(ARIPUC 14, 1980)

AN ACOUSTIC INVESTIGATION OF INTRINSIC VOWEL DURATION IN DANISH

Michael Bundgaard

Abstract: The duration of twelve short and long stressed vowels in disyllabic nonsense words, embedded in a carrier sentence, recorded 10 times by each of five speakers, was measured. The vowels can be grouped (roughly) in five tongue height categories, and duration was found to increase significantly from category 1 (highest) through 5 (lowest), in about 1 cs steps. Unstressed vowels show the same tendency but the increase in duration with lower tongue height is considerably smaller. Variation in speaking rate (even rather considerable) did not significantly influence the relationship between high and low vowels. In contradistinction to earlier investigations, it seems that the difference in duration between long and short vowels is constant over different tongue heights, i.e. $V_{short} = V_{long} - b$, where b is a constant (approximately 5 cs).

1. Introduction

Intrinsic vowel duration is the object of several investigations in several languages: Danish - Fischer-Jørgensen (1955 and 1964), German - Maack (1949), English - Peterson and Lehiste (1960), and Thai - Abramson (1962). All agree that there is a universal tendency for vowels with high tongue position to be of shorter duration than vowels with lower tongue position, everything else being equal.

Various explanations for this correlation between tongue height and vowel duration have been offered. Fischer-Jørgensen (1964) prefers the hypothesis, advanced by Jespersen (1926), that low vowels require more time for the speech organs to reach their target positions than do high vowels. Lindblom (1967) adheres to the same theory, and in his model of lip and jaw co-ordination vowel duration increases as a function of increased jaw opening. The jaw is described as a damped spring-mass system, where the differences in slaphed

duration are due to inertia in the system, but on a higher level of production control, different vowel qualities are presumably normally programmed with equal durations.

The purpose of the present investigation is to examine the relationship between tongue height and vowel duration in Danish vowels in stressed (and unstressed) syllables, and to investigate the relation between long and short vowels. Danish vowels have been accounted for in these respects previously by Fischer-Jørgensen (1955, 1964) and by Holtse (1977), and my experiments are intended as a supplement to their investigations: The speakers in this study are all fairly young, and they speak Advanced Standard Copenhagen Danish (ASC - see Basbøll 1968). Fischer-Jørgensen's subjects all belong to an older generation, and both her and Holtse's subjects represent different dialects. Holtse's study examines only 4 vowel qualities, Fischer-Jørgensen's includes 10, whereas 12 vowel qualities are involved in this material.

2. Experimental procedure

2.1 Material

The material consists of 12 long and 12 short vowels:

[i: e̞: ɛ̞: æ: ɑ: y: ø: œ̞: u: o̞: ɔ̞+: ɒ̞:]

[i e̞ ɛ̞ a ɑ y ø œ̞ u o̞ ɔ̞+ ʌ̝+][1]

According to Thorsen and Thorsen (1978), these vowels can be placed in the Cardinal Vowel Diagram as seen in fig. 1.

Long and short vowels in ASC Danish have not been subjected to acoustic analysis, but the quality differences between long and short is generally considered to be very small in ten of the twelve pairs above. The difference is greater between [æ:] and [a], and between [ɒ:] and [ʌ], but with regard to tongue height I shall assume that they are close enough to be treated as pairs. The listing and pairing is made on purely phonetic grounds. Thus, [æ:, ɑ:] and [a, ɑ] are variants of /a:/ and /a/, respectively,

---

1) The diacritics (except the length mark) are omitted in the following.

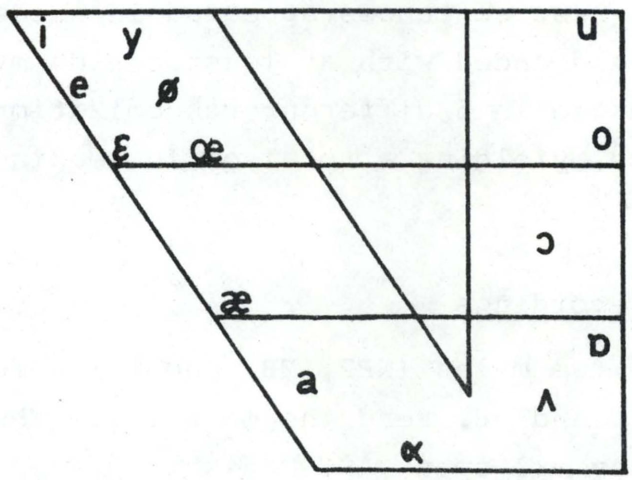

Figure 1

The positions in the Cardinal Vowel diagram of the vowels used in the experiment.

[o, ɔ] are variants of /o/, [ɔ:, ɒ:] are variants of /ɔ:/, and [ʌ] is the phonetic realization of /ɔ/.

The 24 vowels occurred in nonsense words of the type: ˈbVbə and were embedded in a frame sentence:

'Trykket i bVbə ligger på første stavelse.'
(The stress in ˈbVbə is on the first syllable.)

In a small supplementary material one of the subjects read the words biˈbibə and bɑˈbɑbə - in the same frame sentence - to examine the relationship between tongue height and duration in unstressed vowels. To avoid a rhythmic reading of the text, 14 dummy sentences were mixed with the test sentences, in such a way that no more than 5 test sentences appeared in succession, and every page started and ended with at least one dummy. The 24 + 14 sentences were arranged in 5 different randomizations in a list which was read twice, yielding a total of 10 readings by each subject.

2.2 Subjects and recordings

Five subjects, two males (NRP, MB) and three females (KM, ER, NT), aged between 24 and 38, read the material. They are all phoneticians and they all speak ASC Danish.

The recordings took place in a sound treated room at the Institute of Phonetics with professional equipment (Revox A700 tape-recorder, 7½ i.p.s., Sennheiser MD21 microphone, Agfa PE36 tape). The ten readings were obtained in one recording session which lasted from 20 to 30 minutes. This may be an upper limit before fatigue effects occur, but none of the subjects complained and they all declared that it was an easy task.

