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) Redundancy of the phonetic features

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) Contact features

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.

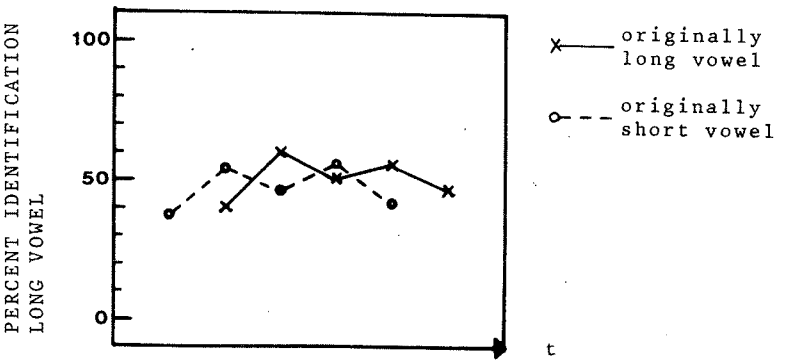
Languages with phonological long and short vowels (quantity languages) very often show some differences in formant structure between both categories of vowel length (cf Fant et al. 1969 for Swedish, Fischer-Jørgensen 1972 for Danish, Lehiste 1970 for Czech and Serbo-Croatian).

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) Redundancy



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):

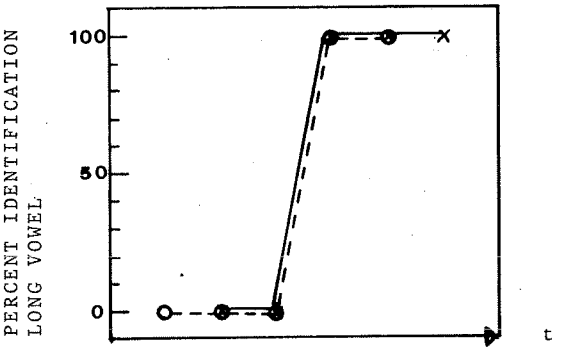
> "... suprasegmental features (including time, RB) can only be identified by comparison of items in sequence, and thus differ in a very essential way from features that may be identified by inspection of a segment (Jakobson, Fant, and Halle 1952; Lehiste 1967a)."

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.

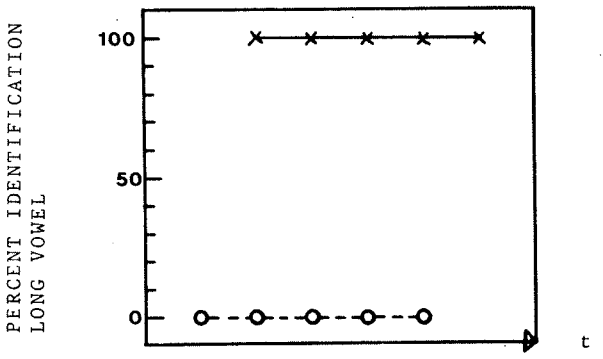
(ii) Segment features

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:

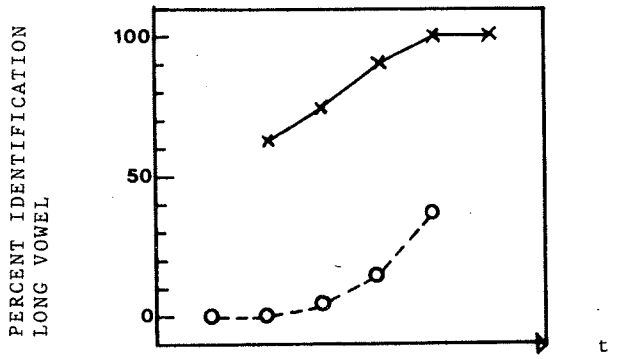
(a) Listeners identify the words exclusively from vowel duration.



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

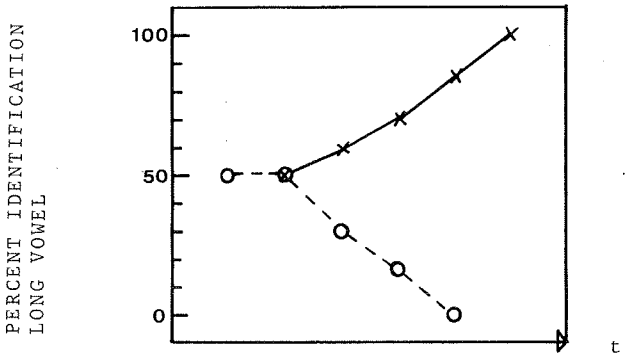


(c) Both features may of course counterbalance each other or interfere with each other, causing at least parts of the curves to deviate from their expected course towards the other alternative as indicated in the following figure:



(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 listen-

ers (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	Duration of vowel stimuli (msec)				
Original	F_1	F_2	F_3	(msec)	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

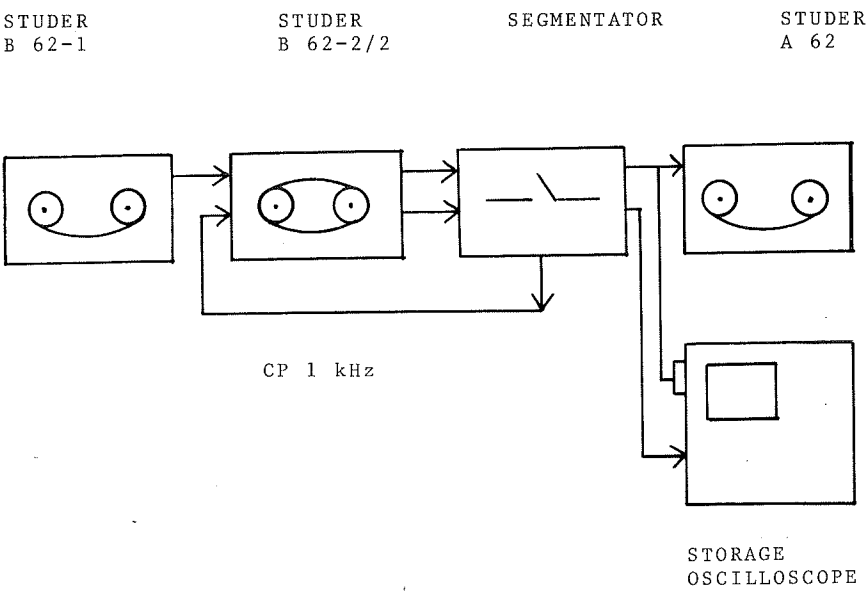
According to the test strategy outlined above the stimuli consist of the initial consonant followed by portions of the stressed vowel differing in duration, while the postvocalic consonant and the rest of the word are eliminated. The stimuli are of the structure /CV/. Starting from the pre-recorded natural 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 suddenly 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?

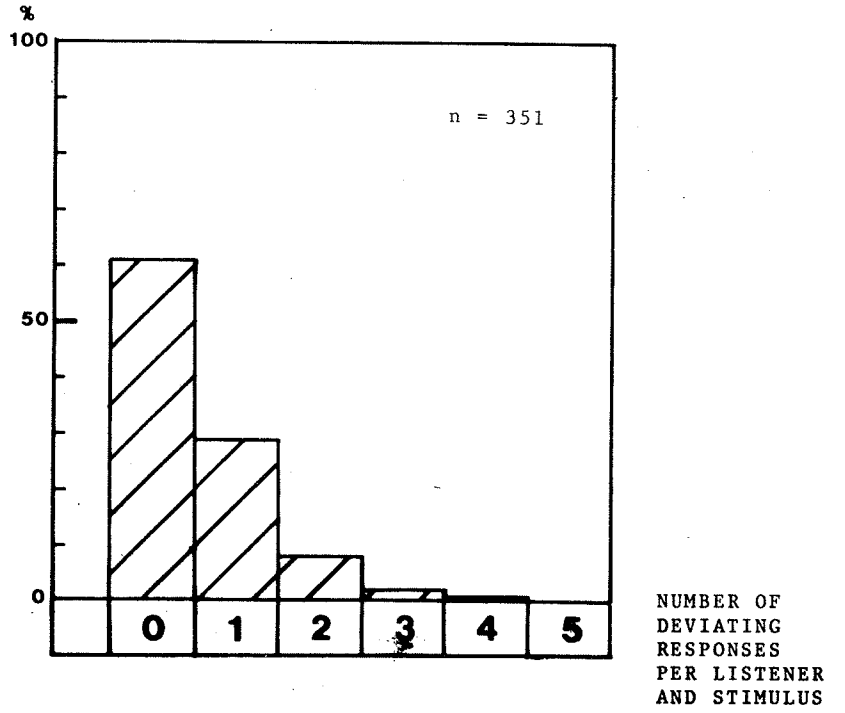
(b) How difficult was the identification of each stimulus by the whole group? Is each vowel identified to the same degree by each listener?

