ARIPUC VOL. 7.

ts

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

COPENHAGEN 1973

ARIPUC VOL. 7.

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

COPENHAGEN 1973

CONTENTS

(continued)

PERSONNEL OF THE INSTITUTE OF PHONETICS
1972

Permanent Staff:

Eli Fischer-Jørgensen (professor, director of the Institute)

Lecturers:

Jørgen Rischel (general phonetics)
Hans-Peter Jørgensen (general phonetics)
Oluf M. Thorsen (general and French phonetics)
Børge Frøkjær-Jensen (experimental phonetics)
Hans Basbøll (general phonetics)
Karen Landschultz (general and French phonetics)
Nina Thorsen (general phonetics)
Hideo Mase (general phonetics), temporarily appointed
Peter Holtse (general phonetics), temporarily appointed
Birgit Hutters (general phonetics), temporarily appointed

Technical Staff:

Carl Ludvigsen (M.Sc.)
Poul Thorvaldsen (B.Sc.) until December 1, 1972
Svend-Erik Lystlund (technician)
Inger Østergaard (secretary)
Aase Thiim (secretary)

Part Time Teachers of General Phonetics:

Preben Andersen
Kirsten Gregersen
Peter Molbæk Hansen
Niels Reinholt Petersen
Pia Riber Petersen

Guest Research Workers:

Marie-Hélène Galvagny (Paris): investigations of German vowels.
Professor Stephen Anderson (Harvard): phonological investigations.
Professor Antti Sovijärvi (Helsinki): segmentation of speech sounds.

Furthermore several phoneticians from European countries, Australia, Canada, and the United States of America have visited the laboratory.

ABBREVIATIONS EMPLOYED IN REFERENCES

ARIPUC Annual Report of the Institute of Phonetics, University of Copenhagen

FRJ For Roman Jakobson

F&S Form and Substance

Haskins SR Status Report on Speech Research, Haskins Laboratories

JASA Journal of the Acoustical Society of America

JL Journal of Linguistics

JSHR Journal of Speech and Hearing Research

Ling. Linguistics

LS Language and Speech

MIT QPR M.I.T. Quarterly Progress Report

NTTS Nordisk Tidsskrift for Tale og Stemme

Proc.Acoust. Proceedings of the ... International Congress on Acoustics

Proc.Phon. Proceedings of the ... International Congress of Phonetic Sciences

STL-QPSR Speech Transmission Laboratories, Quarterly Progress and Status Report

SL Studia Linguistica

UCLA Working Papers in Phonetics, University of California

Zs.f.Ph. Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung

PUBLICATIONS BY STAFF MEMBERS 1972:

Hans Basbøll:	"Some remarks concerning the stød in a generative grammar of Danish", Derivational Processes (Stockholm) (F. Kiefer, ed.), p. 5-30.
Hans Basbøll:	"Remarks on the regular plural formation of English nouns", Language and Literature 1.3 (Copenhagen), p. 39-42.
Eli Fischer-Jørgensen:	"PTK et BDG francais en position intervocalique accentuée", Papers in Linguistics and Phonetics to the Memory of Pierre Delattre, p. 143-200.
Eli Fischer-Jørgensen:	"Formant frequencies of long and short Danish vowels", Studies for Einar Haugen, p. 189-213.
Kirsten Gregersen:	"Junktur på engelsk med en instrumentalfonetisk undersøgelse af kombinationen V+CV contra VC+V i RP", Extracta 4, p. 95-100.
Kirsten Gregersen	review of: David Crystal: "Linguistics", Language and Literature 1.3, p. 67-68.

Hideo Mase and Jørgen Rischel	"A study of consonant quantity in West Greenlandic", Int. Journal of Dravidian Linguistics 1, II, p. 138-195.
Jørgen Rischel	"Consonant reduction in Faroese non-compound wordforms", Studies for Einar Haugen, p. 482-497.
Jørgen Rischel	"Derivation as a syntactic process in Greenlandic", Derivational Processes (Stockholm), p. 60-73.
Jørgen Rischel	"A comment on lexical insertion", Int. Journal of Dravidian Linguistics 1, II, p. 84-99.

LECTURES AND COURSES IN 1972

1. Elementary phonetics courses

One-semester courses (two hours a week) in elementary phonetics (intended for all students of foreign languages except French) were given by Preben Andersen, Kirsten Gregersen, Peter Molbæk Hansen, Peter Holtse, Birgit Hutters, Hans-Peter Jørgensen, Niels Reinholt Petersen, Pia Riber Petersen, and Nina Thorsen.

There were two parallel classes in the spring semester and 17 in the autumn semester.

Courses in general and French phonetics for students of French (two hours a week in two semesters) were given through 1972 by Ole Kongsdal Jensen, Karen Landschultz, and Oluf M. Thorsen. There were 6 classes in the spring semester and 7 in the autumn semester.

2. Practical training in sound perception and transcription

Courses for beginners as well as courses for more advanced students were given through 1972 by Jørgen Rischel, Nina Thorsen, and Oluf M. Thorsen. (The courses which are based in part on tape recordings and in part on work with informants, 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 by Børge Frøkjær-Jensen in the spring semester (experimental acoustic phonetics and experimental physiological phonetics), and in the autumn semester (registration of the intensity and fundamental frequency of speech).

4. Phonology

Hans Basbøll gave a course for beginners. Jørgen Rischel gave two courses for more advanced students. (The courses form a cycle of three semesters with two hours a week. The contents were: elementary phonological method, trends in phonological theory, generative phonology.)

5. Other courses

Eli Fischer-Jørgensen lectured on auditory phonetics and held seminars on experimental phonetics.

Børge Frøkjær-Jensen gave a course in general acoustics.

Carl Ludvigsen and Børge Frøkjær-Jensen gave a three-week course in Fortran-programming in the autumn.

Peter Holtse gave a course in English phonetics.

Birgit Hutters gave a course in the physiology of the speech organs.

Carl Ludvigsen gave a course in mathematics, electronics, and statistics in the spring and autumn semesters, and, furthermore, he gave a course in statistics for graduate students in the autumn semester.

6. Seminars

The following seminars were held in 1972:
Professor Max Wajskop (Bruxelles) lectured on threshold values for the perception and identification of isolated vowels.

Professor Egerod (Copenhagen) lectured on phonation types and tones in some East-Asian languages.

Professor William S-Y. Wang (Berkeley) lectured on tone languages.

Eli Fischer-Jørgensen and Børge Frøkjær-Jensen gave an account of their impressions from congresses and symposia in 1971-72.

Niels Reinholt Petersen presented a study on the influence of the degree of opening on the perception of vowel length in Danish.

Professor A. Sovijärvi (Helsinki) presented an X-ray film of the speech organs.

Oluf M. Thorsen lectured on vowel features in French.

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

Eli Fischer-Jørgensen participated in the following conferences: the "Conference on Speech Production, Speech Perception and Cortical Function" in Vancouver (April), the "Conference on Speech Communication and Processing" in Newton, Mass. (April), the "83rd Meeting of the Acoustical Society of America" (April), at which she read a paper on "Tape-cutting experiments with stop consonants", and the "2nd International

Phonology Conference" in Vienna (September).

Eli Fischer-Jørgensen gave lectures (in April) at several American universities and research institutions (MIT, Ann Arbor, Michigan State University, Berkeley, Los Angeles, Ohio State, and Harvard) and (in September-October) at universities in Czechoslovakia (Bratislava and Praha) and in Belgium (Vrije Universiteit and Université Libre).

Eli Fischer-Jørgensen read a paper on auditory phonetics at: Videnskabernes Selskab, Dansk Selskab for Logopædi og Foniatri, Fysisk Forening, and Lingvistkredsen.

Eli Fischer-Jørgensen visited the phonetic institutions in MIT, Haskins Laboratories, University of Ann Arbor, Michigan State University, University of California, Berkeley, Los Angeles, Ohio State University. Université Libre de Bruxelles, Vrije Universiteit van Brussel (Neurolinguistiek), Instituut voor Perceptie onderzoek, Eindhoven.

Jørgen Rischel participated in the "XIth International Congress of Linguistics" in Bologna (September).

Jørgen Rischel lectured on the Danish stress rules at the University of Lund, and read a paper on The Danish stress from a generative point of view at "Selskab for nordisk filologi".

Hans Basbøll and Jørgen Rischel participated in the "Spring Seminar", arranged by "Forskningsgruppen för kvantitativ lingvistik" in Åbo and Stockholm (April), and each read a paper at this seminar (see publications: Derivational Processes).

Hans Basbøll lectured on the Danish stød at the University of Lund, and at "Selskab for nordisk filologi".

Børge Frøkjær-Jensen read a paper on glottography at the University of Leeds.

Børge Frøkjær-Jensen visited eight phonetic and/or linguistic institutions in Edinburgh, Leeds, Colchester, Reading, Cambridge, and London (June), under the Foreign University Interchange Scheme.

Karen Landschultz participated in the "3rd International Congress of Applied Linguistics" in Copenhagen.

Karen Landschultz, Nina Thorsen, and Hans-Peter Jørgensen participated in various courses at the Pedagogical Institute of the University of Copenhagen.

INSTRUMENTAL EQUIPMENT OF THE LABORATORY BY THE END OF 1972

1. Instrumentation for speech analysis

2 Sona-Graphs, Kay-Electric, type 6061 A

2 amplitude display/scale magnifier units, Kay-Electric, type 6076 A

1 contour display unit, Kay-Electric, type 6070 A

1 fundamental frequency extractor ("Trans Pitchmeter")

1 intensity meter (dual channel, with active variable highpass and lowpass filters)

1 electro aerometer (dual channel)

2 air-pressure manometers, Simonsen & Weel, type HB 66 (modified)

1 photo-electric glottograph

1 Fabre-Glottograph

1 palatoscope with complete outfit for palatography

1 segmentator, model IPO (Eindhoven)

1 Meyer-Schneider pitchmeter

1 stroboscope, Philips, type PR 9103

1 segmentator, type PT

1 electro-aerometer, four channel, type AM 508/4

1 audio frequency filter, type 445

1 data collecting system, type PT

2. Instrumentation for speech synthesis

1 formant-coded speech synthesizer

1 provisional vowel synthesizer

1 voice-source generator

1 larynx vibrator with power supply

3. Filters

1 LC highpass filter (with stepwise variation of cutoff frequency)

1 active RC lowpass filter

4. Instrumentation for visual recordings

1 mingograph, Elema 42 (4 channels)

1 mingograph, Elema 800 (8 channels)

1 kymograph (with electro-motor)

1 automatic frequency response and spectrum recorder, Brüel & Kjær, type 3332

1 oscilloscope, Telequipment (single beam)

1 oscilloscope, Solartron, type CD 1400 (dual beam)

1 oscilloscope, Tektronix, type 502 A

2 oscilloscopes, Tektronix, type 564 storage

1 dual-trace amplifier, Tektronix, type 3A1

1 four-trace amplifier, Tektronix, type 3A74

1 dual-trace differential amplifier, Tektronix, type 3A3

1 time-base, Tektronix, type 3B3

1 time-base, Tektronix, type 2B67

5. Tape recorders

1 professional recorder, Lyrec (stereo, speeds 7.5" and 15")

2 professional recorders, Lyrec (mono, speeds 7.5" and 15")

1 semi-professional recorder, Movic (stereo, speeds 3 3/4" and 7.5")

3 semi-professional recorders, Revox (stereo, speeds 3 3/4" and 7.5")

1 portable semi-professional recorder, Uher, type 4000, stereo

4 recorders, Tandberg, type 92 SL

1 recorder, Tandberg, type 7, stereo

4 semi-professional recorders, Revox (stereo, speeds 3 3/4" and 7.5")

6. Gramophones

1 gramophone, B & O (mono, Ortofon pick-up)

3 gramophones, Delphon (mono, Ortofon pick-up)

7. Microphones

1 microphone, Neuman, type KM 56

1 dynamic microphone, Sennheiser, type MD 21

2 crystal microphones, of different brands

2 microphones, Altec

1 1" microphone, Brüel & Kjær, type 4131/32

1 1/4" microphone, Brüel & Kjær, type 4135/36

1 larynx microphone

8. Amplifiers

1 microphone pre-amplifier, Telefunken

1 microphone amplifier, Brüel & Kjær, type 2603

1 laboratory amplifier (mono/stereo, with matching for different impedances)

9. Loudspeakers/headphones

7 different loudspeaker systems

5 headphones, AKG, type K 58

3 loudspeakers, Beovox, type 2600

10. General-purpose electronic instrumentation

1 oscillator, Hewlett & Packard, type CD 200

1 function generator, Wavetec VGC III (o.oo3 c/s-1 Mc/S)

1 frequency counter, Rochar, type A 1360 CH (5 digits)

2 vacuum-tube voltmeters, Brüel & Kjær, type 2409

1 vacuum-tube voltmeter, Heathkit, type V-7A

1 vacuum-tube voltmeter, Radiometer, type RV 23 b

1 DC millivoltmeter, Danameter, type 205

1 DC nanoammeter, Danameter, type 206

1 universal meter, Philips, type P 817

1 transistor tester, Taylor, model 44

1 Piston-phone, Brüel & Kjær, type 4220

1 component bridge C/L/R, Wayne Kerr, type B 522

1 AC automatic voltage stabilizer, Claude Lyons, type BTR-5F

4 resistance decades, Danbridge, type DR 4

1 condenser decade, Danbridge, type DK 4 AV

6 stabilized rectifiers

1 multi-generator, Exact, type 126 VCF

2 stabilized rectifiers, Danica, type TPS 1d

1 stabilized rectifier, Danica, type TPS 3c

1 impulse precision sound level meter, Brüel & Kjær, type 2204

1 attenuator set, Hewlett Packard, type 350 D

1 band-pass filter set, Brüel & Kjær, type 1615

1 digital multimeter, Philips, type PM 2422

1 timer counter, Advance, type SC3

1 oscillator, Advance, type J2E

Additional oscillators, rectifiers, etc., for special purposes

11. Outfit for photography

1 Minolta camera SR-1 (with various accessories)

1 complete outfit for reproduction (including 1 Liesegang UNI-RAX with frame)

1 Telford oscilloscope camera, type "A" (polaroid)

12. Equipment for EDP

1 tape punch, Facit, type 4060

1 control unit, Facit, type 5106

1 tape reader, GNT, model 24

1 DECwriter, Digital, LA30P

1 computer, Digital, PDP8/E,8K

13. Projectors

1 Liesegang epidiascope

1 Leitz projector for slides

1 Voigtländer Perkeo Automat - J 150

1 16 m/m tone film projector, Bell & Howell "Filmsound 644"

2 overhead projectors with accessories

1 projector, Leitz, type Pradovit color 250

2 overhead projectors, 3M company

INSTALLATION OF A PDP8/E COMPUTER AT THE LABORATORY

Børge Frøkjær-Jensen

The Institute of Phonetics has obtained a grant from the University of Copenhagen to buy a "Digital Equipment Cooperation" computer. The first part of the computer system was ordered in July 1972 and has been installed in January 1973.

This set-up includes a PDP8/e central processor, which is a parallel-transfer computer with single adressing and fixed word length using 12-bits, 2's complement arithmetic. The cycle time of the 8196-words random address magnetic core memory is 1.2 microseconds, which allows 385,000 additions per second. Standard features are indirect addressing, instruction skip, and program interrupt. Five 12-bits registers which can be loaded directly from the programmer's console are used for control of programs and addresses and for intermediate data storing.

The communication with the central processor unit takes place via a DECwriter data terminal type LA30 (electric typewriter) with parallel interface. Output and input devices for punched paper tapes (GNT reader and GNT puncher both with a speed of 70 characters/sec.) have been interfaced.

The programming can be made in two ways: (1) in higher level language such as the interactive FOCAL or BASIC, or in a compiler language such as FORTRAN; (2) in assembly languages which are translated directly into machine language commands.

The total set-up is to include a dual drive DECtape unit with control unit, a hardware arithmetic unit, a programmable digital clock, an 8 channel analog-to-digital converter with multiplexer, a 16 channel digital-to-analog converter with demultiplexer, and an X/Y recorder.

A STUDY OF THE ROLE OF SYLLABLE AND MORA FOR THE TONAL MANIFESTATION IN WEST GREENLANDIC

Hideo Mase

1. Introduction

The purpose of this investigation is to see the relation between syllable and mora in connection with the tonal manifestation of the word in phrase-final position.

A mora is here to be understood as a unit of length which consists of (C)V, i.e. a short vowel segment with or without a preceding consonant. A short vowel is a one-mora vowel and a long vowel is a two-mora vowel. A syllable-final consonant is also taken as a mora, while a syllable-initial consonant is a non-syllabic, non-mora consonant. Thus, a long (i.e. geminate) consonant consists of one mora and one non-mora consonant. The terms "mora vowel" and "vowel mora" will be used synonymously, and so will "mora consonant" and "consonant mora".

2. Earlier studies of Greenlandic prosodics and the purpose of the present investigation

2.1. Stress

The great 19th century scholar Kleinschmidt (Holtved 1964) recognized that stress placement in Greenland Eskimo is tied up with segmental word structure and has merely a demarcative function. Thalbitzer (1904) gave a less clear account, which was to some extent due to his failure to recognize consonant length correctly in all instances (many of Thalbitzer's apparent examples of contrastive stress placement are entirely predictable once the distinction between long and short consonants is made consistently). Petersen (1970) argues that stress is acquiring some relevance because there is a tendency to clip off the final parts of words in running speech without moving the stress from its position in the representations prior to clipping.

In the present study the problems of stress are disregarded.

2.2. Tonal accent and intonation

The following is a brief survey of previous literature on tone in Greenlandic Eskimo.

2.2.1. Thalbitzer's description of musical accent

First it should be mentioned that there is actually no tonal accent in this language, so what Thalbitzer calls "musical accents" are elements of intonation patterns. Further, the tonal patterns which he describes seem to be what is called a phrase-final intonation below. Our concern here is whether the tone base is a syllable or mora in his description.

Thalbitzer (1904) describes "musical accent" in terms of (combinations of) two relative pitches: high and low. He says (1904) that

> "a change of tone often takes place during the pronunciation of a long sound, and this change is chromatic, not sudden. But changes of tone from syllable to syllable seem to take place in sudden leaps." (p. 136)

A combination of pitches (for example, a high-low pattern) may, according to Thalbitzer, be realized on a short vowel, on a long vowel, or on two short vowels divided into two syllables.

As for the distribution of the three types of "tones", he says that in dissyllabic words both rising (i.e. low-high) and compound (i.e. high-low-high) patterns occur, but the former is the more frequent. In three-syllable words the compound pattern is normal. The last high pitch is usually higher than the first one. In polysyllabic words there is an instance where the syllables of a word fall into pairs, each of which is given a low-high pattern. But the realization of this pattern is often blocked by the occurrence of a compound pattern, i.e. a high-low-high pattern. There are no further descriptions of the tonal patterns of polysyllabic words. Anyway, we can see no systematic description of tonal phenomena in his work.

In table 2-1 some of Thalbitzer's word examples are given for illustration.[1]

---

1) Phonetic notation by Thalbitzer, and phonemic notation by me.

TABLE 2-1

Examples from Thalbitzer (1904)

(˥ = high, ˩ = low, ˧ = middle pitches)

```
                      ˩˥
1-syllable word:    [kʋ·k]     /kuuk/,
                     ˩  ˥                  ˩  ˥
2-syllable words:   [uneq]     /uniq/,   [mɑ·ne]      /maani/,
                     ˩  ˥                  ˩  ˥
                    [ik·ɑ]     /ikka/,   [qarλɩ·t]    /qarliit/,
                     ˩ ˥                   ˥ ˩ ˥
                    [ɑs·aq]    /assaq/,  [mɑ·wnɑ]     /maana/,
                      ˥˩ ˥                 ˥  ˩˥
                    [qup·aq]   /quppaq/, [pɩŋɑ··]     /piŋaa/,
                      ˩  ˥                ˧ ˩ ˥
                    [qum·ut]   /qummut/, [ɑwŋ·ɔ·q]    /aŋŋuuq/,
                      ˥ ˩ ˥
3-syllable word:    [täsɑne]   /tasani/,
                      ˩ ˥  ˩ ˥
4-syllable word:    [sinip·ise]   /sinippisi/,
                      ˩ ˥    ˩ ˥
5-syllable word:    [pujɔrs·imɑsɔq]   /pujursimasuq/.
```

(The phonemicization follows Thalbitzer's transcription, which is not entirely correct as regards quantity. Note in particular that the last-mentioned 2-syllable word actually has a long vowel in the first syllable: modern West Greenlandic /aaŋŋuuq/. Note also that Thalbitzer does not have the distinction between two s-phonemes which is made in this paper. This is a matter of dialect difference. Editors' note.)

2.2.2. Collis/Petersen's description of intonation

Collis (1970 p. 277) says that according to Petersen (op. cit.), intonation is more important than stress.

Moreover, Collis makes the following statements, more or less as a paraphrase of Petersen:[1]

When the stress falls on the antepenultimate, it is accompanied by a rising pitch on the final syllable, except in the case of an interrogative intonation (i.e. a 'yes-no' question, HM) in which case the final syllable has a falling pitch. Thus, when the antepenultimate syllable is a closed one, the intonation pattern is

´ ` ´ (statement),

´ _ ` (question),

for example,

/'taanárpìáq/, and /'taannárpiàq/.

However, since the closed syllable always takes stress, the stress accent does not always fall on the antepenultimate, but the intonation contour is unaffected by the existence of closed syllable(s) and remains unchanged. Thus Petersen affirms the priority of intonation as the distinctive feature of the last three syllables of the word.

---

1) As far as I can see in Petersen (op. cit.), stress and intonation is not so explicitly described as Collis will have it. Consequently I refer to the latter's description partly as his own and partly as Petersen's, i.e. as Collis/Petersen.

It is not certain whether the two intonation patterns described by Collis/Petersen are ones that occur in the same position in the sentence. The pattern which is said to be associated with "statement", i.e. the "rise-fall-rise" pattern, seems to be the phrase-final intonation defined in 2.2.3.

According to Collis/Petersen the tonal contours of the last three syllables of words serve to distinguish intonational patterns. However, as will be shown in the following section, it is actually the tonal contour of the last three (vowel) moras which is fixed, and not the contour of the last three syllables of the word.

2.2.3. Rischel's description of intonation

Rischel (1971 and personal communication) considers the single vowel segment rather than the syllable to be the unit of measure in West Greenlandic prosody. A long vowel is analysed as VV, i.e. a succession of two segments. He makes the following tentative analysis of the West Greenlandic intonation pattern in terms of terminal contours:[1]

(a) Phrase-internal contour: a succession of words belonging to one sentence may form a phrase. A word boundary which is non-final in a phrase, is signalled by high pitch on the last vowel segments (i.e. vowel moras), the penultimate one having in most cases a slightly higher pitch than the neighbouring ones ("raised" pitch).

---

1) A similar description of some of the intonation types is found in Petersen (1970). (Cf. 2.2.2.)

(b) Phrase-final contour: the penultimate vowel segment (i.e. vowel mora) of the word has a low pitch, whereas the antepenultimate and ultimate ones have a high pitch. For example, in /akiwuq/ and /akiwuŋa/ the underlined vowels are low-pitched. When the penultimate vocalic segment is part of a long vowel or of a diphthong /ai/ (the diphthong occurs only word-finally), the part in question has a low pitch. Thus, there is a falling pitch on the penultimate syllable of a word like /qaawuq/, and there is a rising pitch on the ultimate syllable of a word like /aŋiqaaq/.

(c) Yes-no question contour: the ultimate vowel segment of the word is low-pitched, the penultimate and antepenultimate vowel segments are high(er)-pitched, the tonal peak being on the antepenultimate one (except in some types of very short words). The relationship may be described as high-mid-low, or perhaps rather: raised-high-low. (If the penultimate vocalic segment is low-pitched, the word will presumably be perceived as having contour (e) below.) Examples of (c) are, schematically:

takuwaaŋā ?

takuwaaŋŋā ?

(d) Wh-question contour: a low pitch occurs on the penultimate vocalic segment, and a higher pitch occurs on the antepenultimate and ultimate vowel segments, for example,

kina takuwījuk ? ("kina" = "who").

It is not clear whether this contour exhibits some systematic difference from (b) above. In (d) the ultimate pitch may not be equal in height to the antepenultimate one.

(e) Absolutely final contour: the antepenultimate vowel segment has a high pitch, but the last two segments are low-pitched. The pattern of distribution of this contour is not well known, and it is difficult to distinguish rigidly between (b) and (e). Presumably the latter has a more marked status, i.e., (e) tends to be used only if finality is to be expressed explicitly. Schematically, the intonation is

takuwaa̅ŋa̅,

takuwaa̅ŋŋa̅.

Other syllables than those involved in the contours mentioned in (a)-(e) have normally a neutral pitch, but three tendencies should be mentioned.

(I) Other conditions being equal, heavy syllables tend to be a little higher-pitched than light ones. (II) Long vowels tend to have less abrupt tone jumps than bisyllabic sequences. (III) There is a certain tendency for pitches over the utterance to go down gradually as the utterance goes on. But this gradual fall of pitches is not always observed.

In Rischel's description we understand that the tone base is definitely a vowel mora. His view is also summarized in Mase-Rischel (op. cit.):

> "the tonal contour of the word is closely bound to the vowel segments. For example, the typical contour before a non-final pause is characterized by nonlow tone on the antepenultimate and ultimate vowels, and low tone on the penultimate, no matter how they are distributed on syllables; schematically:

(a̅)a̅wa̅a 'he fetches it' (a̅)k̅iw̅ar̅a 'I answer him'

The tonal rules are as yet quite imperfectly understood, but there is evidence that the vowel segment (the "syllabic") plays at least as important a role as the syllable in prosodic rules." (p. 193)

It now seems evident that Thalbitzer's 'compound' (i.e. high-low-high) pattern and Collis/Petersen's intonation pattern for the statement (i.e. rise-fall-rise pattern) correspond to Rischel's phrase-final intonation, if and only if each of the last three syllables contains a short vowel.

2.3. Purpose of the present investigation

According to Rischel's scheme we should get the same type of tonal realization on the last three vowel moras of a word, irrespective of their distribution on syllables. For example, the phrase-final contour will appear as follows on three-mora words of three and two syllables:

i-kī-tit; na-kā-taq;

tiī-tit (tii-tit); kaā-taq (kaa-taq);

i-wīit (i-wiit); a-māaq (a-maaq);

Likewise, we should get different pitch distributions on syllables of three-syllable words which have different numbers of moras.

For example,

$\text{na-k}\overline{\text{a}}\text{-taq}$, and

$\overline{\text{a}}\text{-naa-w}\overline{\text{a}}\text{a}$,

but not

$^{*}\text{a-n}\overline{\text{aa}}\text{-waa}$.

The first thing is to see how Rischel's terminal contours are manifested acoustically.

The next question is whether a consonant mora behaves more or less like a vowel mora. By way of definition, every syllable-final consonant is a mora.[1] Two kinds of consonant moras should be distinguished: (I) voiceless, and (II) voiced.

(I,a) Word-medial single versus double (= geminate) consonants: the pitch on the voiceless consonant mora is of course not realized. If, in word pairs such as

{ katak / mattak } and { taqaq / qaqqaq },

we find that the timing of the tonal contour of /mattak/ and /qaqqaq/ with respect to the two vowel segments is different from that of /katak/ and /taqaq/ and is similar to that of, for example, /nakataq/, then we can say that the voiceless consonant mora counts the same as a vowel mora. If, on the other hand, the tonal contour of /mattak/ and /qaqqaq/ is realized in the same way as in /katak/ and /taqaq/, then we can say that the voiceless consonant mora is not counted as a tone base.

---

1) I.e., such consonants are here labelled "moras", not by reference to some potential or actual role in the realization of intonation contours, but on a distributional basis. The category may, however, be found useful both in connection with quantity (cf. Mase-Rischel 1971), stress, and intonation.

(I,b) Word-final consonant mora: Do /illu/ and /illut/ have the same or different tonal contours?

(II) Voiced consonant mora: The same question could be asked about the voiced consonant mora. In the type of material investigated here the word-medial nasal is the only type of voiced consonant mora.[1]

In words such as

1. maa<u>n</u>a,
2. ma<u>nn</u>a, and
3. maa<u>nn</u>a,

it is possible that a voiced consonant mora counts like a vowel mora, so that 1 and 2 are grouped together, or it may not, so that 1 and 3 are grouped together.

A priori, one would certainly consider it fairly probable that a nasal mora behaves in the same way as a vowel mora, whereas reference to voiceless consonant moras in the timing of intonation would seem a most improbable hypothesis. Both Collis' and Rischel's descriptions suggest that consonant moras altogether do not participate in regulating the intonation patterns.

The purpose of this investigation is to see how these things are realized acoustically.

3. Procedure in the acoustic-phonetic investigation

3.1. Subjects and recording

Three native, male speakers of central West Greenlandic (RP, AS, and GT) recorded the material during July and August

1) It must be mentioned that reduction phenomena, and their implications for tonal realizations, are not considered here. According to Petersen (1970) long (voiced) consonants may arise in connected speech by optional deletion of vowels between two like consonants. These long consonants would form a separate object of study from the point of view of intonation.

1971 in the sound treated studio of the Institute of Phonetics, University of Copenhagen. Each subject spoke each of two texts (see 3.2. below) six or seven times, so that twelve or thirteen recordings of each word token were obtained. The word lists were read distinctly at an even speed, but no further instructions were given.

3.2. Word list and text

A list of 90 wordforms (phonological words) was prepared.[1] In the text these words were arranged in random order, but, in order not to disturb the natural rhythm of speech, care was taken not to put next to each other words which contain drastically different numbers of syllables and/or moras, and words which, in sequence, might arouse some unusual or comical association of meaning. Two texts with different order of words were made, in such a way that a word which was put at the beginning or end of the line in one text was put in the middle of the line in the other text.

64 words with different numbers of syllables and moras were chosen for this investigation.[2] (Six five-mora words = nos. 86-91 were added after RP's recording was finished.) But due to difficulties in the precise segmentation, some of the words which include [j] (/j/), [w] (/w/)[3], [ɣ] (/g/) or [ʁ] (/r/) had to be omitted. Consequently, the number of words which are included in the following data is not the same among

---

1) I am very much obliged to Mr. Jørgen Rischel who was kind enough to make the original list.

2) The remaining 26 words contain various voiced continuants in intervocalic position and may be used later, particularly for the purpose of investigating the status of combinations of vowel+semivowel+vowel vis-à-vis long vowels on the one hand and sequences of vowel+fricative+vowel on the other.

3) Segmentation was not difficult when the phoneme /w/ was realized as a [v]- or [β]-like sound.

all three subjects. This explains why there are some missing tokens in the tables below. The word list is shown in table 3-1.

Each separate item of the word list was expected to be read with the terminal contour which is here called "phrase-final" in accordance with Rischel (see 2.2.3. above). It was rather easy to check whether this is actually the case. The phrase-final intonation is supposed to have a high-low-high pitch pattern on the last part of the word. The last vowel segments in the wh-question has a similar pattern, but it is uncertain whether this is phonologically the same intonation. The other patterns,according to Rischel,have descending pitches on the last two vowel moras, i.e. the penultimate vowel mora is higher than the ultimate one. With some exceptions which could be sorted out on a principled basis,[1] all words in the recordings were found to exhibit the first-mentioned relationship, i.e. a rise from the penultimate to the ultimate vowel mora. So, if there is any remaining mixture of intonations in the material of this investigation, it must be a confusion between phrase-final and wh-question intonations. However, no single question word is found in the text, so this possible source of variation can be disregarded.

3.3. Registration

All the material was registered on mingograms displaying a fundamental frequency curve, a duplex oscillogram, and four intensity curves (with an integration time of 5 msec) with different frequency filterings: (1) HP with cut-off frequency at

---

1) A few words which seem to have been spoken with absolutely final intonation because they occurred at the end of a line, were omitted from the measurements.

1000 or 1200 Hz, (2) HP at 500 Hz, (3) LP at 500 Hz, and (4) no filtering. Some wide-band spectrograms were taken to check the segmentation of the mingograms, but actually the difficulty of segmentation was the same both on the spectrogram and mingogram.

TABLE 3-1

List of words
(in phonemic notation)

No.	word	Nos. in texts	words to be compared with
1.	taaq	(18, 64)	
2.	puut	(80, 80)	20. puukka
3.	illu	(2, 79)	4. illut
4.	illut	(6, 72)	3. illu
5.	katak	(4, 78)	9. mattak, 19. kaataq
6.	masak	(7, 74)	7. nasaq, 10. aššak
7.	nasaq	(22, 59)	6. masak, 10. aššak
8.	taqaq	(24, 56)	12. qaqqaq
9.	mattak	(3, 76)	5. katak
10.	aššak	(9, 71)	6. masak, 11. aššaq
11.	aššaq	(20, 62)	7. nasaq, 10. aššak
12.	qaqqaq	(21, 58)	8. taqaq
13.	maniq	(26, 54)	16. anniq, 17. marniq, 23. taaniq, 31. maanniq, 32. maarniq
14.	tuwit	(82, 83)	29. tuwikka
15.	manna	(84, 84)	22. maana, 30. maanna
16.	anniq	(29, 53)	See word 13.
17.	marniq	(44, 36)	See word 13.
18.	tiitit	(50, 28)	24. iwiit, 26. ikitit
19.	kaataq	(28, 52)	5. katak, 25. amaaq, 27. nakataq
20.	puukka	(81, 81)	2. puut
21.	quurqa	(85, 85)	36. quurqaa

TABLE 3-1

(continued)

No.	word	Nos. in texts	words to be compared with
22.	maana	(54, 29)	15. manna, 30. maanna
23.	taaniq	(52, 31)	See word 13.
24.	iwiit	(41, 41)	18. tiitit, 26. ikitit
25.	amaaq	(13, 70)	{ 33. ammaaq, 19. kaataq, 27. nakataq, 28. awataq
26.	ikitit	(25, 57)	18. tiitit, 24. iwiit
27.	nakataq	(42, 38)	19. kaataq, 25. amaaq, 28. awataq
28.	awataq	(47, 35)	{ 42. awataaq 19. kaataq, 25. amaaq, 27. nakataq
29.	tuwikka	(83, 82)	14. tuwit
30.	maanna	(19, 61)	See word 15.
31.	maanniq	(32, 48)	} See word 13.
32.	maarniq	(38, 42)	
33.	ammaaq	(16, 63)	25. amaaq
34.	aataa	(12, 69)	
35.	aappaa	(15, 67)	
36.	quurqaa	(35, 47)	21. quurqa
37.	tuurpara	(61, 18)	
38.	naalagaq	(43, 39)	
39.	kisiisa	(79, 2)	
40.	sukkuutit	(33, 49)	
41.	killuutit	(49, 33)	
42.	awataaq	(53, 30)	28. awataq
43.	naparutaq	(48, 32)	
44.	uqurluni	(36, 44)	
45.	iqqanarpuq	(66, 16)	
46.	iqqarsarpuq	(64, 17)	
49.	ammaqaaq	(89, 90)	
50.	anaawaa	(77, 4)	57. annaawaa

TABLE 3-1
(continued)

No.	word	Nos. in texts	words to be compared with
51.	tikiŋŋilaq	(88, 86)	
52.	ikusimmi	(91, 89)	
53.	aamaliwik	(39, 43)	
54.	amaaliwik	(46, 34)	
55.	aputituqaq	(86, 88)	
56.	akisimawuq	(87, 91)	
57.	annaawaa	(62, 20)	50. anaawaa
60.	iišaššaawuq	(67, 12)	63. ilišaššaawuq
62.	naluumasurtuq	(37, 45)	
63.	ilišaššaawuq	(75, 7)	60. iišaššaawuq
64.	ilutusimapput	(63, 21)	

3.4. Measurements

Only five recordings (= samples) of each word token by each informant were used in this investigation. If possible, these five samples were selected from the first two or three recordings of each text, but if a difficulty arose in these words, other recordings were used. One word, AS's No. 51 /tikiŋŋilaq/ has only four samples, because /i/ between the initial /t/ and the following /k/ was devoiced in all recordings except these four.[1] Other typical reasons for discarding a sample were: (a) straightforward problems of segmentation (see 3.4.1.), (b) extreme shortness of a vowel (after a voiceless consonant) causing a badly defined fundamental frequency, (c) a gradual onset of a word-initial vowel after a period of irregularity (glottal catch), which made it difficult to define the duration of the vowel proper, (d) other occasional occurrences of highly irregular F_o-curves (bad phonation?).

3.4.1. Duration

Segmentation was carried out according to conventions.[2] Accuracy of measurement is $\pm$ 0.25 cs.

---

1) This devoicing seems to have occurred due to strong affrication of the preceding /t/ ([t^s]), and perhaps partly due to devoicing of an unstressed high vowel between voiceless consonants.

2) As for the accuracy of measurement, the disadvantage of a mingogram without the "aid of supplementary broad band spectrograms" (Fant 1957, p. 62) seems to have been overcome to a high degree by the use of four intensity curves with different filtering frequencies.

Long vowels, i.e. two-mora vowels, occur in various consonantal environments, i.e. (-C)(C)_(C)(C-), and in various positions in the words. We cannot a priori decide where the mora boundary lies between two moras of a long vowel. Consequently, a long vowel was arbitrarily divided into two identical halves at the middle of the long segment. As for the long nasal, the first two thirds of the total duration was arbitrarily assigned as a mora-nasal and the last one third as the initial nasal consonant of the following syllable. This arbitrary decision was based on the observation that the minimal or maximal value of the fundamental frequency often occurs after a lapse of approximately two thirds of the total duration.

Initial and final consonants of words were not measured.

As mentioned in 3.2. it was not easy to delimit intervocalic [j], [w], [ɣ], and [ʁ]. In the transcriptions used here, /j, w/ after respectively /i, u/ stands for more or less audible glides. After other vowels /j, w/ are clearly distinct both from zero and from other fricatives. However, even when [j] was clearly heard, the change from [j] to the vowel was smooth and gradual on the acoustic curve. As a result, words containing /j/ were all omitted. [w] (/w/) was just as difficult to segment as [j], but when the phoneme was realized as [v] or [β], segmentation was not difficult. [ɣ] posed just the same problem as [j]. The uvular fricative [ʁ] showed a very low and often very irregular pitch, and the pitch of the adjacent vowel was very much influenced and became rather irregular. The influence was most conspicuous in words spoken by GT, who had a very deep voice, so that most of his words containing /r/ had to be discarded. (And, as shown later, words containing the consonants in question caused a certain irregularity in the results.)

Segmentation was also sometimes difficult in the case of nasal plus vowel or vowel plus nasal.

3.4.2. Fundamental frequency

Fundamental frequencies are measured at the following points: for a short vowel, the beginning and end of the segment; for a long vowel, the beginning, middle, and end of the segment; and for a mora nasal, the beginning, middle and end of the segment. As mentioned in 3.4.1., two thirds of the total duration of a long (i.e. geminate) nasal are assumed to represent a mora nasal, so the middle and end of the mora-segment correspond to the 1/3 and 2/3 points of the total duration of a long nasal, respectively. When a long vowel or a long nasal has a rise-fall or fall-rise pitch, more than three points were measured. Accuracy of measurement is $\pm$ 1.25 Hz.

As is generally known, the fundamental frequency curve of a vowel segment is sometimes distorted and shows some irregularity during the first and/or last one to two centiseconds of the segment, due to the influence of adjacent consonants (especially the preceding one). However, such irregularity might also arise as a consequence of the instrumental processing. In order to test the reliability of the frequency measuring device used for registration, a calibration with an intermittent sine wave of 125 Hz was carried out, and the result shows that the first and last cycles are registered a little lower than the rest.

The problems associated with these boundary phenomena were to some extent circumvented in the measurements: when any irregularities were found at the beginning and/or end of the segment, the tangent was drawn from the stable part and measured as the beginning and/or end. But it was sometimes difficult to decide whether the irregularity was an acoustic reality or an instrumental artifact. (Cf. section 5. and fig. 5-3.)

4. Results

4.1. Method of calculation

The original raw data were fed into a computer and run through a program which calculated the arithmetic mean and standard deviation[1] for each word token.[2] Each word token consists of five samples, except AS's word token No. 51, which has only four samples. Calculations were made for the duration, mean fundamental frequency, and slope of each tone base, and the difference of mean fundamental frequency between two tone bases adjacent to each other.

4.1.1. Mean fundamental frequency

The mean fundamental frequency of a tone base was calculated in the following way. [beg = beginning of a tone base, mid = mid point, end = end; V = vowel mora, N = nasal mora.]

1) short vowel: Mean = $(F_o(\text{beg}) + F_o(\text{end}))/2$,

---

1) The standard deviation as published in the tables below was calculated by the formula:

$$s = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N}}$$

, which actually gives a somewhat too small estimate of the standard deviation in a large population because of the small number of samples of each token. Correction was made for the use of N instead of (N-1) in this formula, when the findings were tested for significance.

2) I am very much obliged to Mrs. B. Auld and Cand. polyt. Carl Ludvigsen. Mrs. Auld was kind enough to undertake the laborious work of punching the cards and to write the first part of the program, and Mr. Ludvigsen very kindly completed the program in a very short time.

2) long vowel:
   i) when counted as one tone base: Mean = (F_o(beg+mid+end))/3,
   ii) when counted as two tone bases: Mean = F_o(beg + end)/2, where for the first tone base, beg = beg, end = mid of a long vowel, and for the second tone base, beg = mid, end = end of a long vowel,

3) nasal mora: Mean = F_o(beg + mid + end)/3, where beg = beg, mid = 1/3 point, end = 2/3 point of a long nasal,

4) a vowel mora + a nasal mora = one tone base: Mean = (VF_o(beg + end) + NF_o(mid + end))/4.

In the data processed by the computer, the mean frequency values were calculated over three points of measurement for a long vowel, but a long vowel sometimes had a fall-rise or rise-fall contour, instead of a unidirectional rise or fall. For those long vowels which showed such rise-fall or fall-rise contours more than three points were measured and a hand-calculation was carried out to get the mean fundamental frequency value. (In the tables the values calculated by the latter method are shown in [ ].)

These values of mean fundamental frequencies of tone bases are compared in order to reveal any high-low relation between neighbouring tone bases. The reason why the mean values are compared is that, since the role of the direction and degree of pitch change of a tone base is not known, we cannot be sure whether a comparison reasonably should be made between beginnings and/or ends of two tone bases.

According to Rossi (1971), when one cannot hear a glissando during a (vowel) segment, it will be perceived as if it has a static pitch corresponding to the frequency at the transition between the second and third portions of the segment. Consequently, a comparison of mean fundamental frequencies seems to be reasonable at the first step, no matter whether a rise or fall is significant.

The absolute values of mean fundamental frequencies of tone bases may vary according to segmental constituents (and other factors). Thus, for example, the fundamental frequency of a penultimate vowel, which is lower than that of the antepenultimate in the same word, may be higher than that of the antepenultimate vowel in another word, which has the same relationship between the antepenultimate and penultimate vowels. This relation may also be seen among the samples within one and the same word token. In the following, our main concern is the difference in fundamental frequency between vowels.

Word tokens are classified not according to their vowel quality nor consonantal environment, but according to the numbers of syllables and/or moras.

4.1.2. Slope

A vowel or a nasal consonant has usually a rising or a falling pitch. That is, a certain fundamental frequency change or interval can be observed during a certain period of its duration. One question about this change is whether the absolute value of the frequency interval of the vowel (or the nasal consonant) is a constant, or whether it increases with increasing vowel duration.

Another question, which is of specific interest to the present investigation, is whether the direction of the fundamental frequency movement in a vowel is determined by the syllable or mora position.

To answer the first question the slope was calculated as:

$$\text{Slope} = (F_o(\text{end}) - F_o(\text{beg}))/\text{Duration}.$$

See the values of samples in the following word token used for illustration.

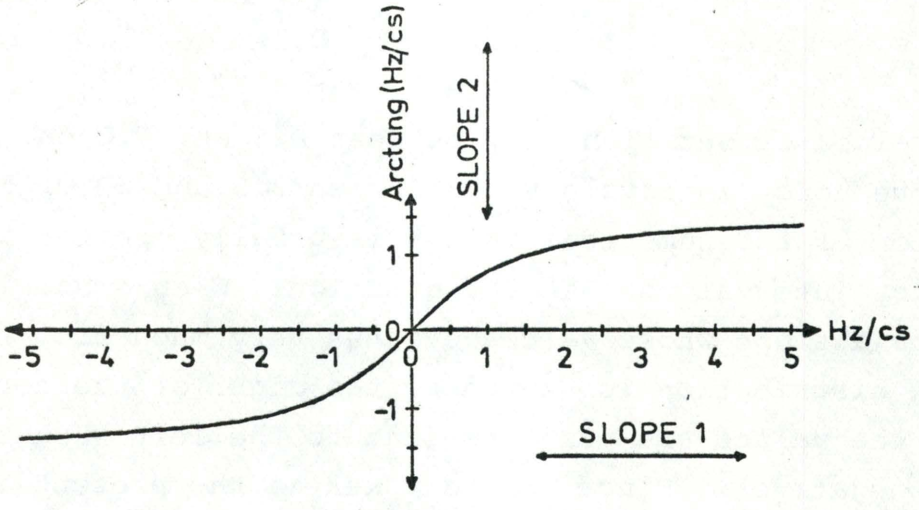
Sample [No.]	Dur. [cs]	Interv. [Hz]	Slope [Hz/cs]	Arctang(Slope) [radians]
1	5.5	12.5	2.27	1.19
2	5.5	2.5	0.46	0.43
3	5.0	20.0	4.00	1.33
4	7.0	5.0	0.71	0.62
5	4.5	5.0	1.11	0.84
$\overline{X}$	5.5	9.0	1.71	0.88
s	0.84	6.4	1.30	0.34

The values of duration vary between 4.5 and 7.0 cs, while the frequency intervals vary between 2.5 and 20.0 Hz. The variation of the numerical values is usually greater in the frequency interval than in the duration. When a sample (or two) has a slope whose value deviates very much from the others, the distribution is skewed to the right of the normal curve when the values are positive (and to the left when the values are negative). Since a word token in the present investigation has only five samples, this kind of deviation affects the results rather strongly.

But it is true that all five samples in the token shown above have positive slopes, i.e. they are all different from zero in positive direction. When the value of the slope is transformed to Arctangent(Slope), i.e. Arctangent(Hz/cs), the distribution of samples becomes less skewed. And we can see that this particular token has a slope whose mean value is significantly different from zero in positive direction. The figure on the next page shows the relation of values in (Hz/cs) and Arctang(Hz/cs).

In the following, the first slope (i.e. Hz/cs) will be called "slope 1" and the second transformed one (i.e. Arctang (Slope)) "slope 2". As for the slope, vowels are classified according to their position in the word and consonantal environments.

Relation between slopes 1 and 2



4.1.3. Other remarks

The position of the syllable or mora in the word is reckoned from the end of the word. Thus, the ultimate syllable or mora is called S1 or M1, the penultimate one S2 or M2, the antepenultimate one S3 or M3, and so on (where "S" and "M" stand for "syllable" and "mora" respectively).

Structurally a syllable or mora consists of (C)V(V)(C). But, since the consonants (except some mora consonants functioning as tone bases) are disregarded in the following, the vowels in S1, S2, M1, M2, etc. are often, for convenience, abbreviated as S1, S2, M1, M2, etc.

When a syllable is considered to be a tone base one vowel is counted as one tone base, whether it is long or short. When a mora is considered as a tone base, a short vowel and a nasal mora are counted as one base each, and a long vowel is counted as two bases.

4.2. Duration

4.2.1. Consonants

The statistic calculations were only carried out for short and long consonants in two-syllable, two-mora words and for all short and long nasal consonants.

As for short and long voiceless consonants in two-syllable, two-(vowel) mora words, a phonemically long consonant is phonetically longer than a phonemically short one in the same or a similar environment. The difference is significant at the 99% confidence level.

In this investigation, a "mora nasal" is arbitrarily considered as 2/3 of the total duration of a long nasal. In this case, a short nasal consonant is shorter than a mora nasal in the same or similar environment. The difference is significant at the 99% confidence level.

4.2.2. Vowels

No exhaustive comparison was attempted, since not all short and long vowels appear in (sub)minimal pairs or sets. It is not altogether certain whether the difference between short and long vowels is always significant at, say, the 95 or 99% confidence level. But the difference (i.e. a short vowel is shorter than a long one) is significant at the 99% level in some of the pairs or sets.

4.3. Mean fundamental frequency

It is assumed in this paper that the phrase-final intonation is signalled by a high-low-high pitch pattern on the last three tone bases. That is, with regard to the mean fundamental frequency, the tone base of S1 or M1 can be expected to be higher than that of S2 or M2, and that of S2 or M2 can be expected to be lower than that of S3 or M3. The tonal contours of the preceding tone bases (i.e. from S4 to S6, or from M4 to M6) are largely unknown.

4.3.1. Is the tonal pattern regulated on a syllable basis?

4.3.1.1. Words containing only short vowels

The number of words are:

["-s" = syllable]

	RP	AS	GT
2-s	14	15	15
3-s	3	4	4
4-s	3	6	4
5-s	0	1	2
6-s	1	1	1
total	21	27	26

See table 4-1.

(1) S2-S1: S2 will be lower than S1: The vowel of S2 is lower than that of S1, and the difference is statistically significant[1] in all the tokens of three speakers.

1) Hereafter, statistically significant means at the 95% confidence level.

(2) S3-S2: S3 will be higher than S2: All the vowels of S2 are lower than those of S3, and the difference is significant in all tokens.

(3) S4-S3: In all four tokens spoken by RP, S4 is lower than S3 in their mean values, but there are some samples where S4 is higher than S3. The result is that the difference between the two vowels cannot be shown to be significant. In all eight tokens spoken by AS, S4 is significantly lower than S3, and so it is in all seven tokens spoken by GT.

(4) S5-S4: In all six tokens (three persons pooled), S5 is lower than S4, and the difference is significant except in GT's token (56) akisimawuq.

(5) S6-S5: There is only one six-syllable word for each speaker. S6 of RP's token is significantly higher than S5, while S6 is significantly lower than S5 in the same tokens spoken by AS and GT.

The mean fundamental frequencies of vowels of both S3 and S1 can generally be expected to be rather high. S3 is generally lower than S1. The number of word tokens in which S3 is significantly lower than S1 is 5 out of 7 of RP's, 11 out of 12 of AS's, and 10 out of 11 of GT's words.

As a whole, the difference is smaller in RP's than in AS's and GT's words. This is due to the fact that the range of variation of the fundamental frequency is smallest in RP's words.

In the tokens where the differences between adjacent vowels are significant, the mean values of vowels which are assumed to be the higher ones are more than 5 Hz greater than those of the lower ones. One exception is AS's token (64) ilutusimapput, where the difference between S4 and S3 is 4.0 Hz.

In addition to the assumption on the pitches of the last three tone bases, when we tentatively assume that S3 is higher than S4, that S4 is higher than S5, and that S5 is higher than S6, the number of tokens where the differences between adjacent tone bases are significant is as follows. [The figure to the left of the slash indicates the number of tokens where the difference is significant; the figure to the right of the slash indicates the total number of tokens.]

	RP	AS	GT
S2 < S1	21/21	27/27	26/26
S3 > S2	7/7	12/12	11/11
S4 < S3	0/4	8/8	7/7
S5 < S4	1/1	2/2	2/3
S6 < S5	0/1	1/1	1/1

[NB : In RP's token S6 is significantly higher than S5.]

TABLE 4-1

Mean fundamental frequency values (in Hz), and the significance of differences of mean fundamental frequency values between adjacent vowels. The calculations are carried out on a "syllable" basis in words containing only "short" vowels.

(+: agrees significantly with the assumption,
-: differs significantly from the assumption,
O: no significance.)

Subject: RP	Mean fund. freq.				Significance		
	S4	S3	S2	S1	4<3	3>2	2<1
5. ka-tak			132.0	152.3			+
6. ma-sak			129.5	151.0			+
7. na-saq			135.5	154.0			+
8. ta-qaq			132.8	155.8			+
3. il-lu			140.5	157.5			+
4. il-lut			139.5	154.8			+
9. mat-tak			138.3	154.0			+
10. aš-šak			140.5	152.0			+
11. aš-šaq			144.8	150.8			+
12. qaq-qaq			135.8	147.8			+
13. ma-niq			138.0	157.5			+
15. man-niq			137.5	154.0			+
16. an-niq			143.3	150.5			+
17. mar-niq			136.5	150.8			+
26. i-ki-tit		151.3	139.3	159.5		+	+
27. na-ka-taq		143.5	133.0	153.5		+	+
28. a-wa-taq		141.8	131.0	156.3		+	+
44. u-qur-lu-ni	141.8	147.3	134.3	155.5	O	+	+
45. iq-qa-nar-puq	148.3	148.3	137.3	155.3	O	+	+
46. iq-qar-sar-puq	140.0	147.3	137.0	153.5	O	+	+
64. i-lu-tu-si-map-put	144.8	152.5	125.5	155.8	O	+	+
			S6	S5		6<5	5<4
64.			145.5	139.3		-	+

TABLE 4-1
(continued)

Subject: AS	Mean fund. freq.				Significance		
	S4	S3	S2	S1	4<3	3>2	2<1
5. ka-tak			125.3	159.5			+
6. ma-sak			117.8	159.5			+
7. na-saq			120.3	154.8			+
8. ta-qaq			120.3	157.3			+
3, il-lu			132.0	167.8			+
4. il-lut			137.3	167.8			+
9. mat-tak			126.8	164.8			+
10. aš-šak			129.5	164.0			+
11. aš-šaq			128.8	162.5			+
12. qaq-qaq			131.5	161.0			+
14. tu-wit			124.5	153.3			+
13. ma-niq			121.8	158.3			+
15. man-na			140.8	161.8			+
16. an-niq			129.8	155.8			+
17. mar-niq			128.0	156.8			+
26. i-ki-tit		142.5	128.0	167.3		+	+
27. na-ka-taq		138.8	121.8	157.8		+	+
28. a-wa-taq		130.5	121.5	156.3		+	+
29. tu-wik-ka		148.3	127.5	161.8		+	+
43. na-pa-ru-taq	126.0	138.3	117.0	153.0	+	+	+
44. u-qur̂-lu-ni	127.3	151.3	122.5	159.8	+	+	+
45. iq-qa-nar-puq	129.8	147.8	126.5	162.5	+	+	+
46. iq-qar-sar-puq	125.8	146.3	119.3	153.5	+	+	+
51. ti-kiŋ-ŋi-laq	138.1	152.5	118.4	148.8	+	+	+
52. i-ku-sim-mi	136.3	148.8	134.0	153.8	+	+	+
55. a-pu-ti-tu-qaq	143.5	150.3	126.3	158.5	+	+	+
64. i-lu-tu-si-map-put	145.0	149.0	124.5	159.5	+	+	+
			S6	S5		6<5	5<4
55.				128.3			+
64.			126.5	132.8		+	+

TABLE 4-1
(continued)

Subject: GT	Mean fund. freq.				Significance		
	S4	S3	S2	S1	4<3	3>2	2<1
5. ka-tak			100.0	138.0			+
6. ma-sak			101.3	139.8			+
7. na-saq			96.5	133.5			+
8. ta-qaq			98.3	130.3			+
3. il-lu			111.3	144.5			+
4. il-lut			114.0	144.3			+
9. mat-tak			101.8	138.5			+
10. aš-šak			106.8	141.5			+
11. aš-šaq			104.0	144.3			+
12. qaq-qaq			97.3	132.8			+
14. tu-wit			103.5	134.5			+
13. ma-niq			97.8	138.8			+
15. man-na			110.0	140.3			+
16. an-niq			103.8	134.0			+
17. mar-niq			101.8	142.0			+
26. i-ki-tit		120.8	107.8	144.0		+	+
27. na-ka-taq		109.8	97.3	130.5		+	+
28. a-wa-taq		110.0	96.8	135.5		+	+
29. tu-wik-ka		122.0	110.3	137.5		+	+
44. u-qur-lu-ni	102.0	130.8	105.0	139.3	+	+	+
45. iq-qa-nar-puq	105.0	120.3	107.5	147.8	+	+	+
51. ti-kiŋ-ŋi-laq	115.8	122.0	105.0	132.8	+	+	+
52. i-ku-sim-mi	109.8	127.0	121.3	140.5	+	+	+
55. a-pu-ti-tu-qaq	111.3	127.3	102.8	132.8	+	+	+
56. a-ki-si-ma-wuq	112.3	132.3	103.3	131.3	+	+	+
64. i-lu-tu-si-map-put	121.3	138.3	110.0	150.5	+	+	+
			S6	S5		6<5	5<4
55.				103.3			+
56.				108.3			O
64.			105.0	116.3		+	+

4.3.1.2. Words containing both short and long vowels

Table 4-2 shows the mean fundamental frequencies of the vowels and the differences of mean fundamental frequencies between the vowels.

The assumption on the pitches of the last three tone bases holds true of two-syllable words of the type VV-V and three-syllable words of the type VV-V-V (where VV is a long vowel and V is a short vowel, and a hyphen shows a syllable boundary, and consonants are disregarded). But in two- and three-syllable words of other types, the results are quite different.[1] Cf. fig. 4-1 which shows the values of the mean fundamental frequencies of vowels in adjacent syllables. The values of vowels which can be expected to be the higher ones are given on the y-axis and the one which can be expected to be the lower ones on the x-axis.

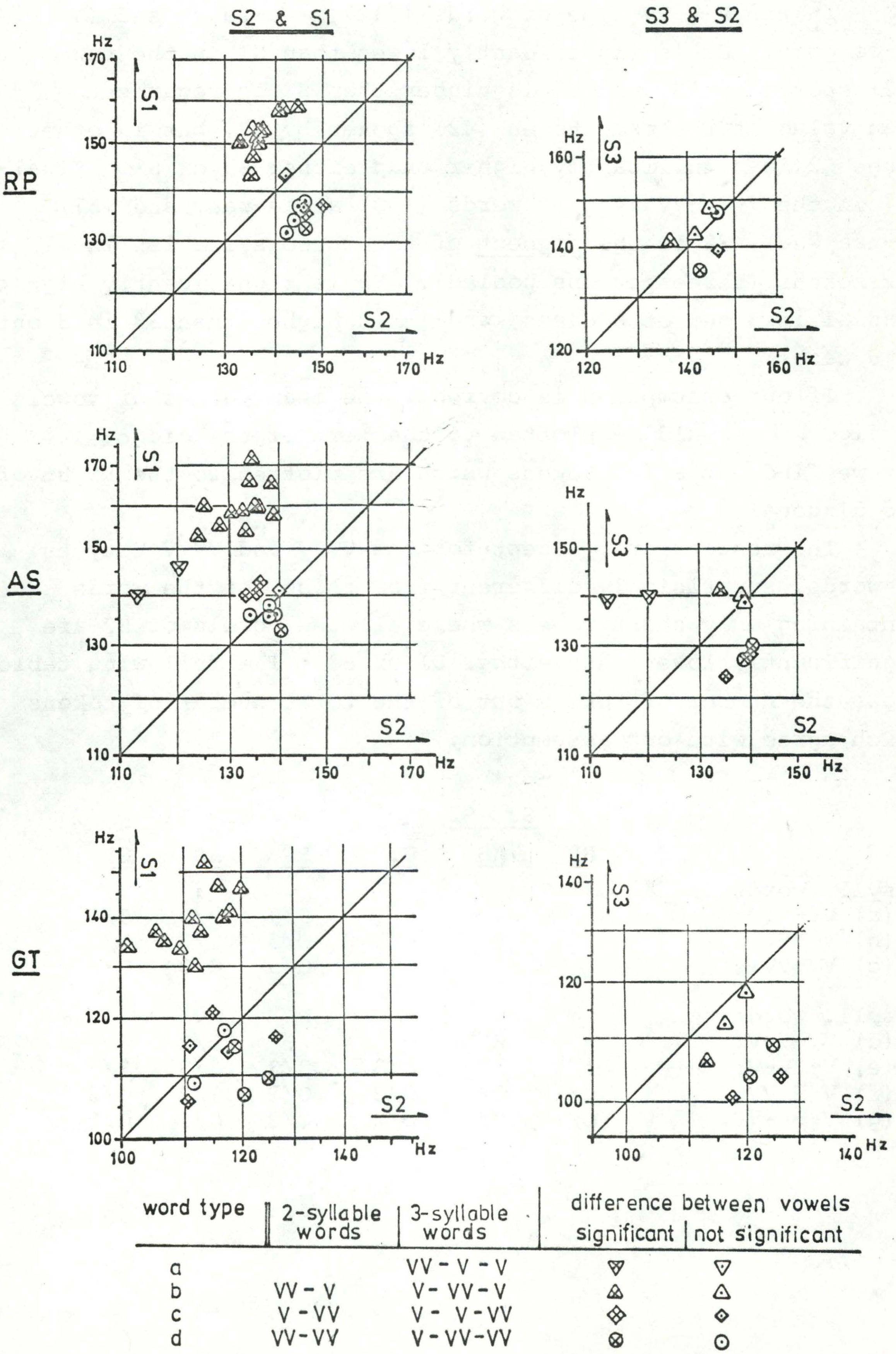
In two-syllable words of the V-VV type of words ((24) i-wiit, (25) a-maaq, and (33) am-maaq), S2 is either higher or lower than S1. This difference is significant in some cases, but not in other cases. In the VV-VV type of words ((34) aa-taa, (35) aap-paa, and (36) quur-qaa), no significant difference is seen between S2 and S1, except that in the tokens (36) spoken by RP and GT, S2 is significantly higher than S1.

With regard to three-syllable words, in the V-VV-V type of words ((39) ki-sii-sa, (40) suk-kuu-tit, and (41) kil-luu-tit), S2 is significantly lower than S1, but S2 is either higher or lower than S3. That is, S2 is significantly

1) We have a few words which have more than four syllables (see table 3-1), but since the last three syllables of these words contain only short vowels, they are irrelevant to the present task of checking the pitches of the last three tone bases.

Fig. 4-1.

Values of mean fundamental frequencies of vowels in two adjacent syllables. [The values of vowels supposed to be the higher of the two are given on the y-axis, and the other supposed to be the lower on the x-axis.]



lower than S3 in the token (39) spoken by RP and in the token (41) spoken by AS. S2 is significantly higher than S3 in the token (39) spoken by GT. In other cases the difference between S3 and S2 is not significant.

In the V-V-VV type of words ((42) a-wa-taaq, and (49) am-ma-qaaq), S2 is significantly lower than S1 in the token (42) spoken by AS, and S2 is higher than S3 as regards the mean value of the same token (42) spoken by RP, but in other cases S2 is significantly higher than either S3 or S1. Finally, in the V-VV-VV type of words ((50) a-naa-waa, and (57) an-naa-waa), S2 is the highest of the three syllables in all six tokens (three persons pooled). S2 is significantly higher than S1 in 3 out of 6 cases, and it is higher than S3 in 5 out of 6 cases.

If our assumption is correct, the mean values of vowels in fig. 4-1 should be plotted to the left of the diagonal, but we find not a few tokens which are plotted to the right of the diagonal.

The results are, except for the VV-V and VV-V-V types of words, surprisingly different from those for the words containing only short vowels where all the vowels of S2 are significantly lower than either S1 or S3. The following table shows the number of tokens out of the total number of tokens which agree with our assumption.

	S3 > S2			S2 < S1		
	RP	AS	GT	RP	AS	GT
2-syll. words:						
(a) VV-V:				9/9	9/9	9/9
(b) V-VV:				0/3	2/2	1/3
(c) VV-VV:				0/3	0/3	0/3
3-syll. words:						
(d) VV-V-V:	-	2/2	-	-	2/2	-
(e) V-VV-V:	1/3	1/3	0/3	3/3	3/3	3/3
(f) V-V-VV:	0/1	0/2	0/2	0/1	1/2	0/2
(g) V-VV-VV:	0/2	0/2	0/2	0/2	0/2	0/2

TABLE 4-2

Mean fundamental frequency values (in Hz), and the significance of differences of mean fundamental frequency values between adjacent vowels. The calculations are carried out on a "syllable" basis in words containing both "short" and "long" vowels.

([ ] indicate values calculated over more than three points of measurements.
+: agrees significantly with the assumption,
-: differs significantly from the assumption,
O: no significance.)

	Mean fund. freq.				Significance		
Subject: RP	S4	S3	S2	S1	4<3	3>2	2<1
18. tii-tit			140.3	157.3			+
19. kaa-taq			134.8	146.8			+
20. puuk-ka			137.5	153.3			+
21. quur-qa			135.0	143.5			+
22. maa-na			132.5	150.3			+
23. taa-niq			132.7	150.3			+
30. maan-na			136.5	150.3			+
31. maan-niq			137.2	153.3			+
32. maar-niq			134.8	152.8			+
24. i-wiit			141.8	143.2			O
25. a-maaq			145.8	[136.3]			O
33. am-maaq			149.8	137.2			-
34. aa-taa			[144.7]	[136.7]			O
35. aap-paa			143.8	134.0			O
36. quur-qaa			146.5	[132.1]			-
39. ki-sii-sa		140.8	136.5	151.0		+	+
40. suk-kuu-tit		147.8	144.3	159.0		O	+
41. kil-luu-tit		142.3	141.3	158.0		O	+
42. a-wa-taaq		139.0	146.3	[135.0]		O	-
50. a-naa-waa		134.8	[142.3]	[131.3]		-	O
57. an-naa-waa		146.8	145.8	137.3		O	O
53. aa-ma-li-wik	142.5	144.3	134.8	156.3	O	+	+
54. a-maa-li-wik	142.8	145.2	133.5	151.8	O	+	+

TABLE 4-2
(continued)

Subject: AS	S4	S3	S2	S1	4<3	3>2	2<1
	Mean fund. freq.				Significance		
18. tii-tit			134.3	171.3			+
19. kaa-taq			[124.1]	152.8			+
20. puuk-ka			134.8	159.5			+
21. quur-qa			135.5	159.0			+
22. maa-na			[128.1]	155.3			+
23. taa-niq			125.2	159.8			+
30. maan-na			132.7	154.0			+
31. maan-niq			130.5	158.3			+
32. maar-niq			132.3	158.8			+
25. a-maaq			133.0	140.0			+
33. am-maaq			136.0	142.3			+
34. aa-taa			137.8	135.5			O
35. aap-paa			133.7	136.0			O
36. quur-qaa			138.2	137.7			O
37. tuur-pa-ra		139.0	112.8	139.8		+	+
38. naa-la-gaq		139.7	120.5	146.0		+	+
39. ki-sii-sa		138.5	138.8	157.8		O	+
40. suk-kuu-tit		140.0	138.5	165.5		O	+
41. kil-luu-tit		141.0	134.2	165.8		+	+
42. a-wa-taaq		124.0	135.0	[140.5]		-	+
49. am-ma-qaaq		127.8	139.8	140.7		-	O
50. a-naa-waa		127.0	[139.1]	[136.0]		-	O
57. an-naa-waa		130.0	[140.3]	[132.7]		-	-
53. aa-ma-li-wik	131.8	146.5	127.8	155.0	+	+	+
54. a-maa-li-wik	123.5	140.8	128.0	157.5	+	+	+
62. na-luu-ma-sur-tuq	138.8	148.0	128.3	161.5	+	+	+
				S5			5<4
62.				125.0			+

TABLE 4-2
(continued)

Subject: GT	Mean fund. freq.				Significance		
	S4	S3	S2	S1	4<3	3>2	2<1
18. tii-tit			113.3	151.8			+
19. kaa-taq			101.2	133.8			+
20. puuk-ka			117.8	140.8			+
21. quur-qa			117.0	139.8			+
22. maa-na			[105.3]	136.8			+
23. taa-niq			106.7	134.8			+
30. maan-na			111.7	129.8			+
31. maan-niq			109.5	133.5			+
32. maar-niq			111.3	140.0			+
24. i-wiit			114.8	[120.9]			+
25. a-maaq			110.5	[105.7]			-
33. am-maaq			110.8	114.7			0
34. aa-taa			117.0	117.7			0
35. aap-paa			110.8	108.7			0
36. quur-qaa			118.7	115.2			-
39. ki-sii-sa		106.3	113.0	137.0		-	+
40. suk-kuu-tit		118.0	120.0	146.0		0	+
41. kil-luu-tit		112.5	116.0	146.5		0	+
42. a-wa-taaq		100.3	117.3	114.0		-	0
49. am-ma-qaaq		103.8	126.3	116.2		-	-
50. a-naa-waa		103.8	[120.4]	[106.8]		-	-
57. an-naa-waa		108.8	125.0	[109.4]		-	-
53. aa-ma-li-wik	105.5	126.0	104.0	137.3	+	+	+
54. a-maa-li-wik	106.5	122.0	106.0	137.3	+	+	+
60. ii-šaš-šaa-wuq	110.5	128.0	113.67	134.0	+	+	+
63. i-li-šaš-šaa-wuq	110.8	126.5	117.5	136.3	+	+	+

	S5	5<4
63.	103.8	+

4.3.2. Is the tonal pattern regulated on a mora basis?

4.3.2.1. The role of the voiceless consonant mora

In section 1., every syllable-final consonant was considered as a mora. If a word-final consonant is taken to count as a tone base, then in the (C)V-CVC type of words the vowel of the final syllable will be M2, and the vowel of the initial syllable will be M3. In this case, M2 is the mora which has the highest mean fundamental frequency. It is higher than that of M3, and this difference is significant! This is contrary to the assumption concerning phrase-final intonation. Further, in one pair of words, (3) illu and (4) illut, there is no significant difference of mean fundamental frequency between the final vowels of the two words. Consequently, the word-final consonant (which is always voiceless) should not be counted as a tone base.