3. Registration and measurements

The same tape recorder was used for recording and replay. The tapes were processed by intensity and pitch meters (F-J Electronics) and registered on an Elema 800 mingograph at a paper speed of 100 mm/sec.[1], cf. fig. 2. One intensity curve is high-pass filtered at 500 Hz, integration time 2.5 ms, logarithmic display, the other has full frequency range, integration time 2.5 ms, and linear display. A short integration time facilitates an accurate segmentation, which is crucial for duration measurements.

Segmentation, however, is not the only problem in vowel duration investigations. One must also decide on a definition of vowel beginning and end: Peterson and Lehiste (1960) regard the aspiration of preceding stop consonants as being part of the vowel. Fischer-Jørgensen (1964) considers the vowel to start where the higher formants appear in the spectrum which is also the procedure adopted by Holtse (1977), and I have used the same criterion so that the previous and present results for Danish may be directly compared. Accordingly, the vowels are considered to begin where the high-pass filtered intensity curve rises sharply. This is approximately 1 cs after the oscillations start - and it corresponds to the point in time on sonagrams where the higher formants appear.

The vowels terminate where the intensity curves start to drop sharply. There were, however, several cases where this point was difficult to establish: when the following b was voiced in the beginning of the closure. I had expected that the high-pass filtered intensity curve would eliminate this voicing, but it did not do so completely, and since this curve did show normal decreases for nasals (in the dummy sentences), this energy cannot be due to deficient filtering (but it can be due, at least partly, to the logarithmic display): Fig. 3 shows mingograms and sonagrams of ˈby:bə and biˈbibə, respectively, which clearly demonstrate that e.g. the interval with lower, but level, intensity in i in fig. 3c must be due to the energy in the 'voice bar'.

Measurements were made in whole millimeters (centiseconds), i.e. the measuring accuracy is ±0.5 cs.

---

1) The paper speed was checked intermittently throughout registration, and turned out to be constant, with distances between the 1 Hz pulses of exactly 100 mm.

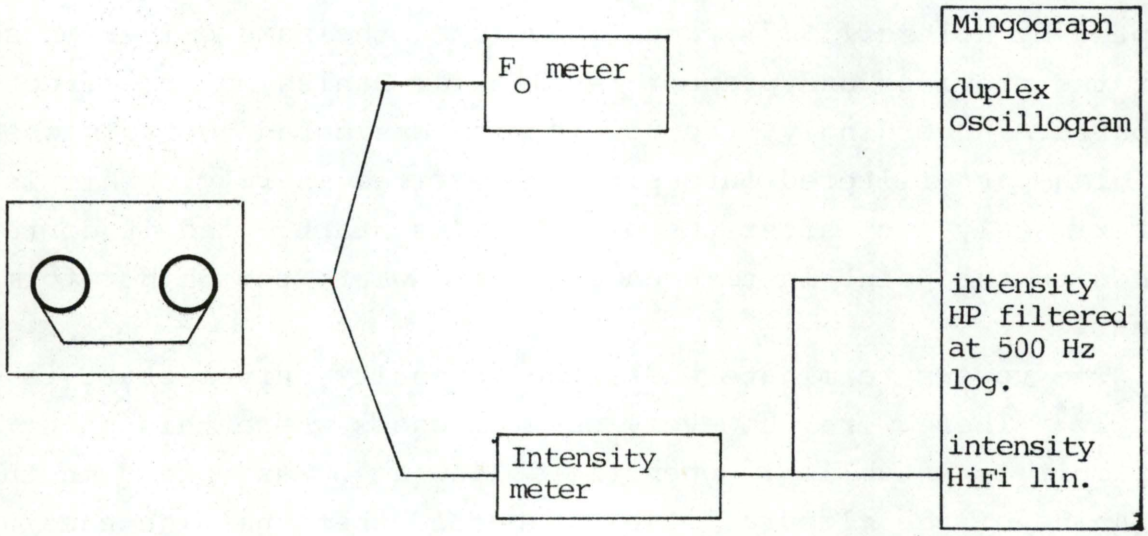

Figure 2

Block diagram of the experimental set up.

