COMPARISONS BETWEEN THE VOWEL FORMANT FREQUENCIES IN SPEECH AND SONG

Børge Frøkjær-Jensen

The following analysis of the formant frequencies of singers is part of a more comprehensive analysis of the Danish vowel formants in speech, shout, and song (7). Different parts of the material from that analysis have been used for investigations on the changes of formant frequencies at high voice effort (6), and for investigations on the relations between male, female, and children's formant frequencies (8).

1. The purpose.

The purpose of the present investigation has been to study why vowels sung by trained singers have a different timbre compared with spoken vowels. This problem should be divided in two parts:

(a) a comparison between the acoustic spectra in speech and song

and

(b) an investigation of the physiological changes in the vocal tract which cause the acoustic changes.

The preliminary examination of the vowels in song discussed in this article deals mainly with the first part of this problem, i.e. the acoustic changes in song compared with normal speech.

2. The material.

The material for the formant measurements consists of tape recordings from seven professional male singers (two basses, three baritones, and two tenors) engaged with the Danish

Radio's choir and church choirs.*) The physiological observations are based on a material consisting of lo male subjects, lo female subjects, and 5 children. Half of the male and half of the female subjects were trained and professional singers. The rest must be classified as untrained singers. During the tape recordings some visual observations concerning the changes of the resonating cavities in song in relation to speech were written down. The observed modifications in song concern: (1) position of the larynx, (2) mouth opening, (3) lip rounding, (4) movements of the hyoid bone, and (5) the dilated pharynx.

All the subjects have contributed with recordings of

(a) normal speech,

(b) song in low-pitched chest register (the fundamental frequency nearly equals the fundamental frequency in speech),

(c) song in high-pitched chest register (the fundamental frequency is about one octave higher than in low-pitched chest register).

All the subjects have spoken and sung eleven Danish bisyllabic words in which all the Danish long vowel phonemes and the lowered r-influenced /a/-variant [ɑ:] occur (6),(7).

The tape recordings have been analysed by means of the Sona-Graph. All the formant frequencies are based upon averages of measurements on narrow band and wide band sonagrams. These data have been converted into punched cards and treated in the electronic data processing center (NEUCC) as discussed

---

*) My material includes hundreds of sonagrams from professional female singers, too. They have, however, been excluded from the formant measurements for this article in order to get a more homogeneous collection of data, and because it is impossible to measure the formant frequencies of the more high-pitched female voices in song. Such sonagrams can only be analysed by a comparison of single harmonics, which is outside the scope of this article.

in (8). Photo copies of the EDP machine outputs are shown in Figs. 1 and 2.

3. The vowel formants in speech and song.

Column AMEAN in Figs. 1 and 2 shows the formant frequencies for the spoken vowels of the seven singers. Column BMEAN in Fig. 1 shows the formant frequencies of the eleven vowels sung in low-pitched chest register. Column BMEAN in Fig. 2 shows the formant frequencies of the eleven vowels sung in high-pitched chest register. Column BAPCT in Fig. 1 and Fig. 2 shows the differences between the sung and the spoken vowel formant frequencies expressed as a percentage of the spoken vowels (the variable "BAPCT" thus equals the parameter K_n in the terminology of acoustic phonetics), and column SDPCT2 indicates ± twice the standard deviation of that difference (which equals a 95.45 per cent significance level).

A more detailed explanation of the different parameters in the tables Fig. 1 and Fig. 2 may be found in reference (8).

4. Significant changes in formant frequencies.

The different formant frequencies for normal spoken vowels and for sung vowels in low- and high-pitched chest register are depicted in Fig. 3. The vowel formants are depicted in the following way: in speech by means of solid lines; in low-pitched chest register song by means of dashed lines; and in high-pitched chest register song by means of chain-dashed lines.

The diagram Fig. 4 shows the K_n-values expressed as percentages (3) for the formants of each vowel sung both at high and low pitch in relation to the same spoken vowel. Twice the standard deviation is indicated for each K_n-value.

Assuming 95 per cent confidence limits we can make the following conclusions:

F1:

Vowels sung at the same pitch as normal speech do not have first formants which are significantly different from those of normal speech, whereas F1 is considerably raised when the pitch goes up. If the pitch is raised one octave in the chest re-

PROGRAM FOR COMPARISONS BETWEEN VOWEL SPECTRA GROUP A AND VOWEL SPECTRA GROUP B WITH STATISTIC CALCULATIONS.
VERSION OF DECEMBER 1967, (BFJ).

INPUT GROUP A AND INPUT GROUP B ARE TWO SETS OF VARIABLE NUMBERS,
MEAN VALUES,
STANDARD ERRORS OF MEANS, AND
NUMBER OF SAMPLES FOR EACH VARIABLE.

NUMBER OF VARIABLES MUST BE THE SAME IN GROUP A AND GROUP B.
THE MAXIMUM NUMBER OF VARIABLES IS 999.
THE MAXIMUM NUMBER OF SAMPLES FOR A VARIABLE IS 999.

