ARIPUC VOL. 14.

ISSN 0589-6681

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

COPENHAGEN 1980

ARIPUC VOL. 14.

ISSN 0589-6681

Ts

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

96, Njalsgade
DK 2300 Copenhagen

COPENHAGEN 1980

CONTENTS

PERSONNEL OF THE INSTITUTE OF PHONETICS January 1, 1979 - June 30, 1980

Professor: Eli Fischer-Jørgensen, dr.phil.h.c. (director of the Institute until April 30, 1980)

Associate professors:

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

Peter Holtse, cand.phil. (from March 1, 1979)

Birgit Hutters, cand.mag. (from October 1, 1979)

Nina Thorsen, cand.phil.

Oluf Thorsen, cand.mag.

Assistant professors:

Birgit Hutters, cand.mag. (until September 30, 1979)

Peter Holtse, cand.phil. (until February 29, 1979)

Niels Reinholt Petersen (from August 1, 1979)

Teaching assistants:

Christian Becker-Christensen, cand.mag.

Peter Molbæk Hansen, stud.mag.

John Jørgensen, stud.mag.

Niels Reinholt Petersen, cand.phil.

Lisbeth Strøjer, M.A.

Engineers:

Preben Dømler, B.Sc.

Mogens Hald Kristensen, M.Sc. (from November 1, 1979, temporarily appointed)

Carl Ludvigsen, M.Sc. (until September 30, 1979, temporarily appointed)

Mogens Møller, M.Sc. (on leave)

Technician: Svend-Erik Lystlund

Secretary: Else Parkmann

Teachers from other institutes lecturing at the Institute of Phonetics:

Una Canger, Ph.D. (Institute of Linguistics)
Jørgen Rischel, dr.phil. (Institute of Linguistics)
Henning Spang-Hanssen, dr.phil. (Institute of Applied and Mathematical Linguistics)

PUBLICATIONS BY STAFF MEMBERS

Eli Fischer-Jørgensen	"Indledning" (Introduction) to Roman Jakobson: Elementer, funktioner og strukturer i sproget, Copenhagen, 1979, p. 3-34
Eli Fischer-Jørgensen	"Fonetik", in Københavns Universitet 1479-1979, vol. IX, Copenhagen 1979, p. 401-420
Eli Fischer-Jørgensen	"Temporal relations in consonant-vowel syllables with stop consonants, based on a Danish material", in Frontiers in Speech Communication Research, Festschrift for Gunnar Fant, (eds. B. Lindblom and S. Öhman), 1979, p. 51-68
Eli Fischer-Jørgensen	"Zu den deutschen Verschlusslauten und Affrikaten", Sprache und Sprechen, Festschrift für Eberhard Zwirner zum 80. Geburtstag (eds. K. Ezawa, K.H. Rensch, and W. Bethge), Tübingen, 1979, p. 79-100
Eli Fischer-Jørgensen	"Viggo Brøndal", in Dansk biografisk leksikon III, Copenhagen, 1979, p. 26-27
Eli Fischer-Jørgensen	25 Years' Phonological Comments, (Wilhelm Fink Verlag) München 1979, 262 pp.
Børge Frøkjær-Jensen	"Akustiske-statistiske analyser af stemmen", Dansk Audiologopædi, May 1979
Børge Frøkjær-Jensen (ed.)	The Sciences of Deaf Singing, (Audiologopedics Research Group), Copenhagen, 1980
Nina Thorsen	"Interpreting raw fundamental frequency tracings of Danish", Phonetica 36 (1979), p. 57-78
Nina Thorsen	"Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish", Proc.Phon. 9, Copenhagen (1979), p. 417-423
Nina Thorsen	"Intonation contours in declarative sentences of varying length in ASC Danish", Autumn Conference 1979, Institute of Acoustics, p. 59-62

Nina Thorsen	"A study of the perception of sentence intonation - Evidence from Danish", JASA 67, 1980, p. 1014-1030
Nina Thorsen and Oluf Thorsen	Fonetik for sprogstuderende, 3rd edition, 3rd printing, Copenhagen 1980, 169 pp.
Oluf Thorsen, Ole Kongsdal Jensen, and Karen Landschultz	Fransk Fonetik, 9th revised edition, Copenhagen 1980, 272 pp.

LECTURES AND COURSES January 1, 1979 - June 30, 1980

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 for students of foreign languages other than English, French, German, Modern Greek, Italian, and Russian) were given by Niels Reinholt Petersen. There was one class in the spring semester 1979, one in the autumn semester 1979, and one in the spring semester 1980.

A course in general and French phonetics including practical exercises in the language laboratory (three hours a week) in the spring semester 1979, and courses in general and French phonetics (two hours a week) in the autumn semester 1979 and in the spring semester 1980 were given by Oluf Thorsen.

Courses in general and German phonetics (two hours a week) were given by John Jørgensen in the spring semesters 1979 and 1980.

A course in general phonetics for students of Modern Greek (two hours a week) was given in the autumn semester 1979 by John Jørgensen.

A course in general phonetics for students of Italian (two hours a week) was given in the autumn semester 1979 by John Jørgensen.

A course in general and Russian phonetics (two hours a week) was given in the autumn semester 1979 by Peter Molbæk Hansen.

2. Practical exercises in ear-training and phonetic transcription

Nina Thorsen gave a course for more advanced students (two hours a week) in the spring semester 1979.

Birgit Hutters gave a course for beginners (two hours a week) in the autumn semester 1979, and a course for more advanced students (two hours a week) in the spring semester 1980.

Oluf Thorsen gave a course for advanced students (two hours a week) in the autumn semester 1979.

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

Lisbeth Strøjer gave a course for more advanced students (two hours a week) in the spring semester 1979.

Jørgen Rischel (Institute of Linguistics) gave a course for advanced students (two hours a week) in the spring semester 1979.

Peter Molbæk Hansen gave a course for beginners (two hours a week) in the spring semester 1980.

Una Canger (Institute of Linguistics) gave an introductory course in general linguistics (two hours a week) in the spring semester 1980.

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, 1979, and a course in the physiology of speech (two hours a week) in the autumn semester, 1979.

Peter Holtse gave a course in instrumental physiological phonetics (two hours a week plus individual exercises) in the spring semester 1980.

5. The acoustics of speech

Nina Thorsen gave a course in the acoustics of speech (two hours a week) in the spring semester, 1980.

Peter Holtse and Niels Reinholt Petersen gave a course in instrumental acoustic phonetics (four hours a week plus individual exercises) in the autumn semester, 1979.

Preben Dømler and Nina Thorsen gave a course in elementary mathematics and electronics (two hours a week) in the autumn semester, 1979.

6. Other courses

Eli Fischer-Jørgensen gave courses in German phonetics (two hours a week), one in the spring semester 1979, and one in the spring semester 1980, and a series of lectures on Sound Change (two hours a week) during the autumn semester 1979 and the spring semester 1980.

Oluf Thorsen gave courses in French phonetics (two hours a week), one in the spring semester 1979, and one in the spring semester 1980, and a course in the theory and practice of the language laboratory (one hour a week) in the spring semester 1980.

Nina Thorsen gave a course in English phonetics (two hours a week) in the autumn semester 1979.

Birgit Hutters and Nina Thorsen presided at a series of seminars for advanced students on topics in experimental phonetics (two hours a week) in the spring semester 1980.

Peter Holtse and Birgit Hutters presided at a series of seminars for advanced students on topics in experimental phonetics (two hours a week) in the spring semester 1979.

Peter Holtse and Niels Reinholt Petersen gave an introductory course to the computer system of the laboratory (two hours a week) in the spring semester 1980.

Henning Spang-Hanssen (Institute of Applied and Mathematical Linguistics) gave a course in statistics (two hours a week) in the autumn semester 1979.

Christian Becker-Christensen gave a course in the phonology and phonetics of Danish (two hours a week) in the spring semester 1979.

7. Seminars

John Jørgensen discussed the concept of marking and its application.

Eli Fischer-Jørgensen gave a lecture titled: "What is characteristic of the human language - particularly in comparison with that of animals?".

Niels Reinholt Petersen lectured on the perception of sound duration.

Poul Erik Spliid and Erik Andersen gave a lecture on the acoustic basis of phoneme perception.

Eli Fischer-Jørgensen gave two lectures on the history of linguistic and phonetic research in Denmark.

Zyun'ici Simada gave a lecture titled: "Physiological correlates of Japanese accent patterns".

Seiji Niimi lectured on problems in EMG recordings and processing and on pharyngeal adjustment in Japanese vowels.

Peter Molbæk Hansen gave a lecture on syllable structure as a factor in sound change.

Linda Waugh gave a lecture titled: "The sound shape of language".

Henning Andersen lectured on marking and hierarchy in phonology.

George Allen gave a lecture titled: "The development of rhythm in children's speech".

Ved Kumari Ghai lectured on vowel length and stress in Dogri.

Eli Fischer-Jørgensen presented and discussed the results of an investigation of the perception of stress in Dogri by Danish listeners.

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

Eli Fischer-Jørgensen participated in "Phonologietagung" in Vienna June 28 - July 7, 1980.

Nina Thorsen participated in "Institute of Acoustics, Autumn Conference", Windermere, November 2-4, 1979, and gave a paper: "Intonation contours in declarative sentences of varying length in ASC Danish", in a symposium on "Internordisk sprogforståelse" (Internordic language comprehension) at Rungstedlund, March 24-26, 1980, in "11èmes Journées d'Etude sur la Parole", Strasbourg, May 28-29, 1980, where she gave a paper: "Perception de l'intonation - Expériences sur le danois (parole naturelle)", and in the "2nd Symposium on the Prosody of Nordic Languages", Trondheim, June 19-21, 1980, and gave a paper: "Intonation contours and stress group patterns in declarative sentences of varying length in Advanced Standard Copenhagen Danish".

Nina Thorsen and Niels Reinholt Petersen participated in "'Time' in the Production and Perception of Speech: An Interdisciplinary Colloquium", held in the Phonetics Department of Kiel University, February 22-24, 1979.

Eli Fischer-Jørgensen, Birgit Hutters, Niels Reinholt Petersen and a number of students participated in the "Seventh Swedish-Danish Symposium" at Lund, April 25-26, 1980. The following papers were given: Niels Dyhr: "A pilot investigation of the Fo pattern in American English", Peter Molbæk Hansen: "Syllable structure as a factor in some Danish sound changes", John Jørgensen: "The terms intensive/extensive in Hjelmslev's theory of language", and Niels

Reinholt Petersen: "Assimilation of Fo".

Børge Frøkjær-Jensen lectured on instrumental registration methods in audiologopedics for speech therapists and phoniatrists at Bruxelles, May 14-16 1979, and in June 1979 he presided at a one-week seminar on registration methods in phonetics at the Phonetics Laboratory, University of Riyadh, Saudi Arabia. In February 1980 he gave 18 lectures and a course in clinical glottography at El Nasr Specialized Hospital, Ain Sham University, Cairo.

The Ninth International Congress of Phonetic Sciences was held in Copenhagen, August 6-11 1979. The congress was organized by a Congress Committee:

Preben Dømler	Institute of Phonetics
Eli Fischer-Jørgensen (president)	Institute of Phonetics
Børge Frøkjær-Jensen	Institute of Phonetics
Peter Holtse	Institute of Phonetics
Birgit Hutters	Institute of Phonetics
Inge Knudsen (administrative secretary)	
Carl Ludvigsen	Institute of Phonetics
Svend-Erik Lystlund	Institute of Phonetics
Mogens Møller	Institute of Phonetics
Else Parkmann (secretary)	Institute of Phonetics
Niels Reinholt Petersen	Institute of Phonetics
Jørgen Rischel	Institute of Linguistics
Nina G. Thorsen	Institute of Phonetics
Oluf M. Thorsen	Institute of Phonetics

Nina Thorsen participated in Symposium No. 7, "The relation between sentence prosody and word prosody" as a member of the discussion panel, and contributed a paper: "Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish".

INSTRUMENTAL EQUIPMENT OF THE LABORATORY

The following is a list of instruments that have been purchased or built during the period January 1, 1979 - June 30, 1980.

1. General-purpose electronic instrumentation

   Digital multimeter, Keitley, type 169

2. Equipment for EDP

   Highspeed plotter, Houston, type 12" DP-11

3. Instrumentation for video

   Video camera, Sony, type AVC-3250 CES

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 Copenhague
TCLP	Travaux du Cercle Linguistique de Prague
UCLA WPP	Working Papers in Phonetics, University of California, Los Angeles
Zs.f.Ph.	Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung.

(ARIPUC 14, 1980)

INTONATION CONTOURS AND STRESS GROUP PATTERNS IN DECLARATIVE SENTENCES OF VARYING LENGTH IN ASC DANISH[1]

Nina Thorsen

Abstract: Four subjects recorded eight non-compound declarative sentences, containing from one to eight stress groups. Acoustic analysis reveals a tendency for fundamental frequency range to increase with increased utterance length, but in a non-linear and seemingly random fashion. The increase is brought about by higher starting points as well as lower ending points in the longer utterances. Concomitant with the range increase we find a decrease in overall downdrift in the longer utterances, but degree of downdrift is not simply inversely related to utterance length. With four and more stress groups the intonation contour is decomposed into prosodic phrase groups, i.e. the contour contains discontinuities in the shape of partial resettings. The prosodic phrase group boundaries are determined by but do not exactly coincide with major syntactic boundaries, and the data present an argument in favour of a hypothesis of prosodic categories as distinct entities with a non-isomorphous relation to syntactic structure.

1. Introduction

The relationship between stress and fundamental frequency (Fo) and the intonation contours of various types of short sentences in Advanced Standard Copenhagen (ASC) Danish have been described elsewhere (Thorsen 1978, 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

---

1) Revised and expanded version of a paper published in ARIPUC 13 (1979), p. 1-7. Instead of referring extensively to vol. 13, I have chosen to repeat some passages from the earlier version. An abbreviated edition of this paper will appear in the proceedings of the "2nd Symposium on the Prosody of Nordic Languages, Trondheim 19-21 June 1980".

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). (A detailed account of the stress group and its tonal properties can be found in Thorsen 1980b - this volume, sections 2 and 7.)

The Fo patterns of SGs are predictable and recurrent entities (though allowing for certain context dependent modifications), wherefore 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 1980a.) This concept of intonation contour is different from the current 'topline' or 'baseline' concepts: To Bruce (1979) and Gårding (1979) (Swedish) the overall contour of an utterance is determined by a topline (connecting local Fo

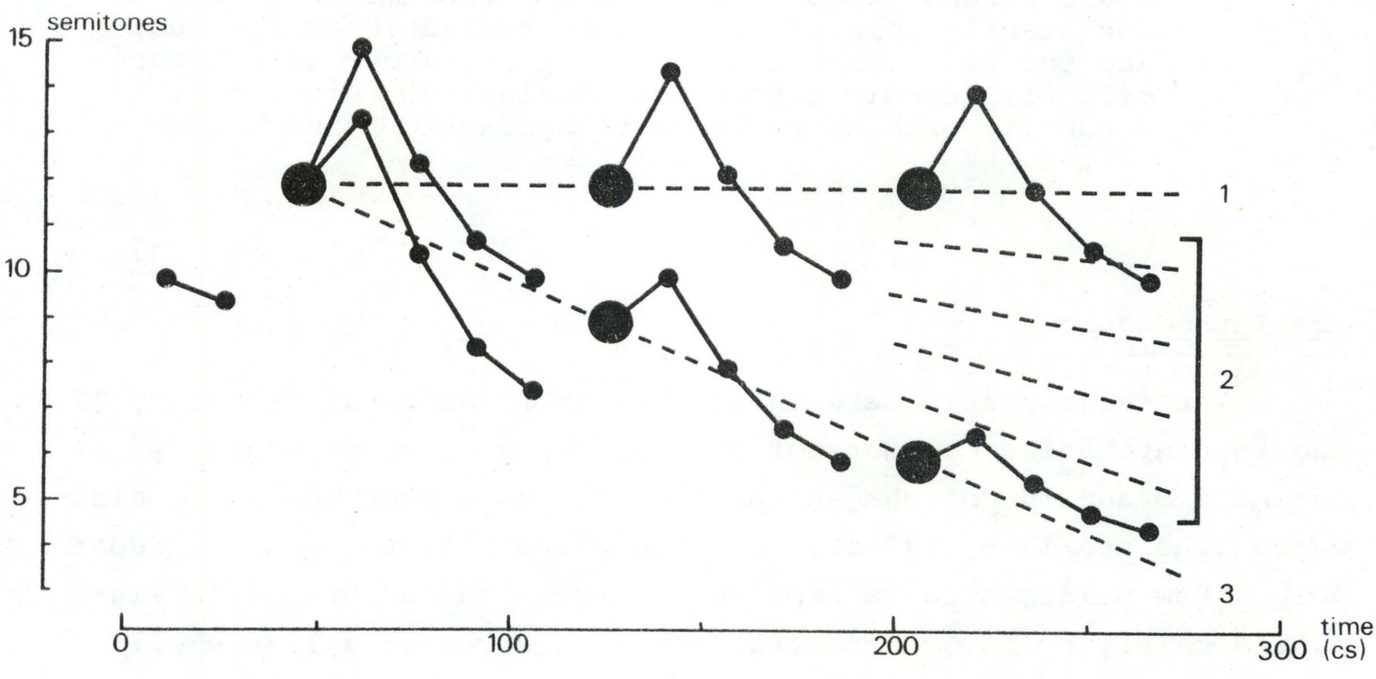


Figure 1

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. The full lines represent the Fo pattern associated with stress groups, and the broken lines denote the intonation contours. Zero on the logarithmic frequency scale corresponds to 100 Hz.

maxima) as well as a baseline (connecting local Fo minima), with the topline declining more rapidly than the baseline. Maeda (1977) (American English) and Fujisaki et al. (1979) (Japanese) attribute the downdrift in declarative sentences to a baseline, but Fujisaki et al.'s baseline (termed the 'voicing component') is an abstraction in the sense that it need have no direct physical representation in the Fo course which is the combined result of the voicing and accent control mechanisms. 't Hart and Collier (1975) (Dutch) have the baseline (connecting Fo minima) as determinant of the declination in declarative sentences. Breckenridge and Liberman (1977), Pierrehumbert (1979), and Sorensen and Cooper (1980) (American English) let the topline (connecting local Fo maxima) determine the downdrift (but, apparently, the maxima always coincide with the stressed syllables of the utterance). Sternberg et al. (1980) (American English) define the downdrift in terms of the stressed syllables, but presumably this would be identical to a topline (their test material contained no unstressed syllables). Neither a topline (connecting local maxima) nor a baseline (connecting local minima) will serve as useful determinants of the overall contours in ASC Danish because they are both highly dependent upon the stress group composition, i.e. on the number of unstressed syllables (if any) in the SG, cf. fig. 1: The baseline may be but is not invariably coincident with a connection of the stressed syllables (if, e.g., the first SG contains more than two unstressed syllables); the topline may be but is not invariably coincident with a connection of the first post-tonic syllables (if, e.g., the second SG contains no post-tonic): in both instances the base- and toplines would steer a zig-zag course downwards. (Furthermore, there are individual differences in stress group patterns: compare fig. 3 (JR) to figs. 2, 4, and 5 and note that JR's unstressed syllables transgress the intonation contour in several instances.) A line connecting the stressed syllables, however, will always exhibit a smooth and gradual course. Furthermore, in an experiment where subjects had to identify utterances as interrogative, non-final, or declarative, solely on the basis of their fundamental frequency course, it turned out that the distribution of subjects' responses was more closely correlated with the stressed syllables in the stimuli (Thorsen, 1980a).

In simple sentences in ASC 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 level contours at the other. In between are found various syntactically marked questions as well as non-final clauses, with a tendency for a trade-off relationship between syntax and intonation contour: The more syntactic information is contained in the sentence about its interrogative or non-final function, the more declarative-like is its intonation contour, and vice versa, cf. fig. 1 (broken lines).

The literature on intonation contours in sentences of varying length is generally only concerned with declarative utterances. There seems to be consensus on an overall downdrift being characteristic of such utterances, i.e. downdrift is a global rather than local phenomenon, but descriptions vary with respect to the extent and shape of the declination. A majority of the authors cited above adhere to the simplest possible model where range is constant over utterances of different length and consequently the rate of the downdrift is inversely proportional to the length of the utterance it spans. This is true of Bruce (1979) and Gårding (1979), the numerous works of Cohen, Collier and 't Hart, explicited in 't Hart (1979), Weitzman (1970) (Japanese, cited from Ohala 1978), Hirose (1971) (Japanese, cited from Ohala 1978), Silverstein (1976) (Hausa, cited from Ohala 1978), Sternberg et al. (1980), and Maeda (1977); Pierrehumbert (1979) finds support for this model in perceptual experiments. McAllister (1971) and Sorensen and Cooper (1980) find that range increases with increased length: the longer utterances start higher than the shorter ones, whereas the lower limit is nearly constant. (An examination of Sorensen and Cooper's data reveals, however, that in addition to the range variation, there is also a slope variation: the longer utterances have less steep slopes than the shorter ones.)

McAllister (1971), Fujisaki et al. (1979), and Sorensen and Cooper (1980) deviate from most other writers on the subject who describe the downdrift in terms of straight lines. Common to their descriptions is a more rapid decline in the early part of the utterance.

For ASC Danish I have previously hypothesized (1979b) that range would be constant over utterances of different length, and that slope would vary inversely with length, and the stressed syllables between the first and last ones would be equidistantly spaced on the (logarithmic) frequency scale. The experiments reported below were designed to test this hypothesis.

2. Material, subjects, and procedures

2.1 Material

Since declarative sentences have the widest range (cf. fig. 1), differences in slope would be most easily detected in them. Accordingly, eight simple statements were made up, containing from one to eight stress groups, all variations on the same theme (´ denotes the stressed vowels and the vertical bar denotes the boundaries between noun phrase and verb phrase, between verb phrase and (compound) complement, and between the two complements):

1. Til Thísted.
2. Túkke | skal til Thísted.
3. Búster | skal med bússen | til Thísted.
4. Kísser | skal med bússen | til "Kílden" i Thísted.
5. Líssi | skal med bússen | til fésten | på "Kílden" i Thísted.
6. Aníta | skal med bússen | til fésten for Kísser | på "Kílden" i Thísted.
7. Hútters | skal med bússen | til fésten for Kísser og Líssi | på "Kílden" i Thísted.
8. Knúdsen og Bítten | skal med bússen | til fésten for Kísser og Líssi | på "Kílden" i Thísted.

