CHANGES IN FORMANT FREQUENCIES AND FORMANT LEVELS AT HIGH VOICE EFFORT.

Børge Frøkjær-Jensen

In combination with the study mentioned page 34 ff. in this report an acoustic study has been undertaken to illustrate how an increase of voice effort affects the spectral composition of vowels.

The analysis comprised 2o subjects, who all were able to shout (not everybody can increase the voice effort to a "real shout"). All the subjects were recorded on tape when reading a word list containing the Danish long vowels in stressed position in two-syllable words of the type /hV-Cə/. Some of the subjects were recorded when reading the words with a higher degree of voice effort, and they were all finally recorded while shouting the words from the list.

The absolute sound level pressure was not observed during the recordings because of lack of equipment for that purpose, and the investigation should therefore be considered as preliminary. The speakers started the recordings with what may be called normal speech level. Then the voice effort was gradually increased to shout by some of the subjects. The remaining subjects only shifted between talking and shouting. The shout level varies for different speakers. Some persons can shout with an absolute sound pressure level of 7o-8o dB, whereas others are only able to shout with a level of 5o-6o dB, or not able to shout at all.

In a future investigation I should like to fix various speaking and shouting sound levels, e.g. 3o, 4o, 5o, 6o, and 7o dB re 0.0002 dyn/cm^2, and only subjects able to phonate and articulate all the vowels at these levels should be included in the investigation.

The entire material consists of 319 spoken vowels and 22o shouted vowels. During the recordings I took notes concerning the position of the larynx and mouth opening. All recordings were analysed on the Sona-Graph with the purpose of finding relations between changes in formant frequencies and changes in formant levels. The high-shaping filter of the Sona-Graph was used, and all the vowels were analysed with both narrow band and wide band filter, and several of the measurements were controlled by means of narrow band section sonagrams.

A. Directly observable changes in the articulation:

When the subjects increased their voice effort the following articulatory changes were observed:

1. The larynx may be either raised or lowered during a shout. Most people raise their larynx when shouting, but singers appear to lower their larynx when the intensity level is increased.
2. Higher intensity level is accompanied by a stronger contraction of the muscles of the larynx and pharynx. The muscles of the floor of the mouth (musculus mylohyoideus) is more contracted, too, i.e. the mouth is more open.

   This causes a better acoustic impedance matching between the mouth cavity and the air outside the mouth: more sound energy is transmitted to the surrounding air, and the damping is reduced. As for the "Radiation Load at the Mouth" see reference no. (1).
3. This impedance matching (A2) is probably improved by pulling the tongue backwards.

B. Changes in formant frequencies:

1. The above-mentioned changes in articulation influence the formant frequencies, and for all male voices, female voices, and children's voices separately, I have calculated the mean values of the formant frequencies both in normal speech and in shout.

   In the diagram shown below I have indicated whether the formant frequencies in shout are higher (+) or lower (-) than those for normal speech. The three symbols in each column indicate the direction of change for male voices, female voices, and children's voices, in that order.

	i:	e:	ɛ:	a:	ɑ:	y:	ø:	œ:	u:	o:	ɔ:
F_1	+++	+++	+++	+++	−++	++−	+++	+++	+++	+++	+++
F_2	−−−	−−−	−−−	−−−	+++	−−−	−−−	−−+	+++	+++	+++
F_3	−−−	−−−	−−−	++−	+++	+++	+++	+++	?	?	++−
F_4	−−−	−−−	−−−	−+−	−+−	−+−	++−	−+−	−+−	−−?	++−

The diagram shows that:

F1 is raised in all vowels in shout as compared to speech.

F2 is raised in the back vowels and lowered in the front vowels (this lowering is more pronounced in the unrounded than in the rounded vowels).

F3 is raised in the rounded front vowels and lowered in the unrounded front vowels.

The material is too restricted for anything conclusive to be said about the back vowels, but possibly they follow the front vowels in the separation of rounded/unrounded.