I = NUMBER OF A VARIABLE IN GROUP A.
AMEAN = MEANS OF SAMPLES IN GROUP A.
K = NUMBER OF SAMPLES FOR AMEAN.
ASDE = STANDARD ERRORS OF MEANS IN GROUP A.
APCT = ASDE IN PER CENT OF MEANS.
J = NUMBER OF A VARIABLE IN GROUP B.
BMEAN = MEANS OF SAMPLES IN GROUP B.
L = NUMBER OF SAMPLES FOR BMEAN.
BSDE = STANDARD ERRORS OF MEANS IN GROUP B.
BPCT = BSDE IN PER CENT OF MEANS.
DIFBA = THE DIFFERENCE BMEAN - AMEAN.
BAPCT = DIFBA IN PER CENT OF AMEAN.
SDABS = STANDARD DEVIATION FOR THE DIFFERENCE DIFBA.
SDPCT = THE RELATIVE STANDARD DEVIATION FOR DIFBA CALCULATED IN PER CENT.
SDPCT2 = TWO TIMES SDPCT.
SDPCT3 = THREE TIMES SDPCT.
NO = THE TOTAL NUMBER OF VARIABLES IN A GROUP.
NVAR = AN INTEGER VARIABLE WHICH DETERMINES WHEN WARNING MESSAGES MUST APPEAR.
NF = A CODE NUMBER. NF = 0, NO CALCULATIONS ON BAPCT-AVERAGES ARE GENERATED,
NF = 1, CALCULATIONS ON BAPCT-AVERAGES ARE GENERATED FOR FORMANT FREQUENCIES,
NF = 2, CALCULATIONS ON BAPCT-AVERAGES ARE GENERATED FOR FORMANT FREQUENCIES AND FORMANT LEVELS.

ERROR MESSAGES OCCUR IN OUTPUT TABLE WHEN INPUT DATA EQUAL ZERO OR ARE ABSENT.
ERROR MESSAGES OCCUR IN OUTPUT TABLE WHEN COMPARISONS ARE MADE FOR WRONG VARIABLE NUMBERS.
WARNINGS OCCUR WHEN NUMBER OF SAMPLES EQUALS OR IS LESS THAN NVAR.

MALE SINGERS. COMPARISON BETWEEN SPOKEN AND SUNG VOWELS, LOW PITCHED CHEST REG.