(Sentence no. 8 translates as follows: Knudsen and Bitten are taking a bus to the party for Kisser and Lissi at "Kilden" in Thisted.) The stressed vowels are all short, high (except [ɛ] in 'festen'), and surrounded by unvoiced obstruents (except [l] in 'Kilden' and 'Lissi', and [n] in 'Anita' and 'Knudsen') in order to facilitate the subsequent interpretation of the tracings (cf. Thorsen 1979a). - Note that the syntactic boundaries all occur after the first post-tonic syllable in the stress groups.

The sentences 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 (on two separate occasions), giving a total of six recordings of each sentence by each speaker.

2.2 Subjects

Four phoneticians recorded the material, three ASC speakers (NRP male, BH and NT female) and one with a slightly more conservative pronunciation (JR male).

2.3 Procedures

The recordings were made with semi-professional equipment (Revox A-77 tape recorder, Sennheiser MD21 microphone, larynx microphone) in a quasi-damped room at the Institute of Phonetics. The tapes were processed by hard-ware intensity and pitch meters (F-J Electronics) and registered on a Mingograph (Elema 800). 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 allowing for a measuring accuracy of 1 Hz for males and 2 Hz for females.

Fo of each of the vowels and syllabic consonants was measured at 2/3 of the distance from vowel/consonant onset (cf. Rossi 1971, 1978) which was an uncontroversial procedure since all the vowels/consonants had monotonically falling movements, excepting a few instances where the first post-tonic was rising-falling and was measured at its maximum. The distance in time of each of these points from the onset of the first stressed vowel was also measured. The average Fo measurements were converted to semitones (re 100 Hz) and a correction made for intrinsic Fo level differences between stressed [ˈu], [ˈɛ], and [ˈi], in accordance with Reinholt Petersen's (1978) results: [ˈɛ] is raised by 1.2 semitones; [ˈu] is lowered by 0.5 semitones with BH and NT, and by 0.25 semitones with NRP and JR. No correction was attempted for the unstressed vowels or syllabic consonants, cf. Reinholt Petersen (1979). - The standard deviations on the mean Fo values are generally small. E.g. in the longest sentence (no. 8) they range between 2.5% and 3.5% of the mean for the stressed vowels and between 3% and 4% of the mean for the unstressed vowels/syllabic consonants. This means that production stability is rather great and the figures to follow must be fairly reliable indications of subjects' behaviour.

3. Results

Stylized Fo tracings of the eight sentences are depicted in figs. 2-5 for individual subjects, the grand mean in fig. 6. - It is immediately clear that range is not constant over the utterances, nor are the stressed syllables equidistantly spaced on the frequency scale. (Note that in some stress groups with NRP and BH, namely the last but one in sentences 4-8, a syllable is apparently missing. These were instances where the Fo maximum was only reached in the second post-tonic vowel, which is the one shown in figs. 2 and 4. The behaviour of syllables with assimilated schwa and syllabic consonants are the object of a separate investigation, see Thorsen, forthcoming.)

3.1 Range

Range may be defined in different manners, but no matter what definition we choose, the hypothesized constancy is lacking.

3.1.1 Range determined by the stressed vowels

Fig. 7a depicts the variation in range as defined by the interval between the first and the last stressed vowel. Although there is an overall trend for range to increase with increased number of SGs, range does not increase monotonically with length. Combining this information with the fact that the longer utterances seem to be composed of two and three gradients, respectively, with partial resettings between them (located at the broken lines in figs. 2-6, see further below), a relationship between range variation and the number of partial resettings of the intonation contour suggests itself: introducing a partial resetting might decrease the range. Sentences 2 and 3 contain no discontinuity, sentences 4-6 contain one, and sentences 7 and 8 contain two par-

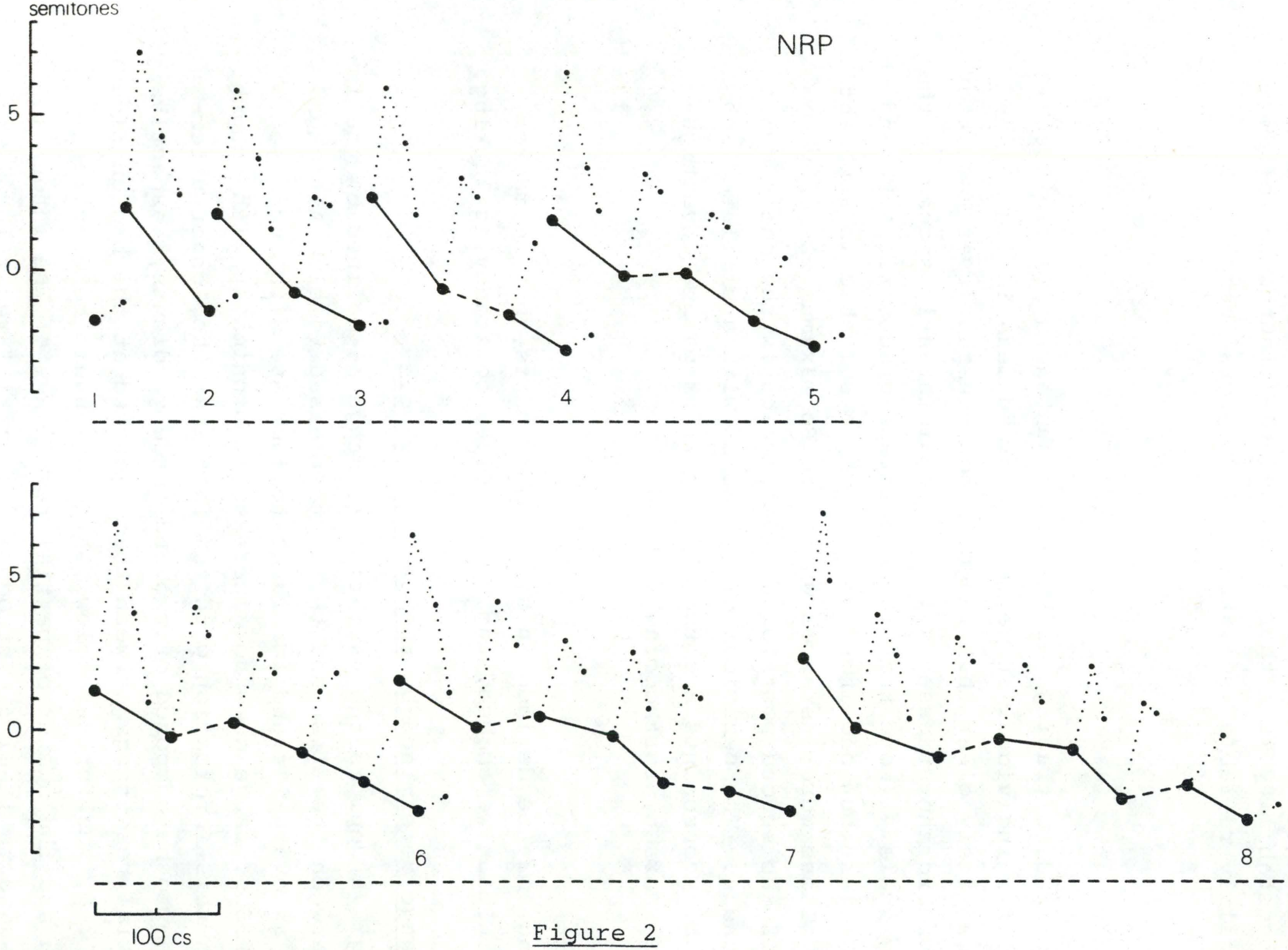


Figure 2

Intonation contours (full and broken lines) and stress group patterns (dotted lines) in declarative sentences containing from one to eight stress groups. Large dots represent stressed syllables, small dots unstressed syllables. Zero on the logarithmic frequency scale corresponds to 100 Hz. Subject: NRP.

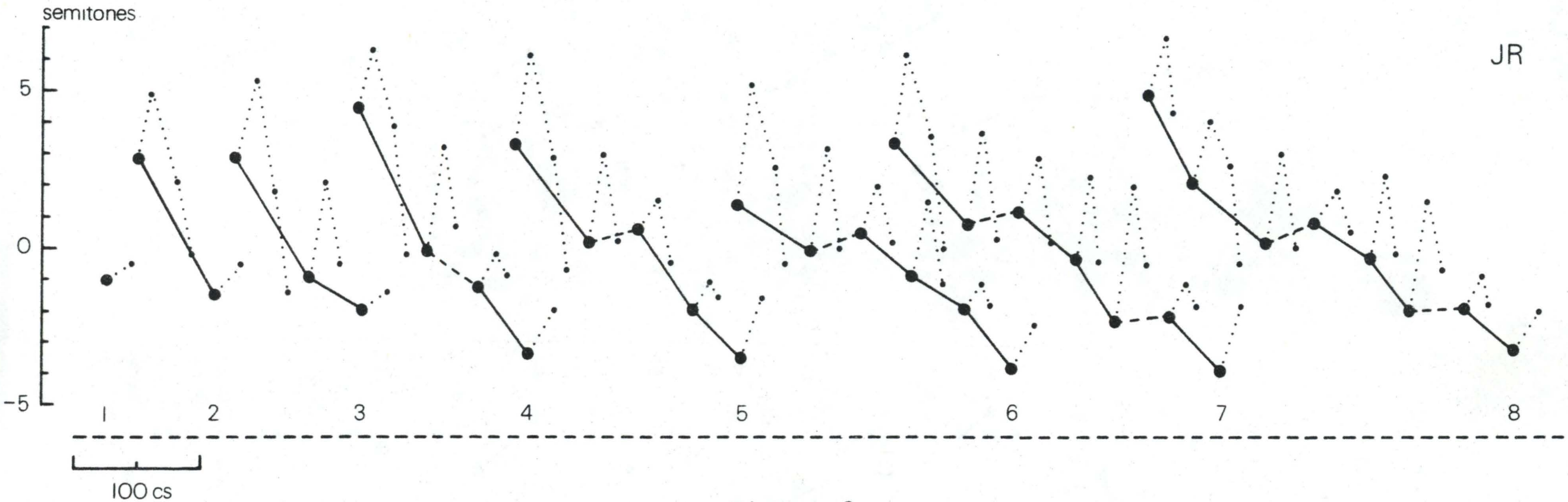


Figure 3

Subject: JR. See further the legend to figure 2.

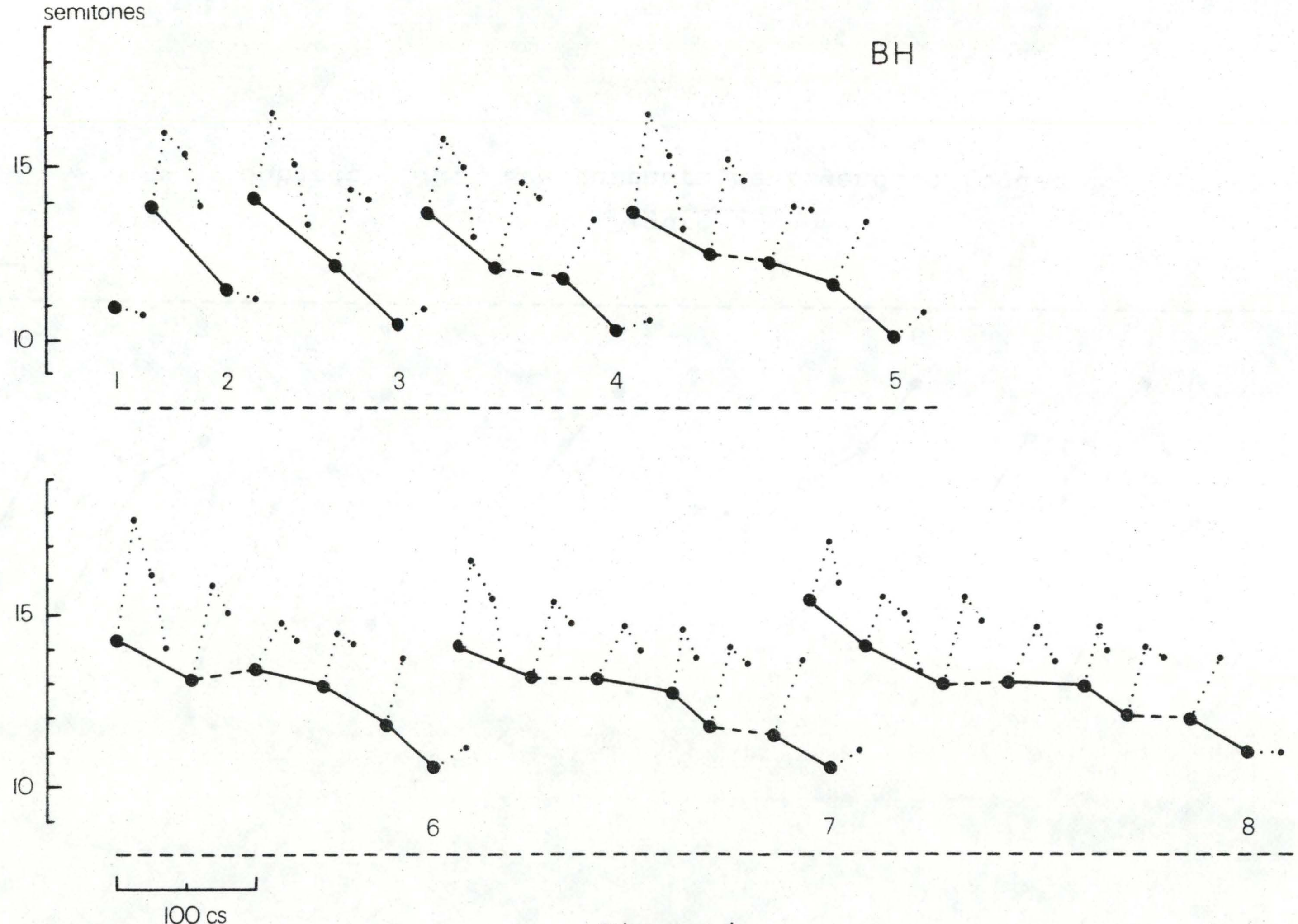


Figure 4

Subject: BH. See further the legend to figure 2.

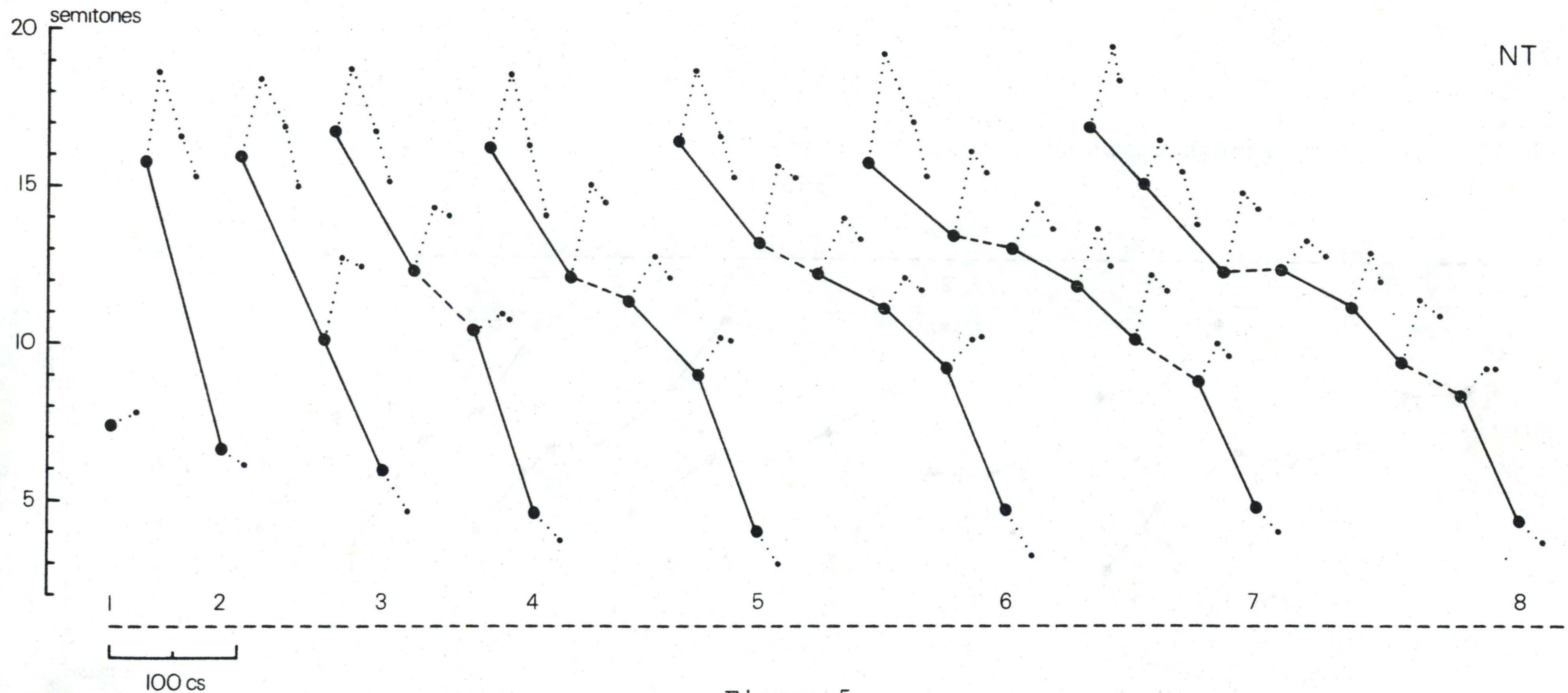


Figure 5

Subject: NT. See further the legend to figure 2.

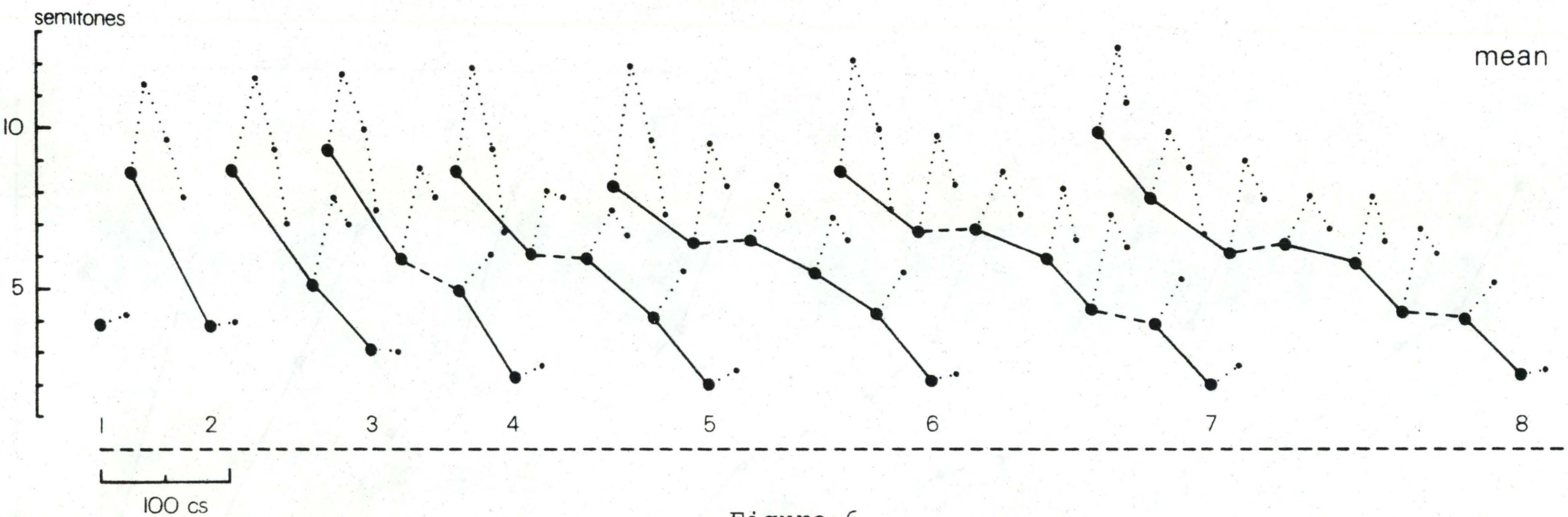


Figure 6

Average over four subjects. See further the legend to figure 2.

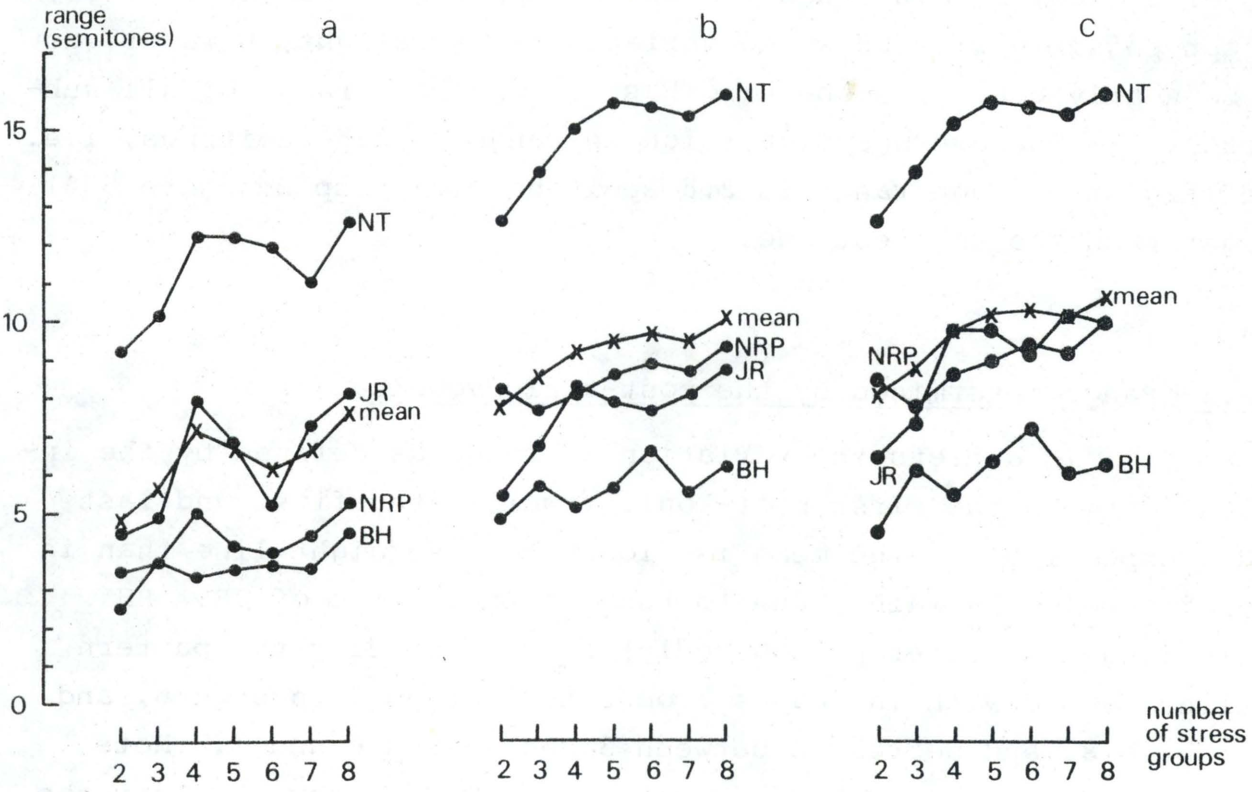


Figure 7

Range of fundamental frequency in seven declarative utterances, containing from two to eight stress groups, depicted as a function of the length of the utterance (i.e. number of stress groups). Four subjects and their grand mean (crosses). In (a) range is defined as the interval between first and last stressed vowel measurement in each utterance; in (b) range is defined as the interval between the first post-tonic vowel in the first and last stress group in each utterance; in (c) range is defined as the interval between the absolute Fo maximum (i.e. the first post-tonic vowel in the first stress group) and the absolute Fo minimum (i.e. the last stressed vowel) in each utterance. (NT's Fo minimum is constituted by the last post-tonic vowel and the curve is identical to the one in (b).) Zero on the logarithmic frequency scale corresponds to 100 Hz.

tial resettings. Range increases, as expected, from 2 to 3; it does not decrease from 3 to 4 (except with BH); it decreases through 4 to 6 (except with BH); it increases from 6 to 7 with NRP, JR, and the mean and decreases (as hypothesized) with BH and NT; it increases, as expected, from 7 to 8.