(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 vow-

el 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.

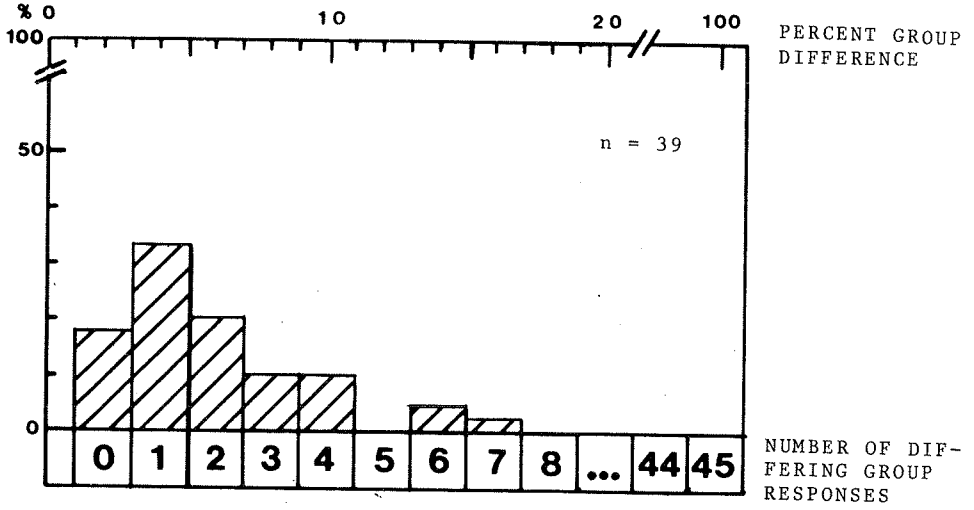
(c) The group

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

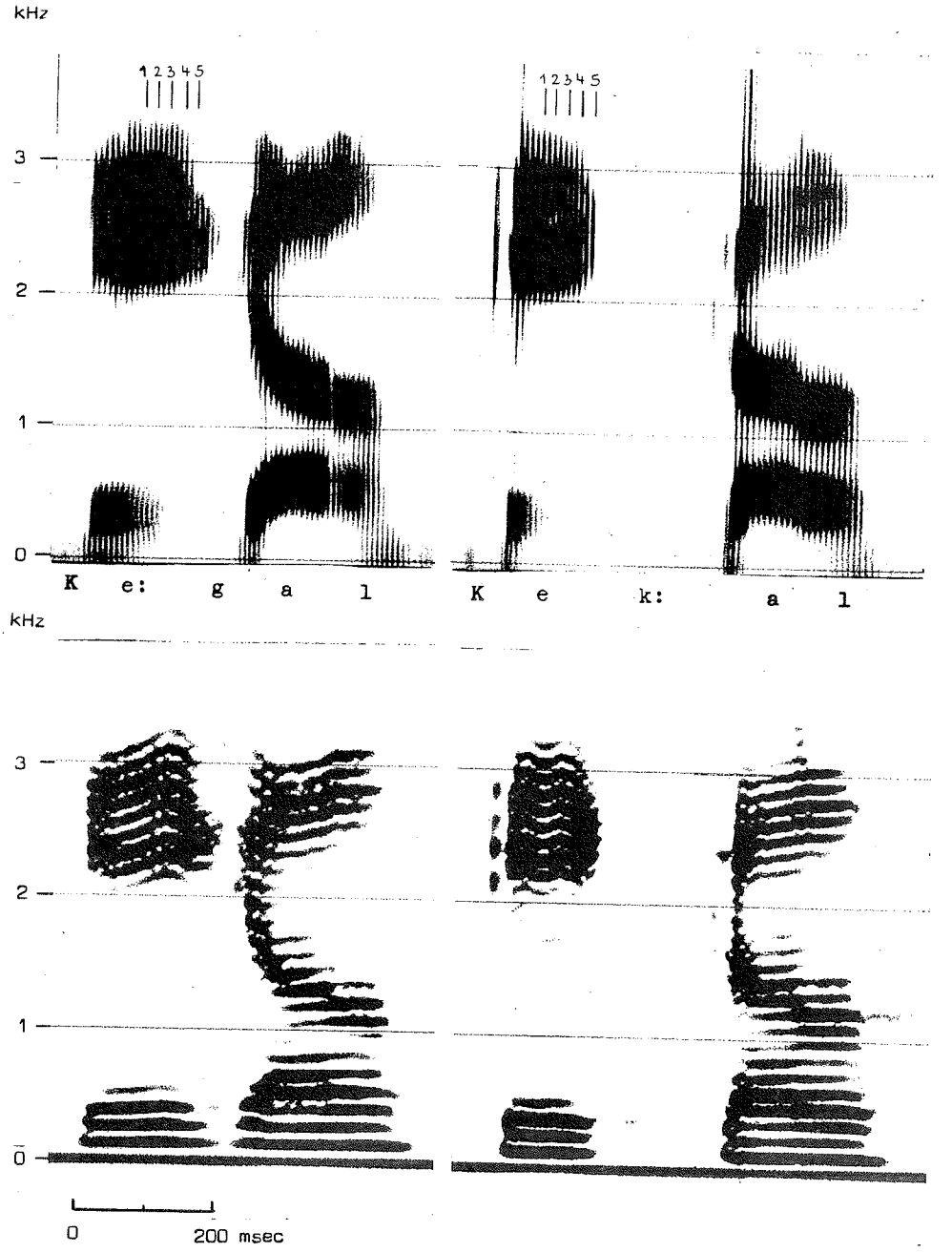


Figure 1a. Broad-band and narrow-band spectrograms of Gegal and Gekal. The vowel portions (stimuli 1-5) are indicated.