If part of the word-medial voiceless consonant is taken as a tone base, the first vowel of (C)VC-CV(C) words will be M3, and the word-medial consonant mora will be M2. As is seen in table 4-1 , there is no difference of mean fundamental frequency between the vowel of M3 (= S2) in the (C)V-C-CV(C) type of words and that of M2 (= S2) in the (C)V-CV(C) type.

Consequently, in the following a voiceless consonant mora is not taken as a tone base. The behaviour of a voiced consonant mora, i.e. a nasal mora, will be taken up later in 4.3.2.4.

4.3.2.2. Words containing only short vowels (and no long nasal)

Since a voiceless consonant mora is not taken as a tone base, the results are the same as those obtained for words with short vowels on a syllable basis, except that words containing a long nasal are, for the time being, not included.

4.3.2.3. Words containing both short and long vowels (and no long nasal)

Here a short vowel is counted as one mora vowel, and a long vowel is counted as two mora vowels. See table 4-3.

(1) M2-M1: M2 will be lower than M1: The numbers of tokens are 20 for RP, 22 for AS, and 22 for GT. In all the tokens M2 is significantly lower than M1.

(2) M3-M2: M3 will be higher than M2: The numbers of tokens are 18 for RP, 20 for AS, and 20 for GT. In all the tokens M3 is significantly higher than M2.

(3) M4-M3: The numbers of tokens are 10 for RP, 13 for AS, and 12 for GT. In five tokens spoken by RP, M4 is significantly lower than M3. In three tokens of the same person M4 is lower than M3, but the difference is not significant. Two tokens of his ((39) kisiisa, and (40) sukkuutit) have significantly higher M4. In the tokens spoken by AS, 11 out of 13 tokens have significantly lower M4. Two tokens ((40) sukkuutit, and (41) killuutit) have either higher or lower M4. In the tokens spoken by GT, M4 is significantly lower than M3 in four- and five-mora words (10 tokens). But in two six-mora words the result is different. In (60) iišaššaawuq, M4 is significantly higher than M3, and in (63) ilišaššaawuq, M4 is either higher or lower than M3. (These six-mora words are found only in GT's list.)

(4) M5-M4: In the token (50) anaawaa spoken by RP, M5 is significantly lower than M4, but in two other tokens spoken by the same person the difference is not significant. In AS's four tokens and in GT's five tokens, M5 is significantly lower than M4.

(5) M6-M5: The numbers of tokens are 1 for AS and 2 for GT. M6 is significantly lower than M5.

As for M3 and M1, M3 is significantly lower than M1 in 12 out of 18 tokens of the material spoken by RP, in 17 out of 20 tokens by AS, and in 14 out of 20 tokens by GT.

The number of tokens which agree with our assumption is as follows. [The figure to the left of the slash indicates the number of tokens in which the difference is significant; the figure to the right of the slash indicates the total number of tokens.]

	RP	AS	GT
M2 < M1	20/20	22/22	22/22
M3 > M2	18/18	20/20	20/20
M4 < M3	5/10	11/13	10/12
M5 < M4	1/3	4/4	5/5
M6 < M5	-	1/1	2/2

[NB: In RP's (39) and (40) M4 is significantly higher than M3; in GT's (60) M4 is significantly higher than M3.]

When words containing only short vowels or both short and long vowels are put together, but words containing long nasals are omitted, the number of tokens which agree with our assumption on the mora basis is as follows.

	RP	AS	GT
M2 < M1	38/38	44/44	43/43
M3 > M2	25/25	30/30	29/29
M4 < M3	5/14	17/19	15/17
M5 < M4	2/4	6/6	7/8
M6 < M5	0/1	2/2	3/3

In the tokens in which the differences between adjacent moras are significant, the following tokens have values of less than 5.0 Hz difference. [Values in Hz.]

M3 > M2:

AS:	(42)	awataaq	(4.3)
GT:	(22)	maana	(5.0)
	(24)	iwiit	(5.0)

M4 < M3:

RP:	(35)	aappaa	(4.2)
	(36)	quurqaa	(4.2)
	(50)	anaawaa	(3.0)
	(53)	aamaliwik	(4.3)

M5 < M4:

GT:	(53)	aamaliwik	(5.0)

M6 < M5:

GT:	(60)	iišaššaawuq	(3.0)

TABLE 4-3

Mean fundamental frequency values (in Hz), and the significance of differences of mean fundamental frequency values between adjacent vowel moras. The calculations are carried out on a "mora" basis in words containing both "short" and "long" vowels (and no long nasal).

([ ]: indicates values calculated over more than two points of measurement,
+: agrees significantly with the assumption,
-: differs significantly from the assumption,
O: no significance.)

Subject: RP	Mean fund. freq. M4	M3	M2	M1	Significance 4<3	3>2	2<1
1. ta-aq			[127.8]	[140.3]			+
2. pu-ut			142.3	156.3			+
18. ti-i-tit		141.5	135.5	157.3		+	+
19. ka-a-taq		134.8	126.3	146.8		+	+
20. pu-uk-ka		141.3	133.8	153.3		+	+
21. qu-ur-qa		135.8	130.5	143.5		+	+
22. ma-a-na		132.5	125.5	150.3		+	+
23. ta-a-niq		133.8	126.5	150.3		+	+
24. i-wi-it		141.8	134.3	147.5		+	+
25. a-ma-aq		145.8	[132.2]	[145.2]		+	+
34. a-a-ta-a	[140.8]	[148.5]	[128.7]	142.5	+	+	+
35. a-ap-pa-a	140.8	145.0	126.3	135.0	+	+	+
36. qu-ur-qa-a	141.3	145.5	[126.3]	[136.2]	+	+	+
39. ki-si-i-sa	140.8	136.5	130.0	151.0	-	+	+
40. suk-ku-u-tit	147.8	144.0	138.3	159.0	-	+	+
41. kil-lu-u-tit	142.3	142.8	135.0	158.0	O	+	+
42. a-wa-ta-aq	139.0	146.3	[129.8]	143.5	O	+	+
50. a-na-a-wa-a	140.3	[143.3]	[126.5]	138.0	+	+	+
53. a-a-ma-li-wik	140.0	144.3	134.8	156.3	+	+	+
54. a-ma-a-li-wik	142.8	144.3	133.5	151.8	O	+	+

	M5	5<4
50.	134.8	+
53.	141.8	O
54.	142.8	O

TABLE 4-3
(continued)

Subject: <u>AS</u>	Mean fund. freq. M4	M3	M2	M1	Significance 4<3	3>2	2<1
1. ta-aq			125.3	145.0			+
2. pu-ut			136.0	154.8			+
18. ti-i-tit		141.3	126.8	171.3		+	+
19. ka-a-taq		[130.5]	[115.5]	152.8		+	+
20. pu-uk-ka		140.3	125.5	159.5		+	+
21. qu-ur-qa		140.0	127.0	159.0		+	+
22. ma-a-na		[131.3]	122.8	155.3		+	+
23. ta-a-niq		129.0	118.5	159.8		+	+
25. a-ma-aq		133.0	125.5	150.8		+	+
34. a-a-ta-a	130.0	143.5	120.8	142.8	+	+	+
35. a-ap-pa-a	126.5	136.5	123.0	146.0	+	+	+
36. qu-ur-qa-a	132.8	140.3	122.8	143.8	+	+	+
37. tu-ur-pa-ra	134.3	139.5	112.8	139.8	+	+	+
38. na-a-la-gaq	133.3	140.8	120.5	146.0	+	+	+
39. ki-si-i-sa	138.5	145.0	127.0	157.8	+	+	+
40. suk-ku-u-tit	140.0	141.0	127.3	165.5	0	+	+
41. kil-lu-u-tit	141.0	139.5	124.3	165.8	0	+	+
42. a-wa-ta-aq	124.0	135.5	[131.2]	153.8	+	+	+
50. a-na-a-wa-a	136.0	[142.2]	[125.7]	150.5	+	+	+
53. a-a-ma-li-wik	134.3	146.5	127.8	155.5	+	+	+
54. a-ma-a-li-wik	136.3	144.0	128.0	157.5	+	+	+
62. na-lu-u-ma-sur-tuq	140.3	148.0	128.3	161.5	+	+	+
			M6	M5		6<5	5<4
50.				127.0			+
53.				125.8			+
54.				123.5			+
62.			125.0	133.5		+	+

TABLE 4-3
(continued)

Subject: GT	Mean fund. freq.				Significance		
	M4	M3	M2	M1	4<3	3>2	2<1
1. ta-aq			96.0	119.3			+
2. pu-ut			113.3	134.5			+
18. ti-i-tit		118.3	108.3	151.8		+	+
19. ka-a-taq		107.5	97.3	133.8		+	+
20. pu-uk-ka		122.0	114.0	140.8		+	+
21. qu-ur-qa		118.5	110.8	139.8		+	+
22. ma-a-na		106.0	101.0	136.8		+	+
23. ta-a-niq		107.5	100.3	134.8		+	+
24. i-wi-it		114.8	[109.8]	133.3		+	+
25. a-ma-aq		110.5	[96.4]	120.5		+	+
34. a-a-ta-a	109.8	123.8	103.3	121.8	+	+	+
35. a-ap-pa-a	103.0	118.3	95.5	115.8	+	+	+
36. qu-ur-qa-a	110.8	126.0	101.3	121.3	+	+	+
39. ki-si-i-sa	106.3	118.5	106.8	137.0	+	+	+
40. suk-ku-u-tit	118.0	127.0	110.0	146.0	+	+	+
41. kil-lu-u-tit	112.5	122.8	106.0	146.5	+	+	+
42. a-wa-ta-aq	100.3	117.3	[98.0]	121.8	+	+	+
50. a-na-a-wa-a	116.5	[124.2]	[99.2]	116.0	+	+	+
53. a-a-ma-li-wik	107.0	126.0	104.0	137.3	+	+	+
54. a-ma-a-li-wik	117.0	126.3	106.0	137.3	+	+	+
60. i-i-šaš-ša-a-wuq	128.0	119.3	103.3	134.0	-	+	+
63. i-li-šaš-ša-a-wuq	126.5	123.8	106.5	136.3	0	+	+

		M6	M5	6<5	5<4
50.			103.8		+
53.			102.0		+
54.			106.5		+
60.		108.0	111.0	+	+
63.		103.8	110.8	+	+

4.3.2.4. Words containing a long nasal

In 4.3.2.1. it was mentioned that a voiceless consonant mora is not a tone base. A nasal mora, on the other hand, may behave differently, since this is a voiced sound with a well-defined fundamental frequency throughout its duration. Three hypotheses should be tested concerning the role of a nasal mora as a tone base: (a) a nasal mora is not a tone base, (b) a vowel mora and the following nasal mora form a single compound tone base, and (c) a nasal mora is an independent tone base.

It should be remembered that a nasal mora is arbitrarily assigned two thirds of the total duration of a long nasal.

As for the results, see table 4-4.

In words in which the nasal mora /N/ occurs as M2 or M3 when counted as a tone base, the results are as follows:

(a) When /N/ is completely omitted, the difference between mean values of M3 and mean values of M2 in (30) maanna, (31) maanniq, and (32) maarniq is less than 5.0 Hz, except in AS's token (32).

(b) When /V/ + /N/ is counted as a compound tone base, our assumption holds true of all the tokens. That is, the differences between M2 and M1, and M3 and M2 are significant, and the mean values of the higher tone base are more than 5.0 Hz higher than those of the lower.

(c) When /N/ is counted as an independent tone base, M2 is significantly lower than M1. But the difference between M3 and M2 is not significant in GT's token (15) manna. And the mean values of M3 are higher than those of M2, but are less than 5.0 Hz higher than those of M2 in RP's tokens (30), (31), and (32), and GT's tokens (32) and (17) marniq. Notice here that M2 is a nasal mora and M3 is a vowel mora. The difference between M4 and M3 is contrary to the assumption in most of the cases. In the following tokens M4 is significantly higher than M3: for RP, (31); for AS, (30), (31), (51) tiki<u>ŋ</u>-ŋilaq, and (52) ik<u>u</u>s<u>i</u>mmi; for GT, (32) and (52). In the rest of the tokens the difference is not significant, except in GT's (33) <u>a</u>-<u>m</u>-maaq, in which M3 (= /N/) is significantly higher than M4.[1]

---

1) In the transcription used here it should be remembered that /rn/ stands for a long nasal with an initial uvular component. Hence /r/ may be referred to as a nasal mora.

TABLE 4-4

Mean fundamental frequency value (in Hz), and the significance of differences of mean fundamental frequency values between adjacent moras in words containing a long nasal.

(tb = tone base)

Subject: RP	M4	M3	M2	M1	4<3	3>2	2<1
	Mean fund. freq.				Significance		
N ≠ tb							
15. ma(n)-na			137.5	154.0			+
16. a(n)-niq			143.3	150.5			+
17. ma(r)-niq			136.5	150.8			+
30. ma-a(n)-na		136.0	133.3	150.3		O	+
31. ma-a(n)-niq		139.8	134.8	153.3		+	+
32. ma-a(r)-niq		135.0	131.3	152.8		O	+
33. a(m)-ma-aq		149.8	131.3	141.3		+	+
57. a(n)-na-a-wa-a	145.0	144.3	128.3	139.3	O	+	+
				M5			5<4
57.				146.8			O
V+N = tb							
15. man-na			131.6	154.0			+
16. an-niq			136.5	150.5			+
17. mar-niq			131.9	150.8			+
30. ma-an-na		136.0	130.5	150.3		+	+
31. ma-an-niq		139.8	132.5	153.3		+	+
32. ma-ar-niq		135.0	130.8	152.8		O	+
33. am-ma-aq		147.0	131.3	141.3		+	+
57. an-na-a-wa-a	M5 (/an/) = 144.6, "5 < 4" = O.						
N = tb							
15. ma-n-na		137.5	128.3	154.0		+	+
16. a-n-niq		143.3	133.3	150.5		+	+
17. ma-r-niq		136.5	129.7	150.8		+	+
30. ma-a-n-na	136.0	133.3	129.0	150.3	O	+	+
31. ma-a-n-niq	139.8	134.8	130.8	153.3	-	+	+
32. ma-a-r-niq	135.0	131.3	129.3	152.8	O	+	+
33. a-m-ma-aq	149.8	145.7	131.3	141.3	O	+	+
57. a-n-na-a-wa-a	M6 = 146.8, M5 (/N/) = 143.7, "6 < 5" = 0, "5 < 4" = 0.						

TABLE 4-4
(continued)

Subject: <u>AS</u>	Mean fund. freq.				Significance		
	M4	M3	M2	M1	4<3	3>2	2<1
<u>N ≠ tb</u>							
15. ma(n)-na			140.8	161.8			+
16. a(n)-niq			129.8	155.8			+
17. ma(r)-niq			128.0	156.8			+
30. ma-a(n)-na		133.8	130.5	154.0		+	+
31. ma-a(n)-niq		132.8	130.8	158.3		O	+
32. ma-a(r)-niq		136.0	130.0	158.8		+	+
52. i-ku-si(m)-mi	136.3	148.8	134.0	153.8	+	+	+
33. a(m)-ma-aq		136.0	127.8	150.5	+	+	+
51. ti-ki(ŋ)-ŋi-laq	138.1	152.5	118.4	148.8	+	+	+
49. a(m)-ma-qa-aq	127.8	139.8	129.3	145.5	+	+	+
57. a(n)-na-a-wa-a	140.5	[141.9]	[123.3]	144.8	O	+	+
				M5			5<4
57.				130.0			+
<u>V+N = tb</u>							
15. man-na			133.3	161.8			+
16. an-niq			123.0	155.8			+
17. mar-niq			123.3	156.8			+
30. ma-an-na		133.8	123.6	154.0		+	+
31. ma-an-niq		132.8	124.4	158.3		+	+
32. ma-ar-niq		136.0	124.5	158.8		+	+
52. i-ku-sim-mi	136.3	148.8	125.8	153.8	+	+	+
33. am-ma-aq		137.8	127.8	150.5		+	+
51. ti-kiŋ-ŋi-laq	138.1	146.9	118.4	148.8	+	+	+
49. am-ma-qa-aq	130.1	139.8	129.3	145.5	+	+	+
57. an-na-a-wa-a	M5 (/an/) = 131.9, "5 < 4" = +.						

TABLE 4-4
(continued)

Subject: AS	Mean fund. freq.				Significance		
	M4	M3	M2	M1	4<3	3>2	2<1
N = tb							
15. ma-n-na		140.8	130.0	161.8		+	+
16. a-n-niq		129.8	120.3	155.8		+	+
17. ma-r-niq		128.0	119.5	156.8		+	+
30. ma-a-n-na	133.8	130.5	119.3	154.0	-	+	+
31. ma-a-n-niq	132.8	130.8	120.7	158.3	O	+	+
32. ma-a-r-niq	136.0	130.0	121.0	158.8	-	+	+
52. i-ku-si-m-mi	148.8	134.0	121.7	153.8	-	+	+
33. a-m-ma-aq	136.0	140.0	127.8	150.5	O	+	+
51. ti-ki-ŋ-ŋi-laq	152.5	144.6	118.4	148.8	-	+	+
49. a-m-ma-qa-aq	131.2	139.8	129.3	145.5	+	+	+
57. a-n-na-a-wa-a	140.5	[141.9]	[123.3]	144.8	O	+	+
			M6	M5		6<5	5<4
52.				136.3			+
51.				138.1			+
49.				127.8			O
57.			130.0	132.7		O	+

TABLE 4-4
(continued)

Subject: GT

	Mean fund. freq.				Significance		
	M4	M3	M2	M1	4<3	3>2	2<1
N ≠ tb							
15. ma(n)-na			110.0	140.3			+
16. a(n)-niq			103.8	134.0			+
17. ma(r)-niq			101.8	142.0			+
30. ma-a(n)-na		112.3	109.3	129.8		0	+
31. ma-a(n)-niq		110.8	108.5	133.5		0	+
32. ma-a(r)-niq		109.5	106.0	140.0		+	+
52. i-ku-si(m)-mi	109.8	127.0	121.3	140.5	+	+	+
33. a(m)-ma-aq		110.8	101.5	121.3		+	+
51. ti-ki(ŋ)-ŋi-laq	115.8	122.0	105.0	132.8	+	+	+
49. a(m)-ma-qa-aq	103.8	126.3	102.8	122.3	+	+	+
57. a(n)-na-a-wa-a	122.0	127.3	[101.2]	118.0	+	+	+
				M5			5<4
57.				108.8			+
V+N = tb							
15. man-na			108.8	140.3			+
16. an-niq			99.0	134.0			+
17. mar-niq			100.1	142.0			+
30. ma-an-na		112.3	102.8	129.8		+	+
31. ma-an-niq		110.8	103.1	133.5		+	+
32. ma-ar-niq		109.5	103.2	140.0		+	+
52. i-ku-sim-mi	109.8	127.0	113.5	140.5	+	+	+
33. am-ma-aq		116.4	101.5	121.3		+	+
51. ti-kiŋ-ŋi-laq	115.8	124.5	105.0	132.8	+	+	+
49. am-ma-qa-aq	106.6	126.3	102.8	122.3	+	+	+
57. an-na-a-wa-a	M5 (/an/)= 110.4, "5<4" = +.						

TABLE 4-4
(continued)

Subject: GT	Mean fund. freq. M4	M3	M2	M1	Significance 4<3	3>2	2<1
N = tb							
15. ma-n-na		110.0	108.3	140.3		0	+
16. a-n-niq		103.8	97.2	134.0		+	+
17. ma-r-niq		101.8	99.2	142.0		+	+
30. ma-a-n-na	112.3	109.0	99.2	129.8	0	+	+
31. ma-a-n-niq	110.8	108.5	99.2	133.5	0	+	+
32. ma-a-r-niq	109.5	106.0	101.3	140.0	-	+	+
52. i-ku-si-m-mi	127.0	121.3	109.2	140.5	-	+	+
33. a-m-ma-aq	110.8	119.3	101.5	121.3	+	+	+
51. ti-ki-ŋ-ŋi-laq	122.0	124.8	105.0	132.8	0	+	+
49. a-m-ma-qa-aq	108.0	126.3	102.8	122.3	+	+	+
57. a-n-na-a-wa-a	122.0	127.3	[101.2]	118.0	+	+	+

			M6	M5		6<5	5<4
52.				109.8			+
51.				115.8			+
49.				103.8			+
57.			108.8	111.0		0	+

4.4. Slope

4.4.1. Slope of short vowels (on a syllable basis)

See 4.4.3. below.

4.4.2. Slope 1 of vowels (on a syllable basis)

Table 4-5 shows the slope 1 (i.e. [Hz/cs]) of vowels in words containing both short and long vowels and containing no long nasal.[1] Both long and short vowels are counted as one vowel segment. Is the slope of the vowel fixed according to the syllable position? The results show that except for the slope of S1, there is no definite tendency. The vowel of S1 has a positive slope. Notice that there are some tokens in which the vowel of S2 has a significantly positive slope.

---

1) Notice that "+" and "-" are used here in a different way than in the presentation of differences of mean fundamental frequency values.

TABLE 4-5

Slope 1 (Hz/cs) of vowels in two- and three- "syllable" words containing both "short" and "long" vowels.

($\overline{X}$ = mean, s = standard deviation)
("+" at the upper right of the X value: slope 1 is significant)

Subject: RP	S3		S2		S1	
	$\overline{X}$	s	$\overline{X}$	s	$\overline{X}$	s
18. tii-tit			-0.6^{+}	0.2	0.1	0.5
19. kaa-taq			-0.4	0.3	-0.5	1.0
20. puuk-ka			-0.6^{+}	0.4	-0.8	0.9
21. quur-qa			-0.5^{+}	0.2	-0.7	1.3
22. maa-na			-0.2^{+}	0.1	1.2^{+}	0.7
23. taa-niq			-0.2^{+}	0.1	2.2^{+}	0.7
24. i-wiit			-1.6^{+}	0.3	1.9^{+}	0.2
25. a-maaq			-0.6^{+}	0.3	1.2^{+}	0.4
34. aa-taa			0.6^{+}	0.4	1.4^{+}	0.4
35. aap-paa			0.4	0.4	1.2^{+}	0.3
36. quur-qaa			0.4^{+}	0.2	1.1^{+}	0.2
39. ki-sii-sa	-2.0^{+}	0.4	-0.4^{+}	0.2	-0.7	1.3
40. suk-kuu-tit	-0.3	0.8	-0.3	0.3	-0.6	1.0
41. kil-luu-tit	-1.6^{+}	1.1	-0.4^{+}	0.1	0.0	0.0
42. a-wa-taaq	-1.1	0.8	0.9^{+}	0.5	0.9^{+}	0.2
50. a-naa-waa	-0.7	0.9	-0.3^{+}	0.2	1.3^{+}	0.3
57. an-naa-waa	-0.6	1.0	-0.3^{+}	0.1	1.0^{+}	0.3

TABLE 4-5
(continued)

Subject: AS	S3 $\overline{X}$	S3 s	S2 $\overline{X}$	S2 s	S1 $\overline{X}$	S1 s
18. tii-tit			-1.2^{+}	0.3	0.3	0.5
19. kaa-taq			-0.5^{+}	0.3	1.6^{+}	0.5
20. puuk-ka			-1.7^{+}	0.4	1.5^{+}	0.6
21. quur-qa			-1.7^{+}	0.3	1.9^{+}	0.4
22. maa-na			-0.7^{+}	0.2	3.4^{+}	0.9
23. taa-niq			-0.4^{+}	0.1	3.8^{+}	0.8
25. a-maaq			0.7	0.6	1.8^{+}	0.3
34. aa-taa			0.9^{+}	0.3	1.9^{+}	0.2
35. aap-paa			1.1^{+}	0.3	2.0^{+}	0.1
36. quur-qaa			0.5	0.4	1.9^{+}	0.2
37. tuur-pa-ra	0.5^{+}	0.3	-0.3	0.4	3.7^{+}	0.6
38. naa-la-gaq	0.3	0.2	0.0	0.3	3.9^{+}	0.1
39. ki-sii-sa	-2.5^{+}	0.9	-1.2^{+}	0.2	1.7^{+}	0.3
40. suk-kuu-tit	-2.2^{+}	0.7	-0.8^{+}	0.1	-0.8	1.2
41. kil-luu-tit	-2.0^{+}	0.4	-0.9^{+}	0.1	0.1	0.6
42. a-wa-taaq	0.0	0.6	1.6^{+}	0.4	1.9^{+}	0.2
50. a-naa-waa	0.2	0.3	-0.1	0.3	1.9^{+}	0.4
57. an-naa-waa	0.2	0.3	-0.6^{+}	0.2	1.9^{+}	0.2

TABLE 4-5
(continued)

Subject: GT	S3 $\overline{X}$	S3 s	S2 $\overline{X}$	S2 s	S1 $\overline{X}$	S1 s
18. tii-tit			-0.6^{+}	0.1	1.1	0.8
19. kaa-taq			-0.6^{+}	0.2	1.3^{+}	0.6
20. puuk-ka			-0.9^{+}	0.1	1.3^{+}	0.5
21. quur-qa			-0.7^{+}	0.1	1.6^{+}	0.5
22. maa-na			-0.3^{+}	0.0	3.4^{+}	0.5
23. taa-niq			-0.3^{+}	0.1	3.6^{+}	0.6
24. i-wiit			0.3^{+}	0.2	1.4^{+}	0.3
25. a-maaq			1.0^{+}	0.4	1.7^{+}	0.1
34. aa-taa			0.4^{+}	0.2	1.3^{+}	0.2
35. aap-paa			1.1^{+}	0.2	1.5^{+}	0.3
36. quur-qaa			1.0^{+}	0.2	1.4^{+}	0.2
39. ki-sii-sa	-1.2^{+}	0.3	-0.7^{+}	0.2	1.5^{+}	0.4
40. suk-kuu-tit	-1.6^{+}	0.3	-0.8^{+}	0.1	1.1^{+}	0.7
41. kil-luu-tit	-1.1^{+}	0.5	-0.7^{+}	0.2	1.6^{+}	0.9
42. a-wa-taaq	0.6^{+}	0.3	1.0^{+}	0.2	1.8^{+}	0.2
50. a-naa-waa	0.4	0.5	-0.1	0.1	1.2^{+}	0.2
57. an-naa-waa	0.1	0.1	0.1	0.1	1.5^{+}	0.2

4.4.3. Slope of short vowels (on a mora basis)

Based on the results in the preceding section 4.3., we tentatively assume that the tonal pattern is regulated on a mora basis. (NB: When a word contains only short vowels, the results obtained on a mora basis are equal to those obtained on a syllable basis. And short vowels of M2 and M1 are always equal to those of S2 and S1.) A nasal mora is not an independent tone base, but it is uncertain as yet whether a vowel before a nasal mora forms a tone base with or without help of the following nasal mora. Therefore, vowels immediately before a nasal mora will not be included here.

As for the results of slope 1 ([Hz/cs]) and slope 2 ([Arctang (Hz/cs)]), see table 4-6. Fig. 4-2 shows the number of significant tokens.

In this section only the significance of the direction of the slope is considered. But in section 5. the degree of the slope and relations between the slope and the mean fundamental frequency and the consonantal environment will be discussed.

The direction of the slope is not clear except for the vowels of M1 and M2. In M1 there is a general tendency for the slope to be positive, i.e. the pitch is rising. When the slope is not significantly positive, the vowel is always preceded by a voiceless consonant. In M2 the slope is generally negative, i.e. the pitch is falling. The slope of M3 shows no clear tendency common to the three persons. Further, in half of all these cases, the slope does not differ significantly from zero. In RP's and AS's tokens, the slope tends to be negative, while in GT's tokens it tends to be positive. In M4 there is no case where either slope 1 or 2 is significantly positive, except in one of GT's tokens (i.e. (42) a̲wataaq), in which the slope is significantly positive. The slopes of M5 and M6 have, as a whole, no significant direction.

Fig. 4-2.

Slopes of short vowels in each mora position.

- Number of significant tokens -

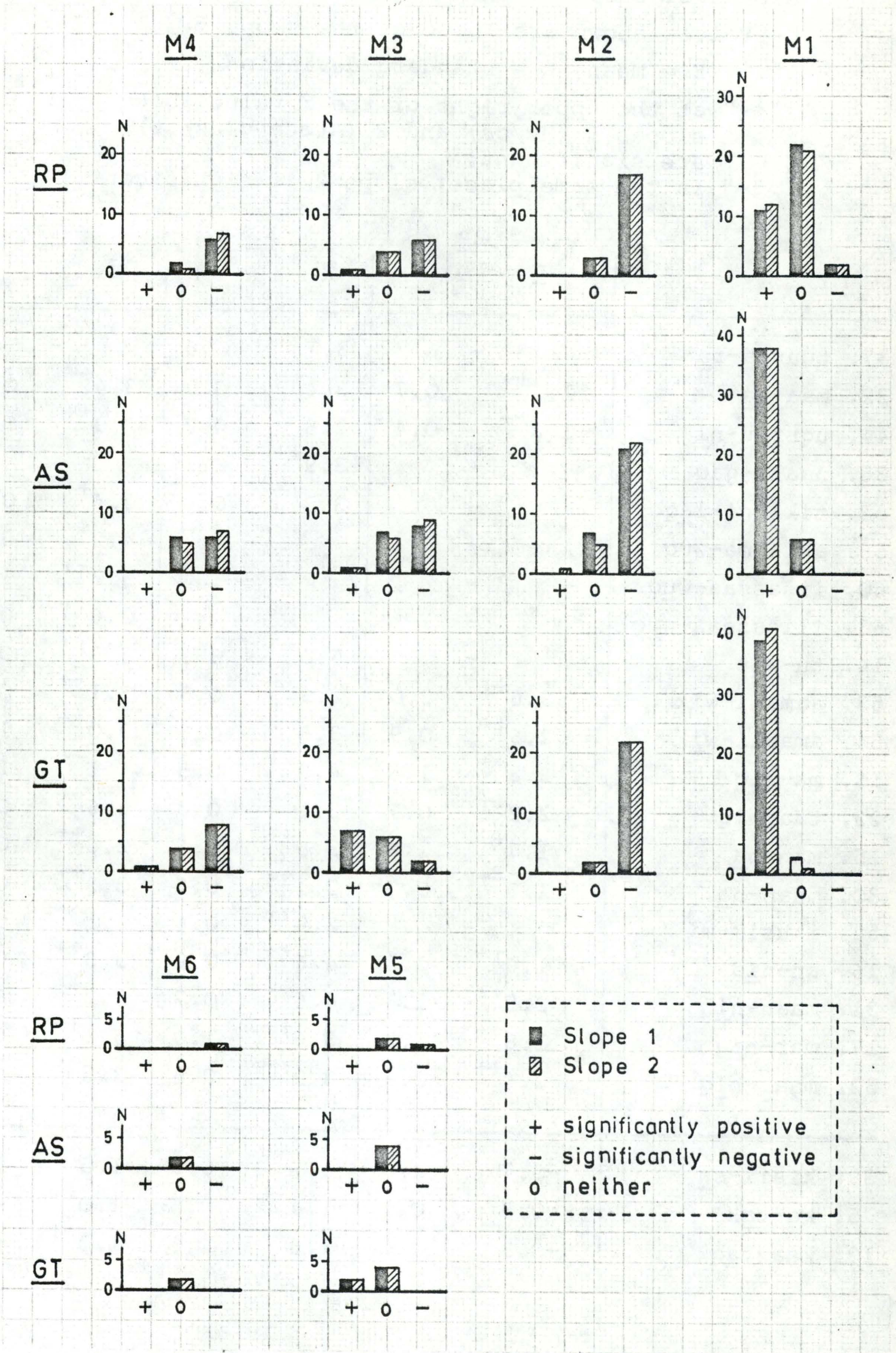


TABLE 4-6

Slopes of short vowels in each mora position.

[Values shown are those of the slope 1.]

($\overline{X}$ = mean, s = standard deviation)

("++" at the upper right of the $\overline{X}$ value: both slopes 1 (Hz/cs) and 2 (Arctang(Hz/cs)) are significant;
"+" in the same place: slope 2 is significant.)

M1	RP $\overline{X}$	RP s	AS $\overline{X}$	AS s	GT $\overline{X}$	GT s
37. tuurpa-ra	-		3.7^{++}	0.6	-	
22. maa-na	1.2^{++}	0.7	3.4^{++}	0.9	3.4^{++}	0.5
44. uqurlu-ni	1.3^{++}	0.4	2.8^{++}	0.7	2.9^{++}	0.6
38. naala-gaq	-		3.9^{++}	0.1	-	
51. tikiŋŋi-laq	-		3.3^{++}	0.8	3.4^{++}	0.4
56. akisima-wuq	-		-		4.2^{++}	0.9
60. iišaššaa-wuq	-		-		3.7^{++}	1.2
63. ilišaššaa-wuq	-		-		3.6^{++}	0.8
14. tu-wit	-		5.0^{++}	0.5	3.3^{++}	0.8
53. aamali-wik	3.6^{++}	2.0	3.9^{++}	0.3	3.1^{++}	0.7
54. amaali-wik	2.6^{++}	0.9	3.3^{++}	0.6	3.0^{++}	1.2
13. ma-niq	2.8^{++}	1.5	4.8^{++}	0.5	4.1^{++}	1.0
23. taa-niq	2.2^{++}	0.7	3.8^{++}	0.8	3.6^{++}	0.6
15. man-na	1.7^{+}	1.3	3.5^{++}	0.6	2.6^{++}	0.5
30. maan-na	1.3^{++}	0.7	3.1^{++}	0.6	3.3^{++}	0.4
52. ikusim-mi	-		2.9^{++}	0.6	3.1^{++}	0.5
16. an-niq	1.6^{++}	0.8	4.6^{++}	0.9	3.3^{++}	0.5
31. maan-niq	2.5^{++}	1.1	4.6^{++}	0.7	3.2^{++}	0.8
17. mar-niq	1.8^{++}	1.2	3.6^{++}	1.0	3.1^{++}	0.5
32. maar-niq	2.4^{++}	0.3	3.9^{++}	0.5	3.5^{++}	0.4
39. kisii-sa	-0.7	1.3	1.7^{++}	0.3	1.5^{++}	0.4
5. ka-tak	-0.5	0.8	1.4^{++}	0.5	1.0^{++}	0.5
19. kaa-taq	-0.5	1.0	1.6^{++}	0.5	1.3^{++}	0.6

TABLE 4-6
(continued)

M1	RP $\bar{X}$	RP s	AS $\bar{X}$	AS s	GT $\bar{X}$	GT s
27. naka-taq	-0.5	0.7	1.6^{++}	0.5	1.2^{++}	0.7
28. awa-taq	0.1	0.7	1.0^{++}	0.6	2.3^{++}	0.4
43. naparu-taq	-		0.7^{++}	0.4	-	
8. ta-qaq	1.3	1.4	1.3^{++}	0.6	2.2^{++}	0.6
55. aputitu-qaq	-		1.0^{++}	0.5	2.5^{++}	0.6
6. ma-sak	0.0	0.6	1.9^{++}	0.6	1.3^{++}	0.6
7. na-saq	-0.8	0.8	1.7^{++}	0.6	1.8^{++}	0.3
18. tii-tit	0.1	0.5	0.3	0.5	1.1^{+}	0.8
26. iki-tit	-0.6	0.7	0.9^{++}	0.6	1.3^{++}	0.4
40. killuu-tit	-0.6	1.0	-0.8	1.2	1.1^{++}	0.7
41. sukkuu-tit	0.0	0.0	0.1	0.6	1.6^{++}	0.9
20. puuk-ka	-0.8	0.9	1.5^{++}	0.6	1.3^{++}	0.5
29. tuwik-ka	-		0.9^{++}	0.5	2.7^{++}	0.3
21. quur-qa	-0.7	1.3	1.9^{++}	0.4	1.6^{++}	0.5
3. il-lu	-0.6	1.3	1.3^{++}	0.6	1.0^{++}	0.7
9. mat-tak	0.0	0.6	1.2^{++}	0.2	0.2	0.3
12. qaq-qaq	-0.2	1.0	1.8^{++}	0.2	1.6^{++}	0.5
10. aš-šak	0.2	1.0	1.1^{++}	0.3	0.4^{+}	0.3
11. aš-šaq	0.1	0.4	1.5^{++}	0.5	0.6^{++}	0.4
64. ilutusimap-put	-0.7	0.6	0.7^{++}	0.5	1.9^{++}	1.2
45. iqqanar-puq	-1.5^{++}	0.8	1.0	0.9	1.7^{++}	0.8
46. iqqarsar-puq	-1.3^{++}	0.2	-0.2	0.6	-	
62. naluumasur-tuq	-		0.0	0.0	-	
4. il-lut	-0.2	1.1	1.9^{++}	0.8	1.1^{++}	0.6

TABLE 4-6
(continued)

M2	RP $\overline{X}$	RP s	AS $\overline{X}$	AS s	GT $\overline{X}$	GT s
10. aš-šak	-2.7^{++}	1.0	-0.9^{++}	0.6	-0.3	0.5
11. aš-šaq	-2.1^{++}	0.8	-1.4^{++}	0.4	-0.2	0.4
3. il-lu	-2.6^{++}	0.9	-1.4^{+}	1.1	-1.6^{++}	0.5
4. il-lut	-1.1^{++}	0.2	-2.1^{++}	0.7	-2.1^{++}	0.4
38. naa-la-gaq	-		0.0	0.3	-	
56. akisi-ma-wuq	-		-		-0.7^{++}	0.5
13. ma-niq	-0.8^{++}	0.4	-0.7^{++}	0.4	-0.3^{++}	0.2
53. aama-li-wik	0.2	0.4	0.4^{+}	0.3	-0.4^{++}	0.1
54. amaa-li-wik	0.1	0.7	0.2	0.3	-0.3^{++}	0.2
28. a-wa-taq	-0.9^{++}	0.5	-1.1^{++}	0.2	-1.4^{++}	0.5
6. ma-sak	-1.2^{++}	0.8	-0.7^{++}	0.4	-0.8^{++}	0.5
7. na-saq	-1.1^{++}	0.5	-1.7^{++}	0.4	-1.0^{++}	0.4
43. napa-ru-taq	-		-0.5^{++}	0.3	-	
64. ilutusi-map-put	-1.8^{++}	0.4	-2.9^{++}	0.4	-2.9^{++}	0.4
9. mat-tak	-0.9	0.8	-1.5^{++}	0.3	-0.6^{++}	0.4
45. iqqa-nar-puq	-1.5^{++}	0.3	-3.8^{++}	0.8	-2.2^{++}	0.6
29. tu-wik-ka	-		-1.9^{++}	0.4	-1.5^{++}	0.6
51. tikiŋ-ŋi-laq	-		-0.1	0.4	-1.5^{++}	0.2
14. tu-wit	-		-2.4^{++}	0.5	-0.7^{++}	0.4
8. ta-qaq	-2.5^{++}	0.7	-1.3^{++}	0.6	-1.0^{++}	0.6
5. ka-taq	-2.1^{++}	0.5	-1.6^{++}	0.5	-1.1^{++}	0.6
27. na-ka-taq	-1.8^{++}	0.6	-1.6^{++}	0.2	-1.0^{++}	0.5
55. aputi-tu-qaq	-		-2.3^{++}	0.6	-1.6^{++}	0.4
26. i-ki-tit	-1.6^{++}	0.5	-2.2^{++}	0.4	-1.7^{++}	0.6
12. qaq-qaq	-2.1^{++}	0.6	-2.2^{++}	0.4	-1.3^{++}	0.2
62. naluuma-sur-tuq	-		-1.6^{++}	0.3	-	
37. tuur-pa-ra	-		-0.3	0.4	-	
44. uqur-lu-ni	-1.9^{++}	0.6	-0.5	0.5	-0.9^{++}	0.3
46. iqqar-sar-puq	-2.5^{++}	0.8	-2.0^{++}	0.6	-	

TABLE 4-6
(continued)

M3	RP $\bar{X}$	RP s	AS $\bar{X}$	AS s	GT $\bar{X}$	GT s
28. a-wataq	-1.1	1.2	0.6	0.5	0.9^{++}	0.1
25. a-maaq	-0.6^{++}	0.3	0.7	0.6	1.0^{++}	0.4
24. i-wiit	-1.6^{++}	0.3	-		0.3^{++}	0.2
26. i-kitit	-2.2^{++}	0.5	0.7	0.7	-0.4	1.1
53. aa-ma-liwik	-0.8^{++}	0.5	0.0	0.0	0.7	0.6
42. a-wa-taaq	0.9^{++}	0.5	1.6^{++}	0.4	1.0^{++}	0.2
27. na-kataq	0.0	0.7	0.3	0.3	0.6^{++}	0.2
62. naluu-ma-surtuq	-		-1.5^{++}	0.4	-	
49. am-ma-qaaq	-		-0.8^{++}	0.3	1.2^{++}	0.5
43. na-pa-rutaq	-		-1.3^{++}	0.6	-	
29. tu-wikka	-		-1.0^{++}	0.2	-0.3	0.8
56. aki-si-mawuq	-		-		-0.7^{++}	0.2
64. ilutu-si-mapput	-0.7	0.6	-2.1^{++}	1.2	-0.3	0.4
52. i-ku-simmi	-		-1.6^{++}	0.5	0.1	0.4
55. apu-ti-tuqaq	-		-1.4	1.1	-0.9^{++}	0.4
44. u-qur-luni	-1.0^{++}	0.7	-1.0^{++}	0.6	1.1^{++}	0.7
45. iq-qa-narpuq	-0.8	0.9	-0.7^{++}	0.5	0.7	1.0
46. iq-qar-sarpuq	-1.4^{++}	0.7	-1.2^{+}	0.9	-	

TABLE 4-6
(continued)

	RP		AS		GT	
M4	X̄	s	X̄	s	X̄	s
42. a-wataaq	-1.1^{+}	0.8	0.0	0.6	0.6^{++}	0.3
44. u-qurluni	-1.4^{++}	0.6	-0.5	1.2	-0.8	0.7
52. i-kusimmi	-		-0.7	0.8	-1.7^{++}	0.7
45. iq-qanarpuq	-1.9^{++}	0.9	0.1	0.9	0.2	0.8
46. iq-qarsarpuq	-1.6^{++}	0.8	-0.4	0.7	-	
43. na-parutaq	-		-1.3^{++}	0.4	-	
55. a-pu-tituqaq	-		-0.5^{++}	0.3	0.3	0.7
64. ilu-tu-simapput	-1.9^{++}	0.5	-1.7^{++}	1.2	-0.9^{++}	0.2
51. ti-kiŋŋilaq	-		-1.6^{+}	1.2	-1.9^{++}	1.2
39. ki-siisa	-2.0^{++}	0.4	-2.5^{++}	0.9	-1.2^{++}	0.3
56. a-ki-simawuq	-		-		-0.7^{++}	0.3
60. ii-šaš-šaawuq	-		-		-0.5^{++}	0.2
63. ili-šaš-šaawuq	-		-		-0.4	0.3
40. suk-kuutit	-0.3	0.8	-2.2^{++}	0.7	-1.6^{++}	0.3
41. kil-luutit	-1.6^{++}	1.1	-2.0^{++}	0.4	-1.1^{++}	0.5
M5						
54. a-maaliwik	-1.6^{++}	0.6	0.5	0.5	0.0	0.3
50. a-naawaa	-0.7	0.9	0.2	0.3	0.4	0.5
55. a-putituqaq	-		0.3	1.2	-0.2	0.3
56. a-kisimawuq	-		-		-0.1	0.1
64. i-lu-tusimapput	0.1	0.8	0.4	0.3	0.7^{++}	0.3
63. i-li-šaššaawuq	-		-		0.6^{++}	0.4
M6						
64. i-lutusimapput	-1.2^{++}	0.2	0.4	0.5	0.2	0.4
63. i-lišaššaawuq	-		-		0.4	0.4
62. na-luumasurtuq	-		-0.3	0.3	-	

4.4.4. Slope of long vowels (on a mora basis)

As opposed to the slopes of the short vowels, the slopes of long vowels seem to have rather fixed directions, depending on the mora positions of their occurrences in the word. See fig. 4-3. Generally, a long vowel of M2 and M1 (i.e. a long vowel which forms M2 and M1) has a positive slope, that of M3 and M2 a negative slope, that of M4 and M3 a positive slope. But the slopes of long vowels forming M5 and M4, and M6 and M5 are not clear. This is due to the fact that the change of fundamental frequency is smaller in the positions of M6, M5, and M4, and that there are only a few examples in the material of this investigation.

When there is a rise-fall or fall-rise during one mora portion of a long vowel, the slope is not significant, since the slope in this case is multidirectional. In fig. 4-3 and table 4-7 slopes which are not significant include both unidirectional ones (i.e. either a rise or a fall) and multidirectional ones (i.e. a rise-fall or a fall-rise). Notice here that the vowel mora in M3 position has a positive slope when it forms the second half of a long vowel, while it has a negative slope (though often not significantly negative) when it forms the first half of a long vowel. The vowel mora in M2 position has a negative slope when it forms the second half of a long vowel, while its slope is positive (though often not significantly positive) when it forms the first half of a long vowel.

Fig. 4-3.
Slopes of long vowels in the last four mora positions.
- Number of significant tokens -

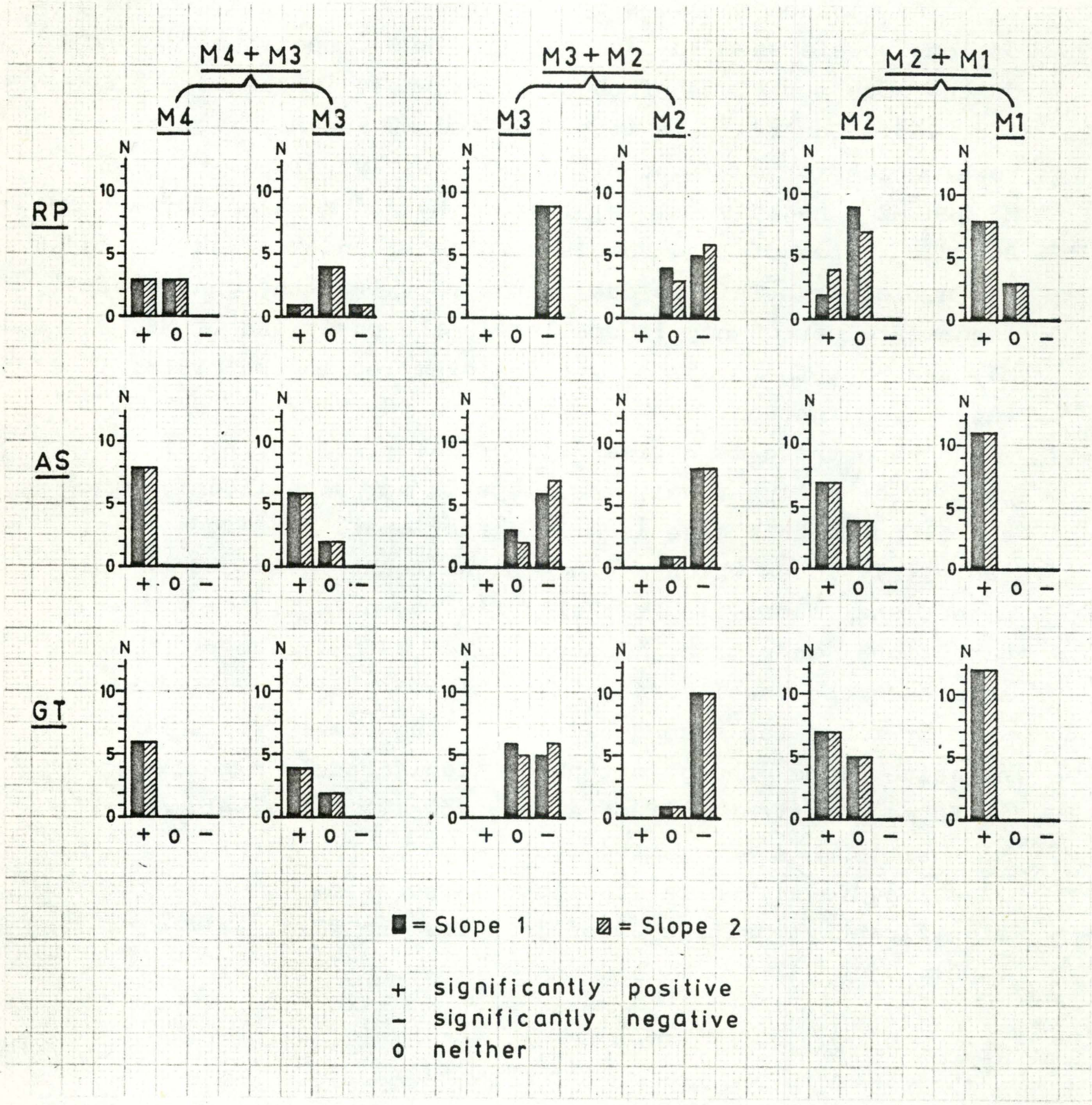


TABLE 4-7

Slopes of long vowels in each mora position

[Values shown are those of the slope 1]

($\bar{X}$ = mean, s = standard deviation)

("...": the pitch of the mora is multidirectional; "++" at the upper right of the $\bar{X}$ value: both slopes 1 and 2 are significant; "+" in the same place: the slope 2 is significant.)

M2+M1		RP		AS		GT	
		M2	M1	M2	M1	M2	M1
50. anaa-waa	$\bar{X}$	...	2.7^{++}	...	3.8^{++}	...	2.4^{++}
	s		0.7		0.7		0.4
57. annaa-waa	$\bar{X}$	1.1^{+}	2.1^{++}	...	3.8^{++}	...	3.0^{++}
	s	0.8	0.5		0.4		0.3
24. i-wiit	$\bar{X}$	1.0^{++}	3.8^{++}	-	-	...	2.8^{++}
	s	0.4	0.4				0.7
25. a-maaq	$\bar{X}$	...	...	2.4^{++}	3.7^{++}	...	3.4^{++}
	s			0.6	0.5		0.2
33. am-maaq	$\bar{X}$	0.7^{++}	2.3^{++}	1.8^{++}	3.6^{++}	0.6^{++}	3.2^{++}
	s	0.5	0.3	0.5	0.4	0.4	0.5
34. aa-taa	$\bar{X}$	...	2.8^{++}	1.3^{++}	3.8^{++}	0.5^{++}	2.6^{++}
	s		0.8	0.4	0.4	0.2	0.3
1. taaq	$\bar{X}$	...	...	0.9^{++}	3.1^{++}	0.9^{++}	3.2^{++}
	s			0.5	0.5	0.4	0.2
42. awa-taaq	$\bar{X}$	...	1.9^{++}	...	3.7^{++}	...	3.5^{++}
	s		0.4		0.4		0.5
49. amma-qaaq	$\bar{X}$	-	-	1.3^{++}	3.6^{++}	1.1^{++}	3.3^{++}
	s			0.5	0.5	0.2	0.9
2. puut	$\bar{X}$	1.0^{+}	3.3^{++}	1.0	4.0^{++}	1.8^{++}	2.5^{++}
	s	0.8	1.1	0.8	1.1	0.5	0.3
35. aap-paa	$\bar{X}$	0.0	2.3^{++}	1.7^{++}	3.9^{++}	0.8^{++}	3.1^{++}
	s	1.1	0.5	0.6	0.3	0.3	0.7
36. quur-qaa	$\bar{X}$	...	...	1.6^{++}	3.7^{++}	0.8^{++}	2.8^{++}
	s			0.4	0.5	0.2	0.5

TABLE 4-7
(continued)

M3+M2		RP M3	RP M2	AS M3	AS M2	GT M3	GT M2
22. maa-na	$\bar{X}$	-1.1^{++}	-0.4	...	-1.3^{++}	-0.0	-0.6^{++}
	s	0.2	0.3		0.3	0.4	0.1
23. taa-niq	$\bar{X}$	-1.0^{++}	-0.4	-1.3^{++}	-0.8^{++}	-0.4^{+}	-0.5^{++}
	s	0.2	0.3	0.2	0.3	0.3	0.2
19. kaa-taq	$\bar{X}$	-1.3^{++}	-0.8^{+}	...	...	-0.2	-1.2^{++}
	s	0.5	0.6			0.3	0.5
18. tii-tit	$\bar{X}$	-0.9^{++}	-1.2^{++}	-1.5^{++}	-2.5^{++}	-0.3	-1.2^{++}
	s	0.4	0.3	0.6	0.5	0.3	0.1
39. ki-sii-sa	$\bar{X}$	-1.2^{++}	-0.8^{++}	-2.2^{++}	-2.5^{++}	-0.7	-1.4^{++}
	s	0.3	0.4	0.6	0.4	0.5	0.3
20. puuk-ka	$\bar{X}$	-1.2^{++}	-1.3^{++}	-1.3^{++}	-3.4^{++}	-0.1	-1.8^{++}
	s	0.5	0.7	0.7	0.7	0.2	0.2
21. quur-qa	$\bar{X}$	-0.8^{++}	-1.0^{++}	-1.2^{+}	-3.4^{++}	-0.5^{++}	-1.3^{++}
	s	0.3	0.4	0.9	0.6	0.3	0.3
60. iišaš-šaa-wuq	$\bar{X}$	-	-	-	-	-1.6^{++}	-0.5^{++}
	s					0.6	0.2
63. ilišaš-šaa-wuq	$\bar{X}$	-	-	-	-	-1.8^{++}	-0.7
	s					0.2	0.4
40. suk-kuu-tit	$\bar{X}$	-1.0^{++}	-0.6	-2.1^{++}	-1.5^{++}	-1.1^{++}	-1.6^{++}
	s	0.4	0.5	0.7	0.2	0.4	0.2
41. kil-luu-tit	$\bar{X}$	-1.5^{++}	-0.7^{++}	-2.3^{++}	-1.8^{++}	-1.1^{++}	-1.4^{++}
	s	0.3	0.2	0.4	0.3	0.4	0.4
30. maan-na	$\bar{X}$	-0.3	-0.6	0.8^{++}	-1.8^{++}	-0.2	-0.8
	s	0.4	0.5	0.4	0.6	0.2	0.8
31. maan-niq	$\bar{X}$	-0.7^{++}	-0.9^{++}	0.8^{++}	-1.4^{++}	0.4^{++}	-0.9^{++}
	s	0.2	0.3	0.5	0.2	0.2	0.3
32. maar-niq	$\bar{X}$	-0.3	-0.8^{++}	-0.3	-1.4^{++}	-0.1	-0.5^{++}
	s	0.3	0.5	0.5	0.1	0.4	0.1

TABLE 4-7
(continued)

		RP		AS		GT	
M4+M3		M4	M3	M4	M3	M4	M3
34. aa-taa	$\bar{X}$	...	...	1.6^{++}	1.7^{++}	1.6^{++}	0.8^{++}
	s			0.6	0.6	0.1	0.3
35. aap-paa	$\bar{X}$	0.9^{++}	0.8	1.8^{++}	2.2^{++}	1.9^{++}	2.2^{++}
	s	0.5	0.8	0.4	0.6	0.3	0.3
54. a-maa-liwik	$\bar{X}$	0.5^{++}	-0.1	1.2^{++}	0.8^{++}	1.2^{++}	0.5^{++}
	s	0.3	0.4	0.5	0.2	0.2	0.3
38. naa-lagaq	$\bar{X}$	-	-	1.2^{++}	0.7^{++}	-	-
	s			0.2	0.3		
50. a-naa-waa	$\bar{X}$	1.2^{++}	...	1.4^{++}	...	1.0^{++}	...
	s	0.6		0.4		0.2	
57. an-naa-waa	$\bar{X}$	0.4	-0.6^{++}	1.1^{++}	...	0.7^{++}	...
	s	0.4	0.2	0.3		0.3	
37. tuur-para	$\bar{X}$	-	-	0.8^{++}	1.0^{++}	-	-
	s			0.4	0.7		
36. quur-qaa	$\bar{X}$	0.6	0.9^{++}	1.7^{++}	1.0^{++}	1.6^{++}	2.0^{++}
	s	0.8	0.5	0.7	0.7	0.6	0.4
M5+M4		M5	M4	M5	M4	M5	M4
53. aa-maliwik	$\bar{X}$	-0.8	0.3	1.2^{++}	0.9	0.4	0.4^{++}
	s	0.7	0.7	0.7	0.8	0.4	0.2
62. na-luu-masurtuq	$\bar{X}$	-	-	1.4^{++}	1.3^{++}	-	-
	s			0.3	0.7		
M6+M5						M6	M5
60. ii-šaššaa-wuq	$\bar{X}$	-	-	-	-	0.3^{++}	0.3^{++}
	s					0.2	0.2

5. Discussion

In this section the question of syllable or mora basis will be taken up first. Then the nasal mora, the slope, and the difference of mean fundamental frequency will be considered, in that order.

5.1. Syllable or mora basis

The phrase-final intonation is not regulated on a syllable basis. As for differences of mean fundamental frequency the assumption often works when the last syllable contains no long vowel, but once it has a long one, the situation is chaotic. That is, when the last syllable has a long vowel, the penultimate syllable (i.e. S2) of a word often has the highest value of mean fundamental frequency; thus the high-low-high relation is exactly the opposite of the relation observed in words containing only short vowels. This is due to the fact that the syllable which contains the penultimate vowel mora (i.e. M2) tends to take the lowest value of mean fundamental frequency. A typical example is a three-syllable word, (50) a-naa-waa. For all three persons the penultimate syllable has the highest mean fundamental frequency in the word.

The slope does not support the syllable basis theory, either. When both S1 and S2 have long vowels, both of them have a positive slope, and the vowel of S2 often ends with a pitch higher than that of S1. Examples are (34) aa-taa, (35) aap-paa, and (36) quur-qaa. Further, if the slope of the vowel is fixed according to the syllable position, there can be no explanation of why S2 of the token (21) quurqa has a negative slope, while that of (36) quurqaa has a positive slope.

Consequently, the syllable is no useful candidate as tone base. So the following discussion will be concentrated on the mora basis.

5.2. The role of the nasal mora

No completely clear picture is seen concerning the role of the nasal mora in forming the phrase-final intonation pattern. This fact may be partly due to the arbitrary assignment of two thirds of a long nasal as a mora nasal. But, disregarding this source of uncertainty about the interpretation, acoustic results favour the assumption that a vowel mora and the following nasal mora form a single tone base.

When the nasal mora is totally disregarded as part of the tone base, the relation between M3 and M2 becomes inconsistent. In some cases the result is that the difference between M3 and M2 is not significant, and further, there are a few samples (recordings) where M2 is higher than M3. This is observed in the following tokens of all three persons: (30) maanna, (31) maanniq, and (32) maarniq. It also holds true of GT's token (52) ikusimmi. In word tokens containing no nasal mora, M3 is always significantly higher than M2.

If, instead, a nasal mora is counted as an independent tone base, the pattern is likewise obscured because the difference between a vowel occurring as M3 and a nasal occurring as M2 is very small (this seems to be due to the fact that both moras are low pitched). Moreover, the relationship between M4 and M3 in AS's and GT's tokens tends to be different in words with and without a nasal mora: in most tokens without a long nasal M4 is significantly lower than M3, whereas it is significantly higher than M3 in most tokens with a long nasal. (RP has a general tendency to have a higher M4.) These findings speak against the assumption of a separate nasal tone base.

When a vowel mora and the following nasal mora are taken to form a compound tone base, the relations of pitch difference between adjacent vowels are clear, and the results agree with the assumption on the phrase-final pitch pattern. M2 is lower than M3 and M1. (The difference between M3 and M2 is not significant in RP's token (32) maarniq, but this lack of significance seems to be caused by other factors than the existence of the nasal, see 5.3.)

The slope of long vowels forming M3 and M2 (i.e. in (30) maanna, (31) maanniq, and (32) maarniq) is similar to that observed in (22) maana, and (23) taaniq.

When we count /V+N/ as one tone base, the numbers of tokens which agree with the assumption are as follows:

	RP	AS	GT
M2 < M1	8/8	11/11	11/11
M3 > M2	4/5	8/8	8/8
M4 < M3	0/1	3/4	4/4
M5 < M4	0/1	1/1	1/1

5.3. Slope and mean fundamental frequency

Quite a few tokens show slopes which are not found to be significantly different from zero. However, the slopes of the vowels, especially those of the last three vowel moras, are hardly quite arbitrary, although the direction of the slope in a vowel mora consisting of a short vowel may differ from the first part of a long vowel.

The phrase-final intonation should be distinguished from other patterns by a combination of high-low-high tones on the last three vowel moras, if the initial assumption is correct. In order to characterize the phrase-final intonation, the penultimate vowel mora should have a low tone, and the final vowel mora should have a high tone. When the vowel mora starts at a lower tone it will reach a high tone at the end. When a short-vowel mora starts at a high tone, it will not necessarily end with a still higher tone. Actually, it may fall a little, but the amount of fall should not be so great that it may be perceived as another intonation pattern. The penultimate vowel mora should have a low tone. When a short-vowel mora in this position starts with a higher tone, it should have a falling tone and end with a lower tone. When it starts low, it will stay low or rise a little, but the rise must be small in order to avoid confusion with the higher tone of the final or antepenultimate vowel mora.

When the last two vowel moras consist of a long vowel, the tone of the long segment should be rising, otherwise it will be perceived as other patterns. When M3 and M2 are formed by one long vowel, the tone of the long segment should be falling and be followed by a noticeable rise or a high tone of M1, otherwise it will be perceived as other patterns.

In the light of this presupposition, the slope of the vowel in each mora position will be discussed. It should now be possible to explain the interference on the fundamental frequency from various phonetic factors. Such interference may for instance be caused by vowel quality and quantity, consonant quality and quantity, stress, or the degree of assimilation present (especially regressive assimilation) in connection with open and close contact. Some of these influential factors may be of a universal character, and others may be specific to Greenlandic. But notice that if it is true that the direction of the slope is rather fixed, this factor might be able to dominate to such a degree as to minimize the influence from various phonetic factors.

When all the influential factors mentioned above are taken into account, a normalization should be possible. Unfortunately, however, the material of the present investigation is not suitable for such an attempt, since the material is too inhomogeneous. Thus, the following account is more like an informal speculation on the factors which may possibly influence the fundamental frequency of the vowel in a particular position. It should be remembered that the accuracy of measurement may be one of the factors which should be taken into account.

The problems concerning the slope of long vowels are less complicated than those concerning short vowels, so this point will be taken up first. The slope of a long vowel segment in M2+M1 position will be positive. The slope of M2 (i.e. the first half of the vowel) is not significantly different from zero in RP's words, and it is significantly positive in half of AS's and GT's words, while the slope of M1 (i.e. the second half of the vowel) is always significantly positive. This means that the tone stays low during the first half of the vowel and rises steeply during the second half of the vowel. A fall-rise is often observed in the first half.

The slope of a long vowel forming M3 and M2 will be negative. In RP's words the fundamental frequency falls during the first half and falls again or stays low during the second half of the vowel. In AS's and GT's words the fundamental frequency falls all the way through the long vowel, or it stays rather high during the first half and falls steeply during the second half of the vowel. A delayed fall is often observed in GT's words.

Thus, the slope of M3 of RP's words is significantly negative, while that of AS's and GT's words is significantly negative only in half of the total set of tokens, and the slope of M2 is sometimes not significantly negative in RP's words, whereas it is significantly negative in the great majority of AS's and GT's words.

The slope of a long vowel forming M4 and M3 will be positive. In RP's words a rise is certainly observed, but the rise is small. In half of all cases it is true of both M4 and M3 that they do not show significantly positive slopes. In AS's and GT's tokens, on the other hand, the tone movement during the long segment is either a simple rise or rise-and-stay-high. Thus, the slope of M4 is significantly positive, but in some cases that of M3 is not significantly different from zero.

The slope of a long vowel forming M5 and M4 is not well-defined, since the change is small, and the number of examples is limited. As for the word token (53) <u>aa</u>-maliwik, which is the only word common to all three persons, there is either a rise or a rise-fall in RP's token, a rise or a rise-and-stay-high in AS's token, and a small rise in GT's token. As for AS's (62) na-l<u>uu</u>-masurtuq, the slope of the long vowel is positive. As for the slope of a long vowel forming M6+M5, the only word we have is GT's (60) <u>ii</u>-šaššaa-wuq. Here the slope of the long vowel is positive, though the rise is small.

Rischel[1] assumes that a long (homosyllabic) vowel tends to have less pitch inflection than a sequence of two heterosyllabic vowels. Although the auditive pitch impression is not equal to acoustic fact, this tendency toward even pitch may be observed in the present material when a long vowel occurs in the positions of M4 to M6. Long vowels of either M4 and M3 or M3 and M2 sometimes have a very small change,

---

1) Personal communication (also mentioned in Rischel (1971)).

but they are seldom quite level. Since they take part in the contour signalling phrase-final intonation, they will exhibit some inflection.

There may be some influence of surrounding consonants on the slope. If there is any, it may be of the same kind as observed in the short vowels. This will be taken up below.

The general outline of the slope of short vowels is as follows. Fig. 5-1,a,b,c shows the relation between duration and amount of fundamental frequency change of vowels in the last three moras classified according to whether the preceding consonant[1] is voiced, voiceless, or zero. Fig. 5-2,a,b shows the relation between the slope and the mean fundamental frequency of the vowel.

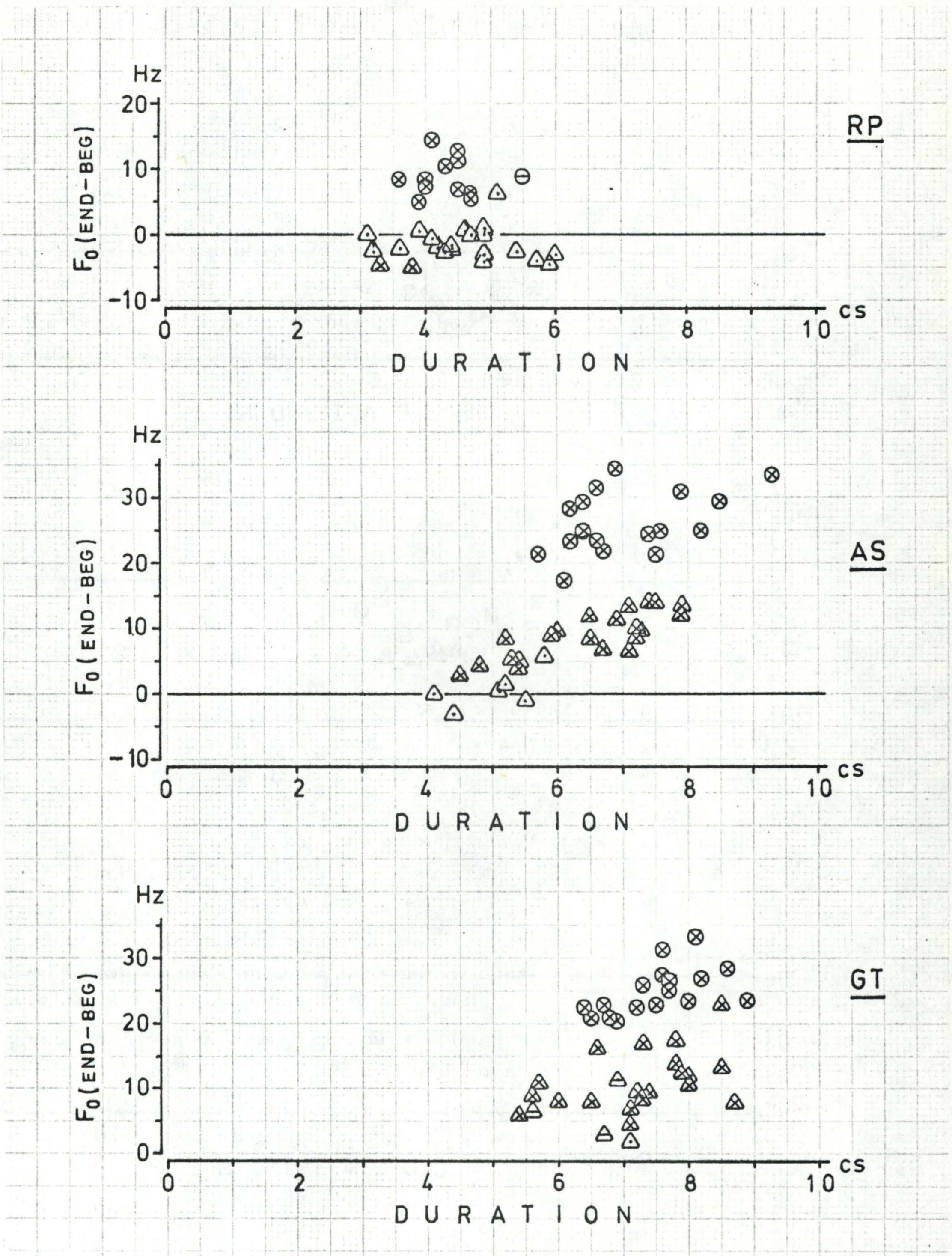
The slope of M1 is positive in the majority of cases. When a vowel of M1 has a positive slope there is a certain correlation between duration and amount of fundamental frequency rise. When the preceding consonant is voiced, both slopes 1 and 2 are significantly positive, no matter whether the vowel occurs in an open or closed syllable. When the preceding consonant is voiceless, the situation is different. Of 23 tokens of RP's, none has a slope which is significantly positive. Two tokens (i.e. (45) iqqanar-pu̲q, and (46) iqqarsar-pu̲q) have significantly negative slopes. In AS's and GT's tokens there are some whose slopes are not significantly different from zero. As a whole, the amount of change

---

1) "Preceding" and "following" consonants are not always the same as "initial" and "final" consonants of the syllable. When a preceding consonant is short, it is the initial consonant of the same syllable, but when it is long, i.e. geminate, it is final in the preceding syllable as well as initial in the syllable in question. A "following" long consonant is final in the syllable and at the same time initial in the following syllable, and a "following" short consonant is initial in the following syllable.