Concludingly, it seems that range increases with increased number of SGs, but in a non-linear and apparently random fashion, except possibly with BH whose variation is so slight, however, as to be nearly constant (she also has the smallest range of all subjects). - The average dispersion in range is 2.8 semitones, i.e. the largest average range is 2.8 semitones (corresponding to 58%) larger than the smallest one.

3.1.2 Range determined by the post-tonic vowels

Fig. 7b depicts the variation in range as defined by the interval between the first post-tonic vowel in the first and last SGs, respectively. The mean is closer to a straight line than in fig. 7a, which is mainly due to the smoother curve of JR. BH again (and now rather pronouncedly) follows exactly the pattern outlined above with increases from 2 to 3, from 4 through 6, and from 7 to 8, and decreases between 3 and 4, and 6 and 7. Note that the range spanned by the post-tonics is greater than for the stressed vowels, which is a reflection of the fact that the rise from stressed to post-tonic decreases from the earlier to the later parts of the utterance, cf. fig. 1, a phenomenon similar to the commonly noted faster decrease of toplines than baselines, cf. Gårding (1979), Breckenridge and Liberman (1977), and Sorensen and Cooper (1980). Interestingly, this faster decrease is carried by the unstressed syllables in ASC Danish, but by the stressed syllables (the topline) in American English. The average dispersion in range is 2.3 semitones, corresponding to 30% of the smallest one.

3.1.3 Range determined by the absolute Fo maximum and minimum

Fig. 7c depicts the variation in range as defined by the interval between the maximum, i.e. the first post-tonic vowel in the first SG, and the minimum, i.e. the last stressed vowel (except

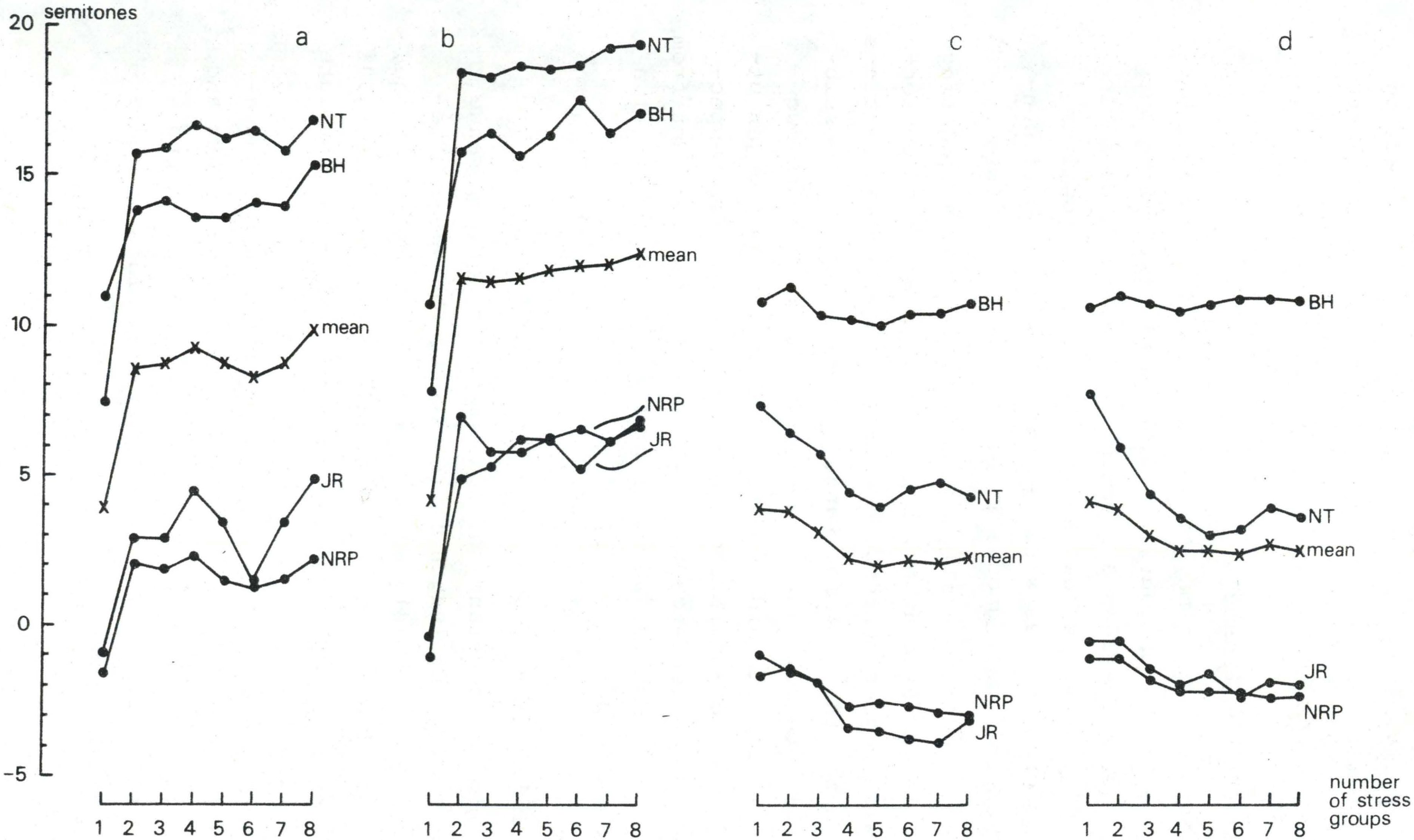


Figure 8

Fundamental frequency starting and end points in eight declarative utterances, containing from one to eight stress groups, depicted as a function of the length of the utterance (i.e. number of stress groups). Four subjects and their grand mean (crosses). In (a) is depicted the frequency of the first stressed vowel in each utterance; in (b) the frequency of the first post-tonic vowel in each utterance. In (c) is depicted the frequency of the last stressed vowel and in (d) the last post-tonic vowel in each utterance. Zero on the logarithmic frequency scale corresponds to 100 Hz.

that NT's minimum is the last post-tonic vowel, cf. fig. 5, and her curve is identical to the one in fig. 7b). The pattern rather resembles that of fig. 7b, cf. above. - The average dispersion in range is 2.5 semitones, corresponding to 31% of the smallest range.

3.1.4 Starting and end points

The variation in range may be due to variation in starting and/or end points of the contours. Fig. 8a depicts the frequency of the first stressed vowel as a function of number of SGs in the utterance, and fig. 8b shows how the first post-tonic vowel in the first stress group varies with length. If we disregard sentence no. 1 which obviously groups itself with the endpoints (cf. fig. 8c and d) we note a slight trend towards higher starting points with the longer utterances, but the curves are not monotonically rising with higher sentence number and the fluctuation seems to be random, as was the case for range variation (except, possibly, again for BH). Concomitant with higher starting points, we find a (stronger) trend towards lower end points in the longer utterances (fig. 8c and d - stressed and post-tonic vowels, respectively) but there is a tendency for endpoints to stay constant from sentence no. 4/5 and upwards, which is probably a reflection of a physiological constraint.

3.1.5 Range - conclusion

The hypothesized constancy of fundamental frequency range over utterances of varying length is refuted by the data. The largest average range is nearly 60% greater than the smallest one, if range is defined in terms of the stressed vowels, and about 30% if range is determined by the post-tonic vowels or by the absolute Fo maxima and minima.

The increase in range with longer utterances is brought about by a combination of higher starting and lower end points and, contrary to the results of McAllister (1971) and Sorensen and Cooper (1980), up to five stress groups the end points decrease more than the starting points increase.

Table 1

Least squares regression line slopes of the stressed vowels and of the post-tonic vowels depicted in figs. 2-6, and their correlation coefficients.

	Sentence no.	NRP Slope/corr.	JR Slope/corr.	BH Slope/corr.	NT Slope/corr.	mean Slope/corr.
stressed vowels	2	-5.10	-7.24	-4.50	-15.11	-7.93
	3	-3.31/ -.98	-4.82/ -.98	-3.57/-1.00	-8.96/-1.00	-5.21/-1.00
	4	-2.95/ -.97	-5.61/ -.98	-2.21/ -.96	-7.14/ -.98	-4.53/ -.99
	5	-1.81/ -.97	-3.66/ -.97	-1.66/ -.96	-5.26/ -.97	-3.13/ -.98
	6	-1.31/ -.95	-2.14/ -.94	-1.37/ -.93	-4.00/ -.96	-2.23/ -.96
	7	-1.26/ -.96	-2.66/ -.97	-1.20/ -.97	-3.23/ -.96	-2.10/ -.97
	8	-1.09/ -.89	-2.42/ -.95	-1.16/ -.95	-3.20/ -.97	-1.95/ -.96
unstressed vowels	2	-10.52	-7.83	-7.27	-18.38	-11.00
	3	-6.43/-1.00	-6.37/-1.00	-5.21/ -.99	-11.76/ -.99	-7.56/ -.99
	4	-4.57/ -.99	-6.03/ -.99	-3.14/ -.95	-8.57/ -.97	-5.64/ -.99
	5	-3.52/ -.99	-4.39/ -.99	-2.55/ -.96	-6.64/ -.96	-4.12/ -.98
	6	-2.95/ -.99	-3.36/ -.99	-2.36/ -.96	-5.25/ -.96	-3.48/ -.98
	7	-2.30/ -.98	-2.91/ -.97	-1.61/ -.93	-4.28/ -.96	-2.82/ -.97
	8	-1.98/ -.95	-2.54/ -.96	-1.49/ -.91	-3.79/ -.96	-2.42/ -.96

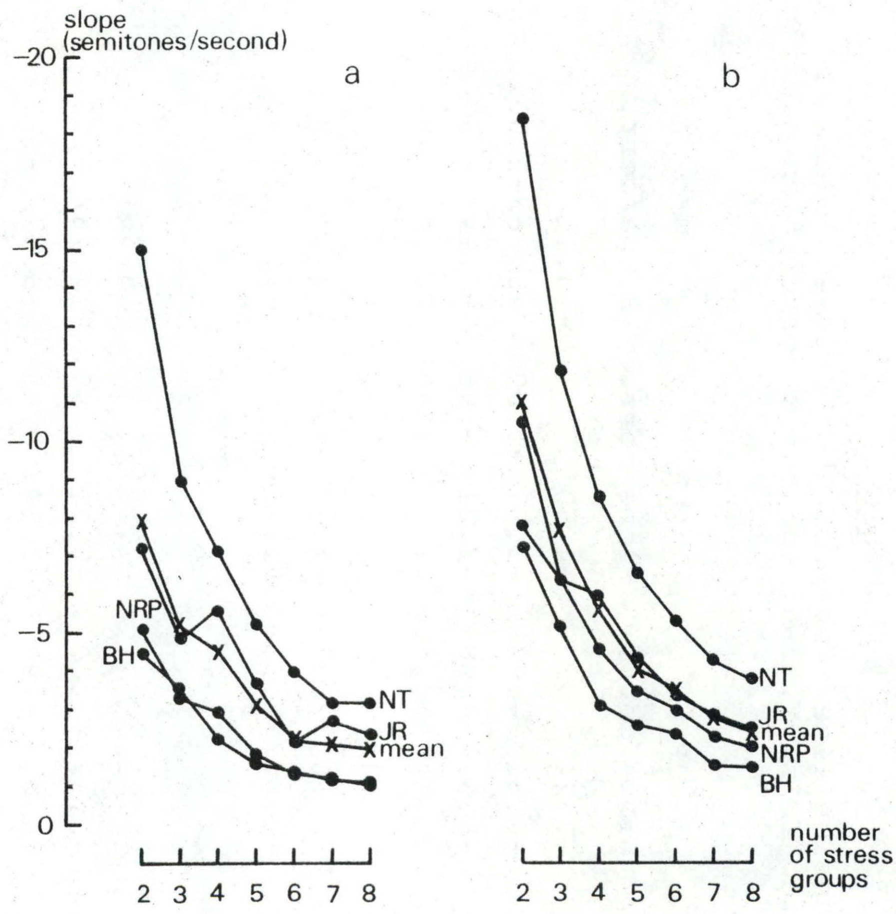


Figure 9

Slope of the overall downdrift of seven utterances, containing from two to eight stress groups, depicted as a function of length of the utterance (i.e. number of stress groups). Four subjects and their grand mean (crosses). In (a) is depicted the downdrift as determined by the stressed vowels; in (b) downdrift is determined by the post-tonic vowels (the "topline" - see further the text).

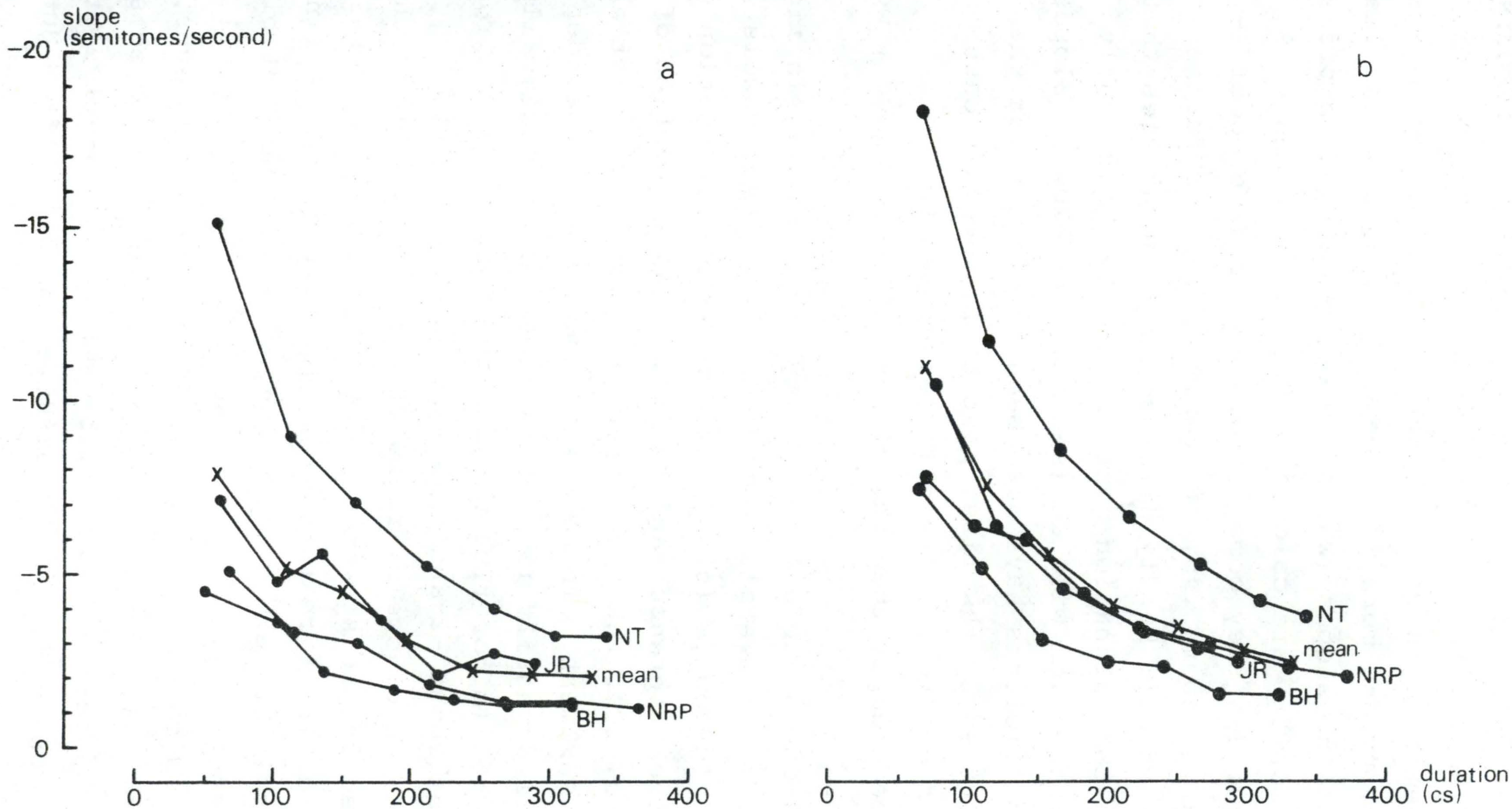


Figure 10

Slope of the overall downdrift of seven utterances, containing from two to eight stress groups, depicted as a function of the duration of the utterance (i.e. the time interval between first and last vowel measurement). Four subjects and their grand mean (crosses). In (a) is depicted the downdrift as determined by the stressed vowels; in (b) downdrift is determined by the post-tonic vowels (the "topline" - see further the text).

4. Intonation contours

4.1 Overall downdrift

It is apparent from figs. 2-6 that the intonation contours are not straight lines, i.e. the medial stressed vowels do not occur with equal semitone intervals, and the irregularity generally sets in with four and more SGs. Nevertheless, when the least squares regression line slopes (which may be taken as an expression of the overall downdrift) are calculated for the stressed vowel data points of figs. 2-6 they come out with correlation coefficients generally at or above .95, cf. table 1, and any further statistical treatment will hardly disclose the regularities that can be observed in the intonation contours any more succinctly than mere visual inspection will do (see further below). Slopes and correlation coefficients have also been calculated for the data points constituted by the first post-tonic vowel in each stress group (the "topline"). These latter slopes turn out (as expected) to be steeper than the corresponding stressed vowel slopes, cf. section 3.1.2.

The information in table 1 is displayed graphically in figs. 9 and 10. In fig. 9 the overall downdrift - stressed vowels (a) as well as post-tonic vowels (b) - is depicted as a function of number of SGs in the utterance, and in fig. 10 as a function of the actual duration, i.e. the time interval between the first and the last vowel measurement. In both instances, but most evidently in fig. 10, overall downdrift is seen to decrease asymptotically with increased length (minor fluctuations occur with JR), approaching a mean value of about -2 and -2.5 semitones per second (stressed and post-tonic vowels, respectively). Presumably, this non-linear decrease in the steepness of the slope of the overall declination is a pertinent feature of longer utterances. A linear decrease would result in zero declination, which is incompatible with declarative sentences. - The preservation of an overall downdrift is secured through a widening of the range. However, there are physiologically determined upper limits to a subjects total range, and theoretically we might envisage declarative utterances that are long enough to jeopardize the downdrift. But I think this problem is purely academic: spontaneous speech rarely contains utterances as long as, say, sentences 7 and 8 that do not have internal clause boundaries or breath group pauses, which presumably both could lead to a decomposition into what might be termed intonational phrases, each with its own intonation contour (downdrift). Even for read speech, sentences 7 and 8 are rather long and it is conceivable that if subjects were compelled to expand them even further, they would indeed introduce pauses at convenient places.

4.2 Prosodic phrase group boundaries

If we term discontinuities those places in the contours where the slope of the line connecting two stressed vowels is less steep than the preceding as well as succeeding slopes, then sentences 2 and 3 contain no discontinuity, sentences 4, 5, and 6 contain one, and sentences 7 and 8 contain two, for individual subjects as well as the mean, denoted by the broken lines in figs. 2-6. (Whether sentence 6 with NT actually has the discontinuity in the indicated place is debatable - according to the definition just given, it occurs rather between the third and fourth stressed vowels.) Thus, the longer utterances seem to be composed of two and three, respectively, prosodic phrase groups, the boundaries between which coincide with major syntactical boundaries, viz. before the (compound) complement and between the purpose and place complements. Note that these prosodic phrase group boundaries are not accompanied by pauses but seem to be caused by the syntactic structure per se, as pointed out also by e.g. Cooper and Sorensen (1977) and Fujisaki et al. (1979). The degree of resetting of the intonation contour at these boundaries is slight with BH and NT, somewhat larger with NRP and JR where it takes the form of actual rises across the boundary in most instances. (Incidentally, the cross-boundary rises 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.)

Due to the particular behaviour of the post-tonic syllables (cf. section 5 below) the prosodic phrase group boundaries cannot be detected in the "topline" in the same manner as in the intonation contour proper.

4.3 The shape of the downdrift

Even though straight lines are good approximations to the overall downdrift of the stressed syllables (cf. the high correlation coefficients in table 1), it is also apparent from figs. 2-6 that with some of the utterances by NRP, JR, and BH (but not NT) the trend is towards greater declination in the early part of the utterances: sentences 3-8 by NRP, 3-5 and 8 by JR, and sentence 8 by BH (and sentences 3-5 and 8 of the grand mean). This is in line with the results of McAllister (1971) and Sorensen and Cooper (1980).

The "topline" does not show the same trend, partly because pre-boundary post-tonics exhibit higher rises (cf. section 5 below), partly because the rise in the final stress group is considerably smaller than in preceding stress groups with NRP, BH, and NT (where it actually falls) and thus the steepest "topline" declination is located towards the end of the utterances.