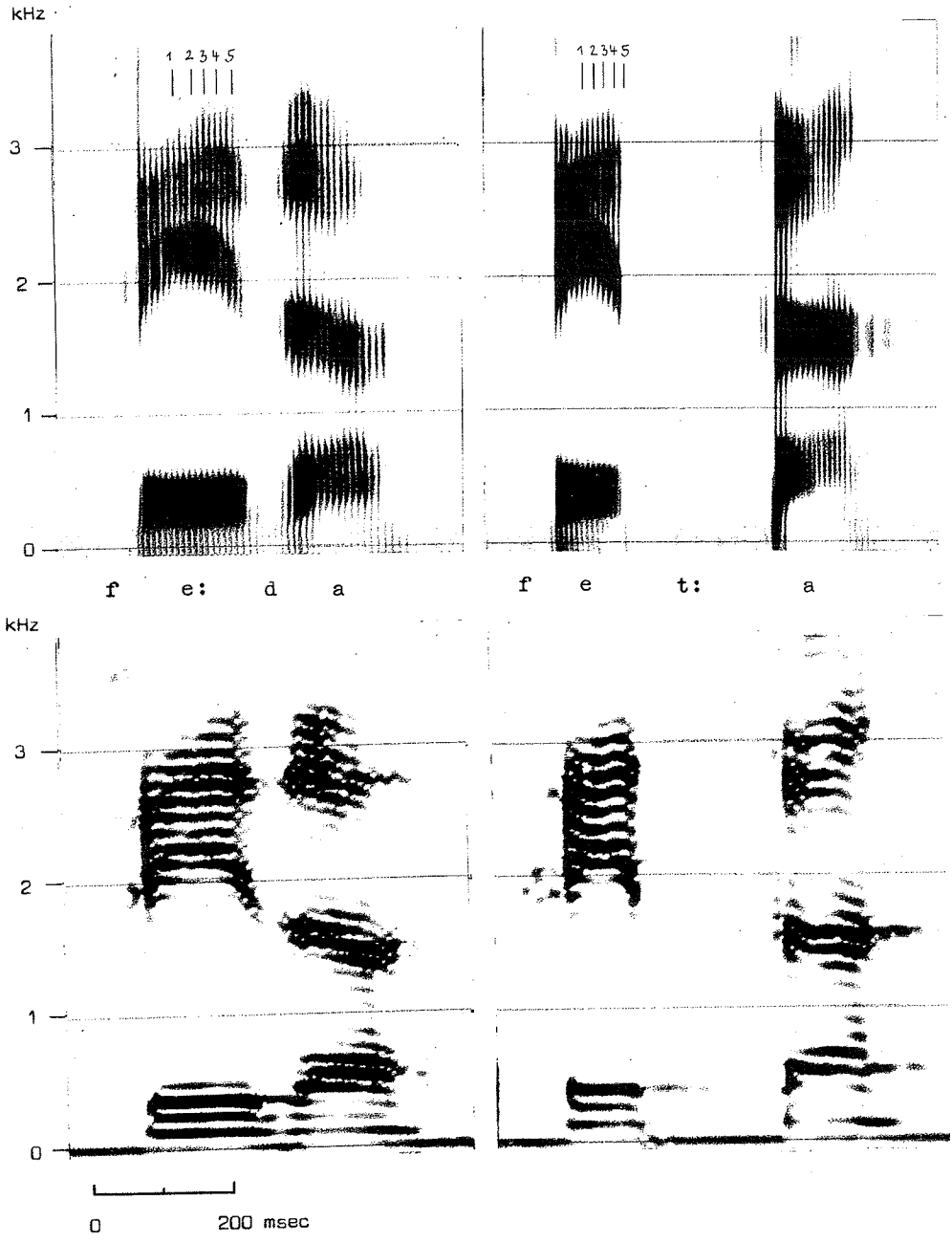


Figure 1b. Spectrograms of Feda and Feta. The vowel portions are indicated.

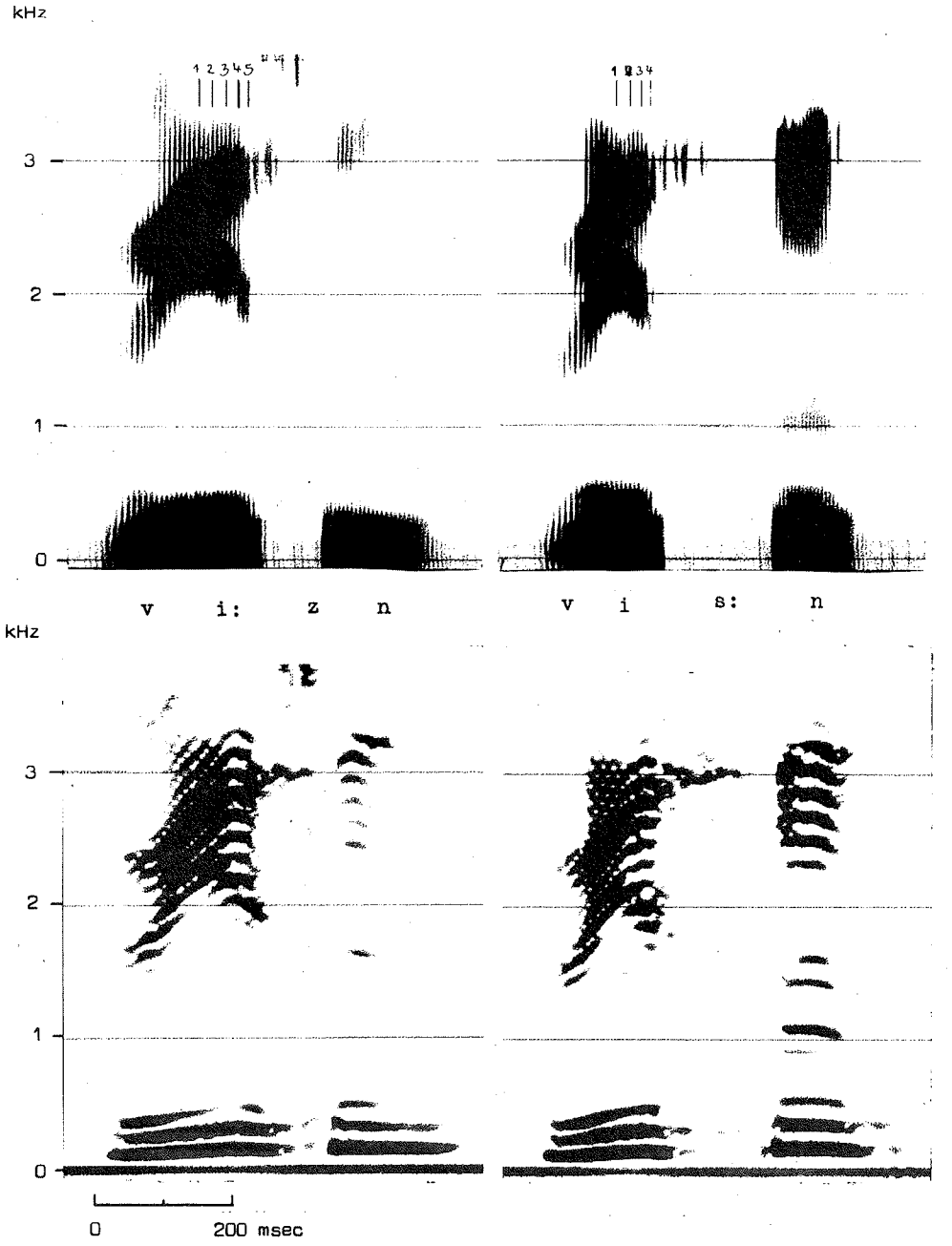


Figure 1c. Spectrograms of Wiesn and wissn. The vowel portions are indicated.