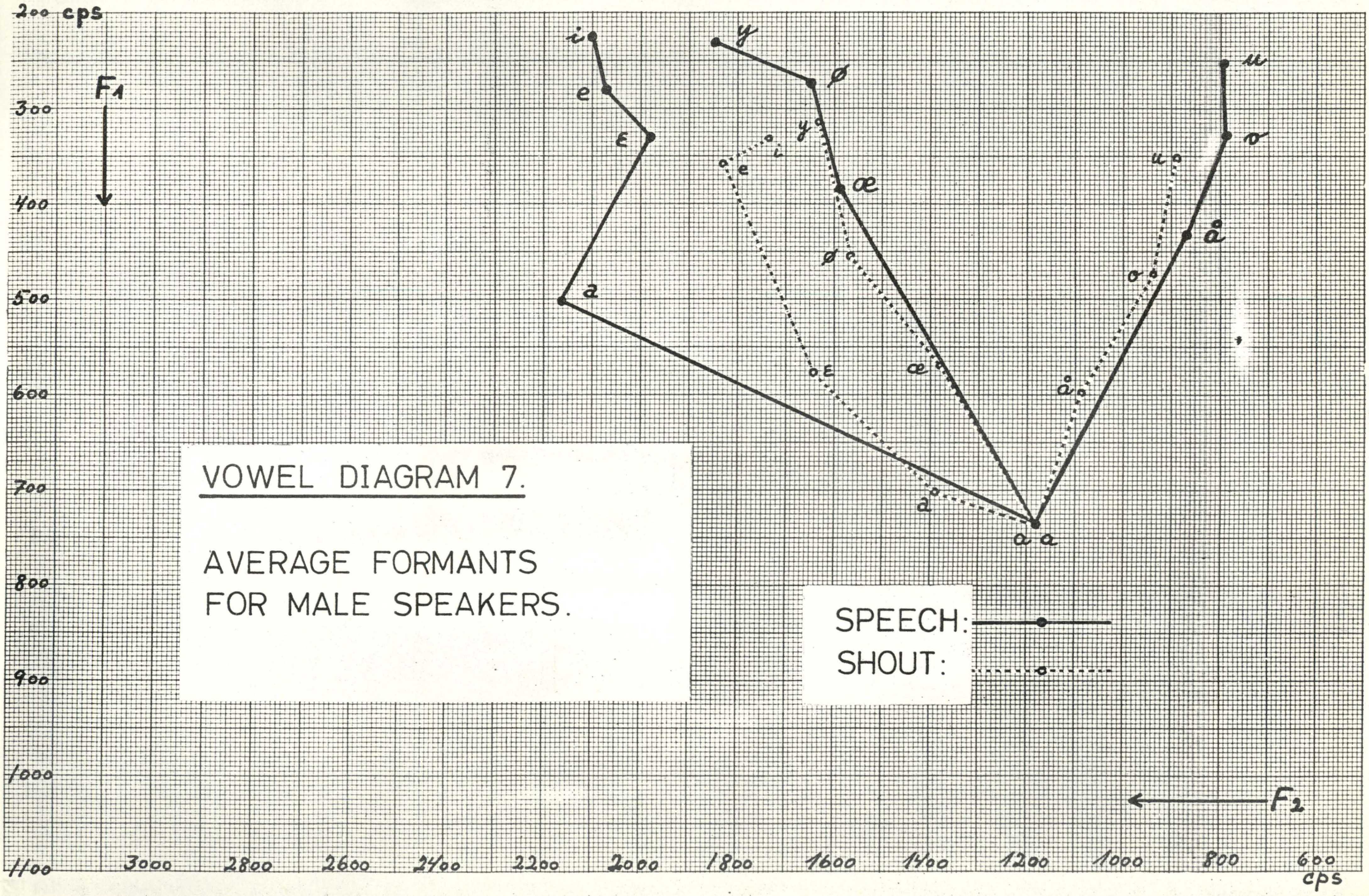
F4 follows F3 to a certain degree, but is probably individually conditioned by the subject.

2. Some of these variations are shown in vowel diagrams "7" and "8" (which are taken out of a larger set of diagrams). Diagram no. 7 shows the male vowel patterns in normal speech and in shout. Only the calculated average frequencies have been plotted into the diagram. The averages are normal arithmetical mean values measured in c/s (it would perhaps have been more significant to calculate the average of the formant frequencies represented in terms of the mel scale).