In order to compare the Danish to Fujisaki et al.'s Japanese data, we should look at individual gradients ('voicing components'): the second prosodic phrase group of sentences 5-7 and the first and second one of sentence 8 contain more than two stressed syllables and could thus exhibit an exponential decay. NRP shows a tendency for such a decay in two instances (sentence 5, and the first gradient of no. 8) but quite the reverse in the second prosodic phrase group of sentences 7 and 8. JR has a tendency for asymptotic declination in sentence 5 and the first gradient of sentence 8, but a reversal in 6 and 7, as well as in the second gradient of no. 8. BH exhibits faster decays at the end of the prosodic phrase groups in sentences 5-7 as well as in the second gradient of no. 8, a pattern repeated by NT. Exponentially decaying "toplines" are equally scarce.

4.4 Intonation contours - conclusion

The hypothesized simple inverse relationship between intonation contour declination and utterance length is not supported by the data. At and above 4 stress groups, the intonation contour, as defined by the stressed syllables, is decomposed into prosodic phrase groups, with partial resettings of the contour between them. However, an overall downdrift <u>is</u> preserved, which <u>does</u> become less steep with increased length (although the relationship is not a linear one), and which, further, exhibits a tendency towards exponential decay, i.e. greater declination in the beginning of the utterance.

5. Stress group patterns

On the basis of the 1978 analysis of ASC Danish, the stress group was defined as a stressed syllable plus all succeeding unstressed syllables, irrespective of intervening syntactic boundaries within the simple (i.e. non-compound) sentence. Thorsen (1980c) corroborated this definition: word boundaries (which were simultaneously noun phrase-verb phrase boundaries) do indeed seem to be immaterial for the Fo patterns of stress groups in ASC Danish. (JR was also a subject for that investigation and a reservation had to be made for more conservative variants of Danish since word boundaries left a clear trace in his utterances.) However, the sentences for the 1978 and 1980c materials were comparatively short, containing three and four SGs, and they exhibited no intonation contour discontinuities. It is conceivable that syntactic boundaries, when they co-occur with prosodic phrase group boundaries, as in sentences 4 through 8 in the present investigation, will break up the regular Fo pattern.

The syntactic boundaries always occur after the first post-tonic syllable in the utterances; accordingly, we might expect the relationship between the first and second post-tonic in SGs before and after phrase boundaries to be different. For instance, the fall from first to second post-tonic could be smaller before the boundary, and thus the second post-tonic could be comparatively higher, in anticipation of the "rise" performed by the succeeding stressed syllable, to which it is affiliated syntactically. Or the fall from first to second post-tonic could be (substantially) larger, bringing the second post-tonic below the level of the succeeding stressed syllable, in imitation of the way sentence initial unstressed syllables behave (cf. fig. 1). - No such differences appear: the fall from first to second post-tonic seems completely unaffected by the discontinuities in the intonation contour and the syntactic boundary per se cannot be made responsible for any changes in stress group patterns. This is not to say that the phrase group boundary does not affect the stress group pattern: it does - only the variation is not located at the syntactic boundary. Inspection of figs. 2-6 reveal a consistent trend for the rise from stressed to post-tonic (which belong to the same word in all instances) to be greater in pre- than post boundary position. In sentences 5 and 6, and at the first boundary in 7 and 8, the rises can be compared directly, since the post-tonic syllable is carried by a syllabic [n̩] in all cases ('... bűssen til féssten ...' - [ˈb̥usn̩ d̥ˢe ˈfɛsd̥n̩]). The average rise is 2.8 semitones before the boundary, as compared to a rise of 1.9 semitones after the boundary, i.e. a difference of nearly 1 semitone. - Now, greater pre- than post boundary rises do not in themselves prove a boundary effect, because progressively decreasing magnitude of the rise to the post-tonic is a feature also of statements without any phrase group boundaries, cf. fig. 1. - Due to differences in the segmental composition of the post-tonic syllable in the stress groups (which entails possible differences in intrinsic Fo levels) part of the following argumentation is qualitative only. First of all, it does seem that a difference in the magnitude of the pre- and post boundary rises of one semitone is rather more than one would expect between neighbouring stress groups in a long statement without any prosodic phrase group boundaries, cf. fig. 1. Secondly, the pre-boundary rises seem to be rather high also in comparison to the <u>preceding</u> rise: In sentence 8, the first prosodic phrase group's second and third stress groups both have syllabic nasals for post-tonics, and the pre-boundary rise is larger than the preceding rise with all subjects (the difference is 0.8 semitones in the grand mean). In the second prosodic phrase group in sentences 7 and 8, the two stress groups under scrutiny both have vowels in their post-tonics, [i] in the pre-boundary stress group, [ʌ] in the preceding one, and although an unstressed [i] may have an intrinsically higher Fo than an unstressed [ʌ], this alone does not seem sufficient to explain the increase we get in the pre-boundary rise (in sentence 8 the pre-boundary rise is 0.5 semitones, in sentence 7 0.8 semitones higher than the preceding one - grand mean). In the first prosodic phrase group in sentences 6 and 7 the two rises are of very nearly the same magnitude. The first post-tonic is carried by a vowel, [a] and [ʌ], respectively, the second one by a syllabic nasal, and if we assume that an unstressed syllabic nasal has a lower intrinsic Fo than unstressed vowels, then the pre-boundary rise is even more "excessive".

On the whole, it seems safe to conclude that the prosodic phrase group boundary results in a comparatively higher rise to the post-tonic in the preceding stress group. The cause of the greater pre-boundary rise may be sought in either of two processes (or in a combination of them): it is a signal for the prosodic phrase group boundary, and thus controlled by the speaker, or it is an automatic consequence of the higher position of the succeeding stressed syllable, i.e. the higher the following stressed vowel, the less of a fall must be executed by the preceding Fo pattern and, consequently, the higher the rise may be from the preceding stressed vowel. The latter explanation would be identical to, and could be taken as further support for, the explanation offered previously (Thorsen 1979b, 1980a) for the fact that the magnitude of the rise from stressed to post-tonic vowel varies with intonation contour, i.e. we get higher rises on less steep contours, cf. fig. 1.

Note that in this investigation JR does not deviate from the three ASC speakers as far as syntactic boundaries and fundamental frequency are concerned. His stress group pattern deviates from those of the other subjects by having more steeply falling unstressed syllables, but there is no trace of a syntactic boundary signalling in his traces. - I am inclined to think, now, that the word boundary signalling he performed in the previous (1979c) experiments does not constitute an example of a difference between ASC and more conservative variants of Danish. Rather, it demonstrates that it is possible for a speaker to signal word boundaries, also with fundamental frequency, if he so desires, a possibility which presumably is also open to ASC speakers.

In summary, if the results of the present investigation can be extended to cover simple sentences in general, they present an argument in favour of a theory expounded in Selkirk (forthcoming) that prosodic categories (in casu: stress groups and prosodic phrase groups) are distinct entities in the phonology that do not have an isomorphous relation to syntactic structure. Rischel (1972) argues in a similar fashion: Danish stress is best represented in a hierarchy (a tree structure) which is not necessarily congruent with the syntactic structure. - The autonomy of prosodic structure does not, of course, deprive it of a relation to syntax (cf. the questions posed below), on the contrary, prosodic categories can be seen as reconciling the syntactic structure to the phonetic output (in casu: the course of fundamental frequency).

6. Discussion

The purpose of the investigation was not to investigate the interplay between syntax and prosodic structure 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 syntactic boundary before the (compound) complement seems to be more manifest than the noun phrase/verb phrase boundary. The constituent which varied most in number of stress groups was the complement: What would the contours have looked like if instead the noun phrase and/or the verb phrase had varied? E.g. is the tendency towards a faster declination early in the utterance an inherent feature of declarative intonation or is it an artefact of the material that would disappear if the noun phrase or verb phrase were longer? With a short complement but a long noun phrase, would a prosodic phrase group boundary be introduced after the noun phrase, and would the verb phrase and complement merge into one prosodic phrase group? If the second of the two complements had consisted of only one stress group, would it have had to merge prosodically with the preceding complement in order to preserve a final declination? And where would the first complement be cut up then, if four stress groups are the maximum in a prosodic phrase group (compare sentence 6 and 7)? Or is the final declination dispensable as long as there is an overall downdrift in the utterance? How do clause boundaries manifest themselves? When unaccompanied by pauses, will they exhibit greater amounts of resetting than do clause internal boundaries? And will they affect stress group patterns?

Acknowledgement

Sincere thanks are due to Jeanette Holtse for her patient and competent graphics work.

References

Breckenridge, J. and M.Y. Liberman 1977: "The declination effect in perception" (unpublished manuscript)

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

Cooper, W.E. and J.M. Sorensen 1977: "Fundamental frequency contours at syntactic boundaries", JASA 62, p. 683-692

Fujisaki, H., K. Hirose, and K. Ohta 1979: "Acoustic features of the fundamental frequency contours of declarative sentences in Japanese", Annual Bulletin, Research Institute of Logopedics and Phoniatrics 13, p. 163-173

Gårding, E. 1979: "Sentence intonation in Swedish", Phonetica 36, p. 207-215

't Hart, J. 1979: "Exploration in automatic stylization of Fo curves", IPO APR 14, p. 61-65

't Hart, J. and R. Collier 1975: "Integrating different levels of intonation analysis", JPh 3, p. 235-255

Hirose, H. 1971 "On Japanese accent: An acoustic analysis", Monthly Internal Memorandum, Phonology Laboratory, Berkeley, May, p. 16-40 (cited from Ohala 1978)

Maeda, S. 1977: "Sur les corrélatifs physiologiques de la fréquence du fondamental", Actes des 8emes Journées du Groupe de la 'Communication Parlée', Aix-en-Provence, 25-27 Mai, p. 43-50

McAllister, R. 1971: "Predicting physical aspects of English stress", STL-QPSR 1, p. 20-29

Ohala, J. 1978: "Production of tone", Tone: A Linguistic Survey (ed.: V. Fromkin), Academic Press, New York, p. 5-39

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

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

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

Rischel, J. 1972: "Compound stress in Danish without a cycle", ARIPUC 6, p. 211-229

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

Selkirk, E.O., (forthcoming): "On the nature of phonological representation", The Cognitive Representation of Speech (eds.: J. Anderson, J. Laver, and T. Myers) North Holland Publishing Company

Silverstein, R.O. 1976: "A strategy for utterance production in Hausa", Studies in African Linguistics, Suppl. 6, p. 233-241 (cited from Ohala 1978)

Sorensen, J.M. and W.E. Cooper 1980: "Syntactic coding of fundamental frequency in speech production", Perception and Production of Fluent Speech (ed.: R.A. Cole), Lawrence Erlbaum Associates, Hillsdale, New Jersey, p. 399-440

Sternberg, S., C.E. Wright, R.L. Knoll, and S. Monsell 1980: "Motor programs in rapid speech: Additional evidence", Perception and Production of Fluent Speech (ed.: R.A. Cole), Lawrence Erlbaum Associates, Hillsdale, New Jersey, p. 507-534

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

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

Thorsen, N. 1980a: "A study of the perception of intonation contours - Evidence from Danish", JASA 67, p. 1014-1030

Thorsen, N. 1980b: "Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish", ARIPUC 14 (this volume)

Thorsen, N. 1980c: "Word boundaries and Fo patterns in Advanced Standard Copenhagen Danish", Phonetica 37, p. 121-130 (also in ARIPUC 13, 1979, p. 121-134)

Weitzman, R.S. 1970: "Word accent in Japanese", Studies in the Phonology of Asian Languages, University of Southern California, 11 (cited from Ohala 1978).

(ARIPUC 14, 1980)

CONTRIBUTIONS TO DOGRI PHONETICS AND PHONOLOGY[1]

Ved Kumari Ghai[2]

Abstract: The present paper deals with some selected problems of Dogri phonetics and phonology. Section 1 contains a general introduction. Section 2 contains a formant analysis of oral vowels (2.2), a discussion of the distinctive features of vowels, in particular the relative constancy of the concomitant features length and centralization (2.3), a statement of the distribution of vowel phonemes with a discussion of the occurrence of peripheral (long) vowels before geminated consonants (2.4), and a discussion of the phonological interpretation and the phonetic manifestation of pairs of the type [ˈgəlɑː] [gəˈlɑː] as a difference in length or stress. Section 3 treats the problems of consonant gemination and consonant weakening with an account of the historical background (3.2), and a discussion of the relation to syllable division (3.3) and of the interpretation of one of the grades as basic (3.4). 3.5 and 3.6 deal with the phonetic manifestation of consonant gemination and weakening, respectively.

1. Introduction

1.1 Previous treatments

Dogri belongs to the Indo Aryan language family and is spoken in the Jammu province of Jammu and Kashmir State in India. According to the Census of India 1971, it has nearly 1.200.000 speakers. Grierson (in "Linguistic Survey of India" vol. IX (1916) has

---

1) This study was undertaken at the Institute of Phonetics, University of Copenhagen, under the supervision of professor Eli Fischer-Jørgensen and was supported by a grant from the Danish Ministry of Education. I am very grateful to professor Fischer-Jørgensen for her able guidance and for her keen interest in the completion of the project. Thanks are also due to professor Jørgen Rischel (Institute of Linguistics), and to staff members of the Institute of Phonetics: to Birgit Hutters and Niels Reinholt Petersen for various helpful suggestions, to Svend-Erik Lystlund for technical help, to Nina Thorsen for improving the style of the paper, and to Else Parkmann for typing the manuscript. Michael Bundgaard, Niels-Jørn Dyhr, Kirsten Morsing, and Eva Rosenvold have helped making and measuring spectrograms, calculating averages, and drawing figures.

2) Professor of Sanskrit, on leave from Jammu University.

given a brief description of Dogri and has grouped it with Panjabi. Gauri Shankar (1931) has given a description of Dogri with a small vocabulary and four pages of connected text in phonetic transcription. The monograph points out some linguistic phenomena in which Dogri differs from Panjabi and lists a number of words which have different phonetic realizations in Dogri and Panjabi, but does not specify which sounds are phonemically distinct. In 1965 appeared a Dogri grammar in Hindi, written by Bansilal Gupta (1965), which gives a brief account of Dogri sounds. A more detailed description of Dogri has been given by Ujjal Singh Bahri in his paper "Phonology of Dogri" (1969). The paper is based on an auditory study of tape recordings made in various parts of Jammu Province. He gives a detailed articulatory, auditory, and phonological description of Dogri sounds. His description, however, differs from the present description on some points, which will be discussed in later paragraphs.

I have chosen to treat some selected problems of Dogri phonology and phonetics more thoroughly, and mention others very briefly. The subjects treated in more detail are the system of oral vowels, the problems of quantity and stress, consonant gemination and consonant weakening. The tonal system and its phonetic manifestation has been treated in a previous paper (Ghai 1968).

1.2 Informants and material for the present investigation

The phonetic part of the present work is based on tape recordings by six informants, three male (RN, SL, and SD, aged 62, 30, and 32, respectively) and three female (CS, AS, and VG, aged 40, 23, and 48, respectively). VG is the author. They all speak Dogri as their first language.

The recordings of the first five informants were made in the studio of Radio Kashmir of Jammu, on a professional tape recorder at a speed of 7½ i.p.s., whereas the recordings of VG were made in a sound-treated room at the Institute of Phonetics in Copenhagen.

The recordings consisted of lists of words which were all spoken in a carrier sentence, generally: ɷs ... sɷṇe:ɑ: 'he ... heard', but in some cases: jəs... kərədi:hi: 'Jes was ... doing' (used for verbs). Moreover, the subjects read a short connected text composed by the author, and a story. Each of the word lists was read five times by each informant. AS, however, only read the lists three times and she did not speak all the words, so she has been left out in various averages.

List I consisted of the following words, quasi-randomized, as shown:[1]

təˈlɑ:	ləˈdɑ:	ˈgəl(ə)	ˈdɑ:bbɑ:	ḍəpph(ə)	gəɟ(ə)
gəˈlɑ:	kəˈrɑ:	ləˈgɑ:	ɟəd(ə)	tu:ph	ˈcəppəl(ə)
pəˈlɑ:	ˈgəlɑ:	lɑ:g(ə)	ˈɟɑ:ddi:	ˈtə̀ppha:	ˈgəɟɟ(ə)
ləˈkɑ:	ˈpəlɑ:	gəll(ə)	ɟədd(ə)	ˈməṭṭhɑ:	ˈkəṭənɑ
bəˈsɑ:	ˈləkkɑ:	ˈlɑ:ggɛ	ˈɟɑ:dɑ:	ˈtò:phɑ:	ɟed(ə)
səˈdɑ:	ˈbəssɑ:	ˈlɑ:gɑ:	kɑ́:d(ə)	ˈtəlɑ:	kəˈtənnɑ
səˈtɑ:	ˈsəddɑ:	gɑ:l(ə)	ˈgo:ḍḍɑ:	ˈpətɑ:	ɟədd(ə)
kəˈsɑ:	ˈləggɑ:	ˈgo:lli:	ɟɑ:t(ə)	ˈkəccɑ:	gətt(ə)
bəˈgɑ:	ˈkəssɑ:	ˈləgg(ə)	ˈɟɑ:ttɑ:	ˈkəcɑ:	
dəˈbɑ:	ˈbəgɑ:	ˈgo:li:	dɑ:kh(ə)	ˈbo:li	
kəˈḍɑ̀:	ˈdəbbɑ:	ləg/ɟəg	dəkh(ə)	pɷll(ə)	
kəˈtɑ:	ˈkə́ḍḍɑ:	bɑ:g(ə)	tɷ̀kh(ə)	sət(ə)	
ɟəˈtɑ:	ˈkəttɑ:	ˈsəti:	ˈtò:khɑ:	ɟəgg(ə)	
dəˈkhɑ:	ˈɟittɑ:	ˈbəggɑ:	khəˈbɑ:r	ˈpəttɑ:	
təˈkhɑ̀:	ˈpə̀khɑ:	ˈbɑ:ggɛ:	təkkh(ə)	ˈbo:lli:	
cəˈthɑ:	ˈtɷkhɑ:	ˈbəg(ə)	ˈkəthɑ:	ɟəg(ə)	
ḍəˈphɑ:	ˈḍəpphɑ:	ɟəˈdɑ:i	ˈcɔ:thɑ:	ˈsətt(ə)	
bəˈthɑ:	ˈpə̀kkhɑ:	ˈɟɷdɑ:	cɔ:th(ə)	pɷl(ə)	
səˈkhɑ:		dəbb(ə)	ˈkətthɑ:	ˈcəpəl(ə)	

List I was mainly intended for the analysis of consonants, vowel length, and stress.

List II consisted of the following words:

ˈkɷli:	ˈko:lli:	kɷl(ə)	tɑ:l(ə)
ˈgo:ḍḍɑ	gɷˈḍɑ:	kɔ:l(ə)	ti:l(ə)
ˈkɷlli:	ˈte:llɑ:	tɛ:ʃ(ə)	ku:l(ə)
ˈtɩllɑ:	ˈbo:bbo:	təl(ə)	tɩl(ə)
ˈkɩlɑ:	ˈke:llɑ:	kɷl(ə)	ˈkɩllɑ:
ˈgɷḍḍɑ:	ˈkɷbbi:	te:l(ə)	

This was a supplementary list intended for a spectrographic analysis of vowel quality and of vowel length. A few words were taken from a third list, set up for a different purpose.

---

1) The transcription used in this list and in the rest of the paper is a phonemic notation with some redundancy, e.g. both vowel quality (centralization) and vowel length are indicated (see the discussion in section 2.) (ə) in parentheses indicates an optional sound. It is always possible to pronounce an [ə] after a final consonant.

2. Vowel system, quantity and stress

2.1 The vowel phonemes of Dogri

Dogri has 10 oral and 5 nasal vowel phonemes, which can be set up in the following preliminary system:

i		u		ĩ		ũ
	ɩ	ɷ		ɛ̃		ɔ̃
e	ə	o			ɑ̃	
ɛ		ɔ				
	ɑ					

For the oral vowels the following commutation series can be given: tɩ:l(ə) 'match stick', te:l(ə) 'oil', tɩl(ə) 'sesam', tɛ:r(ə) 'to swim', təl(ə) 'to fry', tɑ:l(ə) 'to cleanse', tɔ:l(ə) 'to hurry', to:l(ə) 'to weigh', tɷl(ə) 'equal', tu:ṛ(ə) 'to fill'. The distinction between the nasal vowels can be shown by the words sərɑ̃: 'inn', sərĩ: 'name of tree', pərũ: 'last year', [pərẽ: ∿ pərɛ̃:] 'away', [rõ: ∿ rɔ̃:] 'wood', and the distinction between oral and nasal vowels can be shown by the pairs əkkhi: 'eye' vs. əkkhĩ: 'eyes', ku: 'say' vs. kũ: 'warping of the cot', ɟɑ:gətɛ: 'the boy' (oblique form) vs. ɟɑ:gətɛ̃: 'the boys', sɔ: 'hundred' vs. sɔ̃: 'oath', tɑ: 'heat' vs. tɑ̃: 'them'. Bahri states that all ten oral vowels have nasal counterparts, but according to my own observations, there is free variation between [ẽ:] and [ɛ̃:] and between [õ:] and [ɔ̃:], and ɩ ɷ ə do not show phonemic nasalization; when nasalized they are always followed by a nasal consonant, at least in the underlying form. In the case of the other vowels there is commutation between nasalized vowel and vowel plus nasal consonant (see later). The nasal vowels will not be treated in any detail in this paper.

2.2 Acoustic analysis of vowel quality

A spectrographic analysis by means of the Kay Elemetric Sonagraph has been made of the ten oral vowels in the words ti:l(ə), tɩl(ə), te:l(ə), tɛ:ʃ(ə), təl(ə), tɑ:l(ə), kɔ:l(ə), ko:l(ə), kɷl(ə), ku:l(ə) spoken five times in the frame ɷs ... sɷṇe:ɑ: by the four informants RN, SD, SL, and VG.

The reason for choosing k before rounded back vowels and t before front and unrounded open vowels was that these combinations give shorter formant transitions and thus facilitate the measurement of the vowel targets (but l may have raised F_2 of the short vowel ɷ somewhat). Both wide band and narrow band sonagrams were made of all vowels, and in some dubious cases also sections.

For the unrounded vowels the frequency of formants 1, 2 and 3 were measured, for the rounded back vowels only F_1 and F_2 could be measured since F_3 was (as it is normal for these vowels) too weak to be identified. Formants 1 and 2 of ɑ: were sufficiently wide apart to be measured accurately for RN, SD and SL; but VG's vowels presented some difficulties in this respect. The exact location of F_1 in high vowels and the separation of F_1 and F_2 in u:, ɷ and o: presented the expected difficulties, and particularly F_2 in u: ɷ o: was dubious in many cases, so that the measurement of these formants must be taken with some reservation. SL's F_2 was very weak in these vowels. The high F_o of VG's voice also presented difficulties for accurate measurements. The wide band sonagrams in most cases resulted in higher values than the narrow band spectrograms. The frequencies given are averages of the two measurements, except for the higher formants of front vowels which could be measured only on the wide band spectrograms.