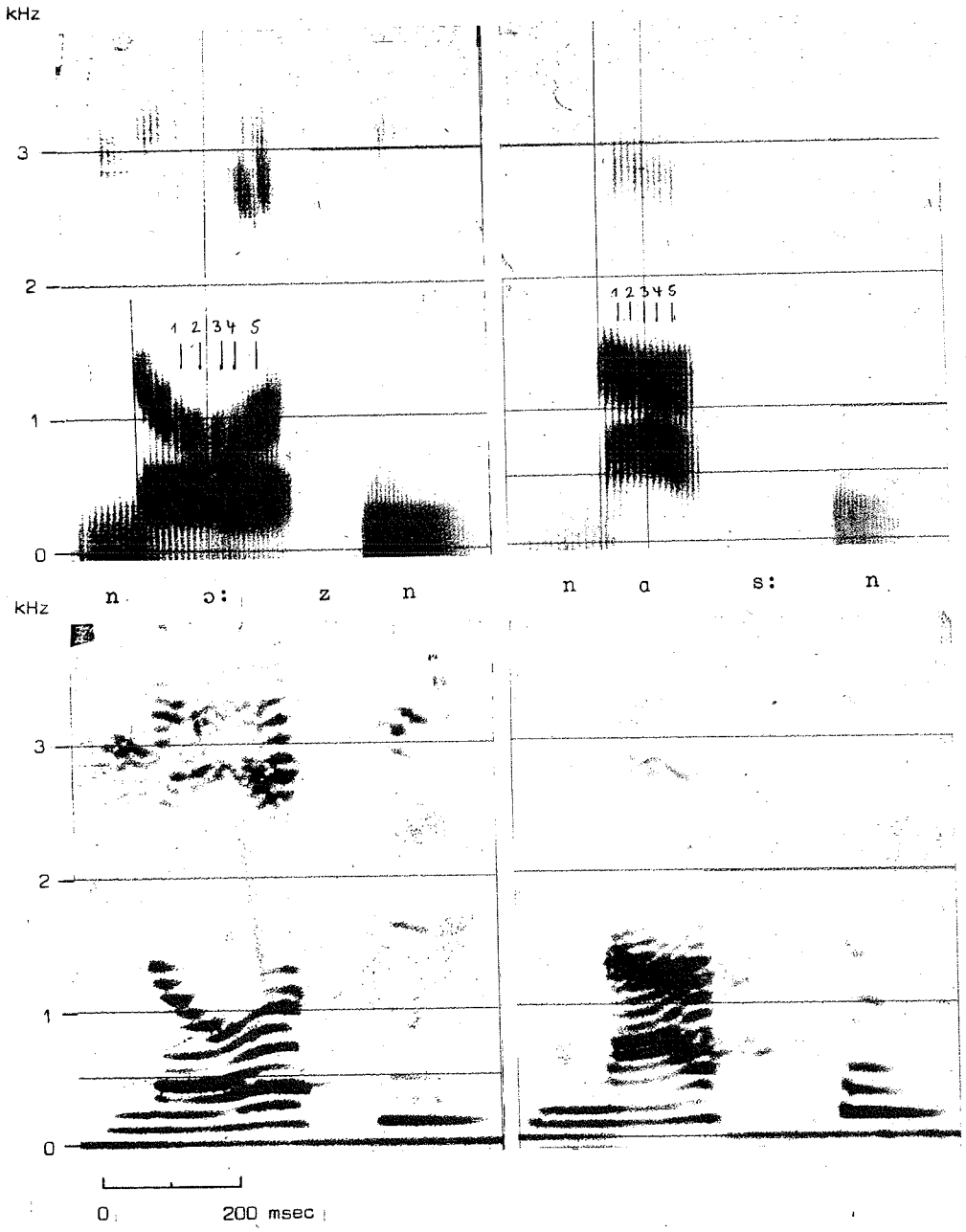


Figure 1d. Spectrograms of Nasn and nassn. The vowel portions are indicated.

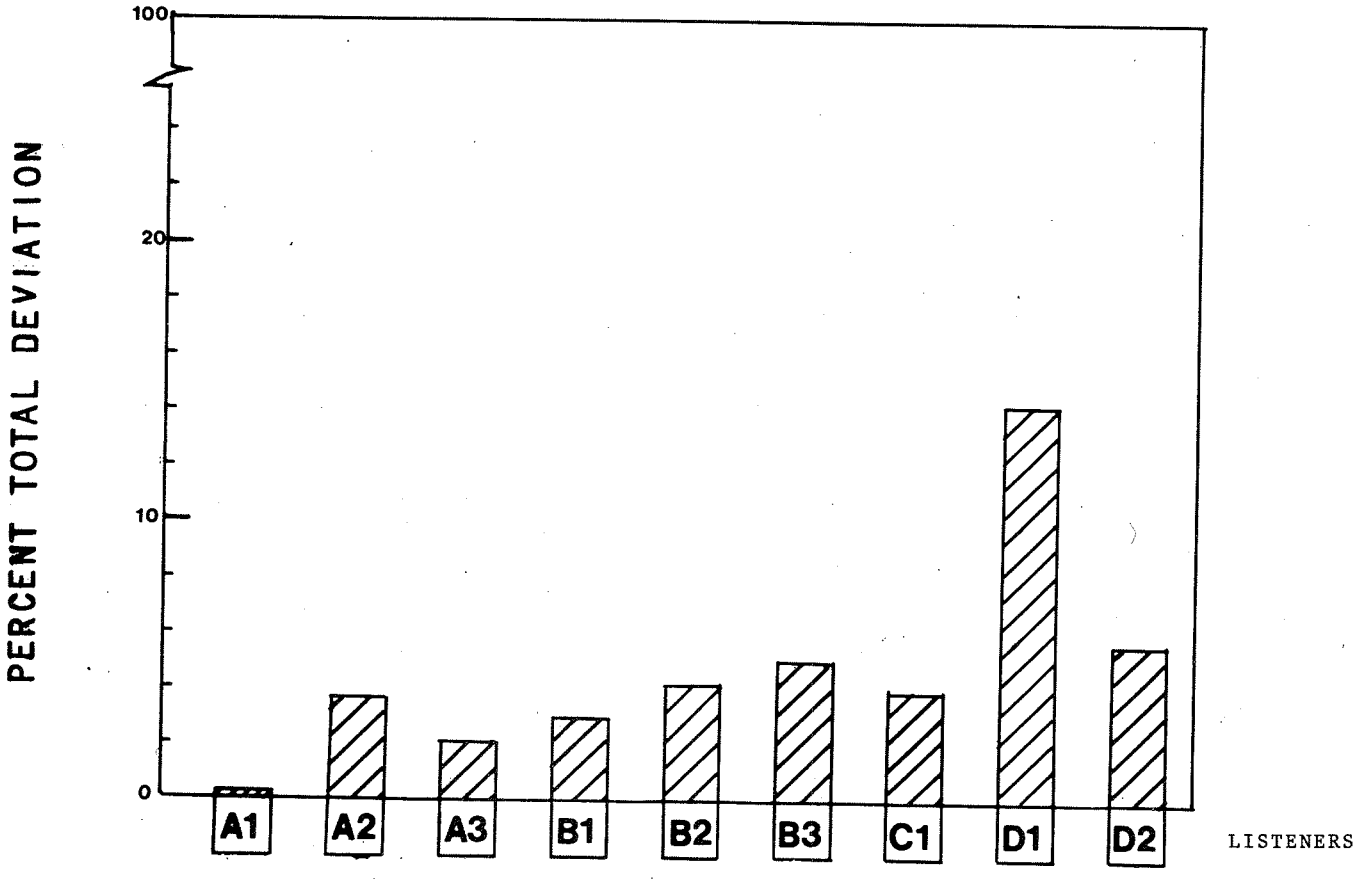


Figure 2. Total deviation of each listener as a measure of reliability of judgment.