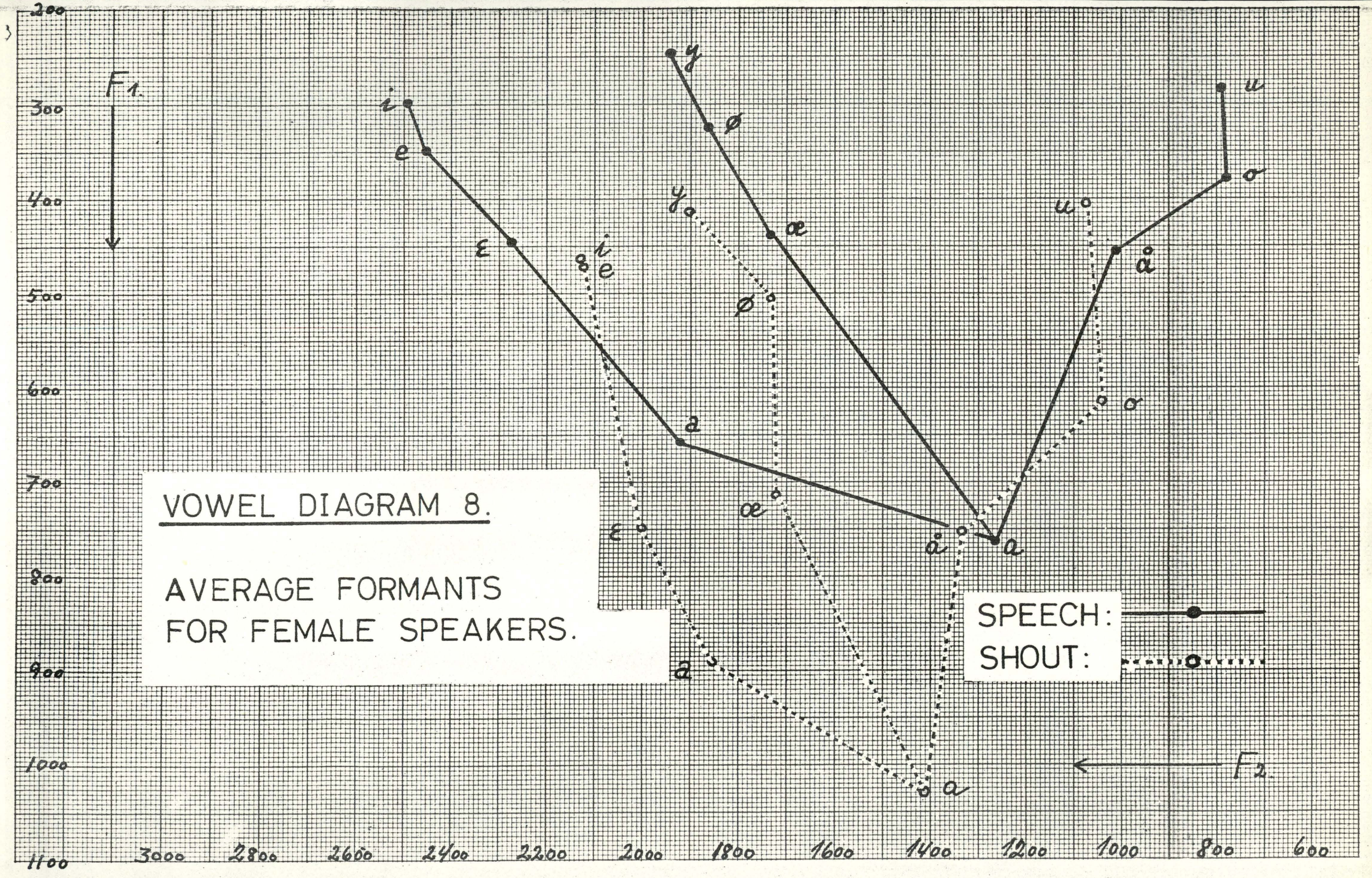
The main tendency to be observed is that the vowels are moved in the direction towards a more open and more central articulation at higher voice effort. Vowel diagram no. 8 for female speakers shows this tendency more clearly. In spite of these changes in phoneme areas when the voice effort is increased, all the shouted vowels utilized in this investigation are clearly identifiable. (The badly identifiable shouted vowels were omitted from the investigation during the tape recordings.)

3. Besides the changes in the formant pattern we can observe a change in the fundamental frequency. The fundamental frequency goes up when the intensity level is increased, sometimes the fundamental frequency is one octave higher in shout than in normal speech. This connexion between intensity level and fundamental frequency has been observed before (2).