The average formants of the ten vowels are given in table 1, and the frequencies of F_1 and F_2 of individual tokens have been plotted in figs. 1-4. A logarithmic frequency scale has been used.

Table 1

Averages of vowel formant frequencies in monosyllables for four subjects (N = 5).

		RN	SL	SD	VG
ti:l(ə)	F_1	285	291	286	266
	F_2	2255	2070	2285	2865
	F_3	3150	2538	3025	3287
tɩl(ə)	F_1	358	423	370	361
	F_2	2095	1785	2140	2520
	F_3	2580	2470	2590	2990

Table 1 (continued)

		RN	SL	SD	VG
te:l(ə)	F_1	343	372	351	363
	F_2	2225	1975	2175	2635
	F_3	2920	2555	2720	3095
tɛ:ʃ(ə)	F_1	673	576	639	804
	F_2	1781	1454	1586	1851
	F_3	2450	2655	2705	2755
tɑ:l(ə)	F_1	684	597	780	860
	F_2	1096	1161	1135	1220
	F_3	2388	2625	2785	2666
təl(ə)	F_1	645	556	720	874
	F_2	1141	1188	1162	1544
	F_3	2562	2688	2725	2688
ku:l(ə)	F_1	279	346	314	321
	F_2	613	663	704	583
kɷl(ə)	F_1	364	392	336	350
	F_2	706	883	705	720
ko:l(ə)	F_1	348	395	352	377
	F_2	649	779	648	715
kɔ:l(ə)	F_1	579	499	539	588
	F_2	956	953	945	991

2.3 Distinctive features of the vowels

On the basis of the auditory impression and the acoustic analysis the following features were set up to distinguish the vowel phonemes of Dogri: ± high, ± back, ± round, ± long (or ± peripheral), and ± nasal. The distribution of the features is given in Table 2.

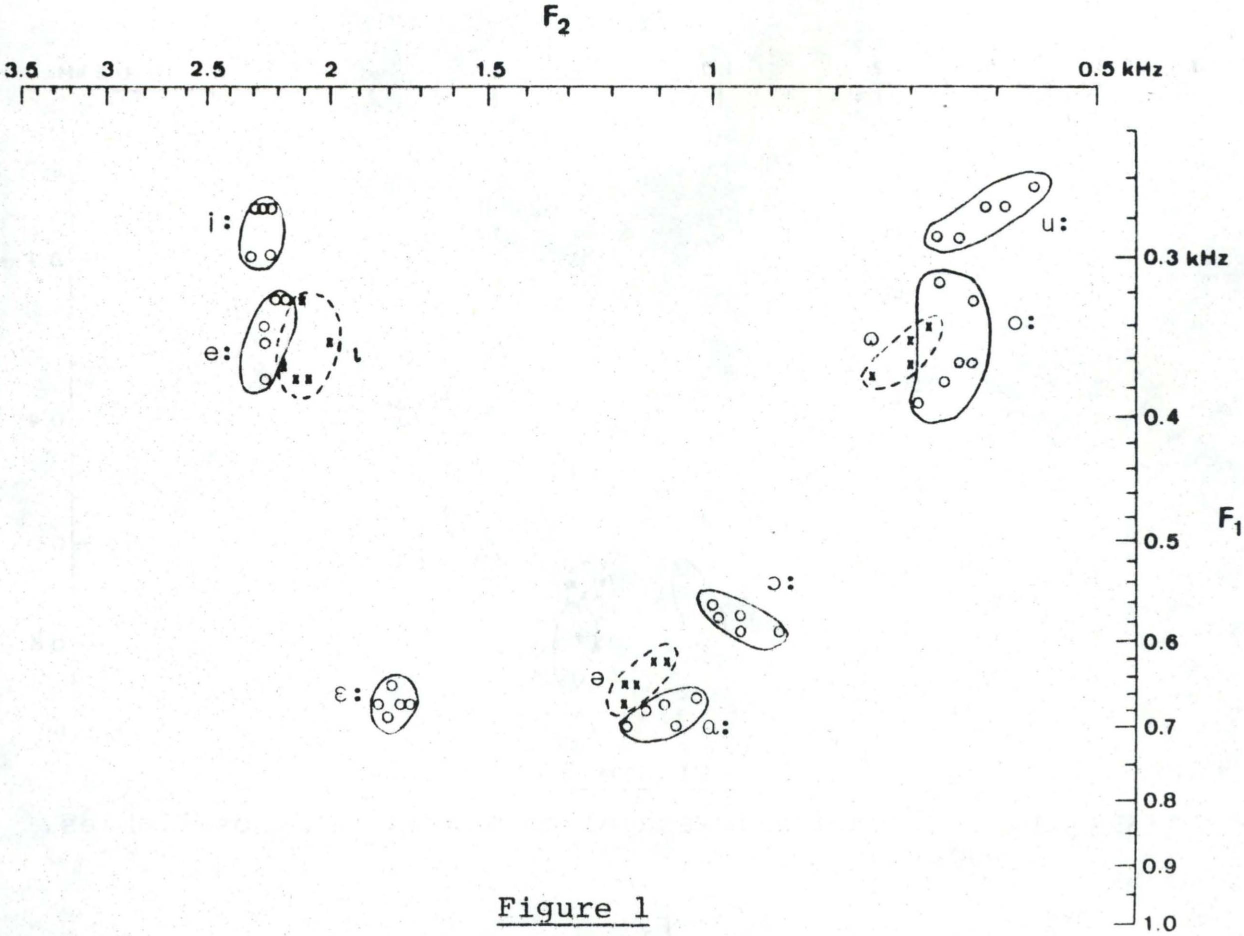


Figure 1

F_1/F_2 plot (logarithmic scale) of vowels in monosyllables. Subject RN.

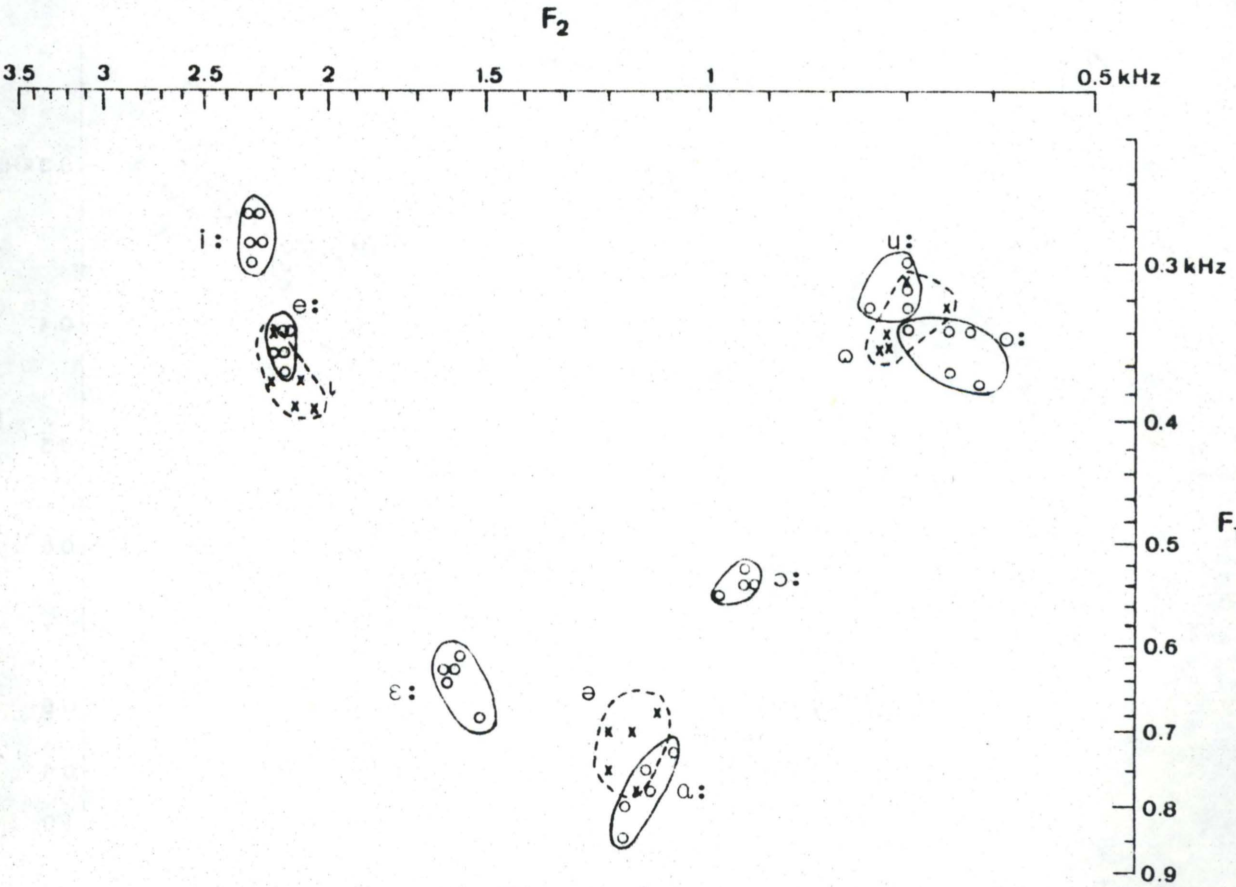


Figure 2

F_1/F_2 plot (logarithmic scale) of vowels in monosyllables. Subject SD.

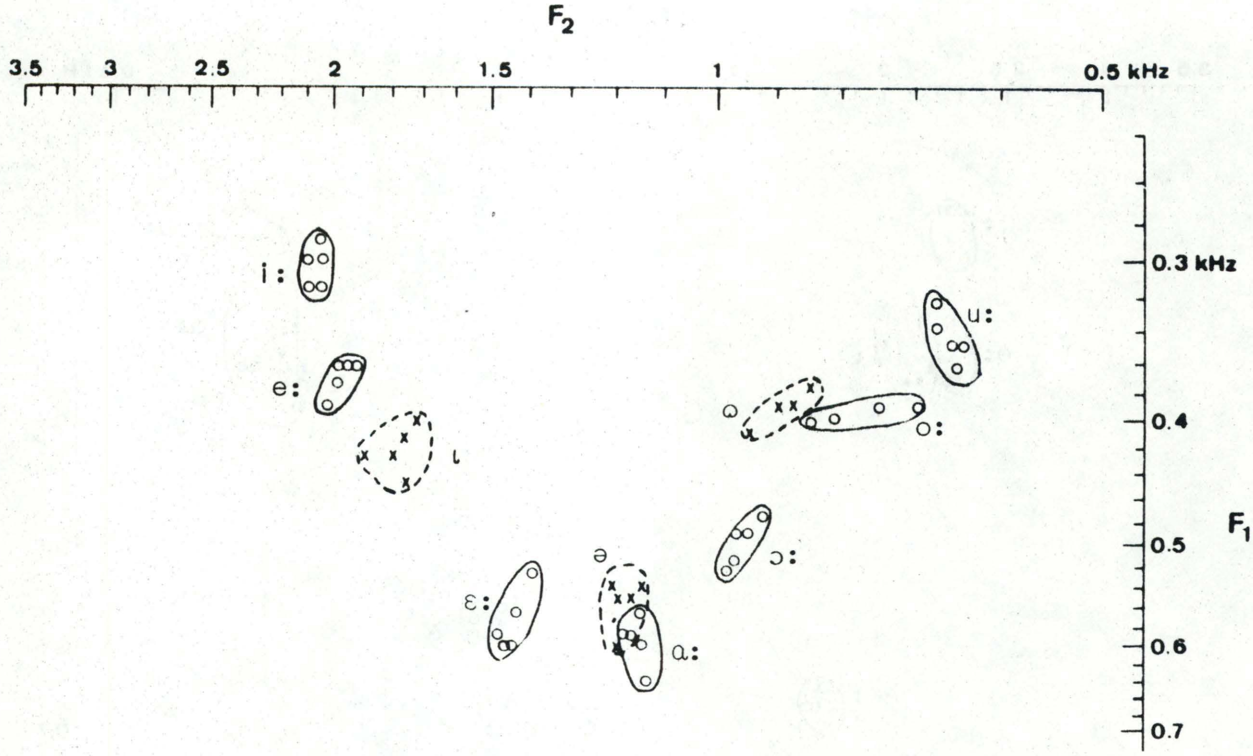


Figure 3

F_1/F_2 plot (logarithmic scale) of vowels in monosyllables. Subject SL.

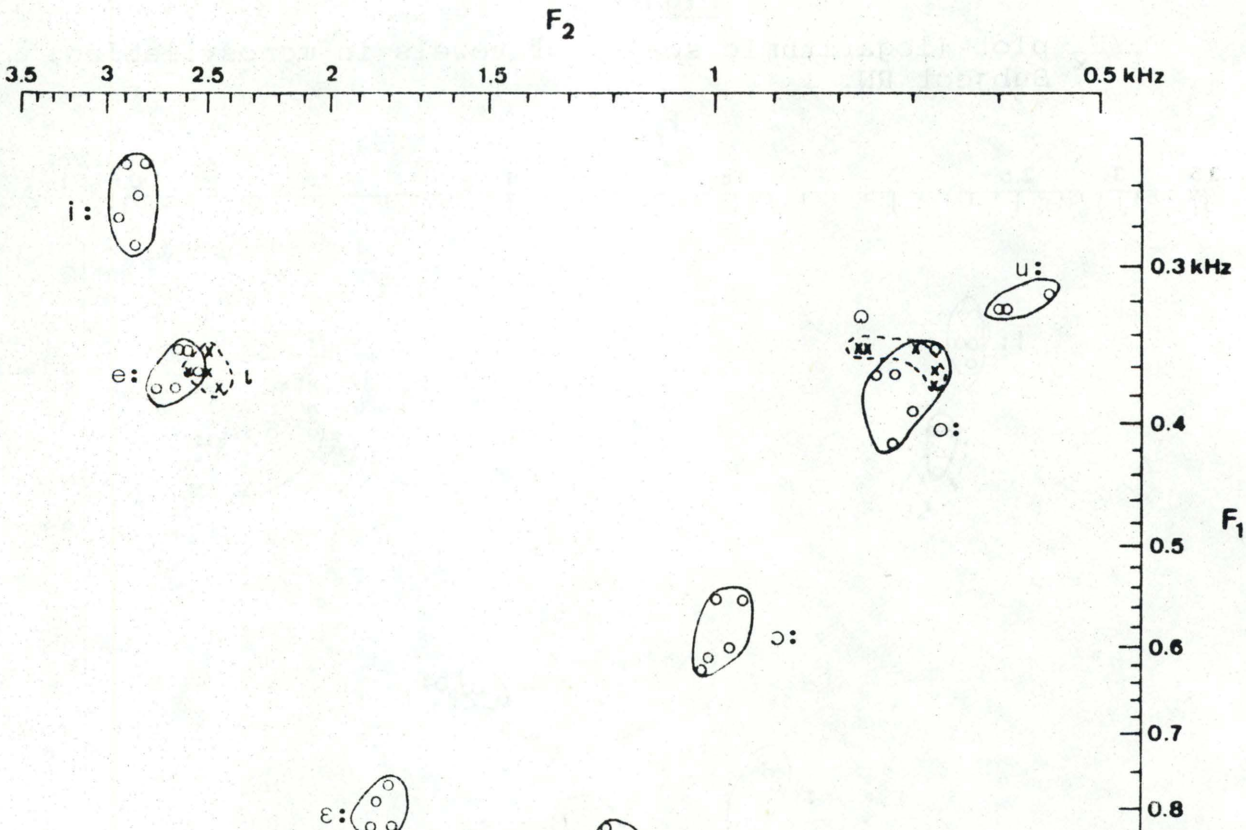


Figure 4

F_1/F_2 plot (logarithmic scale) of vowels in monosyllables. Subject VG.

Table 2

Distinctive features of Dogri vowels

	i	e	ɛ	ɑ	ɔ	o	u	ɩ	ə	ɷ	ĩ	ɛ̃	ɑ̃	ɔ̃	ũ
high	+	-	-	-	-	-	+	-	-	-	+	-	-	-	+
low	-	-	+	+	+	-	-	-	+	-	-	+	+	+	-
back	-	-	-	+	+	+	+	-	+	+	-	-	+	+	+
round	-	-	-	-	+	+	+	-	-	+	-	-	-	+	+
long	+	+	+	+	+	+	+	-	-	-	+	+	+	+	+
nasal	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+

As for vowel height, I follow Chomsky and Halle (1968) in setting up two binary features, +high/-high and +low/-low. Ladefoged (1975) finds that the auditory property vowel height, as used in D. Jones' cardinal vowel quadrangle, is better correlated with the acoustic difference lower vs. higher first formant than with the articulatory feature "highest point of the tongue", whereas Mona Lindau (1977) has shown that the auditory impression correlates well with both the acoustic and the articulatory facts.

I have regarded i: and u: as +high, although u: comes rather close to ɷ and o: on the vowel charts. Auditorily it sounds higher, and for RN and SD it has a clearly weaker F2 than o:, which may make the perceptual difference greater than shown by the frequencies of the formants on the chart. ɛ:, ə, ɑ:, ɔ: are considered to be +low. The back low vowel ɔ: is not as low as the front low vowel ɛ: (which might also be transcribed [æ:]), but it is lower than u:, o:, and ɷ . It is separated from ɑ: by the feature +round. ə is evidently a low vowel in Dogri. It is heard by trained phoneticians as an [ɑ]-sound.[1]

As for ±back, I consider it also as binary. Here there is a clear correlation between the auditory impression and the front/back position of the tongue, and also with the acoustic differences

---

1) Bahri uses the terms high-low. But he also talks of close-open (this seems to be a subdivision) and in some cases of compact-diffuse (the meaning of these latter terms is not clear. They are at any rate not used in the sense coined by Roman Jakobson (1952).

in F2. The more fronted the vowel, the higher F2. Of Dogri vowels, I have regarded <u>i:</u> <u>e:</u> <u>ɩ</u> <u>ɛ:</u> as -back, and <u>ə</u> <u>ɑ:</u> <u>ɔ:</u> <u>ɷ</u> <u>o:</u> <u>u:</u> as +back.

As for rounding, <u>ɷ</u> <u>u:</u> <u>o:</u> <u>ɔ:</u> are rounded. The others are unrounded. The only unrounded and rounded vowels of the same height and backness are <u>ɑ:</u> and <u>ɔ:</u> .

The feature ± long is concomitant with the feature ± peripheral which is normally called tense/lax. I have regarded ± long as the distinctive feature and ± peripheral as a redundant feature which, however, helps a lot in the perception of the difference. On the whole, the short vowels <u>ɩ</u> <u>ɷ</u> <u>ə</u> are generally centralized to some extent in the F1-F2 vowel diagram, but the centralization is not as obvious as in the case of e.g. German lax vowels (Hans-Peter Jørgensen 1969). One informant only (SL) has clearly centralized <u>ɩ</u> and <u>ɷ</u>.

As regards <u>ɩ</u> and <u>ɷ</u> , they obviously have a higher F1 (i.e. they are lower) than <u>i:</u> and <u>u:</u> , respectively: <u>ɩ</u> is much closer to <u>e:</u> for all four informants, and in the case of SL even lower, and <u>ɷ</u> is closer to <u>o:</u> than to <u>u:</u> (perhaps with the exception of SD). As for the second formant, <u>ɩ</u> and <u>ɷ</u> are also more centralized than <u>i:</u> and <u>u:</u> . But they are not so clearly distinguished from <u>e:</u> and <u>o:</u> . Only SL, who has evidently centralized <u>ɩ</u> and <u>ɷ</u> , has no overlapping between <u>ɩ</u> and <u>e:</u> , <u>ɷ</u> and <u>o</u>. RN has a slight overlapping between <u>ɷ</u> and <u>o:</u> , and VG and SL have extensive overlapping both for <u>ɩ</u> and <u>e:</u> and <u>ɷ</u> and <u>o:</u> (SD also <u>u:</u> and <u>ɷ</u>), although the sounds were spoken in test words in the same surroundings (but VG has a stronger F2 in <u>o:</u> than in <u>ɷ</u> . For one informant (SD), who has a particularly extensive overlapping, we tried to replace F2 by F2-F1 (following Ladefoged 1975 and Mona Lindau 1977), but this did not improve matters. As <u>e:</u> has a higher F3 than <u>ɩ</u>, we also tried for one informant to replace F2 by a weighted average of F3 and F2 according to Gunnar Fant's older formula (Fant 1959), but this gave only a very slight improvement. It must therefore be stated that, although there is a clear tendency to centralization for <u>ɩ</u> and <u>ɷ</u>, there is some qualitative overlapping with <u>e:</u> and <u>o:</u> .

As for <u>ə</u> , it is more centralized than <u>ɑ:</u> , vertically for SD and RN, horizontally for VG, and in both dimensions for SL, but both SL and SD show some overlapping.

Moreover, a test carried out with four trained Danish phoneticians listening to the first recording of list II, spoken by RN, SL, SD, and CS, showed that they had obvious difficulties in identifying the centralized vowels, although all are familiar with German. kɷl(ə) was heard as ko:l(ə) in all (3 x 4) cases (with one exception), təl(ə) as tɑ:l(ə) (one listener heard a mid [a-]), and tɩl(ə) as te:l(ə) in 8 out of 12 cases.

Thus the difference centralized-peripheral is not consistent and cannot be regarded as the only determining factor.

The difference of duration is, however, consistent. The so-called non-peripheral or lax vowels are always shorter than the corresponding peripheral or tense vowels in similar environments. In monosyllabic words, the duration of the long vowel is generally more than twice the duration of the short vowel, for SL slightly less than that (see table 3[1]). The relative duration of the vowels <u>ɩ ɷ ə</u> compared to <u>e: o: ɑ:</u> lies between 35% and 42% for RN, SD and VG and at 54%-55% for SL. In disyllabic words, the vowels are shortened (table 4 shows this for some vowels) and still more so before geminated consonants, but the relations are maintained (see table 5), and SL has in this case the same relative difference as the other informants.