VAR NO	AMEAN	K	ASDE	APCT	BMEAN	L	BSDE	BPCT	DIFBA	BAPCT	SDABS	SDPCT	SDPCT2	SDPCT3	
1	236.	7	7.27	3.08	250.	7	11.23	4.49	14.	5.93	13.38	5.45	10.89	16.34	i:
2	2101.	7	47.88	2.28	1962.	7	59.01	3.01	-139.	-6.62	75.99	3.77	7.55	11.32	
3	2974.	7	74.73	2.51	2731.	7	93.17	3.41	-243.	-8.17	119.44	4.24	8.47	12.71	
4	3321.	7	74.55	2.24	3108.	7	89.10	2.87	-213.	-6.41	116.17	3.64	7.28	10.92	
5	281.	7	10.32	3.67	298.	7	4.34	1.46	17.	6.05	11.20	3.95	7.90	11.85	e:
6	2072.	7	40.13	1.94	1926.	7	47.84	2.48	-146.	-7.05	62.44	3.15	6.30	9.45	
7	2744.	7	57.80	2.11	2644.	7	48.82	1.85	-100.	-3.64	75.66	2.80	5.60	8.40	
8	3379.	7	97.84	2.90	3104.	7	90.03	2.90	-275.	-8.14	132.96	4.10	8.20	12.30	
9	365.	7	11.55	3.16	410.	7	21.63	5.28	45.	12.33	24.52	6.15	12.30	18.46	ε:
10	1961.	7	44.45	2.27	1809.	7	42.73	2.36	-152.	-7.75	61.66	3.27	6.55	9.82	
11	2474.	7	79.03	3.19	2350.	7	56.47	2.40	-124.	-5.01	97.13	4.00	7.99	11.99	
12	3341.	7	66.91	2.00	3126.	7	97.81	3.13	-215.	-6.44	118.51	3.71	7.43	11.14	
13	542.	7	29.90	5.52	586.	7	31.93	5.45	44.	8.12	43.74	7.75	15.51	23.26	ɑ:
14	1701.	7	77.12	4.53	1472.	7	70.56	4.79	-229.	-13.46	104.53	6.60	13.20	19.79	
15	2306.	7	75.60	3.28	2372.	7	62.46	2.63	66.	2.86	98.06	4.20	8.41	12.61	
16	3433.	7	95.03	2.77	3162.	7	94.82	3.00	-271.	-7.89	134.24	4.08	8.16	12.24	
17	717.	7	26.61	3.71	671.	7	34.29	5.11	-46.	-6.42	43.40	6.32	12.63	18.95	a:
18	1157.	7	17.11	1.48	1084.	7	40.07	3.70	-73.	-6.31	43.57	3.98	7.96	11.94	
19	2531.	7	98.62	3.90	2586.	7	76.43	2.96	55.	2.17	124.77	4.89	9.78	14.67	
20	3551.	7	184.07	5.18	3154.	7	118.30	3.75	-397.	-11.18	218.81	6.40	12.80	19.19	
21	237.	7	8.08	3.41	260.	7	12.49	4.80	23.	9.70	14.88	5.89	11.78	17.67	y:
22	1863.	7	34.55	1.85	1805.	7	41.69	2.31	-58.	-3.11	54.15	2.96	5.92	8.89	
23	2118.	7	43.67	2.06	2079.	7	44.24	2.13	-39.	-1.84	62.16	2.96	5.93	8.89	
24	3195.	7	81.64	2.56	3070.	7	104.19	3.39	-125.	-3.91	132.37	4.25	8.50	12.74	
25	301.	7	8.41	2.79	359.	7	29.11	8.11	58.	19.27	30.30	8.58	17.15	25.73	ø:
26	1637.	7	40.29	2.46	1620.	7	34.31	2.12	-17.	-1.04	52.92	3.25	6.49	9.74	
27	2051.	7	47.81	2.33	2020.	7	40.74	2.02	-31.	-1.51	62.81	3.08	6.16	9.25	
28	3161.	7	78.55	2.48	3026.	7	106.18	3.51	-135.	-4.27	132.08	4.30	8.60	12.90	
29	385.	7	11.70	3.04	416.	7	16.10	3.87	31.	8.05	19.90	4.92	9.84	14.76	æ:
30	1619.	7	67.14	4.15	1504.	7	53.01	3.52	-115.	-7.10	85.54	5.44	10.88	16.33	
31	2086.	7	43.54	2.09	2079.	7	29.53	1.42	-7.	-0.34	52.61	2.52	5.05	7.57	
32	3170.	7	71.14	2.24	3062.	7	106.01	3.46	-108.	-3.41	127.67	4.13	8.25	12.38	
33	254.	7	7.43	2.93	279.	7	8.78	3.15	25.	9.84	11.50	4.30	8.59	12.89	u:
34	759.	6	24.37	3.21	774.	4	42.30	5.47	15.	1.98	48.82	6.34	12.68	19.02	
35	2105.	5	42.37	2.01	2082.	5	78.73	3.78	-23.	-1.09	89.41	4.28	8.57	12.85	
WARNING ONLY, - NUMBER OF SAMPLES (K) FOR AMEAN AND NUMBER OF SAMPLES (L) FOR BMEAN ARE LESS THAN OR EQUAL														5	
36	3248.	4	149.19	4.59	2865.	5	82.78	2.89	-383.	-11.79	170.62	5.43	10.85	16.28	
WARNING ONLY, - NUMBER OF SAMPLES (K) FOR AMEAN AND NUMBER OF SAMPLES (L) FOR BMEAN ARE LESS THAN OR EQUAL														5	
37	346.	7	10.99	3.18	360.	7	13.09	3.64	14.	4.05	17.09	4.83	9.66	14.48	o:
38	762.	7	13.04	1.71	769.	7	12.09	1.57	7.	0.92	17.78	2.32	4.65	6.97	
39	2316.	4	80.97	3.50	2318.	6	91.71	3.96	2.	0.09	122.34	5.28	10.56	15.84	
40	3116.	6	102.00	3.27	3040.	7	113.88	3.75	-76.	-2.44	152.88	4.97	9.95	14.92	
41	430.	7	9.76	2.27	436.	7	10.62	2.44	6.	1.40	14.42	3.33	6.66	9.99	ɔ:
42	862.	7	25.82	3.00	859.	7	23.97	2.79	-3.	-0.35	35.23	4.09	8.19	12.28	
43	2292.	6	81.66	3.56	2287.	7	81.84	3.58	-5.	-0.22	115.61	5.05	10.10	15.15	
44	3163.	6	120.12	3.80	3038.	7	93.52	3.08	-125.	-3.95	152.23	4.89	9.78	14.67	

THE AVERAGE FREQUENCY OF FORMANT 1 IN GROUP B IN RELATION TO GROUP A IS 7.12 PER CENT,

THE AVERAGE FREQUENCY OF FORMANT 2 IN GROUP B IN RELATION TO GROUP A IS -4.54 PER CENT,

THE AVERAGE FREQUENCY OF FORMANT 3 IN GROUP B IN RELATION TO GROUP A IS -1.52 PER CENT,

THE AVERAGE FREQUENCY OF FORMANT 4 IN GROUP B IN RELATION TO GROUP A IS -6.35 PER CENT.

THE TOTAL AVERAGE PERCENTAGE OF FORMANT FREQUENCIES IN GROUP B IN RELATION TO GROUP A IS -1.32 PER CENT.

CALCULATIONS HAVE BEEN COMPLETED.

FIG. 1

STOP

PROGRAM FOR COMPARISONS BETWEEN VOWEL SPECTRA GROUP A AND VOWEL SPECTRA GROUP B WITH STATISTIC CALCULATIONS.
VERSION OF DECEMBER 1967, (BFJ).

INPUT GROUP A AND INPUT GROUP B ARE TWO SETS OF VARIABLE NUMBERS,
MEAN VALUES,
STANDARD ERRORS OF MEANS, AND
NUMBER OF SAMPLES FOR EACH VARIABLE.

NUMBER OF VARIABLES MUST BE THE SAME IN GROUP A AND GROUP B.
THE MAXIMUM NUMBER OF VARIABLES IS 999.
THE MAXIMUM NUMBER OF SAMPLES FOR A VARIABLE IS 999.

