ARIPUC VOL. 13.

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

COPENHAGEN 1979

ARIPUC VOL. 13.

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

96, Njalsgade
DK 2300 Copenhagen

COPENHAGEN 1979

This annual report is published on the basis of a grant from the Faculty of the Humanities (publications account), University of Copenhagen.

KØBENHAVNS UNIVERSITET
REPRO.AFDELINGEN

CONTENTS

Personnel of the Institute of Phonetics III

Publications by staff members V

Lectures and courses in 1978 VI

Instrumental equipment of the laboratory by the end of 1978 X

Abbreviations employed in references XIV

Intonation contours in declarative sentences of varying length in ASC Danish
(Nina Thorsen) 1

The acquisition of stop consonants by Danish two-year-old children
(John Jørgensen) 9

Variation in inherent Fo level differences between vowels as a function of position in the utterance and in the stress group
(Niels Reinholt Petersen) 27

Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish
(Nina Thorsen) 59

A comment on Basbøll's phonological syllabification as applied to Danish
(Peter Molbæk Hansen) 87

Word boundaries and Fo patterns in Advanced Standard Copenhagen Danish
(Nina Thorsen) 121

A sketch of the history of phonetics in Denmark until the beginning of the 20th century
(Eli Fischer-Jørgensen) 135

Influence of microphone position in the recording of speech signals
(Carl Ludvigsen) 171

Acoustic-statistical analysis of voice quality
(Børge Frøkjær-Jensen) 189

PERSONNEL OF THE INSTITUTE OF PHONETICS IN 1978

Professor: Eli Fischer-Jørgensen, dr.phil. h.c. (director of the Institute)

Associate professors:

Børge Frøkjær-Jensen, cand.mag. (seconded to the Audiologopedics Research Group)

Jørgen Rischel, dr.phil. (until July 31)

Nina Thorsen, cand.phil.

Oluf Thorsen, cand.mag.

Assistant professors:

Peter Holtse, cand.phil. (from September 1)

Birgit Hutters, cand.mag.

Niels Reinholt Petersen, cand.phil. (until August 31, temporarily appointed)

Research fellow:

Peter Holtse, cand.phil. (until August 31)

Lecturer:

Niels Knudsen, cand.mag. (spring semester)

Teaching assistants:

Peter Molbæk Hansen, stud.mag.

John Jørgensen, stud.mag.

Niels Reinholt Petersen, cand.phil.

Lisbeth Strøjer, M.A.

Engineers:

Preben Dømler, B.Sc.
Carl Ludvigsen, M.Sc. (from November 1, temporarily appointed)
Mogens Møller, M.Sc. (on leave from October 1)

Technician: Svend-Erik Lystlund

Secretary: Else Parkmann

Teachers from other institutes lecturing at the Institute of Phonetics:

Niels Davidsen-Nielsen, cand.mag. et art. (English Institute)
Jørgen Rischel, dr.phil. (Institute of Linguistics)
Henning Spang-Hanssen, dr.phil. (Institute of Applied and Mathematical Linguistics)

Guest lecturer: Christopher J. Darwin (Laboratory of Experimental Psychology and Centre for Research on Perception and Cognition, University of Sussex).

PUBLICATIONS BY STAFF MEMBERS

Eli Fischer-Jørgensen	"On the universal character of phonetic symbolism with special reference to vowels", *Sign and Sound, Studies presented to Bertil Malmberg on the occasion of his sixty-fifth birthday* (eds. B. Sigurd and J. Svartvik), *SL* 32/1-2, p. 80-90
Børge Frøkjær-Jensen (ed.)	*Aids for the Handicapped*, *Logos* 12
Niels Reinholt Petersen	"Intrinsic fundamental frequency of Danish vowels", *JPh.* 6, p. 177-189
Jørgen Rischel	"How can phonological descriptions be made more realistic?", *Papers from the Fourth Scandinavian Conference of Linguistics*, Hindsgavl, January 6-8, p. 419-431
Nina Thorsen	"Une explication simplifiée, en termes physiques, des conséquences acoustiques des mouvements de la langue et des lèvres dans la production des voyelles", *Actes des 9èmes Journées d'Etudes sur la Parole du Groupe de la "Communication Parlée"*, Lannion 31 mai - 2 juin 1978, p. 201-210
Nina Thorsen	"An acoustical investigation of Danish intonation", *JPh.* 6, p. 151-175
Nina Thorsen	"Aspects of Danish intonation", *Nordic Prosody* (eds. Gårding, Bruce, and Bannert), Department of Linguistics, Lund University 1978, p. 23-32
Nina Thorsen and Oluf Thorsen	*Fonetik for Sprogstuderende*, 3rd edition, 2nd printing, 169 pp.
Oluf Thorsen, Ole Kongsdal Jensen, and Karen Landschultz	*Fransk Fonetik*, 8th revised edition, 272 pp.

LECTURES AND COURSES IN 1978

1. Elementary courses in general phonetics

One semester courses (two hours a week) in elementary general phonetics (intended for students of phonetics and linguistics and students of foreign languages except English, French, German, Modern Greek, Italian, and Russian) were given by Niels Reinholt Petersen. There was one class in the spring semester and one in the autumn semester.

Courses in general and French phonetics including practical exercises in the language laboratory (three hours a week) were given through 1978 by Oluf Thorsen.

A course in general and German phonetics (two hours a week) was given in the autumn semester by John Jørgensen.

A course in general phonetics for students of Modern Greek (two hours a week) was given in the autumn semester by John Jørgensen.

A course in general phonetics for students of Italian (two hours a week) was given in the autumn semester by Nina Thorsen.

A course in general and Russian phonetics (two hours a week) was given in the autumn semester by Peter Molbæk Hansen.

2. Practical exercises in ear-training and phonetic transcription

Birgit Hutters gave a course for beginners (two hours a week) in the autumn semester.

Nina Thorsen gave a course for more advanced students (two hours a week) in the spring semester.

Oluf Thorsen gave a course for advanced students (two hours a week) in the autumn semester.

These courses form a cycle of three semesters, and are based on tape recordings, as well as work with informants (on the advanced level).

3. Phonology

Jørgen Rischel gave courses in phonology for beginners (two hours a week) and for advanced students (two hours a week) in the spring semester.

Lisbeth Strøjer gave a course in phonology for beginners (two hours a week) in the autumn semester.

Niels Davidsen-Nielsen (English Institute) gave a course in phonology for more advanced students (two hours a week) in the autumn semester.

Jørgen Rischel (Institute of Linguistics) gave an introductory course in general linguistics (one hour a week) in the autumn semester.

4. The physiology of speech

Birgit Hutters gave a course in instrumental physiological phonetics (two hours a week plus individual exercises) in the spring semester, and a course in the physiology of speech (two hours a week) in the autumn semester.

5. The acoustics of speech

Peter Holtse and Niels Reinholt Petersen gave a course in instrumental acoustic phonetics (four hours a week plus individual exercises) in the autumn semester.

Nina Thorsen gave a course in the acoustics of speech (two hours a week) in the autumn semester.

Preben Dømler and Nina Thorsen gave a course in elementary mathematics and electronics (two hours a week) in the spring semester.

6. Other courses

Eli Fischer-Jørgensen and Christopher J. Darwin (University of Sussex) gave a course in speech perception for advanced students (two hours a week) in the spring semester.

Peter Holtse and Niels Reinholt Petersen gave a course in computer programming (BASIC) (two hours a week) in the spring semester.

Birgit Hutters and Niels Reinholt Petersen presided at a series of seminars for advanced students on topics in experimental phonetics (two hours a week) in the spring semester.

Henning Spang-Hanssen (Institute of Applied and Mathematical Linguistics) gave a course in statistics (two hours a week) in the autumn semester.

Oluf Thorsen gave a course in the theory and practice of the language laboratory (one hour a week) in the autumn semester.

Oluf Thorsen gave a course in French phonetics (two hours a week) in the spring semester.

Niels Knudsen gave a course in German phonetics (two hours a week) in the spring semester.

Lisbeth Strøjer gave a course in English phonetics (two hours a week) in the autumn semester.

Christopher J. Darwin (University of Sussex) presided at seminars for advanced students and staff members on topics in speech perception in March, April, and May.

7. Seminars

Nina Thorsen presented and discussed some informal experiments with synthesis of Fo patterns and intonation contours.

Michael Studdert-Kennedy lectured on "Selective adaptation of speech: is the evidence for feature detectors sufficient?" and "Hemispheric specialization of speech, functions and mechanisms".

Robert J. Porter gave a lecture on "Rapid shadowing of vowel sequences".

André Martinet lectured on "The internal conditioning of phonological changes".

Hans Basbøll gave a lecture on recent contributions to phonological theory.

8. Participation in congresses, symposia, meetings, etc.

Eli Fischer-Jørgensen, Birgit Hutters, Nina Thorsen, and Oluf Thorsen participated in a symposium on Nordic Prosody in Lund, June 14-16, and Nina Thorsen gave a paper titled: "Aspects of Danish intonation".

Nina Thorsen participated in "9èmes Journées d'Etudes sur la Parole, Lannion 31 mai - 2 Juin 1978" and gave a paper "Une explication simplifiée, en termes physiques, des conséquences acoustiques de mouvements de la langue et des lèvres dans la production des voyelles", and in "First Scandinavian-German Symposium on the Language of Immigrant Workers and their Children", Roskilde University, March 19-23, and gave an introduction: "The role of phonetics in language teaching".

Børge Frøkjær-Jensen participated in an interdisciplinary workshop "Pausological Implications of Speech Production", Gesamthochschule Kassel, Kassel, BDR, June 13-17, and in the Yugoslavian Symposium on Logopedics, Skopje and Ochrid, October 27-30, and gave an invited paper: "Logometrics used in logopedic diagnosis".

Birgit Hutters, Nina Thorsen, and several students participated in the Swedish-Danish Phonetics Seminar, Stockholm, November 9-10.

Eli Fischer-Jørgensen gave guest lectures at Strassbourg University in January on perceptual phonetics, stop consonant systems, and electromyography, at Kiel University in June on electromyography, and at Århus University in September on the subject "Hvad karakteriserer det menneskelige lydsprog?" ('What are the characteristics of human speech?').

Børge Frøkjær-Jensen gave a lecture: "The audiologopedic education at the University of Copenhagen", at the Social-pediatric Clinic, the Academic Hospital of Copenhagen, October 13.

Jørgen Rischel participated in the Fourth Scandinavian Conference of Linguistics at Hindsgavl, Denmark, January 6-8, and presided at the phonology section, in which he gave an introductory address: "How can phonological descriptions be made more realistic?". During April, May, and June Jørgen Rischel was guest professor at University of California, Los Angeles, teaching phonetics and Greenlandic phonology. During this stay he gave lectures at University of California, Berkeley, on "The diachronic phonology of Greenlandic" (to an "Amerindian Studies Group"), on "The Eskimo word" (to the Berkeley Linguistics Society),and on "The phonetic research at the Institute of Phonetics, Copenhagen" (to the staff of the Phonetics Laboratory at the Department of Linguistics), and at University of California, San Diego, on "The diachronic phonology of Greenlandic", and on "Stress in Danish". He made a journey to Alaska to meet the staff of the Native Languages Center at Fairbanks, and to study the teaching of Eskimo orthography and grammar to native Eskimos at Point Barrow.

Nina Thorsen gave guest lectures at Umeå University, Sweden, in January on "Tryk og tone, og sætningsintonation i moderne københavnsk rigsmål" ('Stress and tone, and sentence intonation in Advanced Standard Copenhagen Danish'), at Stockholm University, Sweden, in January on "The perception of intonation contours in

Danish", at Lund University, Sweden, in March on "Eksperimenter med sætningsintonationen i dansk" ('Experiments on sentence intonation in Danish'), at Uppsala University, Sweden, in April on "Kontrastemfas och satsintonation i danskan" ('Emphasis for contrast and sentence intonation in Danish'), and at Lund University, Sweden, in December on "Leksikalsk tryk, kontrasttryk og sætningsintonation i moderne københavnsk" ('Lexical stress, contrastive stress, and sentence intonation in Advanced Copenhagen Danish'). In April she gave a talk on "Prosody" at the Annual Meeting of the Audiologopedics Association (together with Gösta Bruce, Lund University, Sweden).

Eli Fischer-Jørgensen received the honorary degree of fil.dr. at Lund University in June, and the honorary degree of dr.phil. at Aarhus University in September.

INSTRUMENTAL EQUIPMENT OF THE LABORATORY BY THE END OF 1978

1. Instrumentation for speech analysis

- 2 Sona-Graphs, Kay Elemetrics, type 6061 A
- 2 amplitude display/scale magnifier units, Kay Elemetrics, type 6076 A
- 1 contour display unit, Kay Elemetrics, 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, type PT
- 1 electro-aerometer, four channel, type AM 508/4
- 1 audio frequency filter, type 445
- 1 vocal cords fiberscope, Olympus, type VF
- 1 fundamental frequency meter, type FFM 650

1 Sona-Graph, Kay Elemetrics, type 7029 A

2 intensity meters, Fonema

2 fundamental frequency meters, Fonema.

2. Instrumentation for speech synthesis

1 formant-coded speech synthesizer

1 voice-source generator

1 larynx vibrator with power supply.

3. Filters

1 LC highpass filter (with stepwise variation of cut-off frequency)

1 active RC lowpass filter.

4. Instrumentation for visual recordings

1 mingograph, Elema 800 (8 channels)

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

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

1 oscilloscope, Tektronix, type 465

1 oscilloscope, Tektronix, type 5115

2 dual-trace amplifiers, Tektronix, type 5A18N

1 time-base, Tektronix, type 5B10N

1 mingograph, Siemens-Elema, type 803.

5. Tape recorders

1 instrumentation recorder, Lyrec, type TR 86

1 professional recorder, Lyrec, type TR 42

2 professional recorders, Lyrec, type TR 20

10 semi-professional recorders, Revox, type A 77

1 professional recorder, Revox, type A 700

1 portable semi-professional recorder, Uher, type 4000
1 cassette recorder, Tandberg, type TCD 310 MK-2.

6. Gramophones

2 gramophones, Delphon (mono, Ortofon pick-up).

7. Microphones

1 microphone, Neuman, type KM 56
3 dynamic microphones, Sennheiser, type MD 21
1 1" microphone, Brüel & Kjær, type 4131/32
1 1/4" microphone, Brüel & Kjær, type 4135/36
1 larynx microphone
2 1" microphones, Brüel & Kjær, type 4145
1 microphone, power supply, Brüel & Kjær, type 2807
1 microphone, power supply, Telefunken.

8. Amplifiers

1 microphone amplifier, Brüel & Kjær, type 2603
1 power amplifier, Brüel & Kjær, type 2706
2 microphone pre-amplifiers, Brüel & Kjær, type 2627
2 measuring amplifiers, Brüel & Kjær, type 2607 A.

9. Loudspeakers/headphones

3 loudspeakers, Beovox, type 2600
10 headphones, Sennheiser, type 414
4 headphones, Sennheiser, type 424
2 loudspeakers, Beovox, type M 70
2 headphones, Sennheiser, type 424 X
8 loudspeakers, Philips, type RH 541 MFB.

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, Radiometer, type RV 23b
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

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

1 X-Y recorder, Hewlett & Packard, type 7044 A

1 noise generator, Brüel & Kjær, type 1405

1 stabilized rectifier, Danica, type TPS 21

1 capacitance meter, ECD Corp.

Additional oscillators, rectifiers, etc., for special purposes.

11. Outfit for photography

1 Minolta camera SR-1 (with various accessories)

1 complete outfit for reproduction

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

1 timer, Kaiser, type 4033

1 enlarger, Durst, type A 300

1 rapidoprint, Agfa, type DD 1437

1 dry-machine, Durst, type 400

12. Equipment for EDB

1 computer, Digital, PDP8/E, 8k

1 arithmetic unit, Digital, type KE 8-E

1 bootstrap loader, Digital, type MR 8-EC

1 add-on memory system 24k, Fabri-Tek, type 8/E
1 dectape, Digital, type TD8-EM
1 tape reader, GNT, type 24
1 tape punch, GNT, type 34
1 decwriter, Digital, type LA 30P
1 teletype, Teletype, type ASR 33
1 display terminal, Tektronix, type 4014-1
1 real time clock, Digital, type DK8-EP
1 a/d converter, Digital, type AB8-EA
1 d/a converter, type PD
1 disk drive, Digital, type RK 8J-ED
1 disk drive, Plessey, type DD-8/B
1 line printer, Binder Magnete, type BM 132
3 CTR terminal, Perkin-Elmer, type FOX 1100

13. Projectors

1 Leitz projector for slides
1 16 m/m tone film projector, Bell & Howell, "Filmsound 644"
1 projector, Leitz, type Pradovit color 250.

ABBREVIATIONS EMPLOYED IN REFERENCES:

AJPs.	American Journal of Psychology
AL	Acta Linguistica
ALH	Acta Linguistica Hafniensia
ARIPUC	Annual Report of the Institute of Phonetics, University of Copenhagen
Folia Ph.	Folia Phoniatrica
FRJ	For Roman Jakobson
F&S	Form and Substance (Akademisk forlag), København 1971
Haskins SR	Status Report on Speech Research, Haskins Laboratories
IJAL	International Journal of American Linguistics
IPO APR	IPO Annual Progress Report
JASA	Journal of the Acoustical Society of America
JL	Journal of Linguistics
JPh.	Journal of Phonetics

JSHD	Journal of Speech and Hearing Disorders
JSHR	Journal of Speech and Hearing Research
Lg.	Language
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. Ling. ...	Proceedings of the ... International Congress of Linguists
Proc.Phon. ...	Proceedings of the ... International Congress of Phonetic Sciences
STL-QPSR	Speech Transmission Laboratory, Quarterly Progress and Status Report, Royal Institute of Technology, Stockholm
SL	Studia Linguistica
SPE	The Sound Pattern of English, Chomsky and Halle, 1968
TCLC	Travaux du Cercle Linguistique de Copenhaque
TCLP	Travaux du Cercle Linguistique de Praque
UCLA WPP	Working Papers in Phonetics, University of California, Los Angeles
Zs.f.Ph.	Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung.

(ARIPUC 13, 1979)

INTONATION CONTOURS IN DECLARATIVE SENTENCES OF VARYING LENGTH IN ASC DANISH[1]

Nina Thorsen

1. Introduction

The stress/Fo relationship and the intonation contours of various types of short sentences in Advanced Standard Copenhagen (ASC) Danish have been described elsewhere (Thorsen 1978a, 1979b). For the purpose of the present paper, only a few points need be repeated: Stress in ASC Danish is signalled mainly by Fo. In neutral speech a stressed syllable will be (relatively) low and followed by a high-falling tail of unstressed syllables, i.e. the stressed syllable is one that is jumped or glided up from, depending on the segmental composition, cf. Fig. 1 (full lines). The unit which carries this Fo pattern consists of the stressed syllable plus all succeeding unstressed ones, irrespective of intervening syntactic boundaries within the simple (i.e. non-compound) sentence. It is termed a stress group (SG). Since the Fo patterns of SGs are predictable and recurrent entities (though allowing for certain context dependent modifications), the intonation contour may be defined solely in terms of the stressed syllables. (This does not necessarily mean that the course of the unstressed syllables is irrelevant, e.g. for the perception of intonation contours but it is, strictly speaking, redundant, cf. Thorsen 1978b.)

In simple sentences having no more than three SGs, the intonation contours were found to approach straight lines whose slopes varied according to sentence type. Declarative sentences have the most steeply falling contours at one extreme and syntactically unmarked questions have "flat" contours at the other. In between are found various syntactically marked questions as well as non-

---

1) Summary of a paper read at the Autumn Meeting of the Institute of Acoustics, Windermere 4-6th November 1979. Also to be published in the Proceedings of the Institute of Acoustics.

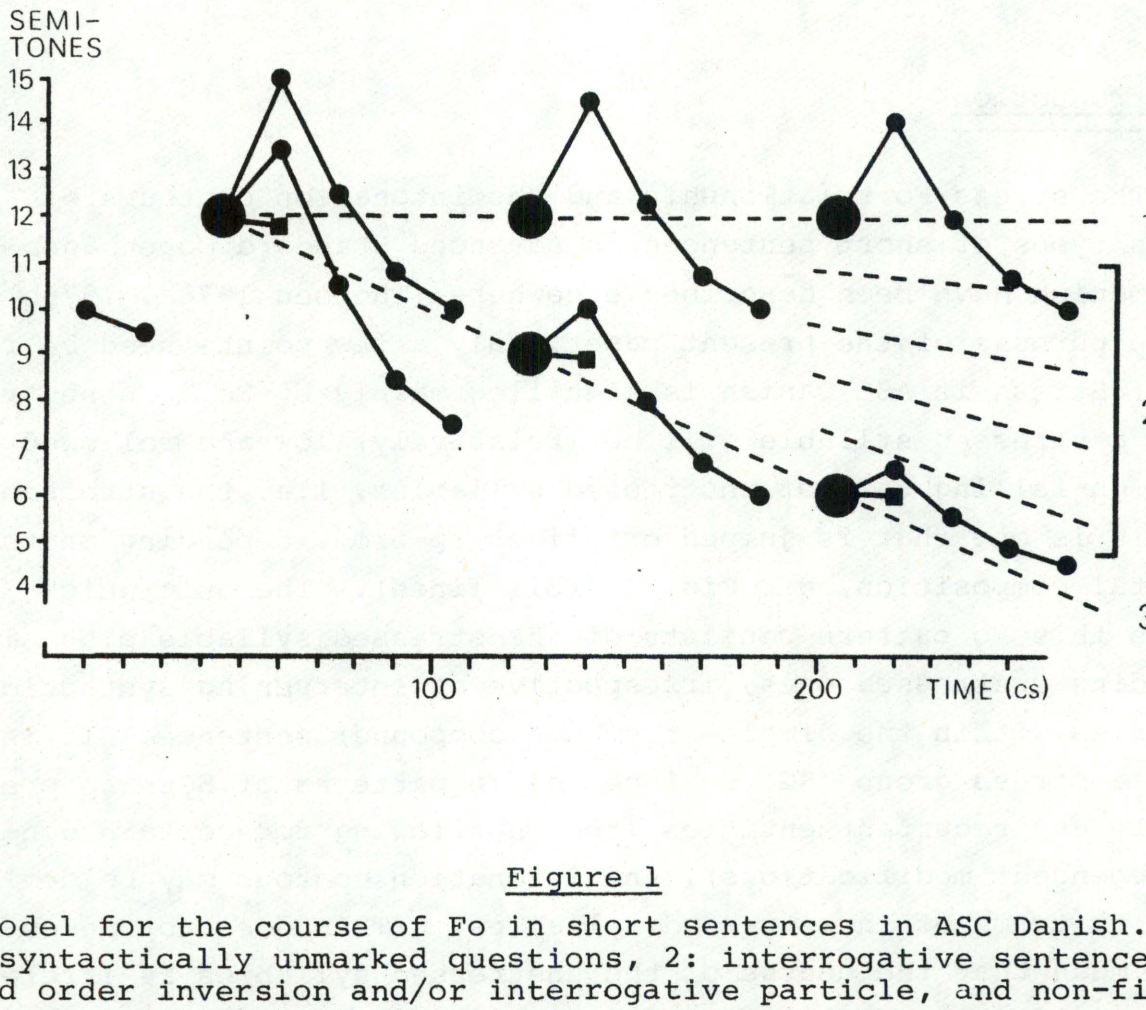


Figure 1

A model for the course of Fo in short sentences in ASC Danish.
1: syntactically unmarked questions, 2: interrogative sentences with word order inversion and/or interrogative particle, and non-final periods (variable), 3: declarative sentences. The large dots represent stressed syllables, the small dots unstressed ones, and the small squares represent an unstressed syllable being the only one between two stressed ones. The full lines represent the Fo pattern associated with stress groups, and the broken lines denote the intonation contours.

final clauses, with a tendency for the contour to be steeper, the more syntactic information is contained in the sentence about its interrogative or non-final function, cf. Fig. 1 (broken lines).

2. The present experiment

Thorsen (1979b) hypothesized that when the number of SGs changes, everything else being equal, so does the slope of a given contour, leaving only the flat ones intact. The constancy presumably lies in the interval between first and last stressed syllable, intervening stressed syllables being evenly distributed between them. (A similar hypothesis is advanced for Swedish, cf. Bruce 1979.) Since declarative sentences have the widest range, differences in slope would be most easily detected in them: eight simple statements were made up, containing from one to eight SGs, all variations on the same theme. They may be deducted from the longest one, though some of the names were moved about for the sake of variety (´ denotes the stressed syllables and | denotes the syntactical boundaries between NP and VP, VP and compound adjunct, and the two adjuncts):

^{2}Hútters (og ^{8}Bítten) | ^{2}skal (med ^{3}bússen) | ^{1}til (^{5}fésten)

(for ^{6}Kísser) (og ^{7}Líssi) | ^{5}på (^{4}Kílden i) ^{1}Thísted.

(Hutters and Bitten are taking a bus to the party for Kisser and Lissi at Kilden in Thisted.) To get sentence no. X, pick out all words from left to right numbered X and lower. The stressed vowels are all short, high (except [ε] in 'festen'), and surrounded by unvoiced obstruents (except [l] in 'Lissi') to facilitate the subsequent interpretation of the tracings (cf. Thorsen 1979a).

The items were mixed with a material recorded for a different purpose, being evenly distributed over two full pages of recording material, which appeared in three different randomizations, each being read twice, giving a total of six recordings of each sentence. Subjects were four phoneticians, three ASC speakers (NRP male, BH female, NT female (the author)) and one with a slightly more conservative pronunciation (JR male).

The recordings were made with semi-professional equipment in a quasi-damped room and were processed by hard-ware intensity and pitch meters, registered on a mingograph and measured by hand. Fo of each of the stressed vowels was measured at 2/3 of the distance in time from vowel onset (cf. Rossi 1971) which was an uncontroversial procedure since the vowels all had monotonously falling movements. The distance in time of each of these points from the first vowel Fo was also measured. The average Fo measurements were then converted to semitones (re 100 Hz) and a correction made for intrinsic Fo level differences between [u], [ɛ], and [i], in accordance with Reinholt Petersen's (1978) results.

3. Results

The results are displayed graphically in Fig. 2. In Table 1 are given the slopes of the intonation contours, as determined by the least squares linear regression on the data points in Fig. 2, their correlation coefficients, and the interval between first and last vowel (range).

First of all, range is not constant throughout the eight sentences with any subject but neither is it monotonously increasing with increasing number of SGs and there is no clear relationship between (smaller) range and (greater) size or number of breaks in the contours. The inconstancy of the range is also reflected by the fact that the regression line slopes, although they do get less and less steep, do not do so in even approximately equal size steps through sentences 2 to 8 - even though the time increments are very nearly constant.

Secondly, the intonation contours are obviously not straight lines, but the correlation coefficients are generally above .95 and any further statistical treatment will hardly disclose the regularities that can be observed in Fig. 2 any more succinctly than visual inspection will do: The medial stressed vowels do not occur with equal intervals and the irregularity generally sets in with four and more SGs. With BH, NRP and JR, the longer utterances seem to be composed of two (sentences 4-6) and three (7-8) prosodic phrase groups, respectively, the breaks between which coincide with major syntactical boundaries, viz. before the (compound)

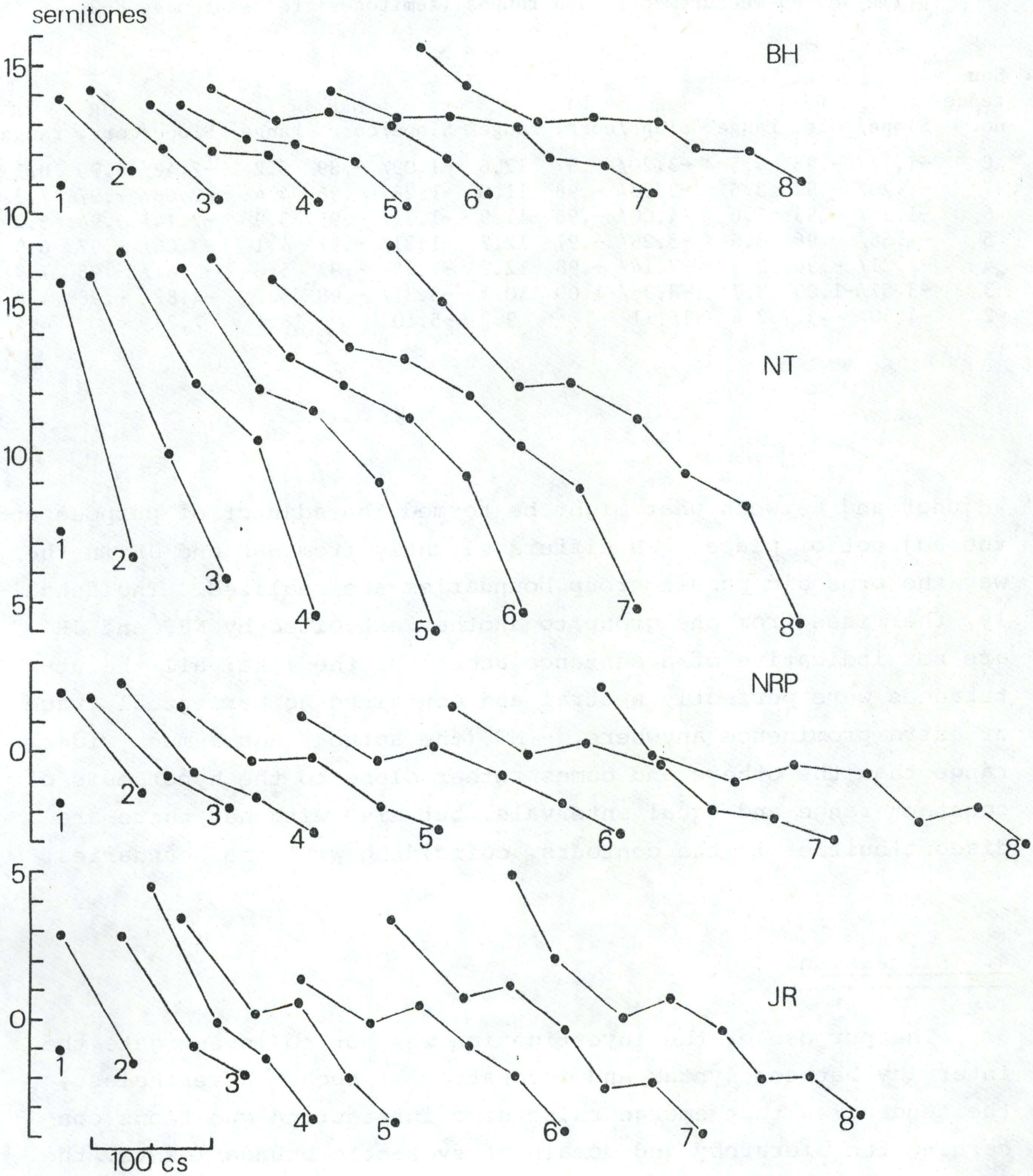


Figure 2

Intonation contours in simple declarative sentences containing from one to eight stress groups.

Table 1

Intonation contour slopes (semitones/sec), correlations between the time and Fo measurements, and ranges (semitones) for sentences 8-2.

Sentence no.	BH Slope/corr.	BH range	NT Slope/corr.	NT range	NRP Slope/corr.	NRP range	JR Slope/corr.	JR range
8	-1.16/ -.95	4.5	-3.20/ -.97	12.6	-1.09/ -.89	5.2	-2.42/ -.95	8.1
7	-1.20/ -.97	3.5	-3.23/ -.96	11.0	-1.26/ -.96	4.4	-2.66/ -.97	7.3
6	-1.37/ -.93	3.6	-4.00/ -.96	11.9	-1.31/ -.95	3.9	-2.14/ -.94	5.2
5	-1.66/ -.96	3.5	-5.26/ -.97	12.2	-1.81/ -.97	4.1	-3.66/ -.97	6.9
4	-2.21/ -.96	3.3	-7.14/ -.98	12.2	-2.95/ -.97	5.0	-5.61/ -.98	7.9
3	-3.57/-1.00	3.7	-8.96/-1.00	10.1	-3.31/ -.98	3.7	-4.82/ -.98	4.8
2	-4.50	2.4	-15.11	9.2	-5.10	3.4	-7.24	4.4

adjunct and between what might be termed the adjunct of purpose and the adjunct of place. BH differs slightly from NRP and JR in the way the prosodic phrase group boundaries are realized. Incidentally, the rises from one group to another exhibited by NRP and JR are not indicative of a sentence accent or the like: all the utterances were perfectly neutral and contained no perceptual trace of extra prominence anywhere. - NT (the author) has a much wider range than the others and comes rather close to the hypothesis of constant range and equal intervals, but also with her there are discontinuities in the contours, coinciding with the boundaries.

4. Discussion

The purpose of the investigation was not to investigate the interplay between syntax and intonation as such. Nevertheless, the tendencies that emerge raise some interesting questions concerning the hierarchy and domain of syntactic boundaries vs. the inherent features of declarative intonation. In this material, the boundary before the (compound) adjunct seems to be more manifest than the NP-VP boundary. Further, 4 SGs in the prosodic phrase group seems to be the maximum (cf. sentences 6-7). The constituent which was varied most in number of SGs was the adjunct. What would the contours have looked like if instead the NP and/or

VP had varied? E.g. is the rather steep slope in the first gradient an inherent feature of declarative intonation or an artefact of the material that would disappear if the NP or VP were longer? With a short adjunct but long NP, would a break occur after the NP and would the VP and adjunct merge into one prosodic phrase group? If the second of the two adjuncts had consisted of only one SG, would it have had to merge with the preceding adjunct in order to preserve a final fall? (and where would the first adjunct be cut up then, if 4 SGs are the maximum in a prosodic phrase group?), or is the final fall dispensable as long as there is a general downdrift as observed in the material? In other words, is overall downdrift the only requirement to be filled in declarative sentences (a downdrift whose slope will be a function, although not a linear one, of the total duration of the sentence)? These, and a heap of other questions, are good candidates for further research.

References

Bruce, G. 1979: "Word prosody and sentence prosody in Swedish", Proc.Phon. 9, p. 388-394

Reinholt Petersen, N. 1978: "Intrinsic fundamental frequency of Danish vowels", JPh 6, p. 177-189

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

Thorsen, N. 1978a: "An acoustical investigation of Danish intonation", JPh 6, p. 151-175

Thorsen, N. 1978b: "On the identification of selected Danish intonation contours", ARIPUC 12, p. 17-73

Thorsen, N. 1979a: "Interpreting raw fundamental frequency tracings of Danish", Phonetica 36, p. 57-78

Thorsen, N. 1979b: "Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish", Proc.Phon. 9, p. 417-423.

(ARIPUC 13, 1979)

THE ACQUISITION OF STOP CONSONANTS BY DANISH TWO-YEAR-OLD CHILDREN

John Jørgensen

Abstract: This paper presents VOT data collected from four two-year-old Danish children. The results show that they differentiate the two categories - i.e. aspirated vs. unaspirated stops - in their production. The duration of aspiration in the unaspirated stops is found to be noticeably longer (the boundary between the two categories is at 50 ms) than what has been found in other investigations. An articulatory model accounting for the priority of a 0-30 ms VOT range for unaspirated stops is discussed and partly rejected.

1. Introduction

The phonological and phonetic aspects of children's acquisition of their mother tongue has attracted increasing interest, especially during the last decade. In spite of this, Danish remains almost unexplored in this field. To my knowledge, only three investigations of Danish child language have been published, viz. Jespersen (1916) which was intended to draw the attention of a wide public to linguistic phenomena, but is mainly of historical interest today, Abrahams (1955), introducing Roman Jakobson's ideas, containing a registration of the phonemes most often mispronounced by normal and impaired French, Danish and English children, and finally, Jørgensen (1971) which contains a registration of the development of selected subsets of the phonetic systems of the author's two children during their very first years of life (the material was collected 30 years earlier - a fact which has to some degree influenced the treatment).

The scarcity of phonetic investigations of Danish child language is regrettable, mainly because the Danish phonetic/phonological system contains several interesting phenomena (a presentation of the Danish phoneme system is given by Basbøll, 1969) among them:

1) Aspiration/voicing and the relation between these two features (Danish stops are distinguished only by aspiration, whereas one fricative-pair is distinguished by voicing).

2) A rich vowel system. Danish has at least 10 qualitatively different vowels which may all appear long and short.

3) Stød (the Danish stød may be considered a phonemic, suprasegmental entity).

The three phenomena are all of interest in relation to the theories concerning children's acquisition of language, among which the theory advanced by Jakobson (1942) is of fundamental importance, and there are few papers on child language that do not refer to this work.

This paper is only concerned with one of the problems pointed to above, viz. the relation between aspirated and unaspirated stops, and the investigation reported below is an attempt to answer the following question: How do Danish children at the age of approximately two years "treat" (phonemically and phonetically) the aspiration distinction in the stop consonants?

2. Other investigations

In any survey, however sketchy, of problems in first language acquisition it seems natural to quote Roman Jakobson. The following quotation may be representative of Jakobson's opinion about children's linguistic behaviour when learning a language with a distinction of aspiration in the stop consonants. Explicitly mentioning Danish, Jakobson (1942, p. 9f) says:

"Solange die Verschlusslaute beim Kinde nach dem Verhalten des Kehlkopfs ungespaltet bleiben, werden sie gewöhnlich stimmlos und ohne Aspiration vollbracht: das Kind verallgemeinert diese Spielart unabhängig davon, ob das landläufige Vorbild (wie z.B. die slavischen und romanischen Sprachen) neben ihr ein stimmhaftes unaspiriertes oder ein stimmloses aspiriertes Gegenstück enthält (wie es im Dänischen der Fall ist);" and later on the same view is repeated (p. 95): "... der Konsonant ist gewöhnlich stimmlos auf der Anfangsstufe der Kindersprache, auch bei der partiellen Lautstummheit verliert er seine Stimmhaftigkeit (...), und bei der partiellen Lauttaubheit werden eher die stimmlosen als die stimmhaften Konsonanten erkannt ...".

Thus, Jakobson does not say anything about the age at which the distinction is acquired nor about the relation between production and perception. What is in focus is the order in which phonemic contrasts are acquired. As a consequence of the phonemic view held by Jakobson, little attention is paid to the precise phonetic realisations. Most often data are given in phonemic interpretation rather than in phonetic transcription.

2.1 When is the unaspirated/aspirated distinction acquired?

Investigations of more recent date have paid special attention to the age at which the aspiration distinction is acquired. Investigating the perceptual acquisition of the "voicing" distinction (which for American English most often is an aspiration distinction), Garnica (1971) finds that this distinction is acquired very late by American children. It appears to be the latest among the 13 distinctions tested, and a closer inspection of the results reveals that not even the oldest subject (age 3;5 years) perceptually masters the voicing distinction. In several respects these results are surprising and in my opinion they seem rather suspicious. In particular, the results for the oldest subject are extraordinary: he only masters 2 out of the 9 distinctions tested for. He does not master the distinction d vs. z which is mastered by all (i.e. 4) the other younger subjects. All these subjects (age 1;9 - 2;10) except one (age 1;9) master more distinctions than does the oldest child. Another example of the surprising results is that one subject does not master the distinction b vs. zero, whereas she does have b vs. m. One possible explanation of the curious findings may lie in the method used. The subjects were asked to remember (for a short time) nonsense names given to toys. The ability to remember newly learned nonsense words is probably highly sensitive to factors such as attention and general interest.

Not surprisingly, other investigations have not confirmed Garnica's results concerning the age at which the distinction is acquired. Edwards (1974) finds that 8 out of 9 children (age 1;8 - 3;11) perceive correctly the t vs. d distinction and that 7 out of 9 have a correct production, too. Unfortunately, the precise ages of those with correct production are not given.

The Garnica results are also disconfirmed by an investigation by Wintercorn et al. (1967). In this investigation 5 children (aged around 3 years) were asked to repeat synthesized syllables with varying VOT values (from -30 to +100 ms). The results are summed up as follows:

(p. 43): "Four of the five subjects gave almost "errorless" repetitions of the test tape. They responded to practically all of the stimuli with VOT values of -30, -5 and +20 as /da/ and to almost all other stimuli as /ta/, just as adult subjects do. Furthermore, their repetitions were made immediately after stimulus presentation. None of these subjects made more than three errors and the errors were not systematically distributed. The errors which did occur were judged to be due to inattention."

The results of an investigation by Gilbert (1977) in which 6 children's production of d and t (spontaneous speech) was examined, are summed up by the investigator as follows (p. 7): "In this investigation we have attempted to support the hypothesis that VOT has assumed adult values for /d/ and /t/ by an average age of 3;0. This attempt has proved partially successful in that the data for initial /d/ VOT values taken from SPONTANEOUS discourse do, for the most part, fall clearly within the adult range of VOT values reported by Lisker & Abramson (1964), with few intrusions into the adult long voicing lag category."

In conclusion, it seems safe to state that the general opinion is that around the age of 2 the child has gained complete control over the aspiration (voicing) distinction perceptually and between the age of 2 and 3 the distinction is also mastered in production.

3. Subjects - collection of the material

The main topic of the present investigation was the development of distinctive aspiration in stop consonants. Since - as has been pointed out by Jakobson (1942), among others - the stops are developed earlier than other members of the consonantal system, it was necessary to choose as subjects children who were very young, but nevertheless capable of producing stop consonants at at least two of the three places of articulation used in Danish stops (bilabial, alveolar, and palato-velar). This stage of linguistic development is reached (it appeared) around the end of the second year. Children whose parents were not speakers of Advanced Standard Copenhagen Danish (see Basbøll, 1969) and children who suffered from disorders of speech or hearing or other physical or psychical impairments were excluded as subjects.

The subjects were selected among 32 children attending a day nursery in the centre of Copenhagen, and the selection was based on careful daily observation during several weeks. Seven children were chosen as potential subjects, but three of them had to be excluded since they could not cope with the unfamiliar experimental situation. The remaining 4 subjects were KR (boy) 2;1;2, MI (boy) 2;1;0, NO (boy) 1;11;1, and TI (girl) 1;10;0.

As it was considered desirable for the linguistic material to be so close to natural spontaneous speech as possible (rather than imitated speech) the recordings were made at the nursery, where the children were familiar with the environment.[1]

At the recording sessions which lasted from 20 to 30 minutes, the child and a staff member (always the staff member preferred by the child) were talking about that particular child's favourite toys or picture books. The experimenter handled the recording and asked additional questions if the situation permitted. The quality of the recordings was not always the best, but with young children it seems to be the only method which will ensure the desired naturalness and spontaneity.

---

1) In the literature on child language the relationship between imitated and spontaneous speech is occasionally discussed (e.g. Templin 1947, Ingram 1976, Faircloth et al. 1970, Danielson et al. 1976, and Edwards 1974). The question is of primary importance when investigating children's active vocabulary. If the child succeeds better when imitating, no true picture of the linguistic system normally used by the child is achieved if imitation is used in the test, and consequently the imitation method must be rejected. The general opinion seems, however, to be that imitation is safe enough as a method. The investigations of the relation between the two methods have some weaknesses, however: (1) none of the investigations involved <u>normal</u> young children (two years or younger), (2) at least one of them states that imitation causes improvement of the speech, (3) it seems counterintuitive that young children should not perform better in imitation (how do they learn new pronounciations if not by imitating?). Accordingly, I chose the safest method - spontaneous speech.

Collecting material in the form of spontaneous speech from very young children has one serious disadvantage, however, in that it is extremely laborious and time consuming to get a representative corpus (even only partly so).

Altogether about four hours of speech were recorded, corresponding to approximately 500 phrases of one or (frequently) more words. The number of phrases recorded for each child varied between 80 and 130. As a rough estimate, the words in the recordings represent about 50 per cent of the children's active vocabulary. The subset of the material to be analyzed here was 270 occurrences of stop consonants in word initial position.[1]

The duration of aspiration (VOT) in these stops was measured from wide-band spectrograms with expanded frequency scale (200 - 4000 Hz). The duration was measured from the beginning of the explosion phase to the beginning of the voice bar in the following vowel.[2] 255 measurements of aspiration were obtained, 15 items being left out for various reasons.

4. Results

The distribution of VOT values measured in the supposedly aspirated and unaspirated stops (i.e. aspirated/unaspirated in the adult language) are displayed graphically in figs. 1 and 2.

The data indicate that the children in their speech production seem to be able to differentiate the two categories with reasonable certainty. A one-tailed Mann-Whitney U-test showed that the VOT values of ptk were significantly longer than those of bdg at the 5% level or better in all cases but one (see table 1).

Notice, that although a visual inspection of MI's distribution of /t/ and /d/ might lead to the conclusion that the two categories are not separated, the statistical test proves that in fact they are. A closer inspection reveals that with a boundary at +70 ms the overlapping is actually quite small - viz. 4 occurrences of /t/ below this value and 2 occurrences of /d/ above it.

---

1) Apart from the measurements of aspiration in the stop consonants, a phonetic transcription was made of the entire material, the purpose of which was to provide the basis for a more general evaluation of the material in relation to Jakobson's theories of the linguistic development in children. Some aspects of this work are briefly reported in Jørgensen (1978).

2) This will yield systematically lower values than if the onset of periodic energy in the higher formants are used as a criterion.

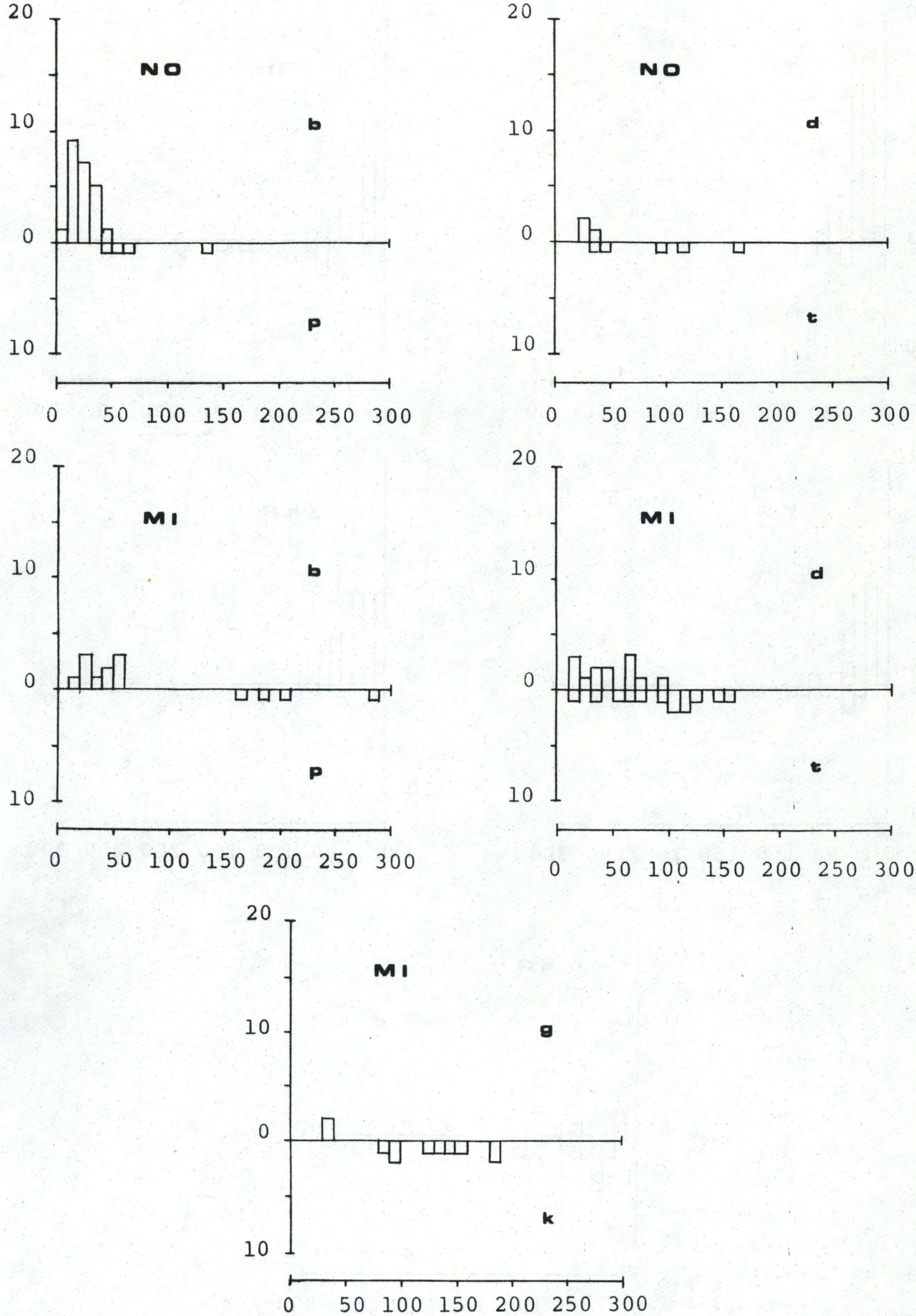


Figure 1

Histograms showing number of occurrences (vertical axis) vs. duration of the aspiration (horizontal axis). Subjects NO and MI.

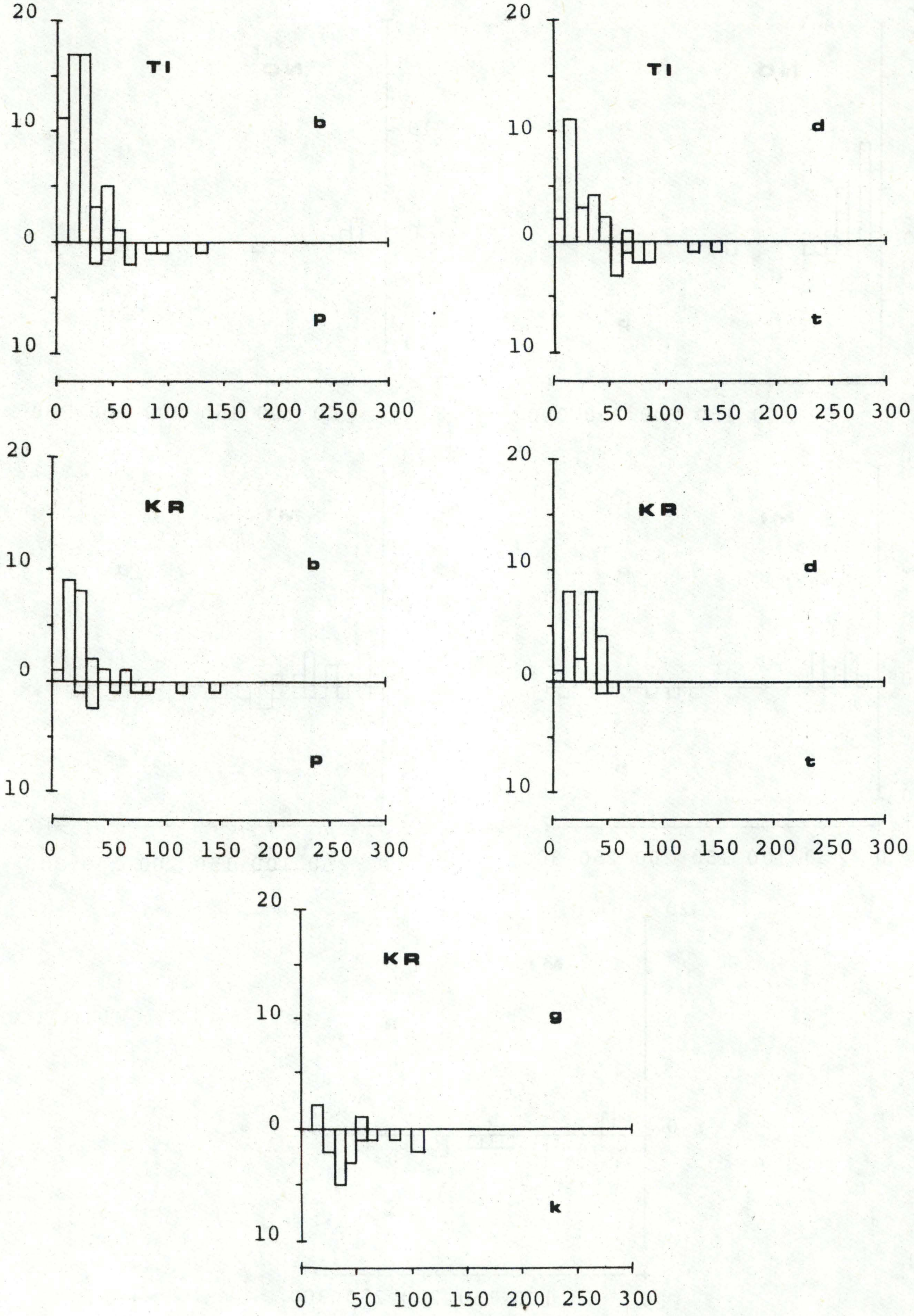


Figure 2

Histograms showing number of occurrences (vertical axis) vs. duration of the aspiration (horizontal axis). Subjects TI and KR.