Table 3

Average vowel duration in monosyllabic words in cs. (N = 5)

	RN	SL	SD	VG
ti:l(ə)	24.5	16.6	34.5	26.4
te:l(ə)	25.5	15.6	35.6	27.1
tɛ:ʃ(ə)	25.4	18.8	35.8	27.8
tɑ:l(ə)	28.1	19.0	38.3	29.8
ku:l(ə)	23.9	17.6	35.0	(42.3)
ko:l(ə)	25.8	17.5	36.1	25.4
kɔ:l(ə)	25.4	17.1	36.8	25.8
tɩl(ə)	9.8	8.7	14.6	10.0
kɷl(ə)	10.6	9.5	14.2	8.9
təl(ə)	10.8	10.4	16.1	12.5
% tɩl(ə)/ti:l(ə)	39	52	42	38
kɷl(ə)/ku:l(ə)	44	54	41	(21)
təl(ə)/tɑ:l(ə)	38	55	42	42

---

1) VG's examples of [ku:l(ə)] were taken from a different recording where she spoke more slowly.

Table 4

Average vowel duration in cs in some disyllabic words (N = 5)

	RN	SL	SD	VG
ɟa:da:	14.0	12.8	18.2	18.8
bo:ḷi:	11.6	12.5	20.0	16.6
kɩla:	8.0	6.2	11.0	6.5
kɷli:	7.2	5.4	10.2	7.5

Table 5

Average vowel duration in disyllabic words before geminate consonants in cs (N = 5)

	RN	SL	SD	VG
tɩlla:- te:lla:	4.5 - 11.3	4.5 - 9.1	6.0 - 14.0	4.0 - 10.0
kɩlla:- ke:lla:	4.3 - 10.6	4.3 - 10.0	5.4 - 14.5	3.5 - 11.0
kɷbbi:- bo:bbo:	5.6 - 11.4	4.3 - 9.8	6.3 - 14.0	4.0 - 11.0
kɷlli:- ko:ḷḷi:	4.2 - 10.4	3.9 - 10.4	6.3 - 14.2	4.0 - 11.0
gɷḍḍa:- go:ḍḍa:	4.8 - 9.8	4.7 - 10.2	6.2 - 13.5	4.0 - 11.0
grand mean	4.7 - 10.7	4.3 - 9.9	6.0 - 14.0	3.9 - 10.8
% short/long	44%	43%	43%	36%

The feature ± long used here corresponds to Bahri's central/peripheral[1]. He describes ɩ ɛ: u: as tense, ɩ o: as lax, but does not state anything about a: ɔ: ə ɷ regarding tenseness/laxness, so it is not clear whether the term tense used by him corresponds to peripheral or not. According to my description, o: is +long and may be regarded as redundantly peripheral and tense, but not as lax.

---

1) He considers this difference as more stable than the difference of duration. In my 1970 paper I described the difference i:-ɩ and u:-ɷ as a length difference but ə-a: as a difference in vowel quality. But Bahri is certainly right in grouping ɩ ɷ ə together. As mentioned above, in this paper both quality and length are indicated (redundantly) in the transcription.

Besides acoustic and auditory factors, there is another reason why I prefer the feature ± long. It appears that vowel quantity plays a role in the formation of rules for stress in Dogri. Stress in Dogri is predictable in some cases and unpredictable in others. All Dogri syllables can be placed in one of the following three grades: Light: a syllable ending in a short vowel; medium: a syllable ending in a long vowel, or ending in a short vowel checked by a single consonant; heavy: a syllable having a short vowel checked by two consonants, or a long vowel checked by one or two non-syllabic consonants.

Rule I: If there is one heavy syllable in the word, it is stressed. Two heavy syllables do not occur consecutively in Dogri in simple words. bəˈɟɑ:r, kəˈpətt, ˈɟɑ:ddi:, cəˈlɑ:kki: . Forms like * bəɟɟɑ:r , * kɑ:pətt , * cɑ:lɑ:kki: cannot occur in Dogri.

Rule II: When there is more than one medium syllable (and no heavy syllable) in a word, the last but one of the medium syllables is stressed. ˈɟɑ:dɑ:, mərəˈɟɑ:dɑ:, ˈbəggɑ:, ˈɟɑ:gət, bəˈɟɑ:rɛ.

Rule III: When there is one medium syllable preceded by one or two light syllables, the stress is unpredictable and phonemic. ˈbəgɑ:, bəˈgɑ:, ˈpəgətɑ:, pəgəˈtɑ:, ˈrəso:, rəˈso.

The fact that in these rules each peripheral vowel (i: e: ɛ: ɑ: ɔ: o: u:) behaves like one non-peripheral vowel (ɩ ɷ ə) + one non-syllabic, speaks in favour of the description of the former as long and the latter as short. Thus, both phonetic and phonological arguments speak for considering length and not centralization as the relevant feature.

As regards the relationship of short vowels to long vowels, they are grouped traditionally as ɩ i:, ɷ u:, ə ɑ: . They are, no doubt, related historically, but now, by being lowered and centralized, ɩ ɷ have come phonetically closer to e: o: than to i: u: .

Moreover, there are various related words in Dogri in which there is morphological alternation of long vowels with short vowels. Here the long, -low, -high vowels (e: o:) alternate with short -low, -high vowels (ɩ ɷ), while the +low long vowel (ɑ:) alternates with the +low short vowel (ə).

Long vowel in transitive verb		Short vowel in intransitive verb	
ḍo:b(ə)	'to dip'	ḍɷb(ə)	'to dive'
kho:l(ə)	'to open'	khɷl(ə)	'to be opened'
ke:r(ə)	'to fell'	kɩr(ə)	'to fall'
ré:ṛ(ə)	'to push'	rɩ́ṛ(ə)	'to crawl'
pɑ:l(ə)	'to bring up'	pəl(ə)	'to be brought up'
gɑ:l(ə)	'to cause to melt'	gəl(ə)	'to melt'

Long vowel in noun		Short vowel in verb	
ḍo:bbɑ:	a dip	ḍɷbbɑ	past of verb ḍɷb to dive
co:r(ə)	a thief	cɷr(ə)	to be stolen
dɑ:l(ə)	grinded beans	dəl(ə)	to grind
bɑ́:ddɑ:	an increase	bə́d(ə)	to increase
so:kkɑ:	drought	sɷk(ə)	to get dry

So it seems more reasonable to group <u>ɩ-e:</u> and <u>ɷ-o:</u> together by regarding them as -low and -high, and to group <u>ə</u> with <u>ɑ:</u> and regard both of them as +low.

<u>2.4 Distribution</u>

On the basis of their behaviour in the syllable structure of the language, long and short vowels also form two distinct groups. The short vowels, viz. <u>ɩ ɷ ə</u> occur initially and medially but do not occur finally, where they may be considered to be neutralized with i:(e:), u:(o:) and zero, respectively. There are a few functional words in which they appear finally, e.g. kɩ 'that', jɩ 'that', kɷ 'about', cə 'in', ʃə 'from', but these words are short forms of ke:, je:, ku:, bɩc, kəʃ and are pronounced frequently in full form. <u>ə</u> appears phonetically as a vocalic release in some cases, especially after geminated consonants or consonant clusters in final position, but it is not phonemic in this position and often alternates with zero. The other short vowels <u>ɩ ɷ</u> do not occur in this position. Long vowels, <u>i:</u> <u>e:</u> <u>ɛ:</u> <u>ɑ:</u> <u>ɔ:</u> <u>o:</u> <u>u:</u>, occur initially, medially, and finally, but they do not occur in the first syllable of simple disyllabics with closed final syllable which have a vowel of the same group in the second syllable, nor in trisyllabic words with a long vowel in the second syllable.

Dogri short and long vowels occur in the following syllabic patterns:

(C)V	+	only in a few functional words (ci, ʃə, etc.)
(C)V:	+	ɑ: 'come', bɑ: 'breeze'.
(C)VC	+	məl(ə) 'paste', mɩl(ə) 'meet', pɷl(ə) 'bridge'.
(C)V:C	+	mi:l(ə) 'mile', me:l(ə) 'union', mɛ:l(ə) 'dirt', mɑ:l(ə) 'herd of animals', tɔ:l(ə) 'hurry', to:l(ə) 'weight', mu:l(ə) 'beginning'.
(C)VCC	+	gəlt(ə) 'wrong', pɷll(ə) 'softness'.
(C)V:CC	+	ɑ̀:lt(ə) 'condition'.
CVCV	-	
(C)VCV:	+	pətɑ: 'address', pɷjɑ:
CV:CV	-	
(C)V:CV:	+	ˈpɑ:lɑ: 'cold'.
(C)VCVC	+	ˈkədər(ə) 'respect'.
(C)VCV:C	+	bəˈzɑ:r(ə) 'market'.
(C)V:CVC	+	ˈɟɑ:gət(ə) 'boy'
CV:CV:C	-	
CVCCV	-	
(C)VCCV:	+	ˈbəggɑ: 'whitish', ˈpɷɟɟɑ: 'arrived'.
CV:CCV	-	
(C)V:CCV:	+	ˈbo:lli: 'speech', go:ḍḍɑ: 'knee'. ke:llɑ
(C)VCCVC	+	ˈkheɟɟəl 'distressed'.
(C)VCCV:C	+	mərˈɟɑ:d 'propriety of conduct'.
(C)V:CCVC	+	ˈɑ:lsən(ə) 'lazy woman', ɑ:kkhən 'saying'
CV:CCV:C	-	
CVCVCV:	+	ˈcəkəlɑ: 'bread-board'.
CVCVCCV:	+	cəˈkənnɑ 'watchful'.

This distributional description is at variance with Bahri's assertion that peripheral vowels do not appear before geminates. However, as shown in table 5 the difference of duration is preserved before geminates, and even if all informants are taken together, there is no overlapping at all between single tokens of the two vowel categories (see fig. 5); exactly the same

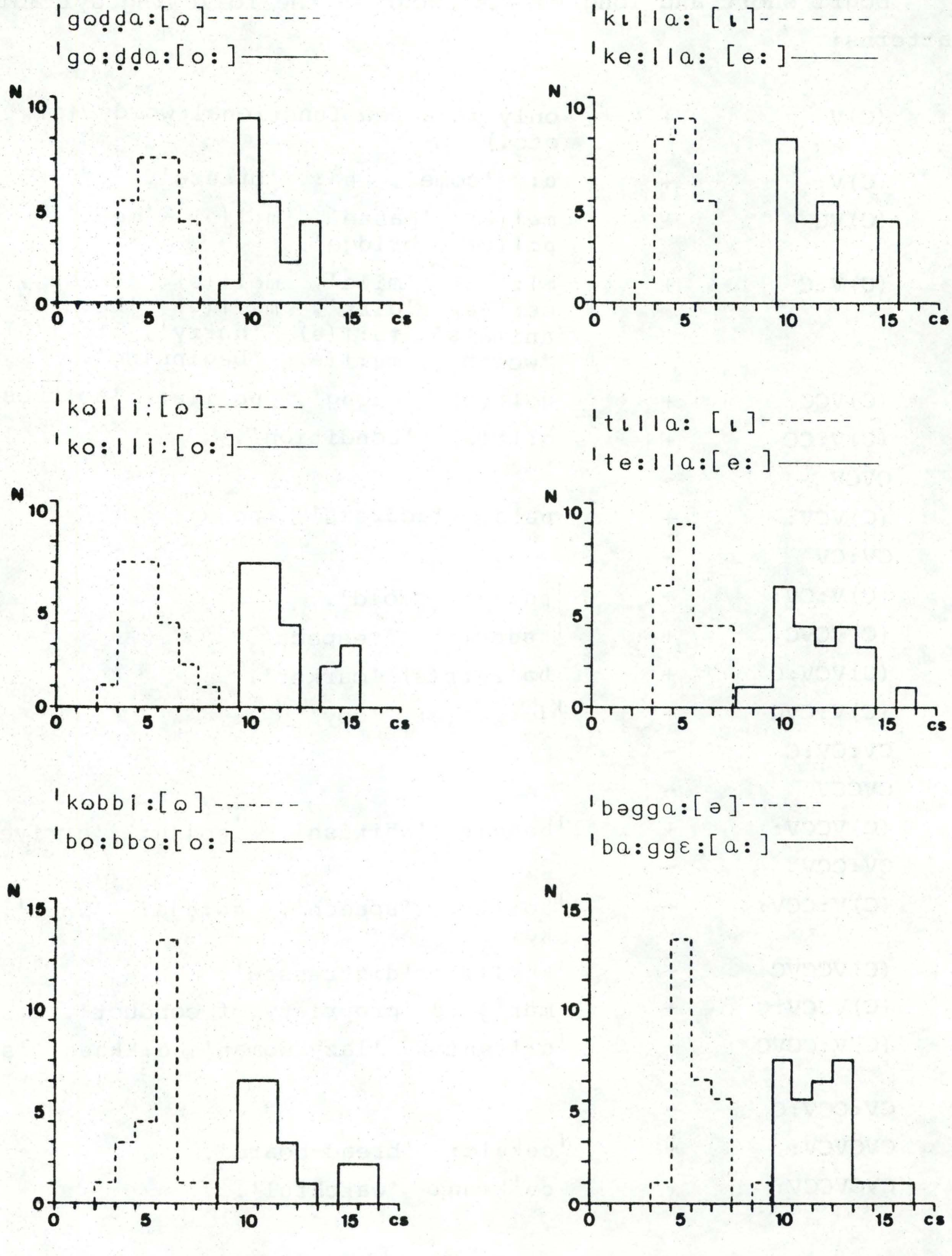


Figure 5

Duration of central (-----) and peripheral (——) vowels before geminated consonants. Five informants (N=23) combined (for ˈbəgga:/ba:ggɛ: six informants) (N=28-26).

picture was found for the word pairs dəbba:/da:bba:, sɷtta:/ so:tta:, gɷtta:/go:tta:, ḍɷbba:/do:bba:. Thus, when length is taken to be the distinctive feature, it must be stated that long and short vowels are kept clearly apart before geminates.

As for the difference in quality, the vowels are not confounded more often than in monosyllabics. The word pairs kɷlli:/ ko:lli:, gɷḍḍa:/go:ḍḍa:, kɷbbi:/bo:bbo:, tɩlla:/te:lla:, kɩlla:/ ke:lla:, and bəgga:/ba:ggɛ: were measured for the informants RN, SL, SD and VG. As far as the words with ɷ and o: are concerned, SL and RN have slightly more overlapping than in the monosyllabic words, VG approximately the same, but she has always a stronger F2 in o: than in ɷ . SD has no overlapping (in contradistinction to his monosyllabic words); the members of all three pairs are kept apart by means of F2, but the relation of F2 in o: and ɷ has been reversed in the pair kɷbbi:/bo:bbo: in which o: has a higher F2 than ɷ.

The words with e: and ɩ are kept apart by means of F2 by all informants (for SD and VG there is overlapping in the monosyllabic words). F2 of kɩlla: is somewhat uncertain but lower than in ke:lla: . The formant averages of tɩlla:/te:lla: are given in table 6 together with the formant averages of the pair bəgga:/ba:ggɛ:. This latter pair is kept more clearly apart than təl(ə)/ta:l(ə) , but ə in ˈbəgga: is shifted upwards in the vowel chart, i.e. it has a much lower F_1 and a somewhat lower F_2 than in təl(ə). The shift in F_2 is due to the place of articulation of the surrounding consonants: in təl(ə) the two dentals have a raised F_2, whereas b in bəgga: lowers F_2. As for F_1, the main reason is that ə is much shorter in bəgga:, and in position between two stops it becomes less open. This lax vowel is very much influenced by the surroundings, particularly when it is shortened. The placement of ə in the charts figs. 1-4 must therefore be taken with some reservation.

In any case, the two types of vowels are kept clearly apart as far as length is concerned, and at least as much as in monosyllabics as far as quality is concerned. Now it is possible that Bahri would not deny that the words mentioned above have peripheral vowels (although he speaks of neutralization before geminates), but that he considers the following consonant not as geminate, but as fortis. For a further treatment of this problem, see section 3.

Table 6

Average formant frequencies in Hz of the first vowels in ˈtɩllɑ:, ˈte:llɑ: and ˈbəggɑ:, bɑ:ggɛ: (N = 5)

		RN	SL	SD	VG
ˈtɩllɑ:	F_1	381	373	345	385
	F_2	1915	1615	1995	2250
	F_3	2560	2515	2570	2917
ˈte:llɑ:	F_1	353	381	349	422
	F_2	2195	1910	2188	2383
	F_3	2750	2580	2713	2942
ˈbəggɑ:	F_1	540	506	546	686
	F_2	1017	1073	1055	1260
bɑ:ggɛ:	F_1	698	678	711	892
	F_2	1100	1120	1190	1365

2.5 Phonological interpretation of the contrastive types ˈgəlɑ: - gəˈlɑ:

2.5.1 The problem

In section 2.3 it was stated that there is a consistent length difference between peripheral and centralized vowels and that this difference is more stable than the difference of quality. There is, however, also a difference of duration within the group of peripheral vowels in open final position, and there is quite a number of minimal pairs, e.g. (in narrow phonetic transcription):

	A.		B.
ˈdəri:	'carpet'	d(ə)ˈri:·	'daughter of husband's younger brother'
ˈpərəti:	'filling of earth'	p(ə)rəˈti:·	'brother's daughter'
ˈtəlɑ:	'sole of shoe'	t(ə)ˈlɑ:·	'water pond'
ˈgəlɑ:	'throat'	g(ə)ˈlɑ:·	'to say'
ˈbəgɑ:	'to flow'	b(ə)ˈgɑ:·	'cause to flow'

ˈpərɑ:	'to fill'	p(ə)ˈrɑ:·	'cause to fill'
ˈpəlɑ:	'to be brought up'	p(ə)ˈlɑ:·	'to make one drink'
ˈrəso:	'live (imp. v.)'	r(ə)ˈso:·	'kitchen'

Besides the difference of vowel length in the second syllable, there is also a difference of stress. Whereas the first syllable is stressed in the words in column A, the second syllable is stressed in the words in column B. Moreover, [ə] of the first syllable is shorter and may be optional in column B, but not in A.

The problem is how to explain these forms phonologically. It is possible to set up four different solutions.

(1) The first vowel (ə) may be regarded as long in words in column A and as short in words in column B. It means setting up a short-long opposition in a central vowel or splitting ə into two phonemes, one in the group of central vowels and another in the group of peripheral vowels.

(2) The first vowel (ə) may be regarded as being present in words in column A, and as being absent in words in column B. This means that ə in words of column B is considered as a non-phonological svarabhakti vowel which may be inserted to avoid heavy consonant clusters in the surface form. Then there will be no need to set up a short/long opposition in peripheral vowels.

(3) The second vowel (ɑ: i: o:) may be regarded as normal long vowels in words in column A and as over-long /ɑ:· i:· o:·/ in words in column B. Thus, central vowels will always be short, but there will be an opposition between longer and shorter peripheral vowels in final, open syllables.

(4) The difference of stress may be regarded as phonemic. Thus, stress is on the first syllable in words in column A and on the second syllable in words in column B.

2.5.2 The solutions 1 and 2 (ə)

Solution (1) is rather impracticable on the following grounds: Firstly, the difference of duration in the two types of ə is small, generally 2-3 centiseconds, i.e. much smaller than the normal minimal difference between long and short vowels (cf. tables 3-5). Phonetically, ə is always short and in this respect behaves like ɩ and ɷ . Moreover, a length difference in ə would be in contradiction to the general concomitance of length and quality. ə is certainly not peripheral phonetically, and phonologically it evidently behaves as a central vowel. The syllabic structure of Dogri does not tolerate two peripheral vowels in a disyllabic word with closed second syllable. As ə appears in such words with a peripheral vowel, it must be regarded as non-peripheral. Words like * ba:da:m, *ja:ga:t cannot occur in Dogri, but bəda:m and ja:gət do occur.

Moreover, there are various related words in Dogri in which a morphological alternation of peripheral vowels with non-peripheral vowels occurs. The vowel ə appears as a non-peripheral vowel in such alternations (see 2.3).

Solution (2) is also faced with difficulties.

First of all, the deletion of ə in the phonological form of words in column B will lead to the occurrence of a large number of strange consonant clusters which do not otherwise occur in Dogri.

Initial clusters in Dogri consist of stop + r, ṛ or l.
Final clusters in Dogri consist of (1) flap or lateral + stop, or sibilant or nasal, (2) sibilant + stop or nasal, (3) nasal + stop.

Deletion of ə will give rise to such initial clusters as

Voiced stop + voiced stop	bəga:	→	bga:
Unvoiced stop + unvoiced stop	təpa:	→	tpa:
Voiced stop + unvoiced stop	gəta:r(ə)	→	gta:r(ə)
Unvoiced stop + voiced stop	təbela:	→	tbela:
Aspirated stop + unaspirated stop	khəta:i:	→	khta:i:
Unaspirated stop + aspirated stop	təkha:	→	tkha:
Stop + lateral	təla:	→	tla:
Lateral + stop	ləta:ṛ(ə)	→	lta:ṛ(ə)
Sibilant + stop	səka:	→	ska:
Stop + sibilant	kəsa:	→	ksa:

Moreover, there seems to be a phonological distinction between stop consonant + r, and stop consonant + ə + r in a few cases, e.g. təra: 'cause to swim' and tra: 'fear'.

Secondly, ə occurs in free variation with ɩ a: in a restricted number of words, but not in others.

ˈphɩraː	past form of verb 'to walk about'	phɩˈraː· ∿ phəˈraː· 'to make one move about'
ˈpoɟɟaː	past form of 'to reach'	poˈɟaː· ∿ pəˈɟaː· 'to make one reach'
ˈlɩkhaː	past form of 'to write'	lɩˈkhaː· ∿ ləˈkhaː· 'to make one write'
ˈtokhaː	past form of 'to smoke'	toˈkhaː· ∿ təˈkhaː· 'to make a thing smoke'

If /ə/ is regarded as absent in words in column B, it cannot be explained how some words may also insert ɩ or o in consonant clusters. ɩ and o cannot be determined by the surroundings. Thus, neither the interpretation of ə in /ˈgəlaː/ as long, nor the interpretation of ə in /gəˈlaː/ as non-phonemic can be considered to give a satisfactory solution of the problem.

2.5.3 Solutions 3 and 4 (stress or quantity)

Solutions (3) and (4) are both plausible. The problem is that stress and vowel length co-occur, and it is difficult to decide which is the independent variable.