I = NUMBER OF A VARIABLE IN GROUP A.
AMEAN = MEANS OF SAMPLES IN GROUP A.
K = NUMBER OF SAMPLES FOR AMEAN.
ASDE = STANDARD ERRORS OF MEANS IN GROUP A.
APCT = ASDE IN PER CENT OF MEANS.
J = NUMBER OF A VARIABLE IN GROUP B.
BMEAN = MEANS OF SAMPLES IN GROUP B.
L = NUMBER OF SAMPLES FOR BMEAN.
BSDE = STANDARD ERRORS OF MEANS IN GROUP B.
BPCT = BSDE IN PER CENT OF MEANS.
DIFBA = THE DIFFERENCE BMEAN - AMEAN.
BAPCT = DIFBA IN PER CENT OF AMEAN.
SDABS = STANDARD DEVIATION FOR THE DIFFERENCE DIFBA.
SDPCT = THE RELATIVE STANDARD DEVIATION FOR DIFBA CALCULATED IN PER CENT.
SDPCT2 = TWO TIMES SDPCT.
SDPCT3 = THREE TIMES SDPCT.
NO = THE TOTAL NUMBER OF VARIABLES IN A GROUP.
NVAR = AN INTEGER VARIABLE WHICH DETERMINES WHEN WARNING MESSAGES MUST APPEAR.
NF = A CODE NUMBER. NF = 0, NO CALCULATIONS ON BAPCT-AVERAGES ARE GENERATED,
NF = 1, CALCULATIONS ON BAPCT-AVERAGES ARE GENERATED FOR FORMANT FREQUENCIES,
NF = 2, CALCULATIONS ON BAPCT-AVERAGES ARE GENERATED FOR FORMANT FREQUENCIES AND FORMANT LEVELS.

ERROR MESSAGES OCCUR IN OUTPUT TABLE WHEN INPUT DATA EQUAL ZERO OR ARE ABSENT.
ERROR MESSAGES OCCUR IN OUTPUT TABLE WHEN COMPARISONS ARE MADE FOR WRONG VARIABLE NUMBERS.
WARNINGS OCCUR WHEN NUMBER OF SAMPLES EQUALS OR IS LESS THAN NVAR.

MALE SINGERS. COMPARISON BETWEEN SPOKEN AND SUNG VOWELS,HIGH PITCHED CHEST REG.