Table 1

Levels of significance obtained in one-tailed Mann-Whitney U-tests of the hypothesis ptk>bdg with respect to voice onset time.

	p - b	t - d	k - g
TI	$p<0.01$	$p<0.001$	
KR	$p<0.01$	$p<0.001$	$p>0.05$
NO	$p<0.001$	$p<0.05$	
MI	$p<0.001$	$p<0.01$	$p<0.001$

Table 2

Tables showing the distribution of the measured VOT values relative to four different boundaries (30 ms, 40 ms, 50 ms, 60 ms).The terms "aspirated"/"unaspirated" in the tables mean +/- aspiration in the adult language in the words measured.

		VOT< +30 ms	VOT> +30 ms
"aspi-rated"	p	1	24
	t	1	29
	k	2	22
"unaspi-rated"	b	84	25
	d	33	29
	g	2	3

		VOT< +40 ms	VOT> +40 ms
"aspi-rated"	p	6	19
	t	3	27
	k	7	17
"unaspi-rated"	b	95	14
	d	48	14
	g	4	1

		VOT< +50 ms	VOT> +50 ms
"aspi-rated"	p	8	17
	t	5	25
	k	10	14
"unaspi-rated"	b	108	1
	d	56	6
	g	5	0

		VOT< +60 ms	VOT> +60 ms
"aspi-rated"	p	10	15
	t	10	20
	k	11	13
"unaspi-rated"	b	108	1
	d	56	6
	g	5	0

In connection with the results, the following points should be noted: Repetitions of a word have been treated as if they were new words, since most often repetitions occurred not because a word was repeated in a parrot like way but simply because the word was an important part of that child's vocabulary and consequently frequently used.

Clusters consisting of s + stop consonant often reduced to single stops. These stops have been treated as unaspirated stops.[1] The reason for doing so is partly that these stops are phonetically unaspirated in the adult language and partly that the children, not surprisingly, treated them like normal, unaspirated stops.

It should also be noted that e.g. TI at the time of recording had no velar stops in her speech, but substituted an alveolar stop for the velar one. These substitutions have been included in the alveolar stop category.

5. Discussion

The results indicate that Danish children at the age of two years, except for very few mistakes, keep aspirated and unaspirated stops apart in spontaneous speech.

In conformity with what has been found in other investigations, e.g. Kewley-Port et al. (1974), the present results seem to show that the marked members of the pairs (for Danish, the aspirated stops) are spread over a wider range in terms of VOT than the unaspirated stops.

At one point, however, there is a noticeable difference between the results of this study and those of others: the duration of the aspiration in the unmarked (unaspirated) stops is remarkably long compared to what has been found by other researchers.

---

1) The phonemic interpretation of the stop consonant (with regard to the voicing/aspiration feature) of /s/ + stop consonant is in Danish complicated by the existence of both /sf/ (found in a few words) and /sv/.

This is shown in table 2 from which it appears that the best separation of the VOT measurements into the aspirated and unaspirated categories is achieved when the boundary between the categories is placed at +50 ms VOT. At values lower than that (viz. +30 and +40 ms) and higher (+60 ms), the separation is poorer.[1] This is in disagreement with Port and Preston (1974), who state (p.198) that: "It is evident that the children's distributions are remarkably similar to one another. Each has a single mode, and the majority of the productions fall in the 0 to +20 ms voicing lag region." and with Preston et al. (1967), who in an investigation of prelinguistic Lebanese and American children summed up their results (p. 120) as follows: "With respect to apical stops for which the most data is available, these children show uni-modal distributions falling mainly in the 0 to 20 msec voicing lag range. This occurs in spite of the fact that the Lebanese children had been exposed to apical stops with voicing lag between 40 and 100 msec or greater."

The present data are also in disagreement with Lisker's (1970) assumption of the priority of the 0 - +30 ms VOT range.[2] Lisker's view is based on data from a cross-language investigation, Lisker et al. (1964), showing that all of the languages investigated possessing a VOT distinction among the stops are characterized by the following structure: short voicing lag plus either voicing lead or long voicing lag, the point being that the short voicing lag is never missing. The view that the short voicing lag range is the preferred one, gives support to the ideas advanced by Jakobson, and, naturally, the wish to find an articulatory explanation for this preference was strongly encouraged by the work of Lisker et al.

---

1) Setting the boundary at +50 ms also has the advantage that most of the deviant productions are found to be in the natural direction: marked (i.e. aspirated) stops are wrongly pronounced as unmarked (i.e. unaspirated) stops. A boundary at +30 ms implies that the mistakes would be going in the opposite, <u>unnatural</u> direction, i.e. from unmarked to marked.

2) This range may even be reduced to a 0 - +25 ms VOT range, according to the data presented in Lisker et al. (1964) and illustrated by the following quotation (p. 403): "...for it appears that the stop categories overall fall generally into three ranges: one from about -125 to -75 msec, one from zero to +25 msec, and a third from about +60 to +100 msec. The median values for these ranges are -100, +10 and +75 msec respectively."

The systematically longer aspiration found in my material gives rise to questions of some theoretical interest. Is it true that the VOT range 0 - +30 ms is the universally unmarked one? At what age does the influence of the adult linguistic system on that of the child make itself felt? Narrowing the scope to the main point of the present paper, we may ask whether it is accidental that the difference in aspiration in the unmarked stops of Danish vs. American children corresponds to the difference between adult Danish vs. American stop consonants? Complete answers will not be attempted in the following but a few points related to the problems presented will be touched upon.

In proposing an articulatory explanation for the priority of the unaspirated unvoiced stop, Kewley-Port et al. (1974) point to the timing complexity as an important factor, and state (p. 203): "The hypothesis is that short voicing lag stops are in specific ways easier for the infant to produce successfully than the other two types", (viz. voicing lead and long voicing lag).

In search for evidence in favour of their hypothesis, the authors focus on the timing relations between oral and glottal closure and taking (absolute) initial position as a point of departure, they argue that the correct production of an unaspirated stop makes relatively weak demands on the timing relations between oral and glottal closure, the only condition to be fulfilled being that at the time of oral release the glottal adduction should be completed - leaving freedom as to exactly when it is finished. The requirements for obtaining a correct production of both a voiced and an aspirated stop are stronger, since the same freedom of timing is not permitted here. For a successful production of an aspirated stop it is required that the glottal adduction starts no earlier than at the time of oral release.

The voiced stop requires glottal adduction to be completed at the time of onset of the oral closure - probably with the addition of an active widening of the pharyngeal and oral cavities. In evaluating which of these three types requires the highest degree of precision, it seems justified to leave aside the voiced one as it demands the use of extra mechanisms not needed in the other two. Now, considering the unvoiced aspirated and unaspirated stops, it is not obvious that the aspirated stop requires more precision than the unaspirated one, if there are no constraints on the degree of aspiration. In fact, it may be argued that the precision of timing required for the successful production of an aspirated stop is no greater than the precision required for an unaspirated stop, since the adduction in the aspirated stop may take place at any time after the release of the oral closure - the only consequence of the delayed adduction being a longer aspiration. Thus the aspirated stop may be said to permit more freedom of timing than the unaspirated one. Only if a fixed degree of aspiration is required can the aspirated stop be said to imply less freedom (i.e. higher precision).

To solve the problem of the different degree of complexity in laryngeal adjustments in aspirated and unaspirated stops, we must also consider them in a wider range of phonetic contexts than just word-initially after a pause. In a position between voiced segments, the glottal abduction should be considered too. As demonstrated by e.g. Kagaya et al. (1975), most often the abduction starts at the same point in time relative to the oral closure in aspirated and unaspirated stops, thus offering no help in deciding which is the more complex.

Finally, EMG data have been referred to in support of the view that the unaspirated stop is more simple than the aspirated one. A detailed discussion of the arguments shall not be attempted here. Instead, it seems appropriate to raise a few fundamental questions concerning the way in which EMG data are interpreted as evidence of phenomena such as ease of articulation and simplicity in general: Does a sudden forceful increase in muscular activity imply more effort (less ease of articulation) than a less sudden and less forceful one? Can the activity of one muscle - several muscles are known to function antagonistically - be viewed separately? If only the interaction of different muscles is of interest, how is then this complex interaction evaluated? Not until questions of this sort have been satisfactorily answered would it be safe to use EMG data in discussions concerning articulatory complexity, general articulatory difficulty, ease, or the like.

Giving unambiguous articulatory substance to expressions such as ease of articulation is evidently a difficult task but also an important one for our understanding of both first language acquisition and language in general.

As demonstrated by data presented by Fischer-Jørgensen (1979), the boundary in terms of VOT between the two categories (i.e. aspirated vs. unaspirated) is often found to be at 50 ms or even 60 ms in the speech of young subjects from the Copenhagen area. The absolute VOT values are, however, greatly influenced by factors such as place of articulation of the stop consonant (velars showing the longest and labials the shortest aspiration)[1], and the quality of the following vowel (the longest aspiration is found before high vowels). The younger subjects of the Fischer-Jørgensen investigation were from 25 to 30 years old, and it may be hypothesized that the tendency towards long aspiration in the unaspirated stops is more pronounced among still younger subjects - this, however, has not yet been systematically investigated.

Comparing the VOT system of relatively young adult speakers (as given by Fischer-Jørgensen (1979)) with the VOT system of the two-year-old children (as presented in this investigation), the two systems turn out to be almost coinciding. It is natural that such a coincidence takes place at some stage of the linguistic development, what is surprising is that it takes place at this early stage considering the "deviant" nature of the adult system.

In conclusion, if there is a universally and naturally unmarked VOT range of 0 - +30 ms, the period of time during which it is preferred by Danish children must be extremely short and occur very early in their linguistic development, since at the age of two it has already moved into the +50 ms range.

On the other hand, it may be hypothesized on the basis of the present data that the unmarked range extends beyond +30 ms VOT. If this is the case, it has implications for the analysis of the adult VOT system of e.g. American English. One consequence is that for that language, all of the unvoiced and (almost) half of the occurrences of the aspirated stops are comprised in the proposed unmarked region. Such a discrepancy between naturally unmarked regions and boundaries compared to what is found in actual languages is hardly acceptable.

---

1) It is interesting to note that this is not always the case in my investigation of the children's speech.

6. Linguistic observations of children younger than two years

Since the children at the day nursery at which the recordings took place were between 6 weeks and three years of age, it was possible to observe the linguistic behaviour of children younger than those selected for the present investigation. The following are impressionistic observations of the aspiration in the stops of very young children:

At the end of the first year of life, the first few words with aspirated stops are produced (these may even be among the very first recognizable words of the child's speech). However, only a minority of the words with adult aspiration is produced with an aspirated stop by the child. The majority of these words contain an unaspirated stop in the child's speech. When aspirated stops do occur they almost always do so in words with aspirated stop in the adult language. The relatively few words with aspirated stop in the child's speech is pronounced with aspiration whenever pronounced - exceptions were found to be very few.

The observations reported above are in conformity with what other investigators have found, e.g. Ferguson et al. (1973).

Finally, it should be pointed out that making linguistic observations of very young children is a difficult task. One of the main problems is to decide whether a given vocalisation uttered by the child is to be characterized as babbling or a (reduced) linguistic word. It is of no help, of course, to ask the child if a given semantic content is attached to what was uttered; and making use of e.g. games always implies the addition of other factors than the linguistic one. The decision is made solely by the observer and not necessarily in conformity with what was intended by the child.

References

Abrahams, H. 1955: "Børns tilegnelse af modersmålets fonologiske stof", NTTS 15, p. 13-26

Basbøll, H. 1969: "The phoneme system of Advanced Standard Copenhagen", ARIPUC 3, p. 33-55

Danielsson, B., C. Nilsson and B. Sahlén 1976: "Fonologiskt status för jämförelse mellan produktion och imitation", Unpublished, Lund

Edwards, M.L. 1974: "Perception and production in child phonology: The testing of a hypothesis", Child Language Development 7, p. 67-85

Faircloth and Faircloth 1970: "An analysis of the articulatory speech defective child in connected speech and in isolated word responses", JSHD 35, p. 51-61

Ferguson, Ch.A. and C.B. Farwell 1973: "Words and sounds in early language acquisition: English initial consonants in the first fifty words", Child Language Development 6, p. 1-61

Fischer-Jørgensen, Eli 1979: "Temporal relations in consonant-vowel syllables with stop consonants based on Danish material", Frontiers of Speech Communication Research (eds.: B. Lindblom and S. Öhman), p. 51-68

Garnica, O. 1971: "The development of the perception of phonemic differences in initial consonants by English-speaking children: A pilot study", Child Language Development 3, p. 1-31

Gilbert, J.H.V. 1977: "A voice onset time analysis of apical stop production in 3-year olds", Journal of Child Language 4, p. 1-8

Ingram, D. 1976: Phonological Disability in Children, Oxford

Jakobson, R. 1942: Kindersprache, Aphasie und allgemeine Lautgesetze, Uppsala

Jespersen, O. 1916: Nutidssprog hos Børn & Voxne, Copenhagen

Jørgensen, J. 1978: ""Stød" and unexpected substitutions in the speech of Danish 2-years-old children", ARIPUC 12, p. S-25 - S-29

Jørgensen, P. 1971: "Mine børns barnesprog", Acta Philologica Scandinavica 29,1, p. 1-17

Kagaya, R. and H. Hirose 1975: "Fiberoptic electromyographic and acoustic analyses of Hindi stop consonants", Annual Bulletin RILP 9, p. 27-46

Lisker, L. and A.S. Abramson 1964: "A cross-language study of voicing in initial stops: Acoustical measurements", Word 20, p. 384-422

Lisker, L. 1970: "On learning a new contrast", Haskins SR-24, p. 1-15

Port, D.K. and M.S. Preston 1974: "Early apical stop production: a voice onset time analysis", J.Ph. 2, p. 195-210

Preston, M.S., G. Yeni-Komshian and R.E. Stark 1967: "A study of voicing in initial stops found in the pre-linguistic vocalizations of infants from different language environments", Haskins SR 10, p. 109-120

Templin, M. 1947: "Spontaneous versus imitated verbalization in testing articulation in preschool children", JSHD 12, p. 293-300

Winterkorn, J.M.S., R.F. MacNeilage and M.S. Preston 1967: "Perception of voiced and voiceless consonants in three years old children", Haskins SR-11, p. 41-45 [illegible], J. (1978). "[illegible] and unexpected substitutions in the speech of [illegible] patients" [illegible] 27, [illegible]

[illegible], P. (19[illegible]). "[illegible] between [illegible]" [illegible] 23:1, p. [illegible]

[illegible], S. and H. Hirose (1976). "[illegible] electromyographic and [illegible] analysis of [illegible]" Annual Bulletin [illegible] 10, [illegible]-45

[illegible] and [illegible] (1984). "[illegible]-language study of [illegible]" Word 20, [illegible]

[illegible] (19[illegible]). "[illegible] contrast" [illegible] 22-24, p. 1-15

[illegible] and [illegible] (1974). "[illegible] production [illegible] analysis" 27, [illegible]-210

[illegible] and R. [illegible] (1977). "[illegible] of voicing in [illegible] consonants [illegible]" 10, [illegible]-128

[illegible] (1977). "[illegible] initiated [illegible]" [illegible]-100

[illegible] and R. [illegible] (19[illegible]). "[illegible]" [illegible] 28:1, [illegible]-41-45

(ARIPUC 13, 1979)

VARIATION IN INHERENT Fo LEVEL DIFFERENCES BETWEEN VOWELS AS A FUNCTION OF POSITION IN THE UTTERANCE AND IN THE STRESS GROUP

Niels Reinholt Petersen

Abstract: The inherent Fo level differences in the vowels u and ɑ were examined in a material which made it possible to vary the position of the vowels in the utterance, keeping their position in the stress group constant, and vice versa. The main finding was that the inherent Fo level differences were statistically significant in both stressed and unstressed syllables throughout the material, but in stressed vowels the differences were larger than in the unstressed ones. The effect of position in the utterance and stress group can be summarized as follows: In vowels in stressed position the differences tend to decrease through the utterance, in vowels in first posttonic position they (surprisingly) increase, and in second posttonic position they do not seem to vary systematically with the position in the utterance. Apart from the differences being larger in stressed than in unstressed syllables, there seems to be no effect in the unstressed vowels from their position in the stress group.

1. Introduction

The course of fundamental frequency in non-tonal languages such as Danish may be described as the superposition of three simultaneous main components (apart from emotional factors): 1) a sentence component, which conveys information about the utterance as a whole (is it declarative, interrogative, continuative?), 2) a stress group component, and 3) a segmental component, which is specified by the segments of the utterance.

The sentence and stress group components are mainly language specific (an analysis of these components in Advanced Standard Copenhagen Danish is reported in Thorsen, 1978 and 1979a). The segmentally determined Fo variation, on the other hand, is assumed to be universal[1] and to be a consequence of physiological, aerodynamic and (possibly) acoustical factors in the production of speech (see, e.g., Ohala (1978) and Reinholt Petersen (1978)).

One of the segmental factors influencing fundamental frequency is vowel height. There seems to be a universal tendency for Fo to be higher in high vowels than in low vowels, other things being equal. This phenomenon, which has been termed inherent Fo, intrinsic Fo, vowel specific Fo, has been shown to exist in a number of languages, e.g. English (Peterson and Barney, 1952, Lehiste and Peterson, 1961), French (Di Cristo and Chafcouloff, 1977), German (Neweklowsky, 1975), Swedish (Löfqvist, 1975), Danish (Reinholt Petersen, 1978), Yoruba (Hombert, 1976), and other West African languages (Ladefoged, 1964). (A summary of the literature is given in Hirst, Di Cristo and Nishinuma, 1979.)

The majority of these experiments, however, employed stressed vowels in words embedded in short carrier phrases, and have thus investigated a rather limited subset of the contexts where inherent Fo level differences may play a role in speech.

Reinholt Petersen (1978) used a test material which permitted inherent Fo to be considered in both stressed and unstressed syllables. In that material, vowels were inserted in nonsense words of the structure CV'CV:CV in the carrier phrase "Stavelserne i _____ forkortes" ('The syllables of _____ are shortened'). For technical reasons no measurements were made of Fo in the unstressed vowels, but tracings of the Fo curves showed a clear tendency towards smaller differences between i/u vs. ɑ in unstressed than in stressed vowels.

The purpose of the experiments reported below was to investigate the inherent Fo level variation under more varied conditions. The questions to be considered were the following: Do we find significant Fo level variation in unstressed as well as in stressed syllables, and can the magnitude of the variation be shown to be

---

1) This applies to non-tonal languages. In Yoruba, a tonal language, Hombert (1976) has found a tendency to actively minimize the effect of preceding consonants on Fo in vowels.

related to (a) position in the utterance, and (b) position in the stress group (the stress group being defined as a stressed syllable plus the following unstressed syllables up to the next stressed one, irrespective of intervening word or morpheme boundaries, or to the end of the utterance)?

The focusing upon the stress group as a frame of reference in the investigation was due to the fact that the stress group seems to be a basic unit in the description of intonation in Advanced Standard Copenhagen Danish (cf. Thorsen, 1978) and also in the description of variation of vowel duration (Peter Holtse, personal communication).

2. Method

2.1 Material

The test vowels chosen were u and ɑ, embedded in nonsense words of the type mVmVmV, because these vowels have been shown (Reinholt Petersen, 1978) to have the largest inherent Fo differences, and m should have a negligible influence on Fo of the vowels.

The test words appeared in three different versions, viz. with the stress on the first, second, and third syllable, respectively, and each version was placed initially, medially, and finally in a carrier sentence. Thus, the material consisted of the following 18 test sentences:[1]

1 I 'mumumu muteres em'erne
2 I mu'mumu muteres em'erne (In ___ the m's are mutated)
3 I mumu'mu muteres em'erne

4 Em'erne i 'mumumu muteres
5 Em'erne i mu'mumu muteres (The m's in ___ are mutated)
6 Em'erne i mumu'mu muteres

7 Em'erne muteres i 'mumumu
8 Em'erne muteres i mu'mumu (The m's are mutated in ___)
9 Em'erne muteres i mumu'mu

10 I 'marmarmar markeres em'erne
11 I mar'marmar markeres em'erne (In ___ the m's are marked)
12 I marmar'mar markeres em'erne

1) Note that the orthographic sequence ar is pronounced [ɑ].

13 Em'erne i 'marmarmar markeres
14 Em'erne i mar'marmar markeres (The m's in __ are marked)
15 Em'erne i marmar'mar markeres
16 Em'erne markeres i 'marmarmar
17 Em'erne markeres i mar'marmar (The m's are marked in _)
18 Em'erne markeres i marmar'mar

The relevant syllables are underlined. Note that the word immediately following the testword varies with the testword: muteres after u-words and markeres after ɑ-words. This made it possible to include the first syllable of these words in the test material and thus expand it without having extremely long nonsense words. On the other hand, the segmental structure of these syllables was different from that of the nonsense words, and - what might possibly be more serious - they differed among themselves: u was followed by t and ɑ by k.

It was not clear to what extent the different segmental conditions in the material might be expected to affect the fundamental frequency of the vowels and particularly the Fo difference between u and ɑ in the first syllables of muteres and markeres in comparison with the Fo differences in the nonsense words. From Johansson (1976) it appears that voiceless stops give rise to a higher Fo in the preceding vowel than do nasal consonants (Johansson does not specify the point of measurement in the vowel). On the other hand, Lehiste and Peterson (1961) and Jeel (1975) found no such effect. To my knowledge, an interaction between vowel height and following consonant, particularly with regard to its place of articulation, that would influence the magnitude of the inherent differences between vowels, has never been attested.

In an attempt to clarify these matters under conditions comparable to those of the present investigation, a small, supplementary experiment was conducted, the details of which are given in Appendix I. In that experiment no significant effect on Fo in u and ɑ from the following consonant (m, t, k) could be shown. Nor was there any significant interaction between t and k and preceding u and ɑ. On this basis it was considered justified to include the first syllable of the words muteres and markeres in the test material.

2.2 Recordings, subjects

The 18 test sentences were combined with 9 distractor sentences in 10 different random orders in a reading list. The list was read once by each subject. The recordings took place in a sound treated room at the Institute, using a Sennheiser MD21 microphone and a REVOX A77 tape recorder. The subjects were instructed to read the sentences with a neutral, declarative intonation and to use a speaking rate natural to them. There were four speakers: three female (ER, SI, and KM) and one male (NR, the author), all phoneticians and all speakers of Advanced Standard Copenhagen Danish.

2.3 Registrations and measurements

The apparatus used for registration was a REVOX A77 tape recorder, an intensity meter, and a fundamental frequency extractor (both F-J Electronics), and a Mingograph (Elema 800). The following acoustic registrations were made: duplex oscillogram, linear HiFi intensity curve (integration time 2.5 ms), logarithmic HP-filtered (500 Hz) intensity curve (integration time 5 ms), fundamental frequency curve.

The duplex oscillogram and the intensity curves were used for segmentation. The accuracy of segmentation was ±0.5 cs. With regard to the Fo measurements, it proved impossible to establish a point of measurement which was defined on the basis of the Fo curve itself (e.g. Fo minimum or Fo maximum) and which could be used consistently throughout the material. Therefore, Fo was measured at the middle of the test vowels. This point has the advantage of being easily determined from the duplex oscillogram and intensity curves, and it is appropriate in the sense that the midpoint Fo measure yields a reasonably accurate description of the Fo movement in the testwords. This is illustrated in fig. 1, where actual Fo curves are compared with curves interpolated between the midpoint Fo values. The maximum difference between the curves is about 3 Hz for subject NR and 7 Hz for subject KM, i.e. about 3 per cent. A few samples had to be dismissed for various reasons, but altogether 2537 Fo measurements were obtained out of 2640 possible (4 subjects x 10 repetitions x (12 sentences x 4 vowels + 6 sentences x 3 vowels)).

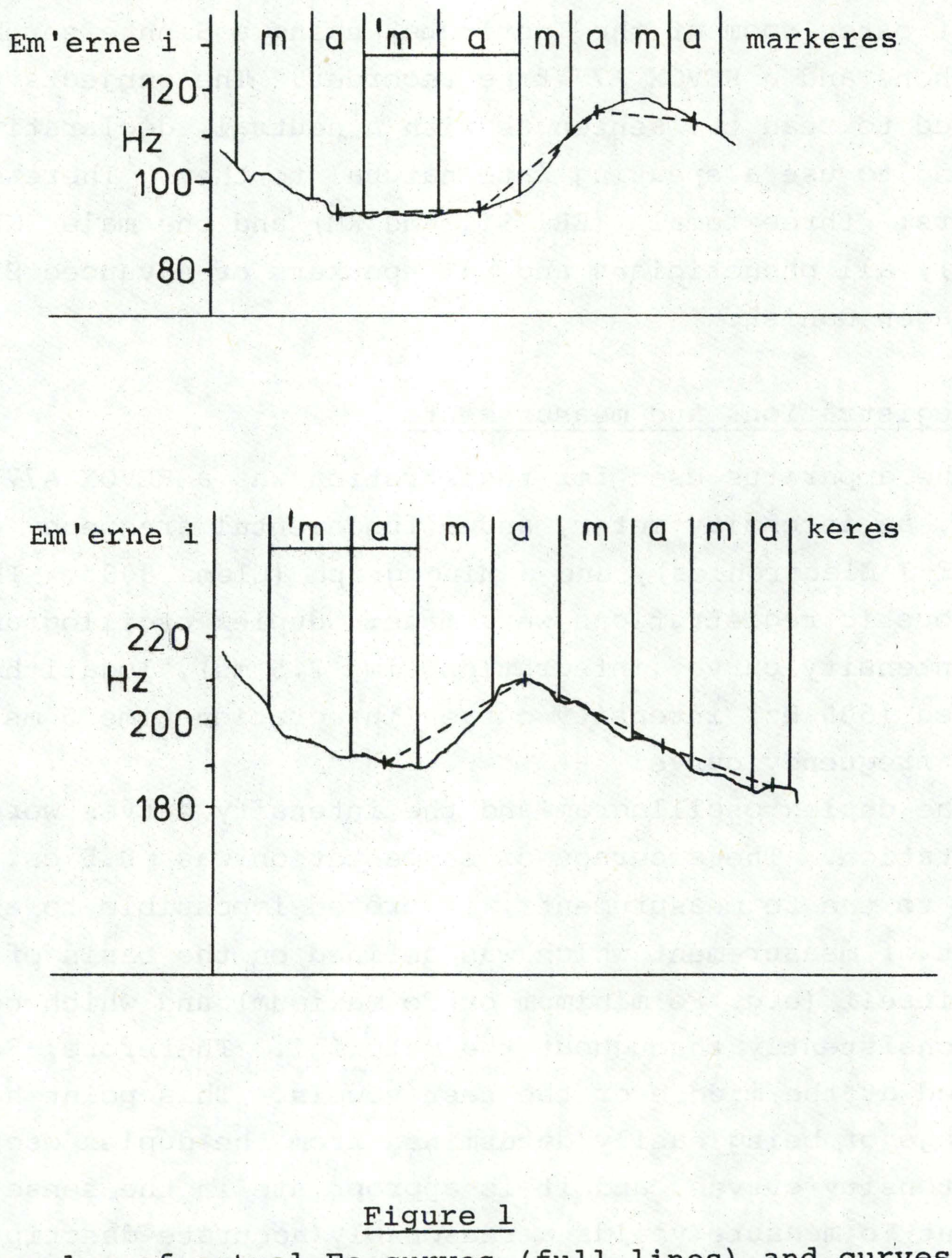


Figure 1

Examples of actual Fo curves (full lines) and curves interpolated between the Fo values measured at the middle of the vowels (broken line). The short vertical lines in the curves indicate the points of measurement. Subject NR (top) and subject KM (bottom).

3. Results

Mean fundamental frequencies and standard deviations for all vowels in the test material are tabulated for the four speakers in tables 1 to 4, and the means are shown graphically in figs. 2 to 5. Furthermore, the means obtained for each subject were converted into a relative measure, viz. the deviation in semitones from that subject's stressed ɑ in sentence 10 (i.e. the ɑ-sentence in which the test word occurs in initial position and is stressed on the first syllable). Fig. 6 shows the means over all subjects expressed in semitones.

It appears that the Fo measures obtained in the present experiment are in good agreement with what should be expected from Thorsen's (1978) model for the Fo course in declarative sentences in Advanced Standard Copenhagen Danish (part of Thorsen's model is shown in fig. 7). Furthermore, it is clear that the magnitude of the Fo variation determined by the stress groups decreases through the utterance, also when expressed in semitones, i.e. the decrease of the absolute variation through the utterance as shown in figs. 2 to 5 does not correspond to a constant relative variation (see also Thorsen, 1979b).

3.1 Inherent Fo variation

Tables 5 to 8 enumerate the mean Fo differences between u and ɑ. The statistical significance of the differences was tested by a series of t-tests. The levels of significance achieved are given in the tables.

In stressed syllables the differences between u and ɑ were significant at a very high level ($p < 0.001$) in all cases except one, namely in utterance final position for subject NR; here the difference was significant at the 1 per cent level.

In unstressed syllables the differences were slightly less clear, but still a level of significance of 5 per cent or better was achieved in 85 out of 96 cases (i.e. 24 unstressed vowels x 4 speakers). The fundamental frequency of ɑ was in no case higher than that of u.

From tables 5 to 8 and figs. 2 to 5 it appears that the magnitude of the differences between the Fo of u and ɑ is correlated

Table 1

Means and standard deviations (in Hz), arranged in accordance with the structure of the material (see the list in section 2.1 above). The stress groups are delimited by vertical lines. Subject ER.

	utterance position									
	1	2	3	4	5	6	7	8	9	lo
u	i	296.8 11.o9	369.7 17.62	31o.8 14.63	285.6 14.75	te	res	em'	er	ne
ɑ	i	248.9 2o.78	351.8 19.59	277.4 2o.71	252.3 14.41	ke	res	em'	er	ne
u	i	249.7 1o.63	269.6 16.81	338.5 12.42	284.3 16.15	te	res	em'	er	ne
ɑ	i	228.3 1o.7o	237.2 14.o6	323.8 24.36	271.5 9.66	ke	res	em'	er	ne
u	i	254.5 12.42	244.6 17.32	3o9.6 15.41	327.1 18.oo	te	res	em'	er	ne
ɑ	i	231.6 9.58	222.o 8.79	252.o 15.1o	3o3.7 26.22	ke	res	em'	er	ne
u	em'	er	ne	i	267.8 15.51	313.2 19.14	272.5 1o.78	267.2 14.26	te	res
ɑ	em'	er	ne	i	225.1 4.48	284.2 12.45	237.3 11.97	23o.3 8.9o	ke	res
u	em'	er	ne	i	26o.8 11.51	269.3 14.38	313.3 12.6o	278.4 14.67	te	res
ɑ	em'	er	ne	i	24o.1 19.72	215.3 8.78	277.7 12.75	254.o 12.9o	ke	res
u	em'	er	ne	i	262.o 11.17	245.3 9.33	282.4 19.35	3oo.7 17.81	te	res
ɑ	em'	er	ne	i	244.1 13.76	224.2 12.o8	238.9 13.85	281.o 9.85	ke	res
u	em'	er	ne	mu	te	res	i	222.3 12.32	24o.1 13.o6	2o7.8 7.4o
ɑ	em'	er	ne	mar	ke	res	i	194.7 9.11	217.7 7.45	193.1 4.14
u	em'	er	ne	mu	te	res	i	246.5 16.86	214.4 6.54	222.o 9.67
ɑ	em'	er	ne	mar	ke	res	i	229.7 18.74	187.4 5.72	205.3 8.34
u	em'	er	ne	mu	te	res	i	241.8 13.15	226.o 15.19	211.1 7.42
ɑ	em'	er	ne	mar	ke	res	i	229.o 13.o3	216.2 13.75	184.5 6.oo

Table 2

Means and standard deviations (in Hz), arranged in accordance with the structure of the material (see the list in section 2.1 above). The stress groups are delimited by vertical lines. Subject SI.

	utterance position									
	1	2	3	4	5	6	7	8	9	lo
u	i	285.o lo.72	326.4 13.88	282.1 12.6o	259.8 9.4o	te	res	em'	er	ne
a	i	242.8 7.27	3o3.3 lo.14	27o.3 7.87	251.5 6.11	ke	res	em'	er	ne
u	i	26o.6 14.35	28o.7 17.62	322.1 16.34	289.1 14.65	te	res	em'	er	ne
a	i	239.7 6.95	241.6 8.92	3o3.1 14.39	286.3 8.54	ke	res	em'	er	ne
u	i	256.6 8.14	250.8 lo.56	284.3 lo.ol	3o8.8 15.3o	te	res	em'	er	ne
a	i	245.o 8.29	238.3 7.23	257.4 12.65	298.3 12.62	ke	res	em'	er	ne
u	em'	er	ne	i	262.8 11.77	288.9 12.79	262.8 7.69	249.8 8.66	te	res
a	em'	er	ne	i	223.o lo.19	274.4 18.55	253.9 11.63	24o.9 8.41	ke	res
u	em'	er	ne	i	253.4 5.17	247.o 5.o6	276.9 6.97	266.2 6.lo	te	res
a	em'	er	ne	i	243.o 6.27	217.8 6.49	257.9 8.49	259.o lo.49	ke	res
u	em'	er	ne	i	264.3 6.11	25o.9 3.45	261.6 6.46	28o.9 7.ol	te	res
a	em'	er	ne	i	246.5 3.34	234.o 5.ol	235.3 4.76	272.3 9.64	ke	res
u	em'	er	ne	mu	te	res	i	242.7 12.o4	264.1 18.o3	232.8 9.o4
a	em'	er	ne	mar	ke	res	i	211.4 7.92	242.8 12.74	227.3 8.44
u	em'	er	ne	mu	te	res	i	254.4 12.47	228.9 14.54	235.7 15.o7
a	em'	er	ne	mar	ke	res	i	248.1 11.39	2o5.9 7.oo	226.7 7.21
u	em'	er	ne	mu	te	res	i	263.1 6.lo	249.5 7.43	224.3 8.65
a	em'	er	ne	mar	ke	res	i	245.2 4.49	233.9 7.19	2o3.6 6.5o

Table 3

Means and standard deviations (in Hz), arranged in accordance with the structure of the material (see the list in section 2.1 above). The stress groups are delimited by vertical lines. Subject KM.

utterance position

	1	2	3	4	5	6	7	8	9	10
u	i	198.6 8.30	230.8 9.19	206.7 5.56	191.7 6.17	te	res	em'	er	ne
a	i	182.0 6.57	223.8 9.75	198.8 5.83	185.0 3.81	ke	res	em'	er	ne
u	i	187.5 5.13	196.8 6.12	228.0 7.67	204.5 4.77	te	res	em'	er	ne
a	i	175.3 4.22	176.1 3.96	216.4 8.54	201.0 6.94	ke	res	em'	er	ne
u	i	182.4 4.86	176.9 5.16	199.6 3.92	214.6 9.13	te	res	em'	er	ne
a	i	171.6 4.93	169.1 3.63	183.0 7.16	208.7 8.98	ke	res	em'	er	ne
u	em'	er	ne	i	188.6 4.72	212.0 6.00	193.3 5.62	182.5 2.17	te	res
a	em'	er	ne	i	173.0 4.78	200.3 4.47	185.4 6.47	177.8 4.73	ke	res
u	em'	er	ne	i	187.7 5.33	186.5 6.22	209.1 6.40	193.4 7.24	te	res
a	em'	er	ne	i	178.2 6.25	166.9 5.32	197.3 10.23	184.6 6.80	ke	res
u	em'	er	ne	i	184.9 4.31	178.0 4.35	192.4 6.33	204.3 4.40	te	res
a	em'	er	ne	i	172.7 4.08	167.7 4.08	176.6 5.58	199.0 6.41	ke	res
u	em'	er	ne	mu	te	res	i	179.2 5.05	185.9 4.26	179.0 6.60
a	em'	er	ne	mar	ke	res	i	165.0 4.67	175.2 6.53	172.1 5.40
u	em'	er	ne	mu	te	res	i	184.0 7.45	172.0 3.16	187.6 4.24
a	em'	er	ne	mar	ke	res	i	173.4 5.78	157.8 2.70	170.3 4.81
u	em'	er	ne	mu	te	res	i	180.6 3.57	171.7 3.83	172.5 4.84
a	em'	er	ne	mar	ke	res	i	172.5 4.22	163.0 4.02	162.3 3.62

Table 4

Means and standard deviations (in Hz), arranged in accordance with the structure of the material (see the list in section 2.1 above). The stress groups are delimited by vertical lines. Subject NR.

utterance position

	1	2	3	4	5	6	7	8	9	1o
u	i	1o1.o 2.45	124.7 3.27	118.o 3.92	1o8.9 3.38	te	res	em'	er	ne
ɑ	i	87.5 2.68	117.3 4.22	11o.5 3.69	95.8 2.1o	ke	res	em'	er	ne
u	i	91.8 2.39	99.9 3.45	123.5 4.45	113.3 2.54	te	res	em'	er	ne
ɑ	i	85.4 3.27	87.6 2.17	117.o 4.83	1o9.1 4.o4	ke	res	em'	er	ne
u	i	93.o 1.63	92.o 1.7o	1o3.5 2.92	125.4 4.o1	te	res	em'	er	ne
ɑ	i	87.1 3.35	86.2 2.57	93.4 2.76	121.o 4.4o	ke	res	em'	er	ne
u	em'	er	ne	i	1oo.o 2.87	117.2 5.57	111.8 4.57	1o4.9 3.45	te	res
ɑ	em'	er	ne	i	86.8 1.o3	1o9.8 7.o8	1o5.o 4.32	95.o 4.88	ke	res
u	em'	er	ne	i	95.6 3.13	99.3 2.55	117.1 4.46	11o.2 2.68	te	res
ɑ	em'	er	ne	i	89.8 1.23	87.3 1.64	11o.o 3.27	1o7.1 3.67	ke	res
u	em'	er	ne	i	98.1 2.88	94.7 2.63	1o2.5 4.4o	119.3 5.5o	te	res
ɑ	em'	er	ne	i	89.o 2.62	84.7 3.oo	9o.9 2.42	11o.9 3.67	ke	res
u	em'	er	ne	mu	te	res	i	89.3 3.27	9o.7 2.45	91.4 4.43
ɑ	em'	er	ne	mar	ke	res	i	82.9 2.28	82.8 1.93	85.3 1.34
u	em'	er	ne	mu	te	res	i	97.1 5.9o	89.3 2.65	88.o 2.65
ɑ	em'	er	ne	mar	ke	res	i	88.4 6.oo	78.6 4.25	8o.o 2.33
u	em'	er	ne	mu	te	res	i	97.2 3.26	94.3 2.83	86.2 3.26
ɑ	em'	er	ne	mar	ke	res	i	91.5 4.79	9o.1 2.69	8o.6 5.2o

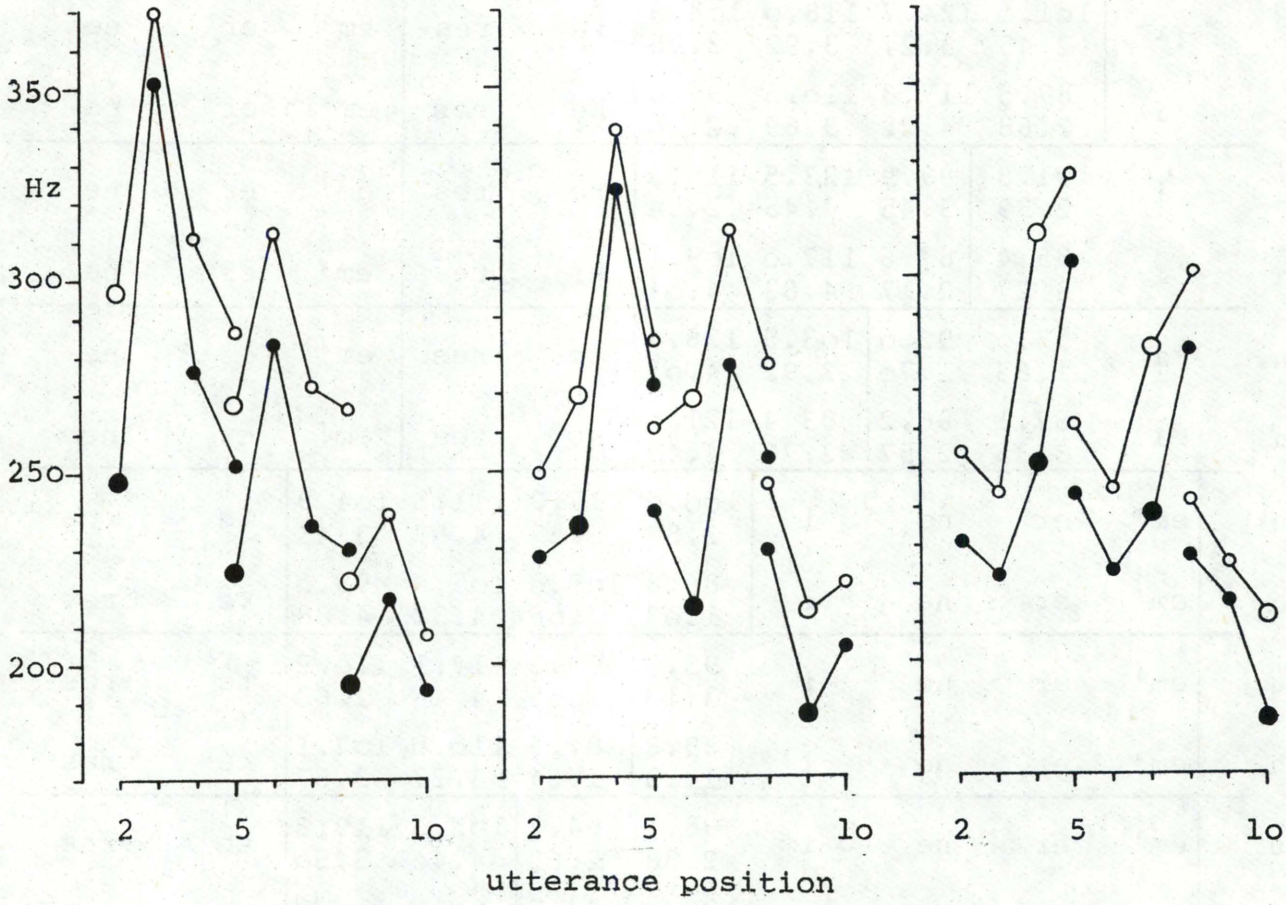


Figure 2

Mean fundamental frequencies (in Hz) in the test words. The data-points belonging to the same word are connected by straight lines. The left, middle, and right graphs display words with the stress on the first, second, and third syllables, respectively. Large circles represent stressed and small circles unstressed syllables. u is indicated by open and ɑ by filled circles. Subject ER.

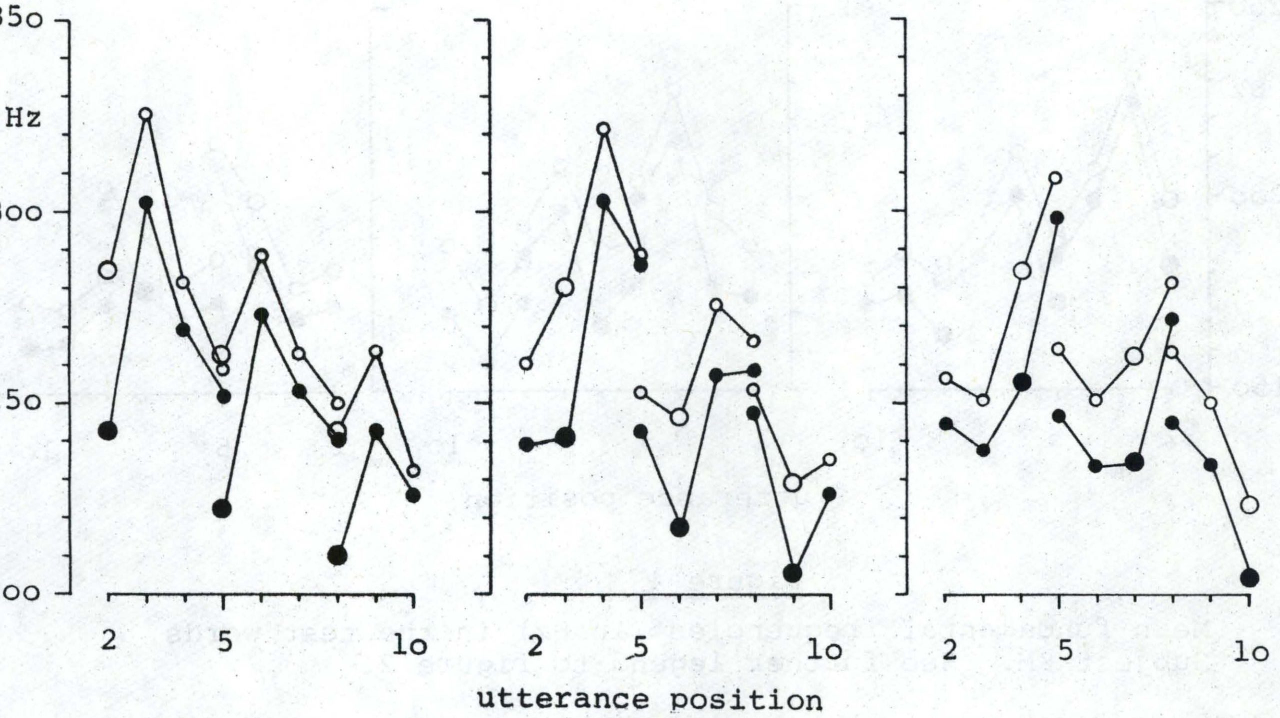


Figure 3

Mean fundamental frequencies (in Hz) in the test words. Subject SI. See further legend to figure 2.

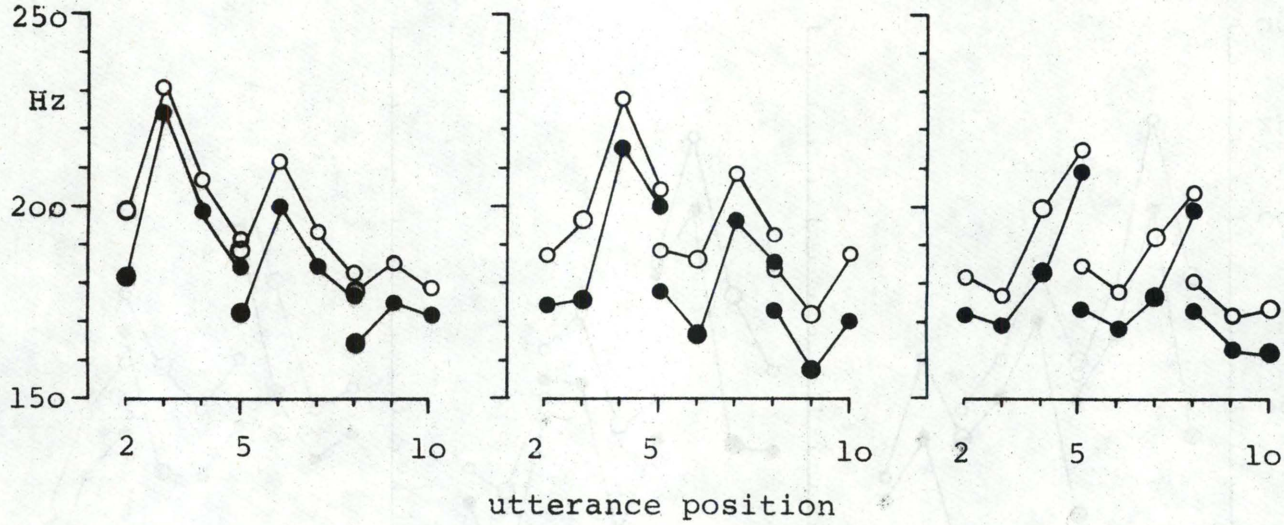


Figure 4

Mean fundamental frequencies (in Hz) in the test words. Subject KM. See further legend to figure 2.

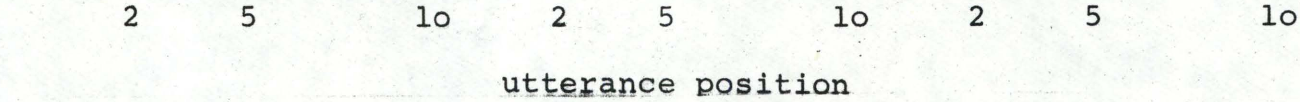


Figure 5

Mean fundamental frequencies (in Hz) in the test words. Subject NR. See further legend to figure 2.

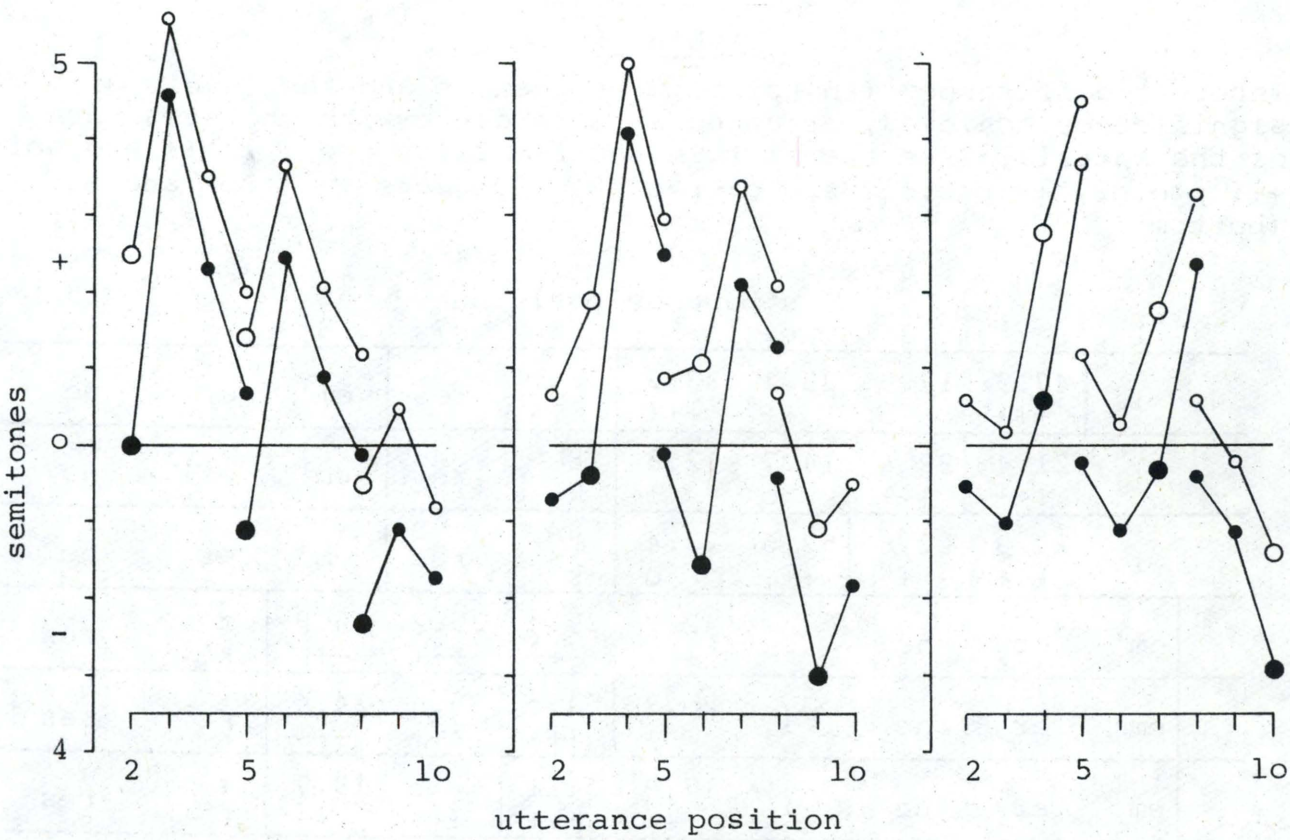


Figure 6

Mean Fo (in semitones) in the test words. All subjects pooled. See further legend to figure 2.