An argument against choosing stress as the relevant property is that stress is in most cases predictable (see 2.3). The only word types where it is not predictable are those under discussion here, i.e. disyllabic words with a light + a medium syllable (dəriː) or trisyllabic words with two light syllables + a medium syllable (pəgətaː). Moreover, non-Dogri speakers, also trained phoneticians, tend to hear the difference as one of length only, perceiving the second syllable as stressed in both types. A test with disyllabic Dogri words of different types has been carried out by Eli Fischer-Jørgensen. It consisted of 26 words said five times each by VG in a frame with non-final intonation, and 16 of the same words said three times as the last words in a frame with final intonation. The listeners were Danish phoneticians and students of phonetics, 14 in the first case and 8 in the second. The result was that, when said with non-final intonation, the words ˈpəlaː, ˈdəriː, ˈkə́ṛaː, ˈkə́riː, ˈkə̀riː, ˈpə̀raː showed a great majority (90%) for stress on the second syllable, whereas with final intonation the words with high tone (ˈkə́riː and ˈkə́ṛaː) were heard as stressed on the first syllable (90%). ˈpəlaː had a small majority for first syllable stress (58%), but ˈpə̀raː

with falling tone was still heard as stressed on the second syllable (92%). Eli Fischer-Jørgensen, who listened to the recordings of all informants, also heard the type ˈCəCɑ: as stressed on the second syllable in all cases. But this, of course, does not prove that, from the point of view of Dogri structure, the stress is not on the first syllable. The reaction of the listeners may be due to the fact that they were not accustomed to hearing the type: stressed syllable with short ə + unstressed syllable with longer full vowel and often rising pitch, which is a very rare type, at least in European languages. (With final intonation, the second vowel was shortened so that the difference of length was small, and it was also somewhat lower in frequency.)

An argument against choosing length is that we would thus get three degrees of length in vowels. This is not a very strong argument, though, since there would only be two degrees in each position: long and overlong in open final syllable, short and long elsewhere, but it is, at any rate, not possible to identify the final vowels in words like ˈdəri: and ˈpəlɑ: ˈrəso: with short centralized vowels.

The strongest arguments, however, come from two facts of Dogri phonological structure, which point to stress as an indispensable part of the phonological system: (a) stress determines the place of the word tone, (b) it must be taken into account in the rules for nasalization. - The present writer has therefore chosen stress as the distinctive difference of the two word types.[1]

(a) Tone and stress: Dogri has tonal distinctions and there is mutual interaction between stress and tone. Dogri has three phonologically distinct tones:

I. Neutral tone, which is unmarked, often slowly rising.

II. Falling tone, which starts at a point generally higher than the middle of the voice range of the speaker and then falls to the lowest point, from where it generally rises somewhat again.

III. Rising tone, which starts at a point generally lower than the middle of the voice range and then rises to the highest level or at least to a level higher than the middle of the voice range (it may fall somewhat again at the end).

---

1) Bahri does not treat this problem, but he puts a stress mark in words of the type dəˈri: (with even tone), which indicates that he is of the same opinion.

I.	ra:	'opinion'	I.	ka:r	'work
II.	rà:	'tune'	II.	kà:r	'house'
III.	rá:	'way'	III.	ká:r	'line'

These are word tones in the sense that only one significant tone occurs on a simple word. They can, however, be described as syllabic tones in the sense that the start of the tone may occur on any one of the syllables in a polysyllabic word, while other syllables are adjusted to the starting point and the end point of the tonebearing syllable. The position of tone is thus distinctive and which syllable of the word will carry the tone is determined by stress. Thus all words which are placed in column A have the start of the tone on the first syllable peak, and words in column B on the second syllable peak. This difference is more clearly perceived in the case of falling and rising tones:

A		B	
ˈpə̀ṛa:	'read'	pəˈṛà:	'teach'
ˈkə́ṛa:	'boil'	kəˈṛá:	'frying pan'

Historically, the tones in Dogri are connected with aspiration and stress accent of classical Sanskrit (OIA) and Prakrits (MIA). If the aspiration (generally voiced) followed a stressed vowel, it gave rise to a rising tone, and if it preceded a stressed syllable, it gave rise to a falling tone. This phenomenon is seen even now in new loanwords which Dogri has borrowed from Hindi or Urdu. The following two words of Hindi are realized with two different tones in Dogri, due to their difference of stress:

Hindi			Dogri
svəˈbhav	'nature'	→	sɒbà:
ˈsəbha	'meeting'	→	sə́ba:

The Urdu word ˈsubəh 'morning' is realized as sɒ́ba: in Dogri. Tone alternations in such word pairs as lə́ba/ləbà, kə́ḍa:/kəḍà:, pə́ṛa:/pəṛà: are easily explained by stress, which preceded aspiration (historically) in one case and followed it in another. From a synchronic point of view, it can be stated that stress is significant in all these cases and is realized phonetically by duration and pitch movement.

It might be possible to set up underlying forms with aspiration in all such words where low or high tone appears in surface forms and derive the surface forms by means of the following rules:

1. The aspirated h occurring as an independent consonant or as part of a voiced aspirated consonant disappears when followed by a stressed vowel in the same syllable or any following syllable, giving rise to low tone on that syllable.

|ˈhɑːr|[1]	ɑ̀ːr	'garland'
|bəˈhɑːr|	bɑ̀ːr	'season'
|gəˈhɑːr|	gɑ̀ːr	'a village given in donation'
|səməˈɟhɑː|	səməɟɑ̀ː	'to make one understand'
|ləˈbhɑː|	ləbɑ̀ː	'to cause to find out'
|pəˈɽhɑː|	pəɽɑ̀ː	'to teach'
|bəˈdhɑːɳɑː|	bədɑ̀ːɳɑː	'to cause to increase'

2. If a voiced aspirate is not preceded by a vowel, the voiced stop also loses its voice.

|ˈghɑːr|	kɑ̀ːr	'house'
|ˈbhɑːr|	pɑ̀ːr	'weight'
|ˈbhrɑː|	prɑ̀ː	'brother'

3. The aspirated h, occurring as an independent consonant or as part of a voiced aspirated consonant, disappears when a stressed vowel precedes it in the same syllable or any preceding syllable, giving rise to a high tone on that syllable.

|ˈkɑːhr|	kɑ́ːr	'line'
|ˈbɑːhr|	bɑ́ːr	'outside'
|ˈlɑːbh|	lɑ́ːb	'profit'
|ˈləbbhɑː|	lə́bbɑː	'found out'
|ˈsəməɟhɑː|	sə́məɟɑː	'understanding'
|ˈpəɽhɑː|	pə́ɽɑː	'to study'
|ˈloːhɑː|	lóɑː	'iron'
|ˈlɑːh|	lɑ́ː	'use'

---

1) Vertical lines are used to indicate underlying forms.

But even if these historically induced underlying forms are not set up, stress must be considered phonemic in Dogri since the place of tone is distinctive.

(b) Nasalization and stress: The phenomenon of nasalization in Dogri is also related to stress. Nasalization is phonemic in final open syllables but is predictable and non-phonemic in other cases.

/bɑ:/	'wind'	/bɑ̃:/	'water reservoir'
/jɑ:/	'go' (2nd person)	/jɑ̃:/	'go' (1st person)
/tɑ:/	'heat'	/tɑ̃:/	'then'
/ku:/	'speak'	/kũ:/	'I should speak'
/mə'nɑ:/	'console' (imp.)	/mə'nɑ̃:/	'I should console'
/kə'mɑ:/	'earn'	/kə'mɑ̃:/	'I should earn'

But nasalization is predictable in other positions. A stressed vowel followed by a nasal consonant belonging to the same syllable or the following syllable is allophonically nasalized, and if it is followed by an unstressed vowel, that vowel too is nasalized.

['jɑ̃:n]	/jɑ:n/	'life'
['jɑ̃:ŋ]	/jɑ:ŋ/	'(I) will go'
['mə̃nɑ̃:]	/'mənɑ:/	'refuse'
['jə̃mɑ̃:]	/'jəmɑ:/	'addition'
['dũ:ṇɑ̃:]	/'du:ṇɑ:/	'double'

But this allophonic nasalization does not occur in a stressed vowel following a nasal consonant.

[də'mɑ:k(ə)]	'brain'
[tə'mo:l(ə)]	'money given at the marriage ceremony'
[sənɛ́:r(ə)]	'brightness'

Moreover, a cluster consisting of a homorganic nasal plus a consonant tends to be simplified word-finally and when preceded by an unstressed vowel and followed by a stressed vowel (the preceding vowel is always nasalized). In the cluster of a homorganic nasal with a velar and a palatal stop, the velar nasal and the palatal nasal are preserved, while the stop is lost (a); in other cases the nasal consonant disappears, while the other consonant is preserved (b).

(a) /ˈḍi:ŋg/ [ˈdĩ:ŋ] 'boasting'
/səɲɟ/ [ˈsə̃ɲ] 'evening'
/rəŋga:i/ [rə̃ˈŋa:i] 'dyeing'
(b) /gənḍɑ̀:i/ [gə̃ˈḍɑ̀:i] 'repairing'

But the cluster is not simplified when a stressed vowel precedes it as in ˈgənḍa:, ˈgənda:, ˈba:ndər.

Thus, both non-phonemic nasalization and simplification of clusters with nasals are determined by stress.

As mentioned above, the place of tone is also determined by stress, and these two factors together form, in my opinion, a strong argument for regarding stress as part of the phonological system of Dogri.

2.6 Phonetic analysis of the problematic word types

2.6.1 Duration

A phonetic analysis of syllable duration in types A (ˈgəlɑ:) and B (gəˈlɑ:) supports the interpretation of the difference as one of stress.

Ten words from list I could be compared pairwise in this respect, viz. ˈgəlɑ:/gəˈlɑ:, ˈpəlɑ:/pəˈlɑ:, ˈbəgɑ:/bəˈgɑ:, ˈpəkhɑ:/pəˈkhɑ:, and ˈpətɑ:/kəˈtɑ: . These words were spoken five times each by the informants RN, SL, SD, CS and VG, and three times each by AS (AS did not read the pair ˈpətɑ:/kəˈtɑ:). Moreover, the following words, spoken five times by VG (in the test list) could be compared pairwise: pə̀rɑ:/pəˈrɑ̀: and ˈkə́ṛɑ:/kəˈṛɑ́:.

All segments of these words have been measured in centiseconds on the basis of mingogram recordings comprising a duplex oscillogram, two intensity curves, highpass filtered at 500 and 2000 Hz, a high fidelity intensity curve, and a pitch curve. Since the words were spoken in the frame ɷs(ə) ... soṇe:ɑ: , the initial consonant could also generally be delimitated. It gave, however, some problems that the speakers sometimes made a short pause after the word ɷs(ə) of the frame, which could not be distinguished from a voiceless closure. For initial stops the measurement of the initial consonant is therefore somewhat uncertain. But the delimitation of vowels and medial consonants did not generally present any serious problems. The start of the medial consonant was taken to be at the point where the oscillations of the duplex oscillogram are diminished abruptly, where the intensity curve, correspondingly, decreases abruptly, and where the pitch curve drops out or decreases abruptly. This point also corresponds to the cessation of the vowel formants on the spectrograms. The closure of stops is reckoned from this point to the explosion, and the aspiration (or, for unaspirated stops, "the open interval") is reckoned from the explosion to the steep rise of the intensity curve for the next vowel.

Averages were taken of the five recordings for each informant and differences between the averages of the word pairs calculated. Table 7 contains the average durations of the segments of each word for all informants combined and the differences between the corresponding averages for each word pair.

The significance of the overall averages can be judged by looking at the separate averages. For each segment of a given position (e.g. the first consonant in type A words), there are 31 averages (viz. 6 informants x 5 words (minus one for AS, plus two for VG), which can be compared with the corresponding 31 averages for word type B.

The main difference obviously lies in the second vowel. It is longer in type B words than in type A words, the average difference for all words being 8.7 cs. This difference is evidently significant. It is valid in 30 out of 31 pairs of averages (CS, who has spoken very quickly and shortened the final vowels, so that she has an average difference of only 2.5 cs, has one average without any difference, but even she has no overlapping in single tokens with the exception of the word pair ˈpəkhɑː/təˈkhɑː).

The preceding consonant of the second syllable is also slightly longer in type B than in type A words. This difference is small, 0.8 cs on the average, but it is stable and is valid for 26 out of 31 pairs of averages (in two pairs there is no difference, and in three cases the relation is inverted). This means that it is not only a question of the duration of the second vowel but one of syllable length, which indicates that it is a manifestation of stress. (The difference in the total length of the second syllable in type A and type B words is 9.7 cs.)

Table 7

Average durations (in cs) of segments and syllables of word types A and B for all informants combined.

		N.	cons.	vowel	1.syll.	cons.	vowel	2.syll.
A	ˈgəlaː	23	10.2	6.6	16.8	4.9	11.3	16.2
B	gəˈlaː	23	10.6	4.5	15.1	6.0	19.4	25.4
	diff. A-B		-0.4	+2.1	+1.7	-1.1	-8.1	-9.2
A	ˈpəlaː	23	13.1	6.2	19.3	4.7	11.9	16.6
B	pəˈlaː	23	12.2	3.4	15.6	6.2	18.5	24.7
	diff. A-B		+0.9	+2.8	+3.7	-1.5	-6.6	-8.1
A	ˈbəgaː	23	11.3	7.9	19.2	4.4	12.3	16.7
B	bəˈgaː	23	10.7	6.2	16.9	5.5	20.4	25.9
	diff. A-B		+0.4	+1.7	+2.1	-1.1	-8.1	-9.2
A	ˈpəkhaː	23	13.0	6.2	19.2	9.3	12.2	21.5
B	təˈkhaː	23	11.8	4.9	16.7	10.1	19.5	29.6
	diff. A-B		+1.2	+1.3	+2.5	-0.9	-7.1	-8.1
A	ˈpətaː	20	14.8	7.0	21.8	6.7	14.5	21.2
B	kəˈtaː	20	11.7	3.8	15.5	7.5	20.8	28.3
	diff. A-B		+3.1	+3.2	+6.3	-0.8	-6.3	-7.1
VG only								
A	ˈpə̀raː	5	12.5	9.8	22.3	2.0	17.7	19.7
B	pəˈrà:	5	12.3	7.0	19.3	2.0	34.2	36.2
	diff. A-B		+0.2	+2.8	+3.0	0	-16.5	-16.5
A	ˈkə́ṛaː	5	14.2	9.2	23.4	2.0	17.7	19.7
B	kəˈṛáː	5	13.2	7.6	20.8	2.2	27.0	29.2
	diff. A-B		+1.0	+1.6	+2.6	-0.2	-9.3	-9.5
Difference A-B between all words combined.			+1.0	+2.2	+3.2	-0.8	-8.9	-9.7

This is supported by the relations between the first syllables. Here the vowel is also longer when stressed, i.e., it is longer in type A than in type B words. But the difference is considerably smaller than in the second syllable (2.2 cs on the average), which may be explained by the fact that the vowel is short and cannot be lengthened so much. There is generally a much larger scatter of the values in long vowels than in short vowels. In a quantity opposition, the short member may be considered as a point, whereas the long member is stretchable (Lehiste 1979, p. 36). The small difference is, however, very stable. It is valid in 28 out of 31 pairs of averages and evidently significant. (It was mentioned in 2.5 that the vowel may disappear completely in type B words. In the present material this has, however, happened in only four single tokens, spoken by VG.)

As for the initial consonant of the first syllable, the difference is not stable. It is longer in type A than in type B words in 21 out of 30 pairs of averages but, as mentioned above, the delimitation has not always been quite certain. The difference adds, however, to the difference of the syllables, the first syllable being longer in type A words than in type B words in 27 out of 30 pairs of averages (in one case the first consonant has not been measured because of a long pause).

Type A (ˈgəlɑ:) is not a very common word type in Dogri. When the first syllable is stressed, the following consonant is often geminated. There are thus more pairs like ˈbəggɑ:/bəˈgɑ: (the type ˈbəggɑ: will be called type C in the following). The duration of the second vowel in 12 such pairs has been measured for the subjects RN, SL, SD, CS and VG.

The difference between the final vowels in type C (e.g. ˈbəggɑ:) and B (e.g. bəˈgɑ:), viz. 9.0 cs (all subjects combined), is practically the same as the difference found between type A (e.g. ˈbəgɑ:) and type B, viz. 8.9 cs.

As for the vowel of the first syllable, there is hardly any difference between B and C. This is, however, due to the fact (mentioned in 2.3) that vowels are shortened before geminated consonants. This appears very clearly from a comparison between triplets such as ˈbəgɑ: (A), bəˈgɑ:· (B), ˈbəggɑ: (C). There are four such comparable triplets in the material, viz.

A	B	C
ˈbəgaː	bəˈgaː·	ˈbəggaː
ˈpə̀khaː	təˈkhà:·	ˈpəkkhaː
ˈpətaː	kəˈtaː·	ˈkəttaː
ˈkəthaː	cəˈthaː·	ˈkətthaː

The duration of the first vowel was measured for the subjects RN, SL, SD, CS, and VG.

For all subjects combined, the difference between A and B is 2.1 cs, and the difference between A and C is 2.2 cs. This means that on the average the vowel is shortened to the same degree when unstressed and when followed by a geminated consonant, and the reason why ə in C (ˈbəggaː) is not shortened compared to B (bəˈgaː) (the difference is 0.1 cs only), although it is followed by a geminated consonant, is that it is stressed. Stress has thus the same influence here as in the preceding cases.

Other examples in the material, where only A and C or B and C can be compared, support this result. The difference in duration between the first vowels in A and C (i.e. vowels in stressed syllables before single or geminated consonants) is on the average 2.7 cs for the pairs ˈkəcaː/ˈkəccaː, ˈkɒliː/ˈkɒlliː and ˈkɩlaː/ˈkɩllaː, and almost the same (2.7 cs) before final geminate (sət(ə)/sətt(ə), ɟəd(ə)/ɟədd(ə), pɒl(ə)/pɒll(ə), ɟəg(ə)/ɟəgg(ə), gəɟ(ə)/gəɟɟ(ə)). In the case of two long vowels it is slightly larger, viz. 3.6 cs for ˈboːliː/ˈboːlliː and ˈɟaːdaː/ˈɟaːddaː. Moreover, as was the case in the triplets, the first vowel has practically the same duration in types B and C, the general average of the difference for the pairs ləˈkaː/ˈləkkaː, bəˈsaː/ˈbəssaː, kəˈsaː/ˈkəssaː, dəˈbaː/ˈdəbbaː, kəˈḍaː/ˈkə́ḍḍaː, gɒˈḍaː/ˈgɒḍḍaː, ḍəˈphaː/ˈḍəpphaː, səˈdaː/ˈsəddaː being only 0.1 cs.

2.6.2 Intensity

As for physical intensity it is a general experience that it does not have much to do with stress. This is also the case here. Both type A and B may have higher intensity either on the first or on the second syllable, and there is no evident relation with stress, except to a certain extent for VG, whose first syllable in type B (gəˈlaː) is often much weaker than the second syllable, whereas the relation is variable and the difference between the syllable peaks smaller in type A. This may be due to the fact that her first vowel is often particularly short in type B. More often there is a certain correlation between intensity and pitch, particularly so that high pitch is combined with high intensity, but there may, on the other hand, be a definite drop in pitch, e.g. in tone 2, without any corresponding drop in intensity.

2.6.3 Pitch

There is often a close correlation between pitch and stress. This is, e.g., the case in Danish. In Dogri, which is a tone language, one cannot expect to find a consistent relation in pitch between stressed and unstressed syllables. But, as mentioned in 2.5, according to the auditory impression the position of tone in disyllabic words is determined by stress, so that the tone movement starts in the stressed syllable. This impression is, on the whole, confirmed by a phonetic analysis (carried out by Eli Fischer-Jørgensen). It is particularly evident for tone 2 and 3. In the list used for testing stress perception, VG had spoken the words ˈkə́ṛɑː, kəˈṛɑ́ː and ˈpə̀rɑː, pəˈrɑ̀ː five times each. In the case of tone 2 (the high rising tone), it is very clear that the pitch movement takes place on the stressed syllable. In ˈkə́ṛɑː the first syllable is high rising, and the second syllable still higher, but level, or slightly rising (or, in final position, rising-falling), in kəˈṛɑː the first syllable is level (or slightly rising or falling), and the second syllable starts slightly higher and is strongly rising (see fig. 6).

In the case of tone 3 (the falling-rising tone), a considerable part of the fall takes place on the stressed syllable. In ˈpə̀rɑː about half of the fall takes place on the first syllable, and the second syllable continues the steep fall and then rises again (there is only one case in which most of the fall takes place on the second syllable). In pəˈrɑ̀ː, on the other hand, which has stress on the second syllable, the first syllable is level or slightly falling, and the second syllable starts at the level where the first ends and shows an extensive and rather slow fall, followed by a steep rise (which is less pronounced in final position) (see fig. 6).

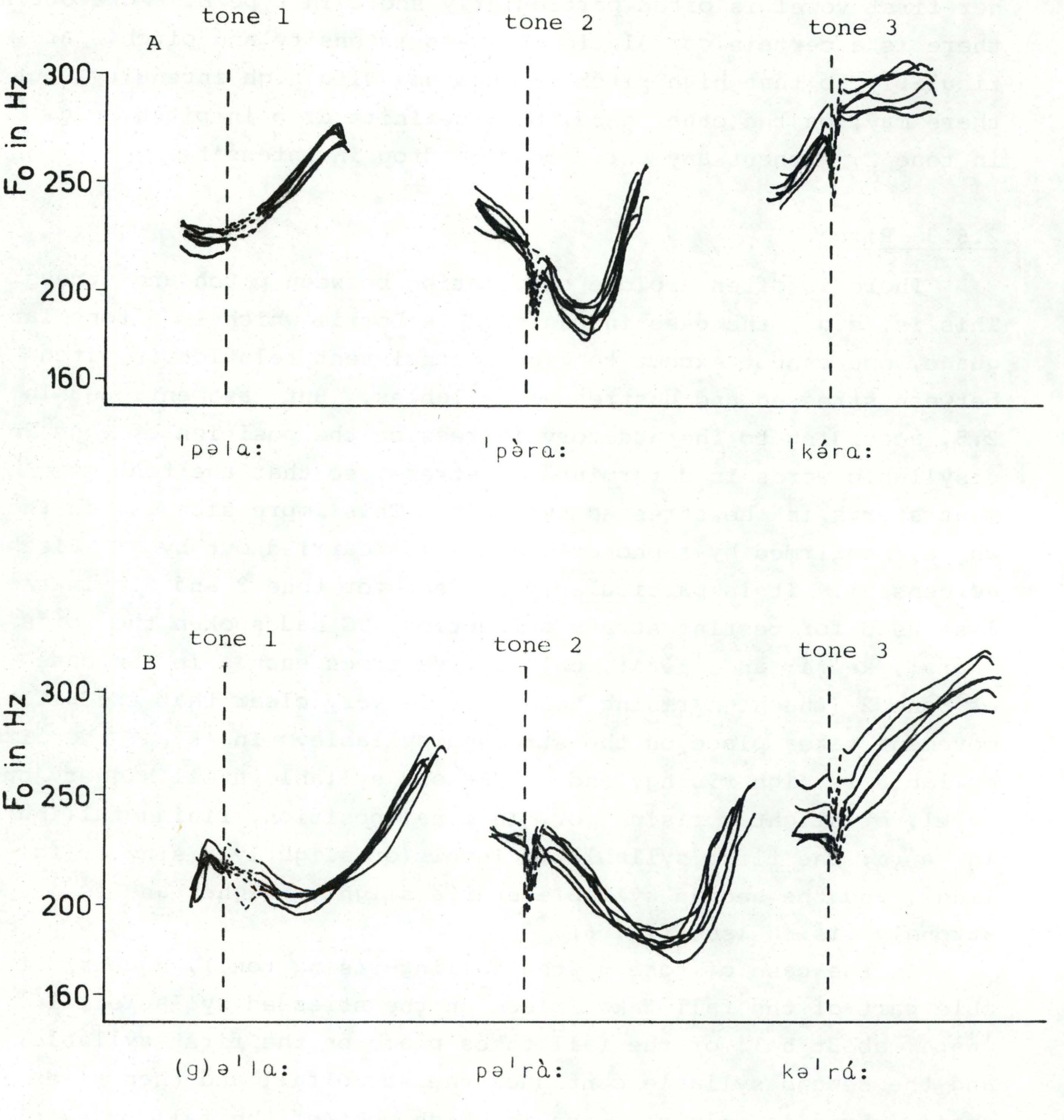


Figure 6

Tracings of words with tone 1,2, and 3 with stress on the first (A) and the second (B) syllable, spoken by VG. The tracings of six vowels are superimposed. Line-up point: start of the second syllable. The medial consonant is traced in broken line.