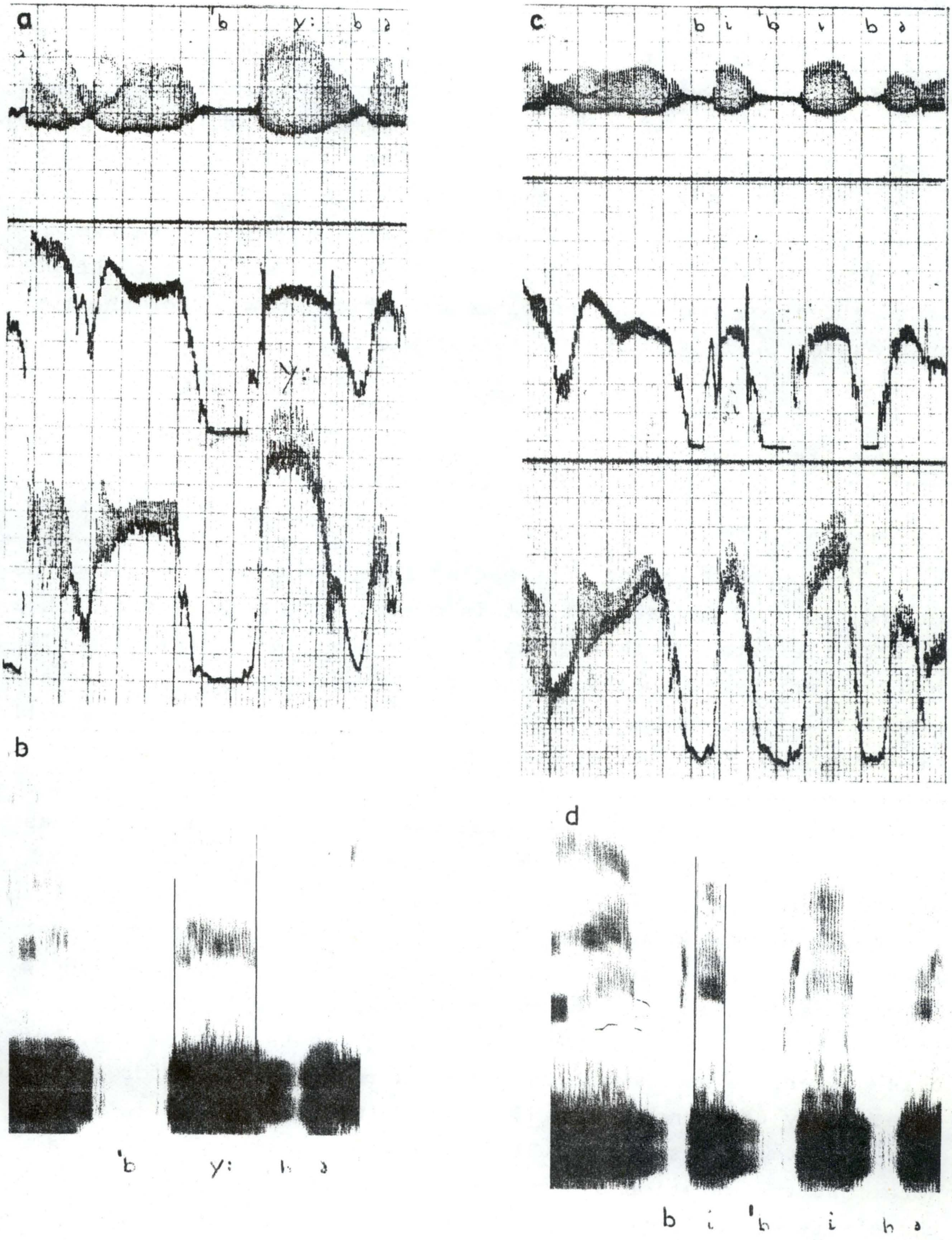

Figure 3

Mingogram (a) and spectrogram (b) of the word ˈby:bə, and mingogram (c) and spectrogram (d) of the word biˈbibə.

4. Statistical treatment

Each subject is treated separately. A one-way analysis of variance was performed to find out whether the durations obtained for the different vowels belong to the same or to different populations. Further, a multiple comparison procedure - Sheffé's method - (Ferguson 1971) was employed in which each vowel quality's mean value is compared to every other mean (within each of the two length categories). This was done in order to single out the possible effect of rounding and place of articulation. For the two unstressed vowels a Student's t-test was run.

5. Results

5.1 Speaking rate

Though no variation in speaking rate was observed during the recordings with any subject, it is of course possible that such a variation exists and that it will influence vowel durations. If vowel duration and speaking rate are correlated, and if it is a linear relationship, then homogeneity of the data (within as well as across subjects) can be achieved by applying a simple scaling procedure to the data.

Rather than measure the total duration of the utterances as an expression of speaking rate, the duration from the explosion of the utterance-initial t to the beginning of the first vowel in "første" was measured, in order to eliminate, to the extent that it is possible, an intra-utterance rate variation effect. There was a clear tendency in the material that the last word varied more in duration than did other words in the test sentences, but this difference has been eliminated in the expression for speaking rate. Only utterances with i, α, and u were measured.

Fig. 4 depicts the results for each subject and each vowel. Note that the duration of the test vowel has been subtracted from the entity which expresses speaking rate in order to eliminate a "double" effect from any correlation between vowel duration and speaking rate. Apparently, duration and rate are not correlated, neither within, nor across subjects. E.g. NRP has a rather great variation in speaking rate, but his vowel durations are nearly

