VOL. 3. 1968

ANNUAL REPORT
OF THE
INSTITUTE OF PHONETICS
UNIVERSITY OF COPENHAGEN

ARIPUC Vol. 3. 1968.

ANNUAL REPORT OF THE INSTITUTE OF PHONETICS

OF THE UNIVERSITY OF COPENHAGEN

Copenhagen, 1969

Kbhvns. Univ. Offset-afd. nr. 2396/69.

CONTENTS.

CONTENTS (continued).

PERSONNEL OF THE INSTITUTE OF PHONETICS

1968.

Permanent Staff:

Eli Fischer-Jørgensen (professor, director of the Institute)
Jørgen Rischel (lecturer and amanuensis of general phonetics)
Hans Peter Jørgensen (amanuensis of general phonetics)
Oluf M. Thorsen (amanuensis of general phonetics)
Børge Frøkjær-Jensen (amanuensis of experimental phonetics)
Poul Thorvaldsen (engineer)
Svend-Erik Lystlund (technician)
Inger Østergaard (secretary)
Britta Olesen (secretary)

Part Time Teachers of General Phonetics:

Mogens Baumann Larsen (amanuensis of Danish phonetics)
Niels Knudsen (lecturer of German)
W.W. Schuhmacher (teaching assistant)
Hans Basbøll (teaching assistant)

Other Teachers :

Nils H. Buch, registrar, M.D.

Guest Research Workers:

Hideo Mase (Japan)

INSTRUMENTAL EQUIPMENT OF THE LABORATORY.

The following is a list of the instruments that have been purchased or built since January 1st, 1968.

Instrumentation for Speech Analysis.

1 Fabre-Glottograph (see this report, pp. 1ff.).

* 1 Kay-Electric Sona-Graph, type 6o61A.

* 1 Philips Stroboscope, type PR 91o3.

Instrumentation for Visual Recording.

* 1 Tektronix Oscilloscope, type 5o2A.

* 1 Brüel & Kjær Automatic Frequency Response and Spectrum Recorder, type 3332.

Tape Recorders.

1 Lyrec Professional Recorder (stereo, speeds 7.5" and 15").

* 1 Movic Professional Recorder (stereo, speeds 3¾" and 7.5").

Microphones.

1 Brüel & Kjær 1" Microphone, type 4131/32.

1 Brüel & Kjær 1/4 " Microphone, type 4135/36.

Amplifiers.

1 Brüel & Kjær Microphone Amplifier, type 26o3.

General-Purpose Electronic Instrumentation.

1 Brüel & Kjær Piston, type 422o.
1 Claude Lyons AC Automatic Voltage Stabilizer, type BTR-5F.
*)1 Brüel & Kjær Vacuum-Tube Voltmeter, type 24o9.
*)1 Radiometer Vacuum-Tube Voltmeter, type RV 23b.
1 Taylor Transistor Tester, model 44.
*)1 Stabilized Rectifier (6.3V, 15oV and 3ooV).

Outfit for photography.

1 Telford Oscilloscope Camera, type 'A',(Polaroid).

Projectors.

1 Voigtländer Perkeo Automat - J 15o.

*) Deposited at the Institute of Phonetics by The State Institute for Speech Defectives, Copenhagen.

LECTURES AND COURSES IN 1968.

1. Elementary phonetics courses.

One-semester courses (two hours a week) in elementary phonetics (intended for all students of foreign languages) were given by Hans Basbøll, Hans Peter Jørgensen, Niels Knudsen, Mogens Baumann Larsen, Jørgen Rischel, W.W. Schuhmacher, and Oluf M. Thorsen. There were 5 parallel classes in the spring semester and 17 in the autumn semester (slightly less than 500 students in all).

2. Practical training in sound perception and transcription.

Courses for beginners as well as courses for advanced students were given through 1968 by Jørgen Rischel and Oluf M. Thorsen. (The courses form a cycle of three semesters with two hours a week.)

3. Instrumental phonetics.

Courses for beginners as well as courses for more advanced students were given through 1968 by Eli Fischer-Jørgensen and Børge Frøkjær-Jensen. (The courses form a cycle of three semesters with two hours a week.)

4. Phonemics.

The standard courses (phonemic method and trends in phonemic theory) were given through 1968 by Eli Fischer-Jørgensen and Jørgen Rischel. In addition, Jørgen Rischel gave a separate course in generative phonology. (In total, the courses form a cycle of three semesters with two hours a week, the subject of the third term being chosen freely.)

5. Other courses.

Eli Fischer-Jørgensen lectured on diachronic phonetics through the autumn semester.

Nils H. Buch lectured on the anatomy and physiology of the speech-organs in the spring semester.

Eli Fischer-Jørgensen gave a course in the preparation of material for perceptual tests.

Børge Frøkjær-Jensen gave a course in elementary electronics and acoustical physics.

Moreover, the students followed courses in Danish, English, French, and German phonetics at the University.

6. Seminars.

The following seminars were held in 1968:

Amanuensis Niels G. Davidsen-Nielsen: sp,st,sk in English.

Professor K.N. Stevens (MIT) gave an informal account of some problems in articulatory and acoustic phonetics.

Dr. odont. Folke Strenger (Stockholm): Physiological investigation of nasal vowels in French.

Professor Eli Fischer-Jørgensen: Interference among French and Danish stops in the speech of a bilingual.

Hideo Mase, M.A.: The Japanese sound system.

Professor M. Onishi (Japan): The phonetic system of Japanese and its special features.

Stud. mag. Hans Basbøll: The phoneme system in a variety of Copenhagen Standard Danish.

Fil. dr. Björn Lindblom (Stockholm): Phonological and phonetic aspects of distinctive features.

Professor Eli Fischer-Jørgensen: Voice, fortis-lenis and aspiration in stops.

7. Participation in congresses and lectures at other institutions by members of the staff.

Eli Fischer-Jørgensen gave a lecture on the murmered vowels of Gujarati at the Royal Technical High School in Stockholm (cp. ARIPUC 2, pp. 35-85).

Jørgen Rischel participated in the Annual Meeting of the Societas Linguistica Europaea in Kiel.

CONSTRUCTION OF A FABRE GLOTTOGRAPH

Børge Frøkjær-Jensen and Poul Thorvaldsen

Introduction.

In the last decade a good deal of research has been done at different laboratories to develope equipment for investigation of the opening and closing movements of the glottis. In the year 1967 we have developed a photo-electric glottograph (3), and recently we have finished the prototype of an electrical glottograph according to the principles of Ph. Fabre (1), (4), (5), (6), (7).

The photo-electric glottograph registers the opening square of the glottis which probably corresponds very well to the primary voice source in the normal phonation range. The disadvantage of this type of glottograph is that not all subjects can or will insert the plastic tube with the light sensitive transducer through the nose down into the pharynx.

The Fabre Glottograph has not this drawback. This apparatus employs two electrodes which are fixed on each side of the thyroid cartilages. On the other hand some problems arise when curves are to be interpreted because of the up and down movements of the glottis and the air pressure variations in the pharynx both of which influence the glottograms. Conclusions are thus more difficult when based upon the Fabre Glottograms than when based upon the photo-electric glottograms (2).

The construction of both types of glottographs has enabled us to make comparisons between the two instruments in order to see to what extent it is possible to use the less precisely registering Fabre Glottograph instead of the better photo-electric glottograph.

2. The principle of the Fabre Glottograph.

The Fabre Glottograph is used for measurements of variations of the electrical impedance between the vocal cords. This impedance variation is measured by means of a high frequency alternating current which is sent through the larynx between two metal plates placed on each side of the thyroid cartilages. When the vocal cords vibrate they modulate the current between the metal plates. The modulated high frequency signal is rectified, filtered and amplified, and the resulting DC-signal which corresponds to the variations in the electrical impedance in the glottis may be recorded e.g. on an oscilloscope or on any form of recording device.

3. Requirements on our Fabre Glottograph.

When planning the construction we put up the requirements that our Fabre Glottograph

(1) should have the capability of being easily matched to the different glottis impedance of different subjects;

(2) should be able to handle frequencies from DC up to 10.000 cps (see Fig. 3);

(3) should have a low impedanced output of a few volts which would faciliate the matching to different recorders.

4. Survey of construction. (Fig. 4.)

The generator oscillates at the frequency 300 kcps. The generator consists of three stages, - a modified Hartley oscillator, a buffer stage, and an emitter follower.

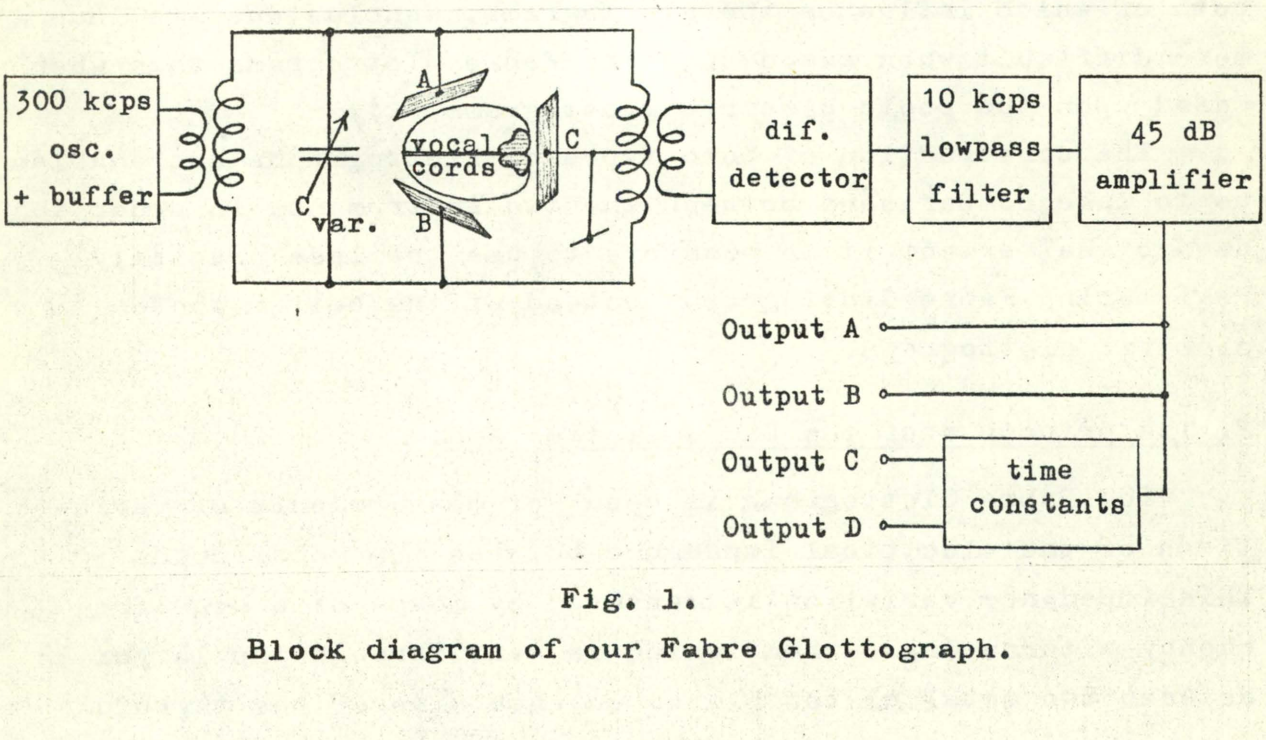


Fig. 1.

Block diagram of our Fabre Glottograph.

The changes in glottis impedance during fonation are sensed by two balanced electrodes placed on both sides of the laminae thyreoidei (Fig. 1). By balancing the electrodes it has been possible to cancel out interferences from electro-magnetic fields such as our local broadcast station.

Some trouble may arise when the glottograph is to be used by many different subjects, because both the electrical parallel-shunting resistance of the fat tissues between the electrodes and the electrical series-resistance of the fat tissues and the laminae thyreoidei vary considerable from one subject to another. This causes a reduction of the effective degree of modulation produced by the glottis which in turn results in a loss in the total sensitivity of the glottograph. In order to avoid some of these physiologically conditioned problems, we have designed the "glottis-circuit" in such a way that it is possible to adjust the circuit for best matching to the glottis impedance. A variable condenser is connected between the electrodes. By means of this condenser it is possible both to adjust the glottis circuit for maximum output (correct matching impedance) and with the same control to adjust the output level to zero volts DC, which is indicated by the meter in the output circuit. An electrical analog to the physiological part of the system is shown in Fig. 2.

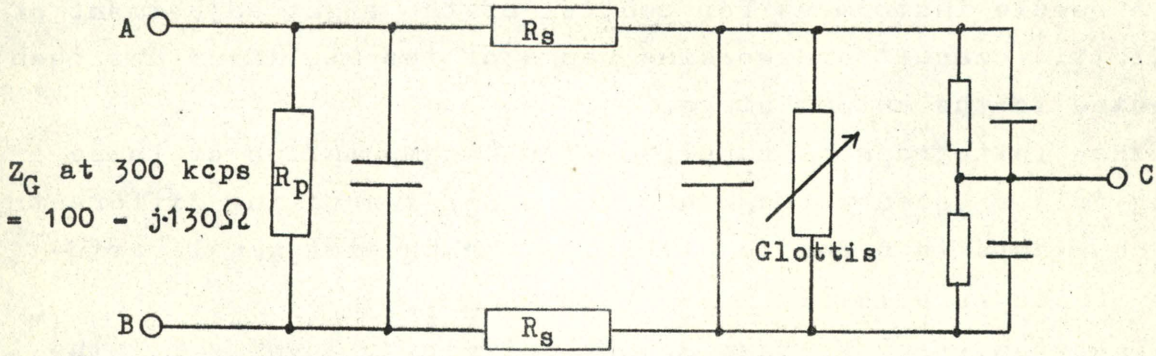


Fig. 2.
An electrical analog simulating the complex impedance of the larynx. (Subject Poul Thorvaldsen. Size of electrodes 2 cm^2 each. Conducting jelly has been applied to the electrodes.)

The resistors called R_p and R_s are represented by short-circuiting and resistive fat tissues which reduce the signal level.

The balanced output from the glottis is fed to a transformer with two symmetrical primary coils which is followed by a differential detector where the two transistors alternate between cut off and active conditions. This detector type is preferred because of its hum and noise rejection. The collectors of the double transistor are limited to a variation between ±8 volts by means of two zener diodes. This is necessary in order to avoid an excessive collector/emitter-voltage in the off condition.

The detector is followed by a 10.000 cps lowpass filter and an integrated amplifier (Philco type PA7709C). The input of this amplifier is also limited to ±8 volts to avoid damage. The condenser of 1 nF limits the frequency response of the amplifier to maximum 15 kcps.

The output stage is able to drive recorders which require some power. When short-circuited the output stage can deliver maximum 150 mA peak to peak. The maximum output voltage is 25 volts peak to peak.

The frequency response of the total electronic circuits from the electrodes to the output terminals is shown in Fig. 3. The frequency response has been controlled by means of a 300 kcps sine wave modulated with the low frequency oscillator in the spectrum analyzer.

A needle instrument for control of the right adjustment of the glottis circuit and working range of the amplifier has been connected to the output stage.

The glottograph is supplied with 4 outputs: two of these having full frequency range and the other two having different time constants in order to stabilize the physiological variations of the zero level.

The stabilized voltage supply delivers ±15 volts for the amplifier stage and for the reference voltages, whereas the generator unit and detector work at +80 volts.

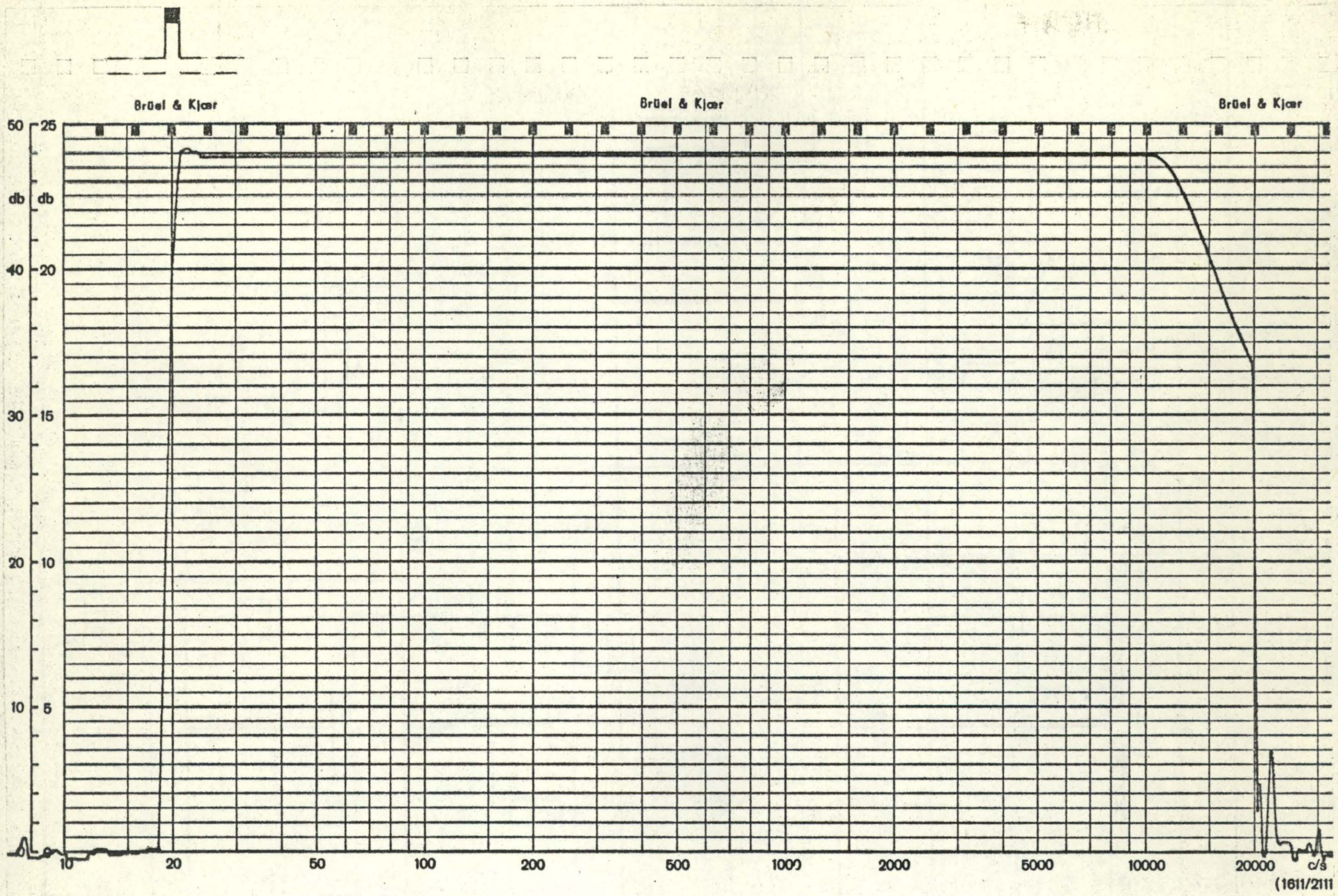


Fig. 3.

Frequency response of the glottograph without the final filters. The level variations from DC to 10.000 cps are below 0.1 dB. (50 dB scale). Above 10.000 cps the response falls at a rate of 18 dB/octave. The spectrum analyzer is not able to analyze below 20 cps.

Obtained specifications:

Oscillator frequency: 300.000 cps.

Generator impedance: 50 ohms.

Generator output, loaded throat electrodes: 0.8 volts RMS.

Generator output, unloaded throat electrodes: 1.6 volts RMS.

High frequency current through larynx: 15 mA RMS.

Voltage amplification of the DC amplifier: 180 times = 45 dB.

Mains supply: 110-130 volts or 220-240 volts/50-60 cps.

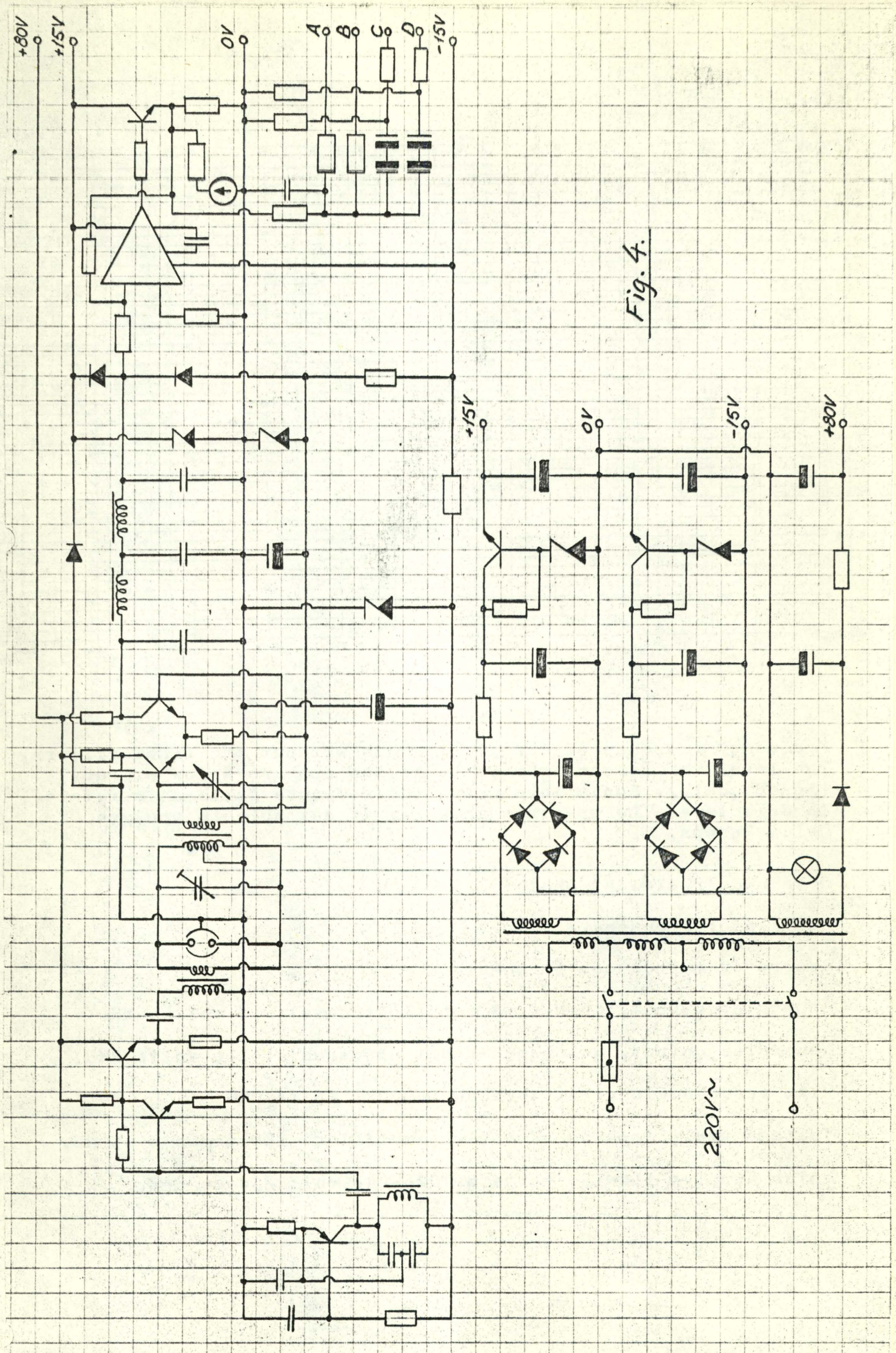


Fig. 4.

	Output A	Output B	Output C	Output C
Output impedance	60 ohms	500 ohms	5.6 kohms	10 kohms
Min. load impedance	0 ohms	0 ohms	0 ohms	0 ohms
Max. output voltage $R_L = R_{out}$, f=100 cps	3.0 V RMS	4.2 V RMS	4.2 V RMS	4.2 V RMS
Max. output current $R_L = R_{out}$, f=100 cps	50 mA RMS	2.4 mA RMS	0.75 mA RMS	0.42 mA RMS
Frequency range (-0.5 dB)	DC - 10 kcps	DC - 10 kcps	12 cps - 10 kcps T = 25 ms	0.08 cps - 10 kcps T = 4 sec.
Integrated noise voltage at output $R_L = R_{out}$	7 mV RMS	7 mV RMS	7 mV RMS	7 mV RMS
Dynamic range $(10 \cdot \log \frac{P_{out}}{P_{noise}})$	53 dB	56 dB	56 dB	56 dB

Specifications of the four outputs.

Fig. 5.

As for the results obtained with the above described Fabre Glottograph see the article in this report: "Comparison between a Fabre Glottograph and a Photo-electric Glottograph".

Acknowledgement:

The construction of the Fabre Glottograph is supported by "Statens Almindelige Videnskabsfond" (The Danish State Research Foundation).

----*****----

References:

(1) Ph. Fabre, "Glottographie respiratoire", Comptes rendus des séances de l'Académie des sciences No. 6 (1961), pp. 1386-1388; id., "La glottographie en haute fréquence, particularités de l'appareillage", Comptes rendus de la Soc. de biologie CLIII, 7 (1959), pp. 1361-1364.

(2) Eli Fischer-Jørgensen, Børge Frøkjær-Jensen and Jørgen Rischel, "Preliminary Experiments with the Fabre Glottograph", ARIPUC No. 1 (1966), pp. 22-3o.

(3) Børge Frøkjær-Jensen, "A Photo-electric Glottograph", ARIPUC No. 2 (1967), pp. 5-19.

(4) R. Husson, Physiologie de la phonation (1962), pp. 32-4o.

(5) Cl. van Michel, "La glottographie. Son intérèt en physiologie phonatoire", Revue Médicale de Liège Vol. XXI, No. 12 (1966).

(6) Cl. van Michel, "Recherches électroglottographiques en physiologie phonatoire expérimentale et en clinique otorhinolaryngologique", Extrait des Acta du premier congrès latin de logopédie, Louvain, Avril 1966.

(7) J. Ondrácková, "Glottographical Research in Czech Sound Groups", Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung, Band 21, Heft 1/2 (1968).

(8) L. Striglioni, Contribution à l'étude de la physiologie laryngée en voix parlée. Apport de l'électroglottographie. Thèse de doctorat, Marseille (1963).

COMPARISON BETWEEN A FABRE GLOTTOGRAPH AND A PHOTO-ELECTRIC GLOTTOGRAPH

Børge Frøkjær-Jensen

1. Introduction.

In this article are published a few preliminary results from the test of our Fabre Glottograph and comparisons between our photo-electric glottograph and the Fabre Glottograph. Glottograms of most of the Danish consonants are compared with the airflow curves and the oscillograms, and the three-dimensional type of vocal cord oscillation is shown by comparing the curves from the two types of glottographs. *)

1.1. Airflow registration.

Some simultaneous recordings of the airflow (trace no.1), the electric impedance of glottis (trace no.2), the square area of the glottis (trace no.3), and the acoustic oscillogram (trace no.4) have been made for this article. The airflow recordings are made by means of the Electro Aerometer and show the outlet airflow measured in volume unit per time unit (cm^3/sec.), - the instrument can be exactly calibrated but has not been calibrated for the illustrations in this article. A raising of the airflow trace in the recordings means that the outlet airflow is increased.

1.2. Photo-electric glottogram.

The photo-electric glottogram shows the square area between the vocal cords in a horizontal plane. The photo-electric glottograms have a well defined zero level for closed glottis. However, the glottis may sometimes be closed though the zero level in the glottogram trace has not been reached, this is because of a certain degree of transillumination of the closed glottis which can be observed especially in the head register.

With some subjects we have experienced trouble when inserting the plastic tube with the photo transistor (light

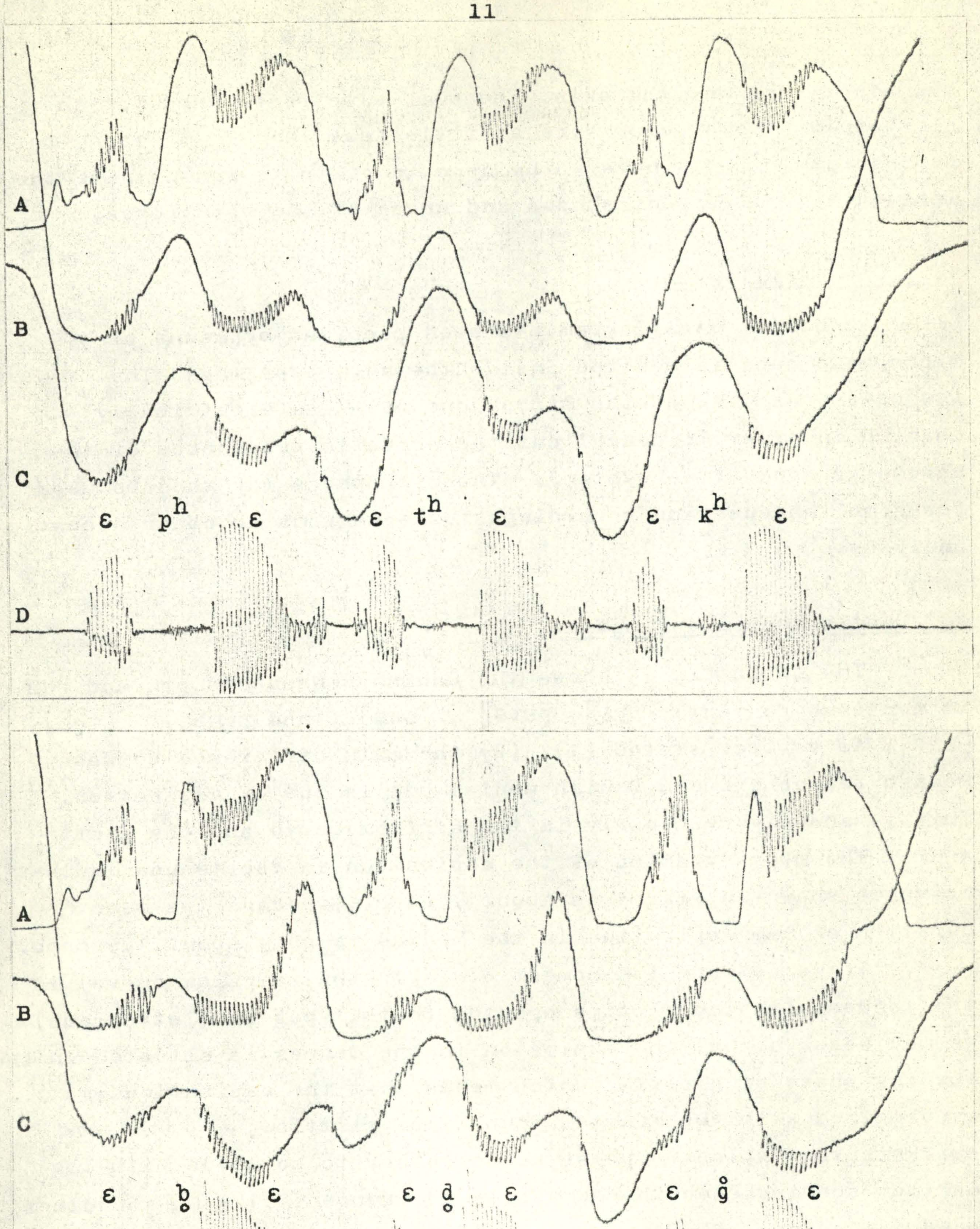
*) Both types of glottographs and the Electro Aerometer used for this article are manufactured for sale by B. Frøkjær-Jensen, Vestre Paradisvej 46, 2840 Holte, Denmark.

transducer) through the nose down into the pharynx. We have to a great extent overcome this problem by mounting the transducer in a 2 mm plastic tube and by glueing the free end of the transducer to the tube by means of liquid plastic mass. Furthermore, both a small lense in the end of the photo transistor and the black painting of the transistor has been removed. These improvements of the transducer makes it easier to insert it through the nostrils without much inconvenience. The transducer may be inserted in position in the pharynx and fixed outside the nose by means of a piece of court plaster. Thus fixed it can be used for hours without discomfort and without disturbing the articulation (no local anaesthesia is necessary). It is also possible to pick up the light in the pharynx without serious interference from the articulatory movements. Only when articulating the most open vowels [ɑː] and [ɒː] the back tongue and the epiglottis push the photo transistor out of position. These vowels can, therefore, normally not be recorded by the photo-electric glottograph.

1.3. The Fabre Glottograph.

The Fabre Glottograph shows the electric contact area between the vocal cords in a vertical plane. However, different articulatory movements of the pharynx and of the larynx may influence the electric impedance between the electrodes, as pointed out earlier (3). These undesirable physiological changes cause a very jumpy zero line which is difficult to stabilize in the registration range of the recorder. Our glottograph has therefore been provided with four different outputs, two of which have different time constants (25 msec. and 4 sec.) for stabilizing the zero line during articulation (6). (These outputs have not been provided for the glottograms shown in this article.)

Some few subjects seems to have a throat unfitted for electric impedance measurements because of too much fat which results in a bad signal-to-noise ratio. The same sort of noise disturbance of the glottograms may be observed if the electrical contact between the skin and the electrodes is insufficient. Therefore we always use electrically conducting jelly between



A: Airflow curve
B: Photo-electric glottogram
C: Fabre glottogram
D: Oscillogram

Speed: 100 mm/sec.

Fig. 1.

the electrodes and the skin. The oval electrodes used for the glottograms shown below were a little less than 2 cm^2 each, but further experiments have to be done in order to investigate the influence of different shapes and areas of the electrodes.

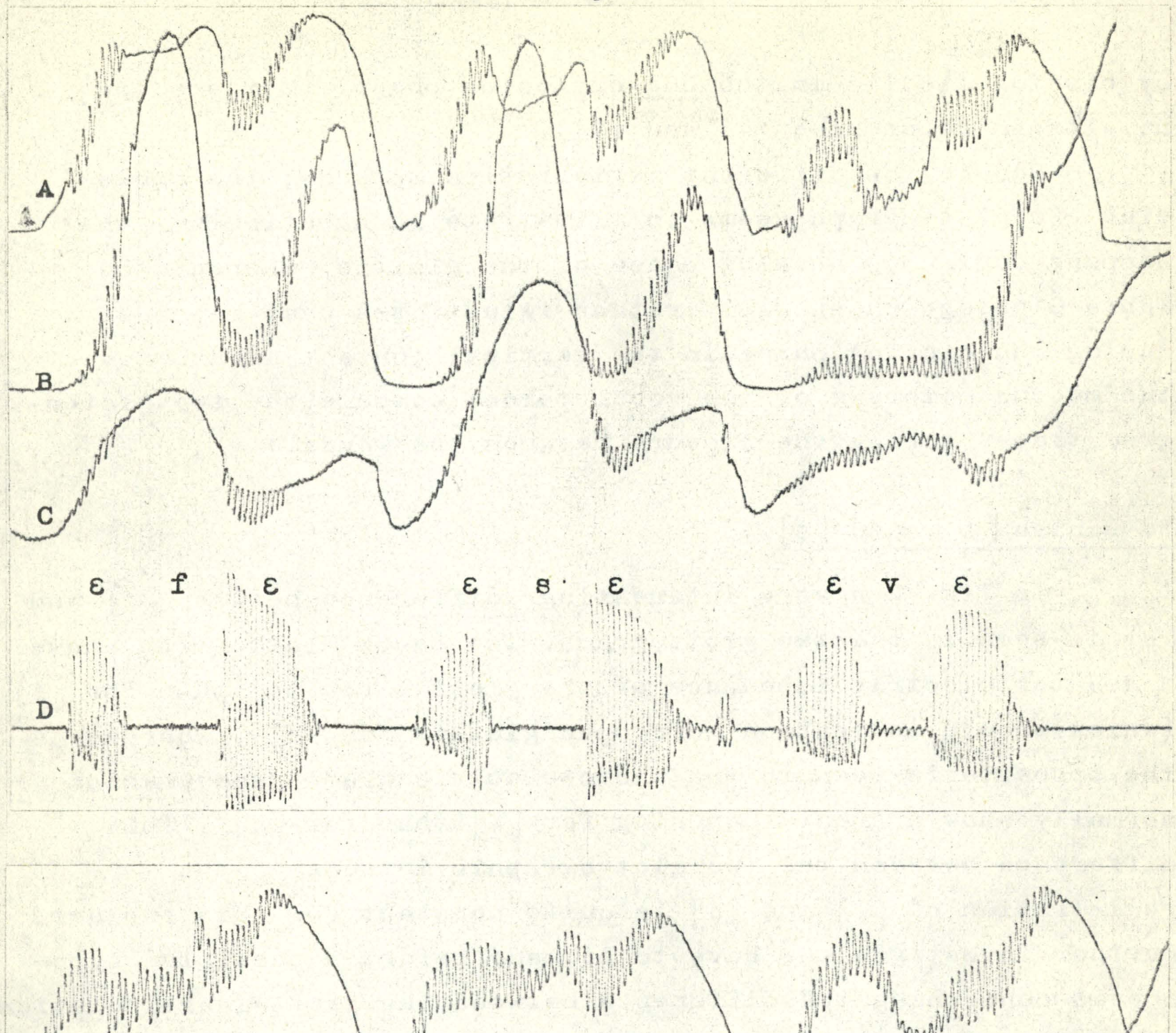
1.4. The oscillogram.

A normal oscillogram has been recorded by means of a microphone capsula mounted inside the aerometer mask. The frequency response of the microphone has been electrically changed in order to cancel out the acoustic resonances of the aerometer mask (tube system). The microphone may thus be used for simultaneous tape recordings for sonagrams or other acoustic analyses.

2. Danish tenues and mediae.

The table Fig. 1 shows the Danish consonants ptk and bdg in nonsense context: [εpε], [εtε] , [εkε] , and [εb̥ε], [εd̥ε], [εg̊ε]. As earlier stated (1), (2) the main difference between Danish tenues ptk and Danish mediae bdg is one of aspiration. This is shown very clearly in the airflow curve and the glottograms. The mean duration of the aspiration in the tenues in the examples shown in Fig. 1 is about 7 csec., whereas the mean duration of the aspiration in the mediae is only about 1.5 csec.

If we compare the moment of explosion (airflow curve) in the tenues with the glottis opening (both types of glottograms) we may observe that the explosion in the tenues is situated in the opening phase of glottis, which means that the aspiration is syncronized with the maximum opening of glottis. - With some speakers the explosion phase seems rather to coincide with the maximum opening, see the curves in reference (3). - On the other hand the curves of mediae show that the moment of explosion in bdg is situated in the closing phase of glottis, which conditions that only a little amount of air is left for a possible aspiration. Fig. 1 thus very clearly proves that the main difference between the Danish ptk and bdg (based on 256 examples for only one speaker BFJ) is one of opening glottis during the ptk-stop closure and fully open glottis during the ptk-aspiration versus medium open glottis during the bdg-stop closure followed



Speed: 100 mm/sec.

A: Airflow curve
B: Photo-electric glottogram
C: Fabre glottogram
D: Oscillogram

Fig. 2.

by closing glottis in the bdg-explosion phase. (Compare the results in references (4) and (7)).

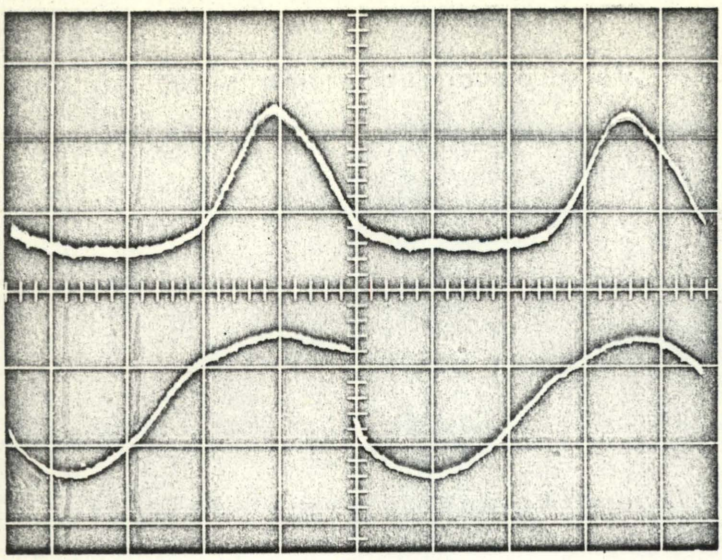
Due to the different registration methods, the photo-electric glottograph seems to illustrate in a better way what happens during the opening phase of the glottis, whereas the Fabre Glottograph probably rather illustrates what happens during the closing phase in the vertical contact area between the mucous membrane of the vocal cords (compare the two glottogram traces during the closure between the words.)

3. Danish [f] and [s].

In Fig. 2 a very interesting difference between [f] and [s] is seen in the two glottograms. The Fabre Glottograph shows a greater electric impedance of the glottis for [s] than for [f] (which should indicate a more open glottis for [s]), whereas the opposite is seen in the photo-electric glottograms, which normally show a greater opening for [f] than for [s] . This difference between the two glottographic methods in the registration of [f] and [s] is quite constant for this subject. Further investigations have to be made before a possible hypothesis concerning the difference between the glottal articulations of [f] and [s] can be advanced.

4. The voiced consonants.

Fig. 2 also shows some examples of voiced consonants. Especially the Danish [ð] shows the difference of phonation between voiced consonants and surrounding vowels. The electric impedance of the glottis is considerably raised during the articulation of [ð] , i.e. the contact area between the mucous membranes of the vocal cords in the vertical plane is reduced. However, the photo-electric glottogram does not show such a change in phonatory conditions (see the article in this report: Fourier Analyses of Photo-electric Glottograms). Once again we may conclude that the two types of glottographs deliver two different sets of parameters concerning the working conditions of the human larynx.



Vowel[ɛ:], spoken by subject BFJ. Chest register, medium voice effort, fundamental frequency 120 cps.

Fig. 3.

5. Recordings of the primary voice source cycle.

Fig. 3 shows a polaroid photo of simultaneous recordings made by the two glottographs. The photo is a good illustration of how we can get information about the three-dimensional opening and closing movements of glottis.

If we compare the slope of the glottal cycle in the Fabre glottogram (the lower trace) with the photo-electric glottogram we can only explain the different forms of the two simultaneously recorded curves by assuming the explosions (excitations) as a slowly opening movement starting from below and moving upwards. When it reaches the edges the vocal cords open up and we get in the upper curve a registration of the opening area. When the glottis closes the Fabre glottogram shows that the two mucous membranes in the vertical plane suddenly come into touching contact with each other, i.e. we have a gradually opening and a suddenly closing movement.

This function of the primary voice is typical for the chest register (8), (9), which operates both in the vertical and the horizontal plane.

Recordings of the typical head register function do not show this three-dimensional type of oscillation.

Acknowledgement.

The construction of the two glottographs has been supported by "Statens Almindelige Videnskabsfond" (the Danish State Research Foundation).

References.

(1) Eli Fischer-Jørgensen, "Acoustic Analysis of Stop Consonants", Miscellanea Phonetica II(1954), pp.42-59.

(2) Eli Fischer-Jørgensen, "Phonetic Analysis of Danish Stop Consonants", ARIPUC 1(1966), pp. 31-33.

(3) Eli Fischer-Jørgensen, B. Frøkjær-Jensen, and J. Rischel, "Preliminary Experiments with the Fabre Glottograph", ARIPUC 1(1966), pp. 22-31.

(4) L. Lisker, A.S. Abramson, F.S. Cooper, and M.H. Schwey, "Transillumination of the Larynx in Running Speech", Status Report on Speech Research SR-5/6(1966), Haskins Laboratories.

(5) C. van Michel, "Les Artefacts en Glottographie", Revue Electro Diagnostic Thérapie 4(1967).

(6) Jana Ondrácková, "Glottographical Research in Czech Sound Groups", Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung, Band 21, Hft.1/2(1968), pp.145-152.

(7) I.H. Slis and P.H. Damsté, "Transillumination of the Glottis during Voiced and Voiceless Consonants", Annual Progress Report 2(1967), IPO, Eindhoven.

(8) Svend Smith, "Remarks on the Physiology of the Vibrations of the Vocal Cords", Folia Phoniatrica 6(1954), pp. 166-178.

(9) Svend Smith, "Vertikale Komponenten bei der Funktion der Stimmlippen", Phonetica 11(1964), pp. 243-247.

CONSTRUCTIONAL WORK ON A FUNCTION GENERATOR FOR SPEECH SYNTHESIS

Jørgen Rischel

1. General status of speech synthesizer project.

As mentioned briefly in last year's Annual Report (6), a formant-coded speech synthesizer is being constructed at the Institute. The apparatus consists of two parts: a function generator supplying the control voltages for synthesis of connected speech and a sound producing system comprising voice source, noise source, formant filters, and variable gain amplifiers associated with the formant filters.

As for voice and noise sources, our present devices exhibit no special features.

The formant filters for vowel synthesis have been designed along the same lines as the circuits employed in a provisional vowel synthesizer built by this author some years ago (5). The filters are high-Q LC series circuits in which part of the capacitance is represented by capacitance diodes (type BA 163), and the resonant frequencies can be varied simply by varying the bias voltage applied to the diodes. A sufficient range of frequency variation for the individual formants is obtained by letting the circuits resonate at rather high frequencies, the frequency spectrum of the voice source pulses being shifted correspondingly by a heterodyning process. This, of course, involves the use of a modulating and a demodulating system (common to all formant circuits). In order to obtain good amplitude and phase linearity down to low frequencies we have designed a transformerless double-balanced modulator, which seems well suited for the purpose.

The new version of this whole system is not yet ready for use, so it would be entirely premature to give any circuit details. (Our "phonetic" experience is limited to the old model.)

Similar circuits (but including antiformant filters) have been planned for nasals and stops/fricatives.

The use of frequency transposition in the formant filter system necessitates a parallel connection of the filters (with phase relationships taken care of). This means that the level of each formant can and must be controlled independently by the experimenter. Our newly constructed formant circuits include voltage controlled attenuators exhibiting a linear relationship between control voltage and output level over a 30 dB range. The freedom of control thus offered is crucial for some kinds of experiments, but there may certainly be other cases in which the experimenter wishes formant levels to be derived automatically from the pattern of formant frequencies. This can perhaps be done by means of nonlinear circuits, possibly a "potential field"(7). However, since the formant filter system and the function generator are functionally independent units it would also be possible to use alternative versions of each (e.g. a series connexion of filters can be added eventually) and to combine these freely according to the immediate purpose. This may be a simpler solution.

In the past year a major part of our efforts within the synthesis project have been devoted to the development of an expedient function generator, which can store an arbitrary configuration of synchronously varying control parameters and, on request, reproduce a corresponding set of time-dependent voltages. The part of the generator that functions by now can deliver eight independent voltage functions which can be used to specify a sound sequence lasting not more than 2.0 seconds (this maximum duration is possible only with a rather coarse resolution in time). The number of parameters is going to be increased (this is simply a matter of adding strings of potentiometers in parallel with those already present), whereas we have no immediate intention of increasing the capacity in the time dimension very much.

Various considerations have influenced our choice of type of function generator, such as: a desire to avoid demanding mechanical constructions, demands for accuracy and good reproducibility, and a wish to be able to synthesize sound sequences at different information rates ("fine surgery" and "synthesis by rule" being the extremes).

The function generators most widely used with analog synthesizers are in a sense simple transducers: they convert an exact graphical representation of each parameter (mostly in the form of a painted trace on a sheet) into an electrical control signal, and ideally, the traces painted by the experimenter should be exact replica of the formant movements etc. to be synthesized. Our function generator belong to another category of devices which represent the configuration of parameters as a sequence of states each lasting for a preset amount of time. In this case the experimenter must somehow specify the parameter values once per temporal segment (the duration of each of these may, of course, be quite short if a good definition of transient phenomena is the research objective). The output voltages will thus be varying stepwise. In the following the function generator (taken in a restricted sense) producing these step functions will be referred to as the "staircase voltage generator".

The most obvious problem in using a staircase voltage generator for analog synthesis control is that the discontinuities (jumps from one voltage to another) must somehow be smoothed out in order not to produce transient bursts in the sound producing system. This smoothing out is in fact a very crucial feature of the system generating the control voltages, since the choice of smoothing function may influence the specification of the input information to the function generator considerably. The problems which have presented themselves to us will be dealt with in some detail in a later part of this paper.

Our staircase voltage generator was designed independently of other similar devices in current or earlier use (a leading principle being to avoid mechanically functioning parts), but it must be stated that the type as such is not at all new (1). Each parameter is represented by a series of potentiometers (one per time segment) whose scales can be calibrated directly in cps or decibels depending on the parameter. The staircase voltage is produced by setting each potentiometer to the appropriate voltage and scanning the whole series by means of a multiplexing system. This approach is supposed to be favourable in the case of experiments in which relatively short sequences are to be synthesized and altered systematically in respect of one or several parameters (typically perception tests). Changing the setting of a knob is likely to be a more agreeable job than deleting a trace and painting a different trace. The difference is most obvious the moment the duration of a single portion of the sound sequence is to be changed. With our function generator this can be accomplished within wide limits by means of a switch, "duration" being a separate parameter of the system.

Electronically, the function generator has been designed in such a way that a fairly high degree of stability is obtained. The multiplexing system is a sequence of gates each of which triggers the next after having been in the "ON" state, the durations of the individual states being determined by adjustable CR circuits. Each section includes a Schmitt trigger ensuring well defined switching times and "ON" voltages (the rise time of the gates is considerably better than 50 microseconds, which is sufficiently short for the purpose). The actual multiplexing is obtained by alternative biasing of diodes as scetched in Fig. 1.

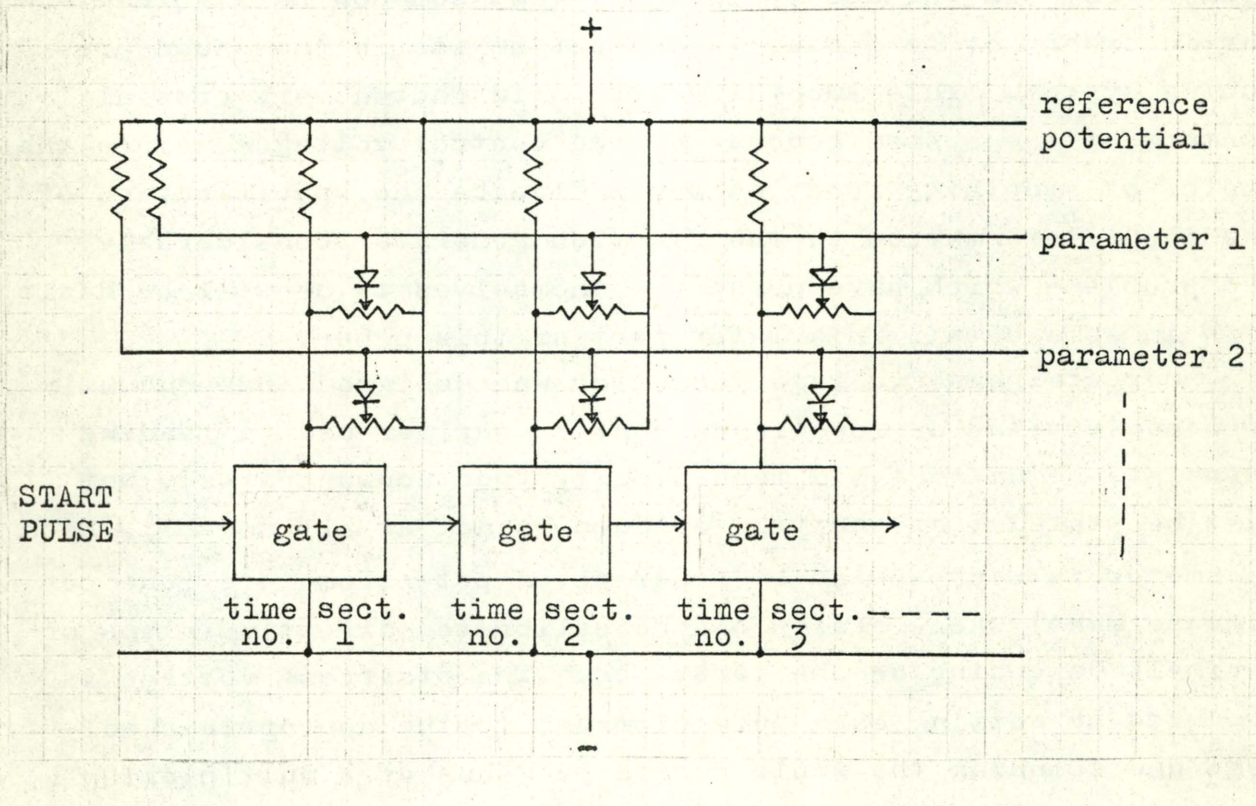


Fig. 1.

It may be argued that a function generator of this kind copies some of the properties of a digital generator (either in the form of a flexowriter or in the form of a small computer) without sharing other of its virtues, such as complete temperature stability,*) extreme versatility and storage capacity, and appropriateness for other purposes than synthesis. It should be noted, however, that the analog function generator is in several respects preferable to a generator working from a punched tape (flexowriter). It shares with the on line operated computer such features as immediate feedback and easy manipulation of stored data. Conceptually, the manual setting of parameters is an extremely simple matter, so the system seems attractive also for teaching purposes. And even if the system may eventually become obsolete with a complete change of emphasis from analog to digital control methods even in work on a small scale, we feel confident that experience with it will furnish a very useful background for later computer-oriented work.

A direct comparison between analog and digital control of the same synthesizer would probably not show any considerable difference in the time it takes to synthesize a <u>short</u> sound sequence by the two methods, provided that the segment durations and the smoothing function (see below) are optimal.

2. Quantizing control signals in time.

When planning our function generator we considered different possibilities of performing a quantization in time, the only prerequisite being that for practical reasons all parameters must be specified synchronously (even though the specification often consists in a repetition of the preceding value for one or several parameters). Although only one of these was deemed really acceptable they will all be described briefly in order to make the point clear:

1) One possibility is to specify parameter values a fixed number of times per second. Fant (3a) has obtained good quality of synthetic speech with a sampling rate of some 40 per second.

---

*) As for temperature-dependent drift the superiority of digital systems is most obvious if the formant circuits are controlled digitally as in OVE III of the Speech Transmission Laboratory in Stockholm (4). If the digital control signals must be converted into analog signals before they can be applied to the circuits some of the calibration problems must return. Nevertheless, this may be a compromise worthy of consideration.

This approach is of interest in analysis-synthesis techniques, but for our purposes is must be immediately discarded because the number of potentiometer settings will be prohibitively high.