VAR NO	AMEAN	K	ASDE	APCT	BMEAN	L	BSDE	BPCT	DIFBA	BAPCT	SDABS	SDPCT	SDPCT2	SDPCT3	
1	236.	7	7.27	3.03	304.	7	12.12	3.99	68.	28.81	14.13	5.04	10.08	15.11	i
2	2101.	7	47.88	2.28	1915.	6	81.08	4.23	-186.	-8.85	94.16	4.81	9.62	14.42	
3	2974.	7	74.73	2.51	2643.	6	119.36	4.52	-331.	-11.13	140.82	5.17	10.34	15.50	
4	3321.	7	74.55	2.24	3187.	7	95.14	2.99	-134.	-4.03	120.87	3.74	7.47	11.21	
5	281.	7	10.32	3.67	386.	7	25.13	6.51	105.	37.37	27.17	7.47	14.95	22.42	e:
6	2072.	7	40.13	1.94	1811.	7	69.04	3.81	-261.	-12.60	79.86	4.28	8.55	12.83	
7	2744.	7	57.90	2.11	2415.	7	61.43	2.54	-329.	-11.99	84.35	3.30	6.61	9.91	
8	3379.	7	97.84	2.90	3195.	7	95.41	2.99	-184.	-5.45	136.66	4.16	8.32	12.48	
9	365.	7	11.55	3.16	543.	7	22.93	4.20	178.	48.77	25.59	5.26	10.52	15.79	ε:
10	1961.	7	44.45	2.27	1656.	7	55.50	3.35	-305.	-15.55	71.11	4.05	8.09	12.14	
11	2474.	7	79.03	3.19	2283.	7	58.76	2.57	-191.	-7.72	98.48	4.10	8.20	12.31	
12	3341.	7	66.91	2.00	3216.	7	77.16	2.40	-125.	-3.74	102.13	3.13	6.25	9.38	
13	542.	7	29.90	5.52	622.	7	23.58	3.79	80.	14.76	38.08	6.69	13.39	20.08	a:
14	1701.	7	77.12	4.53	1292.	7	45.91	3.55	-409.	-24.04	89.75	5.76	11.52	17.28	
15	2306.	7	75.60	3.28	2410.	7	75.74	3.14	104.	4.51	107.01	4.54	9.08	13.62	
16	3433.	7	95.03	2.77	3250.	7	69.51	2.14	-183.	-5.33	117.74	3.50	7.00	10.49	
17	717.	7	26.61	3.71	664.	7	40.63	6.12	-53.	-7.39	48.57	7.16	14.31	21.47	ɑ:
18	1157.	7	17.11	1.48	1062.	7	49.78	4.69	-95.	-8.21	52.64	4.92	9.83	14.75	
19	2531.	7	98.62	3.90	2654.	7	74.49	2.81	123.	4.86	123.59	4.80	9.60	14.41	
20	3551.	7	184.07	5.18	3333.	7	92.37	2.77	-218.	-6.14	205.95	5.88	11.76	17.63	
21	237.	7	8.03	3.41	321.	7	12.07	3.76	84.	35.44	14.52	5.08	10.15	15.23	y:
22	1863.	7	34.55	1.85	1711.	7	38.07	2.23	-152.	-8.16	51.41	2.90	5.79	8.69	
23	2118.	7	43.67	2.06	2146.	7	46.58	2.17	28.	1.32	63.85	2.99	5.99	8.98	
24	3195.	7	81.64	2.56	3090.	7	57.61	1.86	-105.	-3.29	99.92	3.16	6.33	9.49	
25	301.	7	8.41	2.79	409.	7	24.95	6.10	108.	35.88	26.33	6.71	13.42	20.13	ø:
26	1637.	7	40.29	2.46	1551.	7	56.34	3.63	-86.	-5.25	69.26	4.39	8.78	13.16	
27	2051.	7	47.81	2.33	2061.	7	42.66	2.07	10.	0.49	64.08	3.12	6.23	9.35	
28	3161.	7	78.55	2.48	3096.	7	60.64	1.96	-65.	-2.06	99.23	3.16	6.33	9.49	
29	385.	7	11.70	3.04	541.	7	22.22	4.11	156.	40.52	25.11	5.11	10.22	15.33	œ:
30	1619.	7	67.14	4.15	1393.	7	37.51	2.69	-226.	-13.96	76.91	4.94	9.89	14.83	
31	2086.	7	43.54	2.09	2192.	7	57.03	2.60	106.	5.08	71.75	3.34	6.67	10.01	
32	3170.	7	71.14	2.24	3139.	7	67.40	2.15	-31.	-0.98	98.00	3.11	6.21	9.32	
33	254.	7	7.43	2.93	368.	7	27.47	7.46	114.	44.88	28.46	8.02	16.03	24.05	u
34	759.	6	24.37	3.21	849.	7	23.92	2.82	90.	11.86	34.15	4.27	8.54	12.82	
35	2105.	5	42.37	2.01	2139.	7	105.98	4.95	34.	1.62	114.14	5.35	10.70	16.04	
	WARNING ONLY, - NUMBER OF SAMPLES (K) FOR AMEAN IS LESS THAN OR EQUALS 5														
36	3248.	4	149.19	4.59	3065.	7	63.11	2.06	-183.	-5.63	161.99	5.03	10.07	15.10	
	WARNING ONLY, - NUMBER OF SAMPLES (K) FOR AMEAN IS LESS THAN OR EQUALS 5														
37	346.	7	10.99	3.18	455.	7	21.63	4.75	109.	31.50	24.26	5.72	11.43	17.15	o:
38	762.	7	13.04	1.71	812.	6	18.74	2.31	50.	6.56	22.83	2.87	5.75	8.62	
39	2316.	4	81.97	3.50	2623.	6	74.56	2.84	307.	13.26	110.07	4.51	9.01	13.52	
40	3116.	6	102.00	3.27	3111.	7	51.63	1.66	-5.	-0.16	114.32	3.67	7.34	11.01	
41	430.	7	9.76	2.27	529.	7	12.43	2.35	99.	23.02	15.80	3.27	6.53	9.80	ɔ:
42	862.	7	25.82	3.00	984.	7	48.08	4.89	122.	14.15	54.57	5.73	11.46	17.19	
43	2292.	6	81.66	3.56	2583.	6	64.97	2.52	291.	12.70	104.35	4.36	8.72	13.08	
44	3163.	5	120.12	3.80	3181.	7	49.93	1.57	18.	0.57	130.08	4.11	8.22	12.33	