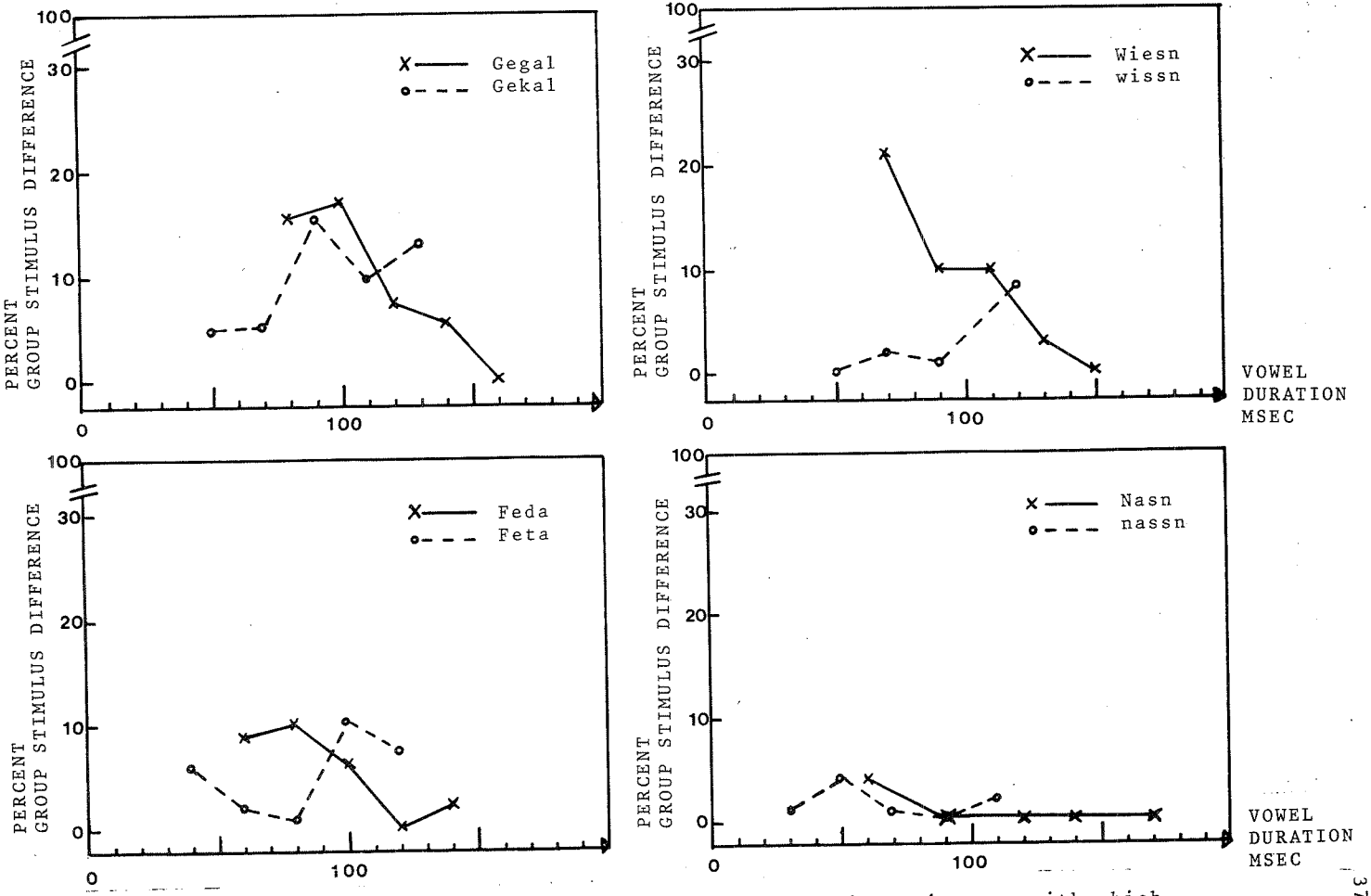


Figure 3. The group stimulus differences expressing the inconsistency with which each stimulus was indentified by all the listeners.

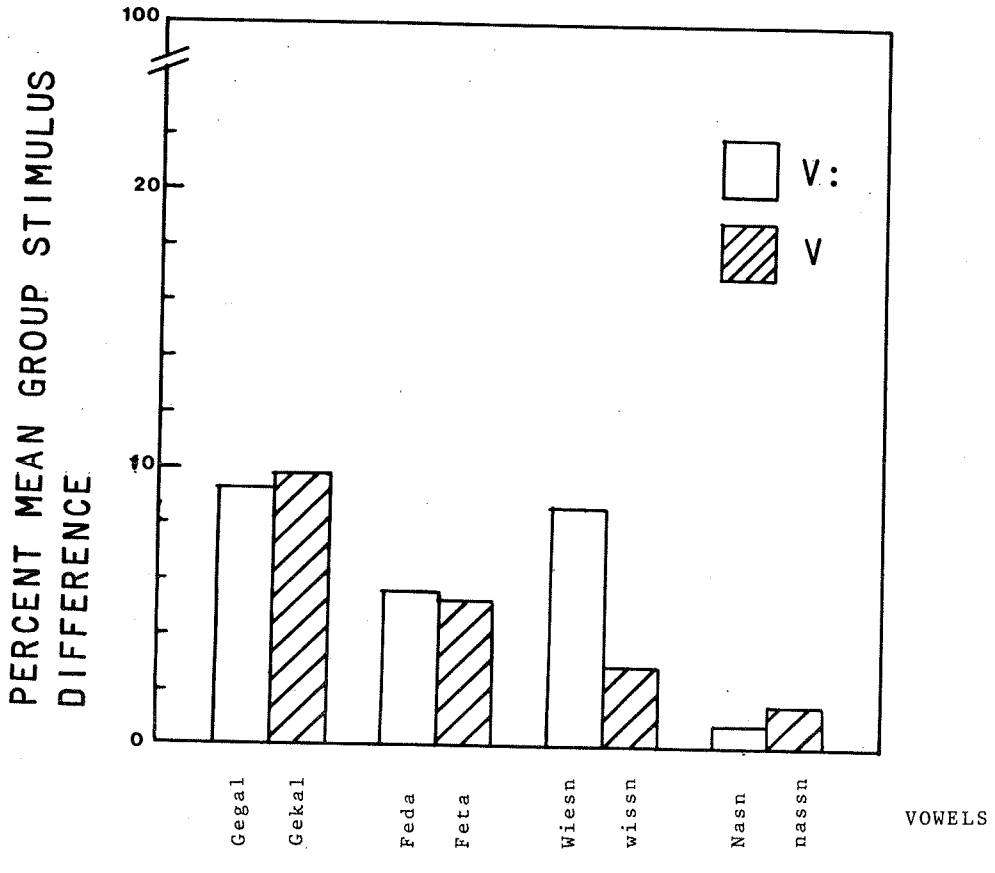


Figure 4. Means of group stimulus differences for each vowel.

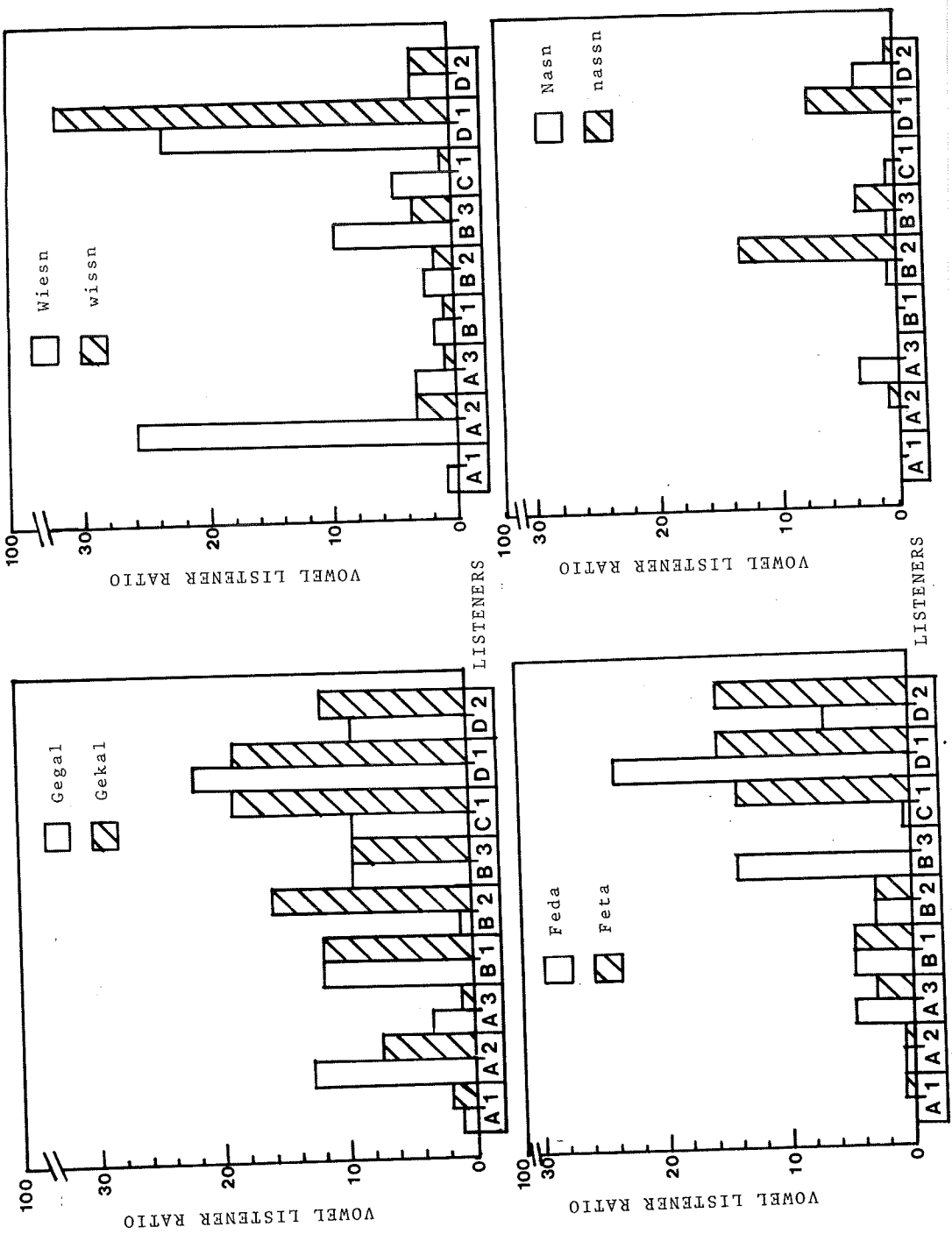


Figure 5. The vowel listener ratios as a measure of uncertainty of identification of each vowel by each listener.