2) A more favourable solution (which forms part of our present "strategy") is to make the duration of each segment independently variable. This means that the "specification rate" can be made quite high in the transient parts of the speech-wave but considerably lower in parts where only slow changes occur. However, since the smoothing function must be chosen in accordance with the highest specification rate used the smoothing will be imperfect in parts with a low specification rate. Since the aberrations from the true parameter slopes resulting from imperfect smoothing can only be tolerated to a certain degree, this problem sets a limit to the gain in economy that can be obtained by varying the specification rate (with the exception of obvious cases like silent intervals).

3) A considerable gain in economy is obtained if the transition rate of the parameters can be reflected by corresponding values of the smoothing function. Ideally this would mean that sequences like ba, wa, ua could be represented by essentially the same stored configuration of parameter values, the differences in transition rates being brought about by varying the time constant of the smoothing circuit. - If only a single CV combination is being synthesized, such switching poses no problem, but in longer sequences the transition rate may change from one part of a syllable to the next. Thus a full exploitation of this method requires that "transient time" be included as a parameter along with "segment duration". The strategy then approaches the concept of "synthesis by rule", since a sequence of two sounds will be represented by two sets of target values (valid for the particular combination), two segment durations, and a transient time (being the time elapsing from the theoretical beginning of the second segment until its target value is reached) which is conditioned by the type of sound combination, but specified independently by the experimenter.

4) Finally it is possible to design the smoothing circuit in such a way that a greater or lesser "overshoot" in the voltage step function speeds up the transition from one target value to the next. It is thus possible to vary the transient

times of the individual parameters quite independently. The drawback of this method is that the possibility of calibrating the potentiometer scales is lost, since an increment in the setting of a potentiometer may function both to speed up a transition and to increase the target value of the parameter at this point. - This loss of a simple relation between potentiometer setting and parameter value is a reality if ordinary integrators are used as (the first stage of) smoothing circuits (cp. below).

Solution "3" above (i.e. parametric control of transient time) has seemed to us the most attractive solution, although it does not remove all inconveniences stemming from the quantization procedure. Technically, the design of electronically adjustable smoothing circuits has turned out to pose very difficult problems, and we have obtained acceptable results only recently. There are two sources of difficulties: one is that the exact shape of the step response of the smoothing circuit becomes quite crucial when slowly varying phenomena have to be generated from step functions, the other is that there must be no interference at all between the switching voltage controlling the transient time and the voltage function to be smoothed out. It is no overstatement that the circuits we are experimenting with are unduly complicated. Nevertheless we consider the implications of the approach so interesting that we are going to build such smoothing circuits for at least some of the parameters.

Although circuit details will not be dealt with at all in this report, the principal questions raised in connexion with the design can be approached from the more general point of view of "synthesis strategy". The major points which we have had to consider will be presented below.

3. Transition shaping circuits.

The simplest way to convert a voltage step into a raising or falling slope is to use an RC circuit. The slopes generated by such a circuit are, however, exponential. Spectrograms of real speech reveal that the formant transitions in consonant-vowel sequences can often be matched very well by means of such exponential slopes, but in vowel-consonant sequences the agreement will often be rather poor. The differences will be clear from the stylized curves shown in Fig. 2:

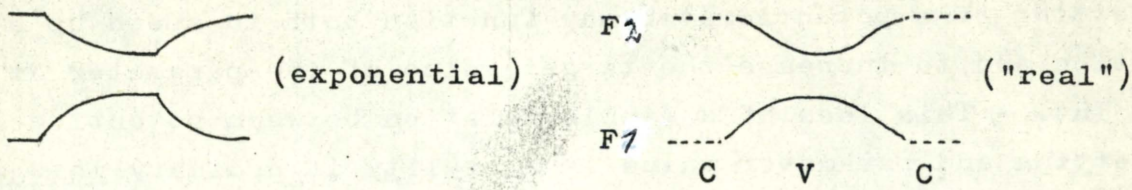


Fig. 2.

A particular inconvenience arising with the use of RC circuits is that theoretically the target values are never reached. The transient time must be defined as the time it takes for the slope to reach 63 per cent of its final value (or rather 63 per cent of the difference between the previous and present target values), but the control voltage occurring at this or any other point in time will never be equal to the calibrated value read off the function generator. With very long time constants this may be a problem.

An oblique slope can also be obtained from a voltage step function by using a lowpass filter with good phase characteristics. This is in fact a very attractive solution, but if very different transient times are to be used in the same piece of synthesized speech difficulties may arise, both because of the switching problems and because of time lag phenomena. The inconvenience caused by the variation of delay time with filter cutoff frequency may be serious if an exact control of the temporal relationships among the events is required.

Conceptually, the simplest solution is to produce a linear slope going obliquely from the point where one time segment ends and reaching the target value of the next segment sooner or later depending on the transient time chosen. Such a linear slope is not an ideal approximation to a formant transition, for example, but it is a simple and well-defined function to work with. If an extreme degree of fidelity is required, it is perfectly possible to generate a formant transition by combining two or several such linear slopes to give the desired curvature, but there is probably seldom any reason to do so. - The discontinuities present at the points where the slope starts and ends can be removed by simple means if this is considered essential. In our experimental set-up the slope-shaping circuit is terminated by a buffer amplifier, which by the addition of a few components can be changed into a simple (active) lowpass filter. If a sufficiently high cutoff frequency is chosen for this filter, it will not have any appreciable influence on the shape of the control voltage passing through it except that disturbing discontinuities are removed. The oscillogram in Fig. 3 below is shown without this additional filtering in order to exhibit the linear slope more clearly, whereas the oscillograms shown later in this paper are representative of the voltages after filtering.

Theoretically the simplest way to produce such oblique slopes is by linear interpolation between values specified at discrete points in time. This approach has been employed in the computer control of OVE III in Stockholm.

We have not investigated into the possibilities of constructing an expedient "interpolator" by analog methods. Our considerations were based on the assumption that the input function must be a step function and that we wish to change this step function into a ramp function with the same minimum and maximum values and with a duration of the oblique slope which can be controlled externally.

According to this method every constant-voltage segment of the input function (the voltage delivered by the staircase voltage generator) is replaced by a succession of an oblique and a horizontal portion.

Phonetically it is implied that a transition is always counted with the <u>following</u> segment. If, for example, the bottom curve of Fig. 3 (next page) is taken as a crude approximation to the first formant movement in a sequence like <u>dad</u>, it is seen that this CVC sequence can be synthesized by means of three segments. Each of these comprises a transition plus a following steady portion (the latter may, of course, be absent if the transition is made equal in duration to the whole segment) but it is specified by the experimenter in terms of three simultaneous parameter values: a target value for the formant in question, a duration of the transition to this target value, and a duration of the total segment.

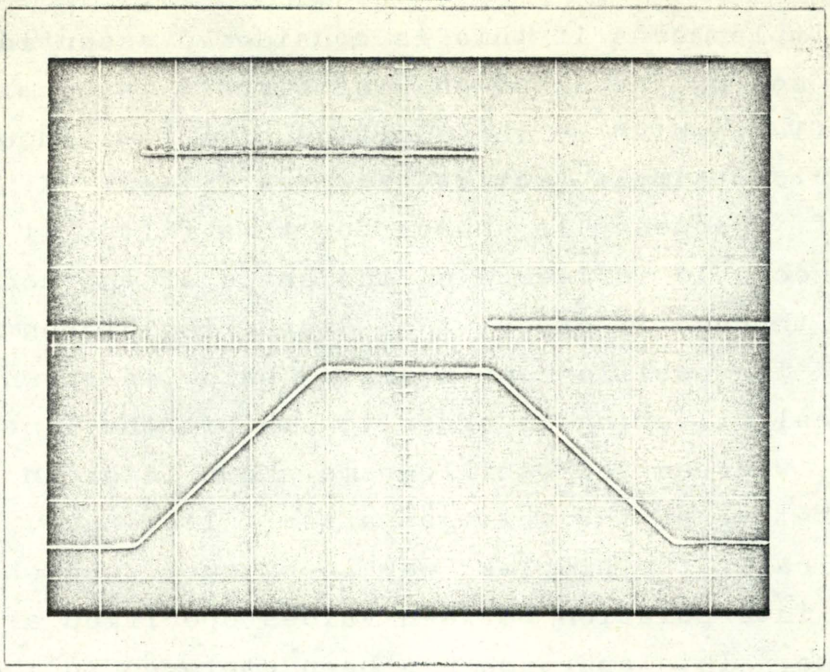


Fig. 3.

Trapezoidal voltage functions like the one shown in Fig. 3 can, of course, be produced by integrating a succession of positive and negative square pulses, cp. Fig. 4, which shows the input to and output from an operational amplifier coupled as an integrator:

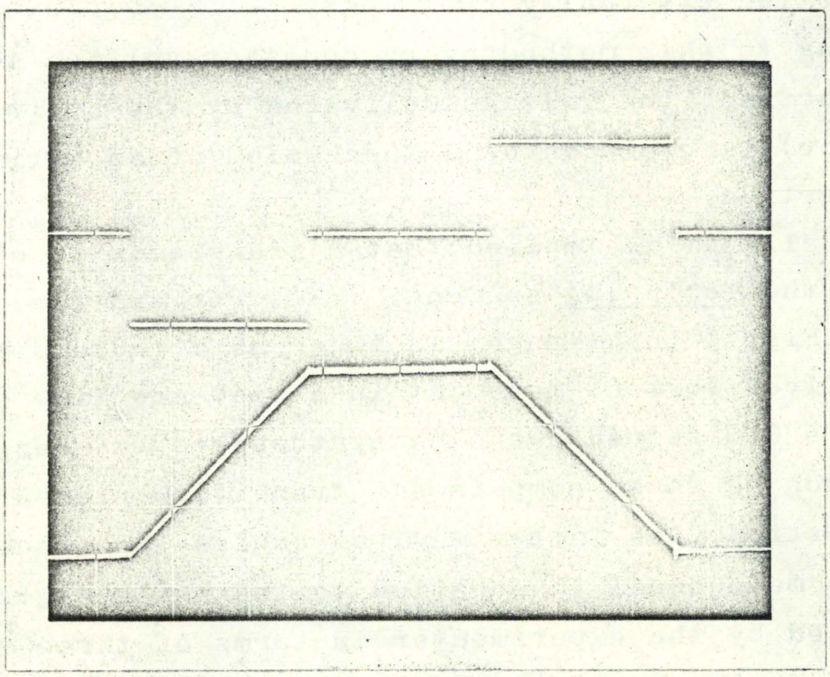


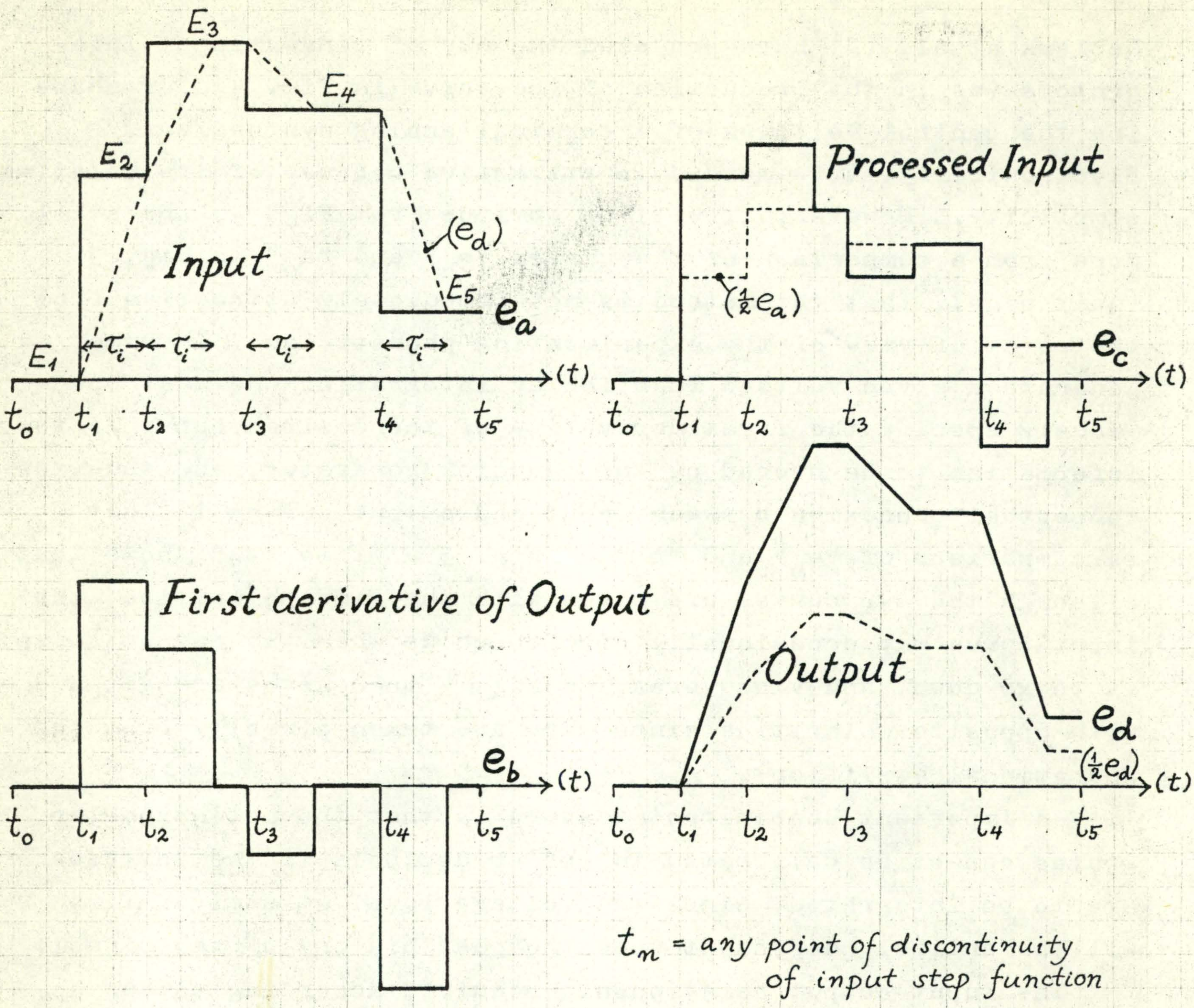
Fig. 4.

DeClerk et al. (2) have proposed the use of conventional integrators (as in the production of the curve in Fig. 4) for shaping the control voltages of a terminal analog synthesizer. Electronically, this method is attractive because of its relative simplicity. However, it will be obvious from Fig. 4, and still more from a comparison of the curves "e_b" and "e_d" in Fig. 5 (next page), that the method is not immediately attractive from the point of view of the experimenting phonetician. In order to generate a given voltage function by integration one must obviously specify the first derivative of the desired curve as the information to be stored by the function generator, and thus the conceptual connexion between input and output curves is lost (a comparison of "e_b" and "e_d" of Fig. 5 at time "t_2" shows that although the two curves are given with the same polarities, the input curve may occasionally have to go up when the output curve is to go down, and vice versa; in Fig. 4 the curves are given with opposite polarities since they are taken directly from the operational amplifier).

An important concomitant feature is that the potentiometer scales cannot be calibrated in cps or decibels if the voltages are to be integrated, since the voltage level at a given time will be entirely dependent upon the past history of the signal.

The input-output relationship required according to our approach is as exemplified by "e_a" and "e_d" of Fig. 5 (for clarity the latter has been projected on the former and shown by a dashed line), except that here the transient time "τ_i" is assumed constant in order to simplify the exposition. The points "t_1", "t_2", etc. indicate the limits of each segment as specified in the input curve. The curve has (arbitrarily) been drawn in such a way that one of the segments (between "t_1" and "t_2") is of duration τ_i (whereas the other segments last longer). In this way the rising transition has been given a certain curvature.

The conversion of e_a into e_d is not a simple matter electronically even if the transient time is kept constant. However, it is possible to modify the input voltage function in such a way that the required output is obtained by integration. The principle underlying our experimental circuit can be stated by reference to the "Processed Input" curve "e_c" of Fig. 5.



t_n = any point of discontinuity of input step function

τ_i = preset transient time of integrating circuit

$$e_a = E_n \qquad t_{n-1} < t < t_n$$

$$e_b \begin{cases} = E_n - E_{n-1} & t_{n-1} < t < (t_{n-1} + \tau_i) \\ = 0 & (t_{n-1} + \tau_i) < t < t_n \end{cases}$$

$$e_c \begin{cases} = E_n - \frac{1}{2} E_{n-1} & t_{n-1} < t < (t_{n-1} + \tau_i) \\ = \frac{1}{2} E_n & (t_{n-1} + \tau_i) < t < t_n \end{cases} \quad \therefore \quad \boxed{e_c = \tfrac{1}{2}(e_a + e_b)}$$

$$e_d \begin{cases} = E_{n-1} + (E_n - E_{n-1}) \cdot \dfrac{t - t_{n-1}}{\tau_i} & t_{n-1} < t < (t_{n-1} + \tau_i) \\ = E_n & (t_{n-1} + \tau_i) < t < t_n \end{cases} \quad \therefore \quad \boxed{e_d = \int e_b \, dt = 2 \cdot \int (e_c - \tfrac{1}{2} e_a) \, dt}$$

Fig. 5.

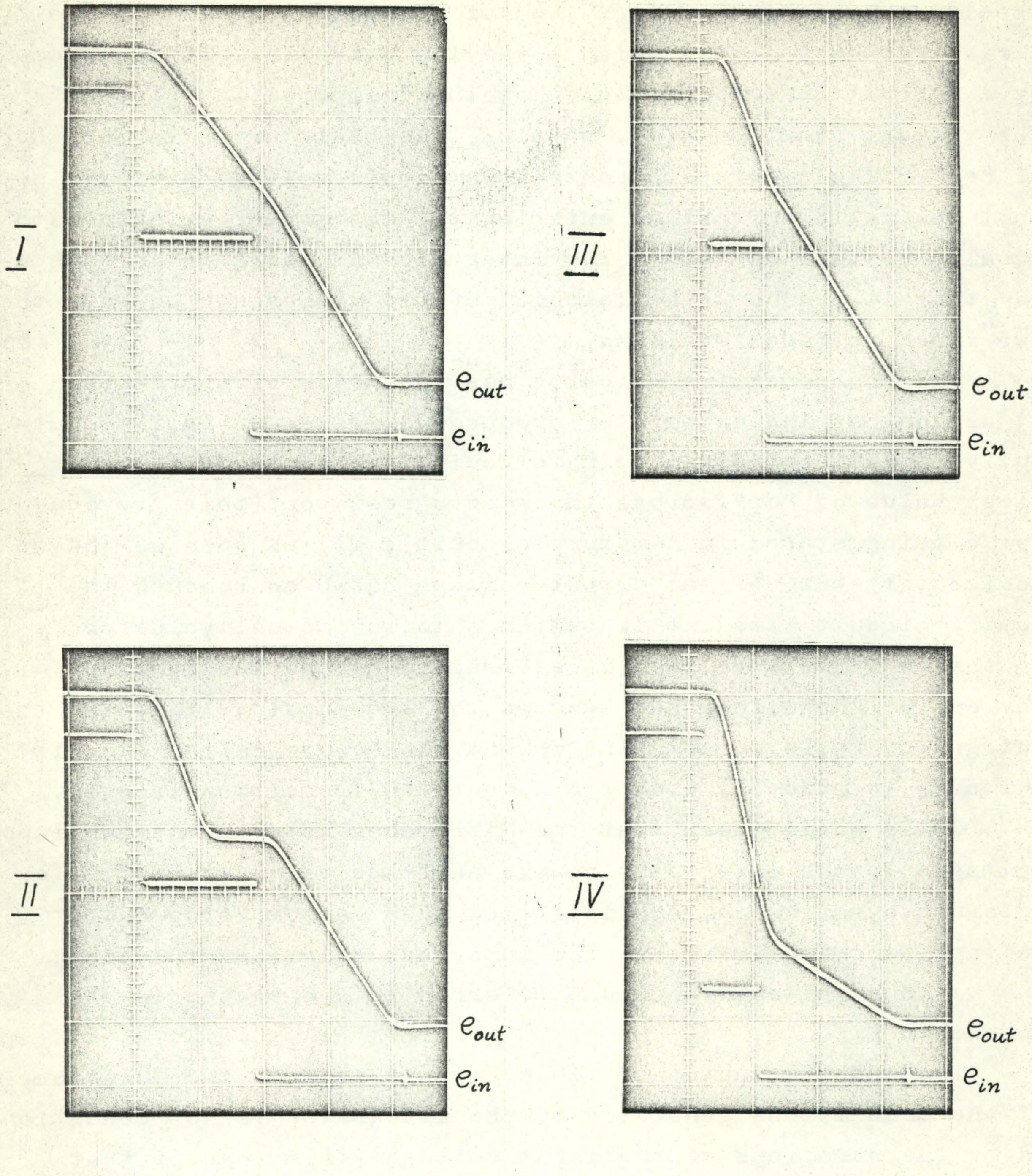
Logic of smoothing circuit.

The input voltage "e_a" is led by two different paths to a special type of integrator. One of the paths contains a small series resistor across which a varying potential drop is caused by a circuit referred to below as the "compensating circuit". The voltage function after passing this resistor is in the shape of "e_c". The other path contains a simple voltage divider, so that the voltage function entering the integrator by this path equals one half of "e_a". The integrator is designed in such a way that it produces the integral of the <u>difference</u> between the two input voltage functions.

The "compensating circuit" has two modes of operation: I) as long as the integrator produces a rising or falling output voltage the "compensating circuit" subtracts half the <u>previous</u> value of "e_a" across the series resistor (this previous value being stored in a memory circuit); II) as soon as the horizontal state of the output voltage has been reached the memory circuit immediately switches to the actual value of "e_a" so that half this voltage is subtracted across the resistor. Obviously, the difference between the two input voltages to the integrator falls to zero the moment the "compensating circuit" switches to mode II, i.e. the moment the transition is over.

It is easily seen that the difference between the two input voltages to the integrator equals half the value of "e_b", i.e. after integration and amplification by a factor 2 the output voltage function is "e_d". The numerical relationships among the various voltage functions referred to are stated in the legend to Fig. 5.

The special feature of this circuit is that the duration of the transitions generated by the integrator is not conditioned by the durations of the input voltage steps (except that these should not be shorter than τ_i). In the integrator a constant comparison between input and output voltages takes place, and the moment the output voltage equals half the instantaneous value of "e_a" the integrator is stopped and the output voltage remains constant until "e_a" assumes a new value. Thus the output voltage is clamped to the input voltage after every transition, which is an important safeguard against long-time DC drift.



I: τ_i of the first step equals the duration of the step - the slope produced is continuous though inflected.
II: τ_i of the first step has been shortened (everything else being equal) - two separate slopes are produced.
III: the duration of the first step has been reduced to $\tau_d = \tau_i$ - the two slopes become contiguous again.
IV: the voltage at the point of inflection has been lowered - the shape of the slope is changed.

Fig. 6.

Oscilloscope display of parameter synthesis with 3 variables.

(The constant offset between input and output is caused by a fixed filter inserted in order to round off the edges.)

Electronically, the basic unit of the integrator is a high gain transistor whose emitter is connected via a suitable resistor to the input terminal where "e_c" occurs, whereas the base is connected directly to the mid point of the voltage divider across which "e_a" occurs. The output is taken across a capacitor connected between collector and ground. Unfortunately this integrator can function only with one polarity of the input step function (with steps of the opposite polarity the output just follows the input though with some distortion). In order to obtain the same behaviour toward voltage steps of both polarities (cp. Fig. 3 above) it is necessary to have two complementary integrators and to combine their outputs in a mixing circuit which at any instant selects the most slowly changing voltage. As shown by Fig. 3 satisfactory results can be obtained in this way but certainly at a price.

The transient time is determined by the values of R and C in the integrator. By means of electronic switches it is possible to change the value of R from one event to the next. In the case of vowel formants a moderate number of different values of τ_i over the range 10 - 100 ms will certainly suffice.

Fig. 6 (preceding page) displays the three "dimensions" in which a control parameter can be varied according to the suggestions given above. It remains to be tested whether this three-dimensionality is a sufficiently attractive feature from the point of view of the experimenter to warrant its existence.

Acknowledgements:

The project is supported by grants from the Danish Research Committee for the Technical Sciences (Danmarks teknisk-videnskabelige Forskningsråd).

I am indebted to the staff of the Speech Transmission Laboratory of the Royal Technical High-School, Stockholm, for a stimulating discussion of some of the problems treated in this paper.

References:

(1) John M. Borst, "The Use of Spectrograms for Speech Analysis and Synthesis", Journal of the Audio Engineering Society 4 (1956), pp. 14-23, see pp. 19 ff.

(2) J. L. DeClerk, D. L. Phyfe, and R. A. Fisch, "Formant Synthesizer Electronically Controlled", Proceed. of the Speech Comm. Seminar, Stockholm 1962 (1963), Paper F4.

(3) J. B. Dennis, "Computer Control of an Analog Vocal Tract", ibid., Paper F3.

(3a) G. Fant and J. Mártony, "Quantization of Formant-Coded Synthetic Speech", STL-QPSR 2/1961 (Stockholm), pp. 16-18.

(4) Johan C. W. A. Liljencrants, "The OVE III Speech Synthesizer", IEEE Trans. on Audio and Electroacoustics AU-16 (1968), pp. 137-140.

(5) Jørgen Rischel, "Instrumentation for Vowel Synthesis", ARIPUC 1/1966 (1967), pp. 15-21.

(6) Jørgen Rischel and Svend-Erik Lystlund, "Speech Synthesizer", ARIPUC 2/1967 (1968), p. 34.

(7) W. Tscheschner, "Ueber ein Steuersystem zur Sprachsynthese", 5e congrès international d'acoustique (1965), rapport A53.

THE PHONEME SYSTEM OF ADVANCED STANDARD COPENHAGEN

Hans Basbøll

1. Introduction.

The purpose of this paper is to give a phonetic and phonemic description of Advanced Standard Copenhagen (ASC). For this description I shall use a taxonomic procedure with two operationally defined levels between the phonetic transcription and the ultimate phonemes, viz. the levels of prephonemes and of phonemes (see section 2.2.).

1.1. What is ASC?

ASC (with a perhaps not too lucky term) is the language spoken by a large group of the younger generation in Copenhagen, whose language is normally considered to be a variety of Standard Danish. In many respects (see section 1.2.) it differs from what might to-day be called Conservative Standard Danish (CSD), the language described by Jespersen (9), Uldall (15), Martinet (11), and Hjelmslev (8); the language described by Rischel (this report, pp. 177ff.) is intermediate between CSD and ASC (see also Diderichsen (3), Hansen (5), and Andersen (1), where many phonetic observations on ASC and other varieties of Standard Danish are mentioned).

ASC is clearly different from both the Copenhagen dialect (sometimes termed "vulgar") and the language (sometimes termed "affected") spoken by the upper class in the northern parts of Copenhagen.

The analysis is based mainly upon my own speech. Thus it is hardly quite valid for ASC in general since probably every individual has his own phoneme system (sounds which in the language taken as a whole are free variants may in the speech of some individuals have such a distribution that they must be considered as belonging to different phonemes; the reason probably is that we learn some words from our parents, others from our playmates - who may use different phoneme-manifestations - etc.).

1.2. Some phonetic differences between ASC and CSD.

CSD and ASC can in general be considered as an older and a newer stage of the same language (Standard Danish). The effect of some of the phonetic changes from CSD to ASC are distributional limitations in ASC, and I shall therefore refer to the rules of section 4.1.1. in the following survey of the main phonetic differences between CSD and ASC.

(1) Whereas CSD has forms like nøden, nødden [ˈnø·ˀðn̩ ˈnøðˀn̩] 'the distress', 'the nut', stød-vocoids never appear before ð in ASC, i.e. both forms are pronounced [ˈnøðˀn̩] in ASC (cp. rule (6)).

(2) The voiced velar or postpalatal fricative [ɣ] (which according to Martinet (11) manifests a separate phoneme /q/) has merged with other phonemes (and in some cases, particularly after narrow vocoids*), with zero). Its velar variant, which in CSD occurs after back vocoids and [ʁ], has merged in ASC with the postvocalic manifestation of /v/ (in CSD the words låg [ˈlɔ·ˀɣ] 'cover' and lov! [ˈlɔ·ˀv ˈlɔ·ˀʊ̯ **)] 'promise' (imperative) are kept apart, whereas in ASC both are pronounced [ˈlɔʊ̯ˀ]). Its postpalatal variant, which in CSD occurs after front vocoids and [l], has merged in ASC with the manifestation of /j/ (in CSD the words balje, galge [ˈb̥aljə, ˈg̊alɣə] 'tub', 'gibbet' do not rhyme, but they do in ASC: [ˈb̥æljə, ˈg̊æljə]). It should be noted that when the phonetic diphthongs arising by the change of CSD [ɣ] to ASC [ɪ̯] or [ʊ̯] occur in stød-syllables, the stød is always on [ɪ̯] or [ʊ̯], whereas [ɣ] in CSD can never have stød.

----

*) In this paper the terms 'vowel-consonant' (used in a functional sense) and 'vocoid-contoid' (used in a phonetic sense) will be distinguished, as suggested by Pike (12).

**) In Rischel's broader transcription, phonetic [-ɪ̯, -ʊ̯] are rendered as [-i̯, -u̯].

The result is that these forms must be analyzed phonemically differently in CSD and ASC. E.g. according to Martinet (11) the opposition between such forms as CSD leg [ˈla-ɪ̯ʔ] 'game' and lag[ˈla·ʔɣ] 'layer' involves phonemically vowel length and the last phoneme, whereas the difference in vowel quality is not considered distinctive. In ASC, on the other hand, the words are pronounced [ˈlɑɪ̯ʔ] and [ˈlæɪ̯ʔ], i.e. the only phonetic difference is in the vowel quality which must then be phonemic.

(3) The 'colouring' (i.e. opening and retracting) effect on the vocoid of preceding [ʁ] is much stronger in ASC than in CSD; the most important effect of this opening is that CSD /e:/ and /ø:/ merge with CSD /ɛ:/ and /œ:/ after [ʁ] in ASC (e.g. CSD tre, træ [ˈtʰʁ̥e·ʔ, ˈtʰʁ̥ɛ·ʔ] 'three', 'wood' are both [d̥ˢhʁ̥ɛ·ʔ] in ASC, and CSD røbe, brøle [ˈʁø·b̥ə b̥ʁœ:lə] 'reveal', 'roar' have the same vocoid in ASC: [ˈʁœ:b̥ə ˈb̥ʁœ:lə]) (cp. rule (3a)).

(4) The contoid [ʁ] never occurs post-vocalically in ASC (cp. Koefoed (10) and Andersen (2)). Corresponding to CSD-forms with short or long /a/ or /ɔ/ plus /r/ (e.g. CSD var, bort [ˈvɑʁ̥, ˈb̥ɒʁ̥t] 'was', 'away'), ASC-forms always have pure vocoids [ɑ(:), ɒ(:)] (e.g. ASC var, bort [ˈvɑ, ˈb̥ɒ:d̥]). The ASC-vowels ɑ and ɒ resulting from this change are short in word final position (except in the word far [ˈfɑ:] 'father') and long in all other positions, i.e. they do not participate in the quantity-opposition (but cp. paragraph (5) below); CSD-forms like båre, borre [ˈb̥ɒ:ʁə, ˈb̥ɒʁə] 'stretcher', 'cockchafer' are both [ˈb̥ɒ:ɒ] in ASC.

Corresponding to combinations of other vocoids +[ʁ] in CSD, ASC has phonetic diphthongs ending in [-ɒ̯] (see section 4.4.; cp. rule (2a)). When such forms are affected by stød, it normally is on the vocoid in CSD (e.g. ser [ˈse·ʔʁ̥] 'sees'), but in ASC it is always on [ɒ̯] (ser [ˈseɒ̯ʔ]) (cp. rule (2b)). In unstressed syllables, CSD distinguishes between [əʁ ʁə ʁəʁ] (CSD gnister, gnistre, gnistrer [ˈg̊nisd̥əʁ ˈg̊nisd̥ʁə ˈg̊nisd̥ʁəʁ] 'sparks', 'sparkle', sparkles'), whereas ASC maximally distinguishes [ɒ ʁɒ] (ASC gnister, gnistre=gnistrer [ˈg̊nisd̥ɒ ˈg̊nisd̥ʁɒ]);

after vowels, ASC has only [ɒ] (CSD tuer,ture,turer [ˈthu:əʁ, ˈthu:ʁə, ˈthu:ʁəʁ] 'tufts', 'trips', 'gads about' are all pronounced [ˈd̥shu:ɒ] in ASC).

(5) Whereas the short /a/ in CSD has three bound variants ([ɑ] before or after [ʁ] (e.g. var [ˈvɑʁ̥]), [a-] otherwise before velars (e.g. lang [ˈla-ŋˀ] 'long'), [a] before dentals, labials and zero (e.g. land,lam,da [ˈlanˀ, ˈlamˀ, ˈd̥a] 'land', 'lamb', 'when')), the vocoids in e.g. lam,lang and var are identical in ASC ([ˈlɑmˀ, ˈlɑŋˀ, ˈvɑ]); this change has the effect that [ɑ] (but not [ɒ]) in ASC participates in the quantity opposition (e.g. ASC lamme,larme [ˈlɑmə , ˈlɑ:mə] 'paralyse', 'make noise') (cp. paragraph (4) above and rules (4a) and (4b)).

(6) In CSD it is generally assumed that there is an allophonic variation between short [i] and [e], [y] and [ø] , [u] and [o] before /r/ (cp. Rischel in this report, p. 181); no such overlapping is found in ASC.

Because of these and other phonetic differences between ASC and CSD, the taxonomic analyses of Standard Danish do not apply to ASC.

2. General approach.

The present analysis has been highly influenced by the contributions to phoneme theory of Twaddell (14), Harris (6), and many others, but is not identical in approach to any of these.

It must be emphasized, however, that I do not consider this analysis (which most linguists will probably find out-of-day) superior in any way to generative phonology. On the contrary, an ordered set of rules will undoubtedly in many cases account for the linguistic data in a simpler way than my analysis does. Above all, this would be true for an over-all description of ASC.

I use a taxonomic procedure mainly in order to facilitate comparisons between the present analysis and the existing descriptions of other varieties of Danish which are all taxonomic.

Furthermore, I hope that this analysis with the operationally defined levels of prephonemes and phonemes will make the relation between phonetic and phonemic entities clearer than has sometimes been the case in Danish phonemics.

In this analysis, stress (section 3.) is handled before the establishment of the prephonemes (section 4.), whereas length and stød are handled afterwards (section 5.1.), although both stress, length and stød will turn out to be accents. The reason for this procedure is that phonetically vowel length and stød are clearly properties of one single sound and therefore cannot - within the framework of the present analysis - be extracted from the sound chain before the segmental entities of which they are properties (i.e. the prephonemes) have been set up; stress, on the other hand, is phonetically a property of the syllable (involving problems of syllable boundary, etc.) and may therefore be handled apart from length and stød.

2.1. Analytical criteria.

I use Harris' (6) distinction between phonemes and morphophonemes; i.e. it is presupposed that there should be a one-one-correspondence between a sound chain and a string of phonemes, but a one-many-correspondence between a sound chain and a string of morphophonemes (disregarding free variants in both cases). E.g. the two words *tre* 'three' and *træ* 'wood' are both pronounced [ˈd̥ˢhʁ̥ɛ·ˀ] and are thus represented by the same chain of phonemes (cp. section 1.2., paragraph (3)), but there is commutation between compounds like *trefod* [ˈd̥ˢhʁ̥ɛˌfoðˀ] 'tripod' and *træfod* [ˈd̥ˢhʁ̥æˌfoðˀ] 'wooden foot'; i.e. the simplex words *tre* and *træ*, although phonemically identical, are morphophonemically different.

In accordance with the taxonomic procedure used in this paper no morphological or syntactic criteria will be used in the analysis, and no morphophonemic description of ASC will be given. In a generative procedure, on the other hand, morphological criteria are basic prerequisites to phonological description.

Morphophonemically, the variety of Standard Danish described by Rischel (this report pp.177ff.) and ASC are very similar, and the reader is therefore referred to Rischel's paper for morphological information also applying to ASC.

'Partial overlapping' is allowed for in the present paper, i.e. the same sound can manifest different phonemes in different phonemic environments (e.g. the sound [œ] can manifest either ø or Œ, see section 4.1.).

The material is limited to isolated words, and the phoneme manifestations (or other linguistic features) in connected speech will not be discussed. Foreign words are excluded from the material.

There is more than one kind of linguistic simplicity (cp. Spang-Hanssen (13)). As already noted, I shall not take morphological simplicity into account, but still both paradigmatic and syntagmatic (or phonotactic) aspects of simplicity remain. Reductions which seem fully justified from a paradigmatic point of view may sometimes complicate the phonotactic description substantially. No account of the phonotactics of ASC will be given here, and I shall normally use the word 'simplicity' for 'paradigmatic simplicity' - but one must be aware of the problem.

2.2. Levels of analysis.

Our starting point is the phonetic transcription; it is evident, however, that such a transcription presupposes some linguistic analysis, including a segmentation of the continuous flow of speech. But the problems involved in the phonetic transcription will not be discussed here.

The phonetic transcription is, of course, a narrow one; e.g. the initial aspirated plosives are transcribed [b̥h d̥ˢh g̊h] which is phonetically more correct than the normal transcriptions $[p^h\ \ t^h\ \ k^h]$. The consequence of this transcription is that e.g. [b̥h] must be considered the manifestation of a cluster of prephonemes /bh/ (see section 4.3.) and not of a single prephoneme /p/ (e.g. in the word pande ['b̥hænə] 'pan', [b̥] is commutable with e.g. [g̊] (kande ['g̊hænə] 'can'), and [h] with e.g. [l] (blande ['b̥lænə] 'mix')).

This interpretation is consequently different from Hjelmslev's (8), who reduces the prephoneme ? p to either hb or bh, etc. The well-known reduction of [ð] to a bound variant of /d/ (proposed by Uldall (15)) is for the same reason not possible in the present analysis.

The prephonemes are commutable units established by a classificatory procedure starting from minimal sound segments. Minimal sound segments are the smallest parts into which the sound chain can be divided under the condition that each sound segment be commutable with at least one other sound segment or with zero. The prephonemes are obtained by identifying the commutable minimal sound segments in different environments by means of both phonetic and systematic criteria. Sounds in complementary distribution are identified according to phonetic similarity (and pattern congruity), and sounds which occur in the same (or partially the same) environments but are never commutable are also identified. In the present phonetic transcription the number of sound symbols is not much greater than the number of prephonemes, and sounds that always occur as free variants are generally not distinguished; the practical procedure for establishing the prephonemes is therefore mainly an identification of sound symbols in complementary distribution according to the phonetic similarity(and pattern congruity)of the sounds in question.

It should be noted that the existence of a minimal pair is here considered a sufficient, but not necessary, condition for establishing a phonemic opposition (e.g. the sounds [ø:] and [œ:] belong to different phonemes although no minimal pair can be found, see section 4.1.). Consequently, if I say that two sounds are 'commutable', this does not necessarily imply the existence of a minimal pair.

A prevowel is a prephoneme which can form a word by itself, or which is commutable with such prephonemes (section 4.1.).

A preconsonant is a prephoneme which can never form a syllable by itself (section 4.3.).

A weak precentral is a prephoneme which can form a syllable, but not a word, by itself, or which is commutable with such prephonemes (section 4.2.).

Syllabic and non-syllabic segmental entities (sounds, phonemes, etc.) are kept distinct throughout the analysis (e.g. ɒ : ɒ̯ ; l̩ : l).

The phonemes are obtained by extracting the 'suprasegmental phonemes' - i.e. the accents - from the chain of prephonemes, and then reducing the remainder by means of the mentioned criteria of phonetic similarity, complementary distribution etc. (section 5.).

The ultimate phonemes (section 6.) are obtained by interpreting one phoneme as a group of successive ultimate phonemes. The criteria for this type of reduction are discussed in section 6.1.

3. Stress.

There are two phonemically distinct degrees of stress in ASC: strong and weak stress. Every word has one syllable with phonetically primary stress, i.e. phonemically with stress accent.

Phonetically secondary stress, e.g. in forgård [ˈfɒːˌg̊ɒ·ˀ] 'forecourt', is only found in words which also have primary stress, and never before the syllable with primary stress, i.e. phonetically secondary stress is in bound variation with phonetically primary stress and can therefore be considered as phonemically strong stress, i.e. as stress accent.
When a word has more than two stressed syllables (in the phonemic sense)(i.e. phonetically more than one secondary stress), there is no phonemic difference in stress between these latter syllables (e.g. in slotsforgård [ˈslʌd̥sˌfɒːˌg̊ɒ·ˀ] 'forecourt of a castle').

The relevance of strong and weak stress can be shown by commutation pairs like forgård [ˈfɒːˌg̊ɒ·ˀ], forgår [fɒˈg̊ɒ·ˀ] 'perishes', phonemically /ˈfɔrˀgɔr, fɔrˀgɔr/ *), the only phonemic difference being one of strong versus weak stress on the first syllable, i.e. one of presence versus absence of stress accent on that syllable.

(In foreign words we have oppositions like plastic, plastik [ˈb̥hl̥æsd̥ig̊, b̥hl̥æˈsd̥ig̊].)

Word and syllable are here taken as axiomatic entities.

In this paper stressed syllables are said to have stress accent, whereas the unstressed syllables have no stress accent.

----

*) The length of the vowel ɒ is never phonemic in ASC.
(The phonemic accent symbols ˈˀ and ː, placed before the syllable in question, are explained in section 5.1.3.)

4. From phonetic transcription to prephonemes.

4.1. Prevowels.

Five stød-vocoids appear as isolated words with different meanings in ASC, and they thus manifest different prevowels: [ˈi·ʔ, ˈɑ·ʔ, ˈø·ʔ, ˈɔ·ʔ, ˈɒ·ʔ] i 'in (adverb)', ar 'cicatrice', ø 'island', å 'small river', år 'year'. A number of other stød-vocoids also occur in isolation but only as names of letters: [ˈe·ʔ, ˈɛ·ʔ, ˈæ·ʔ, ˈy·ʔ, ˈu·ʔ, ˈo·ʔ]; since they are commutable with the above-mentioned vocoids in other positions, they, too, manifest prevowels.

Two other stød-vocoids appear in ASC: [œ·ʔ] is only found after ʁ- (frø 'seed', brøl 'roar', etc.) (and in foreign words before -n (obskøn 'obscene')); [ø·ʔ] is excluded in these positions, and I therefore regard [œ·ʔ] as a bound variant of the prevowel ø·ʔ; [ɶ·ʔ] is found only before unstressed [ɒ], e.g. in the verb udgøre 'constitute', and there is commutation between [ø·ʔ] and [ɶ·ʔ] (but only in quasi-minimal pairs like udkøre [ˈuð˛g̊hø·ʔɒ] 'cart out', udgøre [ˈuð˛g̊œ·ʔɒ]). I regard [ɶ·ʔ] as manifesting the prevowel ɶ·ʔ (the roots in which [ɶː] - and in some types of compounds [ɶ·ʔ] - is found, are listed below in the present section).

The system of stød-prevowels in ASC is then:

i·ʔ	y·ʔ	u·ʔ
e·ʔ	ø·ʔ	o·ʔ
ɛ·ʔ	ɶ·ʔ	ɔ·ʔ
æ·ʔ	ɑ·ʔ	ɒ·ʔ

Long vocoids cannot form words by themselves in ASC; but in other positions they are commutable with stød-vocoids, and they are therefore manifestations of prevowels (e.g. [ˈsg̊uːɒð, ˈsg̊u·ʔɒð] skuret 'scoured (past ptc.)' versus 'the shed'). The following long vocoids can immediatly be shown to manifest different prevowels: [iː, eː, ɛː, æː, ɑː, yː, øː, uː, oː, ɔː, ɒː]. Two other long vocoids are found in ASC: [œː] only appears after ʁ- (røbe 'reveal') and before -n (høne 'hen'); [øː] does not appear in these positions, and [œː] can therefore be regarded as a bound variant of the prevowel øː. [ɶː] is only found before unstressed [ɒ]; we have no minimal pairs for the opposition [ø:]:[ɶ:] (but we have for [y:]:[ø:] and [y:]:[ɶ:]), but the material clearly shows that the gap is accidental:

(1) [-y:ɒ] <u>dyre</u>,<u>fyre</u>,<u>hyre</u>,<u>lyre</u>,<u>myre</u>,<u>nyre</u>,<u>styre</u>,<u>syre</u>,<u>tyre</u>.
(2) [-ø:ɒ] <u>føre</u>,<u>høre</u>,<u>køre</u>,<u>møre</u>,<u>pløre</u>,<u>skøre</u>,<u>sløre</u>,<u>øre</u>.
(3) [-ɶ:ɒ] <u>døre</u>,<u>gøre</u>,<u>røre</u>,<u>smøre</u>,<u>snøre</u>.

The system of long prevowels is then:

i:	y:	u:
e:	ø:	o:
ɛ:	ɶ:	ɔ:
æ:	ɑ:	ɒ:

(Both the system and the mentioned manifestation rules correspond exactly to those of the stød-prevowels.)

A few short vocoids can form words by themselves in ASC: [ˈi ˈæ (ˈʌ)] <u>I</u> 'you' (plur.), <u>à</u> , (<u>og</u> 'and') (the last example is dubious in stressed position); these are, therefore, manifestations of prevowels. Other short vocoids are commutable with these and with each other: [e ɛ y ø u ɔ ɒ] and thus manifest prevowels, too.

Two other short vocoids are found in ASC: [œ] and [ɶ] . Except for the position before [-ɒ̭ˀ], never more than two of the four short rounded front vocoids occur (in contrast), as shown in the following table

	-N	-ɒ̭	-X
ʁ-	[œ]:[ɶ]		[y]:[œ]
X-	[ø]:[œ]	[y]:[ɶ]	[y]:[ø]

where 'N' means 'nasals' (before -ŋ only [ø] occurs) and 'X' means 'other phonemes'.

I.e. if the position before [-ɒ̭ˀ] could be disregarded, the number of prevowels in the short round front-series could be reduced to 2, with the following manifestation rules:
y is manifested by [œ] in the context ʁ-N (e.g. <u>grynt</u>[ˈg̊ʁœnˀd̥] 'grunt'), by [ø] before -N when no ʁ precedes (e.g. <u>skynde</u> (<u>sig</u>) [ˈsg̊ønə] 'hurry'), by [y] elsewhere (e.g. <u>syd</u> [ˈsyð] 'south');
ø is manifested by [ɶ] in the context ʁ-N (e.g. <u>grøn</u> [ˈg̊ʁɶnˀ] 'green') and before -ɒ̭ (without stød) (e.g. <u>smør</u> [ˈsmɶɒ̭]'butter'),

by [œ] after ʁ- when no N follows (e.g. ryste [ˈʁœsd̥ə] 'shake') and before -N when no ʁ precedes (e.g. skøn [ˈsg̊œnˀ] 'beautiful'), and by [ø] elsewhere (e.g. kyst [ˈg̊høsd̥] 'shore').
(This reduction, with even simpler manifestation rules, would be possible for the variety of Standard Danish described by Martinet (11).)

But in ASC we have commutations like dyr, dør, dør [ˈd̥yɒ̯ˀ, ˈd̥øɒ̯ˀ, ˈd̥œɒ̯ˀ] 'animal', 'dies', 'door', and we thus have 3 prevowels in this series: y, ø, Œ; i.e. the above-mentioned reduction is impossible for ASC.

We cannot, however, completely avoid partial overlapping in this series:

y is manifested as [y] (e.g. syd [ˈsyð]).

ø is manifested as [œ] after ʁ- (e.g. ryste [ˈʁœsd̥ə]), as [ø] elsewhere (e.g. kyst [ˈg̊høsd̥]).

Œ is manifested as [œ] before -N when no ʁ precedes (e.g. skøn [ˈsg̊œnˀ]), as Œ elsewhere (e.g. smør, grøn [ˈsmœɒ̯, g̊ʁœnˀ]) (cp. section 4.1.1., rule (5)).

Before ð, the following short rounded back vocoids are found: [u o ɔ ʌ] (gud, fodre, båd, od [ˈg̊uð, ˈfoðʁɒ, ˈb̥ɔðˀ, ˈʌð] 'god', 'feed', 'boat', 'point'). [ɒ] is excluded in this position, and I therefore regard [ʌ, ɔ, o] as manifesting respectively ɒ, ʌ, and ɔ in the position before ð (i.e. the manifestations of short prevowels before ð have the same quality as the corresponding long prevowels).

The system of short prevowels is then

i	y	u
e	ø	ɔ
ɛ	Œ	ʌ
æ	ɑ	ɒ

Each of the 36 prevowels that have hitherto been set up is commutable with all the others, with one important exception: there is no commutation between [ɒ] and [ɒː]; these are bound variants ([ɒ] appearing only in word final position (e.g. vor [ˈvɒ] 'our') where [ɒː] is excluded). Consequently there are only 35 prevowels. [ɒː] occurs in no position where short prevowels are excluded, but [ɒ] occurs finally in monosyllabic words where generally no long prevowels appear (except in one word far [fɑː] 'father'); it therefore seems preferable to exclude ɒː from the long prevowel system and to retain ɒ in the short one, rather than vice versa.

4.1.1. Distributional limitations of the prevowels.

There follows a list of the main limitations in the distribution of the 35 prevowels (in stressed position) (in these rules 'long' means 'long or stød-').

I do not pretend that none of the combinations excluded by the following rules ever occur in actual speech, but only that if one of them does, it is an accidental variation of some 'permitted' combination. Some of the rules are not so absolute as they might seem to be : e.g. rule (1d) might cover i as well, if it were not for the word linje [ˈlinjə] 'line'; and the pronunciation brynje [ˈb̥ʁynjə] 'coat of mail' can in fact be heard as a variant of the normal one [ˈb̥ʁœnjə].

(1a) Long prevowels are not found before [ŋ].
(1b) ɒ is not found before [ŋ].
(1c) i,y,u are not found before [ŋ] or [m].
(1d) y is not found before [n].
(1e) u is not found before [n]+contoid.
(2a) ɑ,ɒ are not found before [ɐ̯].
(2b) Long prevowels are not found before [ɐ̯].
(3a) e,eː are not found after [ʁ].
(3b) æː is not found after [ʁ] except before [ð] (e.g. græde [ˈɡ̊ʁæːðə] 'cry').
(4a) æ is not found before labials and velars except after [ʁ] (e.g. rem [ˈʁæmˀ] 'strap').
(4b) ɑ is not found before dentals except after [ʁ] (e.g. rat [ˈʁɑd̥] 'steering wheel').
(5) Œ is only found before [-n,-m,-ɐ̯ˀ] (e.g. grøn 'green', drøm 'dream', dør 'door' [ˈɡ̊ʁœnˀ, ˈd̥ʁœmˀ, ˈd̥œɐ̯ˀ]).
(6) Stød-prevowels are not found before non-syllabic [ð].
(7a) Short prevowels are not found before [ɣ̞ə].
(7b) Short prevowels are not found immediately before an unstressed vocoid or a syllabic contoid.

It should be added that there is never commutation between ɒ·ˀc and ɒcˀ (where c is a contoid) (e.g. [ˈsd̥ɒ·ˀm] and

[ˈsd̥ɒm?] are free variants for storm 'gale').

(Some of these and other distributional limitations may later (in section 6.) be used for a reduction of the inventory.)

4.1.2. Prevowels in pretonic syllables.

In weak syllables before the first stressed syllable in a word no stød-vocoids and no long vocoids *) occur. The syllabics in these syllables are short vocoids, but the rounded front vocoids only occur in foreign words (like dynamit [d̥ynæˈmid̥] 'dynamite'). These unstressed short vocoids are identified with the corresponding stressed ones of the same quality, i.e. manifestations of short prevowels.

The syllabic segments of unstressed syllables occurring between two stressed ones in the same word can always be recognized as manifestations of either short prevowels or weak precentrals; in cases where both identifications are possible, it does not matter which identification we choose because in the final analysis the short prevowel and the weak precentral in question will be identified as the same vowel, see section 5.2.

4.2. Weak precentrals.

In unstressed syllables after the last stressed syllable in the word the following syllabic vocoids occur in native words: [i, e, ə, ɒ, u] (e.g. in hyppig 'frequent', madding 'bait', hyppe 'hoe', hypper 'hoes', vindue 'window' [ˈhyb̥i, ˈmæðeŋ, ˈhyb̥ə, ˈhyb̥ɒ, ˈvend̥u]). (In foreign words also [æ, o, ɔ] occur, e.g. in kvota, cello, centrum [ˈg̊hvo:d̥ˢhæ, ˈʃɛlo, ˈsɛnd̥ˢhʁ̥ɔm].)
Apart from [e] these vocoids are commutable (except [o] and [ɔ] which are bound variants, [o] occurring only in open syllables and [ɔ] only in closed ones). As [e] occurs only before [ŋ] where none of these other vocoids is found, the possibility of identifying [e] with some other vocoid should be considered; but as it will be shown that [e] occupies a place in the 'weak precentral-hierarchy' different from those of all the other vocoids (see below), this identification will be avoided here.

Furthermore, five syllabic contoids occur in ASC: [l̩ m̩ n̩ ŋ̍ ð̩] (e.g. in mandel 'almond', lampen 'the lamp', manden 'the man', lakken 'the lacquer', huset 'the house'

----

*) Except in a few foreign words of the type maleri [mæ:lɒˈʁi·?] 'painting'.

[ˈmænʔl̩, ˈlɑmb̥m̩, ˈmænʔn̩, ˈlɑg̊ŋ̍, ˈhu·ʔsð̩]). They never have stød.

[m̩] and [ŋ̍] are only found after [b̥] and [g̊], respectively, and are in this position in free variation with [n̩] . [m̩] and [ŋ̍] are therefore considered to be variants of [n̩].

[l̩, n̩, ð̩] are all commutable with each other and with the syllabic vocoids [i, ə, u, ɒ].

Now a hierarchy of the weak precentrals can be set up as follows:

Definition	Level
occurring in unstressed posttonic syllables in native words	4th level: u
never occurring after h; occurring after ð	3rd level: e
never occurring in syllables with obligatory stød	2nd level: i
never occurring in stød-syllables	1st level: ə ɒ l̩ n̩ ð̩

The 'definition' of a level applies to all lower levels as well. Each level is fully specified by the definition of its own level and the negation of the definition of the level just below (the negation is formed by deleting the word 'never' if it occurs, and by inserting the word 'never' if it does not occur); e.g. the third level is defined as comprising the weak precentrals which 'never occur after h but occur after ð', and which 'occur in syllables with obligatory stød', and this can only be e (e.g. in maddingen [ˈmæðeŋʔn̩] 'the bait').