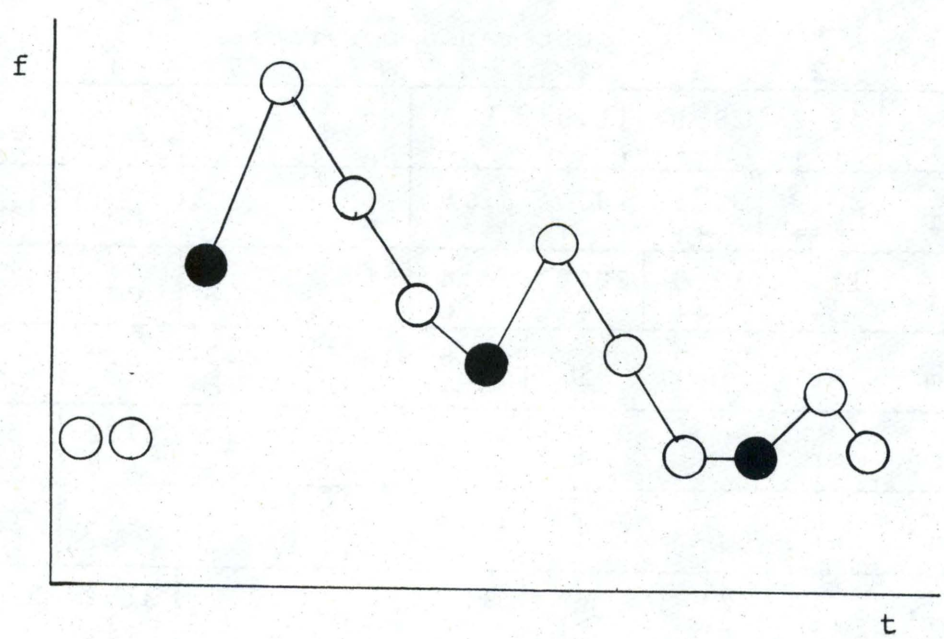


Figure 7

Part of Thorsen's (1978) model for the Fo course in short declarative sentences in Advanced Standard Copenhagen Danish. Filled circles indicate stressed and open circles indicate unstressed syllables.

Tables 5-6

Inherent differences (in Hz) between u and ɑ and the levels of significance achieved, arranged in accordance with the structure of the material (see the list in section 2.1 above). +++: p<o.ool, ++: p<o.ol, +: p<o.o5, and o: p>o.o5. Subjects ER (top) and SI (bottom).

utterance position

1	2	3	4	5	6	7	8	9	lo
i	47.9 +++	17.9 o	33.3 ++	33.2 +++	$^{t}_{k}$e	res	em'	er	ne
i	21.4 +++	32.4 +++	14.7 o	12.8 +	$^{t}_{k}$e	res	em'	er	ne
i	22.9 +++	22.6 ++	57.6 +++	23.4 o	$^{t}_{k}$e	res	em'	er	ne
em'	er	ne	i	42.7 +++	29.o ++	35.2 +++	36.9 +++	$^{t}_{k}$e	res
em'	er	ne	i	2o.7 +	54.o +++	35.6 +++	24.4 +++	$^{t}_{k}$e	res
em'	er	ne	i	17.9 ++	21.1 +++	43.5 +++	19.7 ++	$^{t}_{k}$e	res
em'	er	ne	m$^{u}_{ar}$	$^{t}_{k}$e	res	i	27.6 +++	22.4 +++	14.7 +++
em'	er	ne	m$^{u}_{ar}$	$^{t}_{k}$e	res	i	16.8 +	27.o +++	16.7 +++
em'	er	ne	m$^{u}_{ar}$	$^{t}_{k}$e	res	i	12.8 +	9.8 o	26.6 +++

utterance position

1	2	3	4	5	6	7	8	9	lo
i	42.2 +++	23.1 +++	11.8 +	8.3 +	$^{t}_{k}$e	res	em'	er	ne
i	2o.8 +++	39.1 +++	19.o +	2.8 +++	$^{t}_{k}$e	res	em'	er	ne
i	11.6 ++	12.5 ++	26.9 +++	1o.4 o	$^{t}_{k}$e	res	em'	er	ne
em'	er	ne	i	39.8 +++	14.5 o	8.9 o	8.9 +	$^{t}_{k}$e	res
em'	er	ne	i	1o.4 +++	29.2 +++	19.o +++	7.2 o	$^{t}_{k}$e	res
em'	er	ne	i	17.8 +++	16.9 +++	26.3 +++	8.6 +	$^{t}_{k}$e	res
em'	er	ne	m$^{u}_{ar}$	$^{t}_{k}$e	res	i	31.2 +++	21.3 +	5.4 o
em'	er	ne	m$^{u}_{ar}$	$^{t}_{k}$e	res	i	6.3 o	23.o +++	9.o o
em'	er	ne	m$^{u}_{ar}$	$^{t}_{k}$e	res	i	17.9 +++	15.6 +++	2o.7 +++

Tables 7-8

Inherent differences (in Hz) between u and ɑ and the levels of significance achieved, arranged in accordance with the structure of the material (see the list in section 2.1 above). +++: p<o.ool, ++: p<o.ol, +: p<o.o5, and o: p>o.o5. Subjects KM (top) and NR (bottom).

utterance position

1	2	3	4	5	6	7	8	9	lo
i	16.6 +++	7.o o	7.9 ++	6.7 +	(t/k)e	res	em'	er	ne
i	12.2 +++	2o.7 +++	11.6 ++	3.5 o	(t/k)e	res	em'	er	ne
i	lo.8 +++	7.8 ++	16.6 +++	5.9 o	(t/k)e	res	em'	er	ne
em'	er	ne	i	15.6 +++	11.7 +++	7.9 ++	4.7 +	(t/k)e	res
em'	er	ne	i	9.5 ++	19.6 +++	11.8 ++	8.8 +	(t/k)e	res
em'	er	ne	i	12.2 +++	lo.3 +++	15.8 +++	5.3 +	(t/k)e	res
em'	er	ne	m(u/a)r	(t/k)e	res	i	14.2 +++	lo.7 +++	6.9 +
em'	er	ne	m(u/a)r	(t/k)e	res	i	lo.6 +	14.2 +++	17.3 +++
em'	er	ne	m(u/a)r	(t/k)e	res	i	8.1 +++	7.9 +++	lo.2 +++

utterance position

1	2	3	4	5	6	7	8	9	lo
i	13.5 +++	7.4 +++	7.5 +++	13.1 +++	(t/k)e	res	em'	er	ne
i	6.4 +++	12.3 +++	6.5 ++	4.2 +	(t/k)e	res	em'	er	ne
i	5.9 +++	5.8 +++	lo.1 +++	4.4 +	(t/k)e	res	em'	er	ne
em'	er	ne	i	13.2 +++	7.4 +	6.8 ++	9.9 +++	(t/k)e	res
em'	er	ne	i	5.8 +++	12.o +++	7.1 +++	3.1 o	(t/k)e	res
em'	er	ne	i	9.1 +++	lo.o +++	11.6 +++	8.4 +++	(t/k)e	res
em'	er	ne	m(u/a)r	(t/k)e	res	i	6.4 +++	7.9 +++	6.1 +++
em'	er	ne	m(u/a)r	(t/k)e	res	i	8.7 ++	lo.8 +++	8.o +++
em'	er	ne	m(u/a)r	(t/k)e	res	i	5.7 ++	4.2 ++	5.6 ++

with the over-all Fo level of the speaker. Therefore, the absolute differences were converted into relative ones, so the difference between a given u and corresponding ɑ was expressed in semitones relative to the ɑ. Note, however, that subject ER's inherent differences are somewhat larger than those of SI, although the two speakers have approximately same over-all Fo levels. In this connection it is also worth noticing that the range of Fo deflection in the stress group and intonation contour is larger with ER than with SI. This is interesting, since it could be indicative of a tendency towards agreement between suprasegmental and segmental factors with regard to the magnitude of the Fo variation they give rise to.

3.1.1 Inherent Fo differences in relation to position in the utterance

The material was structured in such a manner that the Fo differences between u and ɑ could be examined in different positions in the utterance while position in the stress group was kept constant. This is illustrated in table 9 which also shows the part of the material considered in the present section.

In fig. 8 the inherent Fo differences, expressed in semitones, are plotted as a function of the position in the utterance in stressed, first posttonic, and second posttonic syllables, respectively. (The third, fourth, and fifth posttonic syllables were left out, because they occurred only in few and unevenly distributed positions in the utterance - cf. table 9). The straight lines in fig. 8 are regression lines fitted to the data points of all speakers pooled using the least squares method.

In the stressed vowels there is a statistically significant negative correlation between the inherent differences (expressed in semitones) and the position in the utterance. The slope of the regression line is -0.09 and the correlation coefficient -0.367 ($p < 0.05$, N=36).

In vowels in the first posttonic syllable the opposite tendency seems to apply, i.e. there is a positive correlation between the differences and the position in the utterance. The slope of the regression line is +0.072 and the correlation coefficient 0.406 ($p < 0.02$, N=32).

Table 9

Schematic description of the test material. The columns correspond to the utterance positions, and each of the rows represent two sentences with identical stress patterns (one u- and one ɑ-sentence). The testwords are underlined, and the stress patterns are given by numerals, 0 being the stressed syllable, 1 the first posttonic, 2 the second posttonic, etc. The syllables considered in the present section are enclosed in rectangles.

Word No.	Utterance position									
	1	2	3	4	5	6	7	8	9	10
1 - 10	-	0	1	2	3	0	1	0	1	2
2 - 11	-	-	0	1	2	0	1	0	1	2
3 - 12	-	-	-	0	1	0	1	0	1	2
4 - 13	0	1	2	3	0	1	2	3	0	1
5 - 14	0	1	2	3	4	0	1	2	0	1
6 - 15	0	1	2	3	4	5	0	1	0	1
7 - 16	0	1	2	3	0	1	2	0	1	2
8 - 17	0	1	2	3	0	1	2	3	0	1
9 - 18	0	1	2	3	0	1	2	3	4	0

In vowels in the second posttonic position the inherent Fo differences are also increasing through the utterance, although to a much lesser degree than in the first posttonic position. The slope of the regression line is +0.004 and the correlation could not be shown to be statistically significant, $r = 0.015$ ($p > 0.05$, N=20).

Thus, it seems reasonable to conclude that in stressed position the inherent Fo differences decrease slightly through the utterance, in first posttonic syllables they increase slightly, and in the second posttonic position they are constant (constant in the sense that the variation cannot be accounted for by the position in the utterance).

3.1.2 Inherent Fo differences in relation to position in the stress group

The structure of the material made it possible to examine the inherent Fo differences in different positions in the stress

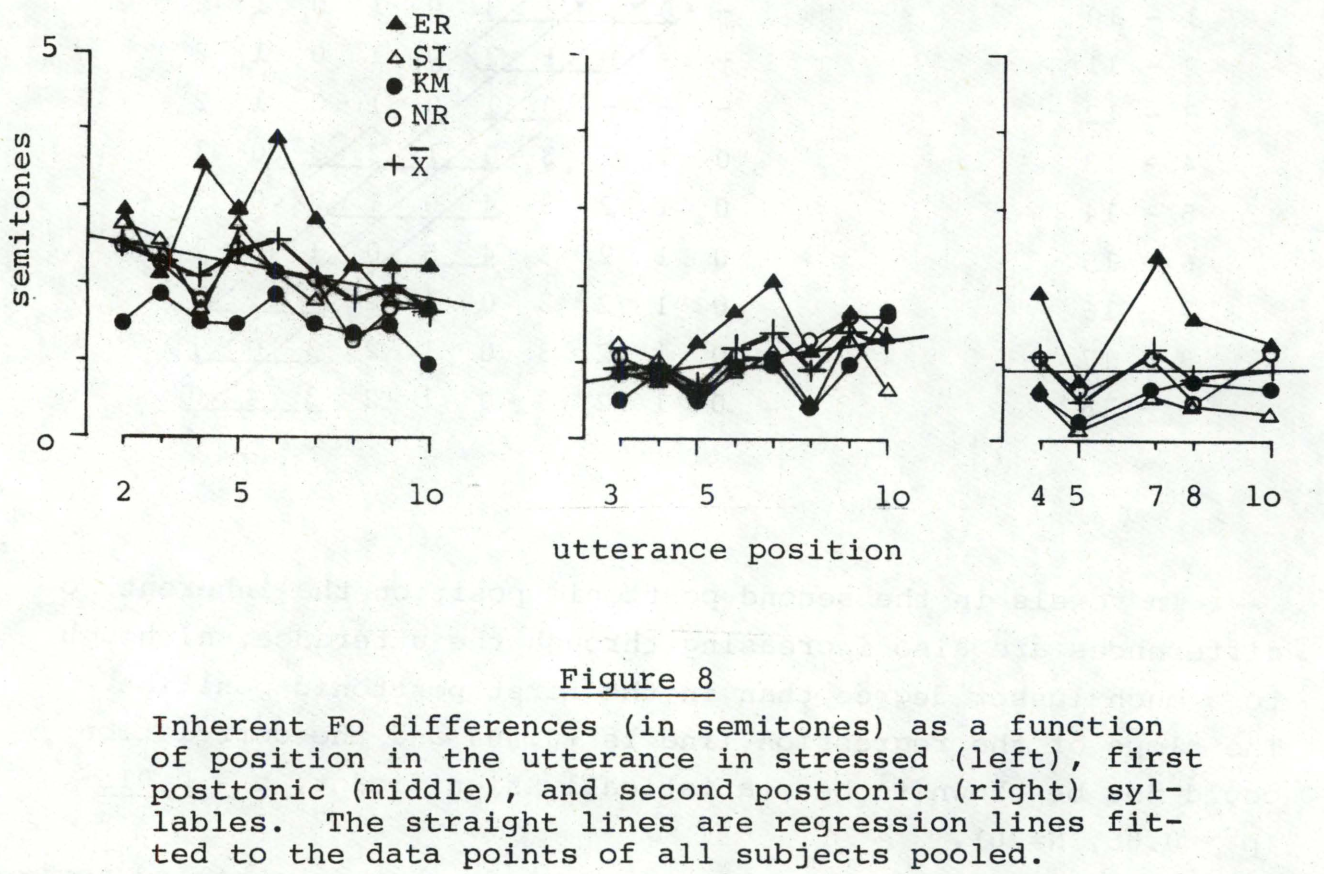


Figure 8
Inherent Fo differences (in semitones) as a function of position in the utterance in stressed (left), first posttonic (middle), and second posttonic (right) syllables. The straight lines are regression lines fitted to the data points of all subjects pooled.

group, keeping the position in the utterance constant. This is shown in table 10 from which it can also be seen that the number of different stress group positions represented in each utterance position differs between utterance positions; in utterance position 4, for example, stress group positions 0, 1, and 2 (i.e. stressed syllable and the first and second posttonic syllables) are represented, while e.g. in position 5 in the utterance the stress group positions 0 through 4 are represented.

Table 10

Schematic description of the test material. The columns correspond to the utterance positions, and each of the rows represents two sentences with identical stress patterns (one u- and one ɑ-sentence). The testwords are underlined, and the stress patterns are given by numerals, 0 being the stressed syllable, 1 the first posttonic, 2 the second posttonic, etc. The syllables considered in the present section are enclosed in rectangles.

Word No.	Utterance position									
	1	2	3	4	5	6	7	8	9	10
1 - 10	-	0	1	2	3	0	1	0	1	2
2 - 11	-	-	0	1	2	0	1	0	1	2
3 - 12	-	-	-	0	1	0	1	0	1	2
4 - 13	0	1	2	3	0	1	2	3	0	1
5 - 14	0	1	2	3	4	0	1	2	0	1
6 - 15	0	1	2	3	4	5	0	1	0	1
7 - 16	0	1	2	3	0	1	2	0	1	2
8 - 17	0	1	2	3	0	1	2	3	0	1
9 - 18	0	1	2	3	0	1	2	3	4	0

In fig. 9 the relative inherent Fo differences are plotted as a function of their position in the stress group in each of the utterance positions examined. Although the variation between subjects is considerable, the general tendency (although not entirely consistent) seems to be that the differences are reduced as a function of the position in the stress group up to and including the second posttonic position. From that point the differences seem to increase, as judged from the utterance positions

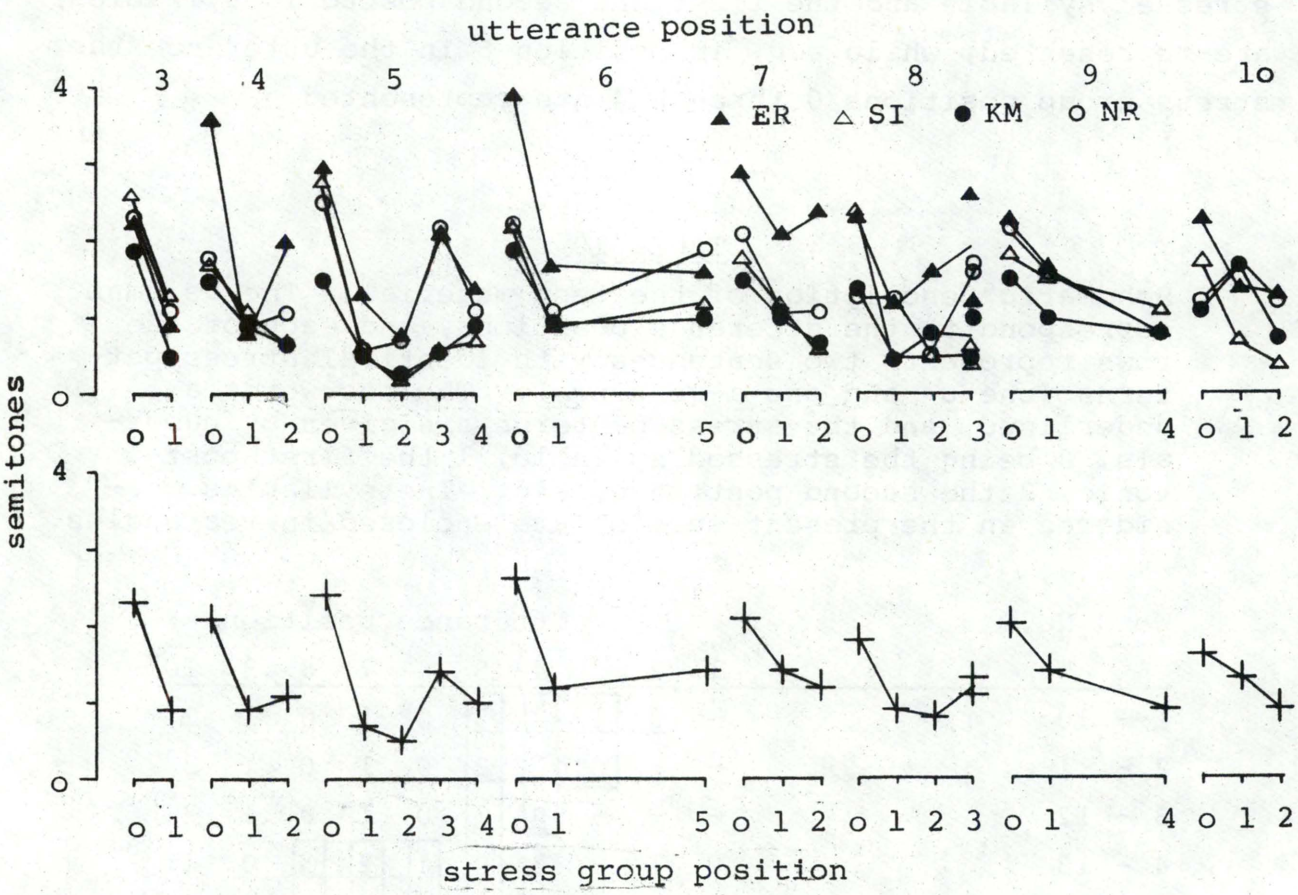


Figure 9

Inherent Fo (in semitones) as a function of the position in the stress group in utterance position 3 through 10. Position o in the stress group indicates the stressed syllable, and 1 through 5 indicate the first through fifth posttonic syllables. The upper graph shows the data for the individual subjects, and the lower shows the averages over all subjects. Note that the third posttonic syllable in utterance position 8 is represented by two data points. This is due to the fact that this syllable occurred twice under equivalent conditions (see table 10). In the graphs the lines connecting the data points end at the mean of the two data points.

in which the stress group positions later than the second posttonic position are represented (viz. positions 5, 6, 8, and 9 in the utterance).

It must be kept in mind, however, that the effect of the stress group position is to some extent blurred by the influence from position in the utterance. Fig. 9 shows that the difference in inherent Fo between the stressed syllable and the first posttonic syllable is reduced through the utterance, and that the relation between those syllables and the others is also changed through the utterance (cf. section 3.1.1 above).

In an attempt to normalize for the utterance position effect the mean inherent differences (fig. 9, lower graph) in stress group positions 0 and 1 were converted to deviations from the regression lines describing the utterance position effect (cf. section 3.1.1 and fig. 8). The data of stress group positions 2, 3, 4, and 5 were not normalized, since the utterance position effect was assumed to be negligible in stress group position 2 and, supposedly, this was also the case in positions 3, 4, and 5. In order to obtain the correct relations between the stressed syllable and the first posttonic one, and between those syllables and the remaining ones, the difference in the middle of the utterance (i.e. in the imaginary position 5.5) computed from the regression lines was added to the deviation found in the stressed and first posttonic syllables. The normalized inherent differences in each stress group position were then averaged over all utterance positions. The normalized and averaged differences are plotted in fig. 10. The tendency towards decreasing/increasing inherent Fo differences through the stress group is also apparent in the normalized data, but it must be pointed out that the increase is represented by the data points corresponding to the third and the fifth posttonic syllables, which have been determined under the assumption of constant inherent differences through the utterance (cf. section 3.1.1 above). Furthermore, the fifth posttonic syllable occurs only once in the material, namely in position 6 in the utterance. In view of this, it may perhaps be too rash to maintain that the inherent Fo differences show a decreasing/increasing tendency through the stress group. On the contrary, it may not be unreasonable to suggest that in the unstressed syllables the magnitude of the inherent Fo differences is constant through the stress group.

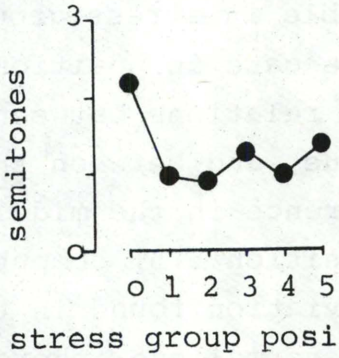


Figure 10
Normalized and averaged inherent differences (in semitones) plotted as a function of position in the stress group.

4. Discussion

As regards the fundamental frequency contributions from the sentence- and stress group components, the results obtained in the present experiment are in good agreement with those of previous investigations of the intonation of Advanced Standard Copenhagen Danish (Thorsen, 1978).

The specific purpose of the experiment reported above was to examine the effect upon the inherent Fo differences in vowels of suprasegmental factors such as stress conditions (stress vs. nonstress), and position of the vowels in the utterance and the stress group in declarative sentences. The result which emerges most clearly is that the inherent Fo differences between u and ɑ are greater in stressed than in unstressed vowels, but also that under both conditions of stress the differences are statistically significant regardless of the position in utterance or stress group. The differences found in stressed vowels in the present material correspond reasonably well with those found in Reinholt Petersen (1978) and, further, the tendency reported in that paper towards larger inherent differences in stressed than in unstressed vowels is also in agreement with the present findings.

With respect to the influence from position in utterance and stress group on the magnitude of the inherent Fo differences, the results seem less clear. Obviously, the two factors cannot be regarded as mutually independent, a fact which makes the description of the inherent Fo differences rather complex. What particularly contributes to the complexity is the behaviour of the first posttonic syllable, in which the differences tend to increase through the utterance. If that syllable had shown a behaviour similar to that of the remaining unstressed syllables, the following description might have been suggested: In stressed syllables the inherent differences between u and ɑ decrease through the utterance (from about 2.5 to about 1.75 semitones), and in unstressed syllables the differences between the two vowels remain constant at approximately 1 semitone in all utterance- and stress group positions.

One explanation for the deviant behaviour of the first posttonic syllable may be related to the fact - which is characteristic for that syllable and distinguishes it from all others - that in a stress group it constitutes the peak Fo and the termination of the very steep Fo rise starting at the Fo minimum located in or immediately before the preceding stressed syllable. In fig. 11 the extent of the Fo rise in a and u words is plotted as a function of the position in the utterance of the stressed syllable. It is seen that the rise is greater in a than in u; and along with the general reduction of the rise as a function of the position in the utterance, the difference between a and u with respect to the extent of the rise is also decreasing through the utterance, so that the extent of the rise is almost the same in the two vowels at the end of the utterance.

Thus, the increasing inherent Fo differences observed in the first posttonic syllable might be explained in the following way: At the beginning of the utterance, where Fo is high and the extent of the rise from the stressed syllable to the first posttonic syllable is large, it is more difficult for u than for a to reach the intended Fo of the first posttonic syllable. At the end of the utterance, where Fo is lower and the extent of the rise smaller, the two vowels differ less in their ability to reach the intended level in the first posttonic syllable.

Whether this explanation is physiologically acceptable will not be discussed in detail here, but if the Fo differences between high and low vowels can be accounted for by a higher degree of tension in the laryngeal tissues in high vowels than in low vowels as a consequence of the tongue pull in high vowels (whether translated into a horizontal tension in the vocal cords as suggested by Ladefoged (1964) and Lehiste (1970), or into a vertical tension as suggested by Ohala (1978)), then it might not be unreasonable to assume that the higher degree of tension in u can reduce the effect of the forces responsible for the Fo rise from stressed to first posttonic syllable, and further, that this impeding effect is more pronounced at the beginning of the utterance than at the end of it.

It must be emphasized that the results presented in the present paper are based on declarative sentences, in which the sentence component contributes to a falling fundamental frequency through the utterance. It will be of interest to see whether the tendencies towards decreasing inherent differences in stressed syllable and increasing differences in first posttonic syl-

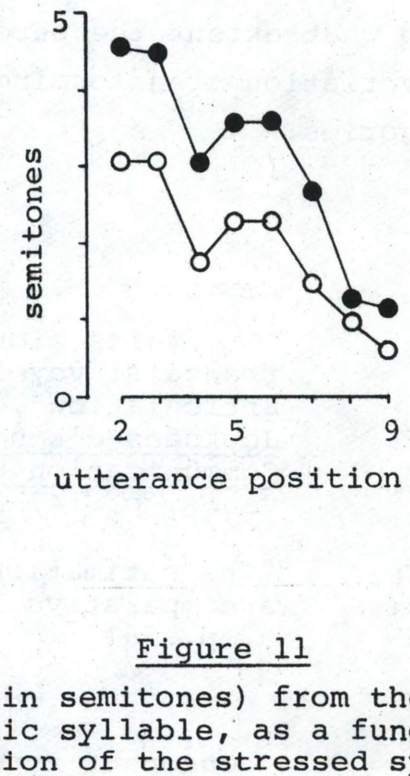


Figure 11

The Fo rise (in semitones) from the stressed to the first posttonic syllable, as a function of the utterance position of the stressed syllable.

lable through the utterance will also be apparent in interrogative sentences (or more correctly: syntactically unmarked questions), which are characterized by a "flat" sentence component at a relatively high Fo level, i.e. whether they can be ascribed to a "position in the utterance" effect (i.e. early/late), or a "position on the intonation contour" effect (i.e. high/low). Data obtained from such sentences might also contribute to the discussion of the physiological explanation attempted above.

For future work on the topic dealt with in the present paper two lines of research suggest themselves: one concentrating upon the physiological causes of the inherent Fo variation, taking into account the findings of the present experiment; the other turning upon the perceptual role of the inherent Fo variation, or, more specifically, to what extent the perceptual process employs knowledge of that variation transforming the acoustic signal into linguistic categories.

References

Di Cristo, A. and M. Chafcouloff 1977: "Les faits microprosodiques du français: voyelles, consonnes, coarticulation", Actes des 8èmes Journées d'Etudes du Groupe de la Communication Parlée, Aix-en-Provence, p. 147-158

Hirst, D.J., A. Di Cristo and Y. Nishinuma 1979: "The estimation of intrinsic Fo: a comparative study", Proc.Phon. 9, p. 381

Hombert, J.M. 1976: "Consonant types, vowel height and tone in Yoruba", UCLA WP 33, p. 40-54

Jeel, V. 1975: "An investigation of the fundamental frequency of vowels after various consonants, in particular stop consonants", ARIPUC 9, p. 191-211

Johansson, I. 1976: Inherenta grundtonsfrekvenser hos svenska vokaler och deras inflytande på satsintonationen, Stadsmål i övre Norrland 9, (Umeå University)

Ladefoged, P. 1964: A phonetic study of West African languages, an auditory-instrumental survey, Cambridge

Lehiste, I. 1970: *Suprasegmentals*, (MIT Press)

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

Löfqvist, A. 1975: "Intrinsic and extrinsic Fo variation in Swedish tonal accents", *Phonetica* 31, p. 228-247

Neweklowsky, G. 1975: "Spezifische Dauer und spezifische Tonhöhe der Vokale", *Phonetica* 32, p. 38-60

Ohala, J.J. 1978: "The production of tone", *Report of the Phonology Laboratory* 2 (Berkeley), p. 63-117

Peterson, G.E. and H.L. Barney 1952: "Control methods used in a study of the vowels", *JASA* 24, p. 118-127

Reinholt Petersen, N. 1978 "Intrinsic fundamental frequency of Danish vowels", *J.Ph.* 6, p. 177-189

Thorsen, N. 1978: "An acoustical investigation of Danish intonation", *JPh* 6, p. 151-176

Thorsen, N. 1979a: "Interpreting raw fundamental tracings of Danish", *Phonetica* 36, p. 57-78

Thorsen, N. 1979b: "Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish", *Proc.Phon.* 9, p. 417-423

APPENDIX

AN INVESTIGATION OF THE EFFECT OF CERTAIN CONSONANTS ON THE FUNDAMENTAL FREQUENCY IN PRECEDING UNSTRESSED VOWELS

1. Method

The vowels u and ɑ were combined with each of the consonants t, k, and m in words and carrier sentences similar to those employed in the main experiment. The following sentences:

Em'erne i mu'mumu muteres
Em'erne i mu'mumu mukeres
Em'erne i mu'mumu mumeres

Em'erne i mar'marmar marteres
Em'erne i mar'marmar markeres
Em'erne i mar'marmar marmeres

were arranged together with 2 distractor sentences in 10 different random orders in a list, which was read by one subject (NR, the author) under conditions identical to those of the main experiment. The following acoustic registrations were made: duplex oscillogram, two intensity curves, and Fo curve. The fundamental frequency was measured at the middle of the unstressed vowels in the test syllables (underlined in the list above).

2. Results

The means of the Fo measures were:

	t	k	m
u	116.3	116.6	115.7
ɑ	112.1	111.3	111.8

The data were submitted to a two-way analysis of variance (2 vowels x 3 consonants) which showed a significant effect of vowel quality ($F = 37.646$, $p < 0.01$), no effect of the following consonant ($F = 0.126$, $p > 0.05$), and no significant interaction ($F = 0.346$, $p > 0.05$).

In view of the structure of the material used in the main experiment, the point of interest was whether there was any interaction between the vowels and the two consonants k and t. Therefore the data were submitted to another two-way analysis of variance, in which only t and k were included. Thus, any interaction could be ascribed to them alone. The results of this analysis did not deviate essentially from the first one, and, specifically, the interaction could not be shown to be statistically significant ($F = 0.442$, $p > 0.05$).

In view of the structure of the material used in the main experiments the point of interest was whether there was any interaction between the vowels and the *two* consonants *k* and *g*. Therefore the data were submitted to another two-way analysis of variance, in which only *g* and *k* were included. Thus, any interaction could be ascribed to them alone. The results of this analysis did not deviate essentially from the first one, and, specifically, the interaction could not be shown to be statistically significant ($F = 0.2441$, $p > 0.05$).

(ARIPUC 13, 1979)

LEXICAL STRESS, EMPHASIS FOR CONTRAST, AND SENTENCE INTONATION IN ADVANCED STANDARD COPENHAGEN DANISH

Nina Thorsen

Preface

A paper with the above title was published in the Proceedings of the Ninth International Congress of Phonetic Sciences, 1979, vol. II, p. 417-423, as a contribution to the symposium on 'The relation between sentence prosody and word prosody'. That paper contained no reference to the literature on intonation in other languages, and no documentation was given for the figures and statements made. - I had intended to publish, in this volume of ARIPUC, a complete account of the material analyzed as well as the results of analyses of questions with emphasis for contrast, conducted after the congress proceedings went to press. However, lack of time prevents this, and a comprehensive version will have to be put off till the next volume of ARIPUC.

On the following pages, the congress paper is reprinted, with a few, very minor, adjustments and with the addition of some footnotes. In an addendum, a couple of points from this paper will be commented upon, and the results of the supplementary analysis of questions with emphasis for contrast are presented.

LEXICAL STRESS, EMPHASIS FOR CONTRAST, AND SENTENCE INTONATION IN ADVANCED STANDARD DANISH

Nina Thorsen

(Reprint from Proceedings of the Ninth International Congress of Phonetic Sciences, 1979, vol. II, p. 417-423.)

Due to lack of space, no references will be made to the very considerable literature on intonation in other languages, nor will any extensive documentation be given.

1. A model of Danish intonation

Intonation in short sentences in Advanced Standard Copenhagen (ASC) Danish may be presented as in fig. 1, which is but a model, with the advantages and shortcomings that modeling almost always entails in terms of simplicity and inaccuracy, respectively. It is based on recordings by six subjects, three males and three females, of a rather elaborate material (cf. Thorsen 1978a and section 5.1 below). The qualitative statements which can be read off the figure are perfectly representative of all subjects, but the quantifications involved are, of course, averages, and no one subject behaves as mathematically neatly as the model would have you believe.

A basic assumption underlying fig. 1 is that the complex course of fundamental frequency (Fo) in an utterance is the outcome of a superposition of several components: (1) A sentence component which supplies the INTONATION CONTOUR (broken lines). (2) On the contour is superposed a stress group component which furnishes the STRESS GROUP PATTERNS (full lines). (3) To the resultant of those two components is added, in words containing stød, a stød component, rendering STØD MOVEMENTS. However, as stød words had been excluded from the material, the model does not include this particular feature. These first three components are language specific and thus "speaker controlled". (4) Finally, intrinsic Fo level differences between segments, and coarticulatory variations at segment boundaries supply a MICROPROSODIC COMPONENT, which, at least in non-tonal languages, is not consciously controlled by the speaker, but due to inherent properties of the speech production apparatus and which, therefore, is superfluous

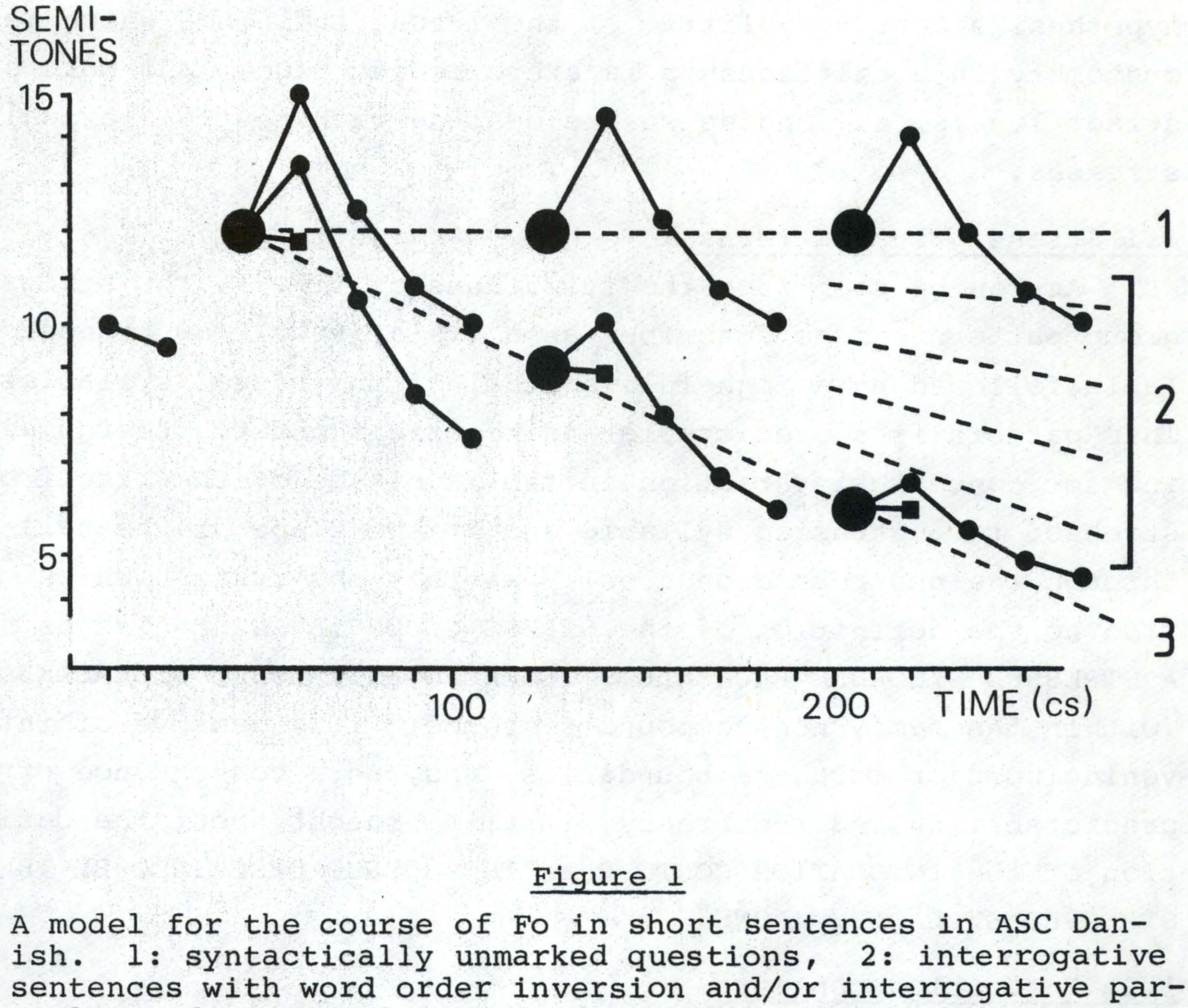


Figure 1

A model for the course of Fo in short sentences in ASC Danish. 1: syntactically unmarked questions, 2: interrogative sentences with word order inversion and/or interrogative particle, and non-final periods (variable), 3: declarative sentences. The large dots represent stressed syllables, the small dots unstressed ones, and the small squares represent an unstressed syllable being the only one between two stressed ones (see further the text). The full lines represent the Fo pattern associated with stress groups, and the broken lines denote the intonation contours.

in the model from the point of view of the human speaker.[1] - This concept of "layers" in intonation is anything but original; the triviality of the statement does not, however, deprive it of its validity or its relevance: firstly, it is tremendously useful in the interpretation of Fo tracings (Thorsen, 1979) and, secondly, it has a very direct bearing on the theme of this symposium, 'The relation between sentence prosody and word prosody':

The relation between word stress and sentence prosody (i.e. sentence intonation: duration and intensity are not considered) is physically a very close-knit and intricate one, but we may hypothesize that very little of the mutual influence which is customary in a relationship is exercised at higher and more abstract levels, as long as we are dealing with neutral lexical stresses.

1.1 Stress group patterns

As can be seen from the full lines of fig. 1, the stress group pattern can be described as a (relatively) low stressed syllable followed by a high-falling tail of unstressed syllables. This pattern is a predictable and recurrent entity, though allowing for contextual variation in the magnitude of the rise from stressed to unstressed syllable and in the slope of the fall through the unstressed ones. It was this observation which gave rise to the definition of the STRESS GROUP in ASC Danish as A STRESSED SYLLABLE PLUS ALL SUCCEEDING UNSTRESSED SYLLABLES (within the same, non-compound sentence), irrespective of intervening word or morpheme boundaries, and, as a consequence of this predictability and recurrency, it also brought about the definition of the INTONATION CONTOUR as THE COURSE DESCRIBED BY THE STRESSED SYLLABLES ALONE.

1.2 Intonation contours

The intonation contours tend to vary systematically with sentence type, declarative sentences having the most steeply falling contours, at one extreme, and syntactically unmarked questions having "flat" contours, at the other extreme. In between these

---

1) It seems as though in tone-languages those perturbations are actively brought within time limits where they will not interfere with the perception of the tonal distinctions in the language, cf. Hombert (1976).

two are found other types of questions as well as non-final periods with slopes that tend to vary in a trade-off relationship with syntax. For a further account of these contours and their perception, see Thorsen (1978a, 1978b).

2. Implications of the model

2.1 Fo movements in syllables

The model does not specify the Fo movements of syllables: the tonal composition of the stress group pattern as one of LOW plus HIGH FALLING allows for a very simple account of Fo movements in vowels and consonants: segments do not carry specific movements (except when stød is involved) but simply float on the Fo pattern, and slight variations in Fo movement would be due then to the fact that segments do not always hit the patterns at exactly the same place.[1]

2.2 The course of the intonation contour

(a) When the number of stress groups changes, everything else being equal, so does the slope of a given contour, leaving only the flat ones intact; the constancy presumably lies in the interval between the first and the last stressed syllable, with intervening stressed syllables evenly distributed between them, and not in a certain rate of change.[2]

(b) When the number of unstressed syllables varies in the stress groups, the stressed syllables will not be equidistantly spaced in time, and the straight lines of fig. 1 break up into a succession of shorter ones with unequal slopes.

Combining the effects of changes of both types leaves us with an infinity of physically different intonation contour configurations. On a higher level in production these variations may not exist, and perceptually they may be obliterated, turning the contours into smoothly slanting slopes, (1) if what we aim at producing and what we perceive are equal intervals between stressed syllables and not the actual slope of the contour, and (2) if we assume that isochrony, be it not a physical reality, is a psychological reality with the speaker/listener.[3]

---

1) I am not happy about this formulation, see further the addendum.

2) This assumption turned out to be wrong, at least for longer utterances, cf. Thorsen, this volume.

3) I am inclined to think, now, that assumption (1), as it is formulated here, does not capture the essential fact about perception of intonation contours, see further the addendum.

2.3 Fo patterns of stress groups

2.3.1 Stress groups with more than one unstressed syllable

(a) In statements, the rise from stressed to unstressed syllable is, on the average 1½, 1, and ½ semitone, respectively, in the first, second, and third stress group. In syntactically unmarked questions, the rises amount to 3, 2½, and 2 semitones, respectively. The difference in magnitude of this rise, between patterns riding on different contours is very likely a direct consequence of differences in the level of the following stressed syllable.

(b) The decrease with time in the rise from stressed to unstressed syllable is the same in syntactically unmarked questions and statements, one semitone. This decrease, which is independent of the particular contour, may be seen as a consequence of either of two distinct processes, or of a combination of them. It may be a "voluntary" decrease, i.e. a signal of finality, and/or it may be a physiological phenomenon: the closer you get to the end of the utterance, the less energy is expended and the less complete the gestures will be; either or both phenomena may also account for the less and less steep falls through the unstressed syllables.

If the variation in the Fo patterns with intonation contour and time is physiologically determined, the speaker may be unconscious of it, and the listener may neglect or compensate for it.

2.3.2 Stress groups with only one unstressed syllable

Stress groups with one unstressed syllable will of course be shorter than those with several, a feature which is not reflected in fig. 1. - A single unstressed syllable does not accomplish a full rise-fall when the following stressed syllable is considerably lower than the preceding one, as is the case in statements. Instead it lands on very nearly the same level or slightly below the preceding stressed syllable and, accordingly, the rise-fall is amputated. A full rise-fall may be intended by the speaker and the amputation be due to a shortcoming in the peripheral speech production mechanism. Accordingly, the listener may well re-introduce a rise-fall (this is, indeed, my own subjective impression). But we have here an indication that time (rhythm) overrides Fo when the two are in conflict. On the other hand, there is definitely a tendency for as complete rise-falls as possible. Two unstressed syllables will traverse more than half the fall exhibited by four, everything else being equal. - These two facts together are yet another reminder that speech is not a card-board structure but a smooth and dynamic process.

2.4 Conclusion

If the assumptions made about production and perception of Fo courses hold water, we are left with two components which physically are highly interactive but on more abstract levels may be invariant, apart from the fact that contours change with sentence type.[1]

3. Emphasis for contrast

By emphasis for contrast is meant the extra prominence on one of the syllables in the utterance, used to denote a contrast which may be implicit or may be explicitly stated in the context. I have deliberately avoided terms like 'focus', 'sentence accent', or 'nucleus' because these terms are used, in a number of languages, to describe a phenomenon different from emphasis for contrast: one of the lexically stressed syllables in the utterance will always have slightly greater prominence (realized, very roughly speaking, as a more elaborate Fo movement within or in the environment of that syllable), and if nothing else is specified by the context, it will fall on the last stressed syllable. - A similar phenomenon does not exist in ASC Danish as a thing apart from emphasis for contrast. Whenever and wherever such a slightly heavier stress is introduced, it invariably invokes the impression of contrast. Insofar as we are not faced with incomplete evidence or with a false dichotomy, i.e. one due to differences in concepts, Danish seems to be markedly different from e.g. English, German, and Swedish.

3.1 Contrastive stress and Fo

The following account is based on a material of sentences, uttered in dialogues, where the contrasts were all explicitly stated in the context (but I strongly believe that they would have looked no different had they been implicit). - When emphasis for contrast occurs, it affects the intonation contour as well as Fo patterns.

---

1) Thus, the sentence and stress group components of the model in fig. 1 may be simply additive from a productional point of view. But for synthesis purposes, the modifications in Fo patterns with contour and time must of course be specified.

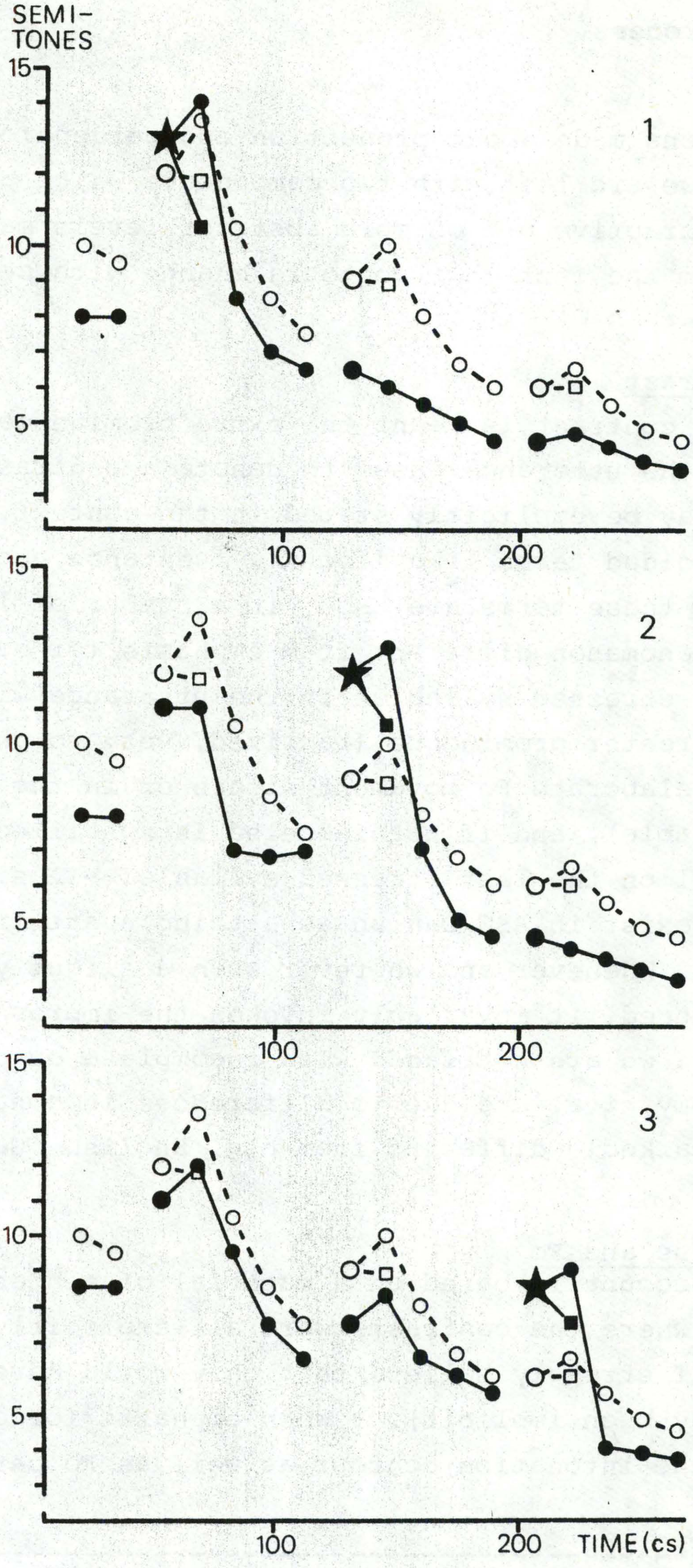


Figure 2

Models for statements with emphasis for contrast on
(1) the first,
(2) the second,
(3) the last stressed syllable,
compared to the neutral edition (broken lines and empty dots).

Fig. 2 compares the neutral edition with three statements where the emphasis lay on the first, second, and third lexically stressed syllable. (Durational differences between the neutral and emphatic editions are very slight and there is no doubt that Fo is the prime cue to contrast, as it is to neutral lexical stress.) The obvious changes introduced in the Fo course by emphasis for contrast is a raising of the syllable in question (represented by a star), a drastic fall from first to second unstressed syllable, plus a not inconsiderable shrinking of the surrounding Fo patterns:

(1) When emphasis is on the first stressed syllable, it is higher, the rise to the first unstressed syllable is smaller, and the fall through the following unstressed syllables is steeper than for the neutral case. The levels of the second and third stressed syllables are considerably lowered and the LOW+HIGH FALLING pattern is annihilated in the second and shrunk in the third stress group. The syllables of the second and third stress groups look, tonally, more like a series of unstressed syllables continuing the fall in the first stress group.

(2) Emphasis on the second stressed syllable repeats the pattern of (1), and we also get a certain reduction of the first stress group with a very steep fall from first to second unstressed syllable.

(3) The pattern repeats itself in the last stress group, with a shrinking of the preceding ones as well.

Again we note that a single unstressed syllable does not accomplish a full rise-fall but instead drops well below the preceding emphatic one.