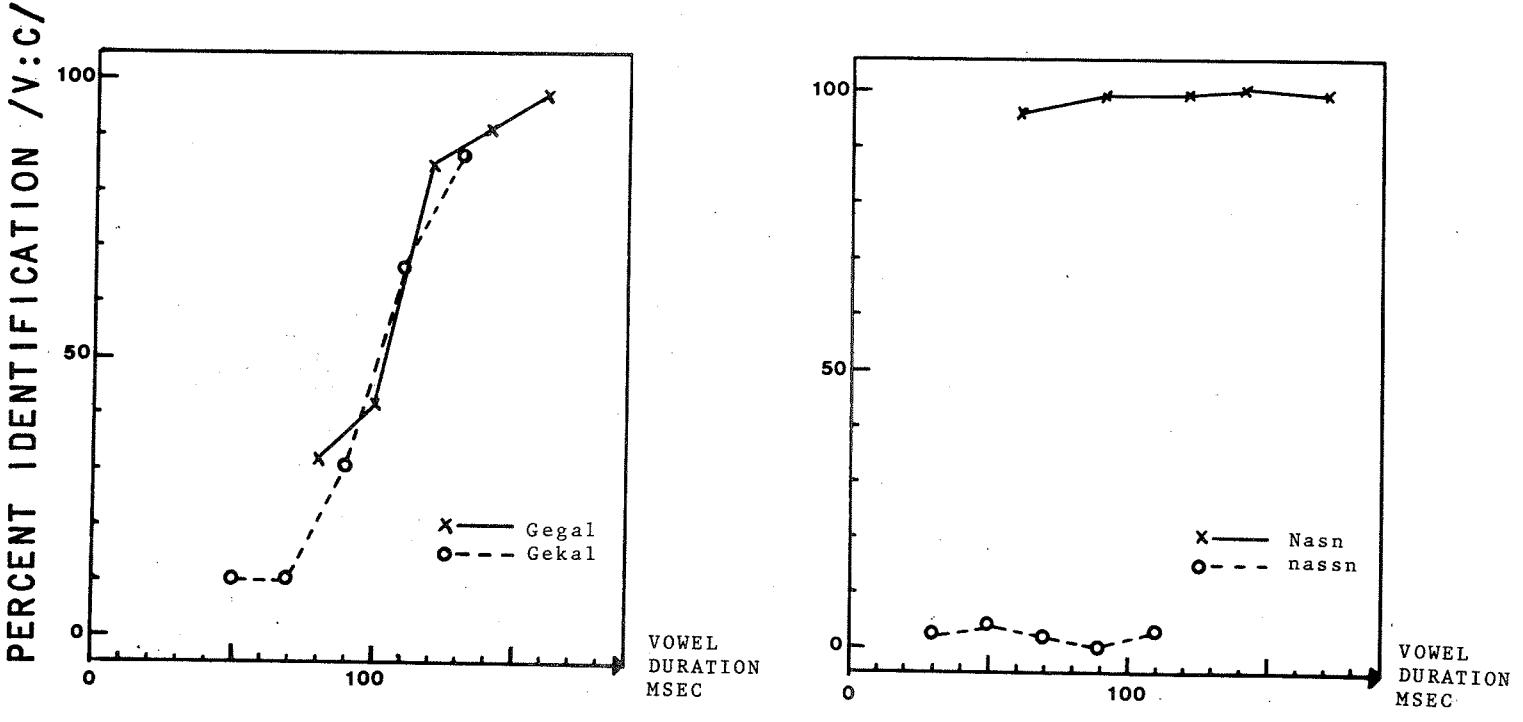


Figure 6a. Identification of the vowels of Gegal-Gekal (left) and Nasn-nassn (right) as /V:C/.

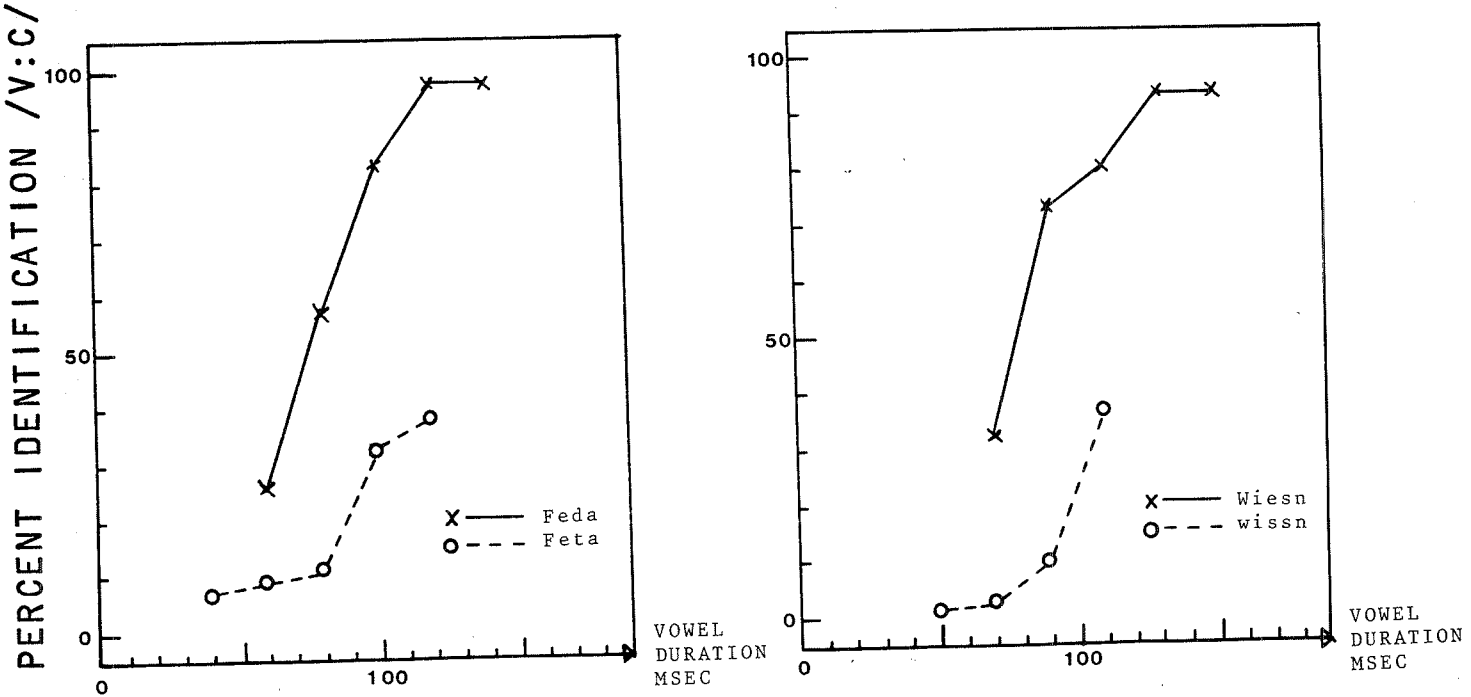


Figure 6b. Identification of the vowels of Feda-Feta (left) and Wiesn-wissn (right) as /V:C/.

