fig. 11. L'ouverture glottale relative des consonnes étudiées. La valeur 1.00 est assignée arbitrairement à la consonne [p^h] et à partir de ce chiffre les valeurs relatives ont été calculées pour les autres consonnes.

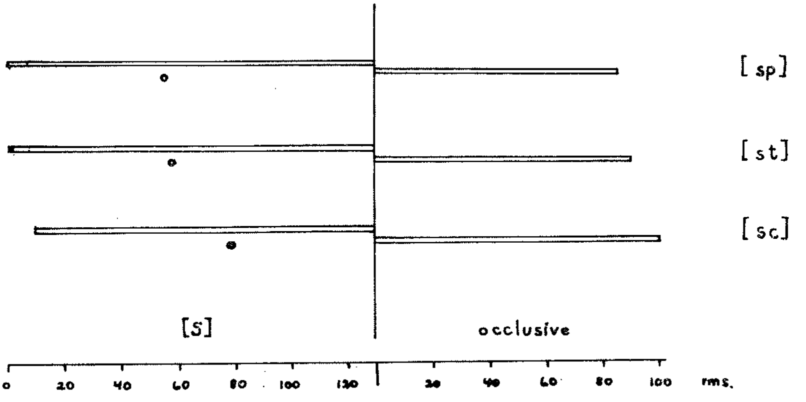


Fig. 12. Le moment de l'ouverture glottale maximale (indiquée par ∘) par rapport au début de la consonne s dans les groupes de s+occlusive.

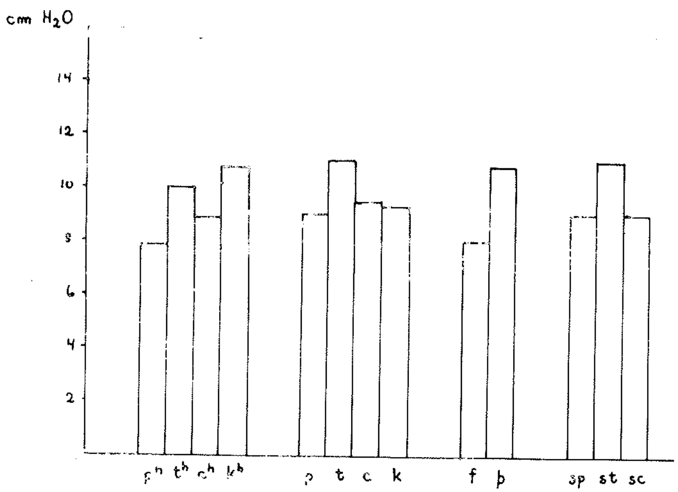


Fig. 13. Pression intrabuccale moyenne en cm H_2O pour les consonnes étudiées.

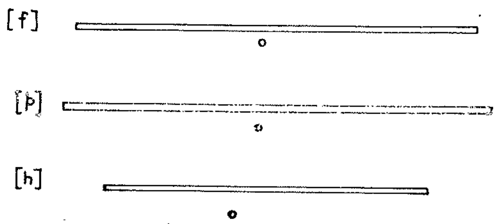


Fig. 14. Le moment de l'ouverture glottale maximale par rapport à la durée moyenne des constrictives [f þ h].