4.3. Preconsonants.

Initially, the following 13 commutable contoids occur:

[b̥, d̥, g̊, h, m, n, l, ʁ, j, f, v, s, ʃ].

[d̥ˢ] is only found before [h] where [d̥] is not found; [d̥ˢ] and [d̥] are therefore bound variants.

The plosives may be voiced, especially in intervocalic position, and the voiced nasals and fricatives are partly or wholly devoiced after [h,f,s]; in neither of these cases is the degree of voicing phonologically relevant.

Finally, the following ten stød-less commutable contoids are found:[b̥,d̥,g̊,m,n,ŋ,l,ð,f,s] (final [ʃ] is only found in foreign words). In utterance-final position these contoids are often aspirated and devoiced. In intervocalic position in simple, native words we find the same inventory as in final position. (Intervocalically [ɾ] occurs as a stilistically conditioned variant of d, i.e. in rapid colloquial speech.)

Furthermore, the following five stød-contoids occur in ASC: [mˀ nˀ ŋˀ lˀ ðˀ] . They are commutable with each other and with all the stød-less contoids.

By identifying phonetically identical final and initial contoids we obtain the following eight preconsonants occurring both initially and finally: b d g m n l f s . The following five contoids occur only initially: h ʁ j v ʃ (but if we include the foreign vocabulary, ʃ occurs both initially and finally). The following seven contoids occur only finally: ŋ ð mˀ nˀ ŋˀ lˀ ðˀ; there is no possibility of identifying any of these with h- ʁ- j- v- or ʃ-.

4.4. The phonetic diphthongs.

A great number of phonetic diphthongs occur in ASC, all ending in [-ɪ̯, -ʊ̯, -ɒ̯] or [-ɪ̯ˀ, -ʊ̯ˀ, -ɒ̯ˀ]. In one-syllable words their first component is always a short stød-less vocoid, in other word types also long vocoids and stød-vocoids occur in this position. The vocoids occurring as the first component of the diphthongs can be identified with the corresponding monophthongal manifestations of prevowels.

It might seem natural to consider the last component as a manifestation of a short prevowel and later to reduce the long prevowels to combinations of two identical short prevowels. But the parallelism between long prevowels and diphthongs implied by this analysis is contradicted by the fact that no long prevowels occur finally in one-syllable words (except in the word <u>far</u> [fɑ:]) whereas stød-less diphthongs do occur here (e.g. in <u>tøj</u>

'clothes', hav 'sea', smør 'butter' [ˈd̥ˢhœɪ̯, ˈhɑʊ̯, ˈsmɶɐ̯]).
On the contrary, these stød-less forms correspond exactly to combinations of short prevowels + certain preconsonants *) as in man 'one', til 'to', gud 'god' [ˈmæn, ˈd̥ˢhel, ˈg̊uð].
We therefore consider the second components to be manifestations of preconsonants, i.e. we identify -ɪ̯ with j- (belonging to the preconsonant j), -ʊ̯ with v-(belonging to the preconsonant v), and -ɐ̯ with ʁ-(belonging to the preconsonant r). We must then recognize three further only-final preconsonants, i.e. - jˀ, -vˀ, -rˀ (manifested as [- ɪ̯ˀ - ʊ̯ˀ - ɐ̯ˀ]).

4.5. The inventory of prephonemes.

The inventory of prephonemes has now been found, viz.

35 prevowels:

12 stød-prevowels:

i·ˀ	y·ˀ	u·ˀ
e·ˀ	ø·ˀ	o·ˀ
ɛ·ˀ	œ·ˀ	ɔ·ˀ
æ·ˀ	ɑ·ˀ	ɒ·ˀ

11 long prevowels:

i:	y:	u:
e:	ø:	o:
ɛ:	œ:	ɔ:
æ:	ɑ:	

12 short prevowels:

i	y	u
e	ø	ɔ
ɛ	œ	ʌ
æ	ɑ	ɒ

----

*) I.e. preconsonants which after the reductions in section 5.1. turn out as consonants able to occur immediately after the vowel in syllables having the stød accent but not the length accent.

8 weak precentrals:

ə ɒ l̩ n̩ ð̩ i e u

23 preconsonants:

Initially and finally 11:

b d g m n l f s r j v

Only initially 2:

h ʃ

Only finally 10:

ŋ ð mʔ nʔ ŋʔ lʔ ðʔ rʔ jʔ vʔ

In total, 66 prephonemes.

5. From prephonemes to phonemes.

5.1. Accents.

5.1.1. The stød accent.

There is never more than one stød in a syllable in ASC, and it can only be at one place in the syllable, namely on the syllabic vocoid if this is half-long, otherwise on the following (non-syllabic) sound. The stød is therefore considered as an accent belonging to the syllable as a whole. This means that mʔ, nʔ, ŋʔ, lʔ, ðʔ, rʔ, jʔ, vʔ are bound variants of m, n, ŋ, l, ð, r, j, v occurring in stød-syllables. (These eight consonants are the only phonemes found immediately after short prevowels in syllables having the stød accent, i.e. (after the reduction made in section 5.1.2.) immediately after the vowel in syllables having the stød accent but not the length accent (cp. section 4.4.).)

Half-long vocoids are now reduced to bound variants of long vocoids, conditioned by the occurrence in stød-syllables (v· and v: being in complementary distribution); this analysis is confirmed by the similarity of the manifestation rules for stød-prevowels and for long prevowels. But we must then restate ɒ: as a long vowel, cp. the opposition between ɒ and ɒ: in vor 'our', vår 'spring' [ˈvɒ ˈvɒ·ʔ].

5.1.2. The length accent.

The systems of long and short vowels now exhibit the same number of units and the same pattern. We shall briefly discuss the possibility of reducing the two systems to one by considering the long vowels as short vowels plus something else (the remainder being an identical short vowel, h, or a length-unit).*)

If we were to interpret the long vowels as consisting of two identical short vowels, each of these would be commutable with consonants or zero, but with no other vowels (because the diphthongs were interpreted as a short prevowel plus a preconsonant), a situation which does not agree with the fact that vowels and consonants belong to different categories. This implies that if one vowel in such a two-vowel group is commutated, the other vowel will enter into a paradigm comprising the other vowels but no consonants (nor Ø), i.e. it will entirely change its paradigm. This description seems highly artificial and is therefore rejected.

If we consider the long vowels as consisting of a short vowel + h, the phoneme h would be given a "normal distribution" without increasing the number of phonemic elements. But the phonetic description of /h/ would be highly complicated, so this description, too, must be rejected.

To consider the long vowels as short vowels + length is phonetically the most satisfying description, but length is then a new phonemic entity which can be taken in several senses:
(1) as a segmental phoneme /:/ with a number of bound variants, [æ] after /æ/ etc. /:/ is then commutable with consonants and Ø (cp. that /:/ can never form a syllable by itself);
(2) as an entity characterizing the vowel, i.e. as a distinctive feature extracted from the other distinctive features which together constitute the vowel quality.

---

*) This discussion belongs to the present section (5.1.2.) because length will turn out to be an accent. But it should be noted that a reduction of long vowels to two identical vowels or to short vowels plus h would belong to section 6. (after the phonemes have been established).

The difference between (1) and (2) is of theoretical but not of practical consequence in so far as the phonemic notation and the manifestation rules will remain the same;
(3) **as an entity characterizing the syllable, i.e. as an accent.** This interpretation is supported by the fact that the length is often lost when the syllable becomes the first part of a compound. In this respect it behaves like the stød accent but unlike all segmental phonemes (the same loss of stød and length occurs in connected speech when a syllable loses its stress). I therefore choose solution (3), i.e. vowel length is accentual.

5.1.3. The accentual system of ASC.

No unstressed syllables can take the length accent.

The only unstressed syllables with obligatory stød are those with the weak precentrals e and u, whereas those with the weak precentral i can have facultative stød (e.g. in maddingen; vinduet; hyppige [ˈmæðeŋˀn̩ ˈvendu·ˀð̩ ˈhybi·ˀə or ˈhybiə]). This stød may be considered purely automatic, i.e. non-phonemic, according to the following rule: an unstressed syllable with the weak precentral e or u before an unstressed syllable with ə ɒ l̩ n̩ or ð̩ (i.e. the weak precentrals on the 1st level (section 4.2.))have obligatory stød; and an unstressed syllable with the weak precentral i has facultative stød under the same conditions. Now only stressed syllables can have phonemic stød, i.e. a stød accent; the stød and length accents therefore presuppose the presence of the stress accent, but not inversely (in the phonemic transcription, the symbol for the stress accent can therefore be omitted before syllables with length and/or stød accent).

According to this analysis there are four types of stressed syllables:

	long	non-long
stød	ˀ: [ˈmæ·ˀn] /ˀ: mæn/ man 'conjure' (imperative)	ˀ [ˈmænˀ] /ˀmæn/ mand 'man'
non-stød	: [ˈmæ:n(ə)] /:mæn(ɛ)/ *) mane 'conjure'	ˈ [ˈmæn] /ˈmæn / man 'one'

Cp. the scheme of Ege (4).

---

*) The location of the syllable border is not considered here.

5.2. From weak precentrals to phonemes.

The weak precentrals in native words are l̩ n̩ ð̩ ə ɒ i e u (section 4.2.); the weak precentrals ɒ i e u are identified with the vowels ɒ i e u, the only difference being the absence versus presence of the stress accent. ə should not be identified with stressed æ or o because æ and o occur as weak precentrals commutable with ə in foreign words; nor should we identify ə with stressed ɔ or ɑ, because the weak precentral ɒ can vary in the whole range [ʌ-ɒ-ɑ]. Is it reasonable to identify ə with one of the other stressed vowels, i.e. y, ø, œ, or ɛ ? The simplest solution is an identification with ɛ which permits us to state the rule that y, ø, œ, i.e. the vowels of the rounded front series, are never found in unstressed syllables in native words.

The remaining weak precentrals l̩ n̩ ð̩ cannot be identified with any of the other syllabic phonemes, i.e. the vowels; we shall not identify syllabic with non-syllabic phonemes, and l̩ n̩ ð̩ are therefore separate phonemes, i.e. weak centrals.

5.3. The inventory of phonemes.

The resultant phonemes of ASC are:

12 vowels

i	y	u
e	ø	o
ɛ	œ	ɔ
æ	ɑ	ɒ

3 weak centrals

l̩ n̩ ð̩

15 consonants

Both initially and finally 11:

b d g m n l f s r j v

Only initially 2:

h ʃ

Only finally 2:

ŋ ð

In total, 3o phonemes plus 3 accents: stress, stød and length.

6. From phonemes to ultimate phonemes.

6.1. Criteria for the reductions.

To reduce the number of phonemes a further operation must be used, i.e. dissolving one sound into two simultaneous components each manifesting an ultimate phoneme. This presupposes, of course, that the sound is not commutable with the successive manifestation of the two ultimate phonemes.

If we make no complete analysis in distinctive features (or simultaneous components à la Harris), we are in lack of safe phonemic criteria telling us where to stop these reductions. I shall, therefore, make only few of these and only in cases where it leads to a clear simplification in the phonemic description of ASC as a whole and not only a decrease in the number of units. (It should be noted that 'Hjelmslev's Law', which says that a cluster xyz can only exist if also xy and yz exist in that language, is only found to appear as a tendency in ASC, and it can therefore not be used as a criterion telling us where to stop the reductions of the phoneme inventory.)

First, the criterion of "maximally differentiated allophones" is used; i.e. when particularly distinct pronunciations are required, e.g. in the presence of noise, some sounds (e.g. [l n ð]) can always be replaced by certain other sounds ([ˈəl ˈən əð])(their "maximally differentiated allophones") which do not otherwise occur together. If the "maximally differentiated allophones" are groups of sounds where each sound is the normal manifestation of an already established phoneme, a clear simplification of the description as a whole can be obtained by analyzing the "maximally differentiated allophones" instead of the sounds in question (see below).

6.2. The reductions.

(1) l̩ n̩ ð̩ to ɛl ɛn ɛð. [əl ən əð] are the "maximally differentiated allophones" of [l̩ n̩ ð̩] (only occurring in unstressed syllables and manifesting the phonemes l̩ n̩ ð̩). As [ə] is the normal manifestation of the vowel ɛ in unstressed syllables, and [l n ð] of the consonants l n ð, syllabic l̩ n̩ ð̩ are reduced to the ultimate phoneme groups ɛl ɛn ɛð (in unstressed syllables, i.e. occurring without the stress accent).

(2) ɑ: to ɑr. Until this reduction is made ɑ is never found before r, and the word far [ˈfɑ:] stands out as the only exception to the rule that no long vowels occur finally in monosyllabic words. After the restatement, however, oppositions like var, far [ˈvɑ , ˈfɑ:] are accounted for as /ˈvɑ /versus

/ˈfɑr/ . Furthermore, oppositions like pak, park [ˈb̥hɑg̊, ˈb̥hɑːg̊] 'mob', 'park' are now /ˈbhɑg/versus /ˈbhɑrg/ etc., so that we can state the important rule that no long stød-less vowels occur in monosyllabic words. Thus, in monosyllables the length accent presupposes the presence of the stød accent, but not inversely. After the present reduction , ɑ is the only vowel which never occurs in syllables with the length accent. (It should be noted that (2) is not a reduction in the proper sense, the number of phonemes and accents being unchanged, but merely a simplification.)

(3) ɒ to ɔr. ɒ occurs with the length accent in stød syllables only,*) and after the application of (2) ɒ and ɔ are the only phonemes which never occur before r. After reduction (3), a vowel phoneme with a peculiar distribution (ɒ) has disappeared, and another vowel (ɔ) has been given a more normal distribution; the opposition vor, vår [ˈvɒ ˈvɒ·ˀ] which before (3) involved both length and stød, is now /ˈvɔr/ versus /ˀvɔr/; i.e. it involves only stød.

(4) ʃ to sj. In ASC as a whole, both [ʃ] and [sj] can be heard, although [ʃ] is by far the most general pronunciation. Reduction (4) seems justifiable because j will be given a more normal distribution without complicating the distributional description of s.

6.3. The inventory of ultimate phonemes.

The system of ultimate phonemes in ASC is then:

11 vowels:

i	y	u
e	ø	o
ɛ	œ	ɔ
æ		ɑ

14 consonants:

Both initially and finally 11:

b d g m n l f s r j v

---

*) Cp. the reduction at the end of section 4.1.

Only initially, 1 consonant: h

Only finally, 2 consonants: ŋ ð

In total, 25 ultimate phonemes plus 3 accents: stress, length and stød.

----****----

References:

(1) Poul Andersen, "Dansk fonetik", Chapter XV of Nordisk Lærebog for Talepædagoger (1954), pp. 3o8-354.

(2) Poul Andersen, "Om det finale r i Dansk", Danske Studier 1965, pp. 75-84.

(3) Paul Diderichsen, "Udtalen af dansk rigssprog", Danske Studier 1957, pp. 41-79.

(4) Niels Ege, "The Danish Vowel System", Gengo Kenkyu (Journal of the Linguistic Society of Japan) No. 47 (1965), pp. 21-35.

(5) Aage Hansen, Udtalen i moderne dansk (1956).

(6) Zellig S. Harris, Methods in Structural Linguistics (1951).

(7) Louis Hjelmslev, "On the Principles of Phonematics", Proc. Sec. Int. Congr. Phon. Sc. (1936), pp. 49-54.

(8) Louis Hjelmslev, "Grundtræk af det danske udtrykssystem med særligt henblik på stødet", Selskab for nordisk filologis årsberetning 1948-49-5o (1951), pp. 12-24.

(9) Otto Jespersen, Modersmålets fonetik 2 (1934).

(1o) H.A. Koefoed, "Er det finale r en vokal i rigsdansk?", Danske Studier 1964, pp. 85-112.

(11) André Martinet, La phonologie du mot en danois (1937).

(12) Kenneth L. Pike, Phonetics (1943) (p. 145 of the 1966-edition).

(13) Henning Spang-Hanssen, "On the Simplicity of Descriptions", Recherches structurales = Travaux du Cercle linguistique de Copenhague V (1949), pp. 61-7o [Reprinted in Hamp et al.: Readings in Linguistics II (1966), pp. 234-241].

(14) W. Freeman Twaddell, On Defining the Phoneme, Language Monograph no. 16 (1935) [Reprinted in M. Joos: Readings in Linguistics I (1957), pp. 55-79].

(15) H.J. Uldall, "The Phonematics of Danish", Proc. Sec. Int. Cong. Phon. Sc. (1936), pp. 54-57.

ENGLISH STOPS AFTER INITIAL /s/ *)

Niels Davidsen-Nielsen

1. Introduction.

After initial /s/ the English stops present difficulties which are both phonetic and phonemic. With the former it is a question of the exact articulation and phonetic properties of these sounds. With phonemics, it is well-known that the opposition between /p,t,k/ and b,d,g/ is suspended in this position. Thus words like spin, steam, and scold may be interpreted without ambiguity as /spin, sti:m, skould/ as well as /sbin, sdi:m, sgould/. The question is thereby raised as to which of the two phoneme series is realized. In such cases, when from a structural point of view a sound may be included under either of two phonemes with equal justification, it has been the usual practice to choose the solution which is phonetically the more realistic. A large majority of linguists who have applied this criterion have preferred the interpretation /sp,st,sk/. It is the object of this paper to examine whether this preference is phonetically justified.

2. A perceptory experiment. **)

In order to investigate whether the stops after /s/ are auditorily closer to /p,t,k/ or /b,d,g/ a perceptory experiment was carried out.

2.1. The first part of this experiment consisted in letting native speakers of English identify a number of words which begin with sp, st, sc but whose initial [s] had been removed.

----

*) This paper is an abbreviated version of an article in the special Danish issue of "English Studies" (to appear in 1969). I am indebted to professor E. Fischer-Jørgensen for her kind assistance and many valuable suggestions.

**) For American English the results of this experiment are supported by two previous investigations: Lotz, Abramson, Gerstman, Ingemann, Nemser (4) and Reeds & Wang (5).

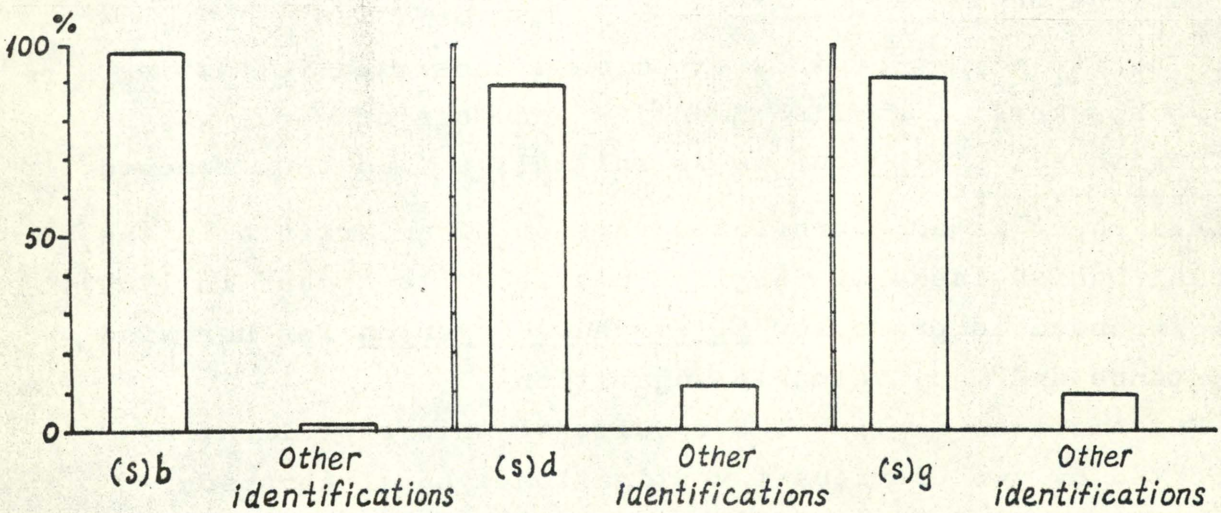
The following words were selected:

I	II	III
spear	pier	beer
spat	pat	bat
steam	team	deem
sty	tie	dye
scold	cold	gold
score	core	gore

The words were inserted into sentences and recorded by four persons (two English and two American) at a tape speed of fifteen inches per second. The test words from column I were then cut out of their environments. In the six words thus isolated the initial [s] was removed. By hand, the tape was moved slowly past the play-back head and was subsequently cut immediately after the friction of [s] had ceased, i.e. during the first part of the following stop.

The perceptory experiment consisted in letting 32 test persons (24 English and 8 American) identify 52 recordings of these six truncated words. Each of the words, which had been randomized, was played twice to the test persons, who were then asked to write down the English word they thought they heard.

The experiment showed that in 92% of the cases the test persons heard words beginning with /b,d,g/, i.e. the words which appear in column III, while in 8% of the cases they heard words beginning with another sound, which was nearly always /p,t,k/, i.e. the words in column II. It turned out that the words with <u>sp</u> were identified more consistently than the words with <u>st</u> and <u>sc</u>. This is shown in the following bar chart:



In order to examine British and American English separately the material was subsequently divided into two parts. The 24 Englishmen's judgment of the British English recordings showed 92% /b,d,g/ identification. The 8 Americans' judgment of the American English recordings displayed 98% /b,d,g/ perception.

2.2. In the second part of the perceptory experiment the test persons were asked to identify two recordings, one American and one British English, of the following sentence:

'Thanks, Stan, that'll be all.'

It was anticipated that this sentence in a number of cases would be confused with:

'Thanks, Dan, that'll be all.'

This was expected because the first part of the test had proved that the stop in st is perceived as a /d/. In this way the phonetic difference between the two sentences becomes very small and is probably restricted to a difference in the length of the [s].

The presumption that this distinction is precarious was supported clearly by this test. In only 58% of the cases the test persons identified the sentence correctly, whereas 'Dan' was heard in the remaining 42% of the cases.

3. Acoustic investigations. *)

Of all the recorded words in the three columns there were taken one duplex oscillogram, one pitch curve, and two intensity curves (logarithmically and linearly registered respectively). On these three curves and on the oscillogram the stops could be precisely delimited. In this way it was possible to measure the duration of the three stop series. Furthermore voicing and release burst could be examined.

----

*) B. Frøkjær-Jensen, of the Institute of Phonetics, has been of great assistance at the execution of the acoustic and articulatory parts of the investigation.

3.1. Duration of release stage.

The duration of explosion plus aspiration was measured in all the recorded words of the three types (216 words). It appeared that the average duration of this element of [(s)p, (s)t,(s)k] was 2 cs., while in the case of [p,t,k] it was 8 cs. As regards [b,d,g] the average duration of this element turned out to be 3 cs. On the basis of these measurements the only reasonable conclusion is to consider the stops after [s], as well as initial [b,d,g], unaspirated. [p,t,k], on the other hand, can only be termed aspirated.

3.2. Duration of hold stage.

On measuring this stage of the stops in the above mentioned 216 words it turned out that there was no significant difference between [p,t,k] and [b,d,g] in this respect. The average duration of the hold stage of [p,t,k] was 11 cs.; for [b,d,g] the mean length was 10 cs. In pairs the averages were the following: b/p: 10/12, d/t: 8/9, g/k: 12/11. The hold stage of [(s)p,(s)t,(s)k] was somewhat shorter (7 cs.). This, however, is to be expected in consonant clusters. On the basis of this material, consequently, no conclusions regarding force of articulation in English can be drawn from the duration of the hold stage.

3.3. Voicing.

An investigation of voicing showed very clearly that both [p,t,k] and [(s)p,(s)t,(s)k] were unvoiced. The traditional notion of [b,d,g] being articulated with a certain amount of voicing proved largely correct with regard to the two Americans. As far as the two Englishmen are concerned, however, divergent results were obtained. In the first place, all these sounds were unvoiced initially in a sentence. Secondly they were not fully voiced between vowels. It appears, then, that [b,d,g] in English may be realized in two different ways: with or without vibration of the vocal cords during the hold stage. It may be concluded that voicing cannot be regarded as any constant feature of these stops in English.

4. Articulatory investigation.

The last part of this investigation consisted in a physiological examination of intra-oral air-pressure. In the case of the two Americans and one of the two English test persons this was measured by inserting a thin plastic tube into the mouth behind the point of articulation. This tube was connected with an electric manometer,*) which registered the air-pressure in cm H_2O by means of a mingograph. At the same time a duplex oscillogram and two intensity curves were registered. The air-pressure of the fourth test person was measured by inserting the tube through the nose into the mesopharynx. In this way it was possible to measure also [k], [g], and [(s)k].

The experiment showed that two Americans (J.W. and M.W.) had a very clear difference between [p,t,(k)] and [b,d,(g)], which consisted in [b,d,(g)] having considerably lower pressure than [p,t,(k)]. Converting absolute into relative values by stating the percentage value of the voiced stops in relation to the unvoiced stops the following figures were arrived at:

	J.W.	M.W.
b/p	73	57
d/t	76	69

These results support the traditional view that [p,t,(k)] are fortes whereas [b,d,(g)] are lenes. This, however, was not the case with the English informants (N.S. and R.D.), who showed almost no difference of air-pressure between the two stop series:

	N.S.	R.D.
b/p	96	96
d/t	90	109
g/k	87	

----

*). Constructed by A. Tybjærg Hansen and described in Eli Fischer-Jørgensen & A. Tybjærg Hansen (2).

The figures from the last table may seem surprising; but in an investigation by Eli Fischer-Jørgensen (1), voicing and intra-oral air-pressure have been demonstrated to be closely connected: voiced sounds have relatively low intra-oral air-pressure, unvoiced sounds relatively high pressure. Considering this negative correlation it could be anticipated that almost no difference of air-pressure could be demonstrated with the English test persons, for their [b,d,g]'s were predominantly unvoiced.

It was now to be expected that [(s)p,(s)t,(s)k], as regards intra-oral air-pressure, would be similar to [p,t,k], and that both these series would show relatively high pressure. The assumption proved correct for three out of four persons. The differences registered here were inconsiderable:

	N.S.	R.D.	J.W.
(s)p/p	98	98	99
(s)t/t	100	111	106
(s)k/k	96		

With regard to the fourth informant, who found it difficult to articulate with a plastic tube in her mouth, the pressure of [(s)p,(s)t] was somewhat lower than that of [p,t].

5. Summary and phonemic evaluation.

5.1. On the basis of this auditory, acoustic, and physiological investigation of stops after initial /s/ in English the following points may be underlined:

The perceptory experiment demonstrates that these sounds are significantly closer to /b,d,g/ than to /p,t,k/. This has been proved with American English in two previous investigations. The results obtained here, however, also suggest that British English is similar to American English in this respect.

The acoustic investigation shows that the stops after /s/ are unaspirated and unvoiced. They have the first of these features in common with /b,d,g/, the second with /p,t,k/. But the experiment also demonstrates that voicing is no constant feature of English /b,d,g/. There is, therefore, a hierarchic organization among the features of the stops in this position according to which aspiration is more important than voicing.

Lack of aspiration, consequently, forces speakers of English to evaluate the stops after /s/ as /b,d,g/.

It appears from the physiological investigation that [b,d,g] when voiced have lower intra-oral air-pressure than [p,t,k] and [(s)p,(s)t,(s)k]. When [b,d,g] are unvoiced, however, the three stop series have nearly the same intra-oral pressure.

5.2. With regard to the phonemic interpretation of these consonant groups this investigation demonstrates that the interpretation /sb,sd,sg/ is preferable if the criterion of "phonetic similarity" is applied. If, on the other hand, the criterion of "pattern congruity" is used, Hockett (3,p.159) is of the opinion that /sp,st,sk/ is the better solution, for as /p,t,k/ in codas are found together with /s/ (e.g. 'past'), while /b,d,g/ go with /z/ (e.g. 'razzed'), we arrive at greater symmetry if this is also considered to be the case initially. This argument, however, is not really convincing. In the first place, the structural possibilities of initial and final consonant clusters in English are clearly different, and it is therefore less obvious to generalize from final to initial position. Secondly, Hockett's interpretation of the final cluster of obstruents in 'passed' as /pa:st/ is not the only possible one:/pa:sd/ is also conceivable. It may be held *) that in a final cluster of two consonants in English the opposition between voiced and unvoiced consonant is suspended finally if voicing is distinctive with regard to the prefinal consonant. In a pair like 'raised' and 'raced' the two words are distinguished by means of vowel length and voicing of the sibilant. The final consonants, however, are respectively unvoiced ('raised') and unaspirated ('raced').**)

----

*) This solution has been pointed out to me by Jørgen Rischel, of the Institute of Phonetics, from whom I have received many valuable suggestions.

**)The possibility of final aspiration in 'raced', however, can hardly be excluded.

The form 'raced' may therefore be interpreted unambigously as either /reist/ or /reisd/, and also be kept apart from 'raised' /reizd/ if the second solution is chosen. This implies that it is possible to operate with final /sb,sd,sg/, e.g. 'lisp' /lisb/, 'past' /pa:sd/, 'cask' /ka:sg/. If the initial consonant groups written <u>sp</u>, <u>st</u>, <u>sc</u> are also interpreted as /sb,sd,sg/ the demand for symmetry is satisfied. This solution does not in any way conflict with the principle of "phonetic realism" as progressive assimilations of voicing are to be expected in such environments. On these grounds /sb-,sd-,sg-/ seems to be a legitimate analysis.

<u>References:</u>

(1) Eli Fischer-Jørgensen, "Beobachtungen über den Zusammenhang zwischen Stimmhaftigkeit und intraoralem Luftdruck", <u>Zeitschr. f. Phonetik</u> 16 (1963), pp. 19-36.

(2) Eli Fischer-Jørgensen & A. Tybjærg Hansen, "An Electrical Manometer and its Use in Phonetic Research", <u>Phonetica</u> 4 (1959), pp. 43-53.

(3) Ch. F. Hockett, <u>A Manual of Phonology</u> (1955).

(4) Lotz, Abramson, Gerstman, Ingemann, Nemser, "The Perception of English Stops by Speakers of English, Spanish, Hungarian, and Thai", <u>Language and Speech</u> (1960), pp. 71-77.

(5) J.A. Reeds & W.S.-Y.Wang, "The Perception of Stops after s", <u>Phonetica</u> 6 (1961), pp. 78-81.

(A more detailed list of references will be given in my forthcoming article.)

VOICING, TENSENESS AND ASPIRATION IN STOP CONSONANTS, WITH SPECIAL REFERENCE TO FRENCH AND DANISH

Eli Fischer-Jørgensen

1. Introduction.

1.1. The present report is, to a great extent, a summary of two papers which will be published in 1969, (a) "Les occlusives françaises et danoises d'un sujet bilingue" (to appear in the Festschrift to A. Martinet, vol. II) and (b) "Observations sur les traits phonétiques distinguant ptk de bdg en français".*) Both of these papers contain an account of speakers, texts and instrumental set-up and a detailed presentation of the results in the form of tables of averages and diagrams. These facts will be presented rather briefly here, in section 2, with close line spacing. Instead, the report will mainly be concentrated on the general problems associated with the features of voicing**), tenseness and aspiration, their mutual relations (particularly the relation between tenseness and aspiration) and the problem of attributing the physiological and acoustic phenomena to these features. A special discussion is devoted to intra-oral pressure and F_1 transitions. (These problems are treated briefly at the end of the first-mentioned paper, and more extensively in the second paper.)

---

*) The investigations have been carried out in the laboratory of our institute. I am grateful to Børge Frøkjær-Jensen for help. The work has been subventioned by Statens almindelige Videnskabsfond.

**) The term "voicing" is used here in its traditional sense, indicating vibrations of the vocal cords. In many papers by members of the Haskins Group "voiced and voiceless stops" are simply used to indicate respectively bdg and ptk. But as bdg may be voiceless (in the traditional sense of the word) in many languages, and ptk are occasionally voiced in some languages, this terminology does not seem to be very practical, and it may sometimes be rather confusing.

Readers mainly interested in the general problems may skip section 2 containing the results of the measurements (with the exception of Table I, p. 80 below, which gives a summary of the results). Section 3 gives a brief survey of the stability of the differences, and section 4 contains a summary and discussion of the perceptual tests. In section 5 (general problems) references will be given to the more important points of section 2.

Some books and articles to which I did not have access when finishing the two papers are mentioned in this report, and the titles are added to the references at the end of the report, namely: Chomsky and Halle (1968), Delattre (1968), Halle and Stevens (1967), Ladefoged (1967), and Perkell (1965 a and b).

On some points the present report contains new facts and modifications of earlier hypotheses. It thus represents a later stage in the discussion of the problems than the two papers in print, although it appears before these.

1.2. The problem from which I started was the difficulty of describing tenseness and aspiration as belonging to the same feature, e.g. the difficulty of describing the phonetic differences between ptk and bdg as belonging to the same opposition in French and Danish, as required by Jakobson-Fant-Halle (1952, p. 38). Moreover I was puzzled by the tenseness difference in the narrower sense of the word, e.g. by the difficulty of giving a precise phonetic description of the difference between French ptk and Danish (voiceless) bdg, which are obviously different from a perceptual point of view (Danish bdg are not accepted as good French ptk), and which seem to differ only in respect of tenseness (fortis-lenis).

1.3. One way to throw some light on these questions is to analyse the pronunciation of bilingual speakers. It is, however, not quite easy to find persons who speak both French and Danish perfectly. The bilingual subject used in the first investigation, CHH, daughter of A. Martinet, speaks both languages fluently, and she has a normal Copenhagen pronunciation of the Danish stops (her bdg are perhaps slightly too "strong"); the only evident French influence found in her pronunciation of Danish concerns the rhythm. However, her French bdg are almost totally voiceless, perhaps due to Danish influence, and her French ptk are in some cases strongly aspirated and affricated. This pronunciation of ptk does not seem to be due to Danish influence. Exactly the same type of affrication was found in the speech of one of the other French subjects (SRO), and moreover it differs from the normal Danish aspiration and affrication in three respects: (a) it includes affrication of d before i, (b) it does not include aspiration before open vowels, (c) the aspiration before u is relatively weak, i.e. there is only affrication before front vowels. (In Danish ptk are strongly aspirated before all vowels, and t is always affricated.) This type of affrication is characteristic of Parisian (and Canadian) French, and, as a matter of fact, both CHH and SRO are from Paris.

This Parisian pronunciation makes the comparison more complicated. Of the three vowels used almost exclusively in this investigation (a, i and u) only a is preceded by really unaspirated ptk. On the other hand, the fact that CHH's French bdg are voiceless, has the advantage for the investigation that tense and lax voiceless consonants are found within one language.

1.4. CHH's French and Danish consonants must be seen on the background of normal Danish and French stops. For Danish I can refer to extensive but not yet fully processed data, and to a few papers (Abrahams (1949), EFJ (1954) and (1966), EFJ, Frøkjær-Jensen and Rischel (1966), and Frøkjær-Jensen (1967)).

Instrumental analyses of French stops are more numerous, but all of very limited size (see Rousselot (1891), (1897), (1899), Roudet (19oo a), (19oo b), (191o), Chlumsky (1922), Evertz (1929), Marguerite Durand (1936), (1956), Brunner (1953), Delattre (194o a), (1949 b), Straka (1942), (1953), (1965), P. Simon (1961), (1967), Thorsen (1967), and K. Landschultz (1968)).

Experiments with synthetic stops identified by French listeners have been carried out by Marguerite Durand (1956). In the experiments with ptk versus bdg undertaken by the Haskins group the listeners were mainly Americans (cp. Libermann, Delattre and Cooper (1958), see also the summaries of the results given by Delattre (1958), (1961), (1964)). Delattre (1968) gives the perceptive cues for French phonemes, but nothing is said about the number of listeners.

The number of subjects used in my analysis of French stops is very restricted, partly because it was difficult to find good subjects, partly because I have preferred to undertake a thorough analysis of a few subjects, which would make it possible to find relations between the different phonetic factors and to draw some conclusions about the physiological mechanism.

The main subject, apart from CHH, was SRO. The supplementary subjects are indicated by the initials EH, JPP, JT, Sch, and MAS. Moreover I have had the opportunity to use material recorded by O. Thorsen and K. Landschultz. The stops of CHH and SRO were analysed from many points of view: position of the glottis, intra-oral pressure, airflow (only CHH), lip pressure, duration, voicing, intensity of the explosion and formant transitions. The stops of the supplementary subjects were only analysed in some of these respects. The stops of five Danish subjects have, with a few exceptions, been analysed in all the respects mentioned.

The text material used consists almost exclusively of words containing stop consonants in initial stressed position after unstressed vowel and before the vowels a, i or u, e.g. la panne, la balle, les pistes, les boules etc. The choice of this position was conditioned by the fact that in Danish ptk and bdg are not distinguished in final position, nor in medial position before ə.

2. Results of the measurements.

2.1. Physiological measurements.

2.1.1. Position of the glottis.

An endoscopic examination of the main French subject SRO gave the result that p was spoken with slightly open glottis. This result was corroborated by glottograms taken with the Fabre glottograph, showing higher electric resistance in the glottis for ptk than for bdg. Glottograms for JPP gave the same result (the number of word pairs were 99 for SRO and 84 for JPP).

Endoscopic examinations of Danish stops have shown that ptk are spoken with widely open glottis, bdg with slightly open glottis (the intercartilaginous part being normally closed in bdg). The difference in degree of opening has been corroborated by glottographic recordings.

Fabre glottograms of 72 French and 56 Danish word pairs with the vowel a spoken by CHH show a corresponding difference between ptk and bdg in the two languages. There is no clear difference between the languages as far as the maximum opening is concerned. However, it is not quite certain whether the plates were removed in between the two sessions, and since their placement may influence the maximum amplitude of the curve, this result must be taken with some reservation.

The place of the maximum in the two languages is, however, of interest. Both French and Danish subjects show a predominantly falling curve for bdg. French ptk have a predominantly rising-falling curve during the closure period, whereas Danish ptk have a rising curve up to the moment of explosion, followed by a fall (see EFJ, Frøkjær-Jensen and Rischel (1966), and Frøkjær-Jensen (1967)). Aspirated Danish ptk thus have a much wider glottis at the moment of explosion compared to unaspirated French ptk and to bdg.*) CHH's glottograms show this difference clearly for French and Danish t and k, whereas her Danish p has its maximum at the start of the closure period, although it is also aspirated (see the schematic average glottograms Fig. 1).

The curves of SRO and JPP are more irregular than those of CHH and the Danish speakers, and the French material needs corroboration. There may be individual differences (see Kloster Jensen (1956)).

2.1.2. Airflow.

The maximum airflow after the explosion of French stops has only been measured for JPP and CHH. They both have a stronger airflow after ptk than after bdg, but CHH (118 pairs) has a much smaller difference (t/d no difference, p/b k/g + a 2.9 l/m, + u, i 5.1 l/m) than JPP (38 pairs, 11.8 l/m, only a). In her Danish stop consonants CHH has a considerably greater difference between ptk and bdg (100 pairs, 15.0 l/m), and her Danish bdg have a somewhat weaker airflow than her French ptk (5.3 l/m). The curves of the Danish subjects have not been measured, but the difference is evident, and it may even be greater than that found for CHH. When the consonant is affricated, the airflow is relatively slow at the start.

2.1.3. Intra-oral air pressure.

The intra-oral air pressure of labial and dental stops has been measured (in cm H_2O) for the French subjects SRO (92 pairs) and EA (40 pairs). The peak pressure is evidently and significantly higher for pt than for bd, the pressure for the latter amounting to 49 and 45 per cent of that for pt for the two subjects. In an earlier kymographic recording of 15 pairs spoken by JT the percentage is 64. Thorsen found, for his 5 speakers, an average percentage of 61, and the measurement of

---

*) See also Frøkjær-Jensen in this report p.12.

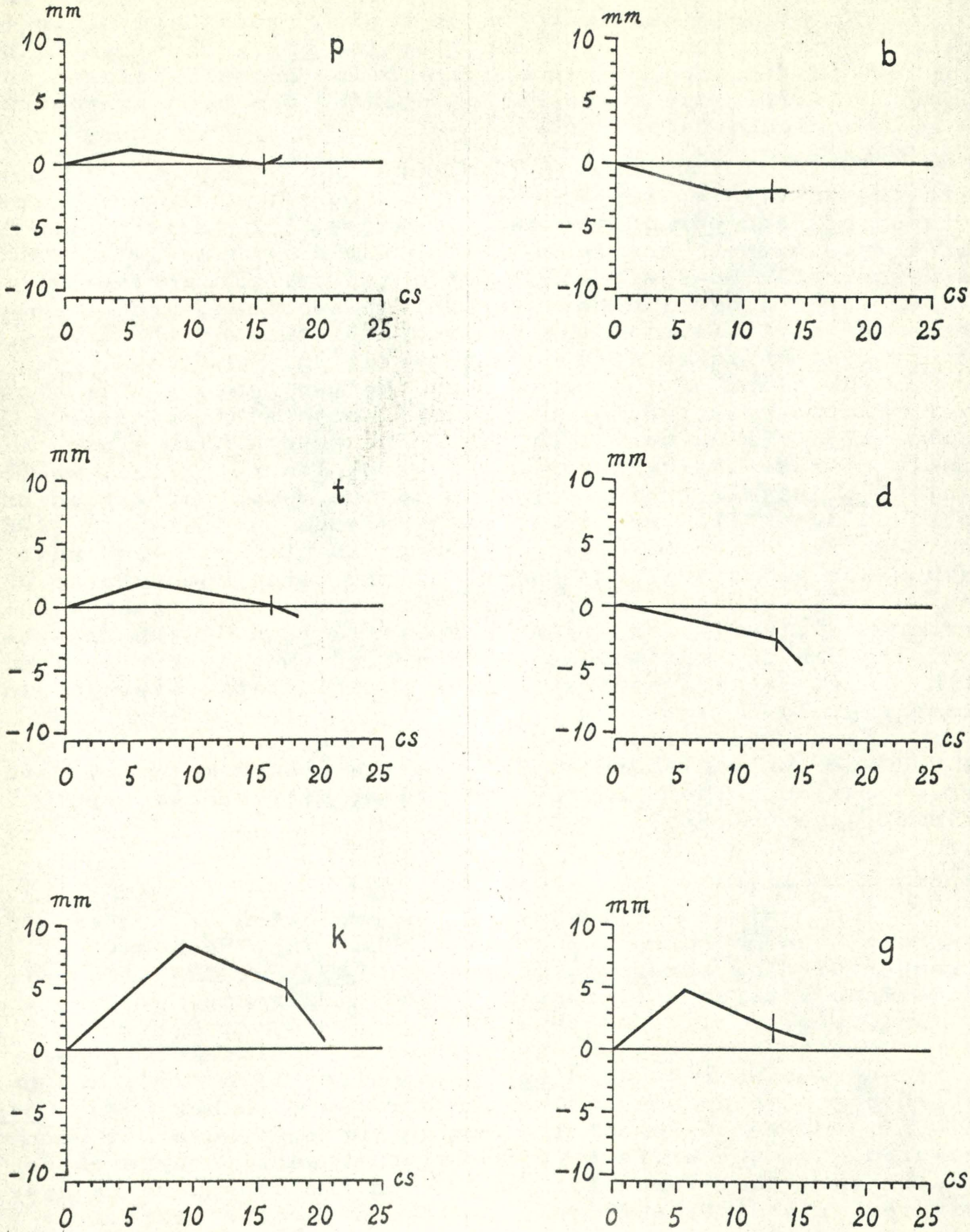


Fig. 1 a.

Average glottograms of CHH's ptk and bdg in French

| = explosion

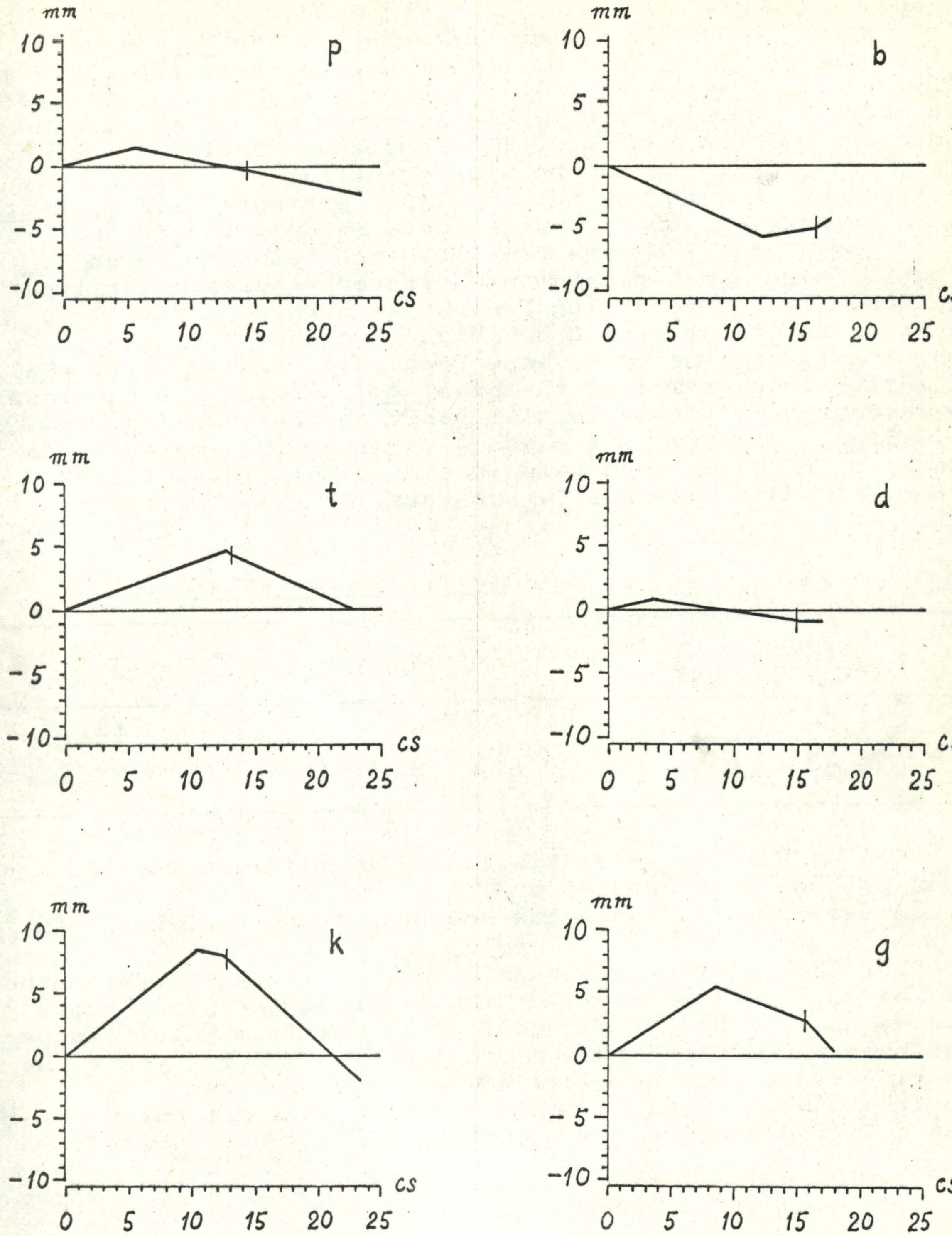


Fig. 1 b.

Average glottograms of CHH's <u>ptk</u> and <u>bdg</u> in Danish

| = explosion

81 further examples from his material gives an average of 71 per cent.

In Danish the difference between ptk and bdg is very small; some speakers have no difference, others have bdg slightly lower (around 92 - 96 per cent of ptk), and in these cases there is generally only a difference at the end of the closure, ptk showing an increase of pressure during the closure period, probably owing to the wider glottis.

CHH has a tendency toward higher pressure in French ptk than in French bdg, but the difference is only significant for p/b (92 per cent). She has no difference between Danish ptk and bdg. There is no consistent difference between the two languages. In one recording French has higher peak values, in another recording Danish has higher values.

Apart from the peak value French pt and bd differ significantly in the tempo of the rise, bd showing a much slower rise (except in sentence initial position where also pt have a slow rise). The rise at 2 and 4 cs after the implosion in percentage of the maximum value has been calculated for SRO and for Thorsen's material (4 subjects, 81 examples). The result is:

SRO:			Thorsen's subjects:		
	2 cs per cent	4 cs per cent		2 cs per cent	4 cs per cent
p	62	72	pt	68	92
b	48	52	bd	43	63

The difference is statistically significant for SRO and for two of Thorsen's subjects.

A similar difference has not been found in Danish air pressure curves.

For CHH the rise has been measured by a different method, namely as the distance from the implosion to the point where the curve has reached 85 per cent of its maximum value (in her curves this point very often corresponds to a point where the quick rise stops and the curve becomes more horizontal).

She has, both in French and Danish, a small but significant difference between ptk and bdg, the averages being:

French ptk	2.6 cs
French bdg	4.5 cs
Danish ptk	2.4 cs
Danish bdg	3.7 cs

It appears from these averages that there is practically no difference between her French and Danish ptk, and that Danish bdg is in between ptk and French bdg.

These values cannot be compared directly to the values given for the French subjects, but it is evident that 85 per cent rise in 4.5 cs for CHH's French bdg is considerably quicker than 52 and 63 per cent in 4 cs for the French subjects, and it is also evident that her ptk have a quicker rise than those of the French subjects.

The so-called pressure impulse, i.e. the area under the curve (measured by Lisker (1965) and Malécot (1966 a)) has not been calculated for these recordings, because I consider it a rather complicated measure combining three parameters (peak value, rise time and duration), which I prefer to give separately, and because it depends on the scales used. It would show a pronounced difference between French pt and bd, because all three parameters go in the same direction, and it would give a somewhat greater difference between CHH's French ptk and bdg than the difference in peak value for the same reasons. For Danish it would in many cases give slightly higher values for bdg than for ptk (depending on the scales used for pressure and duration), because bdg have a longer closure period, but often slightly lower peak pressure.

Both SRO, EA and CHH show an increase in pressure during the closure period of ptk with the maximum at or very close to the explosion. This may be due to the open glottis.

Unaspirated stops have an abrupt fall after the explosion, aspirated stops normally have a short abrupt fall followed by a slowing down, and affricated stops exhibit a slow fall from the start.

2.1.4. Lip pressure.

Lip pressure has been recorded by means of a rubber bulb and measured in mm on the curve for SRO (71 word pairs), Sch. (16 pairs) and JT (18 pairs). SRO and Sch. show a significant difference between p and b, the latter reaching 72 per cent (SRO) and 66 per cent (Sch.) of the peak of p. In JT's examples the stop consonant was in initial position in isolated words, and in this position the lip pressure is very irregular. Thorsen has found the lip pressure of b to be 7o - 89 per cent of that of p for four subjects.

CHH has only a slight tendency toward stronger lip pressure in her French p, but a significant, though small, difference in Danish, where b has stronger lip pressure than p, in accordance with the general tendency in Danish.

2.2. Acoustic measurements.

2.2.1. Voicing.

The French material (5 subjects) contains 41o examples of bdg; 87 per cent of these examples are voiced throughout, whereas 13 per cent of the examples are voiceless in the last part (on the average 23 per cent) of the closure period. There is only one example of overlapping with ptk. 29 examples in final position are fully voiced.

This is in accordance with earlier measurements (see particularly Evertz (1929) and Brunner (1953)). Thorsen's five subjects (Thorsen (1967)) also have fully voiced bdg both in intervocalic and in final position. I have also measured 174 examples of b from K. Landschultz' material (4 subjects, among whom also CHH). In intervocalic position these are all fully or partly voiced, and only CHH has less than 4o per cent voicing.

As already mentioned CHH has practically voiceless French bdg. There is, however, a slight but significant difference between ptk and bdg in that bdg show a slightly longer voiced interval in the beginning through assimilation to the preceding vowel:

French ptk	1.4 cs
French bdg	3.1 cs
Danish ptk	1.5 cs
Danish bdg	2.2 cs.

As it was the case with the tempo of the pressure rise, Danish bdg are in between ptk and French bdg (there are approximately 4oo examples of each set). The variations are small except for French bdg, which vary between 1.o and 13.5 cs. Four examples of b are fully voiced.

In curves of Danish intervocalic stops initially in a stressed syllable one finds generally 1-2 cs voicing, with a small, but unstable tendency for bdg to show more voicing than ptk. If such a syllable loses its stress in the sentence, bdg and sometimes ptk may become voiced, but they are still kept apart by means of the aspiration of ptk. Before weak syllables with ə there is only one series of stops. These often have (relatively weak) voicing.

2.2.2. Fundamental frequency of the following vowel.

The fundamental frequency of the following vowel has been measured for SRO. The vowel starts on a lower tone after bdg than after ptk in 75 out of 78 pairs (spoken in alternating order), the average difference being 27 cps.

A tendency toward a similar difference is found in one of the recordings of CHH, but not in the other recordings. The fundamental frequency has not been recorded for the other subjects.

In Danish no such difference has been found.

2.2.3. Intensity of the explosion.

The difficulty of separating the intensity of the explosion from the intensity of the voicing makes a comparison between ptk and voiced bdg rather problematic. An attempt to separate the two factors in SRO's stops by means of a high-pass filter with a cut-off frequency of 5oo cps was not quite successful. On the one hand there was often some voicing left, on the other hand the filtering may have removed some of the explosion noise in labials, and in velars before u. The high-pass filtered intensity curve showed a significantly higher intensity for k than for g (although with much variation), but no clear difference for labials and dentals. A restriction to the curves in which all voicing had been removed did not give a better result. (The integration time was 2.5 ms.)

CHH's curves and the Danish curves did not present this difficulty, but they did not show any consistent difference either, except that Danish t has a weaker explosion than d. On the other hand the aspiration of p, k and particularly t is strong. CHH's Danish b and g tend to be slightly stronger than her French p and t.

A comparison between the explosions of French ptk and Danish bdg spoken by a larger number of subjects has been planned by the author. Such a comparison is possible when the intensity of the explosion is measured in relation to the following vowel provided that the words are said with normal voice effort (the quality of many Danish and French vowels is quite similar).

A preliminary comparison between the author's bdg and SRO's ptk gave the result that the Danish b was somewhat weaker than the French p, but nothing could be said about d/t and g/k.

On the whole the intensity of the explosion seems to be of very restricted importance.

(The number of explosions measured was rather restricted, because most of the intensity curves have been taken with a linear scale. A logarithmic scale is necessary when both explosions and vowels are to be measured.)