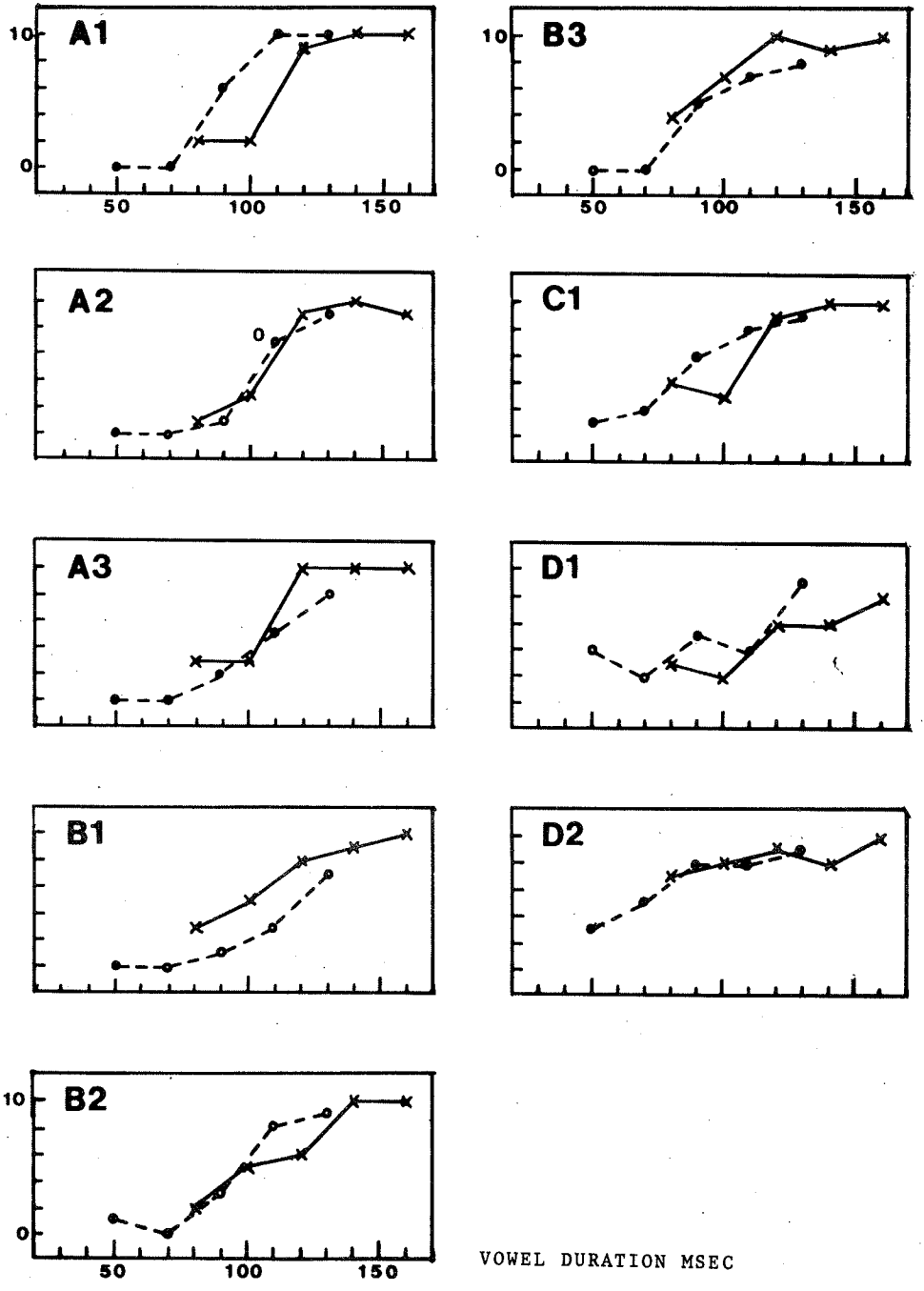


Figure 7a. Identification curves of Gegal and Gekal as /V:C/ for eac listener (both runs, totally 10 responses per stimulus). Dashed line: originally short vowel, solid line: original long vowel.

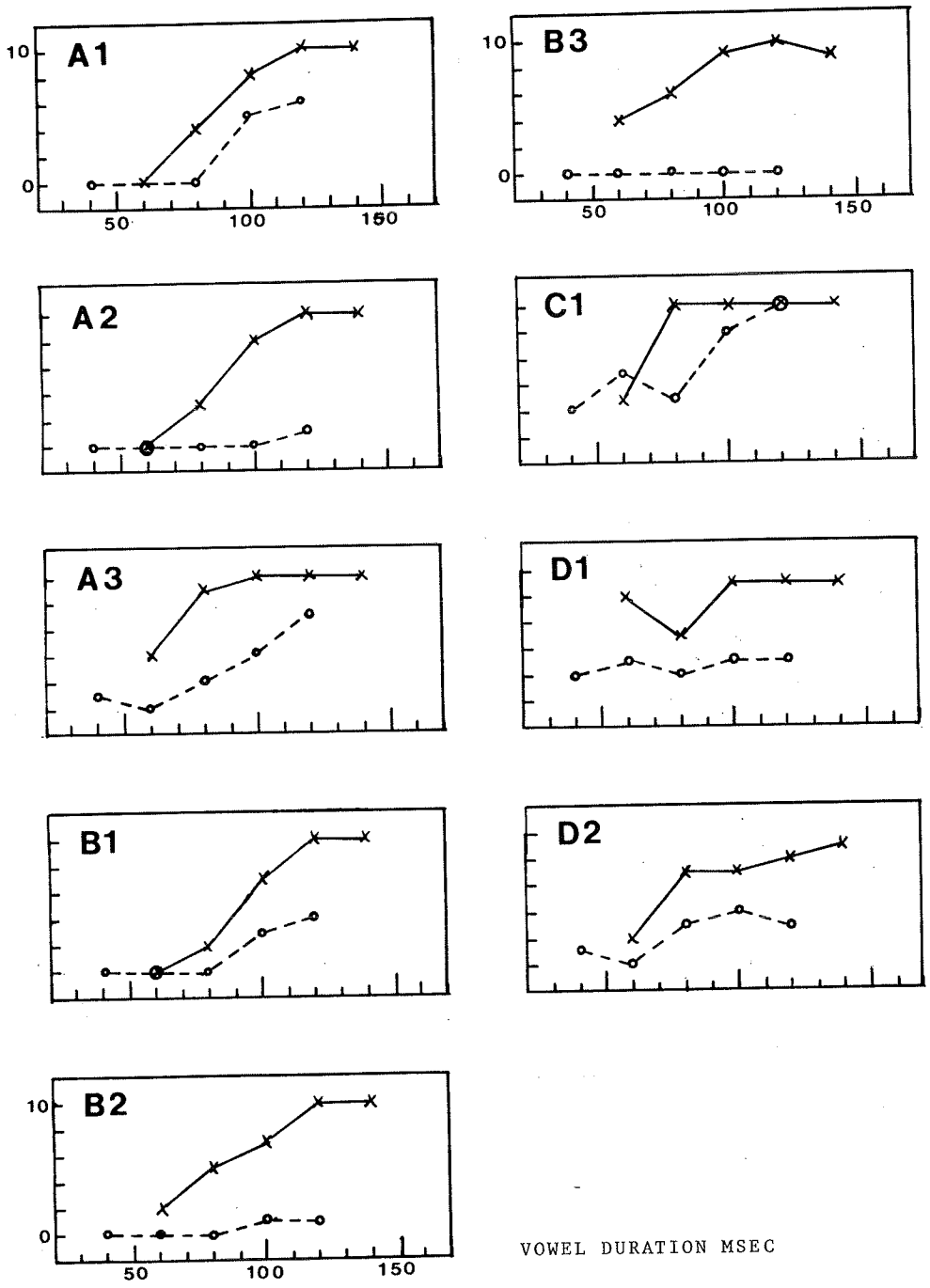


Figure 7b. Identification curves of Feda and Feta as /V:C/ for each listener.