Fig. 5-1a.

Intervals of duration and frequency between the beginning and the end of a short vowel in M1 position.



preceding cons.

Ç_	Ç̥_	Ø_	
⊗	⟁	⊠	: slopes 1 and 2 are significantly different from zero,
⊖	△	⊟	: slope 2 is significantly different from zero,
⊙	△	⊡	: neither of the above two cases.

Fig. 5-1b.

Intervals of duration and frequency between the beginning and the end of a short vowel in M2 position.
(As for symbols, see Fig. 5-1a.)

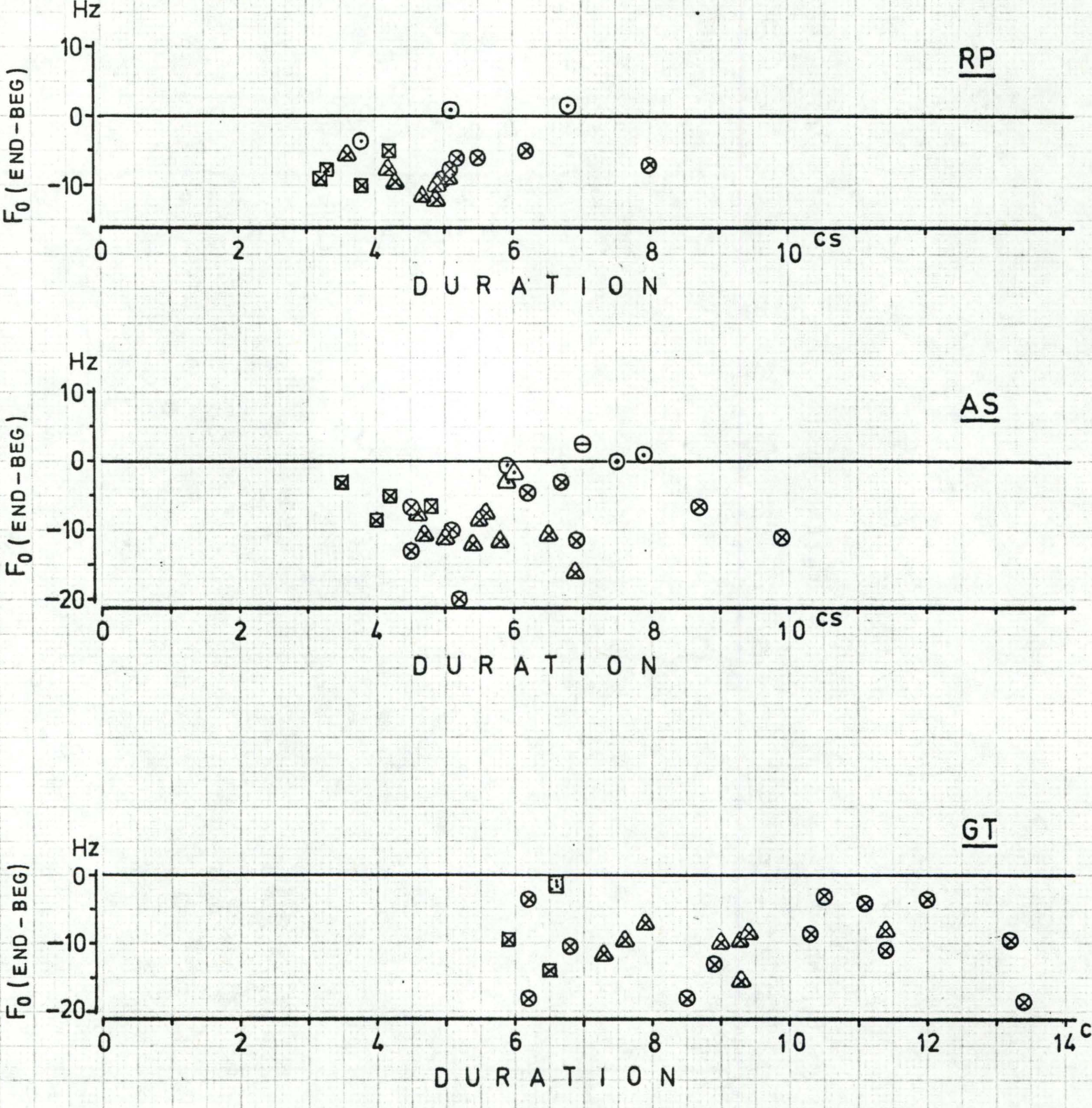


Fig. 5-1c.

Intervals of duration and frequency between the beginning and the end of a short vowel in M3 position.

(As for symbols, see Fig. 5-1a.)

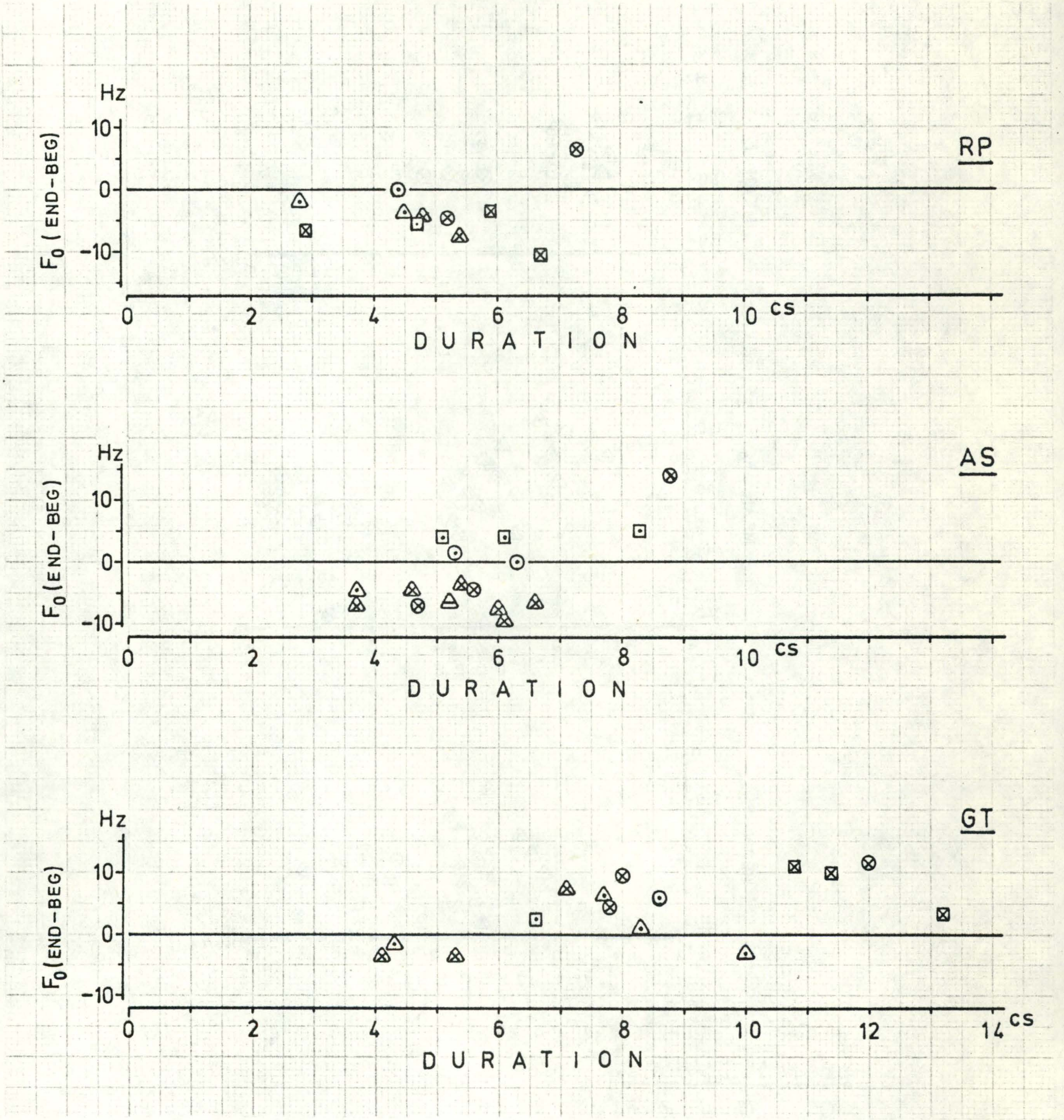
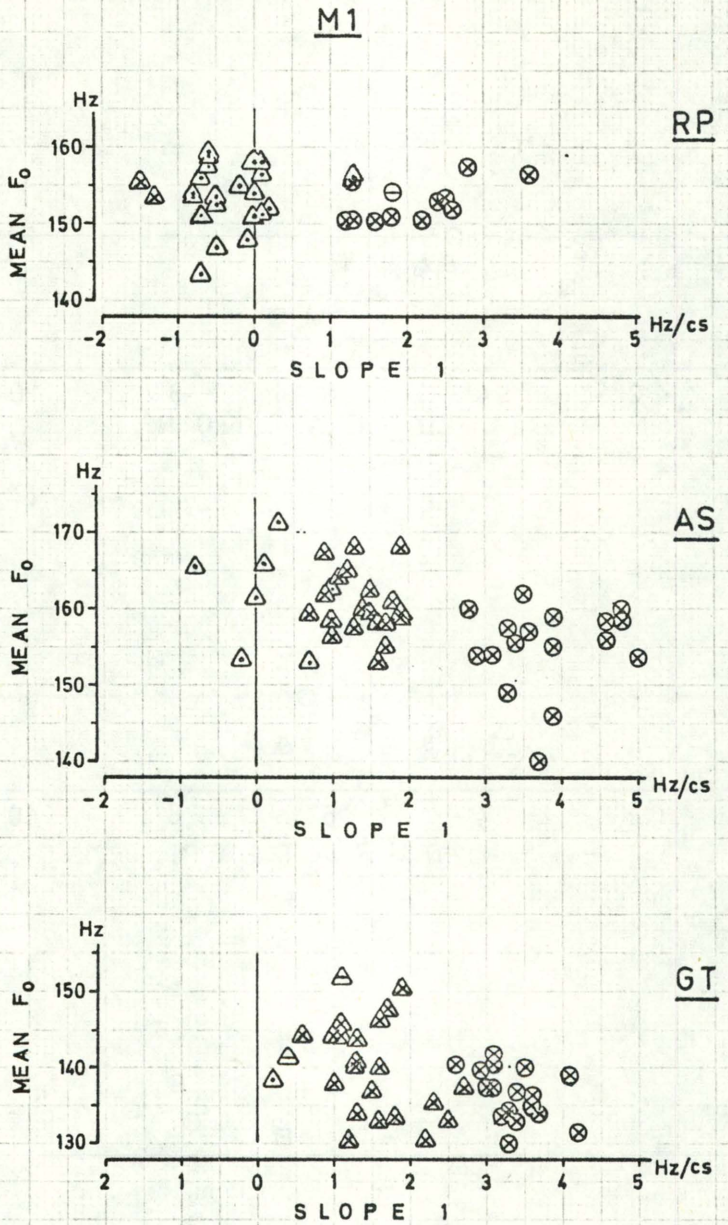


Fig. 5-2a.

Slopes and mean fundamental frequencies of short vowels in M1 position.

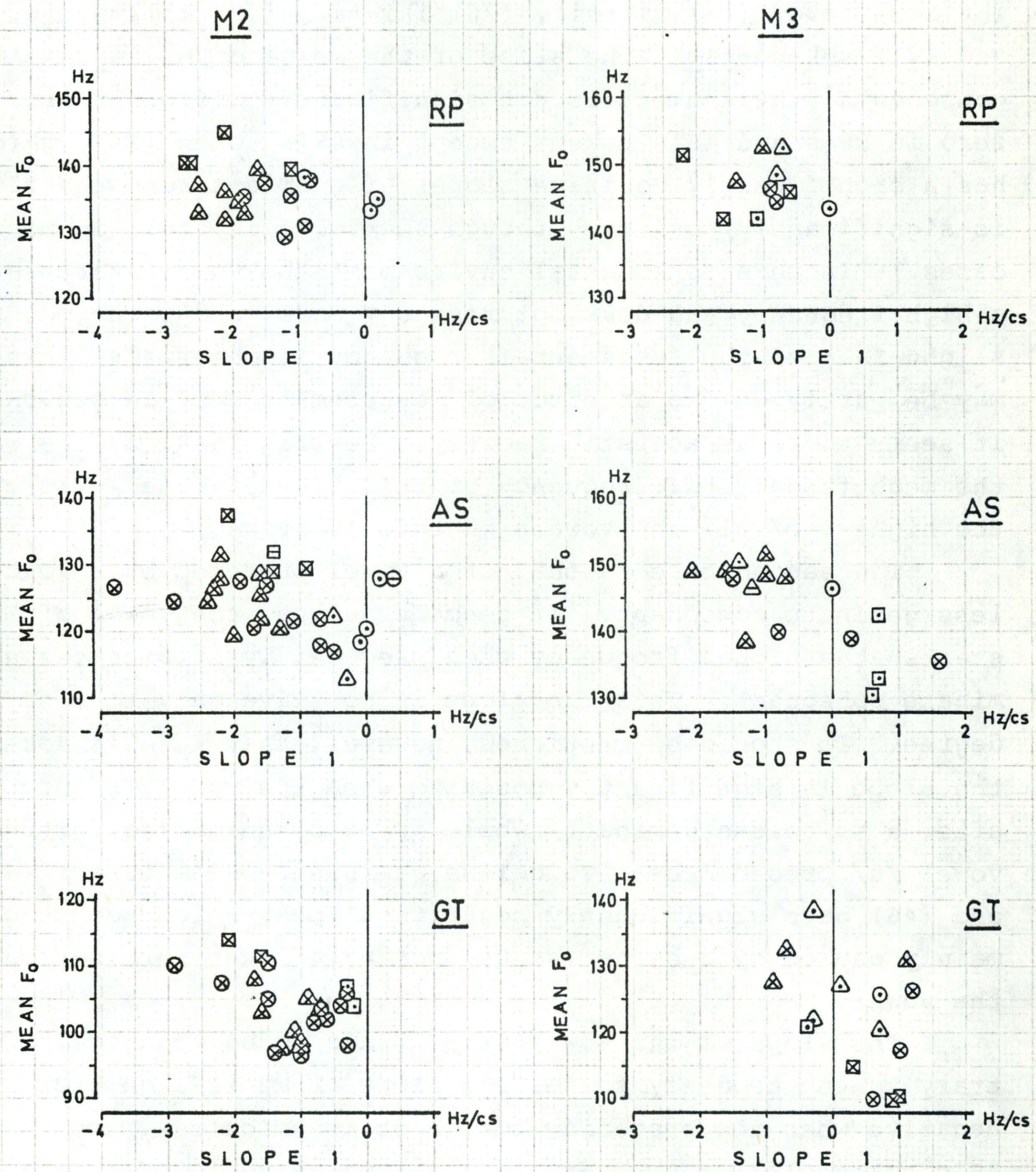


preceding cons.

Ç_	Ç̥_	Ø_	
⊗	▲ (hatched)	⊠	: slopes 1 and 2 are significantly different from zero,
⊖	△ (barred)	⊟	: slope 2 is significantly different from zero,
⊙	△ (dotted)	⊡	: neither of the above two cases.

Fig. 5-2b.

Slopes and mean fundamental frequencies of short vowels in M2 position (left column) and in M3 position (right column).
(As for symbols, see 5-2a.)



is smaller in vowels preceded by voiceless consonants than in vowels preceded by voiced ones. But this does not necessarily mean that the mean fundamental frequency is lower in the former case than in the latter; it is rather the other way round.

There are four tokens in which M1 consists of -tit ([t^sit]), i.e. (18) tii-tit, (26) iki-tit, (40) killuu-tit, and (41) sukkuu-tit. The slope of the vowel of M1 in this consonantal environment is not significantly different from zero in RP's and AS's tokens except in AS's token (26), which has a significantly positive slope. (In GT's tokens the slope is significantly positive, though the value is small in most cases.) In this consonantal environment the vowel starts at a high frequency and does not rise any more, and sometimes a slight fall of the fundamental frequency is observed. This may be partly due to an error of measurement, but in general it seems to be an acoustic reality. See fig. 5-3.(By the way, the mean fundamental frequency of this "level"-tone /-tit/ is the highest of all the vowels in this position.)

The second case is that of a vowel preceded by a voiceless geminate consonant. It seems true that the vowel of M1 starts at a higher frequency when preceded by a voiceless geminate consonant. Thus the slope is positive only to a small degree. It should be mentioned, however, that in AS's tokens the slope is significantly positive when the vowel /a/ occurs after a voiceless geminate, while it is often not so for the vowel /u/ preceded by a voiceless geminate. RP's tokens (45) and (46) have significantly negative slopes, which seems to be mainly caused by a higher start, not by a lower termination of the vowel.

The slope of the vowel of M2 tends to be negative. Contrary to the tendency for M1 the slope of M2 is significantly negative when the preceding consonant is voiceless, except in AS's tokens (37) tuur-pa-ra and (44) uqur-lu-ni. When the

Fig. 5-3a.

Narrow-band spectrograms with a frequency expanded scale, and mingograms (traces from top to bottom: two intensity curves, fundamental frequency curve, and duplex oscillogram).

Words: /taqaq/ and /ikitit/. Speaker: RP.

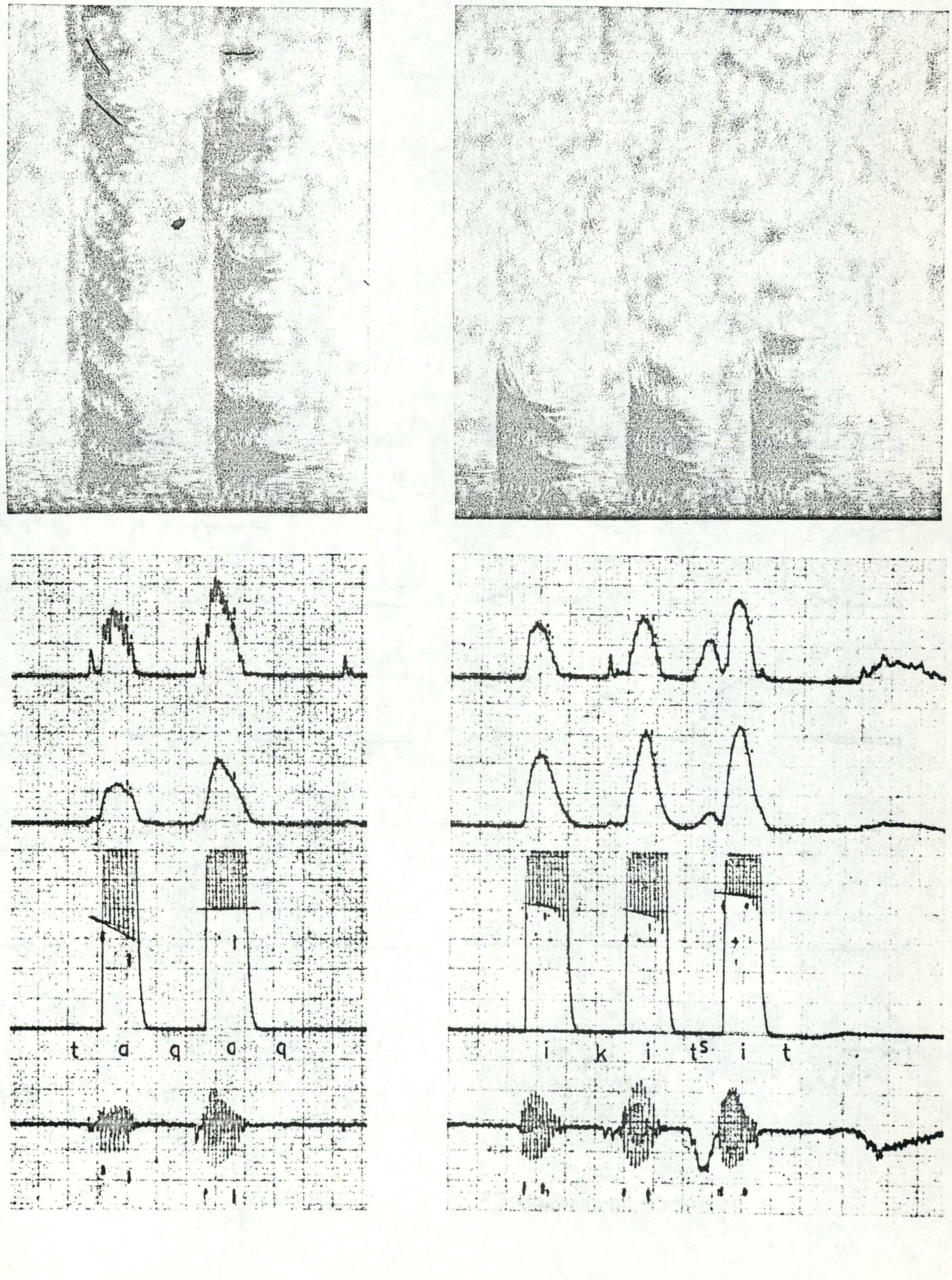
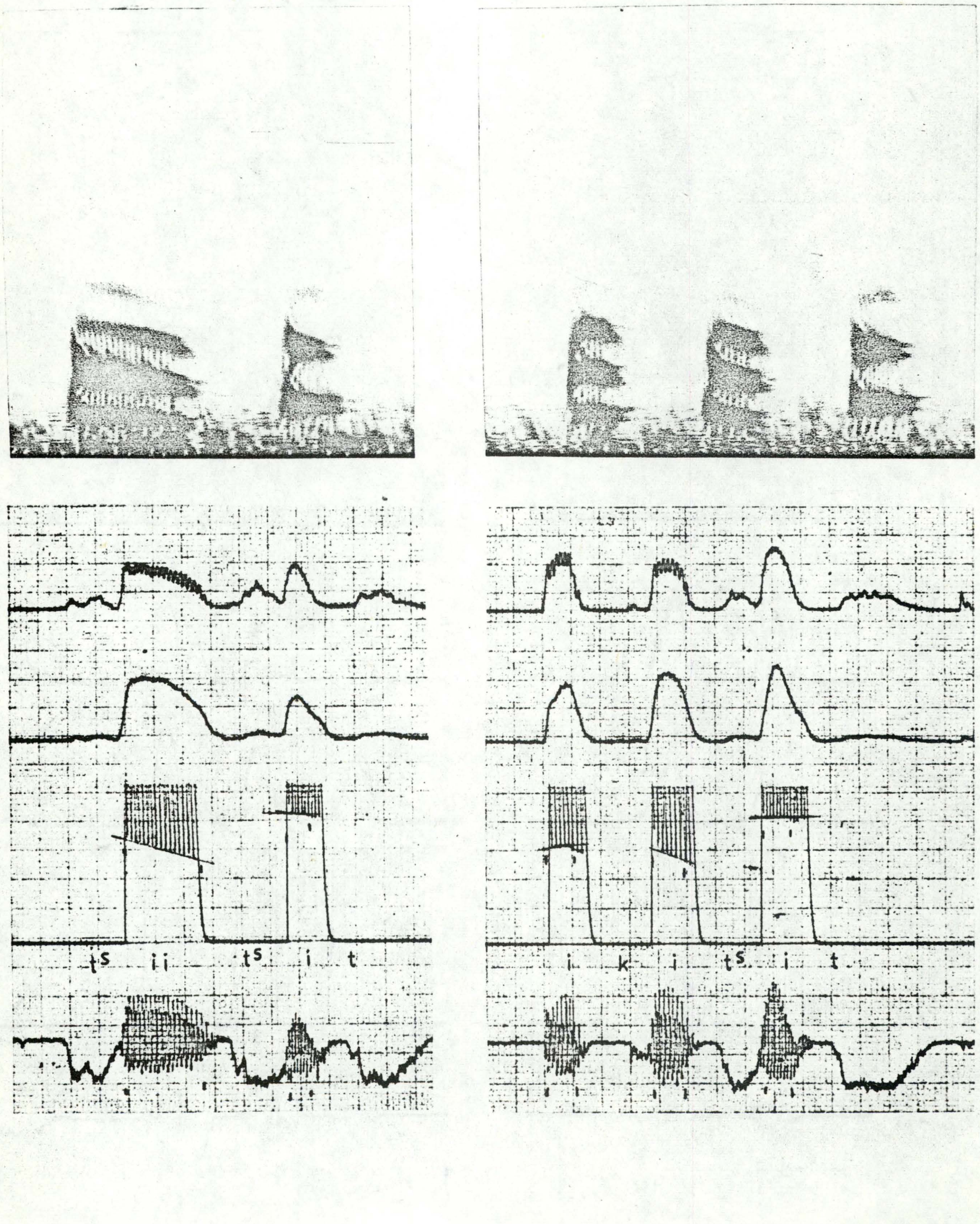


Fig. 5-3b.

Narrow-band spectrograms with a frequency expanded scale, and mingograms (as for traces, see Fig. 5-3a).

Words: /tiitit/ and /ikitit/. Speaker: AS.



slope is not significantly different from zero the preceding consonant is voiced, or the vowel is preceded by no (initial) consonant and followed by a voiceless geminate (no initial consonant in the case of two-syllable, two-(vowel) mora words). The tokens (53) aama-l<u>i</u>-wik, and (54) amaa-l<u>i</u>-wik spoken by RP and AS are the only ones in which the slope is slightly positive.

When all the tokens of three persons are pooled, there are 10 cases where the vowel of M2 is preceded by a short voiced consonant and followed by a short voiceless consonant, and there are 12 cases where the vowel of M2 occurs between a short voiced consonant and a long voiceless consonant. In all, except one, of these 22 cases, the vowel of M2 has a significantly negative slope. This may mean that the following consonant is not so influential as the preceding one upon the direction of the slope.

The results for the slope of M2 show that when the vowel of M2 starts higher, as it does after a voiceless consonant, the fundamental frequency of the vowel should be lowered in order to achieve a relatively low mean value. As the results show, the mean fundamental frequencies of vowels are not significantly different, whether the slopes are significant or not.

Whereas M1 and M2 show a rather clear tendency, the remaining moras do not. The first question is whether there is a fixed direction of the slope for the vowels from M3 to M6. <u>M3</u> is the least promising one; there are too many tokens in which the slope is not significantly different from zero. But a weak tendency for the slope to be negative is seen in RP's and AS's tokens, and a weak positive tendency is observed in GT's tokens. The token (42) a-w<u>a</u>-taaq is the only case in which the slope of M3 is significantly positive for all the three persons. This is the only one among RP's tokens in which the slope is positive, and it is the only one among AS's tokens in which the slope is significantly positive. No consistent influence from surrounding consonants can be observed.

The slope of M4 will be negative, since there is no case where the slope is significantly positive, except for GT's (42) a-wataaq, though the change here is small. The slopes of M5 and of M6 are not clear, since the number of tokens is small and since the change becomes smaller. The influence from surrounding consonants and the influence of stress will be greater here.

Mohr (1971) says that in vowels spoken with an "intended" level pitch,[1] the following results were obtained:

> "After voiceless consonants, onset frequencies are slightly higher than offset frequencies. After voiced consonants, onset frequencies are more than 15 cps below offset frequencies, for CVC syllables. For VCV sequences, onset frequencies are only about 5 cps lower than offset frequencies in the case of obstruents, and less than 2 cps lower in the case of liquids." (p. 7o-71)

---

1) B. Mohr made a study of the (acoustic) intrinsic variations in F_o and the durations of vowels (/a i u/) and consonants (/p t k b d g f v l r ? h/). Three speakers of different languages (a Chinese, a Russian, and a German) with "sufficient experience in English" spoke "language-independent" nonsense CVC syllables and VCV sequences. "Speakers were instructed to read the lists at normal speed and uniform pitch. All syllables were to have level and equal stress." (p. 67)

_If_ these vowels with "intended" level pitches are perceived as level pitched,[1] our results will be as follows. A slight rise in fundamental frequency of the vowel of M1 after a voiceless consonant may correspond to an auditively more rising pitch, and a slightly falling fundamental frequency may correspond to an auditively level pitch. A great rise in the vowel of M1 after a voiced consonant will be not so great auditively. In the case of the vowel of M2, a falling fundamental frequency after a voiceless consonant may not be auditively very noticeable, and the (almost) level frequency of a vowel after a voiced consonant may be perceived as falling pitch. But we have not carried out any perception test. The only thing we know is that the words of the present investigation sounded normal to the ears of the native speakers.

---

1) It is not explicitly mentioned in Mohr whether the vowels were perceived to be level pitched. But it is true that the author carried out a perception test using the same stimuli. The procedure of the test is not completely clear to me, but anyway, the author's concern is the distinction of "voiced" and "voiceless" categories of consonants and/or the "naturalness" of the pitch contour of the vowel after a certain consonant, but not the pitch perception of the vowel in a certain consonantal environment. We can interpret the results as follows. A vowel preceded by a voiceless consonant sounded "natural" if the vowel had a (slight) falling contour, while a vowel preceded by a voiced consonant sounded natural if the vowel had a rising contour. (cf. Mohr, p. 85 ff.)

As for the pitch change of the nasal consonant, the author says: "Kyungnyun Kim (1968) found that the average pitch level of nasals is lower than that of vowels, but that the pitch contour during the nasal is rising" (p. 8o).

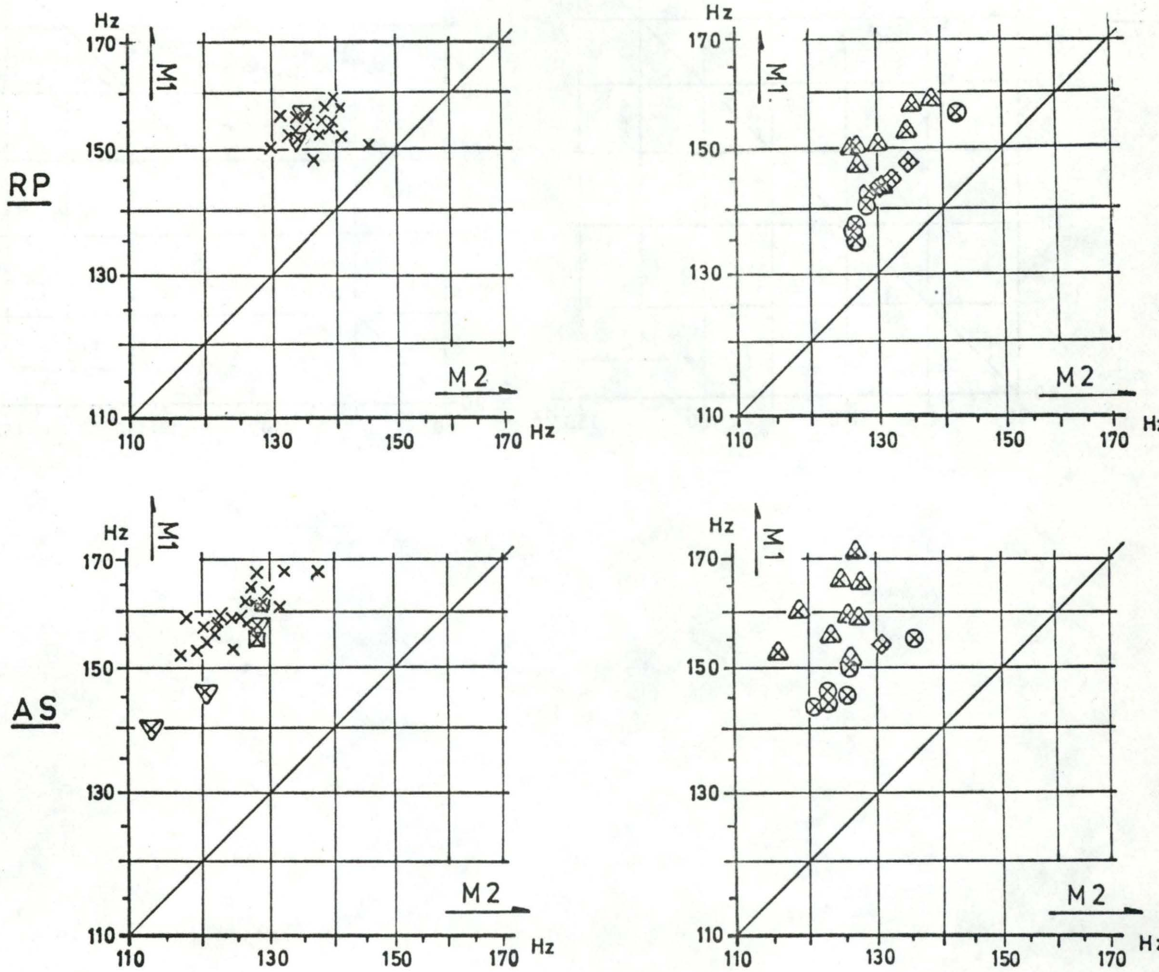
In their research on Swedish Accent 2, Erikson-Alstermark (1972) report that if an accented vowel (with a falling fundamental frequency) varies in duration, the value of the (negative) slope is constant, and the vowel truncation occurs at the final part of the vowel. Thus, the shorter the vowel is, the higher the F_o value at the end of the vowel. This is called the "truncation" type, as against the "rate adjustment" type, by which is meant that the value of frequency interval is constant without regard to the vowel duration.

In our case, the vowel of M1 has generally a positive slope. If the direction of change is relevant, the vowel of M1 will be more like the "truncation" type than the "rate adjustment" type. The truncation occurs at the beginning of the vowel. Thus, if there is any difference of mean fundamental frequencies between the vowels after voiceless and voiced consonants, the values after a voiceless consonant are a little higher than those after a voiced consonant. But generally there is no significant difference. This means that the longer duration of the vowel after the voiced consonant is necessary to get a sufficiently high mean fundamental frequency, no matter whether the longer duration of the vowel is physiologically conditioned or not.

Now we proceed to the question of the mean fundamental frequency. From the above discussion, it will be fairly clear that the fundamental frequency of the vowel may vary according to the particular consonantal environment. In addition to this factor, the difference of mean fundamental frequency values between adjacent vowel moras may vary according to the distribution of short and long vowels in a word. Fig. 5-4,a,b,c shows the mean fundamental frequencies of adjacent vowel moras. Both absolute values and differences between adjacent vowel moras are often great when both of the moras are short vowels. This is especially true of short vowels in M2-M1 and M3-M2 positions.

Fig. 5-4a.

Values of mean fundamental frequencies of two adjacent vowel moras. [The value of the vowel mora of M2 is on the x-axis, and that of M1 is on the y-axis.]



Vowel distribution in words (- = syllable boundary) (long vowels in [ ]) M4 M3 M2 M1	difference between moras: significant	difference between moras: not significant
Type 1: (-) V - V - V - V	×	•
" 2: (-[V) V] - V - V - V	⊠	⊡
" 3: (-) [V V] - V - V	▽ (with ×)	▽ (with •)
" 4: (-) V - [V V] - V	△ (with ×)	△ (with •)
" 5: (-) V - V - [V V]	◇ (with ×)	◇ (with •)
" 6: (-) [V V] - [V V]	⊗	⊙

Fig. 5-4a.
(continued)

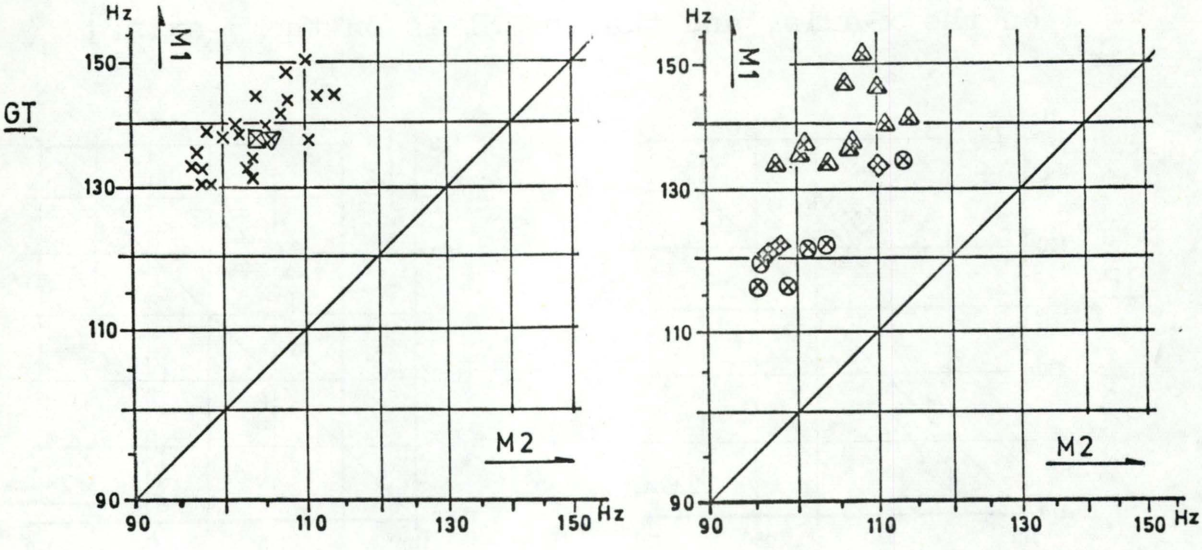


Fig. 5-4b.

Values of mean fundamental frequencies of two adjacent vowel moras. [The value of the vowel mora of M2 is on the x-axis, and that of M3 is on the y-axis.] (As for symbols, see Fig. 5-4a.)

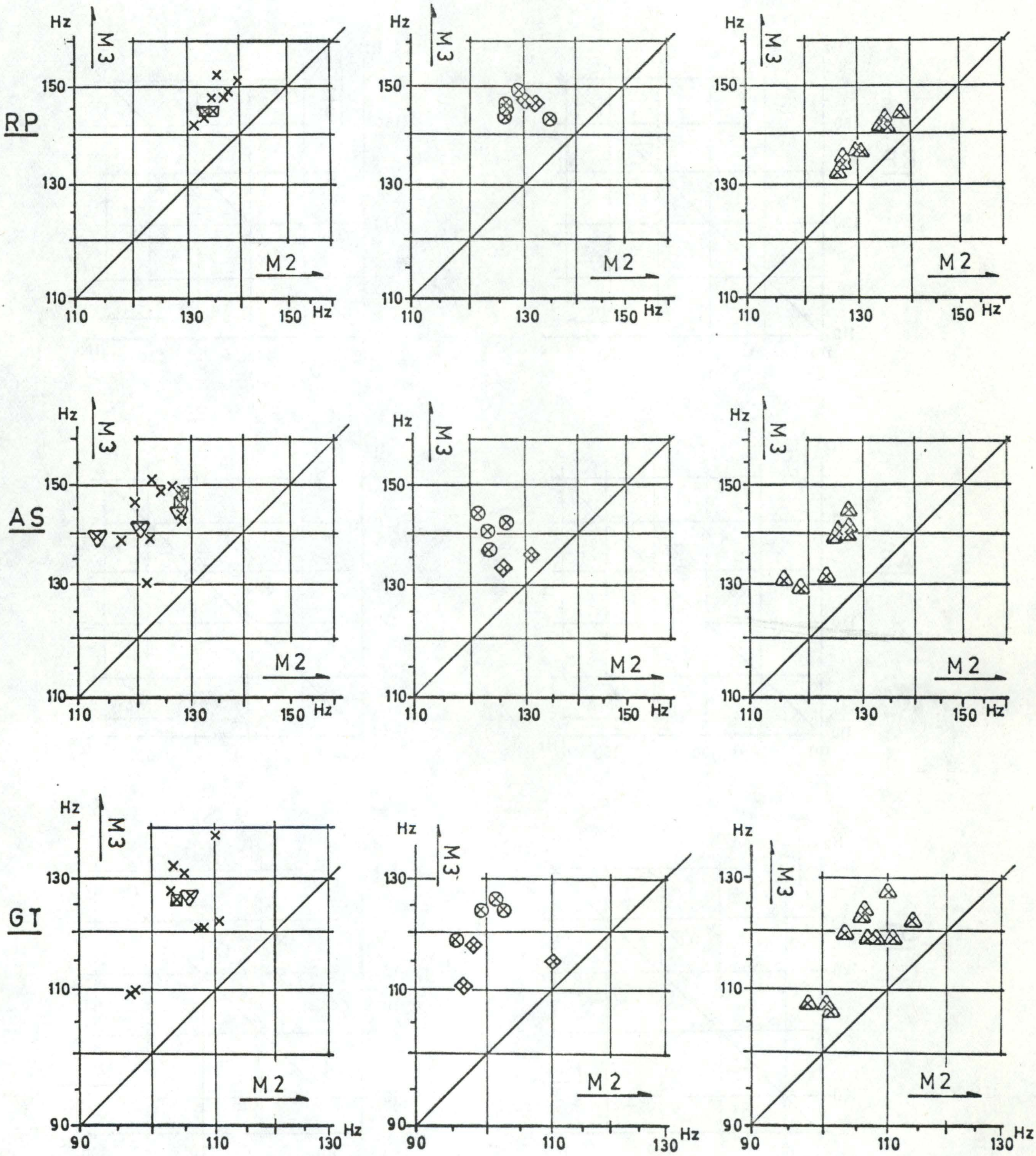
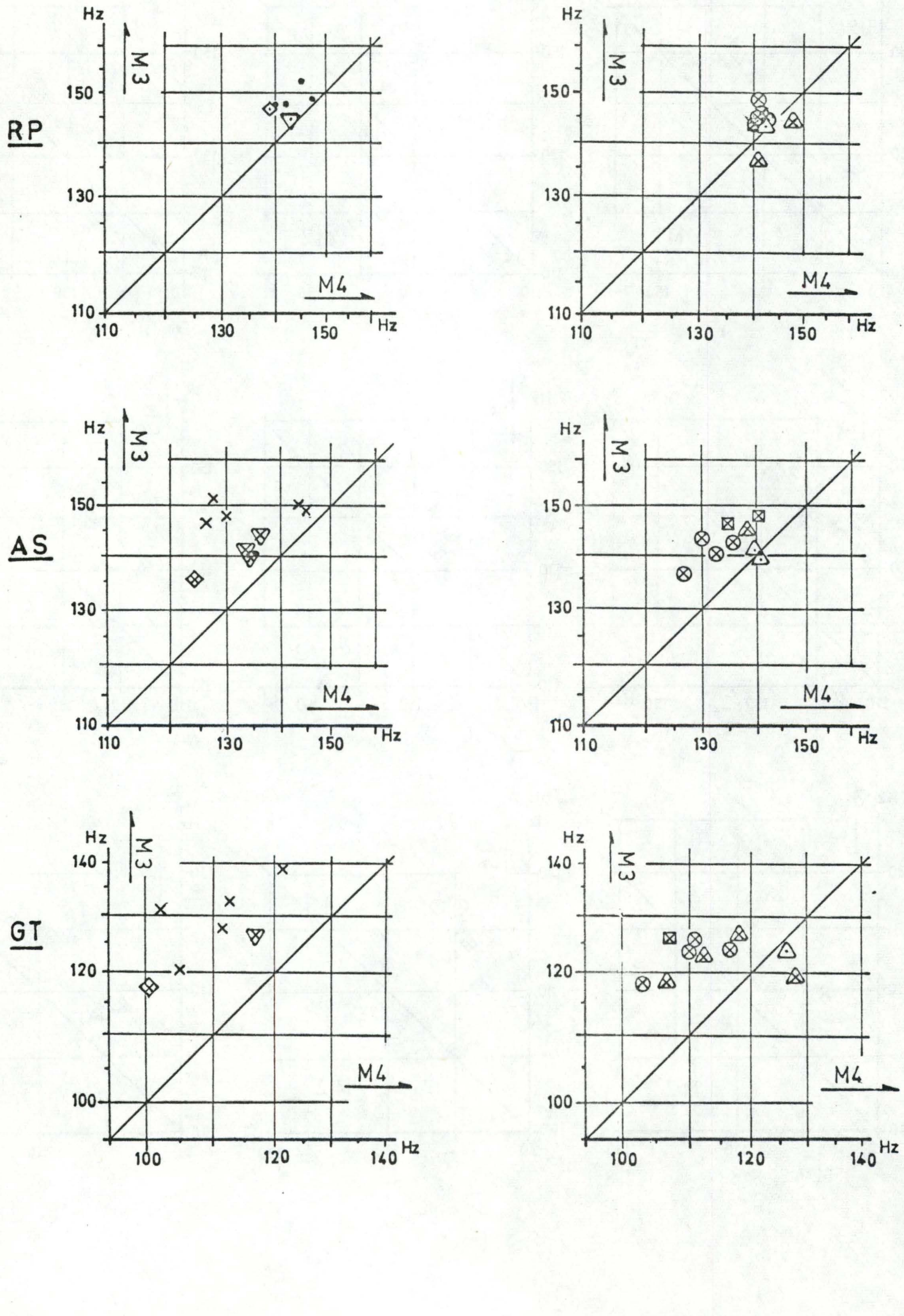


Fig. 5-4c.

Values of mean fundamental frequencies of two adjacent vowel moras. [The value of the vowel mora of M4 is on the x-axis, and that of M3 is on the y-axis.] (As for symbols, see Fig. 5-4a.)



Both absolute values of M1 and the difference between M2 and M1 are great when M2 is the second half of a long vowel and the vowel of M1 is short. On the other hand, both absolute values and the difference are smaller in M2-M1 position when the two vowel moras consist of a long vowel.

Another fact emerges from fig. 5-4,a,b,c. Though the absolute values of the mean fundamental frequency vary among the words depending upon the vowel quality, the quality and duration of surrounding consonants, the differences of mean fundamental frequency values are rather constant among words which have the same distribution of short and long vowels.

The relation between M4 and M3 is that M4 is generally lower than M3, but this is not true of three-syllable, four-mora words of the type CV(C)-CVV-CV(C), if the consonants are voiceless. In RP's tokens (39) ki-si-i-sa and (40) suk-ku-u-tit M4 is significantly higher than M3; in the token (41) kil-lu-u-tit of the same person and in AS's tokens (40) and (41), the difference between M4 and M3 is not significant. In the same tokens spoken by GT, M4 is significantly lower than M3. It is different with his two six-mora words, i.e. (60) ii-šaš-ša-a-wuq, and (63) ili-šaš-ša-a-wuq (these tokens occur only in the material spoken by GT): M4 is significantly higher in the former, and it is either higher or lower in the latter. The vowel qualities of M4 and M3 are almost the same here. The choice of dividing a long vowel into two moras of equal length may not have been appropriate. It is very regrettable that the factor of duration was not taken up in this paper, and is completely left to be analysed.

In words of AS and GT, except in the tokens above mentioned, M4 is always lower than M3. On the other hand, in RP's tokens there are only 5 out of 14 tokens in which M4 is significantly lower than M3. Thus, it may be one of the characteristics of his words that there is no significant difference between M4 and M3. In his words the initial mora of the word is often comparatively high-pitched. This fact may have some connection with stress.

As for the differences between M5 and M4, and M6 and M5, a given mora will be lower than the following one, though the difference is small.

Almost all the way through we have been discussing the high-low relations between tone bases or moras, but we have only briefly touched upon the absolute values of the differences. Since the measurements ought to be normalized with respect to various factors the values shown in the tables in section 4. do not necessarily correspond to the auditive distances between tone bases. As mentioned before, the values of the differences between M2 and M1, and between M3 and M2 are usually more than 20 Hz and 10 Hz, respectively. The values of the differences between tone bases from M3 to M6 are smaller than those found between M2 and M1 and between M3 and M2. But all values vary according to the consonantal environment and the distribution of short and long vowels in the words.

Summing up the above discussion, the slope is not as important as the mean fundamental frequency. A large change in a short vowel will be required only if the fundamental frequency starts higher or lower than the target mean.[1]

---

1) If our speculation on possible consonant influence upon adjacent vowels is right, the particular example of phrase-final intonation in Collis (1970) will get, as he describes, a rise-fall-rise pattern.
Collis (1971, p. 142) shows pitch shapes of vowels in different segmental environments. It seems that in this latter work Collis is inclined to say that the direction of pitch change depends upon the segmental environment. However, the pitch shapes shown in the work do not agree with the tendencies of frequency change observed in the material of the present investigation. This discrepancy cannot be due to the fact that in the present investigation only two points (i.e. beginning and end) are taken into account to measure the values of fundamental frequency.

In the cases of long vowels, the directions of the slopes are rather fixed according to the mora positions. M2 as part of a long vowel, for example, has a positive slope when it is the first half of a long vowel, and it has a negative slope when it is the second half of a long vowel. M3 as part of a long vowel has a positive slope when it is formed by the second half of a long vowel, while it has a negative slope when it is the first half of a long vowel.

Fig. 5-5 shows schematic contours of various types of words which are found in the material of the present investigation. These contours were drawn on the basis of the results shown in the preceding sections. Since the target fundamental frequency values of the vowel moras are supposed to be fixed according to the mora positions, the contours are supposed to be roughly the same within the words of one contour type without regard to the length of the words.

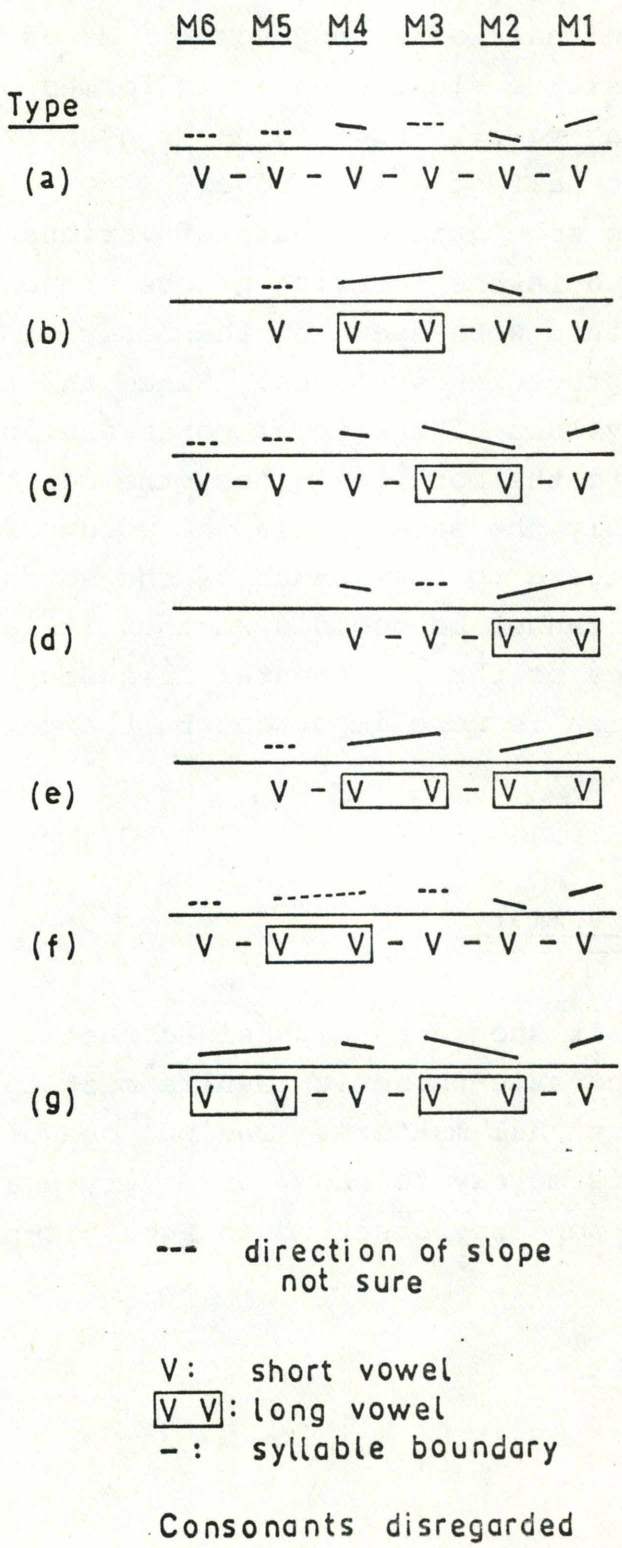
At present it cannot be decided whether it is the mean fundamental frequency or the fundamental frequency at the end of the vowel which is more important to the perception of intonation patterns.

6. Final remarks - summary

First of all, it should be emphasized that I have no illusion that the acoustic-phonetic results must decide on phonological or functional matters. The purpose of the present investigation is merely to show, as a very preliminary research, a certain tendency observed in acoustic-phonetic results.

Fig. 5-5.

Schematic pitch contours of vowel moras in each mora position.



Though there are many unsolved and unclear points in the results of this investigation, the argument that tone is regulated not on a syllable basis but with reference to vowel moras - is supported on various grounds.

The phrase-final intonation is signalled by a "high-low-high" pattern on the last three vowel moras, since a comparison of the mean fundamental frequencies shows that the second vowel mora from the end has a lower frequency than the first and the third in all the tokens of three speakers. (If we say that it is not the mean fundamental frequency but the fundamental frequency at the end of the vowel mora which signals the phrase-final intonation, we can at least say that M2 is lower than M1.) Thus, the instrumental results here agree with what Rischel's phonological analysis shows. The relations between the vowel moras from M3 to M6 are not so clear, which may be partly due to the fact that they do not contribute to characterize the phrase-final intonation, and partly due to the fact that the influences of surrounding consonants and stress (if any) will be comparatively great here, since the fixed movements are smaller. But it is quite probable that in the positions from M3 to M6, the preceding vowel mora is lower than the following one.

With regard to the consonants, a voiceless consonant mora cannot be a tone base, though it may have a certain influence upon the pitch of adjacent vowels. A nasal mora cannot be a tone base, either. But if the nasal mora is completely disregarded, the tonal pattern of words containing a long nasal does not fit well with the patterns of other words which do not contain a long nasal. The first (= syllable-final) portion of a long nasal seems to form one compound tone base together with a preceding vowel mora.

References

Collis, Dermot R.F. 1970: "Etudes philologiques et linguistiques des langages esquimaux", Internord N° 11, p. 263-282

Collis, Dermot R.F. 1971: Pour une sémiologie de l'esquimau (Paris)

Erikson, Y. and M. Alstermark 1972: "Fundamental frequency correlates of the grave word accent in Swedish: The effect of vowel duration", STL-QPSR 2-3, p. 53-60

Fant, Gunnar 1958: Modern instruments and methods for acoustic studies of speech (Oslo)

Flanagan, J.L. and M.G. Saslow 1958: "Pitch discrimination for synthetic vowels", JASA 30, p. 435-442

Hadding-Koch, Kerstin 1961: Acoustico-phonetic studies in the intonation of southern Swedish (Lund)

Holtved, Erik 1964: Kleinschmidts Briefe an Theodor Bourquin = Meddelelser om Grønland 140, 3

Lehiste, I. and G.E. Peterson 1961: "Some basic considerations in the analysis of intonation", JASA 33, p. 419-425

Lehiste, Ilse 1970: Suprasegmentals (Cambridge, Mass.)

Mase, H. and J. Rischel 1971: "A study of consonant quantity in West Greenlandic", ARIPUC 5, p. 175-247

Mohr, B. 1971: "Intrinsic variations in the speech signal", Phonetica 23, p. 65-93

Petersen, Robert 1970: "On phonological length in the Eastern Eskimo dialects", Folk 11-12, p. 329-344

Rischel, Jørgen 1971: Lecture notes, esp. "Udkast til formuleringer af nogle regler vedrørende grønlandsk intonation" (København, 4 pp)

Rossi, M. 1971: "Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole", Phonetica 23, p. 1-33

Thalbitzer, W.A. 1904: A phonetical study of the Eskimo language = Meddelelser om Grønland 31

Editors' note:

The editing of this paper for ARIPUC was not finished when Mr. Mase left for Japan. With his consent a final revision has been made by the editors.

There are some essential changes in the sections dealing with Rischel's phonological analysis, to the extent that this presentation was based on "personal communication": the terminology in Rischel's analysis has been modified in accordance with the most recent formalization (to appear in Rischel's forthcoming monograph on West Greenlandic phonology). These modifications do not affect the theoretical issues of Mr. Mase's paper.

INFLUENCE FROM DIALECTAL BACKGROUND ON THE SCORE IN HEARING TESTS EMPLOYING SPEECH STIMULI

Carl Ludvigsen

1. Introduction

The purpose of an audiometric test is, of course, to investigate certain aspects of the hearing ability of the person taking the test. In practice, however, the result of an audiometric test will also depend on factors which are not directly related to the auditory mechanism. Thus, when speech segments such as single words or sentences are used as stimuli in the test, the test-score will also be determined by e.g. the subject's vocabulary, intelligence, education, and dialect. In extreme situations these factors may even be dominating.

1.1. Procedure

In the winter of 1971/72 a series of experiments were carried out at The State Hearing Center, Bispebjerg Hospital, Copenhagen. A stimulus material consisting of 100 short interrogatory sentences had been constructed and a TV video-recording of this material read by a female speaker[1)] was made. The stimuli could be presented audio-visually (by means of a TV-monitor and a loudspeaker) or auditorily (by means of the loudspeaker only). The stimuli were presented in one loudspeaker and a noise signal in 4 others placed round the subject. The speech level was kept constant throughout the experiment (approximately 55 dB) and different signal to noise ratios were obtained by attenuating the noise signal.

---

1) The speaker's dialect was close to Standard Copenhagen.

1.2. The subjects

48 subjects participated in the experiment. These were selected from 80 volunteers[2] on the basis of audiological, optometrical, and psychological tests. These tests assured that all the subjects participating in the experiment were of normal hearing, sight, and intelligence. All subjects had lived for a period of at least one year in the Copenhagen area, but a substantial part of the subjects were born and grown up outside this area.

1.3. Variation of test-scores obtained

In the light of the careful pre-testing of the subjects, the test-scores obtained showed a considerable degree of variation: Subjects tested with exactly the same stimuli, order of presentation, signal to noise ratios etc. obtained radically different scores. As we were firmly convinced that this variation was not attributable to organic differences in the hearing mechanisms, other possible sources of variation had to be examined. One of these was the dialectal background of the subjects.

2. Dialectal background of the subjects

2.1. Collection of data

Inquiry forms were sent to each of the 48 subjects, and 24 of these returned the forms duly filled in. The inquiry forms gave information about the place of birth and childhood of the subjects and their parents, and of the places where the subjects had lived for longer periods. On the basis of these

---

2) Mainly students at a teachers' training college or a nursing school.

inquiry forms the 24 subjects were divided into two groups. The first contained subjects born and grown up in Copenhagen and the second, subjects born and grown up outside the Copenhagen area (in cases of doubt, information on the subject's parents was used to determine the grouping.

2.2. Statistical treatment

In the test the subjects were divided in groups each containing 6 subjects, and the members of each group were tested with the same stimuli and the same S/N. The subjects were ranked from 1 to 6 according to their test results: In each group the subject who answered the largest number of questions correctly was given number 1, and the subject who obtained the lowest score was given number 6.

TABLE 1 shows how the rankings were distributed among the two groups. Mann-Whitney's U test shows that at the 99.9% level subjects from GROUP 1 are higher ranked than subjects from GROUP 2.

TABLE 1

The ranks of the 12 subjects born in Copenhagen (i.e. GROUP 1) and the ranks of the 12 subjects born outside Copenhagen (i.e. GROUP 2)

	GROUP 1	GROUP 2
RANK	NUMBER OF SUBJECTS	NUMBER OF SUBJECTS
1	5	0
2	3	2
3	2	2
4	2	2
5	0	4
6	0	2

2.3. Conclusions

The main difference between the groups 1 and 2 seems to be that the persons in group 1 speak the same dialect as the speaker employed in the test (i.e. Standard Copenhagen) and the remaining subjects, presumably, speak dialects differing from this in various degrees.

As persons born in Copenhagen are ranked higher than persons born outside of Copenhagen it seems likely that a person's dialectal background influences his ability to understand speech presented in noise.

The clear tendency shown in table 1 must, however, be viewed in the light of the extremely homogenous group with respect to other factors influencing the test score. In practice a major part of the variation will be caused by hearing disorders and the influence of other factors may be neglible.

However, the subjects employed in this test were familiar with the dialect of the speaker. A stronger influence will inevitably be found among the clientele of a hearing center.

This problem will be given further attention in subsequent tests.

Acknowledgements

This work was supported by William Demant and wife Ida Emilie's Foundation, the Danish Research Council for the Humanities, and the Danish Medical Research Council.

NOTES ON DANISH CONSONANT COMBINATIONS

Hans Basbøll

1. Introduction[1]

The "classic" paper on Danish consonant combinations is Vestergaard 1967, which is in turn based upon methods developed by Henning Spang-Hanssen (1959) and Bengt Sigurd (especially 1965). Of course, Danish consonant combinations have been treated in earlier works on Danish phonetics and phonology as well (e.g. Jespersen 1934 and Martinet 1937), but since this literature has been discussed by Vestergaard I shall only mention it in passing.

The present notes are of a highly tentative nature. Nevertheless, I think it worth while to present them for discussion despite the existence of Vestergaard's valuable paper, since there are some important respects in which Vestergaard's treatment is inadequate. As far as I can see it misses linguistically significant generalizations (there are also a number of minor disagreements between Vestergaard's presentation and mine, but I shall not burden the exposition unnecessarily by stating these):[2]

---

1) I am very indebted to Jørgen Rischel for many improvements of my English style, as well as for stimulating suggestions on several of the topics of the present paper (cf. note 2).

2) Vestergaard's paper was discussed in about two of a series of seminars on "The syllable", conducted by Jørgen Rischel, in the autumn of 1969. A number of concrete criticisms of Vestergaard 1967 were brought up on that occasion, not least by Jørgen Rischel, but also by the participants (including Eli Fischer-Jørgensen (and myself)). Much of the criticism of Vestergaard 1967 which follows is inspired by these discussions.

(i) Vestergaard describes the combinations of Danish consonant phonemes, operating with a rather abstract phoneme inventory. It is one of the main claims of the present notes that several generalizations concerning Danish consonant combinations can only be captured by using an inventory of considerably less abstract consonantal segments (cf. section 2.). Note that it is not the question at issue whether Vestergaard's phoneme analysis is reasonable from one or another point of view, but rather it is an empirical issue on what "level" distributional restrictions are best described. I think that Vestergaard has made a bad choice in this respect, which obscures the regularities to be observed concerning maximal syllabic structure (section 5.).

(ii) Although he makes some interesting remarks on the distinction between accidental and systematic (or structural) non-occurrence, Vestergaard has not in all cases followed Spang-Hanssen's fundamental insight that the material must be considered open in order to allow a structural classification of its elements. To take a particularly revealing example, it can only be because the material is considered to be given once-and-for-all that an initial cluster /mj/ but not /nj/ is registered. The former only occurs in the interjection mjav and the loan word (from Old Icelandic) mjød, the latter in the famous Old Icelandic names Njal and (the god) Njord, which are in fact both pronounced with initial [nj] by the Danes. The two clusters seem to be quite analogous to Danes, and it would have been motivated either to exclude /mj/ by Spang-Hanssen's criterion of generalizability, or to include /nj/. (The latter solution is the more revealing of order restrictions, since /nj-/ is infinitely more possible than /jn-/ which is simply excluded.) The distinction in question leads Vestergaard to the conclusion that it is scarcely distributionally relevant to distinguish a class of "nasals", on which I disagree.

(iii) Vestergaard writes (p. 50 f): "The object of the present paper is the simple word, and all inflected forms are accordingly left out. This restriction is obviously necessary; a theoretically possible formation like aspskt (Asp+sk+t) being clearly in conflict with the phonotactic structure of the language - it would never be pronounced." This is a strange argument, since our task must be (among other things) to give rules accounting for which morpheme combinations are allowed within the same word and which are not; the example aspskt is an example of what is not allowed and not of what is "theoretically possible", and is therefore void of significance in Vestergaard's context. Furthermore, Vestergaard's inclusion of forms like tabt [tɑbd] (past participle of tabe [tæ:bə]) as monomorphemic because of the difference in vowel quantity between the participle and the infinitive (the vowel quality is uniquely dependent on the vowel quantity and the following consonant), is unwarranted. The vowel quantity is governed by rule (see Rischel 1970), and the relationship between the mentioned forms is quite transparent (see further sections 4.1. and 4.3. below).

(iv) Vestergaard has deliberately, as shown in the title of his paper, excluded medial clusters from examination. For the sake of completeness a brief survey of these is included here (section 6.). This is of course no criticism of Vestergaard's paper.

The preceding points of criticism should in no way be taken as indicating that Vestergaard's paper is not a good one. On the contrary, his exposition is both careful and extremely clear, and his treatment has a high degree of observational adequacy.

2. Inventory of consonants

The language described here is meant to be a neutral variety of Standard Danish, much like the language described in Rischel 1969, something in between what was called "Conservative Standard Danish" and "Advanced Standard Copenhagen" in Basbøll 1969.

In the language under consideration, what might be called the "diphthongization" of short vowel plus a following homosyllabic r or ɣ is carried through phonetically (similarly with that of short vowel plus v or j, but this is true in all varieties of Danish). Words like bær, steg, sagn are thus pronounced [bæɒ̯, sdai̯ʔ, sɑu̯ʔn], the latter two rhyming with maj and havn respectively. (This diphthongization is general in contemporary Standard Danish, except in Jutland.) Note that the diphthongization in question has not in all cases led to a change in underlying forms, i.e. to restructuring; e.g. alternations like [gœɒ̯, gœ:rə] 'do(es)', 'to do' point to an underlying /r/ as the source of [ɒ̯], and compounds like sagfører ['sɑu̯ˌfø:ɒ] 'lawyer' or savklinge ['sɑu̯ˌkleŋə] 'saw blade' have phonetically identical first parts which (in the norm in question) in isolation are pronounced [sæ·ʔɣ] (sag 'case') and [sæ·ʔv] (sav 'saw'), respectively.

2.1. "Level" of description

As already mentioned, it is one of the main claims of the present paper that the distributional restrictions of consonants get their most general description when the consonantal segments in question are considered as considerably less abstract entities than in Vestergaard's paper. E.g. forms like bær, steg, sagn above will be considered to contain diphthongs in the present distributional description, although the phonetic diphthongs are derived from underlying /VC/-combinations.

On the other hand, the consonantal segments will not be those in the phonetic surface structure either. The forms used here will be independent of certain late rules, notably of the rule which deletes [ɒ̯] after a low back vowel (whether this is a correct description of the rule is unessential here), e.g. ta'r [tɑ·ʔ] 'take(s)', present of ta' [tæ·ʔ] 'to take' will be considered as ending in a vowel plus [ɒ̯], i.e. as ending in a diphthong (cf. ser [se·ʔɒ̯]). Also the rule which devoices voiced non-vowels after an aspirated consonant seems to have no effect on the principles for consonantal distribution, i.e. the forms used here will have voiced l in plaske, phonetically [pl̥asgə], and sj- will be considered s plus voiced j, although in this combination j is generally unvoiced and (more or less) coalesced with s to [ʃ]. (These facts can of course be expressed by saying that the level of distributional description employed here "comes after" the "diphthongization rule", but "before" the "ɒ̯-deletion rule" and the "unvoicing-after-aspirates rule".)

Finally it should be pointed out that certain optional rules, which are to my knowledge of the "variable" sort discussed in Labov 1970 (the application of such rules seems to be correlated with socio-economic class membership of the speaker as well as with stylistic factors) seem also to have no influence on the distributional principles. This is true of the rule that turns [v] into [u̯] after long vowels and /r/, and the rule that shortens long vowels before voiced non-lateral continuants, especially in stød-syllables. The forms used here are thus [sgæɒ̯ʔv, læ·ʔv, u·ʔð, bo·ʔɒ̯] and not [sgæɒ̯ʔu̯, læ·ʔu̯ or læu̯ʔ, uðʔ, boɒ̯ʔ]. The relation between these rules and the distributional description is different from that concerning the late rules mentioned above, since phonetic forms like those we use here are in fact found in the language under consideration, and therefore they do not constitute any proof that the description is not given on the phonetic surface. A form like [vɑɒ̑], on the other hand (the input to the "ɒ̑-deletion rule") is non-existent phonetically and thus constitutes evidence that the description is not given on the phonetic surface (but it is of course possible that forms like [vɑɒ̑] should not be used in the distributional description).