The feature common to the three cases seems to be that the syllable on which the emphasis for contrast occurs must stand out clearly from the surroundings, which is brought about by a raising of that syllable as well as by a lowering of the immediate surroundings, except for the first of several post-tonic syllables. The change is slightly greater in the succeeding than in the preceding Fo course. During some informal experiments performed with the ILS-system for analysis and synthesis at the Institute of Linguistics, Uppsala University, it appeared that shrinking the Fo course in the surroundings is sufficient to create the impression of emphasis for contrast. To get emphasis on the word 'sidste' in the statement 'Det er sidste bus til Tiflis.' (It's the last bus for Tiflis.) it is sufficient to change the rise from 'bus' to 'til' to a level or a slight fall, whereas just raising the stressed syllable of 'sidste' will not do the trick. Likewise, to get emphasis on 'bus', lowering the unstressed syllable of 'sidste' will do and just raising 'bus' does not accomplish anything.

The three Fo courses in fig. 2 look widely different and only vaguely resemble fig. 1 (3), although the utterances still sound declarative. What constitutes the intonation contour in utterances with emphasis for contrast, I hesitate to say at present. They may resemble one-word utterances in that the difference between statement and question lies in the level of <u>and</u> movement within the emphatically stressed syllable as well as in the course of the succeeding unstressed ones (Thorsen, 1978a), or the intonation contour may be extrapolated from, and thus still be definable in terms of, the lexically stressed syllables surrounding the emphatic one. The first solution would be interesting, because it implies that in utterances with emphasis, word prosody takes precedence over sentence prosody, whereas the second solution would make the definition of intonation contour apply to a wider range of utterances.

(The references are assembled after the addendum.)

ADDENDUM

4. Comments and revisions

4.1 Re section 2.1 above 'Fo movements in syllables'

The formulation that "segments do not carry specific Fo movements ... but simply float on the Fo pattern" might give the impression that the variation in intravocalic Fo movement is greater than is actually the case. In fact, there are rather narrow limits to the amount of free floating involved and the following account is a more adequate presentation of intrasyllabic Fo movements. A description of vowel movements is sufficient, since the Fo course in voiced consonants seems to be simply a smooth interpolation between the vowels, modified by specific intrinsic Fo properties of the consonant, as the case may be: The course of the unstressed vowels is directly determined by the stress group pattern, being generally falling. The first post-tonic, however, may be rising-falling or even purely rising, depending on the exact timing of the syllable with respect to the Fo maximum. - Short stressed vowels are also generally falling. Long stressed vowels are falling-rising, but the fall is normally as great as or greater than the rise (Reinholt Petersen, 1978). We might ask: why are stressed vowels generally and mainly falling, even though the Fo pattern is on the way <u>up</u> for the post-tonic syllable? and we might seek the answer in an influence from the intonation contour, since we also observe a tendency for a positive correlation between stressed vowel movement and intonation contour slope (the less steep the intonation contour, the less falling the vowel, although the variation in vowel movement is slight). However, I think that the explanation for the mainly falling stressed vowels is to be found in the way stressed syllables are signalled in ASC Danish, being ones that are jumped <u>up</u> from, and a fall will maximize the perceptual distance to the succeeding higher post-tonic. The less steep falls in stressed vowels on less steep intonation contours may be accounted for in terms of the greater rise to the post-tonic, and thus only indirectly to the intonation contour proper.

4.2 Re section 2.2 above 'The course of the intonation contour'

In view of the results of the analysis of intonation contours in utterances of varying length (cf. Thorsen,this volume), I do not think it reasonable to assume that intonation contours must necessarily be perceived as smoothly slanting slopes, i.e. I do not think that the boundaries between prosodic phrase groups, when such grouping occurs, go unnoticed. - Instead, I think that the invariance of intonation contours associated with the same type of sentence or with the same linguistic function and which may be physically different will lie in the degree of over-all downdrift, the perception of which is not crucially dependent on the exact location of the stressed syllables on the frequency and time scales. (See also Pierrehumbert, 1979.)

5. Emphasis for contrast in declarative sentences and syntactically unmarked questions

5.1 Material, subjects, and recordings

The same type of material as was used for the original analysis (Thorsen, 1978a) served here, i.e. nonsense words embedded in initial, medial, and final position in short sentences which are as much alike semantically, syntactically, and rhythmically as possible (for a discussion of the choice of words, see Thorsen 1978a p. 153). The sentences are listed below (they translate as follows: '_______ has shorter syllables', 'The syllables of _______ are shortened', and 'There are shorter syllables in _______', respectively):

1. 'pipi giver kortere stavelser.
2. 'pipipi giver kortere stavelser.
3. pi'pipi giver kortere stavelser.
4. pipi'pi giver kortere stavelser.

5. <u>'pipipi</u> giver kortere stavelser.
6. <u>pi'pipi</u> giver kortere stavelser.
7. <u>pipi'pi</u> giver kortere stavelser.

8. 'pipi giver kortere stavelser?
9. 'pipipi giver kortere stavelser?
10. pi'pipi giver kortere stavelser?

11. Stavelserne i 'pipi forkortes.
12. Stavelserne i pi'pi forkortes.
13. Stavelserne i 'pipipi forkortes.
14. Stavelserne i pi'pipi forkortes.
15. Stavelserne i pipi'pi forkortes.

16. Stavelserne i 'pipipi forkortes.
17. Stavelserne i pi'pipi forkortes.
18. Stavelserne i pipi'pi forkortes.

19. Stavelserne i 'pipi forkortes?
20. Stavelserne i pi'pi forkortes?
21. Stavelserne i pipi'pi forkortes?

22. Det giver kortere stavelser med 'pipi.
23. Det giver kortere stavelser med 'pipipi.
24. Det giver kortere stavelser med pi'pipi.
25. Det giver kortere stavelser med pipi'pi.

26. Det giver kortere stavelser med 'pipipi.
27. Det giver kortere stavelser med pi'pipi.
28. Det giver kortere stavelser med pipi'pi.

29. Det giver kortere stavelser med 'pipi?
30. Det giver kortere stavelser med 'pipipi?
31. Det giver kortere stavelser med pi'pipi?

The questions, together with the neutral statements, were intended for a control and quantification of one aspect of the model as presented in Thorsen (1978a p. 174): that the magnitude of the Fo rise from stressed to post-tonic decreases with time, but is independent of the particular contour (falling or flat) that the stress group rides upon. But they also serve as a frame of reference for the questions with emphasis for contrast, that were recorded at a later stage, cf. below.

In order to make sure that the utterances with nonsense words do not behave aberrantly, the following sentences were recorded as well ('There are many buses out of Tiflis'):

32. Der går mange busser fra Tiflis.
33. Der går mange busser fra Tiflis.
34. Der går mange busser fra Tiflis.
35. Der går mange busser fra Tiflis.

These 35 utterances, embedded in small dialogues, were mixed with a rather large material that served a completely different purpose (five full pages of reading material in all) and were recorded six times by three phoneticians (two males and one female) and ten times by the author. One of the males speaks a slightly conservative variant of Standard Copenhagen Danish, the other subjects speak Advanced Standard Copenhagen Danish.

This material is the foundation for the paper reprinted above.

Later on, three of the four subjects (excluding the ASC-speaking male) recorded the following material:

36. 'pipipi giver kortere stavelser.
37. pi'pipi giver kortere stavelser.
38. pi'pipi giver kortere stavelser.

39. Stavelserne i pi'pipi forkortes.
40. Stavelserne i pi'pipi forkortes.

41. Det giver kortere stavelser med pi'pipi.
42. Det giver kortere stavelser med pi'pipi.

43. pi'pipi giver kortere stavelser?
44. pi'pipi giver kortere stavelser?
45. pi'pipi giver kortere stavelser?

46. Stavelserne i pi'pipi forkortes?
47. Stavelserne i pi'pipi forkortes?
48. Stavelserne i pi'pipi forkortes?

49. Det giver kortere stavelser med pi'pipi?
50. Det giver kortere stavelser med pi'pipi?
51. Det giver kortere stavelser med pi'pipi?
52. Det giver kortere stavelser med 'pipipi?

The statements and questions where emphasis occurs in the surroundings of the nonsense words would yield a sounder basis for the quantification of the shrinking of Fo patterns that surround an emphasis for contrast (cf. fig. 2 above).

These 17 sentences were embedded in dialogues and again mixed with a material that served a different purpose, which gave two pages of reading material, that occurred in three different randomizations, each being read twice by the three subjects, i.e. six readings of each sentence were obtained.

The recordings were made with semi-professional equipment in a quasi-damped room and were processed by hard-ware intensity and pitch meters (F-J Electronics), registered on a Mingograph (Elema 800) and measured by hand. - In unidirectional Fo courses with constant slope, only the beginning and end points were measured. In more complex Fo courses, three to six points were measured, in a manner so that the traces could be accurately reconstructed by smooth interpolation through the measuring points. The duration of each segment was likewise measured, to the extent that segmentation could be reliably performed. Fo and time measurements were averaged over the six (ten) recordings of each sentence by each subject, and average tracings drawn. The Fo averages were later converted into semitones (re 100 Hz) and these values were turned into account for figs. 1 and 2 above.

5.2 Results

The bulk of the material has been accounted for in sections 2.3(1,2) and 3 above. What remains is the analysis of questions with emphasis for contrast. This analysis has been completed for only one subject (BH - female), in the sense that the appropriate stylized figures have so far only been turned out for her, but inspection of the "raw" average tracings of the other subjects leave no doubt that the conclusions that may be drawn on the basis of BH can be extended to cover the other speakers as well, and the (slight) deviations among subjects will only affect the quantification involved.

Figs. 3-11 are stylized tracings that compare (a) neutral statements and statements with emphasis for contrast on the first through third stress group (figs. 3-5), (b) neutral questions and questions with emphasis for contrast on the first through third stress group (figs. 6-8), (c) questions and statements with emphasis for contrast on the first through third stress group (figs. 9-11). These figures differ from fig. 2 in that they are the outcome of a combination of information from several average tracings: each Fo pattern is based on the average of the nonsense words in the appropriate condition. E.g. the neutral statement in figs. 3-5 is pieced together of the over-all averages of the nonsense words in sentences 1-4 (initial position), 11-15 (medial position), and 22-25 (final position), and the statement with emphasis in the first stress group is composed of the nonsense words in sentences 5-7, 39, and 41. The averages over several words are calculated only after the words have been lined up according to the stressed syllables, and they include the following word/syllable ([˳g̊iʌ̯] and [˳fʌ]) when necessary to obtain two post-tonics.

On a couple of points BH deviates slightly from the average trends that can be read off the model (fig. 1) and fig. 2: In the neutral question (cf. figs. 6-8), the stressed syllable is slightly higher and the rise to the post-tonic slightly smaller than expected in the second stress group. Therefore, the intonation contour is not completely flat and the progressive decrease in the rise from stressed to post-tonic from first through last stress group is not apparent. (This might be a side-effect of the way the figures have been constructed, being pieced together of average tracings of the nonsense words in initial, medial, and final position in different sentences - although on the whole this procedure seems to be quite satisfactory.)

It is clear from figs. 6-8 that the changes in Fo introduced by emphasis for contrast in questions is analogous to that found in statements: the emphatic syllable gets its extra prominence by an annihilation of the HIGH+LOW FALLING pattern in the surrounding stress groups and by a lowering of preceding ones, if any, as well, cf. figs. 7-8. The raising of the emphatic syllable itself, that we observe in statements (cf. figs. 3-5) is hardly noticeable.

The most obvious difference between questions and statements with emphasis for contrast (figs. 9-11) is in the course of the syllables succeeding the emphatic one: they run higher in questions than in statements. In fig. 9 and fig. 11 there is also a difference in the level of the emphatic syllable, but the order and magnitude of this difference is subject to a high degree of inter-speaker variation (apparent from inspection of the average tracings of the other two subjects who recorded questions with

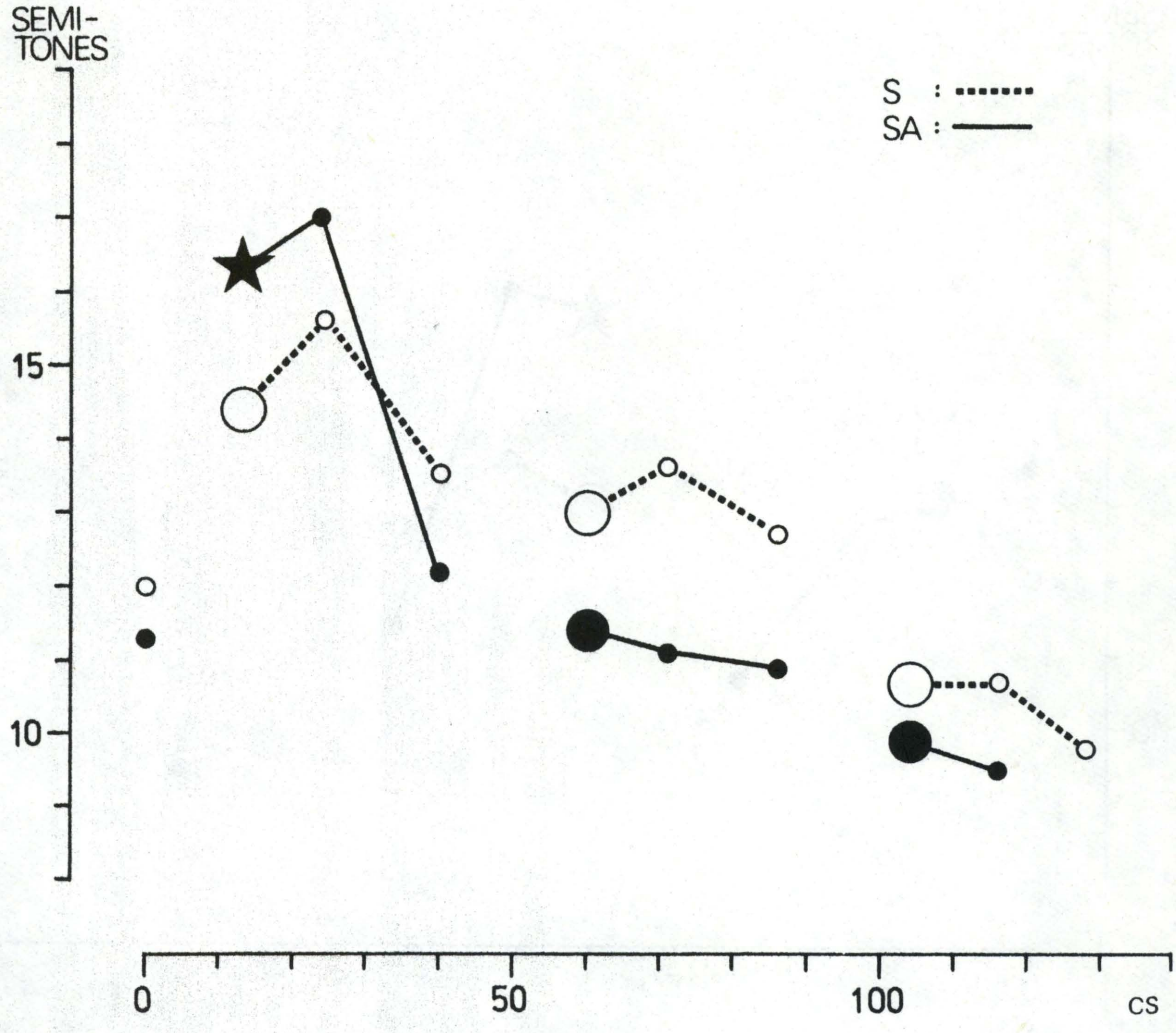


Figure 3

Stylized tracing of a neutral statement (S) and one with emphasis for contrast on the word of the first stress group (SA). S is a combination of the nonsense words in sentences 1-4, 11-15, and 22-25. SA is a combination of the nonsense words in sentences 5-7, 39, and 41.

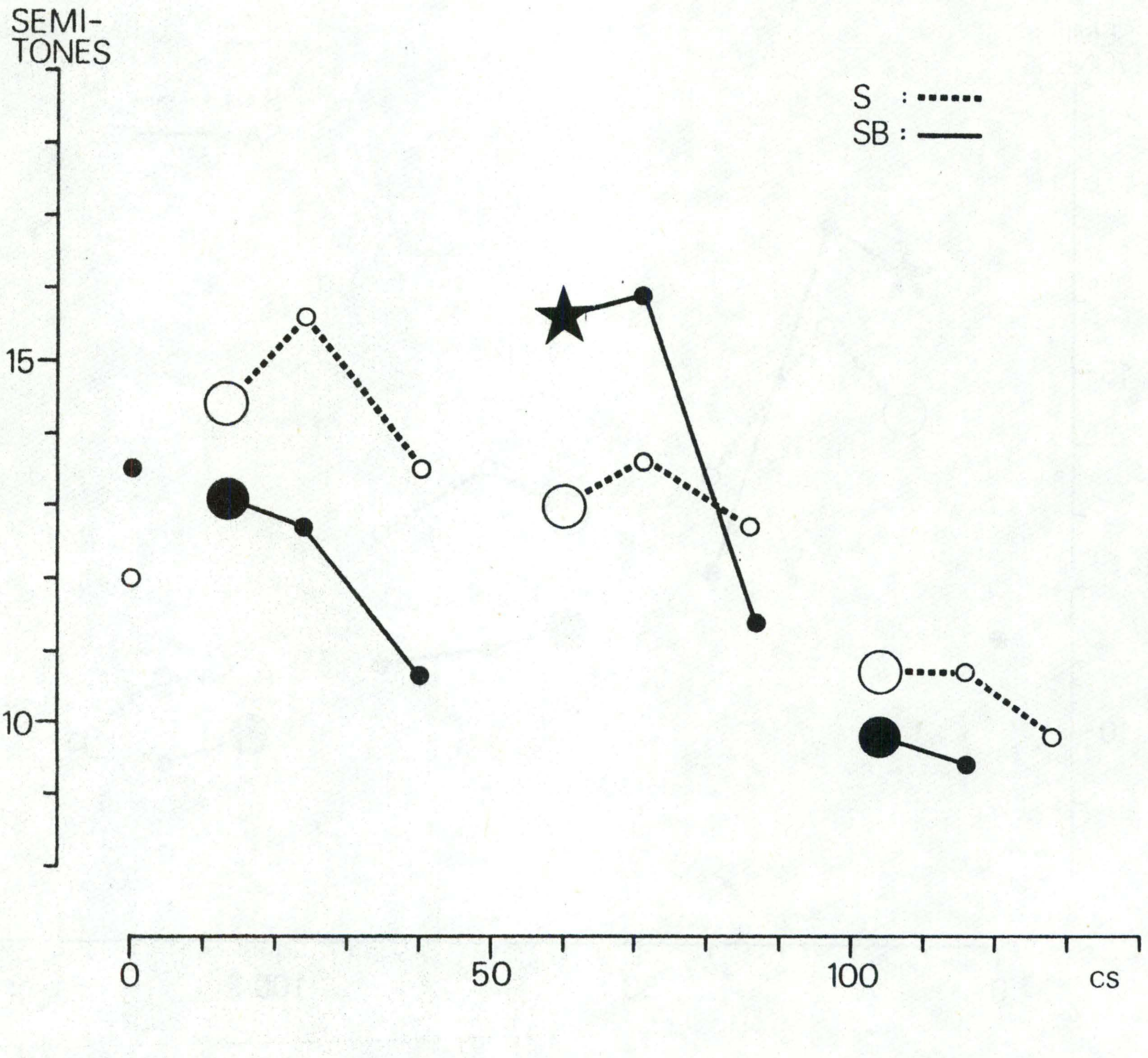


Figure 4

Stylized tracing of a neutral statement (S) and one with emphasis for contrast on the word of the second stress group (SB). S is a combination of the nonsense words in sentences 1-4, 11-15, and 22-25. SB is a combination of the nonsense words in sentences 36-37, 16-18, and 42.

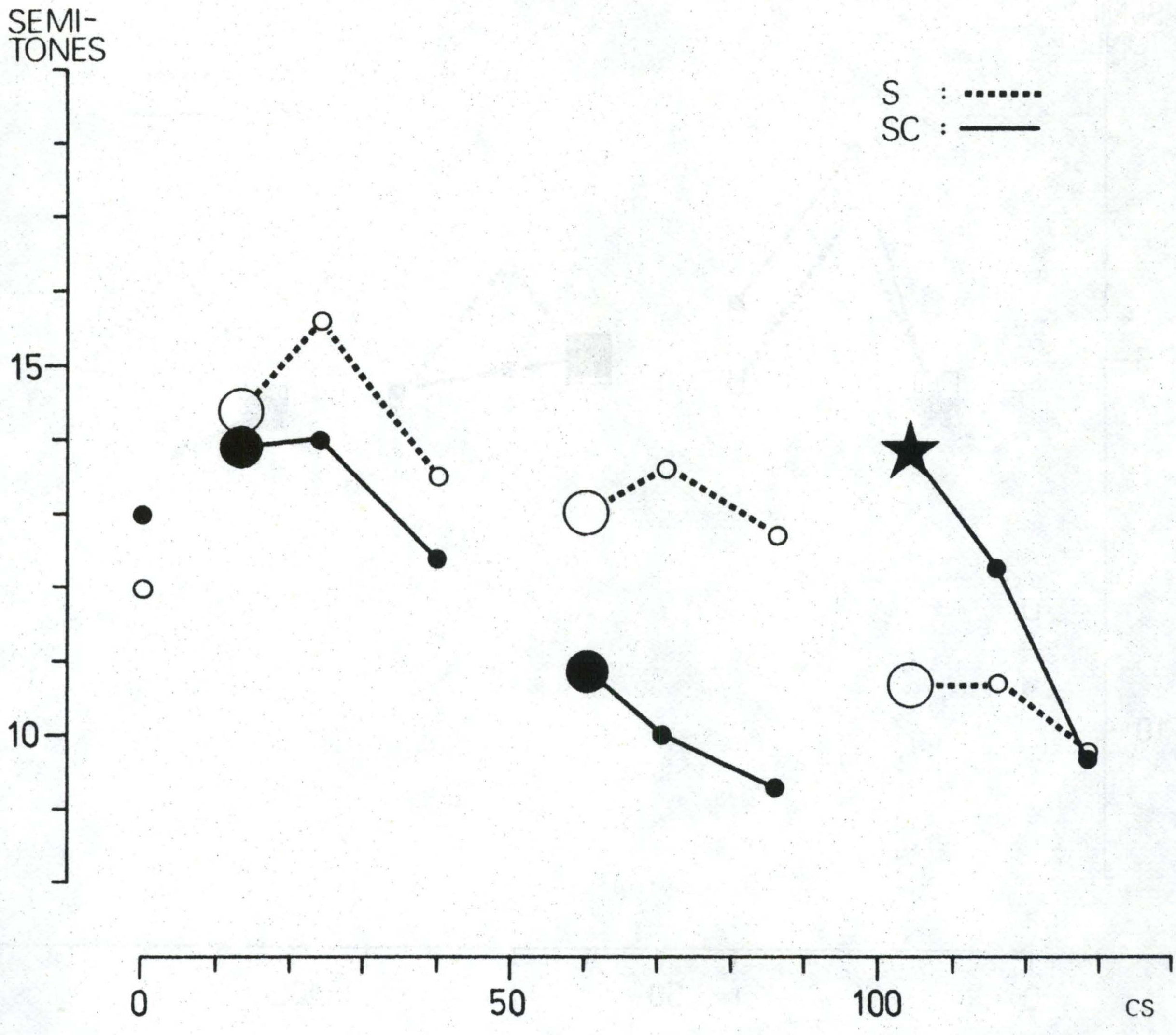


Figure 5

Stylized tracing of a neutral statement (S) and one with emphasis for contrast on the word of the third stress group (SC). S is a combination of the nonsense words in sentences 1-4, 11-15, and 22-25. SC is a combination of the nonsense words in sentences 38, 40, and 26-28.

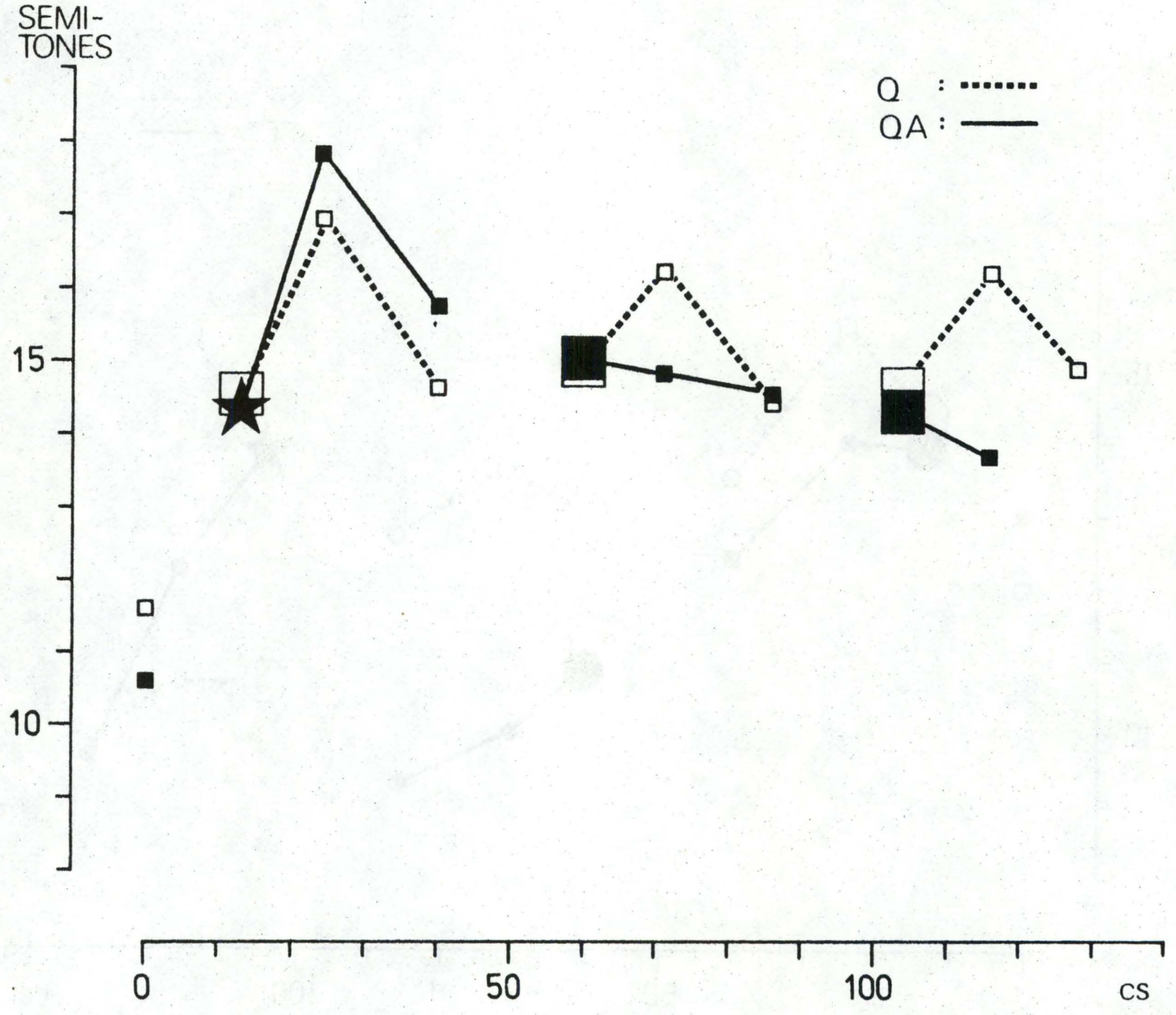


Figure 6

Stylized tracing of a neutral question (Q) and one with emphasis for contrast on the word of the first stress group (QA). Q is a combination of the nonsense words in sentences 8-10, 19-21, and 29-31. QA is a combination of the nonsense words in sentences 43, 46, and 49.

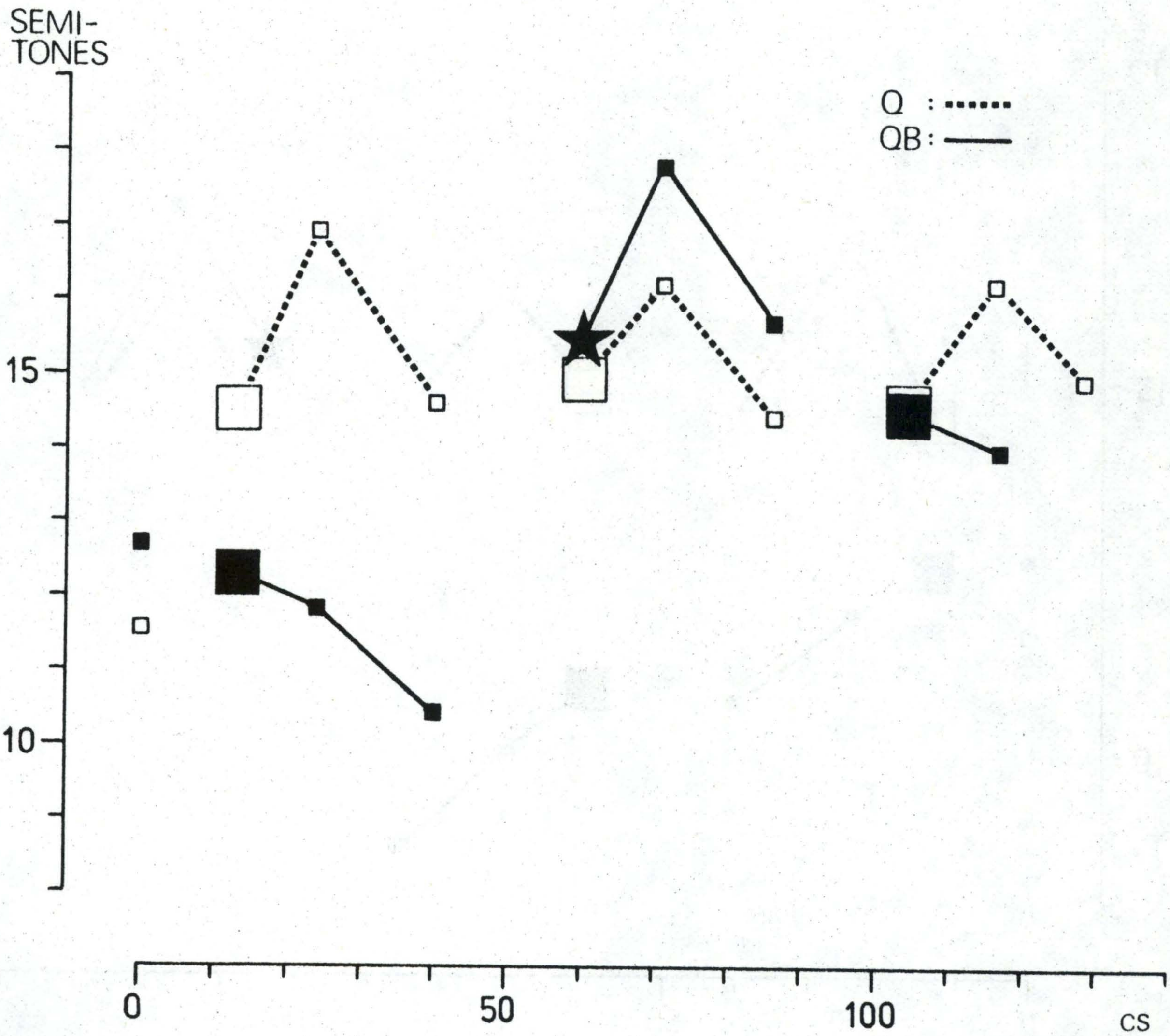


Figure 7

Stylized tracing of a neutral question (Q) and one with emphasis for contrast on the word of the second stress group (QB). Q is a combination of the nonsense words from sentences 8-10, 19-21, and 29-31. QB is a combination of the nonsense words in sentences 44, 47, and 50.

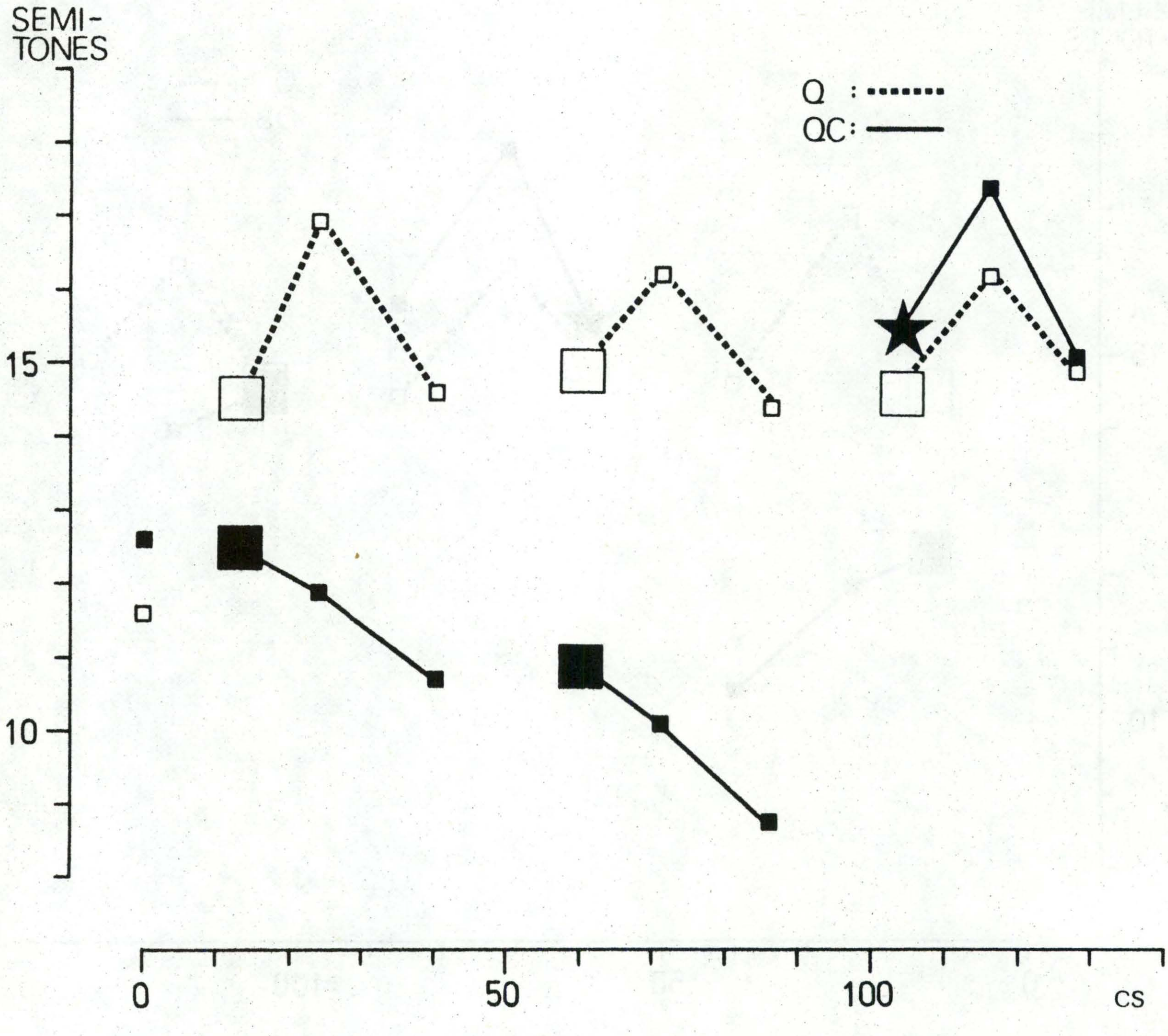


Figure 8

Stylized tracing of a neutral question (Q) and one with emphasis for contrast on the word of the third stress group (QC). Q is a combination of the nonsense words from sentences 8-10, 19-21, and 29-31. QC is a combination of the nonsense words in sentences 45, 48, and 51-52.

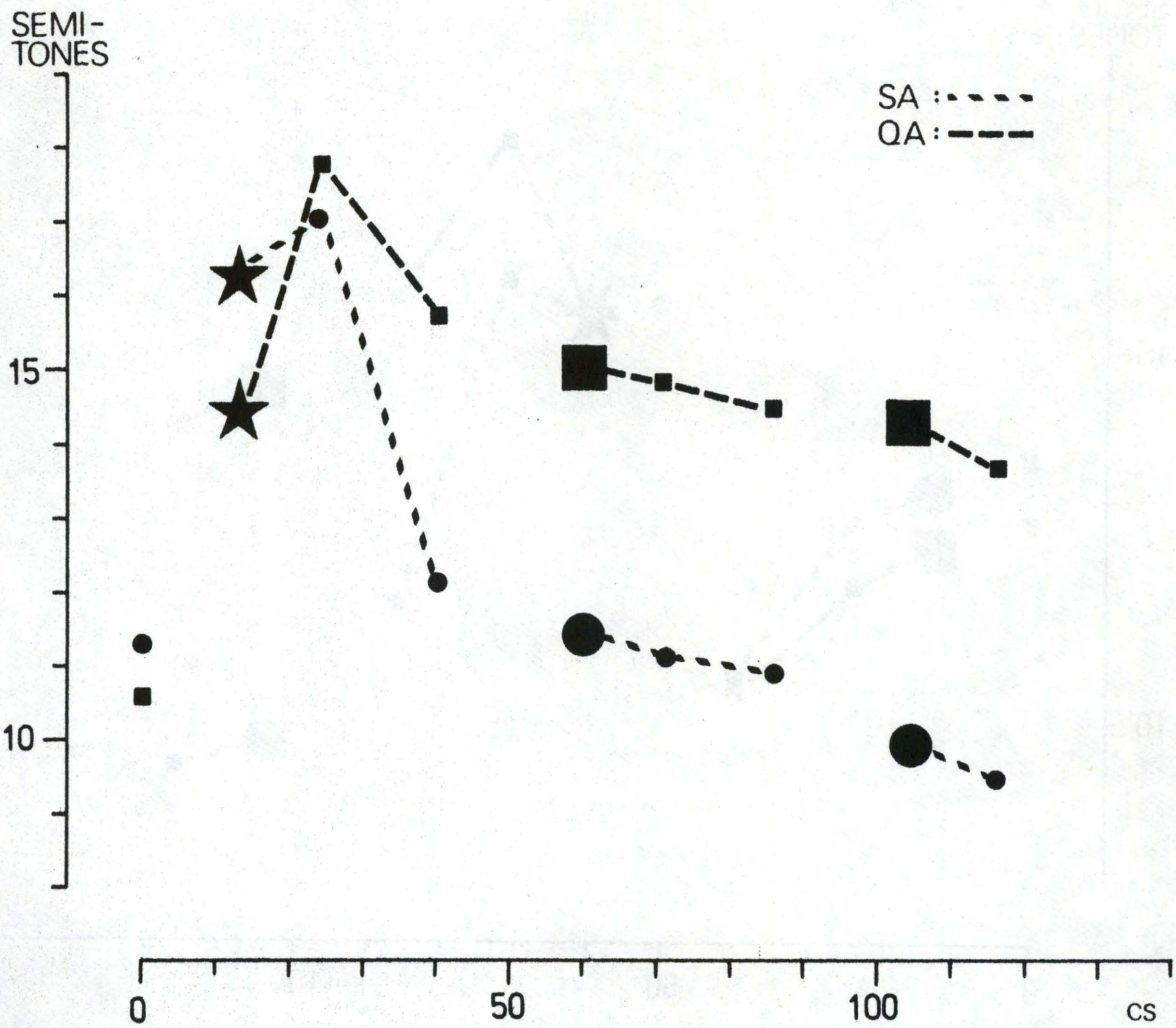


Figure 9

Stylized tracing of a question (QA) and a statement (SA) with emphasis for contrast on the word of the first stress group. QA is a combination of the nonsense words in sentences 43, 46, and 49. SA is a combination of the nonsense words in sentences 5-7, 39, and 41.

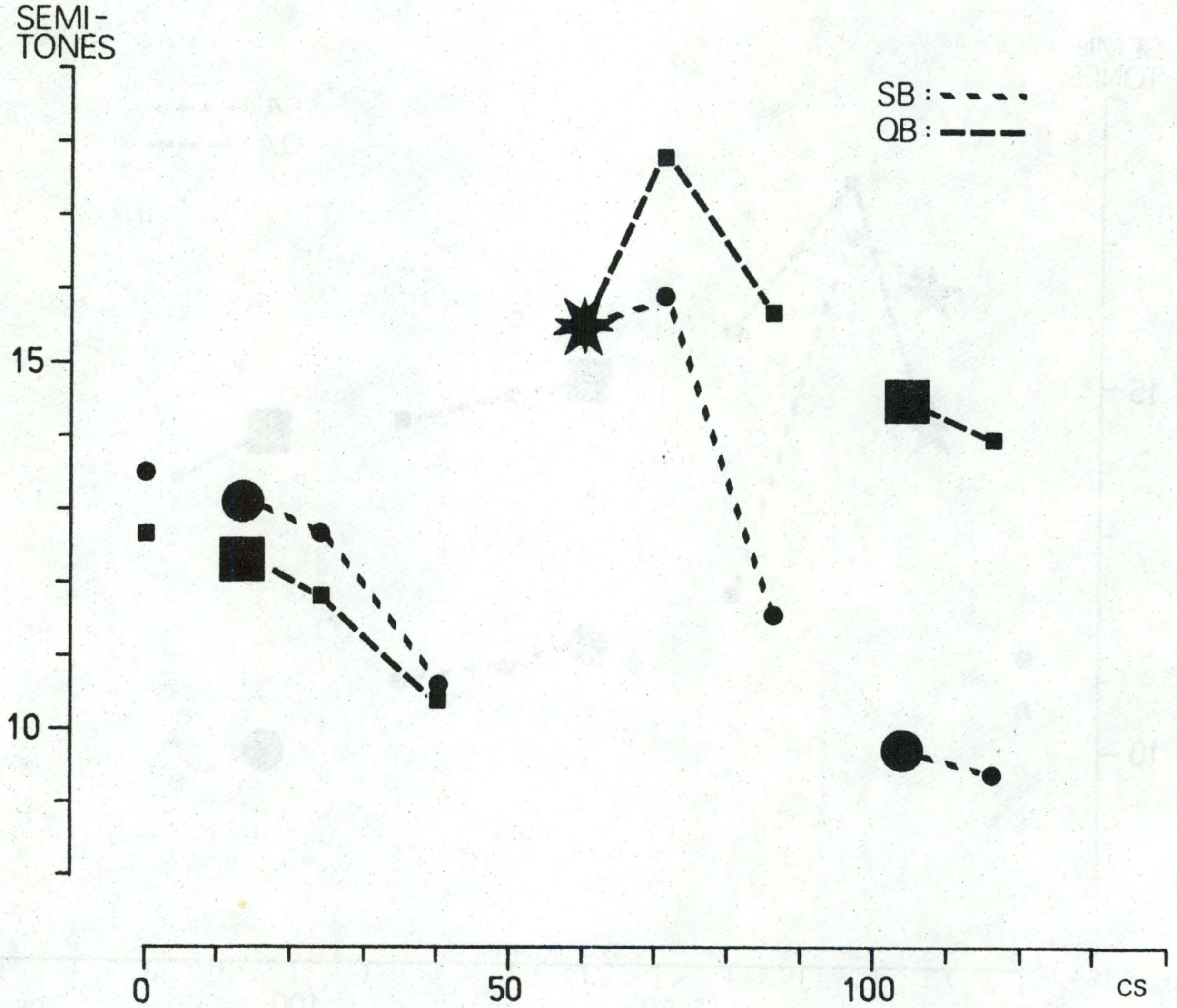


Figure 10

Stylized tracing of a question (QB) and a statement (SB) with emphasis for contrast on the word of the second stress group. QB is a combination of the nonsense words from sentences 44, 47, and 50. SB is a combination of the nonsense words from sentences 36-37, 16-18, and 42.

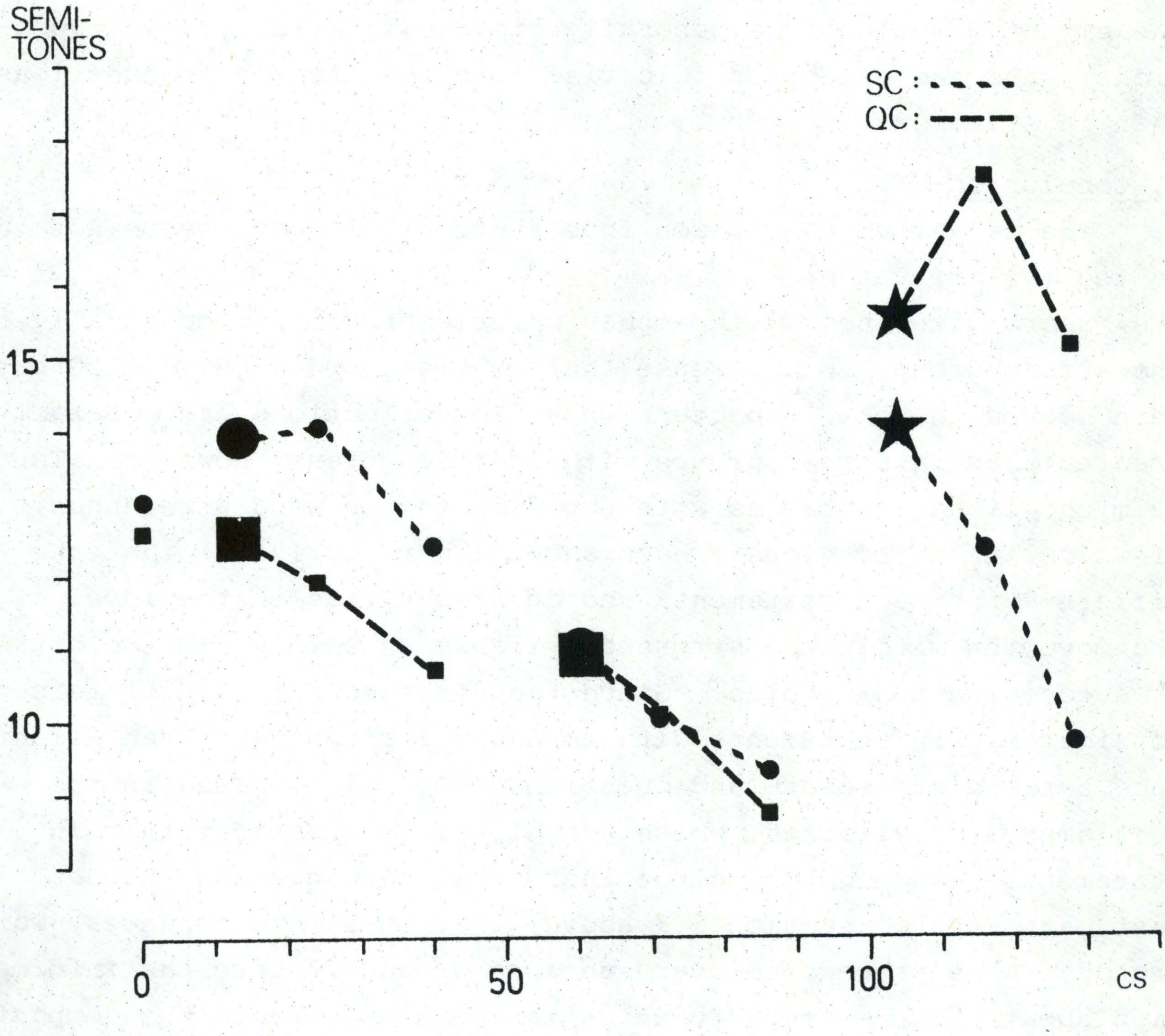


Figure 11

Stylized tracing of a question (QC) and a statement (SC) with emphasis for contrast on the word of the third stress group. QC is a combination of the nonsense words from sentences 45, 48, and 51-52. SC is a combination of the nonsense words from sentences 38, 40, and 26-28.

emphasis for contrast), and this is hardly a salient feature.

The figures do not specify Fo movement within the vowels, but inspection of the raw average tracings of all three (four) subjects who recorded utterances with emphasis for contrast suggests the following account: The Fo movement in the vowels preceding and following the emphatically stressed one is invariably falling. The emphatic syllable is generally rising (invariably so in questions), and the extent of this rise is often greater in questions than in statements.

6. Conclusion

The tentative conclusion from section 3.1 above seems not to be too far off the mark:

Short utterances with emphasis for contrast reduce tonally to one stress group, in the sense that only one LOW+HIGH FALLING pattern occurs in them, a pattern whose Fo deflections are generally greater than in stress groups with neutral stress, however. Thus, we might liken utterances with emphasis to one word utterances, of which it has previously been shown (Thorsen, 1978a) that the difference between statements and questions lies in the level of and movement within the stressed syllable as well as in the course of succeeding post-tonics. Since the stressed syllable is raised considerably in statements with emphasis for contrast, the difference between statements and questions is mainly located in the post-emphatic syllables, which run higher in questions than in statements. We might conclude therefore (as suggested in the last paragraph of section 3.1 above) that in such utterances, word prosody takes precedence over sentence prosody, since the information about sentence type/function is contained within the "emphatic stress group" and not in the ensemble of lexically stressed syllables, as is the case for neutral utterances. But this is neither here nor there. If short utterances with emphasis for contrast are regarded and treated as one word utterances, then the argument about what takes precedence over what is meaningless, precisely because the "word" is simultaneously a complete utterance, i.e. its Fo course will contain information about the distribution of stressed and unstressed syllables as well as about the function of the utterance, and one does not exclude or take precedence over the other.

References

Hombert, J.-M. 1976: "Consonant types, vowel height, and tone in Yoruba", UCLA WPP 33, 40-54

Pierrehumbert, J. 1979: "The perception of fundamental frequency declination", JASA 66, 363-369

Reinholt Petersen, N. 1978: "Intrinsic fundamental frequency of Danish vowels", JPh 6, 177-189 (also in ARIPUC 10, 1976, 1-27)

Thorsen, N. 1978a: "An acoustical investigation of Danish intonation", JPh 6, 151-175 (also in ARIPUC 10, 1976, 85-147)

- 1978b: "On the identification of selected Danish intonation contours", ARIPUC 12, 17-73

- 1979: "Interpreting raw fundamental frequency tracings of Danish", Phonetica 36, 57-78

- this volume: "Intonation contours in declarative sentences of varying length in ASC Danish" (paper read at the Autumn Meeting of the Institute of Acoustics, Windermere 4-6th November 1979).

(ARIPUC 13, 1979)

A COMMENT ON BASBØLL'S PHONOLOGICAL SYLLABIFICATION AS APPLIED TO DANISH

Peter Molbæk Hansen

Abstract: Basbøll's principles for syllabifying phonological strings in Danish are discussed. Certain counter-examples to the predictions of these principles are presented, and it is argued that certain structural properties of Danish words cannot be adequately accounted for by Basbøll's syllabification strategy in its present form.

1. Introduction

In an important paper, Basbøll (1972) launched the idea that the inclusion of a syllable concept in a (generative) phonology of Standard Danish (henceforth SD) would explain (or at least connect in a natural way) quite a few seemingly disparate facts of the surface or near-surface phonological structure of SD. Basbøll's main finding was that a unique set of syllabification principles would supply the sufficient conditions for a simple account of the distribution of allophones of certain phonemes (or, within a generative phonological framework, the surface phonological behaviour of certain phonological segments). Of the distributional statements permitted by the introduction of syllabification, the most important were the following: Short /a/ is pronounced [α][1] before homosyllabic non-coronal consonants, otherwise it is [a]. Short /o/ is pronounced [o] in open syllables and [ɔ] in closed syllables.

---

1) My phonetic transcription of Danish word-forms is in agreement with Basbøll (1969), except that I use [p t k] for [b̥h d̥ˢh g̊h] and [b d g] for [b̥ d̥ g̊]. [α] denotes a low vowel half-way between cardinal vowels 4 ([a]) and 5 ([ɑ]).

/p t k d g r/ are pronounced [p t k d g ʁ] in syllable initial position and [b d g ð ɣ ɒ̯][1] in syllable final position, except that /d/ is dropped after homosyllabic sonorant consonants, and /g/ is dropped after homosyllabic nasals. (For further details, see Basbøll (1972, p. 187 ff, and 1974, p. 40ff.)