2.2.4. Intensity rise of the following vowel.

SRO's intensity curves of stop consonants before the vowels i and u (54 pairs) display a slower rise of the vowel after bdg than after ptk in 53 cases. JT has 10 pairs before u and Sch. 9 pairs. All of JT's pairs and 8 of Sch.'s 9 pairs have a slower rise after bdg. In the case of a (SRO 27 pairs, JT and Sch. 19) and e (JT and Sch. 19 pairs) one finds, however, rather the opposite tendency. (But from an auditory point of view a may perhaps also have a slower increase in loudness after bdg than after ptk, because in the former case F_1 starts at a lower frequency where the ear is less sensitive.)

Neither CHH nor the Danish subjects show any such difference. The rise of the vowel is in all cases rather abrupt, much like that of the vowels after ptk in French.

2.2.5. Duration of the closure period.*)

It has been shown in earlier work on the subject that French ptk have a longer closure period than bdg (e.g. Marguerite Durand (1936), Evertz (1929, pp. 22 ff.), P. Simon (1967, pp. 174 ff.)). The same has been found in the present investigation. The differences are small but relatively stable and statistically significant: SRO (221 pairs) 3.6 cs, JPP (82 pairs) 2.3 cs, EH (69 pairs) 4.8 cs, JT (18 pairs) 3.5 cs, Sch. (18 pairs) 4.0 cs. In Thorsen's material the difference for JT is 3.6 cs. On the whole the closure period of bdg is about 80 per cent of that of ptk.

In Danish stops the relation is the opposite, bdg having a longer closure period than ptk. For labials and velars the difference is 2-3 cs, for dentals 4.5 cs (this difference according to place of articulation is obviously due to the particular shortness of the closure period of the strongly aspirated and affricated Danish t).

In accordance with the general difference between French and Danish, CHH has a longer closure period in ptk than in bdg in French (300 pairs, 2.3 cs), and a shorter closure period in ptk than in bdg in Danish (283 pairs, 1.5 cs).

----

*) Duration is here treated as an acoustic phenomenon because of the method of measurement.

The averages for the four sets are:

French <u>ptk</u>	16.2 cs
French <u>bdg</u>	13.9 cs
Danish <u>bdg</u>	13.7 cs
Danish <u>ptk</u>	12.2 cs

Danish and French <u>bdg</u> are almost alike, but the other differences are stable and significant, though small.

Thorsen has found the differences in French to be particularly large and stable in sentence final position, where the closure of <u>bdg</u> is only 65 per cent of that of <u>ptk</u>. In the present material there is only a small number of final stops, they show a greater difference than the initial stops. In the interior of a sentence the difference between word final <u>ptk</u> and <u>bdg</u> does not seem to be stable, however (this appears from K. Landschultz' material).

2.2.6. The duration of the preceding vowel.

It has often been observed that the vowels in French are longer before voiced fricatives than before voiceless fricatives. Detailed measurements have been made by K. Landschultz (1968). P. Delattre (1939),(194o),(1962) has pointed out that the difference is also valid before stops, and this appears clearly from O. Thorsen's and K. Landschultz' materials. In the present texts there is only one comparable case (SRO 16 pairs, difference 2.5 cs).

2.2.7. Duration of the following vowel.

In SRO's recordings there are four comparable word pairs comprising 95 single pairs, and EH has one word pair (16 ex.). In all cases the vowel is longer after <u>bdg</u> than after <u>ptk</u> (with average differences from 1.4 to 3.8 cs). Similar relations have been found for Danish (EFJ (1964, p. 186) and English (Peterson-Lehiste (196o)) where it might be explained as a consequence of the aspiration. But in the French examples the difference in vowel duration is larger than the difference in duration of the open interval; thus it cannot be explained by the latter, but perhaps, as suggested by Delattre, by compensation of force of articulation. No measurements have been made for CHH.

2.2.8. Duration of the open interval.

The term "open interval" is used to designate the distance from the explosion to the beginning of the vowel. In French <u>bdg</u> this interval is usually voiced, in <u>ptk</u> and in Danish stops it is always voiceless. In the latter cases it is the same as "voicing lag".

In French stops there is a significant difference between the open intervals after <u>ptk</u> and <u>bdg</u>, but it is difficult to summarize the results in a few indications of averages because the variation is rather large. The interval is relatively long after velars and relatively short after labials, and it is longer before close vowels than before open vowels (these are general tendencies found in many other languages). The single

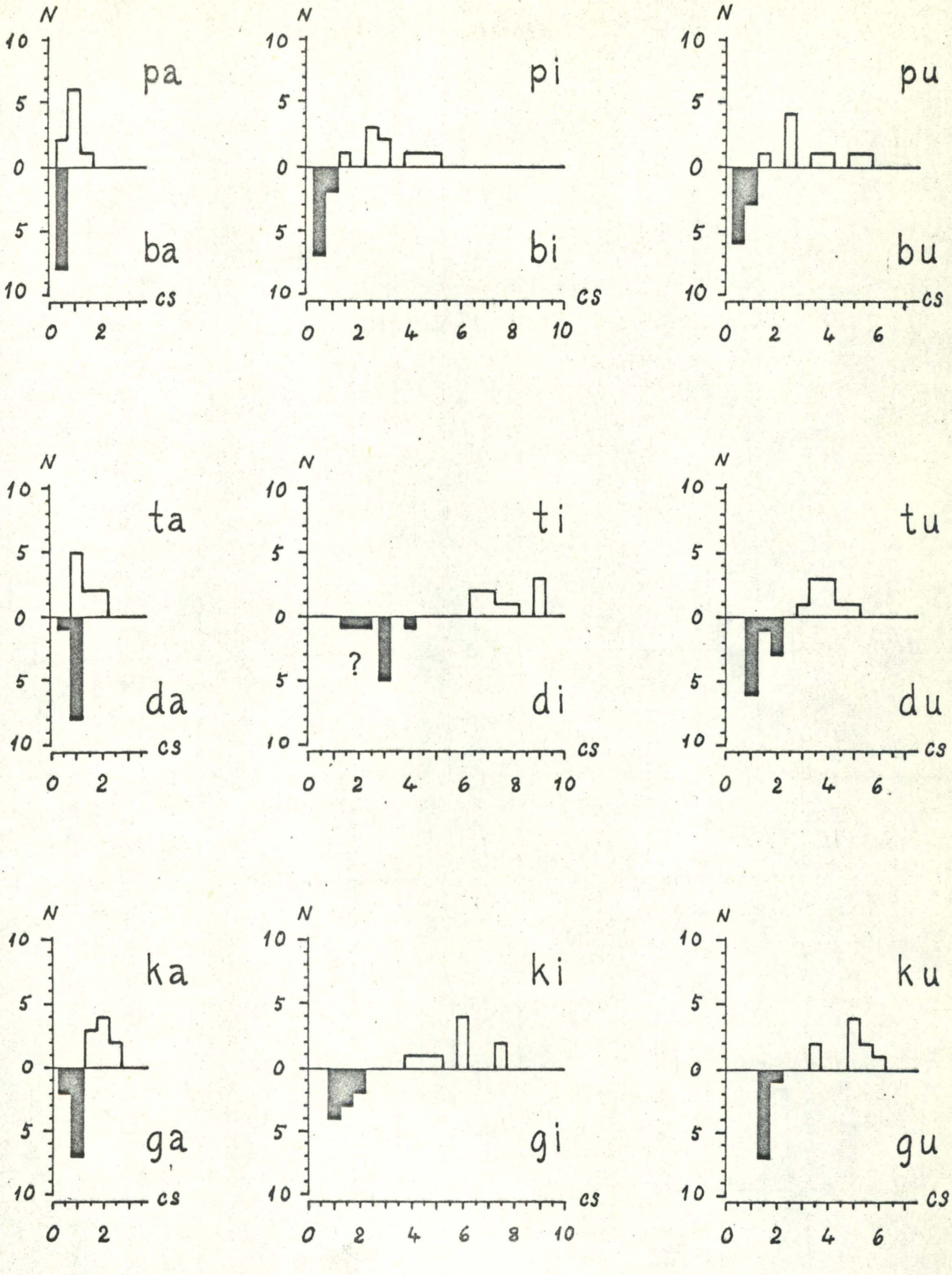


Fig. 2.

Duration of the open interval (SRO)

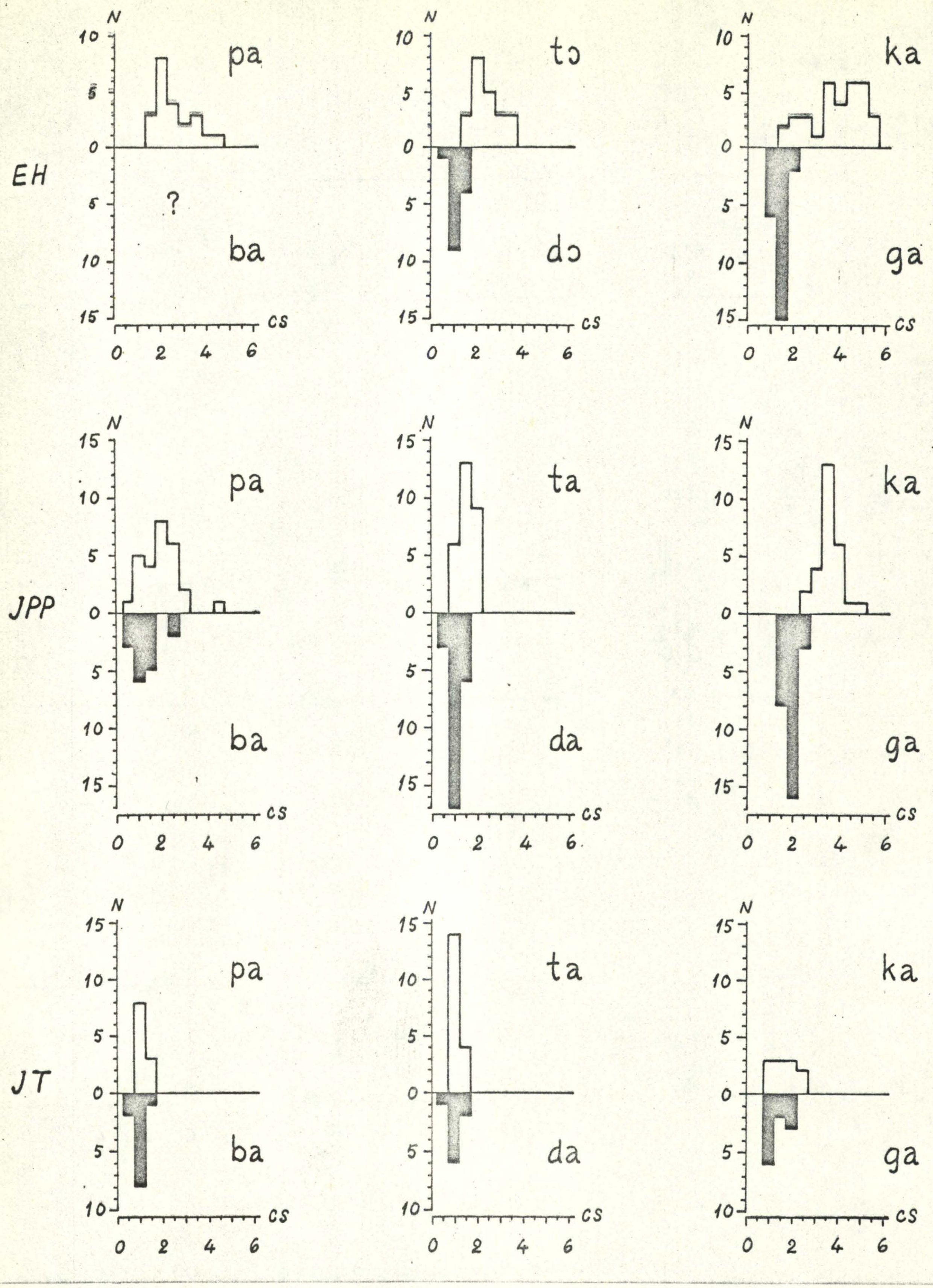


Fig. 3.

Duration of the open interval (EH, JPP, JT)

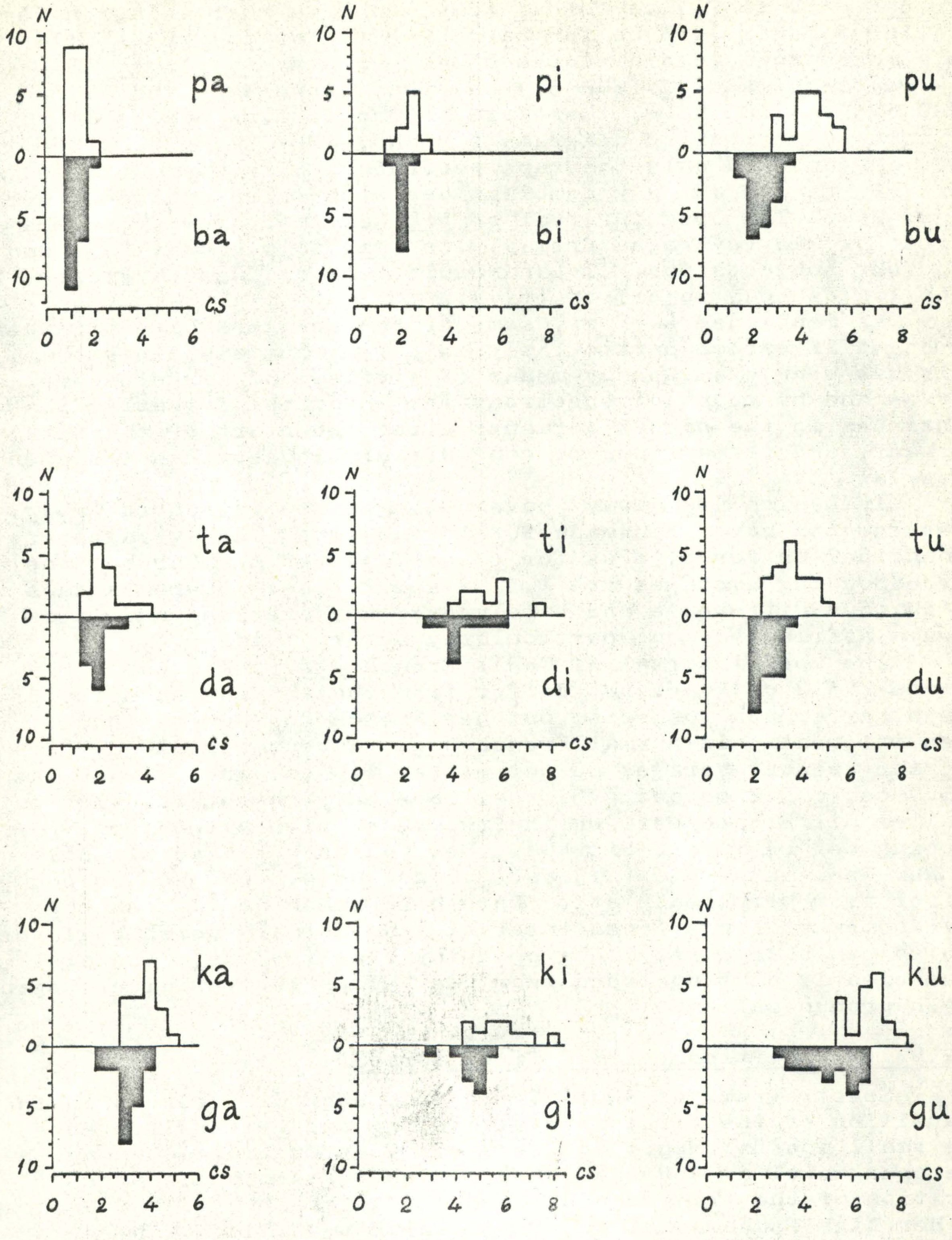


Fig. 4.

Duration of the open interval (CHH, French)

(only tape recordings)

averages of the open interval of ptk thus vary between o.9 cs for pa (MAS) to 7.7 cs for ti (SRO), and as the variation is smaller for bdg (o.5 cs for ba to 2.3 cs for di) there will be a more pronounced difference between e.g. ki and gi (SRO 4.4 cs) or ti and di (SRO 5.4 cs) than between pa and ba (SRO o.5 cs). I have, therefore, found it more practical to give the distribution diagrams for SRO, JPP, EA, JT and CHH (for SRO and CHH only the tape recordings have been utilized, for JPP, EA and JT only combinations with a are included), see Figs. 2-4. The open interval of ptk has, moreover, been measured for MAS (averages from o.9 cs (pa) to 5.4 cs (ku)), and for four subjects from K. Landschultz' recordings (averages from 1.2 cs (pa) to 6.7 cs (ti)).

As mentioned earlier, some of the speakers have a rather strong affrication particularly in ti, di but also in ki. This has been examined by means of airflow and air pressure curves and by means of spectrograms. In di the fricative phase continues in the higher formants after the start of the first formant, and it may thus be considerably longer than the open interval.

In Danish ptk always have a longer aspiration than bdg. Measurements have been made for 1o subjects. The averages are around 6-7 cs for p, 8 cs for t and 7 cs for k. For bdg the corresponding averages are 1.5, 2 and 2.5 cs. There is thus an obvious difference and hardly ever any overlapping. t is always affricated, but particularly before i and y.

The open interval of CHH's French ptk (p 2.6 cs, t 3.9 cs and k 5.2 cs) does not differ from that of the majority of the other French subjects, but her French bdg have a longer open interval than normal French bdg. Except for SRO di (2.4 cs) the French averages do not exceed 1.9 cs, whereas CHH has b 1.9 cs, d 3.2 cs and g 3.8 cs. She has, therefore, also a smaller difference between the two sets, with extensive overlapping before a (see Fig. 4). Her Danish bdg have slightly longer open intervals (2.3 cs, 3.9 cs and 4.8 cs) and do not differ from her French ptk. Both her French and Danish bdg have somewhat longer open interval than normal Danish bdg. Her Danish ptk have a very long open interval (9.7 cs), and are thus clearly distinguished from bdg. This is typical of Copenhagen pronunciation.

2.2.9. Transitions of the first formant.

Spectrograms of SRO, JT and Sch. show differences in the transition of the first formant. For i and u the differences are small and difficult to measure. They are more obvious in the open vowel a. SRO has here a higher start and a shorter duration of the transition after ptk than after bdg, but because of her high fundamental an exact measurement has not been possible. The measurements have thus been restricted to 71 spectrograms with initial stops and 31 with final stops from the recordings of JT, and 2o spectrograms with initial stops from the recordings of Sch., all containing the vowel a.

Before final stops the F_1 transition differs significantly in four respects (the differences are given in parentheses): before ptk the transition stops at a higher frequency (+ 189 cps), the interval is thus smaller (- 214 cps), the duration shorter (- 3.7 cps), and the tempo (measured in cps/cs) slower

(- 16) before ptk than before bdg. After initial stops the same four differences are found: ptk start at a higher frequency (JT +79 cps, Sch. +72 cps), the interval is smaller (JT -59 cps, Sch. -80 cps), the duration is shorter (JT -0.8 cs, Sch. - 1.2 cs), and the tempo is slower (JT and Sch. -4 cps/cs). But all these differences are much less pronounced initially than finally. In addition a fifth difference can be seen initially: the distance from the explosion to the top of the transition is slightly shorter after ptk (JT and Sch. - 0.5 cs).

In Danish spectrograms of stop consonants before the vowel a a rising transition is generally seen after bdg, but only sometimes after ptk, and in the latter case it is normally rather short (except after strongly affricated t), but the distance from the explosion to the top of the transition is not shorter than in bdg.

Velars generally have longer transitions than dentals, and these in turn have longer transitions than labials. No difference can be seen in CHH's recordings. - Table I, p. 80 gives a survey of the results described in section 2.

3. Stability of the differences.

It appears immediately from Table I that French ptk and bdg are distinguished by a large number of cues, at any rate in the position investigated. This position has been chosen because it allows a comparison with Danish. In final position before obstruents (e.g. une robe courte) there will not be much more left than a small difference in the duration of the closure and of the preceding vowel (see Thorsen 1967), and perhaps a difference in the transition (but this has not been investigated).

The stability of the differences seen as the percentage of word pairs distinguished by each difference is shown in Table II. Sonority and closure duration seem to be very stable cues. - But this need not be the same as perceptual relevance.

4. Perceptual relevance of the acoustic cues.

4.1. French stops.

For the French stops no testing has been undertaken by the author, but the problem has been investigated by Marguerite Durand (1956) and by P. Delattre (1968). The main result of M. Durand's experiments is that the tempo and length of the transition and the duration of the closure are more important cues than voicing. But the problem is how she has painted the "voice bar". Voicing in French stops is normally quite strong and not restricted to frequencies around 120 cps, and it is

Table I

Summary of the phonetic differences between <u>ptk</u> and <u>bdg</u> in French and Danish

		N Fr.	French S. excl. CHH ptk-bdg	Danish S. incl. CHH $p^h t^h k^h$-b̥d̥g̊	CHH(Fr.) ptk-ḅḍg̣	CHH Fr.Dan. ptk-ḅḍg̣
A. <u>Physiological</u>						
1. glottis max.open	VA	2	>	>	>	> ?
2. airflow	VA	1	>	>	>	>
3. intra-oral press.peak	V	2+	>	(>) or =	(>)	?
4. intra-oral press.rise	V	2+	>	=	>	>
5. lip pressure	T	3+	>	<	(>)	=
B. <u>Acoustical</u>						
1. voicing	V	5+	< D	(<)	<	<
2. Fo of vowel	V	1	<	=	=	=
3. intensity of explosion	T	1	=(k>g) D	=	=	(<)
4. intensity rise of vowel	T?	3	>iu =a	=	=	=
5. duration of closure	T	5+	> D	<	>	>
6. duration of preceding vowel	T	1+	<			
7. duration of following vowel	T/A	2+	<	<		
8. duration of open interval	A	5+	>iu D >a	>	>iu (>)a	=
9. transition of F1	T/A	3	< (D)	(<)	= ?	

V means voicing, T = tenseness, and A = aspiration. These letters refer to the subsequent discussion. VA means: both V and A, T/A means: either T or A. N(Fr.) indicates the number of French subjects of the present investigation besides the bilingual subject CHH, who is included in the Danish subjects because of her voiceless <u>bdg</u>. + after a figure means that further materials (from Thorsen and K. Landschultz) are taken into account. > means that <u>ptk</u> have a higher degree of the given parameter than <u>bdg</u>. A small > indicates a very small difference. ( ) indicates that the difference is not significant. D means: mentioned by Delattre.

Table II

Stability of the differences between French ptk and bdg

Percentage of pairs distinguished.

	Voicing			Duration of closure period			Duration of open interval		
S	p/b	t/d	k/g	p/b	t/d	k/g	p/b	t/d	k/g
SRO	loo	loo	loo	loo	loo	92	loo	86	loo
JPP	loo	loo	loo	97	96	75	94	85	97
EH	loo	loo	loo	86	93	89	-	93	83
JT	loo	loo	loo	loo	loo	loo	78	74	78
Sch.	loo	loo	loo	89	loo	loo	84	89	loo

	Duration of preceding vowel			Duration of following vowel			Intensity of explosion			Fo of following vowel		
S	p/b	t/d	k/g	p/b	t/d	k/g	p/b	t/d	k/g	p/b	t/d	k/g
SRO	89	76	89	8o	82	loo	52	35	85	loo	loo	97

Transitions of formants (initial position)

S	Frequency Start	Frequency Interval	Duration	Distance from explosion
JT	89	89	78	89
Sch.	89	loo	89	67

Intensity rise of vowel

S	i,u	a
SRO	98	15
JT	loo	47
Sch.	89	33

	Resistance in the glottis			Airflow			Lip Pressure	Intra-oral Pressure Peak		Intra-oral Pressure Rise
S	p/b	t/d	k/g	p/b	t/d	k/g	p/b	p/b	t/d	p/b
SRO	loo	loo	loo				96	loo	loo	83
JPP	96	loo	loo	88	loo	75				
EH								loo	loo	

very clearly audible in normal French bdg. New experiments should therefore be undertaken with French listeners and with precise indication of the frequency and intensity of the "voice bar".

Delattre (1968) sets up 6 acoustic cues which determine the identification of stops as ptk rather than bdg; (a) a longer hold (closure period), (b) shorter preceding vowel, (c) a cutback in the first-formant transition (i.e. later start - or weak start - of F_1 compared to higher formants), (d) a stronger turbulent noise in the explosion, (e) absence of "voice bar", (f) sometimes aspiration. He does not give any ranking of these six factors.

F_1 cutback (c) is shown to be an effective cue for American listeners by Libermann, Delattre, Cooper (1958), but the authors identify this cue with aspiration (f), and they find that the cue is more effective when the beginning of the higher formants is noisy and suggest that the pattern playback may give good results without this noise because of its general background noise. This sounds probable. F_1 cutback without noise in the higher formants is an unrealistic cue. Natural speech has never earlier start of higher voiced formants.

Delattre correlates the F_1 cutback with "the unusual degree of pressure that prevails as the organs come into contact" (1968, p. 214). Probably he thinks of air pressure, but the correlation does not seem convincing. In the earlier paper (1958) the authors quoted Fant, who has said that F_1 is weakened by the large resonance chamber below the open glottis, which seems to be a better explanation. Delattre correlates "aspiration" with a delay in the vibration of the vocal cords, and as this is also due to the open glottis, we come again to the conclusion that F_1 cutback and aspiration are not two different cues.

If F_1 cutback (c) is left out, the five remaining cues (a b d e f) correspond to the acoustic differences Nos. B 5, 6, 3, 1 and 8 of Table I. As for B 2 (fundamental frequency start of the following vowel) Fujimura (1959) has found it to be a very effective cue for one Japanese listener and to have some effect for five American listeners. It might thus be worth while to try it out on French listeners. Intensity rise and duration of the following vowel have, as far as I know, not been tested.

As for point 9 of Table I (F_1 transition), I do not think that this is covered by "F_1 cutback" or "aspiration". I shall return to this problem in section 5.1.

4.2. Danish stops.

The number of cues is more restricted for Danish stops, but the difference of aspiration is so stable and so clear and audible that it is quite sufficient. Danes are therefore inclined to identify unaspirated stops of other languages (e.g. French, Dutch, Hindi) as bdg and to hear aspirated bdg in Hindi as ptk. (A test showing this reaction has been undertaken by the author.)

The importance of the aspiration has also been shown by a tape-cutting experiment using Danish words with initial ptkbdg followed by the vowels a, i and u (1 to 3 examples of each combination). 21 Danish listeners (all phoneticians or dialectologists) were asked to identify the word as an existing or (in some rare cases) possible Danish word. When the explosion was removed from words with initial ptk, they were still correctly identified as ptk in 95 per cent of the cases (with the exception of one word with pu heard as fu), and there were only 4 bdg-answers out of 378 (1 per cent). When, on the other hand, the aspiration was removed, there were 86 per cent bdg-answers and only 7 per cent ptk-answers. When the explosions in p and b, t and d, k and g were interchanged, there was no change in the perception. The explosion is thus of very little importance in this respect.

Aspiration is partly a question of distance from the explosion to the start of the vowel (open interval), partly a question of noise during the interval. For the Danish affricated t the noise is very important. If this noise is replaced by a pause of 6-7 cs, the t is only identified in 10 per cent of the cases; it is heard as p in 15 per cent and as d or b in 27 per cent of the cases. - For p and k an explosion followed by a pause of 6-7 cs before the start of the vowel is still identified as p-k in 5o per cent of the cases (with 26 per cent b-g-answers). - If the explosion of g and b is placed at a distance of 6-7 cs from the vowel, the most common answer is b or g (62 per cent), whereas only 18 per cent of the examples were heard as p-k. This is probably not a question of the explosion but of the transitions of the vowel, for a k-explosion placed at this distance before -ilə (cut from gilə) gives 86 per cent answers in favour of g and none in favour of k, and a g-explosion at 6 cs from a u originally preceded by k gives 34 per cent answers in favour of k and 38 per cent in favour of g (in these cases only two different words were used).

The opposite experiment, i.e. the introduction of aspiration noise in the brief interval between bdg-explosions and the vowel, has not been tried, but cutting from the beginning in words with initial ptk or fsh gives the result that there is a limit somewhere between 4 and 2 cs, where the majority of the listeners start hearing bdg instead of ptk. This means that a short noise interval is not sufficient to provoke ptk-answers.

4.3. CHH's stops.

In CHH's pronunciation of French the number of acoustic cues is strongly restricted compared to normal French. The length of the preceding vowel is not relevant for initial stops, and the differences in initial voicing and in the duration of the closure and of the open interval are very small, and they are still smaller when French ptk and Danish bdg are compared. One may, therefore, ask whether it is at all possible to distinguish these types of stops. To check this a test was set up consisting of "words" of the type apa, aba, etu etc. cut out of one of the CHH's recordings of French and Danish words (French: la balle, la panne, les tours, etc., Danish: [a panən, a baljən] etc.). There were two examples with each consonant before a, i and u from each language, i.e. in total 36 French and 36 Danish "words". There was no audible difference between the French and Danish vowels. The words were played back in random order and repeated three times each. In test No. 1 all words were mixed. Test No. 2 contained only French words (and there was a forced choice between French ptk and bdg), test No. 3 contained only French ptk and Danish bdg (and the choice was between these consonants), test No. 4 contained only French words, but given in pairs apa - aba, etc.

Unfortunately only a small number of listeners have up till now listened to the tapes: 4 Danish phoneticians who know French well and who also know the pronunciation of CHH, CHH herself, and (for tests 2 and 3) two French listeners (teachers of French, but with no phonetic training). The different groups of listeners agreed to a very large extent, but the French listeners were somewhat less successful than the others.

On the whole the result was positive. When all four types were mixed, both Danish and French ptk were identified correctly in 78 to 100 per cent of the cases, but French bdg were very often heard as Danish bdg, and Danish bdg as French ptk. In tests 2, 3 and 4 the percentages of correct identification were the following:

Table III

Listeners	Test 2		Test 4	Test 3	
	French ptk	French bdg	French ptk/bdg	French ptk	Danish bdg
CHH	100%	83%	94%	94%	67%
Danish	92%	79%	90%	93%	60%
French	89%	61%	75%		

It appears from the table that CHH's Danish bdg are often heard as French ptk (they have a somewhat longer open interval than normal Danish bdg). Apart from this the identification is good. It is interesting to note that the identification of words presented in pairs is not very much superior to the identification of words in random order. This is in agreement with Libermann (1957).

As the differences in initial voicing and the duration of the closure seemed to be the most stable cues, it was tried to remove the preceding vowel and the beginning of the closure, and these new tests 2a and 3a were presented to the same listeners:

Table IV

Listeners	Test 2a		Test 3a	
	French ptk	French bdg	French ptk	Danish bdg
CHH	78%	89%	72%	83%
Danish	83%	74%	88%	67%
French	67%	67%		

The main difference from tests 2 and 3 is that more bdg have been identified correctly and that some ptk have been heard as bdg. Probably the long duration of the closure in French apa ata etc. in tests 2 and 3 has favoured ptk-responses. In tests 2a and 3a only the explosion and the open interval were left as cues, but the identification is still quite good. A detailed analysis of the single cases has shown that the intensity of the explosion does not play any role, whereas the duration of the open interval seems to be important. Most of the mistakes concern examples of bdg with relatively long open interval and ptk with relatively short open interval. There is, however, a number of cases which cannot be explained in this way, e.g. ta-da. One of the examples of French ta having 1.5 cs open interval is heard as ta by the majority, whereas the other, with 2.o cs open interval, is heard as da by the majority, and Danish da (2.5 cs) is heard as da in all cases.

It is possible that oscillograms taken at great speed might reveal small differences in the explosions or in the noise level of the open interval, which do not appear clearly in intensity curves taken with an integration time of 2.5 ms.

At any rate the large overlapping between the durations of the open interval before a, which is found in CHH's recordings as well as in those of the other French subjects (see Figs. 2-4) makes it necessary to look for other cues.

In these voiceless stops "open interval" is synonymous with "voicing lag". And the importance of this cue has been demonstrated by Liberman, Delattre, Cooper (1958) for American listeners and by Abramson-Lisker (1965) and Lisker-Abramson (1967) for Spanish and Thai listeners. The fact that the Spanish listeners have their crossing point to the right of the limit in their natural speech is partly paralleled by the French listeners in the present test, at any rate for pa/ba, but both the Spanish and the French listeners have probably been familiar with languages having voiceless bdg initially (Danish and American English), and this may have influenced their judgements.

5. General problems.

5.1. Aspiration and tenseness.

5.1.1. As mentioned above aspiration and tenseness have been combined into one feature by Jakobson-Fant-Halle (1952). On the basis of the results of the measurements of French and Danish stops this does not seem advisable. One would e.g. have to say that tense consonants are characterized (a) by a longer closure period (in French) or by a shorter closure period (in Danish), (b) by higher lip pressure (in French) or by lower lip pressure (in Danish), (c) by a higher intra-oral pressure (in French) or not by this difference (Danish), (d) perhaps by a more abrupt opening of the closure (French)*) or by a relatively slow opening (Danish)**), etc. Item (c) may be due to the combination with voicing in French, but the other cues are really contradictory. Tenseness and aspiration must, therefore, be separated as two different features. Da-

---

*) if the interpretation of the transitions of F_1 is correct.

**) obvious for the affricated t, dubious for p/b.

nish stops are distinguished by aspiration, French stops by a combination of tenseness and voicing. In Chomsky-Halle (1968, pp. 324-326) aspiration and tenseness have also been separated as two different features. They suppose that aspirated stops have heightened subglottal pressure. This may be possible, but it has not been proved. It seems to have been inferred from differences in supraglottal pressure. When the glottis is wide open as in Danish ptk there will probably not be much difference between supraglottal and subglottal pressure. But in the cases where the vocal cords are closer together, it is more difficult to draw any conclusions. In the many cases where Danish b has the same supraglottal pressure as p one may well conclude that the subglottal pressure can hardly be higher than in p (for otherwise the vocal cords would start vibrating), but one cannot know whether it is lower than in p. If it is lower bdg may be said to be less tense as far as the expiratory muscles are concerned, but as noticed by Chomsky-Halle, the articulatory mechanisms for subglottal and supraglottal tension are different and independent, and the two types must be kept apart.

The use of the term tenseness (or fortis-lenis) as restricted to the supraglottal cavities is in accordance with Rousselot (1897, p. 583) and Straka (1963, pp. 60-61), whereas Delattre uses the term in a wider sense.

The description given by Fant (1960, pp. 224 and 279) of tense and lax stops is evidently a description of aspirated and unaspirated stops. He underlines the importance of an open glottis for the airflow, and of a slow opening of the constriction which causes a longer noise interval.

5.1.2. Duration of following vowel. In some respects aspiration and tenseness may produce similar results. One of these is the duration of the following vowel. Shortness of the vowel after ptk may simply be due to the aspiration which delays the vibration of the vocal cords (e.g. in Danish and English), or it may perhaps be a compensation of effort (e.g. in French, where the difference in vowel length is greater than the difference in open interval, whereas it is the opposite in Danish).

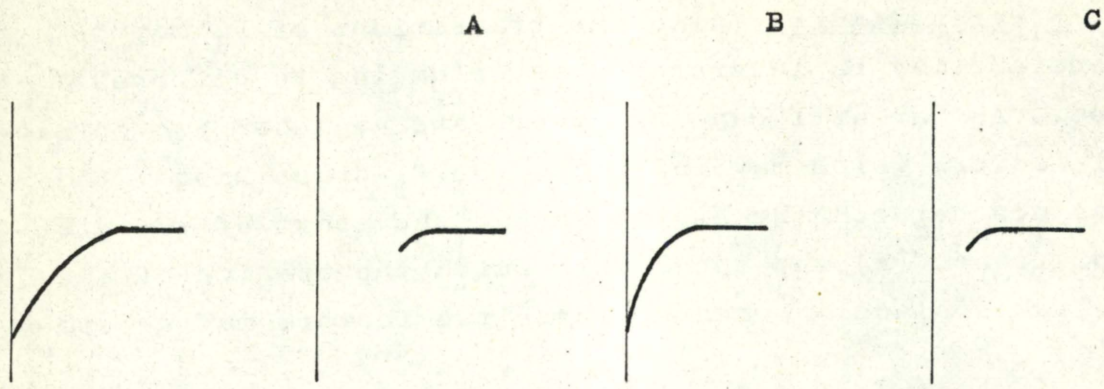
But a combination of tenseness and moderate aspiration is probably not excluded. It may be true of British English.

5.1.3. F_1-transitions. Also the transitions of F_1 may be ambiguous and may be interpreted as belonging to different features. As far as I can see, there are at least two physiological factors which may influence the F_1-transitions: (A) the distance between the explosion and the start of the glottal vibrations, (B) the speed with which the opening of the closure takes place. Moreover these two factors may be combined (C).

In Fig. 5 a sketch is given of the possible acoustic consequences of these two physiological factors. In case A, where there is a longer distance from the explosion to the start of the glottal vibrations (a longer "voicing lag") in p than in b, the F_1-transition after p will display (a) a shorter duration, (b) a higher start (and consequently a shorter frequency interval - this is not an independent factor); but there will not be any difference in the distance to the top of the transition (c) nor any difference in the tempo of the transition (d). In case B, where a quicker movement of the speech organs is assumed for p, there will be (a) a shorter duration of the transition, (c) a shorter distance to the top of the transition, (d) a quicker tempo of the transition, but not a higher start (b).

The experiments with F_1-cutback correspond to possibility A. This situation is found in languages with aspirated ptk, e.g. Danish, where the aspiration is probably the only cause of the difference in F_1-transition. M. Durand's experiments correspond to possibility B. This situation might be found in a language without any difference in open interval between ptk and bdg, but with fortis (or tense) articulation of ptk.

Possibility C corresponds approximately to what was found in French (see section 2.2.1o). This will give (a) a shorter duration, (b) a higher start (and consequently a smaller frequency interval), (c) a shorter distance to the top of the transition, and (d) in the case of straight transitions a quicker tempo of the transition after p than after b;



b	aspirated p	unaspirated p	slightly aspirated p
	movement = b	quicker movement	quicker movement

F_1 transition for p:			
a) shorter duration	+	+	+
b) higher start	+	0	+
c) shorter distance from the explosion to the top of the transition	0	+	+
d) quicker tempo of transition	0	+	−(+)

Fig. 5. Sketch of possible F_1 transition

but as the transitions are normally curved, only the upper flat end of the curve will remain, and the tempo of the remaining part of the transition may thus be slower than that of the whole transition after b.

The difference in distance to the top is very small in the French examples; on the other hand the difference in open interval is also very small, and the high start of ptk is therefore probably due to a very quick movement of the speech organs in the first centisecond after the explosion. Fujimura (1961), who has filmed the movement of lip opening in p and b spoken by an American subject, has found that the lips may move very quickly during the very first centisecond.

In final position we have the same situation, since both the cessation of the vibrations and the quick movement of the organs influence the transitions.

There may also be cases where a long aspiration conceals the transitions completely so that nothing can be said about the movement of the speech organs. This is often the case in Danish.

The obvious differences found finally in French are in good agreement with the observations of P. Simon (1967), who has found that X-ray films point to a more energetic closing movement of French p than of b, and also with the electro-myographic analysis carried out by Öhman, Leanderson and Persson (1966), showing a stronger muscular activity at the implosion of p than at the implosion of b (cp. also Harris et al. (1965)). On the other hand, the differences found initially in the French spectrograms and the results of M. Durand's perceptual tests are in conflict with P. Simon's observation that the opening movement is slower in p than in b, and with the corresponding findings of Öhman et al. for initial p, which shows less muscular activity than b. To this it must be added that Fujimura has found a more rapid movement for p only initially in an isolated word, not after unstressed ə; but it appears from Fujimura's information about the open interval after p that the examples after ə were spoken with aspiration (4.7, 5.2, and 5.4 cs), whereas the examples in initial position were practically unaspirated (1.2, 1.9 and 3.5 cs). Öhman's speaker was apparently Swedish, and has also had aspirated p, and both Rousselot (1897) and M. Durand (1956) have observed a slower opening movement in aspirated stops. The only cases contradicting the results obtained for JT and Sch. are thus the observations of P. Simon. It is possible that the number of frames (5o per second) has not been sufficiently high to allow quite precise observations of these very quick movements.

One might perhaps think of other factors influencing the F_1-transitions, e.g. the difference in the size of the pharyngeal cavity, ptk apparently having a smaller pharyngeal cavity than voiced bdg (see later section 5.3.3.). According to Gunnar Fant, this could, however, only cause a difference of about lo per cent in the starting frequency of F_1 after ptk and bdg. The tension of the cavity walls might also be assumed to have some influence, but too little is known about this factor.

It thus seems that similar, though not in all respects identical differences in F_1-transitions may be ascribed to two different features: (1) aspiration, (2) tenseness, reflected in the rapidity of the movements of the speech organs.

5.2. Voicing and aspiration.

Lisker and Abramson (1964, cp. also Abramson-Lisker 1965) have proposed to combine voicing and aspiration into one feature: voice onset timing (VOT), the important thing being the timing between voice onset and explosion. In voiced stops the voicing starts well before the explosion (they have "voicing lead"), in unaspirated voiceless consonants the voicing starts at the explosion or immediately afterwards, and in aspirated consonants voicing starts well after the explosion (they have "voicing lag").

This is a possible solution for a good number of languages, and it is a very simple solution (but it can, of course, not be combined with the binary feature theory). Some cases, however, make difficulties: (a) in some languages bdg are voiceless in initial position after a pause and after voiceless sounds, but often partly voiced after voiced sounds (e.g. North German, English, and, sometimes, French). Probably Lisker-Abramson would consider this voicing as an example of weak "edge vibrations", which do not count. But the voicing may be quite strong, and there is often free variation with full voicing. This type is evidently different from unaspirated ptk. As the voicing is always in the first part of the closure in this type of voicing, it cannot be a case of voicing lead, but it might perhaps be fitted into the scheme as a more complicated fourth type. (b) More problematic is the case of the Indian aspirated mediae. Lisker-Abramson consider these as having breathy voice or murmur, not ordinary voicing (1964, pp. 4o3 and 419). Ladefoged is of the same opinion (1967, pp. 1o and 74 ff.), but he still considers voicing and aspiration as belonging to two different dimensions: (a) glottal constriction (including murmur) and (b) glottal timing. In any case this type cannot be fitted into Lisker-Abramson's VOT-dimension.

5.3. Voicing and tenseness.

5.3.1. Lisker and Abramson also want to get rid of tenseness. I do not think this is possible. One may mention (a) CHH's distinction between ptk and bdg in French, which before open vowels is based on the duration of the closure period, and on the very small difference in initial voicing, and only before close vowels also on the open interval ("voicing lag"). (b) The case of ptk in Swiss German, which differ from bdg mainly by the duration of the closure period and the organic pressure (bdg may be weakly voiced, but they may also be voiceless). (c) The Korean stops (Kim 1965), which are also difficult to describe by means of voicing and aspiration only.

I am, however, inclined to consider the intra-oral pressure as belonging to the voicing feature and not to the tenseness feature, although it has often been considered as the most important aspect of tenseness.

5.3.2. Intra-oral pressure and organic pressure. It is often assumed that there is an intimate relation between the intra-oral air pressure and the organic pressure at the point of articulation: the higher the air pressure, the stronger must the organic pressure be in order to maintain the closure (e.g. Otto Jespersen (1914), and more recently Malécot (1956 and 1968) and Delattre (1965)). As the intra-oral air pressure might also influence the airflow, and thus indirectly the intensity of the explosion, it must be considered as a principal factor of the tenseness feature.

This hypothesis sounds very plausible, but it is contradicted by various findings of the present investigation and also by the results of other investigations. It is true that e.g. French stops are characterized both by a higher intra-oral pressure and by a higher lip pressure, but this need not be a mechanical dependency.

Firstly, there are some languages which have a difference in air pressure between ptk and bdg, but no clear difference in lip pressure, see e.g. for American English the electro-myographic investigations of Harris, Lysaught, Schwey (1965) and the measurements of mechanical pressure made by Malécot (1966 b). However, in this case the differences in air pressure are not too clear either.

Moreover, in Danish b has generally slightly lower air pressure, but higher lip pressure than p.

This inverse relation is also found in Gujarati, which has four types of stops. In 16 word series, pronounced with alternating order of the words (subject RD), ph had higher air pressure than p in 13 cases, and bh higher air pressure than b in 16 cases, but the aspirated consonants had lower lip pressure than the unaspirated ones (see Table V A 3-4 and Fig. 6).

Of particular interest is the behaviour of the nasal consonants. These have practically always a much lower intra-oral air pressure than the oral stops, but the lip pressure of m is almost the same as that of p or b. This is seen in the above-mentioned American investigations, and I have found these relations often in recordings from various languages.

It is also interesting that air pressure maximum and lip pressure maximum do not coincide in time, the maximum lip pressure occurring normally in the first half of the closure and the air pressure maximum at the end of the closure period. Table VI,A contains the average values for the two main subjects of this investigation. For SRO the two pressure curves were recorded simultaneously, for CHH the averages are from two different recordings. When the lip pressure is at its maximum, the

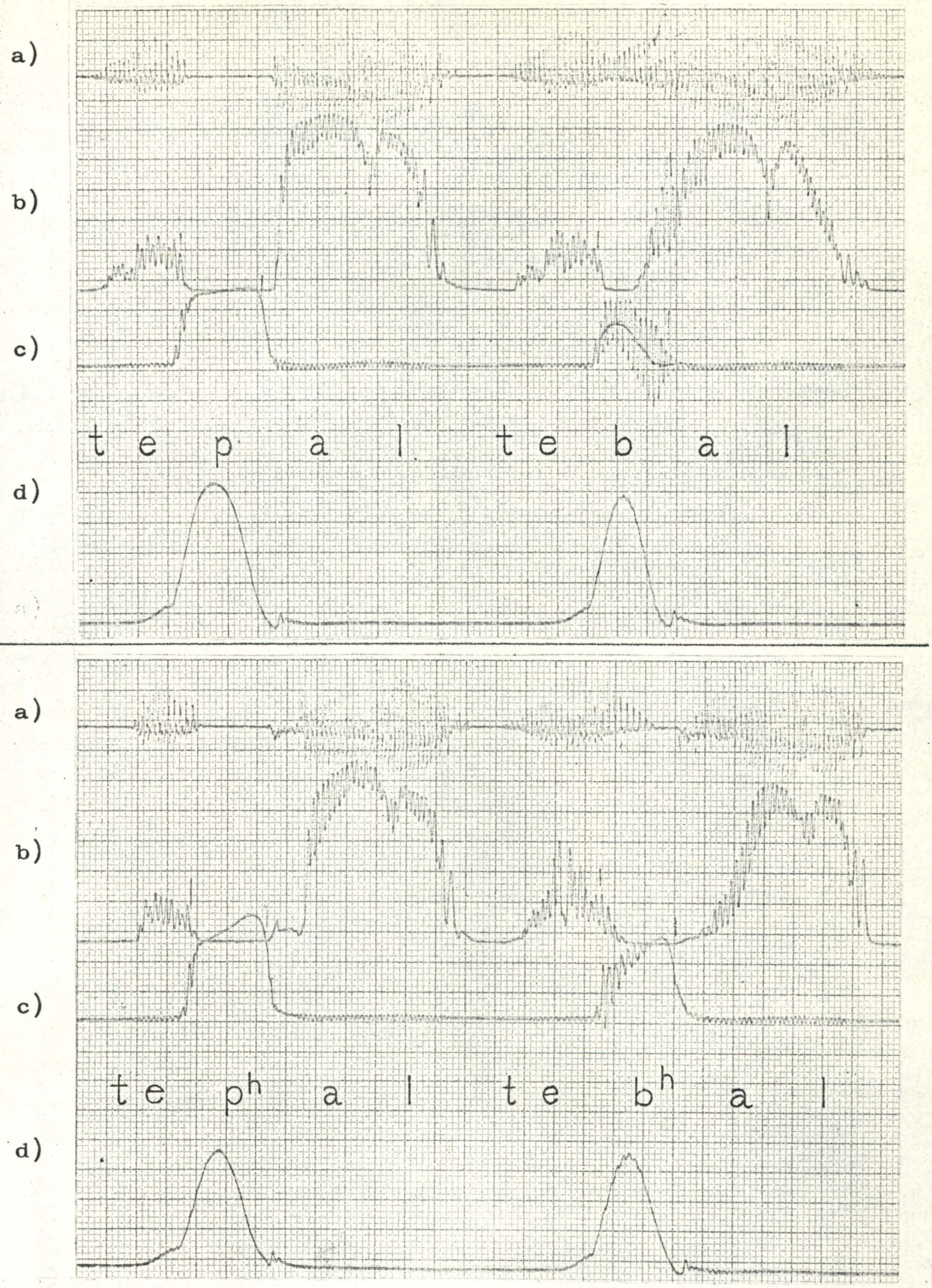


Fig. 6.

Mingograms of labial stops in Gujarati (RD)

a) duplex oscillogram
b) intensity curve
c) intra-oral air pressure
d) lip pressure

Table V

Relations between p ph b and bh in 16 word series pronounced in alternating order (Gujarati).

A.

1	Degree of voicing	ph-p $\overset{(16)}{<}$ bh $\overset{(12)}{<}$ b
2	Intra-oral air pressure	ph $\overset{(13)}{>}$ p $\overset{(15)}{>}$ bh $\overset{(16)}{>}$ b
3	Lip pressure	p $\overset{(11)}{>}$ ph $\overset{(12)}{>}$ b $\overset{(14)}{>}$ bh
4	Duration of closure period	p-ph $\overset{(12)}{>}$ b-bh

B.

	Intensity of voicing		Intra-oral air pressure	
	b	bh	b	bh
rising	16	0	0	16
falling	0	16	16	0

Table VI

Time relations between intra-oral air pressure and lip pressure.

A. Distances

	N	SRO (list B) p 38	SRO (list B) b 38	CHH p 36/27	CHH b 36/27
1. Distance from the implosion to the lip pressure maximum		6.9	6.o	7.3	5.9
Distance in percentage of the duration		34%	38%	41%	37%
2. Distance from the implosion to the air pressure maximum		18.3	13.o	17.9	15.1
Distance in percentage of the duration		95%	82%	1oo%	1oo%

B.

	N	p 38	b 38
Intra-oral air pressure measured at the point of maximum lip pressure, and indicated in percentage of the maximum air pressure		79%	68%

C. Air pressure and lip pressure in percentage of the corresponding maximum pressures, measured at 7o% of the distance from implosion to explosion.

	SRO pa	SRO pi	SRO pu	SRO ba	SRO bi	SRO bu	CHH pa	CHH pi	CHH pu	CHH ba	CHH bi	CHH bu
Air pressure (rising)	92%	93%	87%	8o%	83%	8o%	98%	96%	95%	95%	95%	97%
Lip pressure (falling)	74%	72%	62%	71%	71%	61%	72%	56%	43%	63%	42%	2o%

air pressure has only reached 79 per cent (p) or 68 per cent (b) of its maximum value. The fact that the lip pressure maximum is reached earlier than the air pressure maximum is not in itself sufficient to contradict the hypothesis, but the crucial point is that it decreases rapidly again, while the intra-oral air pressure is still rising. Table VI,C gives the values for the two pressures at an arbitrary point of the closure period (at 70 per cent distance from the implosion). At this point the lip pressure has decreased considerably, particularly before u, and particularly in CHH's recordings, which are more typical than those of SRO (the rubber bulb used for SRO had somewhat more inertia, and her curves do not always reach zero at the explosion). Similar relations have been found in curves of Danish, German and Gujarati. The lip pressure seems to increase and decrease independently of the air pressure (see also Figs. 6 and 8a).*)

From all these facts one must draw the conclusion that intra-oral air pressure and lip pressure are independent factors. And this has two consequences: (1) that lip pressure, when it is not a mechanical result of the air pressure, must be an independent factor of tenseness, (2) that air pressure need not belong with this feature.

5.3.3. Intra-oral air pressure and voicing. The intra-oral air pressure seems to be much more intimately related to voicing.

A comparison between different languages shows that in languages where bdg are voiced, these have a much lower intra-oral air pressure than ptk (e.g. French), whereas in languages where they are voiceless the difference is very small (e.g. Danish and the French spoken by CHH). Also within the same

---

*) In a discussion at the Technical High School in Stockholm, where I gave a lecture on these problems in February 1969, Sven Öhman and J. Liljencrantz objected that the mechanical lip pressure does not only depend on the muscular activity of the lips, but also on the air pressure , and the early decrease in lip pressure might be due to the air pressure. This is possible, but a glance at Fig. 6 shows that the lip pressure decreases in exactly the same way in voiced b as in the other consonants, although the air pressure is very low and reaches zero before the explosion.