4. The formant structure becomes less regular because of many small uncontrolled contractions of the muscles in the speech organs.



VOWEL DIAGRAM 7.
AVERAGE FORMANTS
FOR MALE SPEAKERS.
SPEECH:
SHOUT:
F1
F2
200 cps
300
400
500
600
700
800
900
1000
1100
3000
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
cps
i
e
ε
a
y
ø
œ
u
o
å
ɑ


VOWEL DIAGRAM 8.
AVERAGE FORMANTS
FOR FEMALE SPEAKERS.
SPEECH:
SHOUT:
F1.
F2.
200
300
400
500
600
700
800
900
1000
1100
3000
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
i
e
ε
a
y
ø
œ
u
o
å
a

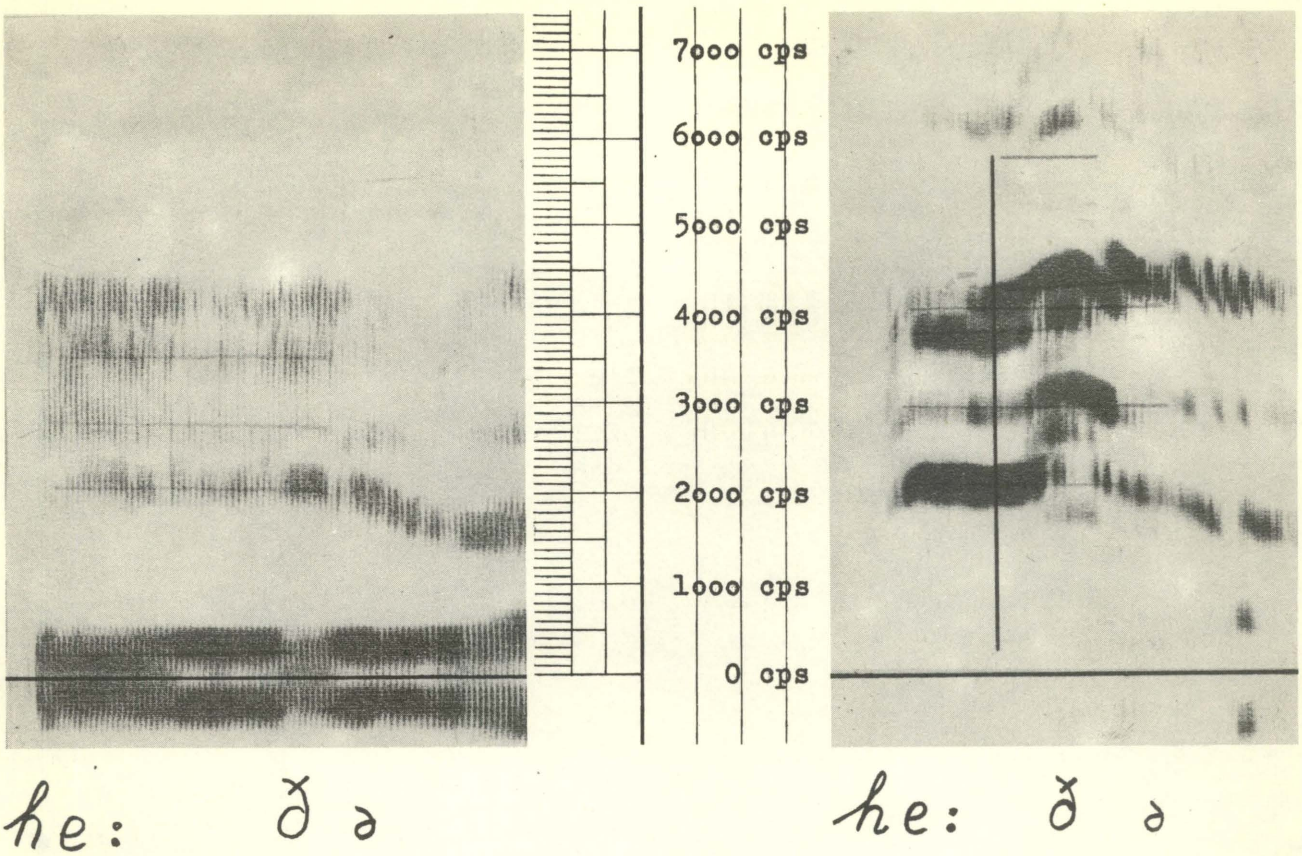
C. Changes in formant levels:

On the basis of the present material it is possible to measure the relative changes in the spectral distribution of energy when the subjects change their voice effort.