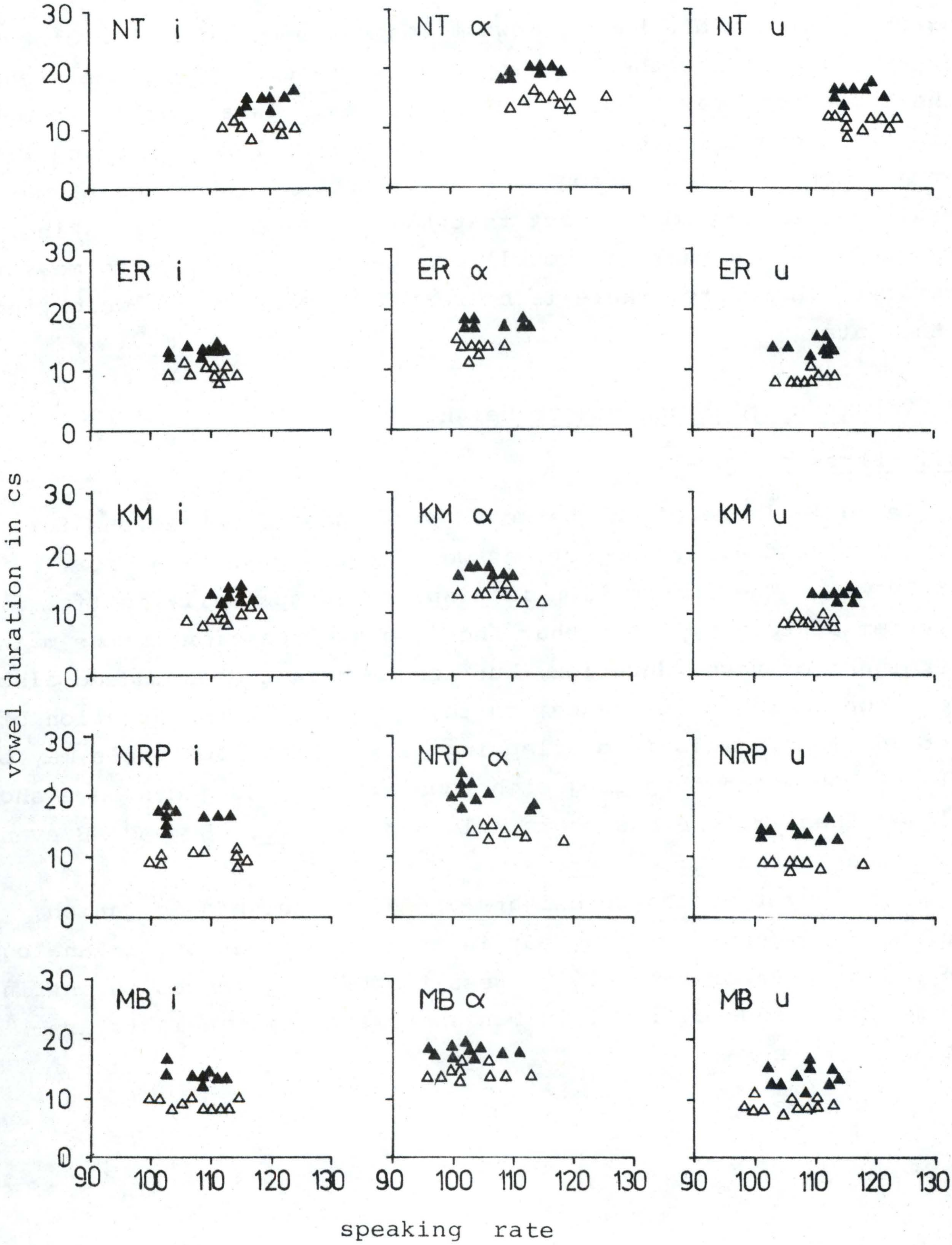

Figure 4

Vowel duration plotted as a function of speaking rate for the vowels i, α, and u. Filled and open triangles indicate long and short vowels, respectively. The derivation of the expression for speaking rate is explained in the text.

constant. NT and NRP have somewhat longer vowels than the other subjects, but NT has the highest speaking rate. KM has the second highest speaking rate, but the shortest vowel durations. Of course, it is true generally that vowels are longer at lower speaking rates. That no such correlation could be established in the present material must be due to the fact that the differences in speaking rate (and thus in the test vowel) are small in relation to measuring accuracy. Therefore, there is no reason to perform any weighting of the data.

5.2 Vowel duration and tongue height

5.2.1 Stressed vowels

In table 1 are given the means and standard deviations for every vowel and every subject, as well as the grand mean for every vowel. The information is displayed graphically in fig. 5. It is immediately apparent that the length opposition is the main determinant of vowel duration, but tongue height does indeed also have a considerable influence, to the effect that the duration of short and long vowels is overlapping, i.e. short low vowels may be of longer duration than long high vowels: Four subjects have short ɑ longer than i:, one has short ʌ longer than u:. ER and MB even have ɑ longer than o: and ø:.

Normally, four tongue heights are distinguished in Danish, high/semi-high/semi-low/low, but in this investigation, in analogy with Reinholt Petersen's (1977) results for differences in intrinsic fundamental frequency level in Danish vowels, five heights are distinguished here:

1.	i -	y -	u
2.	e -	ø -	o
3.	ɛ -	œ -	ɔ
4.	a/æ:	-	ʌ/ɒ:
5.		ɑ	

The vowels have been grouped, in table 1 and fig. 5, according to length (short/long) and in three series of unrounded front, rounded front, and back vowels, and within each series according to tongue height. The tendency for duration to increase with increased

Table 1

Means and standard deviations for short and long vowels for the individual subjects and for all subjects pooled. The leftmost column gives the F-values obtained in the one-way analyses of variance applied to the data of each subject.

			i	e	ɛ	a/æ	ɑ	y	ø	œ	u	o	ɔ	ʌ/ɒ	F
NT	short	$\overline{X}$	9.9	11.1	12.3	13.8	14.6	10.5	12.2	12.8	10.7	11.7	12.6	13.7	30.64
		s	0.74	0.74	0.82	0.63	1.08	0.71	1.40	0.63	0.95	0.68	0.70	0.68	
	long	$\overline{X}$	14.6	15.9	17.4	17.8	19.3	15.4	16.3	17.4	15.7	16.2	18.1	18.5	12.60
		s	o.97	0.88	0.70	1.14	0.82	1.90	1.25	0.70	0.82	1.32	0.99	0.85	
NRP	short	$\overline{X}$	9.6	10.5	12.6	13.6	14.0	10.0	10.6	12.4	9.3	9.9	12.1	14.1	28.75
		s	0.70	0.53	1.27	0.52	1.05	0	0.70	0.52	1.25	0.32	0.57	1.10	
	long	$\overline{X}$	16.1	16.1	17.1	18.3	19.6	15.7	16.5	17.3	14.0	15.5	17.8	17.6	13.17
		s	1.10	0.88	0.88	1.16	1.57	0.68	1.18	0.68	0.94	1.43	1.40	0.97	
KM	short	$\overline{X}$	9.4	10.5	10.9	12.6	13.1	9.4	9.9	10.8	8.9	9.7	9.9	12.6	32.75
		s	0.97	0.71	0.32	0.52	0.86	0.70	0.32	0.63	0.74	0.68	0.74	0.52	
	long	$\overline{X}$	12.7	13.2	14.3	15.2	16.3	12.5	13.4	14.9	12.8	14.3	13.9	14.5	17.12
		s	0.95	0.92	0.48	0.78	0.52	1.08	0.97	0.57	0.63	0.48	0.57	1.43	
ER	short	$\overline{X}$	9.6	9.8	11.3	12.7	13.8	9.2	9.5	11.8	9.2	8.5	10.9	12.7	30.51
		s	0.70	0.42	0.48	0.48	0.79	0.63	0.97	0.79	0.92	1.18	0.74	0.68	
	long	$\overline{X}$	13.0	14.0	14.9	15.7	17.3	13.3	14.4	15.0	13.5	13.5	15.3	15.2	10.26
		s	0.67	0.94	0.99	0.68	0.48	1.16	1.43	1.25	1.08	1.18	1.83	1.03	
MB	short	$\overline{X}$	9.0	1o.0	11.6	13.5	14.3	8.8	9.5	11.0	9.1	10.7	11.6	12.9	26.64
		s	0.94	1.25	0.52	0.85	0.82	0.79	1.08	0.82	0.88	0.68	1.35	0.74	
	long	$\overline{X}$	13.8	14.5	15.6	17.4	17.6	12.8	13.7	16.7	13.9	14.8	16.1	17.4	13.47
		s	1.o3	1.52	1.35	1.07	1.o8	0.42	0.67	0.68	1.45	1.14	1.60	0.84	
Grand mean	short	$\overline{X}$	9.5	10.4	11.7	13.2	14.0	9.6	10.3	11.8	9.4	10.1	11.4	13.2	
		s	0.33	0.51	0.70	0.55	0.57	0.67	1.13	0.87	0.72	1.14	1.06	0.63	
	long	$\overline{X}$	14.4	14.7	15.9	16.9	18.0	13.9	14.9	16.3	14.0	14.9	16.2	16.6	
		s	1.37	1.24	1.35	1.35	1.37	1.50	1.45	1.23	1.07	1.05	1.75	1.70	