THE AVERAGE FREQUENCY OF FORMANT 1 IN GROUP B IN RELATION TO GROUP A IS 30.32 PER CENT,

THE AVERAGE FREQUENCY OF FORMANT 2 IN GROUP B IN RELATION TO GROUP A IS -5.82 PER CENT,

THE AVERAGE FREQUENCY OF FORMANT 3 IN GROUP B IN RELATION TO GROUP A IS 1.18 PER CENT,

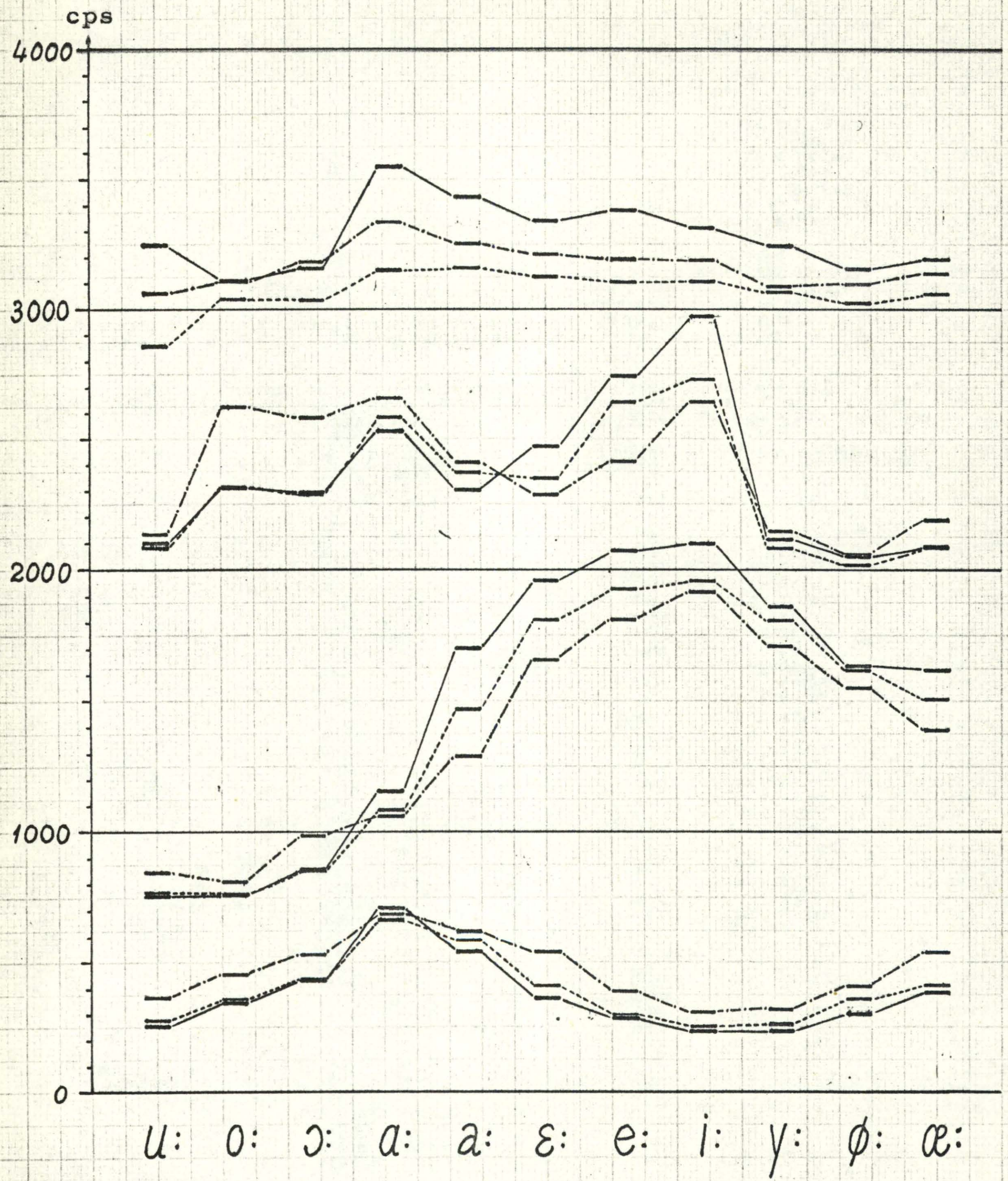
THE AVERAGE FREQUENCY OF FORMANT 4 IN GROUP B IN RELATION TO GROUP A IS -3.29 PER CENT.

THE TOTAL AVERAGE PERCENTAGE OF FORMANT FREQUENCIES IN GROUP B IN RELATION TO GROUP A IS 5.60 PER CENT.

CALCULATIONS HAVE BEEN COMPLETED.

FIG. 2

STOP



Comparison between spoken and sung vowels.

Normal speech (chest register)

Song in low-pitched chest register

Song in high-pitched chest register

Fig. 3

gister, F1 becomes significantly higher except for the open vowels [a:] and [ɑ:] (the average increment for F1 in song is: low-pitched chest register 7 per cent, and high-pitched chest register 3o per cent).*)

F2:

All the front vowels show a lowered F2 in song. The lowering is more pronounced the higher the pitch is, but the difference from speech is only big enough to be considered significant for singing in the high-pitched chest register. The rounded back vowels exhibit nearly the same formant frequencies as in speech when sung in low-pitched chest register. When sung in high-pitched chest register they have a significantly higher F2 than in speech.

F3:

The changes of F3 in song are very small, and the few changes are most pronounced at high pitch. Thus the unrounded front vowels [i:], [e:] and [ɛ:] are significantly lowered, and the rounded back vowels [o:] and [ɔ:] are significantly raised in relation to speech.

F4:

F4 is lowered during song in the whole pitch range of the chest register.

It may be seen from the above-mentioned formant changes that the vowels in high-pitched song are more centralized than in speech. In this respect the formant changes equal what happens when the voice effect is increased (6). However, it is

---

*) This article only takes into account song at nearly the same pitch as normal speech and song at about one octave higher. An examination of song at lower pitch than normal speech (9 subjects have been involved in this analysis) has given the result that all the formants are lowered in relation to speech. However, the individual deviations are too pronounced, and it is not possible on the basis of my material to speak about a significantly relevant formant lowering in very low-pitched song.

clearly seen from the sonagrams that the main difference between shout and song is a changed energy distribution in the spectrum. F1 is very weak in shout and often very prominent in song. The upper part of the spectrum is strongly intensified in shout, and often fully cut off in song.

5. The physiological changes.

As previously mentioned this article will not deal with the physiological changes of the vocal tract because I have not had the possibility of making X-ray photos yet. A more detailed investigation of the articulatory/acoustic problems in song must naturally involve calculations on the resonatory cavities based on X-ray photos.*) Therefore only a few primitive visual observations carried out during the tape recordings of the subjects will be mentioned here with suggestions of possible explanations.

With 23 subjects singing at the same pitch as that of normal speech I have observed:

Lowered larynx	1o subjects	
Raised larynx	1 subject	
Increased mouth opening	9 subjects	
Decreased mouth opening	2 subjects	
Increased lip rounding	9 subjects	
Decreased lip rounding	0 subjects	
Lowered and advanced hyoid bone .	9 subjects	) mainly) with) trained) singers
Dilated pharynx	1o subjects	

*) Sundberg (10) has shown by means of X-ray photos of 4 male singers that the vocal tract has about the same length in rest and in speech, whereas the length of the vocal tract is increased in song. The increase is caused by a 1 cm lowering of the larynx. Furthermore he has shown that the jaw is lowered and the distance between the Sinus Morgagni is increased by 5o per cent in song (Fo $\approx$ 12o cps).