2.2. Fragment of a distinctive feature matrix of Danish non-vowels

On the level discussed in the preceding section we have the following non-vowels (i.e. consonants and glides, see section 2.2.1. below): b d g p t k s f v ɣ ð m n ŋ l r u̯ ɒ̑ i̯ h. Since h is limited to occur in one position, viz. as the only non-vowel in the initial part of the syllable, and thus never exhibits any order relation to any other consonant, h is not very interesting from our point of view once this fact has been stated. As a consequence, h will not be treated in the present notes, and it is therefore not included in the matrix below (its distinctive feature composition is controversial, but irrelevant for the present matter). In the matrix features referring to place of articulation are not taken into account:

	bdg	ptk	s	f	v	ɣ	ð	mnŋ	lr	u̯ɒ̯i̯
syllabic	-	-	-	-	-	-	-	-	-	+
sonorant	-	-	-	-	-	?	?	+	+	+
voiced	-	-	-	-	+	+	+	+	+	+
aspirated	-	+	+	+	-	-	-	-	-	-
sibilant	-	-	+	-	-	-	-	-	-	-
nasal	-	-	-	-	-	-	-	+	-	-
continuant	-	-	+	+	+	+	+	-	+	+

The meaning of the features "voiced", "sibilant", "nasal" and "continuant" is self-explanatory. The other features will be discussed below (the question marks are explained under the feature "sonorant", section 2.2.2.). Note that the feature matrix is intended to facilitate our discussion of distributional facts; it is not claimed that exactly the features above are those which are relevant for underlying forms or for instructions to the speech apparatus. (The relevance of the features in question will hopefully become clear in the following sections, not least section 5.)

It is clear from what was said above that the registered segments are not all in contrast. E.g.[u̯]in [hɑu̯]and [v] in [læ·ʔv] are underlying the same, namely /v/ (and in a taxonomic description they would be classified as members of the same phoneme /v/ since they are - in the language under consideration - in complementary distribution). But they are phonetically clearly distinct (although related), [v] being a labio-dental consonant, [u̯] a labio-velar glide, and this difference seems to be related to distributional factors (see sections 4. and 5.).

In cases where the syllable initial and syllable final realizations of an underlying segment are phonetically clearly distinct, they have not been classified as one segment (e.g. [t-] and [-d], [d-] and [-ð], [r-] and [-ɒ̯], see the following section).

2.2.1. The feature "syllabic"

A segment constituting the peak of a syllable is called a "vowel". Other segments (which may be called "non-vowels") are either syllabic (called "glides") or non-syllabic (called "consonants"). Two syllabic segments constitute a "diphthong", three a "triphthong", etc. Thus the feature <u>syllabic</u> expresses the common observation that if there is more than one phonetic vowel in a syllable, they are adjacent. This use of the feature "syllabic" has been proposed by Stephen R. Anderson (unpublished notes), and in fact it underlies Hjelmslev's (1951) use of the term "vowel" (although this use is not consistent with his own definitions, cf. e.g. Spang-Hanssen 1959, p. 38).[1]

According to the matrix above there are three glides in Danish: [u̯ ɒ̯ i̯]. These are all found postvocalically,[2] (e.g. in [hɒu̯], [bæɒ̯], [tʌi̯]).[3] [i̯] also occurs prevocalically (e.g. <u>jeg</u> [i̯ai̯]), whereas the prevocalic segments corresponding to (i.e. phonologically identical with) postvocalic [u̯, ɒ̯] are both phonetic consonants: [v, r].

---

1) However, if one prefers to keep the traditional sense of the term "syllabic", some other feature should be used for designating the voiced phonetic vowels (whether they constitute the peak of a syllable or not).

2) The terms "pre-" and "post-vocalically" refer to the position in the <u>syllable</u> (the syllable division being in accordance with Basbøll 1972a).

3) Although the symbols for "extreme" (phonetic) vowels are used here, it is well known that the last segment when artificially isolated is considerably more central.

2.2.2. The feature "sonorant"

Although the Chomsky and Halle-definition of the feature "sonorant" (SPE p. 302) is not wholly satisfactory, there seems to be a high degree of consistency in the use of it (except for some sounds which do not concern us here, notably the h-sounds). The generalization is that vowels, glides, liquids, and nasals are sonorants, whereas fricatives, affricates, and plosives are non-sonorant (i.e. obstruents).

It seems reasonably clear that the labio-dental [v] is an obstruent (in distinction to [u̯] which is a sonorant), although it is generally frictionless. This is supported by the fact that [v] never has stød in contradistinction to [u̯] (cf. [læ·?v, hal?v] versus[grʌu̯?, sɑu̯?n]), and there seem to be good reasons for characterizing the class of segments that are able to receive the stød as "sonorant" (cf. Austin 1971 and Basbøll 1972b).

The case of [ð, ɣ] is less clear (hence the question marks in the matrix), and it is not even clear whether these have the same coefficient for the feature "sonorant". Phonetically they are frictionless, [ð] with a very loose constriction (much looser than that of English [ð]),[ɣ] probably with a little more constriction.[1] The criterion of ability to receive the stød points to the solution that [ð] should be sonorant (e.g. spid! [sbið?]) but [ɣ] non-sonorant (in a form like dag the last segment may receive the stød, but only

---

1) Note that [ɣ] only appears after liquids: [l,ɐ̯] and after long vowels (e.g. elg, dværg, dag [ɛl?ɣ, dvæɐ̯?ɣ, dæ·?ɣ]). After short vowels [ɣ] has vocalized to [i̯] (after front vowels) or [u̯] (after back vowels), see section 2. above.

if it has vocalized). This may not be phonetically nonsensical, and it is assumed here since this makes the distributional facts easier to understand (sections 4. and 5.).

2.2.3. The feature "aspirated"

It is well known that the phonetic distinction between [bdg] and [ptk] is mainly one of aspiration (and in the case of [d] and [t] also of affrication), both series being voiceless. In the matrix we have classified [ptkfs] together as a class of "aspirated" sounds (it is uncertain whether [h] should be included in this class too). There are the following reasons for this analysis (it should be noted that [ptk] occur prevocalically; postvocalically they are only found after a short vowel plus r):

(i) After exactly these five consonants a voiced non-vowel in the same syllable is devoiced, or, stated in another way: some of the aspiration phase occurs contemporarily with the (underlyingly) voiced non-vowel. (However, this effect may be greater after the aspirated plosives than after the fricatives.)

(ii) A postvocalic r followed by one of these five consonants does not constitute a stød-basis in Conservative Standard Danish (cf. Basbøll 1970 p. 19 ff).

(iii) The important glottographic study of some Danish consonants by Frøkjær-Jensen, Ludvigsen and Rischel (1971) shows that the laryngeal gesture of aspirated plosives and voiceless fricatives is suggestive of a common articulatory command for these sounds, as opposed to the much weaker gesture of the unaspirated plosives. (It is probably a consequence of this command that [fs] are often aspirated (Jørgen Rischel, personal communication).)

Of course, not all of the arguments given above indicate that "aspiration" is the relevant feature, but we consider the class ptkfs to be well-founded, and this is the important issue for our purpose.

3. Initial clusters

3.1. Three member clusters

All initial three member clusters in Danish are of the following structure:

$$s \begin{Bmatrix} b \\ d \\ g \end{Bmatrix} \begin{Bmatrix} l \\ r \\ v \\ \text{i̯} \end{Bmatrix} V$$

Note that the columns are both "natural classes" in a phonological sense, since [bdg] are the only unaspirated plosives[1] and [lrvi̯] the only voiced non-nasal non-vowels (or, equivalently, the only voiced continuant non-vowels). (Furthermore, [s] is the only sibilant.)

---

1) Vestergaard (1967 p. 48 f) considers s plus plosive to be phonemically /sp, st, sk/ (there is only one series of plosives after s). He claims that phonetically there is doubt as to the identification of the plosive with either bdg or ptk, and then chooses the latter possibility (in accordance with Uldall 1936), because of what he calls "Hjelmslev's law", i.e. the principle that the existence of a cluster xyz presupposes the existence of xy and yz, but not inversely (cf. Fischer-Jørgensen 1952 p. 35). This principle demands the interpretation /skv/ in skvat, etc. since /kv/ can be found (kvist, etc.), but not /gv/. I find this argument objectionable. There is absolutely no doubt as to the phonetic identification of the plosive after s since the only stable distinction between the two series of plosives is one of aspiration (cf. section 2.2.3.), and aspiration is absent after s. Furthermore, the non-occurrence of /gv/ is certainly an accidental gap and should thus not be used as a structural argument. And a further argument against Vestergaard's position is the fact (noted by himself p. 48) that /sv/ exists, but not /sf/, and this distinction is certainly systematic. (Fischer-Jørgensen points out that /gv/ occurs medially in the foreign word lingvist.)

The formula predicts the following initial three-member clusters in addition to the ones registered in Vestergaard's material (which for practical reasons is reproduced in toto as an appendix (section 8.) to the present paper): sbv, sdl, sdv, sgl.

The question arises whether these clusters are accidental or systematic gaps (cf. Vestergaard 1967 p. 49 f and Fischer-Jørgensen 1952, chapter VI). I think the most important fact for deciding on this matter is the observation that neither bv- nor pv- exist in isolation, and similarly neither dl- nor tl-, whereas both dv- and tv-, and both gl- and kl- occur. I.e., we can formulate the rule that /v/ never occurs after labial plosives, and /l/ never after dental plosives, and hence sbv- and sdl- are excluded, whereas no such rules prohibit the other two combinations, and hence sdv- and sgl- are only accidentally non-occurring (cf. the foreign word sklerose). This is in accordance with my intuitions. The two mentioned non-combination rules may lead to the hypothesis that homorganic non-vowels do not combine initially. This hypothesis will be examined in section 3.2.2. below.

3.2. Two member clusters

3.2.1. Order restrictions

When one tries to set up "order classes", two points of view can be in conflict with each other:

(i) If the material is considered closed, or phonotactically homogeneous, one tries to establish as few order classes as necessary to account for the observed order relations, the philosophy being that the fewer the order classes, the less over-generating the model, other things equal. (This can be done in a rather mechanical manner, although in some cases there will be doubt as to where to place a given consonant.) This is the line taken by Vestergaard.

(ii) If the material is considered open and heterogeneous from a phonotactic point of view, one may be interested in characterizing as many ordering relations as possible, viz. to establish as many order classes as possible. To take an example, Vestergaard's material (from Dansk Retskrivningsordbog) does not give any reason for separating the phonemes /lrjn/ into several order classes. Since the group /nj/ seems to be just as well established as /mj/ (as mentioned in section 1.), /n/ should belong to a more vowel-remote class than /j/. But even apart from the nasal, the phonemes /lrj/ could be split up since /lj/ is certainly more possible than /jl-/ which is clearly excluded. Thus, if one is interested in characterizing as many order relations as possible, preliminary order classes should be split up until it is true for any two members of one order class that both configurations of them are equally impossible (this presupposes that there is at most one acceptable order of any two consonants, which is true for the initial position). That is, although none of the initial combinations of v, l, and j belong to Vestergaard's material, one could nevertheless set up three order classes for those three consonants as follows: v l i̭
since the following groups are in fact found, outside Vestergaard's material (they are indeed quite possible for Danes in distinction to the "reverse" clusters): /vl-/ (Vladimir), /lj-/ (Ljubljana), and /vj-/ (Vietnam).

It will be clear from the above considerations and the introduction that in principle we consider the latter procedure the more linguistically interesting. It will lead to the following order classes initially:

V	IV	III	??	II	I	
s	bdg fptk	v	mn	lr	i̯	VOWEL

Note that the order classes are rather homogeneous from a phonetic point of view. For a phonetic characterization of the order classes, see section 5.

This scheme shows that we have (at most) five order classes initially. The question marks above the nasals signify that on the present basis it is not possible to make a non-arbitrary decision on whether /mn/ belong to II or III (to my intuition, at least, groups like /vn-/ and /ml-/, etc., seem equally impossible), but that they do <u>not</u> constitute a separate order class.

In a more restricted material which does not include groups of a liquid plus [i̯], it is sufficient to operate with four order classes (as does Vestergaard, but he splits up the nasals), and then there is no doubt that the nasals should go together with <u>v</u>:

IV	III	II	I	
s	bdg fptk	v mn	lr i̯	VOWEL

This order scheme generates all Danish initial two- and three-member clusters. Initially there are at most three non-vowels in a row, and when there are three, these are subjected to strong restrictions (section 3.1.). The model is, however, over-generating for two member clusters as well and should thus be supplied with restrictions of other sorts, as will be discussed next.

3.2.2. Other restrictions

(i) In section 3.1. we advanced the hypothesis that homorganic non-vowels do not combine initially. This hypothesis is confirmed for two member clusters as well, with the following restatement of Vestergaard's phonemic classification: The velar stops are neither homorganic with (the uvular) r, nor with (the palatal) j,[1] whereas Vestergaard operates with a class of "palatals" including k g j ɣ ŋ r (p. 55). And since s does combine with dentals (sd, sl, sn), s does not "count as" homorganic with any other non-vowel.[2] Furthermore it should be noted that labiodentals (f v) belong to the labial class (in agreement with Vestergaard), which agrees well with the fact that the influence of a following homosyllabic labiodental on a preceding short a is like that of a bilabial but different from that of a dental (cf. Basbøll 1972a). By this principle (i) the following non-occurring initial two member clusters, which are in accordance with the order restrictions, are excluded: bv,[3] pv, fv, bm, pm, fm, dn, tn, dl, tl, nl.

---

1) In our framework this latter fact could of course be interpreted as indicating that the restriction concerns only consonants, not glides.

2) Since r is the only non-vowel which does not combine with s initially (except for h, of course, which as mentioned is completely non-combinable and therefore ignored in the present notes), one might treat initial s and r as constituting a functional "homorganity class" ("queer dentals", cf. the fact that r distributionally acts like dentals (including s) in several respects, see the end of section 6.).

3) The group bv is found in the interjection bvadr, used in the Danish translation of Charles M. Schultz' (comic strip) "Peanuts" (in Danish: "Radiserne"). Note that its final consonant group is aberrant too.

(ii) *s* does not combine with aspirated consonants or, more generally perhaps, *aspirates do not combine initially*. This restriction excludes the following non-occurring initial two member clusters which are in accordance with the order restrictions: *sf*, *sp*, *st*, *sk*.

(iii) *m does not combine with plosives initially*. This restriction excludes *bm*, *pm*, *dm*, *tm*, *gm*, *km* (*bm*, *pm* were already excluded by virtue of principle (i)).

(iv) *Nasals and liquids do not combine initially*. This restriction excludes *ml*, *mr*, *nl*, *nr* (of which *nl* was already excluded by principle (i)). However, if the nasals and liquids belong to the same order class, (iv) should be replaced by the restriction that *voiced obstruents and nasals do not combine initially* (the same is true finally, cf. 4.2.2.1. and 5.).

Now, there are only two non-existent initial two member clusters satisfying the order restrictions which are not excluded by principles (i)-(iv), namely *bn-* and *sr-*, both of which seem impossible to me.[1] Concerning *bn* it should be noted that *pn-* is found (thus pronounced) in foreign (Greek) words like *pneumatisk* (cf. Vestergaard p. 50), so maybe *bn-* is not a structural impossibility. For the non-existence of *sr-*, cf. note 2 of the preceding page.

4. Final clusters

4.1. Mono- or polymorphemic?

It was mentioned in the introduction that Vestergaard requires that both parts of a suggested bimorphemic word be found independently in other words in exactly the same phonemic

---

1) /nj/, /vj/, and /vl/ are here considered to be possible clusters. /kj/, which according to Vestergaard (p. 40) is not found in any monosyllable, occurs in my language in *kiosk* [ki̯ʌsg] (like [ki̯o:lə] which Vestergaard mentions).

shape, and furthermore that the suggested stem occurs as an isolated word, in order to recognize a morpheme boundary between them. I agree that the distinction between mono- and polymorphemic clusters is important, but I recognize more morpheme boundaries than Vestergaard, i.e., some of his monomorphemic clusters will be broken up here (section 4.2.1.).

To fully understand the implication of this disagreement with Vestergaard, the following remark should be made. As for the distinction between accidental and systematic gaps, I do not think the subclass of final clusters appearing in native monomorphemic words is a particularly interesting one, since I would like to suggest the hypothesis that final clusters appearing in polymorphemic words can be freely introduced in monomorphemic words too.[1] (Something resembling this idea can, of course, be found many places in the literature, see e.g. Vestergaard p. 57 with reference.) This principle has the further consequence that a statement about the clusters occurring finally in native monomorphemic words is less of a hypothesis and more of a simple registration than was the case in section 3. above. In my view, this also justifies a higher degree of exclusion of registered material by the criterion of generalizability (cf. Spang-Hanssen 1959 p. 110 ff) than before. On the other hand, this conclusion forces us to try to state the principles determining the occurrence of sequences of morphemes that do not contain any vowel (which Vestergaard has not aimed at doing), e.g. to state which constraints are violated by a sequence like aspskt. This will be attempted in section 4.3.

One further remark: I fully agree with Vestergaard that imperatives should be left out of the primary material. On

---

1) Thus new monomorphemic words can be introduced without any cluster simplification even if they end in a "polymorphemic" cluster (the reservation should be added that the number of consonants may not exceed that found in monomorphemic words).

the other hand, Spang-Hanssen (p. 218 ff) includes even the oddest imperatives, like mejsl! or hindr!. Spang-Hanssen is not much concerned with order phenomena, and so this procedure is of minor consequence, whereas it would seriously affect our treatment. It should be pointed out that disregarding the imperatives at the outset (which does not prevent that they may eventually be classified according to the criteria established for the rest of the material, cf. Basbøll 1970 p. 21-23) is not free of commitment: An infinitive which ends in shwa preceded by a cluster, and which is not excluded from the material as "foreign" or the like, should either have a medial cluster which is acceptable as a final cluster too, or a medial cluster which can be resolved into a final cluster (that is acceptable) plus one of the consonants /l r n/ (see section 6.). E.g., if the verb sløjfe is not excluded from the material, jf must be recognized as a possible final cluster (in monosyllables it is only found in imperatives and the name Leif, cf. Vestergaard p. 51), since f cannot occur in a shwa-syllable according to section 6. (If the conditions on shwa-syllables should turn out to be different than claimed here, this might of course invalidate the argument). Note that the special status of the imperatives is not expressed simply by classifying them as bimorphemic (the second morpheme being a sort of "subtraction morpheme", cf. Uldall 1936), since their phonotactic structure is often contradictory to the phonotactics not only of monomorphemic but also of polymorphemic words.

4.2. Monomorphemic clusters

4.2.1. Three member clusters

In Vestergaard's material (p. 57, reproduced in section 8.) there are only three clusters which do not begin with a sonorant, viz. (in his notation) /psk/ glubsk, /kst/ takst, /tsk/ skotsk. However, the examples given are clearly bimorphemic, cf. glubende (appetit), taksere, Skotland. They have been included as monomorphemic by Vestergaard's criterion that in order to be bimorphemic the stem of a word must be found as an isolated word (no word glub exists, and taks, skot are only found as clear homonyms to the stems in question). Furthermore, glubsk has, quite regularly, a short vowel in distinction to glubende, and as already mentioned this will suffice for Vestergaard as a criterion for not recognizing them as instances of the same stem. According to Spang-Hanssen's table 5 (1959, at the end of the book), the only occurrences of such monomorphemic final clusters are 3 examples of kst (which are not listed). These probably include vækst (which according to our criteria is bimorphemic, cf. vokse) and tekst (a foreign word, whose t does not belong etymologically to the root, cf. Latin textum, texere). In my view, this indicates that the clusters in question do not belong to the core of native monomorphemic final clusters.

When we consider the rest of Vestergaard's material, the following clusters are in some way deviant from the core of final monomorphemic three member clusters (cf. Spang-Hanssen 1959 p. 167, from where the following data is taken): /rts/, found only in five (foreign) words, viz. erts, marts, sirts (rare), terts, kvarts, /rft/, found only in one (foreign) word, viz. værft, /lkt/, found only in one (archaic) word, viz. mulkt, often pronounced without /k/, /lft/, found only in two (archaic) words, viz. hælvt, tylvt, of which the latter is often pronounced without /f/ and the former with either /v/ or /f/ (cf. its near homonym helt [hɛl?d]). The only cluster of any importance among these is /rts/. The reason why the foreign words ending in /rts/ seem to have been so readily accepted phonotactically (i.e. without simplification of their final cluster) is probably that the cluster is very common in morphemically complex forms, viz. genitive forms in rt+s, cf. section 4.1. above.

The remaining clusters of three consonants all consist of a sonorant plus s plus a non-labial plosive. Furthermore, every form containing a cluster of this kind which ends in a velar (viz. /rsk, lsk, msk, nsk/) is an adjective. Among Vestergaard's examples dansk is clearly bimorphemic according to our criteria (cf. Danmark, danificere, etc.), and the other three: harsk, falsk, lumsk might be called "pseudo-derived" even though there is no synchronically probable root for them to be derived from. This is not meant as an exact description, let alone an explanation, but only as a suggestion that the examples with -sk are peculiar in some way.

One need not be bothered by the fact that there seems to be no fixed borderline between clusters which we accept as monomorphemic, and clusters which must be polymorphemic[1]. On the contrary, it is a consequence of the assumption made in section 4.1. that no such clear-cut borderline should be expected.

Note that the non-vowels which occur in the first position of a native monomorphemic final three member cluster form a "natural phonological class", comprising the sonorant non-vowels. There is one exception: [ð] is not found in this position. However, this need not lead to a revision of our classification of [ð] as a sonorant (see sections 2.2.2. and 4.2.2.1.), nor does it force us to give up the idea that phonotactic facts should be expressed by means of "natural classes" in the phonological sense. Instead, the non-occurrence of [ð] can be illustrated by referring to the rule which deletes [ð] before a dental stop belonging to the same word (cf. Rischel 1970). John Austin (1971 p. 46 ff) has proposed to enlarge the scope of the rule to apply before all dental obstruents, i.e. also before s, in order to account for pronunciations like bidsel [bisəl], historically derived from

1) although one may, of couse, (like Vestergaard) choose one fixed criterion to decide all cases.

bide [bi:ðə]. Similar examples are *fødsel*, *ødsel* [føsəl, øsəl] 'birth', 'extravagant', derived from *føde*, *øde* [fø:ðə, ø:ðə] 'give birth to', 'waste (money etc.)' with the non-productive and rare nominalizing suffix *sel* (the same as in *bidsel*).[1] One could adduce examples like *bedst* [bɛsd], cf. *bedre* [bɛðrə], and alternative pronunciations in rare forms like *nådsensbrød* ['nʌsəns,brø˙ʔð] and *bådsmand* ['bʌs,manʔ].

However, the deletion of *ð* before *s* only occurs in some fixed forms (like *bedst*) and in some forms in which the morpheme boundary is not very transparent (like *bidsel*), but never before e.g. the genitive ending *s*. The connexion between the transparency of the morpheme boundary and the tendency to retain [ð] before [s] can be seen by considering the word *rødspætte* ['rœs,bɛdə] 'plaice', which despite its accent structure as a normal compound is clearly lexicalized (semantically), probably because the second part of the "compound" exists in isolation only with the meaning 'woodpecker'. It should be compared with a possible, but to my knowledge unexploited form like *rødspætte* constructed as a "woodpecker-compound" like *flagspætte*, which would most certainly be pronounced with [ð], i.e. ['rœð,sbɛdə], because of its transparent analysis into *rød* and *spætte*. (Notice that it will be clearly distinct from *rød* *spætte* in a NP like *en rød spætte* [en 'rœðʔ 'sbɛdə] 'a red woodpecker'.)

Since the deletion of *ð* before a dental stop (the rule proposed by Rischel) also occurs before perfectly transparent morpheme boundaries (e.g. in *fødte* [fø:də], preterite of *føde*), the rule could be formulated like this:

$$\eth \longrightarrow \emptyset \;/\; \underline{\quad\quad} \; \big\langle (\,\#\,) \big\rangle \begin{bmatrix} -\text{son} \\ \langle -\text{sib} \rangle \end{bmatrix}$$

1) A form like *rædsel* [ræðʔsəl] 'horror' might be explained by its long or geminated *ð* which receives the stød (as opposed to all the other examples which have a short or single *ð*), since it is derived from the adjective *ræd* [ræðʔ] 'scared' with stød and *short* vowel.

(i.e. the rule ignores whether or not the juncture # is present, unless ð is followed by a sibilant, i.e. s). This formulation presupposes that e.g. the genitive ending s is preceded by # , but there is no space here for a discussion of boundaries. The existence of this rule agrees well with the fact that ð is in general not found before any obstruent within the morpheme (the words snedker, bødker 'joiner', 'cooper' are normally pronounced [sneˑʔgɒ, bøˑʔgɒ], although alternative pronunciations exist, manifesting a tendency towards "spelling pronunciations" in such cases). Since sequences like ð+p, ð+g, etc. never arise in the concatenation of morphemes within the word (all the relevant endings starting with a sonorant or a dental), the rule has been given the most general environment possible in that it does not mention that the following obstruent should be a dental.

The conclusion is that the non-occurrence of [ð] as the first member of monomorphemic three member clusters is not a special fact about such clusters but a consequence of the mentioned rule with the effect of restricting the occurrence of [ð] before [s] to clearly bimorphemic clusters.

4.2.2. Two member clusters

4.2.2.1. Order restrictions

The endings (relevant to the present discussion) which can be added to a monosyllabic root either contain a vowel, or consist entirely of voiceless consonants. For this reason the order restrictions of voiced segments in the final part of the stressed monosyllable will be true for both mono- and polymorphemic words. This has the further consequence that it will be revealing to set up as many order classes as possible for voiced segments in order to account for all the order restrictions which exist for Danish speakers (cf. section 3.2.1.), as the following scheme shows:

	I	IIa	IIb	III	
VOWEL	i̯ u̯ ɒ̯	ð	l	m n ŋ v ɣ	VOICELESS SEGMENTS

Note that classes I and II are "natural classes", and that III includes two such classes (cf. section 5. below).

ð is not placed in class I because of the existence of such forms as arbejd! which may be pronounced [ˈɑːˌbai̯ʔð] (together with [ˈɑːˌbai̯ʔd]), and similarly the old Latin term for a university grade: haud (illaudabilis) [hɑu̯ʔð] (together with the more normal pronunciation [hɑu̯ʔd]). Uldall (1936 p. 54) quotes the form byrd [byɒ̯ʔð] (which is normally pronounced [byɒ̯ʔd]); the form ending in [ð] is very rare, but there is no doubt that if pronounced it will always be a monosyllable (in distinction to hædr! [hɛðʔɒ] with the reverse cluster, which is generally bisyllabic). It is true of all the mentioned clusters ending in [ð] that they are much more possible than the reverse clusters in true monosyllables.

The distinction between IIa and IIb is only motivated by imperative forms like padl! [paðʔl], which are certainly more possible as monosyllabic forms than anything ending in [lð]. However, this may be due to the rules for the pronunciation of the underlying segment /d/ (cf. Rischel 1970 and Basbøll 1972a), and thus need not be decisive of order classes.

The above suggests that the evidence for separating IIa and IIb is rather weak, whereas the separation between I and II (a-b) is quite firmly established, cf. e.g. avl, jarl [ɑu̯ʔl, jɑˑʔl].

Non-nasal sonorants always precede nasals in syllable final clusters, cf. halm, vidn! [halʔm, viðʔn], and the nasals therefore belong to the order class "after" that of l. The voiced obstruents ([v,ɣ]) never combine with nasals and occupy the same place as the nasals in relation to other non-vowels; they are therefore placed in the same order class.[1] Note that the voiced obstruents are only found after long vowels and after liquids (as a consequence of the phonetic "diphthongization", cf. section 2. above).

Finally, one might wonder why [ŋ] is placed together with the other nasals despite the fact that it is always vowel adjacent (which is Vestergaard's motivation for placing /ŋ/ in his class I). The reason is that [ŋ] is never followed by other voiced segments and thereby agrees with the other nasals (this fact would have to be stated by a special restriction if [ŋ] were placed in class I). But why, then, is it always vowel adjacent? This is due to the fact that [ŋ] is derived from /ng/ (see e.g. Austin 1971 p. 54 f, Basbøll 1972a p. 199 and 1972b p. 11 f for arguments in that direction), and there are never more than two voiced[2] non-vowels in a final cluster (mono- or polymorphemic, see section 4.3.), which excludes the possibility of an intervening voiced segment between the vowel and [ŋ]. This restriction seems to operate on a "higher level" than most of the other restrictions we have discussed, and it may in fact be a morpheme structure condition.

---

1) This does not apply to [ð] (cf. vidn! and rødm! [røðʔm] which are quite possible and clearly monosyllabic in all pronunciations), and this fact supports our decision to distinguish between [ð] and [ɣ] by a "higher" feature, viz. sonorant, than one which only accounts for their different place of articulation.

2) The /g/ of /ng/ is the same underlying segment which shows up as [ɣ] or [i̯, u̯] in other syllable final contexts, and is probably best described as being underlyingly voiced, as argued by Rischel (1970).

Concerning the voiceless segments, the order restrictions accounting for these can only be described in a very clumsy manner by means of order classes, since we have both sequences like [sg] (fisk) and [gs] (fiks). All the order restrictions operating here can be condensed into the statement that sequences of non-sibilant obstruents end in a dental. This statement permits the following sequences of non-sibilant obstruents: [vd, ɣd, fd, bd, gd] which are all found, and which are the only such sequences (e.g. stout, lægd, loft, recept, vægt). Notice that this restriction applies to both mono- and polymorphemic clusters, and that it is at the same time a restriction on order and on class membership.

4.2.2.2. Other restrictions

The scheme which implies that there are three order classes of voiced non-vowels is clearly over-generating, also with regard to sequences including sonorants. The absence of the "over-generated" clusters cannot in all cases be considered accidental. Thus some additional restriction(s) must be at work here (cf. above on the non-occurrence of [ð] plus obstruent).

In the core of native monomorphemic monosyllables, it is found that all sequences of non-dental non-vowels are /r/-combinations (cf. the end of section 6. where it is pointed out that /r/ in several more respects "counts as" a dental). Within Vestergaard's material (see section 8.), only /mp, mf, ŋk, jk/ violate this criterion. The last-mentioned group /jk/ is only found in the word sheik, which is clearly foreign.

In view of the non-existence of the groups /np, nf, nk/, the other three groups can be derived from an unspecified nasal followed by /p, f, k/ respectively, by a well-known rule of articulatory place assimilation of nasals before obstruents (it must be noted that in Danish this rule only applies before non-dental obstruents, cf. words like amt, punkt, vams, gængs). Since the maximally unspecified nasal is a dental, the general restriction can thus be claimed to be more generally applicable: not only does it hold for the mentioned nasal-plosive-groups as well as for other groups; it also explains the non-existence of phonetic [np, nf, nk]. Note that the mentioned assimilation rule applies before certain consonants can drop, e.g. a word like sfinks can be pronounced either [sfeŋʔgs] or [sfeŋʔs] (similarly with [sv-]), but never [sfenʔ(g)s], i.e. the rule must have applied before the optional dropping of the velar stop. Similarly a word like hingst [heŋʔsd] is never pronounced [henʔsd].

If the scope is widened to include foreign words, names, etc., it can be argued that the mentioned restriction only applies to consonants and not to glides, since forms like Hauch, sløjf! [hɑu̯ʔɡ̊, slʌi̯ʔf] are perfectly possible for Danish speakers (cf. Vestergaard p. 60). This formulation (e.g. finally, non-dental consonants do not combine directly) also removes the necessity of ascribing a special status to /r/. Note that that part of the restriction "sequences of non-sibilant obstruents end in a dental" (mentioned at the end of the preceding section) which is not an order restriction, makes exactly the prediction for obstruents which we have discussed in general in the present paragraph.

Also in the final part of the syllable (cf. section 3.2.2.), it seems to be the case that within the morpheme, aspirates do not combine (remember that only [f, s] and p t k preceded by a short vowel plus /r/ (e.g. vært, kork, etc.) are aspirated in the final part of the syllable, cf. section 2.2.3.). The only exception to that generalization is rts in erts, etc. which may be considered marginal (see section 4.2.1.).

4.3. Polymorphemic clusters

There are a number of Danish suffixes which do not contain any vowel (cf. Spang-Hanssen 1959 p. 204 and Diderichsen 1953 p. 180 ff). Among these there are some (generally non-productive) derivational suffixes making verbs out of non-verbal roots, and consisting of sonorants: *l* (e.g. *samle*, cf. *sammen*), *r* (e.g. *bladre*, cf. *blad*), and *n* (e.g. *blegne*, cf. *bleg*). Since these are not found as suffixes in monosyllabic words except in imperatives, they will be excluded here and only briefly mentioned in section 6.

The remaining ones are (generally productive) inflexional and derivational suffixes composed of voiceless obstruents and beginning with a dental: *s*, *t*, *st*, *sk*. They have a very restricted mutual combinability, and in general they seem to conform to the following maximal scheme:[1]

$$\begin{array}{l} \text{Initial} \\ \text{consonant} \\ \text{(cluster)} \end{array} \quad \text{VOWEL} \left\{ \begin{array}{l} \begin{array}{l}\text{Single}\\ \text{consonant}\end{array} + \left\{ \begin{array}{l} \text{(s) t} \\ \text{sk} \end{array} \right\} \\ \\ \begin{array}{l}\text{Final cluster ending}\\ \text{in } \underline{\text{s}} \text{ plus plosive}\end{array} \\ \\ \begin{array}{l}\text{Other final}\\ \text{clusters}\end{array} \quad + \text{ t} \end{array} \right\} + \text{ s}$$

This seems to be remarkably close to the structure of monomorphemic clusters, but it deviates in the following two respects: (i) any final cluster can be followed by +*s*; (ii) whereas monomorphemic final three member clusters consist of a sonorant plus *s* plus a (non-labial) stop, the polymorphemic ones (disregarding final +*s*) need only satisfy the restriction for their third member and for one of their first two members

---

1) An exception is the derivative *skælmsk*; cf. Diderichsen, *loc. cit.* + in the scheme indicates morpheme boundaries.

(i.e. their first member may be an obstruent, e.g. glubsk, or they may not have s as their second member, e.g. stærkt); but they should, of course, obey the order restrictions.

The scheme presupposes that identical obstruents are degeminated. It excludes, as it should, sequences like Pingstsk (Diderichsen), aspskt (Vestergaard), and aspskts (Jespersen).

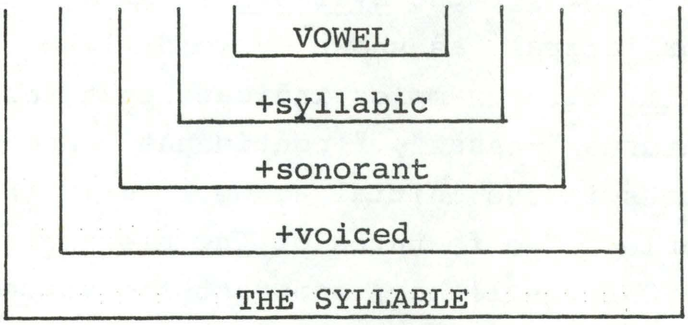
Since all the suffixes in question contain at least one dental, it will be seen that polymorphemic words obey the same basic combinability restriction (mentioned in the preceding section) as monomorphemic words, viz. that non-dental consonants do not combine directly in the final part of the syllable. However, they do not always obey the additional criterion that aspirates do not combine, since there are combinations with +s: hof+s, fork+s, etc.

It should be added that the present tense ending ([ɒ̯,ɒ]) forms a syllable by itself except after a vowel (e.g. ser, falder [se·ʔɒ̯, falʔɒ]). Thus its addition to a stem does not create any new consonant clusters.

5. Maximal syllabic structure

In the preceding we have argued in favour of the establishment of order classes which were assumed to be more homogeneous phonetically than Vestergaard's. It is tempting now to compare the insights of sections 3.2.1. and 4.2.2.1., and to try to generalize the results to a hypothesis on syllabic structure in Danish, applying at the "level" discussed in section 2.1. It will be the maximal syllabic structure, where the term "maximal" both implies that the syllable is stressed and that the vowel is short (long vowels having a much more restricted repertoire of consonant clusters after them). If either of these criteria for "maximality" is not satisfied, the "maximal syllabic structure" model in a way still applies, but then it presupposes additional restrictions (to those mentioned in the preceding sections), i.e., it will then be vastly over-generating.

The syllable will thus be postulated to have the following "hierarchical" structure:



The figure should be read as follows: The peak of the syllable is the vowel. The vowel and possible adjacent glides constitute the syllabic part of the syllable. The syllabic part of the syllable together with possible adjacent non-syllabic sonorants constitute the sonorant part of the syllable. The sonorant part of the syllable together with possible adjacent voiced obstruents constitute the voiced part of the syllable. And, finally, the voiced part of the syllable plus possible adjacent voiceless consonants constitute the syllable (or syllabic theme, in Hjelmslev's terminology). It is evident that in concrete cases the extent of contiguous (in the "hierarchy") features can be identical.[1] This amounts to the following hypothesis on the Danish syllable (and that is where the "hierarchical" nature of the model comes in): The following "implication chain" holds true without exception: vowel ⊇ [+syllabic] ⊇ [+sonorant] ⊇ [+voiced] and, similarly, the chain with switched ±values holds true in the opposite direction:

[-voiced] ⊇ [-sonorant] ⊇ [-syllabic] ⊇ non-vowel.

---

1) e.g. in the word kat, the voiced part, the sonorant part, and the syllabic part all equal the vowel a.

This is the reason why we should not insert features like "nasal", "continuant", and "lateral" in the model, although it holds true that <u>within the sonorant part</u> of the syllable the center is non-nasal and continuant, and the margin nasal and non-continuant, and that <u>within this non-nasal center</u> the center is non-lateral and the margin lateral. For it is not true that the margin of the <u>syllable</u> is always nasal and non-continuant, or lateral, as shown by words like <u>fnat</u>, <u>flamsk</u>, <u>plejl</u>, <u>halm</u>. Such examples indicate that the distribution of the features "-nasal", "+continuant", and "-lateral" can be discontinuous in the initial as well as in the final part of the syllable. The features of the hierarchy, on the other hand (i.e. VOWEL, syllabic, sonorant and voiced), can never be discontinuous in the syllable, neither in their positive nor in their negative values.

In fact, the claim is that the features of the "hierarchy" are distributed around the peak of the syllable, so that each feature may spread continuously over several segments in the way indicated by the hierarchy. This could be formulated so that "one instance of" e.g. the feature [+sonorant] "belongs to" several segments at the same time.[1]

But our hypothesis on syllabic structure does not end here. This syllable model also explains (in a vague sense) nearly all of the observed order relations in the monosyllable. In fact, the hierarchy is, in all relevant respects, a notational variant of the following scheme which expresses the order relations more directly:

-son -voi	-son +voi	-syl +son	+syl	VOWEL	+syl	-syl +son	-son +voi	-son -voi
sf ptk bdg	v	rl mn	i̯		u̯ɒ̯i̯	ðl mnŋ	vɣ	sf bdg (ptk)

---

1) This was, in fact, Stephen Anderson's conclusion concerning the feature "syllabic" (but he did not include any of the other features of our "hierarchy" in his discussion, cf. section 2.2.1. above).

It will be seen that this scheme correctly predicts the order of *any* permitted (unordered) set of *either* initial *or* final voiced non-vowels with the reservation that *in the final part of the syllable, non-nasal sonorants precede nasals*. The order of voiceless consonants conforms to the following principle (which in fact excludes all wrong orderings of such consonants): *in the initial part of the syllable, only /s/ may precede a voiceless consonant*, and *in the final part of the syllable, any sequence of non-sibilant obstruents ends in a dental*.

Note especially that none of the order classes predicted by the model is "descriptively" superfluous. Moreover, as we saw in section 4.2.2.1., the restriction on final clusters that "non-nasal sonorants precede nasals" becomes a direct consequence of the order class model if nasals are taken to "belong to" the order class of voiced obstruents instead (syllable initially such a "readjustment" of the order classes will have no substantial consequences, see section 3.2.2. above). However, such an order class consisting of nasals and voiced obstruents can hardly be said to be a "natural class"[1] (except in the vague sense that both voiced obstruents and nasals form a somewhat "less peak-forming" class of voiced segments in distinction to the continuant sonorants, but this intuition may be of a circular nature, viz. depending on distributional knowledge), and I find this a very serious objection to such an alteration of the model.

---

1) Of course the class of nasals and voiced obstruents can be given a common definition in distinctive features when using Greek letter variables, viz. as [α sonorant, $-\alpha$ continuant, +voiced], but this seems to me completely "hocus-pocus" (one should probably in principle refrain from using such variables when defining "natural classes", cf. Wheeler 1972 p. 90 f).

Everything which has been said so far in this section applies both to mono- and polymorphemic monosyllabic words (see section 6. on polysyllabic words). Notice that in order for the model not to be over-generating, the general restrictions discussed in the preceding sections also apply here, of course (monomorphemic clusters contain at most three non-vowels, and if there are three, exactly one of them is voiced, etc.).

6. Medial clusters

The material of this section is furnished by the infinitive forms of Danish verbs registered in *Dansk Rimordbog* (2nd edition, Politiken 1963). It turned out that all medial clusters in this material (i.e. sequences of non-vowels between the stressed vowel and the word-final shwa) could be described as consisting of a possible monomorphemic final cluster or of such a cluster plus one of the non-vowels /l, r, n/. This was true regardless of whether postulation of a morpheme boundary before /l, r, n/ was possible (cf. *angre* without a morpheme boundary before *r*,[1] but *blomstre* with a morpheme boundary before *r*, cf. *blomst*).

It should be pointed out that the (phonological) syllable boundary need not go directly before /l, r, n/. For example, *fordre* and *hærde* have the syllable boundary before *d*, which is therefore pronounced (as a plosive) and not deleted, but they conform to the general structure (concerning medial clusters) only if the "final" cluster is considered to be *rd* (cf. Basbøll 1972a p. 199 f).

---

1) However, the consonantal manifestation of /r/ shows that the syllable boundary goes *before r*. (Actually, this is a better example of the importance of "Hjelmslev's law" for syllabification - viz. that the syllable boundary must go before *r* since *ŋr* is an impossible termination of a Danish monosyllable - than the one given in Basbøll 1972a p. 187 (*Ad (B)*).

Only very few of the registered final clusters in Vestergaard 1967 (cf. section 8.) fail to occur medially in the material (with or without a following /l, r, n/). Those which do not are probably accidentally missing except in a few cases where the "original" cluster is so rare that no conclusions on accidental vs. systematic non-occurrence can be drawn ("sporadic occurrence" in Spang-Hanssen's terminology, cf. [ɣd, ŋd, msg], some of which may occur in inflected forms like hængte, lumske, as mentioned in the end of the present paragraph). On the other hand, according to our section 4.2.1. it is no accident that three member clusters consisting only of voiceless consonants (Vestergaard's /psk, kst, tsk/) never occur medially in monomorphemic words (with or without a following /l, r, n/): it is simply because such clusters are not monomorphemic (but it is evident that final clusters can generally be turned into medial clusters by adding an ending which starts with shwa, but then there will, of course, always be at least one morpheme boundary between the stressed vowel and shwa).

Medial clusters may also throw light upon the distinction between accidental and systematic gap in another way: certain medial clusters may fill accidental holes in final combinations (e.g. sløjfe, strejke, dogme, sødme).

Finally, one class of deviating medial clusters should be mentioned, viz. examples like balje, midje, linje. Since j does not belong to the class /l, r, n/, and since final groups like lj, ðj, nj seem to be systematically missing in monosyllables, we should look for another explanation. We propose that j be derived from the vowel i which is thus "reduced" in the position between a consonant and shwa. This rule cannot be shown invalid by means of examples like villige [viliə], commutable with [vili̯ə], since villig should end underlyingly in shwa plus a high consonant (a /ɣ/ or the like), cf. Eric P. Hamp's proposal that the derivative ending ig be derived from /əj/ (unpublished hand-out).

The existence of the class /l, r, n/ (which is well established, cf. Diderichsen, loc. cit.) confirms the claim that r in some respects "counts as" a dental (since the segments in question are then exactly those of the initial non-vowels which are sonorant and dental, and in the position in question /l, r, n/ are of course syllable initial). Further, it could be added that r is the only non-vowel which is phonetically non-dental and occurs in the final part of shwa-syllables (the others are [s, d, ð, l, n], i.e. the entire set of final dental consonants), which supports the same conclusion. (It is well known that r is historically derived from a dental.)

7. Conclusion

In the present notes we have tried to find and discuss some principles which can be said to account for systematic restrictions in Danish consonant combinations.

Most importantly, it was found that nearly all order restrictions could be explained by reference to a very general model of maximal syllabic structure in Danish, applying at a level near the phonetic surface (but not identical to it; cf. the rather unexplored notion of "shallow phonological structure", which may be relevant here), see section 5.

We have also found some further restrictions which are not quite identical in initial and final position, but nevertheless exhibit a high degree of parallelism: There are at most three non-vowels in monomorphemic initial and final clusters, and if there are three, exactly one of these is voiced whereas the others are s plus a plosive; furthermore, aspirates do not directly combine within the morpheme, neither initially nor finally.

Some further non-combination rules have been discussed, and the basic rules may be formulated as follows: (i) initially, homorganic consonants do not combine, whereas (ii) finally, non-dental consonants do not combine directly (that the restrictions are really different can be seen by comparing the excluded initial groups dl, tl with the perfect final group lt [ld]). This latter restriction seemed to apply on a rather "high level" compared to the other restrictions (since groups like [ŋg, mf] were explained as being derived from /nk, nf/), possibly it is some kind of a morpheme structure condition (in distinction to this, cf. that [ŋ] seems to "count as" one velar segment in the final group [ŋsd]). The non-occurrence of [ð] before obstruents belonging to the same morpheme is probably due to a morpheme structure condition too. (We have not discussed the extent to which such morpheme structure conditions could or should be viewed as "blank filling rules", nor other questions concerning the relation between the present distributional survey and different models of generative phonology and morphology.)

Whether distributional facts, for example of Danish, should after all be described as we have tried to outline, viz. partly as due to morpheme structure conditions, partly (and this aspect was particularly emphasized in the present paper) as due to general principles of syllabic structure on a rather phonetic level of description, is of course a completely open issue. But at least I think it should be interesting to investigate further (and hopefully less superficially) into these matters.

8. Appendix: Vestergaard's material

For ease of reference of the reader, Vestergaard's examples of each cluster are reproduced here, in his phonemic notation and in the order in which he gives them.

8.1. Initial clusters (p. 40 f)

Two consonants

/sj/ sjæl	/fr/ frem	/sv/ svær
/tj/ tjørn	/dr/ drik	/tv/ tvær
/pj/ pjat	/br/ brun	/kv/ kvik
/kj/ kjole	/gr/ grov	/dv/ dværg
/fj/ fjols	/vr/ vred	/sn/ sno
/dj/ djærv	/sl/ slag	/kn/ kno
/bj/ bjørn	/pl/ plads	/fn/ fnat
/gj/ gjord	/kl/ klo	/gn/ gny
/mj/ mjød	/fl/ fløjl	/sm/ smal
/tr/ træ	/bl/ blandt	/st/ stå
/pr/ præst	/gl/ glad	/sp/ spå
/kr/ kro		/sk/ skal

Three consonants

/str/ strå	/stj/ stjært	/skv/ skvat
/spr/ spring	/spj/ spjæld	/spl/ splint
/skr/ skrog	/skj/ skjold	

8.2. Final clusters

Two consonants (p. 51 f)

/rl/ jarl	/lf/ alf	/jk/ sheik
/rp/ skarp	/ls/ hals	/jn/ degn
/rt/ vært	/mp/ damp	/js/ majs
/rk/ værk	/mt/ amt	/jl/ gøgl
/rg/ dværg	/mf/ trumf	/gt/ lægd
/rm/ arm	/ms/ vams	/gn/ vogn
/rn/ ørn	/ŋt/ punkt	/gl/ hagl
/rv/ arv	/ŋk/ flink	/ft/ gift
/rf/ skærf	/ŋs/ gængs	/pt/ recept (tabt etc.)
/rs/ vers	/nt/ splint	/ps/ gips
/lp/ skalp	/ns/ dans	/ts/ skyts

/lt/ filt	/vt/ støvt	/kt/ vægt
/lk/ folk	/vn/ ovn	/ks/ straks
/lg/ valg	/vs/ snavs	/sp/ bisp
/lm/ halm	/vl/ svovl	/st/ hest
/lv/ ulv	/jt/ sløjd	/sk/ flæsk

Three consonants (p. 57)

/rst/ tørst	/lsk/ falsk	/nst/ kunst
/rts/ erts	/lft/ tylvt	/vst/ provst
/rft/ værft	/mst/ blomst	/jst/ gejst
/rsk/ harsk	/msk/ lumsk	/psk/ glubsk
/lkt/ mulkt	/ŋst/ hingst	/kst/ takst
/lst/ svulst	/nsk/ dansk	/tsk/ skotsk

References

Austin, John S. 1971: Topics in Danish Phonology (unpublished thesis, Cornell University)

Basbøll, Hans 1969: "The phoneme system of advanced standard Copenhagen", ARIPUC 3/1968, p. 33-54

Basbøll, Hans 1970: "Notes of the phonology of Danish imperatives with a digression on vowel quantity", ARIPUC 4/1969, p. 15-42

Basbøll, Hans 1972a: "Some conditioning phonological factors for the pronunciation of short vowels in Danish with special reference to syllabification", ARIPUC 6, p. 185-210

Basbøll, Hans 1972b: "Some remarks concerning the stød in a generative grammar of Danish", in: F. Kiefer (ed.), Derivational Processes (Stockholm), p. 5-30

Diderichsen, Paul 1953: "Bidrag til en analyse af det danske skriftsprogs struktur", Selskab for nordisk filologi. Årsberetning for 1951-52, p. 7-22, reprinted in: Diderichsen, Paul, Helhed og Struktur (Copenhagen 1966), p. 169-187

Fischer-Jørgensen, Eli 1952: "On the definition of phoneme categories on a distributional basis", Acta Linguistica VII, p. 8-39, reprinted in: Hamp, E., F. Householder, and R. Austerlitz (eds.), Readings in Linguistics. II (1966), p. 299-321

Frøkjær-Jensen, B., C. Ludvigsen and J. Rischel 1971: "A glottographic study of some Danish consonants", F&S, p. 123-140, also published in identical form in this issue of ARIPUC

Hjelmslev, Louis 1951: "Grundtræk af det danske udtryks-system med særligt henblik på stødet", Selskab for nordisk filologi. Årsberetning for 1948-49-50 (København), p. 12-24, to be published in English in: Hjelmslev, L., Essais linguistiques II

Jespersen, Otto 1934: <u>Modersmålets fonetik</u> (København)

Labov, William 1970: "The study of language in its social context", <u>Studium Generale</u> 23, p. 30-87, reprinted in: Fishman, J: (ed.), <u>Advances in the Sociology of Language</u> I (1971), p. 152-216

Martinet, André 1937: <u>La phonologie du mot en danois</u> (= <u>Bulletin de la Société linguistique de Paris</u> 38, p. 169-266)

Rischel, Jørgen 1969: "Notes on the Danish vowel pattern", <u>ARIPUC</u> 3/1968, p. 177-205

Rischel, Jørgen 1970: "Consonant gradation: A problem in Danish phonology and morphology", in: Benediktsson, H. (ed.), <u>The Nordic Languages and Modern Linguistics</u>, p. 460-480

Sigurd, Bengt 1965: <u>Phonotactic Structures in Swedish</u> (Lund)

Spang-Hanssen, H. 1959: <u>Probability and Structural Classification</u> (Copenhagen)

Uldall, H. J. 1936: "The phonematics of Danish", <u>Proc.Phon.</u> 2, p. 54-57

Vestergaard, Torben 1967: "Initial and final consonant combinations in Danish monosyllables", SL XXI, p. 37-66

Wheeler, Max W. 1972: "Distinctive features and natural classes in phonology", JL 8, p. 87-102

PERCEPTION OF GERMAN AND DANISH VOWELS WITH SPECIAL REFERENCE TO THE GERMAN LAX VOWELS /I, Y, U/

Eli Fischer-Jørgensen

1. Introduction

The investigations reported in this paper were carried out at large intervals from 1955 to 1973. They had a double purpose:

(a) The main purpose was to find out how German lax vowels, especially /I, Y, U/, were perceived when heard in isolation. Some phoneticians, e.g. E.A. Meyer (1910 and 1913) and R.M.S. Heffner (1949 p. 97-98), consider these vowels to be more closely related perceptually to [i, y, u] than to [e, ø, o], cp. e.g. Heffner (1949 p. 97): "It is still a fact that [i] [I] [y] and [Y] are to our perception [i]-type vowels rather than [e]-type vowels". Both Meyer and Heffner emphasize that this auditory impression is in contradiction to the results of E.A. Meyer's palatographic and plastographic investigations which showed that North German /I/ and /Y/ have a much lower tongue position than /i:/ and /y:/, and that it may even be lower than that of /e:/ and /ø:/. An explanation which immediately suggests itself is that the auditory impression may be influenced by orthography. This hypothesis might be tested by investigating the perception of isolated vowel segments.

(b) A second purpose was to compare the perception of vowels in words and cut out of words, in which the surrounding consonants affect not only the transitions, but also the formant frequencies of the central part of the vowel. Such cases are found particularly in short lax vowels, e.g. in short front unrounded vowels between labials and short rounded back vowels between alveolars. How, for instance, is the vowel of Dutzend heard in isolation?

2. General characteristics of German lax vowels

It is not the intention to enter here into any detailed discussion of the opposition tense/lax, only to draw attention to a few major points. The discussion has mainly been centered around English vowels. It might be more useful to concentrate on North German, where the distinction is much more clear, and where the lax vowels /I, Y, U, ɛ, ɔ, œ, a/ can be compared to a set of monophthongal tense vowels /i:, y:, u:, e:, ø:, o:, a:/ (/ɛ:/ is often absent in natural North German speech).

2.1. Physiological characteristics

It is obvious that the lax vowels have a relatively lower tongue height than the tense vowels. It is also obvious that they do not simply constitute intermediate steps of vowel height. E.A. Meyer's finding that /I/ and /Y/ may have lower tongue height than /e:/ and /ø:/ has been confirmed by later X-ray photos of German vowels, cp. e.g. the X-ray photos of German /i:, I, e:/ in Russell 1929, Chiba and Kajiyama 1958, and Wängler 1961 (second edition, in the first edition (1958) /e:/ has been reproduced a second time instead of /I/). I have found the same relation in X-ray photos of /y:, ø:, Y/ spoken by a North German subject.

The tongue is, however, not only lower in lax vowels, it is also flattened out so that the tongue root is closer to the pharynx wall. Stewart (1967) has found the difference in pharynx width to be the essential feature in some West African vowel systems, partly based on Ladefoged (1964). Halle-Stevens (1970) have proposed to use the feature "advanced tongue root" also in English and German instead of tense/lax.

The elevation of the tongue in [i:] compared to [I] is considered to be a consequence of drawing the tongue root forward, the tongue being bunched up by this movement. The tension which Sievers (1901 p. 98) could feel in the muscles under the lower jaw (an observation which Hockett has repeated for English, e.g. 1955 p. 31) should thus be due to a contraction of the geniohyoid and mylohyoid muscles involved in advancing the tongue root, and the feature tense/lax could be dispensed with.

It is true that the high tense vowels /i:, y:, u:/ have a more advanced tongue root than the lax vowels /I, Y, U/ in German (and English), and sometimes it is also true of /e:, ø:, o:/ compared to /I, Y, U/; but this might be part of a general difference between tense and lax articulation. The description given by Jakobson and Halle (1956 p. 30) "greater (vs. smaller) deformation of the vocal tract away from its rest position" may still be a better way of formulating the difference. This question could be settled by means of electromyographic recordings of the various muscles involved in vowel articulation. Preliminary investigations of the tongue muscles in English show higher activity in the genioglossus for the tense vowels (McNeilage-Sholes 1964, Hirano-Smith 1967, and Lawrence J. Raphael 1971). Raphael is puzzled by finding more activity in /e:/ than in /I/ in his own pronunciation, but this is not astonishing since the tongue may be lower in /I/ than in /e:/, also in American English (cp. the X-ray photos in McNeilage-Sholes 1964).

German vowels do not seem to have been investigated by means of electromyography. Simple observations of lip and jaw movements give, however, a good deal of information, and such observations could be multiplied and supported by measurements without great difficulties. First, it is quite evident that the lip movement is less pronounced in lax vowels.

Already Sievers (1901 p. 86) drew attention to this fact. /Y/ and /U/ have less rounding than /y:/ and /u:/, and often less than /ø:/ and /o:/ (see, e.g., the lip photos in Wängler 1961), and /I/ has normally less lip spreading than /i:/. Moreover, although /I/ may have lower tongue height than /e:/, it generally seems to have less jaw opening (see Russell 1929, Chiba and Kajiyama 1958 p. 150, and Wängler 1961). These features can certainly not be a consequence of tongue root advancement, but point to a general laxness of the muscles involved in the formation of these vowels compared to tense vowels. Cp. also that Louise Kaiser (1941 p. 186 ff) has found that in Dutch /I/ has lower tongue height, but smaller lip and jaw opening than /e:/.

One more difference between tense and lax vowels which is fairly well established is that lax vowels have a stronger air stream (see E.A. Meyer 1913, Schuhmacher 1970, and EFJ 1969). E.A. Meyer explains this by a looser contact between the vocal cords in lax vowels, and this could be seen as part of the general laxness. Sievers (1901 p. 98 ff) is of the same opinion, Halle-Stevens (1970) have advanced the opposite hypothesis. They assume the vocal cords to close less firmly in tense vowels which should give them a breathy character. This may be true of the so-called tense vowels in West African languages, but sounds astonishing as a characteristic of tense vowels in German.

2.2. Acoustic analysis

Very few acoustic investigations of German vowels have been undertaken. Barczinski-Thienhaus (1935) found relatively more partials in lax vowels, but as these vowels were sung and continued for five minutes, the results are pretty uncertain.

Vierling-Sennheiser (1937) found some amplification of higher partials above 5000 Hz, particularly in /U/.

A detailed acoustic analysis of the vowels of six informants has been undertaken by Hans Peter Jørgensen (1969). The following perceptual investigation is mainly based on his material. Two of his informants were from the Rheinland. They had the first formant of /I, Y, U/ placed in between those of /i:, y:, u:/ and /e:, ø:, o:/. The other four informants were genuine North Germans, the first formants of their lax vowels /I, Y, U/ were either below or at the level of the first formants of /e:, ø:, o:/. As for the higher formants, lax unrounded front vowels had a lower F_2 and F_3 than the corresponding tense vowels, lax rounded front vowels had a lower F_2, but a higher F_3 (evidently because of the smaller amount of lip rounding), and lax back rounded vowels had a higher F_2 than the corresponding tense vowels. On the whole, the lax vowels were found to have a more central position in the F_1-F_2 vowel diagram, which is in complete agreement with what should be expected on the basis of the articulatory differences.

A. Fliflet (1962) has described the acoustic difference between German tense and lax vowels on the basis of visual inspection of spectrograms. He lists six characteristics of lax vowels: (1) centralized formant frequencies, (2) a more equal distribution of energy in the whole frequency domain, (3) less sharp formant contours, (4) less regular structure of formants, (5) longer transitions, (6) a general shimmering ("Unruhe") of the entire spectral picture, for instance more noise components.

These observations are in good agreement with the finding that lax vowels have a stronger air stream and with E.A. Meyer's hypothesis about laxness in the vocal cords; cp. also that Joos 1948 p. 97 states that synthetic vowels with broader formants sound more like American English /I, U/, whereas vowels with narrower formants sound more like German and French /e:, o:/. In support of their (opposite) hypothesis Halle-Stevens (1970) mention that the weakness of the higher formants in [i:] is so pronounced that it cannot be explained sufficiently by the low position of the first formant. Glottographic investigations are needed on this point.

E.A. Meyer supposed that the stronger air stream in lax vowels might give rise to high fricative noise, which might make /I/ sound more like [i:]. This is not very probable, and at any rate it does not explain why /U/ should sound more like [u:].

On the whole the articulatory and acoustic descriptions of tense and lax vowels are in good agreement and give no motivation for assuming that German /I, Y, U/ should be closer to [i:, y:, u:] than to [e:, ø:, o:] from an auditory point of view. If they sound like [i:, y:, u:] to Germans it must be due to other (phonemic and/or orthographic) factors.

3. Perceptual tests[1]

3.1. The material

The material consisted of a number of vowel segments cut out of words. A set of stimuli was prepared in 1955 and used for an identification test and for a small discrimination test. A new set of stimuli was prepared in 1963 and used for identification tests in 1963, 1964, 1969, and 1973.

---

1) I am grateful to Birgit Hutters, Hans Peter Jørgensen, and Henny Pontoppidan Lauritzen for helping me with measurements and calculations, and to Peter Holtse and Hideo Mase for making the graphs.

3.1.1. The material for the identification test 1955

The material comprised 47 vowel segments from various languages, Danish (7 items), German (33), Dutch (4), and English (3).

The vowels were taken from the following words:
Danish: (Speaker EFJ, female) /ki:lə, he:lə, hɛ:lə, pibə, vebə, nesə, sdødə/.
German: (Speaker ED, male) lieben, Lippen, Ibis, eben, immer, älter, Schluss
(Speaker GR, male) lieben, leben, Lippe, läppisch, Kehle, killen, kühle, Hölle, Hülle, tun, Ton, Tunnel
(Speaker CH, female) Kiele, Kehle, killen, Keller, Biber, Riff, Güte, Goethe, Kütte, Götter, fühlen, bündig.

The words spoken by ED were taken from a record, the words spoken by GR and CH were taken from a tape recording made at the Royal Technical High School in Stockholm for the purpose of investigating close and loose contact.

ED was a speech teacher speaking a clear and distinguished standard German, GR has an obvious Berlin accent. His [I, Y, U] are very low. CH was from Hamburg. Her speech is characterized by very close vowels and a high fundamental. The higher formants of her vowels were very weak.
Dutch: (Speaker Sch., male) tien, dienen, denen, tinnen.
English: (Speaker F.I., Amer. Engl., female) did, pip.
(Speaker B., Brit. Engl., male) soot.

The central part of the vowel (50-90 msec) was cut out so that (practically) no transitions remained. The cut was sharp.

The vowel segments were combined into two test tapes. The vowels from the same speaker were kept together, and the female voices came last, but there was no indication of the start of a new speaker. The vowels of the same speaker were given in random order. Each vowel was repeated three times and a number was said by the present author before each new vowel.

The tape recordings were made on a professional tape recorder, but for the dubbing and listening a semi-professional tape recorder was used, and the listening took place in a class room via loudspeaker. I was therefore not sure that the quality was sufficiently good, and as it turned out that the vowels of Güte and bündig spoken by CH were heard as [u:] and [o:] respectively by the majority of the listeners, I put the material aside.

The mistakes made in these two words may, however, have been due to the unusually weak higher formants of CH. The other results were confirmed by later experiments, and they will therefore be taken into consideration in the section on the results, but they will be mentioned relatively briefly, and no lists of the formant frequencies are given.

3.1.2. The material used for the listening tests 1963-1973

In 1963 a new series of stimuli was prepared. Vowel segments of 60-80 msec duration were cut out of the central part of the sounds as in the preceding experiment. The material comprised whole sets of vowels from one Danish and four German speakers, with the exclusion of the low vowels /a/ and /a:/ for which the difference tense/lax is rather dubious. Moreover there were some extra items, mainly vowels which were strongly influenced by the surroundings. There were 19 Danish and 57 German vowel segments.

All the speakers were male.

The Danish speaker was PD, born 1906, speaking a distinct and relatively old-fashioned Standard Danish. Compared to the mean of eight male speakers (EFJ 1972) he has a relatively low first formant of the high vowels and a relatively low second and third formant of most of the front vowels. The frequencies of the three first formants of his vowels are given in table 1.

The words are given in broad phonetic transcription.

TABLE 1.

Formant frequencies (in Hz) of the Danish vowels used in the identification tests 1963-1973.

Danish speaker PD

	F_1	F_2	F_3		F_1	F_2	F_3
/i:lə/	200	1975	3000	/ilə/	225	1900	2950
/he:lə/	290	2050	2750	/sbelə/	300	1925	2575
/hɛ:lə/	375	1975	2650	/hɛlə/	400	1850	2525
/hy:lə/	210	1825	2050	/hylə/	230	1625	2000
/ø:ðə/	275	1625	1925	/øləð/	290	1550	2025
/hœ:nə/	375	1550	1925	/hœnsə/	450	1475	2275
/ku:bə/	220	775		/kubəl/	270	750	
/ho:bə/	300	650		/sgobə/	500	950	2150
/hɔ:bə/	410	975		/hɔbə/	550	1150	2275

One of the four German speakers, ED, was also employed in 1955. The words used in 1963 were taken from the same record and were partly the same words, in order to obtain a certain control on the results of the old experiment. ED's long vowels are sufficiently long to present a minimal influence from surrounding consonants. Of the short vowels [U] in Schluss was not a very good example of a typical /U/, since it stood between two alveolars.

The other German speakers were the informants of Hans Peter Jørgensen, used for his acoustic analysis of German vowels (1969), and the items were taken from his tape recordings. The speakers were NB (from Schleswig), HT (from Berlin), and HL (from the Ruhr).

The words were chosen so as to present a minimal influence from surrounding consonants. The formant frequencies of these vowels (mainly based on Hans Peter Jørgensen's measurements) are given in table 2.

TABLE 2.

Formant frequencies (in Hz) of the German vowels used in the identification tests 1963-1973.

Speaker ED:

	F_1	F_2	F_3		F_1	F_2	F_3
lieben	275	2300	2950	Distel	300	2125	2750
leben	300	2175	2575	älter	450	1925	2425
müde	275	1825	2125	nützen	375	1575	2400
Flöte	325	1600	2150	öfter	525	1475	2350
suchen	260	600	2950	Schluss	325	950	2700
Ostern	350	750		hoffen	575	1000	2875

Speakers:

	NB			HT			HL		
	F_1	F_2	F_3	F_1	F_2	F_3	F_1	F_2	F_3
hiessen	220	2100	2900	225	2400	3075	275	2025	2700
Esel	300	2075	2525	275	2425	2750	350	1950	2500
hissen	300	1875	2325	325	2125	2750	340	1875	2450
essen	450	1825	2400	500	2025	2600	450	1775	2450
hüten	220	1550	1900	210	1550	2000	280	1650	1975
Höhlen	300	1350	2000	325	1525	1975	320	1500	1950
Hütten	300	1375	1975	320	1575	2125	300	1575	1950
Höllen	450	1325	2175	435	1450	2175	500	1350	1925
hupen	240	720		260	700		300	800	
hoben	340	750		325	700		400	800	
hupfen	340	825	2425	400	850		375	980	
Hopfen	500	950	2400	525	1035	2100	650	1025	2600

Some extra items were added, mainly to show maximal influence from surrounding consonants. The formants of these vowels are given in table 3.

TABLE 3.

Formant frequencies (in Hz) of the extra vowels used in the identification tests 1963 - 1973.

	F_1	F_2	F_3		F_1	F_2	F_3
ED:				NB:			
Lippen	300	1700	2425	bibbern	300	1700	2350
Ziege	250	2450	2975	Dutzend	340	960	2400
Minne	275	2200	2825				
				HL:			
				Dutzend	280	1100	2025
				Dotter	550	1150	2150

The vowel segments were combined into a test tape consisting of five distinct series, one for each speaker. Each vowel was repeated three times with an interval of one second. There was a five second interval between different vowels. Before each new vowel there was a brief tone signal of 300 Hz, which was found to be a pleasant pitch. Before numbers 5, 10, and 15 there were two tone signals. This was done in order to avoid spoken numbers, which might influence the perception of the following vowels.

A small extra test containing the words: hissen, hütten, hupfen spoken by NB, HT and HL, Dutzend and bibbern spoken by NB, Dutzend and Dotter spoken by HL and Lippen spoken by ED was played to the German listeners only.

The dubbing was done on a professional tape recorder (Lyrec) of the same type as the one used for the recording. The listening took place via loudspeaker in a class room with some damping of the walls (Danish listeners) or via earphones (German listeners in Kiel and Köln).

3.1.3. Materials used for the discrimination test 1955

A small number of German vowel segments were used in a discrimination test in 1955. The formant frequencies of these words are given in table 4. The measurements of CH's vowels are not very exact since she has a fundamental above 250 Hz.

TABLE 4.

Formant frequencies (in Hz) of the vowels used in the discrimination test 1955.