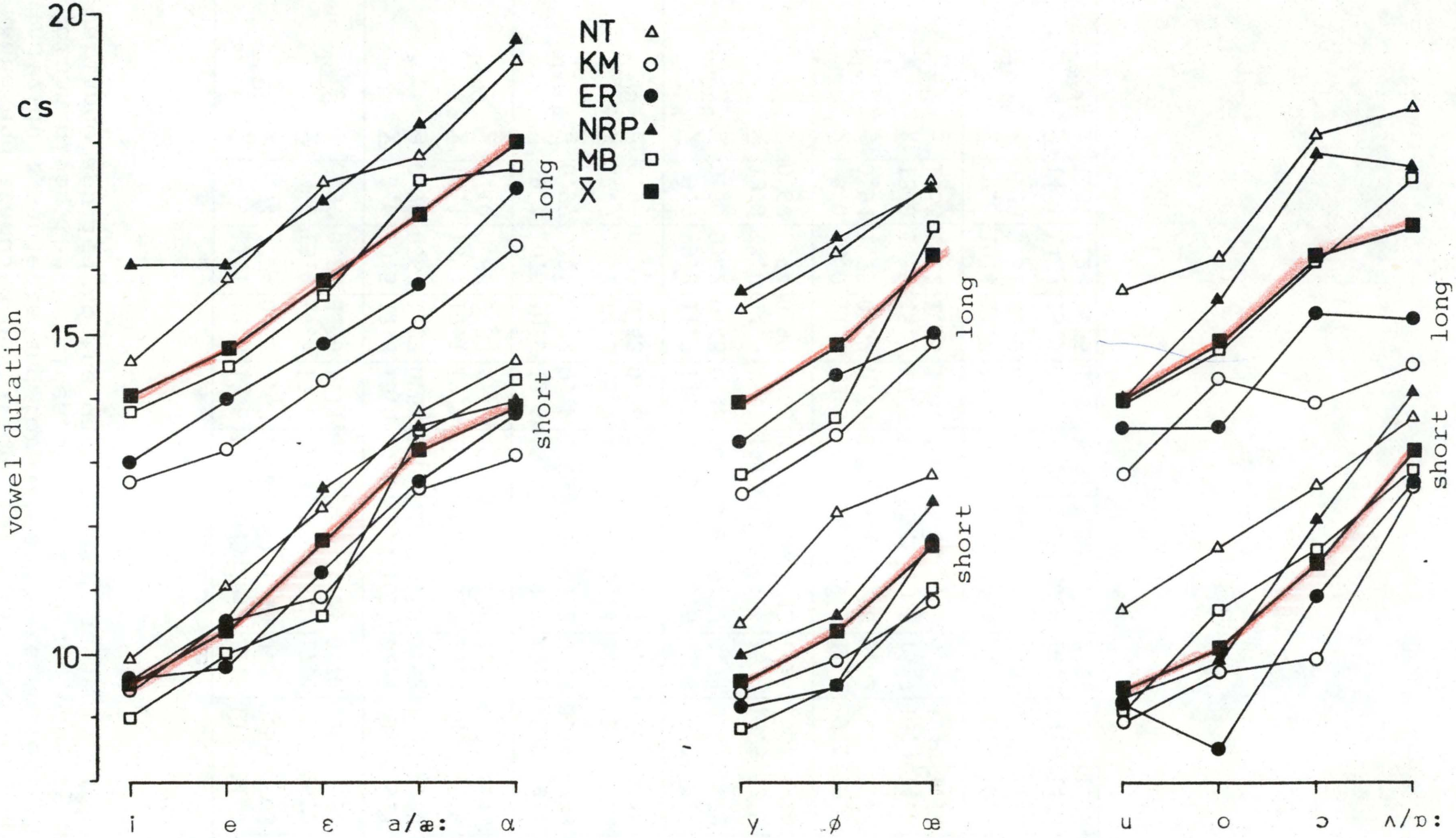

Figure 5

Mean durations for long and short vowels plotted for the individual subjects and for all subjects pooled (termed $\overline{X}$ and indicated by filled squares in the graph).

tongue height number is clear within each series, - there are only 6 mean values out of 120 that oppose this tendency, and no exceptions are found in the grand mean gennemsnit.

The linear dependency of vowel duration upon tongue height may be expressed as follows:

$$\text{Duration}_V = b + a \cdot \text{(tongue height)}$$

where a and b are constants, and tongue height is an integer ranging from 1 to 5. To simplify matters, the grand mean of long and short vowels, respectively, with the same tongue height was calculated, cf. table 2 and fig. 6. The slopes of the regression lines (least squares method) in fig. 6 are 1.0 and 1.2 for long and short vowels, respectively, and the scatter is small (the correlation coefficients are 0.995 and 0.992 for long and short vowels, respectively. Thus, the increase in duration is constant, about 1 cs per tongue height step, which is in agreement with Holtse's (1977) results, and he concludes that (my translation): "Roughly speaking one could say that the vowels are lengthened by 1 cs for every degree they are lowered". Note, however, that this formulation and conclusion imply that the jaw is lowered in equidistant steps from e.g. i through α, which we do not really know to be true, but we may, of course, accept the durational data as an indication of such med samme afstand equidistant tongue heights.

Holtse (1977), Fischer-Jørgensen (1964), and Reinholt Petersen (1974) find that the duration of long vowels increases more with lower tongue height than short vowel duration does. If anything, the tendency is in the opposite direction in this material. This question will be treated in further detail in section 5.3.