With 2o subjects singing in high-pitched chest register I have observed:

Lowered larynx	7 subjects	
Raised larynx	4 subjects	
Increased mouth opening	8 subjects	
Decreased mouth opening	(1) subject	
Increased lip rounding	8 subjects	
Decreased lip rounding	0 subjects	
Lowered and advanced hyoid bone .	1o subjects	trained singers only
Dilated pharynx	1o subjects	trained singers only

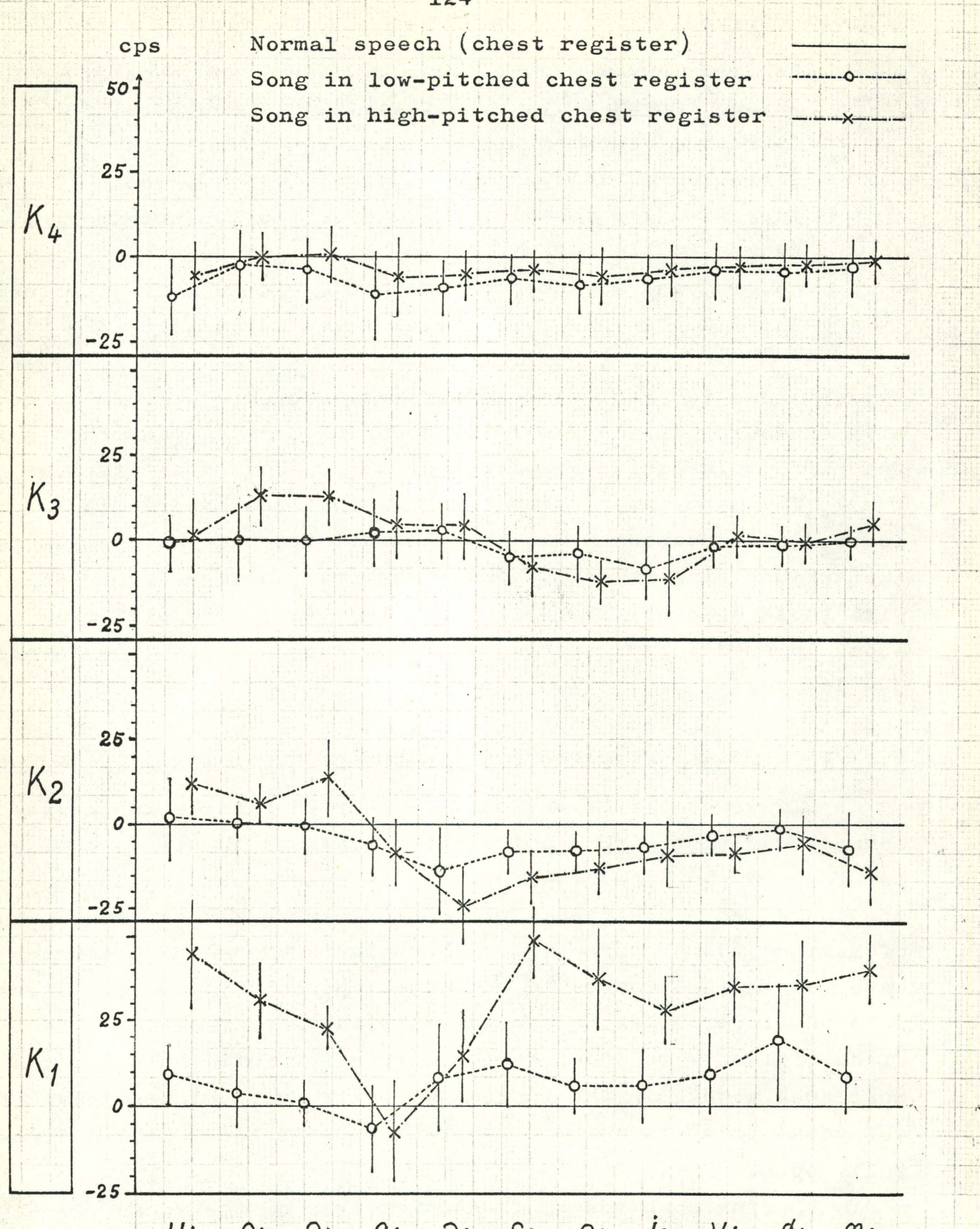
These two tables show that many subjects perform an active expanding of the vocal tract during singing, and that very few perform an active reduction of the resonance cavities.

The most interesting thing is that the observation of lowered and advanced hyoid bone and dilated pharynx applies to the trained singers only. These observations agree very well with those of other investigators (10), (11), who all speak about lowered larynx, increased resonance cavities, and lowered jaw in song.