ED:	F_1	F_2	F_3	CH:	F_1	F_2	F_3
Ibis	250	2325	3050	fühlen	270	1800	2600
eben	290	2175	2700	Goethe	325	1825	2600
immer	290	2125	2675	Kütte	350	1825	2700
älter	450	1925	2425	Götter	400	1750	2700

GR:			
lieben	225	2336	3300
leben	340	2375	2975
Lippe	375	1775	2625
läppisch	525	1800	2650

The vowels of each speaker were combined in groups for comparison of the type later labelled 4IAX, for example "i-e, i-I", where the listener had to decide which pair, number one or number two, showed the higher degree of similarity between the members. Each double pair was repeated three times, in order to avoid memory effects as far as possible. The distance between double pairs was 1.5 sec., and the interval between different double pairs was 5 seconds. A number was spoken before each new double pair.

3.2. Listeners and instructions for listeners

The listeners taking part in the test in 1955 were 22 (for some stimuli 28) Danish students, attending a course of general phonetics for students of foreign languages.

The test tape made in 1963 was played both to Danish and German listeners. The Danish listeners were (1) two groups of students of foreign languages (24 in each), attending a course of general phonetics in the autumn 1963, (2) two groups of students of Danish (20 and 21), attending a course of Danish phonetics in the autumn 1963, 5 phoneticians and 5 dialectologists who listened individually in January 1964.

The German listeners were groups of phoneticians and students from Bonn, Hamburg, Kiel, and Köln.[1] The test in Bonn took place in 1969. Unfortunately, some high frequency components must have got lost when the test was played back, since many rounded front vowels were heard as rounded back vowels, and some unrounded front vowels were heard as rounded front vowels. That something was wrong also appears from a comparison between two answers to the test made by the same person, in Bonn in 1969 and in Kiel in 1973. In 1969 he had 12 mistakes of the type mentioned, in 1973 one. The material from Bonn was therefore discarded.

In Hamburg the test also took place in 1969. 19 listeners, mostly speech therapists, took part in the test. The results showed a very great dispersion, and as long /e:/, /ø:/, and /o:/ were identified more often with [i:, y:, u:] than with [e:, ø:, o:] (in a number of cases also with [ɛ, œ, ɔ]), I concluded that the listeners had not understood the instruction

---

1) I am grateful to the leaders of the Phonetic Institutes of Bonn, Hamburg, Kiel, and Köln for running the tests in their institutes.

(see below), and put the material aside for the time being.

In 1973 a group of 10 phoneticians and students in Kiel and 23 phoneticians and students in Köln listened to the test tape. The results did not differ very much from those obtained in Hamburg, although the dispersion was smaller. I therefore concluded that this type of answer could not be avoided and might give some information about German vowels. The results will be described in detail below.

The answer sheets for the test (as for the test in 1955) contained five vertical columns, one for each series, indicating the numbers 1-17 for series I, etc. There was a star before each number, and two before the numbers 5, 10, and 15, to indicate the tone signals.

At the top of the sheet was the diagram reproduced in fig. 1.

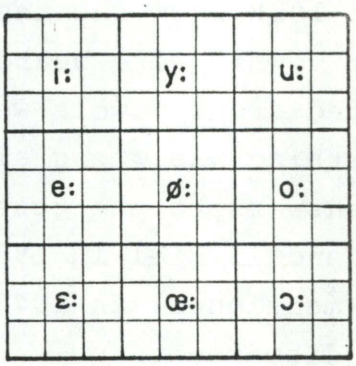


Fig. 1. Diagram placed on the answer sheet.

The instructions for the Danish listeners, which were spoken on the tape, contained the following information:

"You will now hear a series of vowels. Your task is to identify them by ear and take them down in phonetic transcription on the answer sheet to the right of the running numbers. The vowels are taken from different languages and are cut out of words. They are all very brief. In some cases a weak <u>p</u> or <u>b</u> may be heard after the vowel. Please, do not take notice of this. Each vowel is repeated three times ... (instructions about numbers and tone signals) ...

In order to obtain as precise an indication as possible of your auditory impression, the Danish long vowels have been used as a basis for the transcription. The vowels you hear are to be compared with the Danish long vowels [iː, yː, uː, eː, øː, oː, ɛː, œː, ɔː], which are indicated in the diagram on the top of the answer sheet. Each vowel symbol is placed in the small center square of a larger square. This is meant to be the placement of the normal Danish long vowels, whereas the surrounding smaller squares are intended for the placement of finer shades. If, for instance, you hear a vowel which has just the same quality as the long Danish [eː], you simply write the symbol e. If you hear an [e] which sounds somewhat [i]-like, you write an e with an arrow pointing upwards in order to indicate that you would place it in the small square just above the center square of eː, towards iː. If you find that it is an [e] which sounds somewhat [ø]-like, you must write an e with an arrow pointing to the right in order to indicate that you would place it in the small square to the right of the center square of eː, towards øː. An [ɛ]-like [e] should be written with an arrow pointing downwards, and an [œ]-like [e] with an arrow pointing obliquely downwards to the right. Similarly an [o] which sounds [ø]-like should be indicated by an o with an arrow pointing to the left, etc. Please write the arrows clearly, so that it is possible to distinguish between horizontal and oblique arrows."

The comparison with Danish vowels was chosen because the listeners could not be expected to have sufficient training in Cardinal Vowels, and the procedure with arrows was chosen because, otherwise, there should have been a large square on the sheet for each stimulus. Of course, the listeners might make the mistake of comparing with Danish short vowels instead. In most cases this would not make much difference, since Danish short vowels are only slightly lower than the corresponding long vowels and some speakers have no difference whatever. The pairs /oː/o/, /ɔː/ɔ/, and /aː/a/, however, are exceptions to this rule. The answers to the Danish vowel stimuli showed that the listeners had really used the long vowels as standards, since long /iː/, /yː/, and /uː/ were placed in the center squares, whereas short /i/, /y/, and /u/ were indicated to be slightly lowered. Moreover, the short /o/ was placed in the square of long [ɔː], which is in agreement with its quality.

Some of the students had not understood that they were allowed to use the peripheral squares (e.g. raised i, lowered ε). This is relevant for the judgement of short German /ε, œ, ɔ/. Only the experts have placed these sounds as lowered Danish [ε:, œ:, ɔ:]. The answers about Danish and German /ε, œ, ɔ/ should therefore be judged with caution.

For the German listeners there were similar instructions, but as German does not possess a set of long low vowels, it was not possible to use long vowel qualities as standards in all cases. The vowel symbols in the lowest row had therefore no length marks, and the German listeners were asked to take the vowel symbols in the central squares to indicate the long German vowels in bieten, hüten, hupen, beten, Höhle, hoben, and the short vowels in Betten, öffnen, hocken. The sheets used in Kiel and Hamburg contained an extra diagram with these key words placed in the center square.

This mixture of long and short key vowels seems to have caused some difficulties.

In order to give a rough impression of the formant frequencies of the vowels to be used as standards by Danish and German listeners, the average formant frequencies of the long vowels (except /a:/) of eight Danish male speakers (from EFJ 1972) and the average formant frequencies of the relevant German vowels spoken by six German male speakers (Hans Peter Jørgensen 1969) are indicated in fig. 2. The short German /I, Y, U/ are indicated by crosses for comparison. The vertical axis indicates formant 1, the horizontal axis formant 2 for back vowels and formant 2' for front vowels according to Fant's formula (Fant 1959):

$$F_{2'} = F_2 + 1/2\,(F_3 - F_2)\;\frac{F_2 - F_1}{F_3 - F_1}$$

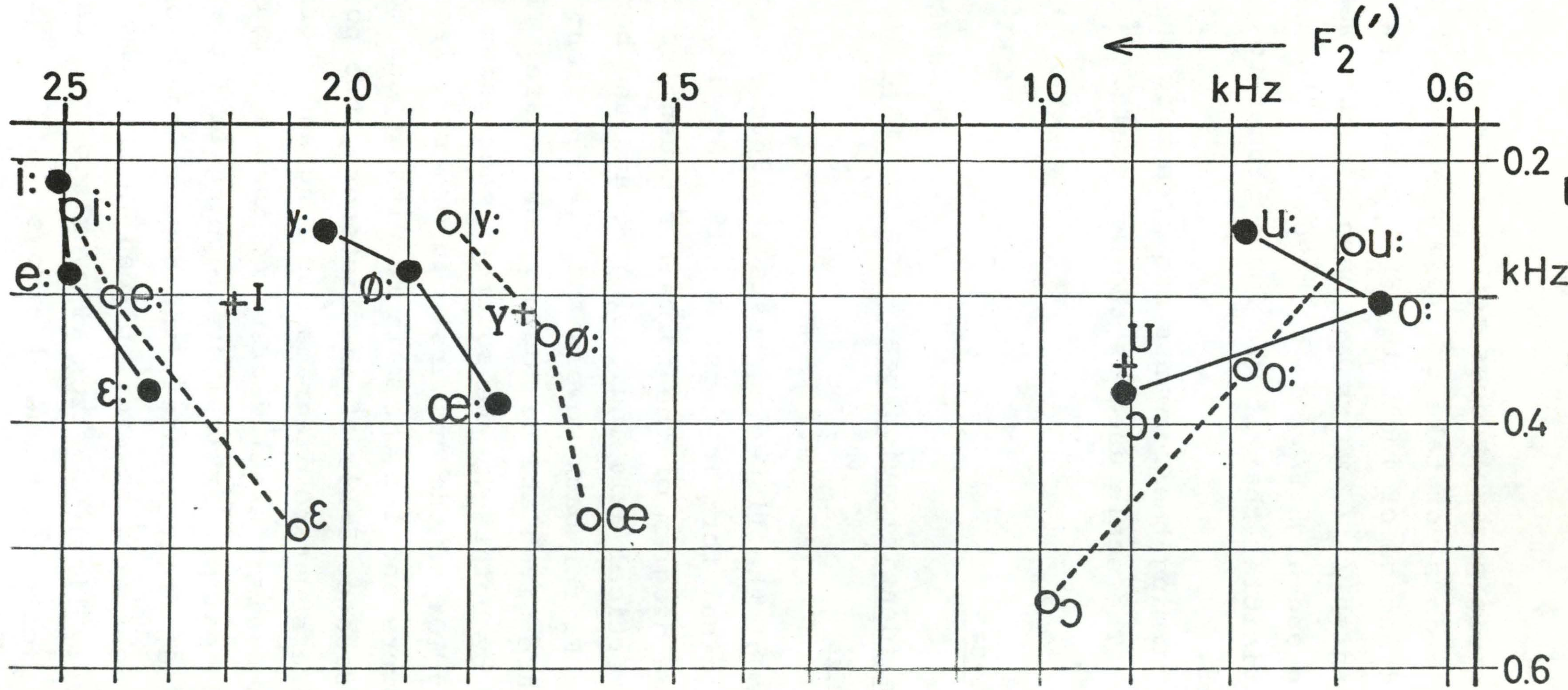


Fig. 2. Average formant frequencies of Danish and German vowels indicating approximately the frequencies of the vowels used as standards by the listeners

●——● The Danish long vowels /i:, y:, u:, e:, ø:, o:, ε:, œ:, ɔ:/ spoken by 8 Danish male speakers.

o- - - -o The German vowels /i:, y:, u:, e:, ø:, o:, ε, œ, ɔ/ spoken by 6 German male speakers.

+ The German vowels /I, Y, U/ for comparison.

Although this formula probably does not correspond precisely to perception (it does not take the formant levels into account), it has the advantage of including F_3, which is particularly important in Danish /i:/, which may have a rather weak and low F_2 and a strong and high F_3. In back vowels F_3 was so weak (and sometimes invisible) that it was found better to leave it out altogether.

This way of diagramming the formants has been used in all subsequent graphs. The scale used is the mel scale.

3.3. Results of the tests

3.3.1. Results of the identification tests for the main stimuli (tables 1 and 2)

The main results of the listening tests in 1963-64 and 1973 are given in graphical form in figs. 3 - 18. At the top of each figure is a diagram of the formant frequencies of the speaker with the vertical axis indicating F_1 and the horizontal axis indicating F_2 for back vowels and F_2- for front vowels. The stimuli are given at the top of each square. The answers are indicated by points in the squares, each point indicating approximately 10% of the answers. In order to leave out quite erratic answers and let the main results appear more clearly, one point indicates 5-14% of the listeners, two points 15-24%, etc. There are separate diagrams for long and short vowels of the two (in Danish three) degrees of tongue height, since there is a good deal of vertical overlapping in the answers. Front unrounded, front rounded, and back rounded vowels of the same tongue height are, however, placed in the same diagram since there is very little horizontal overlapping. Cases of overlapping are indicated by means of arrows. If, for instance, a front unrounded vowel is heard as rounded, the point is placed in the column for front rounded vowels, but an arrow pointing to the left indicates that the stimulus was a front unrounded vowel. Similarly, if a back vowel is heard as a front rounded vowel, the point is placed in the mid column with an arrow pointing to the right.

A. Danish listeners

a. Test 1963-64

The results for the three listening groups have been combined in the graphs (figs. 3 - 7). The percentages were calculated for each group separately, and the average taken of the three percentages. The answers of the three groups show very few divergencies. There is less dispersion in the expert group, which might be expected, but this may also be due to the fact that the number of listeners was smaller and they were allowed to listen twice. As mentioned above only the experts used the peripheral squares for the low vowels. Therefore, we will not treat the answers to these vowels in detail. Moreover, for some reason inexplicable to me the students of Danish have often placed German long /i:, y:, u:/ somewhat lower than the other two groups.

<u>The Danish stimuli</u> were mainly included in order to have a control of the procedure. The answers are given in fig. 3. They show that on the whole the vowels have been identified correctly. Short /i, y, u/ are heard as somewhat lowered compared to long /i:, y:/ and /u:/, which is in agreement with the formant frequencies, and short /o/ has been indicated as <u>ɔ</u>, which is also in agreement with the formants. The only disagreement between the acoustic chart and the answers is that some have heard /ɔ:/ and /ɔ/ as somewhat [o]-like, although they are not raised, but somewhat fronted compared to the average and particularly compared to the

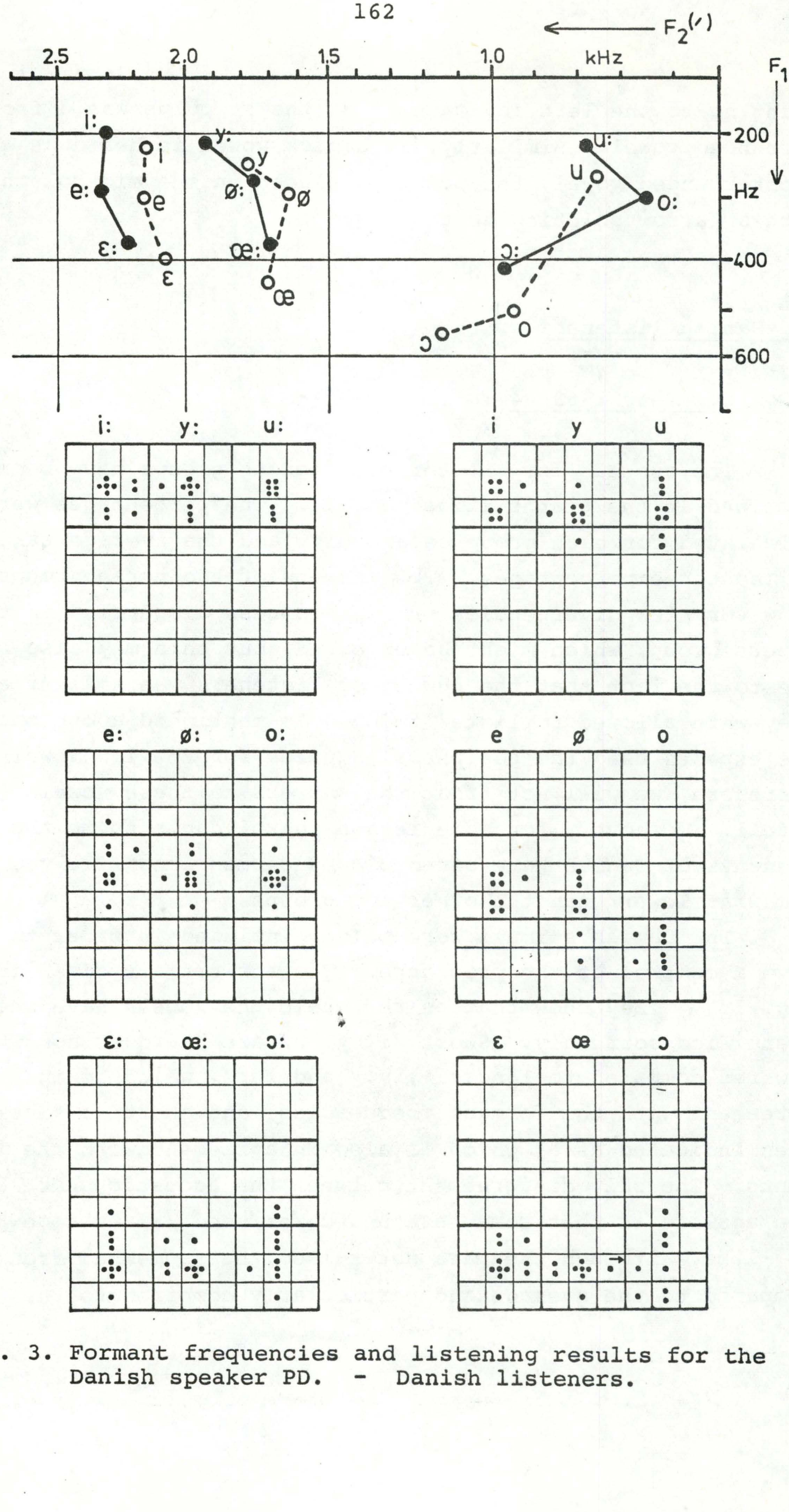


Fig. 3. Formant frequencies and listening results for the Danish speaker PD. - Danish listeners.

pronunciation of the younger generation. This fronting has been indicated by the expert group. (There may be some context effect here since /ɔ/ comes after /œ/, and /ɔ:/ after /y:/.)

<u>German /i:, y:, u:/</u> are generally heard as lower than the corresponding Danish vowels. This is in agreement with the general subjective impression. HL's /i:, y:, u:/ are heard as particularly low, and this is in accordance with his unusually high formant 1. It is also in agreement with the position of formant 1 that HT's /u:/ is heard as lower than his /y:/ and /i:/. Generally, the differences in F_1 are not sufficient to explain the differences in auditory impression of tongue height. The higher formants are also of importance. German /u:/ has a somewhat lower F_2, /y:/ has a lower F_2 and F_3, and /i:/ a lower F_3 than the corresponding Danish sounds. This brings German /u:/ closer to Danish /o:/ (which has a lower F_2 than the somewhat fronted Danish /u:/), and German /y:/ and /i:/ closer to Danish /ø:/ and /e:/ (and the high and strong F_3 in Danish /i:/ is probably still more dominant for the auditory impression than Fant's formula shows). It should not be forgotten that lower vowels have not only a higher F_1, but also a different F_2, lowered in front vowels and raised in back vowels (see also fig. 2.).

<u>German /e:, ø:, o:/</u> are also, except for HT, heard as lower than the corresponding Danish vowels. Here again it is not only due to F_1, but also to the higher formants. German /ø:/ has lower formants 2 and 3, and /e:/ has a lower formant 3, and this brings them closer to Danish /œ:/ and /ɛ:/, and German /o:/ has a higher formant 2, which brings it closer to Danish /ɔ:/ (see fig. 2.). HL's /e:, ø:, o:/ are heard as particularly low in accordance with his high F_1. NB's /ø:/ and /o:/ are also heard as particularly low, but this seems to be

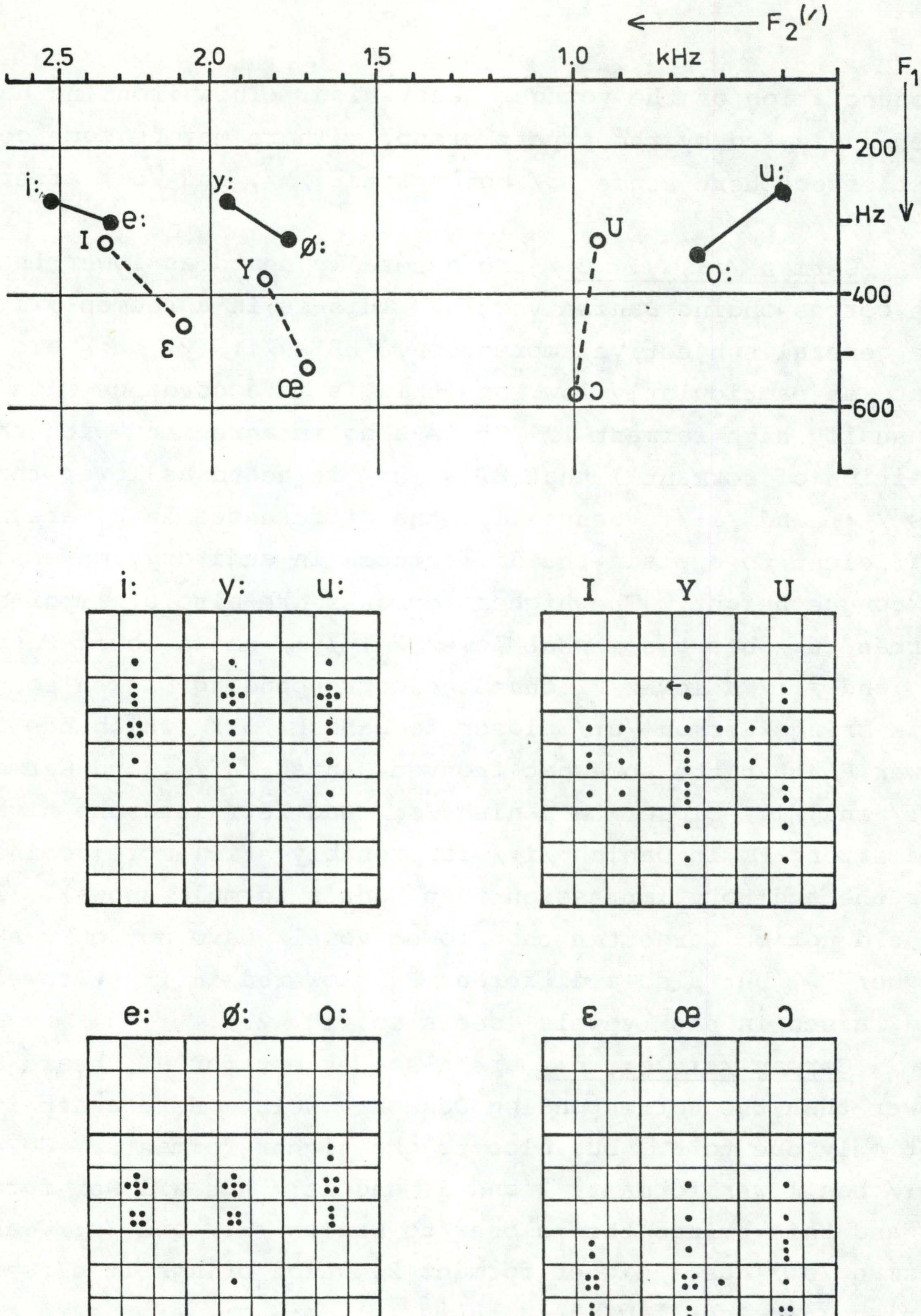


Fig. 4. Formant frequencies and listening results for the German speaker ED. - Danish listeners.

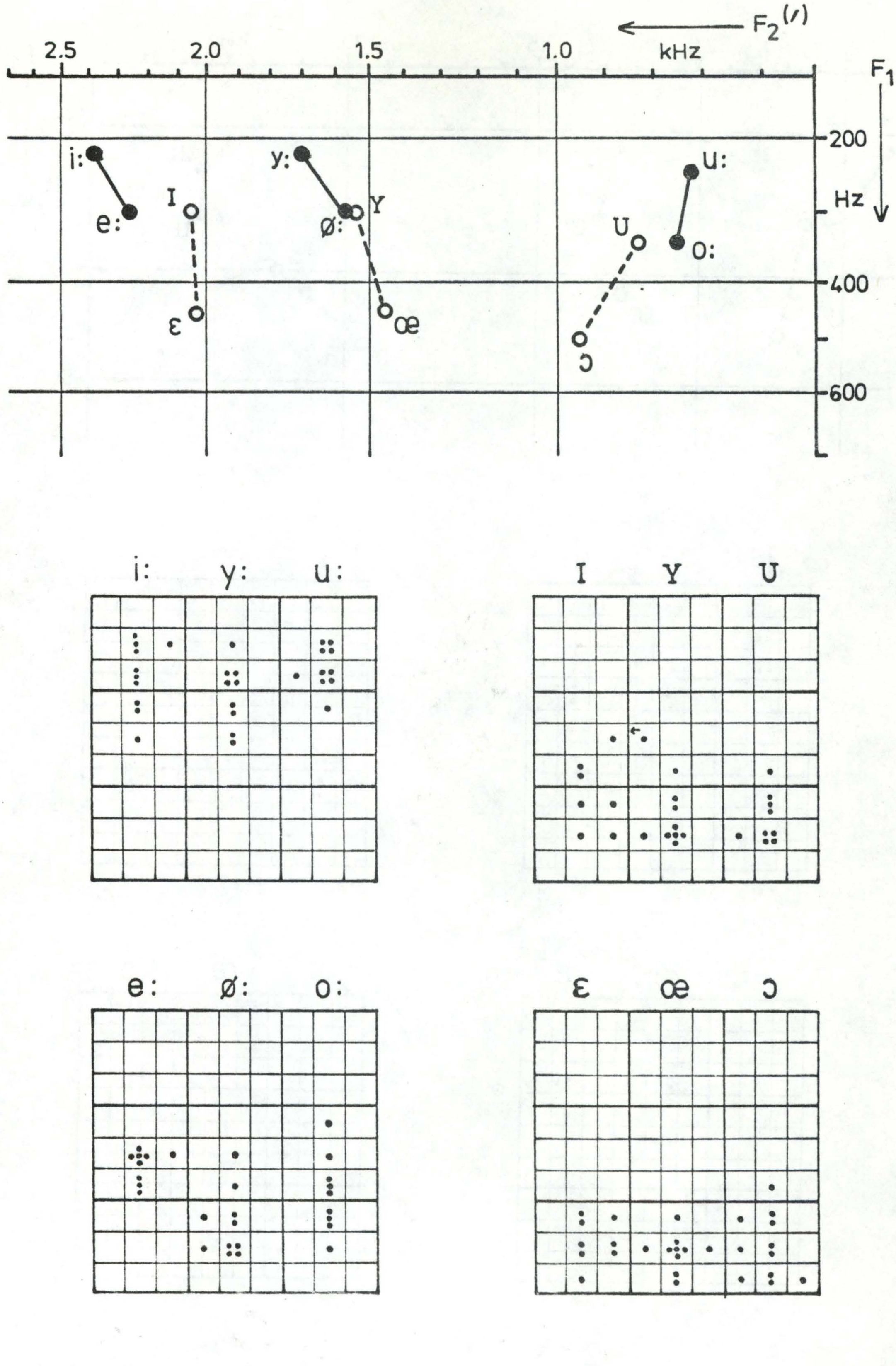


Fig. 5. Formant frequencies and listening results for the German speaker NB. - Danish listeners.

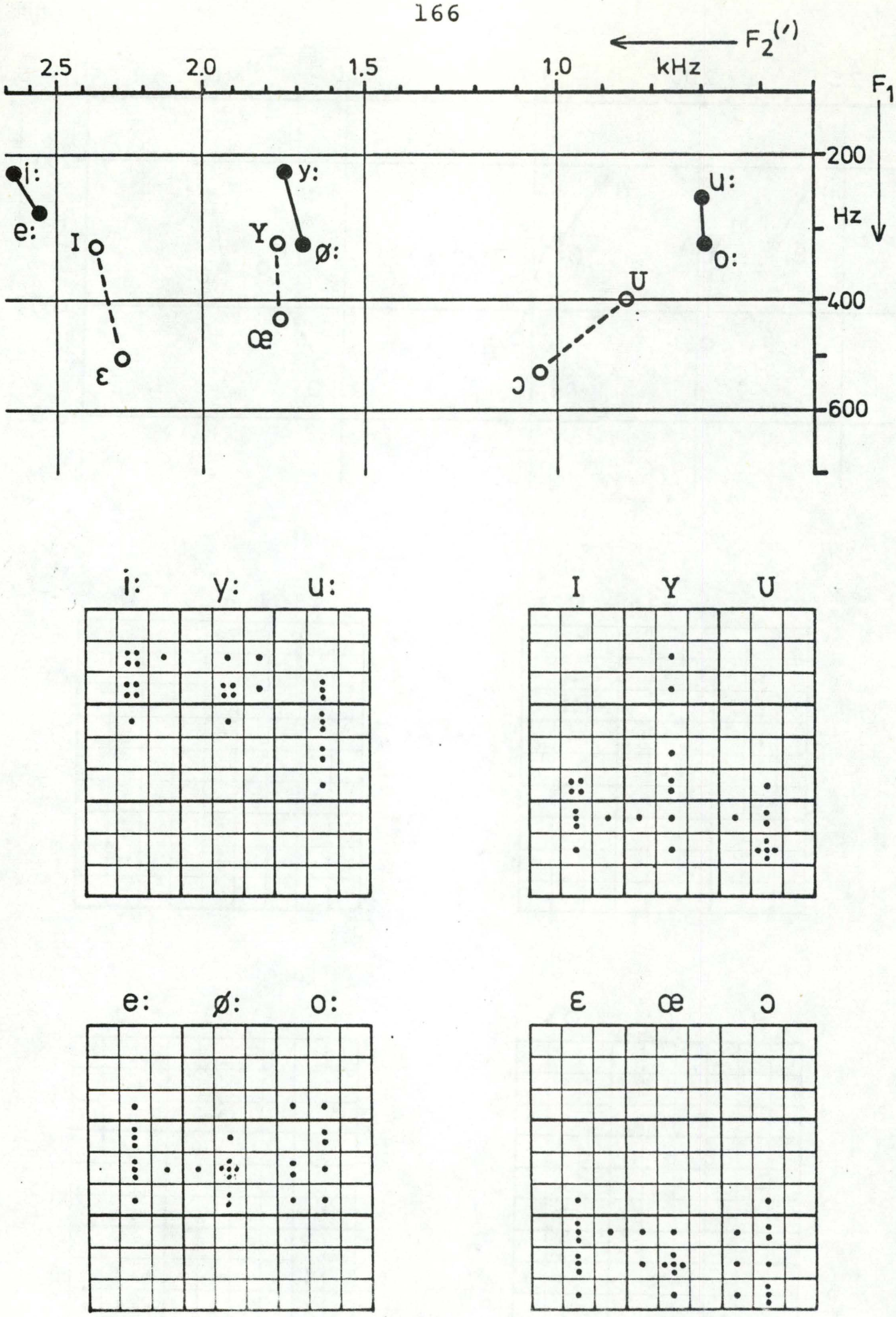


Fig. 6. Formant frequencies and listening results for the German speaker HT. - Danish listeners.

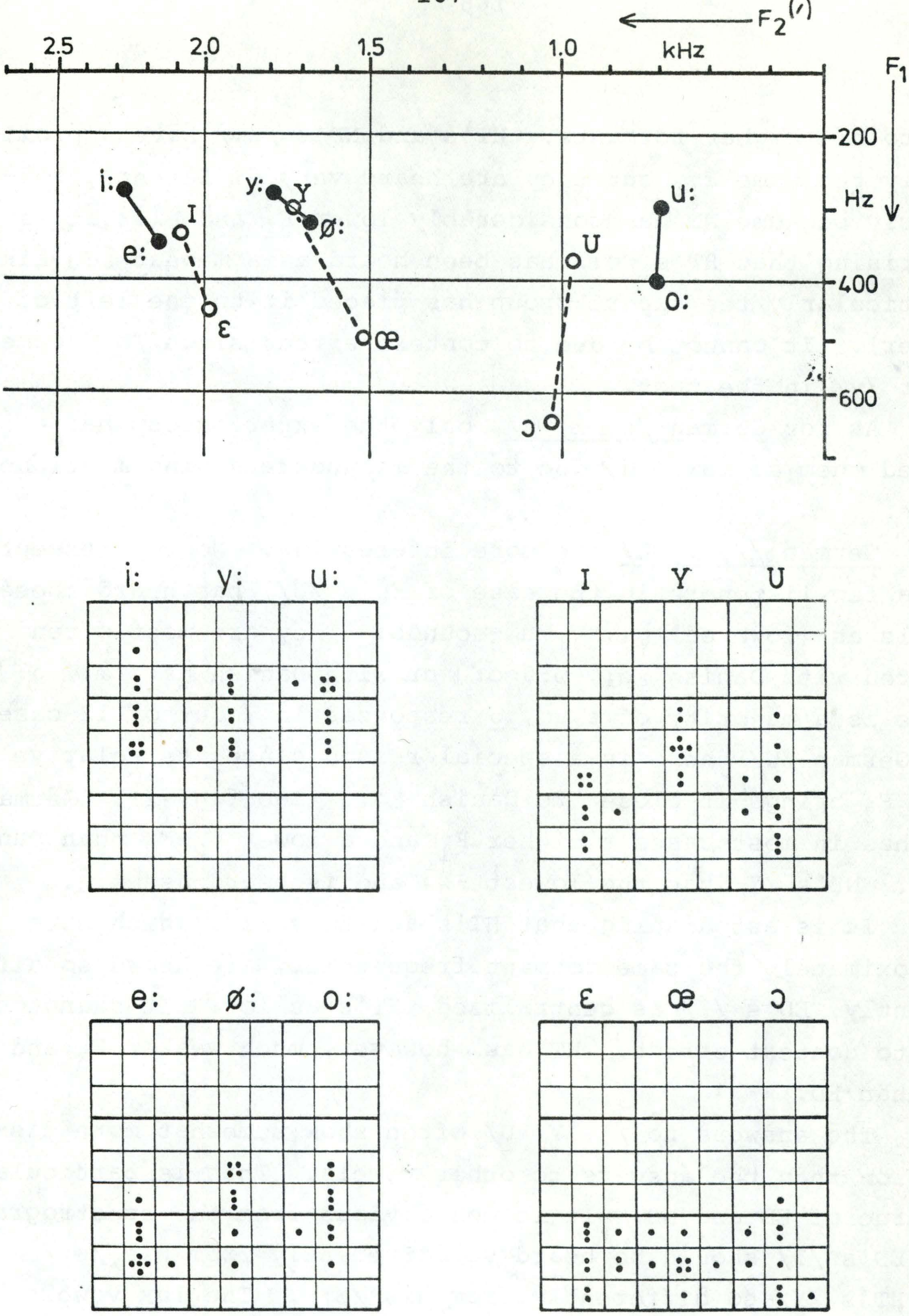


Fig. 7. Formant frequencies and listening results for the German speaker HL. - Danish listeners.

due to his higher formants. HT's and NB's /ø:/ have approximately the same F_1, but they are heard very differently, obviously because NB has considerably lower F_2 and F_3. It is surprising that HT's /o:/ has been heard as somewhat [ø]-like (particularly the expert group has placed it to the left of the center). It cannot be due to context effect since /o:/ comes after /y/ in the test.

As for German /ɛ, œ, ɔ/, only the expert group has placed them as lowered, due to the misunderstanding mentioned above.

German /I, Y, U/ are more interesting. Nobody (except for a few listeners in the case of ED's /U/) has heard these vowels as (lowered) [i, y, u]-sounds. They are most often equated with Danish [e:, ø:, o:] or with Danish [ɛ:, œ:, ɔ:]. There is a majority of ɛ, œ, ɔ responses in 7 out of 12 cases. For German /U/ there is a special reason since its relatively high F_2 brings it close to Danish [ɔ:] (see fig. 2). German /Y/ has in most cases a higher F_1 and a lower F_2-F_3 than Danish /ø:/. NB's /Y/ has the lowest F_2- and is heard as [œ:].

It is astonishing that HT's and ED's /I/, which have approximately the same formant frequencies, are heard so differently, ED's /I/ as centralized, HT's as low. It cannot be due to context effect. HT has, however, much weaker F_3 and F_4 than ED.

The answers to /I, Y, U/ often show somewhat more dispersion than the answers to other vowels. This is particularly true of ED and NB. It is not obvious from the spectrograms why ED's /I/ should be heard so differently from his /e:/ and HT's /Y/ so differently from his /ø:/. The lax vowels seem to make a less neat impression. But, on the whole, the differences in perception between /e:, ø:, o:/ and /I, Y, U/ for the different listeners can be explained from the formant frequencies.

b. Test 1955

The results of the 1955-test are in agreement with the results of the 1963-test. The seven Danish vowel segments were placed correctly. Short /i/ and /e/ were not placed lower than long /i:/ and /e:/, which is in accordance with the formants of the speaker.

German tense vowels were not heard as lowered compared to the Danish vowels, with the exception of the vowels in ED lieben (as in 1963), GR Höhle, and CH fühlen. But most of the examples were spoken by GR and CH, who both have unusually high higher formants in these vowels.

Of the lax /I, Y, U/, ED's vowel in immer and CH's in killen were heard as Danish [e], but GR killen as an [ø]-like [e]. The vowels of CH Kütte and GR Hülle were heard as [ø] and lowered [ø], respectively, GR Bulle was heard as [o], and Tunnel, with some dispersion, as [u] or [o].

The vowel of English did was heard as lowered [e] or raised [ε]; the response to Dutch tinnen were spread over [e], [ø] and [œ] (whereas the tense vowel in denen was heard as a clear [e]).

It should be noticed that the tendency found in Holtse's test with synthetic vowels (see his article in this volume) and also in Rischel's unpublished test with synthetic vowels, viz. to perceive vowels as higher than would be expected according to known formant averages, is not found in the present tests.

B. German listeners

The answers from Kiel and Köln have been combined into the same graphs (figs. 8 - 12). In this case the average has been taken of all answers as one group. The Köln results were divided into experts (6), and students (17), but the Kiel results (10 persons) were not, although the group contained several phoneticians. A division into groups would have given

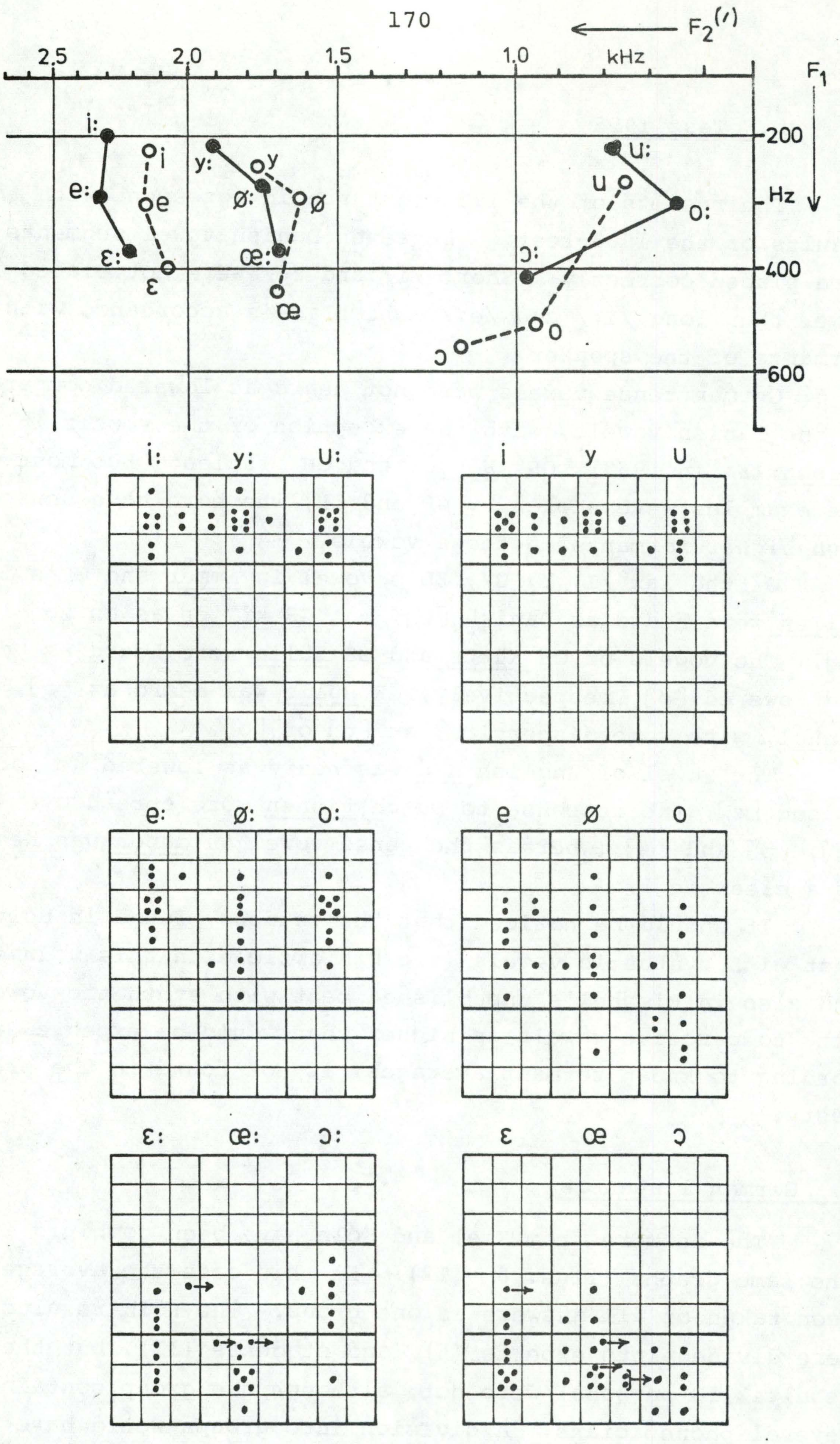
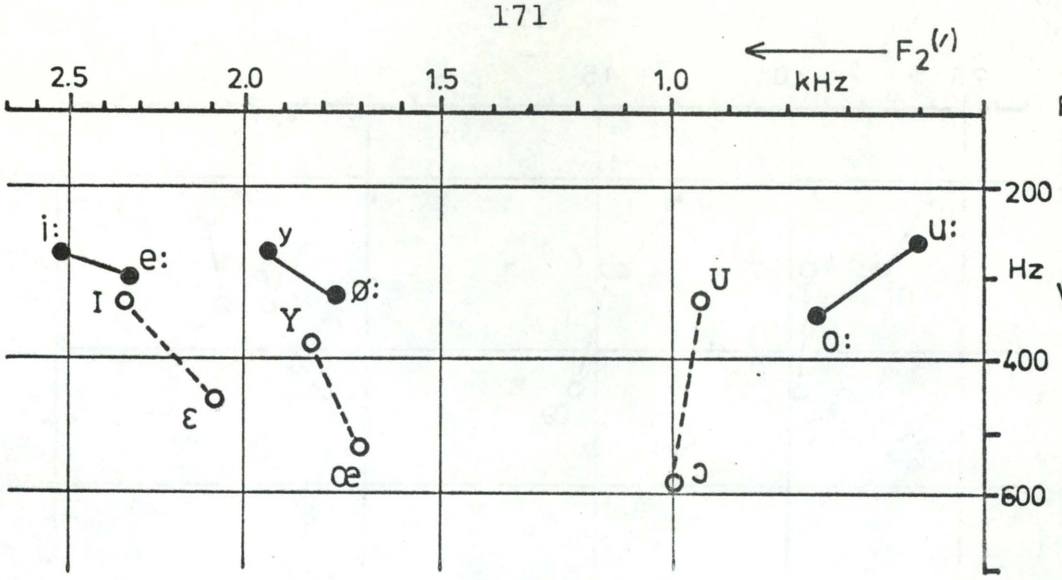


Fig. 8. Formant frequencies and listening results for the Danish speaker PD. - German listeners.



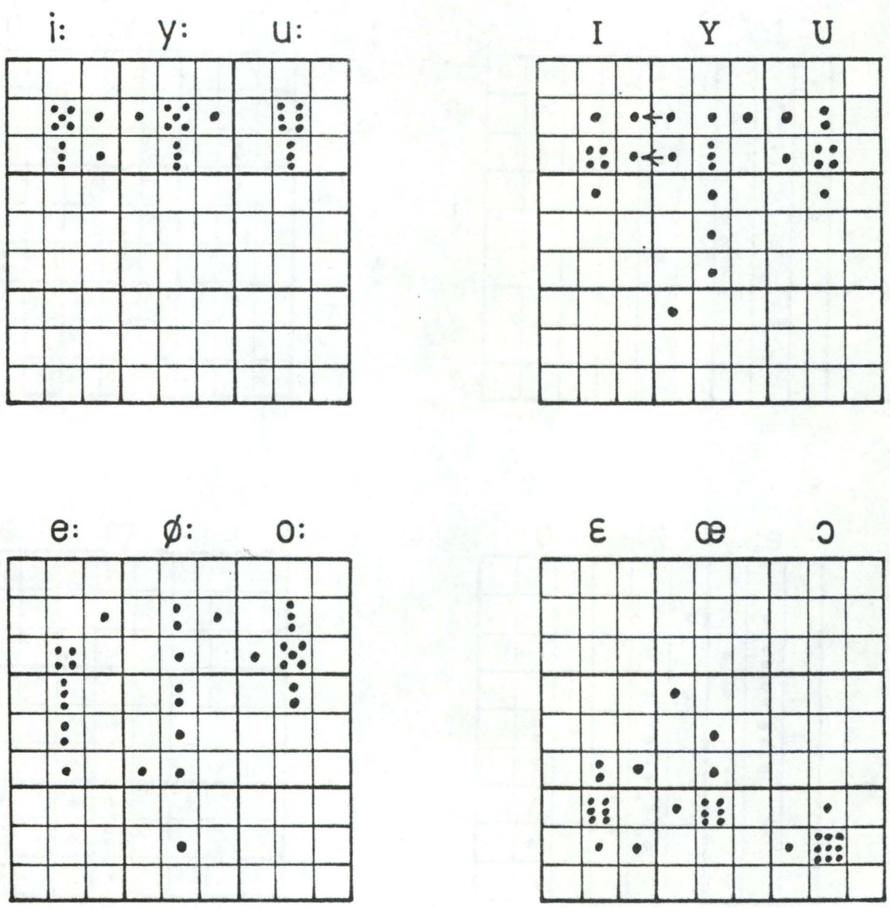


Fig. 9. Formant frequencies and listening results for the German speaker ED. - German listeners.

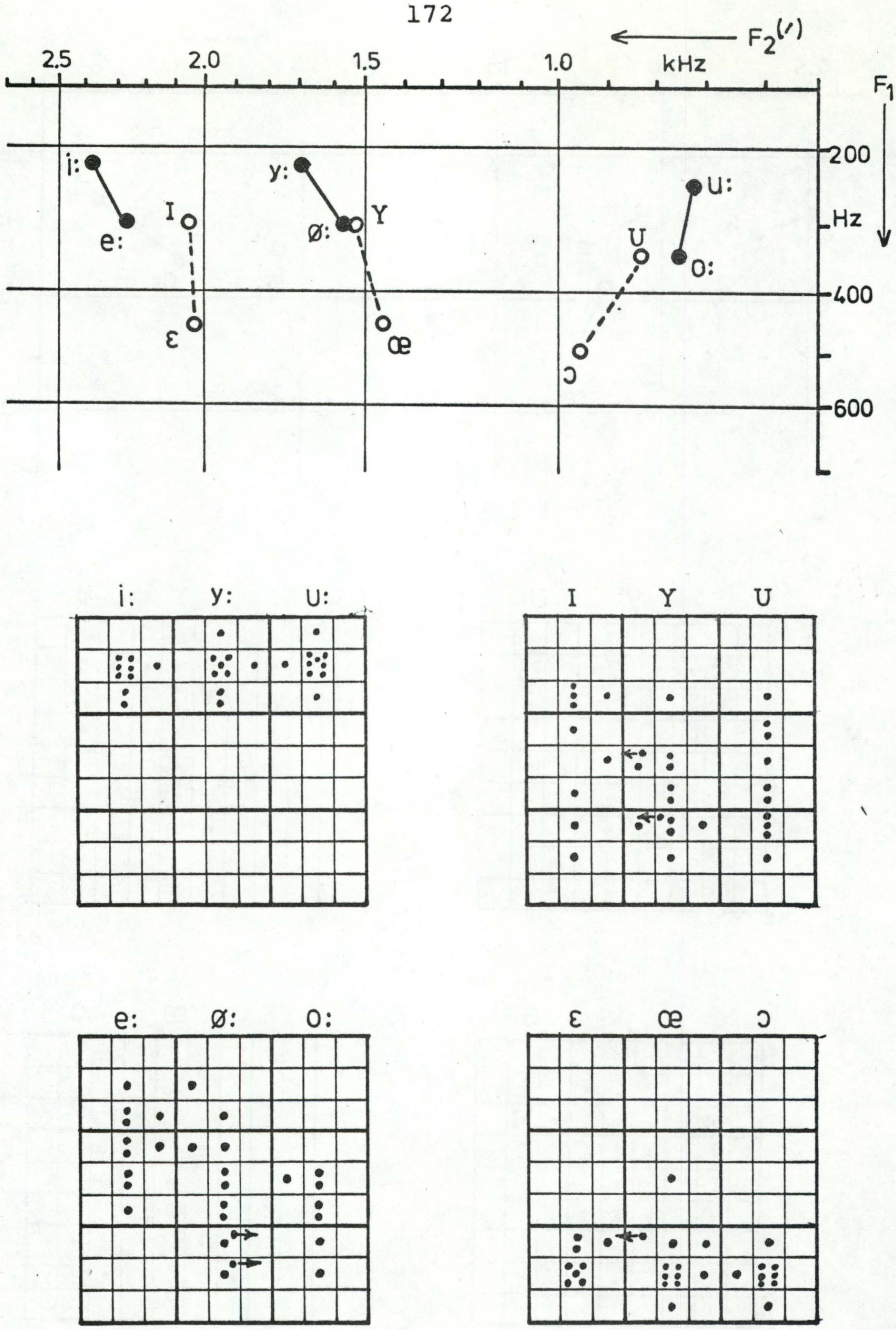


Fig. 10. Formant frequencies and listening results for the German speaker NB. - German listeners.

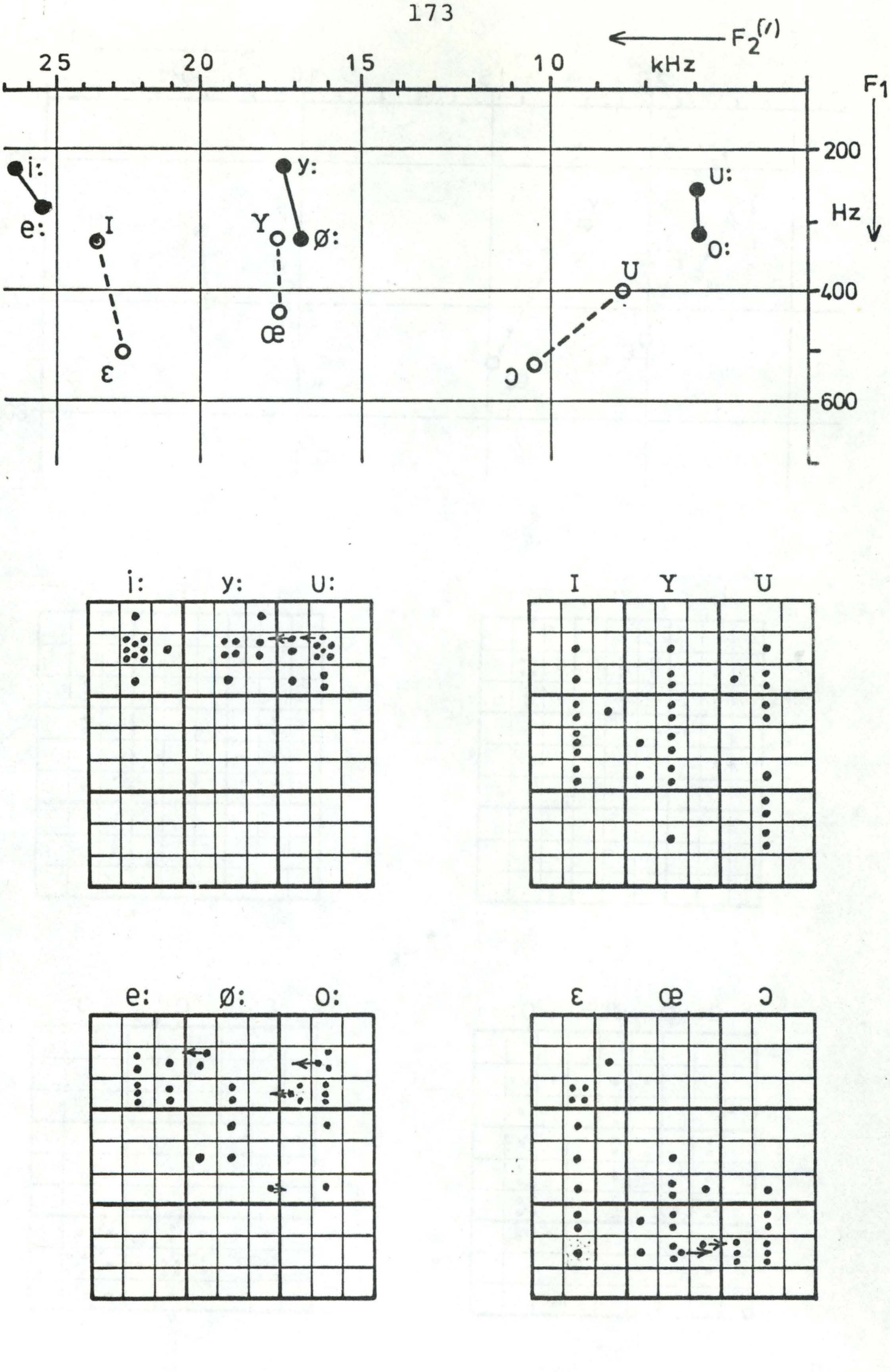


Fig. 11. Formant frequencies and listening results for the German speaker HT. - German listeners.

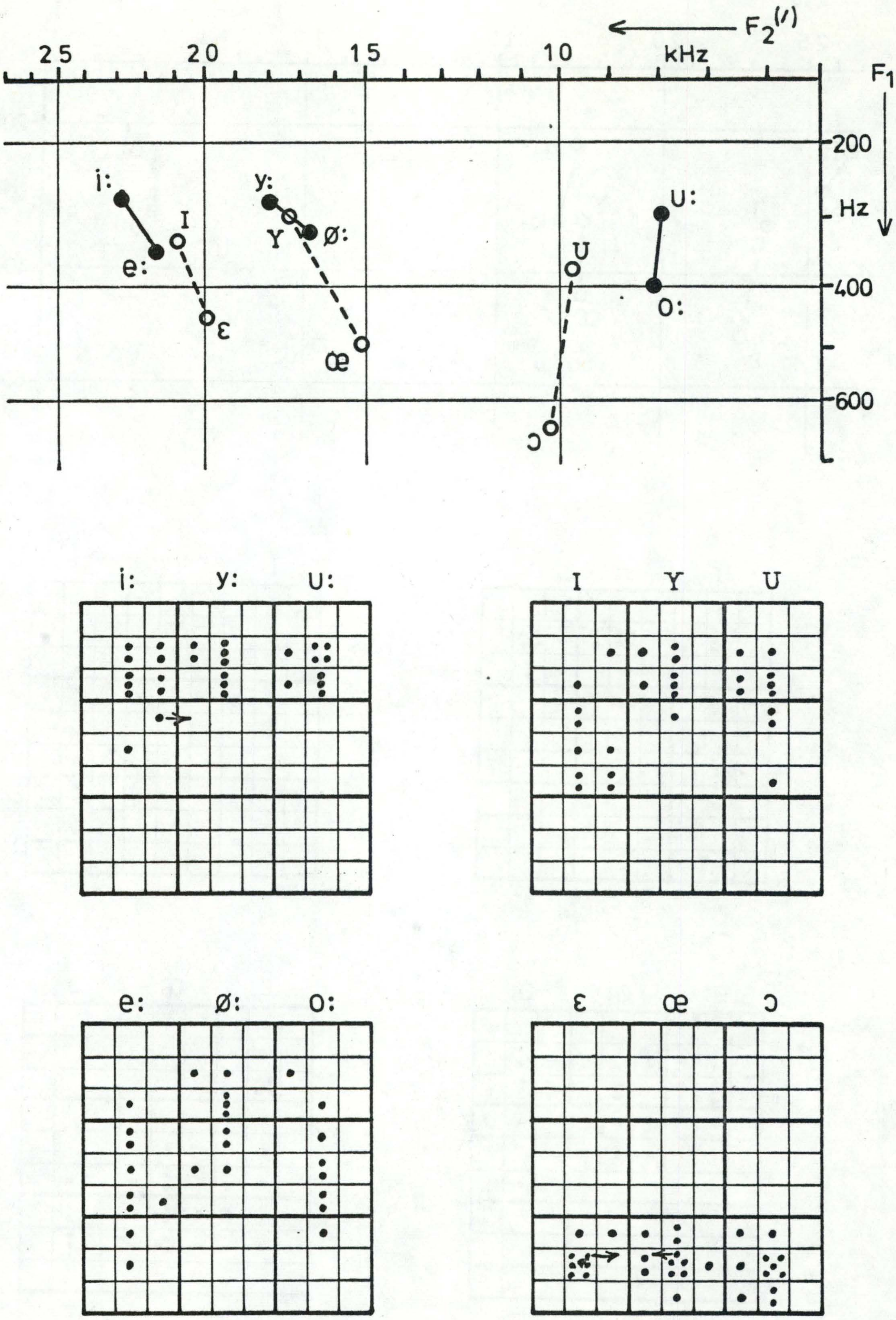


Fig. 12. Formant frequencies and listening results for the German speaker HL. - German listeners.

more weight to the expert answers than for the Danish groups. Moreover, it was easier to treat them as one group. It could also be done because there were no consistent differences between the groups.

There is, on the whole, more dispersion among the German listeners than among the Danish listeners, and the inclusion of the answers from Hamburg in the graphs would have made this still more apparent. There is even more dispersion than can be seen on the graphs, since percentages below 5% have not been included. In some cases with three single answers in the same direction from the central square, but in different subsquares, one point has, however, been put in the middle in order to give a better survey of the distribution. If there are many subsquares with only two answers, i.e. 6% of the listeners (which is just above the limit for one point), the number of points may exceed 10.

There is also more horizontal overlapping than in the Danish answers, so that more arrows have been necessary. This uncertainty may be due partly to the fact that the tape was 10 years old, and that it was played back on a different type of tape recorder. Some high frequencies may therefore have been lost, and the signal/noise level may have been less favorable, although the listening took place via earphones.

The answers to long /i:, y:, u:/ are, however, quite clear. The short /i, y, u/ of the Danish speaker and the long /i:, y:, u:/ of the German speakers, which were heard as lowered by the Danish listeners, are only heard as very slightly lowered by the German listeners (with the exception of HL, who has a very high F_1). There is a marked difference between the German and Danish listeners on this point, which supports the impression of a small difference between Danish and German high vowels, mainly depending on the higher formants.

The answers to /ɛ, œ, ɔ/ present few problems. German /ɛ, œ, ɔ/ are generally placed in the expected squares. It is not clear why ED's /œ/ is heard as [ø]-like (this was also done, to some extent, by the Danish listeners). That HT's /ɛ/ is heard as raised, and even in some cases as [i], by more than one third of the listeners must be due to his extraordinarily high F_2-F_3, which are even a good deal higher than the higher formants of NB's and HL's /I/.

It is understandable that the Danish low vowels are heard as raised. One might even have expected a more pronounced displacement. The fact that Danish /ɔ:/ and /ɔ/ are often heard as front rounded vowels can be explained by their relatively high F_2. When all single answers are taken into account, /ɔ:/ is heard as a front vowel by 24% and /ɔ/ by 39% of the German listeners (the corresponding percentages for Danish listeners are 3% and 17%).

The real problems concern the answers to /e:, ø:, o:/ and /I, Y, U/.

It is not astonishing that Danish /e:, ø:, o:/ are heard as raised, and /e:, o:/ even predominantly as /i/ and /u/. This corresponds well with the opposite tendencies for the perception of German sounds by Danish listeners and can be explained from the different formant positions. But it is astonishing that German /e:, ø:, o:/ have often been heard as [i, y, u]. These are even the predominant answers to HT's and ED's /e:/ and to HL's /ø:/. HT has very high F_2-F_3 for /e:/, and the Danish listeners have also heard his /e:/ as somewhat raised, but there are no obvious reasons in the other cases. It also happens that German /e:, ø:, o:/ are heard as open vowels, e.g. HL's /e:/ and NB's /ø:/. They have relatively low F_2-F_3. But the general dispersion over all three vertical fields compared to the concentrated responses of Danish listeners to

Danish /e:, ø:, o:/, and even, though less pronounced, to the corresponding German vowels, is surprising.

The explanation is probably that, since German long /e:, ø:, o:/ are qualitatively very close to /I, Y, U/, some listeners will spontaneously identify a short vowel of this quality with /I, Y, U/; it is therefore difficult for them to make a purely qualitative comparison with long vowels. And as /I, Y, U/ are the highest vowels of the lax vowel category and are written i, ü, u, they will be inclined to transcribe the stimuli as i, y, u and forget that these symbols were meant to represent the quality of long /i:, y:, u:/.

In this connection it is of interest that Bennett (1965) found that a synthetic vowel which is heard by German listeners as [o:] when long, is heard as [u] when shortened. And Lindner (1966) found that synthetic vowels heard as [e:, ø:, o:] when long, were heard as [i, y, u] when short. One might assume a general tendency to perceive shorter vowels as higher, but the results for the Danish listeners show that this is not the case. The different results for the German listeners must be due to the phonemic and orthographic system of German.[1]

The same difficulty arises, of course, for the responses to /I, Y, U/. The dispersion for these sounds is somewhat greater. In some cases, e.g. ED's /I/ and /Y/, there are more i, y-responses than to /e:/ and /ø:/. This does not seem to be justified by the formant frequencies; and, except for one case (HL /Y/), the lax vowels /I, Y, U/ are always acoustically closer to /e:, ø:, o:/ than to /i:, y:, u:/. It might be due to a specific quality of lax vowels (cp. that ED's /I/ is heard as centralized). But HT's /e:/ and /I/ show the opposite tendency: there are more e-responses to his /I/ and more i-responses to his /e:/, which might be explained by the very high F_2 of his /e:/. His /U/ is mostly heard as [ɔ], probably because of the high F_1.

---

1) see postscriptum, p. 194

In any case, there are (just as for /e:, ø:, o:/) a good many i, y, u-responses, which are not justified by the formant positions.

The results from Hamburg have not been included in the graphs because the somewhat larger dispersion would make the picture more blurred. The greater dispersion is partly due to the fact that both German /e:, ø:, o:/ and /I, Y, U/ are more often identified with /ɛ, œ, ɔ/ than in the other groups, so that the answers are spread over the whole vertical scale, partly to the fact that in a few cases there is more horizontal overlapping (e.g. in HT's /ɛ, œ, ɔ/). The Hamburg listeners have, just as the other groups, been inclined to identify the short stimuli of /I, Y, U/-quality with short vowels, but some have apparently found them too open to be [i, y, u] and have chosen [ɛ, œ, ɔ].[1] Apart from this, the answers look very much like those of the other two groups. Almost all the listeners were from North Germany.

The very sharp cuts may have reinforced the impression of short vowels. Cohen et al. (1963) report that an abrupt decay favours short-vowel-responses for Dutch listeners. On the other hand, measurements of German vowels (EFJ 1941, Hans Peter Jørgensen 1969, EFJ and H.P. Jørgensen 1970) have not supported Sievers' old theory that the decay of short vowels should be more abrupt than that of long vowels in German.

Five of the listeners from Köln repeated the test later. The most consistent listener gave exactly the same response the second time in 53% of the cases, the least consistent in 21% of the cases, i.e. in most cases there was a change, but, generally, it was only one step in the diagram. This means, however, that one should not try to explain very small differences.

1) Why this identification has been made especially by the listeners in Hamburg is difficult to explain. It may perhaps have something to do with the symbols used. On the answer sheets used in Hamburg, [ɛ œ ɔ] were written ĕ, ö̆, ŏ, and in the key words they were spelt e, ö, o (Betten, öffnen, hocken).

3.3.2. Comparison between perception of vowels in isolation and in words in various consonant surroundings

The German listeners in Kiel, Köln, and Hamburg listened to a small subsidiary test containing some of the same vowels in words. The task was exactly the same as for isolated vowels, but the result was very different.

The responses to HT hissen, hütten, hupfen are shown in fig. 13, and the responses to the corresponding isolated vowels by German and Danish listeners are presented again for comparison. It appears from the graphs that whereas /I, Y, U/ in isolation are often heard as [e, ø, o] and sometimes even as [ɛ, œ, ɔ] by the German listeners (and generally as somewhat lower by the Danish listeners whose standard vowels are somewhat higher), they are identified almost exclusively as [i, y, u], although somewhat lowered, when heard in words.

The corresponding vowels spoken by NB give a quite similar picture: the vowels in words were heard as higher. Moreover, whereas 24% of the German listeners heard the vowel in hissen as a front rounded vowel when presented in isolation, nobody made this identification when the vowel was presented in the word hissen. For HL the differences are less drastic since his /Y/ and /U/ are perceived as rather high, also in isolation.

In these cases the transitions were of a very small extent since the front vowels were found between /h/ and an alveolar consonant and the back vowels between /h/ and a labial consonant. The main reason for the different responses to the vowels in isolation and in words is no doubt that isolated segments are heard by most listeners as sounds, whose quality can be determined freely, whereas the same vowels spoken in words are identified immediately with the phonemes /I, Y, U/,

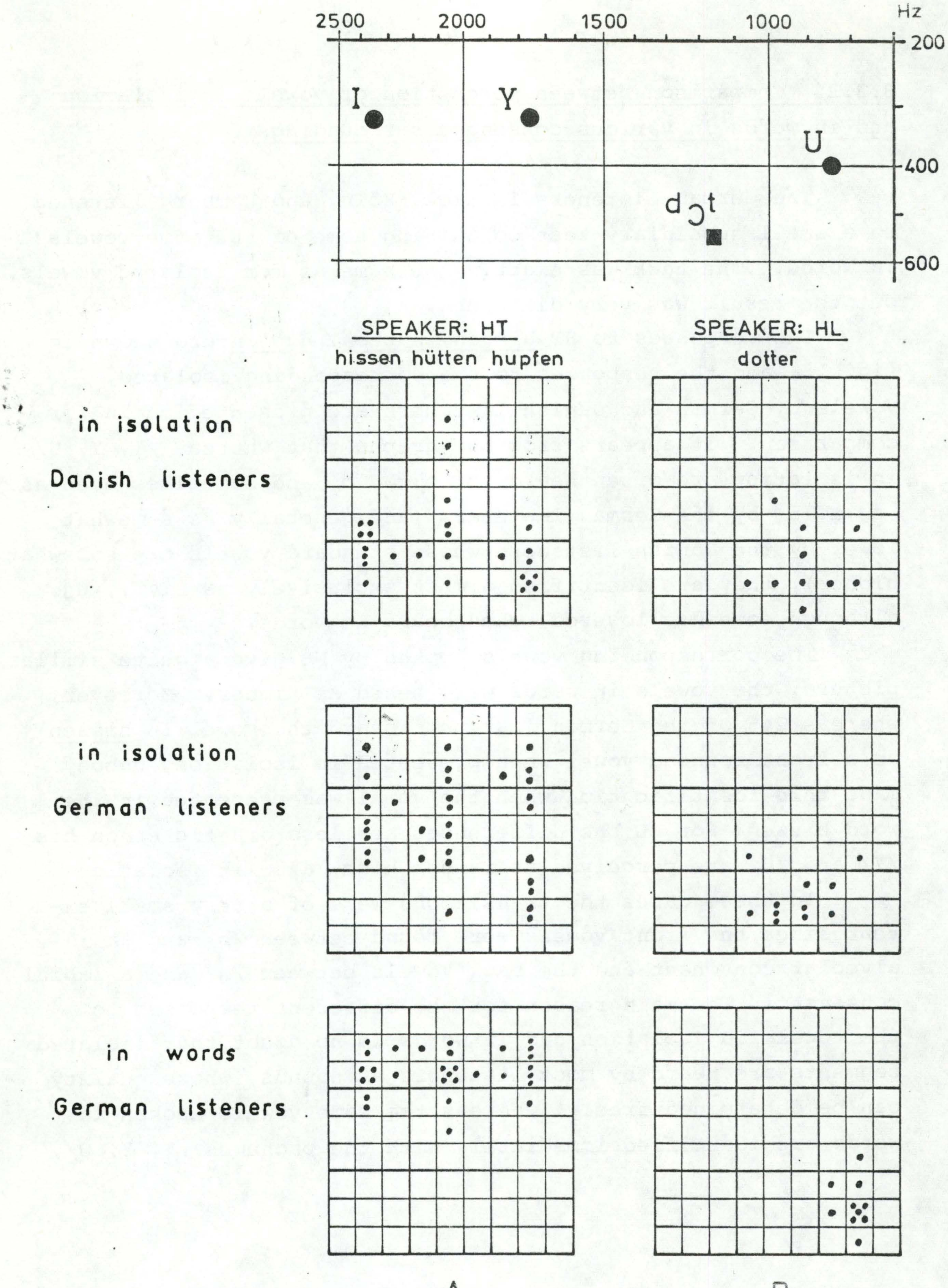


Fig. 13. Comparison between responses to vowels presented in isolation and in words.

i.e., they are identified in a categorical way, and it is difficult to abstract from this identification and listen to them as sounds.