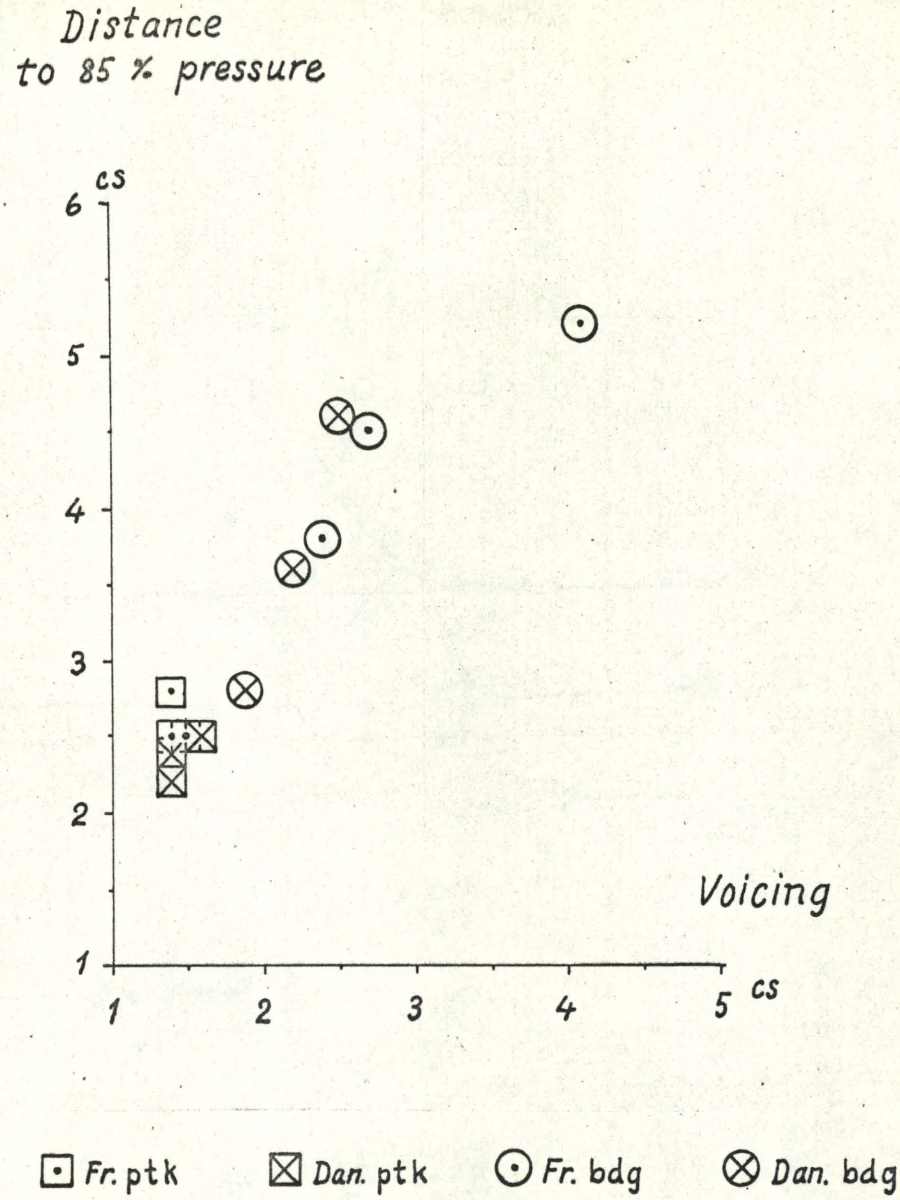


Fig. 7.
Correlation between voicing and intra-oral air pressure

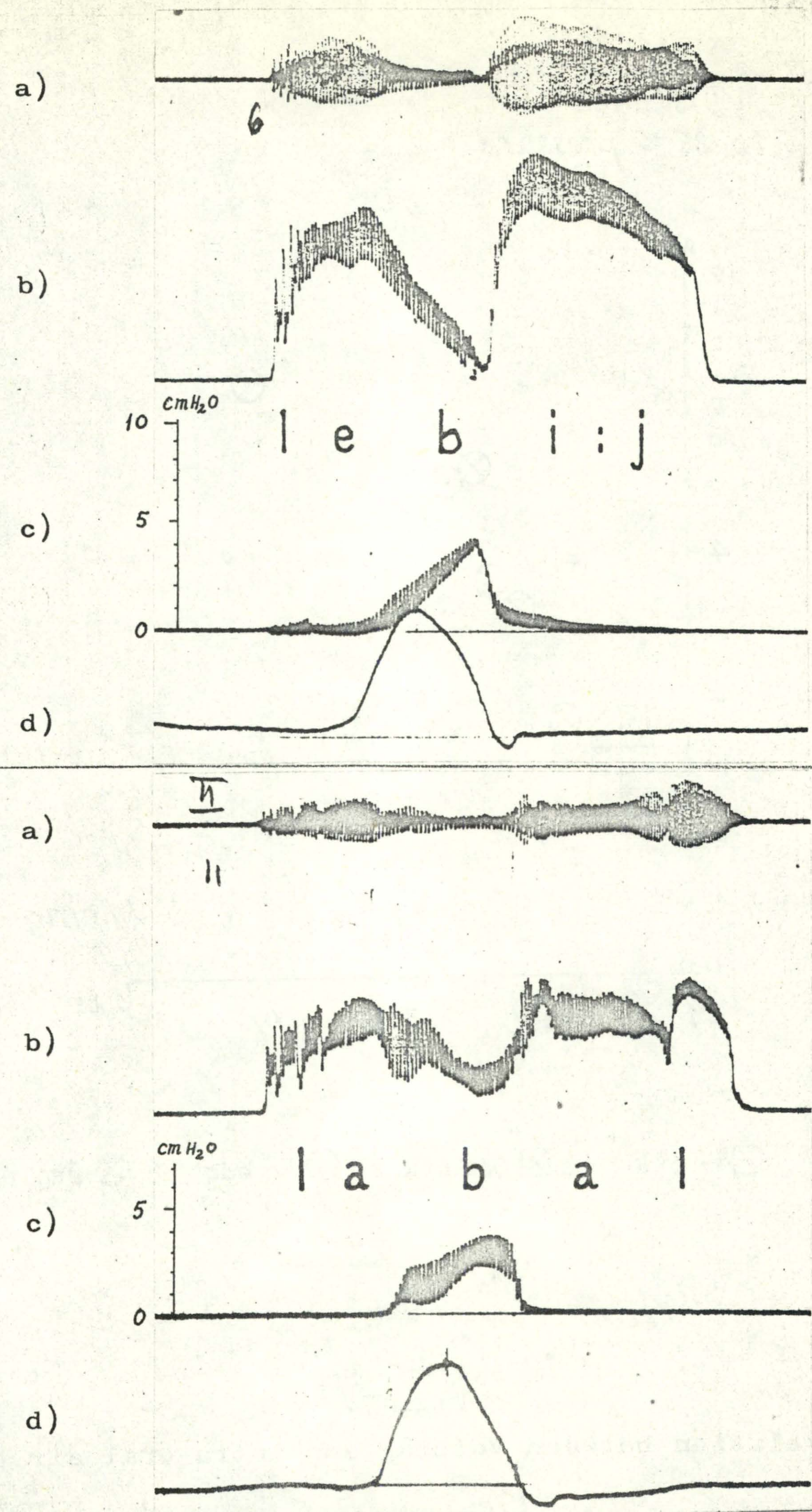


Fig. 8 a.

Mingograms showing the relation between voicing and intra-oral air pressure (SRO)

a) duplex oscillogram
b) intensity curve
c) intra-oral pressure
d) lip pressure

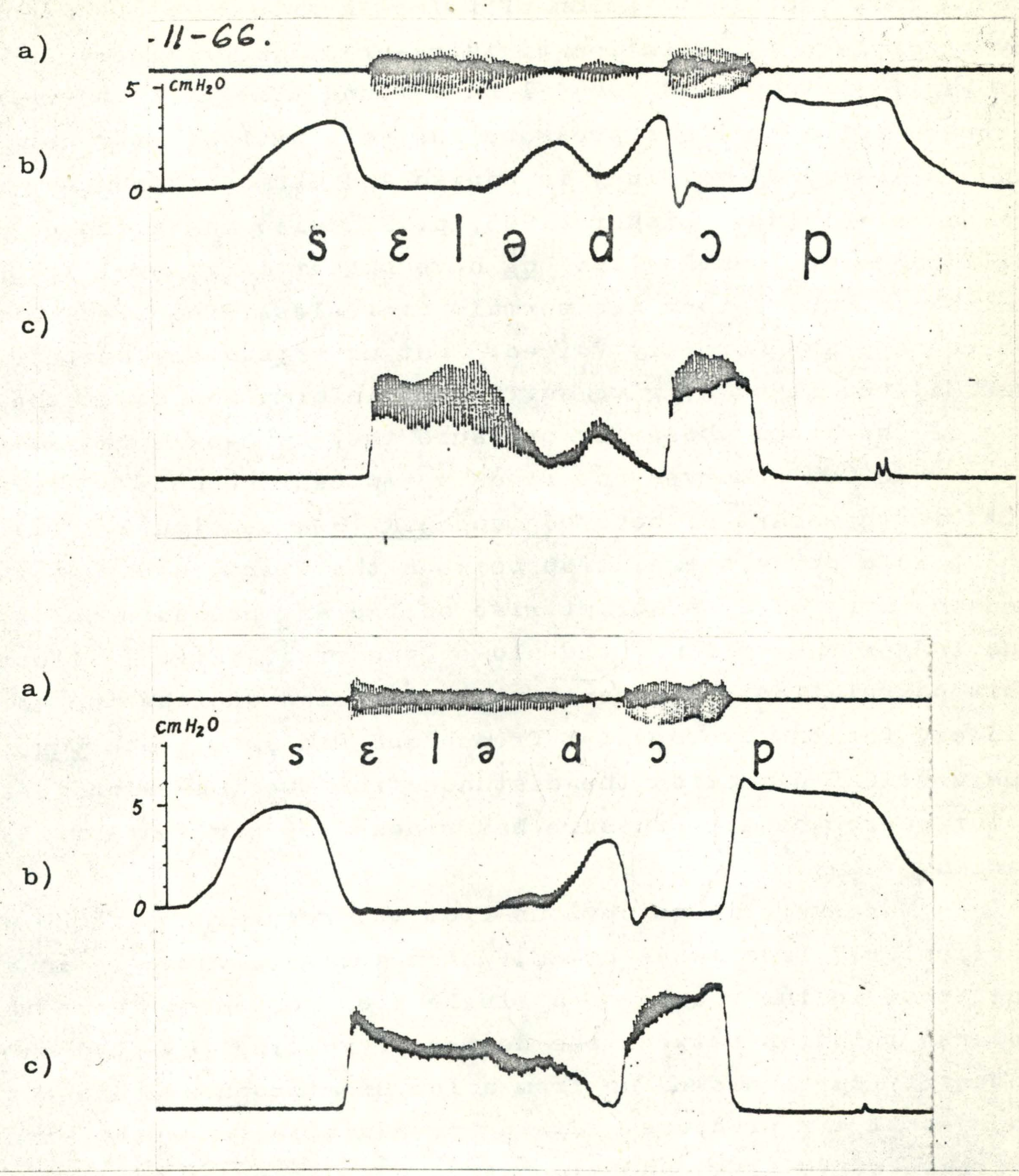


Fig. 8 b.

Mingograms showing the relation between voicing and intra-oral air pressure (EH)

a) duplex oscillogram
b) intra-oral air pressure
c) intensity curve

language this relation can be observed. O. Thorsen (1967) has found that the assimilation of voice in French is accompanied by a complete assimilation of intra-oral air pressure, and in an earlier study (EFJ 1964) I have found a negative correlation between intra-oral air pressure and voicing (both for duration and intensity of voicing) in Danish and German consonants with variable voicing. Lisker (1965, p. 12) also makes the observation that American English bdg have higher intra-oral pressure initially, where they are normally voiceless, than medially, where they are normally voiced. But unfortunately neither he nor Malécot (1966 a) give sufficient information about the voicing of the stops whose air pressure they have examined.

In CHH's curves one finds a few centiseconds of voicing just in the start of both bdg and ptk (see section 2.2.1.), and there is a close correlation between the averages of voicing and the more or less abrupt rise of the air pressure curve: the longer the voicing, the slower the rise. Fig. 7 illustrates this correlation. The horizontal dimension depicts the voicing (in cs) for the averages of French and Danish ptk and bdg, and the vertical dimension the distance from the implosion to the point where the air pressure has reached 85 per cent of its maximum value.

Moreover, it is obvious from the recordings of SRO and particularly from those of EH, which are more variable in voicing, that in the course of a single stop consonant there is an inverse relation between the degree of voicing (as seen on the intensity curve picked up from a larynx microphone) and the intra-oral air pressure. Rising pressure is connected with a decrease of voicing, and vice versa. A number of mingograms demonstrating this inverse relation are shown in Fig. 8. The curves form almost perfectly symmetrical figures.

Finally the same inverse relation is found in the curves of Gujarati. It appears from Table I,A,1-2 that the voiced stops have a lower air pressure than the voiceless stops, and that bh has less voicing but a higher air pressure than b. Moreover, Table VI,B shows that in all cases b has increasing voicing and falling air pressure, whereas bh has decreasing voicing and rising air pressure (see also Fig. 6).

The intimate relation between voicing and air pressure is thus quite obvious. The problem is what is cause and what is effect.

Abramson-Lisker try to derive all the differences from the position of the glottis and the timing relations between glottis movement and oral closure. This is certainly very important, but the mechanism seems to be somewhat more complicated.

(A) It is evident that if the glottis is wide open it cannot vibrate, and the pressure during the oral closure will be high. And if the closure is released before the glottis is closed the stops will be aspirated.

But perhaps "closure lead" would be a better terminology than "voicing lag". For, according to Rothenberg (1968, p. 74) (see also Ventsov 1966), the opening and closing movement of the glottis requires a certain time (the minimum being 8-1o cs, and the average somewhat longer), whereas the oral closure duration can be varied more freely. This view can be supported by the observation that Danish aspirated stops have very short closure periods, often 6-9 cs in stressed syllables in running speech, and the strongly aspirated t has a shorter closure period than p and k. CHH's glottograms in Fig. 1 also show that, at any rate for k, the difference between her French and Danish stops is rather a difference in the duration of the closure than in that of the glottal movement.

(The Gujarati speaker RD has no consistent difference in closure duration between ph and p, although the tendency toward shorter closure in ph is obvious.)

However, if the glottis is relatively narrow, which is probably often the case in unaspirated ptk e.g. in French or Russian, the movement can hardly be assumed to require the same time as when it is wide open, and the reasoning of Rothenberg may not hold.

(B) If, on the other hand, the glottis is narrow, voicing is not excluded, but apart from a narrower glottis it requires a subglottal overpressure.

In Danish bdg the glottis seems to be relatively open at the implosion, and the supraglottal pressure is built up in the course of 2-3 cs at the same rate as the supraglottal pressure in ptk. Therefore, although the glottis becomes more closed during the oral closure period, the consonant will remain voiceless. It is the degree of opening of the glottis during the implosion phase which is important.

Even if there is only a slight opening in the start of bdg, it will not take many centiseconds before the supraglottal pressure becomes too high to allow voicing, if no other mechanism intervenes. This development is accelerated by the effect of the rising supraglottal pressure on the vibratory mode of the vocal cords: the opening phase becomes longer and more air comes through (see Fant 1960, pp. 266-267 and Halle-Stevens 1967, p. 268).

If the supraglottal pressure is drastically decreased, the vocal cords will immediately start vibrating. This has been shown by Ventsov (1966) by means of the following set-up: The subject speaks into a wide tube, which is closed at the opposite end. During the closure of a stop consonant the experimenter can open the tube by means of a valve, so that the air escapes. Under these conditions Russian p becomes immediately voiced, and Ventsov draws the conclusion that the vocal cords must be relatively close. During a stay in Leningrad I had the opportunity to use the instrument, and it turned out that Danish b becomes voiced when the valve is opened, whereas Danish p remains voiceless, except when the opening takes place in the beginning of the closure period. This confirms the observations of the different degree of glottal opening in Danish p and b.

In normal speech the supraglottal pressure can be reduced by an enlargement of the cavities. This mechanism has been observed for Bengali voiced stops by Senn (1935), for Dutch by Slis (1967), for English by Perkell (1965 a and 1965 b), and it can also be observed in the X-ray pictures of Russian stops by H. Koneszna and Zawadowski (1956). The enlargement is normally achieved by an expansion of the pharynx or a lowering of the larynx. This allows voicing to continue in bdg. If the supraglottal air pressure is very much reduced, the stop may become implosive. In Gujarati implosive stops occur as free variants of voiced stops.

Perkell (1965) and Chomsky-Halle (1968) consider the difference in supraglottal cavity size between ptk and bdg as an effect of tenseness: in ptk the walls are tense and stiff and cannot be expanded, in bdg they are lax and are expanded passively by the air pressure. Rothenberg (1968) does not consider a passive expansion as sufficient for a longer continuation of voicing, and assumes that there is also an active expansion, and moreover there is the possibility of letting some air escape through a leakage in the velic closure. I am inclined to think that Rothenberg is right. Danish bdg do not display any expansion, but they are not tense consonants. I think the expansion is due to muscular activity and that it has the purpose of making voicing possible, i.e. that it is part of the mechanism of voicing. The reduction of the pressure is a means to an end. It is also difficult to understand why Perkell considers the widening of the pharynx in bdg as passive and typical of lax consonants, whereas the widening of the pharynx in i versus I is considered as a feature of the tense vowel (1965 a).

5.3.4. Duration. The longer duration of the closure period in e.g. French ptk is normally considered as belonging to the tenseness feature, and the shortening of the preceding vowel as a consequence of the force of the consonant (e.g. Delattre 1939, 1940, 1941). This seems plausible, since the difference is particularly stable in languages which have a clear difference in organic pressure too, e.g. French, Swiss German, and Korean. Cp. also that Danish b has both stronger lip pressure and longer closure duration than p. Rothenberg, on the other hand, considers the length of the consonant closure as a means to avoid aspiration (1968, pp. 83-84). But, as observed above, this argument presupposes the assumption of a widely open glottis for these consonants, and this can hardly be assumed.

Halle-Stevens (1967, p. 269) and Chomsky-Halle (1968, p. 301) explain the longer vowel before voiced stops and fricatives as a delay caused by the time required for adjusting the vocal cords to the configuration appropriate for consonant vibrations. But this adjustment may be a mechanical consequence of the articulatory constriction, which does not require particularly long time. It is not quite convincing either why this adjustment cannot start during the vowel just as the opening before voiceless consonants; as a matter of fact many airflow curves show a rise at the end of the vowel before voiced stops, although it is not as high as before voiceless consonants. Chomsky-Halle mention that before nasals, which do not require any adjustment, the vowels are shorter than before voiced stops. I have not found any confirmation of this for German, nor is such a difference found in the measurements of Czech vowels by Per Jacobsen (this volume pp. 135-142).

The assumption that it is the adjustment of the vocal cords for consonant vibrations which lengthens the vowel is also contradicted by the following two facts: (1) When French bdg vzʒ become voiceless through assimilation to a following consonant they still preserve (at least part of) their lengthening effect on the preceding vowel (Thorsen 1967). (2) Also French speakers who normally have voiceless vzʒ (CHH and her sister ThM) have longer vowels before these consonants than before fsʃ, although the difference is somewhat smaller (K. Landschultz 1968). This is also valid for bdg (EFJ 1969). Sometimes these consonants have partial voicing, but the preceding vowel is not longer in these cases than in cases with completely voiceless vzʒ (oral communication from K. Landschultz).

6. Final remarks.

6.1. As it appears from the preceding pages, I am inclined to keep voicing, aspiration and tenseness as three separate phonetic features. By phonetic feature I mean the general phonetic dimensions which may be utilized for distinctive purposes in various languages. Two different phonetic features may be combined to one complex phonemic feature (e.g. tenseness and voicing in French stops, rounding and front-back in Spanish vowels).

The criterion for keeping phonetic features apart must be their independency. One must not be a mechanical consequence of the other.

In 5.1. above arguments were given for keeping tenseness and aspiration apart, the differences being exemplified by a comparison between French and Danish ptk. They seem to be independent, and they may be combined in different ways.

As for tenseness and voicing, there seems to be a certain affinity in the sense that voiced stops are generally lax; however, according to the most common description of voice assimilation in French, French ptk may become voiced while remaining tense. At any rate the assimilation of duration does not seem to be complete. The total assimilation of intra-oral pressure may be due to the fact that intra-oral pressure is part of the voicing feature (Thorsen 1967).

Voicing and aspiration may be combined as proposed by Abramson-Lisker as steps in the same dimension if Indian aspirated bdg are left out of consideration, but somehow these must belong to the voicing dimension, which means that in any case aspiration can be combined with two different steps in this dimension. - Moreover, it is too simple to consider voiced stops, unaspirated stops and aspirated stops as steps in one dimension. Voicing requires an extra mechanism, not only a certain position of the glottis.

6.2. The phonetic qualities described for French and Danish can be tentatively distributed on these three features in the following way (see also Table I):

Voicing.

Voiced stops - as compared with voiceless stops - are characterized physiologically by having a narrow glottis (whereas the glottis in voiceless stops may be wide or narrow); moreover they have a larger pharynx cavity, and consequently a lower supraglottal pressure permitting vibrations of the vocal cords, whereas the higher supraglottal air pressure in the voiceless stops prevents the vocal cords from vibrating. Voiced stops have also less airflow.

Acoustically, voiced stops are characterized by periodic sound at a low frequency during the closure, and by a lower start of the fundamental of the following vowel, whereas voiceless stops have silent closure and a higher start of the vowel.

Aspiration.

Physiologically, aspirated stops are characterized (as against unaspirated stops) by having a wider glottis opening and by the fact that the release of the oral closure takes place at a time when the glottis is still relatively wide open, which causes a strong airflow after the explosion.

Acoustically, aspirated stops are characterized by a long open interval with noise at the frequencies of or leading toward the formants (above F_1) of the following vowel; moreover they exhibit a high start and short duration of the F_1-transition.

Tenseness.

Physiologically, tense stops are characterized by having a stronger organic pressure and a longer closure period than lax stops, and probably by a quicker and more precise movement of the articulating organs.

Acoustically, they are characterized by a longer closure period and a shortening of the preceding vowel and, to some extent, of the following vowel, perhaps also by a stronger intensity of the explosion (this is still to be proved) and a quicker intensity rise of the following vowel, and finally by a shorter distance from the explosion to the top of the transition and by a shorter transition.

In CHH's French stops there was a certain difference of aspiration before i and u - but before a tenseness seemed to be the only distinguishing feature. However, no difference was found neither in the explosions nor in the transitions. As mentioned above, it cannot be excluded that differences in explosion type might be found by a more appropriate technique. Spectrograms did not reveal any difference in the frequency of the explosions (cp. Halle, Hughes, Radley 1957).

It is obvious that this distribution is very preliminary, and, on the whole, there are still so many unsolved problems that much of what has been said on the preceding pages has the character of speculations. We need more investigations of subglottal pressure, of glottis opening, of muscular activity and organic pressure, of the intensity and type of the explosions and of the transitions.

P.S.: After having completed this report my attention was drawn to the paper by Abramson and Lisker "Laryngeal Behavior, the Speech Signal and Phonological Simplicity", Status Report on Speech Research, Haskins Laboratories 11 (1967), pp. 23-33, in which the authors draw some general conclusions from their various experiments with stop consonants.

I agree in considering voicing and supraglottal air pressure as closely related and in pointing to the possibility that the widening of the pharynx in voiced stops may be an active process, but not in the general conclusions, since the authors maintain that voice onset timing may be the only fundamental feature distinguishing categories of stops in various languages.

References:

Abramson, Arthur S. and Leigh Lisker. 1965. "Voice Onset Time in Stop Consonants: Acoustic Analysis and Synthesis", Status Report on Speech Research, Haskins Laboratories 3.1.

Brunner, Rudolf. 1953. "Die Stimmhaftigkeit der französischen und zürichdeutschen Lenislaute", Zeitschrift für Phonetik 7, pp. 153-196 und 289-319.

Charbonneau, abbé René. 1955. La palatalisation de t/d en Canadien-Français: Etude de phonétique expérimentale (Montreal).

Chlumsky, J. 1922. "Les consonnes anglaises comparées aux consonnes françaises à l'aide du palais artificiel et de l'observation directe", Revue de phonétique 4, pp. 189-218.

Chomsky, N. and M. Halle. 1968. The Sound Pattern of English (New York, Evanston, London).

Cooper, F.S., P. Delattre, A.M. Liberman, J.M. Borst and L.J. Gerstman. 1952. "Some Experiments in the Perception of Synthetic Speech Sounds", JASA 24,6, pp. 597-606.

Delattre, Pierre. 1939. "Durée vocalique et consonnes subséquentes", Le Maître Phonétique 67 (= Studies in French and Comparative Phonetics 1966, pp. 130-132).

Delattre, Pierre. 1940a. "Anticipation in the Sequence: Vowel and Consonant-Group", The French Review 13 (= Studies in French and Comparative Phonetics 1966, pp. 122-.27).

Delattre, Pierre. 1941. "La force d'articulation consonantique en français", The French Review 14, pp. 220-232, (= Studies in French and Comparative Phonetics 1966, pp. 111-120).

Delattre, Pierre. 1958. "Les indices acoustiques de la parole", Phonetica 2, pp. 108-118 and 226-251.

Delattre, Pierre. 1961. "Le jeu des transitions de formants et la perception des consonnes", Proceedings of the IV International Congress of Phonetic Sciences, pp. 407-417.

Delattre, Pierre. 1962. "Some Factors of Vowel Duration and their Cross-Linguistic Validity", JASA 34 pp.1-2. (= Studies in French and Comparative Phonetics 1966, pp. 133-140).

Delattre, Pierre. 1964. "De la hiérarchie des indices acoustiques pour la perception de la parole", Proceedings of the V International Congress of Phonetic Sciences (Münster), pp. 244-251.

Delattre, Pierre. 1968. "From Acoustic Cues to Distinctive Features", Phonetica 18, pp. 198-230.

Denes, P. 1955. "Effect of Duration on the Perception of Voicing", JASA 27, pp. 761-764.

Dieth, E. und R. Brunner. 1943. "Die Konsonanten und Geminaten des Schweizerdeutschen experimentell untersucht", Festschrift Jud. Romanica Helvetica 20, pp. 737-762.

Durand, Marguerite. 1936. Etude expérimentale sur la durée des consonnes parisiennes. Bibliothèque du "français moderne" (Paris).

Durand, Marguerite. 1956. "De la perception des consonnes occlusives: questions de sonorité", Word 12, pp. 15-34.

Evertz, Erich. 1929. Beiträge zur Phonetik der südfranzösischen Plosive: eine experimentelle Untersuchung (Bonn).

Fabre, Ph. 1961. "Glottographie respiratoire", Comptes rendus des séances de l'Académie des sciences 6, pp. 1386-1388.

Fairbanks, Grant et Murray S. Miron. 1958. "Effects of Vocal Effort upon the Consonant-Vowel Ratio within the Syllable", JASA 29, pp. 621-626.

Falc'hun, F. 1965. "L'énergie articulatoire des occlusives", Phonetica 13, pp. 31-36.

Fant, Gunnar. 1958. "Modern Instruments and Methods for Acoustic Studies of Speech", Acta Polytechnica Scandinavica.

Fant, Gunnar. 1960. Acoustic Theory of Speech Production (Mouton).

Fischer-Jørgensen, Eli. 1954. "Acoustic Analysis of Stop Consonants", Miscellanea Phonetica II, pp. 42-59.

Fischer-Jørgensen, Eli. 1963. "Beobachtungen über den Zusammenhang zwischen Stimmhaftigkeit und intraoralem Luftdruck", Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung 16, pp. 19-36.

Fischer-Jørgensen, Eli. 1964. "Sound Duration and Place of Articulation", Zeitschrift für Phonetik etc. 17, pp. 175-208.

Fischer-Jørgensen, Eli. 1966. "Phonetic Analysis of Danish Stop Consonants", Annual Report of the Institute of Phonetics, University of Copenhagen 1, pp. 31-33.

Fischer-Jørgensen, Eli. 1969. "Les occlusives françaises et danoises d'un sujet bilingue", Festschrift Martinet 2 (in print).

Fischer-Jørgensen, Eli, Børge Frøkjær-Jensen and Jørgen Rischel. 1966. "Preliminary Experiments with the Fabre Glottograph", Annual Report of the Institute of Phonetics, University of Copenhagen 1, pp. 22-3o.

Frøkjær-Jensen, Børge. 1967. "A Photo-Electric Glottograph", Annual Report of the Institute of Phonetics, University of Copenhagen 2, pp. 5-19.

Fujimura, Osamu. 1959. Some Characteristics of Stop Consonants. (Cleveland). 4 pp.

Fujimura, Osamu. 1961. "Bilabial Stop and Nasal Consonants: A Motion Picture Study and its Acoustical Implications", Journal of Speech and Hearing Research 4, pp. 233-247.

Gendron, Jean-Denis. 1966. Tendances phonétiques du français parlé au Canada (Paris).

Gimson, A.C. 1962. An Introduction to the Pronunciation of English (2nd ed. 1964).

Halle, M., G.W. Hughes and J.P.A. Radley. 1957. "Acoustic Properties of Stop Consonants", JASA 29, pp. 1o7-116.

Halle, M. and K. Stevens. 1967. "On the Mechanism of Glottal Vibration for Vowels and Consonants", MIT QPR no. 85, pp. 267-27o.

Harris, Katherine S., Gloria Lysaught and Malcom M. Schvey. 1965. "Some Aspects of the Production of Oral and Nasal Labial Stops", Status Report on Speech Research, Haskins Laboratories 2.1.

Jakobson, Roman, G. Fant and M. Halle. 1952. Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates (Technical Report no. 13. Massachusetts Institute of Technology).

Jakobson, Roman and M. Halle. 1962. "Tenseness and Laxness", Selected Writings, pp. 55o-555.

Jespersen, Otto. 1914. Lehrbuch der Phonetik, 5. Auflage (1932).

Kaneko, Naomichi. 1961. "An Experimental Observation and Phonetic Analysis of Opposite Consonants such as /p/ : /b/", Study of Sounds IX, pp. 53-64.

Kim, Chin-W. 1965. "On the Autonomy of Tensity Feature in Stop Classification (with special reference to Korean Stops)", Word 21, pp. 339-359.

Kloster Jensen, M. 1956. "Zur Glottalität der reinen Tenues", Zeitschrift für Phonetik 9, pp. 344-359.

Koneczna, H. et W. Zawadowski. 1956. Obrazy rentgenograficzne glosek rosyjskich (Warszawa).

Ladefoged, P. 1967. Linguistic Phonetics. Working Papers in Phonetics, UCLA 6.

Landschultz, Karen. 1967. "Duration of French Vowels Before Fricatives", Annual Report of the Institute of Phonetics, University of Copenhagen 2, pp. 1o9-118.

Liberman, Alvin M. 1957. "Some Results of Research on Speech Perception", JASA 29,1, pp. 117-123.

Liberman, Alvin M., P.C. Delattre and F.S. Cooper. 1958. "Some Cues for the Distinction between Voiced and Voiceless Stops in Initial Position", Language and Speech 1, pp. 153-167.

Liberman, Alvin M., F.S. Cooper, K.S. Harris and P.F. Macneilage. 1962. "A Motor Theory of Speech Perception", Speech Communication Seminar, Stockholm.

Lindblom, B.E.F. 1967. "Vowel Duration and a Model of Lip Mandible Coordination", Quarterly Progress and Status Report, Speech Transmission Laboratory Stockholm 4/1967, pp. 1-29.

Lisker, Leigh. 1957. "Closure Duration and the Intervocalic Voiced-Voiceless Distinction in English", Language 33, pp. 42-49.

Lisker, Leigh. 1965. "Supraglottal Air Pressure in the Production of English Stops", Status Report on Speech Research, Haskins Laboratories 4.3.

Lisker, Leigh. 1966. "Measuring Stop Closure Duration from Intra-oral Pressure Records", Status Report on Speech Research, Haskins Laboratories 7/8.5.

Lisker, Leigh, and Arthur S. Abramson. 1964. "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements", Word 2o, pp. 384-422.

Lisker, Leigh, and Arthur S. Abramson. 1966. "Some Effects of Context on Voice Onset Time in English Stops", Status Report on Speech Research, Haskins Laboratories 5/6.1.

Lisker, Leigh, A.S. Abramson, F.S. Cooper and M.H. Schvey. 1966. "Transillumination of the Larynx in Running Speech", Status Report on Speech Research. Haskins Laboratories 5/6, 4.1.

Lisker, Leigh, and Arthur S. Abramson. 1967. "The Voicing Dimension: Some Experiments in Comparative Phonetics", Status Report on Speech Research. Haskins Laboratories 11. 9-16.

Lotz, John, Arthur S. Abramson, Louis J. Gerstman, Frances Ingemann and William J. Nemser. 196o. "The Perception of English Stops by Speakers of English, Spanish, Hungarian and Thai: A Tape-Cutting Experiment", Language and Speech 3, pp. 71-77.

Malécot, André. 1955. "An Experimental Study of the Force of Articulation", Studia Linguistica 9, pp. 35-44.

Malécot, André. 1966a. "The Effectiveness of Intra-Oral Air-Pressure-Pulse Parameters in Distinguishing Between Stop Cognates", Phonetica 14, pp. 65-81.

Malécot, André. 1966b. "Mechanical Pressure as an Index of 'Force of Articulation'", Phonetica 14, pp. 169-180.

Malmberg, Bertil. 1943. Le système consonantique du français moderne: Etudes de phonétique et de phonologie=Etudes Romanes de Lund VII.

Öhman S., R. Leanderson and A. Persson. 1966. "EMG Studies of Facial Muscles Activity in Speech II", Speech Transmission Laboratory, Stockholm. Quarterly Progress and Status Report 1/1966, pp. 1-5.

Perkell, J. 1965a. "Studies of the Dynamics of Speech Production", MIT QPR no. 76, pp. 253-257.

Perkell, J. 1965b. "Cineradiographic Studies of Speech: Implications of a Detailed Analysis of Certain Articulatory Movements", Congrès International d'Acoustique, Liège, Paper A 32.

Peterson, Gordon E. and Ilse Lehiste. 1960. "Duration of Syllable Nuclei in English", JASA 32, pp. 693-703.

Rothenberg, Martin. 1968. The Breath-Stream Dynamics of Simple-Released-Plosive Production. Bibliotheca Phonetica 6.

Roudet, L. 1900a. "De la dépense d'air dans la parole", La Parole 2, pp. 201-230.

Roudet, L. 1900b. "Recherches sur le rôle de la pression sousglottique dans la parole", La Parole 2, pp. 599-612.

Roudet, L. 1910. Eléments de phonétique générale.

Rousselot, J.-P. 1891. Les modifications phonétiques du langage etudiées dans le patois d'une famille. (Paris).

Rousselot, J.-P. 1897. Principes de phonétique expérimentale I (Paris). 2e édition 1924.

Rousselot, J.-P. 1899. "Etude de prononciations parisiennes", La Parole 1, pp. 481-549.

Senn, A.C. 1935. "An Experimental Study of Bengali Occlusives", Proc. II Congr. Phon. Sc., pp. 184-193.

Simon, Pela. 1961. "Les consonnes françaises et leurs variations dans la chaine parlée d'après des films radiologiques", Proceedings of the IV International Congress of Phonetic Sciences (Helsinki), pp. 241-249.

Simon, Pela. 1967. Les consonnes françaises: mouvements et positions articulatoires à la lumière de la radiocinématographie (Paris).

Slis, I.H. 1967. "What Causes the Voiced-Voiceless Distinction?", IPO Annual Progress Report 2, pp. 71-76.

Slis, I.H. and P.H. Damsté. 1967. "Transillumination of the glottis during voiced and voiceless consonants", IPO Annual Progress Report 2, pp. 1o3-1o9.

Straka, Georges. 1942. "Notes de phonétique générale et expérimentale", Bulletin de la faculté des lettres de Strasbourg 1.

Straka, Georges. 1952. "La Prononciation Parisienne", Bulletin de la faculté des lettres de Strasbourg 1 -45.

Straka, Georges. 1963. "La division des sons du langage en voyelles et consonnes peut-elle être justifiée?", Travaux de linguistique et de littérature 1, pp. 17-99.

Straka, Georges. 1965a. Album Phonétique (Quebec).

Straka, Georges. 1965b. "Naissance et disparition des consonnes palatales dans l'évolution du latin au français", Travaux de linguistique et de littérature 3,1, pp. 117-167.

Thorsen, Oluf M. 1966. "Voice Assimilation of Stop Consonants and Fricatives in French and its Relation to Sound Duration and Intra-Oral Air-Pressure", Annual Report of the Institute of Phonetics, University of Copenhagen 1, pp. 67-76.

Ventsov, A. 1967. O rabote golosovykh svyazok pri glukhikh smychnykh intervokalnykh soglasnykh. Mekhanizmy recheograzovaniya i vospriatiya slozhnykh zvukov.

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 [ɑː] 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	80.75	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	a
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	æ:
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	3.19	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	80.97	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,

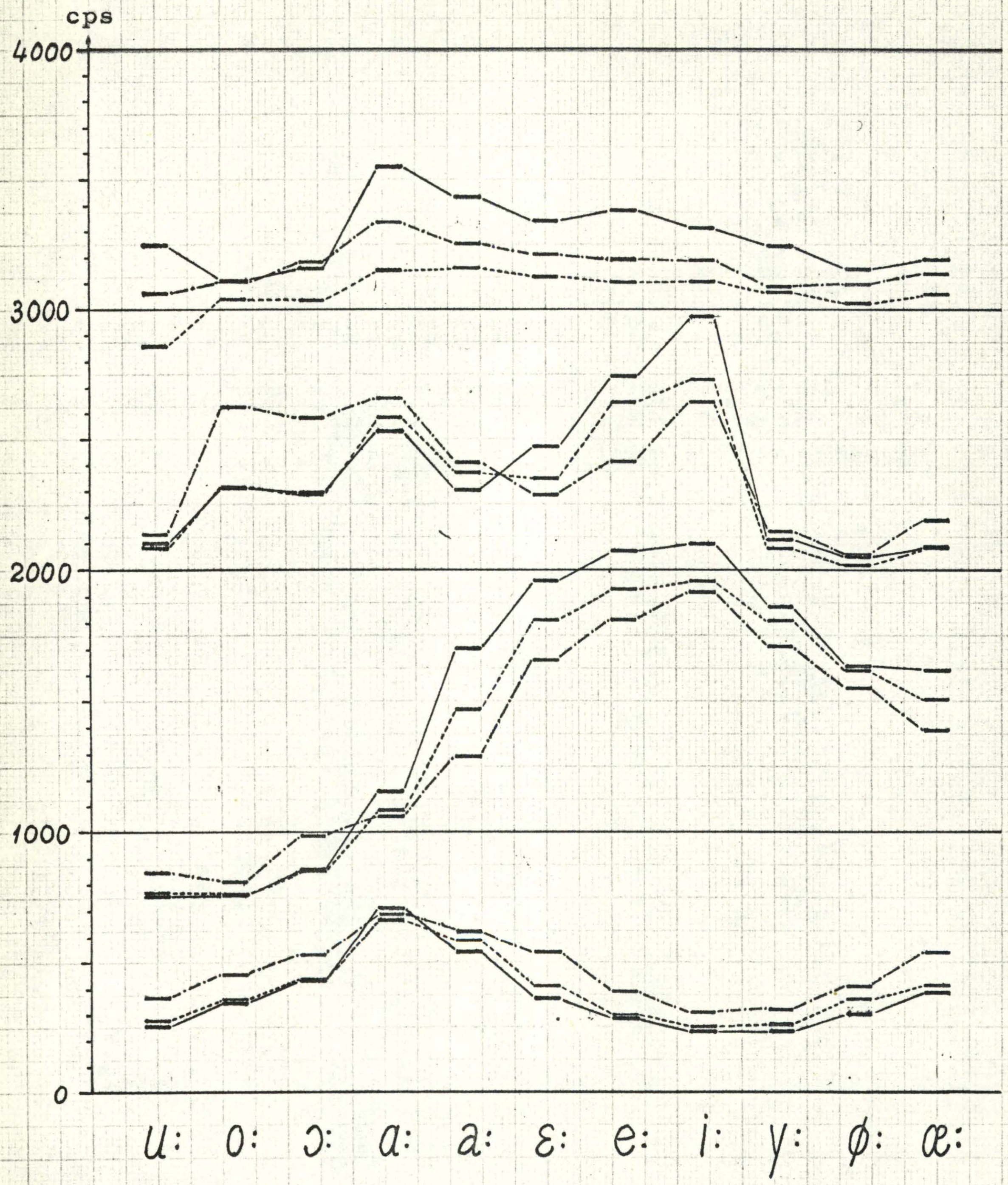
THE AVERAGE FREQUENCY OF FORMANT 4 IN GROUP B IN RELATION TO GROUP A IS -3.29 PER CENT.

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



<u>Comparison between spoken and sung vowels.</u>

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

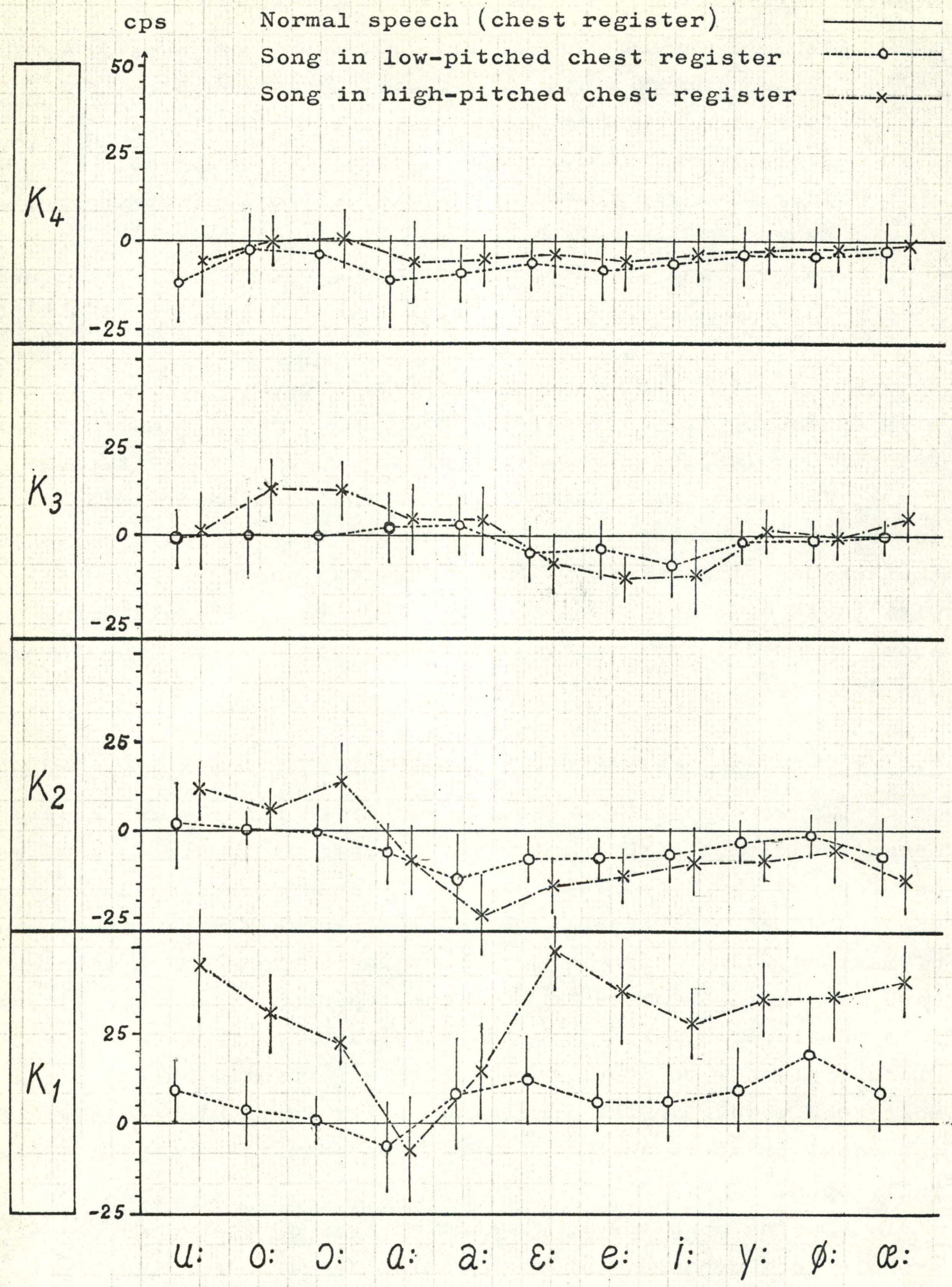
These two tables show that many subjects perform an active expanding of the vocal tract during singing, and that very few perform an active reduction of the resonance cavities.

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.

FOURIER ANALYSES OF PHOTO-ELECTRIC GLOTTOGRAMS

Børge Frøkjær-Jensen and Jørgen Rischel

1. Glottis area versus volume velocity curve.

According to Flanagan (2), (3) there seems to be a good agreement between the glottal area and the volume velocity wave when operating in a constant and moderate subglottal pressure range of about 4 cm H_2O, - the effects of inertance neglected. It should therefore (within the mentioned restrictions) be possible to calculate the acoustic spectrum of the primary voice source based upon photo-electric glottograms (but not upon the Fabre glottograms)*) by means of a digital computer.

According to Fant and Sonesson (1) simultaneous recordings of photo-electric glottograms and inverse filterings seems to give nearly the same waveforms, which supports the theoretical calculations of Flanagan and Fant.

2. Expected differences in the glottal waveforms of vowels and consonants.

A problem of great current interest is the possibility of interaction between poles and zeros of the vocal tract transfer function and the glottal volume-velocity waveform.

The supraglottal impedance relationships are analysed by Flanagan (reference 3, pp. 46-47) on the basis of transmission line theory. He finds that the driving point impedance of the uniform vocal tract (set in position for the neutral reference vowel) reaches maxima at those frequencies where the imaginary term of the propagation constant **) $x\ell$ equals $\pi/2$, $3\pi/2$, etc. The impedance at these maxima, which coincide with F_1, F_2, etc. is determined as the square of the characteristic impedance of the tract divided by the oral radiation impedance. It is clear that the driving point impedance reaches its highest maximum when the radiation impedance is low, i.e. at the F_1 frequency,

---

*) This has actually been done by C. Bordone-Sacerdote & G. Righini, "Glottal wave detected as a high frequency modulation", 5th Int. Congr. of Acoustics (Liège 1965), Paper A 34.

**) These formulae are based on the simplifying assumption that the real term is zero, i.e. the vocal tract is lossless.

where the driving point impedance is so high that tract-source interaction occurs (reference 3, p. 47).

Halle and Stevens (reference 4, p. 268) calculate the conditions at different first formant frequencies. They take the tract as a resonant circuit tuned to F1 and find by considering the impulse response of this supraglottal filter that it will react to the glottal wave with a pressure peak which at low F_1 values (F_1 below 250 cps) is of the order of the steady sub-glottal pressure. Thus the glottal vibrations will clearly be impeded at very low F_1 frequencies.

The authors presume that this effect is counteracted by adjusting the vocal cords towards a more open state so that they can vibrate with a lower pressure drop across them. This adjustment is assumed to condition the longer duration of vowels before voiced than before voiceless obstruents (in English). *)

In order to test whether the Fourier analysis of photo-electric glottograms constitutes a usable method for empirical investigation of these problems we have made a few analyses of typical glottograms. These are, of course, to be taken as pilot experiments only.

3. Experimental procedure.

The two sustained sounds [æ:] and [z:] spoken several times by the authors have been recorded by means of the photo-electric glottograph. Because of lack of a high-speed analog/digital converter at our laboratory we have photographed the oscilloscope screen and by hand measured 281 ordinates per cycle from large scale photographic enlargements of two typical glottograms. The ordinate values have been used as input data for a Fourier analysis made by computer. The computer outputs, parts of which are shown in Fig. 1 and 2, contain the sine and cosine coefficients, the phase differences, the absolute amplitude, and the logarithmic amplitude (in dB). Based on these outputs the line spectrums in Fig. 3 have been drawn. These compare the two primary voice source functions: that of the vowel and that of the consonant.

---

*) For discussion of this problem: see the article in this report by Eli Fischer-Jørgensen.

HARMONICS NO.	FREQUENCY IN CPS	AMPLITUDE IN MVOLTS	AMPLITUDES DECIBELS
DC	0	1090.11833	40.75
1	120	796.87701	38.03
2	240	306.75773	29.74
3	360	40.70412	12.19
4	480	33.45657	10.49
5	600	14.34248	3.13
6	720	15.20397	3.64
7	840	3.03825	-10.35
8	960	4.10174	-7.74
9	1080	1.71341	-15.32
10	1200	4.02491	-7.90
11	1320	1.63957	-15.71
12	1440	0.84083	-21.51
13	1560	0.78087	-22.15
14	1680	1.79724	-14.91
15	1800	0.59086	-24.57
16	1920	0.98869	-20.10
17	2040	0.59233	-24.55
18	2160	0.67564	-23.41
19	2280	0.58004	-24.73
20	2400	0.37273	-28.57
21	2520	0.55248	-25.15
22	2640	0.12864	-37.81
23	2760	0.65609	-23.66
24	2880	0.35398	-29.02
25	3000	0.51560	-25.75

Fig. 1.

HARMONICS NO.	FREQUENCY IN CPS	AMPLITUDE IN MVOLTS	AMPLITUDES DECIBELS
DC	0	1314.15202	42.37
1	120	805.75991	38.12
2	240	145.58764	23.26
3	360	86.47672	18.74
4	480	14.93569	3.48
5	600	9.15732	-0.76
6	720	5.18039	-5.71
7	840	3.07824	-10.23
8	960	2.38402	-12.45
9	1080	3.46416	-9.21
10	1200	0.91052	-20.81
11	1320	1.27400	-17.90
12	1440	1.15195	-18.77
13	1560	1.31342	-17.63
14	1680	0.99831	-20.01
15	1800	1.06320	-19.47
16	1920	0.36361	-28.79
17	2040	1.28072	-17.85
18	2160	0.89437	-20.97
19	2280	0.45730	-26.80
20	2400	0.65082	-23.73
21	2520	0.69761	-23.13
22	2640	0.44947	-26.95
23	2760	0.50149	-25.99
24	2880	0.33686	-29.45
25	3000	0.74337	-22.58

Fig. 2.

4. Results.

4.1. The glottograms.

The photographs in Fig. 3 show typical differences in the waveform of the glottal pulse. The vowel has a very symmetrical pulse (this is not always the case) with steep slopes, whereas the consonant pulse has a less steep slope in the closing phase. If, arbitrarily, we replace the lowest 7 per cent of the amplitude of each glottogram by a straight line (some of the slight bottom curvature is probably associated with transillumination and nonlinearity phenomena, compare reference 3, p.44), the "open" phase (triangular portion) of the vowel pulse is 4.2 ms = 51 per cent of the total duration of the cycle, whereas the corresponding part of the consonant pulse is 5.6 ms = 68 per cent of the total duration of the cycle. The fundamental frequency is the same in both analyses: 120 cps, i.e. the total duration of the cycle is 8.3 ms (1.4 ms per division on the oscilloscope screen). - The curves shown in Fig. 3 do not give any measure of the absolute degree of closure obtained in the two types of sounds, so we are not making any deductions concerning the degree of leakage through the glottis from these curves (cp., however, the curves of Danish voiced fricatives on p.13 of this report).*)

The differences observed may condition a weaker excitation and a greater damping of formants in the consonant.**)

4.2. The line spectra.

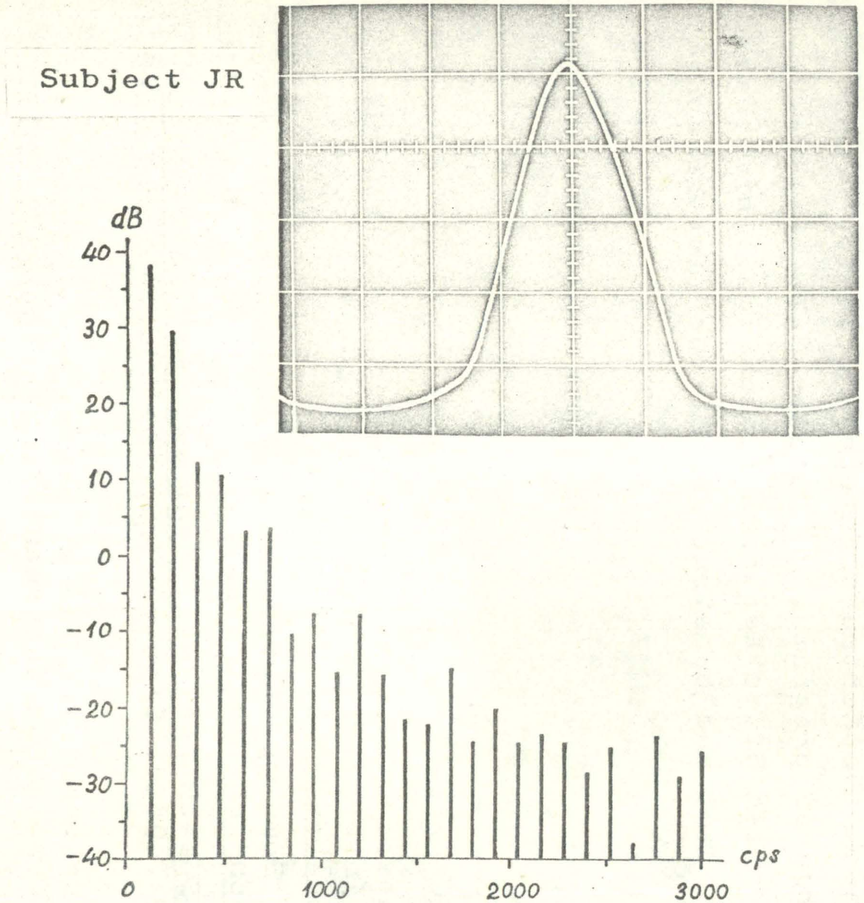
The line spectra for the two sounds are shown below the photographs in Fig. 3.

Up to about 1000 cps (the 9th harmonic) the slope of the primary voice source rolls off at a rate of 12-15 dB/octave, which agrees well with other measurements. Above 1000 cps the line spectrum is not reliable because the dynamic range is limited to about 60 dB. In future experiments we are going to use some kind of high-shaping network, which may extend the usable frequency range considerably.

---

*) See also Lisker, Abramson, Cooper, and Schvey in Status Report (SR-5/6), Haskins Laboratories, 1966, pp. 4.1-8.

**) The absolute amplitudes for the two sounds in Fig. 3 cannot be directly compared, since the recordings have no calibration.



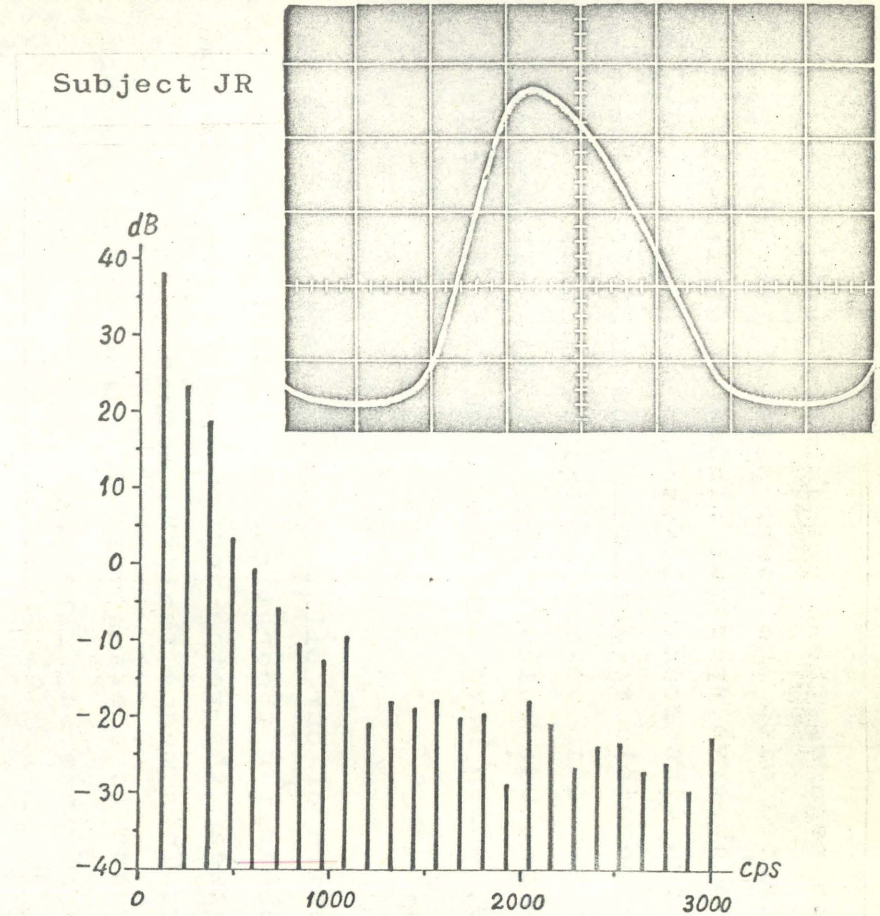


Fig. 3.