In Basbøll (1973c and 1974), this idea is further developed, and some principles for syllabifying phonological strings (at a certain level, cf. below) in SD are given (in the 1974 paper in connection with a more general discussion of the role of syllables and syllabification within the framework of generative phonology).

The present paper is mainly inspired by Basbøll's work. In section 2 we shall discuss the principles of syllabification given in Basbøll (1974) and Basbøll's use of "phonological" syllables and syllabification from a general point of view.

In section 3 we shall develop our main theme, viz. an evaluation of the observational adequacy of Basbøll's principles.

2. Basbøll's principles of syllabification

As mentioned above, several surface-phonological phenomena can be accounted for in a rather simple way if syllable boundaries (marking the domain, i.e. the syllable, of the relevant rules) are inserted in phonological strings according to a judiciously selected set of syllabification principles. The advantage of referring to syllables and syllabification in the description of the so-called consonant gradation in SD - i.e. roughly the alternations p-b, t-d, k-g, d-ð, g-ɣ, and ʁ-ɒ̯ - and certain others, cf. Basbøll (1974, p. 42-43) and Rischel (1970) - is described by Basbøll in the following way: "Now it is very interesting that all these alternations can be subsumed under a single principle, viz. the well-known one of consonant weakening in syllable final position" (1974, p. 43). On p. 44 he goes on to state that "...it is clear

---

1) In most younger standards, [ɣ] has developed to [i̯] after front vowels and [l] and to [u̯] after back vowels and [ɒ̂]. Where this development is relevant to the topic of this paper, it will be further discussed.

that the value of the above mentioned principle depends on whether there can be given explicit (and not unnatural) principles of the location of syllable boundaries that can account for all of the above mentioned phenomena [consonant gradation and the behaviour of short /a/ and /o/; PMH] without giving rise to complications elsewhere in the phonology. This will be attempted ... below."

In this section, Basbøll's principles of syllabification, as formulated in his 1974 paper, will be evaluated both from the point of view of their explicitness and from the point of view of their naturalness.

In section 2.1 we shall discuss whether Basbøll's principles of syllabification are sufficiently explicit for their predictions to be compared with phonological data in SD, i.e. whether or not ambiguities may arise when they are actually applied to phonological strings; this discussion thus concerns neither the naturalness nor the motivation of Basbøll's principles. These matters will be discussed in section 2.2.

In the remainder of this paper the symbol \$ designates a syllable boundary.

2.1 Are Basbøll's principles sufficiently explicit?

After discussing several factors (grammatical boundaries, initial and final segment combinations in grammatical units of different size, stress, surrounding vowels, sequences of consonants) which may interact in different ways in different languages to determine the location of syllable boundaries, Basbøll sets up a hierarchy of three principles "which seem, by and large, to account for the phonological syllabification in Danish" (1974, p. 82). These principles may be paraphrased as follows (but the reader should check this paraphrase against Basbøll's own formulations on p. 84-91):

(i) Let \$'s coincide with grammatical boundaries before words, before stems, before stressed native suffixes, and before zero. If, by the application of this principle, any two consecutive +syllabic segments have a \$ between them, the string is syllabified; otherwise, proceed to (ii).

(ii) In all cases of two consecutive +syllabic segments having no \$ between them, mark off all the places where a \$ can occur without giving rise to a syllable-initial or a syllable-final cluster which is impermissible word-initially or word-finally, respectively; if there is only one such place, a $ is inserted at that place. If, by the application of this principle, any two consecutive +syllabic segments have a $ between them, the string is syllabified; otherwise, proceed to (iii).

(iii) In all cases of two consecutive +syllabic segments having no $ between them, insert a $ at the rightmost place marked off by principle (ii) if the second of the two +syllabic segments is shwa and in the leftmost place marked off by principle (ii) if the second of the two +syllabic segments is a "full" vowel, i.e. any vowel other than shwa, cf. below.

According to Basbøll, the application of these principles at the level of representation which is input to (the first of) the rules that have the syllable as their domain (or refer to a $ in their structural description) will supply the $'s needed by these rules with one important exception, "viz. when a consonant cluster which contains a stop other than /g/ preceded by either a (underlyingly) voiced continuant or a nasal, occurs before shwa. In that case the syllable boundary goes before the stop" (p. 90).

It should be noted that stressed native suffixes comprise both suffixes with main stress and suffixes with reduced main stress (secondary stress) as e.g. -inde and -dom, cp. skuespillerinde ('actress') [sgu·əsbelɒˈenə] and barndom ('childhood') [ˈbɑ:nˌdʌmʔ]. It should also be noted that the unstressed native suffixes -ig ([i]), -(n)ing ([(n)eŋ]), and certain occurrences of -isk ([isg]) have shwa at the level at which syllabification takes place.

The following examples may illustrate the function of these principles: mados ('smell of cooking') and godhed ('goodness') are syllabified /mad$oʔs/ and /go:d$heʔd/ (principle (i)); angre ('regret') is syllabified /ang$ʁə/ (principle (ii), cf., however, below); bade ('bathe') and Ida (a personal name) are syllabified /ba:d$ə/ and /i:$da/ (principle (iii)); verden ('world') and værten ('the landlord') are syllabified /vɛʁ$dən/ and /vɛʁ$tən/ (the exception to principle (iii)).

Now, let us try to determine the meaning of the expression "principles of syllabification". Since they have the effect of changing phonological strings at a certain point in a generative phonological derivation, Basbøll's "principles" of syllabification must apparently be interpreted as informal descriptions of (the function of) a set of phonological rules which apply at a certain point in the generative phonological derivation, namely just before the (first of the) rules which have the syllable as their domain or which mention a $ in their structural description, (cf. Basbøll, 1974, p. 84). In other words, if the derivational machinery is to work, the principles must be spelled out as phonological rules with the combined effect of inserting $'s (hopefully) at the places where they are needed in order to take care of the above-mentioned processes (consonant gradation and vowel adjustment). Such rules can undoubtedly be formulated (on the rule(s) corresponding to principle (ii), however, see below), and I will not discuss here the (mainly formal) problems which probably arise in connection with the formulation of such rules.

Taking for granted that principles (i)-(iii) can be translated (some way or other) to phonological rules, I find them explicit to the extent that the entities they refer to are well-defined. From that point of view principles (i) and (iii) are impeccable: the grammatical boundaries mentioned in (i) and the full vowels and shwa referred to in (iii) are reasonably identifiable at the level at which syllabification takes place, cf. Basbøll's careful descriptions (1974, p. 84-86 and p. 88-89).

The reference in principle (ii) to structurally possible word-initial and -final clusters seems also unambiguous in view of Basbøll's reference (1974, p. 87) to his own work on Danish consonant combinations (Basbøll, 1973a) for a phonotactic description allowing one to distinguish between structurally motivated and accidental gaps in the corpus of consonant combinations. But the phonotactic description given in Basbøll (1973a) concerns a level which presupposes that the $-dependent rules have already applied, cf. e.g. the existence on that level of segments like /ð ɣ ɒ̂/, and cf. Basbøll's explicit statement that "The terms "pre-" and "post-vocalically" refer to the position in the <u>syllable</u> (the syllable division being in accordance with Basbøll (1972))." (Basbøll, 1973a, p. 110, footnote 2). The reference to Basbøll (1972) in this quotation is not of much help either, since the principles for syllable division given in that paper (p. 194) are "intra-level" principles, i.e. they state where the

$'s <u>are</u> (at the level of Basbøll (1973a)), not how they came to be there. Since the rules to which principle (ii) must be translated cannot refer to segments which only arise through the application of later rules, cf. below, we need to know what are the structurally motivated restrictions valid for the level which is input to principle (ii) (= the phonemic level in Basbøll's structuralistic description in 1973c). As far as I can see, this level must come close to the level used in Vestergaard (1967), since the idea underlying Basbøll's whole strategy of syllabification is to account for the allophones of certain phonemic segments (= the segments found at the level which are input to principle (ii) in a generative derivation). But the level in question is not <u>identical</u> to that of Vestergaard (as regards its inventory of consonants and consonant combinations). How, then, can we know what is structurally permissible or excluded on that level if we are not referred to a description of phonotactic constraints valid for such a level? The only answer seems to be that the structurally permissible word-initial and word-final consonant clusters on the level in question are those which - through the application of syllabification rules and syllable final weakening - would <u>yield</u> structurally permissible word-initial and word-final consonant clusters at the level described in Basbøll (1973a). The inventory of permissible clusters on Basbøll's phonemic level (= the level which is input to syllabification) probably only deviates from Vestergaard's inventory in cases where the <u>derivationally</u> corresponding clusters on lower levels are permitted by Basbøll but excluded by Vestergaard (in the sense that Basbøll's rules, if applied to Vestergaard's phonemes, would not generate the desired output). To take an example: In Vestergaard's material (which is reproduced in Basbøll (1973a, p. 138-139)) final clusters containing sonorant + /d/ or /b/ are excluded; but it can be inferred from Basbøll's exceptions to principle (iii) (1974, p. 90) that such final clusters are permissible (and distinct from sonorant + /t/ or /p/) on the level which is input to syllabification, since otherwise it would be superfluous to mention words like <u>ændre</u> ('change' (vb.)) and <u>jambe</u> ('iambus') as examples of such exceptions (they would be unequivocally syllabified by principle (ii)).

I conclude that the clusters permitted word-initially and word-finally on the phonemic level (= the level which is input to syllabification) are such as would <u>produce</u> - through the application of syllabification and syllable final weakening - word-initial and word-final clusters permitted on the level described in Basbøll (1973a).

This means that Basbøll's principle (ii) in a way "looks ahead", i.e. its correct application is conditioned by the possible outcome of tentative applications of principle (ii) and (iii) and other later rules. This is neither circular nor meaningless as long as principle (ii) is only given as an informal description of the function of a phonological rule (or of a set of phonological rules), but as suggested above, such rules must probably be formulated in terms of the segments found at the level at which they apply, since the application of a generative phonological rule can only depend on its own structural description which, as far as I know, is an "intra-level" concept. This means either that the phonological rules corresponding to principle (ii) must refer (in their structural description) to possible word-initial and -final segments on <u>that</u> level or that the rules must refer <u>directly</u> to the sequences of segments found at that level, as Basbøll actually suggests (1974, p. 79); cf. also the formulations of $-insertion rules in other generative phonological works, e.g. Hooper (1972) and Vennemann (1972).

If the rules corresponding to principle (ii) are formulated in terms of the segments found at the level at which syllabification takes place (and this seems to be technically necessary, cf. above) they will, for instance, have to prohibit the potential syllabification /havg$ə/ of the name <u>Hauge</u> [ˈhɑu̯gə], since otherwise the output would be *[ˈhɑu̯ɣə], whereas they will have to allow the potential syllabification /alg$ə/ of <u>alge</u> ('seaweed') [ˈalɣə], since otherwise the output would be *[ˈalgə]. These examples illustrate an important property of the whole strategy: Basbøll has decided that (most of) the phonotactic restrictions valid for consonants should be stated by reference to the phonological syllable and that they should refer to the segments occurring at a level at which "syllable final" weakening has already applied (this is one of his main claims in Basbøll (1973a)).

It is his task, therefore, to supply a syllabification which makes this possible, and this means that the insertion of $'s at a (slightly) higher level is functionally determined rather than structurally motivated from the point of view of the higher level.

In the following, I shall assume that the above-mentioned interpretation of principle (ii) is in accordance with Basbøll's intentions (otherwise one would need a reference to a phonotactic description valid for the phonemic (= pre-syllabificational) level). If this is correct it may be concluded that Basbøll's principles of syllabification are sufficiently explicit in the sense that their application to phonological strings do not give rise to ambiguities. In the next section we shall evaluate these principles from a more general point of view.

2.2 Basbøll's use of syllable boundaries

Basbøll's use of syllables and of syllable boundaries is of interest not only from the point of view of the generative phonological framework in which it was launched. Since it can also be seen as an interesting attempt to account for the distribution of surface phonological segments (see Basbøll, 1973c), it may be compared with other ways of dealing with the syllable in structuralistic and pre-structuralistic as well as in generative phonology and descendants of the latter trend, e.g. natural generative phonology. Since, however, this is not the place for a general discussion of the syllable, I shall confine myself to a few theoretical remarks concerning Basbøll's concept of the role of syllable boundaries.

It may be expedient to consider first Basbøll's motivation for referring to his syllables as "phonological" rather than "phonetic". The following quotation gives us the cue: "In some cases (e.g. Grüsse /grȳs·ə/ the postulated syllable boundary may not coincide with the intuitively felt syllable boundary or with some experimentally established syllable boundary (or better: experimental data may seem to contradict the proposed syllable boundary). This may indicate that the syllable we are dealing with is a more abstract entity than the phonetic syllable, viz. a "phonological syllable". Nevertheless I dare use the term "syllable" since it is an entity which has, in Danish at least, exactly one phonological vowel and whose boundaries can be posited in accordance with some generally recognized principles for syllabification [roughly: principles (i) and (ii) above; PMH]" (1972, p. 193). This quotation and other passages of his clearly show that the "phonological" aspect of Basbøll's syllable lies in its being a descriptive device which need not be motivated by phonetic or typological "naturalness". In other words, Basbøll's phonological syllables (and, in particular, their boundaries) are primarily motivated by what they accomplish (together with the $-dependent rules), viz. by the fact that certain surface-near structural constraints may be economically stated by reference to them. Superficially, Basbøll's principles of syllabification and the later $-dependent rules may thus look like a rule conspiracy in Kisseberth's (1970) sense. The similarity is only apparent, however: postulating a set of rules which have the combined effect of rescuing some (real) surface constraint which is statable in terms of segments in their relation to each other and to independently motivated grammatical boundaries is very different from postulating a set of rules (like Basbøll's) whose application merely permits the description of some surface-near constraints by (direct or indirect) reference to "phonological" syllable boundaries which have, of course, no physico-phonetic (segmental) existence at all, but may, at best, in <u>some</u> cases (e.g. probably not in a word like <u>bade</u> ('bathe') [ˈbæ:ðə]) correspond to psycho-phonetic syllable boundaries.

Incidentally, I find Basbøll's three-way distinction between "phonological", "phonetic", and "psychological" syllables and syllable boundaries (1974, p. 72-73) somewhat dubious. If "the fact that in German, word-initial vowels have a glottal attack, e.g. <u>ein Esel</u> [ʔaɪn ʔe:zl̩]" is an example of "phonetic syllabification" (p. 74), then the terminology is at best confusing. In my view (and in my terminology), the interpretation of glottal stops in German as manifesting syllable boundaries is no less <u>phonological</u> than the interpretation of a Danish word like <u>gade</u> ('street') [ˈgæ:ðə] as /ga:d$ə/, but the former interpretation is obviously more <u>concrete</u> in the sense that the relation between perceptual data (which are conditioned by both physical and psychological factors) and interpretation is more direct.

In order to illustrate the theoretical status of Basbøll's "phonological" syllable it may be fruitful to compare it with two main trends in the use of syllables and syllabification by various phonologists: one of these is characteristic of glossematic theory and may be characterized as follows: once the number of syllable peaks is established at a certain level of representation, then boundaries may be inserted between them at places where the phonologist can use them to provide an economical description of the distribution of segments. In many cases, this gives the phonologist a high degree of freedom to distribute syllable boundaries (with the obvious restriction, though, that only one syllable boundary may be posited between two consecutive syllabic peaks). As is well known, this freedom is typically used (when possible) to insert syllable boundaries at places where they will split up intervocalic consonants or consonant clusters in a possible word-final + a possible word-initial consonant or consonant cluster, the need for making special statements about possible medial clusters thus being eliminated (cf. Basbøll's principle (ii)). In cases where - at a certain level of abstraction - this can be done in several ways, the syllable boundaries are (whenever possible) inserted at such places as to make the manifestations of consonants on lower levels statable in terms of the position in the syllable (cf. Basbøll's principle (iii)). Within structuralistic schools this relatively free use of syllable boundaries often amounts to the postulation of phonemic syllable boundaries which may permit a reduction of the inventory of phonemes, cf. e.g. Hjelmslev's (1938) treatment of [z] and [s] in German as allophones of the same phoneme.

The other main trend is characteristic of natural generative phonology (Hooper, 1976; Kahn, 1976) and differs crucially from the first in that $'s are inserted in phonological strings according to phonetically and typologically based hypotheses of natural syllabification of sequences of segments, leaving relatively little room for language specific deviations.

Basbøll's use of syllable boundaries is in important respects a compromise between these two conceptions: it resembles the former in that its motivation is primarily descriptive economy and in that relatively large freedom is allowed in connection with the placement of syllable boundaries, but it resembles the latter in that these syllable boundaries are nevertheless inserted by explicit principles which are translatable to phonological rules proper.

From the point of view of the latter trend, Basbøll's principles may be criticized on the grounds that some of his $'s are definitely located at unnatural (or at least unexpected) places. This is true, in particular, of the $'s occurring before shwa, especially when the preceding vowel is long as in bade ('bathe') [ˈbæːðə] which is syllabified /baːd$ə/. Basbøll is, of course, aware of that, and in fact he makes a point of claiming that these "unnatural" locations of $'s constitute the "phonological" or "marked" aspect of his syllabification (cf. the quotations above). I shall not go much deeper into such theoretical issues here, but I would like to point to some implications of Basbøll's principles which seem to be of general methodological interest.

In the following we shall discuss each of Basbøll's principles separately.

Ad (i). The general content of this principle, viz. the claim that certain (transparent) grammatical boundaries are also syllable boundaries is uncontroversial, as far as I can see. It is a traditional insight that languages may differ as to the types and ranks of grammatical boundaries which have phonological effects of the sort usually related to syllable structure and syllabification. Pulgram's (1970) distinction between word-languages, nexus-languages, and cursus-languages is probably of typological significance. From this point of view, Danish is definitely not a cursus-language (like e.g. French), but probably a nexus-language, i.e. the morphophonemic string to be syllabified on purely phonological criteria (in Pulgram's terminology: the section) is a non-compound wordform, perhaps including certain clitic extensions. In short, I find Basbøll's principle (i) and also the fact that it is ordered before the other principles both meaningful and typologically plausible.

Ad (ii). This principle is, of course, well known and indeed accepted by many phonologists. I nevertheless think that the application of this principle to Danish (and certain other languages, e.g. German) may be criticized on the following grounds: the splitting up of intervocalic consonant clusters in a possible word-final cluster + a possible word-initial cluster is well motivated in languages in which any vowel may occur in any syllable. In Danish, however, there are certain restrictions in the inventory of vowels found in certain syllable positions, the most important restriction being that shwa does not occur in a word-initial syllable, and this means that principle (ii) becomes somewhat arbitrary: the idea underlying this principle would appear to be that the word should be regarded as composed of an integral number of syllables which are all potential words or minimal utterances. As is well known, such an analysis is not possible in all languages, either because any location of the syllable boundary would lead to a syllable-final or -initial cluster which is structurally excluded word-finally or -initially, as in Finnish, or because certain vowels (notably shwa) do not occur word-initially or word-finally. In Danish and German, shwa does not occur as the first segment of a word (in Danish shwa is even impossible as the vowel of the first syllable of a word). This does not mean that principle (ii) is completely arbitrary, of course; but the application of this principle to a language like Danish seems to be based upon the somewhat questionable assumption that the consonantism of a word is more relevant to potential divisions than its vocalism. Since e.g. the sequences [mnə], [ə], and [ðə] are all structurally excluded as word-initial segment combinations in Danish, neither of the syllabifications /ba:$də/ and /ba:d$ə/ of the word <u>bade</u> ('bathe') [ˈbæ:ðə] (which are both <u>permissible</u> as far as principle (ii) is concerned) are <u>a priori</u> more motivated from the point of view of Danish <u>word structure</u> than the syllabification /ɛ$mnə/ of the word <u>emne</u> ('subject') [ˈɛmnə] (which is <u>excluded</u> by principle (ii)). I do not deny the descriptive advantages of principle (ii) in its combination with certain parts of principle (iii); but the appeal to the conception of syllables as possible words implicit in principle (ii) is not straight-forward in a language like Danish, and if the structure of words must be used as a criterion for excluding certain locations of syllable boundaries, then the reason for considering consonants more important than vowels should be stated explicitly. As mentioned above, these remarks are not meant as a criticism of the descriptive value of principle (ii), and I only mention this problem because the tacit acceptance of principle (ii) is likely to conceal what I consider an important structural property of Danish words, viz. that they cannot always be regarded as composed of syllables which are potential words. Basbøll's statement (1972, p. 194): "Syllables always begin with a "full vowel" or with a possible word-initial consonant or consonant cluster..." must be a lapse, since it is incompatible with his statement later on the same page that "One intervocalic consonant belongs to the syllable of the preceding vowel if the following vowel is /ə/...", but it may nevertheless reflect the above-mentioned neglect of the importance of vocalism in principle (ii).

Ad (iii). This principle is, of course, crucial for Basbøll's whole strategy. However, in addition to being responsible for the major part of the observational inadequacies which will be exemplified in the next section, it gives rise to some minor problems which ought to be taken into account in the evaluation of Basbøll's syllabification. The postulation of principle (iii) was inspired, I think, by four important facts of Danish surface phonology: 1) the inventory of word-final consonants is identical to the inventory of single intervocalic consonants occurring before shwa; 2) the inventory of word-initial consonants is identical to the inventory fo single intervocalic consonants occurring before a full vowel. 3) /a/- and /o/-adjustment invariably take place when the vowel of the following syllable is shwa and when the /a/ or the /o/ is the last vowel of a word. 4) consonants with phonetic stød-basis, i.e. [m n ŋ l ɒ̭ ð ɣ i̭ ṷ], may receive the stød irrespective of whether they are followed by shwa, consonants or word boundary. Thus, by postulating that the syllable boundary goes to the right of a single intervocalic consonant if it is followed by shwa and to the left if it is followed by a full vowel, the need for making special statements about single intervocalic consonants is eliminated. The descriptive advantage of this is obvious, and the principle of locating the syllable boundary to the right of (at least some) consonants before shwa can be traced back at least to Hjelmslev (1951) and has also been adhered to in many Danish dialect monographies (e.g. Jul Nielsen (1968)). In most Jutlandic dialects, this sort of description is even more motivated than in SD because in these dialects the gradation series k-g-ɣ and t-d-ð are parallelled by the labial series p-b-β.

Thus, the description of single intervocalic consonants before shwa as syllable final and also the description of single intervocalic consonants before full vowels as syllable initial must be given credit for a certain elegance.

The situation is somewhat different when it comes to intervocalic consonant clusters. In the case of clusters before shwa, it is clear that it is the behaviour of only certain of these which has lead Basbøll to locate the $ as far to the right as permitted by principle (ii) instead of locating it, say, to the right of the first intervocalic consonant (which would cover the cases with single intervocalic consonants as well): by locating the $ as far to the right as possible it becomes possible to ascribe the behaviour of stops preceded by obstruents and /g/ preceded by sonorants to their position in the syllable in the same way as with single intervocalic stops and /r/. In all other cases, however, there is no descriptive motivation for preferring the rightmost location of the $ compatible with principle (ii) rather than the location one segment to the left of that position; as a matter of fact, the latter location would in most cases be preferable from the point of view of naturalness. Thus, the only motivation I can see for syllabifying a word like salme ('hymn') /salm$ə/ rather than /sal$mə/ is economy of formulation: from a purely descriptive point of view, such a syllable boundary will do no harm within Basbøll's framework, and it permits principle (iii) to be formulated in a simple way; but it has no function whatsoever, and it is such syllable boundaries and also the fact that clusters containing a sonorant + a stop other than /g/ must be excepted from principle (iii) which makes Basbøll's whole strategy less convincing. Consider, e.g., the derivation of the words Hauge, Frauke (personal names), alge ('seaweed'), and malke ('milk' (vb.)) (starting with the level which is input to syllabification (the "phonemic" level); [u̯] is here derived from /v/; it is immaterial in the present context whether it should be derived from /u/ instead, cf. the discussion in section 3 below. The semicolons denote the potential locations of syllable boundaries marked off by principle (ii)):

	Hauge	Frauke	alge	malke
input	havgə	fravkə	algə	malkə
(i)	-----	------	----	-----
(ii)	hav$gə	frav;k;ə	al;g;ə	mal;k;ə
(iii)	-----	frav$kə	alg$ə	mal$kə
weakening	hau̯$gə	frau̯$kə	alɣ$ə	mal$kə
output	ˈhɑu̯gə	ˈfʁɑu̯kə	ˈalɣə	ˈmalkə
(by later rules:	[ˈhɑu̯(.)gə	ˈfʁɑu̯gə	ˈal(.)ɣ/i̯ə	ˈmalgə]).

It is a good illustration of Basbøll's strategy that the similar clusters /vg/ and /vk/ are divided in the <u>same way</u> but for <u>different reasons</u>, i.e. according to two different principles, whereas the equally similar clusters /lg/ and /lk/ are divided in <u>different ways</u> according to <u>one principle</u>, viz. principle (iii) and its exception. Having chosen beforehand to ascribe the pronunciation [ɣ] of /g/ to its position in the final part of the syllable, Basbøll must claim that /g/ is final in <u>alge</u> but initial in <u>Hauge</u> and select his syllabification rules accordingly.

The above-mentioned draw-backs may not be too serious, after all. At least their significance depends on one's conception of the function of syllables and of the factors determining syllabification. If their observational adequacy can be considered satisfactory, Basbøll's principles of syllabification must be considered an interesting attempt to account for consonant gradation and vowel-adjustment.

In the next section we shall see, however, that certain surface phonological facts of SD are incompatible with Basbøll's description.

3. Are Basbøll's principles successful?

Since we concluded above that Basbøll's principles of syllabification are sufficiently explicit to be empirically interpretable, we are now in a position to confront their effects with surface-phonological data from SD to see whether they are observationally adequate in the Chomskyan sense.

First, however, the following remarks should be made: It was natural for Basbøll to emphasize the impressive lot of data which are accounted for by his hypothesis (and which, of course, inspired him to formulate that hypothesis) and to leave it to others to search for data which are not. [Data of the first kind can be found in Basbøll (1972 and 1974).] The data to be presented below belong to the latter category. Such data have not, to my knowledge, been published before (at least not from the point of view of their being counter-examples to Basbøll's findings), and I therefore consider them relevant to an evaluation of the empirical success of Basbøll's syllabification principles, e.g. they may give rise to questions like the following ones: are such data sufficiently marginal to be ignored? if not, can Basbøll's principles of syllabification be amended to cover these facts without becoming extremely ad hoc or without entailing more complications than they were invoked to remove?

The validity of many of the examples in this section depends on the following assumptions concerning Basbøll's descriptive framework (as outlined in Basbøll (1973a, 1973b, 1973c, and 1974))[1]:

(1) Non-alternating as well as alternating occurrences of [ɣ ð ɒ̯ u̯] are derived from /g d r v/ (the alternation v-u̯ raises special problems, though, cf. below). The correctness of this assumption - which is only crucial from a generative viewpoint for the validity of some of the types of counter-examples mentioned below - can be inferred from several passages in Basbøll's writings, cf. e.g. his examples of exceptions to principle (iii) (1974, p. 90) from which it can be seen that e.g. a word like alge [ˈalɣə] contains a stop /g/ at the level which is input to syllabification ([ɣ] does not alternate in the morpheme in question). Anyway, since Basbøll's principles of syllabification are also meant to be part of a structuralistic description of Danish phonology, cf. in particular, Basbøll (1973c, p. 32), it cannot

be decisive, at least in such a description, whether or not certain segments alternate. Since the phonemic level in such a description seems to be identical - as far as the inventory and arrangement of segments is concerned - to the level which in a generative phonology is input to the syllabification rules, I shall refer to such a level as the phonemic level except in certain cases where the generative aspect calls for special comment.

(2) Possible word-initial and -final consonant clusters are those described in Basbøll (1973a) (note especially that /vl/ and /vj/ are taken to be structurally possible word-initial clusters).

(3) The "syllable final weakening processes" comprise the consonant gradation phenomena mentioned in Basbøll (1974, p. 42-43) and the manifestation of /v/ as [u̯] after short vowels, cf. Basbøll (1973c, p. 32) (on the relation between [v], [u̯], and [u], however, cf. below). Note especially that homo-morphemic [ŋ] and [ŋg] are both taken to manifest /ng/, [ŋ] being the syllable final manifestation.

(4) The manifestation of /g/ after homosyllabic vowels and non-nasal sonorants is [ɣ] in conservative standards, whereas it is [u̯], [i̯] or zero in younger standards (cf. also below).

(5) The types of morpheme boundaries obligatorily coinciding with syllable boundaries are those mentioned in principle (i) above.

The correctness of these assumptions can be inferred from Basbøll's writings on Danish phonology (cf. the references above), and I only mention them explicitly, because they are crucial for many of my counter-examples. In some cases, however, some (parts) of these assumptions will be further discussed.

It should be noted that Basbøll's principles "should ... be taken as exemplifying the preceding discussion [of various grammatical and phonological criteria for syllabification; PMH] rather than explaining the location of every syllable boundary in Danish" (1974, p. 83). Thus, if e.g. some of my examples depend crucially on the restricted list of boundary types given in principle (i), this is not in itself fatal to Basbøll's general strategy: Basbøll's principles are tentative (as he himself remarks) and to a certain extent independent of his general use of syllabification. In other words, some of the following types of examples do not per se seriously undermine Basbøll's syllabificational strategy, but they are all problematic in some way or other if principles (i)-(iii) are interpreted rigoristically.

Many of the examples in this section are foreign words and proper names. However, I do not think that this can be used as an argument against their validity as counter-examples to the predictions of Basbøll's principles: most of the material used by Basbøll himself to illustrate the dependence of his syllabification upon the distinction between "full" vowels and shwa consists of foreign words and proper names (like *Amanda*, *Hulda*, *Gerda*), and it is one of the main claims in Basbøll (1972) that it is the presence of full vowels in posttonic syllables, not e.g. a deviant syllabification, which signals a foreign word structure, a view which I fully share.

The structural types which are problematic in connection with Basbøll's principles of syllabification fall into several groups:

(I) Most words in which, according to assumptions (1) and (3) above, one of the phonemic clusters /vj vr vl/ occurs between a (preceding) short vowel and a (following) full vowel which belongs neither to a stressed native suffix nor to one of the suffixes *-ing*, *-ig*, and *-isk*. Since a $ before any of the clusters in question would yield a structurally possible word-initial cluster, and since they are followed by a full vowel (not derivable from shwa), the application of principles (ii) and (iii) will place the $ before these clusters; they will accordingly be treated as syllable-initial, and the syllable-final weakening rules will not apply to them, but in the following words the intervocalic clusters are pronounced with a "weak" first member, i.e. the pronounced clusters are [u̯i̯ u̯ʁ u̯l]: *Sovjet* ('Soviet'), *aura* ('aura'), *aurikel* ('auricula'), *aurora* ('aurora'), *Laura*, *Laurids*, *Europa*, *manøvrere* ('manoeuvre' (vb.)), *Euripides*, *Povla*, *Paulus*, *Aulin* (name of a firm), *aula* ('aula', 'hall'), *paulun* ('tent', 'pavilion'), *støvlet* ('bootee'), *Pauline*.[1]

To be sure, the validity of these words as clear counter-examples to Basbøll's predictions is disputable: they would be more detrimental to Basbøll's syllabification if their [u̯] were

---

1) But note that *Sovjet* and *manøvrere* may also be pronounced with [vi] and [vʁ], respectively, i.e. in *some* words [v] and [u̯] are in (a kind of) free variation.

in *alternation* with [v], but this is not the case in the words above, and I am not sure that examples of alternations between [v] and the [u̯] of *homo-morphemic* [u̯i̯]-, [u̯ʁ]-, and [u̯l]-clusters can be found. The significance of the above examples of [u̯C]-clusters thus hinges upon two assumptions: a) that they are derived from /vC/-clusters, and b) that [u̯] is the manifestation of /v/ in the final part of the syllable, cf. assumptions (1) and (3) above. Although this seems to be Basbøll's position (1973b, p. 42), his suggestion (1973c, p. 76ff) that [Vu̯]-diphthongs before homo-morphemic obstruents (e.g. in a word like *sovs* ('sauce') [sʌu̯ˀs]) be derived from underlying /Vu/ could probably easily be generalized to comprise all occurrences of non-alternating [Vu̯], and in that case the words above would not be counter-examples. I nevertheless think they are worth mentioning, since Basbøll's principles of syllabification are clearly relevant to a structuralistic interpretation of [v] and [u̯] as allophones (syllable initial and -final, respectively) of one phoneme (/v/). Even under this analysis, however, the intervocalic [u̯l]- and [u̯i̯]-clusters before full vowels might be considered marginal. They are listed here because /vl/ and /vj/ are considered possible word-initial clusters in Basbøll (1973a), cf. also the (foreign) names *Vladimir* and *Vietnam*. (The pronunciation [sʌˈvi̯ɛd] of *Sovjet* is, of course, in agreement with Basbøll's predictions.)

In short, what the examples with intervocalic [u̯C] (at least those with [u̯ʁ]) actually show is that if Basbøll's machinery is to work, then *either* [u̯] cannot be derived from /v/ in such cases (phonemically speaking: [u̯] and [v] cannot be interpreted as allophones of the same phoneme) *or* the distribution of [v] and [u̯] cannot be accounted for by reference to their position in the *syllable*. Basbøll's own suggestion that [u̯] be derived from /u/ (phonemically speaking: that [u̯] be the postvocalic allophone of /u/ (at least after short vowels; after long vowels the situation is more complex)) would probably be a satisfactory solution of this problem. /v/ would then simply be defectively distributed (it would not occur after a homosyllabic short vowel), and there would be no discrepancy between the behaviour of the intervocalic consonant cluster in *livré* ('livery') [liˈvʁe] and e.g. *manøvrere* ('manoeuvre' (vb.)) [manøu̯ˈʁeːˀɒ] (these words would be underlyingly and/or phonemically /liˈvre/ and /manøuˈreːˀrə/, respectively).

But of course, a rule which de-syllabifies postvocalic /u/ would have to be postulated.

(*II*) Some words in which the phonemic clusters /gr gl gn gv/ occur in environments like those mentioned under (I) above; although principles (ii) and (iii) will insert a $ before these clusters they are pronounced with a weak first member ([ɣ] or, in younger standards, [u̯] or [i̯] according to the specific context): (*Sigrid*), *Børglum*, *Ragna*, *kognitiv* ('cognitive'), *inkognito* ('incognito'), *magnet* ('magnet'), *magnat* ('magnate'), *stagnere* ('stagnate'), *stagnation* ('stagnancy'), *magnium* ('magnesium'), *magnesium* ('magnesium'), *magnum* ('big-'), *sphagnum* ('peat moss'), *Magnus*, *Agnete*, *Dagny*, (*Sigvald*, *Sigvard*).

Most of these examples are clear counter-examples in those conservative standards in which postvocalic written *g* is *pronounced* [ɣ] in the words above. (In some very conservative standards the pronunciation [ŋn] is possible for intervocalic written *gn* in these words, cf. e.g. Hansen (1956, p. 82). In such standards, the words *magnet*, etc. are counter-examples of the type illustrated in (IV) below.) In most younger standards, however, the words with intervocalic written *gr*, *gl*, *gn*, *gv* contain (non-alternating) [u̯C]- and [i̯C]-clusters, and if, in these standards, such words have *underlying* /jC/ and /vC/, then the words with [i̯C]- and [u̯n]-clusters are definitely not valid as counter-examples, since /jC/ and /vn/ are impossible word-initial clusters; furthermore, the remarks above (concerning non-alternating [u̯C]-clusters) will, in these standards, apply to the words in which written *gr* and *gl* are pronounced [u̯ʁ] and [u̯l].

If, however, these words contain underlying (phonemic) /gC/-clusters (and within Basbøll's framework this seems inescapable in standards with *surfacing* [ɣ]) they are incorrectly processed by Basbøll's rules, since the $ will be located before the /g/, whereas words like *agrar* ('farmer') [aˈgʁɑːˀ] and *agglutinerende* ('agglutinating') [agluti ˈneˀɒnə] are in agreement with Basbøll's treatment. The pronunciation of the words *prognose* ('prognosis') and *diagnose* ('diagnosis') is of particular interest in this connection. These words *may* be pronounced [pʁoˈgno:sə] and [diaˈgno:sə], respectively. Within Basbøll's framework, these pronunciations could be due to two different interpretations: 1) a "pre-stem" boundary may be felt by some speakers to occur before /g/ (this would be etymologically "correct"), or 2) the words may be interpreted as mono-morphemic. In either case, Basbøll's principles will locate the syllable boundary before /g/, albeit for different reasons. In standards with [ɣ], however, the pronunciations [pʁoɣˈno:sə] and [diɑɣˈno:sə] are by no means rare. It is tempting to assume that the pronunciations with [g] in the latter ([ɣ]-)standards reflect the bi-morphemic analysis, whereas the pronunciations with [ɣ] represent the mono-morphemic analysis which leaves the syllabification to be determined by purely segmental criteria. Anyway, the pronunciation with [ɣ] can hardly be compatible with a morpheme boundary before this [ɣ]. (It should be mentioned that the pronunciations [pʁoˈno:sə] and [diaˈno:sə] are also common.)

The proper names Sigrid, Sigvald, and Sigvard are parenthesized because a morpheme boundary (of the "strong" kind mentioned in principle (i)) might be postulated to occur after Sig-, cf. names like Ingrid, Thorvald, and Edvard, but they are of interest because they are probably the only existing words with intervocalic [ɣʁ] and [ɣv] before full vowels, and it is not unlikely that such clusters could be freely introduced in Danish monomorphemic words (this question concerns the descriptive (predictional) adequacy of Basbøll's principles); it seems inescapable that the other words are incompatible with Basbøll's analysis, at least if they are pronounced with [ɣC]-clusters.

(III) Some words with postvocalic /dj dr dv/ before a full vowel: Gudrun, Edvin, Edvard, klodrian ('clumsy person') and some words of Latin origin with the Latin prefix ad-, e.g. adjudant ('adjutant'), adjektiv ('adjective'), advent ('advent'), advokat ('lawyer'). If these words are to be processed correctly by Basbøll's rules, a $ must be inserted after /d/ in order to generate the pronunciation [ð] of this phoneme, and this must be taken care of by principle (i), since principles (ii) and (iii), if allowed to apply, would locate the $ before the /d/ (for the reasons mentioned above). This means that a transparent morpheme boundary (rigoristically: a morpheme boundary belonging to one of the categories mentioned in principle (i)) must be postulated to occur after /d/. As for the latter type (with ad-) one can, of course, claim that such a boundary exists, but the transparency of such a morpheme boundary to other than linguists or latinists is highly questionable in cases where the "stem" does not occur in isolation (or in combination with other "prefixes") and/or the ad- is not interpretable (to the naive speaker) as a prefix, cf. that there are no such "stems" as *jektiv, *judant, *vent, *vokat. If such a morpheme boundary is nonetheless postulated in order to rescue Basbøll's principles, then it will be difficult to explain why there is apparently no such boundary if the latin stem begins with a vowel as in adaptere ('adapt') [adɑbˈteːˀɒ], adept ('adept') [aˈdɛbd], adoptere ('adopt') [adʌbˈteːˀɒ]. It is tempting to assume that the pronunciation of written d as [d] or [ð] in such words is conditioned by purely phonological, i.e. non-grammatical, criteria.

The word adækvat ('adequate') is particularly instructive in this connection. This word is pronounced [adɛˈkvæːˀd] by some people (primarily, I think, by people who are not aware of the fact that from a latin point of view it is morphemically complex). However, those who use this word actively and frequently almost always pronounce it with [ð]. This may in some cases be due to a morphemic analysis, but it is highly significant that the pronunciation of the written d as [ð] in this word is often correlated with the pronunciation [ə] of the following written æ, i.e. [aðəˈkvæːˀd] (at least there seems to be no opposition between [ə] and [ɛ] in such a context). This pronunciation is quite regular if no morpheme boundary is felt to occur after /d/.

As for the words Gudrun, Edvin, Edvard, and klodrian, I would not claim that the postulation of a (transparent?) morpheme boundary after /d/ is entirely ad hoc, cf. words like Gudmund, Ervin, Sigvard, dumrian ('stupid person'), grimrian ('ugly person'); but the addition of such types of morpheme boundaries to the categories listed in principle (i) obviously makes Basbøll's whole strategy considerably more complicated and less attractive, and like the words with intervocalic [ɣʁ] and [ɣv], such words are of interest to the descriptive adequacy of Basbøll's principles.

(IV) Words like jonglere ('juggle'), jonglør ('juggler'), Ingrid, pingvin ('penguin'). At the level which is input to syllabification, these words must apparently contain the intervocalic clusters /ngl ngr ngv/ since the only source of [ŋ] within Basbøll's framework seems to be an underlying nasal followed by /g/ or /k/, the post-nasal /g/ being deleted (via lenition to [ɣ]) in the final part of the syllable. (It is possible that the nasal is at that level already specified as velar; anyway, the rule that velarizes a nasal before /k/ and /g/ seems to have a domain larger than the syllable, but it must at least apply before the syllable-final deletion of post-nasal /g/, and this means that /g/ must be present at the level at which syllabification takes place.) Since these words are normally pronounced [ʃʌŋˈleːʔɒ, ʃʌŋˈløɒ̭ʼ,ˈeŋʁið, peŋˈviːʔn], the $ must be inserted after /g/, but principles (ii) and (iii) will (for the reasons mentioned under (I) above) insert it before /g/, thus yielding *[ʃʌŋˈgleːʔɒ, ʃʌŋˈgløɒ̭ʼ, ˈeŋgʁið, peŋˈgviːʔn] as the final output. (Of these pronunciations at least [ˈeŋgʁið] and [peŋˈgviːʔn] are hardly ever heard; the others are *possible* (according to ODS) but extremely rare, as far as I know.) To state the problem in nuce: it is difficult, within Basbøll's framework, to account for the fact that a word like *lingvist* ('linguist') [leŋˈgvisd] has a pronounced [g] while a word like *pingvin* has not.

(*V*) In words like *gamma* (name of the Greek letter), *Hammurabi* *Kamma*, *mammon* ('mammon'), *mammut* ('mammoth'), *Abba* (name of a popular Swedish song group), *Pablo*, *kappa* (name of the Greek letter), *Afrika*, (*akkumulator*) ('accumulator'), *akkurat* ('accurate'), *Malacca*, *Bacchus*, *khaki* ('khaki'), (*akvavit*) ('aquavit'), *Jakob*, *Ajax*, *Maja*, the first and/or stressed vowel is [ɑ] in my speech and also in the speech of most other speakers of SD (I have interviewed a few speakers of Copenhagen Standard Danish with this in mind and asked my phonetician colleagues), and words like *gummi* ('rubber') and *Gunna* are invariably pronounced [ˈgɔmi] and [ˈgɔna], as far as I know. Within Basbøll's framework these pronunciations would presuppose a $ after the (first) intervocalic consonant, but the principles of syllabification will place it before the (first) intervocalic consonant (for the reasons mentioned under (I) above). Although some words of this type vacillate between [a] and [ɑ] (in my material, this seems to be true of the parenthesized words), I think there is sufficient evidence that the correlation in $C_0aC_0VC_0$-structures between frontness (acuteness) of /a/ before an intervocalic non-coronal (grave) consonant (or before an intervocalic consonant cluster whose *first* member is non-coronal) and the presence in the following syllable of a full vowel is not so high as Basbøll seems to presuppose. This is particularly significant in cases where the intervocalic consonant is an exclusively "syllable-initial allophone", i.e. [p] or [k]: if <u>akkurat</u> is pronounced [ɑkuˈʁɑːˀd] as seems to be the normal case, then the domain of /a/-adjustment must be larger than the syllable (but smaller than the word, cf. below). Incidentally, Basbøll mentions this particular problem (1974, p. 67); he interprets such cases as a symptom of a phonological change in progress, the /a/-adjustment rule being in the process of enlargening its domain. This may be true (although I am somewhat sceptical about his claim that the pronunciation [pɑˈpiɒ̂ˀ] is a <u>new</u> phenomenon (to the extent that it occurs)), but the important thing is that if /a/- and /o/-adjustment (cf. the examples <u>gummi</u> and <u>Gunna</u> above) do not have the syllable as their domain, then two important arguments for Basbøll's syllabification are seriously weakened.

(<u>VI</u>) Words like <u>Harry</u>, <u>Lorry</u>, <u>karry</u> ('curry'), <u>sherry</u> ('sherry'), <u>terrier</u> ('terrier'), <u>terrakotta</u> ('terra-cotta'), <u>paritet</u> ('parity'), <u>Karoline</u> are hardly ever pronounced with the consonantal, "syllable initial" allophone [ʁ] of /r/ as predicted by Basbøll's rules (which will treat the /r/ as syllable initial). For distributional conditions on the pronunciation of etymological /r/, see Brink and Lund (1975, p. 261ff). One could, of course, claim that the unstressed, posttonic [i] in <u>sherry</u>, <u>paritet</u>, etc. is /əj/ at the level of syllabification, cf. that such a strategy is used by Basbøll to account for the unstressed suffixes <u>-ig</u>, <u>-(n)ing</u>, and certain occurrences of <u>-isk</u> (1974, p. 88), but still the words with posttonic [y], [o], and [a] make trouble.

(<u>VII</u>) The words <u>Canada</u>, <u>Malaga</u>, <u>Paludan</u>, <u>annuum</u> ('annual grant'), <u>kognak</u> ('brandy') are to my knowledge invariably pronounced [ˈkanˀada, ˈmalˀaga, ˈpalˀudan, ˈanˀuɔm, ˈkʌnˀi̯ɑg]. If, for a moment, we ignore the stød, their segmental structure after syllabification must be /kanada, malaga, paludan, anuom, kɔ$njag/, according to Basbøll's principles, and this syllabification raises serious problems for the stød, as far as I can see.

Basbøll considers the stød to be a syllable prosody which is assigned to certain syllables by rule (1974, p. 46ff), but it is far from clear where, in the generative phonological derivation, the stød assignment rule is meant to apply; nor is it clear how it applies. In Basbøll (1974, p. 53) a stød assignment rule (applicable to words of the type represented by the words above) is formulated thus:

$$\begin{bmatrix} \text{S} \\ \\ +\text{stress} \end{bmatrix} \rightarrow [+\text{stød}] / \ \_\_ (\text{SS})\ \#\#$$

Does the symbol S in this rule refer to a syllable complete with peak and boundaries? Apparently not, for in his general discussion of syllabification Basbøll states that "...where the syllable functions as a unit in phonological rules (i.e. typically in rules concerning prosodic features like stress, tone, and stød....), syllabification is *not* required for the correct application of the rules: what is necessary is only that the *number* of syllables be known, and this information can possibly be given with an identification of the syllabic *peaks*" (1974, p. 68). I take this and other passages of Basbøll's to mean that an abstract stød-prosody is assigned (at a relatively early step in the derivation, at least before syllabification) to certain syllabic peaks. The *manifestation* of the stød, as a glottalization of the first segment after the syllabic peak if that segment is sonorant (including the final part of a (phonetically) long vowel), must, however, be taken care of by a later rule which must be ordered *after* syllabification and syllable final weakening, since, otherwise, words like *Agner* (a personal name) [ˈɑɣ’nɒ] and *edder* ('venom') [ˈeð’ɒ] could not be realized with stød (before syllabification, the postvocalic segment would be an obstruent ([g] or [d]) which cannot receive the (phonetic) stød). A word like *Malaga* thus has the segmental structure /malaga/ when it is input to the stød manifestation rule, and this means that such a rule will have to disregard $'s in its structural description (or, within Basbøll's framework, its domain must be larger than the phonological syllable), since the syllable initial /l/ is the stød segment. Although such an arrangement of rules would generate the correct output, I sincerely doubt that such consequences are in accordance with Basbøll's intentions: it is hardly meaningful to claim that the stød is a property of the *syllable* if it is first assigned to an entity (the syllabic peak) which may be smaller than a syllable and then specified, by a later rule whose domain is larger than the syllable, as glottalization of a segment which in words like those above belongs to the following syllable. If the conception of the stød as a syllable prosody is not discarded, then it seems inescapable that words like those above must be listed as exceptions to principle (iii).

We shall now take an overall view of these counter-examples and try to evaluate their importance both from the point of view of Basbøll's syllabificational strategy and from the point of view of Danish phonology in general.

Let us consider first the types (I) - (IV). These types have three things in common: a) they contain an intervocalic consonant cluster which on the phonemic level is a possible word-initial cluster (or better: the location of a $ before such clusters would give rise to a syllable initial cluster which is possible word-initially); b) there is no (transparent) morpheme boundary between their first member and the following full vowel (at least this holds true in the majority of cases); c) they are followed by a full vowel. What is at stake here is thus that part of principle (iii) which locates the $ in the leftmost position compatible with principle (ii), before a full vowel. From the point of view of Danish phonology, the significant thing about this part of principle (iii) is that it amounts to postulating that a) in homomorphemic V1C1C2V2-structures where V1 is a short vowel, where V2 is a full vowel, and where C1C2 is one of the phonemic clusters /(vj vr vl) gj gr gl gn gv dj dr dv/, these clusters must be manifested as [(v̂i̯ vʁ vl) ĝi̯ gʁ gl gn gv d̂i̯ dʁ dv], respectively, and b) in homo-morphemic V1ngC2V2-structures /ngj ngr ngl ngn ngv/ must be manifested as [ŋĝi̯ ŋgʁ ŋgl ŋgn ŋgv], respectively. The above examples show that this does not hold true in all cases. It is obvious, however, that these structures are not of equal importance. As mentioned above, the examples with intervocalic [ûî ûʁ ûl] before a full vowel need not per se be detrimental to Basbøll's analysis, and I shall not discuss this type further here, since the analysis of labials poses special problems which are, at least in part, irrelevant to Basbøll's principles of syllabification. They are therefore parenthesized above. Of the remaining clusters, some could perhaps be discarded as marginal or not even valid as counter-examples: the only examples I have found with "unexpected" manifestations of phonemic /gr/- and /dr/-clusters are Sigrid and Gudrun which might be analyzed as morphemically complex (quasi-compounds or the like) and the same may be said of /ngr/ in Ingrid. The /Cr/-clusters are nevertheless of some interest, because they - like the remaining clusters - have some bearing on the question of possible contrasts between consonants in medial clusters before full vowels, and in the following discussion the relevant /Cr/-clusters will be discussed along with the remaining ones (including the /(n)gj/-clusters for which I have found no counter-examples. The only example with this cluster I can think of is the name Ingjald which is, to my knowledge, invariably pronounced [ˈeŋˌgi̯alʔ] in accordance with Basbøll's predictions). The behaviour of these (phonemic) clusters gives rise to the following questions: (1) in the homo-morphemic structures V1C1C2V2 where V1 is a short full vowel, where C2 is one of the non-syllabic segments [ʁ i̯ l n v], and where V2 is a full vowel, is there in position C1 a contrast k-g or a contrast g-ɣ or perhaps a three-way distinction k-g-ɣ? (2) in the structure V1C1C2V2 where V1 and V2 are specified as above and where C2 is one of the non-syllabic segments [i̯ ʁ v], is there a contrast t-d or a contrast d-ð or perhaps a three-way distinction t-d-ð?

My answers to these questions must be split up in the following way (and these answers must be considered tentative hypotheses rather than established truths):

A. The velars.

a) There is a contrast k-g before liquids followed by a full vowel, cp. mikroskop ('microscope'), lukrativ ('lucrative'), acryl ('acrylic'), cyklamen ('cyklamen'), cyklon ('cyclone') [mikʁoˈsgoʔb, ˈlukʁɑˌtiu̯ʔ, aˈkʁyʔl, syˈklæːmən, syˈkloʔn] vs. agrar ('farmer'), hygrometer ('hygrometer'), agglutinerende ('agglutinating'), Tekla, Hekla [aˈgʁɑʔ, hygʁoˈmeʔdɒ, a/ɑglutiˈneʔɒnə, ˈtegla, ˈhɛgla]. The status of [ɣ] in this context is uncertain, cf. the names Sigrid and Børglum.