The following graphs (figs. 13 - 16) give some illustrations of the perception of vowels in various consonant surroundings. Fig. 13 gives a picture of the perception of the vowel of Dotter spoken by HL and presented in isolation and in the word. When listening to this vowel in isolation, 69% of the German listeners and 77% of the Danish listeners identify it as a rounded front vowel because of its high F_2. When heard in the word, none of the German listeners hear it as a front vowel.

The formants of HL's Dotter are almost exactly the same as those in PD's /ɔ/, which was heard as a front rounded vowel by 39% of the German listeners, but not by the Danish listeners. The difference may partly be due to context effects. As mentioned above, PD's /ɔ/ comes after an /œ/; HL's /ɔ/, on the other hand, comes after an /U/ in the test.

Fig. 14 gives a picture of the perception of German /I/ in labial surroundings. Heard in isolation the vowel of NB bibbern is identified as a front rounded vowel by 35% of the Danish and 33% of the German listeners. In the word, none of the German listeners hears it as rounded. Presented in isolation, the vowel of ED Lippen is identified as a front rounded vowel by 78% of the Danish and 91% of the German listeners.

Lippen and bibbern have a lower F_2 than other words with /I/, close to Danish /ø:/, but the formant frequencies do not explain the difference in the results for NB and ED. It should, however, be noticed that ED's vowel has a higher fundamental.

In figs. 15 and 16 the responses to the vowels in hupfen and Dutzend, spoken by NB and HL, can be compared. NB's vowel

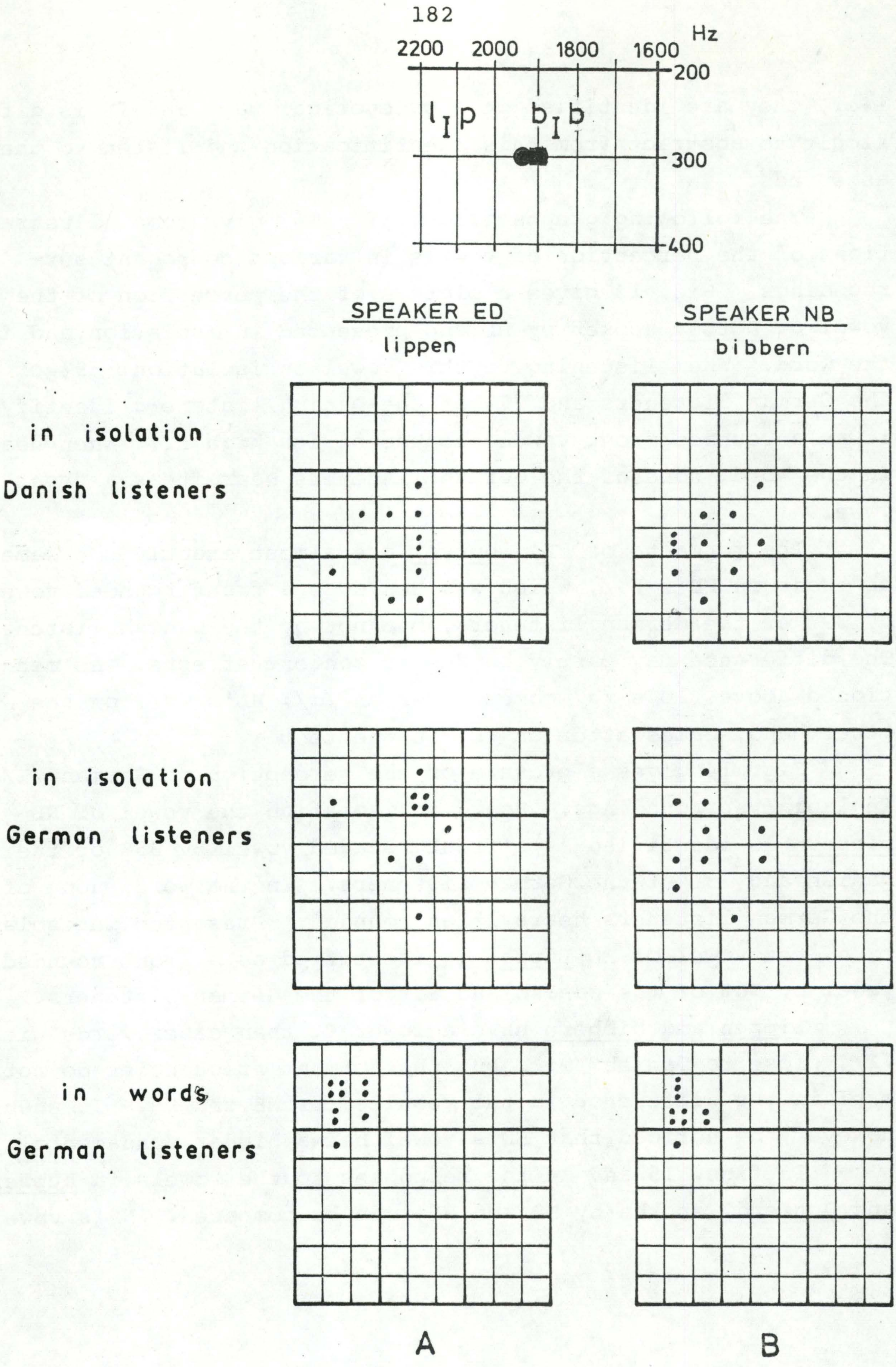


Fig. 14. Comparison between responses to vowels presented in isolation and in words.

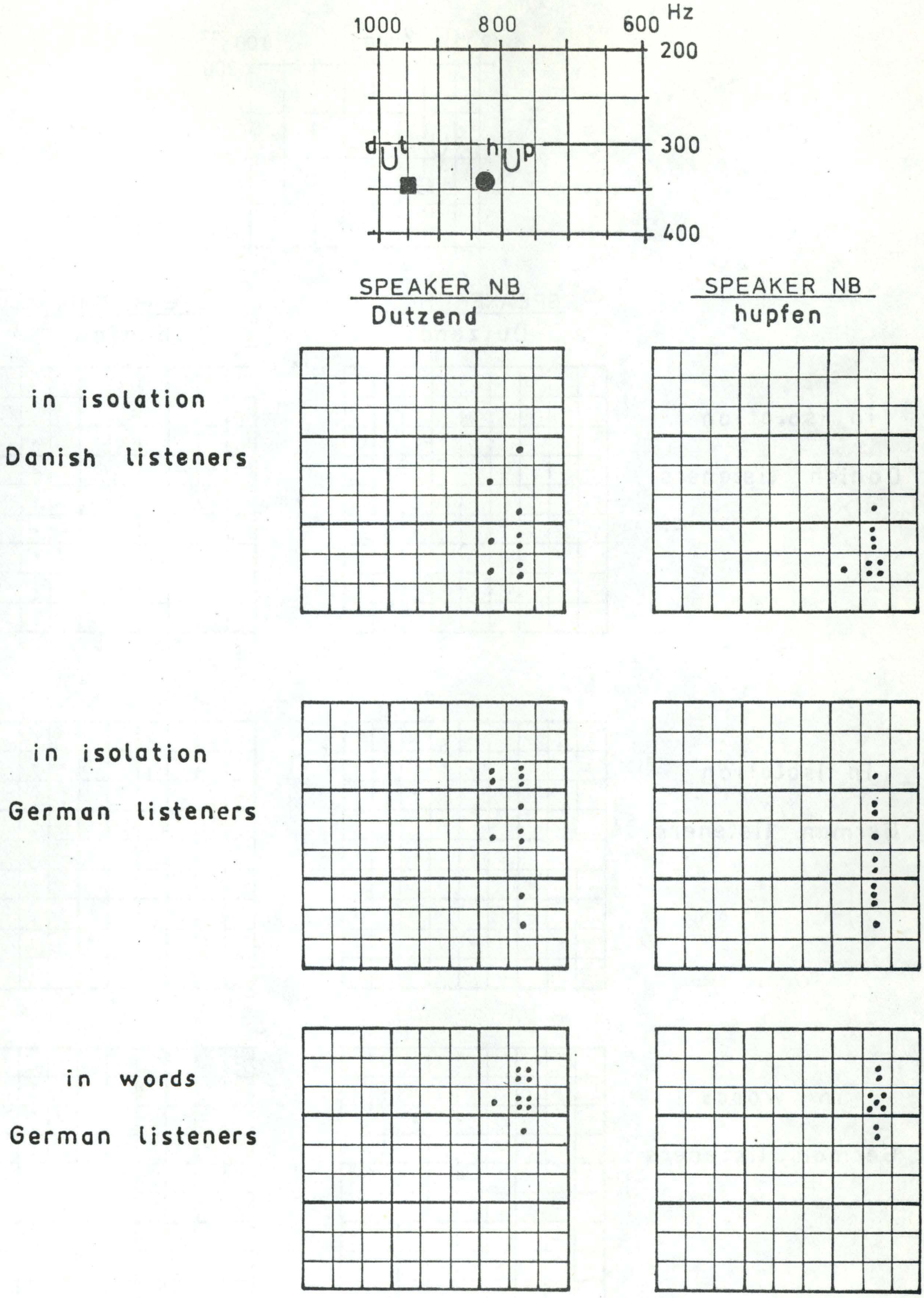


Fig. 15. Comparison between responses to vowels presented in isolation and in words.

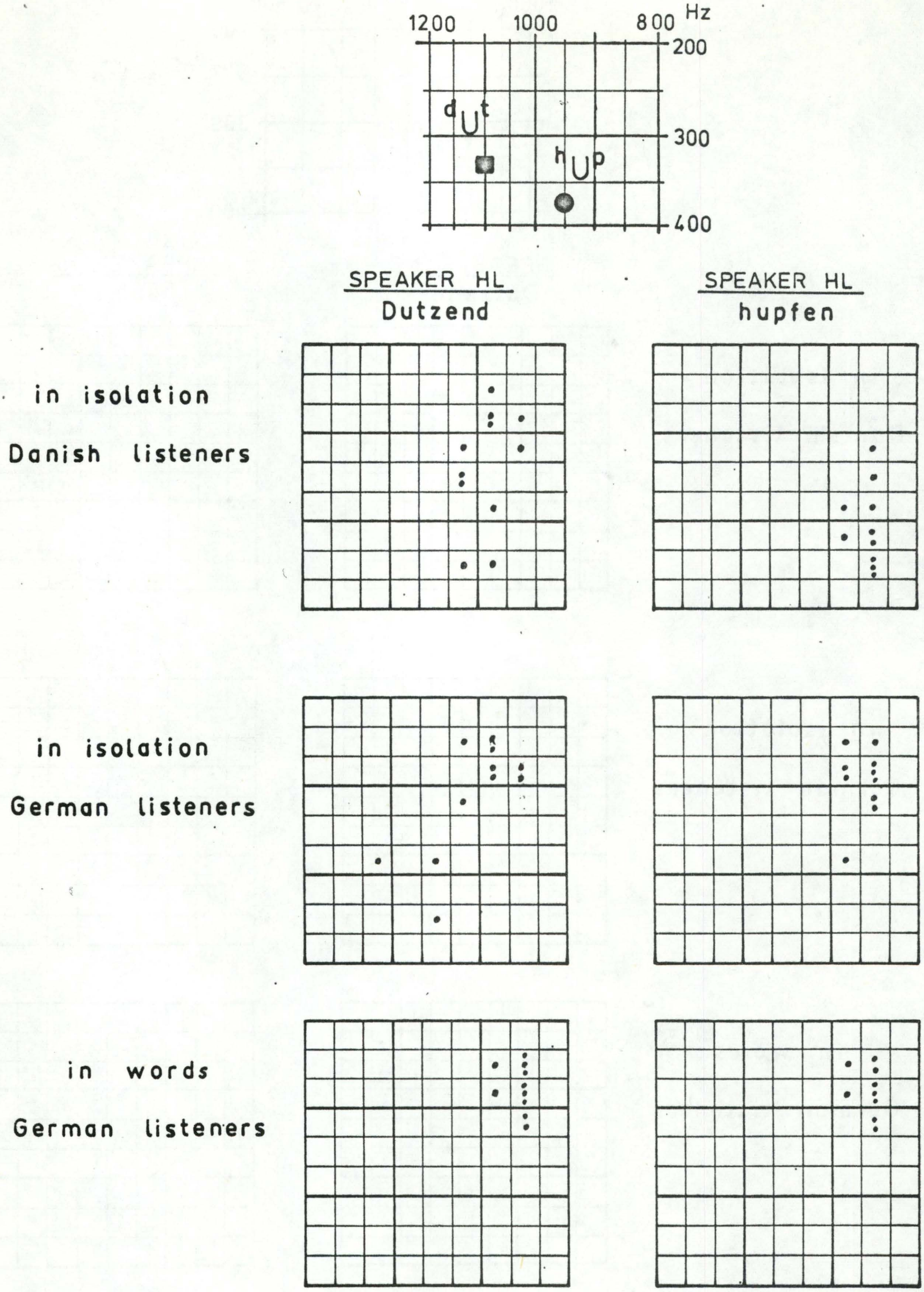


Fig. 16. Comparison between vowels presented in isolation and in words.

in Dutzend is not much fronted. It is rather so that his F_2 in hupfen is somewhat lower than normal, which makes the vowel resemble German [o:], whereas the vowel in Dutzend is identified as [u] by two thirds of the German listeners. The vowel in Dutzend comes close to Danish [ɔ:], and that of hupfen is in between [ɔ:] and [o:], and the listening results are in agreement with this.

HL has a much more fronted vowel in Dutzend with a high F_2. In isolation it is heard as a front rounded vowel by 37% of the Danish and 30% of the German listeners. In words this is never the case.

The results from Hamburg show the same differences between vowels in isolation and in words. The difference is even greater, since 60% of the Hamburg listeners heard HT's /I, Y, U/ as [ε, œ, ɔ], when isolated, and 90% as [i, y, u] in words. As for the vowels of Dutzend and Dotter, Lippe and bibbern, they were only presented in isolation. 45% heard a front rounded vowel in HL Dutzend and 84% in Dotter. The vowel of Lippe was heard as a front rounded vowel by 84% and that in bibbern by 16%.

The 1955-test contained some vowels of the same type. The vowel of ED Lippen was also used in this test. It was heard as [ø] (84% rounded front vowel). GR Lippe was heard as [œ], [u] in Tunnel was not heard as fronted.

The vowel of American English did was heard as [e] or [ε]; but the vowel of pip was heard as [œ] (83%), and the same was true of the /U/ of British English soot for most listeners (70%).

The different reaction to the vowels of hissen, hütten, hupfen in isolation and in words could be explained as the result of a phonetic vs. a phonemic perception. As /I, Y, U/ are the highest vowels of the lax category and are written i, ü, u, they are heard as (lowered) [i, y, u] in words, although they are approximately of the same quality as [e:, ø:, o:]. This explanation is hardly sufficient in the cases Lippe, Dutzend, Dotter. Here the vowels differ from normal /I, U, ɔ/ in quality, and this difference should be noticeable, particularly to phoneticians, but when the vowels occur in words nobody hears these vowels as front rounded vowels, and only about 20 % hear a fronted variety of [U] and a retracted variety of [I]. But here a second factor comes into play: the change in F_2 is due to an automatic coarticulation feature, the vowel is influenced by the surrounding consonants, and there are also pronounced transitions to these consonants. The special features of the vowel are probably perceived as due to the consonants, i.e. the change of formant frequencies is compensated for in the perception.

It is not possible to decide on the relative importance of the two mechanisms (the categorical perception and the compensation for the influence of surrounding consonants) on the basis of the present experiments. For that purpose the tests should have included vowels in consonant surroundings in nonsense words besides the real words. But the reaction to hissen, hütten, hupfen shows at any rate that the categorization effect is strong. That the compensation effect is also at work appears from the experiments made by Lindblom and Studdert-Kennedy (1967). They placed synthetic [I] and [U]-vowels in symmetrical [w-w] and [j-j] surroundings and found a marked compensation effect in the perception of these stimuli.

3.3.3. The effect of relatively small formant differences

A few other vowels were included to see whether relatively small differences in formant frequencies would show up in the listening results. ED Ziege has a lower F_1 (25 Hz) and a higher F_2 (150 Hz) than ED lieben. This difference is clearly heard. 84% of the Danish listeners heard the vowel of lieben as [e:] (partly as a raised [e:]), whereas only 6% heard the vowel of Ziege as [e:], and the German listeners show a somewhat stronger tendency to consider /i:/ in lieben as lowered [i]. - ED Minne has a slightly lower F_1 (25 Hz) and a slightly higher F_2 (75 Hz) than Distel, and F_3 is 275 Hz higher. This difference is reflected in the answers of the Danish listeners. 24% heard an [i] in Minne, but only 5% in Distel, and 10% heard a front rounded vowel in Distel. All the other answers were e (and there were more raised e's for Minne and more lowered e's for Distel). There were also some small differences in the German answers. The vowel of Distel was sometimes heard as [y].

3.3.4. Results of the discrimination test

This test was very restricted; it comprised only three sets of four vowel segments of 50-80 msec.: /i:, e:, I, ε/ spoken by ED and GR, and /y:, ø:, Y, œ/ spoken by CH, and not all possible comparisons were made. It is therefore no use to apply any statistics to the material, and, moreover, it would be complicated, since the subjects were allowed to give three different answers: (1) highest degree of similarity in the first pair, (2) in the second pair, and (3) no difference. 31 Danish students participated in the test.

The results are given in table 5, and the formant frequencies of the vowels are shown in fig. 17.

TABLE 5

Results of a comparison test of the type i-I, i-e. The numbers in the first and second column indicate the percentage of listeners who considered the similarity to be greater in the first and second pair respectively.

ED				GR				CH			
i	I 19	e 23	= 58	i	I 19	e 81	= 0	y	Y 23	ø 26	= 51
e	i 10	I 87	= 3	e	i 65	I 35	= 0	ø	y 19	Y 71	= 10
I	e 97	i 3	= 0	I	e 90	i 7	= 3	Y	y 10	ø 87	= 3
I	e 97	ɛ 3	= 0	I	e 32	ɛ 61	= 7	Y	ø 97	œ 3	= 0
ɛ	I 23	e 13	= 64	ɛ	I 74	e 13	= 13	œ	ø 35	Y 7	= 58

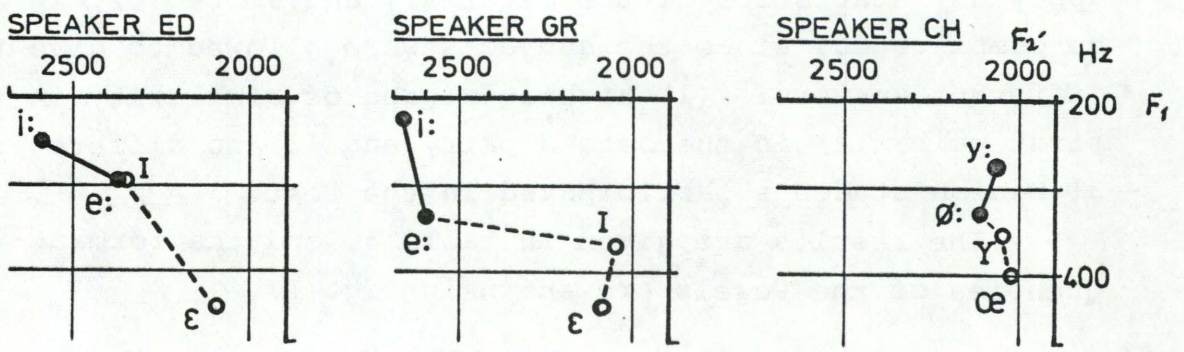


Fig. 17. Formant frequencies of the vowels used in the discrimination test.

It is obvious that the answers are in agreement with the acoustic diagrams, for instance the differences between the formant frequencies of ED and GR are clearly reflected in the answers. There is some indication that the vertical distance is more important than the horizontal distance when measured on the mel scale, cp. that the proximity of GR I-ε compared to I-e is only perceived by 61% of the listeners, whereas the proximity of I-e compared to I-i is perceived by 90%. This supports the assumption of a logarithmic scale as more appropriate than the mel scale (cp. EFJ 1972, footnote 5).

4. Conclusion

The following conclusions can be drawn from the experiments:

(1) For isolated vowel segments the placement of the stimuli in an acoustic F_1-F_2 (or F_1-$F_{2'}$) diagram gives, on the whole, a good basis for predicting the auditory impression, even as regards finer shades.

(2) The responses of German listeners to short segments of German /I, Y, U/ and /e:, ø:, o:/, however, cannot be explained on the basis of formant frequencies alone. The dispersion is considerable. Probably some subjects are able to listen for the quality and write them as e, ø, o, others are influenced by their short duration and tend to perceive them categorically as the phonemes /I, Y, U/, the highest lax vowels, and they write them accordingly as i, y, u. Still others identify them as short /ε œ ɔ/. The fact that /I, Y, U/ and /e:, ø:, o:/ are treated alike shows, however, their auditory similarity.

(3) The answers of Danish listeners to these sounds are more concentrated and in agreement with the formant frequencies. There is, however, a somewhat greater dispersion in their responses to /I, Y, U/ than to the other vowels (particularly for the speaker ED). This points to a more indefinite quality of /I, Y, U/, which is in agreement with their more central placement in the acoustic vowel diagram and with the indications of physiological laxness mentioned in section 2. This quality may present certain difficulties to listeners who have not got this type of vowels in their mother tongue.

(4) When vowels are presented in words, finer shades are not perceived, and the perception is dominated by phonemic categorization, automatic compensation for coarticulation effects, and influence from orthography.

References

Barczinski, L. and E. Thienhaus 1935: "Klangspektren und Lautstärke deutscher Sprachlaute", Archives néerlandaises de phonétique expérimentale 11, p. 47-68

Bennett, D.C. 1965: "Duration and spectral form as cues in the recognition of English and German vowels", Progress Report, Phonetics Laboratory, Univ. College London, p. 1-10

Chiba, T. and M. Kajiyama 1958: The vowel, its nature and structure, p. 150 and 153

Cohen, A., I.H. Slis and J. 'tHart 1963: "Perceptual tolerances of isolated Dutch vowels", Phonetica 9, p. 65-78

Fant, G. 1959: Acoustic analysis and synthesis of speech with applications to Swedish, p. 80

Fischer-Jørgensen, Eli 1941: "Løs og fast Tilslutning", NTTS 5, p. 41-69

Fischer-Jørgensen, Eli 1969: "Untersuchungen zum sogenannten festen und losen Anschluss", Kopenhagener germanistische Studien I (= Festskrift Peter Jørgensen), p. 138-64

Fischer-Jørgensen, Eli 1972: "Formant frequencies of long and short Danish vowels", Studies for Einar Haugen, p. 189-213

Fischer-Jørgensen, Eli and H.P. Jørgensen 1970: "Close and loose contact" ("Anschluss") with special reference to North German", ARIPUC 4/1969, p. 43-80

Fliflet, A.L. 1962: "Gespannte und ungespannte Vokale", Studia Linguistica 16, p. 24-28

Halle, M. and K.N. Stevens 1970: "On the feature "advanced tongue root", MIT QPR No. 94, p. 209-215

Heffner, R.M.S. 1949: <u>General phonetics</u>, p. 97-98

Hirano, Minoru and Timothy Smith 1967: "Electromyographic study of tongue function in speech, a preliminary report", <u>UCLA</u> 7, p. 35-45

Hockett, Ch.F. 1955: <u>A manual of phonology</u>, p. 31

Jakobson, Roman and M. Halle 1956: <u>Fundamentals of language</u>, p. 30

Joos, Martin 1948: <u>Acoustic phonetics</u>, p. 97

Jørgensen, H.P. 1969: "Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantfrequenzen", <u>Phonetica</u> 19, p. 217-45

Kaiser, Louise 1941: "Biological and statistical research concerning the speech of 216 Dutch students III", <u>Archives néerlandaises de phonétique expérimentale</u> XVII, p. 92-118

Ladefoged, P. 1964: <u>A phonetic study of West African languages</u>, p. 38

Lindblom, B.E.F. and M. Studdert-Kennedy 1967: "On the role of formant transitions in vowel recognition", <u>JASA</u> 42, p. 830-43

Lindner, G. 1966: "Beurteilung synthetisch erzeugter vokalartiger Klänge durch deutschsprachige Hörer", <u>Zs.f.Ph.</u> 19, p. 45-90

Meyer, E.A. 1910: "Untersuchungen über Lautbildung". Die neueren Sprachen 18 (= Festschrift Viëtor), p. 166-248

Meyer, E.A. 1913: "Das Problem der Vokalspannung", Die neueren Sprachen 21, p. 65-86 and 145-171

MacNeilage, Peter F. and George N. Sholes 1964: "An electromyographic study of the tongue during vowel production", JSHR 7, p. 209-232

Raphael, Lawrence J. 1971: "An electromyographic investigation of the feature of tension in some American English vowels", Haskins SR 28, p. 179-192

Russell, G.O. 1929: The vowel

Schuhmacher, W.W. 1972: Beitrag zur Bestimmung des physiologischen Korrelates des deutschen Vokalgegensatzes, Ling. 90, p. 35-78

Sievers, E. 1901: Grundzüge der Phonetik, p. 98-100

Stewart, J.M. 1967: "Tongue root position in Akan vowel harmony", Phonetica 16, p. 185-204

Vierling, O. and F. Sennheiser 1937: "Der spektrale Aufbau der langen und kurzen Vokale", Akustische Zeitschrift 2, p. 93-106

Wängler, H.H. 1958: Atlas deutscher Sprachlaute 2, 1961

Postscriptum

After having finished this report I became aware of a paper by Rudolf Weiss, read at the seventh international congress for phonetic sciences in Montreal 1971 (Proceedings 1972, p. 633-636) which is relevant for some of the problems treated here.

Weiss has undertaken a perceptual test with 287 examples of German vowels spoken by a German phonetician in the environment b-tən and varied according to quality and quantity. The listeners were asked to identify the word, i.e. the decisions were phonemic, but in contradistinction to the test described in the preceding report they were not asked to identify by quality alone.

According to the graph p. 635, /e:/ and /ø:/ were perceived as /I/ and /Y/ respectively when shortened below approximately 150 msec, and /o:/ was perceived as /U/ when shortened below approximately 200 msec. This is in accordance with the interpretation of the answers to the test reported here.

Conversely lengthened /I/ was heard as /e:/, lengthened /Y/ as /ø:/ or /y:/, and lengthened /U/ as /o:/ or /u:/ according to the quality. As might be expected, North German subjects listened more for quality, South German subjects for quantity.

AN INSTRUMENTAL INVESTIGATION OF THE DANISH "STØD"

Pia Riber Petersen

1. Introduction

In this study the stød is regarded as a phonologically distinctive element. For a discussion of this point, see Hjelmslev (1951) and Martinet (1937).

By means of the commutation test 4 contrasting types of words are established. These differ in the length of the vowel, and in the presence of the stød in the vowel or in the following consonant vs. absence of the stød. Examples: vener [ve·nəʁ] 'veins', Wegner [veʔnəʁ] (a name), vinder [venəʁ] 'winner', vinder [venʔəʁ] 'win' (present tense).

An acoustic analysis of these four types has been undertaken, comprising measurements of vowel and consonant length and of the pitch and intensity movements of the stressed vowel.

Opinions have differed regarding the length of vowels accompanied by the stød (henceforth referred to as stød-vowels) as in [veʔnəʁ]. Martinet (1937) gives arguments for considering a stød-vowel as a long vowel because it has the same quality as the (stød-less) long vowel in cases where there is a difference of quality between long and short vowels (e.g. ɑ-a: and ɔ-o:). In addition the stød-vowel has the same distribution as a long vowel (it is for example never found before [-ŋ]), and this analysis makes the place of the stød predictable: when a syllable has stød it appears in the vowel if the vowel is long: pæn [pεʔn], otherwise in the following consonant: pen [pεnʔ].

Jespersen (1949) is of the opinion that the stød-vowel is a long vowel; Uldall (1936), on the other hand, considers it to be short.

The length of the vowel has been investigated by Abrahams (1943) who found that the length of the stød-vowel is always intermediate between that of long and short vowels. He is of the opinion that the length of the stød-vowel is a cue for the stød.

Fischer-Jørgensen (1955) also states that the length of the stød-vowel is between long and short, but normally closest to long; she observes that there is overlapping in absolute duration between long vowels and stød-vowels, but not between stød-vowels and short vowels, which supports the assumption that the stød-vowel is acoustically long.

Lauritsen (1968) finds that a long vowel is twice as long as a short vowel and that a stød-vowel is one and a half times as long as the short vowel.

The stød has been considered as originating from a special pitch movement since the words with stød correspond to words with "accent 1" in Norwegian, Swedish, and some southern Danish dialects, and in accordance with this, some scholars have argued that there is a musical difference also in the standard language between words with stød and words without stød (Verner 1903).

The question whether the stød is a glottal stop or not was investigated by Rousselot (1924) who found that a closure is rare. The same result is obtained by Ekblom (1933) who in addition discovered a steep increasing pitch followed by a decrease. Selmer (1925) found often a real closure, but his recordings were made with emphatic pronunciation.

The most important instrumental investigation of the stød was undertaken by Svend Smith (1944). His investigation includes electro-myographic, oscillographic and kymographic registrations. He discovered that there is a particular kind of innervation in the expiratory muscles in words containing a stød: it seems that the stød-word has a more discontinuous character in comparison with the stød-less word. Pitch and intensity show the same movement: In the stød-word there is a decrease in the intensity accompanied by a decrease in pitch, often ending in irregular oscillations, but his recordings do not give any example of a complete closure. Sometimes there is a reappearance of regular oscillations after the irregular interval. Smith does not observe any difference in the pitch level between stød-words and stød-less words. His conclusion is that the stød is a stress accent: a special "marking movement" made by a thrust-like emphasis of sounds. The stød often appears to be three-phased: 1) a ballistic contraction of the expiratory muscles, 2) cessation of this activity, which causes a lack of balance in the reaction of the vocal cords, 3) sometimes a new activity in the expiratory muscles.

2. The present investigation

2.1. The informants

The speakers comprise 3 males and 3 females, designated 22,25,26 and 11,13,14 respectively. They all speak Copenhagen Standard Danish (Rigsdansk), and all of them have normal healthy voices.

2.2. The material

The material consists of 8 groups of words where the frames are approximately identical; only the length of the vowel and the stød varies.

hyle '(to) yell'[hy·lə]	hyler '(I) yell' [hyʔl̩əʁ]
hylde 'shelf' [hylə]	hylden 'the elder' [hylʔən]
kæle '(to) caress' [kʰɛ·lə]	kæler '(I) caress' [kʰɛʔl̩əʁ]
kælder 'cellar' [kʰɛləʁ]	Keller, a name [kʰɛl̩ʔəʁ]

bene '(to) run' [be·nə]	benet 'the leg' [beʔnəð]
binde '(to) tie' [benə]	binder '(I) tie' [benʔəʁ]
hane 'cock' [ha·nə]	afganer '(an) afghan'[ɑu'gaʔnəʁ]
Hanne, a name [hanə]	hannerne 'the males' [hanʔəʁnə]
læser 'reader' [lɛ·səʁ]	læser '(I) read' [lɛʔsəʁ]
læsser '(I) load' [lɛsəʁ]	
gyse '(to) shiver' [gy·sə]	gyser '(I) shiver' [gyʔsəʁ]
gysser 'dough' (colloquial) [gysəʁ]	
taber '(I) lose' [t^{h}a·bəʁ]	tabet 'the loss' [t^{h}aʔbəð]
tapper '(I) drain' [t^{h}abəʁ]	
piber 'pipes' [p^{h}i·bəʁ]	piber '(I) squeak' [p^{h}iʔbəʁ]
pipper '(I) chirp' [p^{h}ibəʁ]	

The test words were all placed in sentences which had an identical rhythm[1], for example:

Sirenen plejer at hyle om aftenen
Sjakalen sidder og hyler i ørkenen
Jeg tror der mangler en hylde i bogskabet
Jeg tror vi plantede hylden i foråret.

2.3. The recordings

All the sentences were read aloud in random order 6 times by each of the informants. The recordings took place in the recording studio of the Institute of Phonetics, University of Copenhagen, using a Lyrec professional Recorder, 1/1 track at 7^{1}/2"/sec on Scotch Magnetic Tape.

---

1) It is seldom possible to place a stød-word and a stød-less word in exactly the same environment in the same sentence, because the stød in many cases distinguishes classes of words or inflected forms.

2.4. The tracings

For the reproduction the same tape recorder was used as in the recording operation. The signal passed a pitch-meter and an intensity meter, and the curves were registered by means of a 8-channel Elema Mingograph. The traces (see fig.1) were:

1. a pitch curve
2. a logarithmic intensity curve, high-pass filtered at 500 Hz and with an integration time of 2,5 msec for speakers 11,22,13,14, and 5 msec for speakers 25 and 26.
3. a linear, unfiltered intensity curve with an integration time of 5 msec for speakers 11,22,13,14, and 10 msec for speakers 25 and 26.
4. a duplex oscillogram.

All measured values were fed into a computer using a simple program which produces the mean, the standard deviation, and the 95 and 99% confidence limits of the mean.

3. Results

3.1. The manifestation of the stød

In the recorded material the stød behaves in many different ways, different for the various speakers and different in the recordings of the different readings of the same word by one speaker. One extreme is complete closure (occurring a few times). The other extreme is constituted by cases without any visible difference between the curves of words containing stød and words without stød. Between these two extremes all intermediate steps are represented. Figs. 2a-5d

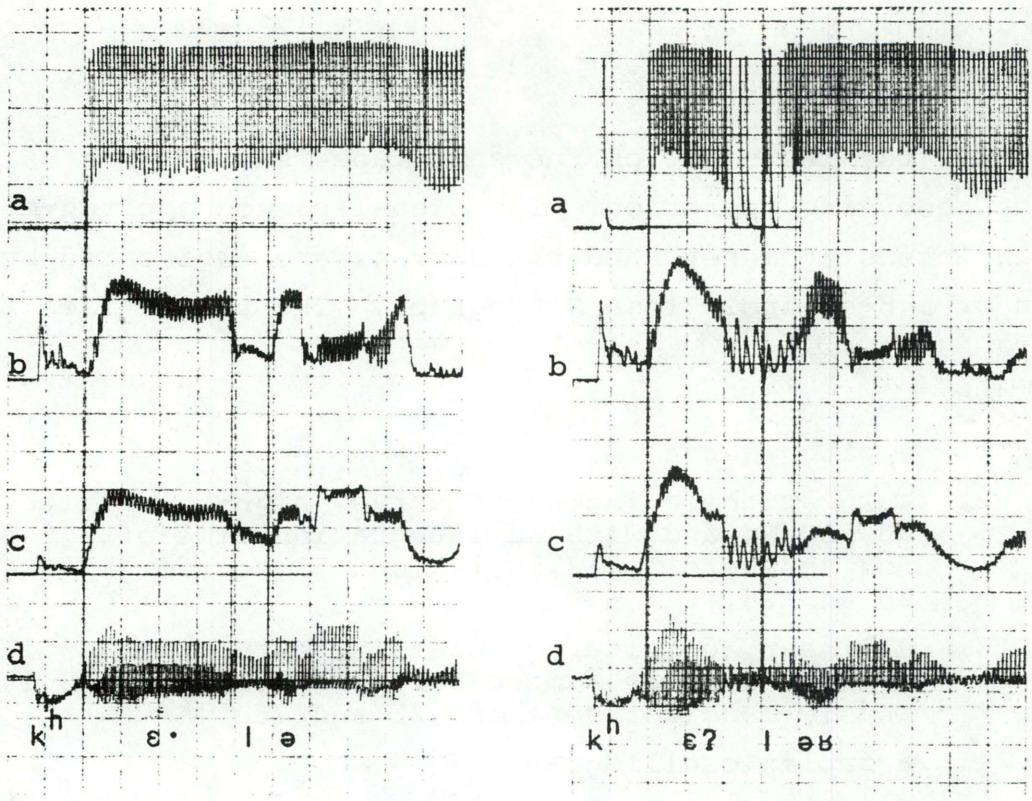


Fig. 1 Mingogram of [kʰε·lə] and [kʰεʔləʁ].
The traces are:
a. a pitch curve
b. a logarithmic intensity curve, highpass filtered at 500 Hz.
c. a linear, unfiltered intensity curve
d. a duplex oscillogram.

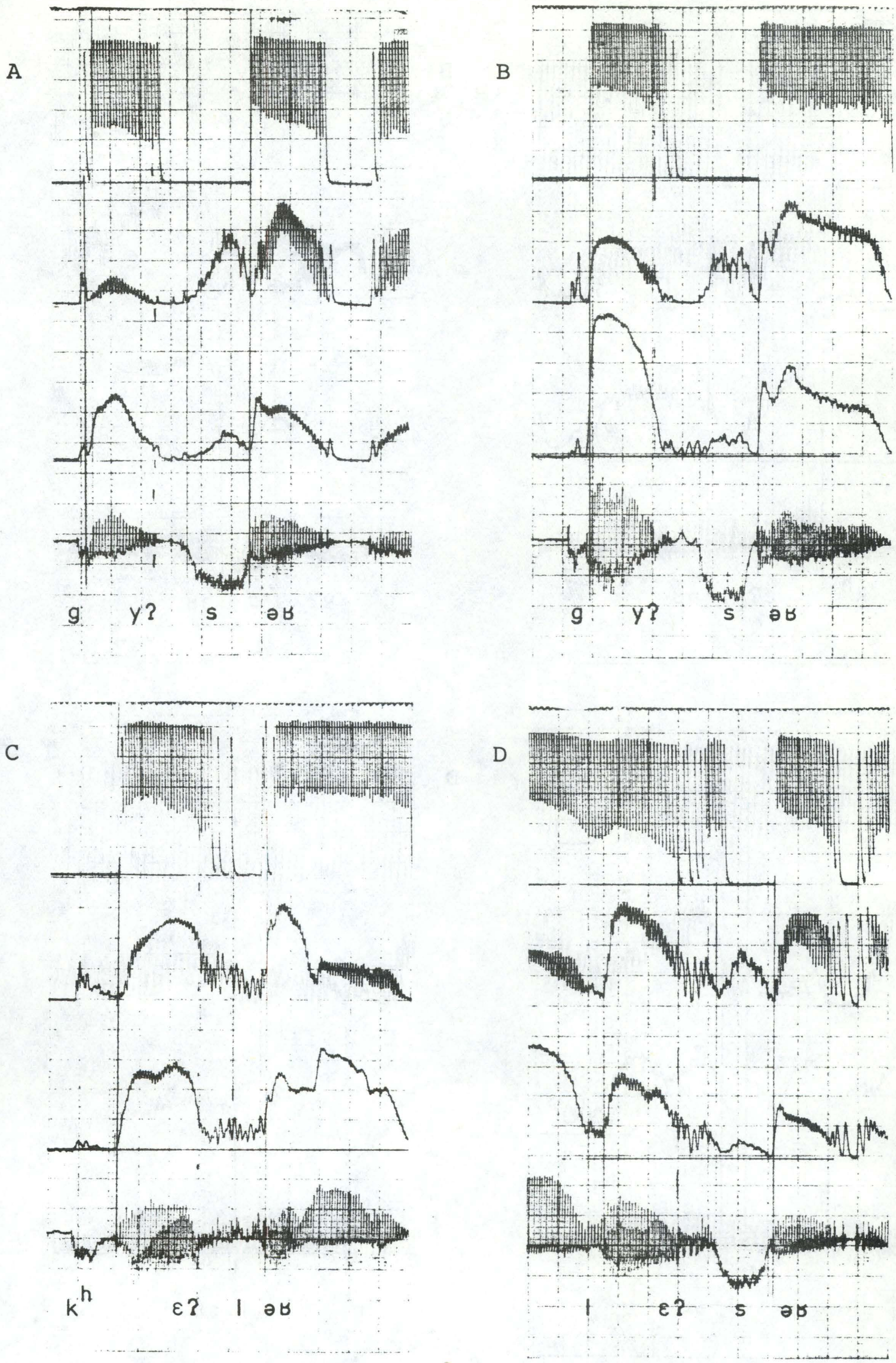


Fig. 2 A-D   Different types of stød in the vowel.

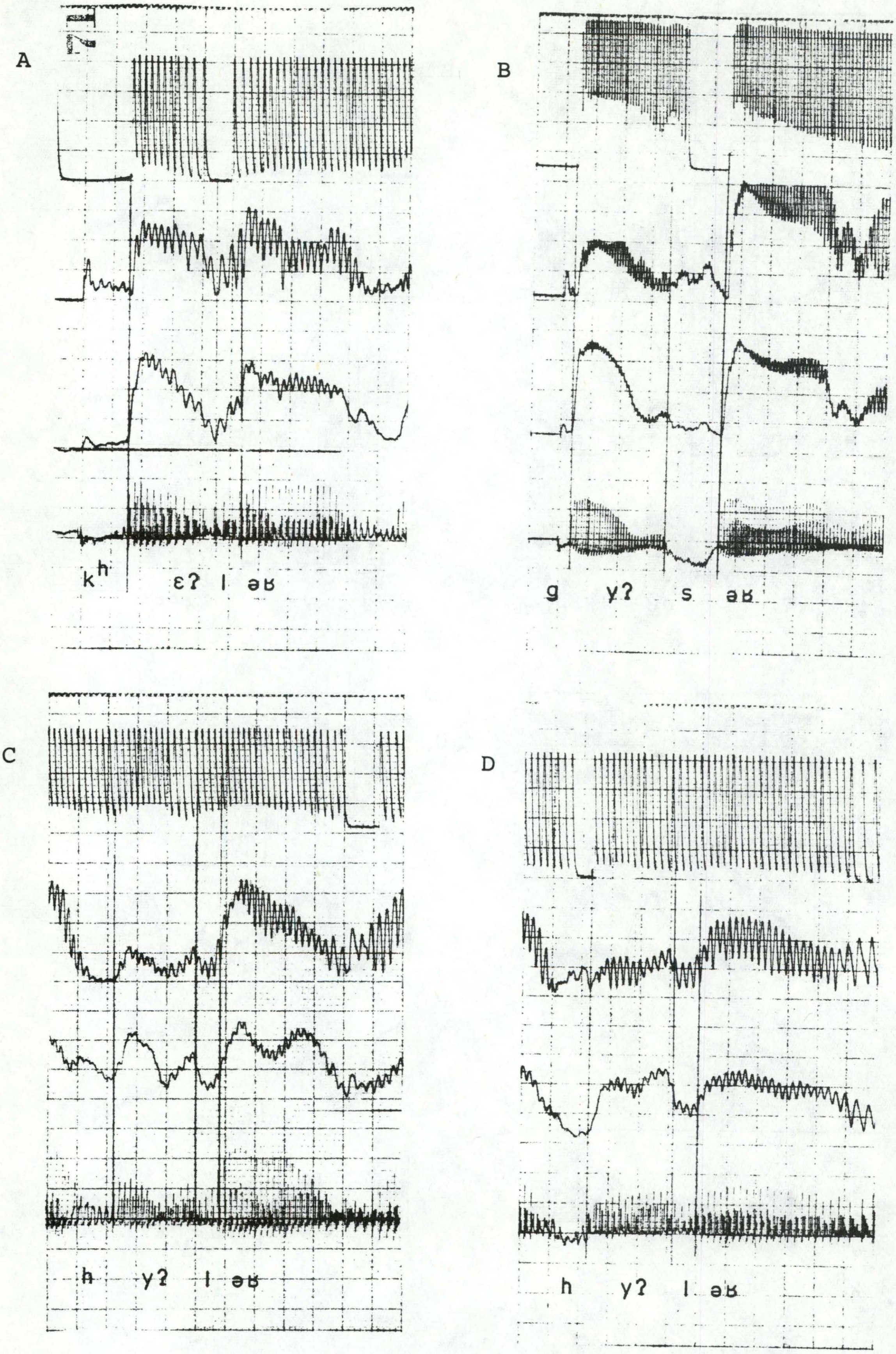


Fig. 3 A-D   Different types of stød in the vowel.

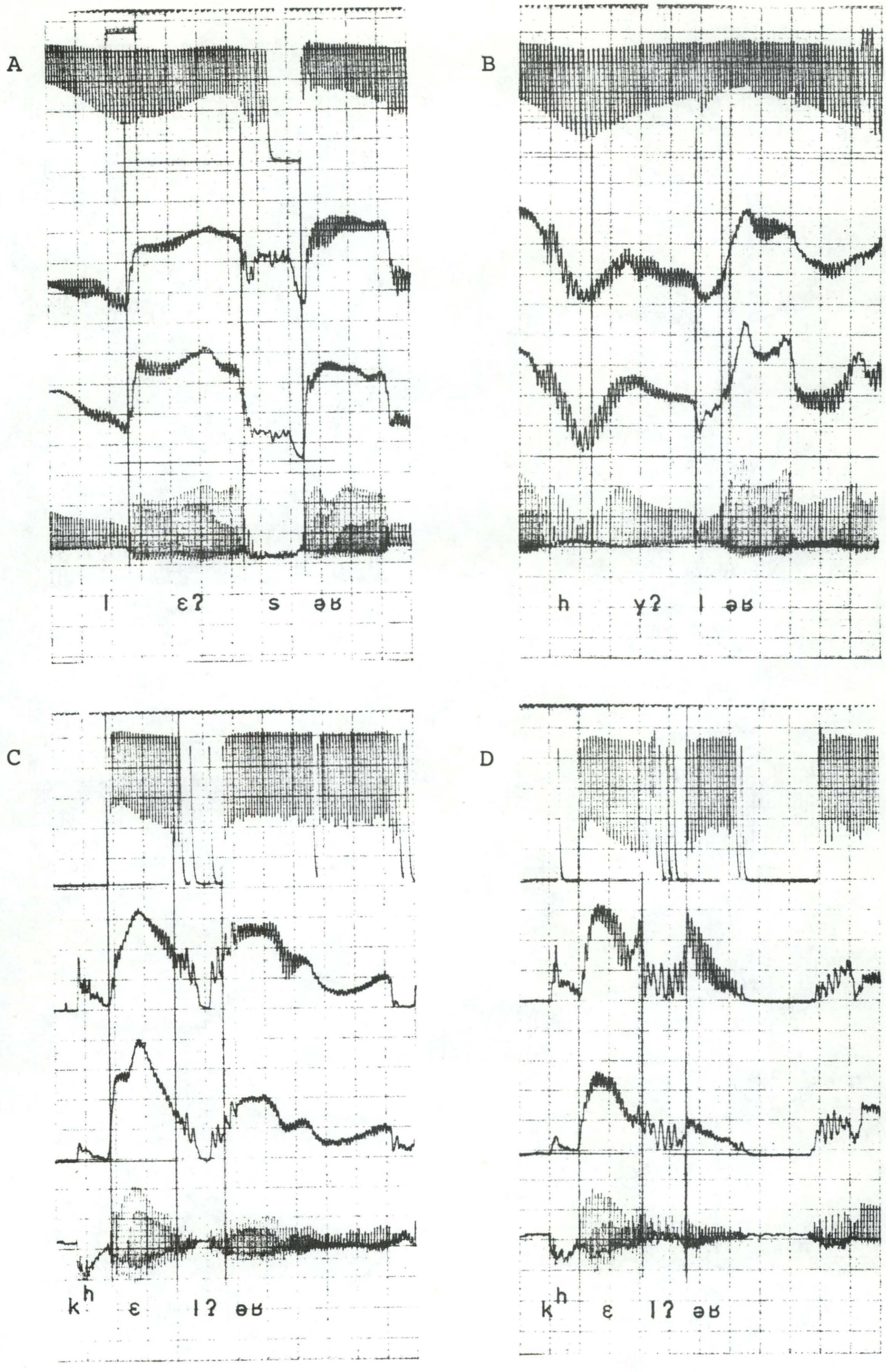


Fig. 4 A-B Different types of stød in the vowel
C-D Different types of stød in the consonant.

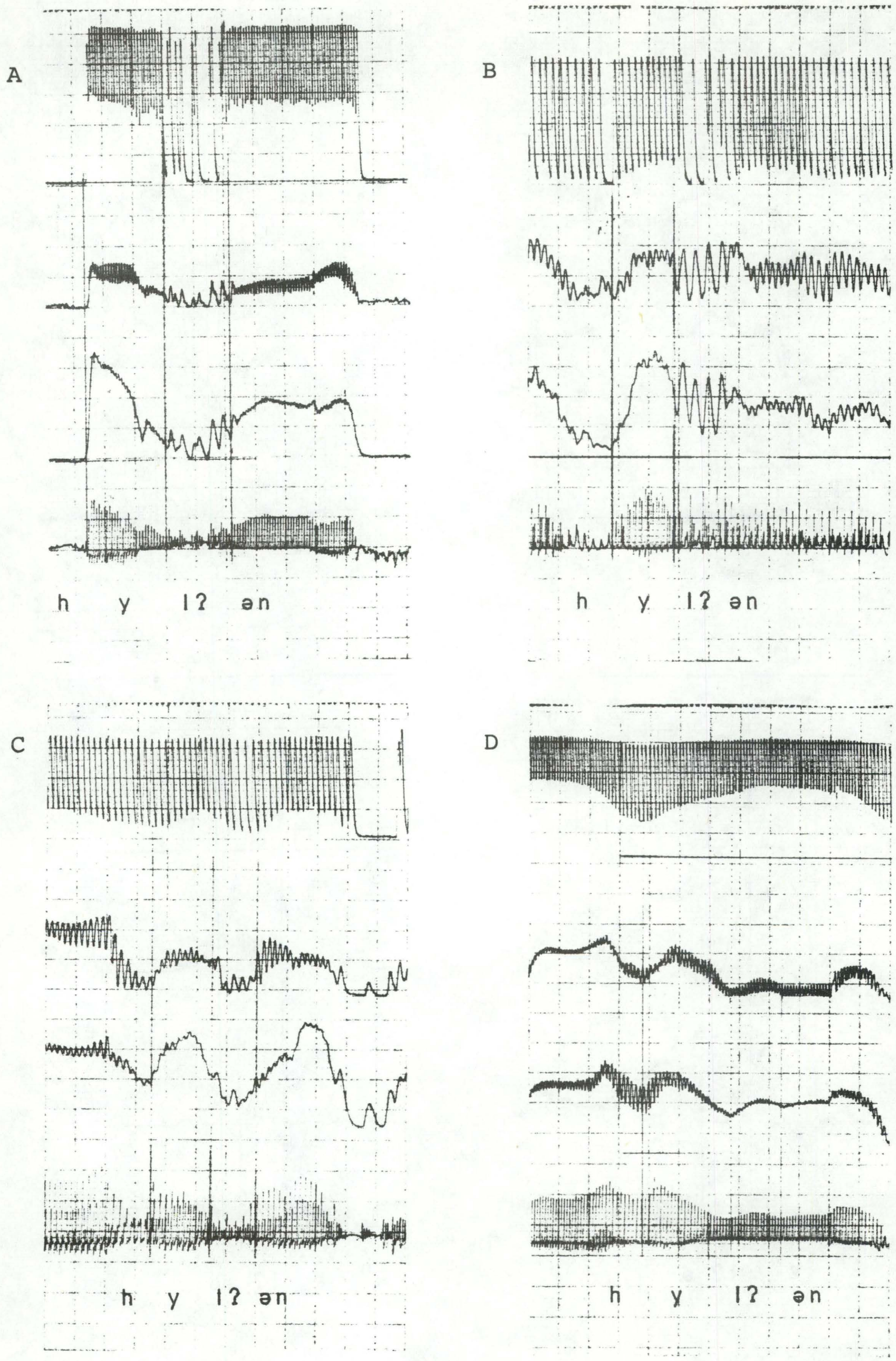


Fig. 5 Different types of stød in the consonant.

illustrate different types of the stød; these are representative of the whole material.

Fig. 2a. An example of closure. This is not found very often.

Fig. 2b. Almost complete closure at the end of the stød phase.

Fig. 2c. Strong stød with complete change of the oscillatory pattern. Steep decrease in intensity and pitch.

Fig. 2d. The same as in 2c, the regularly jagged oscillation in the beginning of the stød phase is remarkable.

Fig. 3a. A stød of minor strength in contradistinction to the preceding types, which is not distinctly localized but seems to have a gradually increasing effect on the vowel. The jagged oscillations continue into the following consonant.

Fig. 3b. Weaker stød, influencing pitch and intensity, both of which show a steep decrease.

Fig. 3c. Stød in the middle of the vowel. All the preceding types of stød are placed at the end of the vowel and there is sometimes an influence also on the oscillations in the following consonant. The present type where the stød is in the middle of the vowel, is characterized by a reduction and a jagged form of the oscillations after which the oscillations gradually return to their normal regular form. As a rule there is an accompanying decrease and increase of the intensity, whereas the pitch does not rise again.

Fig. 3d. Illustration of the same phenomenon as in fig. 3c, but weaker.

Figs. 4a-4b. These are examples where there is no visible difference between words with and without stød. In 4a the pitch and intensity curves show the same movement, in 4b the pitch is increasing and the intensity decreasing.

The figures referred to so far are all illustrations of stød in the vowel; the following figures show stød in the consonant. In contradistinction to the stød in the vowel the stød in the consonant often extends over the whole segment and can also influence the surroundings.

Fig. 4c. Closure in the stød phase.

Fig. 4d. Locally concentrated stød, with complete change of the form of the oscillations.

Fig. 5a. Strong stød in the whole consonant. The clear influence on the last 5 cs of the preceding vowel is remarkable.

Fig. 5b. Strong stød in the beginning of the consonant and a return to regular oscillation at the end of the consonant.

Fig. 5c. Weak stød.

Fig. 5d. Here the stød is not visible, although it can be heard on the tape.

The manifestations of the stød in the consonant and in the vowel of the same speaker are more or less alike. All the types of manifestations are not represented by the same speaker, the variation comprising 2 or 3 degrees of the scale illustrated by figs. 2a to 4b, or figs. 4c to 5d. For speaker 11 the most common type is the one shown in fig. 3b, for speaker 22 those shown in figs. 4b and 5d are the normal types. Only speakers 13 and 14 may have a closure in the stød phase, although rarely. The most common type used by these two speakers is the locally concentrated type, represented by figs. 2c-2d. Speakers 25 and 26 have seldom the local stød, more often the type shown in fig. 3b. The stød in the middle of the vowel is found only in the recordings of these two speakers (figs. 3c-3d).

3.1.2. The type of stød which is not visible in the curves

Figs. 3a-3b show that the stød may be almost invisible in the curves. This type of stød was observed for speaker 22 in particular, only in rare cases for others. Some spectrograms were taken of the words without stød, and the corresponding words with stød for this speaker. See fig. 6. In the words with stød there seems to be a weakening of the higher formants at the end of the vowel and some irregular transitions of the higher formants which do not seem to be caused by the steeply increasing pitch alone. But the material is at this point rather limited, and further investigations are required.

These very small differences between words with and without the stød demanded an investigation of the audibility of the stød, when the stød-word has been removed from its context.

A listening-test was made with one word-pair: <u>læser</u> 'reader' [lɛ·səʁ] and <u>læser</u> '(I) read' [lɛʔsəʁ], where the surrounding sounds were so similar that they could not influence the two words differently. The words were cut out of the original material by means of a segmentator (Thorvaldsen 1969). For the sake of comparison a second test was set up with the same words spoken by speaker 26, whose stød was clearly manifested in the curves, although he did not have the strongest stød phase among the speakers.

The material comprises 6 different recordings of the word with stød and the word without stød, all used twice in the test, that is, 24 test words in total. Each test word was presented just once, and after every test word there was a pause of 5 msec. The subjects were asked to identify the test word with one of the underlined words in the sentences:

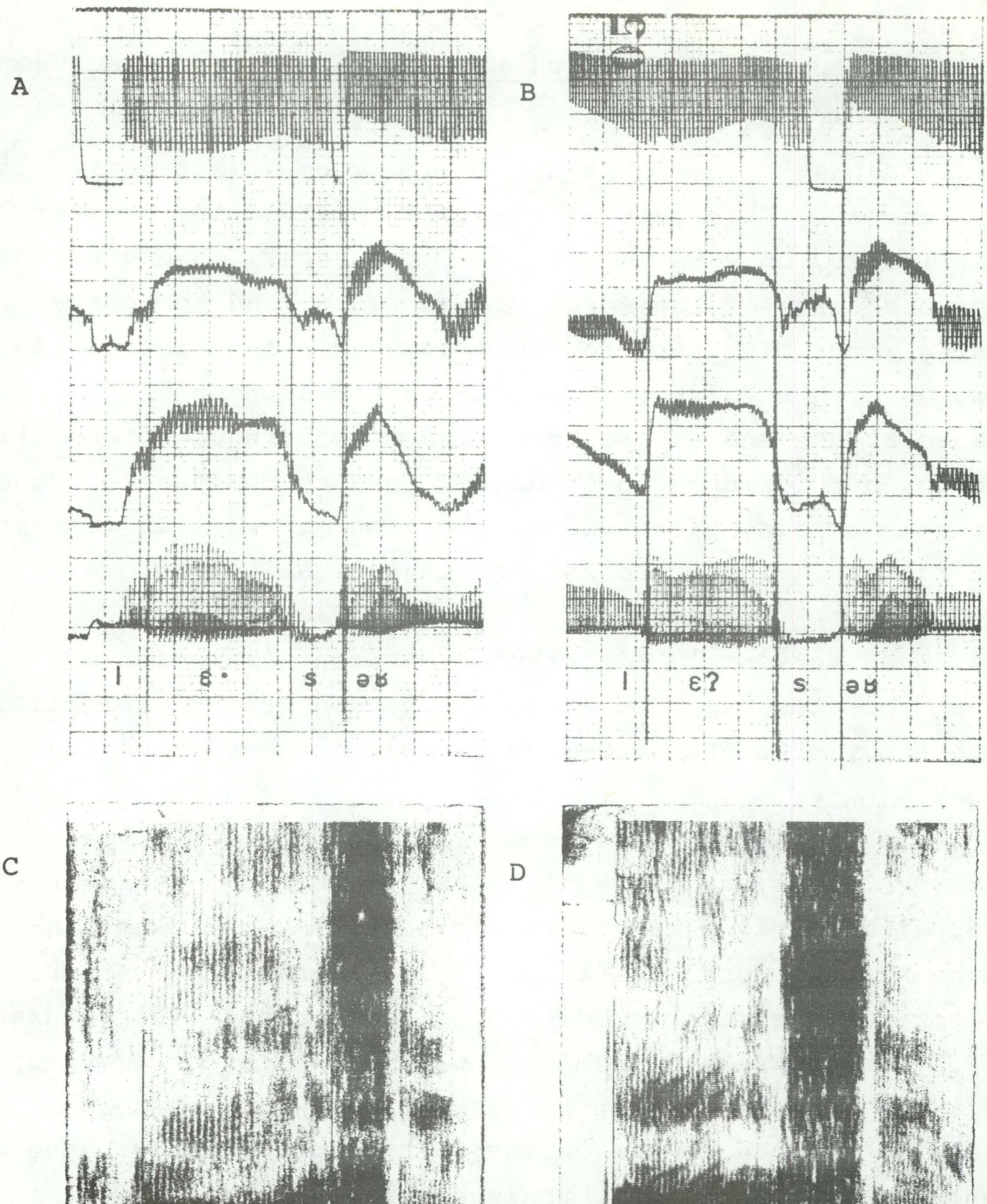


Fig. 6 Mingograms (A-B) and spectrograms (C-D) of non-stød (A-C) versus stød (B-D) at the speaker 22.

Bogens læser undrer sig

Drengen læser om aftenen.

There was forced choice. The tests were played first in the order speakers 26-22 for 7 subjects and next in the reverse order for another 7 subjects. There were thus 168 answers in all.

3.1.2.1. Results

The stød is identified rather badly, but not worse for speaker 22 than for speaker 26, in fact the stød is heard best for the former. It should be noticed that most of the mistakes are made in the words with stød: these are often identified as words without stød, rarely vice-versa.

TABLE 1

Errors in the listening test.

speaker order	26	22	22	26
number of errors	50	16	25	31
errors in words with stød	42	7	20	29
errors in stød-less words	8	9	5	2
N=168				

TABLE 2

The errors in %

speaker order	26	22	22	26
number of errors	30	10	15	19
errors in words with stød	50	8	24	36
errors in stød-less words	10	11	6	2

The stød words of the same speaker are identified better when he comes last in the test, but the difference between speakers 22 and 26 is clear in both cases. The errors for speaker 22 are below 25%.

3.2. Length

The length of the vowel and the length of the following consonant have been investigated in all four types of words which from now on are designated as V:C, V?C, VC, and VC?.[1] In addition, the length of the stød phase was measured in the cases where it was possible, keeping in mind that the stød cannot always be delimitated. The length was measured with an accuracy of 1 cs.

3.2.1. The length of the vowel

The reader is referred to fig. 7 where the average length of each vowel for two typical speakers and the 95% limits are shown.

---

1) ? indicates stød, : length, V vowel, and C consonant.

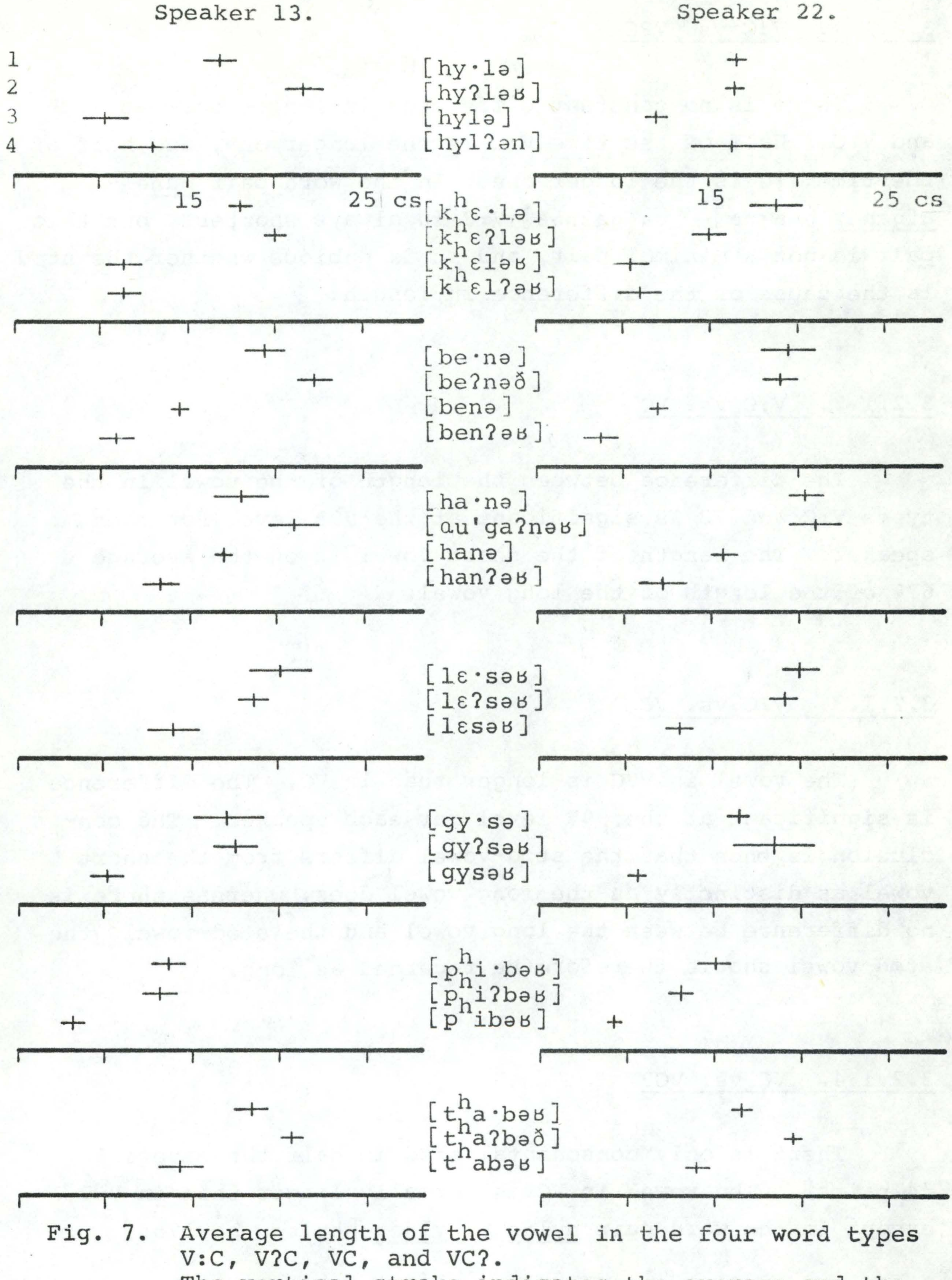


Fig. 7. Average length of the vowel in the four word types V:C, VʔC, VC, and VCʔ.
The vertical stroke indicates the average and the endpoints of the horizontal line indicate the 95% confidence limits.

1 = V:C, 2 = VʔC, 3 = VC, 4 = VCʔ.

3.2.1.1. V:C vs. V?C

There is no constant difference in length between V:C and V?C. Half of the time V:C is the longer one, and half of the time V?C is the longer one. In the word pair hane - afganer [ha·nə]- [au'gaʔnəʁ] V:C is always shortest, but this pair is not a minimal pair, and it is dubious whether the stød is the cause of the difference in length.

3.2.1.2. V:C vs. VC

The difference between the length of the vowel in the types V:C and VC is significant at the 99% level for each speaker. The length of the short vowel is on the average 67% of the length of the long vowel.

3.2.1.3. V?C vs. VC

The vowel in V?C is longer than in VC. The difference is significant at the 99% level for each speaker. The conclusion is then that the stød-vowel differs from the short vowel as distinctly as the long vowel does, whereas there is no difference between the long vowel and the stød-vowel, the stød-vowel should therefore be regarded as long.

3.2.1.4. VC vs. VC?

There is only consonantal stød in half the material (4 pairs). The vowel in VC is normally longer than in VC?, except in the word pair hylde - hylden [hylə]- [hylʔən] 'shelf' - 'the elder', where the vowel in VCʔ is the longest for 4 of the 6 speakers. The difference between the four word types was calculated, and the significance of the average difference for all speakers was tested. See table 3.

TABLE 3

The difference between the length of the vowel in VC and VCʔ (VC-VCʔ). Measured in cs.

speaker	[benə/ benʔəʁ]	[hanə/ hanʔəʁ]	[k^hɛləʁ/ k^hɛlʔəʁ]	[hylə/ hylʔən]		
11	4,3	1,2	3,2	-2,9		
22	4,4	3,6	-0,4	-4,5		
13	3,7	2,1	0,1	-2,6		
14	3,6	2,8	0,8	-0,9		
25	3,0	1,5	1,3	-2,1		
26	1,2	2,3	1,3	0,1		
average	3,4	2,3	1,1	-2,2	total	1,2
standard deviation	1,2	0,9	1,3	2,0		2,2
significance level	99%	99%	90%	90%		-

The result is that the difference between VC and VCʔ is significant at the 99% level in [benə/benʔəʁ] [hanə/hanʔəʁ], but only 90% in [k^hɛləʁ/k^hɛlʔəʁ] and 90% in [hylə/hylʔən] where VCʔ is the longest. When all word pairs are taken together the difference is not significant.

It is evident that the stød on the consonant causes a reduction of the short vowel in most cases, but the length of the vowel also seems to depend on the type of the following consonant, the difference being larger before [n] than before [l].

3.2.2. The length of the following consonant

Knowing that in other languages (e.g. Swedish and Italian) there may be an inverse relation between the length of the vowel and that of the following consonant, i.e. a long vowel is followed by a short consonant and a short vowel by a long consonant, also the length of the consonant in the different word types (V:C, VʔC, VC, VCʔ) was measured.

3.2.2.1. VC vs. V:C

The differences are very small, and there is no constant difference between the length of the consonant in VC and V:C. See table 4.

TABLE 4

The difference in length of the consonant in VC and V:C (VC-V:C). The difference is calculated for each speaker and the average is reproduced below.

kæle	pibe	læse	bene	tabe	gyse	hyle	hane
0,2	-0,2	0,3	0,2	-0,4	-2,9	0,7	0,2

3.2.2.2. VC vs. VʔC

The same appears for a comparison between VʔC and VC, the consonant of VʔC having almost the same length as in V:C. The length of the consonant is thus independent of the length of the preceding vowel.