Glottogram and line spectrum of [ɛ:]

Glottogram and line spectrum of [z:]

Results of Fourier analysis of the glottograms are shown in the photocopies of the computer analysis.

Therefore, as regards these preliminary Fourier analyses we can only be concerned with the lower part of the spectrum. In this range we can observe spectral minima at different frequencies in the two line spectra. With an extensive material it will probably be possible to determine the different locations of the voice source zeros in vowels contra consonants.

Our analyses include calculations of phase differences among the harmonics, which have not been used for this article.

5. Conclusions.

The line spectra of our pilot experiments seem to be in fair agreement with earlier theoretical data and to confirm that differences between the voice source spectra of vowels and consonants do occur. This means that the method used here seems to be a usable approach to the analysis of voice source spectra parallel to the use of inverse filtering.

Acknowledgement.

We want to thank the Northern Europe University Computing Center (NEUCC) for run of programs, and "Statens Almindelige Videnskabsfond" (the Danish State Research Foundation), which has supported the development and construction of the glottograph.

References.

(1) G. Fant, B. Sonesson, "Indirect Studies of Glottal Cycles by Synchronous Inverse Filtering and Photo-electrical Glottography", STL-QPSR 4(1962), pp.1-3.

(2) J.L. Flanagan, "Some properties of the glottal sound source", Journal of Speech and Hearing Research 1(1958), pp. 99-116.

(3) J.L. Flanagan, Speech Analysis, Synthesis and Perception. Springer Verlag (1965). Chapter 3.5, The Source for Voiced Sounds.

(4) M. Halle, K.N. Stevens, "On the Mechanism of Glottal Vibrations for Vowels and Consonants", QPR 85(1968), MIT.

(5) R.L. Miller, "Nature of the Vocal Cord Wave", Journal of the Acoustical Society of America 31(1959), pp. 667-677.

VOWEL QUANTITY IN CZECH

An instrumental study

Per Jacobsen

1. Introduction.

The purpose of this investigation is to examine the relationship between long and short stressed vowels in Czech.

Data on the duration of Czech vowels have been published by others, particularly Josef Chlumsky (1) and L. Kaiser (2) (based on research by P. Janota). A brief comparison between the conclusions reached by these authors and those of the present paper will be given at the end of the paper.

In order to get as uniform a material for the present investigation as possible some preliminary investigations were made with the purpose of determining whether the duration of vowels is influenced by the number of syllables and/or the length of a following syllable.

Word lists of the type host, hosti, hostitel, hostitelka, hostitelkami were set up and the duration of the first vowel were measured and compared. These series of words are referred to as α-words. In a number of word pairs of the type zubu, zubů ['zubu] ['zubuː] the duration of the first vowel was measured. These examples are referred to as α-words. Lists of α-words and β-words together with the results of this analysis are shown in Tables 1 and 2.

As it is seen, it is necessary to consider the number of syllables and, to a certain degree, also the length of the following syllable when vowel duration is to be investigated. In the last and main part of this study only dissyllables with short second syllable are analysed. As the purpose of the investigation is to examine the relationship between long and short vowels, only words with long and short a, e, i, o, u in the stressed syllable are considered, since the opposition long: short is neutralized in vocalic r and l.

Since the following consonant may influence vowel duration all vowels are examined in positions before (1) nasal/liquid, (2)(3) voiced and unvoiced fricatives, and (4)(5) plosives.

The structure of the investigated syllables in this part of the study is: (c)cv(:) (This is not necessarily the case with the α- and β-words.)

As the physical duration usually depends on vowel quality, the material is so arranged that long a is compared with short a, long e with short e, etc.

The utterances were recorded in a sound treated room on a professional tape recorder at the Institute of Phonetics. The words were read in a quasi-random order, the only restriction being that words of the type host,hosti,hostitel ...etc., zubu,zubů or zakop,zákop did not occur adjacent to each other. Each word was read as an isolated utterance with a pause between each utterance, and the informants were asked not to make any rhythmic groups in their reading.

Five informants were used:

JL (male) born in Prague in 1943

OL (female) born in Prague in 1943

LM (female) born in Prague in 1944

BU (female) born in Prague in 1944

AKK (female) born in Danmark in 1944; bilingual (mother Czech, father Danish).

The complete material consists of 117o utterances (117 examples each recorded twice by all 5 informants).

2. Recording of acoustic curves.

The utterances were subjected to acoustic analysis by means of an intensity meter (KTH type) and Frøkjær-Jensen's pitch meter and recorded on the mingograph. Four synchronous traces were recorded on the ink writer:(1) a pitch curve, (2) a logarithmic intensity curve with an integration time of 2,5 ms for female voices and 5 ms for male voices (the high-pass filtering in connection with the logarithmic scale conditions that the consonants stand out rather distinctly, (3) an intensity curve (linear scale) with an integration time of 5 and lo ms, respectively, and with linear frequency response, and (4) a duplex oscillogram (i.e. a combination of an ordinary oscillogram and a high-pass filtered intensity curve).

The logarithmic intensity curve and the duplex oscillogram were used to delimitate the sounds before measurements were made. The speed of the paper was loo mm/sec.

Table 1.a.

Informant	JL					OL				
Number of syllables	1	2	3	4	5	1	2	3	4	5
pan	2o	12,5	12			19	1o,5	9		
dám	3o,5	24	2o,5			3o,5	25	19		
demokrat			1o	9	8,5			9	9,3	7
let	17	12	1o			15,5	1o,3	9,5		
krém	26,5	21	18	15,5		27	22	16,8	16	
lék	25,5	22	15			25,5	2o,5	15,6		
bych	16	9,3				15	1o			
list	14	9,5	7,8	7		15	1o	9	7,5	
mír	25,5	24,5				28	25			
lít	19	14,5	12,5	1o,5		21	14,5	14,5	11,8	
mýt	22,5	16	14,5			24	15,3	14,5		
host	17,8	12,3	11	9,5	7	17	1o,3	1o	8,5	6,3
dóm	34	25,5				31,8	24,3			
tuš	13	8,5				16,5	9,5			
muk	12	9	8,5			14	1o	9		
-										

Table 1.b.

Informant	LM					BU				
Number of syllables	1	2	3	4	5	1	2	3	4	5
pan	2o,3	14	13,5			17	1o	7		
dám	47,5	36	32			3o	25	21		
demokrat			11,3	1o	9,3			7,5	7	5,5
let	15	9,8	8,5			1o,5	6,5	6		
krém	38	34,5	25,5	27,8		24,5	24	19,2	19	
lék	35	24,3	23,8			23	15	16		
bych	15	9,6				13,2	8,3			
list	13	9	8,5	7,8		11,5	8	7,5	7	
mír	43,5	36,3				26	22,5			
lít	3o,8	26,8	25,3	17,5		16	13,8	12	11	
mýt	36,5	27,5	23,5			17	14,5	13,5		
host	13	8,5	8,5	7,8	8,5	12,5	11	8	1o	9,5
dóm	45,5	36,4				22	19			
tuš	13	1o				1o	8,5			
muk	12	9,8	7,3			12,5	1o	8		
-										

Table 1.c.

Informant	AKK				
Number of syllables	1	2	3	4	5
pan	14,3	12,8	1o		
dám	27	21	18,3		
demokrat			7,3	7	6,6
let	12	1o,5	8,5		
krém	23	19,8	16,5	15,3	
lék	24,5	15,5	12,9		
bych	11	8,8			
list	1o	8,1	7,5	7	
mír	26,3	19			
lít	19,8	13,5	11,5	1o	
mýt	17,8	12,8	11		
host	13,5	12	9,5	12	9,3
dóm	28	2o,3			
tuš	12,3	9,8			
muk	8	7,3	7,5		
-					

α-words. Vowel duration of first syllables in cs. All five informants. (Averages of two recordings.)

pan - panel- Panama
dám - dáma - dámami
demokrat - demokrata - demokratičnost
let - letec - letectvo
krém - kréma - krémový- krémovými
lék - lékař - lékaře
bych - bychom
list - listu - listina - listinami
mír - míra
lít - lítost - lítosti - lítostivý
mýt - mýtit - mýtina
host - hosti - hostitel - hostitelka - hostitelkami
dóm - dóma
muk - muka - mukami
tuš - tušit

Table 2.

	JL	OL	LM	BU	AKK
draho	14,3	13	17	13,5	12,5
drahý } drahá	12,3	1o,9	12	8	11,7
našich } našim	9,5	11,5	1o,4	1o,3	11,6
naším	9	1o,8	11	9,5	11,5
cesta	9,5	9,1	1o	7,8	11,5
cestář	7,5	8	9,2	8	8,3
cena	11	1o	14,3	8,8	9,5
ceník	9,5	9	13	7	9,5
ryba	9,3	1o,3	12,8	13	1o,3
rybám	9	9,5	1o,8	1o,5	1o,5
syto } syta	8,9	8,2	8,3	6,8	7,6
sytá	7,8	8	7,3	6,5	7,5
doma } domu	13	12,5	16,3	12,5	11,8
domů	12	1o	16	12	12
dostat	11,5	11,5	9,5	1o	1o,8
dostát	1o	9	11	9	11,8
dobro	11	9	12	11	1o,3
dobré	9,5	9,8	12,3	9,3	9,8
zubu	9,5	1o	12	9	9,3
zubů	9,5	9,5	11	8,3	1o,8
dáma	24	25	36	25	21
dámách	24,5	21	33	21	21,5
léto	17,5	2o,5	28	17,5	17,3
létům	16,5	19,8	28,8	15,5	17,3
lípa	16	16,5	33	14	16,3
lípách	15	16	29	11	16
-					
půda	2o,5	18,8	37,5	2o	21,6
půdách	2o	15,8	34	21	16

β-words. Vowel duration of first syllables in cs. All five informants. (Averages of two recordings.)

Table 3.

nasal/liquid		fricatives		plosives	
		unvoiced	voiced	unvoiced	voiced
short	nanos panel	nasyp našim našich	draho	zakop	zabal
long	nános dáma	násyp	dráhy	zákop	zábal
short	cena	nechat	sever	peče letec	cedit
long	kréma	-	réva	péče lékař léto	-
short	vina	bychom	lyže	syto syta	lidi ryba
long	vína míra	východ	líže	síto lítost mýtit lípa	líbit
short	doma domu	vosa	lože	copak	voda
long	dóma dómu	-	-	kóta	móda
short	suma	duše tušit	muže	kupec muka	tuba zubu
long	vůle	způsob	můze	důkaz	půda

Words used for the main part of the present investigation (as it is seen, some α- and β-words are included).

Table 4. Vowel duration in cs.

	a	e	i	o	u
JL	11,4	1o,3	9,3	11,2	9,9
OL	1o,7	1o,1	9,9	1o,9	9,2
LM	12,1	12,1	11,4	12,5	12,4
BU	1o,3	9,o	9,o	1o,7	9,4
AKK	11,6	1o,2	9,6	11,4	9,9
average:	11,2	1o,3	9,8	11,3	1o,2

	á	é	í	ó	ú
JL	22,4	21,6	18,o	22,2	18,9
OL	22,6	22,5	16,4	23,1	17,5
LM	31,8	33,7	3o,2	32,1	34,6
BU	22,o	22,2	19,2	2o,6	18,2
AKK	18,7	2o,2	17,3	18,3	16,2
average:	23,5	24,o	2o,2	23,3	21,1

As might be expected the close vowels are shorter than the more open ones.

The influence of the following consonant on vowel duration is seen in the following table.

Table 5.

	nasal/liquid	fricatives		plosives	
		unvoiced	voiced	unvoiced	voiced
short vowel	11,2	9,7	12,3	9,o	1o,4
long vowel	25,o	18,8	25,5	2o,2	22,o

The following consonant clearly exerts an influence on vowel duration. Vowels followed by unvoiced consonants are shorter than vowels followed by voiced consonants (and liquids/nasals).

The material is clearly divided into long and short vowels. There is no overlapping at all.

3. Conclusion.

As to the main question raised in this paper it seems difficult to stipulate a permanent ratio between long and short vowels. The informants' long vowels are two to three times as long as their short vowels, although the average seems to be about twice as long. Further investigations are necessary in order to decide whether this is valid for unstressed vowels as well.

According to Chlumsky's investigation from 1928 i and u are shorter than other vowels, especially shorter than a. Long vowels were found to be twice as long as the corresponding short ones, and the number of syllables was found to influence vowel duration in the same way as found in the present investigation. The following consonant does not, according to Chlumsky, influence the duration of the preceding vowel. This statement disagrees with the result reached in this investigation. Chlumsky has no conclusion about influence of long and short vowels on the duration of the preceding vowel (the β-words of this paper).

L. Kaiser quotes some measurements of the duration of Czech vowels spoken in isolation by 5 subjects. These measurements also show long vowels to be twice as long as short vowels in ordinary speech tempo, and open vowels to be longer than close vowels.

-----****-----

References:

(1) Josef Chlumsky, Česka kvantita, melodie a prizvuk (1928).

(2) L. Kaiser, "Phonetic Similarity Apart From Linguistic Affinity", Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung 17 (1964), pp. 243-249.

ACOUSTIC CUES FOR THE PERCEPTION OF ACCENT IN JAPANESE WHEN THE ACCENTED VOWEL IS DEVOICED

Hideo Mase

0. Introduction.

0.1. Phonologically or functionally speaking, Japanese accents are of so-called "register" type. There are only two contrastive registers, high(er) and low(er). In the standard dialect there is a distinction between words (or words with particles) containing a syllable with a high tone immediately followed by a syllable with a low tone, and words which contain no significant fall. In words of the former type the syllable with high tone is said to be accented. Both the place and the presence versus absence of a significant pitch drop are relevant. If, for example, a string of segmental phonemes constitutes three syllables we can distinguish three commutable accent units: (i) the first syllable is significantly high, (ii) the second syllable is significantly high, (iii) there is no syllable that is significantly high:

(i) /óoo/, (ii) /oóo/, (iii) /ooo/ (o = syllable*).

As there is only one significant pitch drop in a word the number of units that can be distinguished by accent depends on the number of syllables. If bisyllabic words are combined with a particle so that they form trisyllabic sequences, then obviously three different types can be distinguished as shown above, but if the particle is removed so that we get

(i) /óo/, (ii) /oó/, (iii) /oo/

it is hardly possible to distinguish (ii) from (iii) though it is of course possible to distinguish (i) from the others). We

---

*) Long nasals and voiceless geminated consonants are not counted as syllables in this paper, although they form morae of their own.

may say that potentially there are three different types of bisyllabic words (three different underlying forms), but hardly more than two different realizations.

<u>0.2.</u> Phonetically speaking the pitch of the first syllable is usually lower than that of the second syllable whenever the accent is not on the first syllable. This is true of /OÓ/-words as well as /OO/-words. However, the pitch difference is greater, i.e. the second syllable is a little higher, in the former type than in the latter, especially if the words are spoken with emphasis (Neustupný (12 a), p. 84, Han (3), p. 112). This difference may serve to establish a phonological contrast between bisyllabic words of the two types for some speakers who either keep up or try to keep up the difference. But this is probably rather rare (cp. Neustupný). It is my impression that there is generally no difference between (ii) and (iii) when words of the two types are spoken isolatedly.

<u>0.3.</u> We shall consider briefly the major phonetic aspects of Japanese accentuation:

<u>A. Pitch.</u>

The most important cue for accent production and perception seems to be pitch variation. According to Han (3) "a pitch rise of 1.9 semitones is perceived as an emic change, while a drop of 1.9 semitones is often not interpreted as an emic drop. This means that native speakers of Japanese are sensitive to a smaller pitch rise than to an equivalent pitch drop, and use a wider range of actual frequency change in signalling emic pitch fall." (p. 1o9), "the pitch change of more than 2 semitones functions as an emic change." (p. 128). I think that native speakers will be more <u>attentive</u> to the pitch drop than to the pitch rise, as far as accent is concerned, because the difference between fall and no fall is phonologically contrastive. (Kawakami (5) interpretes the low pitched first syllable as a manifestation of an intonation.)

Kobayashi (6) says that the degree of fall is greater for words with accent on the first or second syllable than for those with accent on the third or later syllable. In the former case, the high pitch gives an impression of 'strong stress'.

This observation of his agrees to some extent with the characteristics of English stresses, though the latter language is not a tone language. In English fundamental frequency (both pitch level and pitch variation) is reported to play a greater part than intensity. (Fry (2), Lieberman (8), Morton & Jassem (10), etc.)

The observation by Kobayashi means that the pitch fall is gradually decreasing the farther it is from the beginning of the accent unit. A similar statement has been made by other investigators, for example, by Kawakami (op.cit.), but he does not give any quantitative data on pitch variation. Shohara (14) says that only 2 % of the phrases with auditorily level tone have acoustically level tone.

B. Intensity and duration.

As for intensity and duration, Han (op.cit.) says

> "the higher pitch is usually accompanied by a slight increase in intensity and duration. ...if we keep the pitch constant, the change of intensity and duration does not affect the meaning of the word: therefore the pitch is identified as the distinctive feature and the accompanying intensity and duration as redundant features." (p. 104).

On the other hand, Neustupný (12 b) points out the importance of the intensity factor in addition to the pitch factor. He says that the pitch and intensity factors are relatively independent of each other, and that sometimes the intensity factor determines the accent pattern, though in general "pitch seems to be more important than intensity" (p. 1). I think that his statement should be taken with some reservation. He gives special mention to Fig. 4 [a⌈na⌉ta] (the pitch curve is taken with a kymograph, but it seems that the intensity curve is from a high-speed level recorder or some other instrument) and Fig. 11 [a⌈ta⌉takana] (taken by means of a pitch recorder). For the evaluation of these examples, the following considerations may be relevant: (1) Though the pitch of the first syllable in such a word as [a⌈na⌉ta] is usually lower than that of the second syllable (cf. 0.2.), it may not be the high pitch of the first syllable in his Fig. 4 but the clear pitch fall between the second and third syllables that is decisive for the accent perception. (2) There may be some influence of neighbouring consonants upon the fundamental frequency of the vowel, though Han (op. cit.) says

> "These differences are, however, minor and do not exceed 1 or 2 centiseconds in duration and 10 cycles per second in frequency. These are etic differences, and they are kept separate from the data pertinent to the structural analysis." (pp. 1o3-1o4).

House & Fairbanks, in their investigation of American English vowels (4), say

> "In the comparisons of voiced and voiceless consonant environments, vowels in voiced environments, with few exceptions, were longer in duration, lower in fundamental frequency, and greater in relative power." (p. 113).

In their article we find that the fundamental frequency of the vowel is considerably higher between [t]'s than between [n]'s, whereas the intensity of the vowel is considerably lower between [t]'s than between [n]'s. Accordingly, although I do not think the influence of the consonant upon the vowel is quite the same in English and Japanese, it is not improbable that the difference between the two consonants in Neustupný's examples, viz. [n] in Fig. 4 and [t] in Fig. 11 may have a certain influence upon the following vowel, whether the influence be subsidiary (or "etic" in Han's terminology) or not.

0.4. Devoicing of vowels.

Of the five Standard Japanese vowels /i,e,a,o,u/, /i/ and /u/ are often devoiced. Han (3, pp. 17-34) mentions four factors which cause devoicing of vowels:

(1) The intrinsic length of the five vowels: since /u/ is the shortest and /i/ the next shortest of the five vowels, /u/ is devoiced most easily and then /i/, while the other three vowels are rarely devoiced. (Both of them are, of course, high or close vowels.)

(2) Speech tempo: Except for slow speech tempo, the two vowels are easily devoiced. Slow tempo is that which is "used only when the speaker wishes to be precise, or when talking to a child or a foreigner, or in public speaking." (As for speech tempo, M. Temma (16) says in her study on the change in speech tempo within three generations that the first generation pro-

duces approximately 10 to 12 phonemes, 5 syllables, and 3 words per second, while the third generation speaks about 16 phonemes, 9 syllables, and 5.3 words per second. The third generation speaks 31 per cent, and the second generation 19 per cent faster than the first generation. This is the case of the Tokyo dialect, but almost the same tendency is seen in the Osaka and the Kyoto dialects.

(3) The pitch accent: vowels are easily devoiced except when they are 'accented'. Even if they are accented, they may sometimes be devoiced.

(4) The neighbouring consonants: the duration of the vowel is shorter between (generally also after) voiceless consonants than in voiced consonant environments. Of the voiceless consonants, [s] and [ʃ] give the strongest effect. (Nakano (11) reports that in fast speech voiced consonants may sometimes be devoiced, and then they can contribute to the devoicing of vowels, for example: [kab̥-to-tʃo·], [kab̥-ki-za] in the genuine Tokyo dialect, and [biʃi̥b̥iʃi̥], [batʃi̥b̥atʃi̥] (p. 8) (note that [b̥] = /b/, not /p/, [kab̥] = /kabu/).

A question is how the accent is perceived when the accented vowel is devoiced. (Sometimes phonological accent shifts occur, but this matter is out of concern here.) Han (3) says that "in this case where the vowel of the accented syllable is unvoiced, the very low pitch of the second syllable seems to signal the "virtual" high pitch of the preceding syllable." (p. 34).

Sugitow (15) says that in two-syllable words where the vowel of the first syllable is devoiced, the first syllable is perceived as having 'accent' when the pitch of the second vowel falls down very quickly, for example from 300 cps to 250 cps (a young female voice) after 1 centisecond. On the other hand the first syllable is perceived as not having 'accent' when the pitch of the second vowel goes down very slowly. (The informant is a young lady born in Osaka, but her mother was born in Tokyo. Accordingly, we cannot expect that this phenomenon is necessarily the same in the Standard dialect.

1. Procedure.

The investigation summarized here was undertaken in order to find some acoustic cues for the accent perception when the accented vowel is devoiced.

1.1. Texts and subjects.

I chose three-syllable accent units for my analysis. The text (written in Japanese script) consists of words with segmental phonemes: /hasi/ or /kaki/ followed by a particle: /wa/, /ga/, /o/, /mo/ or /to/ with different accent placement. The words were given in different orders, and both in isolated position and in short sentences. Only the words followed by the particle /to/ were used. In this environment the vowel /i/ in /hasito/ and /kakito/ is devoiced in speech of normal tempo. Both /hasito/ and /kakito/ make three commutable units by the difference of accent placement, namely,

Acc. 1: /hásito/; /kákito/ ("chopstick (and)"; "oyster (and)"),
Acc. 2: /hasíto/; /kakíto/ ("bridge (and)"; "fence (and)"),
Acc. 3: /hasito/; /kakito/ ("edge (and)"; "persimmon (and)").

(/to/ is also used in the meaning "to".)
Accent 1 words have accent on the first syllable, Accent 2 words have it on the second, and Accent 3 words have no accent, or, in other words, no significant drop.

The main subject person was a middle aged professor in natural science, who was born and has been living in Tokyo (M.O.). Two other subjects were born in Tokyo and Kanagawa respectively, and have been living in Tokyo. Both of them have completed university courses some years ago (K.T., H.M.). M.O. recorded isolated words 11 times in two sessions and words in sentences 6 times. Two other subjects recorded only isolated words 2 times. A short pause was made after the recording of one text.

1.2. Recording.

The recording was done in the sound-treated studio of the Institute of Phonetics, University of Copenhagen, in July and August, 1968.

1.3. Control of material.

In the speech of the main subject M.O., some words were spoken with different accents from the intended ones by mistake. (In two other persons' speech there was no confusion, but one occurrence of /kakíto/ (Acc. 2) in K.T.'s record was spoken with a different particle.) Besides, some Accent 2 and Accent 3 words have been spoken with a voiced second vowel, which is quite natural because the second syllable is rather high-pitched (higher in pitch than the first syllable). Specimens in which it was not clear which accent had been used, were left out of the material.

In order to find whether the intended accents can be perceived by other persons, all the words were rearranged in different order and given as a listeners' test. The subjects (4 in all) who listened to the test words are others than those who spoke (except H.M.). Each subject listened two times with a time interval of a week. (But a Japanese phonetician who visited Copenhagen listened two times in one day.) The 'perceived accent' is counted positive if the agreement of the answers is 7/8 or 8/8.

Table 1-a.

hasito perceived accents

intended accents		Acc. 1	Acc. 2	Acc. 3	?
Acc. 1	21	21			
Acc. 2	19		13 (+2)	1	3
Acc. 3	22		1	17 (+1)	3

kakito perceived accents

intended accents		Acc. 1	Acc. 2	Acc. 3	?
Acc. 1	27	27			
Acc. 2	19		13 (+2)	1 (+1)	2
Acc. 3	16			14	2

(The numbers in parentheses refer to words with voiced second vowel.)

Of kakito (Accent 2) spoken by K.T. only one example could be used in the table above. H.M. kakito (Accent 2) was spoken with voiced second vowel in both instances.

The number of words with devoiced second vowel is as follows:

Table 1-b.

hasito

	s.	w.	total
Acc. 1	6	15	21
Acc. 2	4	1o	14
Acc. 3	4	14	18

kakito

	s.	w.	total
Acc. 1	7	2o	27
Acc. 2	4	9	13
Acc. 3	4	11	15

(s. = words in sentences. w. = single words.)

1.4. Registration.

Mingographic curves were taken of all the words. Some spectrograms were also taken for the delimitation of segments (and in order to see the formant structures). Pitch contours of some words were also taken by means of a Schneider's Tonhöhenschreiber.

On the mingogram were registered (1) duplex oscillograms, (2) intensity curve (logarithmic, integration time 5 ms, HP-filtering with cutoff at 500 cps, (3) intensity curve (linear, integration time 10 ms, no filtering), and (4) fundamental frequency curve. As for the spectrograms, wide-band, narrow-band and section spectrograms were taken.

1.5. Methods of measurement.

All data were obtained from mingograms. Spectrograms and pitch curves made with the Tonhöhenschreiber were used only to control the mingographic curves.

Length of the vowel: the vowel begins at the point where regular vibrations are seen on the oscillogram. On the intensity curves, it begins at the rising point of intensity. The vowel ends at the point where the amplitude of vibrations, the intensity, and the fundamental frequency all go down abruptly. In most cases the beginning and end points of the vowel as defined above agree on the four curves within the accuracy of $\pm$ 0.25 cs. Consonants and devoiced vowels of the second syllable are measured according to the same principle. The duration of the open phases of consonants has been measured mainly on the logarithmic intensity curve. As for the intensity, I have measured peak values only. The fundamental frequency curve shows a certain irregularity during the first 0 - 1 centiseconds (and sometimes up to 2 or 3 cs), so that the very beginning point is not measured. The accuracy of measurement is ± 0.25 dB for the intensity and ± 2.5 cps for the fundamental frequency, provided that the measuring scales are accurate.

2. Results.

The material is too small in size to be treated statistically; in most cases, therefore, all pertinent data are shown in figures and tables.

2.1. Words with voiced second vowel.

These have not yet been analyzed, but they will soon be, together with other words such as hana + particle and turu + particle. According to the traditional description one should expect

Acc. 1: high - low - low
Acc. 2: low - high - low
Acc. 3: low - high - high,

but a look at the curves for hasi and kaki with voiced second vowel before particles reveals the following tendencies:

Fundamental frequency:

Acc. 1: the second vowel is high falling rather than mid or low falling.

Acc. 2: the third vowel is not low but falling from mid to low or from high to low.

Acc. 3: the second vowel is often lower than in Accent 2 words (cp. 0.2. above).

Intensity:

Acc. 1: the first vowel is relatively stronger than in Accent 2 and Accent 3 words.

Acc. 1 & 2: the intensity of the third vowel decreases abruptly.

Acc. 3: the third vowel is relatively strong, or keeps a certain level of intensity longer than with Accent 1 & 2.

2.2. Words with unvoiced second vowel.

2.2.1. Length.

The subjects of my experiment belong to the second generation and to the generation between the second and the third generation. The words spoken by three subjects are between 30 and 50 centiseconds in length. This means that the subjects spoke the words with a normal speech tempo where devoicing of vowels occurs (cp. 0.4. above).

The duration of words is considerably different among different recordings of the same person and among different persons. The mean value of segments and the percentual distribution of segments are given in Table 2 and Figs. 2 and 3, respectively.

The open phase of medial [t] is usually between 1.5 - 2.5 cs in length. It seems that there is no great difference in the duration of the open phase according to word accentuation or speaker. The open phase of initial [k] in kakito shows rather interesting results. The open phase is shorter before vowels with greater intensity and longer duration (e.g. the first vowel of kákito (Acc. 1)) than before the vowels in the syllable ka of Accent 2 and Accent 3 words, which are shorter in duration and weaker in intensity. If we measure the combined length of open phase + vowel, the difference among three differently accented words with identical segmental structure (i.e. kakito) becomes smaller (see Fig. 1). The reason may be that the first vowel occurs between two voiceless consonants (i.e. k - k), so that when the vowel is unaccented, the tendency of devoicing of the vowel is greater for Accent 2 and Accent 3 words than for Accent 1 words. (NB.: It is quite rare that the vowel /a/ is devoiced completely. Cf. 0.4.)

As for vowel length, the first vowel of Accent 1 words tends to be a little longer than those of Accent 2 and Accent 3 words. It may mean that the accented vowel has a slightly longer duration than the unaccented one. The duration of the third vowel does not seem to differ so much. In hasito the third vowel is often shortest in Accent 1 words (hásito), and often longest in Accent 3 words (hasito). In kakito the third

Table 2. Average Length.

hasito

hasito	Accent type	MO. sentence	MO. word-1	MO. word-2	KT. word	HM. word
whole length minus initial consonant	1	40.4	40.8	34.6	51.3	45.3
	n.	(6)	(6)	(5)	(2)	(2)
	2	39.1		32.7	48.0	45.8
	n.	(4)		(6)	(2)	(2)
	3	40.6	42.1	32.5	50.0	48.8
	n.	(4)	(6)	(3)	(2)	(2)
[a]	1	8.0	7.8	6.7	9.0	8.3
	2	5.9		5.6	5.8	6.5
	3	5.8	6.5	4.8	7.3	7.3
[ʃi̥]	1	11.9	14.7	11.3	16.8	18.3
	2	11.9		10.5	17.5	18.3
	3	13.0	14.7	10.8	15.8	17.5
[t]	1	10.4	11.5	11.0	18.5	13.0
	2	9.9		9.6	15.0	14.3
	3	11.3	11.4	10.2	16.0	14.8
[o]	1	10.1	6.8	5.6	7.0	5.8
	2	11.6		7.0	9.8	6.8
	3	10.6	9.4	6.6	11.0	9.3

kakito

kakito	Accent type	MO. sentence	MO. word-1	MO. word-2	KT. word	HM. word
whole length minus initial consonant	1	43.1	44.6	36.1	50.0	43.0
	n.	(7)	(6)	(8)	(2)	(2)
	2	38.8	41.5	32.9	50.0	
	n.	(4)	(4)	(4)	(1)	
	3	38.8	44.8	36.0	50.5	43.8
	n.	(4)	(6)	(1)	(2)	(2)
[a]	1	8.0	7.5	6.6	7.0	5.8
	2	5.9	6.0	5.6	7.0	
	3	5.9	6.1	5.5	5.3	4.5
[ki̥]	1	13.1	16.3	13.1	18.0	16.3
	2	13.9	16.4	10.5	16.0	
	3	12.7	16.8	12.5	20.0	17.5
[t]	1	10.9	13.1	11.3	15.5	14.8
	2	9.9	10.8	11.1	17.0	
	3	10.2	12.0	11.0	14.3	12.3
[o]	1	11.2	7.6	5.1	9.5	6.3
	2	9.1	8.4	5.6	10.0	
	3	10.0	9.8	7.0	11.0	9.5

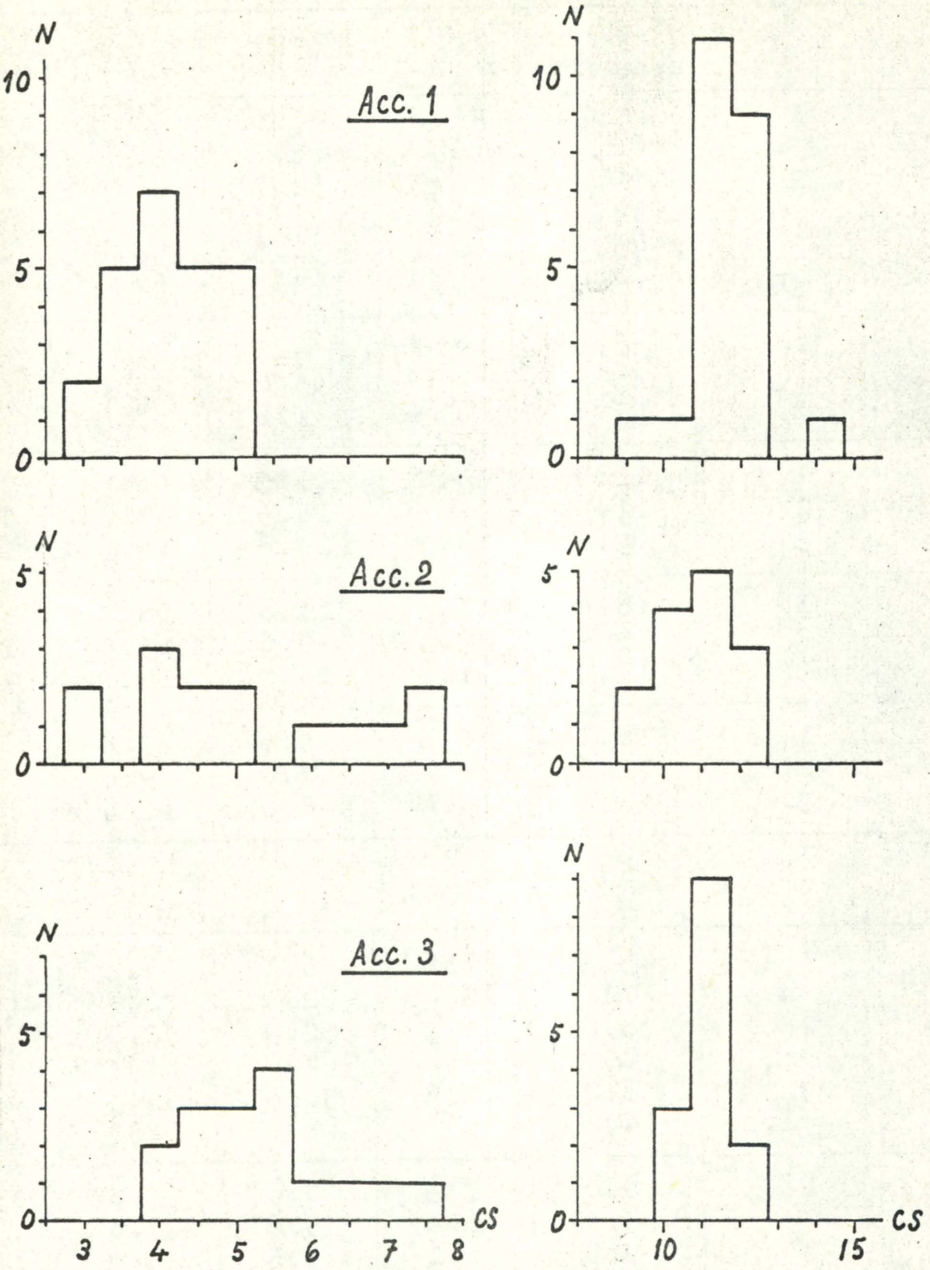


Fig. 1. Lengths of open phase of initial k (left) and of open phase of k + 1st vowel (right).

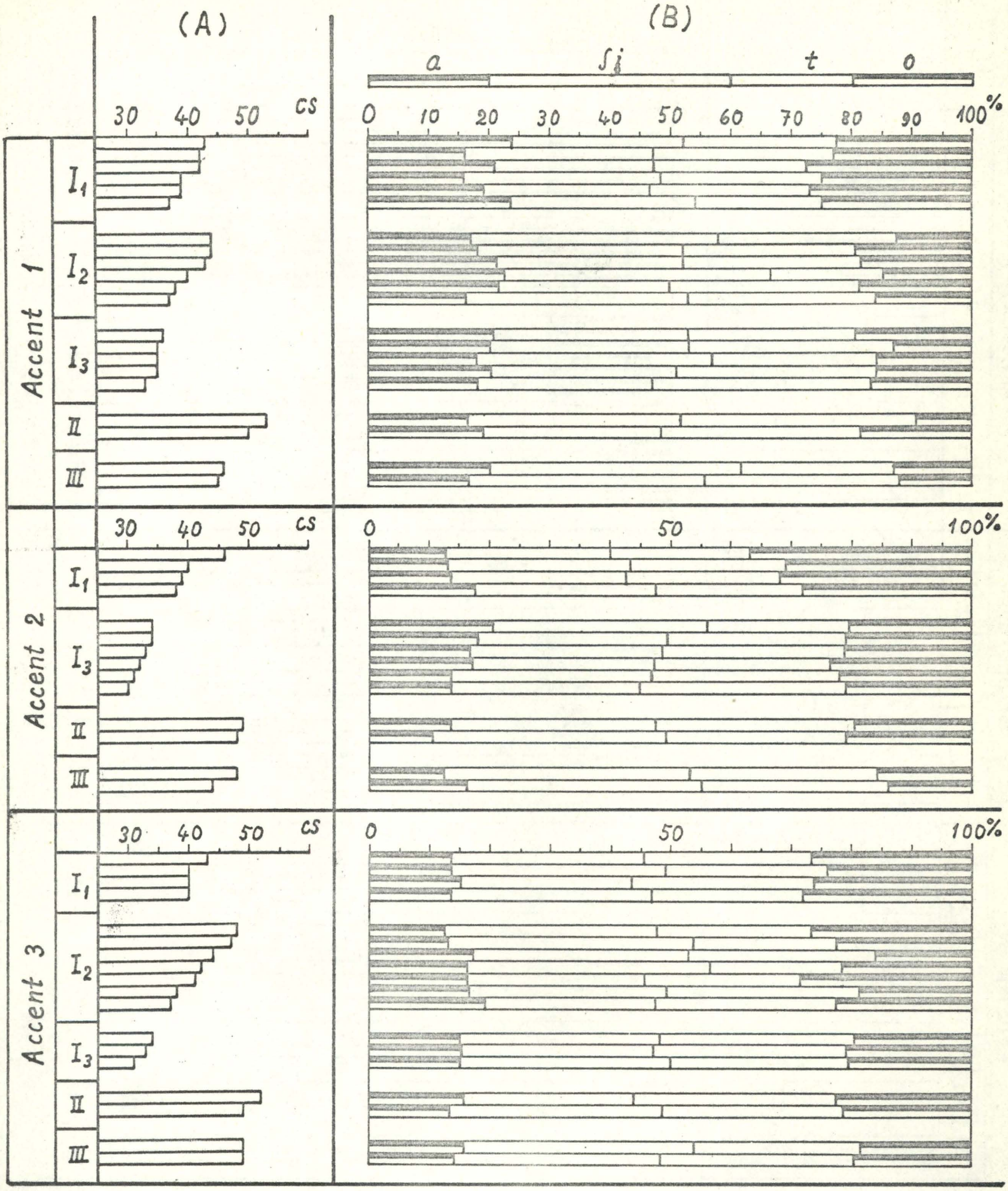


Fig. 2. Length (<u>hasito</u>)

(A) Length of the whole word (without initial h).

(B) Percentage length of segments in the word.

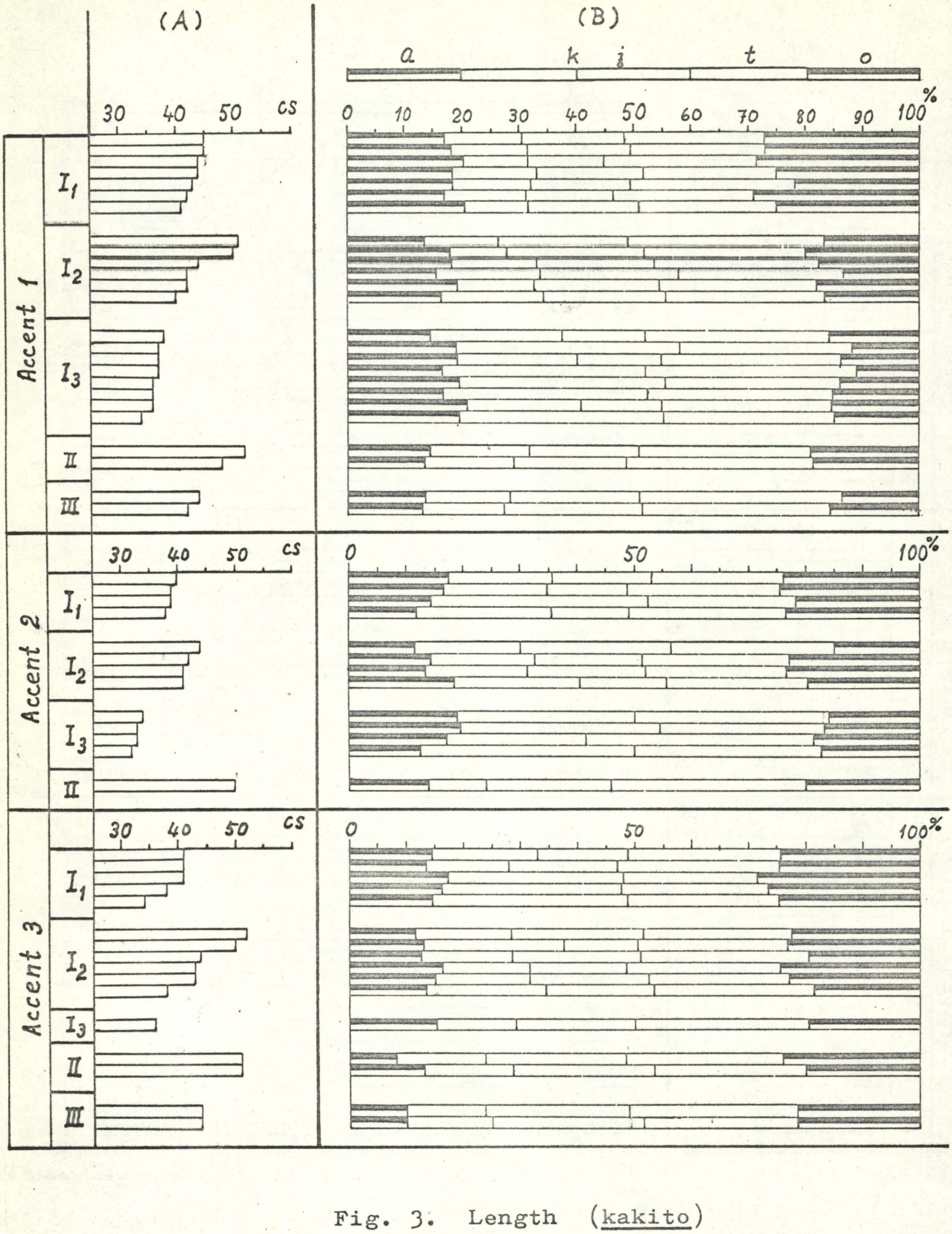


Fig. 3. Length (<u>kakito</u>)

(A) Length of the whole word (without initial k).

(B) Percentage length of segments in the word (the beginning of the syllable ki is here defined as the point of explosion).

vowel of Accent 1 (<u>kákito</u>) and Accent 2 (<u>kakíto</u>) words is often shorter than that of Accent 3 words (<u>kakito</u>), as one might expect. The difference, however, is not great. The duration of the second syllable does not seem to show any significant difference, either.

On the whole, what seems rather constant is that the accented first vowel of Accent 1 words (<u>hásito</u>, <u>kákito</u>) is a little longer than the unaccented first vowel of Accent 2 and Accent 3 words.

<u>2.2.2. Intensity (see Figs. 4-6).</u>

The intensity of the accented vowel of Accent 1 words (<u>hásito</u>, <u>kákito</u>) is a little higher than that of Accent 2 and Accent 3 words. Roughly speaking, the difference amounts to some 6 dB. The weaker vowel of Accent 2 and Accent 3 words tends to have its intensity peak in the first half of its duration, and the intensity falls down quickly, but not always. The relation between the first and the third vowels seems a more stable feature. As is seen in Figs. 4 and 5, the relative intensity of the third vowel compared to the first vowel of Accent 1 words is smaller compared with those of Accent 2 and Accent 3 words. (I should like to emphasize here that this is not a comparison between the inherent intensity of /a/ and /o/, but a comparison of the relation between the first and the third vowels in each type of words.) This difference in the intensity relation between the first and the third vowels, together with the duration factor, may contribute to some extent to distinguish accent placements.

The intensity of the second syllable is almost the same in all types of words. The position of the intensity peak of [ʃi̥] in <u>hasito</u> varies according to which part is stronger, the [ʃ] part or the [i̥] part. On the whole, I have not found any constant difference in peak distribution among the three types of words, but one can observe a slight tendency for the peak in Accent 1 words to come earlier and fall down more quickly than in Accent 2 and Accent 3 words. The intensity peak of the second syllable (either the consonant part or the devoiced vowel part), when compared to that of the first vowel, is lowest in

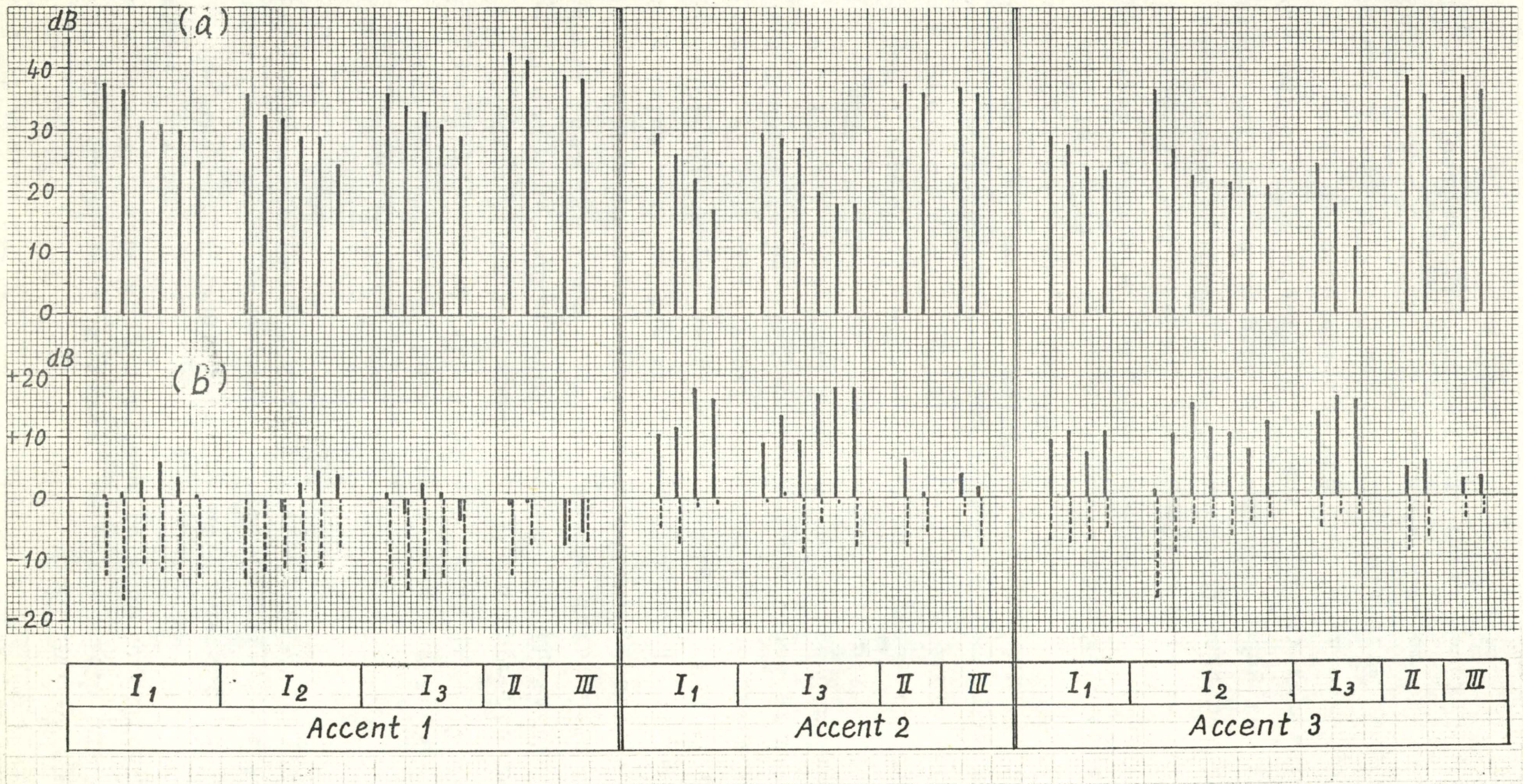


Fig. 4. Intensity (Peak Value) of <u>hasito</u>

(A) Intensity of 1st vowel.

(B) ——— 3rd vowel minus 1st vowel,

----- [ʃi̥] minus 1st vowel.

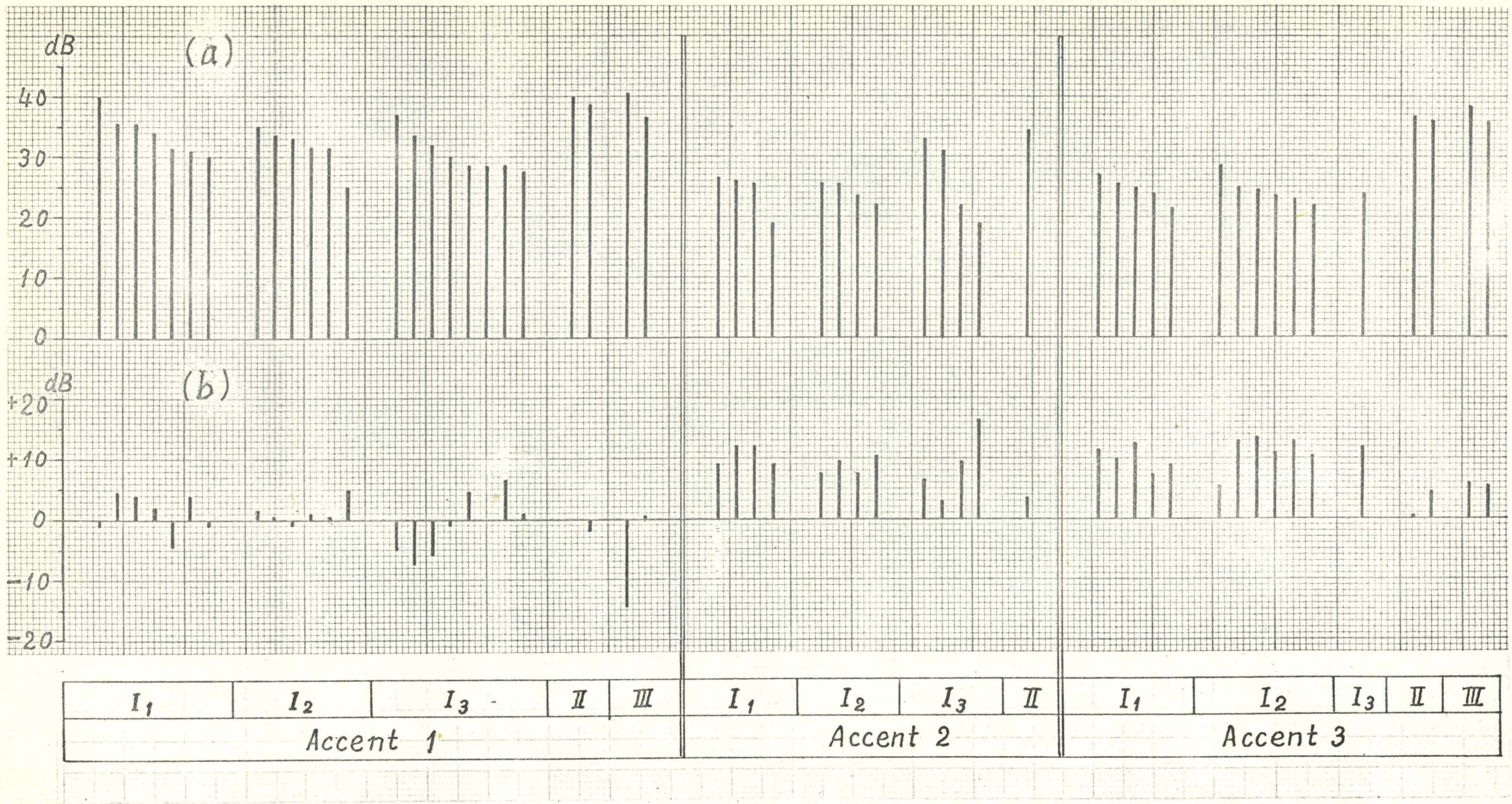


Fig. 5. Intensity (Peak Value) of <u>kakito</u>

(A) Intensity of 1st vowel.

(B) 3rd vowel minus 1st vowel.