It appears from the multiple comparison test, cf. fig. 7, that place of articulation and rounding do not bidrager contribute to the durational differences between the vowels. This is in agreement with Fischer-Jørgensen (1964), who found no significant differences according to place of articulation or rounding in the position before b, but she did find such significant differences within vowels with the same tongue height in the position before d. The number of significant differences ($P < 0.05$) with the five subjects between every vowel compared to every other vowel within the same length category (short and long) are listed. Then the figures for

Table 2

Mean durations (in cs) for vowels having the same tongue height. All subjects pooled.

	tongue height number				
	1	2	3	4	5
Short vowels	9.51	10.27	11.64	13.22	13.96
Long vowels	13.99	14.82	16.12	16.76	18.04

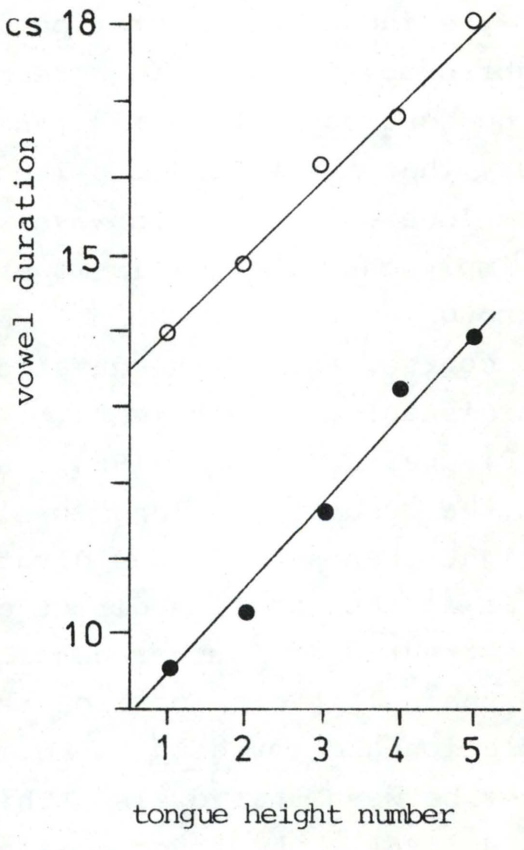

Figure 6

Vowel duration as a function of tongue height. Each data point represents an average over all subjects and all vowels having the same tongue height. Open and filled circles indicate long and short vowels, respectively. The straight lines are regression lines fitted to the data points.

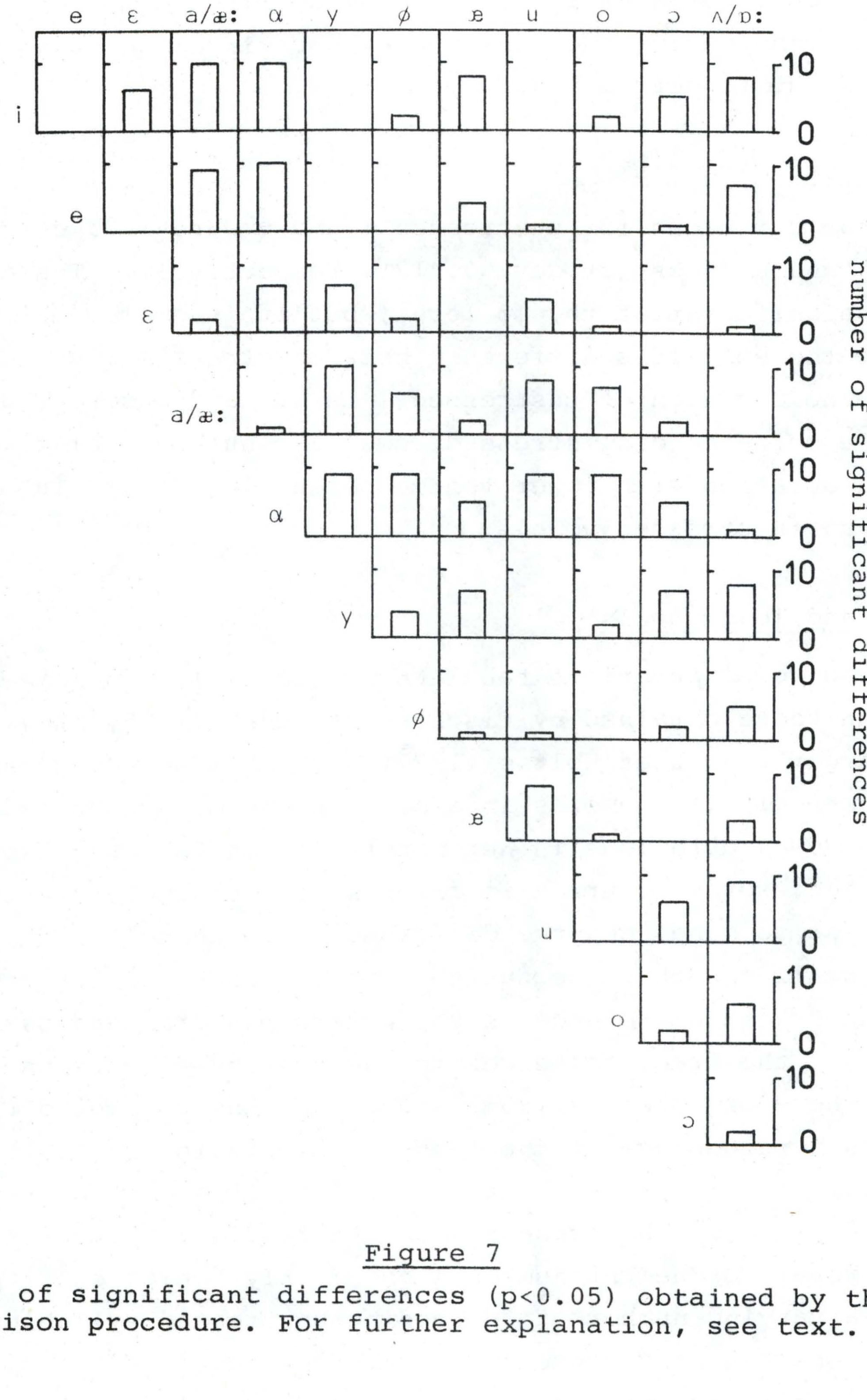

Figure 7

Number of significant differences ($p<0.05$) obtained by the multiple comparison procedure. For further explanation, see text.

long and short vowels are added, i.e. the maximum possible number of significant differences for each vowel quality is 10. Sheffé's method is a conservative one, which does not yield many significant differences, but it serves its purpose here.