6. Formant variations explained by the physiological changes.

If we compare the visually observed changes in the resonance cavities with the formant variations in Fig. 4 we may put forward the hypothesis that F1 is raised in song because of the increased mouth opening which influences F1 most (the point of maximum volume current for the main resonance of a quarter wave length is situated in the mouth opening) (2), (4). Therefore the lowered larynx and the dilated pharynx do not affect F1 sufficiently in order to lower this formant. Only for the most open vowel [ɑ:] the increased cavity in the lowest part of the vocal tract is able to lower F1 because the mouth is already fully open.

During song the cross-section area of the vocal tract as a whole is increased. This must also be valid for the point of maximum constriction. According to Fant (4) a widening at the place of maximum constriction causes a lowering of F2 and F3 in



Comparison between the formant frequencies for sung vowels in low-pitched and high-pitched chest register in relation to speech.

The vertical lines indicate twice the standard error of means.

Fig. 4

the unrounded front vowels, and a raising of F2 in the rounded back vowels.

F4 is lowered in all vowels in song because of the lowering of the larynx. This lowering of the larynx (which probably is accompanied by a widening of Sinus Morgagni) actually causes a lowering of all the formants, but the effect is most pronounced at higher frequencies where the point of maximum pressure in the closed part of the resonator is most influential upon the resonant frequencies.

7. Trained versus untrained singers.

A comparison between the formants of lo trained and lo untrained singers (material not presented here) indicates that F3 in the rounded vowels is higher in song than in speech with trained singers, whereas the opposite is true of untrained singers. According to the articulatory tables we must conclude that the lowered and advanced hyoid bone, the active dilated pharynx and the lowered larynx must in some way be responsible for the changes in F3.

Again, according to Fant's calculations on the horn-shaped three parameter model of the vocal tract, a widening of the lowest 3 cm in the vocal tract (i.e. mainly a widening of the Sinus Morgagni cavity) causes a raising of F3. We may expect that trained singers are able (to a higher degree than untrained singers) to perform such an active widening of the lowest part of the vocal tract. The visible manifestation of such a change would be a lowering of the musculus geniohyoideus, an advanced and lowered hyoid bone, a dilated pharynx, and a lowered larynx. Most of these changes can be observed by visual inspection of the trained singers.

The acoustic observations agree with that of Sundberg (10) who mentions that "F3 and F4 are closer in singing than in speaking and appear to be closer in trained voices than in untrained speaking voices", and with that of McGinnis, Elnik and Kraichman (9), who speak about the more prominent F3 and F4 of trained singers (the intensity level is increased when the distance between the formants is reduced).

As previously mentioned these conclusions have to be verified by means of X-ray photos.

8. Summary.

The changes in song in relation to speech tend to develop a certain voice quality and mood which is determined by the context and by the individual. Also, the trained singer will try to obtain the best unity of sound in accordance with our (i.e. the Western World's) way of conceiving song art. These artistic shapings of the sound influence the vocalisation. The acoustic structure of the vowels is, therefore, changed: there is less variation in formant levels among the different vowels in song compared to speech, and the first formant is often intensified. The analyses show an energy band in the acoustic spectrum between 25oo cps and 35oo cps common for all the vowels.

As for the formant frequencies F1 is drastically raised as a consequence of the increased mouth opening in all vowels except [ɑ:]. The higher formants are lowered because of the lowered larynx and the active dilated vocal tract, except that F2 is raised in the rounded back vowels because of a widening at the point of maximum constriction. The raising of F3 with trained singers may be due to a dilation of the lower part of the pharynx.

These changes are most pronounced with bass singers singing in high-pitched chest register, and least prominent with tenors singing in low-pitched chest register.

Acknowledgement:

We want to thank the Northern Europe University Computing Center (NEUCC) for run of programs.

References:

(1) W. T. Bartholomew, "A Physical Definition of Good Voice Quality in Male Voices", Journal of the Acoustical Society of America, vol. 6 (1934), pp. 25-33.

(2) T. Chiba and M. Kajiyama, The Vowel (Tokyo, 194o).

(3) Gunnar Fant, "Acoustic Analysis and Synthesis of Speech with Applications to Swedish", Series Ericsson Technics 1 (1959), p. 73.

(4) Gunnar Fant, Acoustic Theory of Speech Production (196o).

(5) D. B. Fry, Lucie Manén, "Basis for the Acoustic Study of Singing", Journal of the Acoustical Society of America, vol. 29 (1957), pp. 69o-692.

(6) B. Frøkjær-Jensen, "Changes in Formant Frequencies and Formant Levels at High Voice Effort", ARIPUC No. 1 (1966), pp. 47-55.

(7) B. Frøkjær-Jensen, En undersøgelse af de danske vokaler i tale, råb og sang. (Institute of Musicology, University of Copenhagen (196o)).

(8) B. Frøkjær-Jensen, "Statistic Calculations of Formant Data", ARIPUC No. 2 (1967), pp. 158-17o.

(9) C. S. McGinnis, M. Elnick, M. Kraichman, "A Study of the Vowel Formants of Well-Known Male Operatic Singers", Journal of the Acoustical Society of America, vol. 23 (1951), pp. 44o-446.

(1o) J. Sundberg, "Formant Frequencies of Bass Singers", Quarterly Progress and Status Report No. 1 (1968), pp. 1-7, Speech Transmission Laboratory, Stockholm, Sweden.

(11) M. Flach, "Ueber die unterschiedliche Grösse der Morgagnischen Ventrikel bei Sängern", Folia Phoniatrica, vol. 16 (1964), pp. 67-74.