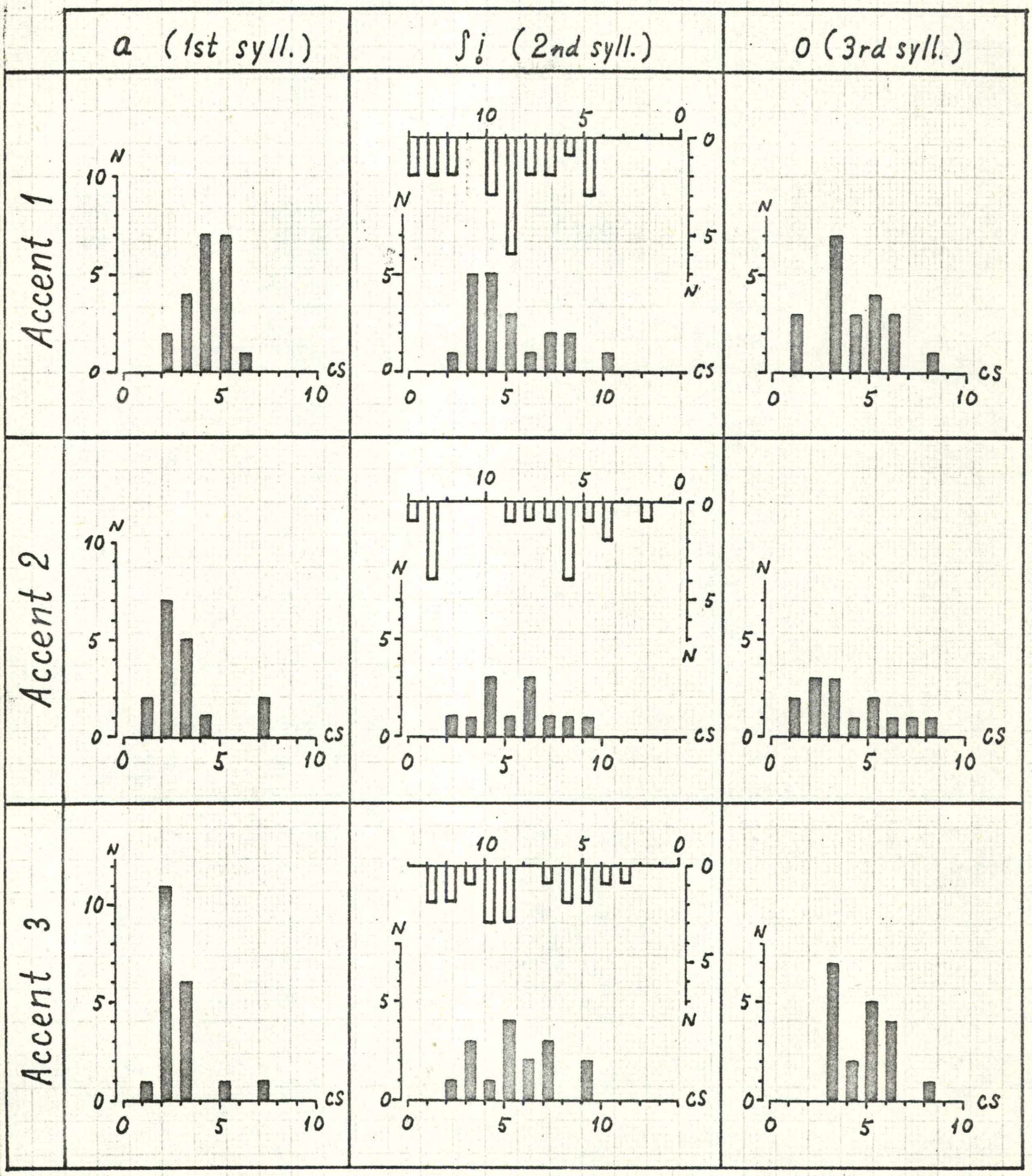


Fig. 6. Position of intensity peak (hasito)

Distance from the beginning (▬) and the end (▭).

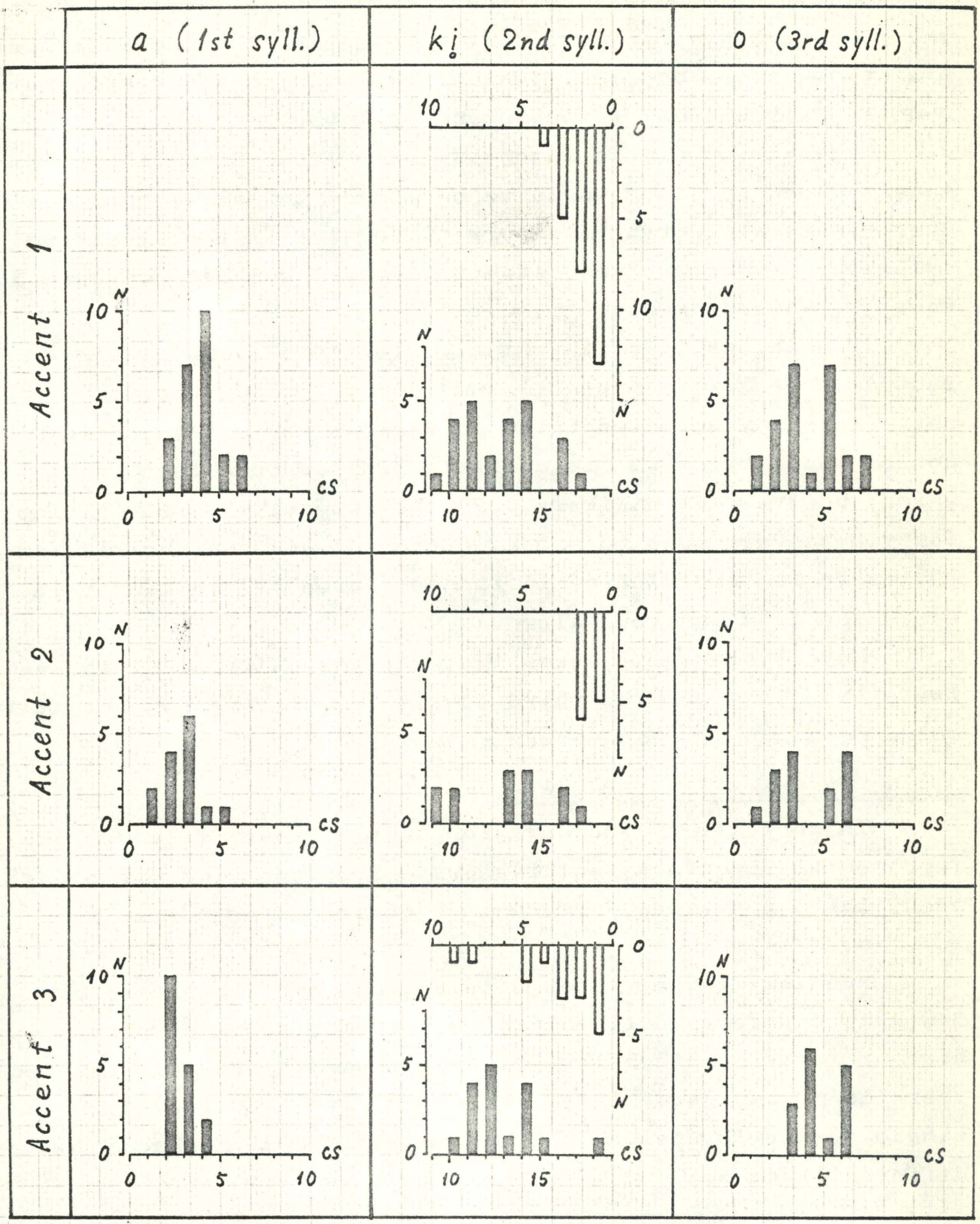


Fig. 7. Position of intensity peak (kakito)

Distance from the beginning (▬) and the end (▭).

Accent 1 words. The intensity of [ki̥] in kakito was sometimes too weak to be measured precisely. As a whole, it holds true that the peak often occurs just before the end of the second syllable. (The intensity peak of the open phase of k is not counted as the peak of the second syllable.)

We can expect the intensity of the second syllable of Accent 2 and Accent 3 words to be almost the same, because there is no noteworthy auditory difference in the first two syllables of Accent 2 and Accent 3 words, whether the second vowel be voiced or not (cp. 0.2.).

The intensity of the final, i.e. third vowel was expected to be low in Accent 1 and Accent 2 words, but it was not. This may be interpreted in various ways: (1) the inherent sonority of [o] is a little greater than that of [a] which, however, is unlikely, (2) the intensity level may not contribute very much to accent differentiation after all, and (3) there is some kind of influence caused by the manner in which the text was read aloud.

Nevertheless, it seems that the intensity relation between the first and the third vowels can distinguish Accent 1 words from Accent 2 and Accent 3 words.

2.2.3. Fundamental frequency.

In contrast to the duration and intensity factors, the fundamental frequency factor shows a very positive difference among the three types of words. A clear difference is seen in Figs. 8-14.

Needless to say, though the pitch level, high(er) or low(er), has certain fixed ranges in which it is perceived as high or low, the pitch level is not absolute but relative, so there is a considerable difference in pitch level even in one single speaker's speech, not to mention the differences among different speakers.

If I may jump to the conclusion or result, I find the following tendency to distinguish the three types of words by pitch:

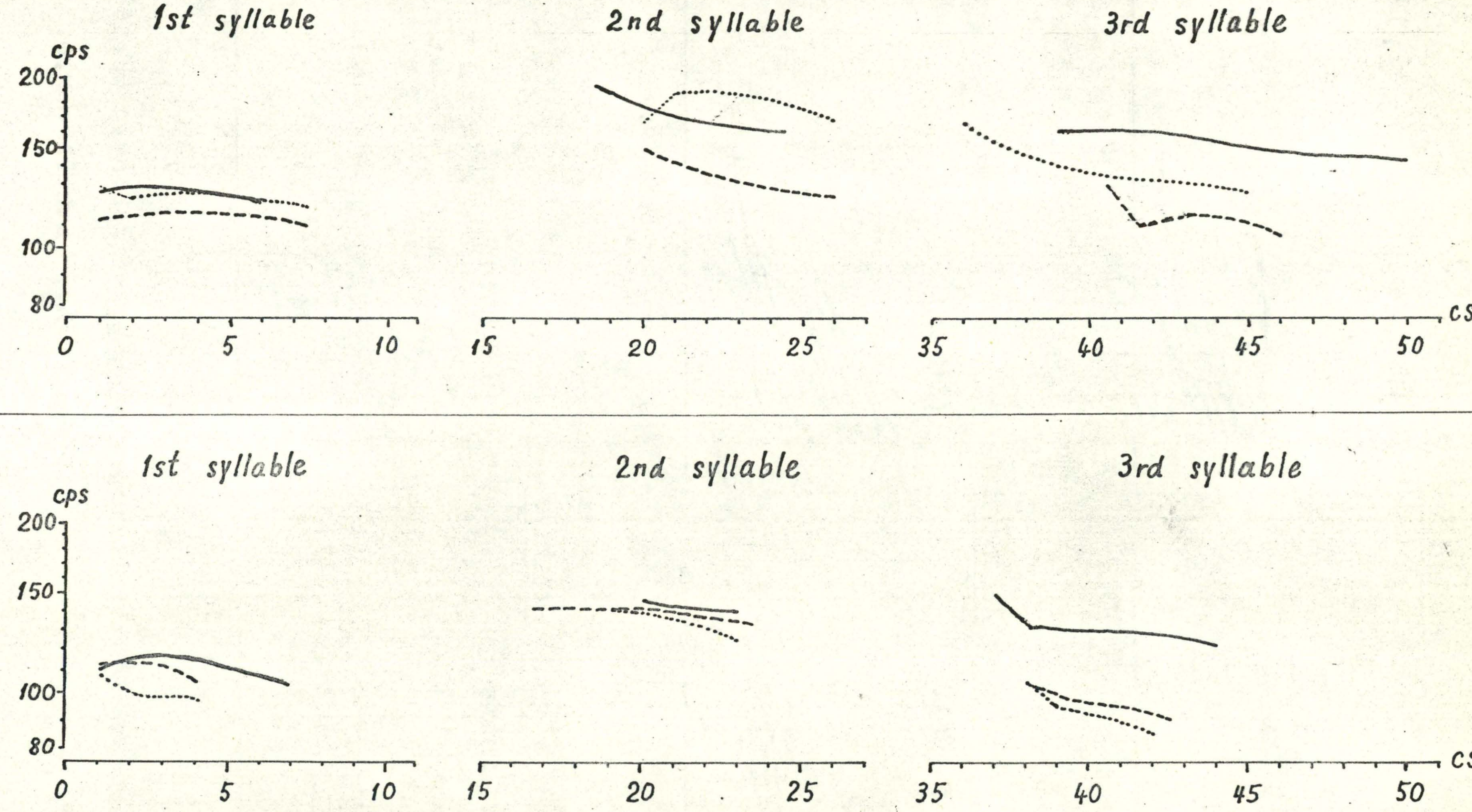


Fig. 8. Pitch contours of words with voiced 2nd vowel.

Upper curve: hasito ··········, ------ Acc. 2 (M.O.), —— Acc. 3 (M.O.)
Lower curve: kakito ··········, ------ Acc. 2 (H.M.), —— Acc. 3 (M.O.)

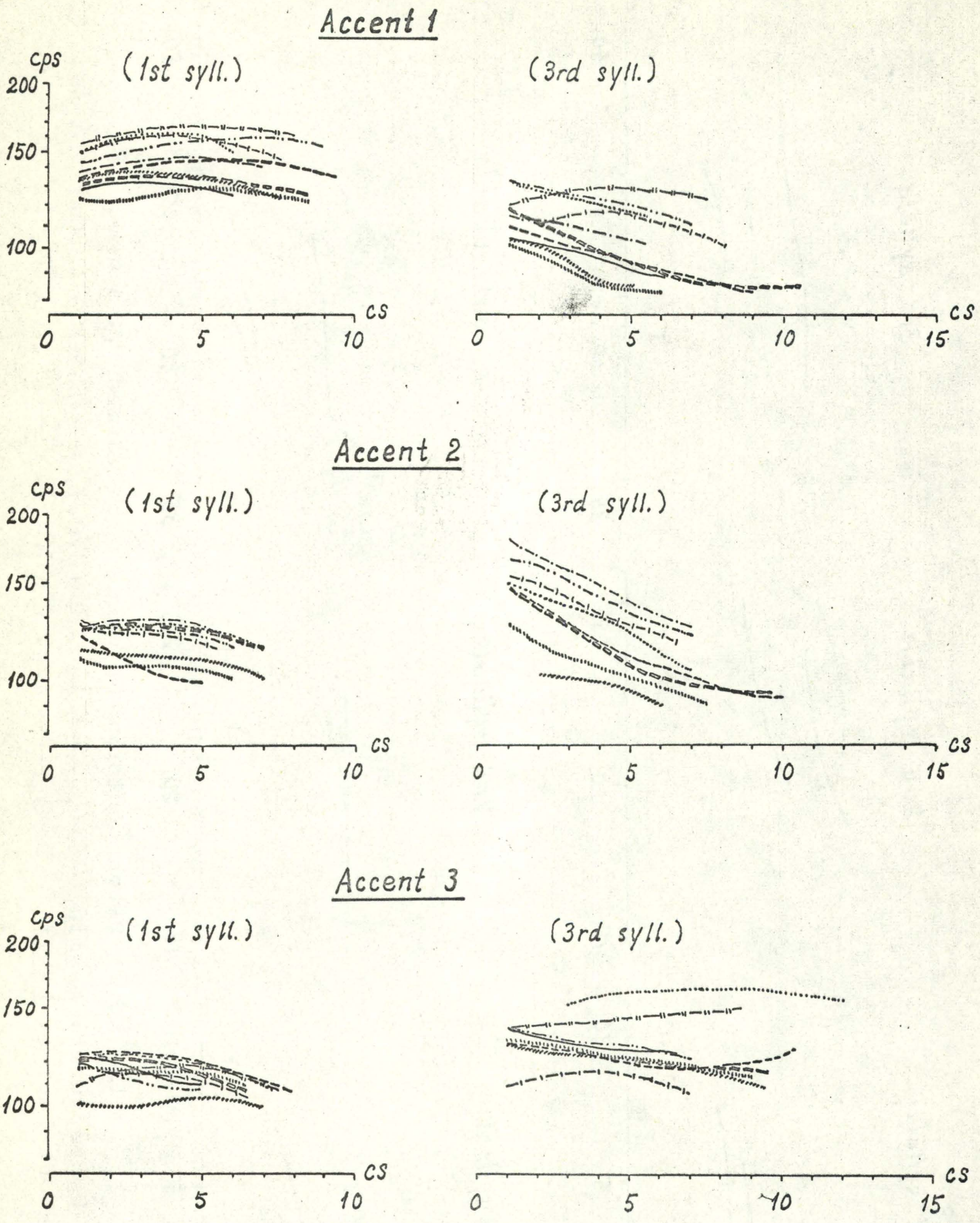


Fig. 9. Pitch contours of <u>hasito</u> (isolated words)

▭▭▭▭ ; - - - - = K.T., ////////// ; ########## = H.M., others = M.O.

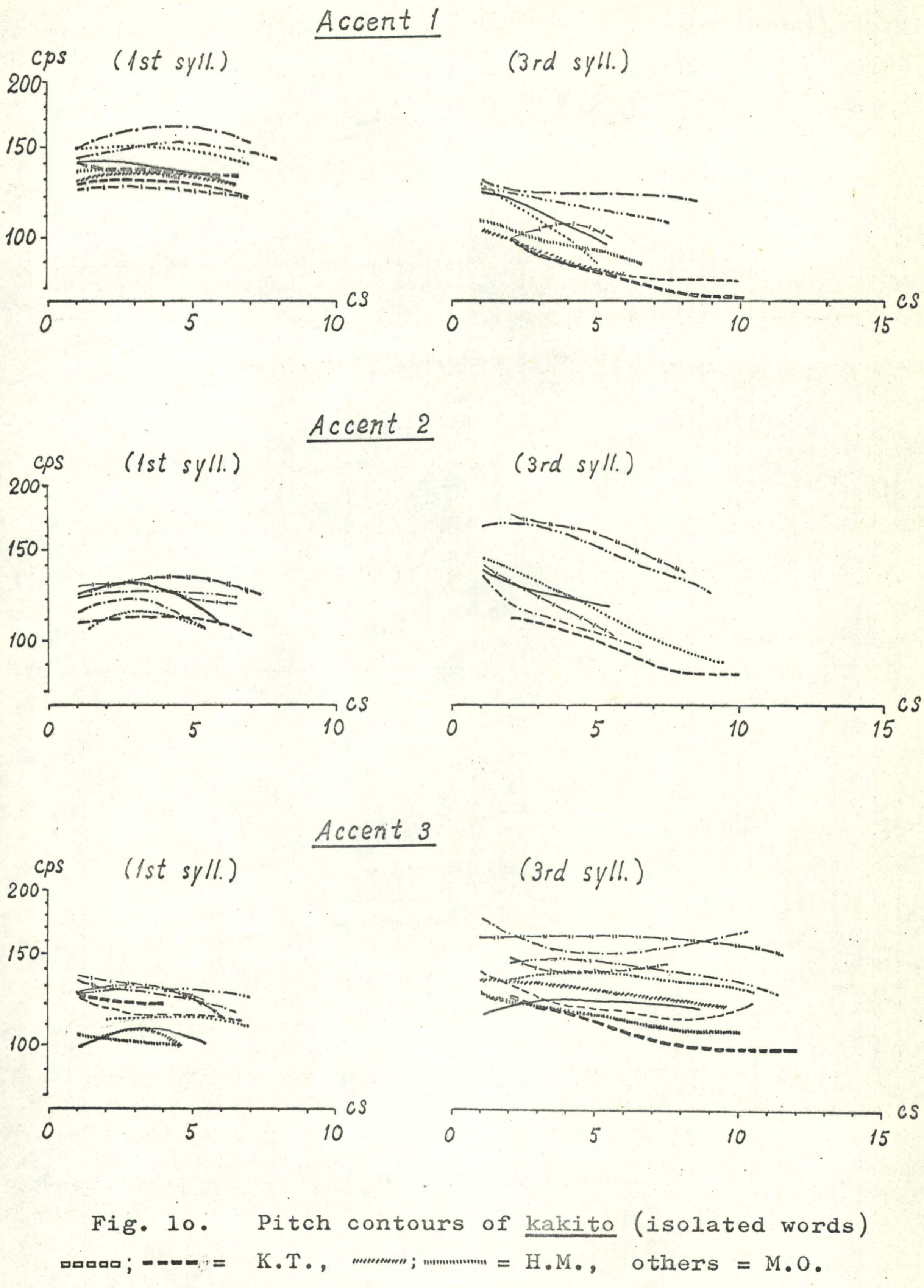


Fig. 1o. Pitch contours of kakito (isolated words)
▭▭▭▭▭ ; ▬▬▬▬ = K.T., ⋯⋯⋯ ; ⋯⋯⋯ = H.M., others = M.O.

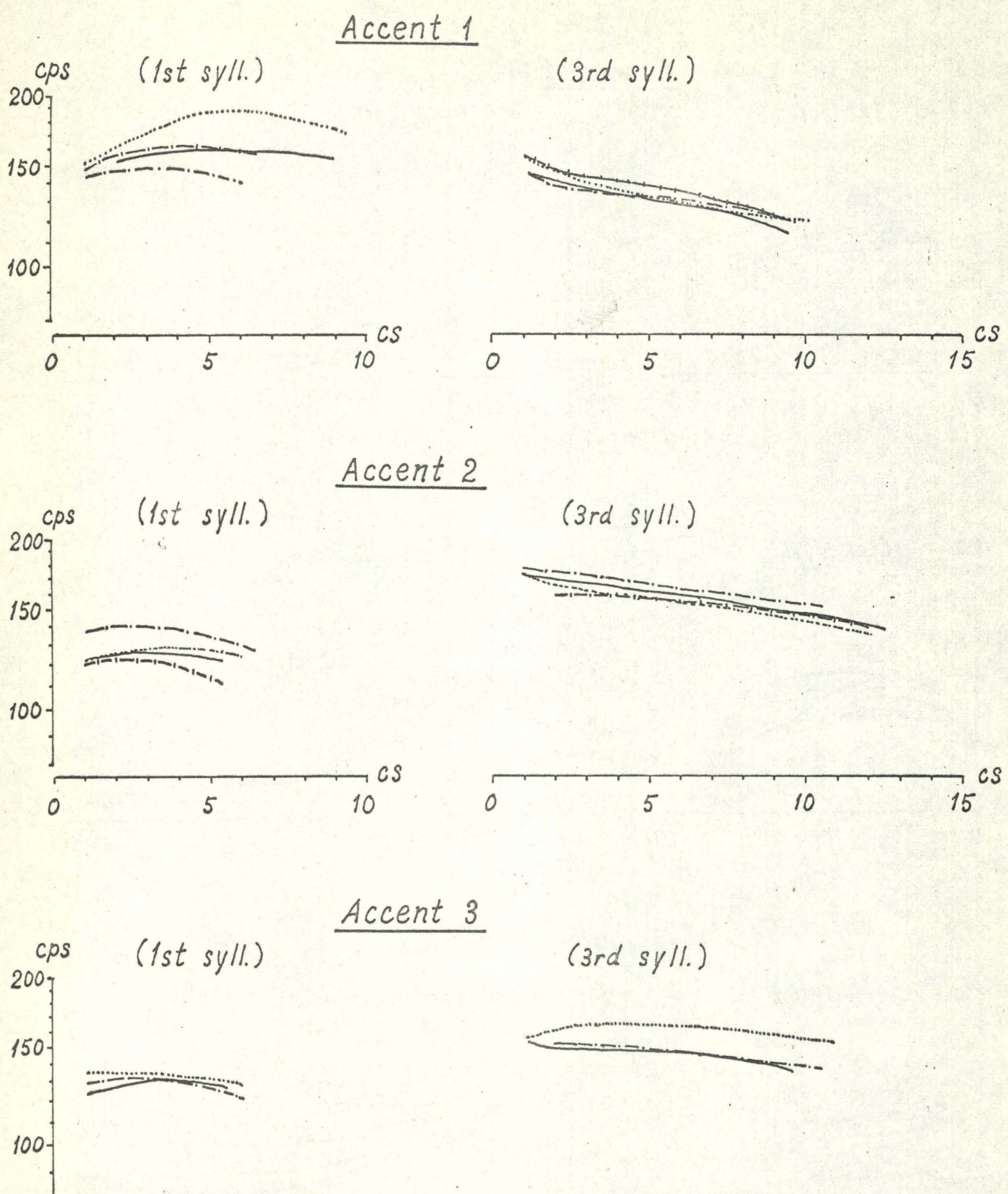


Fig. 11. Pitch contours of hasito (M.O. in sentences).

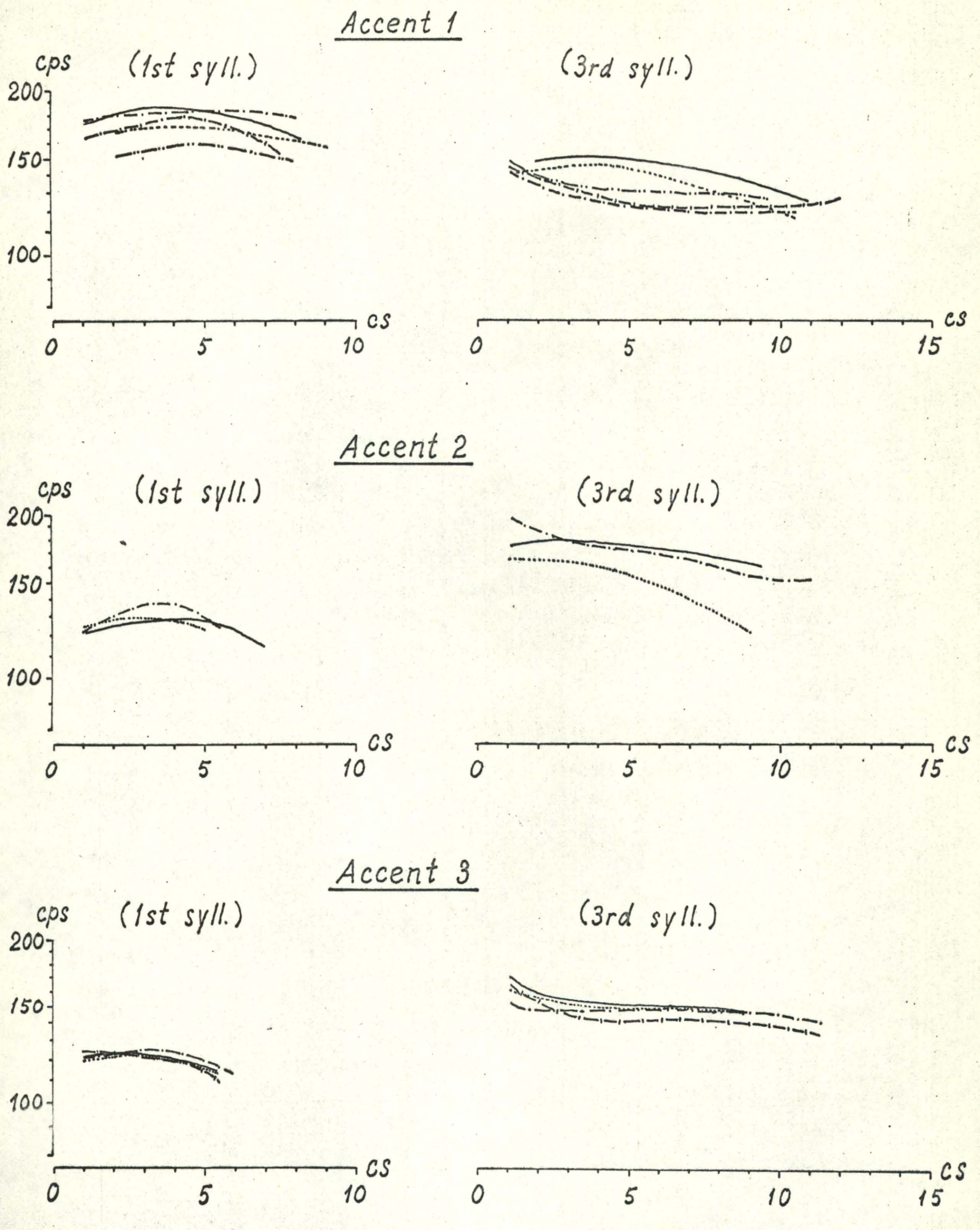


Fig. 12. Pitch contours of kakito (M.O. in sentences).

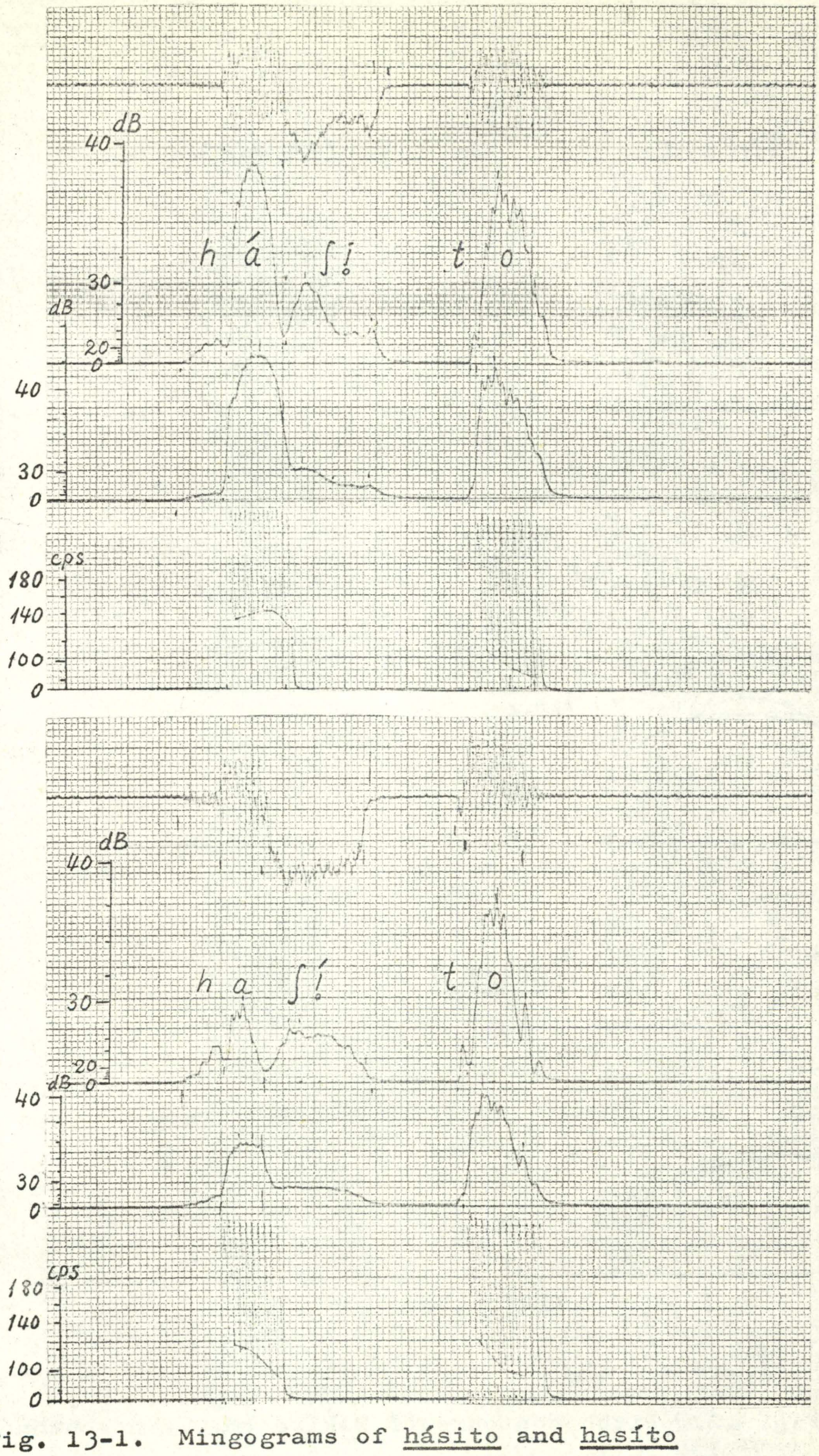


Fig. 13-1. Mingograms of <u>hásito</u> and <u>hasíto</u>
(K.T., isolated words)

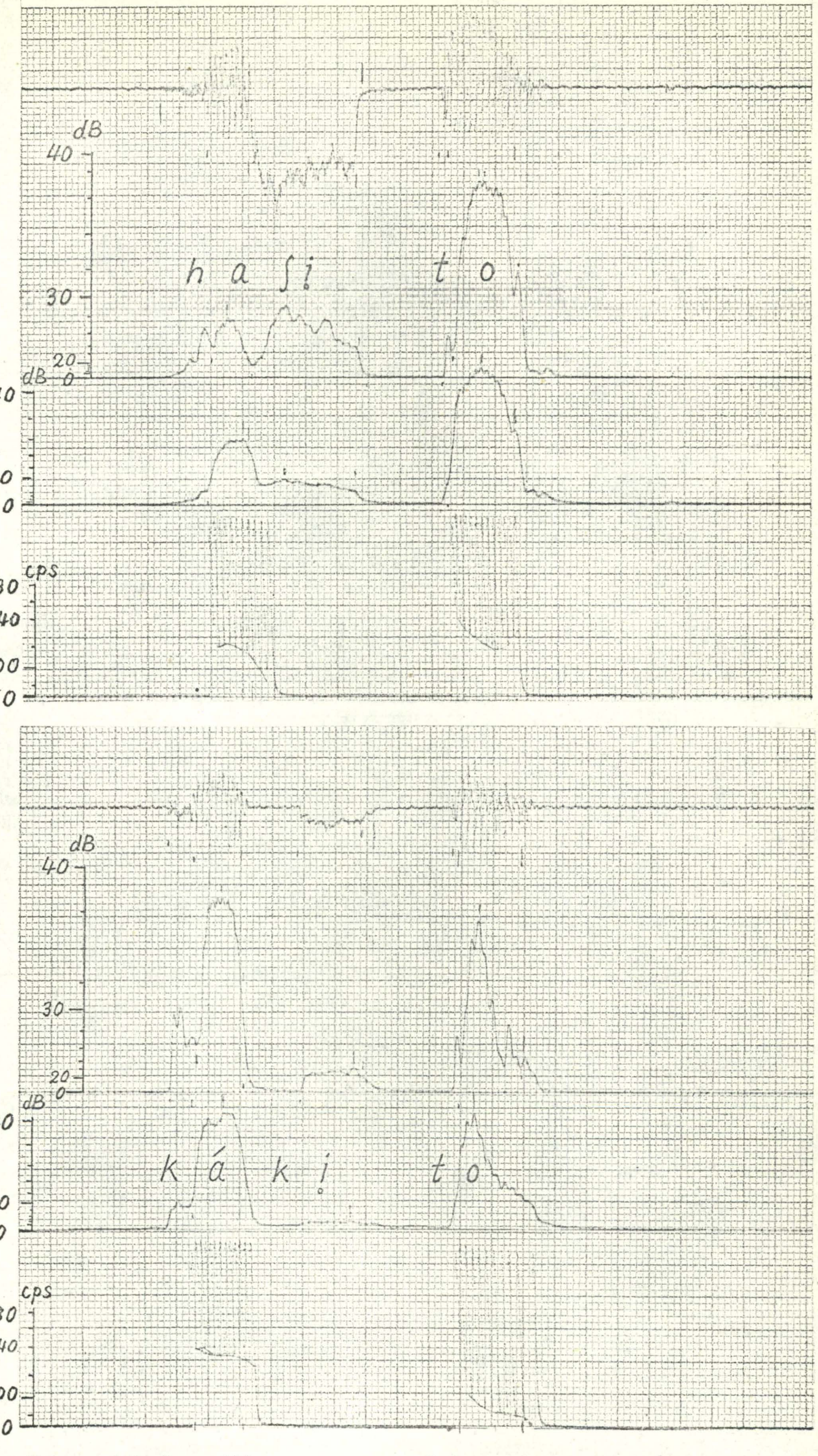
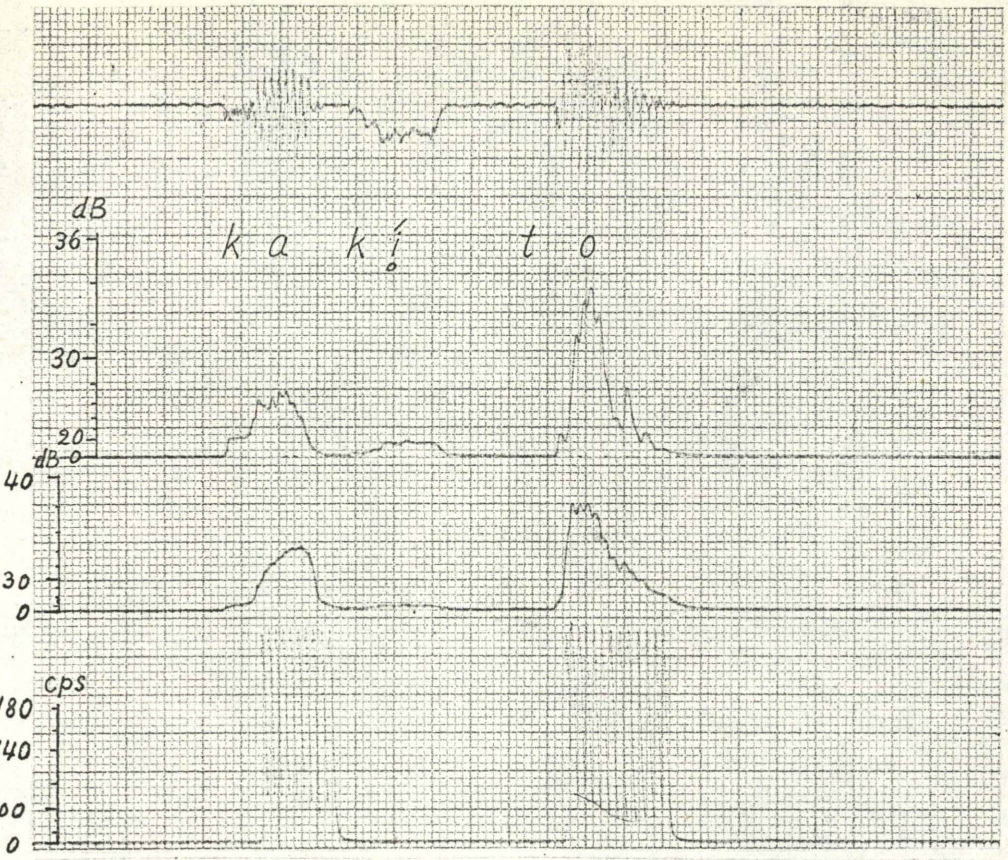


Fig. 13-2. Mingograms of <u>hasito</u> and <u>kákito</u> (K.T., isolated words)



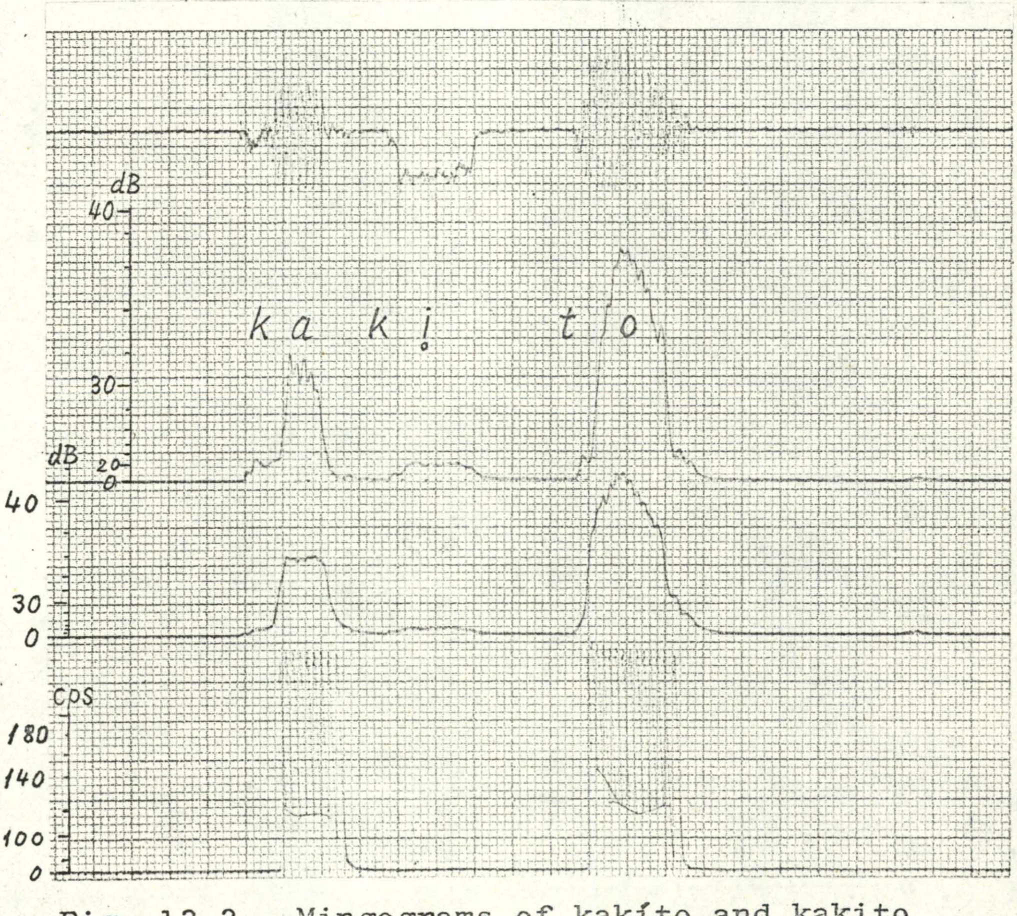


Fig. 13-3. Mingograms of <u>kakíto</u> and <u>kakito</u>
(K.T., isolated words)

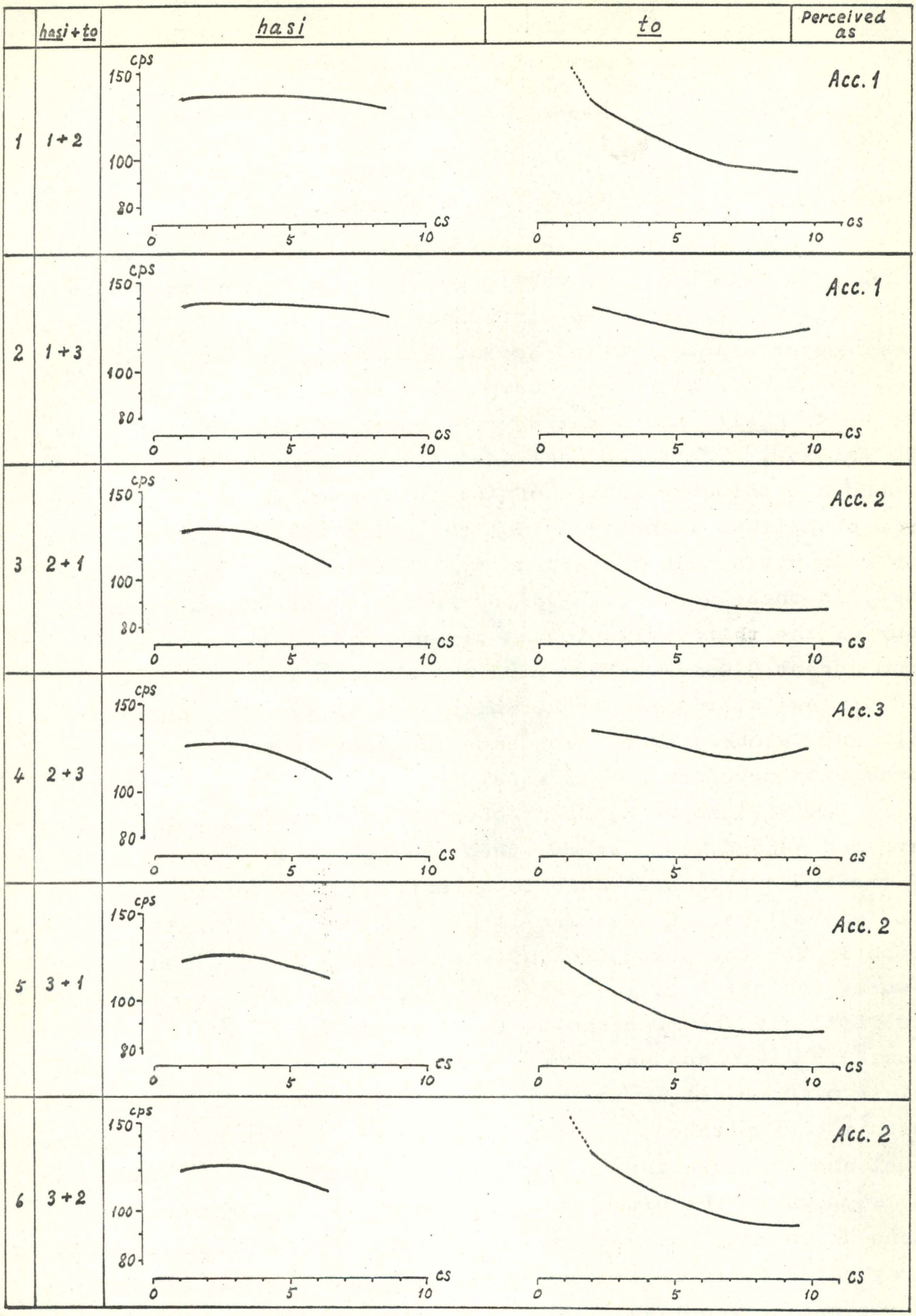


Fig. 14. Perception of accent in words after tape-cutting and recombination.

	1st syllable	3rd syllable
Acc. 1:	high	low falling;
Acc. 2:	low	high (or mid) falling, sometimes, low falling);
Acc. 3:	low	high even.

A problem is how to quantify the difference. In the words with voiced second vowel, we find the following tendencies (see Fig. 8): there is no difference in pitch level of the first vowel between Accent 2 and Accent 3 words in these examples (but cp. 2.1.). Among the examples of Accent 2 words one instance of <u>hasito</u> has a third vowel which is not very low, but mid falling. The functional identity of level between the second and third syllables of Accent 3 words is not reflected by a phonetical identity, since the third syllable is a little lower in pitch than the second syllable (cp. 0.2.). In any case, in these words with voiced second vowel the pitch contour of the third syllable can clearly distinguish Accent 2 from Accent 3 words. The third syllable of Accent 2 words is mid or low falling in pitch, while that of Accent 3 words is high and (almost) even. The same tendency is seen in the words with devoiced second vowel.

The difference between Accent 2 and Accent 3 words with devoiced second vowel seems, then, to lie in that the pitch of the third syllable exhibits a fast and considerable drop in the former type. A question is how important this falling pitch is for the perception of the accent. In isolated Accent 2 words the pitch is lowered about 5 cps per centisecond within the first five centiseconds. In a few cases the falling rate is about 3.5 - 4.0 cps per centisecond, but in this case the absolute pitch level is the same as or lower than the first vowel which is low pitched. Accent 3 words have a high pitched third vowel and the lowering of pitch is less than 10 cps. Mieko Han (3) says that "the pitch change of more than 2 semitones functions as an emic change" (p. 128). This may be quite crucial. In my material it is true of words with and without devoicing of the second vowel that the pitch change within the third syllable of Accent 2 words exceeds two semitones, and besides the pitch ends at a lower level than that of the first syllable. In some cases, the third vowel has low falling pitch. As for the Accent 3 words, all the third vowels have even pitch contours. Their pitch levels are, except for one instance of hasito, higher than those of the first syllable.

As mentioned just above, the pitch contour of the third vowel is similar in words with voiced second vowel and those with devoiced second vowel. (The amount of pitch drop is a little greater in the latter case.) Then, a question is raised: which is more important for the accent perception, the pitch contour of the accented syllable, or that of the syllable following after the accented syllable? However, it is evident that the relation between the accented syllable and the immediately following syllable is most important. Words with devoiced vowel should probably be considered a special case (see, however, 3.2. below).

Accent 1 words do not seem to pose any problem. The accented first vowel is high in pitch, as well as longer in duration and greater in intensity, than the third vowel.

A little more problematic is the pitch contour of the words spoken in a sentence. The general tendency is the same as for isolated words. The pitch contours of Accent 1 words are just as clear as in the case of isolated words. The problem is that there is no great apparent difference between Accent 2 and Accent 3 words, especially in the case of hasito. This problem has not yet been completely solved. The formant structure of these words may give some cues, but minute spectrographic analyses have not yet been done.

2.4. Tapecutting experiment.

Three examples of hasito with different accents were chosen. They have almost the same intensity level and duration of segments. They were cut on the tape between hasi and to at a point 5 centiseconds before the explosion of t and treated in three alternative ways: (a) the final to was removed, (b) a 50 cs pause was inserted between hasi and to, and (c) different parts of hasi and to were combined. The results of a small perception test (three subjects) show that

a) when there is no final syllable Accent 1 is easy to perceive, while there is no difference between the parts of hasi of Accent 2 and Accent 3,

b) when there is an internal pause Accent 1 is again easy to perceive. Accent 2 and Accent 3 are also perceived correctly. (It seems that the timing between syllables does not influence accent perception in Japanese. However, it is a matter of course that the information given by the third syllable is lost if the distance between the second and the third syllables is too great, which fact may be reflected in the case of (a).

c) when the parts are recombined at random the high pitch of the first syllable is the cue for the perception of Accent 1 (cf. (a)) which is further confirmed by the low falling pitch of the third syllable (see Fig. 14, curves 1 & 2). The falling pitch of the third syllable is a cue for Accent 2 (see Fig. 14, curves 3, 5 & 6), and the (almost) even and high pitch of the third syllable in addition to the lower pitch of the first syllable is a cue for Accent 3 (Fig. 14, curve 4). In this small experiment the perception of Accent 3 worked only with a combination of the hasi of Accent 2 and to of Accent 3. The result obtained here shows the same tendency as the above mentioned results of the whole experiment, but this should be confirmed further by tests including several words.

3. Final remarks.

3.1. In summary, taking all the three acoustic factors together, the accents tend to show the following acoustic differences:

1. The pitch factor plays the greatest role of the three factors.
2. Intensity and duration may also be of some importance, especially for Accent 1 words.

The result that the pitch movement in the third syllable differentiates between Accent 2 and Accent 3 words is similar to the result obtained by Sugitow (15) in her investigation of the pitch movement of two-syllable words where the first vowel is devoiced, though she investigated another dialect. As for the degree of pitch movement and the pitch of the third vowel as cues, the results here agree to a large extent with Han's statements.

As for the role of intensity for the accent perception, Neustupný (op.cit.) says that in some cases accent perception does not work without the help of the intensity factor. In my experiment, the case of hasito in the sentence might be explained as a compensation, if we do not take formant structure into account. In other cases, however, the pitch movement does not show any negative result for pitch. The pitch factor may be supported by the intensity factor, but not vice versa.

3.2. There are many unsolved matters, such as the evaluation of the area integral of intensity, the relation between acoustic and auditory level tone, etc. But these problems are of a more general phonetic character.

In this experiment formant structures were not analyzed. At the first look at spectrograms, it is hard to find the difference between three differently accented words. However, if minute spectrographic analyses are undertaken, other factors may be found which are relevant to accent perception. Meyer-Eppler (9) investigated whispered German tense vowels sung in "the first five tones of a diatonic scale", and found some extra noises for whispered /i,e,o/ and noises in addition to formant transitions for whispered /a,u/. I do not know whether such extra noises occur in Japanese devoiced /i/ and /u/, which make the intensity higher and bring about differences in pitch perception. It is said that there is a difference in quality between accented and unaccented vowels (Onishi (13), Koizumi (7), etc.). Koizumi says that a certain correlation between the accent and the tension of muscles has been found:

> "The high pitch accent shows itself in the tense form of lower vowel variants [ɑ,ɔ,ɛ] or higher ones [i,u]. The low pitch accent shows itself in the lax form of higher vowel variants [ə,o,e] or of lower ones [I,U]." (p.9)

If this phenomenon is constant there may be some difference in devoiced vowels too, both acoustically and auditorily. That is, there may be some difference in formant structure and in auditory impression.

References:

(1) Bernard Bloch, "Studies in Colloquial Japanese IV, Phonemics", Language 26 (195o), pp. 86-125; Readings in Linguistics (I), ed. Joos (1958), pp. 329-348.

(2) D.B. Fry, "Experiments in the Perception of Stress", Language and Speech I (1958), pp. 126-152.

(3) Mieko Shimizu Han, Japanese Phonology, An Analysis Based upon Sound Spectrograms (Tokyo, 1962).

(4) A.S. House & Fairbanks, "The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels", JASA 25 (1953), pp. 1o5-113.

(5) Shin Kawakami, "Japanese Accent, Observed by the Aid of Pitchgraph", The Study of Sounds 1o, pp. 115-129.

(6) Ryuji Kobayashi, "Nihongo (Tōkyō hōgen) akusento no nijūsei ni tsuite" ("On the Duplex Nature of Pitch Accent in Japanese (Tokyo Dialect)"), Nihon onkyō-gakkai kōen ronbun-shū (1967), pp. 281-282.

(7) Tamotsu Koizumi, "Nihongo no akusento to boin no onka" ("On the Japanese Accent and Vowel Quality"), Bull. Phon. Soc. Japan 126 (1967), pp. 9-11.

(8) Philip Lieberman, "Some Acoustic Correlates of Word Stress in American English", JASA 32 (196o), pp. 451-454.

(9) W. Meyer-Eppler, "Realization of Prosodic Features in Whispered Speech", JASA 29 (1957), pp. 164-166.

(1o) John Morton & Wiktor Jassem, "Acoustic Correlates of Stress", Language and Speech 8 (1965), pp. 159-181.

(11) Kazuo Nakano, "Shiin no yūseika to museika" ("Consonantal Vocalization and Devocalization"), Bull. Phon. Soc. Japan 125, pp. 7-8.

(12a) J.V. Neustupný, "Haná ("flower") to Hana ("nose") no hatsuon wa hatashite chigau ka" ("Is there any difference of pronunciation between /hana˥/ and /hana/ ?"), Gengo Seikatsu (Tokyo) 172, Part 1, pp. 8o-87.

(12b) J.V. Neustupný, "Nihongo no akusento wa kōtei akusento ka" ("Is the Japanese Accent a Pitch Accent?"), Bull. Phon. Soc. Japan 121, pp. 1-7.

(13) Masao Onishi, "Onshitsu to onchō-suii no sōkankankei" ("Relativity of Sound Quality and Pitch Transition in the Japanese Language"), The Study of Sounds 13, pp. 51-81.

(14) Hide Helen Shohara, "Some Characteristics of Pitch Variations in Japanese", Lingua I (1948), pp. 4o5-4o9.

(15) Miyoko Sugitow, "Musei boin to akusento" ("Reciprocal Action between Japanese Accent and Devocalization"), Bull. Phon. Soc. Japan 123, pp. 4-7.