5.2.2 Unstressed vowels

The mean durations for unstressed i and ɑ were 5.25 cs (st. dev. 0.635) and 6.35 cs (st.dev. 0.412), respectively. The difference is significant at the 1% level, but it is small (1.1 cs) compared to the subject's difference between stressed i and ɑ (4.7 cs). The duration of unstressed i is 53% and unstressed ɑ 43.6% of the corresponding stressed vowels. Thus the relative increase in duration with lower tongue height is smaller in unstressed than in stressed vowels.

5.3 Short and long vowels

The mean vowel durations tabulated in table 1, are somewhat smaller than those obtained by Fischer-Jørgensen (1955, 1964) and slightly larger than what Holtse (1977) got. Fischer-Jørgensen's test words were isolated words in a list, which may account for the greater vowel durations in her material. In Holtse's experiments, the test words occurred in frame sentences under conditions rather similar to mine, and the difference between Holtse's and my results is small enough to be due to differences between individuals and/or to a difference in the stress patterns and segmental composition of the frame sentences in the two investigations.

In Fischer-Jørgensen's (1955) study, the short vowel duration expressed as a percentage of the long vowel duration is 50.6%, in her 1964-study it is 57%; in Holtse's (1977) investigation the figure is 67%, and in the present study it is 71%. Holtse investigated vowels of height number 1 and 3 only, which occurred in words with a varying number of succeeding unstressed syllables and in varying consonant environments. The percentage for the corresponding vowels in my material is 69.5%, so Holtse's and the present results are in good agreement. It would have been tempting to explain the difference between Holtse's and my results as an indication of a relative lengthening of short vowels in ASC Danish, if it were not for the fact that Holtse's subjects were <u>not</u> ASC speakers, and the difference between his short:long fraction and the results that I have obtained (2.5%) is too small to be indicative of a relative lengthening of short vowels in ASC Danish. Fischer-Jørgensen's small values (as compared to Holtse's and mine) are due to the fact that long vowels are considerably longer in isolated words than in sentences, whereas short vowels are only slightly lengthened, cf. Fischer-Jørgensen's values (average over all vowels) of 13.9 cs and 26.9 cs (short and long vowels, respectively) as against 11.9 cs and 15.9 cs in the present material: the difference is tydelig evidently in the long vowels.

The relation between short and long vowels may be illustrated graphically as in fig. 8, where duration of the short vowels is depicted as a function of the corresponding long vowels. The regression line (least squares method) has the equation:

$$\text{short vowel} = 1.19 \cdot (\text{long vowel}) - 7.3$$

The fald slope of 1.19 deviates considerably from Holtse (1977) and Lindblom (1967), who both arrived at a slope of 0.75. Lindblom had only one subject, Holtse employed three, but Holtse calls attention to inter-individual differences and assigns the identity with Lindblom's result to a coincidence.

Even though I have treated the <u>a/æ:</u> and <u>ʌ/ɒ:</u> as short long pairs with the same tongue height, the short member of each pair does have a lower tongue position than the long one (cf. fig. 1), and thus presumably a longer intrinsic duration, which makes for relatively larger <u>y</u>-values in fig. 8. In fig. 9 those two vowel pairs have been excluded, and the regression line equation is now:

$$\text{short vowel} = 1.075 \cdot (\text{long vowel}) - 5.6$$

The slope is so close to unity that a simplification as follows seems justified:

$$V_{short} = V_{long} - b$$

where b is a constant, approximately 5 cs in this case. Thus, the relation between short and long vowels is constant over different vowel pairs. In fig. 10, the short:long difference of the 12 pairs, for individual subjects as well as the grand mean, is de-

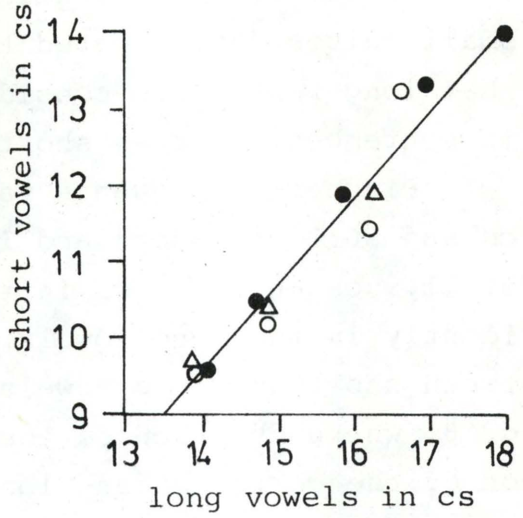

Figure 8

Short vowels plotted as a function of long vowels (all subjects pooled). Unrounded front vowels are indicated by filled circles, rounded front vowels by triangles, and back vowels by open circles.

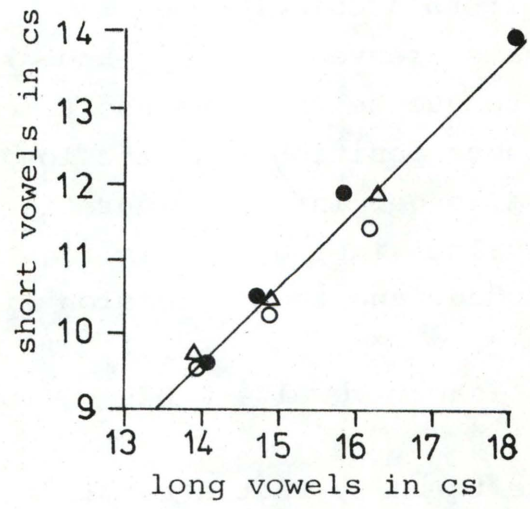

Figure 9

Same as fig. 8, but with the pairs a/æ: and ʌ/ɒ: left out.