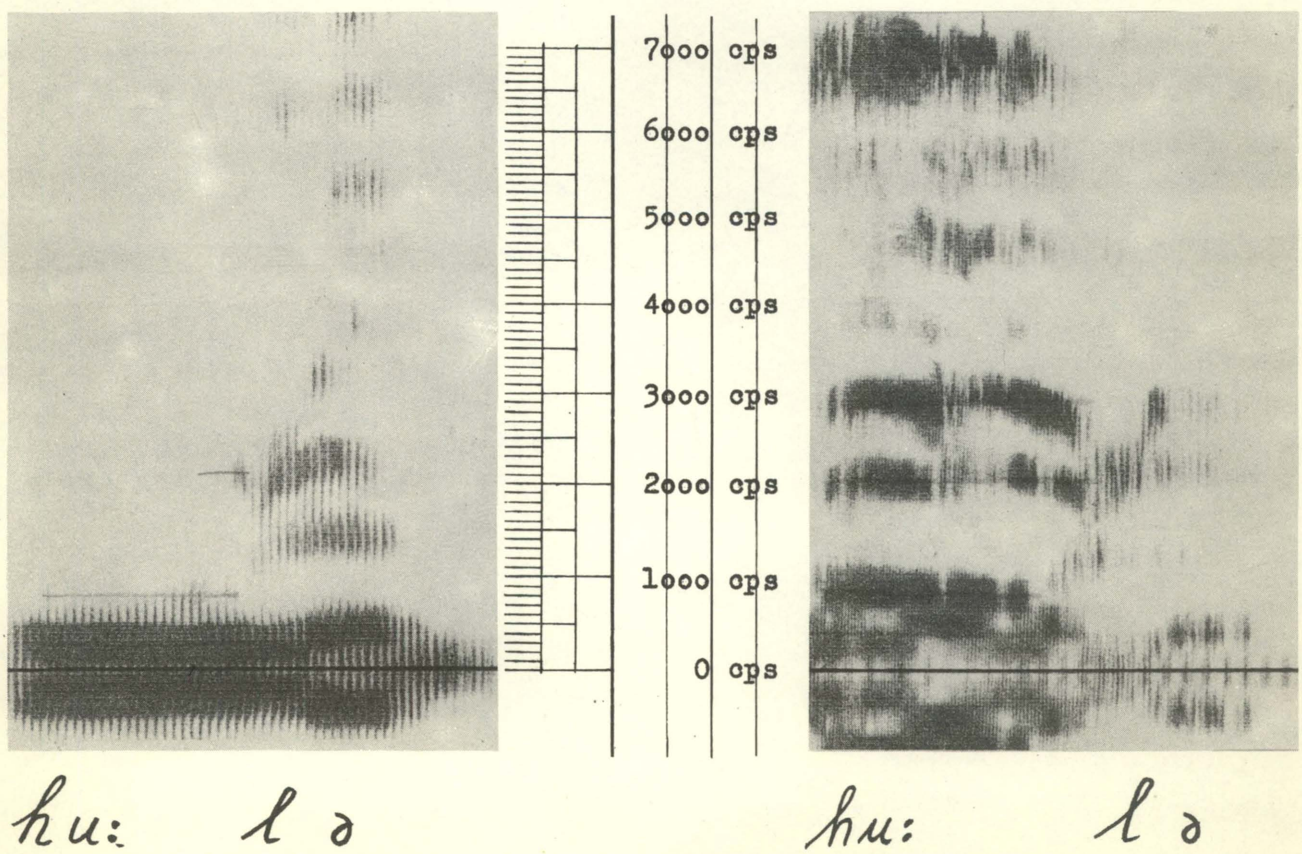
1. The spectrum above 2500 c/s is intensified and the spectrum below 1500 c/s is weakened. The frequency limits depend upon sex and age.
2. As a result of (C1) the harmonics in the range of the first formant may be reduced in intensity level in relation to the rest of the spectrum. In some cases F1 does not show up at all in the sonagrams.
   I have measured a 15-2o dB reduction of L1 (level of F1) in relation to the remaining part of the spectrum. (3)
3. F3 and F4 often contain more than half of the total energy (F5 is often prominent, and may even be the strongest formant).
4. The auditive result should be that the energy center of the upper formant areas thus moves upwards.