(16) Michoko Temma, "An Experimental Study on the Phonetic Changes in Japanese Speech", The Tsuda Review, No. 6, November 1961, Tsuda College, Tokyo, pp. 23-4o.

NOTES ON THE DANISH VOWEL PATTERN

Jørgen Rischel

The vowel pattern of Danish exhibits several features which are interesting from the point of view of general phonological theory. And indeed, Danish is among the languages that have attracted the attention of phonologists already at an early time. The literature on Danish phonology will not be listed in its entirety, let alone discussed in detail, in this report (a critical survey has been given quite recently by Basbøll in an unpublished thesis (2)), but it goes without saying that the results of the various contributions to Danish phonology have been utilized as a basic source of knowledge about contrastive and distributional facts. The reader may be referred to the monograph by Martinet (14) and to the most recent papers on the vowels by Basbøll (in this report) and by Ege (5) for additional, more or less taxonomically oriented information. Aage Hansen (9) is a rich source of general information on the Modern Danish language.

The purpose of this report is to discuss quite informally some controversial aspects of the Danish vowel pattern. Definite solutions to the problems dealt with are offered only occasionally, mainly because the morphophonemics of Danish has not been analyzed in sufficient detail for safe conclusions to be made, neither by this author*) nor in the avaible literature on Danish. Nevertheless, it is the conviction of the present author that there is a widespread interest in data on Danish phonology even in this modest form. To my knowledge there are only two papers that deal essentially with the morphophonemic aspect, viz. Hjelmslev's paper from 1948 (10) and an unpublished paper by Hamp (8) held before the Linguistic Circle of Copenhagen in 1966.

----

*) The contents of this paper are largely based on my preparatory notes to courses in Danish phonetics (given by me at the University of Odense in 1967-68) and in generative phonology (Copenhagen, 1968).

Both of these are of course quite crucial to the present study, but as might be expected the summaries available are more oriented towards solutions (in Hjelmslev's case: reduction of the phoneme inventory, in Hamp's case: positing of feature matrices and rules) than towards a lucid presentation of data. On the whole one finds that in spite of the fact that considerable research has been done on certain aspects of the Danish language, general information on the gross features of the "standard language" is not very easily accessible, at least to nonnative scholars.

1. Preliminary survey of the phoneme system.

This report will contain three main sections: firstly, a tentative scetch of the phoneme system in "traditional" terms, secondly, a (likewise entirely tentative) analysis of the phonemes into distinctive features, and thirdly, some scattered reflexions on the vowel morphophonemics. - The introductory, essentially taxonomic, survey is kept very brief, since most of the relevant phenomena are dealt with at length in Basbøll's paper. (Notice, however, that he describes a local, or even individual, variety of Danish which differs on some points from the perhaps somewhat arbitrary "general" norm dealt with here.)

1.1. Tentative phoneme inventory.

Danish may be assumed to have ten long and ten short vowel phonemes:

i:	y:	u:	i	y	u	
e:	ø:	o:	e	ø	o	(plus shwa)
ɛ:	œ:	ɔ:	ɛ	œ	ɔ	
	a:			a		

since all of these entities are mutually commutable (or at least potentially commutable).

The phonemicists have been and still are in disagreement about several points including the phonemic status of length, the interpretation of [œ] , the interpretation of the various a-sounds found in Danish, and the interpretation of shwa, but the arrangement given here is an expedient starting-point for the subsequent discussion.

1.2. Main rules of (alleged) allophonic variation.

The alphabet of the International Phonetic Association is used in this paper, but it must be said expressly that Danish is badly suited for representation in the IPA alphabet, particularly because most of the vowel are situated nearly halfway between the cardinal vowels, so that narrow transcriptions require an extensive use of diacritic marks, and broad transcriptions involve a certain amount of arbitrariness in the choice among symbols (also compare Poul Andersen (1) and Hjelsmlev (1o)). It is necessary, therefore, to define the use of some of the IPA symbols in this paper:

e ø o are IPA e$^{\perp}$ ø$^{\perp}$ o$^{\perp}$; ɛ œ are roughly IPA e$^{\top}$ ø$^{\top}$

ɔ is almost mid and rather centralized, i.e. IPA ɔ$^{\perp+}$

ɒ is roughly IPA ɔ$^{\top}$or ɒ$^{\perp}$but generally rather centralized, and nearly unrounded in a widespread pronunciation (i.e. ʌ$^{\top+}$ or ɑ$^{\perp+}$). — A more retracted (pharyngealized) [ɒ] is here given as [ɒ̠].

ɑ is a back to central vowel of individually varying quality. A more retracted (pharyngealized) ɑ is given as [ɑ̠].

æ is rather close to IPA æ.

ɶ denotes a very open rounded front vowel.

i̯ u̯ ɒ̯ are nonsyllabic; the last one (from underlying /r/) is more or less pharyngealized.

Long vowels are symbolized as V:, and stress is indicated by an acute mark on the vowel V́.– Stød (accentuation characterized by a glottal constriction or sometimes closure) is indicated by ʔ after the vowel or consonant in which it is heard. When the stød "falls on" a vowel, the vowel is half-long (Vʔ).

The exact phonetic values of consonants are of little concern in the present study, and these are given in a very rough transcription (which disregards allophonic variation of voice and aspiration in the stops). It should be noted that the symbol "r" is used to denote a somewhat fricative uvular occurring syllable-initially (the"vocalic r" is represented by the symbol "ɒ"). A narrower transcription is used in Basbøll's paper.

1.2.1. Main allophones.

The main allophones of /i: y: u: e: ø: ɛ: œ:/ are respectively [i: y: u: e: ø: ɛ: œ:]. The corresponding short vowels are quite similar in quality, i.e. there is no easily perceptible difference in height or in "tenseness".

The main allophones of /a: o: ɔ:/ are respectively [æ: o: ɔ:], whereas short [æ o] have a highly limited distribution; [æ] occurs in some cases where it represents a shortened /a:/ (see 1.2.6. below), and [o] occurs in pretonic and posttonic syllables and in a few other cases (e.g. [fóto]'photo', [sort] 'black'), otherwise the two sounds are found only in special cases where they represent shortened /a:/,/o:/ (see later). Some people have short [o] as a commonly occurring sound, but it has been widely replaced by [ɔ] (quite commonly in the younger generation), so that it is reasonable to assume the latter pronunciation (e.g. [hɔl] rather than [hol] 'hole') in a description of modern Danish.

The three short vowels readily avaible to be taken as main allophones of /a o ɔ/ are respectively [a ɔ ɒ]. This classification postulates a common lowering rule for short /o/ and /ɔ/ (note that /o/ does not go with /e,ø/ in this respect), as well as a rule stating the difference between /a:/ and /a/.- It is clear, however, that one must seriously consider the possibility of considering this lowering as a purely diachronic phenomenon, a <u>phonemic shift</u>, which the implications that /o/ is lacking in most environments, /ɔ/ is [ɔ] in most environments, and there is an additional phoneme which is phonetically [ɒ]. In fact this may be the only tenable analysis from a strictly taxonomic point of view.

Long vowels are (regularly or optionally) shortened in certain types of environments. The grammatical conditions for this will not be dealt with in any detail in this paper, but it must be mentioned that the presence of shortened long vowels may pose awkward problems for a taxonomic analysis.- Shortened /ɔ:/ is [ɔ] in some cases but [ɒ] in others. If, for the sake of "symmetry",[ɔ] is considered the (main) allophone of /o/ it seems imperative to formulate a morphophonemic alternation rule /ɔ:/→/o/ in order to account for forms like [pɔ́ðn] 'on it' from [pɔʔ] 'on' (notice that this form is distinct from [póðn] 'the pot' in most people's speech), although the quoted rule is phonetically vacuous in respect of vowel quality change. See further 2.4. and 3.5. below.

1.2.2. r-conditioned allophones.

Several of the vowels have special allophones adjacent to /r/. The short close vowels are somewhat lowered before [ɒ̯], which is assumed here to be an allophone of /r/, and it is generally contended that i/e, y/ø, and u/o are neutralized and may vary more or less freely in this position (cp. Diderichsen (4)), although there seem to be some words that are invariably pronounced with a close vowel, and others that are invariably pronounced with a half-close vowel. In the same position /ɛ œ/ have very open allophones [æ ɶ] (of which the former closely resembles the main allophone of long /a:/). /a ɔ/, both long and short, fuse with vocalic r and are actualized as back vowels [ɑ̱(:) ɒ̱(:)], cp. [fɑ̱ʔt] 'speed', [tɒ̱ʔn] 'tower' (orthographically fart, tårn).

After /r/ the vowels /e ø ɛ œ a/ are lowered and/or retracted to a greater or lesser extent. In particular, short /ɛ œ a/ and long /a:/ are [æ ɶ ɑ] and [ɑ:]. With some speakers of Danish (particularly in the Copenhagen variety of the younger generation) these vowels (including long /e: ø: ɛ:/) are lowered so much that it is possible to make a restatement in the taxonomic analysis to the effect that these persons have /ɛ œ/ instead of /e ø/, and /a/ instead of /ɛ/. This analysis probably necessitates that two additional phonemes be posited, viz. /ɶ, ɑ/.

Since /a ɔ/ fuse with immediatly following /r/ to form one phonetic segment, viz. [ɑ ɒ] or [ɑ̱ ɒ̱] it can be, and indeed has been, argued that /ɑ ɒ/ must be set up as extra vowel phonemes (cp. Koefoed (12) and Basbøll elsewhere in this report), which further may imply that "vocalic r" is phonemically a vowel in all instances.

There are two features which I shall discuss separately here. Firstly one might argue that [ɑ ɒ] are single segments phonemically. This can hardly be maintained quite generally (cp. contrasting forms like [pɑk] 'scum' versus [pɑ·k] or [pɑ̠·k] (less commonly [pɑr̥k]) 'park', of which the latter can hardly be analyzed as a long vowel phoneme since long vowels in words of this structure invariably take the stød). From a generative point of view the analysis is definitely untenable, cp. section 3.3. below. - Secondly one might argue that "vocalic r" is a vowel also phonemically, i.e. that forms like [bæɒ̯] 'berry' contain a phonemic diphthong. This latter point can probably be defended with much more efficiency. In the following it is assumed that [ɒ̯] represents /r/, but most of the discussion will be almost unchanged even if the restatement mentioned is made.

1.2.3. Allophones of short /a/.

As stated above, short /a/ may be assumed to have at least two allophones: [ɑ] or [ɑ̠] in the sequences /ar/ and /ra/ (probably with many speakers more retracted before than after /r/), and [a] elsewhere. However, with some speakers of Danish this description does not suffice. It is common to have a rather retracted (central or even back) vowel before labials and velars, and in eastern varieties of Danish (cp. the Copenhagen dialect treated by Basbøll elsewhere in this report) this vowel is similar to the allophone found adjacent to /r/ (and before clusters containing /r/), cp. [kɑfə] 'coffee', [tɑk] 'thank'. (Other local varieties have [kafə] but [tɑk] or even [kafə], [tak].) - If a separate phoneme /ɑ/ is set up some of the a-sounds before labials or velars must of course be assumed to belong to this phoneme, too, with the perhaps disturbing result that slight differences in the pronunciation of words with a before labials or velars must be interpreted as phonemic differences.

1.2.4. Short vowels before nasals.

The close vowels [i,y] occur only sporadically before nasals (in words like pinje, hymne which can obviously be considered "[+foreign]"). Since the opposition between [ø] and [œ] is largely confined to this position several scholars (cp. Martinet (14), pp. 18-19, Spang-Hanssen (15), p.66) have considered the possibility of reducing the short vowel series [y ø œ] to two phonemes whose allophones are [y ø] in most environments but [ø œ] before nasals. By generalizing the lowering rule before nasals one might further postulate that [e] before nasals is phonemically /i/. However, with the back vowels this distribution does not hold, since there is a handful of common words with [u] before [n], e.g. [hunʔ] 'dog'.

1.2.5. Rounded front vowels before /r/.

The distinction of [y:] vs. [ø:] vs. [œ:] is found only before [r] or [ɒ̯] (i. e. in sequences containing the phoneme /r/). Example are ([sdy:rɒ] or normally) [sdy:ɒ] 'steer' versus ([kø:rɒ] or normally) [kø:ɒ] 'drive' versus ([gœ:rɒ] or normally) [gœ:ɒ] 'do'. An often cited series with stød is more problematic because the vowel is often pronounced short: [dyʔɒ̯] or [dyɒ̯ʔ] 'deer' versus [døʔɒ̯] or [døɒ̯ʔ] 'dies' versus [dœʔɒ̯] or [dœɒ̯ʔ] 'door'. Otherwise only two vowels are distinguished, viz. [y:] and [ø:] ([œ:] before /n/ and possibly after /r/). Some scholars, therefore, wish to reduce this series to two phonemes /y:/, /ø:/ just as the short series (Martinet (14), pp. 12-13, Hjelmslev (10)). The distributional relationships are quite complex (both with regard to stød and vowel-length: many speakers of Standard Danish do not have long vowels in all of the forms cited above), and the possibility of such restatements will not be considered further here.

1.2.6. Vacillating vowel-length.

The contrastiveness of vowel-length in surface forms is beyond dispute, at any rate in words of two or more syllables, cp. [vi:lə] 'rest' vs. [vilə] 'wild (plur.)'. In monosyllables we find instead an opposition between half-long vowels with stød and short vowels, cp. [duʔ] 'be of value' vs. [du] 'you', [lεʔs] 'read (imperative)' vs. [lεs] 'load'. (The phonemic status of the stød is not at issue here.)

In monosyllables with vowel followed by [ð ɣ i̯ u̯ ɒ̯] many people have little or no distinction of length. In numerous cases there is, however, a difference of stød (stød "on the consonant" versus no stød), which from a generative point of view signals an underlying difference of vowel quantity, cp. [uʔð] or [uðʔ] 'out' vs. [buð] 'message'.

In those cases where short vowel phonemes and long vowel phonemes have different allophones, the qualitative difference is generally preserved even if the vowels are all phonetically short, cp. [bɛʔɒ̭] or [bɛɒ̭ʔ] 'carry (imperative)' vs. [bæɒ̭] 'berry', but not in all cases by all people, cp. [bæʔð] or [bæðʔ] or [baðʔ] 'begged' vs. [bað] 'bath'. For an interesting attempt to give a taxonomic description of an idiolect with generalized short vowels in all of these cases see Basbøll's paper elsewhere in this report.

1.3. Vowels in unstressed syllables.

In pretonic and posttonic syllables we find a number of distinct vowels in words of foreign origin (sofa, brutto, pari etc.), but in posttonic syllables of the genuine vocabulary including productive suffixes of native origin only four vowel qualities occur, viz. [ə, ɒ, i, e] ([u] in [vendu] 'window' is quite exceptional). The first three of these are commutable, cp. [fadi] 'poor' vs. [fadə] '(to) comprehend' vs. [fadɒ] 'comprehends', whereas [e] occurs only before [ŋ] ([raineŋ] 'bill') where the other vowels do not occur.*) It is thus possible to consider this unstressed [e] as phonemically identical with [i] or with [ə].

[ə] can be considered a separate phoneme, a variant of one of the half-close or half-open front vowels, or a neutralization of several vowels. Unstressed [ɒ] can be interpreted in the same way as stressed [ɒ], cp. 1.2.1&2. above, or it can be taken to be a combination of vowel plus /r/, possibly of shwa +/r/ so that its phonemic identity depends on the interpretation of [ə].**)

----

*) It can be added, though, that family names with the suffix spelled -ung are pronounced with [ɔ] (also pronounced [o]), e.g. Hartung [ˈhɑːtoŋ].

**) As Basbøll has pointed out to me this latter identification involves morphophonemic considerations, i.e. there is hardly any conclusive "surface" evidence in favour of it. See section 3.3. below for a restatement in generative phonological terms.

1.4. Diphthongs.

There are several phonetic diphthongs in Danish. The final component of these diphthongs is either palatal, labiovelar, or velar/pharyngeal. Examples are [ai̯ ɑu̯ æɒ̯] , e.g. [mai̯]'me', [hɑu̯] 'sea', [bæɒ̯] 'berry'. It is hardly worth while giving an exhaustive list of these diphthongs here.

The three final components [i̯ u̯ ɒ̯] can be interpreted either as vowel phonemes (/i/, /u/, and /ɔ/ or /ɒ/) or as consonant phonemes (viz. /j/, /v/, and /r/). A hybrid solution with /i, u, r/ is of course feasible, too. It does not appear easy to argue entirely convincingly for any of these solutions on the basis of purely phonetic-distributional evidence (see Basbøll's report for a detailed account within that framework.)

2. Feature analysis.

In this section it will be assumed that the table of vowel phonemes presented in section 1.1. is adequate. Some different versions of the distinctive feature theory will be tried out with reference to this set of phonemes, viz. those found in Jakobson-Fant-Halle, Preliminaries (11), Ladefoged, Linguistic Phonetics (13), and Chomsky-Halle, The Sound Pattern of English (3). These three sources will be referred to as Prl., LP, and SPE.

2.1. Length as a feature.

Apparently the most natural way to account for differences like [lɛ:sə] 'read' versus [lɛsə] 'load' is to posit a distinctive feature of length. Both Prl. and SPE refer to long versus short as a prosodic opposition, apparently because it is based on the temporal relation between phonemes in a sequence, not on their absolute duration (in SPE no explanation is offered). LP suggests that oppositions of length belong to a parameter called "rate", which also serves to account for such distinction as flapped versus trilled sounds.

A very common way of distinguishing between long and short in modern phonological work (including SPE) is to posit a classificatory feature "tense", whose phonetic correlates may include differences of duration as well as differences of formation. This analysis does not make much sense phonetically for Danish since the opposition in this language is strictly durational and phonetically restricted to vowels. If there is a feature "tense" which is relevant in the consonant system, it is not matched by anything in the vowel system.

A third possibility is to assume underlying geminates: /ii/→[iː], etc. This solution may seem a costly one, since it requires that all feature values of a long vowel be specified twice. However, it is possible to take the second member of these geminates as an unspecified vowel, the unspecified features being specified by a redundancy rule of the form

$$\begin{bmatrix}+\text{vocalic}\\ -\text{consonantal}\end{bmatrix} \rightarrow \begin{bmatrix}\alpha F_1\\ \beta F_2\\ \gamma F_3\\ \vdots\end{bmatrix} \Bigg/ \begin{bmatrix}+\text{stress}\\ \alpha F_1\\ \beta F_2\\ \gamma F_3\\ \vdots\end{bmatrix} \underline{\qquad}$$

Where F_1, F_2 ... F_n are distinctive features.

Such a rule can be formulated in perfectly general terms provided that there are no bisyllabic sequences of stressed short vowel plus short vowel (in <u>zoologi</u>, and the like, we must assess [-stress] on the initial syllable before the rule applies, since here [o-o] can be heard), and no phonemic diphthongs (with different first and second member). Thus the formulation hinges on the interpretation of forms like [ai̯ən] 'own', [hɑu̯] 'sea', [bæɒ̯] 'berry', whose postvocalic segment must be defined as a consonant or glide in order not to invalidate the rule. Such a solution is possible and indeed correct (for the forms cited, but perhaps not for all cases of phonetic diphthongs), see later.

The interpretation of vowel-length as gemination is phonetically reasonable, since long vowels are roughly twice as long as short vowels in Danish, see Eli Fischer-Jørgensen (6).

2.2. Lip-rounding.

The distinction between /i e ɛ a/ and the other vowels is beyond dispute: the former are [-flat] or [-round] depending on the terminology, the latter [+flat] or [+round]. - <u>LP</u> has a ternary distinction of spread lips, neutral lips, and close rounding; it is not entirely clear to me how the Danish vowels should be placed along this parameter, since for example /o:/ and /ɔ/ differ very much in their degree of rounding.

2.3. Place of articulation.

Expect for /a/ and shwa, the Danish vowels fall in two fairly well-defined categories: front and back, i.e., in the terminology of Prl., "acute versus grave" or, according to SPE, [-back]versus [+back]. LP suggests that vowels may be placed in four or five distinct regions along a parameter called"articulatory place". For instance, i may be defined as postalveolar or palatal, u as velar, o as uvular, and ɔ as pharyngeal if there are phonological criteria supporting this subdivision. On the classificatory level there is probably no need for such differentiation in Danish, although /a/ and shwa pose certain difficulties which may eventually be solved by recognizing a ternary distinction of front-mid-back, see next section.

2.4. Tongue-height.

Phonetically speaking, Danish has a clear distinction of four degrees of tongue-height or aperture. The phonemic distinctness of four values is plainly demonstrated in the unrounded front series, cp. these examples from Fischer-Jørgensen (7):

[mi:lə] 'miles'	[lit] 'suffered'
[me:lə] 'sprinkle with flour'	[let] 'a little'
[mɛ:lə] 'speak'	[lɛt] 'light'
[mæ:lə] 'paint'	[lat] 'loaded'

Extensive examples of oppositions among Danish vowels may be found in Ege (5).

Prl. distinguishes "diffuse" and "compact" vowels, and it has been customary in more recent work to split this distinction into two in order to get a ternary opposition. It is, however, hard to see how four degrees of tongue-height can be distinguished in this way, since no phonetic meaning can be attached to a distinction between $\begin{bmatrix}\text{+diffuse}\\ \text{+compact}\end{bmatrix}$ and $\begin{bmatrix}\text{-diffuse}\\ \text{-compact}\end{bmatrix}$.

This restriction on the combinability of the two features is expressly stated in SPE, which speaks of "high" instead of "diffuse", and "low" instead of "compact". According to SPE it is a universal constraint (entering the marking conventions) that segments cannot be $\begin{bmatrix}\text{+high}\\ \text{+low}\end{bmatrix}$.

It is a possible objection against this whole treatment that the constraint on the combinability of "high" and "low" is not a characteristic of the speech organs of man, or at all a constraint on the way he uses language, but simply part of the general definition of the words "high" and "low" as forming a contradictory opposition. The authors of SPE, however, contend that "high" and "low" are different parameters and define a neutral tongue-position such that [+high] and [+low] both stand for deviations from this position, although in different directions.

Wang (17) has suggested a feature "mid" instead of "low", since "mid" can combine with "high" to form a four-way opposition:

$$/i/= \begin{bmatrix}\text{+high}\\ \text{-mid}\end{bmatrix} \quad /e/= \begin{bmatrix}\text{+high}\\ \text{+mid}\end{bmatrix} \quad /\varepsilon/= \begin{bmatrix}\text{-high}\\ \text{+mid}\end{bmatrix} \quad /a/= \begin{bmatrix}\text{-high}\\ \text{-mid}\end{bmatrix}$$

Wang has used this type of analysis to account for tone systems with four registers. However, this use is criticized by Ladefoged (LP, pp. 67-69) on the grounds that the difference between [+high] and [-high] is relatively the same as between [+mid] and [-mid], i.e. the features cannot be defined phonetically independently of each other and can only be used in a purely abstract sense. The same objection may be raised against the use in connexion with tongue-height unless it is possible somehow to define a "mid" tongue position in absolute terms.

LP suggests instead a multi-valued parameter of "auditory height". It is assumed that 3 values will suffice to characterize the vowel phonemes of any language, additional oppositions being taken care of by oppositions like "tense-lax". As said above this makes no sense for Danish, and thus the proposal of LP is no more adequate for Danish than the others. - LP refers elsewhere to a ternary parameter "articulatory stricture" (normally used to distinguish stops, fricatives, and approximants including vowels), which can be made four-valued so that it provides a distinction of "near vowels", i.e. vowels with a pronounced constriction somewhere in the vocal tract (including the pharyngeal region), and "far vowels" without any pronounced constriction. /i/ and possibly /a/ would be "near vowels", /e/ and /ɛ/ would be "far vowels", and so this parameter might solve the problem by taking into account the pharyngeal constriction of [a]. However, there seems to be nothing in the rules of the language to support a classification of /a/ together with /i/ as against /e ɛ/, and the phonetic realization of /a:/ as a very advanced and not entirely low vowel [a:] may contradict the classification of /a:/ as "near".

A remaining possibility is to rearrange the system in such a way that only three distinct degrees of tongue-height are assumed:

i	y		u
e	ø		o
ɛ	œ	a	ɔ

Since /a/ is actualized with phonetic values ranging from front over mid to back it would be entirely reasonable to define /a/ as central and to account for the variation by means of fronting and assimilation rules. This means that a ternary opposition "front-central-back" must be posited for Danish. However, the values "central" and "back" are not minimally distinctive in the resulting classificatory matrix, so it may be claimed that we should end up with a binary distinction "front-back" by classifying /a/ as [+back]. It will be shown later that there may indeed be some evidence in favour of this solution.

We can now set up a (fully specified) matrix for the vowels entirely according to the feature theory of <u>SPE</u>:

	i	e	ɛ	y	ø	œ	a	u	o	ɔ
high	+	−	−	+	−	−	−	+	−	−
low	−	−	+	−	−	+	+	−	−	+
back	−	−	−	−	−	−	+	+	+	+
round	−	−	−	+	+	+	−	+	+	+

This matrix is not very satisfactory, however.

Phonetically it seems rather strange to define /e, ø, o/ as $\begin{bmatrix}-\text{high}\\-\text{low}\end{bmatrix}$ i.e. neutral, since these vowels are clearly produced with the tongue raised to a considerable extent (the vowel /ɛ/ would conform better to the description of "neutral position" given in SPE, but /ɛ/ cannot possibly be [-low], since there is then no way to distinguish /e/ from both /i/ and /ɛ/).

Moreover, it is clear that if /a/ is [+back] this vowel must become [-back] in most environments by the operation of several rules. Clearly these cannot be early rules, since they would then make /ɛ/ and /a/ fall together, i.e. they must be preceded by another phonological rule making /a/ more open than /ɛ/, so that the resulting difference along the vertical dimension can take over the differentiating function as soon as the front-back difference is deleted. But this looks conspicuously like a trick made with the sole purpose of avoiding to posit more than three degrees of tongue-height on the systematic phonemic level.*)

If instead we posit a four-valued parameter of tongue-height with values from "1" (highest) to "4" (lowest), /ɛ/ and /a/ (long or short) are distinguished as "3" versus "4". Moreover, it is possible that the half-open and open back vowels can be given a more satisfactory treatment in this way. As stated above (section 1.2.1.) it is not altogether clear how the long vowels [oː ɔː] and the short vowels [ɔ ɒ] should be interpreted in relation to each other. If [oː] is defined as tongue-height "2", [ɔː ɔ] as "3", and [ɒ] as "4", the actual pronunciations of forms with rounded back vowels can be deduced by rather simple rules. We may well symbolize [ɔ] by /o/ (if this is convenient) but define it as tongue-height "3" in contradistinction to long /oː/, which is "2".

----

*) The rule making /a/ "extra-low" is in principle a redundancy rule

$$\begin{bmatrix} \text{+low} \\ \text{+back} \\ \text{-round} \end{bmatrix} \longrightarrow \text{"extra-low"}$$

but it cannot be stated as a redundancy rule in Stanley's (16) sense, i.e. operating to specify the input to the phonological component, without making "extra-low" a phonemic feature. This shows the fallacy of the approach suggested above.

Similarly [ɒ] may be symbolized by /ɔ/ but defined as "4" in contradistinction to /ɔ:/, which is "3". If /ɔ:/ is shortened it either stays "3" or becomes "4" according to rules that can be stated with reference to grammatical structure (see 3.6. below).*)

The matrix now looks like this:

	i	e	ɛ	a	y	ø	œ	u	o	ɔ	ɒ
height	1	2	3	4	1	2	3	1	2	3	4
back	-	-	-	-	-	-	-	+	+	+	+
round	-	-	-	-	+	+	+	+	+	+	+

although I do not insist on this arrangement.

3. Morphophonemic considerations.

The following remarks are for the most part confined to problems associated with the quality of open vowels in stressed and unstressed syllables, not because these problems are necessarily the most interesting ones but because they constitute a (limited) field which is rather inadequately accounted for in current textbooks and dictionaries.

3.1. The status of /a/.

According to the analysis outlined above /a(:)/ is the only unrounded vowel with tongue-height "4". We are thus free to define it as [-back] or [+back] according to which choice serves us best. It seems immediately obvious that it should be defined as [-back] in the classificatory matrix, but the problem is not quite simple after all. The kind of evidence that is particularly relevant here can be indicated rather briefly:

There is in Danish a dorsal fricative or rather frictionless continuant which appears after long vowels as more or less palatal or velar depending on the quality of the vowel, cp. [e:ɣən] or [e:i̯ən] 'own' (slightly old-fashioned pronunciation, except in certain compounds), [rø:ɣə] or [rø:i̯ə] 'smoke (meat or fish)', [dæʔɣ] or [dæʔi̯/dæʔ] 'day', [tɔʔɣ] or [tɔʔu̯] 'train'. After short vowels it often appears as the second part of a diphthong: [ai̯ən] 'own' (more common pronunciation), [rɒi̯ʔ] 'smoke', [daɣli] or [dau̯li] 'daily'. - The diphthongs given as "ai̯, ɒi̯"

----

*) This symbolization is used in the morphophonemic discussion later in this paper but not in the matrices below.

vary much in quality: in general Standard Danish the first component is central or rather front in the former and centralized back in the latter, but the latter may also be heard with a more front first component: [œi̯]. In the Copenhagen dialect the first component of both diphthongs may be quite back: [ɑi̯ ɒi̯].

Note that the postvocalic consonant becomes [i̯] after those short vowels which alternate with long non-open front vowels (i.e. with front vowels whose tongue-height coefficient is less than "4"), but not after those short vowels which alternate with long open or long back vowels. This holds true no matter whether the short vowel is front or back (cp. Copenhagen speech), and it is thus quite clear that the quality of the second component is not at all determined by the surface quality of the first component, cp. (Copenhagen Danish) [ɑi̯(ə)n] versus [dɑu̯li] . This suggests that it must be an underlying difference of [-back] versus [+back] in the vowel (first component) that determines whether the following dorsal approximant appears as [i̯] or [u̯] when the vowel is short, i.e. [ɑi̯, ɒi̯] must have underlying front vowels, and [ɑu̯] must have an underlying back vowel. The quality of the first component of [ɑi̯,ɒi̯] must then be determined by two successive rules: (1) a rule changing the tongue-height to "4" (i.e. [a œ]), (2) a rule according to which each vowel is retracted more or less depending on the idiolect or style of speech.

According to this analysis /a/ is [+back] in the underlying matrix and should thus be written /ɑ/. This implies that the long vowel must be fronted and somewhat raised

ɑ: → æ:

by a later rule, which nevertheless is early enough to make fricative [ɣ] front after this vowel if it remains long.

As a result of these considerations the classificatory matrix may be given in the following alternative form:

	i	e	ɛ	y	ø	œ	ɑ	u	o	ɔ	ɒ
height	1	2	3	1	2	3	4	1	2	3	4
back	-	-	-	-	-	-	+	+	+	+	+
round	-	-	-	+	+	+	-	+	+	+	+

The status of [ə] is a difficult problem even within this arrangement of the matrix, since there is still no way to characterize a "neutral" vowel.

3.2. Phonetic diphthongs and underlying forms.

It will be apparent from some of the examples given in section 3.1. that the diphthongs in -i̯ and -u̯ can be shown to alternate in some forms with sequences of long vowel plus palatal or velar approximant. Moreover, the dorsal approximant alternates in some forms with a stop (when followed by a stop), cp. [sbø:ɣəlsə] or [sbɒi̯əlsə] 'ghost' versus [sbøgt] 'haunted (past participle)'. In other cases the diphthongs in -u̯ alternate with sequences of long vowel plus [v], cp. [hau̯] 'sea' - plural [ha:və], which moreover may alternate with [f] (when followed by a stop), cp. [sdi?v] or [sdiu̯?] 'stiff' (with the derivation [sdiu̯nə] 'stiffen') versus neuter [sdift].

Such examples indicate that some of the diphthongs at least must be generated from underlying vowel-consonant sequences, but it is certainly not obvious that this is true of all diphthongs, or even of all occurrences of a specific diphthong. This cannot be discussed without a detailed analysis of the behaviour of vowel length (and of stød) and of the morphophonemic relationship between continuants and stops in Danish.

The diphthongs in -ɒ̯ are mostly taken to be phonemically sequences of vowel plus /r/ (see, however, Koefoed (12) and Basbøll in this report for alternative analyses). The evidence supporting the phonemic identification of prevocalic [r] and postvocalic [ɒ̯] is hardly decisive as far as the inflected and derived forms belonging to the genuine vocabulary go, but in foreign words with alternating stress pattern the syllable division may alternate accordingly, and in such cases [ɒ̯] and [r] do alternate, cp. [kloɒ̯?] 'chlorine' - [kloríð?] 'chloride'. In some (now less common) varieties of Danish an r-sound (with a constriction exceeding those generally found with vowels) is heard also in syllable-final position ([klo?r] etc., often with unvoiced r); for idiolects of this type the identification raises no problem at all.

Although the evidence is not of the same kind as that adduced for diphthongs in -i̯, -u̯, it may suffice to show that diphthongs in -ɒ̯, too, may be generated from vowel-consonant sequences, the consonant in this case being /r/.

3.3. The two-segment status of ɑ(:), ɒ(:).

3.3.1. Root syllables.

In the discussion in section 3.1. above it was tacitly assumed that there is no phonemic opposition between front /a(:)/ and back /ɑ(:)/. However, surface forms like [gæ:və] 'gift' and (gɑ:və] or) [gɑ:və] 'tan (verb)' must somehow be distinguished in their underlying representations. The general solution is to analyze the back vowel as a sequence of two phonemic segments; this solution is confirmed by alternations like

[ɑ̠ʔm] 'arm' (- plur. [ɑ̠:mə]) versus [æɒ̯mə] 'sleeve'
[bɑ̠ʔn] 'child' - plur. [bœɒ̯ʔn]

Such examples further show that the second component of the sequence underlying [ɑ̠:] is phonemically identical with [ɒ̯], i.e. it represents the consonant /r/.

It was suggested earlier that [æ:] may perhaps be generated from an underlying back vowel "/ɑ:/". If this is to work, the rules must operate in such way that (disregarding the problem of underlying or secondary length)

1. ɑ(:) → æ: (in some contexts)
2. ɑ(:)r → ɑ̠:

which would probably be fairly close to the historical development.

The long vowel [ɒ:] can similarly be shown to represent an underlying sequence of two segments, cp. the parallel between

[gɔʔ] 'go' - present tense [gɒ̠ʔ]
[seʔ] 'see' - present tense [seɒ̯ʔ]

that is, [ɒ̠:] must be derived from underlying vowel plus /r/; thus the rules say (disregarding the status of length in the underlying matrix):

1. ɔ(:) → ɔ:
2. ɔ(:)r → ɒ̠:

Unstressed [ɒ] in root syllables has a multiple origin, which may be illustrated by a few examples:

(1) Some monosyllabic interjections and particles ([nɒ] 'well', [sɒ] 'now then; then; so') must be assumed to have an underlying single segment

(ɔ →) ɒ stressed and unstressed

The actualization being invariant in the presence or absence of stress.

(2) The conjunction og 'and' has a diphthong [ɒu̯] when pronounced distinctly*) but normally loses the second component. This reduction

(ʔɔɣ) → ɒu̯ → ɒ

may be compared to a similar, fairly common reduction in pronominal forms with ai̯, ɑi̯→ɑ, cp. [dai̯/dɑi̯] 'you' (accusative) but often [mέdɑ] 'with you' in the Copenhagen variety of Danish.

(3) The preposition /adverb for has a retracted (pharyngalized) [ɒ̠] when stressed, but the vowel is reduced to [ɒ] when unstressed, cp.[sbɛnə fɒ̠́] 'harness' but [sbɛnə fɒ vɒ́u̯ʔnən] 'hitch (a horse) to the carriage'. - Note that in its stressed form the vowel is distinct from that of [nɒ́], [sɒ́] by being more back; this indicates that for has underlying /ɔr/ and reduction

ɔr → ɒ when unstressed

(4) Present tense forms like [gɒ̠ʔ] 'goes' - with underlying ɔ: + r as shown above - may be more or less reduced when unstressed, cp. [han gɒ̠ jέmʔ]~[han gɒ jέmʔ] 'he goes home', i.e.

ɔ:r → ɒ̠ → (opt.) ɒ when unstressed

(5) The adverbs [voʔɒ̯/voɒ̯ʔ] 'where', 'how', [heʔɒ̯/heɒ̯ʔ] 'here', [deʔɒ̯/deɒ̯ʔ] 'there' (the last two also occur with [ɛʔɒ̯/ɛɒ̯ʔ]) are reduced in unstressed position to [vɒ̠ hæɒ̯ dæɒ̯] and often further to [vɒ hɒ dɒ], cp. der kommer nogen,

----

*) The pronunciation [ɒɣ], which may be found in the official dictionary: Ordbog over det danske Sprog, is hardly current usage (cp. Diderichsen (4), p.55).

literally 'there comes somebody'[dɒ kɔ́m'ɒ nɔ́:un]. This reduction

e(:)r → æɒ̯ } (opt.)→ɒ
o(:)r → ɒ̯ }

is related to processes found in words with alternating stress placement (see 3.3.2. below), but as a process conditioned by syntactically conditioned stress reduction it seems confined to the forms just cited. Such forms as [seʔɒ/seɒ̯ʔ] 'see (present tense)', [troʔɒ/troɒ̯ʔ] 'think (present tense)' do not have reduction to [ɒ] no matter how much the stress is reduced, cp. [han seɒ̯ ɛ́fdɒ] 'he checks (it)', [troɒ̯ du dé] 'do you think so?'. It is possible that the three adverbs should be listed in the dictionary with [-stress], the full forms occurring only when the words are emphasized.

3.3.2. Sources of posttonic ɒ.

There are numerous instances of [ɒ] in unstressed syllables after the stress-syllable of the word. These cases in which [ɒ] is generated from an underlying matrix with [-stress] must be distinguished from those treated above where the weak stress was due to a syntactically conditioned reduction of the word stress.

A comparison of monosyllabic and bisyllabic verb forms:

[seʔ]	'see'	- present tense		[seʔɒ/seɒ̯ʔ]
[tɛlə]	'count'	- "	"	[tɛlɒ]

shows that [ɒ] may be taken as equivalent to [ə] + [ɒ̯]. Since the latter segment was shown above to be generated from /r/, we may generate unstressed [ɒ] from shwa plus /r/:

ər → əɒ̯ → ɒ

(I disregard the various problems associated with present tense forms: whether the morpheme border is before or after shwa, whether shwa is epenthetic, etc., since these are not immediately relevant here.)

The behaviour of the nomen agentis suffix [ɒ] before the feminine suffix [énə] may further support the identification of unstressed [ɒ] as a vowel-consonant sequence, cp.

[mæ:lə]	'paint'
[mæ:lɒ]	'painter'
[mæ:lɒénə]/[mæ:lɒrénə]	'woman painter'

In the speech of some people the underlying /r/ is distributed over two syllables (as if the spelling were malerrinde instead of malerinde). The final suffix is unquestionably [énə], cp.

[vɛ́n]	'friend'
[vɛnénə]	'girl-friend'

so the r-sound in derivations from nomen agentis forms can only be explained if these contain a final /r/.

However, it can be shown that unstressed [ɒ] may also represent underlying consonant-vowel or even consonant-vowel-consonant sequences.

Verbs whose stems end in a consonant form their infinitive by adding shwa: [tɛlə], [mæ:lə], etc. A comparison of such infinitive forms as

[tɛlə]	'count'
[ku:ə]	'cow'
[ɛndrɒ]	'change'
[sbæɒ̯ɒ]	'block'
[ku:(r)ɒ] or [ku:ɒ]	'slide'

shows that some occurrences of unstressed [ɒ] must be generated from final shwa preceded by /r/, since the last three verbs obviously contain stem-final /r/ (in accordance with the orthography: ændre, spærre, kure). - In order to get the correct output we must set up three rules, two of which (the last two) may be optional only:

1. ə → ɒ/r__ ku:rə → ku:rɒ
2. r → ɒ/V__ (sbærə →)sbærɒ → sbæɒ̯ɒ
3. rɒ → ɒ/V__ ku:rɒ → ku:ɒ

If now we consider the present tense forms of the verbs, we find in ordinary conversational pronunciation

[tɛlɒ]
[ku:ɒ]
[gnisdrɒ]
[sbæp̥ɒ]
([ku:rɒ] or) [ku:ɒ]

It is quite common to pronounce the present tense of kue and the infinitive and present tense of kure alike: [ku:ɒ]. If now we consider the verbs to have the same structure in their infinitive and present tense forms, we get

tɛlə → tɛlə	tɛlər → tɛlɒ
ku:ə → ku:ə	ku:ər → ku:ɒ
ku:rə → ku:ɒ	ku:rər → ku:ɒ

i.e.

$$\left\{\begin{matrix} \text{rə} \\ \text{ər} \\ \text{rər} \end{matrix}\right\} \rightarrow \text{ɒ}$$

in a certain (colloquial) style of speech. This fusion and merger takes place in other forms as well, cp. the nouns

bu?r → bu?ɒ̯	'cage'	- plur.	bu:rə → bu:ɒ
bu:ə → bu:ə	'bow'	- "	bu:ər → bu:ɒ
fu:rə → fu:ɒ	'furrow'	- "	fu:rə → fu:ɒ

(the output forms given here are still strictly colloquial).

In the cases above [ɒ] was generated from underlying forms containing shwa but the status of this vowel was not considered further. In words with alternating stress placement it can be seen that posttonic [ɒ] alternates with stressed [e?r e:r] and [o?r o:r] (partly also with pretonic [or]). This is probably best accounted for if one posits underlying /er/ and /or/, cp.

[έ?dɒ] (possibly [έ?tɒ]) 'ether' and [ɛté?risg]
or [ɛtép̥?isg] 'ethereal'

[fágtɒ] 'factor' — [fagtó:rɒ] or [fagtó:ɒ] 'factors'
and [fagtoré?ɒ] 'factorize'

It is not immediately evident whether posttonic / er or / become [ə] via a reduced common form or by lowering and retraction rules, i.e. whether posttonic $\left.\begin{matrix}\text{er}\\ \text{or}\end{matrix}\right\}\rightarrow$ ər → ɒ

or

posttonic $\left.\begin{matrix}\text{er} \rightarrow \text{ær} \rightarrow \text{æɒ}\\ \text{or} \rightarrow \text{ɒr} \rightarrow \hat{\text{ɒ}}\end{matrix}\right\}\rightarrow$ ɒ

Poul Andersen (1b, p.82) has pointed out that words with underlying full vowel in a posttonic syllable may be distinct from similar words with underlying shwa if the vowel is preceded by a stop, cp.

fakt+ər ⟶ fɑgdɒ 'gestures'
faktor ⟶ fɑgtɒ 'factor'

since the preceding stop is aspirated or not according to a gradation rule which must be formulated with reference to several factors including the quality of the immediately following vowel (Poul Andersen formulates the relationships rather differently: he considers words like faktor as phonological compounds: 'fak,tår).

If we have the same underlying consonants in the two words above (and it is not unreasonable to assume that we do), it seems clear that shwa must be phonologically distinct from the full vowel /o/ in the underlying matrix (the example might suggest that the consonant quality is a matter of morpheme border, but it is not that simple), unless words like faktor are marked as [+foreign] in the dictionary. In words with underlying /er/ the aspiration of the stop before [ɒ] does not seem to be used very consistently: æter is [ɛʔdɒ] rather than [ɛʔtɒ] in ordinary usage, so the evidence against a common underlying matrix for /e/ and shwa is not as strong.

Words with underlying /or/ may be assimilated to the native vocabulary and pronounced with unaspirated stop before [ɒ], e.g. [dɒgdɒ] instead of [dɒgtɒ] 'doctor'. In a case like this the surface form has been reinterpreted and assumed to contain underlying shwa, which is evident from the formation of a secondary plural of native type:

	[dɒ́gdɒ] -	plural	[dɒ́gdɒɒ]
like	[mǽːlɒ] -	"	[mǽːlɒɒ]

as against the "correct" inflection

[dɒ́gtɒ] - plural [dɒgtó:rɒ]

3.4. Alternation full vowel ~ shwa.

/e/ and to some extent also /ɛ/ frequently reduce to shwa when the stress is reduced. The problems raised by this alternation are less complex than those associated with unstressed [ɒ], and a few typical examples may suffice to give an idea of the pattern:

(1) "Small words" like [de] 'it', [dɛm] 'them' have the vowel reduced to shwa in enclitic position: [də, dəm], cp.

[ædə sán?t] 'is it true?'
[fɔdəm én?] 'get them in!'

(2) In foreign words [e ɛ] alternate with shwa in unstressed syllables, most regularly in noninitial pretonic syllables. The underlying form is clearly seen in words with alternating stress placement, e.g.

[foné?tigɒ] 'phonetician' versus [fonetík]/[fonətík]

It is clear that this phenomenon is related to the alternation between [e(:)r] and [ɒ] discussed above, although the distribution is not the same. The generality of these alternations between [e(:)] and the reduced vowel qualities point to the identification of shwa with /e/ in terms of classificatory features as a reasonable solution.

In colloquial (particularly "non-educated") speech also other vowels vacillate or are replaced by shwa in non-initial pretonic syllables (in connection with a general assimilation of the word to the native pattern). This is typically found with /a/ followed by /r/:

apɑrá?t → abɑrá?t → abərɑ?t (or abɒrá?t)

The pronunciation with [ə] is given by Martinet (14, p. 22) as the normal one but by Hansen (9, p. 39) as substandard speech. - I presume that all the pronunciations in the series above can be found in contemporary Danish usage.

Conversely, we find (sporadically) a shift

e, ə → ɑ

when adjacent to /r/, cp.

pɑrəntéʔs → pɑrɑntéʔs → pɑ̰ɑ̰ntéʔs

3.5. Some further remarks on length.

Although length is unquestionably phonemic in Danish the relationship between underlying quantity and phonetic length is not a simple one-to-one correspondence. A glance at paradigms of inflected forms reveals that both vowel lengthening and vowel shortening rules seem to operate in Danish:

lengthening takes place typically before a voiced approximant that is followed by a vowel, cp. [hɑu̯] 'sea' - plural [hæ:və], [glað] 'glad' - plural [glæ:ðə],

shortening occurs both as a process conditioned by the internal structure of the word and as a process conditioned by the syntactical relationships among words. There are two main types of intra-word conditioned shortening:

(a) before certain suffixes with initial consonant, e.g. neuter or past participle /t/, long vowels are shortened in some cases, cp. [sdiʔv](or more commonly [sdiʔu̯, sdiu̯ʔ]) 'stiff' - neuter [sdift]. The rule capturing this alternation between long and short vowels must be stated differently for verbs and adjectives (cp. [ly:sə] 'give out light, publish the banns for somebody' - past part. [lyʔsd] versus [lyʔs] 'light' - neuter [lysd]), and also for derivations from such words (cp. [ly:sneŋ] 'banns' versus [lysneŋ] 'dawn, clearing'). Moreover, the application of the shortening rule is to some extent dependent upon the quality of the vowel.

(b) As mentioned in 1.2.6. above long and short vowels are not consistently distinguished in monosyllables containing a postvocalic voiced approximant. We may speak here of optional vowel shortening ([sdiʔv, sdiʔu̯] ~ [sdiu̯ʔ] being an example of this). This phenomenon can be taken care of by a late rule. (According to Aage Hansen (9), p. 87 the individual vowels behave somewhat differently before [ð] ; however, the descriptions he gives do not altogether correspond to the usage familiar to me.)

Finally, vowel shortening may occur in the first part of compounds and conditioned by stress reduction,cp. 3.6. below).

A paradigm like [bað] 'bath' - plural [bæ:ðə] can be described as a case of vowel lengthening in "open syllable" under the conditions summarized above. The verb [bæ:ðə] 'bathe' with its preterite and past participle forms [bæ:ð(ə)ðə], [bæ:ð(ə)t] can apparently be explained in the same way, as derived from underlying forms with short vowel. However, the imperative of this verb is [bæʔð], which rather points to underlying <u>long</u> vowel in the verb. Thus it may seem that vowel-length is generated by a simple rule in plural forms like [bæ:ðə] but is due to a stem formation feature of length in infinitive forms like [bæ:ðə].

There is, however, some evidence that the behaviour of the imperative is due to special formation features and thus should not be taken as decisive in assessing the underlying quantity. It is necessary here to point to the fact that stød, too, functions (on the surface) to distinguish imperative forms from otherwise phonetically similar noun forms, cp. the noun [sbel] 'play' (definite form [sbelˀəð]) versus the imperative [sbelˀ] 'play!' (infinitive [sbelə]).

On the whole the quantity problems are too complex to be handled in this brief paper.

<u>3.6. The quality of shortened ɔ:</u>

As mentioned in 2.4. above /ɔ:/ may appear with two different qualities when it is shortened. It will be apparent from the remarks below that the distribution of these can be put into rule form without too much difficulty.

(1) In combinations of vowel plus voiced fricative or "approximant" the vowel may be alternatively short (see 1.2.6.), but the quality of /ɔ:/ (half-open, i.e."3") is retained: [vɔʔð] or [vɔðˀ] 'wet'.

(2) When shortened before suffix /t/, however, the vowel is opened to "4":
[vɔʔð] - neuter [vɒt].

(3) When the shortening is otherwise conditioned, the distribution of the two qualities, "3" and "4", is at first sight confusing. There is certainly much individual variation, and Poul Andersen (1, p. 338) describes the variation as if it were entirely a matter of usage: most people have [ɔ], but some have [ɒ], when a long ɔ is shortened because the word containing it occurs "in a special position in the sentence or in compounds". It is, however, my impression that there is a regular alternation between the two short reflexes in the usage of probably most speakers of Standard Danish, and this regularity is interesting because it is not - as might be expected a matter of "stressed" versus "unstressed" position but rather a matter of grammatical type.

Put briefly, it holds true for the usage most familiar to me that (A) monosyllables (essentially prepositions or auxilary verbs) which have their vowel shortened <u>before the word they govern</u>, retain quality "3", whereas (B) such words occurring as the <u>first part of a compound</u> exhibit a shift "3" → "4" if the vowel is shortened. The latter is true particularly of open syllables.*)

Thus (A):

[ta d(ə)n pɔ́ʔ] 'put it on'

[pɔ́d(ə)n] 'on it' (less commonly: [pɒ́d(ə)n] i.e. like <u>potten</u> 'the pot')

[pɔ dέnˀ] 'on that one' (- " - : [pɒ dέnˀ])

but (B):

[pɒ́menˀə] 'remind'

[pɒlíʔðəli] 'reliable'

Aage Hansen (9, p. 25) claims that <u>blågrå</u> 'blue - grey', <u>blålig</u> 'blueish' should be pronounced with the closer sound, i.e. [ɔ] in the first syllable. Such pronunciations are quite alien to me.

(4) Note that [ɒ] occurs as an unstressed representation of underlying /ɔr/, which thus (in the usage described here) remains distinct from shortened /ɔ:/, cp.

----

*) I disregard here a number of compounds like <u>bådsmand</u>, <u>rådsmedlem</u>, which do not follow a fixed pattern.

fɔr de færsdə → fɒdə fæɒ̯sdə 'firstly'

fɔ: de gjorˀt → fɔdə gjoɒ̯ˀt 'have it done!'

4. Final remarks.

As stated in the beginning this paper does not contend to give a definitive account of any major aspect of Danish phonology. The main problems, viz. the stød, vowel-length, umlaut and ablaut, and the relationship between stops and fricatives, have on purpose been left almost fully aside. Nevertheless, it seemed to me worth while to discuss some features of the vowel system within the framework of modern feature and rule theory, especially because this account can be contrasted directly with that of Basbøll, which, as an experiment (less so a manifestation of Mr. Basbøll's personal preferences!), he has kept as rigidly within the limits of taxonomic phonemics as practically possible. It may be of some interest that the results obtained by the two approaches agree on many details, after all.

----------*****----------

References:

(1) Poul Andersen, Dansk fonetik (Copenhagen, 1954).

(1a) Poul Andersen, "Om det finale r i dansk", Danske Studier 1965, pp. 75-84.

(2) Hans Basbøll, Unpublished thesis-work (Copenhagen, 1968).

(3) Noam Chomsky & Morris Halle, The Sound-pattern of English (New York, Evanston and London, 1968).

(4) Paul Diderichsen, "Udtalen af dansk rigssprog", Danske studier 1957, pp. 41-79.

(5) Niels Ege, "The Danish Vowel System", Gengyo Kenkyū 47 (1965), pp. 21-35.

(6) Eli Fischer-Jørgensen, "Om vokallængde i dansk rigsmål", Nordisk tidsskrift for tale og stemme 15 (1955), pp.33-56.

(7) Eli Fischer-Jørgensen, Almen fonetik (Copenhagen, 1962).

(8) Eric Hamp, "Ever again on Danish phonology", unpublished paper read before the Linguistic Circle of Copenhagen, 1966.

(9) Aage Hansen, Udtalen i moderne dansk (Copenhagen, 1956).

(10) Louis Hjelmslev, "Grundtræk af det danske udtrykssystem med særligt henblik på stødet", published in Selskab for nordisk filologi, årsberetning for 1948-49-50, pp. 12-24.

(11) R. Jakobson, C.G.M. Fant, and M. Halle, Preliminaries to Speech Analysis (Fourth Printing, 1961).

(12) H.A. Koefoed, "Er det finale r en vokal i rigsdansk?", Danske studier 1964, pp. 85-112.

(13) Peter Ladefoged, Linguistic Phonetics (= Working Papers in Phonetics 6, UCLA 1967).

(14) André Martinet, La phonologie du mot en danois = BSLP 38, 2 (Paris, 1937).

(15) H. Spang-Hanssen, "On the Simplicity of Descriptions", Recherches Structurales = TCLC V (1949), pp. 61-70.

(16) R. Stanley, "Redundancy Rules in Phonology", Language 43 (1967), pp. 393-436. Cp. POLA Reports 6: April 1968 (Univ. of California, Berkeley), pp. S1-S31.

(17) William S.-Y. Wang. "Vowel Features, Paired Variables, and the English Vowel Shift", POLA Reports 6: April 1968 (Univ. of California, Berkeley), pp. 1-40.

PUBLICATIONS BY STAFF MEMBERS

Eli Fischer-Jørgensen, "Perceptual Dimensions of Vowels", To Honor Roman Jakobson (1967), pp. 667-671 (also in Zeitschrift für Phonetik 21 (1968), pp. 94-98).

Jørgen Rischel, "Diphthongization in Faroese", Acta Linguistica Hafniensia XI (1967/68), pp. 89-118.