In di-syllabic trochaic mono-morphemic words, postvocalic [kl] seems to be excluded, cp. the words Hekla, Tekla, cyklus ('cycle') (with [gl]); it would be interesting to test whether naive speakers of the conservative variant of SD with surfacing [ɣ] would pronounce a fancy name like *Ragla as [ˈʁɑɣla] or [ˈʁɑgla]. (My own variant of SD (which has a distinct Jutlandic flavour) belongs here, and I would not hesitate to pronounce this word as [ˈʁɑɣla]). It will be discussed below whether such trochaic di-syllabic words (or parts of words) can in any sense be considered phonological units of structural significance.

b) There is a contrast g-ɣ before [n] + full vowel, cp. teknik ('technique'), stryknin ('strychnine'), teknikum ('technical school'), ignorant ('ignorant') [tɛgˈnig, sdʁygˈniˀn, ˈtɛgnikɔm, ignoˈʁɑnˀd] vs. magnet ('magnet'), magnat ('magnate'), magnium ('magnesium'), Ragna, Dagny [mɑɣˈneˀd, mɑɣˈnæˀd, ˈmɑɣˀniɔm, ˈrɑɣna, ˈdɑɣny]; [k] is excluded in this context. These structures are thus clearly at variance with Basbøll's principles.

c) There is not much material on which to base a hypothesis concerning the behaviour of velars between a short vowel and [v] or [i̯] + a full vowel. It would seem plausible, though, to disregard the names Sigvard, Sigvald and to postulate that there is only a contrast k-g in this context, but it is questionable whether [k] is possible in this context in trochaic words (personally, I would pronounce a fancy word like *nokva as [ˈnʌgva]).

B. The alveolars.

a) Before [ʁ] there is a contrast t-d, cp. nitroglycerin ('nitroglycerine'), nitrat ('nitrate') [ˈnitʁoglysəˌʁiˀn, niˈtʁɑˀd] vs. hydrogen ('hydrogen'), hydrat ('hydrate') [hydʁoˈgeˀn, hyˈdʁɑˀd].

b) Before [i̯] the material is sparse. The names Katja and Nadja need not be distinguished (except of course through their initial consonants), but if they are, it is definitely as [ˈkati̯a] vs. [ˈnadi̯a]. Very few, if anybody, would pronounce the latter name [ˈnaði̯a]. A few speakers of Copenhagen standard Danish whom I have interviewed with this in mind would tend to pronounce both names with [-adi̯a] but could distinguish them (in the above mentioned manner) in careful speech. The words adjektiv and, in particular, adjudant and a few more pose the problems discussed under (III) above.

c) Before [v] the situation is probably somewhat different. It is not unlikely that the contrast is d-ð (cf. (III) above), and that [t] is excluded at least in trochaic words.

I conclude that the behaviour of at least the clusters /kn gn/ and perhaps also /kl gl/ is at variance with Basbøll's principles (note, e.g., that cyklist ('bicycle rider') is normally pronounced [sygˈlisd] (the pronunciation [sy/iˈklisd] given in

ODS is extremely rare, as far as I know); this would presuppose phonemic or underlying /sy'glisd/, but the basic forms cykel, cykle ('bicycle', 'to ride a bicycle') are pronounced ['sygəl, 'syglə], presupposing underlying /sykəl, syklə/). If Basbøll's principles are to be amended to cover such facts as the manifestation of /gn gl/ they would probably have to be made sensitive to stress and to the sequences of intervocalic consonants per se.

If the absence of aspirates in many of the above mentioned structures is not accidental, this ought somehow to be stated in a phonology of SD. Within Basbøll's generative framework, such restrictions would probably have to be stated as morpheme structure conditions: a word like bekneb ('trouble') [be'kneˀb] shows that the absence of [k] in words like teknik and stryknin is not due to word structure conditions; but in trochaic structures, the absence of aspirates might be due to such word structure conditions. I have not discussed the labial stops in this connection because they do not participate in the gradation process, except in a few words, under specific phonological and stylistic conditions, cf. Basbøll (1975), but it is highly significant that in di-syllabic trochaic words there is probably no contrast of aspiration in the context V1(-long)__CV2 in labial stops either, irrespective of whether V2 is a full vowel or shwa, cf. that the name Pablo must be phonemically /'pablo/ within Basbøll's framework (just as teknik, teknologi, cyklus must be /tɛ'gnik, tɛgnolo'giˀ, 'syglus/).

The examples in (V) above show that the distinction between full vowels and shwa is not decisive for the behaviour of /a/ and /o/ in the preceding syllable, or rather: even though /a/- and /o/-adjustment before the relevant consonants invariably takes place when the vowel of the following syllable is shwa, it also takes place in many instances when the following vowel is a full vowel. It is probably significant that in most of these cases (e.g. in almost all the examples listed in (V) above), the full vowel is unstressed (see also Brink and Lund (1975, p. 73)). This is also characteristic of the full vowels before which /r/ is treated as "syllable-final", cf. (VI) above.

The words with stød in segments preceding a full vowel (cf. (VII) above) also seriously weaken Basbøll's claim that the distinction between full vowels and shwa is a (directly) conditioning factor with regard to syllabification, at least if the stød is said to belong to the syllable.

The words pingvin and jonglør (vs. lingvist), cf. IV above, give rise to a particular problem: within Basbøll's framework the only way to account for this distinction seems to be to accept /ŋ/ as an underlying (phonemic) segment distinct from /n+ velar obstruent/.

In this section I have only mentioned problems which are not mentioned by Basbøll. I have thus omitted the problems posed by words like ordne ('manage') [ˈɒ:dnə] and tordne ('thunder' (vb.)) [ˈtoɒ̂dnə]; but Basbøll's suggestion that they are underlyingly /ɔrdənə/ and /tordənə/ (1972, p. 200) seems to me somewhat ad hoc: the late shwa-deletion required by such an analysis is not motivated by Danish word structure, cp. vordende ('prospective') [ˈvɒ:dənə]. In fact, I would find it much more meaningful to claim that the words orden ('order') and torden ('thunder(storm)') [ˈɒˀdən, ˈtoɒ̂dən] from which the verbs ordne and tordne are derived have underlying /ɔrˀdn, tordn/ with phonologically motivated shwa insertion (no Danish word can end in [Vɒ̂dn]), whereas the verbs do not need such a rule.

4. Concluding remarks

We may conclude that although Basbøll's attempt to account for consonant gradation and vowel adjustment by reference to "phonological syllables" is in many respects empirically successful, there are quite a few phenomena which cannot be adequately described within this framework unless his (admittedly tentative) principles of syllabification are considerably modified. Although many of the counter-examples mentioned above may be considered marginal in the sense that they are lexically sporadic, I think that most of them are phonologically significant, and at least they give rise to some interesting questions concerning Danish word structure.

As mentioned in section 1, the aim of this paper has been the modest one of pointing to some problematic consequences of Basbøll's principles of syllabification. This may seem somewhat unconstructive, but I have nevertheless refrained from discussing whether or not Basbøll's principles could or should be amended to cover the above mentioned phenomena, because such a discussion would require theoretical considerations (concerning, among other things, the functional and structural status of the syllable) which could be only superficially dealt with within the scope set for this paper.

For the sake of clarity, I shall recapitulate what I consider to be the main results and perspectives of my investigation.

1) It is doubtful whether both vowel adjustment and consonant gradation can be satisfactorily accounted for by reference to phonological syllabification in Basbøll's sense, even if syllable boundaries be inserted quite arbitrarily, cf. the examples in (V) above.

2) In homo-morphemic strings with intervocalic clusters there is a contrast gn-ɣn irrespective of whether the following vowel is a full vowel or shwa (in younger standards, this need not be a problem for the observational adequacy of Basbøll's principles of syllabification, but it ought to be stated somehow in any phonology of SD that [kn] is systematically excluded in such positions).

3) In di-syllabic trochaic words, the occurrence of intervocalic clusters of the type aspirated stop + sonorant (probably apart from [ʁ]) is probably heavily restricted in normal styles, irrespective of whether the following vowel is a full vowel or shwa. The status of [ɣʁ, ɣl, ɣv, ðʁ, ði̯, ðv] in this context is uncertain; most of these clusters may not be systematically excluded before full vowels; Basbøll's analysis presupposes that they are excluded in homo-morphemic strings.

These hypotheses ought to be tested; e.g. it would be interesting to present naive speakers with written fancy words like *Ragla,[1] *kidvas, *kaglo, etc.

4) If the stød is considered a syllable prosody manifested as a glottalization of the first post-peak segment (including the last portion of a long vowel) of a syllable, the syllabification of Malaga, annuum, etc. must probably be /mal$a$ga, an$u$om/ etc. (unless the manifestation of the stød is considered to be independent of phonological syllable boundaries?).

---

1) In fact, the word raglanfrakke ('raglan') is pronounced [ˈʁɑɣlan-] (in younger standards: [ˈʁɑu̯lan-]).

5) It might be of some interest to investigate the possibilities of recognizing other phonologically significant units than segments and syllables. We have seen that di-syllabic trochaic words seem to have rules of their own, and it may be fruitful to consider the Danish word to be more hierarchically structured than usually assumed. What I have in mind is a unit larger than the syllable but smaller than the word. Any word consists of one or more such units, and any such unit consists of one salient syllable or of one such syllable followed by one or more subordinate syllables whose vowels may be shwa or one of the full vowels /a o i y u/ but not /e ɛ ø æ ɔ/. Such a unit would be internally consolidated by certain obligatory structural properties: /a/- and /o/-adjustment, the restricted occurrence of medial aspirated stops before sonorants, the occurrence of at most one stød, and probably some more. According to this conception, variable pronunciations of a word would in some instances be due to different hierarchizations: a-kva-vit [akvaˈvid] or akva-vit [ɑgvaˈvid], cy-klo-tron [sykloˈtʁoˀn] or cyklo-tron [sygloˈtʁoˀn], etc.

I am fully aware that such a description would also have its costs; e.g. it would presuppose the hierarchization of each word to be a phonemic property or at least a property not exclusively predictable from the sequence of segments. Nevertheless, this hierarchical treatment may be worth while exploring.

I think the data and hypotheses presented in this paper deserve consideration in future work on Danish phonology, irrespective of whether or not syllables or larger units are recognized as phonological units in their own right.

References

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

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

Basbøll, H. 1973a: "Notes on Danish consonant combinations", ARIPUC 7, p. 103-142

Basbøll, H. 1973b: Diftonger i dansk, Mimeographed notes.

Basbøll, H. 1973c: Konsonanter I+II, Mimeographed notes.

Basbøll, H. 1974: "The phonological syllable with special reference to Danish", ARIPUC 8, p. 39-129

Basbøll, H. 1975: "On the phonological interpretation of the falling diphthongs in Danish", ARIPUC 9, p. 49-108

Brink, L. and J. Lund 1975: Dansk rigsmål. Lydudviklingen siden 1840 med særligt henblik på sociolekterne i København, Gyldendal, Copenhagen

Hansen, Å. 1956: Udtalen i moderne dansk, Det danske Sprog- og Litteraturselskab, Copenhagen

Hjelmslev, L. 1938: "Neue Wege der Experimentalphonetik", NTTS 2.10, p. 153-194

Hjelmslev, L. 1951: "Grundtræk af det danske udtrykssystem med særligt henblik på stødet", Selskab for nordisk Filologi. Aarsberetning for 1948-49-50, p. 12-24

Hooper, J. 1972: "The syllable in phonological theory", Lg. 48, p. 525-540

Hooper, J. 1976: An introduction to natural generative phonology, Academic Press, New York/San Francisco/London

Jul Nielsen, B. 1968: Et bjerreherredsmål, Akademisk forlag, Copenhagen

Kahn, D. 1976: Syllable-based Generalizations in English Phonology, Indiana University Linguistic Club, Bloomington

Kisseberth, C. 1970: "On the functional unity of phonological rules", Linguistic Inquiry 1/2, p. 291-306

ODS: Ordbog over det danske Sprog, Gyldendalske Boghandel, Nordisk Forlag, Copenhagen

Pulgram, E. 1970: Syllable, Word, Nexus, Cursus, Mouton, The Hague

Rischel, J. 1970: "Consonant gradation: A problem in Danish phonology and morphology", in: Benediktsson, H. (ed.), The Nordic Languages and Modern Linguistics, p. 460-480, Vísindafélag Íslendinga, Reykjavik

Vennemann, T. 1972: "On the theory of syllabic phonology", Linguistische Berichte 18,1, p. 1-18

Vestergaard, T. 1967: "Initial and final consonant combinations in Danish monosyllables", Studia Linguistica XXI, p. 37-66

(ARIPUC 13, 1979)

WORD BOUNDARIES AND Fo PATTERNS IN ADVANCED STANDARD COPENHAGEN DANISH

Nina Thorsen

Abstract: In a previous analysis word boundaries have been denied any influence upon the course of fundamental frequency in the stress groups of Advanced Standard Copenhagen Danish. A small experiment has been carried out, which supports this contention, but which also suggests that in more conservative standards, word boundaries do affect Fo patterns.

1. Introduction

In many languages linguistic stress and Fo (or pitch) are interrelated. This is true of e.g. Danish (Thorsen, 1978), Dutch ('t Hart and Cohen, 1973), English (Fry, 1958; Lieberman, 1960), and Swedish (Bruce, 1977). The nature of this relationship is language and dialect specific and so is probably also the weight which pitch has among other prosodic cues for the perception of stressed vs. unstressed syllables. Further, the domain, i.e. the specific combination of stressed and unstressed syllables, of Fo patterns may vary between languages. Thus, Esser (1978) hypothesizes that in German the word is the unit which governs Fo, whereas in English it is the foot (Halliday, 1967). Bruce (1977) implies that in Swedish the word does not seem to constitute the basis for fundamental frequency patterning, and the analysis reported in Thorsen (1978) renders support for a contention that in Advanced Standard Copenhagen (ASC) Danish, word boundaries are immaterial for the Fo patterning.

For a brief summary of the 1978 analysis, the reader is referred to Thorsen (1979 - this volume, p. 60-65). - The establishment of the stress group, i.e. a stressed syllable and all succeeding unstressed syllables, as a relevant unit for the description of ASC Danish is corroborated by Reinholt Petersen's investigations on intrinsic Fo level differences between vowels (1979, this volume). Further, an analysis of utterances with emphasis for contrast also strongly suggests that word boundaries are "deleted" in the Fo course (Thorsen, forthcoming). However, none of the Danish material so far analyzed has been designed expressly with regard to an interplay between word boundaries and Fo patterns, and conclusive, unambiguous evidence seemed called for. The present experiment was intended as a pilot study.

2. Material, subjects, and registrations

2.1 Material

It is difficult to construct utterances containing stress groups which are different only in the placement of word boundaries. Still, the utterances presented below are advantageous from two points of view: The stress group under scrutiny is the first one in the sentence, thus rendering any Fo differences more noticeable than they would be in later parts of the utterance, because Fo patterns are more elaborate in initial than in final parts of declarative utterances (cf. fig. 1 on page 60, this volume). Further, the stress group contains only one unstressed syllable, and any word boundary signal would thus have only one syllable to manifest itself on, which might make it more easily detectable than in longer tails of unstressed syllables. Thus, we could expect that the extent of the (incomplete) rise-fall exhibited by a single post-tonic syllable (cf. p. 64-65, this volume) be a function of word boundary location, so that higher rises are found when the post-tonic belongs to the same word as the preceeding stressed syllable, lower rises are found when the unstressed syllable is a separate word, and still lower rises (or falls) when it belongs to the next lexically stressed word.

The test sentences are:

(1) Pipiku kukker bedre end gøgen.
[b̥ʰiˈb̥ʰig̊ʰu ˈg̊ʰug̊ʌ ˈb̥ɛð·ʌ n̩ ˈg̊øˀn̩]
(Pipiku calls better than the cuckoo.)

(2) Pipi ku kukke bedre end gøgen.
[b̥ʰiˈb̥ʰi g̊ʰu ˈg̊ʰug̊ə ...]
(Pipi could call better than the cuckoo.)

(3) Pipi kukkukker bedre end gøgen.
[b̥ʰiˈb̥ʰi g̊ʰuˈg̊ʰug̊ʌ ...]
(Pipi calls better than the cuckoo.)

Pipi is, for the purpose of the present experiment, a French girl's name, and thus stressed on the last syllable, whereas Pipiku is an Indonesian girl's name, and stressed on the second syllable (or so the subjects were told). To kukke is to call like a cuckoo. A cuckoo's call is a kukkuk (or a kuk) and the verb form derived from this noun would be to kukkukke. (None of the subjects protested about the unlikeliness of the sentences, but they were, admittedly, all phoneticians and therefore used to all sorts of things.)

The three sentences were mixed with 30 others that served a completely different purpose and they only occurred once each on every two pages of reading material, which further occurred in three different randomizations (each being read twice by each subject), so any direct comparison of the test items was avoided and, likewise, any list reading effect. (However, if subjects cared to think about it, the purpose of the three sentences was of course very transparent.)

2.2 Subjects

Two subjects who also served for the 1978 analysis (NRP, male, and BH, female) recorded the material, and so did the author (NT). All three speak ASC Danish. Further, a male subject (JR) who speaks a slightly more conservative variant of Standard Danish was recorded, but for technical reasons, only three recordings of each sentence were obtained from him.

2.3 Recordings

The recordings took place in a quasi-damped room with semi-professional equipment (Revox A77 tape-recorder, Sennheiser MD21 microphone, larynx microphone) in two recording sessions, spaced at least one day apart.

2.4 Registration and measurements

The recordings were processed by hard-ware intensity and pitch meters (F-J Electronics) and registered on a mingograph (Elema 800). (The intensity curves serve segmentation purposes only.) The signal from the larynx microphone was processed in the hold mode. This, in combination with adjustment of the zero-line to the lower limit of the subject's voice range and full exploitation of the record space of the mingograph galvanometer, yields a good solution of the frequency scale, generally making a measuring accuracy

Table 1

Average values of Fo (in Hz) in the first four vowels in the test sentences and of the distance in time (in cs) of the Fo measuring points from the first stressed vowel (negative in ₒi). Standard deviations are given in parentheses beneath each value.

	Pipiku kuk- ...				Pipi ku kuk- ...				Pipi kukkuk- ...			
	ₒi	ˡi	ₒu	ˡu	ₒi	ˡi	ₒu	ˡu	ₒi	ˡi	ₒu	ˡu
Fo:												
NRP (N=6)	96.3 (3.82)	133.8 (7.65)	145.0 (8.37)	103.8 (1.72)	97.2 (2.40)	137.8 (4.40)	147.3 (3.61)	103.2 (2.14)	99.2 (1.33)	140.3 (3.61)	147.5 (3.15)	105.3 (1.51)
JR (N=3)	97.3 (3.21)	117.0 (4.36)	134.3 (3.79)	101.3 (2.52)	100.3 (4.16)	127.0 (3.46)	133.7 (7.64)	111.0 (6.08)	104.0 (3.00)	134.7 (4.93)	127.7 (8.33)	111.0 (1.73)
BH (N=6/6/5)	216.3 (5.85)	253.7 (4.27)	272.7 (3.27)	229.0 (3.95)	220.0 (8.67)	258.3 (6.50)	279.0 (4.86)	235.0 (6.90)	215.2 (9.12)	252.0 (12.7)	273.6 (8.41)	227.6 (11.9)
NT (N=6)	228.0 (7.38)	263.3 (11.3)	291.0 (11.7)	234.3 (9.24)	233.3 (10.6)	269.0 (10.9)	290.3 (7.53)	239.3 (6.89)	231.7 (9.75)	268.7 (7.45)	292.0 (10.7)	237.3 (7.76)
Time:												
NRP (N=6)	-18.7 (1.21)	0	14.0 (1.41)	35.8 (0.75)	-18.8 (1.72)	0	15.0 (1.26)	37.7 (1.21)	-17.8 (0.98)	0	14.5 (0.55)	36.5 (0.55)
JR (N=3)	-16.0 (2.00)	0	13.0 (1.00)	36.7 (2.52)	-12.7 (1.53)	0	15.7 (3.79)	36.0 (5.29)	-14.3 (0.58)	0	16.0 (1.00)	35.3 (2.08)
BH (N=6/6/5)	-16.2 (1.72)	0	12.8 (1.94)	34.2 (2.71)	-16.0 (2.10)	0	13.3 (1.21)	33.2 (2.04)	-14.8 (1.48)	0	13.8 (1.30)	33.2 (2.39)
NT (N=6)	-16.8 (2.14)	0	16.3 (2.42)	37.3 (3.39)	-16.3 (1.21)	0	14.8 (1.33)	36.0 (0.89)	-17.2 (0.75)	0	14.7 (0.52)	35.7 (0.82)

of 1 Hz possible for males and 2 Hz for females. - The first four vowels in the sentences were measured, as follows: Monotonously falling or rising Fo movements (and this description covers most of the vowels) were measured at a point in time 2/3 of the distance from vowel onset, because this point corresponds to the perceived level pitch of the vowel, if its movement is not perceived (cf. Rossi, 1971 and 1978). The post-tonic vowel is often rising-falling and is then measured at its peak. - The distance in time of each measuring point from the first stressed vowel was also measured.

3. Results

3.1 Fundamental frequency

The averages are given in table 1 (with standard deviations), and in fig. 1 stylized tracings of the Fo course in the first four syllables of the test sentences are depicted. To facilitate the comparison between stress group types and between subjects, the Fo averages were converted to semitones (re 100 Hz) which values were then weighted (by simple multiplication) to make all intervals between the two stressed vowels equal (3 semitones) and then they were normalized (by addition or subtraction) to have the stressed vowels coincide at 6 and 3 semitones, respectively; the durations were normalized to a stressed vowel distance of 38 cs, cf. fig. 2.

It seems clear (particularly from fig. 2) that the three ASC speakers (NRP, BH, and NT) do not distinguish the three stress group types in the Fo course, except that NT has the post-tonic slightly higher in Pipiku kukker ... than in the other two sentences. JR is evidently different, and he follows the pattern outlined in section 2.1: the highest rise to the post-tonic is found in Pipiku kukker ... where the post-tonic belongs to the same word as the preceding stressed syllable, lower in Pipi ku kukke ... where the post-tonic constitutes a separate word, and lowest in Pipi kukkukker ... where the post-tonic belongs to the next stressed word.

Fig. 2 is informative, but the normalization and weighting that it took to produce it may blur some differences of a fundamental kind (at least as far as JR is concerned): It appears from

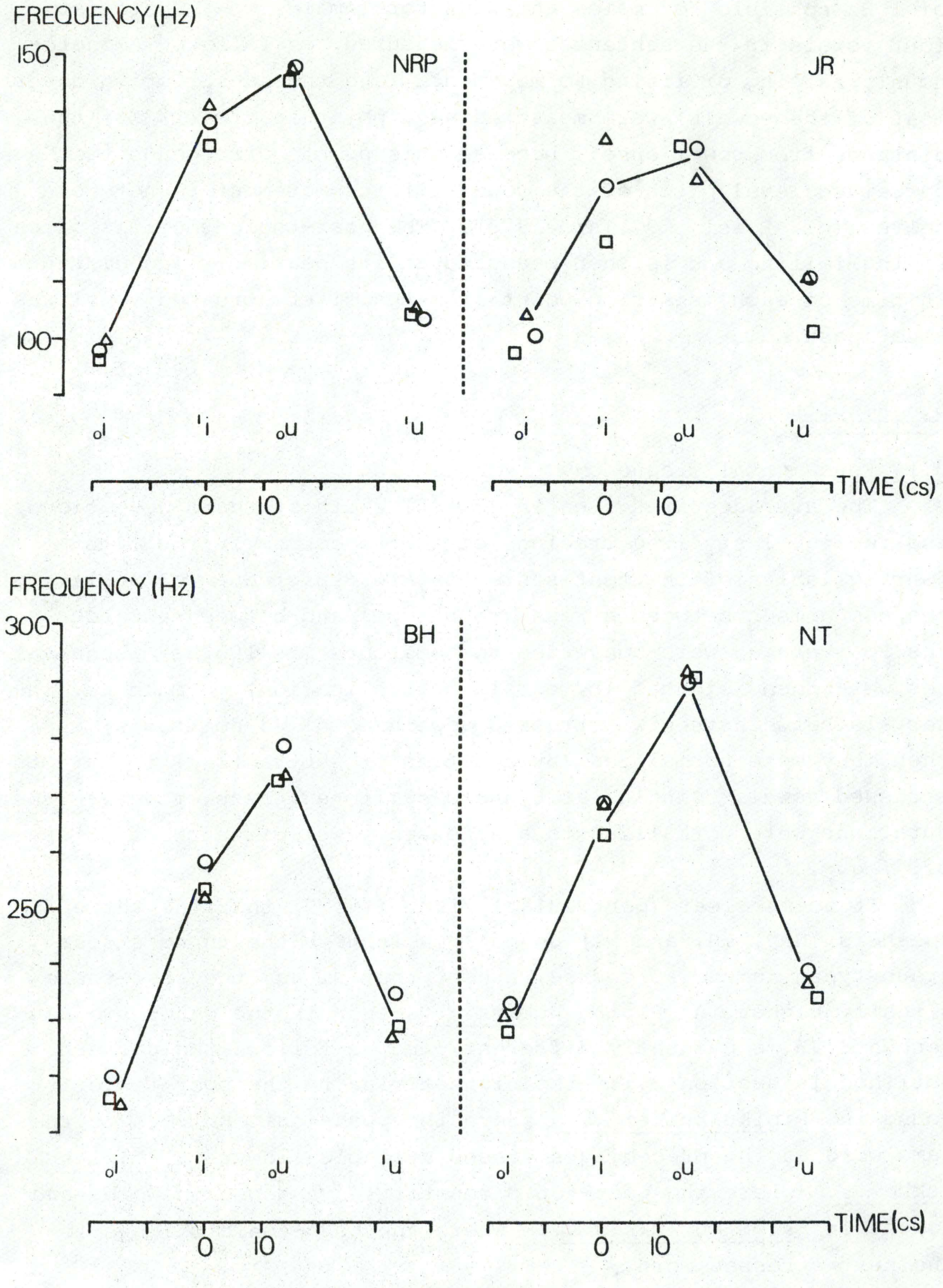
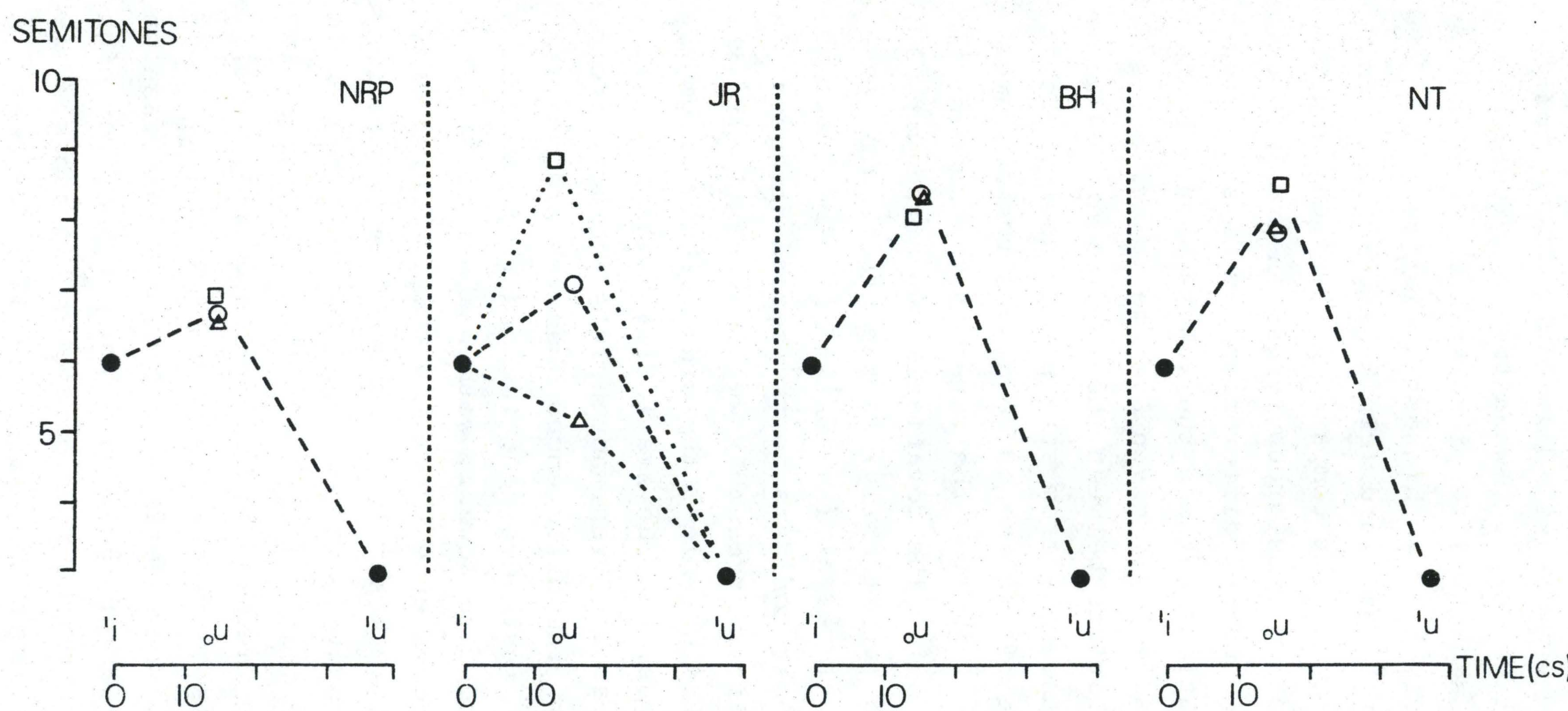


Figure 1

Stylized tracings of the course of fundamental frequency in the first four syllables of the test sentences.

□ : Pipiku kukker bedre end gøgen. O: Pipi ku kukke bedre end gøgen.

△ : Pipi kukkukker bedre end gøgen.



<u>Figure 2</u>

Stylized, normalized, and weighted tracings of the course of fundamental frequency in the first stressed, the post-tonic, and the second stressed syllables of the test sentences. Filled circles pertain to the stressed syllables.

□: Pipi<u>ku</u> kukker bedre end gøgen. O: Pipi <u>ku</u> kukke bedre end gøgen.

Δ: Pipi <u>kuk</u>kukker bedre end gøgen.

fig. 1 (and table 1) that the post-tonic is not the only vowel which varies - so does the level of the first stressed vowel and (to a smaller degree) the second stressed vowel. In particular, JR's patterns in fig. 1 make one suspect that the level of [ˈi] and that of [$_\circ$u] covary: [ˈi] is highest and simultaneously [$_\circ$u] is lowest (relative to [ˈu]) when the post-tonic belongs to the succeeding stressed vowel (Pipi kukkukker ...), and [ˈi] is lowest and simultaneously [$_\circ$u] is highest (relative to [ˈu]) when [ˈi] and [$_\circ$u] belong to the same word (Pipiku kukker ...). This means that it is not sufficient to just compare the average values of the post-tonic vowel across stress group types, in order to establish to what extent the observed differences may be statistically significant: the relations between the vowels should be taken into account. - In table 2 are given average values and standard deviations of the relations between the vowels (excluding the sentence initial one): the rise from stressed to post-tonic (ˈi-$_\circ$u), the fall from post-tonic to the succeeding stressed vowel, numerically ($_\circ$u-ˈu), and the fall from the first to the second stressed vowel, numerically (ˈi-ˈu). The covariation observed with JR, cf. above, may be expressed (arbitrarily) as the fraction of the $_\circ$u-ˈu-interval to the ˈi-ˈu-interval ($_\circ$u-ˈu/ˈi-ˈu): the fraction decreases the closer the post-tonic is to the second stressed vowel and/or the higher the first stressed vowel is (relative to the second stressed vowel). In table 3 these entities are compared across stress group types, and the statistical significance, if 10% or better, is indicated. Even though JR's averages are based on only three recordings, the differences observed are statistically significant in most instances, and the order of the stress group types, from highest to lowest, with respect to the four parameters is as expected: Sentence 1 (Pipiku kukker ...) scores highest and sentence 3 (Pipi kukkukker ...) scores lowest with respect to the rise to the post-tonic (ˈi-$_\circ$u), the fall from the post-tonic ($_\circ$u-ˈu), and the fraction ($_\circ$u-ˈu/ˈi-ˈu). The order is reversed, as it should be, in the inter-stress interval (ˈi-ˈu), except that the difference is nearly nil between Pipi ku ... and Pipiku ... . - The other subjects do not behave in a fashion uniform with JR: the order of the stress group types from highest to lowest, with respect to the four parameters in table 3, is only identical to that of JR in two instances (NRP: ˈi-$_\circ$u and $_\circ$u-ˈu/ˈi-ˈu).

Table 2

Average values of the Fo relationships between the first stressed, the post-tonic and the second stressed vowel in the test sentences, i.e. the rise to the post-tonic (ˈi-˳u), the fall from the post-tonic to the succeeding stressed vowel, numerically (˳u-ˈu), the fall from the first to the second stressed vowel, numerically (ˈi-ˈu), and the ratio between ˳u-ˈu and ˈi-ˈu. Standard deviations are given in parentheses beneath each value. Note that the fraction is not calculated from the average values of its nominator and denominator but from the raw data, which is why there are apparent discrepancies between this fraction and one calculated from ˳u-ˈu and ˈi-ˈu as they appear in the table.

	Pipiku kuk-				Pipi ku kuk-				Pipi kukkuk-			
	ˈi-˳u	˳u-ˈu	ˈi-ˈu	˳u-ˈu / ˈi-ˈu	ˈi-˳u	˳u-ˈu	ˈi-ˈu	˳u-ˈu / ˈi-ˈu	ˈi-˳u	˳u-ˈu	ˈi-ˈu	˳u-ˈu / ˈi-ˈu
NRP (N=6)	11.2 (7.44)	41.2 (6.97)	37.0 (6.63)	1.43 (0.391)	9.5 (3.99)	44.2 (3.31)	34.7 (5.54)	1.29 (0.173)	7.2 (2.71)	42.2 (3.31)	35.0 (4.15)	1.21 (0.090)
JR (N=3)	17.3 (6.51)	33.0 (2.65)	15.7 (6.66)	2.48 (1.33)	6.7 (7.77)	22.7 (3.21)	16.0 (5.00)	1.55 (0.639)	-7.0 (8.89)	16.7 (6.66)	23.7 (4.51)	0.755 (0.347)
BH (N=6/6/5)	19.0 (4.69)	43.7 (4.80)	24.7 (5.01)	1.82 (0.327)	20.7 (6.28)	44.0 (3.10)	23.3 (5.75)	1.98 (0.523)	21.6 (5.55)	46.0 (9.49)	24.4 (9.84)	2.12 (0.787)
NT (N=6)	27.7 (5.72)	56.7 (8.45)	29.0 (7.13)	2.01 (0.403)	21.3 (9.00)	51.0 (8.17)	29.7 (7.31)	1.80 (0.454)	23.3 (7.66)	54.7 (12.37)	31.1 (9.18)	1.81 (0.377)

Table 3

A comparison across stress group types of the Fo relationships tabulated in table 2. The stress group types are designated "1" (Pipiku kukker ...), "2" (Pipi ku kukke ...), and "3" (Pipi kukkukker ...). They are listen in decreasing order of magnitude, from left to right, and the statistical significance, if 10% or better, is indicated in parentheses.

ˈi - ˳u

NRP 1 > 2 > 3

JR 1 > 2 > 3: 1>2 (10%); 2>3 (10%); 1>3 (1%)

BH 3 > 2 > 1

NT 1 > 3 > 2: 1>2 (10%)

˳u - ˈu

NRP 2 > 3 > 1

JR 1 > 2 > 3: 1>2 (5%); 1>3 (5%)

BH 3 > 2 ≃ 1

NT 1 > 3 > 2

ˈi - ˈu

NRP 2 ≃ 3 > 1: 3>1 (10%)

JR 3 > 2 ≃ 1: 3>2 (10%); 3>1 (10%)

BH 1 ≃ 3 > 2

NT 3 > 2 ≃ 1

˳u - ˈu / ˈi - ˈu

NRP 1 > 2 > 3

JR 1 > 2 > 3: 1>3 (5%); 2>3 (10%)

BH 3 > 2 > 1

NT 1 > 3 ≃ 2

We may regard the sentence numbers as ranks (and reverse them for the inter-stress interval, ˈi-ˈu) and calculate the Kendall coefficient of concordance (W) between the "rankings" of NRP, BH, and NT across the four parameters: W = 0.007, i.e. the concordance is nil.

The statistical analysis confirmed the initial observation, that the location of word boundaries is not reflected in the Fo patterns of the ASC speaking subjects, but they are very manifest with JR. JR was brought up on Funen but the word boundary/Fo relations found with him are hardly a Funish influence, because this is not otherwise perceptible, but they may be a characteristic of the conservative variant of Standard Danish. - To resolve this issue would take a separate investigation. (On the other hand, if different varieties of Danish turn out to be fundamentally different with respect to a relation between word boundaries and Fo patterns, then this difference may be confined to neutral speech: When emphasis for contrast occurs on a word in the utterance, JR seems to delete the word boundaries in much the same fashion as do ASC speakers, cf. Thorsen, forthcoming.)

3.2 Duration

Differences in timing may of course exist independently of the lack of differences in fundamental frequency (co-exist in JR's case). In table 4 the durational differences tabulated in table 1 are compared across the three stress group types. Hypotheses about these differences might run as follows: (1): The inter-stress interval (ˈi-ˈu) should be longer in Pipi ku kukke ... where two word boundaries intervene, than in the other two types. (2): The distance from stressed to post-tonic (ˈi-$_{\circ}$u) should be smaller in Pipiku kukker ... where no word boundary intervenes, than in the other two types. (3): The distance from post-tonic to the succeeding stressed vowel ($_{\circ}$u-ˈu) should be smaller in Pipi kukkukker ... where no boundary intervenes, than in the other two types.

Ad (1): NRP is the only one who has a longer inter-stress distance in Pipi ku kukke ..., but with him this distance is also significantly longer in Pipi kukkukker ... than in Pipiku kukker ... for which it is hard to find an explanation. Ad (2): Three subjects have a shorter distance to the post-tonic in Pipiku kukker ..., but only significantly so with JR; with NT

Table 4

A comparison across stress group types of the durational differences tabulated in table 1. The stress group types are designated "1" (Pipiku kukker ...), "2" (Pipi ku kukke ...), and "3" (Pipi kukkukker ...). They are listed in decreasing order of magnitude, from left to right, and the statistical significance, if 10% or better, is indicated in parentheses.

ˈi - $_{\circ}$u

NRP	2 > 3 > 1	
JR	3 > 2 > 1:	3>1 (2.5%)
BH	3 > 2 > 1	
NT	1 > 2 > 3:	1>3 (10%)

$_{\circ}$u - ˈu

NRP	2 > 3 > 1	
JR	1 > 2 > 3:	1>2 (5%); 1>3 (2.5%)
BH	1 > 2 > 3:	1>2 (5%); 1>3 (2.5%)
NT	2 > 1 = 3	

ˈi - ˈu

NRP	2 > 3 > 1:	2>3 (5%); 2>1 (0.5%); 3>1 (5%)
JR	1 > 2 > 3	
BH	1 > 2 = 3	
NT	1 > 2 > 3:	1>3 (10%)

the order is reversed and, further, this distance is significantly longer in in Pipiku kukker ... than in Pipi kukkukker ..., which seems counter-intuitive. Ad (3): BH and JR have significantly shorter distances from post-tonic to the succeeding stressed vowel in Pipi kukkukker ... but it is also significantly shorter in Pipi ku kukke ... than in Pipiku kukker ..., for which it is hard to find an explanation. NRP and NT differ from this pattern.

On the whole, the durational differences cannot be said to reflect word boundary location in any succinct way, except maybe with JR who has a shorter distance from stressed to post-tonic in Pipiku kukker ... and a shorter distance from post-tonic to the succeeding stressed vowel in Pipi kukkukker ... .

4. Conclusion

The results of the present experiment confirm the results of the 1978 analysis, that in Advanced Standard Copenhagen Danish word boundaries leave no trace in the course of fundamental frequency in the stress group - but they also suggest that this may not be generalizable to all varieties of Danish. If this suggestion is corroborated by further analyses, then the definition of the stress group, as it applies to ASC Danish, namely 'a stressed syllable plus all succeeding unstressed syllables, irrespective of intervening word boundaries' may not be entirely adequate, because it cannot be said to be the unit which governs the predictable, recurrent, and in certain respects invariant, Fo patterns (as is the case in ASC). (But then again, there may be other reasons to stick with the stress group: Peter Holtse (personal communication) has performed an analysis on the variation of vowel duration, with speakers from various dialects,and found that this variation seems to be determined by the stress group - in the ASC-sense.)

The investigation was intended as a pilot study. However, I think that the results are clear enough for the ASC speakers to be conclusive - but they may serve as a basis for further analyses of other kinds of Danish.

References

Bruce, G. 1977: "Swedish word accents in sentence perspective", Travaux de l'Institut de Linguistique de Lund 12, p. 1-155

Esser, J. 1978: "Contrastive intonation of German and English. Problems and some results", Phonetica 35, p. 41-55

Fry, D.B. 1958: "Experiments in the perception of stress", L&S 1, p. 126-152

Halliday, M.A.K. 1967: Intonation and Grammar in British English, Mouton, The Hague

't Hart, J. & A. Cohen 1973: "Intonation by rule: a perceptual quest", JPh 1, p. 309-327

Lieberman, P. 1960: "Some acoustic correlates of word stress in American English", JASA 32, p. 451-454

Reinholt Petersen, N. 1979: "Variation in inherent Fo level differences between vowels as a function of position in the utterance and in the stress group", ARIPUC 13, p. 27-57 (this volume)

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

Rossi, M. 1978: "La perception des glissandos descendants dans les contours prosodiques", Phonetica 35, p. 11-40

Thorsen, N. 1978: "An acoustical investigation of Danish intonation", JPh 6, p. 151-175

Thorsen, N. 1979: "Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish", ARIPUC 13, p. 59-85 (this volume)

Thorsen, N. (forthcoming) "Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish" (an enlarged version of Thorsen, 1979), in preparation.

(ARIPUC 13, 1979)

A SKETCH OF THE HISTORY OF PHONETICS IN DENMARK UNTIL THE BEGINNING OF THE 20TH CENTURY

Eli Fischer-Jørgensen

1. Introduction

On the occasion of the 500th anniversary of the University of Copenhagen, I was asked to write a brief contribution to the jubilee publication on the history of phonetics in Denmark. The present paper is a considerably enlarged version of the first part of that contribution.

Danish phonetics from the end of the 16th century till around 1900 cannot be seen as a continuous development. Some of the early phoneticians did not even know each other, and were not read by later scholars either. From Høysgaard on, it is possible to speak of a tradition in the sense that everybody knew most of his predecessors, but it was not a closed tradition. They were all more influenced by what was going on in other European countries than by their Danish predecessors or teachers, and each of the scholars to be mentioned had his own marked personality and his own approach. It will therefore be a story of a number of individual scholars and their works rather than a history of the development of a discipline.

2. Isolated precursors

The medieval grammarians were much more interested in the sign function than in the material manifestation of the signifier. It was not until the time of the Renaissance with its emphasis on empirical observation and its growing interest in the mother tongue that the sound matter of language was given more attention.

2.1 Jacob Madsen Aarhus (1538-1586)

One of the most important books on phonetics in the 16th century (De literis libri duo) was written by a Dane, Jacob Madsen from Aarhus (Jacobus Matthiae Arhusiensis). Jacob Madsen went to school in Aarhus, and then studied at the University of Copenhagen 1559-65. He received the baccalaureus philosophiae degree in 1563. 1565-66 he was rector at Aarhus. In 1566 he went to Germany, where he spent eight years studying theology, philosophy, history, law, medicine, and, in particular, languages at the universities of Wittenberg, Leipzig and Heidelberg. In 1674 he returned to Copenhagen, where he was appointed professor, first of Latin, then of Greek, and finally of theology. The book was published in the year of his death, 1586. A large part has been republished by Techmer in Internationale Zeitschrift für Sprachwissenschaft V, 1890, and a Danish translation of the whole work with an introduction and comments was published in 1930-31. As was normal at that time, Jacob Madsen uses the word "litera" (letter) to indicate a general concept covering the sound, the written symbol, and its name. But Jacob Madsen considers the sound to be central in the concept of "litera", whereas the written symbol and the name are mere accidentia. He emphasizes that in order to explain the sounds, it is necessary to study their production and that it is important to find out what is universal in human speech sounds, and he starts out with a description of the human speech organs, which was quite uncommon at that time. He is the first to have set up the sounds in an articulatory system that was meant to be universal.

Jacob Madsen is strongly influenced by the French philosopher Petrus Ramus and his Scholae grammaticae (1569) and, as was usual practice at that time, he simply copied large parts of that work word by word, also the learned quotations from Latin authors, but it is evident that he also had a good, direct knowledge of the grammatical literature.

Jacob Madsen's vowel system is taken over from Ramus. Ramus distinguishes two types of vowels: a e i and o u y. The first are called "diducta", the second "contracta", indicating that the lips are separated in the first and constricted in the latter group. Madsen sticks to the two groups but calls them "tongue vowels" and "lip vowels" (linguales and labiales), and he gives a slightly more precise description at some points. He follows Ramus in indicating that the tongue is gradually raised from *a* to *e* to *i*, but adds that for *i* the upper lip is retracted in a slight smile (this was, however, not his own invention - older authors have also distinguished the smiling *i* from the "tragic" *u*). As for *o u y*, Ramus gives a very incorrect description of the position of the tongue, mentioning that it is lowered for all three but most for *y*. Jacob Madsen is more cautious. He states that it is not possible to observe the position of the tongue for these vowels, but that it is sufficient to describe the position of the lips, which are more constricted and protruded from *o* to *u* to *y*. *ε* is called a more coarse *e*, and *ø* a sort of *o*.

Jacob Madsen shows more originality in the description of the consonants. Ramus has the traditional distinction between mutae (stops) and semivocales (the other consonants). Jacob Madsen attacks this distinction which, he says, is simply based on the position of the vowel in the name of the letters (*pe*, *te*, but *em* and *ef*). The consonants must be divided according to the articulatory organs involved.

He first makes a distinction between labiales and linguales (as in the vowels). The lip consonants are the bilabials *p*, *b*, *m*, and the labiodentals *f* and *v*. The others are lingual sounds. His further divisions of this group are peculiar. He makes the first division between linguopalatal and linguodental sounds. In the linguopalatal sounds the tongue tip approaches the palate. In the "mobile" linguopalatals (*s* and *r*), the tongue does not touch the palate but remains free and mobile so that the airflow can pass out, whereas in the "fixed" ones (*n* and *l*) the tongue tip touches the palate and stops the airflow, so that it must escape through the nose or at the corners of the mouth (but a difference between *l* and *n* in this respect is not specified). In the linguodental consonants the tongue tip touches the teeth, either the upper teeth, as in *t* and *d*, or the lower teeth as in *k*, *h*, *j* and *g* (*h* is called a weakly articulated *k*). In *j* and *g* the tongue is more fronted than in *k* and *h*, and the tongue tip is pressed firmly against the lower teeth.

Jacob Madsen has a fine observation concerning the difference between *f* and *v*. He remarks that in *v* the inner edge of the lower lip touches the upper teeth lightly, whereas in *f* the outer edge of the lower lip is pressed against the teeth and the air is blown out vigourously. He has also observed that Danish has aspiration after short utterance-final vowels (dah, dih).

He is not aware of the difference between nasal and oral sounds, and thinks that air escapes through the nose in most stops, but he mentions that the nose channel is used for breathing "in order to prevent the mouth from always being open and thus disfiguring the face."

In the first part of his work he gives the "true" description of sounds, in the second part he attacks the "false" descriptions, among them the division in mutae and semivocales, the description of h as an aspiration and not a true consonant, and above all, the description of diphthongs as consisting of two vowels which, he says, is in contradiction to the true definition of vowel, consonant, and syllable: A vowel is a sound which is pronounced by its own force, whereas a consonant is pronounced together with a vowel, and the syllable is a combination of a vowel and a consonant or may consist of a vowel alone. Ramus tries to escape the contradiction by defining the vowel as a sound which is "capable of" forming a syllable alone, but this is a bad type of definition. A definition should be absolute, and the syllable should not be a presupposed concept. Therefore Jacob Madsen will describe diphthongs as consisting of a vowel and a consonant. And he thinks that the whole confusion has arisen because the letters i and u have not been distinguished from j and v. He therefore writes the Danish diphthongs with j and v, and this way of writing them has been preserved in present-day Danish orthography.

Jacob Madsen was not only known in Denmark: he also influenced e.g. the Dutch scholar Petrus Montanus van Delft and his remarkable "Spreekkonst", which was published in 1635.

2.2 Jens Pedersen Høysgaard (1696-1773)

The Danish grammarians of the 17th century did not give any noteworthy contributions to phonetics. But in the 18th century two scholars must be mentioned for their original contributions.

One of them was Jens Pedersen Høysgaard. He is generally considered to be the greatest Danish linguist before Rask. He was born in Jutland, near Aarhus, and studied in Copenhagen, where he acquired the baccalaureus degree, which gave admittance to teach at larger schools. He preferred, however, the very modest job of third porter at the University, which consisted of cleaning the class rooms and reading a passage from the Bible during meals. It was a rather light job since there were only three class rooms at the university; and since they could not be heated, the professors generally gave their lectures at home. Høysgaard even declined to be promoted to porter of the next degree, because he wanted to have time for his studies. Later, however, he got a somewhat better post as sexton at the university church.

Høysgaard's main work is a Danish syntax, but he also gave a very remarkable contribution to Danish phonetics in his *Accentuered og ræsonnered Grammatica* ('Accented and reasoned Grammar'), 1747.

Høysgaard sets up a very original vowel system:

		i				(IPA:
	e		y			æ = ε
æ		ö		ø		ö = œ
a	å		o		u	a = ɑ
						å = ɔ)

It is based partly on alternations between vowels in related words, but at the same time the vertical dimension has a striking similarity to a dominating auditory dimension, which appeared in my experiments with perceptual dimensions of Danish vowels[1]: [i ey εøæ œuoɔɑ], the only difference being that *œ* was generally closer to the last group.[2]

His consonant system is more traditional, with a first division into mutae and semivocales, but he also mentions that *bdgv* are "weak" compared to *ptkf*. He has also given a system according to place of articulation.

---

1) Eli Fischer-Jørgensen: "Perceptual dimensions of vowels", *To Honor Roman Jakobson*, 1967, p. 667-671.

2) In my list, [æ] and [ɑ] are variants of the phoneme /a/. At Høysgaard's time, /a/ was always pronounced [ɑ].

The real merit of his book is, however, the description of the Danish prosodic system. Earlier grammarians often showed a total confusion of length and stress, mainly because Latin metrics was based on length, but in reading aloud, long syllables were rendered as stressed syllables. Høysgaard avoids this confusion and sets up four prosodic properties:[1]

1) accent (or tone). This is explained as high or low pitch. In a later work he states expressly that pitch does not change the word meanings in Danish, and he therefore does not treat the topic in any detail.

2) "Tonehold", which is defined as the stress (emphasis) one syllable has compared to another. He distinguishes three degrees: strong stress, weak stress, and lack of stress, e.g. Forhuset (pronounced [ˈfɒˌhuʔsəð]). But he adds that besides stress they also differ by the amount of time from the start of one syllable to the next, and that the syllable with strong stress normally has the highest pitch and the weak syllable the lowest pitch. Thus, Høysgaard has seen that stress is manifested not only dynamically but also by duration and pitch, whereas most later phoneticians, among them Jespersen, only talk of "expiratory" stress (which was opposed to musical accent). Høysgaard's description is based on a much finer observation of the real pronunciation as it was at that time, particularly in Jutland.[2]

3) Quantity (or length in verse). He remarks that this is in fact the same as stress, since Danish metrics is based on stress. "Tonehold" is the stress a syllable has according to its nature. "Quantity" is the stress it may have in verse, so-called "long" or "short". -- Here he uses the traditional terminology of metric descriptions, but his distinctions are quite clear.