3.2.2.3. VC vs. VCʔ

The consonant in VCʔ (7,2cs) is longer than in VC (5,4cs). The consonant in VCʔ is thus 134% of the consonant in VC. The difference is calculated for each word pair and it is computed that the difference is significant at the 99% level. See table 5.

TABLE 5

The difference between the length of the consonant in VC and VCʔ (VCʔ-VC). Measured in cs.

speaker	[benʔəʁ/ benə]	[hanʔəʁ/ hanə]	[k^hɛlʔəʁ/ k^hɛləʁ]	[hylə/ hylʔən]
11	0,7	2,3	1,9	1,9
22	1,9	-1,0	0,9	5,0
13	0,5	1,3	2,3	3,6
14	0,6	0,9	1,2	6,1
25	2,1	-0,9	1,3	5,0
26	0,9	-0,5	1,8	3,3
average	1,1	0,3	1,5	4,1

The difference is biggest in [hylə/hyl?ən] and smallest in [hanə/han?əʁ]. As stated above (3.2.1.4.) the vowel is generally longer in VC than in VC?, i.e., there is a certain (weak) inverse relation between vowel and consonant length. The pair [hylə/hyl?ən] constitutes an exception to this tendency, however: both vowel and consonant are longer in the stød word [hyl?ən] than in [hylə].

3.2.3. The length of the stød phase (defined by irregular oscillations)

Only examples in which the stød phase is clearly visible in the curves, as in figs. 2a-b-c (stød in the vowel) or in figs. 4c-d (stød in the consonant), have been measured. The types 3c and 3d (with the stød in the middle of the vowel) have been left out. These types are very rare.

3.2.3.1. The length of the stød in the vowel

Table 6 shows the length of the stød phase in percentage of the total length of the vowel. The raised numbers indicate the cases in which it has been possible to determine the limits of the stød phase.

The stød phase has a rather constant relative length of about 1/3 of the total length of the vowel.

TABLE 6

The length of the stød phase in percentage of the total length of the vowel.

speaker	11	22	13	14	25	26	all
kæler	-	-	37^{6}	28^{6}	57^{1}	-	33^{2}
piber	38^{3}	-	25^{6}	42^{6}	-	-	34^{2}
læser	-	-	39^{6}	30^{6}	-	-	35^{2}
tabet	32^{6}	-	27^{6}	34^{6}	47^{2}	34^{4}	33^{3}
gyser	-	-	41^{6}	34^{5}	-	40^{1}	38^{2}
hyler	35^{6}	-	37^{6}	42^{6}	43^{2}	40^{6}	38^{4}
benet	-	-	47^{6}	33^{2}	37^{6}	33^{3}	42^{2}
afganer	30^{6}	-	36^{6}	30^{6}	-	38^{2}	32^{3}

3.2.3.2. The length of the stød phase in the consonant

There is a tendency toward a longer interval of irregular oscillations, but the material is too small for any definite conclusion to be made.

3.2.3.3. The place of the stød phase compared to the onset of the vowel

The average distance from the onset of the vowel to the beginning of irregular oscillations is 11,2 cs for the type V?C, and 11,9 cs for the type VC?.

4. Results concerning pitch and intensity

4.1. Measuring the curves

The pitch and intensity movements were only measured in the vowel. (They were measured according to the same principles.) The following points in the curves were measured: The first maximum (max_1), the following minimum (min), and the second maximum (max_2). The positions of these points were indicated by the distance from the start of the vowel. The curves fall into three classes according to the presence or absence of either of the two maxima:

A: The falling-rising type[1] (max_1-min-max_2)
B: The purely rising type (min-max_2)
C: The purely falling type (max_1-min)

Rising-falling curves are not represented in this classification as an independent type, but will be registered as purely falling. This may give a wrong impression of the movement if max_1 occurs late in the vowel; the type was, however, excluded because it was a very rare type and would have cost another measuring point, which would complicate the calculations.

4.2. The accuracy of the measurements

The pitch was measured with an accuracy of $\pm$ 5 Hz, intensity with $\pm$ 1 dB, and length with $\pm$ 1 cs.

1) The level type, which is very rare, has been included in type A, with identical values for max_1, min, and max_2.

4.3. The pitch

In figs. 8-13 the averages of max_1, min and max_2 of 6 recordings of each word have been calculated for all speakers and placed in a co-ordinate system, where the axis of the abcissa is the length measured in cs, and the axis of the ordinate is the pitch measured in Hz. For considerations of space only words with voiced consonants are included.

4.3.1. V:C vs. V?C

4.3.1.1. The shape of the curves of V:C vs. V?C

There is no constant difference between the shapes of the curves in stød words and stød-less words. All types of curves can be observed in both word types, if the speakers are taken en bloc. The individual speakers show some differences.

Speakers 13 and 14 have all three types of curves (A-B-C) in V:C, but only the falling type in V?C. Speaker 26 has generally a rising curve in V:C and a falling curve in V?C. Speaker 11 has a falling-rising curve as the predominant type both in V:C and V?C. Speaker 25 has generally a falling-rising curve in V:C, and either the same or a purely falling type in V?C. Speaker 22 has no clear difference.

In general it is not rare that the fall starts later in V?C, and in almost all cases the minimum is closer to the end of the vowel than it is in V:C.

4.3.1.2. The pitch of max_1, min and max_2

When both V:C and V?C exhibit a max_1 (i.e., the curves are falling or falling-rising), this max is clearly highest

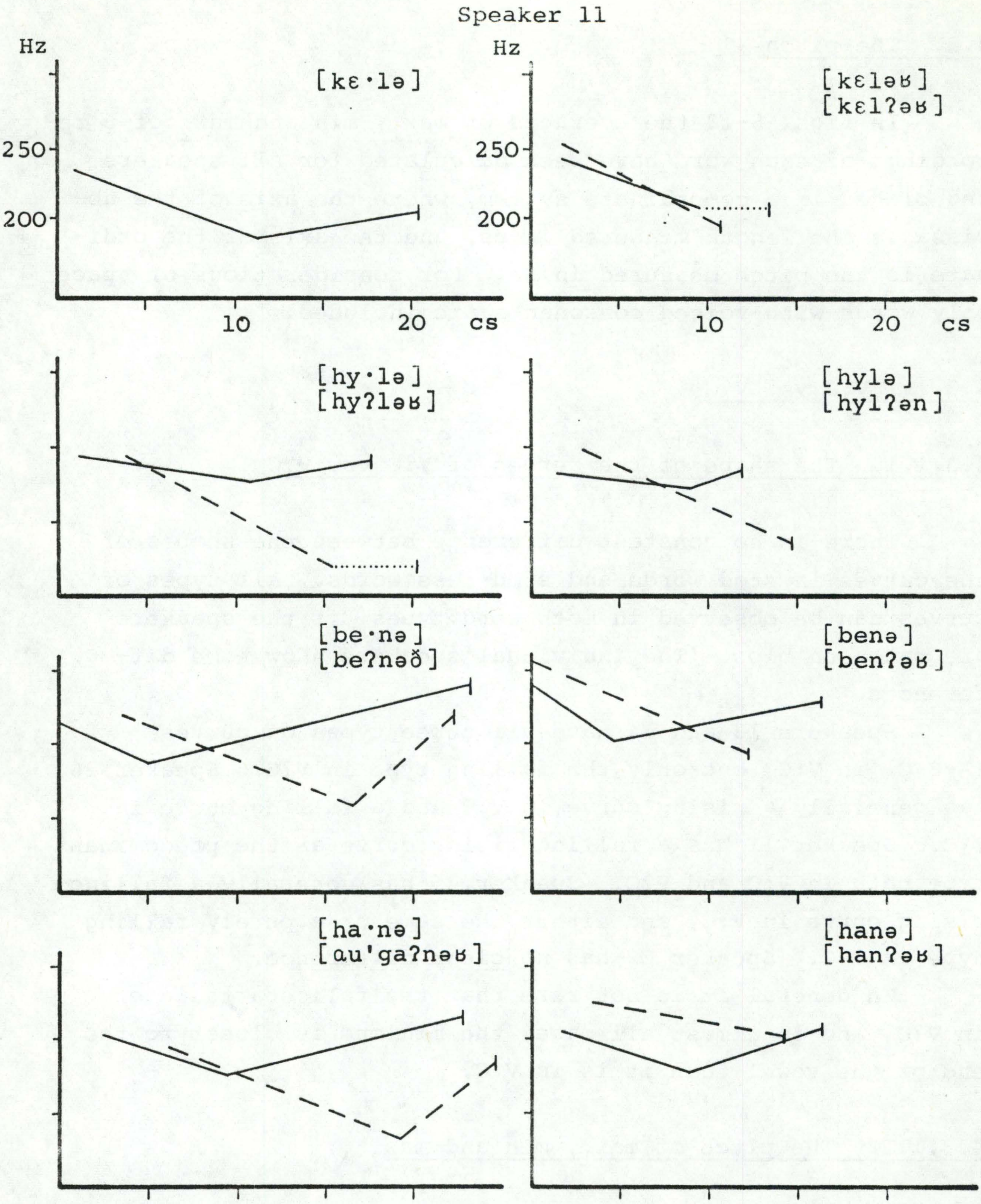


Fig. 8. Average pitch movement of the vowel in the word types with voiced consonants for speaker 11.

V:C and VC ————

V?C and VC? -------

The vowel begins at the vertical axis and ends at the vertical stroke.

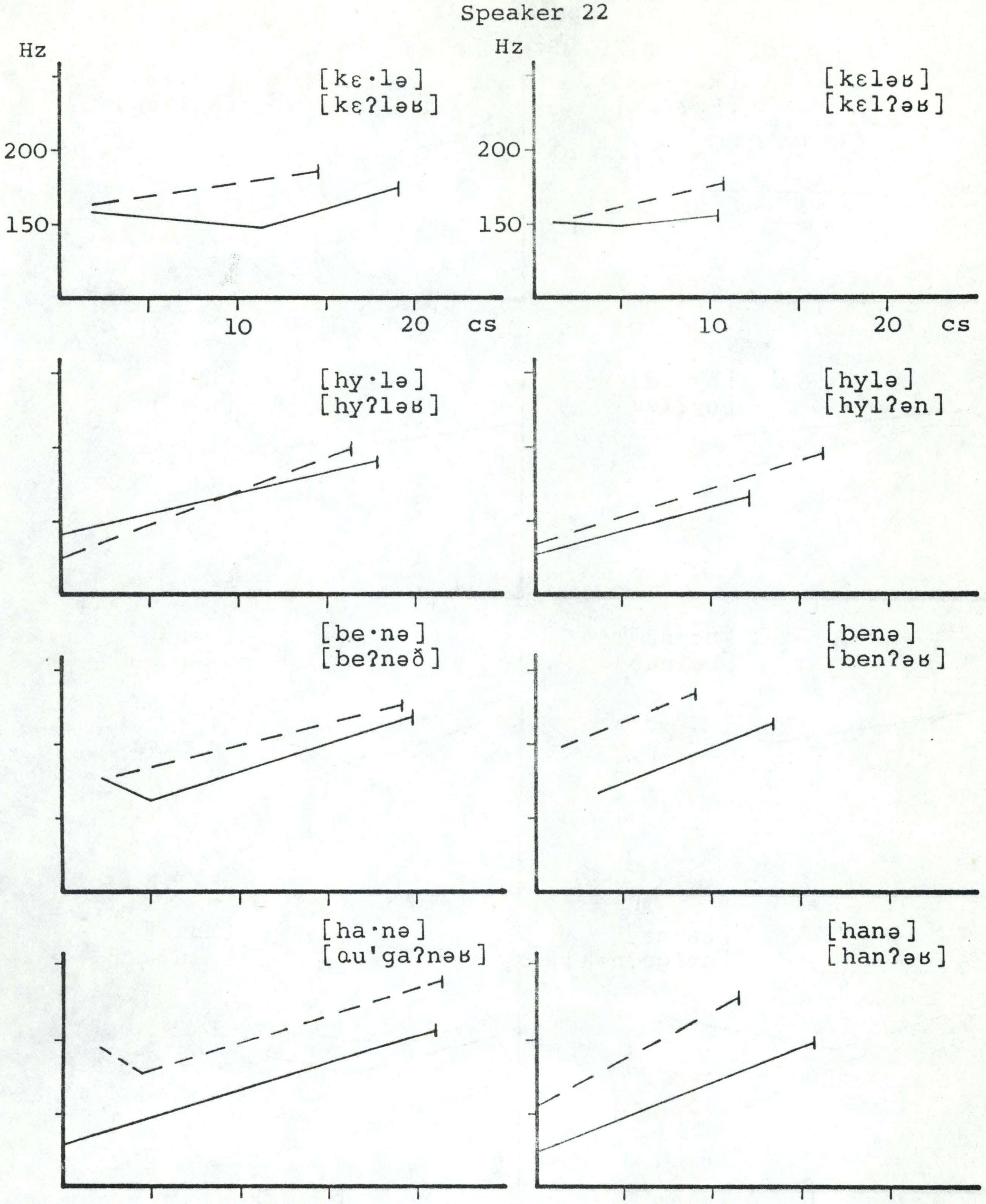


Fig. 9. Average pitch movement of the vowel in the word types with voiced consonants for speaker 22.

V:C and VC ———

VʔC and VCʔ -------

The vowel begins at the vertical axis and ends at the vertical stroke.

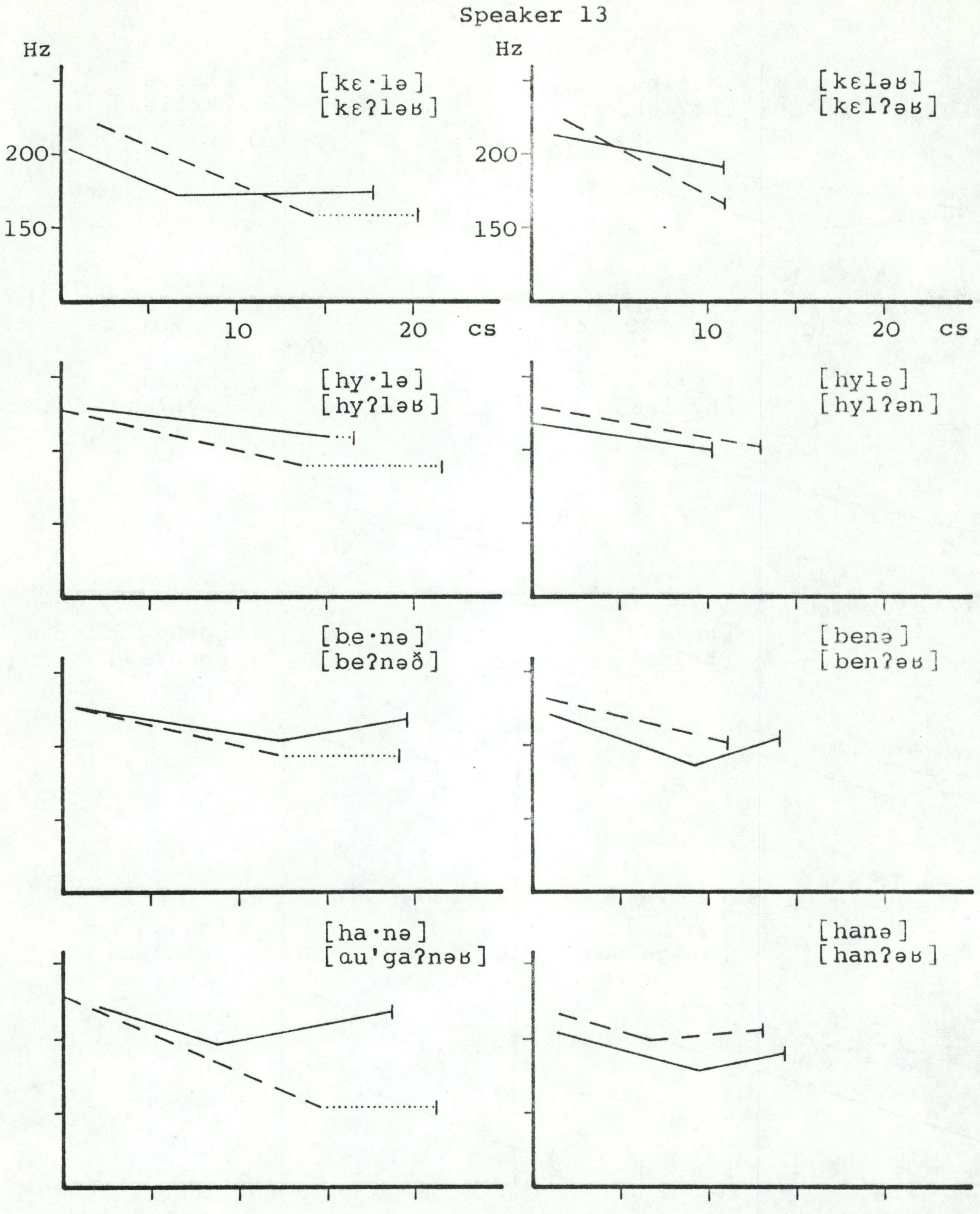


Fig. 10. Average pitch movement of the vowel in the word types with voiced consonants for speaker 13.

V:C and VC ————————

VʔC and VCʔ --------

The vowel begins at the vertical axis and ends at the vertical stroke.

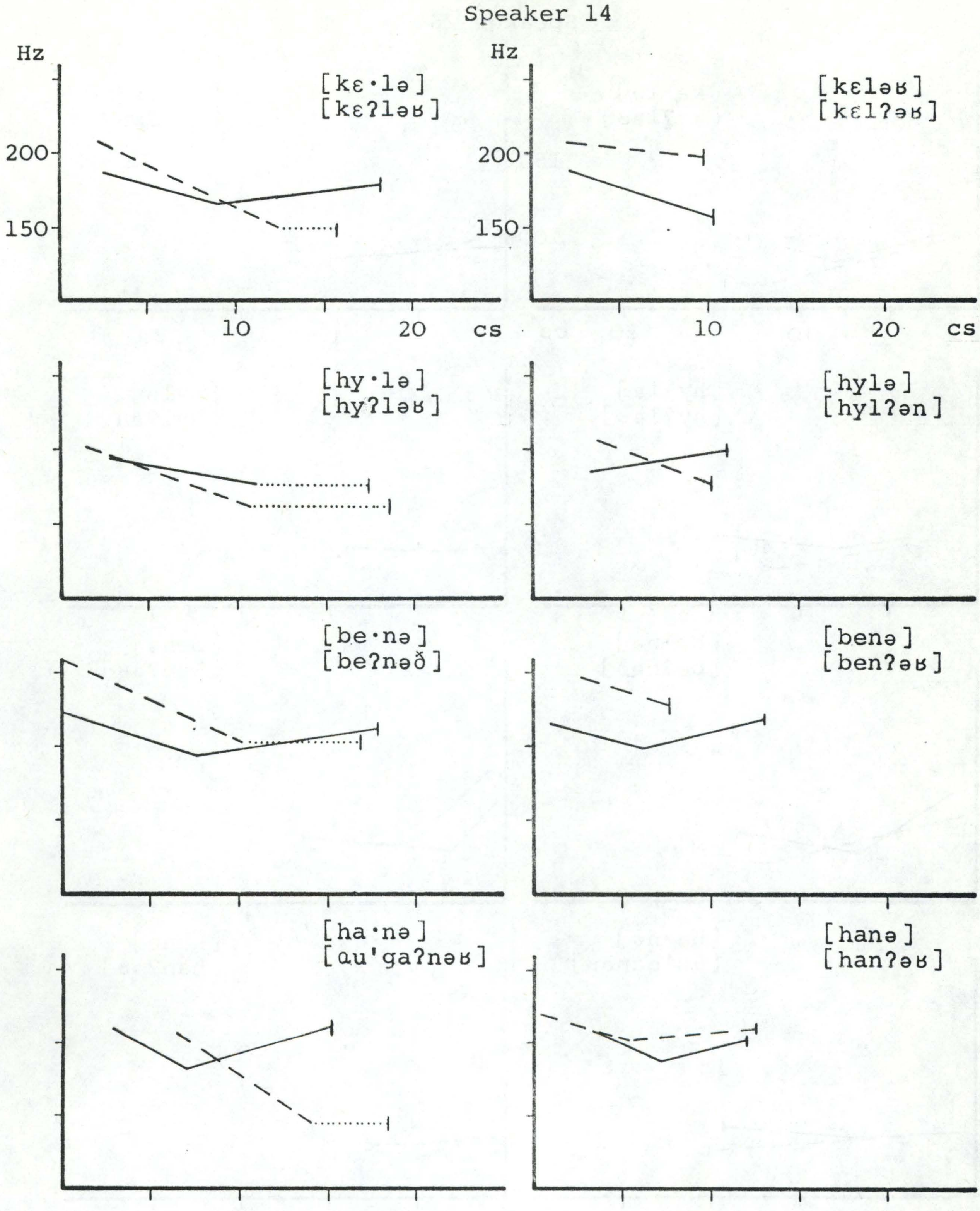


Fig. 11. Average pitch movement of the vowel in the word types with voiced consonants for speaker 14.

V:C and VC ————
VʔC and VCʔ --------

The vowel begins at the vertical axis and ends at the vertical stroke.

Speaker 25

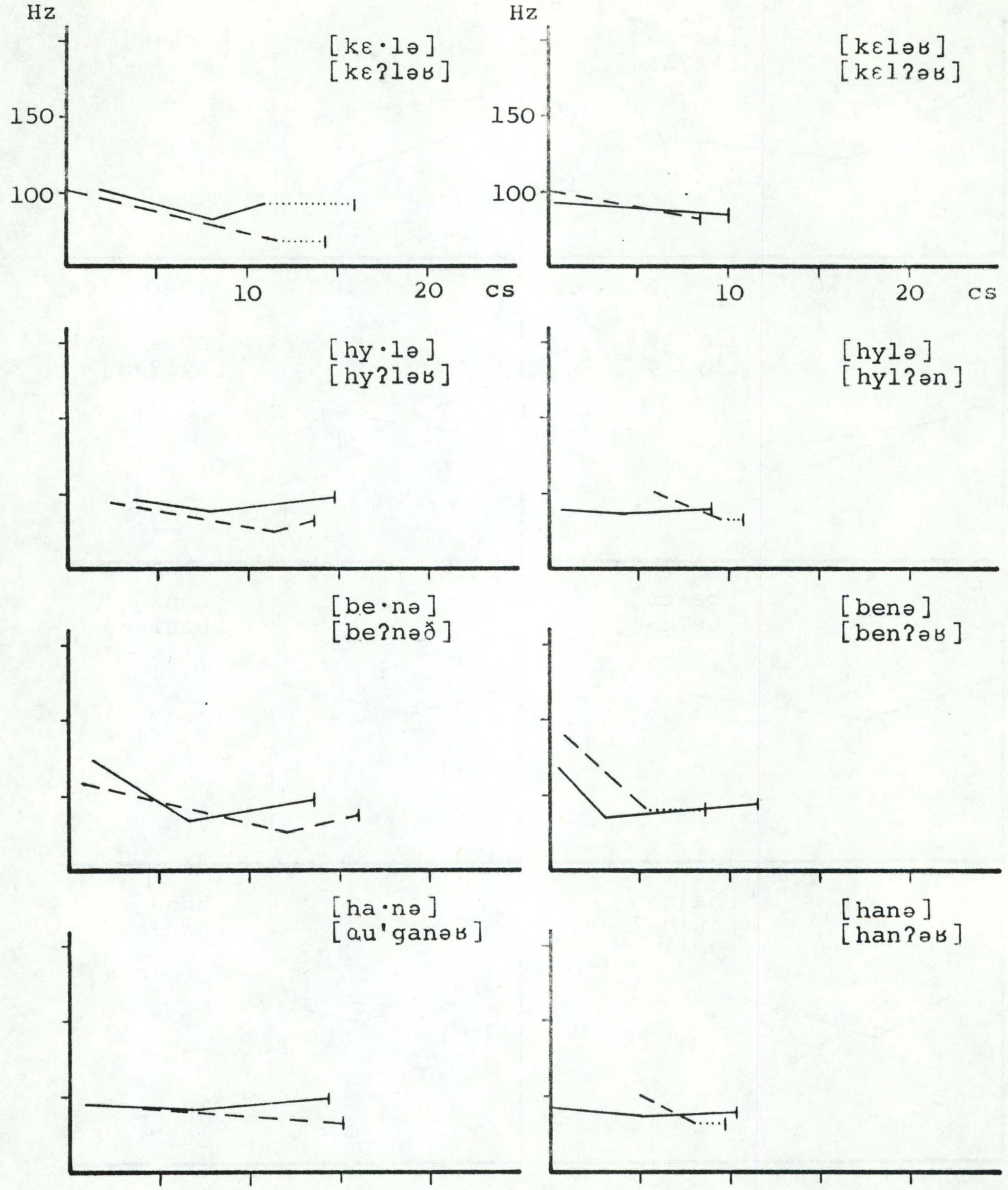


Fig. 12. Average pitch movement of the vowel in the word types with voiced consonants for speaker 25.

V:C and VC ———

VʔC and VCʔ --------

The vowel begins at the vertical axis and ends at the vertical stroke.

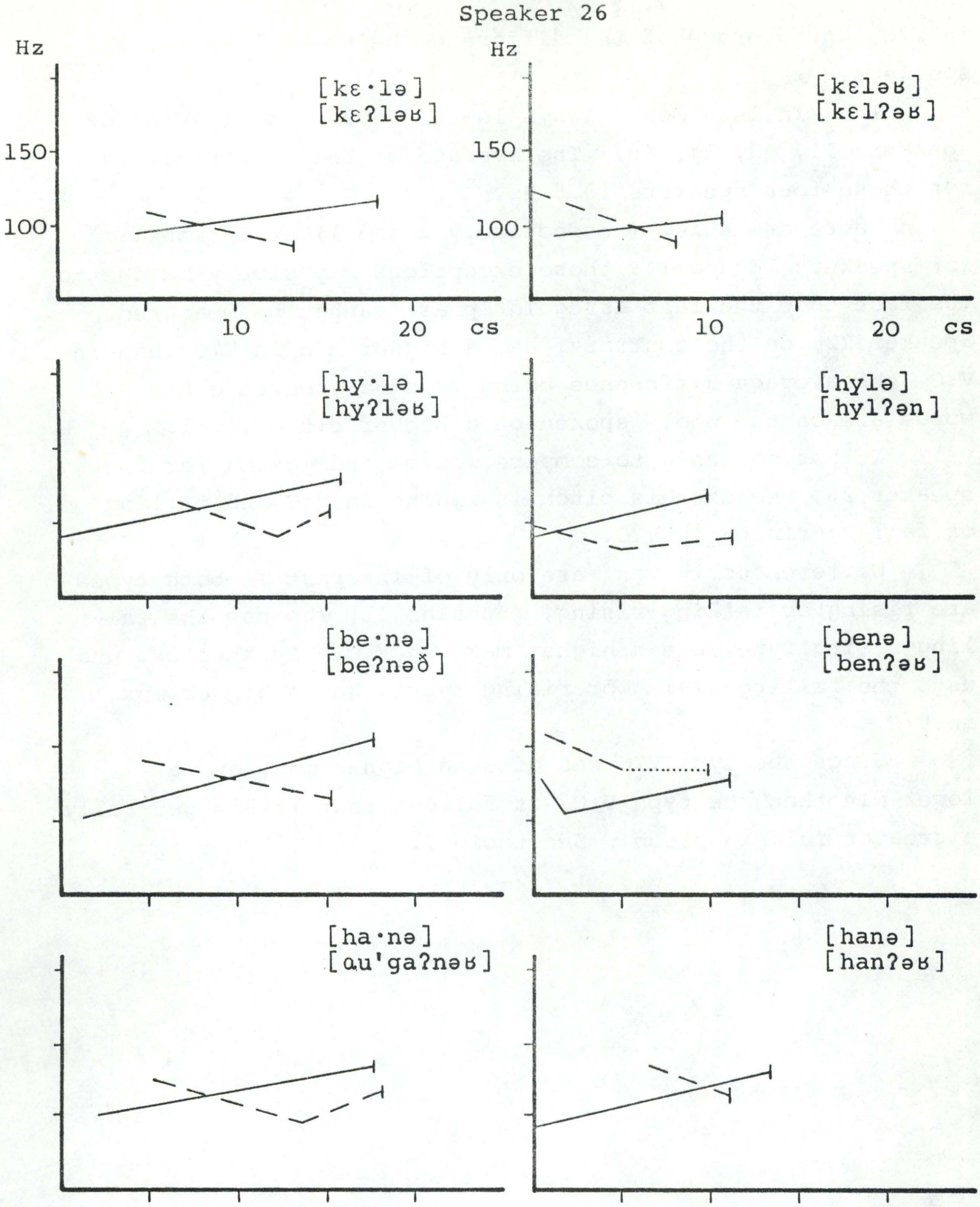


Fig. 13. Average pitch movement of the vowel in the word types with voiced consonants for speaker 26.

V:C and VC ————

VʔC and VCʔ -------

The vowel begins at the vertical axis and ends at the vertical stroke.

in V?C, the average of the difference being 10,2 Hz for all speakers.

The min is almost always lower in V?C than in V:C for speakers 11, 13, 14, 25. The average of the difference is for these four speakers 18,8 Hz.

There are a few (respectively 2 and 1) exceptions for speakers 13 and 14; these exceptions may simply be due to the fact that the last irregular phase cannot be measured. Speaker 22, on the contrary, has a higher min in V?C than in V:C (the average difference being 25,8 Hz), because his V?C words are on the whole spoken on a higher pitch level.

It has no sense to compare minima and maxima for speaker 26, because his pitch is rising in V:C and falling or falling-rising in V?C.

Differences in max_2 are only of interest if both types are rising or falling-rising. Speaker 11, who has the falling-rising type, has a higher max_2 in V:C. Speaker 22, who uses the falling-rising or rising types, has a higher max_2 in V?C.

Since the type V?C has often a higher max_1 and a lower min than the type V:C, it follows that it has generally a greater fall in pitch. See table 7.

TABLE 7

The average fall in V:C and V?C words (in Hz). N is the number of words which have a fall in V:C and V?C.

<u>Fall in V:C</u>						
speaker	11	22	13	14	25	26
average	25,5	14,3	29,3	20,8	13,9	-
N	45	24	48	48	48	-
<u>Fall in V?C</u>						
average	66,6	20,3	45,9	64,6	24,1	22,4
N	42	29	48	48	48	42

The fall is obviously greater in V?C than in V:C; the average difference is 26,2 Hz. The significance has been calculated by means of a pair test, and the difference was found to be significant at the 99% level.

The rise, which in some cases follows after the fall may be steeper in V?C (cp sp. 11), but it does not reach the same height as in V:C.

Speaker 22 differs from the others. In general the whole pitch curve in V?C is placed higher than in V:C. The curves are mainly rising, and the rise is greatest in V?C.

4.3.2. The pitch in VC and VCʔ

VC and VCʔ words were analysed in the same way as the V:C and VʔC words. It must be kept in mind that only the vowels have been measured. For illustrations see figs. 8-13.

4.3.2.1. The shape of the curves

In general the VC words resemble the V:C words, but the curve is often simpler and the rise and fall less extensive.

Speakers 11, 13, 14, 25 have the falling-rising or the falling type in VC words, the falling type dominates in VCʔ. Speaker 22 generally has the rising type in both VC and VCʔ, and speaker 26 has the rising type in VC and the falling type in VCʔ. The differences are thus approximately the same as for V:C vs. VʔC.

4.3.2.2. The pitch of max_1, min and max_2

Speakers 11, 13, 14, 25 have a higher max_1 in VCʔ, the average difference between VCʔ and VC being 14,8 Hz. No constant difference in min and max_2 was observed.

The fall will, therefore, be greater in VCʔ than in VC (see table 8). The fall is often followed by a rise of the same size in VC, but this is rarely the case in VCʔ.

TABLE 8

The pitch fall in VC and VCʔ. N is the number of words showing a fall. The total number of words is 24.

Fall in VC						
speaker	11	22	13	14	25	26
average	23,2	-	30,0	21,0	12,2	23,0
N	24	-	23	19	23	4
Fall in VCʔ						
average	52,2	-	30,5	23,2	22,2	19,0
N	24	-	23	21	22	21

Speaker 22 has not only a higher pitch level, but also a stronger rise in VCʔ. Speaker 26 has a rise of 25,8 Hz in VC and a fall of 19,0 Hz in VCʔ.

On the whole the general pitch level is higher in VCʔ than in VC for all speakers, as was also the case with the pitch level in the beginning of long vowels. An inclusion of the pitch of the following consonant might have shown a further fall in VCʔ words.

4.4. The intensity

There is a tendency toward a difference in intensity movement between V:C and VʔC. An exclusively rising intensity movement has not been observed in VʔC, but it can be found in V:C, and an exclusively falling movement is dominating in VʔC.

whereas both the falling and falling-rising types are found in V:C. See table 9.

TABLE 9

The tokens representing the different types of movements in V:C and VʔC.

	V:C	VʔC
exclusively rising	23	0
falling-rising	128	91
exclusively falling	137	187

It should be remembered that the falling type includes the rising-falling type, which is more common for intensity than for pitch.

Normally the first max is a little higher in VʔC. The average difference is 1,2 dB. The min is nearly always lower in VʔC, the average difference is 4,9 dB, i.e. the stød word starts higher and falls more, and there is rarely any rise.

The average fall in VʔC is 9,2 dB and in V:C 3,2 dB. See table 10. The significance has been calculated by means of a pair test. The difference was found to be significant at the 99% level. This is also valid for speaker 22.

TABLE 10

The fall in V:C and VʔC.

speaker	11	22	13	14	25	26
V:C	4,1	1,6	5,0	5,0	2,4	3,2
VʔC	11,6	5,4	14,1	11,8	5,0	7,5

In the cases where the fall in V?C is followed by a rise the end point will not be as high as in V:C.

4.4.1. The intensity in VC and VC?

The intensity movements of the vowel in VC and VC? words show differences similar to those found for V:C vs. V?C. There are more often falling curves in VC?, the max_1 is placed higher, and the fall is greater in VC? than in VC, except for speaker 22.

5. Summary and conclusion

5.1. The manifestation of the stød

There is a great difference between the manifestation of the stød for the various speakers. The two extremes are complete closure in the stød word and almost no visible difference between the stød word and the stød-less word.

5.2. The length

There is a significant difference between long and short vowel in words without stød. The stød vowel does not differ from the long vowel, but there is a tendency to reduce the length of the short vowel when it is followed by a consonant with stød.

There is no difference in the length of the consonant after long and short vowels. The length of the consonant with stød is greater than the length of the consonant without stød.

5.3. Pitch and intensity

There is a pitch difference between the stød words and the stød-less words (both VʔC vs. V:C and VCʔ vs. VC). One speaker (22), who had rising pitch in both stød words and stød-less words, had a higher pitch-level in the stød word than in the stød-less word. One speaker (26) had falling pitch in the stød word and rising pitch in the stød-less word. The rest of the speakers have a smaller fall in the stød-less words, often followed by a rise, whereas the stød words had a more extensive fall which is rarely followed by a rise. If there is a rise the end point is not as high as in the stød-less word. In addition the first maximum in the stød word is higher than in the stød-less word. The difference between the fall in V:C and VCʔ is significant.

The intensity movement shows less variability among the speakers. There is a significantly greater fall in VʔC.

Both pitch and intensity movement differ in the same way in V:C vs. VʔC and VC vs. VCʔ, which supports the assumption that the stød is a phenomenon belonging to the syllable (Martinet 1937).

This is confirmed by the fact that the distance from the start of the vowel to the start of the irregular oscillations is approximately the same in the two cases.

Pitch and intensity have normally a parallel movement, but the fall does not always start at the same time. Speaker 22 who has rising pitch both in VʔC and V:C, has falling intensity in VʔC.

According to Sv. Smith stød versus non-stød is due to a difference in the expiratory muscles accompanied by a passive reaction in the larynx.

This explanation does not fit very well with the instances of rising pitch combined with falling intensity found in the present material. Here the pitch movement seems to be independent of the intensity, and this may be an independent action of the larynx.

Abrahams, H. 1943: Nogle fonetiske iagttagelser vedrørende stødet i moderne dansk, In memoriam Kr. Sandfeld (København)

Ekblom, R. 1933: Om de danska accentarterna, Uppsala Univ.s Årsskrift 1933, III,5

Fischer-Jørgensen, Eli 1955: Om vokallængde i dansk rigsmål, NTTS XV, (Århus), p. 33 ff

Hjelmslev, L. 1951: "Grundtræk af det danske udtrykssystem med særligt henblik på stødet", Selskab for nordisk filologi. Årsberetning for 1948-49-50 (København), p. 12-24

Jespersen, O. 1949: Modersmålets fonetik

Lauritsen, M. 1968: A phonetic study of the Danish stød, Project on Linguistic Analysis, 7, 1968, University of California, Berkeley

Lehiste, Ilse 1972: Some observations concerning the third tone in Latvian. Papers in Linguistics and Phonetics to the memory of Pierre Delattre 1972, p. 309-15

Martinet, A. 1937: La Phonologie du mot en danois (= Bulletin de la Société linguistique de Paris 38, p. 169-266)

Rousselot, P.J. 1924: *Principes de phonétique expérimentale*

Selmer, E.W. 1925: Noen bemerkninger om stød og tonelag. *Heidersskrift til Marius Hægstad*, p. 94

Smith, Svend 1944: *Bidrag til løsning af problemer vedr. stødet i dansk Rigssprog*

Thorvaldsen, P. 1970: Construction of a segmentator. *ARIPUC* 4/1969, p. 8-15

Uldall, H.J. 1936: The phonematics of Danish, *Proc. Phon.* 1935, p. 54-57

Verner, K. 1903: *Afhandlinger og breve* ved M. Vibæk

IDENTIFICATION AND DISCRIMINATION OF CLOSELY SPACED SYNTHETIC VOWELS[1]

Peter Holtse

1. Introduction

It is commonly reported that human listeners will discriminate smaller differences in quality between vowels than between consonants. Further it is found that whereas there is a strong tendency for consonants to be perceived in a categorical way, i.e. discrimination is poor between qualities which cannot be identified absolutely, the tendency among vowels is towards continuous perception similar to the perception normally found with non-speech sounds. It has, however, been objected (Fischer-Jørgensen (1970-71)) that the relatively high discrimination between vowels might be caused by a difference in the auditory distances between vowel and consonant stimuli in the experiments. Thus it has been assumed by Stevens, Liberman, Studdert-Kennedy and Öhman (1969), and by a number of other authors who have used the same stimulus material, that if thirteen stimuli are placed along a continuum containing three phoneme categories the auditory distances between the stimuli must be approximately equal irrespective of the nature of the continuum. But it seems likely that the auditory distances between different vowel phonemes are actually somewhat larger than the distances between consonants since greater allowance must be made for social and personal variations in the pronunciation of vowels.

---

1) I am grateful to Eli Fischer-Jørgensen for much valuable criticism and many helpful suggestions during the writing of this article.

The aim of the present experiment was therefore to repeat the earlier identification/discrimination experiments with a set of closely spaced vowel stimuli in order to judge the possible influence from interstimulus distances. A second aim was an attempt to eliminate the effect from unevenly spaced stimuli which had been noticed in some earlier experiments (Hutters and Holtse (1972)).

2. The stimulus material

2.1. Formant frequencies

As material for the experiment was chosen a series of front unrounded vowels ranging in quality from [i] to [a]. The formant frequencies of the stimuli were chosen so that a line drawn through the stimuli in the F1/F2-diagram would pass through the areas normally taken up by the Danish long vowels /i:/, /e:/, /ε:/ and /a:/, (see Frøkjær-Jensen (1967) and Fischer-Jørgensen (1972)). In practice this was done by calculating the best fitting curves which would describe F2 and F3 of all the four vowels as polynomial functions of F1.

The stimuli were placed with equal logarithmic steps along F1. And the corresponding F2 and F3 frequencies were determined by the calculated functions. The formant frequencies of the stimuli are listed in table 1. And in fig. 1a F1, F2, and F3 of the stimuli are compared with mean formant frequencies from the data given by Fischer-Jørgensen (1972). In the calculation of these mean values those persons were excluded whose mean F2 of /i:/ was lower than the same person's F2 of /e:/. This particular formant pattern is characteristic

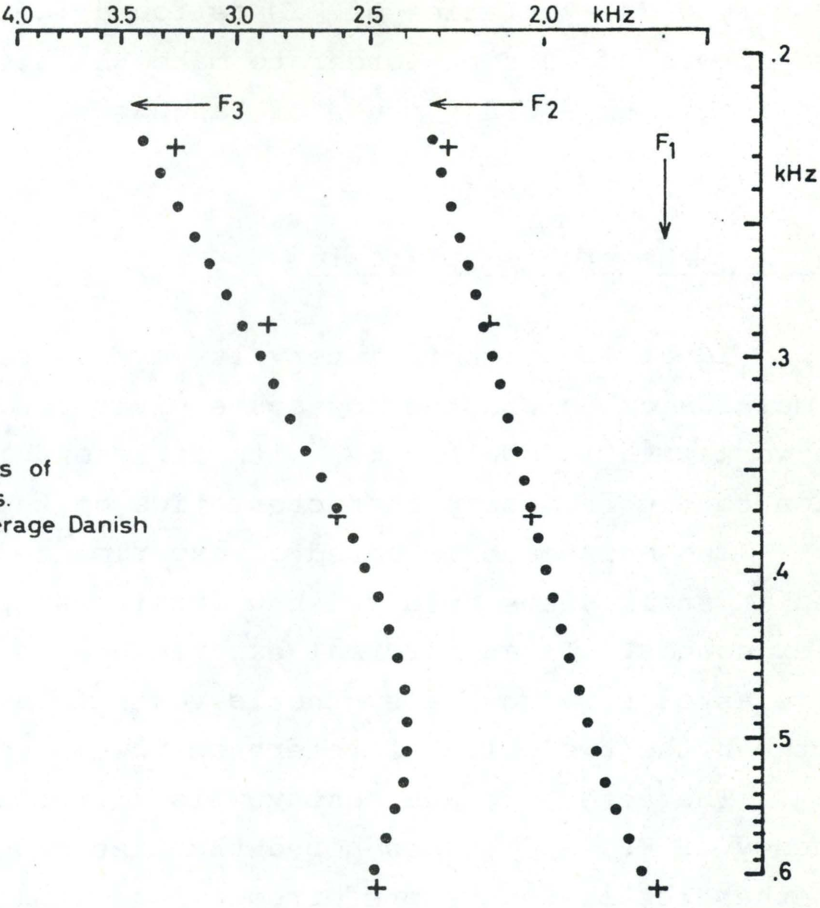


Figure 1a

Formant frequencies of 25 synthetic vowels. Crosses indicate average Danish vowel positions.

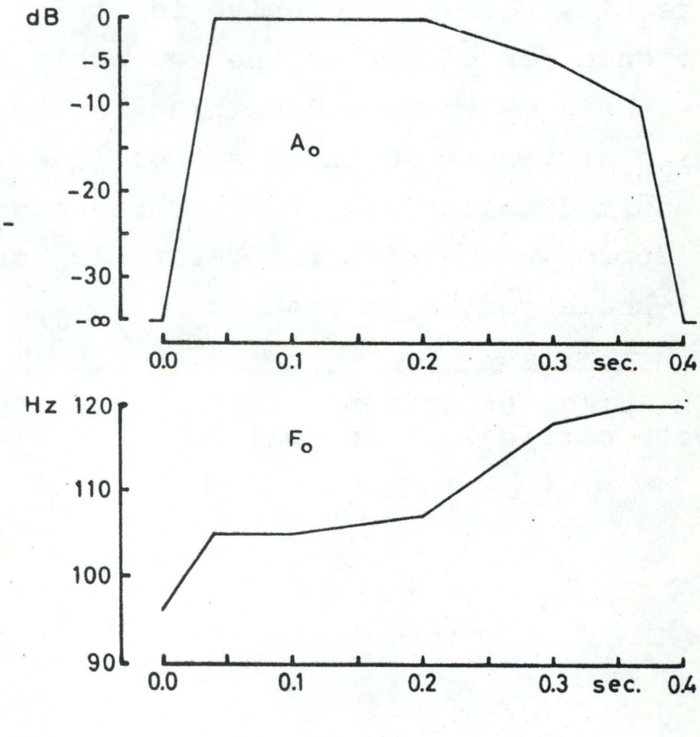


Figure 1b

Variations in total amplitude (A_o) and fundamental frequency (F_o)

of very close varieties of [i] as found in Swedish and Danish, and it was excluded in order to make the stimulus material applicable to a wider range of languages.

2.1.1. The problem of scale

Equal logarithmic intervals were chosen rather than e.g. a mel-scale because the log-scale gives convenient figures to work with in conjunction with a reasonably close approximation to the frequency characteristics of the ear.

The mel-scale is based on experiments with simple sounds and it still seems doubtful how far it is applicable to complex sounds. In an informal experiment,[1] a group of listeners were asked to adjust a synthetic vowel to a quality halfway between the two pairs of reference vowels [œ] - [ɔ] and [ɑ] - [a]. The reference and test vowels differed only in the frequency of F2, and it was hoped that the experiment would show whether the listeners preferred a logarithmic or a mel-scale. The results were inconclusive in so far as they revealed no preference for either of the two scales but showed rather a vague tendency towards something between the two. The patterning of vowels on the basis of judgements of similarity, as reported tentatively by Fischer-Jørgensen (1970-71), are best approximated with a logarithmic scale rather than with a mel-scale for the formants.

---

1) The pilot experiments reported in sections 2.1.1. and 2.1.2. were carried out in collaboration with Birgit Hutters.

TABLE 1

Control parameters of 25 synthetic vowels

(a) Fixed parameters:

Bandwidths: B1 = 50 Hz

B2 = 65 Hz

B3 = 100 Hz

B4 = 120 Hz

F5 = 4400 Hz

L5 = -25 dB

(b) Variable parameters:

Vowel No.	F1	F2 (L2)	F3 (L3)	F4 (L4)
1	226	2326 -26.2	3391 -19.2	3800 -20.2
2	235	2302 -25.7	3320 -20.3	3800 -21.6
3	245	2279 -25.0	3249 -21.7	3800 -22.6
4	255	2255 -24.2	3178 -21.7	3800 -23.6
5	266	2231 -23.4	3108 -22.1	3800 -24.3
6	277	2207 -22.6	3039 -22.3	3800 -25.1
7	288	2183 -21.8	2971 -22.4	3800 -25.7
8	300	2159 -20.9	2905 -22.4	3800 -26.2
9	313	2135 -20.1	2841 -22.2	3800 -26.7

TABLE 1

(continued)

Vowel No.	F1	F2 (L2)	F3 (L3)	F4 (L4)
10	326	2112 -19.2	2779 -21.9	3800 -27.1
11	339	2088 -18.3	2720 -21.6	3800 -27.4
12	353	2066 -17.4	2665 -21.1	3800 -27.7
13	368	2043 -16.5	2613 -20.6	3800 -27.9
14	383	2021 -15.7	2566 -20.1	3800 -28.1
15	399	2000 -14.8	2524 -19.5	3800 -28.1
16	416	1979 -14.0	2487 -18.9	3800 -28.1
17	433	1958 -13.2	2457 -18.4	3800 -28.0
18	451	1937 -12.7	2433 -18.0	3800 -28.0
19	470	1917 -12.0	2416 -17.6	3800 -27.7
20	489	1896 -11.5	2407 -17.2	3800 -27.4
21	509	1874 -11.2	2405 -17.1	3800 -27.0
22	530	1850 -10.9	2411 -17.1	3800 -26.6
23	552	1825 -10.7	2426 -17.1	3800 -26.1
24	575	1796 -10.5	2447 -17.3	3800 -25.6
25	599	1763 -10.2	2476 -17.9	3800 -24.6

2.1.2. Inter stimulus differences

As a pilot experiment the stimuli of Stevens, Liberman, Studdert-Kennedy and Öhman (1969) had been tried in a discrimination test with Danish listeners. The test was a 4IAX test where the listeners are asked which of two pairs, AA or AB, are different. According to Pisoni (1971) this type of test yields a higher proportion of correct discriminations than the traditional ABX test. This was confirmed in our experiment where the discrimination score was almost 100 pct correct.

The results of this pilot experiment clearly indicated that if the auditory distance between the individual vowel stimuli were to be of the same order of magnitude as the distances between the consonant stimuli of other experiments, the physical differences between the vowels must be very small. In the end the frequency difference between F1 of the individual stimuli was set approximately equal to 4 pct, which turned out to be rather close to a just noticeable difference. This small distance ensured that even comparatively efficient test forms could be employed. It was found that twenty-five vowels with a 4 pct difference in F1 would cover the range from /i:/ to /a:/. This number is quite high but necessary if the whole range is to be covered in one test.

2.2. Synthesis

The speech synthesizer of the institute is of the parallel type in which not only the formant frequencies but also the formant levels must be specified separately. In order to make sure partly that the synthetic vowels were close approximations to natural vowels and partly that there was a continuous transition from one stimulus to the next, the levels of the formants relative to the first formant were precalculated after the formulae of Fant (1956). The results of these calculations are listed in table 1.

The source spectrum of the parallel synthesizer is essentially flat in order to allow for independent variation of the individual formant levels. The glottal spectrum of a normal voice was therefore simulated by making appropriate adjustments in the precalculated formant levels. The simulated glottal source had a fall of 14 dB per octave rather than the usual 12 dB per octave, since this appears to be a better approximation to the natural voice. In the frequency region below F1 of any given vowel the glottal spectrum was shaped by a suitably chosen subformant (which is of course no true subformant).

The synthetic vowels were made with a total duration of 400 msec and had a rising pitch as shown in fig. 1b.

The changing fundamental frequency in conjunction with the relatively narrow formant band widths caused some trouble. Thus in some cases the movements of the harmonics through the formants gave the impression of a slight diphthongization which could possibly make these vowels more easily distinguishable than others. This difficulty was taken care of by synthesizing five series of 25 vowels with slightly different F_o-contours. For the production of the test tapes was chosen that series of 25 vowels which on the whole was the most uniform. The chosen series was, however, not perfect in all respects. Thus a slight irregularity in the middle of stimulus no. 4 was most annoying. More will be said about this later.

3. Procedure

3.1. The identification test

The 25 test vowels were arranged in two different quasi random orders. Care was taken to avoid having two vowels of fairly similar quality follow immediately after each other. In order to further diminish any possible context effect a flute signal was played immediately before the presentation of a new vowel stimulus.

Every stimulus was played twice with a pause of about one second between followed by a pause of about four seconds. No numbers or other identification marks were spoken on the test tape, but the flute signal between every fifth and sixth vowel stimulus was of a different character. Before the test tape proper were put five dummy stimuli which were not counted in the scoring. The tape contained a total of 55 stimulus "units" and lasted twelve minutes.

The copying from master tape to test tape was done on semiprofessional Revox tape recorders. The test was presented to the listeners from the same recorders over earphones (AKG, type K58).

The listeners were four trained phoneticians (one male), two first year phonetics students (female), and one (male) technician. They were all speakers of Standard Danish with no known hearing defects. The subjects were asked to identify the vowels with the Danish long vowel phonemes /i:/, /e:/, /ɛ:/, /aː/ as they appear in the words "mile, mele, mæle, male" ['mi:lə, 'me:lə,'mɛ:lə 'mæ:lə]. The subjects noted their identifications on specially prepared answer sheets. Every subject listened ten times to the test, thus giving a total of twenty identifications per stimulus for every person.

3.2. The discrimination test

The testform chosen was the so-called AX-type where two vowels are presented to the listeners who are asked to judge whether the test items are the same or different. According to Pisoni (1971), who compares various test-procedures, listeners could be expected to respond to smaller stimulus differences with the AX-test than with any other procedures. This of course was what the experiment was intended to reveal. One major difficulty was, however, that with this procedure it is quite difficult to control what differences the listeners are detecting, i.e. any accidental physical peculiarities in the test tape will be given undue prominence. This did cause some problems as will become apparent later.

A test tape was prepared in which each vowel was paired (a) once with each of the vowels one and two steps removed and (b) twice with itself. The stimulus pairs were randomized according to the same principles as in the identification test. In the beginning of the tape five dummy stimulus pairs were recorded, giving a total of 105 pairs in the test. There was a pause of about one second between the two vowels in a pair, and the stimulus pairs were separated by flute signals in a manner similar to that employed in the identification test.

Six listeners (identical with those who took the identification test minus one (male) phonetician) wrote their answers on specially prepared answer sheets. The test was rather long (23 minutes) and the listeners were therefore allowed to pause at a certain point in the middle of the test. Every subject listened ten times to the test, thus giving a total of ten judgments on each one and each two step discrimination pair.

In their judgments of differences the listeners were asked to choose between four categories: (++) "certain they are different", (+) "don't know, believe they are different", (-) "don't know, believe they are the same", and (--) "certain they are the same". As it turned out the test was quite difficult which gave some bias towards the minus side, the range of answers was, however, sufficiently wide for all four points to be used in the calculations.

For each discrimination pair was calculated the area under the ROC-curve, P(A). In this case the ROC-curve is the function which relates the probability of a given answer if the vowels were different to the probability of the same answer if the vowels were identical.

This means that for any pair of vowels, AB, the score of correct discriminations is corrected for "false alarms", i.e. judgements "different", to the control pairs AA or BB. If the correction was not included in the scorings of the AX-test the best discrimination score would be found with the listener who simply answered "different" to all the vowel pairs. When the correction is included this strategy of the listener will yield a P(A) value of 0.5, exactly the same as one would get by answering at random. In order to obtain maximum P(A) = 1.0 the listener must not only discriminate correctly between all AB pairs but must also answer "same" to all AA or BB pairs. (See e.g. McNicol (1971) or Robinson and Watson (1972).)

In view of the very limited number of judgements on each pair the non-parametric measure P(A) was chosen rather than the more generally used d'.

4. Results

In fig. 2 are shown the individual results from the two listening tests for each of the seven listeners. In fig. 3 are shown the mean results for the seven listeners. Through the points of this identification curve have been drawn least squares approximations to the best fitting normal ogives. And in table 2 are listed the corresponding 50 pct cross-over points and the standard deviations of the estimated normal distributions.

TABLE 2

Average identification curves. The table shows 50 pct cross-over points (L) relative to stimulus numbers and standard deviation of the corresponding normal distribution (s) expressed in stimulus steps.

	L	s
/i/-/e/	6.85	1.33
/e/-/ε/	12.77	0.97
/ε/-/a/	18.96	0.44

4.1. The identification test

It is interesting to note that the dispersion of the identification scores is markedly smaller than found in earlier identification experiments, e.g. Stevens et al. (1969) or Pisoni (1971). This might of course be due to the choice of listeners, although there was no systematic difference

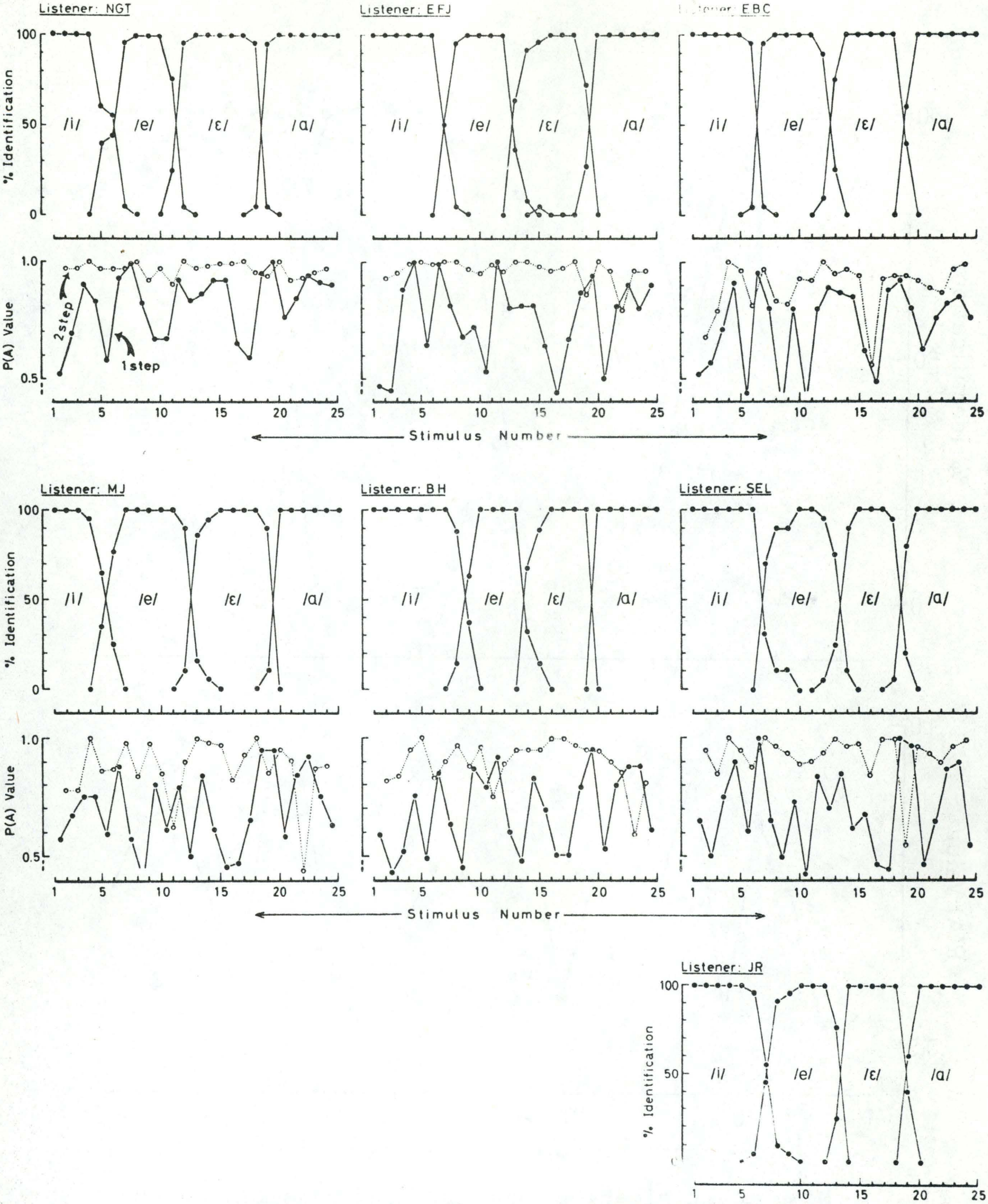


Fig. 2 Individual results from identification and discrimination tests for seven listeners. Listener JR did not take the discrimination test.

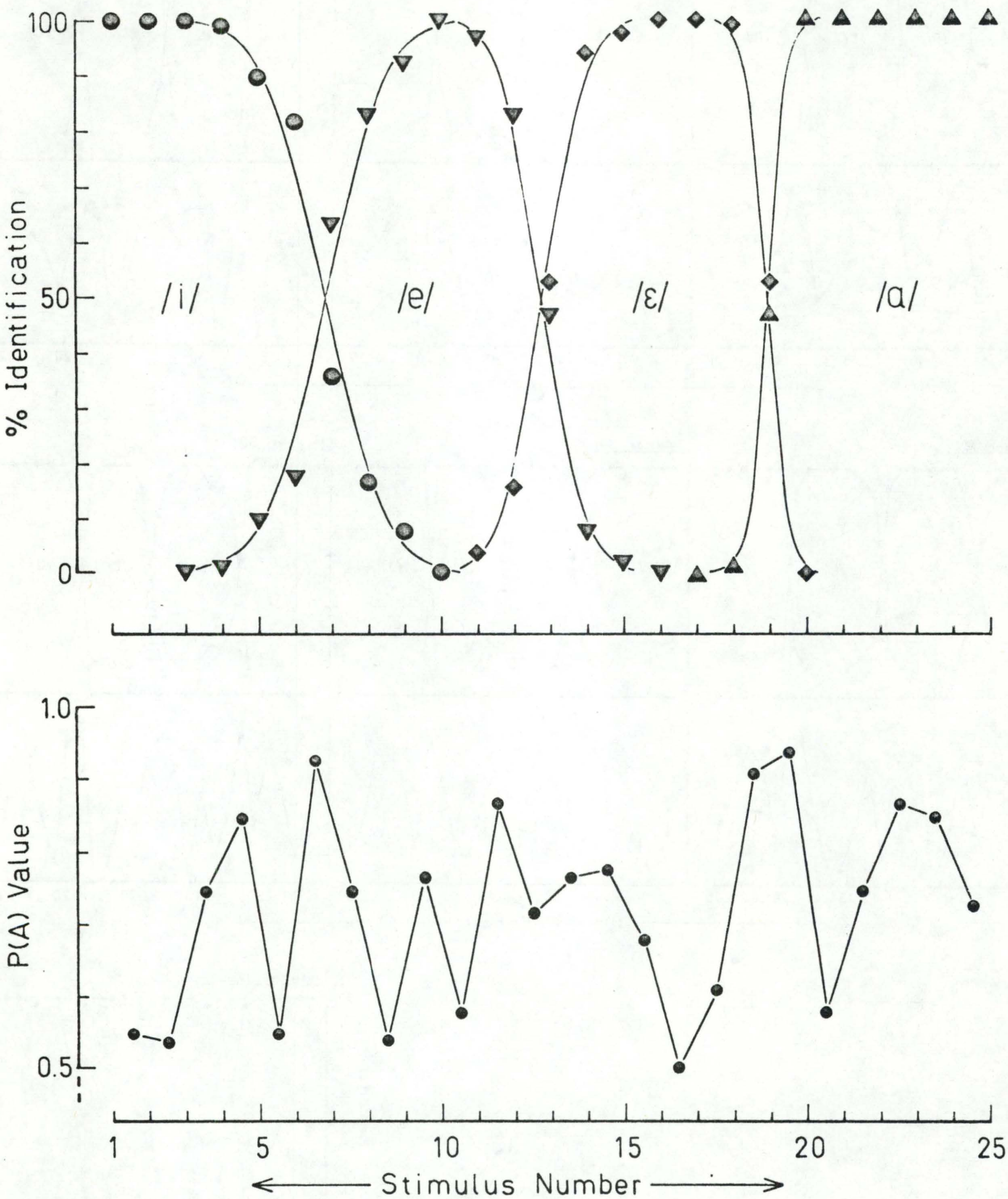


Fig. 3 Mean identification (top) and discrimination results (bottom) for seven listeners.

between the answers from the trained phoneticians and the other listeners. Another explanation could be that the stimuli of this experiment have sounded slightly more natural than the stimuli used with English speaking listeners. Thus in the experiments with English the difference in vowel length has generally been disregarded, although the difference in quality is always accompanied by a difference in duration. Therefore the short vowels /i/ and /ɛ/ may have been unnaturally long for their qualities and this may have disturbed the English listeners. In this respect Danish is an easier background to work with, since all the front vowels occur both long and short with the same qualities (Fischer-Jørgensen (1972)). The only exception is /a:/ which is a front vowel, while /a/ has a more retracted quality. The possible influence from this will be mentioned later.

In fig. 4 the 50 pct cross over points of the seven listeners are shown as lines in the F1/F2-F3 plot of the vowel stimuli. As in fig. 1 the crosses indicate average Danish vowel formant frequencies. Fig. 4 shows clearly that the listeners are in reasonably good agreement on the placing of their phoneme areas. Further it is evident that the phoneme areas of the identification experiment do not correspond 100 pct to the phoneme areas established on the basis of spectrographic measurements. Thus the border between identified /i:/ and /e:/ falls exactly in the middle of the measured /e:/-area. Similarly the /e:/-/ɛ:/ border coincides with the measured /ɛ:/-centre. Only the border between /ɛ:/ and /a:/ is situated in the expected area.

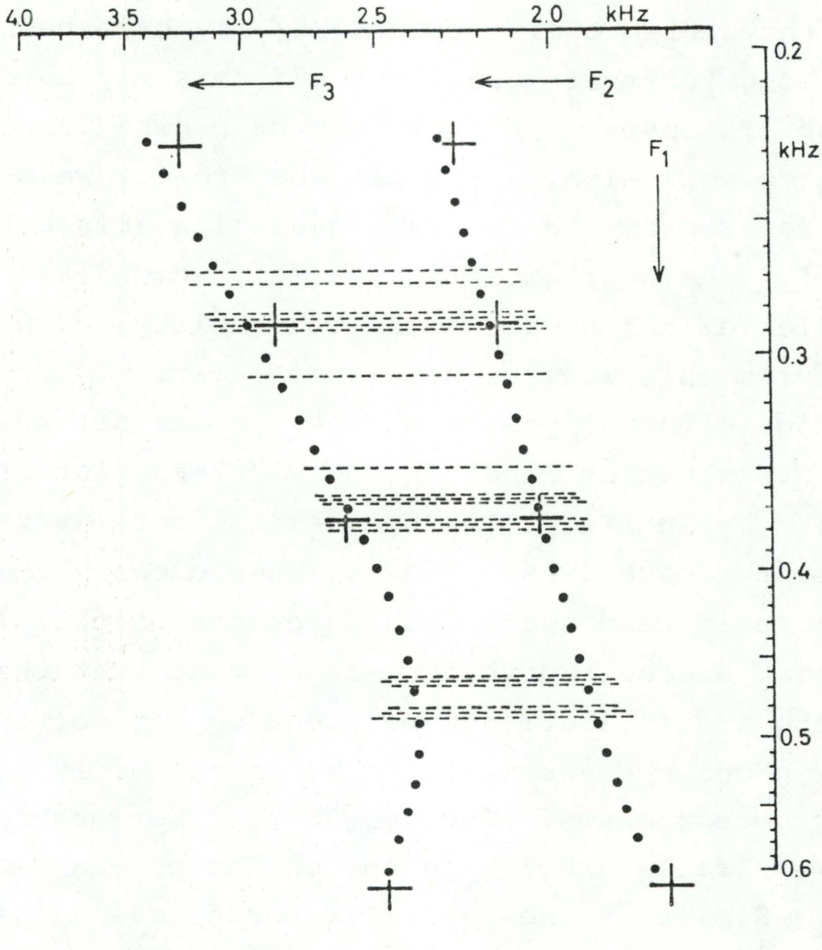


Fig. 4 Comparison between mean formant frequencies of natural vowels and identification of synthetic vowels. The broken lines in the F1/F2-F3 plot indicate 50 pct cross-over points of seven listeners. The crosses indicate average formant frequencies of Danish /i:/, /e:/, /ε:/, and /a:/.

4.1.1. The deviation of identified areas from measured areas

There are several possible explanations to this discrepancy between auditory and acoustic results. One would be that the listeners (of which the majority were women) are influenced by their own formant frequencies in their interpretations of the acoustic signal. This does not sound particularly likely since in natural speech people will interpret correctly almost any sort of voice no matter what their own voice is like. However, in order to investigate this point spectrograms were made of the four vowels in question as spoken by five of the listeners. The results of the measurements are shown in fig. 5. together with the 50 pct cross over points of the same persons.

For two persons, NGT and EFJ, the first formants of the identified and measured areas are in good agreement, for the remaining three persons the discrepancies are at least as great as between identified areas and average formant frequencies. Besides, the comparison of formant frequencies is, strictly speaking, only meaningful with the two male listeners SEL and JR, since the series of synthetic vowels do not pass through the phoneme areas of the female listeners. The fact that both male listeners have identified a series of synthetic vowels as /e:/ which in their own pronunciations would have been divided between /e:/ and /ɛ:/ and on the whole show the greatest discrepancies between identified and measured areas clearly refutes the hypothesis that identification should be affected by the listener's own formant positions.