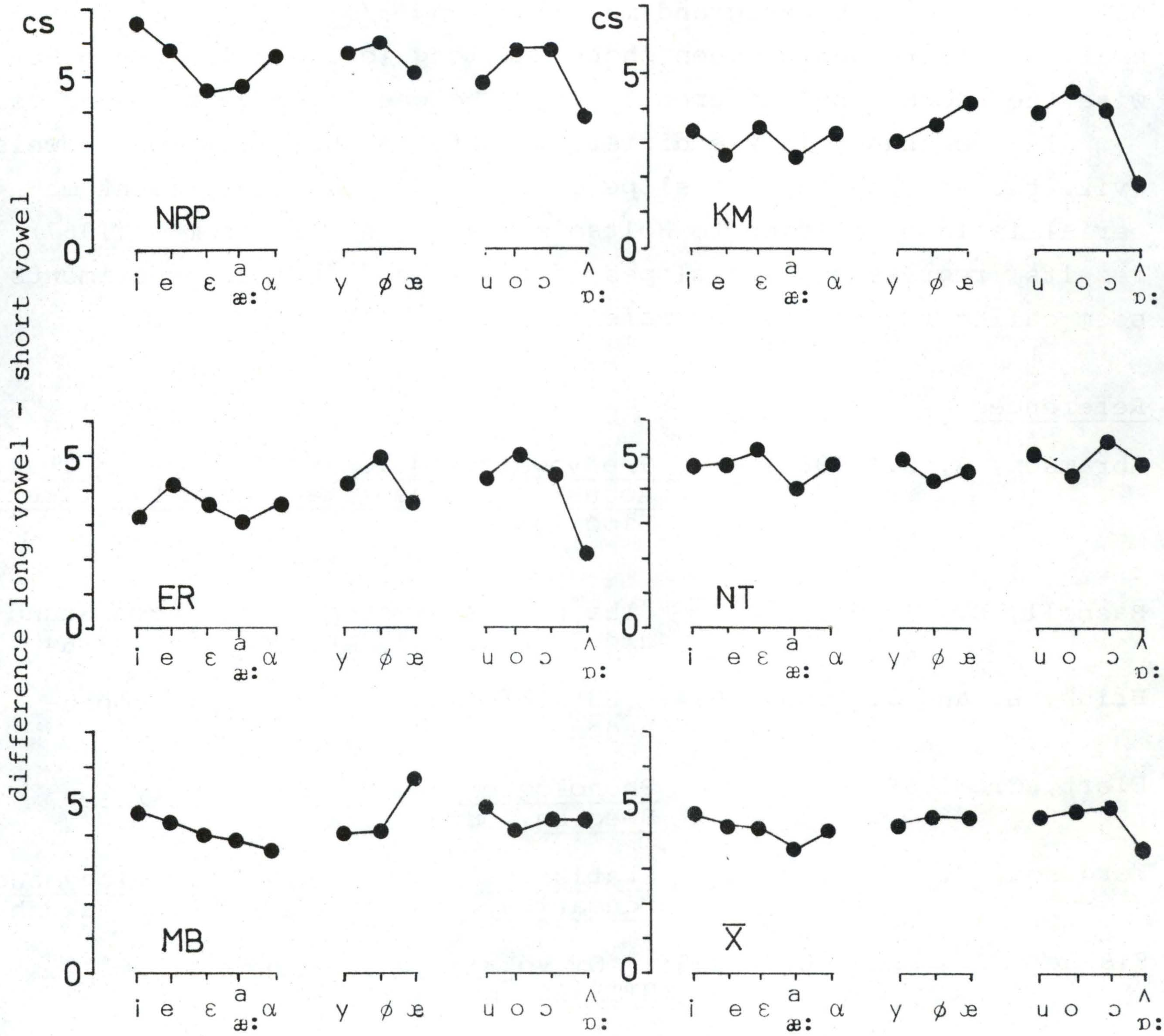

Figure 10

Differences between long and short vowels for the individual subjects, and averaged over all subjects.

picted. The constancy is not very apparent with individual subjects, but on the other hand, the variation seems to be tilfældige random, and it is possible that it would disappear with a larger number of readings. With the grand mean a/æ: and ʌ/ɑ: stick out with smaller differences between short and long (cf. above), but otherwise the short/long difference ranges between 4.08 cs and 4.82 cs.

The constancy in the difference between short and long vowels (viz. the regression line slope of 1.0) found in the present material is in opposition to Holtse's (1977) and Lindblom's (1967) results (regression line slopes of 0.75), and further experiments seem called for before any safe conclusions can be drawn.

References

Abramson, A. S. 1962: The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments, Bloomington, Ind.

Basbøll, H. 1968: "The phoneme system of Advanced Standard Copenhagen", ARIPUC 3, p. 33-54

Brink, L. and J. Lund 1974: Udtaleforskelle i Danmark, Copenhagen

Elert, C. 1964: Phonologic Studies of Quantity in Swedish, Stockholm

Ferguson, G. 1976: Statistical Analysis in Psychology and Education, (fourth edition), Kogakusha

Fischer-Jørgensen, E. 1955: "Om vokallængde i dansk rigsmål", NTTS 15, p. 33-56

Fischer-Jørgensen, E. 1964: "Sound duration and place of articulation", Zs.f.Ph. 17, p. 175-207

Fischer-Jørgensen, E. 1972: "Formant frequencies of long and short Danish vowels", ARIPUC 6, p. 189-213

Holtse, P. 1977: "Variationer i vokallængde på dansk", (unpublished manuscript)

Jespersen, O. 1926: Lehrbuch der Phonetik, Leipzig

Lindblom, B. 1967: "Vowel duration and a model of lip mandible coordination", STL-QPSR 4, p. 1-29

Lindblom, B. 1976: "Durational pattern of Swedish phonology" (mimeographed), Stockholm

Maack, A. 1949: "Die spezifische Lautdauer deutscher Sonanten", _Zs.f.Ph._ 3, p. 190-232

Peterson, G. and I. Lehiste 1960: "Duration of syllable nuclei in English", _JASA_ 32, p. 693-703

Reinholt Petersen, N. 1974: "The influence of tongue height on the perception of vowel duration in Danish", _ARIPUC_ 8, p. 1-10

Reinholt Petersen, N. 1977: "Intrinsic fundamental frequency of Danish vowels", _JPh_ 6, p. 177-190

Thorsen, N. and O. Thorsen 1978: _Fonetik for sprogstuderende_, Copenhagen.