---

1) His terminology varies somewhat in different works.

2) In modern Copenhagen pronunciation, the stressed syllables have low pitch.

4) The fourth prosodic category is called "<u>åndelav</u>" (approximately "breathing type") and is defined as "some peculiar thrusts or puffs of breath with which the syllables or their vowels and other sounds are pronounced". This category comprises a combination of length and "stød" and thus has four members, viz. syllables with short or long vowels, with or without stød. Høysgaard proposes to indicate them orthographically by accent marks. The two types of "åndelav" without stød are called "det kort-jævne" ('the short-even'), e.g. <u>Stàd</u> [sd̥ɑð], and "det lang-jævne" ('the long-even'), e.g. <u>Bro̓r</u> [bro:r], and it is said that the short-even type permits the breath to pass smoothly and the long-even is drawn out. The two types with stød are called "det stødende" ('the thrusting'), e.g. <u>Sáng</u> [sɑŋʔ], and "det dobbelte" ('the double'), e.g. <u>Flôr</u> [flo:ʔʁ]. "Det stødende åndelav" is said to stop the breath flow, whereas "det dobbelte" first stops the breath and then lets it continue. He is well aware of the common feature of the two latter "åndelav" and sometimes mentions both as "stødende". This is the first time the Danish "stød" has been described in the grammatical literature.[1]

The "stød" is also compared to "a very little hiccup", and in a later work Høysgaard says that the pharynx is closed and stops the breath. He thus seems to assume that there is a complete closure, which nowadays is only found in very emphatic forms of the "stød" in standard Danish and as the normal manifestation in some dialects, but it is not improbable that it was produced with a real closure at Høysgaard's time, cp. that Otto Jespersen also describes the Danish stød as a glottal stop.

Høysgaard did not only set up the prosodic system and describe the pronunciation of the stød, he also gave detailed rules for its use in Danish words. His system was used in the grammatical literature until the middle of the 19th century. Later phoneticians have preferred to set up vowel quantity and stød/non-stød as two distinct categories.

---

1) Jespersen (1897-99, p. 298) mentions that a Swedish author in the 16th century describes the Danish language in the following way: "they press the words out as if they were going to cough", and he thinks that this impression might have been due to the 'stød'.

Høysgaard has also written some pamphlets on Danish orthography (1743a and b). He has introduced the sign ö for /œ/, a sign which has been adopted by e.g. Rask and Jespersen, but not in the official orthography. Nobody has followed his proposal to use accents for the four "åndelav", i.e. for vowel length and stød.

2.3 C.G. Kratzenstein (1723-95)

C.G. Kratzenstein is the other scholar from the 18th century whose contribution to phonetics deserves special mention.

C.G. Kratzenstein's father was German, but his mother Danish. He studied medicine in Halle, where he took his degree in 1748. He then got a chair in St. Petersburg, where he succeeded in improving various nautical instruments. In 1753 he was appointed professor of experimental physics and designated professor of medicine at the university of Copenhagen, and from 1763 he was ordinary professor of medicine. He was a highly esteemed scholar, and during four periods he was rector of the University.

Through his physical studies Kratzenstein got interested in vowel production, and in 1780 he wrote a prize essay for the Imperial Academy of St. Petersburg. The requirement was a description of the vowels a e i o u and the construction of a machine that could produce them. Kratzenstein won the prize, and his essay (written in Latin) was published in 1781 in St. Petersburg. A French translation was published in a physics journal in Paris (1782).

Kratzenstein has a long discussion of voice production in the human larynx and how it can best be imitated. Among others, he quotes the famous physiologist Haller, who describes how the vocal cords are set in vibration by the air flow from the lungs, and how they give a higher tone when they are tight and tense than when they are relaxed. However, Kratzenstein does not believe that the small vocal cords are able to vibrate at the lower frequencies of the human voice, nor that they can produce a tone of sufficient loudness. He thinks that they rather function as a valve, and that the vibrations are mainly produced by the epiglottis, at least for the vowels a,e,o, whereas the vocal cords play a greater role for u and i. As is well known, this description has not been confirmed by later studies, but Kratzenstein accounts correctly for differences in fundamental frequency. As for the differences in vowel quality, he thinks that they are due to the different forms and openings of the cavities above the larynx, which shape the passing wave differently, and to interference between parts of the wave.

The really interesting part of his essay is, however, his description of the position of the tongue, lips and teeth in different vowels. This description is quite up to date and far ahead of his time. He has measured the exact distances in inches and lines between tongue and palate and between upper and lower teeth as well as the lip opening horizontally and vertically, for all vowels. He has also investigated the position of the larynx by palpation with his fingers, and the position of the epiglottis by his index finger. He thus finds that the larynx is elevated for i, and that the tongue is retracted and elevated for o and u, and more so for u than for o. He is apparently the first European phonetician to have given a correct description of these back rounded vowels. In the 18th and far into the 19th century they were normally described as "labial vowels", and the tongue position was not mentioned, or was assumed to be as in a. It is also interesting that he has found a smaller distance between the teeth for e and i than for o and u (and particularly for i compared to u). (I have found this observation confirmed by informal questioning, at least for some speakers, including myself). - But although Kratzenstein gives exact measures for the normal pronunciation, he adds that these measures are not absolute. It is, e.g. possible to say a e i with the same distance between the teeth.

On the basis of these investigations he has constructed a machine for vowel synthesis, in which the voice is produced by means of a metal reed which he compares to the epiglottis, and a pair of bellows, to which are added cavities of different form for the different vowels (see Fig. 1). - He reports that the machine could say the vowels i e a o u and the words mama and papa. This seems to have been the first serious attempt at constructing a speech synthesizer. Von Kempelen's machine is slightly later (1791). Von Kempelen's could also pronounce consonants, but his description of vowel formation is much less accurate than that of Kratzenstein.

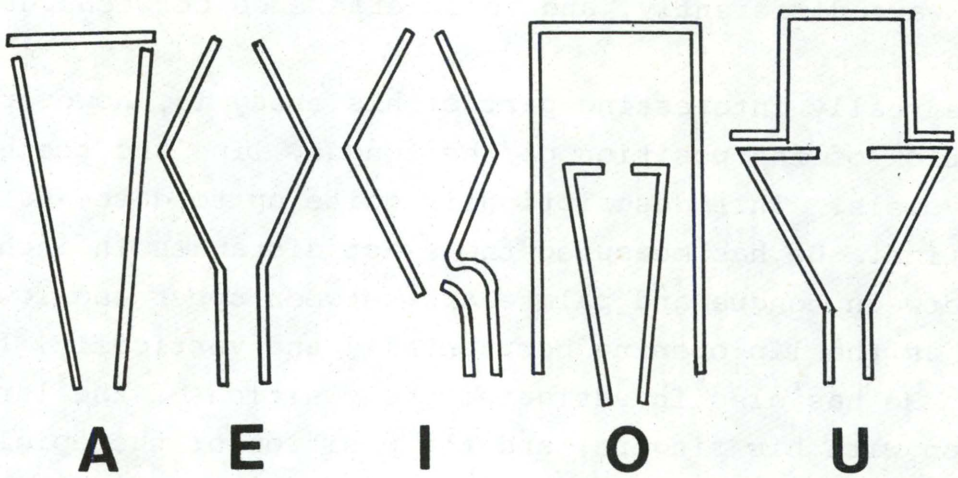


Figure 1

The cavities producing the vowels
a e i o u in Kratzenstein's synthesizer.

Kratzenstein is very rarely cited in the phonetic literature. The fact that his prize essay was published in St. Petersburg may be one of the reasons, and the French translation in a physics journal apparently went unnoticed by phoneticians. Jespersen, who treats von Kempelen at some length (1897-99, p. 24-29), does not know Kratzenstein. But he is mentioned by Wheatstone (1879), who criticizes him for not being aware of the decisive importance of the length of a resonating tube, and by Ungeheuer (1962).[1]

Høysgaard and Kratzenstein, the two Danish scholars who have given essential contributions to phonetics in the 18th century, both worked at the University of Copenhagen, and they had administrative relations (we know that Kratzenstein got into trouble with his colleagues because he had reduced the dues Høysgaard had to pay to the University as a sexton), but the physics professor probably did not know the grammatical works of his porter, and his own treatise on vowels was not published until after Høysgaard's death, and thus it is improbable that they were ever aware of the fact that they had common interests.

---

1) Jørgen Rischel first drew my attention to Kratzenstein's Latin essay. But it was not until quite recently that Preben Dømler drew my attention to a paper by V. Aschoff in Acustica 1979, where the French translation is quoted. I am grateful to professor I. Fónagy for having produced a copy for me from the Bibliothèque Nationale in Paris.

Rask's main work only belongs to the history of phonetics in so far as it is to a large extent based on sound comparisons; some of his other works are closer to descriptive phonetics. He collected a long word list from his Funish dialect which, together with his notes on Funish pronunciation and grammar, and a text in his own phonetic transcription, was edited by Poul Andersen (1938). His transcription shows a fine and very reliable intuition concerning the phonological system of his dialect and, according to Poul Andersen (1937), even modern phonetically well trained dialectologists with a profound knowledge of the dialect could not do better. His descriptions of the standard language were less reliable, probably because his pronunciation had retained many Funish features.

His interest in the Danish sound system was closely connected with his interest in the principles of orthography. As far back as 1808 he constructed a new, radically changed orthography for Danish, which he used in his own writings, and in a manuscript dating from 1808-09 he claims that there should be one and only one letter for each sound (evidently in the sense of phoneme). In 1826 he published a comprehensive treatise, Forsøg til en videnskabelig dansk retskrivning ('An attempt at a scientific theory of Danish orthography'). His claims are here less radical, but he maintains among other things the distinction of ø and ö and the use of å instead of aa, proposals which he took over from Høysgaard. In his description of the Danish system he also takes over Høysgaard's "åndelav", although with some modifications.

Rask has systematized Danish vowels in two different ways. The system he sticks to in his book on Danish orthography (1826) is the following:

a	æ	(IPA:
å	ö	a = ɑ
o	ø	æ = ɛ
u	y	ö = œ
e	i	å = ɔ)

He motivates it (as did Høysgaard) by alternations between related words, caused by ablaut and umlaut, - e.g. kamp-kæmper ('(a) fight' - (he) fights'). The two columns are called hard and weak, respectively, and these designations are motivated by the palatalization of k and g before the weak series. This was true of the pronunciation at Rask's time, except that k and g were also palatalized before e, as objected by his critics. Rask's answer was that they were not palatalized before -e in weak syllables, e.g. hakke (he would not admit that the pronunciation in weak syllables was [ə]). He would therefore write the palatalized k before e as a cluster, kj-, as before "hard" vowels (kjole), whereas the automatic palatalization of k and g before "weak" vowels should not be indicated in the orthography. In a letter of 1819 he writes that in setting up this system he was inspired by the Finnish vowel system (where a,o,u alternate with ä,ö,ü according to the rules of vowel harmony, and i,e do not change). In this way he achieves a regular and harmonious system which, particularly in his later years, was a very essential motive. His consonant system is less interesting. It is mainly based on that of his teacher, S. Bloch. In his prize essay (1818) he distinguished between mutae and liquidae, and the mutae were set up in a two-dimensional system according to place and manner of articulation.

	hard	weak	breathy
labials	p(v)	b(w)	f
linguals	t	d	þ
palatals	k(j)	g(j)	x,h

In his later book (1826) he goes back to an older system, closer to Bloch:

palatals	k	k’	g	g’	g̃
linguals	t	(þ)	d	ð	n
labials	p	f	b		m
semivowels		j	v w		
liquids	l	r			
sibilants	s				
breathing sound	h				

k’, g’, g̃ indicate [x] [ɣ] [ŋ], respectively. The designations are phonetically correct (perhaps with the exception of the "linguals"), but as a system it is not quite satisfactory, since the vertical dimension is a mixture of place and manner of articulation (see Marie Bjerrum (1959), p. 197-203).

3.2 Jacob Hornemann Bredsdorff (1790-1841)

Rask's contemporary and friend Jacob Hornemann Bredsdorff was a brilliant mind and a very productive scholar both in natural sciences and in linguistics. Like Rask he grew up in Funen, where his father was vicar. He was taught at home, but at the age of 17 he was sent to the grammar school in Nykøbing, where Rask's former teacher Bloch had just been appointed rector. Two years later Bredsdorff enrolled at the University, where he received a bachelor's degree in divinity; in 1817 he received the gold medal of the University for an essay on paleontology, and the same year he defended a thesis on taxonomy in the natural sciences. From 1817-19 he taught Greek and Latin at the school in Roskilde. In 1821 he got an appointment at the department of mineralogy at the University, from 1823 as reader. From 1828 to his death in 1841 he taught botany and mineralogy at the Academy of Sorø.

He has given lasting contributions to geology and botany, but at the same time he wrote important papers on comparative Germanic linguistics, Danish phonetics, runology, and general linguistics.[1]

In 1817 he published a Prøve på en efter Udtalen indrettet dansk Retskrivning ('Sample of a Danish orthography based on pronunciation') which contains the first published phonetic transcription of Standard Danish.[2] In 1833 he published a paper, Om Tegn for de enkelte Lyddele i de europæiske Sprog ('On symbols for individual sounds in European languages') in which he presents a transcription system intended for a comparative analysis of European languages. In these works he appears as a very fine phonetic observer and analyst, in this respect superior to Rask. He sets up four degrees of "hardness" for stop consonants, exemplified by Danish p, French p, Danish b and French b; the difference between

---

1) His works within linguistics and runology have been republished in 1933, edited by Jørgen Glahder.

2) Including samples of different styles.

French *p* and Danish *b* is, however, described as problematic. Danish *bdg* must thus have been voiceless also at Bredsdorff's time.

Bredsdorff's main linguistic work is, however, the short treatise *Om årsagerne til sprogets forandringer* ('On the causes of linguistic change') 1821. In his linguistic work Bredsdorff certainly has received inspiration from Rask, but he was a very independent mind. Henning Andersen, in his excellent essay on Bredsdorff (1979), mentions his relations with Bloch and Rask, but at the same time the differences in Rask's and Bredsdorff's approaches. Rask sets up sound correspondences between languages by means of which he proves their relationship; Bredsdorff describes historical developments and tries to explain their causes.

Bredsdorff starts out by stating that the fact that language changes is easy to understand since we learn language from others, but the problem is why it changes in definite directions. He rejects the cause most frequently adduced in older works on linguistic change, viz. euphony, for we generally find that euphonic which we are accustomed to hear. Instead he mentions a number of other causes, of which he considers indolence to be the most important. This may lead to weakening and finally omission of sounds. Vowels may change to *ə*, and as far as consonants are concerned, he gives an instructive schematic account of the most frequently found weakenings, viz.

$$p - \left.\begin{matrix} b \\ f \end{matrix}\right\} - v - u \qquad t - \left.\begin{matrix} d \\ þ \end{matrix}\right\} - ð \qquad k - \left.\begin{matrix} g \\ x \end{matrix}\right\} - ɣ - \begin{matrix} j \\ u̯ \end{matrix}$$

$$\left.\begin{matrix} m \\ n \\ ŋ \end{matrix}\right\} - \text{nasal vowel,} \quad s - z \quad š - ž \qquad \left.\begin{matrix} f \\ x \\ s \end{matrix}\right\} - h$$

Another aspect of indolence is assimilations between neighbouring sounds, often so that they are pronounced with the same speech organ.

But an opposite tendency is also at work, the tendency to distinctness. As examples he mentions *venio* to *vengo*, Icelandic *ll* to *dl*, and, as the opposite of assimilation, the tendency to differentiation of neighbouring sounds, e.g. *ei* to *ai*, etc. Under the heading "The imperfection of the speech organs" he has some interesting observations concerning sounds which seem to be more difficult than others, e.g. ɣ ð ɫ r ž š þ x and the vowels y ø ö. These sounds are learned late by children, and some never learn them. He also mentions the factor of analogy and the possibilities of mishearing and of faulty memory. But these two latter factors are almost exclusively active in foreign words and thus particularly in the special situation of interference between languages, i.e. in borrowings, and in cases of language shift which may involve very radical changes in the adopted language.

In this brief paper, Bredsdorff has succeeded in mentioning most of the causes which have been alleged up to the present times, and his presentation is very clear and well balanced.

Kristian Sandfeld (1934) has characterized this treatise as the first rational theory of linguistic change, and Vilhelm Thomsen (1902, p. 55) called it a masterpiece of observation and penetrating insight which was 50 years ahead of its time. - Only the fact that it was published in Danish in a provincial school programme prevented it from influencing the development of historical linguistics.

4. The period of classical phonetics

4.0

In the second half of the 19th century phonetics developed quickly from spread and fumbling beginnings into a well established scientific discipline, a development which was partly due to the progress of physics and physiology, but was also conditioned by the increasing refinement of the methods of the dominating linguistic trend, comparative historical linguistics, which was based on regular sound correspondences and which required phonetic knowledge. This development occurred, above all, in England and Germany but quickly spread to the Scandinavian countries.

In 1881 Vilhelm Thomsen gave the first course in general phonetics at the University of Copenhagen, but his main achievements belong to other areas of linguistics.

4.1 Karl Verner (1846-96)

The Danish linguist Karl Verner was interested in phonetics in all its aspects. In his school days in Aarhus he read Rask's work on Danish orthography, and during his studies of Slavic languages at the University of Copenhagen he was particularly attracted by the problems of accent.

After having obtained his M.A. in 1873, he returned to his home town because he felt tired and ill and had no prospect of getting a job. It was here that he discovered what was later called "Verner's law". Verner's own story of how he made this discovery, as related by Jespersen in his obituary of Verner (1897), is not without interest for the understanding of the progression of science. Verner felt that he needed an afternoon nap, and when looking for a book to send him to sleep, he happened to take Bopp's Comparative Grammar, where the Sanskrit words are printed very prominently, and as he turned a page, the two words pitár and bhrátar stared him in the face, and it struck him that it was strange that they had different medial consonants in the old Germanic language. He then looked at the accent marks, and it occurred to him that the original Indo-european accent might be the cause of the difference. The next day when he was again going to take a nap, he hit upon an obviously wrong explanation of the difference, and so he sat down and went through the material and found one example after the other confirming his idea. In the course of the year 1874 he wrote the paper Eine Ausnahme der ersten Lautverschiebung, which was published in 1876. The final version of this paper did not retain any traces of its casual origin. It is brilliantly written, proceeding with inexorable logic and as convincing as a mathematical proof. He goes through all of the relevant material, comparing the forms of five older Germanic languages with the corresponding Sanskrit forms. According to W. Lehman (1967, p. 132) it "may be the single most influential publication in linguistics". It is of particular importance, (1) because it showed how deeply prosodic phenomena may influence sound development, and (2) because a large number of hitherto inexplicable exceptions to the first Germanic sound shift were now reduced to one simple rule, saying that the Germanic obstruents f þ x in medial position remained voiceless only after an immediately preceding Indo-european accent, otherwise they were voiced. And it was in particular this elimination of a large number of exceptions from a well-known law that permitted the neo-grammarians to assert that sound laws were not subject to exceptions ("die Ausnahmslosigkeit der Lautgesetze"), a thesis which was much debated in the following years. Verner himself, however, never accepted this doctrine in its strictest form.

As for the phonetic explanation of the rule, Verner suggested that the old Indo-european accent, which was assumed to be a pitch accent, had retained its place in the words in the oldest stage of the Germanic languages but had probably become partly expiratory. The extra airflow of the accented syllable thus preserved the voicelessness of the following consonant (which according to Germanic metrics belonged to the preceding syllable). This may be the right explanation, but now that we know of the relation between voiceless consonants and high pitch, the assumption of a change in the manifestation of the accent may perhaps not be considered necessary.

Verner was only 29 years old when he published this paper, but during the rest of his life he published hardly anything, and the whole body of his works comprises less than a hundred pages. This was not due to lack of knowledge or ideas; he had planned to write comprehensive treatises both on Slavic and on Scandinavian accents, but he was extremely conscientious and modest and without any ambition. He enjoyed struggling with linguistic problems but had no interest in publishing the results, and he would not have published his first paper either if it had not been for the strong pressure from his friends, particularly Vilhelm Thomsen. It was also his modesty which made him refuse a chair in Graz and a post as leading librarian in Jena, but he accepted a more subordinate post as librarian in Halle, and later, under pressure from his friends at home, applied for a post as reader of Slavic at the University of Copenhagen, which he received in 1883 and which was changed into a full professorship in 1888.

Although Verner did not like to publish, he was eager to expose his ideas in long letters to his friends, particularly Julius Hoffory, C.W. Smith, and Vilhelm Thomsen, and a good many of these letters have been published after his death together with his papers (1903). Among the few reviews he wrote, the most important is his review of Axel Kock's book on Scandinavian accents, in which he also develops his own views on this subject, including his observations on the Danish stød. He had previously (e.g. in a letter to C.W. Smith 3/7 1872) described his own production of the Danish stød. He feels an increased energy in his larynx and also, by palpation, finds a sudden movement of the larynx muscles. He therefore thinks that there is a sudden and strong closure or at least constriction of the glottis. In the above mentioned review he adds to this description that, whereas words without stød are characterized by a falling tone, the words with stød have a quickly rising tone in the first syllable, ending in a sudden closure of the vocal cords and a lower tone in the following syllable. Verner now suggests that the stød has developed out of such a rising tone, which has led to an "überschnappen" of the vocal cords. His explanation was accepted by Jespersen (1897-99, p. 608, see 4.3). He also refers to the rising accent I in Norwegian (it had been known since the middle of the century that the Danish stød is historically related to accent I), and to similar phenomena in Baltic languages. He found in his own speech the same rising tone in words which have accent I in Norwegian and Swedish, but which do not have stød in Danish (e.g. kat) because they lack the so-called stød-base (i.e. a long vowel or a short vowel + voiced consonant). The origin of the Danish stød is still under debate, but Verner has given interesting contributions to its solution. His description of his own pronunciation is probably also quite correct, and not an individual feature, but particularly the pitch characteristics vary very much according to dialect.

In his later years Verner developed a certain scepticism with regard to articulatory descriptions and hoped for a much more precise and objective description from the acoustic analysis of speech sounds. In 1878 he had occasion to see the new Edison phonograph and this gave him the idea to construct an ingenious instrument for acoustic measurements of vowels. He used the Edison phonograph in its first form in which it is driven by hand and the impressions of the sound are made vertically on tinfoil. To the small pin which followed the tracings he attached a small mirror which was tipped by the movements of the pin, and these movements were enlarged approximately 2000 times by means of a beam of light reflected from the mirror on to a large measuring yard placed 4-5 meters from the instrument. Through a telescope placed close to the mirror it was possible to see the mirror image of the enlarged measure and to read off the exact value. The horizontal axis of the vibrations was enlarged by means of a system of gear wheels and could be read off a measuring disc which turned 1344 times for one revolution of the drum. In this way he could make exact measurements of the first ten harmonics and he also found a simplified method of making the mathematical computations. However, when he heard that Otto Pipping in Helsinki had built a similar instrument, he discontinued his experiments and did not publish anything about them. But he has given some information about his methods in two letters to Otto Pipping, which were published in 1913, and in a letter to Georg Forchhammer, now in the Royal Library. A short description of the instrument by his brother Rudolf Werner, who was an engineer, was published as an appendix to the collection of Verner's papers and letters in 1903.

Even if he had continued his experiments, the technical development of his time would not have allowed him to reach lasting results. All the attempts at acoustic analysis at the end of the 19th century were soon outdated. Verner's interests were ahead of his time.

4.2 Julius Hoffory (1855-97)

Karl Verner's friend Julius Hoffory must also be mentioned in a history of phonetics in Denmark. He had a broad knowledge and an acute intelligence. His main works are, besides his thesis on Old Scandinavian consonants, a paper on general phonetics, "Phonetische Streitfragen" 1877 and a book in which he criticizes the views of the German phonetician E. Sievers, "Professor Sievers und die Principien der Lautphysiologie" 1889. In his view of phonetic problems he was particularly influenced by the German physiologist Brücke. He was particularly interested in systematizing the sounds of speech and wanted to keep physiological and acoustical classifications apart. In 1883 he lectured on the principles and methods of the physiology of sounds at the University of Copenhagen. Shortly afterwards he went to Berlin, where he lectured at the University, from 1887 as professor of phonetics and Scandinavian languages. But at the age of 34, severe illness forced him to give up his work.

Both Verner and Hoffory were excellent phoneticians, but they buried almost all their knowledge in their private correspondence. Therefore they did not have much direct influence on the development of phonetic studies in Denmark. In this respect, the work of Otto Jespersen was much more important.

4.3 Otto Jespersen (1860-1943)

Otto Jespersen was born in Jutland in 1860. In 1870 his father, who was a judge, died and the family moved to Zealand. At the age of 14 Jespersen also lost his mother, and he had to finance his university studies himself by means of a job as stenographer in parliament. He started studying law (1877), following a family tradition, but after four years' study - shortly before the final examination - he changed to Romance philology, and in 1887 he passed the examination intended for teachers at senior schools which had recently been instituted, with French as a main subject and Latin and English as secondary subjects. He also attended courses in Russian with Verner and in linguistics with Vilhelm Thomsen, whom he admired very much. Jespersen was in opposition to the classical tradition of the University. He hated Latin, which was at that time compulsory as secondary subject, and was much more interested in modern languages and in practical problems of language teaching.

Almost from the start he took a lively interest in phonetics. He attended the first courses given in this subject at the University, in 1881 by Vilhelm Thomsen and in 1883 by Hoffory, and he read the works of Sweet, Passy, Viëtor and Storm and was very much influenced by their ideas concerning better methods of language teaching using the spoken language and by means of texts in phonetic transcription. In 1884 he translated a treatise by Felix Franke on language teaching, and for some years Jespersen and Franke corresponded intensively on linguistics, phonetics, and language teaching. They had the same aims and out of the correspondence grew a friendship which meant much to both of them. But they never met. Franke died of tuberculosis at the age of 25. In 1885 Jespersen published a small English grammar where all examples were in phonetic transcription. In 1886 he published a paper "Zur Lautgesetzfrage", which he reprinted in 1904 and 1933, in both cases with long postscripta, elaborating his view in more detail. At its first appearance in 1886, the paper was a contribution to an ongoing heated debate on the possibility of exceptions to sound laws. Jespersen was not willing to accept the dogma of the "Ausnahmslosigkeit der Lautgesetze". He considers "sound laws" as generalisations which are only valid in a "telescopic" but not in a microscopic perspective. In details there are exceptions which may be due to the communicative function of the words (words which are easily understandable from the situation, e.g. greetings, may be weakened more than other words), and to their meaning.

After his exam in 1887, Jespersen studied abroad and met the influential phoneticians of that time, whose works he already knew: Sweet, Passy, Viëtor, Sievers. About that time Vilhelm Thomsen proposed to him that he should specialize in English because there would soon be a chair in that subject. So Jespersen studied English in Berlin with Zupitza. His thesis on English cases was accepted in 1891, and in 1893 he was appointed professor of English at the University of Copenhagen. In 1889 he published a small book, *Articulation of speech sounds*, in which he proposed a new type of phonetic transcription which he called "analphabetic" and in which each sound was indicated by a combination of symbols. It was inspired by Bell's "visible speech" system, but differed from the latter by the use of well known letters and numbers which could be found in every printing office. Jespersen used Greek letters to indicate the articulating organ, Latin letters to indicate the place approached by this organ, and numbers to indicate the degree of constriction. Thus a French voiced dental is, e.g., written $\beta 0^{e}\delta 0\varepsilon 1$ which means that the tongue tip (β) makes a closure (0) with the teeth (e), the velum (δ) is closed (0), and the vocal cords (ε) are constricted in position 1 (for voicing), whereas a Danish voiceless *d* is written $\beta 0^{ef}\delta 0\varepsilon 2$. Jespersen mentions as one of the advantages that the same components found in different sounds have the same symbols (as a matter of fact, the subgroups of the formulae (e.g. $\beta 0^{e}$) come close to distinctive features). Jespersen uses the system in his subsequent textbooks of phonetics, but of course it cannot be used to render connected texts.

In 1889 he worked out a proposal for a more traditional, alphabetic phonetic transcription (later called Dania, because it was published in the periodical Dania), which was intended particularly for the transcription of Danish, and which has been used since in Danish dialectology. It will also be used in the forthcoming Danish pronouncing dictionary, edited by Jørn Lund et al. One of the principles was that symbols without diacritics are used for normal Danish sounds, which e.g. means that *e ø o* indicate raised values compared to the IPA symbols. The most confusing discrepancies with IPA are probably that r is used for an unrolled uvular ʁ, that œ indicates a low, unrounded front vowel, and (worse) that Dania a is IPA ɑ and ɑ is a.

In 1897-99 Jespersen published his principal phonetic work, Fonetik, en systematisk fremstilling af læren om sproglyd, i.e. a systematic presentation of general phonetics, a work of more than 600 pages. In 1904 two German books appeared which together covered the content of the Danish work, viz. Phonetische Grundfragen and Lehrbuch der Phonetik. Compared to the Danish work they are somewhat abbreviated and brought up-to-date. Most of the Danish examples found in Fonetik were left out, but they were utilized in his book Modersmålets Fonetik, a book on Danish phonetics which appeared in 1906 and has been reprinted and revised several times since then.

Jespersen is not a pioneer in phonetics like Bell and Sweet. His work should rather be seen as the culmination of the trend which may be called classical phonetics, and which may be characterized by a mainly physiological description of speech sounds, built on what can be observed without instruments, except for a mirror, and perceived kinaesthetically, and what can be concluded about articulation on the basis of auditory perception by means of known correlations between articulation and perception.

Like Sievers and Sweet, Jespersen considers the new instrumental methods, as they were described and used e.g. by Rousselot, with a good deal of scepticism. He objects that the instruments may make the informant speak in an unnatural way, that the curves may not mirror the pronunciation exactly, that it is not possible to control afterwards what has been said, that the material is generally not comprehensive enough for generalizations, and that the communicative function of speech is not always respected. It must be admitted that this criticism was in most respects justified at Jespersen's time. But the best experimentalists, like Rousselot, were well aware of the limitations of their methods. And the negative attitude of the leading figures in classical phonetics contributed to delaying the progress in methods and the establishment of phonetic laboratories. This was also true in Denmark.

On the other hand, by his emphasis of the communicative aspect Jespersen prepared the soil for phonological points of view. In his descriptions of sound systems in different languages he very often quotes minimal pairs, and in his treatment of quantity he distinguished sharply between "inner" length, i.e. phonological length, and "outer" length, i.e. length differences conditioned by the environment.

Jespersen's Fonetik contains all the knowledge of his time with elimination of assertions and descriptions which he considered to be dubious. He had a critical mind and a good deal of common sense, combined with a gift for acute observation. And he was very careful not to mention sounds which he had not heard himself and, generally, learnt to produce to the satisfaction of native speakers. His examples are therefore mostly taken from Danish, English, French, and German, and in a few cases from other Romance languages and from Russian. This involves a restriction of the scope of the book, but at the same time it means that it is unusually reliable and that it can still be read with profit. It should be kept in mind, however, that the languages described have developed further during the 80 years that have passed since his Fonetik was published. In particular, his description of Danish pronunciation - correct as it was for his time - is in several points at variance with the pronunciation of the younger generation in Copenhagen. This is particularly true of the variants of the phoneme /a/, the short /ɔ/, the consonant /ɣ/, which has disappeared in the pronunciation of the young generation, and the place of the stød in words with /ð/, /j/, /v/ and /ʁ/.

The description of the individual sounds, as also of the whole system, is based exclusively on sound production. In a short chapter he gives an account of what was known at that time about the acoustics of speech, and motivates his choice of production convincingly by the fact that much more is known about this aspect. He adds that sound production is more important for the explanation of sound change, and he believes that speech sound perception is closely connected to and conditioned by sound production. It is well known that the latter point is still under debate.

More problematic, but characteristic of Jespersen's empirical approach, are his arguments against Pipping, who had maintained that children produce the same vowels as adults and, since their speech organs have much smaller dimensions, that they must articulate differently, and thus the acoustic aspect of sounds is more stable than articulation. In order to refute this assertion, Jespersen has measured the length of the lower jaw for different age groups and found that the differences are small compared to the more general differences of body height. This is correct, but he would have got different results from measuring the length of the pharynx. However, he could not be expected at that time to be aware of the importance of the pharynx in sound production and acoustics.

As far as the general system of speech sounds and the individual characteristics of the sound types are concerned, Jespersen's book does not contain new points of view compared to his immediate predecessors (except for the distinction between "groove" and "slit" consonant types). But he has many fine observations on details (for instance in the long chapter on r-sounds), and the chapters on the syllable and on stress contain original contributions to the description of some controversial concepts in phonetics. In the description of the syllable and of prosodic phenomena, Jespersen - like most other phoneticians - abandons the physiological point of view and takes his starting point in perception. The syllable is considered to be a peak of relative sonority. This is not a new definition, but Jespersen goes into more concrete detail. He sets up a scale of sonority for speech sounds: 1. voiceless sounds, (a) stops, (b) fricatives, 2. voiced stops, 3. voiced fricatives, 4 (a) nasals, (b) laterals, 5. voiced r-sounds, 6. high vowels, 7. mid-vowels, 8. low vowels. -- In each sound chain there are as many syllables as there are relative peaks of sonority. This is exemplified graphically in the following way:

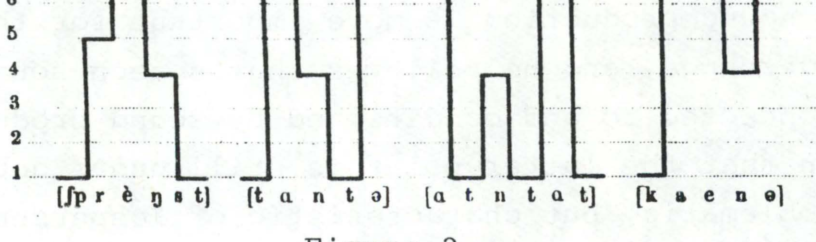


Figure 2

From Lehrbuch der Phonetik p. 192.

He adds the rule that between a given sound and the syllabic peak only sounds of the same or a higher sonority class are permitted. Sonority is defined somewhat vaguely by saying that the most sonorous sounds can be heard at the longest distance. The problem of s and ʃ is solved by putting voiceless stops and fricatives in the same class. In an interjection like pst, s forms the syllabic peak, but compared to a neighbouring vowel (papst), the difference is too small to give the impression of a separate peak in s. - The fact that, e.g., words like Cimbria are heard as having three syllables is accounted for by the principle that a relatively sonorous sound following a less sonorous sound is often heard as a separate syllable even if the following sound is still more sonorous, particularly if it is relatively long, as in Danish [sulldn̩ə] ('sultende', vb. (starving)) (which differs from [sulldnə] ('sultne', adj. (hungry)).

A reduction in intensity between two sounds of the same class is at the same time a valley in sonority and may also give a syllable boundary, like i Italien [i i...], or bonden [bɔnn̩].

The difference between Danish [viʔn] ('vin' (wine)) and [biʔn̩] ('bien' (the bee)) is explained by the n being weak and almost voiceless in the first case, so that it does not give a new peak, but fully voiced in the second case. -- This is certainly right, but Jespersen is very bold when he asserts that if the apparently same consonant has different syllabic functions in different languages, one can be sure that it is pronounced differently. Jespersen's sonority principle functions quite well for the languages he has chosen to include in his description, and may perhaps be considered as a definition of the optimal syllable type, but if the principle is considered to be universal, as he seems to intend, he would get into trouble already with Russian, to say nothing of many African and American Indian languages.

In the chapter on stress he introduces some useful new concepts. He distinguishes four types of stress: (1) traditional stress (i.e. stress that is bound to a particular syllable of a word), (2) "value" stress (which may be either stress of "novelty" or stress of "opposition"), (3) "unity" stress, which combines parts of a word or a syntactic group by stressing the final part, and (4) "rhythmical" stress. The syntactic unity stress has a particularly clear function in Danish compared to German, English, and French. The frequently found stress of the first part of a compound he considers to be stress of opposition.

As mentioned above, Jespersen was very much interested in language teaching and particularly in applying phonetics to language teaching. As early as 1891 he started writing text books in English with phonetic transcription of both texts and word lists, and this approach was generally accepted and applied in text books for schools in the following years. In 1901 he published a book on language teaching.

After 1900, however, he concentrated his efforts more on the study of English grammar and the history of the English language and on the theory of grammar. His main work in the former field, Modern English grammar, appeared in seven volumes from 1909-42, and in the latter field his main work is Philosophy of grammar from 1924. He also wrote a book on child language and an often quoted paper on the symbolic value of the vowel i.

Another of his main works is Language (1922), which deals with language and linguistics in general. In this book he also treats the idea of progress in language which was one of his favourite ideas. In his general attitude to life and to language in particular he was influenced by Darwin, Spencer, and Stuart Mill. He believed firmly in progress, both in politics and in the development of language. The introduction to his thesis (1891) starts with a chapter on progress in language, and 50 years later he published a book with the title Efficiency in linguistic change.

4.4 Georg Forchhammer (1861-1938)

Jespersen's contemporary Georg Forchhammer, originally an engineer, later a teacher of the deaf, deserves to be mentioned briefly, partly for his vowel cubus with three dimensions: tongue height, rounding, and place of articulation, and partly for his theory of vocal intensity. He is opposed to the traditional view that intensity is due to expirative force and points to the function of the vocal cords. A more precise closure will give the most economic utilization of the air pressure, and thus more force, whereas a lax and unprecise closure will give a stronger airflow but a relatively weak voice. Forchhammer now thinks that in singing the intensity is mainly governed by the expirative muscles, but in speech the unstressed syllables are often pronounced with less compression of the vocal cords and more airflow. -- This is probably a correct description of the unstressed syllables in the Germanic languages (schwa). Jespersen quotes this theory as an alternative to the traditional description of weak stress as due to less expiratory force. However, lack of stress is hardly due to lower intensity caused by this type of phonation, but rather to factors of pitch and duration combined with the neutral vowel quality.

Forchhammer also constructed a phonoscope by means of which the oscillations of speech sounds were transferred to a gas flame and recorded on a revolving drum.

5. Brief outlook on the 20th century

Due to Jespersen's influence, language teaching in Danish schools was based on phonetic transcriptions. This meant that also the teachers were supposed to have phonetic knowledge, and from the twenties onwards phonetic training played a greater role in language studies at the Danish universities than in most other countries, and phonetics was part of the requirements for the MA in modern languages. But there was not much phonetic research going on until the middle of the forties. In the following years three main trends may be distinguished: (1) a continuation of the Jespersen tradition (descriptions based on auditory observation) represented by Aage Hansen and, in recent time, by Brink and Lund; (2) instrumental phonetics, starting with Svend Smith's thesis on the Danish stød from 1944, and since 1966 mainly concentrated in the Institute of Phonetics in Copenhagen and its laboratory; (3) phonological research, inspired by the various European and American trends and by Hjelmslev's glossematic theory, whose main representatives, after Hjelmslev, are Jørgen Rischel and Hans Basbøll. It should also be specially mentioned that due to a strong influence from phonology, Danish dialectology has been based on methods of structural linguistics since the early thirties. Among its chief representatives one may mention Poul Andersen and Anders Bjerrum.

A brief sketch of the external history of this period is found in ARIPUC 10. A detailed history and evaluation of its scholarly achievements may better be left for the future.

Bibliography

The bibliography lists the references to each section of the paper separately, first the works by the author treated in the section in question, and then the works dealing with or mentioning this author:

Section 2.1

Madsen Aarhus, Jacob 1586: (Matthiae Arhusiensis, Jacobus) De literis libri duo. Basel. Book I reprinted in Internationale Zeitschrift für Sprachwissenschaft 1890, V, p. 84-132, with an introduction by F. Techmer. Book I-II reprinted in Acta Jutlandica II, 1930; a Danish translation with ith introduction and comments by Christen Møller, Peter Skautrup, and Franz Blatt is printed in Acta Jutlandica III, 1931.

Section 2.2

Pedersen Høysgaard, Jens 1743a: Tres faciunt collegium. Copenhagen.

Pedersen Høysgaard, Jens 1743b: Concordia res parvae crescunt eller Anden Prøve af Dansk Orthographie. Copenhagen.

Pedersen Høysgaard, Jens 1747: Accentuered og Raisonered Grammatica. Copenhagen.

Pedersen Høysgaard, Jens 1752: Methodisk Forsøg til en Fuldstændig Dansk Syntax. Copenhagen.

Pedersen Høysgaard, Jens 1769: Første Anhang til den Accentuerede Grammatica. Copenhagen.

Bertelsen, Henrik 1926: Jens Pedersen Høysgaard og hans Forfatterskab. Copenhagen.

Section 2.3

Kratzenstein, C.G. 1781: Tentamen resolvendi problema ab academia scientiarum imperiali petropolitana ad annum 1780 publice propositum. 1. Qualis sit natura et character sonorum litterarum vocalium a,e,i, o,u tam insigniter inter se diversorum. 2. Annon construi queant instrumenta ordini tuborum organicorum, sub termino vocis humanae noto simila, quae litterarum vocalium a,e,i,o,u sonos exprimant. praemio coronatum. Petropoli.

Kratzenstein, C.G. 1782: "Essai sur la naissance et la formation des voyelles", Observations sur la physique, sur l'histoire naturelle et sur les arts, Supplément 1782. Tome XXI. Paris. Bureau du Journal de physique. (French translation of the Latin prize essay.)

Aschoff, V. 1979: "Phantasie und Wirklichkeit in der Frühgeschichte der Akustik", Acustica 42, p. 121-132.

Ungeheuer, G. 1962: Elemente einer akustischen Theorie der Vokalartikulation. Berlin.

Wheatstone, Ch. 1837: "Reed organ pipes, speaking machines...", London & Westminster Review 28 = The scientific papers of Charles Wheatstone. London 1879. p. 348-367.

Section 3.1

Rask, Rasmus K. 1818: Undersøgelser om det gamle Nordiske eller Islandske Sprogs Oprindelse. Copenhagen. Reprinted in Udvalgte Afhandlinger.

Rask, Rasmus K. 1826: Forsøg til en videnskabelig dansk Retskrivningslære. Copenhagen.

Rask, Rasmus (1932-35): Udvalgte Afhandlinger I-III, ed. (with commentary) by Louis Hjelmslev, with an introduction by Holger Pedersen. Copenhagen.

Rask, Rasmus (1938): Om de fyenske Bønders Sprog. Edited and introduced by Poul Andersen. Udvalg for Folkemaal's Publikationer, Serie A Nr. 2. Copenhagen.

Rask, Rasmus (1941): Breve fra og til Rasmus Rask I-II. ed. by Louis Hjelmslev. III Commentary 1968 by Marie Bjerrum.

Andersen, Poul 1937: "Rasmus Rask", Fra Rask til Wimmer, p. 7-33. Copenhagen.

Bjerrum, Marie 1959: Rasmus Rasks afhandlinger om det danske sprog. Copenhagen.

Diderichsen, Paul 1960: Rasmus Rask og den grammatiske tradition. Copenhagen. German translation: Rasmus Rask und die grammatische Tradition, 1976. München.

Hjelmslev, Louis 1951: "Commentaires sur la vie et l'oeuvre de Rasmus Rask", Conférence de l'institut de linguistique de l'Université de Paris, p. 143-157.

Jespersen, Otto 1918: Rasmus Rask, in Folkets Førere. Copenhagen - Kristiania.

Section 3.2

Bredsdorff, Jakob Hornemann 1817: Prøve af en efter Udtalen indrettet dansk Retskrivning. Copenhagen. Reprinted in Udvalgte Afhandlinger.

Bredsdorff, Jakob Hornemann 1821: Om Aarsagerne til Sprogenes Forandringer (school programme, Roskilde). Reprinted in Udvalgte Afhandlinger.

Bredsdorff, Jakob Hornemann 1833: "Om Tegn for de enkelte Lyddele i de europæiske Sprog", Blandinger fra Sorø 5, p. 17-23. Reprinted in Udvalgte Afhandlinger.

Bredsdorff, Jakob Hornemann (1933): Udvalgte Afhandlinger indenfor Sprogvidenskab og Runologi, edited by Jørgen Glahder. Copenhagen.

Andersen, Henning 1979: "Jakob Hornemann Bredsdorff: On the Causes of Linguistic Change" (1821). English translation, with commentary, and an essay on J.H. Bredsdorff. Rask-Hjelmslev Symposium, Copenhagen 3-5 Sept. 1979 (manuscript).

Glahder, Jørgen 1937: "Christian Molbech og Jakob Hornemann Bredsdorff", Fra Rask til Wimmer, p. 34-43.

Sandfeld, Kristian 1934: Article on J.H. Bredsdorff in Dansk Biografisk Leksikon IV, p. 54-56.

Thomsen, Vilhelm 1902: Sprogvidenskabens Historie. Copenhagen, p. 55.

Section 4.1

Verner, Karl 1876: "Eine Ausnahme der ersten Lautverschiebung", Zeitschrift für vergleichende Sprachforschung 23, p. 97-130. Reprinted in Afhandlinger og Breve 1903. English translation in P.W. Lehman: A Reader in Nineteenth Century Indo-European Linguistics 1967, p. 134-163.

Verner, Karl 1881: Review of Axel Kock "Undersökningar om svensk akcent 1878", *Anzeiger f. deutsches Altertum* VII, p. 1-13. Reprinted in *Afhandlinger og Breve* 1903.

Verner, Karl (1903): *Afhandlinger og Breve* (with a biography by Marius Vibæk). Copenhagen.

Verner, Karl (1913): "To Breve fra Karl Verner, udgivne af Vilh. Thomsen og J.P. Gram", *Oversigt over det kgl. Videnskabernes Selskabs Forhandlinger* No. 3, p. 161-189.

Jespersen, Otto "Karl Verner 1846-96", *Tilskueren* 1887 and *Linguistica* 1933, p. 12-24.

Lehmann, Winfred P. 1967: Introduction to Verner's paper in *A Reader in Nineteenth Century Indo-European Linguistics*, p. 132-134.

Smith, Svend 1944: *Stødet i Dansk Rigssprog*, p. 66-76. Copenhagen.

Vibæk, Marius 1903: Biography in Karl Verner: *Afhandlinger og Breve*, p. XI-XCI.

Werner, Rudolf "Beskrivelse af Karl Verners apparat til fonometriske Undersøgelser", *Afhandlinger og Breve* 1903, p. 365-372.

Section 4.2

Hoffory, Julius 1877: "Phonetische Streitfragen", *Zeitschrift für vergleichende Sprachwissenschaft* 23, p. 525-558.

Hoffory, Julius 1884: *Professor Sievers und die Principien der Sprachphysiologie*. Berlin.

Heusler, Andreas 1898: "Julius Hoffory", *Arkiv för nordisk filologi* 14, p. 206-212.

Kristensen, Marius 1936: Article in *Dansk Biografisk Leksikon* X, p. 306-307.

Section 4.3

Jespersen, Otto 1886: "Zur Lautgesetzfrage", Zeitschrift f. allgemeine Sprachwissenschaft III, and Nordisk tidsskrift for filologi VII. Reprinted, with a long addition in Phonetische Grundfragen, 1904, and, with a further addition in Linguistica 1933.

Jespersen, Otto 1889: Articulation of Speech Sounds, represented by means of Analphabetic Symbols. Marburg.

Jespersen, Otto 1891: Studier over engelske kasus. Copenhagen.

Jespersen, Otto 1894: Progress in Language. London.

Jespersen, Otto 1897-99: Fonetik. Copenhagen.

Jespersen, Otto 1901: Sprogundervisning. Copenhagen.

Jespersen, Otto 1904: Lehrbuch der Phonetik. Leipzig.

Jespersen, Otto 1904: Phonetische Grundfragen. Leipzig.

Jespersen, Otto 1904: How to Teach a Foreign Language. London.

Jespersen, Otto 1905: Growth and Structure of the English Language. Leipzig.

Jespersen, Otto 1906: Modersmålets fonetik. Copenhagen.

Jespersen, Otto 1909-49: Modern English Grammar I-VII. London - Copenhagen.

Jespersen, Otto 1912: Elementarbuch der Phonetik.

Jespersen, Otto 1913: Sprogets logik. Copenhagen.

Jespersen, Otto 1922: Language. London.

Jespersen, Otto 1922: "Symbolic value of the vowel I" Philologica I, reprinted in Meisterwerke der romanischen Sprachwissenschaft, ed. L. Spitzer 1929, and (with additions) in Linguistica, 1933.

Jespersen, Otto 1924: The Philosophy of Grammar. London.

Jespersen, Otto 1933: Linguistica. Copenhagen.

Jespersen, Otto 1938: En sprogmands levned (Autobiography). Copenhagen.

Jespersen, Otto 1941: *Efficiency in Linguistic Change*. Copenhagen.

Hjelmslev, Louis 1942-43: "Otto Jespersen", *Acta Linguistica* III, p. 119-130.

Smith, Svend 1945: "En dansk Fonetikers Fødsel", *Nordisk Tidsskrift for Tale og Stemme* 8, 1944, p. 27-49.

Section 4.4

Georg Forchhammer 1898: "Artikulationslære", *Tidskrift för Dövstumskolan*.

(ARIPUC 13, 1979)

INFLUENCE OF MICROPHONE POSITION IN THE RECORDING OF SPEECH SIGNALS

Carl Ludvigsen

Abstract: Sound pressure levels of various speech sounds are measured simultaneously at different distances from the mouth. The observed values for low vowels and [s] differ especially close to the mouth from those predicted from the distance law for sound radiation. The variation of sound pressure with distance seems to depend on the speech sound in question. Some consequences hereof are pointed out. The results are compared with calculated values of the sound pressure from a sound source on a rigid sphere. Some of the observed deviations from the distance law seem attributable to the different frequency composition of the speech sounds. However, some of the observations (e.g. the difference between the variation of low and high vowels at positions close to the mouth) cannot be accounted for by the model.

1. Introduction

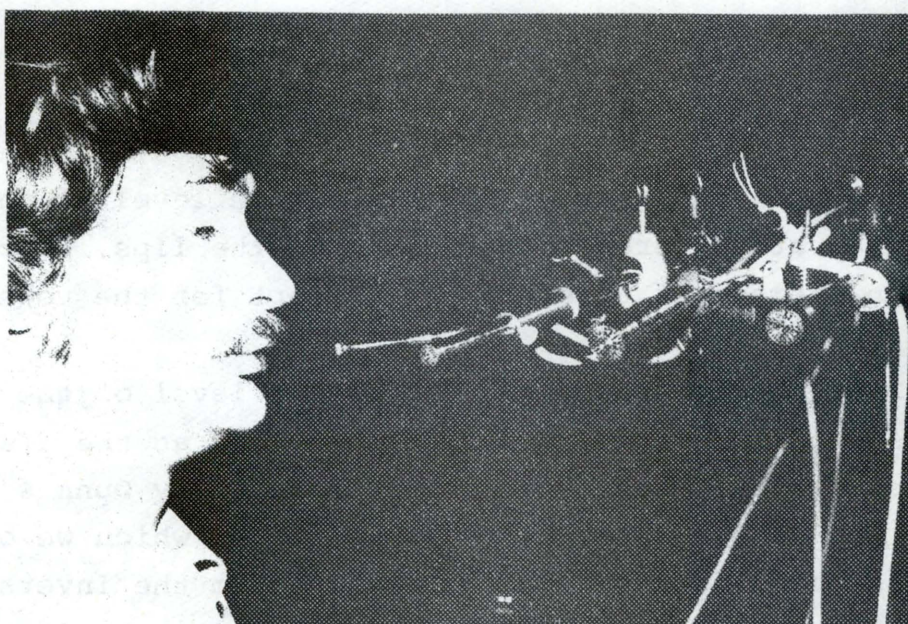
Recording of speech signals is a standard procedure in all phonetic laboratories. The microphone is typically placed at a distance of 25 - 31 cm directly in front of the mouth (see, e.g., Lehiste & Peterson 1959, Fairbanks et al. 1950, Strevens 1960, and House & Fairbanks 1953). Obviously, there are good reasons for that choice. If the microphone is placed too close to the mouth (e.g. closer than 10 cm), the expirated air will tend to generate turbulent noise, when passing the microphone; on the other hand, if the microphone is placed far from the mouth (e.g. more than 1 m away), problems with the signal-to-noise ratio may arise. The main reason for placing the microphone directly in front of the lips is that high frequency sounds are mainly radiated in this direction. Only one detailed discussion concerning this question has been published until now, namely Dunn & Farnsworth (1939).