Another possible explanation was of course a systematic bias in the formant frequency measurements. This seemed a likely hypothesis since the discrepancy between measured and identified vowels is greatest among the close vowels which are

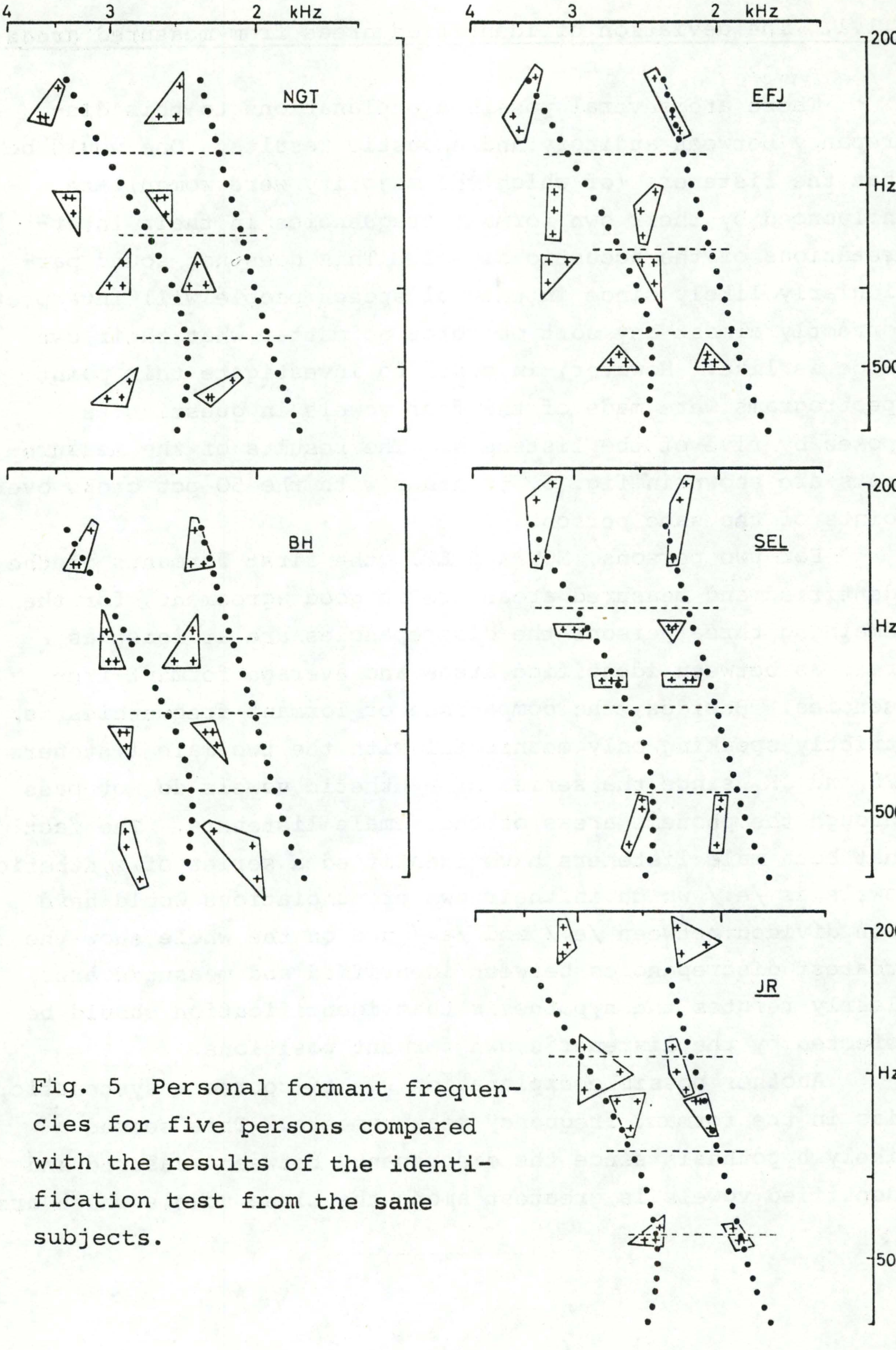


Fig. 5 Personal formant frequencies for five persons compared with the results of the identification test from the same subjects.

known to be difficult to measure accurately. In order to test this hypothesis Professor Eli Fischer-Jørgensen, who had measured the spectrograms on which the synthetic vowels were based, kindly consented to measure the formant frequencies of the stimulus vowels. The deviations of these measurements from the original formant frequencies were very small indeed and showed no systematic variations which could have caused the discrepancy between acoustic and auditory data.

Similar differences between formant measurements and identification tests can be seen in the results of Fry, Abramson, Eimas and Liberman (1962) where the border between English /I/ and /e/ in the identification test is found at about 500 Hz (Holtse (1972)). J. Rischel (unpublished) has found an identification border between Danish /i/ and /e/ at 270 Hz and between /e/ and /ɛ/ at 375 Hz. These findings are in good agreement with the results of the present experiment and equally hard to explain.

The apparent shift in the identification may have been caused by the balance of the frequency spectrum of the synthetic vowels. Thus too much energy in the frequency region below F1 would have the effect of shifting the perceived centre downwards, especially with low first formants. Another possibility, suggested to me by E. Fischer-Jørgensen, is that the listeners simply divide the whole series of stimuli into four parts. These problems will be looked into some time.

4.2. The discrimination test

The individual P(A) values for the six listeners are found in fig. 2. And the average values are shown in fig. 3. Generally speaking the two-step discrimination is good with a P(A) very close to 1.0 in most cases. There are no marked peaks to be seen, but some listeners have a few unmotivated valleys. However, these valleys are found in different places with different listeners. And except for one case (EBC) they do not correspond to valleys in the one-step discrimination. It seems likely therefore that these valleys are artifacts of the test. Because of the very high discrimination score of the two-step test this part of the results has been left out in the following discussion.

The one-step discriminations are more varied but also rather difficult to interpret. The curves show both peaks and valleys but one very disturbing observation is that the agreement seems to be better between the discrimination functions of the six listeners taken together than between the discrimination and identification functions of each individual listener. This may be due either to unnoticed inaccuracies in the test tape (close examination of the tape has failed to reveal any such errors) or to some inherent universal constraint as suggested by Stevens et al. (1969). The material of the present experiment is, however, limited and further experiments appear to be needed before this question can be answered, although the former solution seems the more probable. Especially since the peaks and valleys are not found in the same places in the two experiments. One possible source of error would be some sort of context effect in the test tape. Obviously the experiment should be repeated with several different orders of test items.

4.2.1. Learning effect

Two listeners were examined for learning effects in the discrimination test. This was done by calculating the average P(A) value of all the stimulus pairs from each listening session. The results of these calculations are listed in table 3.

TABLE 3

Average P(A) values from individual listening sessions for two listeners.

Listening session	Listener EFJ	NGT
1	0.71	0.81
2	0.80	
3	0.71	0.75
4	0.68	
5	0.72	0.86
6	0.72	
7	0.72	0.81
8	0.79	
9	0.80	0.83
10	0.82	

No systematic trend is revealed from table 3 but the variation in the average P(A) values is an indication of the maximal amount of reliability to be expected from the individual discriminations.

5. Discussion

5.1. The connection between discrimination and identification

In the average discrimination curves of fig. 3 there is a marked peak about stimuli 18-19. This peak corresponds well with the border between identified /ε/ and identified /a/. The peak is quite steep corresponding to the good agreement in the identification.

Between /i:/ and /e:/ there is another peak in the discrimination function: stimuli 6-7-8. But the picture is complicated by an extra peak between stimuli 3-4-5. However, as previously mentioned there was a small irregularity in stimulus No. 4. This irregularity has made No. 4 so easily recognizable that it would produce a peak in the discrimination.

There is no evident peak in the discrimination between /e:/ and /ε:/ but the discrimination is generally good from stimulus 9 to 15. Similarly the discrimination is rather poor in the middle of the phoneme areas in several cases, approaching P(A) = 0.5. However, we are left with a few unexplained peaks, e.g. between stimuli 9 and 10, and between 22 and 25. There appears to be no explanation for the peak between 9 and 10 which again may have been caused by external disturbances. The peak between 22 and 25 is situated in the area where the [a]-quality changes from that associated with long /a:/ to the quality of short /a/. Therefore this particular peak may well have been caused by influence from the identification system. Because of the questions asked in the identification test this peculiarity does not show in the identification.

5.1.1. The number of identifiable stimuli

In this connection it should be noted that the theory of categorical perception predicts that it is impossible to discriminate between qualities which cannot be identified in isolation. It has, however, always been tacitly assumed that only phonemes could be identified absolutely. This has never been proved. On the contrary it is a common observation that, at least for vowels, people will identify several different variants of the same phoneme, especially if these variants serve some social function.

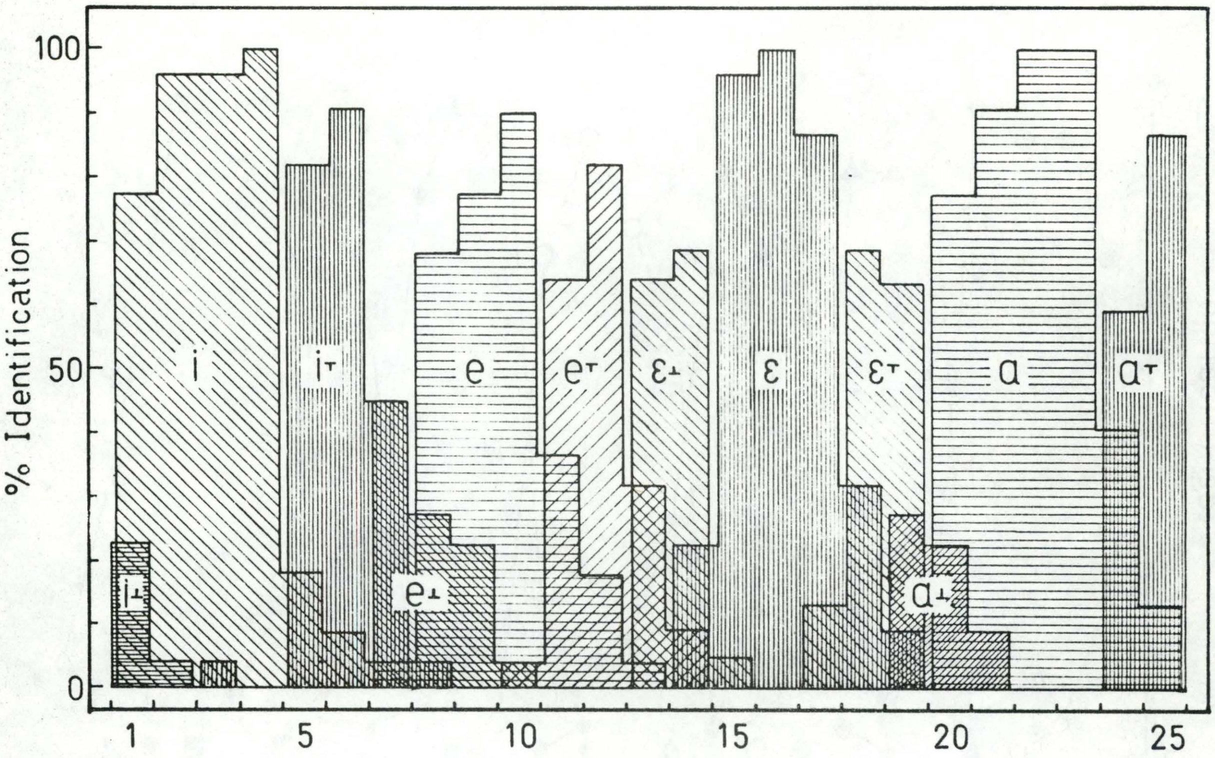
As a very preliminary investigation of this problem one listener (a trained phonetician) identified both a raised and a lowered quality of each of the four vowel phonemes. The results of this identification experiment are shown in fig. 6a.

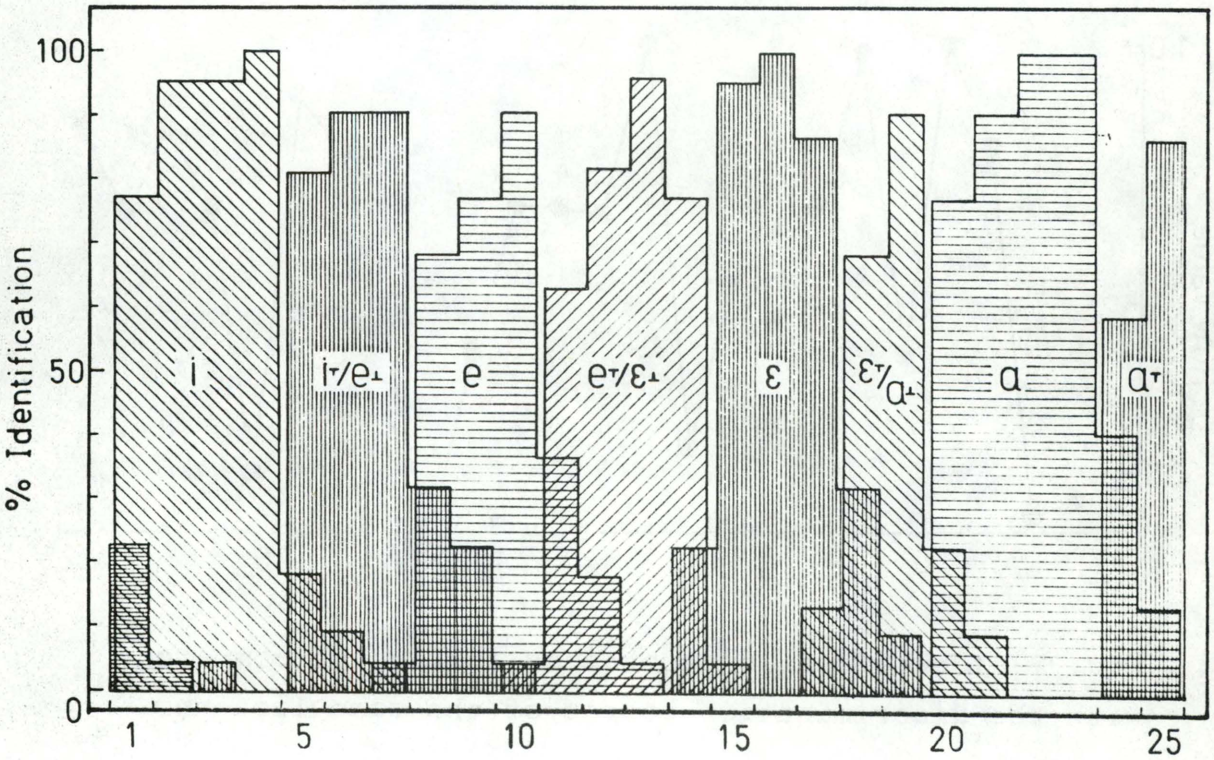
On the whole the identification of twelve different categories appears to have been quite difficult, and some categories have not been much used. Thus category [e˔] has a maximal identification score of 45.0 pct (stimulus 7), while [a˔] only reaches 27.2 pct (stimulus 19). The other between-phoneme categories all exceed 50 pct identification in their maxima but [ɛ˕] and [ɛ˔] do not reach 75 pct. (Category [i˔] is not included since the low score obtained was caused by inconsistent use of this category. The other extreme quality [a˕] was more frequently used, suggesting that, for this listener at least, stimuli 24 and 25 could well have been dispensed with in the original identification test.)

There appears to be a general tendency to use the raised category less often than the lowered category of the neighbouring phoneme, cf. e.g. [i˕] and [e˔]. In fig. 6b the neighbouring raised and lowered categories are combined. This results in a series of homogeneous identification categories all of which have maxima of at least 90 pct identification. This suggests that the listener was unable to identify consistently more than one category between phonemes. This observation may have implications for fine phonetic transcription but the problem is very much in need of further investigation. The possible effect of extending or reducing the range of stimuli should for instance be examined.

A comparison between the maximal identification function and the discrimination function (fig. 7) reveals little connection between these two functions. This is another point in need of investigation.

Fig. 6 a (top): Identification scores for twelve vowel categories, i.e. raised (⊥) and lowered (⊤) and normal variants of Danish /iː/, /eː/, /ɛː/, and /aː/.
b (bottom): Combination of neighbouring raised and lowered categories of fig. 6a.


100
50
0
% Identification
i
iᵀ
e
eᵀ
ε⊥
ε
εᵀ
a
aᵀ
i⊥
e⊥
a⊥
1
5
10
15
20
25


100
50
0
% Identification
i
iᵀ/e⊥
e
eᵀ/ε⊥
ε
εᵀ/a⊥
a
aᵀ
1
5
10
15
20
25

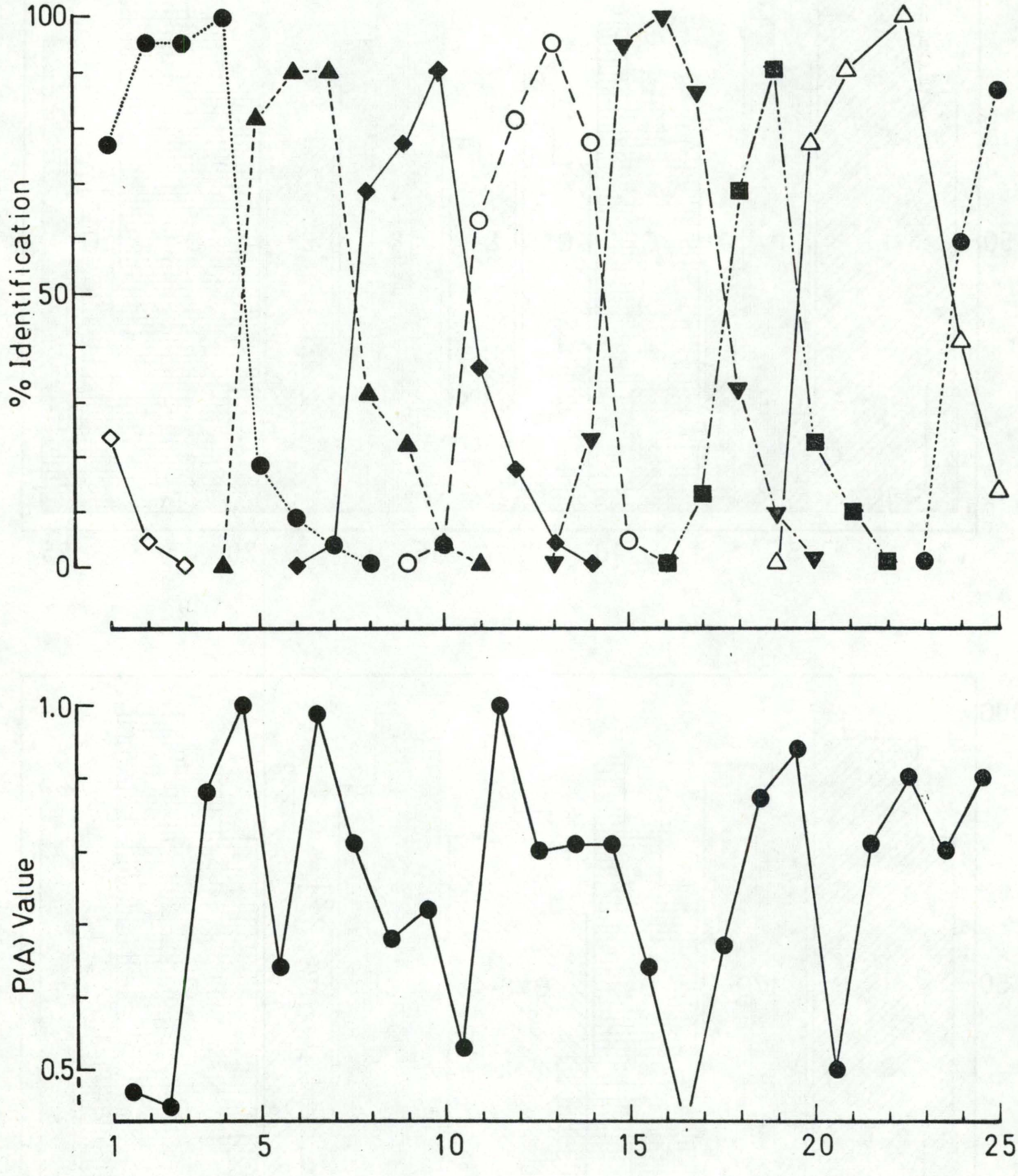


Fig. 7 Identification scores as in fig. 6b compared with the discrimination scores of the same listener.

5.2. Categorical perception

One interesting observation is that the valleys of the discrimination curve of fig. 4 are quite deep. In this respect the results of the present experiment differ from earlier discrimination experiments. Thus in the cases where the ABX-test had been used the percentage "correct discrimination" was usually about 70 for the valleys of vowels and considerably lower for the valleys of the consonants. The theory was therefore advanced that only consonants are perceived categorically. Pisoni (1971) quotes a theory by Sawashima and Fujusaki which suggests that the question whether a given sound continuum shows signs of categorical or continuous perception is determined by the way the incoming information is stored and processed by the brain. Thus most of the acoustic signal is kept for some time in a sort of temporary analogue storage. Some of the information, more particularly the information concerning consonants, is immediately converted to categorical information and only stored in this form. Now, in a discrimination task the listener will, unless he is forced to do otherwise, compare the stored signals which are closest to the original acoustic signal. In the case of consonants the closest approximation appears to be the categorized signal, which of course means that the results of a discrimination test will show signs of categorical perception. In the case of vowels the closest approximation is apparently the signal kept in temporary analogue storage. And the discrimination function will only be determined by the general difference limen for quality differences (if any such thing exists).

In order to make the discrimination of vowels show signs of categorical perception it would be necessary to block admission to the analogue storage. This seems to be what happens when vowels are put in context, as shown by Stevens

(1968). Another way would be to make the difference to be examined so small that even the signal stored in analogue fails to reveal this difference. In this case it will only be possible to detect a difference if the two vowels to be compared have been given different representations in the categorized storage.

The latter procedure is the one adopted for the present experiment. The procedure was successful in so far as there were tops and valleys in the discrimination function. But the expected picture of categorical perception was seriously contaminated by minute physical irregularities in the test tape. In the suggested model any such irregularity will of course be available in the analogue signal and thus produce a high discrimination without passing through the categorized storage. It appears therefore that for the procedure of the present experiment to be successful a very high precision is required in the synthesized material. At present it is difficult to distinguish categorization from minor accidents in the synthesis process.

References

Fant, G. 1956: "On the predictability of formant levels and spectrum envelopes from formant frequencies" FRJ p. 109-120

Fischer-Jørgensen, E. 1970-71: Resume af forelæsninger over psykoakustik og auditiv fonetik, (Abstract of lectures on psychoacoustics and auditory phonetics) Institute of Phonetics, Copenhagen (mimeographed)

Fischer-Jørgensen, E. 1972: "Formant frequencies of long and short Danish vowels" Studies for Einar Haugen, The Hague, p. 189-213

Fry, D.B., A.S. Abramson, P.D. Eimas, and A.M. Liberman 1962: "The identification and discrimination of synthetic vowels" LS 5, p. 171-189

Frøkjær-Jensen, B. 1967: "Statistic calculations of formant data" ARIPUC 2, p. 158-169

Fujisaki, H. 1971: "A model for the mechanisms for identification and discrimination of speech sounds" Proc. Acoust. 9, p. 56-59

Guilford, J.P. 1954: Psychometric Methods, London

Holtse, P. 1972: "Spectrographic analysis of English vowels" ARIPUC 6, p. 1-48

Hutters, B. and P. Holtse 1972: "On universals in vowel perception" ARIPUC 6, p. 177-184

Lane, H.L. 1965: "The motor theory of speech perception: A critical review" Psychological Review 72, p. 275-309

Liberman, A.M., K.S. Harris, H.S. Hoffman, and B.C. Griffith 1957: "The discrimination of speech sounds within and across phoneme boundaries" J. Exp. Psychol. 54, p. 358-368

Mattingly, I.G., A.M. Liberman, A.K. Syrdal, and T. Halwes 1970: "Discrimination in speech and non-speech modes", Haskins SR 21/22, p. 99-131

McNicol, D. 1970: A primer of signal detection theory London

Pisoni, D.B. 1971: "On the nature of categorical perception of speech sounds", Supplement to Haskins SR

Robinson, D.E. and C.S. Watson 1972: "Psychophysical methods in modern psychoacoustics", Foundations of modern auditory theory vol. II (J.V. Tobial ed.), p. 99-131

Stevens, K.N. 1968: "On the relations between speech movement and speech perception", Zs.f.Ph. 21, p. 102-106

Stevens, K.N. A.M. Liberman, M. Studdert-Kennedy, and S.E.G. Öhman 1969: "Cross-language study of vowel perception", LS 12, p. 1-23

Studdert-Kennedy, M., A.M. Liberman, K.S. Harris, and F.S. Cooper 1970: "Motor theory of speech perception: A reply to Lanes's critical review", Psychol. Review 77, p. 324-249

FALLING WORD TONES IN SERBO-CROATIAN (II)

Per Jacobsen

In an earlier paper (ARIPUC 4, 1969) I examined the falling word tones in Serbo-Croatian. The purpose of that investigation was to examine whether wordforms with two falling tones differ first and foremost in length, or whether the highly different tone contours are in themselves sufficient to establish a distinction between the tones.

At the time when my first investigation was made, there was no possibility of making experiments with synthetic speech at the Institute of Phonetics, so it was made in a rather primitive way, the stimuli being produced by a linguistically trained native speaker. However, it is now possible to produce intelligible synthetic words and thereby to produce sets of stimuli for listening tests by keeping certain parameters constant whilst others are varied.

Thus 24 synthetic words of the type grad were produced, the parameter of tone being kept constant, whilst the parameters of duration and intensity were varied, and vice versa. The word grad in Serbo-Croatian means, when pronounced [grȃ:d], town, and when pronounced [grȁd], hail. The tone movement of the former is a rising-falling curve with a peak in the first third of the vowel. The latter has no characteristic tone movement, but may be rising-falling (with a peak in the middle of the vowel), rising, falling, or level according to the sentence intonation.

The 24 synthetic words were recorded on tape in a quasi-random order and played back to a number of native speakers who were to decide on the meaning of each test item. The tone movements, durations and intensities of the synthetic words are given in table 1.

TABLE 1

[grἀd] with various duration, tone, and intensity movements

	duration in cs.	tone in cps. beg.	peak	end	intensity movement
1.1	10	100	130	75	level
1.2	15	-	-	-	-
1.3	20	-	-	-	-
1.4	10	110		100	-
1.5	15	-		-	-
1.6	20	-		-	-
2.1	10	100	130	75	+ 3 db
2.2	15	-	-	-	-
2.3	20	-	-	-	-
2.4	10	110		100	-
2.5	15	-		-	-
2.6	20	-		-	-
3.1	10	100	130	75	- 3 db
3.2	15	-	-	-	-
3.3	20	-	-	-	-
3.4	10	110		100	-
3.5	15	-		-	-
3.6	20	-		-	-
4.1	26	100	130	75	level
4.2	-	-	-	-	+ 3 db
4.3	-	-	-	-	- 3 db

TABLE 1
(continued)

	duration in cs.	tone in cps. beg.	peak	end	intensity movement
4.4	26	110		100	level
4.5	-	-		-	+ 3 db
4.6	-	-		-	- 3 db

As for the stimuli presented under 4.1 - 4.6, all listeners decided without hesitation on the meaning town, whereas there was a clear uncertainty about the items 1.2, 1.5, 2.2, 2.5, 3.2, and 3.5, that is, the words with a duration of 15 cs, but with varying tone and intensity movements. There was a general agreement on the meaning hail of the words with a duration of 10 cs (i.e. items 1.1, 1.4, 2.1, 2.4, 3.1, and 3.4) and on the meaning town of the words with a duration of 20 cs (i.e. items 1.3, 1.6, 2.3, 2.6, 3.3, and 3.6) (both with a variety of tone and intensity movements).

The conclusion to be drawn from these data must be the same as in my earlier paper, that is, that the other parameters do not seem to have any effect on the identification. The distinctive feature of the falling word tones in Serbo-Croatian is duration, not tone movement. The intensity movement does not play any role in the identification, which was one of the open questions in the earlier investigation.

References

Jacobsen, Per 1968: "The word tones of Serbo-Croatian, an instrumental study" ARIPUC 2, p. 90-108

Jacobsen, Per 1970. "Falling word tones in Serbo-Croatian" ARIPUC 4, p. 81-88

Lehiste, Ilse and P. Ivić 1972: "Experiments with synthesized Serbocroatian tones" Phonetica 26, p. 1-15

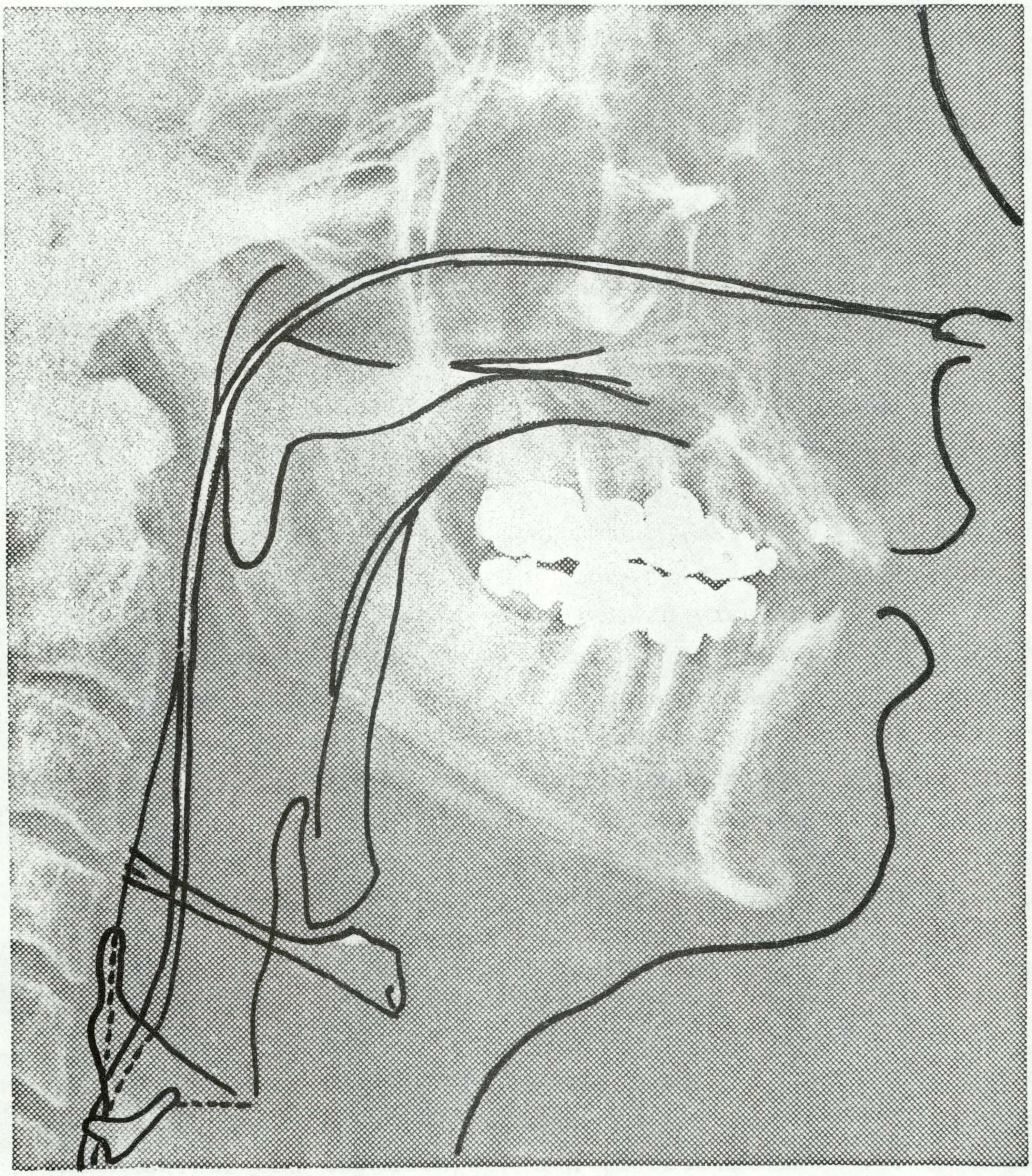
A GLOTTOGRAPHIC STUDY OF SOME DANISH CONSONANTS

Børge Frøkjær-Jensen, Carl Ludvigsen, and Jørgen Rischel
(1971)

This is a reprint of a paper published in the festschrift to Eli Fischer-Jørgensen (Form and Substance, 1971). We have considered it expedient to include the paper in ARIPUC, since other papers on our glottographic research at the Institute of Phonetics have been published in this series. For ease of reference the paper is reprinted virtually without any changes. This means, of course, that the most recent research is not taken into consideration.

It is necessary now to state that the assumptions put forward in the last part of the paper must be taken with a good deal of reservation. Thus, EMG-recordings of the glottis muscles made by Eli Fischer-Jørgensen at the Haskins Laboratories give evidence that there is some activity also in Danish bdg (personal communication). The terms "passive gesture" and "active gesture", used to distinguish between the glottal conditions of bdg and ptk, may thus not be quite appropriate.

In order to throw further light on the problems dealt with in our papers on glottography we are going to work with fiberoptics so that the two approaches can be confronted with each other. Our laboratory has now also got the instrumentation necessary for EMG-recording.



X-ray photograph of a subject (BFJ) with the photo-transistor in position (some displacement of the photo-transistor either upwards or downwards from the position shown here can be tolerated without serious decrease of the signal amplitude).

1. Introduction

The acoustical and physiological properties of Danish stops and - to a lesser extent - fricatives have been the subject of several papers by Eli Fischer-Jørgensen.[1] Thanks to her research, Danish is among the languages in which such features as voicing and aspiration have been thoroughly examined. This is fortunate since Danish is typologically remarkable as far as the system of stop consonants is concerned. It has been found that the stops, i.e. the sets ptk and bdg, are all essentially unvoiced. The alveolar t is typically somewhat affricated but otherwise the distinction between ptk and bdg is in all essentials a matter of aspiration only. Measurements (including glottography) reveal a considerable and stable difference in the duration of the aspiration phase in the two sets of consonants.

The present paper is intended to contribute to this field of research by giving some data on the glottal behaviour in the different kinds of stops. This of course involves some kind of estimate of the degree of glottis aperture in the production of the different sounds. However we have found it especially worth-while to focus our interest on the temporal relationships among the various phases of the oral and glottal gestures involved, since these relationships have not received much attention in previous glottographic work.

---

1) Eli Fischer-Jørgensen 1954, 1959, 1963, 1967, 1969, 1970.

In the present paper we give data of the kinds mentioned above for some Danish consonants, and we also venture some hypotheses on the minimum set of laryngeal "commands" required in order to generate the observed phenomena. The study is limited to labials, partly because this simplifies the problems since there is no affrication present, and partly because the choice of labials minimizes the risk of artifacts in the glottograms stemming from tongue movements (cf. section 3 below).

2. The instrumental set-up

The glottograms dealt with in this paper were made by means of the photo-electric glottograph described by Børge Frøkjær-Jensen (1968, 1969). Like those used by Malécot (1965) and Ohala (1966) it has the light source placed below the glottis, the light variations being picked up by a phototransducer inserted through the nose into the pharynx. The light from a light source fed with D. C. current is lead through a 40 cm long plastic rod, which can be tilted so as to stand in any desired angle to the larynx. The transducer together with its terminals was contained in a protecting polyethylene tube which was flexible and transparent and had an outer diameter of only 3 mm. This enabled us to perform the experiments without local anaesthesia.

Information concerning the timing of the maximum glottis aperture in relation to the lip closure and the offset/onset of vocal fold vibrations was obtained by using the instrumental set-up shown in Fig. 1.

Some recordings were made to check the set-up for synchronization of the different channels. These recordings were supplemented by simultaneous recordings of intra-oral air pressure.

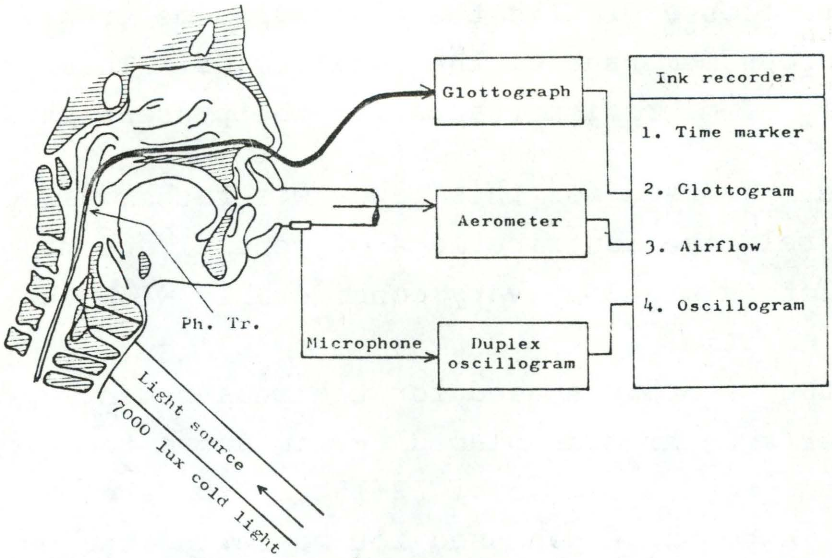


Fig. 1 Block diagram showing the instrumental set-up.

By means of the air pressure curve we have been able to register the time delay of the airflow curve caused by the inertia of the mechanical transducer in the aerometer. We have found this delay to be of the order of a few milliseconds, which is of the same order as the exactitude of the measurements.

The glottogram and duplex oscillograms were also checked for synchronization. We observed that the moment of first formant excitation was somewhat less than one millisecond delayed in relation to the moment of vocal fold closure as seen in high speed glottograms. If we assume, in accordance with Miller (1959), that the first formant excitation coincides with the closing of the glottis, this difference agrees with the acoustical transmission time form glottis to microphone.

3. Positioning of light source and transducer

We state below some of the more important sources of uncertainty in glottographic observations.

1. With closed glottis there occurs some transillumination of the mucous membrane of the vocal folds. This problem is presumably of minor importance for the present study.

2. The PhTr used for this paper was rather directional. Thus the relative contributions of central and marginal parts of the glottis area may vary considerably with the distance from the glottis.

In the recordings used for the measurements presented in this paper the PhTr was placed 6-7 cm above the glottis. This corresponds to a distance of 12-13 cm from the outer nostril. An X-ray photo of JR was used for measuring the position of the PhTr.

3. Since the variations in the width of the glottis slit differ in the front and the back parts of the entire glottis, the results may be expected to vary according to the location of the PhTr.

4. Tongue movements often misplace the PhTr in vowels with pharyngeal narrowing, which can reduce the signal level considerably. In the case of unfortunate positioning of LS and PhTr, artifacts may also arise in the articulation of velar consonants.

5. A further series of complications have to do with the location of the LS and the direction of the light-beam entering the trachea. With light entering the trachea well below the larynx (the light conducting rod being pressed against the upper edge of the sternum) and directed along the neck towards the glottis, we get traces which agree with our expectations about the gross variations in aperture of voiceless consonants. This, however, requires that the PhTr is fixed in a relatively stable position in the pharynx.

In accordance with Ohala (1966) we did this by fixing the PhTr somewhat above the lower end of the transparent protecting tube which was swallowed into the oesophagus. If the PhTr is fixed in this way the output is reduced considerably, but without such fixing in the oesophagus the output is extremely sensitive to tongue movements.

On the other hand, with a nearly frontal light beam entering the front side of the thyreoid cartilage we get a trace with relatively bigger deflections for each vibratory cycle of the voiced sounds, and with less disturbance from tongue articulations even if the PhTr hangs freely. The difference between these two extremes is illustrated in Fig. 2.

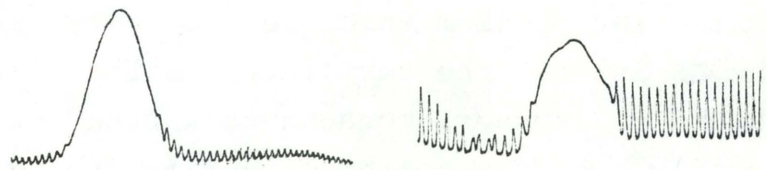


Fig. 2 Intervocalic p with LS placed in the two positions referred to in the text. (Subject JR)

This difference is understandable considering that the nearly horizontal light will pass between the vocal folds along their longitudinal axis and therefore be strongly modulated by small movements of the vocal folds and probably to a somewhat lesser degree by movements of the cartilaginous glottis. The opposite will be true with the light passing through the glottis from below at approximately right angles to the slit, especially if the PhTr hangs close to the back wall of the cavity, as it will probably do when the tube is fixed in the oesophagus. If the various aspects of the glottis behaviour are to be observed, it seems essential not to have the PhTr too close to the glottis. With the PhTr placed just above the cartilaginous glottis it is possible to register recordings where intervocalic voiced h shows a ripple-free peak of glottis aperture of the same size as that for aspirated stops and fricatives, whereas, in the more "normal" traces like those used for the present study (see below) h gives a curve with voice ripple and with less rise than for p and f.

In the present recordings the LS was placed in such a way that the light entered the neck slightly below the thyreoid cartilage. The curves obtained in this way are very similar to those obtained with a more "vertical" light beam, e.g. with a high degree of symmetry between the rising and falling parts of the curve for p, and we feel convinced that this is a reasonably realistic reproduction of the temporal variation in glottis aperture.

One particular problem which we face with certain placements of the LS is that the zero line[1] differs from one type of voiced sound to another, consonants and narrow vowels having generally a higher bottom line than open vowels. This phenomenon, which is most noticeable with frontal light, may partly be due to modulation of the light by slow up-and-down displacements of the vocal folds accompanying different degrees of oral constriction.

If this is true, it means that with small deflections of the glottogram there is no one-to-one relationship between glottis aperture and glottogram curve: a rise in the latter need not necessarily reflect an increase of glottis aperture. However, the shape of the vibratory cycle may be of help here. If the vibrations assume a nearly sinusoidal shape (as they do in voiced h), this is evidence that each fold vibrates freely, without touching during the vibratory cycle.

---

1) i.e. the level of the glottogram supposed to represent closed glottis.

3.6. The output level of the glottograph may vary due to slight variations in the exact position of the subject in relation to the light source. Care should therefore be taken that the subject is comfortably seated so that e.g. excessive heating from the light source does not cause changes in his position.

4. Extraction of parameters

On the basis of the three traces recorded on the ink-writer (see Fig. 1 above) we have extracted seven parameters which may serve to give details about the glottal behaviour.

1) The total duration of each consonant: lip closure, explosion, and aspiration in the case of stops; the interval of fricative noise in the case of f. - The beginning of the consonant was taken to be the point in the oscillogram where the amplitude of periodic vibrations is rapidly reduced to a very small value, or these have practically ceased (this definition of the starting point makes sense with all the unvoiced obstruents in question, and in any case the possible error in determining the moment of implosion will be largely the same for p and b). In stops the implosion is of course seen in the airflow curve, but this does not suffice with f. - The end of the consonant was defined as the moment when normal vocal fold vibration is resumed. This point is seen clearly in the oscillogram.

2) Duration of lip closure in the case of stops. This parameter was extracted from the airflow curve, and the results were checked by means of the duplex oscillogram.

3) Duration of the glottal opening phase measured from the moment of lip closure or labiodental constriction to the point of maximum height in the glottogram.[1]

4) Glottis aperture.

As for glottis aperture the information is qualitative only: roughly speaking, the curve goes up if the glottis opens, and vice versa, but the set-up cannot be calibrated with living larynges (van den Berg 1968). In order to get a numerical measure we have given the peak of each (fully

---

1) This is not a measure of the total duration of the glottal opening gesture. The duration of a complete abduction of the vocal folds is rather long (according to our glottograms it lies in the range 60-120 ms for sounds produced with large glottis aperture such as p, t, f, s), and it goes without saying that in rapid speech this comparatively slow movement is anticipated in the preceding sound. In some recordings of rapid speech by JR the entirety of the preceding voiced sound (vocoid) is characterized by a rise of the glottogram, i.e. increased leakage through the glottis. This anticipation of glottis opening has been observed earlier by Eli Fischer-Jørgensen (1963, p. 36) on the basis of flow and pressure measurements. - In the present paper we have made no attempt to determine the total durations of the glottal opening and closure movements (these are very difficult to define by mere inspection of curves), but it goes without saying that such measurements will be of paramount importance for the solutions of problems like these: (1) how is the closing gesture "triggered"?, (2) is the opening movement braked by the closing movement, i.e., is there a possible undershoot effect in the peak value? The answers to these questions must await future research.

aspirated) p in the text the prescribed value of 1.00 and normalized the peak value measured for other consonants with reference to the nearest p. This enables us to get an impression of the relation in size between the curves for different consonants.

Moreover, the following differences have been calculated and used as parameters:

5) Duration of lip closure minus duration of glottal opening phase (as defined above). This parameter gives a negative value if the glottal maximum occurs after the explosion; it gives a positive value if the maximum occurs before the explosion. The parameter refers to stops only.

6) We have calculated the relative values for parameter No. 5, i.e. as fractions of the total durations (parameter No. 1).

7) Finally, we have calculated the glottal closing phase of the consonant segment as the total duration minus the glottal opening phase (in the sense of parameter No. 3).[1]

With the fricative f the parameters reduce to:

1) Total duration (= par. No. 1 of stops).

2) Duration of glottal opening phase (= par. No. 3 of stops).

3) Glottis aperture (= par. No. 4 of stops).

4) Duration of glottal closing phase (= par. No. 7 of stops).

---

1) Like parameter No. 3 this is not a measure of the total lapse of time from one extreme of glottal state to the other. The end point of the closing movement cannot be seen clearly since the curve rounds off smoothly in the beginning of the following vowel.

5. The linguistic material

The opposition between aspirated stops (ptk) and unaspirated stops (bdg) in Danish is confined to certain positions, which may be subsumed under the common term "strong position". In other positions there is only one series of stops, but in return there is a distinction of stop : fricative not only in labials but also in alveolars and velars. For details see Rischel (1970).

We have considered it our primary task to provide reliable data on the distinction between p and b word initially in a stressed syllable, which is the strong position par excellence. As these consonants are both unvoiced, it is interesting to compare their glottal gestures to that found in a voiceless continuant, and we have therefore included words with f in the material presented in this paper. The consonants were represented by three series of minimally different, meaningful words: Peder-beder-feder (with long [e:]), pinde-binde-finde (with short [e]), and Pelle-bælle-fælle (with short [ɛ]). The words within each series were said in random order, each word occurring six times. The reason for confining the material to sequences with front vowels was that this minimizes the risk of disturbing effects from tongue articulation. The differences among the vowels in the three sets did not seem to influence the parameters of the consonant significantly.

Since we were interested in studying the combined opening and closing glottal gesture of the consonants, each word was said after a carrier sentence (Det hedder: "it is called") ending in a vocoid sound. Since this preceding sound is open and rather back, the choice of carrier sentence is questionable (see section 3 above on coarticulation artifacts). We made some control recordings with a carrier sentence ending in a front vowel (Du skal sige: "you must say"). These generally gave a more pronounced voice ripple for the preconsonantal segment, but the glottograms of the consonants were very similar in both cases, and there seems to be no reason to suspect the carrier sentence of distorting the recordings in any essential way.

In addition to these sets of words we have recorded a considerable amount of material with stops, fricatives and h in other positions. These include clusters of consonants with or without an intervening juncture, single consonants in word internal position, and consonants in word final position followed by a vowel. The position before juncture is a weak position; in absolutely final position there is more or less free variation between aspirated and unaspirated stop (the aspiration here can be explained as a boundary signal). We have symbolized the neutral labial stop in this type of position by B. Some minor parts of the additional material will be commented upon later in this paper.

TABLE 1

Subject: BFJ

	Parameter	Mean value	S.D.	No. of tokens	99% conf.
p	1 total duration	174.00 ms	12.08 ms	17	8.49
	2 dur. of lip closure	107.23 ms	10.79 ms	17	7.58
	3 dur. of glot. open.	117.47 ms	12.47 ms	17	8.76
	4 glottis aperture	1.00	0.00	17	0.00
	5 par. 2 minus par. 3	-10.23 ms	9.12 ms	17	6.41
	6 par. 5 div. by par.1	-0.05	0.05	17	0.03
	7 par. 1 minus par. 3	56.52 ms	8.00 ms	17	5.62
b	1 total duration	142.17 ms	21.30 ms	17	14.97
	2 dur. of lip closure	130.29 ms	21.71 ms	17	15.25
	3 dur. of glot. open.	51.94 ms	10.84 ms	17	7.62
	4 glottis aperture	0.20	0.05	18	0.03
	5 par. 2 minus par. 3	78.35 ms	17.32 ms	17	12.17
	6 par. 5 div. by par. 1	0.54	0.06	17	0.04
	7 par. 1 minus par. 3	90.23 ms	17.59 ms	17	12.36
f	1 total duration	173.33 ms	10.66 ms	18	7.23
	2 dur. of glot. open.	75.83 ms	6.52 ms	18	4.42
	3 glottis aperture	1.04	0.14	18	0.10
	4 par. 1 minus par. 3	97.50 ms	8.52 ms	18	5.78

The whole material was recorded (parts of it several times) with each of the three authors of this paper acting as a subject. We all speak Standard Danish, though with some minor (to some extent dialectal) differences. On visual inspection our glottograms are on the whole very similar.

6. Statistical treatment of glottis behaviour in initial p,b,f

In accordance with the research plan outlined above we measured a number of photo-electric glottograms of word initial p, b, and f and submitted the data to extensive statistical treatment. The results reveal certain properties of the temporal relationship between movements of the glottis and of the upper speech-organs, and they also show the durations of various phases of glottal activity. These general findings are further used as a reference framework in the study of unvoiced consonants in other environments.

TABLE 2

Subject: CL

	Parameter	Mean value	S.D.	No. of tokens	99% conf.
p	1 total duration	166.94 ms	15.82 ms	18	10.73
	2 dur. of lip closure	93.61 ms	8.36 ms	18	5.67
	3 dur. of glot. open.	96.66 ms	10.28 ms	18	6.98
	4 glottis aperture	1.00	0.00	18	0.00
	5 par. 2 minus par. 3	-3.05 ms	6.21 ms	18	4.21
	6 par. 5 div. by par. 1	-0.02	0.04	18	0.03
	7 par. 1 minus par. 3	70.27 ms	8.65 ms	18	5.87
b	1 total duration	130.00 ms	15.81 ms	18	10.72
	2 dur. of lip closure	113.05 ms	12.96 ms	18	8.79
	3 dur. of glot. open.	46.94 ms	11.89 ms	18	8.07
	4 glottis aperture	0.41	0.14	18	0.09
	5 par. 2 minus par. 3	66.11 ms	7.58 ms	18	5.14
	6 par. 5 div. by par. 1	0.51	0.06	18	0.04
	7 par. 1 minus par. 3	83.05 ms	9.09 ms	18	6.17
f	1 total duration	161.66 ms	15.99 ms	18	10.85
	2 dur. of glot. open.	75.27 ms	11.81 ms	18	8.01
	3 glottis aperture	0.86	0.14	18	0.09
	4 par. 1 minus par. 3	86.38 ms	13.59 ms	18	9.21

In order not to mix data from possibly different populations we carried out the statistics for each subject separately. Furthermore, we estimated whether data pertaining to the same consonant but taken from different words (e.g. Peder, pinde, Pelle) belonged to the same statistical population. As this seemed to be the case we did not find it necessary to maintain a statistical distinction between them.

Tables 1-3 show the statistical output for each of the three subjects, BFJ, CL, and JR.

TABLE 3

Subject: JR

	Parameter	Mean value	S.D.	No. of tokens	99% conf.
p	1 total duration	136.94 ms	7.88 ms	18	5.34
	2 dur. of lip closure	76.66 ms	6.41 ms	18	4.35
	3 dur. of glot. open.	81.94 ms	7.09 ms	18	4.81
	4 glottis aperture	1.00	0.00	18	0.00
	5 par. 2 minus par. 3	-5.27 ms	6.52 ms	18	4.42
	6 par. 5 div. by par. 1	-0.03	0.04	18	0.03
	7 par. 1 minus par. 3	55.00 ms	8.22 ms	18	5.57
b	1 total duration	100.55 ms	7.45 ms	18	5.05
	2 dur. of lip closure	84.16 ms	9.27 ms	18	6.29
	3 dur. of glot. open.	35.00 ms	7.85 ms	18	5.33
	4 glottis aperture	0.32	0.13	18	0.08
	5 par. 2 minus par. 3	49.16 ms	10.03 ms	18	6.80
	6 par. 5 div. by par. 1	0.48	0.08	18	0.05
	7 par. 1 minus par. 3	65.55 ms	9.05 ms	18	6.14
f	1 total duration	120.55 ms	13.16 ms	18	8.92
	2 dur. of glot. open.	44.16 ms	6.00 ms	18	4.07
	3 glottis aperture	1.01	0.17	18	0.11
	4 par. 1 minus par. 3	76.38 ms	9.04 ms	18	6.13

7. Results of the statistics

The data given in the preceding section show that the glottis adjustment in p differs in several respects from that of b:

(1) the aperture is much bigger in p than in b,

(2) p shows typically a nearly symmetric opening-closing change of the state of the glottis, whereas in b a relatively greater part of the total duration of the consonant segment is occupied by the closing movement,

(3) in p the moment of maximum glottis aperture falls rather close to (generally slightly after) the moment of explosion of the oral closure, whereas in b it comes much before the moment of explosion.

The shape of the glottogram for p is very similar to that for f (or s), i.e. aspirated stops and voiceless fricatives can be classified together as consonants with more open glottis, as against the unaspirated stop b, which is unvoiced like the others but produced with less open glottis.

Typical curves are shown in Fig. 3.

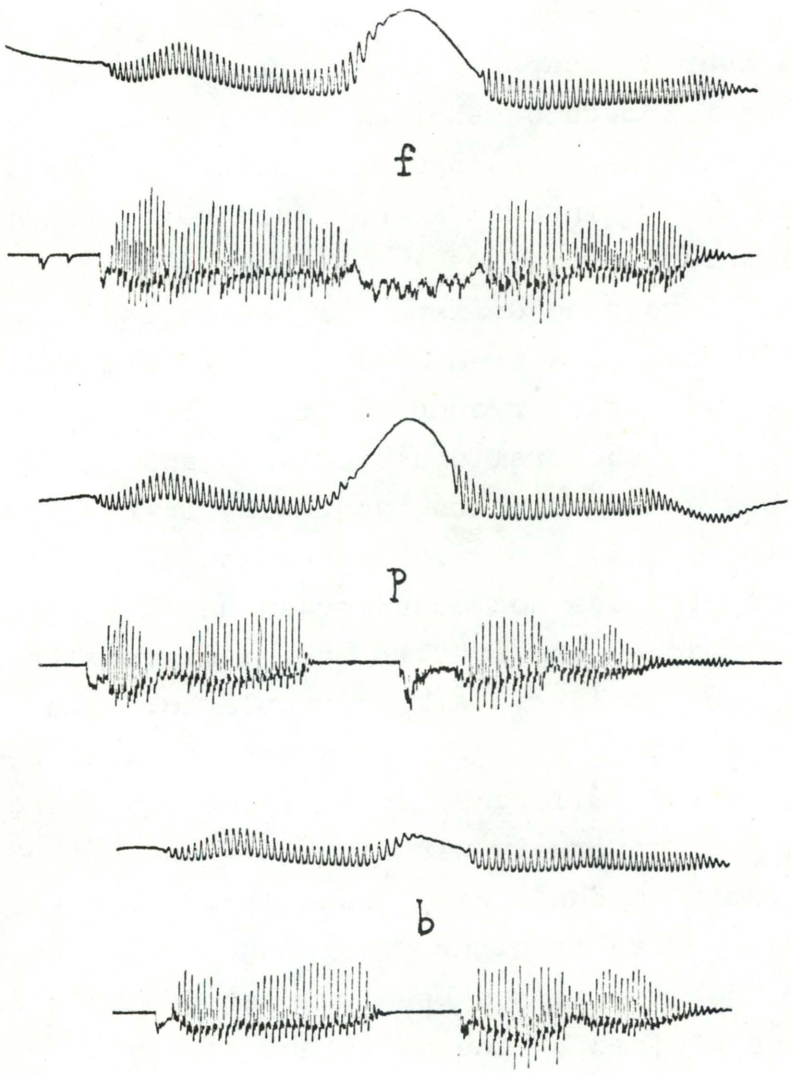


Fig. 3 Glottogram and duplex oscillogram of f, p, and b. Subject CL.

8. Degree of voicing

Inspection of glottograms of h in voiced environments shows that the overall shape of the glottogram is similar to that for p or f though of shorter duration and - under our "normal" conditions of recording - with less amplitude. It differs from p, b f in that it has full voice ripple throughout its duration.

The degree of voicing differs for the consonants in question. In h there is a free oral passage, and voicing is retained even with rather open glottis. In f the airstream is impeded by the oral constriction, and the voice ripple diminishes and disappears eventually in the rising phase of the glottogram, i.e. with abduction of the vocal folds. Voice ripple does not occur again until the glottogram has nearly reached the bottom line, i.e. until the adduction of the vocal folds is almost completed. Thus, although the glottogram on f is rather symmetric with regard to rise time and fall time, there is more voicing in the rising phase. Conversely with p. Here the voice ripple ceases early in the rising phase (shortly after the implosion), whereas it often starts again well before the closing of the glottis has been completed. This difference can be explained by the different aerodynamic conditions: in the glottis opening phase there is oral closure in p and therefore more braking of the airstream than in f; in the glottis closing phase there is almost unimpeded oral passage in p and hence less resistance to the airstream than in f. Voicing with partly open glottis is apparently a matter of a sufficient velocity of air (H. von Leden, 1961, p. 668), cf. voicing of intervocalic h.

In b the voice ripple vanishes quickly, as in p, which suggests that b is no more "lax" than p, i.e. there is no apparent adjustment of the conditions in either the vocal folds or the pharyngeal cavity allowing the vocal folds to vibrate in spite of the oral closure.[1] The onset of voice follows almost immediately after the explosion. This agrees with the fact that the glottis is practically closed at this moment.[2]

Obviously the degree of vocal fold vibration in the Danish consonants in question (in the position referred to) is a symptom rather than an independent physiological feature. It can be explained in a straightforward way as conditioned by the interplay between glottis aperture and oral constriction (also cf. Fischer-Jørgensen 1970, p. 151).

9. Hypotheses about laryngeal commands in p, b, f, h

The glottis aperture in b is much smaller than the aperture found in p, f, h. Moreover, the duration of the opening movement is very short compared to the closing movement. There is no apparent reason to assume that this gesture is effectuated by neural commands; it can probably be explained by the aerodynamic conditions at the transition from vowel to consonant. Hence the opening-closing gesture in b will be referred to as the passive O-C gesture.

The glottis aperture in p is much greater than in b, the two curves differing significantly already during the lip closure phase. Thus the greater opening in p cannot be

---

1) As proposed for voiced obstruents by Halle and Stevens (1968).

2) Eli Fischer-Jørgensen (1969, 1970) has found that voicing starts spontaneously if the closure is artificially released in Danish b, but not in Danish p.

explained by aerodynamic factors, and we therefore assume that it is due to neural commands. Hence the opening-closing gesture in p will be referred to as the active O-C gesture. Gestures similar to those of p are found in f.

The gestures of h are similar in shape to those of p and f, but the aperture is generally smaller. This applies to intervocalic voiced h as well as unvoiced h. In comparing the recordings of f and h we find that the airstream is greater for the second consonant, and hence the lesser amplitude of the glottogram might be explained as a Bernouilli effect. We therefore do not assume a difference in neural commands between h and p, f.

If we go back to our starting point, viz. that p is essentially the aspirated equivalent to b, it is interesting to see that this can be accounted for in terms of the O-C gestures. Provided that it makes sense to divide b roughly into two portions, one characterized by the (possibly aerodynamically conditioned) opening gesture, and the other by gradual adaption to conditions for voicing, then p may, to a first approximation, be conceived as a complex of two such portions, of which the first corresponds to the first part of b, whereas the second corresponds to h, i.e. with active glottis opening.

Similarly, f may be characterized by some kind of combined effect of aerodynamic force on the vocal folds and active glottis opening, but it remains to be found out when the active opening movement starts. There is of course no direct indication in the curves for p or f that the active opening should start later than the passive opening assumed to account for b.

An obvious way to check whether there is any likelihood that opening gestures add, is to compare sequences with p to sequences with B+h. The difficulty here is that there will be junctural phenomena involved, which may make the temporal relationships of the oral articulation of B+h different from those of p.

As a provisional test of this kind we recorded some compound words with post-junctural p, viz. søpølse, trepattet versus compound words with B+h, viz. kæphest, snaphaner.

This gave the following mean durations of the parameters (each representing six tokens, subject JR):

	søpølse	trepattet	kæphest	snaphaner
	ms	ms	ms	ms
parameter 1:	126	99	118	104
parameter 2:	65	55	56	46
parameter 3:	59	49	54	51
parameter 5:	6	6	2	-5
parameter 7:	67	50	64	53

These results show that the variation of respectively p and B+h as conditioned by the environments (and by free variation) exceeds the absolute difference between the averages for p and B+h (it must be admitted that the material is small). Similar tendencies have been found for BFJ and CL.

The measurements above refer to consonants between a vowel with strong stress and a vowel with reduced ("secondary") stress. When the second syllable has strong stress there is often some differences between p and B+h, the latter sequence being characterized by more voice ripple in the falling part of the glottogram than p. Moreover, the peak of maximum aperture tends to occur slightly later in B+h before stress, all of which points toward a more separate h gesture.

Examples of p versus B+h are shown in Fig. 4.

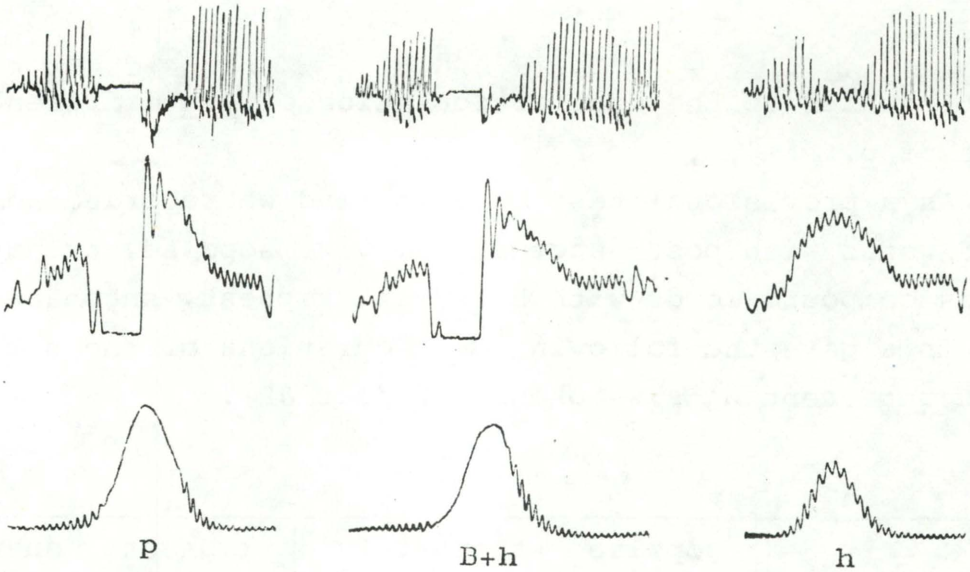


Fig. 4 Intervocalic p, B+h, and h (following syllable stressed). Subject JR. From top to bottom: Duplex oscillogram, aerometer curve, and glottogram.

We conclude from this that there is no reason to assume that the laryngeal commands for B+h under conditions of minimum juncture effect differ significantly from those for p. We have found in other recordings that (word final) B immediately followed by a vowel is similar to (word initial) b, but B and h add to give largely the same curve as p. It is not obvious that B+h should be able to give the same timing of the point of maximum glottis aperture as p unless we assume that the gestures of p is a summation of gestures similar to B and gestures similar to h. It goes without saying that these gestures, whose phonological counterparts are sequential, may overlap or even coincide in time, but this would then be true both of p and of B+h.

Additional types of sequences, which we have inspected are stop plus stop, i.e. B+p or B+b with or without stress on the following vowel, and s+ (aspirated) p versus p alone.

The words with two stops show that the longer hold of oral closure as compared to single stops does not cause the glottal aperture (considerably) to exceed that for respectively single p and single b, on the contrary we sometimes find a tendency toward two-peakedness (which one might wish to take as a boundary phenomenon). This suggests that the differences in glottis aperture observed by comparison of single consonants are not in the first place undershoot phenomena. Sequences like s+p sometimes show a greater aperture than single s or p in similar environments, but the difference is not stable.

The most important finding in connection with consonant combinations of various kinds is that in cases where two consecutive consonants form one peak of glottis aperture (this is the case in normal speed of pronunciation with all the sequences we have studied) this peak occupies the position one would expect if some kind of summation of opening took place. Thus in the sequence s+p the peak comes clearly before the explosion of p, whereas in single p it comes very close to (and mostly after) the explosion. This not only corroborates the summation hypothesis, but also shows that the turning point of the glottogram is not "triggered" by the oral explosion, which one might otherwise be tempted to believe by looking at the glottograms for single p.

In slow pronunciation sequences like s+p and B+p gradually assume a two-peaked shape, cf. Fig. 5. We do not know to what extent this should be explained by an intervening junctual phenomenon, or to what extent each consonant retains its opening and closing commands in such sequences.

We have even noticed slight tendencies towards such two-peakedness in the sequence B+b in very emphatic speech, but this may be due to a sudden rise of subglottal pressure in connection with the emphatic stress. It need not mean that there is an active opening gesture in b.

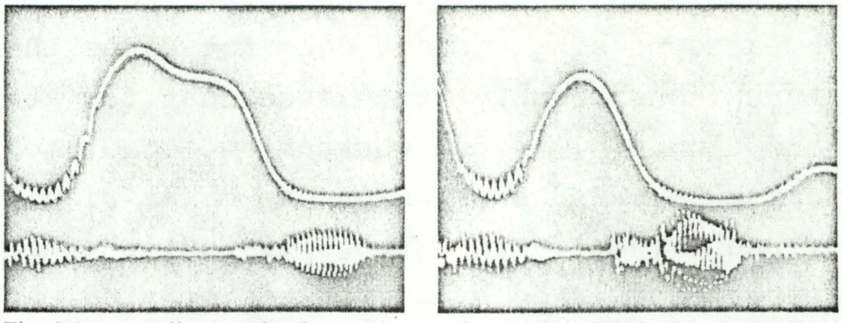


Fig. 5 Intervocalic s+p in slow and normal pronunciation (in the sequence *pas 'på!* "be careful"). Subject JR.

10. Conclusion

Glottographic investigation has corroborated the current conception of Danish p and b as differing essentially in aspiration. The consonants p, f, and h are all characterized by an opening-closing movement of the glottis with fairly equal duration of the opening and closing phases, and we assume the basic "command" to be the same for these three types of consonants. The consonant b which is unvoiced like p, f, has only a slight initial movement of the glottis toward open position, the rest of its duration being characterized by a slight closing movement. The maximum aperture is considerably smaller than for p, f,h. We assume the glottal behaviour of b to be a passive movement caused by the aerodynamic conditions. Inspection of glottograms of various consonant sequences suggests that if unvoiced consonants like those mentioned above occur in succession, there is a summation of the - passive or active - gestures belonging to the different consonants. Thus the sequence B+h exhibits a glottogram that is similar and often identical to that of p in analogous environments.

References

Berg, Jw. van den 1968: "Mechanism of the larynx and the laryngeal vibrations", _Manual of Phonetics_, ed. Bertil Malmberg, p. 278-308

Fischer-Jørgensen, Eli 1954: "Acoustic analysis of stop consonants", _Miscellanea Phonetica_ II, p. 42-59

Fischer-Jørgensen, Eli, and A. Tybjærg Hansen 1959: "An electrical manometer and its use in phonetic research", _Phonetica_ 4, p. 43-53

Fischer-Jørgensen, Eli 1963: "Beobachtungen über den Zusammenhang zwischen Stimmhaftigkeit und intra-oralem Luftdruck", _Zs.f.Ph._ 16, p. 19-36

Fischer-Jørgensen, Eli 1967: "Phonetic analysis of Danish stop consonants", _ARIPUC_ 1, p. 31-33

Fischer-Jørgensen, Eli 1969: "Voicing, tenseness and aspiration in stop consonants", _ARIPUC_ 3, p. 63-114

Fischer-Jørgensen, Eli 1970: "Les occlusives françaises et danoises d'un sujet bilingue", _Word_ 24, p. 112-153

Frøkjær-Jensen, Børge 1968: "A photo-electric glottograph", _ARIPUC_ 2, p. 5-19

Frøkjær-Jensen, Børge 1969: "Comparison between a Fabre glottograph and a photo-electrical glottograph", ARIPUC 3, p. 9-16

Halle, M. and K.N. Stevens 1968: "On the mechanism of glottal vibration for vowels and consonants", MIT QPR 85, p. 267-270

Leden, Hans von 1961: "The mechanism of phonation", Archives of Otolaryngology, vol. 74, No. 4

Lisker, L., A.S. Abramson, F.S. Cooper, and M.H. Schwey 1966: "Transillumination of the larynx in running speech", Haskins SR, 5/6

Malécot, André and Kenneth Peebles 1965: "An optical device for recording glottal adduction-abduction during normal speech", Zs.f.Ph. 18, 6

Miller, R.L. 1959: "Nature of the vocal cord wave", JASA 31, p. 667-677

Ohala, John 1966: "A new photo-electrical glottograph", UCLA 4, p. 40 ff.

Rischel, Jørgen 1970: "Consonant gradation: A problem in Danish phonology and morphology", The Nordic Languages and Modern Linguistics, ed. Hreinn Benediktsson, p. 460-480

Slis, I.H. and P.H. Damsté 1967: "Transillumination of the glottis during voiced and voiceless consonants", <u>IPO</u> No. 2 (Eindhoven)

Smith, Svend 1964: "Vertikale Komponenten bei der Funktion der Stimmlippen", <u>Phonetica</u> 11, p. 243-247