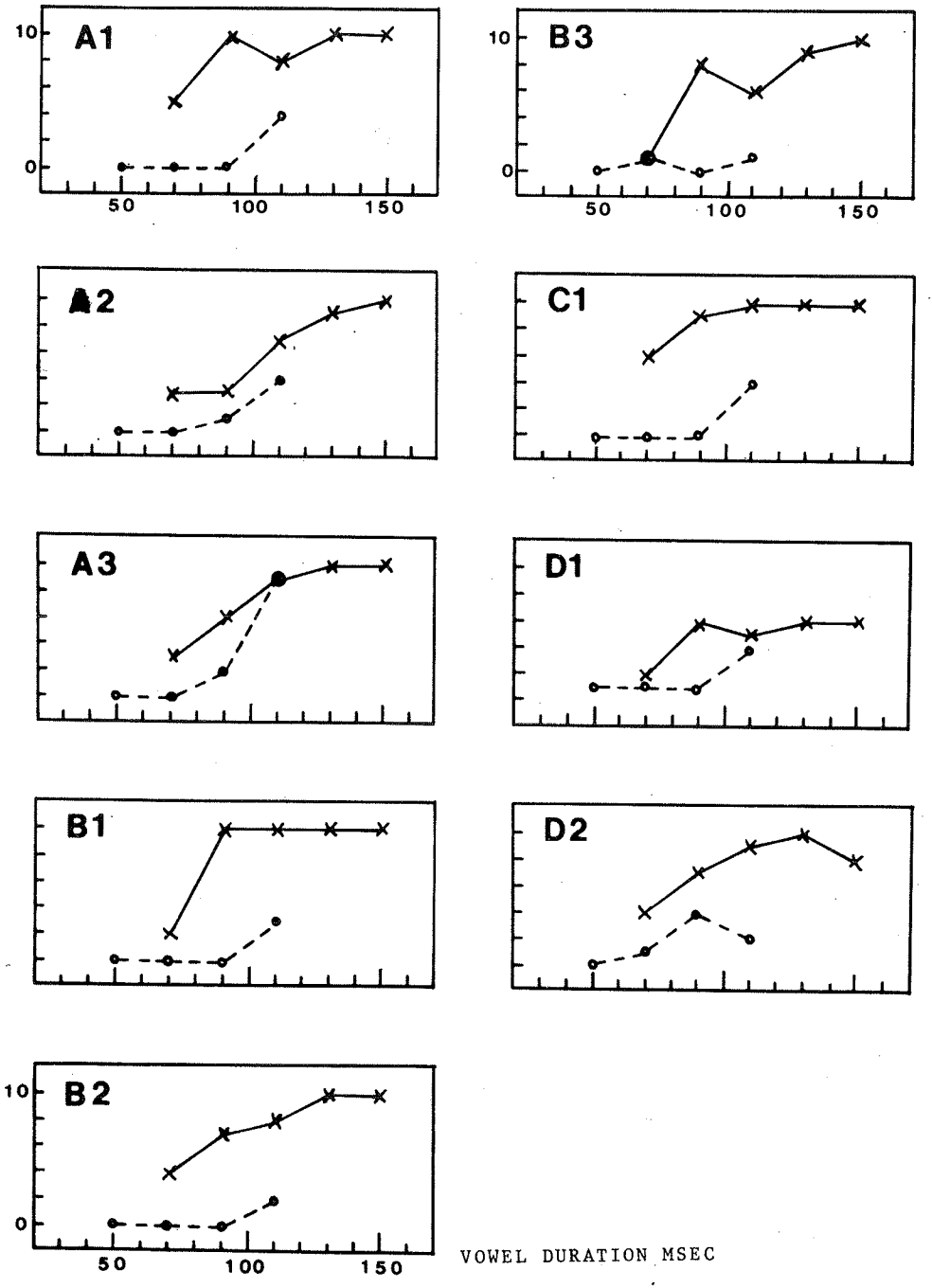


Figure 7c. Identification curves of Wiesn and wissn as /V:C/ for each listener.

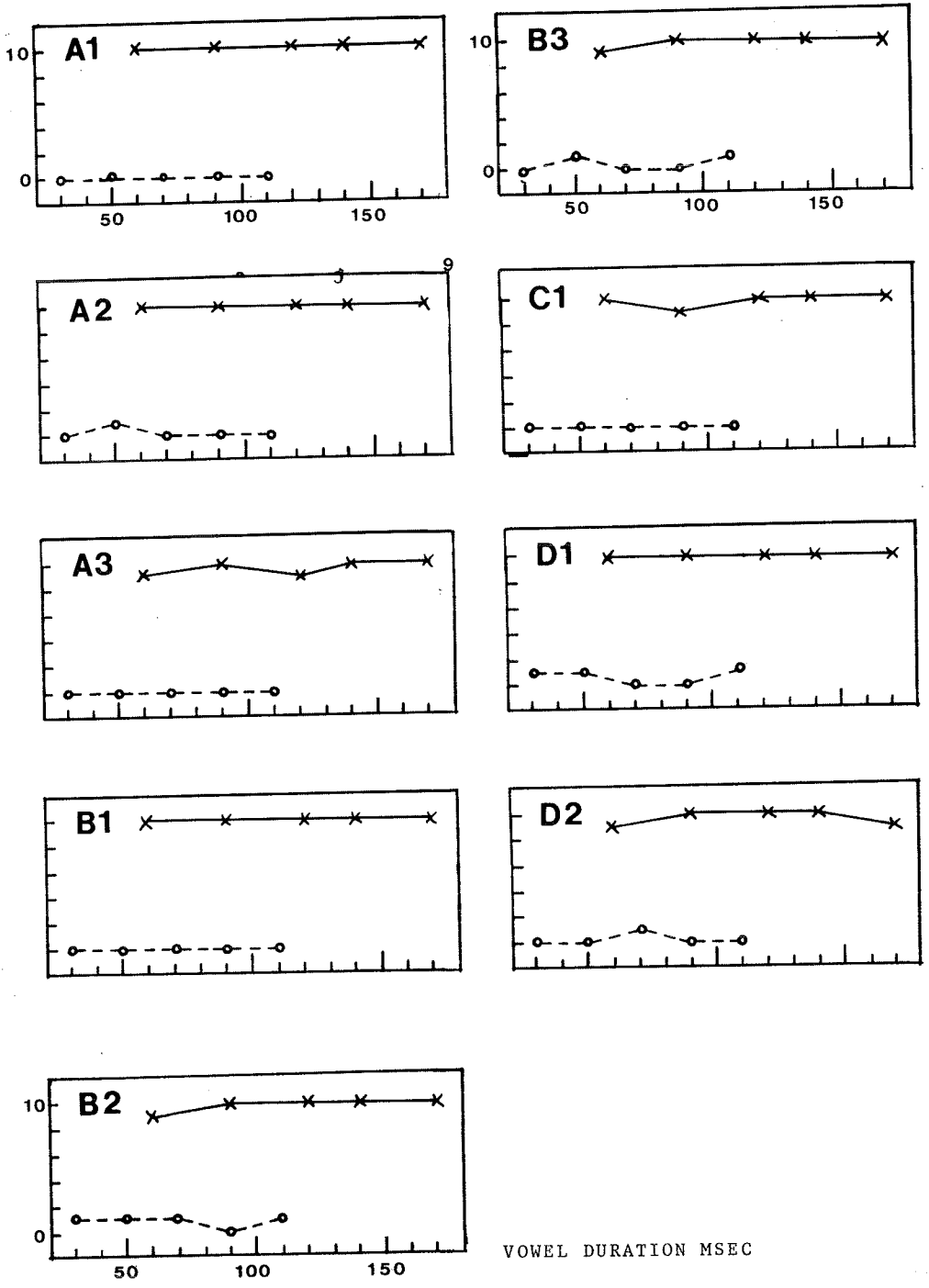


Figure 7d. Identification curves of Nasn and nassn as /V:C/ for each listener.