A few sonagrams illustrate this:

5. The next page shows sonagrams of the Danish words "hede" [he:ðə] and "hule" [hu:lə] . The first sonagram [he:ðə] is made from the recording of a female voice at normal speech effort. F1 is clearly most prominent; F2, F3, and F4 are all considerably weaker.
   The next sonagram shows the word [he:ðə] shouted by the same speaker. It is highly conspicuous that F1 does not show up at all in this sonagram of shout, and that the higher formants F2, F3, and F4 are very strong and well defined.
   The second sonagram of the word [hu:lə] shows the same shift in spectral energy. In normal speech nearly all the spectral energy is concentrated in F1. F2 is barely visible, and the same is true of F3. In the sonagrams of shout F2, F3, and F4 are stronger than F1.
6. The change in energy distribution of the shouted vowels is mainly due to a shift in poles and zeros in the primary voice spectrum (4),(5),(6),(7),(8). However, we are not yet equipped with instruments for acoustic investigation of the voice source (inverse filtering).



Normal speech - - - - - - - contra - - - - - - - shout



I have only shown a few of the most typical modifications that can be observed from my material. Other features of interest have not been investigated in sufficient detail for presentation.

D. Conclusion.

1. Because of the upward shift (see section B) of F_1 and the downward shift of F_2, the vowel quality of the front vowels changes in the direction close → open, whereas the back vowels, because of the constant relation between F_1 and F_2, change perceptually from the quality associated with male voices to that of female voices and ultimately to that of children's voices.
2. Because of the shift in intensity level (see section C) with the upper part of the spectrum emphasized and the F1-region weakened, we must get a perceptual change of the vowel quality in the direction open → close for the front vowels, whereas the back vowels will only get a brighter timbre.
3. In spite of the fact that both the formant frequencies (dependent on articulation) and the distribution of energy in the spectrum (dependent on primary voice source) are changed, the vowels are clearly identifiable.

   The nearest explanation seems to be that the changes in articulation are oppositely directed to the changes in voice source spectrum in such a way that the outcome of the two changes is that the identity of the vowels is preserved.

References:

(1) J.L. Flanagan, Speech Analysis, Synthesis and Perception (1965), pp. 32-34.

(2) Peter Ladefoged, "Sub-glottal Activity During Speech", Proc. IV. Int. Congr. of Ph. Sc., Helsinki (1961), p. 73.

(3) B. Frøkjær-Jensen, Comments of the closing session of "Speech Communication Seminar" in Stockholm (1962): "Variations in the Glottal Spectrum", Proc. of the SCS, Vol. III, Discussions, pp. 12-17.

(4) J. L. Flanagan, Speech Analysis, Synthesis and Perception (1965), pp. 37-44.

(5) J. L. Flanagan, "Some Influences of the Glottal Wave upon Vowel Quality", Proc. of the IV. Int. Congr. of Ph. Sc., Helsinki (1961), pp. 34-49.

(6) Gunnar Fant, Acoustic Analysis and Synthesis of Speech with Applications to Swedish (Ericsson Technics 1/1959) Chapter 1.25: "The Effects of Voice Level on Speech Spectra".

(7) C. Cederlund, A. Krokstad, M. Kringlebotn, "Voice Source Studies", Quarterly Progress and Status Report 1 (1960), Speech Transmission Lab. RIT, Stockholm.

(8) J. Lindqvist, "Inverse Filtering, Instrumentation and Technics", Quarterly Progress and Status Report 4 (1964), Speech Transmission Lab. RIT, Stockholm.

****

STUDIES OF THE PROSODIC PROPERTIES OF DANISH WORDS.

Jørgen Rischel.

The prosodic features characterizing Danish word structures are being analysed. It is to be examined in detail for Standard Danish how tonal contours, stød, etc. serve to signal on the one hand the syllable number (and the structure of syllables), on the other hand the grammatical structure of the word. An acoustical analysis is made alongside with a structural study of Danish word phonology.

Acknowledgements:

My Nordic studies are supported by a subvention from Ludvig F. A. Wimmer's Legacy.