They measure the average sound pressure at seventy-six positions, in different distances and directions from the mouth of a single speaker. They find that the variation of the intensity of speech measured on a horizontal line directly in front of the mouth is similar to that of a single source placed 0,6 cm inside the lips, i.e. the sound pressure at two points at distances L and 2L from the acoustic centre 0,6 cm inside the lips differs with 6 dB. This relationship is often called the 1/R-law. The measurements are average RMS-pressures of a 15 sec. utterance and measured in different frequency bands. The over-all average intensity conforms to the 1/R-law with a high degree of accuracy, whereas deviations are observed within some of the frequency bands.

In recording speech material in our own laboratory, we have noticed a tendency for the intensity difference between low and high vowels to be less pronounced if the microphone distance is short (e.g. 5 cm) and for the intensity of unvoiced s-sounds compared to the intensity of vowels to depend on the microphone distance as well. These observations are partly in agreement with Dunn & Farnsworth (1939), if we assume that the intensity of a vowel is mainly determined by the intensity of the first formant. However, since the measurements of Dunn & Farnsworth are average values for 15 sec. of connected speech, only a gross estimate of the variation for single speech sounds can be derived from their data. In order to throw further light on this problem, a series of measurements were carried out.

2. Measurements

A male speaker, who is a trained phonetician, sat on a specially constructed chair in the anechoic chamber of the Institute of Phonetics. The distance law (1/R-law) is complied with in this room within ± 1.0 dB in the frequency range 100 - 10.000 Hz in the space used for measurements. The subject's neck was supported by a headrest. In front of the subject's lips was placed a row of 5 microphones: 5 cm from the lips a 1/4-inch B & K condenser microphone, 10 cm from the lips a ½-inch B & K condenser microphone, and 20, 30 and 100 cm from the lips 1/1-inch B & K condenser microphones, see figs. 1 and 2.



Figures 1 and 2

Microphone set-up in the anechoic chamber. The microphones are placed at 5, 10, 20, 30, and 100 cm from the average position of the centre of the lips.

The subject read a list of nonsense syllables and isolated vowels at a constant and comfortable speaking level. Each item appeared 6 times in the list, in random order. The list consisted of the vowels:

[i, e, ɛ, æ, a, ɑ, y, ø, œ, ɶ, u, o, ɔ, ʌ, ɒ]

and the nonsense syllables:

[mi, mu, ma, si, su, sa].

The signals from the five microphones were recorded simultaneously. For calibration purposes, a 200 Hz pure tone was recorded before and after the reading of the list. The maximum RMS vowel amplitudes were then registered on a B & K level recorder, type 2305. The amplitudes of the initial [m] and [s] segments were obtained by feeding the signal to an intensity meter (with a double-linear rectifier) and registering the output on a fast ink-writer (Mingograph). Identification of different segments was performed by comparing the intensity curve with an intensity curve of the high-pass filtered signal (500 Hz) and the duplex oscillogram.

3. Results

3.1 Vowels

Figs. 3-5 show the average sound pressure level of the 15 vowels measured at different distances from the lips. Each point represents an average of 6 recordings, except for the vowel [æ], which is based on 5 recordings only.

The curves show that the sound pressure level of the vowels at distances from 10 cm to 1 m decreases 5-6 dB as the distance doubles, i.e. slightly less than the 6 dB found by Dunn & Farnsworth (1939) and the theoretical value of 6 dB, which we obtain from a simple point source. The deviation from the inverse law is statistically the same for all vowels at distances of 10-100 cm from the lips.

At distances of 5-10 cm from the lips, the deviation from the inverse law is greater and is not the same for all vowels. Typically, the sound pressure for low vowels [ɒ], [ɶ] and [ɑ] is only 1-2 dB higher 5 cm from the lips as compared to 10 cm from the lips, while the difference for high vowels [i], [e], [y], [ø], [u] and [o] is 3-5 dB. This difference between low and high vowels is statistically significant beyond the 0.1 per cent level.

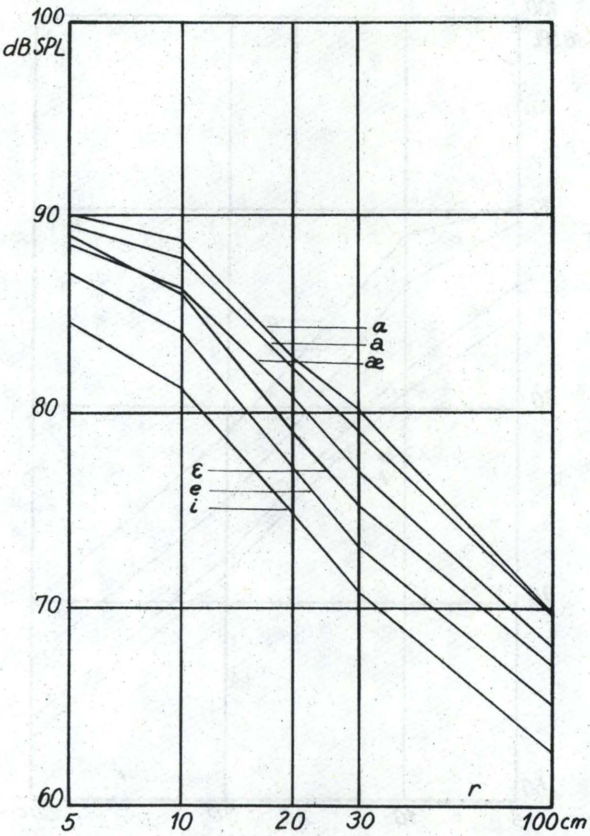


Figure 3

Average sound pressure levels for the vowels [i, e, ɛ, æ, a, ɑ] registered at 5, 10, 20, 30, and 100 cm in front of the lips. The standard deviation is close to 1 dB for all registrations.

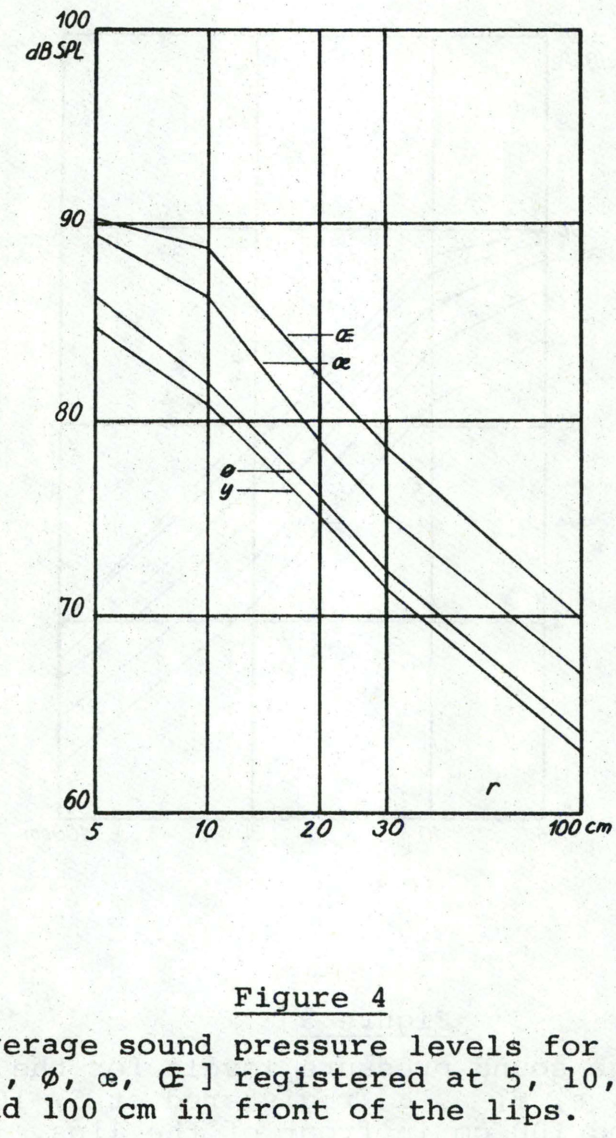


Figure 4

Average sound pressure levels for the vowels [y, ø, œ, Œ] registered at 5, 10, 20, 30, and 100 cm in front of the lips.

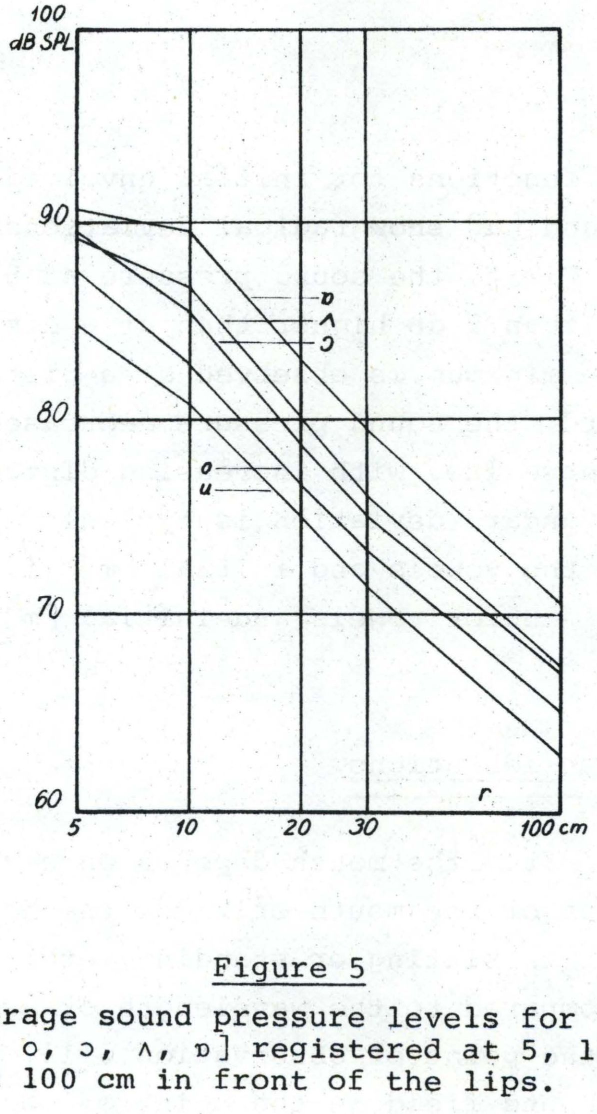


Figure 5

Average sound pressure levels for the vowels [u, o, ɔ, ʌ, ɒ] registered at 5, 10, 20, 30, and 100 cm in front of the lips.

3.2 Consonants

3.2.1 Initial [m]

The observed distance functions of initial [m] before the vowels [i], [u], and [ɑ] resemble those found for narrow vowels, except that the deviation from the inverse law seems to be smaller, cf. fig. 6.

3.2.2 Initial [s]

The distance functions for initial unvoiced [s] before the vowels [i], [u], and [ɑ] show radical deviations from the inverse law, cf. fig. 6. First, the sound pressure at 5 cm from the lips is typically less than 1 dB higher than at a distance of 10 cm. Second, a relative minimum is observed at a distance of 30 cm from the lips, and third, the sound pressure decreases more than predicted by the inverse law, with increasing distance. The unbiased estimate of the standard deviation is typically 3-5 times greater than observed for the vowels and initial [m] (i.e. 3-5 dB compared to approximately 1 dB for vowels and initial [m]).

4. Theoretical considerations

The radiation from the mouth depends on several parameters, e.g. the dimensions of the mouth orifice, the head and the body, whether the person is sitting or standing, etc. However, objects which are small compared to the wavelength or are far from the sound source and the point of observation will only have a small influence on the sound field in the point of observation. Therefore, fairly simple models such as a piston in an infinite rigid wall or a rigid sphere can be used as a basis for the calculation of the radiation from the mouth. In order to explain the observed deviations from the distance law, a model with a piston in an infinite wall is too simple, since it predicts that sound pressure varies according to the inverse law with amplitudes proportional to the frequency, provided that the dimension of the piston is small compared to the wavelength. This will be the case for piston diameters less than 4 cm in the frequency range up till approximately 2-3 kHz.

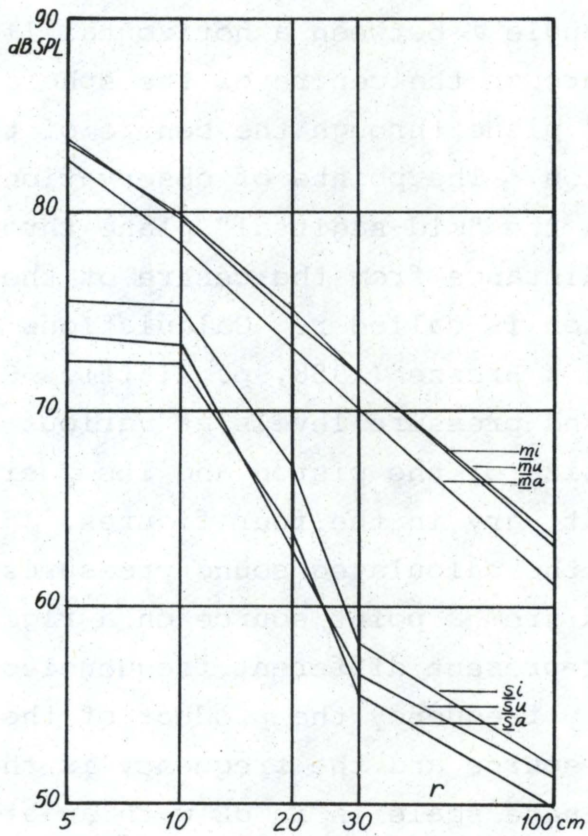


Figure 6

Average sound pressure levels for initial [s] and [m] registered at 5, 10, 20, 30, and 100 cm in front of the lips. The standard deviation for initial [m] is typically 1 dB, whereas the standard deviation for initial [s] is 3-5 dB.

The model with a piston in a rigid sphere gives less trivial results. For the calculations, a sphere with a diameter of 20 cm has been chosen as an approximation to the subject's head. The dimension of the piston is specified by the angle γ_0 between a line from the centre of the sphere to the centre of the piston, and a line from the centre of the sphere to a point on the periphery of the piston. The position of the piston on the sphere is specified by the angle φ_0 between a horizontal line in the "mid-sagittal" plane through the centre of the sphere, and a line in the "mid-sagittal" plane through the centre of the sphere and the centre of the piston. The points of observation are placed on a horizontal line in the "mid-sagittal" plane through the centre of the piston. The distance from the centre of the piston to the point of observation is called r. Calculations are performed according to Stenzel & Brosze (1958, p. 116ff). Figs. 7-10 show the calculated sound pressure levels at various distances from the piston. The size of the piston and its position on the sphere are parameters that vary in the four figures.

Fig. 7 shows the calculated sound pressures at various points in the sound field from a point source on a rigid sphere. The different curves represent different frequencies from 125 Hz to 8000 Hz. For every frequency the product of the volume displacement of the sound source and the frequency is the same arbitrary value and the ordinate scale is in dB with an arbitrary reference. On the average, the calculated sound pressure drops by 6 dB with a doubling of the distance, as predicted by the inverse law. The slope for the higher frequencies is, however, slightly less than 6 dB and slightly more than 6 dB at low frequencies.

Fig. 8 is essentially the same as fig. 7, except that the point source is replaced by a piston ($\gamma_0 = 15^o$) with the same volume displacement. At distances greater than 20 cm, the calculated sound pressures are almost identical for the two types of sound source. Closer to the sound source, a difference is observed. The sound pressures calculated with a piston source are somewhat smaller for a given, small, distance than when calculated on the basis of a point source. This difference is less pronounced at lower frequencies.

In fig. 9, the calculations are again based on a point source, but its position on the sphere is lowered by 30^o compared to the

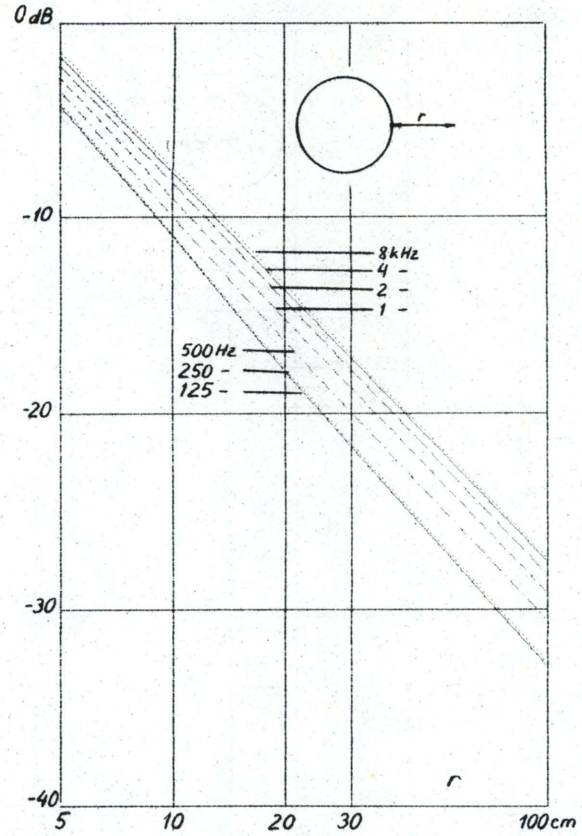


Figure 7

Calculated sound pressure levels from a point source on a rigid sphere in DB, relative to an arbitrary reference.

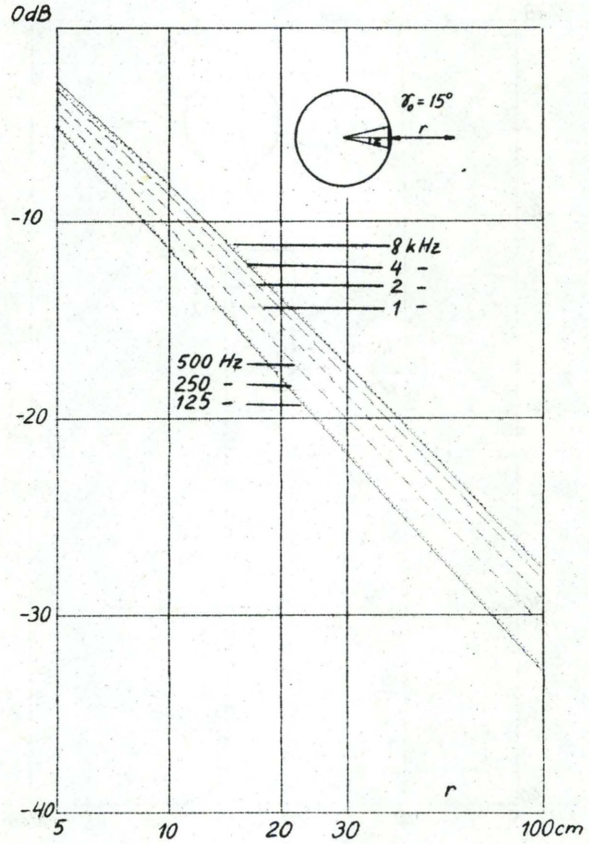


Figure 8

Calculated sound pressure levels from a piston (γ_o = 15°) on a rigid sphere.

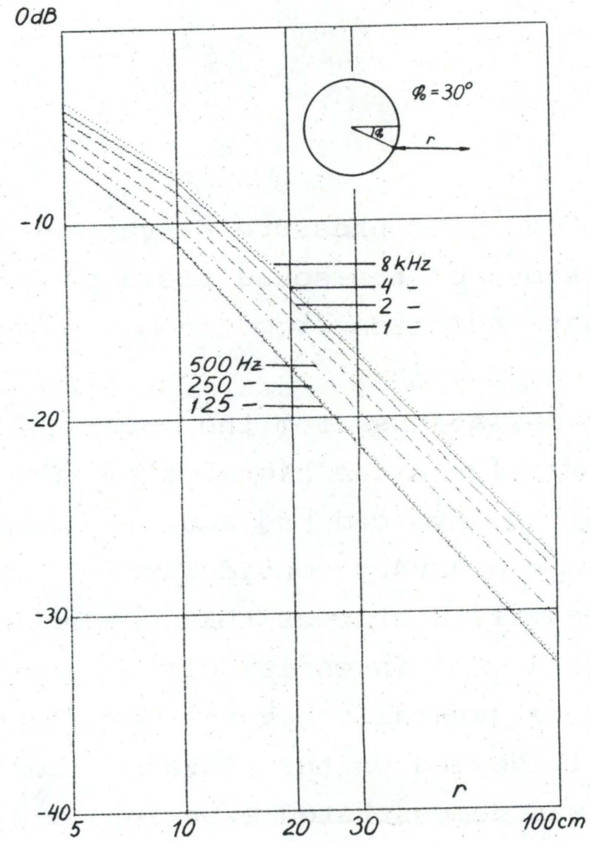


Figure 9

Calculated sound pressure levels from a point source on a rigid sphere. The point source is lowered by 30°, compared with the source in fig. 7.

sound sources in figs. 7 and 8. This causes the curves to be less steep close to the sound source.

In fig. 10 the sound source is a piston ($\gamma_0 = 15^{\circ}$). The centre of the piston is lowered by 30° compared to the situation described in fig. 8. This causes a pronounced deviation from the inverse law at short distances from the sphere, especially for higher frequencies.

5. Discussion

Measurements of sound pressure at various distances from the mouth and calculations of the sound field from a source on a rigid sphere suggest that radiation from the mouth cannot be adequately described by the inverse law.

The greatest deviations from the inverse law are found close to the mouth and mainly in the higher frequency bands. Correspondingly, intensities of unvoiced [s] and, to a somewhat higher degree, of open vowels did vary considerably less with the distance from the mouth at small distances than predicted by the inverse law. This means that the intensity difference between high and low vowels, and more generally between low frequency and high frequency sounds, will depend on the distance from the microphone to the mouth. This was demonstrated experimentally by measuring the intensity of the nasal [m] and the unvoiced fricative [s] simultaneously at various distances from the mouth.

A comparison of the measured and the calculated distance functions shows that these findings can be partly accounted for by a model for the radiation, where a piston is placed on a rigid sphere as described in fig. 10. At short distances, the 4 kHz curve in fig. 10 resembles the curves for [s] in fig. 6 and the 250 Hz curve of fig. 10 resembles the curves for [m] in fig. 6. However, the observed low values of the intensity of [s] at a distance of 30 cm cannot be explained by the model and neither can the tendency for the over-all slope of [s] to be steeper than the slope of [m]. Also, the tendency for the 500 Hz curve in fig. 10 to be less steep at short distances than the 250 Hz curve is inadequate to explain the difference in slope close to the lips which were observed between low and high vowels.

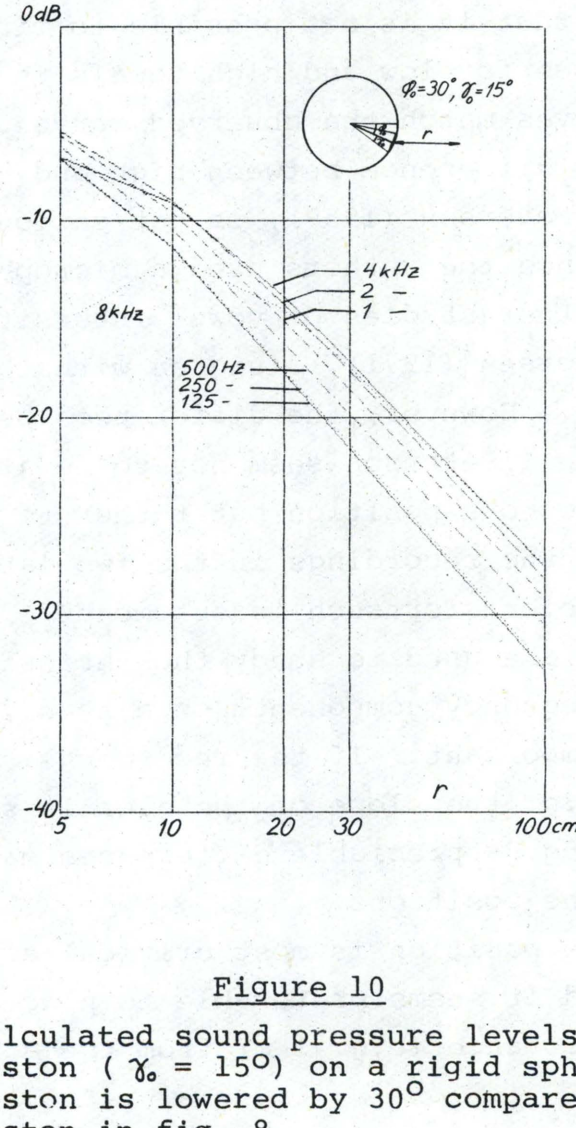


Figure 10

Calculated sound pressure levels from a piston (γ_0 = 15°) on a rigid sphere. The piston is lowered by 30° compared with the piston in fig. 8.

If we assume that the acoustic centre for radiation of speech sounds is placed somewhat inside the mouth, the calculated slopes will be less steep at short distances from the lips. If we further assume that the position of the acoustic centre is closer to the lips for high vowels compared to low vowels, we would expect a less steep slope for low vowels than for high vowels. However, calculations show that it is not possible to find positions for the acoustic centres for low and high vowels in such a way that the calculated curves match the observed curves.

The intensity difference between high and low vowels found by e.g. Lehiste & Peterson (1959), is not seriously affected by these findings, since the authors used a microphone-to-mouth distance of 30 cm. Other studies of vowel intensity, Sharf (1966) and Ludvigsen & Thorsen (1971), disagree with the findings of Lehiste & Peterson. However, the differences between the latter studies and Lehiste & Peterson seem not to be attributable to a difference in microphone position but rather to the dimensions of the rooms used for the recordings in the two latter studies, which were insufficient for such measurements.

If recordings are used to study the intensity relation between different frequency components of a sound, the microphone position will be important. If the recordings, furthermore, are used to estimate e.g. the slope of the glottis spectrum by means of inverse filtering, appreciable differences may appear due to different microphone positions.

The microphone position is most critical at short distances from the mouth, and it seems preferable to place the microphone at a rather long distance (e.g. 1 m) from the mouth to avoid these problems.

References

Dunn, H.K. and D.W. Farnsworth 1939: "Exploration of pressure field around the human head during speech", JASA 10, p. 184-199

Fairbanks, G., A.S. House and E.L. Stevens 1950: "An experimental study of vowel intensities", JASA 22, p. 457-459

House, A.S. and G. Fairbanks 1953: "The influence of consonant environment upon the secondary acoustical characteristics of vowels", JASA 25, p. 105-113

Lehiste, I. and G.E. Peterson 1959: "Vowel amplitude and phonemic stress in American English", _JASA_ 31, p. 428-435

Ludvigsen, C. and N. Thorsen 1971: "Comparison of sound pressure level and loudness (GF) measurements on speech sounds", _ARIPUC_ 5, p. 163-174

Stenzel, H. and O. Brosze 1958: _Leitfaden zur Berechnung von Schallvorgängen_, Springer Verlag, Berlin

Sharf, D.J. 1966: "Variations in vowel intensity measurements", _L&S_ 9, p. 250-256

Strevens, P. 1960: "Spectra of fricative noise in human speech", _L&S_ 3, p. 32-49

(ARIPUC 13, 1979)

ACOUSTIC-STATISTICAL ANALYSIS OF VOICE QUALITY

Børge Frøkjær-Jensen

Abstract: The investigation deals with long-time analyses of:
(1) fundamental frequency,
(2) intensity above 1000 Hz relative to the total intensity, and
(3) total intensity.
Data are sampled 20 times per second. For each parameter, the mean and the coefficient of variance are calculated. The method seems useful for comparisons of changes in prosody taking place e.g. during voice training. Especially the coefficient of variance seems to be a good measure for the dispersion of pitch, "voice quality"[1], and intensity.

1. Purpose

The primary purpose of this investigation is to get an idea about which significant changes in fundamental frequency range, "voice quality", and intensity we can extract from conventional registrations of the fundamental frequency and intensity carried out as long-time analyses. From a phonological point of view the changes are not important, but for the description of speech characteristica for individuals and groups of individuals they are extremely important.

2. Method

2.1 Material

In order to get a normal material, it was decided to use a class of 16 students with normal, healthy voices at a teachers'

---

1) Parameter 2 is in the following text referred to as "voice quality" in quotation marks in order to indicate that there is no unambiguous correspondence between the subjective term voice quality and parameter 2.

college (Poul Hansen, 1979). In connection with their instruction the students were given 15 45-minute lessons of voice and speech training. Tape recordings of the students were made before and after voice training. The tape recordings consisted of easy texts, each of approximately 1 minute's duration. The same text, recording room, etc. was used for both recordings. One student was left out of the investigation because of some noise which made the recordings incomparable.

2.2 Acoustic analysis

When listening to the tape recordings before and after training, some changes are clearly audible in intonation and "voice quality". These changes must be caused by corresponding changes in the acoustic spectrum, primarily in the fundamental frequency, the spectral energy distribution (i.e. the relation between the energy above 1000 Hz and the total energy[1]), and the total intensity[2]. Computing their mean values and dispersions, we may get some statistical measures which will correlate with the audible changes. Only in cases with a breathy and aspirated voice quality parameter 2 is unapplicable because the value of the parameter would be too high. For a good voice the "voice quality" parameter will shift between -10 and -15 dB.

The whole text has been analyzed by a fundamental frequency meter and an intensity meter (F-J Electronics), extracting the mentioned parameters, and the resulting curves have been recorded on a mingograph (Elema 800). The measurements have been carried out on two randomly chosen sentences in the middle of the text (duration of the sentences: approximately 6 seconds).

In these sentences all voiced passages have been measured every 1/20 second (approximately 100 measures per parameter), and transferred to punched cards.

---

1) In other papers (Frøkjær-Jensen and Prytz 1976, Wedin et al. 1977) "voice quality" has been defined as the relation between the energy above and below 1000 Hz, but this was not possible in this investigation because of lack of instrumentation.

2) For practical reasons all intensities have been measured as sound pressures, i.e. a doubling of the intensity does not correspond to 3dB but to 6dB. The term intensity has been preserved because the signals are rectified and integrated over 20 ms.

2.3 Statistical analysis

The data have been processed by an EDP-program which gave the following statistical information:

1. A histogram showing the number of observations per 10 Hz interval for the fundamental frequency, and per 1 dB interval for the remaining two parameters: the intensity relation and the total intensity.
2. Arithmetic mean.
3. Standard deviation.
4. Coefficient of variance ($SD/\overline{X}$).
5. Test of normality based upon a χ^2-test.
6. Skewness with calculated divergence from the normal distribution.
7. Kurtosis with calculated divergence from the normal distribution.
8. Geometric mean, and 20, 25, 50, 75, and 80 percentile points.

The test of normality must be considered a necessary prerequisite for the application of statistical methods based on the normal distribution. The data samples were tested as to normality by means of χ^2-tests. Skewness and kurtosis were also calculated. Table 1 shows the results of the normality test of all data samples. Out of 81 data samples which could be χ^2-tested, only six samples diverged so much that they could not be accepted as belonging to normally distributed data populations.

The coefficient of variance is a measure of the modulation span irrespective of the average being high or low. It appears to be well suited as a measure for the modulation, both for the fundamental frequency, the "voice quality", and the total intensity. The higher the coefficient of variance, the greater the modulatory span of the three parameters.

Concerning the fundamental frequency I have observed (based on more than 100 subjects) that the coefficient of variance varies from below 0.1 for voices with a poor modulation to more than 0.2 for voices with a lively fundamental frequency modulation.

Since the samples in the majority of cases did not deviate significantly from the normal distribution, the comparison of the recordings before and after training were carried out using a series of t-tests.

Table 1

Normality of data populations
+ indicates $P>0.05$ (normal distribution)
0 indicates $0.01<P<0.05$ (borderline case)
- indicates $P<0.01$ (data not normally distributed)

Parameter No. 1 (fundamental frequency):

Subject	1	2	3	4	5	6	7	14	8	9	10	11	12	15	16
before	-	0	+	+	0	-	+	+	+	0	0	+	0	+	+
after	+	+	+	0	+	+	+	+	-	0	+	0	+	0	+

Parameter No. 2 ("voice quality"):

Subject	1	2	3	4	5	6	7	14	8	9	10	11	12	15	16
before	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
after	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Parameter No. 3 (total intensity):

Subject	1	2	3	4	5	6	7	14	8	9	10	11	12	15	16
before	0	+	+		0		0		0	+	+	+	+		
after	-	-	-		0	+	0		+	0	+	+	+		

Table 2

Calculations for subject No. 1
Fo: Fundamental frequency
Alfa: Intensity above 1000 Hz relative to total intensity
Int.: Total intensity

Calculations for subject No. 1	Before voice training			after voice training		
	Fo	Alfa	Int.	Fo	Alfa	Int.
Arithmetic mean	248.1	-15.5	30.3	300.5	-13.5	32.7
Standard deviation	29.1	6.59	6.0	43.6	6.17	5.3
Coefficient of var.	0.117	-0.426	0.199	0.145	-0.456	0.161
χ^2	38.5	9.80	19.2	9.22	13.20	35.8
Probability	0.0001	0.2000	0.0140	0.3241	0.0674	0.0001
Degrees of freedom	8	7	8	8	7	9
Skewness	1.26	-0.38	-1.14	0.27	-0.17	-1.35
Kurtosis	1.20	-0.49	0.96	-0.86	-0.83	1.58
Geometric mean	246.8	-16.1	30.1	297.9	-14.1	32.5

The results of the t-tests were checked, however, by two different non-parametric tests (the Walsh test and the Wilcoxon test). Apart from one instance the three test methods gave identical results regarding significance levels achieved.

3. Results

3.1 Results for an individual subject

3.1.1 Fundamental frequency:

Table 2 shows the results for an individual subject before and after voice training (the corresponding frequency distributions are shown in Figure 1A and Figure 1B). The most important calculations are the arithmetic mean and the coefficient of variance. Comparing the fundamental frequency before and after voice training it is found that the mean frequency has been raised from 248 to 300 Hz. Furthermore, the coefficient of variance has been increased from 0.117 to 0.145[1]. This means that not only has the fundamental frequency been raised, but the dispersion of the fundamental frequency (the intonation range) has also been considerably increased.

3.1.2 "Voice quality"

Table 2 shows the calculations for the "voice quality" (the corresponding histograms are shown in Figure 2A and Figure 2B). The χ^2-tests show that the populations for "voice quality" cannot be considered significantly different from the normal distribution. We may therefore compare the means which show that the "voice quality" is increased +2 dB (from -15.5 to -13.5 dB) during training.

Furthermore, the coefficient of variance has increased by 7% (from 0.426 to 0.456) indicating a tendency to make better use of variations in "voice quality" after the voice training.

---

1) This comparison is not statistically justified because the data before voice training are not normally distributed (see table 1). Instead, we may base the comparison on the geometric mean. However, if we calculate the corresponding coefficient: $(P_{75}-P_{25})$/geometric mean, we get an even better result: 0.121 before compared with 0.268 after voice training.

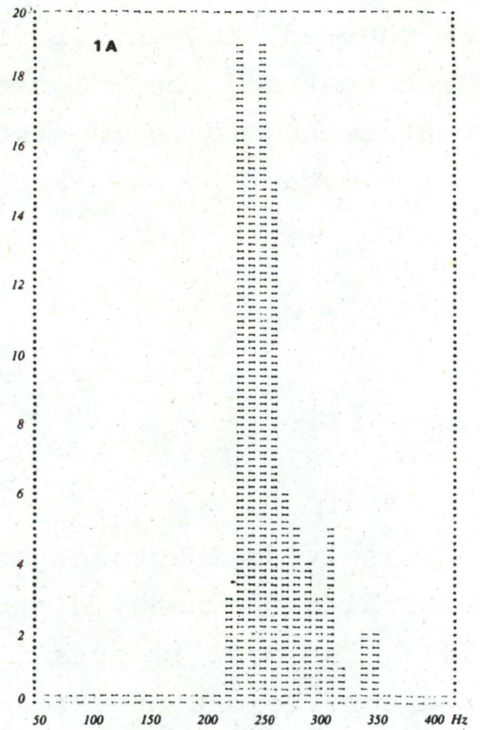


Figure 1A

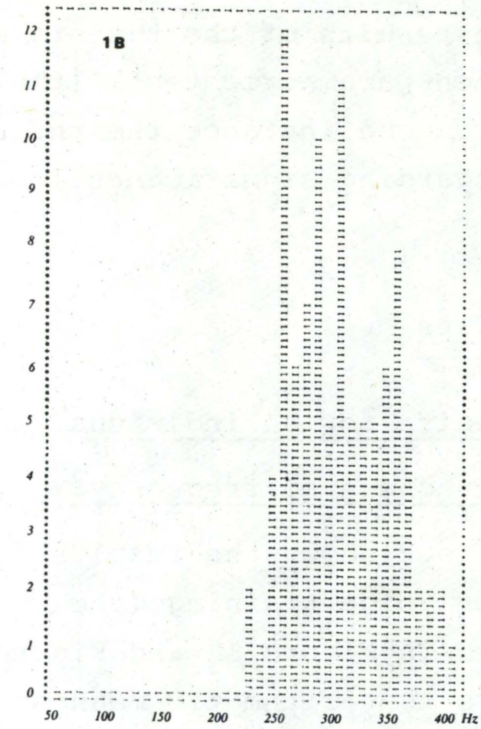


Figure 1B

Histograms showing the distribution of fundamental frequency measurements before and after voice training.

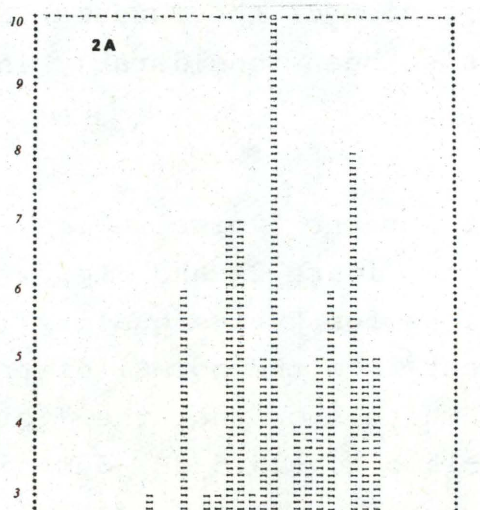


Figure 2A

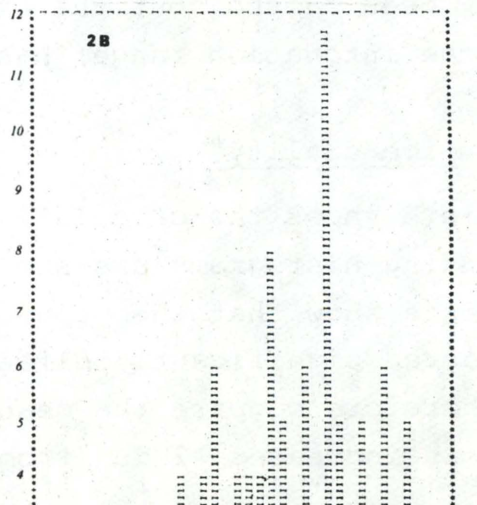


Figure 2B

Histograms showing the distribution of "voice quality" measurements before and after voice training.

3.1.3 Total intensity

Again, refer to table 2 for the calculations for parameter 3. (The corresponding histograms are shown in Figure 3A and Figure 3B.) As shown by the tests, the data populations for the total intensity are not normally distributed. Therefore, we will have to compare the geometric means. Both the geometric mean and the percentiles show that the intensity level has been increased approximately 2 dB during training. These results are only to be considered as a tendency because the two samples cannot be regarded as significantly different. The coefficient of variance is reduced by 19% for this single subject. This is opposite to the normal findings which show an increase of approximately 10% on the average.

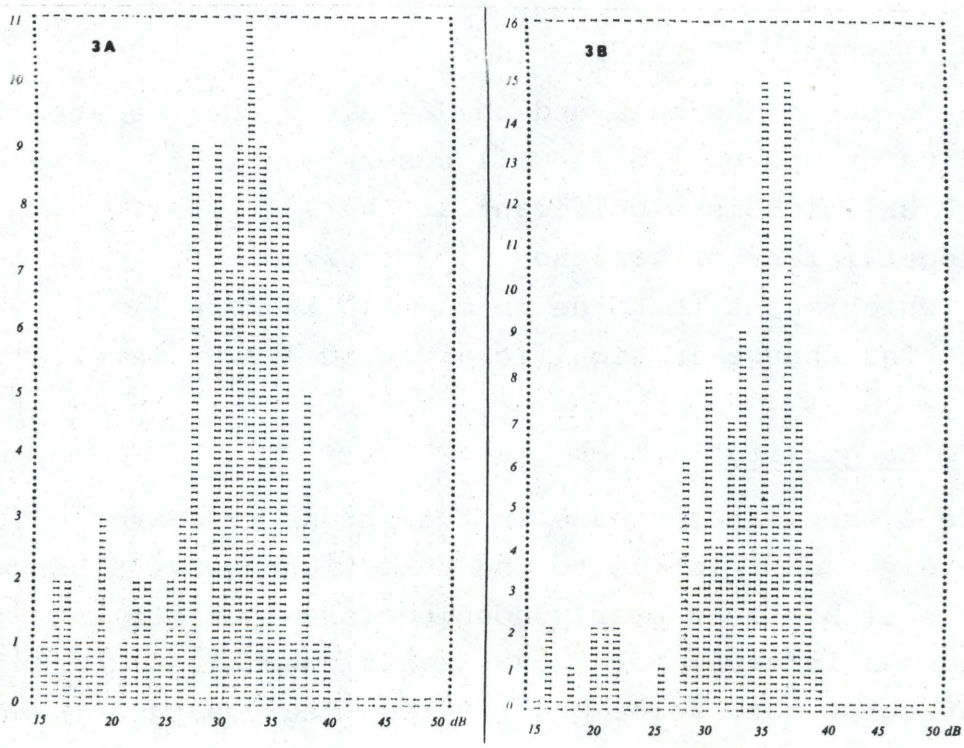


Figure 3A   Figure 3B

Histograms showing the distribution of total intensity measurements before and after voice training.

3.2 Comparisons of the results before and after voice training for all 15 subjects

Tables 3, 4, and 5 show the changes in each of the three parameters during voice training. Most important are the changes in the mean and in the coefficient of variance. There is a clear tendency to change the mean value and to increase the coefficient of variance.

In order to test whether this change is significant, a Student's t-test has been applied to both statistic parameters for all subjects pooled (the two extreme right columns in each table). Roughly, a good agreement has been found between non-parametric tests such as the Walsh and the Wilcoxon test and the Student's t-test, though a few of the data populations differ from the normal distribution.

3.2.1 Fundamental frequency

Table 3 shows the male and the female voices separately. For the two data groups pooled we will observe an increase in F_o of 12.2 ± 10.7 Hz which is significant at the 95% level.

The coefficient of variance is increased by 24% (from 0.122 to 0.151) which means that the intonation is more lively after the training. The change is significant at the 99.5% level.

3.2.2 "Voice quality"

Table 4 shows an increase in the intensity above 1000 Hz of approximately 5 dB relative to the total intensity. This means that the level has been nearly doubled (see footnote 2 on page 190). The confidence interval for 99.9% significance is 5.2 ± 3.2 dB, which means that the "voice quality" has been changed towards a brighter timbre and consequently a better intelligibility.

Also the coefficient of variance has been increased (39% from 0.297 to 0.414). This means that the "voice quality" after the training disperses over a greater voice quality range, - the subjects are using the fluctuations towards the hypo- and the hyperfunctions as part of their natural modulatory pattern. The observed changes are significant at the 99.0% level.

Table 3

Fundamental frequency: comparisons before and after voice training.

$\overline{X}$: average of N_x measurements

SD: standard deviation

SD_x^2: variance of $\overline{X}$

VC_x: coefficient of variance for $\overline{X}$

Df: degrees of freedom

SD_d: standard deviation of the difference

SE_d: standard error of the difference

t: average/standard error (Student's t)

P: probability

K: female subject

M: male subject

	Before training				After training				Difference	
subject	$\overline{x}$	N_x	SD^2_x	VC_x	$\overline{y}$	N_y	SD^2_y	VC_y	$\overline{y}$-$\overline{x}$	VC_y-VC_x
1 K	248,2	100	849	0,117	300,5	99	1897	0,145	52,3	0,028
2 K	239,3	81	942	0,128	276,0	85	941	0,111	36,7	−0,017
3 K	230,8	78	845	0,126	214,6	66	736	0,126	−16,2	0,000
4 K	237,1	92	524	0,097	240,8	72	1053	0,135	3,7	0,038
5 K	230,0	96	465	0,094	239,6	78	874	0,123	9,6	0,029
6 K	245,5	89	720	0,109	227,5	71	1233	0,154	−18,0	0,045
7 K	221,4	86	225	0,068	232,6	84	1178	0,148	11,2	0,080
14 K	201,3	33	1501	0,192	242,3	33	1734	0,172	41,0	−0,020
mean	231,7			0,116	246,7			0,139	15,0	0,023
8 M	110,4	92	99	0,090	130,2	87	500	0,172	19,8	0,082
9 M	111,3	75	163	0,115	112,9	91	274	0,147	1,6	0,032
10 M	150,5	106	513	0,151	155,0	99	1003	0,204	4,5	0,053
11 M	135,5	106	427	0,153	137,3	97	516	0,165	1,8	0,012
12 M	127,1	62	457	0,168	133,0	66	616	0,186	5,9	0,018
15 M	125,5	50	131	0,091	139,6	45	254	0,114	14,1	0,023
16 M	158,6	39	456	0,135	173,6	42	858	0,169	15,0	0,034
mean	131,3			0,129	140,2			0,165	9,0	0,036
mean M+K	184,8			0,122	197,0			0,151	12,2	0,029
								Df	14	14
								SD_d	19,32	0,2947
								SE_d	4,988	0,00761
								t	2,446	3,828
								P <	0,05	0,005
								significance level % >	95,0	99,5
								confidence interval	12,2±10,7	0,029±0,025

Table 4

"Voice quality": comparisons before and after voice training.

subject	Before training $\bar{x}$	N_x	SD^2_x	VC_x	After training $\bar{y}$	N_y	SD^2_y	VC_y	Difference $\bar{y}-\bar{x}$	VC_y-VC_x
1 K	−15,45	100	43,4	0,426	−13,54	99	38,1	0,456	1,91	0,030
2 K	−20,56	81	14,1	0,183	−14,33	85	15,6	0,275	6,23	0,092
3 K	−18,61	76	24,2	0,264	−13,73	66	24,7	0,362	4,88	0,098
4 K										
5 K	−24,71	96	51,4	0,290	−16,08	78	43,2	0,409	8,63	0,119
6 K	−18,26	89	38,9	0,341	−15,20	71	37,6	0,403	3,06	0,062
7 K	−19,64	86	36,4	0,307	−16,45	84	30,6	0,336	3,19	0,029
14 K	−17,91	33	42,8	0,365	−11,00	33	27,5	0,476	6,91	0,111
8 M	−16,48	92	27,1	0,316	−11,91	87	24,7	0,418	4,57	0,102
9 M	−18,95	75	12,5	0,186	−13,70	91	18,8	0,317	5,25	0,131
10 M	−14,05	106	29,6	0,387	−11,56	99	20,3	0,390	2,49	0,003
11 M	−16,35	106	18,5	0,263	−13,88	97	17,1	0,298	2,47	0,035
12 M	−18,69	62	26,9	0,277	−15,98	66	13,8	0,232	2,71	−0,045
15 M	−21,30	50	23,0	0,226	−12,42	45	71,7	0,683	8,88	0,457
16 M	−20,33	39	43,8	0,326	−9,43	42	49,1	0,743	10,90	0,417
mean	−18,66			0,297	−13,52			0,414	5,15	0,117
SE_d									0,751	0,0386
t									6,85	3,036
P <									0,001	0,010
significance level % >									99,9	99,0
confidence interval									5,2 ± 3,2	0,117 ± 0,116

Table 5

Total intensity: comparisons before and after voice training.

subject	Before training $\bar{x}$	N_x	SD^2_x	VC_x	After training $\bar{y}$	N_y	SD^2_y	VC_y	Difference $\bar{y}-\bar{x}$	VC_y-VC_x
1 K	30,3	100	36,2	0,199	32,7	99	27,8	0,161	2,4	−0,038
2 K	34,5	81	17,9	0,122	33,2	85	23,6	0,146	−1,3	0,024
3 K	33,6	76	34,2	0,173	28,2	66	29,3	0,192	−5,4	0,019
4 K										
5 K	34,7	96	39,8	0,182	28,4	78	29,2	0,190	−6,3	0,008
6 K					(28,9)	71	23,9	(0,169)		
7 K	31,0	78	20,0	0,144	28,9	84	26,1	0,176	−2,1	0,032
14 K										
8 M	30,1	92	16,5	0,135	32,4	87	25,2	0,155	2,3	0,020
9 M	32,8	75	42,9	0,200	31,2	91	38,6	0,199	−1,6	−0,001
10 M	34,2	106	40,6	0,186	33,5	99	41,2	0,191	−0,7	0,005
11 M	32,7	106	31,7	0,172	28,6	97	43,3	0,230	−4,1	0,058
12 M	33,3	62	31,6	0,169	33,5	66	37,6	0,183	0,2	0,014
15 M										
16 M										
mean	32,7			0,168	31,1			0,182	−1,7	0,014
SE_d									0,9930	0,00779
t									1,779	1,809
P <									0,15	0,15
significance level % >									85	85
confidence interval									1,7 ± 1,5	0,014 ± 0,013

3.2.3 Total intensity

Table 5 shows a tendency to a minor reduction of the intensity level (2 dB), but the changes are not statistically significant. The reduced intensity may be explained psychologically as a result of the quite audible improvement of the "voice quality" where formant 2 and 3 become stronger. Consequently a significantly clearer voice with a better intelligibility and a greater ability to penetrate the background noise is obtained. The result could be that the subject does not need to talk as loud as before.

No significant change in the coefficient of variance is observed, though there is a tendency to greater variations (8%) in the intensity after voice training.

4. Summary

The results reported show that the dispersion of the single parameters expressed in terms of the coefficient of variance contains so clear information about the changes in the modulation that the three described coefficient of variances could be used as a standardized description of:

for parameter (1): the modulatory span of the fundamental frequency,

for parameter (2): the modulatory span of the "voice quality", and (to a lesser degree, perhaps)

for parameter (3): the span of the intensity modulation.

References

Davis, Steven B. 1958: "Acoustic characteristics of normal and pathological voices", Haskins SR 54

Davis, Steven B. 1976: "Computer evaluation of laryngeal pathology based on inverse filtering of speech", Monograph No. 13, Speech Communications Research Laboratory, Inc., Santa Barbara, California

Frøkjær-Jensen, B. and S. Prytz 1976: "Registration of voice quality" Technical Review 3, Brüel & Kjær.

Hansen, Poul 1979: "Åndedrætsfunktion og stemmedannelse", Dansk Audiologopædi, May.

Kitzing, Peter 1979: Glottografisk frekvensindikering (doctoral dissertation) Malmö, Sweden.

Smith, Svend and Kirsten Thyme 1976: "Statistic research on changes in speech due to pedagogic treatment", Folia Phon. 28, p. 98-103

Snedecor, Georg W. and W.G. Cochran 1976: Statistical methods, 6th edition/ 8th printing (The Iowa State University Press)

Wedin, S., R. Leanderson and L. Wedin 1977: "Evaluation of voice training. Spectral analysis compared with listeners' judgement", 17th International Congress of Logopedics and Phoniatrics (Copenhagen)