This analysis is corroborated by other words of the list. There are four more words with high tone and stress on the first syllable: ˈkhíṛa: looks like ˈkə́ṛa, having rising pitch on the first syllable and high pitch on the second, ˈká:la: has strong rising pitch on the first a:, while the second syllable starts at the level where the first ends and is slightly falling, and in ˈkílla: the first vowel and the following l have rising pitch and the second syllable falling or level pitch. There are also four words with low tone and stress on the first syllable. ˈpìṛa: and ˈpà:ra: have the whole of the fall on the first syllable (it should be noted that r and ṛ are very short), whereas pò:li: and tìlla: have fall on the first vowel, continuing in l, and the rise starting in the second vowel. Of the words from list I which were compared for duration, two have falling tone, namely ˈpə̀kha: and təˈkhà:. Here VG has not much of a fall in the first syllable of ˈpə̀kha:, and also the second syllable is level, but it is obviously lower, whereas in təˈkhà: the first may be slightly falling, but the second syllable starts at the same level as the first and is falling. SD has a clear difference, most of the fall being on the first syllable in ˈpə́kha:, the second being rising or falling-rising, whereas in təˈkha: it is true that the first syllable is also falling, but the second syllable starts higher and has a much more extensive fall. SL has a similar, but less clear difference (the second syllable of təˈkha: may start at a lower point than the first). RN has a clear fall in the first syllable and a lower start of the second syllable of both words, but in ˈpəkha: the second syllable is rising, whereas in təˈkha: it continues the fall before rising. As for AS and CS, they do not have much pitch movement in these words, and the first syllable is generally level, so that the fall is on the second in both cases, but in type B the pitch generally starts higher in the second syllable than in the first, whereas in type A it is the opposite. This difference is, however, not consistent.

As for tone 1 (the neutral tone), the relation to stress is less clear. In some cases it is, however, evident. VG's neutral tone normally has a rising pitch movement, starting at a much lower pitch than the high rising tone, and generally beginning with a small fall. If the first syllable is stressed (type A), the lowest point is in the vowel of that syllable; if it is unstressed, the lowest point is in the second syllable, i.e. the rise has its starting point in the stressed syllable (see fig. 6). This is particularly clear in the words from the test list. This means that if the first syllable is stressed, the whole second vowel is rising; if the second syllable is stressed, the second vowel starts with a fall and then rises. If the intervening consonant is voiced and has a certain length (e.g. l), then it can be clearly seen that it is rising after a stressed first vowel and falling after an unstressed first vowel. (This, by the way, is in complete accordance with the pitch movement of l in Danish (ˈbilist vs. biˈlist).) The same description is valid for the words ˈgəlɑː/gəˈlɑː and ˈpəlɑː/pəˈlɑː in list I. It is also valid for the same words spoken by RN, and partly for SD and SL; for the latter two informants there is some overlapping, and SL's curves are, on the whole, very flat. For CS it is only valid for ˈgəlɑː/gəˈlɑː, for AS no clear difference can be seen. In the pair ˈbəgɑː/bəˈgɑː, VG, RN, and SD show falling-rising pitch on the last ɑː when it is stressed, exclusively rising pitch when it is unstressed. For ˈpətɑː/kəˈtɑː the difference is that the first syllable is lower than the second in ˈpətɑː, whereas it is higher in kəˈtɑː, and the second syllable is falling-rising in the latter word, which again means that the lowest point is in the stressed syllable. This is valid for VG, RN, and SD, but for SL and CS it is only a tendency, and for AS no difference could be seen.

On the whole, it must be stated that the most adequate description of the relation between stress and tone is that the pitch movement starts in the stressed syllable, whereas the most extensive movement may be in the second, unstressed syllable if the first syllable is stressed and has a short vowel. In a few cases, however, e.g. ˈpə̀khɑː/təˈkhɑ́ː in RN's pronunciation, where the first syllable has falling pitch in both words, it might be more adequate to say that it is the nucleus of the tone that is on the stressed syllable. This problem deserves a more thorough investigation.

3. Consonant gemination and consonant weakening in Dogri

3.1 Introduction

There exists in Dogri an opposition between geminated[1] consonants and single consonants in intervocalic and final positions. Some single consonants are weakened in these positions to such an extent that they change into fricatives, flaps, or semivowels. This phenomenon, which has been referred to casually in previous descriptions of the language, has been studied further in this paper.

Dogri has twenty-eight consonants:

	Velar	Palatal	Retroflex	Dental	Labial	
Stops, voiceless	k	c	ṭ	t	p	
Stops, voiceless, aspirated	kh	ch	ṭh	th	ph	
Stops, voiced, unaspirated	g	ɟ	ḍ	d	b	
Nasals	ŋ	ñ	ṇ	n	m	
Flaps			ṛ	r		
Lateral				l		
Sibilants		ʃ		s		h
Semivowels		j				

All consonants except ṇ j w can occur in initial position. All consonants except the semivowels j w can occur in final position. All consonants except ʃ r ṛ ṇ j w can be geminated in medial and final positions. In final positions they are often followed by [ə], which is not phonemic. The phonemic status of h is doubtful as it occurs only in a few forms of the verb 'to be' in the past tense, where it can be regarded as an allophone of th, and in one word, pɛha, where it occurs in free variation with the dental sibilant s. The following minimal pairs show the contrast between geminated and single consonants after short (central) and long (peripheral) vowels.

---

1) The term "geminated consonant" is used in this paper as a common designation both for long consonants distributed on two syllables (geminates in the narrow sense) and for final long consonants.

After short vowels:

ˈkəccɑ:	'unripe'	ˈkəcɑ:	'bad' (man)
ˈpəttɑ:	'leaf'	ˈpətɑ:	'address'
ˈbəggɑ:	'whitish'	ˈbəgɑ:	'to flow'
ˈkətthɑ:	'catechu'	ˈkəthɑ:	'story'
ˈsəddɑ:	'invitation'	ˈsədɑ:	'always'
ˈrəssɑ:	'rope'	ˈrəsɑ:	'soup'
pɒll(ə)	'softness'	pɒl(ə)	'bridge'
ˈcəppəl(ə)	'sandal'	ˈcəpəl(ə)	'naughty'

After peripheral vowels:

ˈɟɑ:ddi:	'freedom'	ˈɟɑ:di:	'much'
ˈsi:ttɑ:	'was sewn'	ˈsi:tɑ:	a name
ˈbo:lli:	'speech'	ˈbo:li:	'deaf' (woman)
ˈkɛ:ddɑ:	'in prison'	ˈkɛ:dɑ:	'good manners'
ˈrɑ:ɟɟɛ:gi:	'to the kingdom'	ˈrɑ:ɟɛ:gi:	'to the king'

3.2 Historical background

It may be of some interest to know the historical development of geminated and weakened consonants. Ancient Indian works on phonetics, while describing various types of combinations of consonants, refer to two types: "dārupiṇḍa", which refers to the combination of stop + semivowel (literally: 'block of wood which can easily be broken', and "abhinidhāna" ('close contact'), which refers to the non-release of a consonant when followed by a stop or a pause. Abhinidhāna is described as the checking of a consonant making it obscure, weakened, deprived of breath and voice. The sound which follows the closure of an unreleased stop is called "dhruva" (literally: 'continuance'); it is audible in the case of voiced stops but not audible in the case of a voiceless stop (Allen, 1965).

This indicates that in Old Indo-Aryan there existed consonant clusters of the type stop + stop, and it was observed that the first member of the cluster could not be released and hence lacked some of its features like aspiration and voice. This obscure and unreleased pronunciation of the first consonant of the consonant cluster resulted in assimilation of consonant clusters in Middle

Indo-Aryan languages like Pali and Prākrits. The explosive consonant completely overpowered the implosive if the latter originally belonged to a similar category or a weaker category. If the first member belonged to a stronger category, then the second member was assimilated to the first one. The order of dominance which is found in consonant assimilation is: 1. Stop, 2. Sibilant, 3. Nasal, 4. l, 5. w, 6. j, 7. r. Thus we find such assimilation as:

OIA	MIA	
bhakta	bhatta	'cooked rice'
utpala	uppala	'lotus'
sapta	satta	'seven'
sarpa	sappa	'serpent'
cakra	cakka	'wheel'
raśmi	rassi	'rope'
ātmā	atta	'self'
kalya	kalla	'tomorrow'
mūlya	mulla	'price'
puṇya	puṇṇa	'good deed'
catuṣka	caukka	'square courtyard'

These geminated consonants have been simplified in Hindi and many other modern Indian languages with a compensatory lengthening of the preceding vowel if it was short in Middle Indo-Aryan. Panjabi, Lehanda and Dogri have retained these geminated consonants. In Dogri, this tendency to gemination has been further extended to certain morphophonemic forms. A single consonant of the base form is generally changed into a geminated consonant in inflectional forms if the suffix begins with a vowel and if the syllabic pattern of the word permits consonant gemination, i.e., if the consonant is preceded by a stressed vowel. Thus, historically, consonant gemination is to be traced back to assimilation of consonant clusters, but now it has become a part of the syllable structure of Dogri.

As regards weakening of intervocalic consonants, the tendency is seen even in OIA. Ṛgvedic spelling "ḷ" for intervocalic "ḍ" and "ḷh" for ḍh and occurrence of "ṇ" for n in Sthaṇu indicate this weakening (Bloch, 1965). This tendency of weakening intervocalic single consonants increased in the MIA stage along with the tendency of assimilation of consonant clusters. Weakening appeared in the form of voicing of intervocalic unvoiced consonants and vocalization of intervocalic voiced consonants.

OIA	MIA	
jānāti	jānādi	'knows'
ākāsa	āgāsa	'sky'
karoti	karadi	'does'
mati	madi	'intelligence'

Change of dental n to retroflex, flapped ṇ as a form of weakening is also very common.

In some Prākrits like Mahārāṣṭrī, the original unvoiced stops were first voiced and thereafter vocalized, thus merging with the original voiced stops. OIA kapi, Mahārāṣṭrī kai.

This development does not seem to have affected the Prākrits from which Panjabi and Dogri are derived.

As regards Dogri, it shows a peculiar development of weak consonants to fricatives. Sanskrit had no fricatives except v and the voiceless sibilants. As regards MIA, the voiced consonants which were changed to glides might have passed through an intermediate fricative stage which is not recorded in writing (Bloch, 1965). North Western Prākrits of the Kharoṣṭhi documents found in Central Asia, however, show the change of the intervocalic consonants k c ṭ t p into voiced consonants g j ḍ d b, and k c ṭ p (and also g j ḍ b (?)) further become spirants ǵ ś (j́) đ v (Burrow, 1937). It appears that just as Dogri has preserved some of the peculiarities of the Prākrits of Aśokan inscriptions discovered in North West India at Shāhbāzgaṛihi and Mānsehrā, similarly it has developed fricative-like pronunciations of weak consonants due to its relationship with North Western Prākrits. The area was under the influence of Iranians for some time, and it is possible that North Western Prākrits developed this tendency due to the influence of Iranian. Opposite to the Middle Indo-Aryan in which there is a tendency to assimilation, the Iranian preserves both consonants by changing stop + stop into fricative + stop. The following related words show this difference:

OIA	MIA	Iranian
bhakta 'cooked rice'	bhatta 'cooked rice'	baxt 'fate'
sapta 'seven'	satta 'seven'	haft 'seven'
pakva 'ripe'	pakka 'ripe'	puxt 'strong'
dugdha 'milk'	duddha 'milk'	duxta 'milk'
vikta 'empty'		bixt 'empty'
kṣubdha 'angry'		asufta 'angry'.

Dogri has borrowed various Persian words. While initial fricatives in such words are changed to stops in Dogri, medial fricatives are generally preserved.

It is interesting to find that some Dogri words show a two-way treatment of OIA consonant clusters, one showing assimilation of the unreleased consonant to the released one, and the other showing preservation of both consonants by fricativizing the first member.

OIA	bhakta	
Dogri	pə̀tt(ə) (cooked rice)	pə̀ɣ(ə)t(ə) (worshipper)

OIA	rəkta	
Dogri	rətt(ə) (blood)	rəɣ(ə)t(ə) (a disease)

3.3 The relation to stress and syllable division

An initial consonant evidently belongs to the following vowel, and a final consonant belongs to the preceding vowel. An intervocalic consonant can belong either to the preceding vowel or to the following vowel or to both vowels. In a word like bəga:, the consonant may belong to ə and thus behave like a final consonant detached from the last vowel bəg/a: [ˈbəɣa:]. It may belong to the final vowel a: and behave like a syllable-initial consonant: bə/ga: [bəˈga:]. It is also possible that it may belong to both vowels. In such a case, the consonant is geminated and the syllabic boundary falls in the middle of the geminated consonant: bəg/ga: [ˈbəg:a:].

Intervocalic consonants in Dogri, when preceded by a stressed vowel, are generally simultaneous members of two syllables and belong to the preceding as well as to the following vowel. Thus, consonant gemination is closely related to the problem of stress in Dogri.

Stress in Dogri is predictable in some conditions and unpredictable in others (see above). In disyllabic words, the consonant of the second syllable can be geminated if stress falls on the first syllable.

In trisyllabic words, the initial consonant of the second syllable may be geminated if the first syllable is stressed; the initial consonant of the third syllable may be geminated if the second syllable is stressed; but no consonant in the word can be geminated if the final syllable is stressed.

Thus the favourable environment for gemination of consonants is:

ˈ(C)V	ˈpɷll(ə)	'softness'
ˈ(C)V - V:	ˈbəggɑ:	'whitish'
ˈ(C)V - VC	ˈcəkkər(ə)	'circle'
ˈ(C)V: - VC	ˈgɑ:ɟɟər(ə)	'carrot'
ˈ(C)V - VCV:	ˈcəkkərɛ:	'circle' (oblique form)
(C)VˈCV -	cəˈtənn(ə)	'conscious'
(C)VˈCV - V:	kəˈṭə̀nnɑ:	'trousers'
(C)VˈCV: - V:	cəˈlɑ:kki:	'cleverness'

A consonant cannot be geminated in such an environment as:

(C)V ˈ- V:	bəˈgɑ:	'flow'
(C)V ˈ- V:C	bəˈɟɑ:r(ə)	'market'
(C)V ˈ- V:CV:	bəˈɟɑ:rɛ:	'market' (oblique form)
(C)VCV ˈ- V:	pəgəˈtɑ:	'pay off'

Words longer than trisyllabics are generally made up of two or more units and each unit bears stress according to the pattern to which it belongs.

3.4 Basic grade

Another problem is to decide as to which should be considered the basic or unmarked consonant grade in Dogri? Are (geminated) consonants shortened under certain conditions, or are (single) consonants lengthened under certain conditions? Generally, that form is considered unmarked which occurs in a wider environment, while the marked form occurs in a limited environment (Fischer-Jørgensen, 1975). This is difficult to decide for Dogri consonants since we find them in strong grade (phonetically) in initial position, in weak grade in final position (with a few exceptions in which the non-phonemic ə generally appears at the end), and in both strong and weak grades in medial position. Phonetically, [g] of ˈgəlɑ and ˈbəggɑ is strong, but [ɣ] of bɑ:g and ˈbəgɑ: is weak. When we examine the problem by taking stress into consideration, we find that single consonants can occur both after stressed vowels and unstressed vowels, while the geminated consonants can occur after stressed vowels only. According to this criterion, the single consonant can be considered basic, but it is notable that intervocalically, geminated consonants are far more numerous than single ones.

3.5 The phonetic manifestation of consonant gemination

The duration of single and geminated consonants of words in lists I and II has been measured on the basis of mingographic recordings for 5 informants (RN, SD, SL, CS, VG), who read each list five times. For details of the recordings and measurements, see 1.2 and 2.5. The difference of duration between geminated and single consonants is presented in table 8. Table 8 presents average durations of single and geminated consonants for the subjects RN, SL, SD, CS and VG. The table shows that geminated consonants are approximately twice as long as single consonants. The averages are based on the following words, spoken five times each:

Voiceless, unaspirated stops:

single: ˈcəpəl(ə), ˈsəti:, ˈpətɑ:, kəˈtɑ:, səˈtɑ:, ɟəˈtɑ:,
ˈkəcɑ:, ləˈkɑ:, sət(ə)

geminated: ˈcəppəl(ə), ˈpəttɑ:, ˈkəttɑ:, ˈɟɩttɑ:, ˈkəccɑ:,
ˈləkkɑ:, ˈsətt(ə)

Voiced, unaspirated stops:

single: dəˈbɑː, ˈɟɑːdɑː, ləˈdɑː, səˈdɑː, kəˈḍɑ̀ː, goˈḍɑː, ˈbəgɑː, bəˈgɑː, ɟəd(ə), gəɟ(ə), ɟəg(ə)

geminated: ˈdəbbɑː, ˈdɑːbbɑː, ˈsəddɑː, ˈɟɑːddiː, ˈkə́ḍḍɑː, ˈgoḍḍɑː, ˈgoːḍḍɑː, ˈbəggɑː, ˈbɑːggɛː, ɟədd(ə), gəɟɟ(ə), ɟəgg(ə)

Voiceless aspirated stops:

single: ˈtòːphɑː, ḍəˈphɑː, ˈkəthɑː, cəˈthɑː, bəˈṭhɑː, ˈpə̀khɑː, ˈtòkhɑː, dəˈkhɑː, səˈkhɑː, təˈkhɑ̀ː, tuːph(ə), cɔːth(ə), tokh(ə), dɑːkh(ə)

geminated: ˈḍəpphɑː, ˈtə̀pphɑː, ˈkətthɑː, ˈcɔːtthɑː, ˈməṭṭhɑː, ˈpə̀kkhɑː

Laterals:

single: ˈboːliː, ˈkoliː, ˈkɩlɑː, pol(ə)

geminated: ˈboːlliː, ˈkəːllɑː, ˈkoːllı̃, ˈkolliː, kɩllɑː, poll(ə)

The averages are thus general averages comprising words of different types. They have been combined because there did not seem to be any consistent differences between them. There are, however, some exceptions: RN's single final aspirated stops are shorter than his single medial aspirated stops. The former have an average duration of 7.7 cs, the latter 12.0 cs. The same is true of SL. His final aspirated stops have an average duration of 5.8 cs, and his medial aspirated stops of 8.6 cs. All subjects except CS have a somewhat longer duration of aspirated medial stops in stressed than in unstressed syllables. As all examples of the geminated stops are medial, and these words are stressed on the first syllable, it might therefore be more adequate to compare them with single medial stops following a stressed vowel only. This will not, however, make a great difference. The averages for single aspirated stops would be for RN: 11.1 cs (vs. 10.1), SL: 8.1 cs (vs. 7.8), SD: 8.0 cs (vs. 9.7), CS: 10.5 cs (vs. 10.4), and VG: 9.5 cs (vs. 10.8), and the general average will be 9.5 cs (vs. 9.9). SD, SL, and CS also have a shorter d after a stressed vowel, but, as there is only one word of this type, it changes the average only slightly. The averages given in the table can therefore be considered as quite representative of the relations be-

Table 8

Duration of single and geminated consonants in cs for five informants

	Unvoiced unaspirated		Voiced unasp.		Unvoiced aspirated		l	
	sing.	gem.	sing.	gem.	sing.	gem.	sing.	gem.
N	45 (SL 40)	35	55 (SL 45)	60	70	30	20	30
RN	7.9	14.6	5.4	9.2	10.8	18.1	4.8	11.6
% single /gem.	54		59		60		41	
SL	8.0	13.3	4.3	8.1	7.8	12.7	5.0	11.2
%	60		53		61		45	
SD	7.7	16.1	5.4	12.5	9.7	16.7	4.3	13.9
%	48		43		58		31	
CS	7.6	13.7	4.9	10.3	10.4	15.8	5.5	10.6
%	55		48		66		52	
VG	8.6	15.3	5.2	13.6	10.6	17.9	5.8	17.2
%	56		38		59		34	
grand mean	8.0	14.6	5.1	10.7	9.9	16.2	5.1	12.9
%	55		48		61		41	

tween single and geminated stops. A comparison with a number of the same words found in the continued text for RN and SD showed very similar values.

It appears from table 8 that voiceless stops are longer than voiced stops for all informants, and aspirated voiceless stops longer than unaspirated ones, except for SL (the aspiration has been included in the measurements. The closure of the aspirated stops is slightly shorter than that of the unaspirated stops). There is very little overlapping between the individual means for voiced and voiceless consonants, more for aspirated vs. unaspirated ones.

According to Bahri (1969, p. 83), geminated consonants appear after central vowels (ɩ̲ ɷ̲ ə̲) only, and on p. 92 he states that stops "are lenis in medial position after class I vowels (ɩ̲ ɷ̲ ə̲) unless geminated, while all the stops are comparatively tense when occurring under high or low tone or with peripheral vowels". This seems to indicate that he would only recognize one type of stop (relatively fortis) after peripheral vowels. But, as mentioned in 3.1, there are - according to my observations - word pairs of the type ˈɟaːdiː/ˈɟaːddiː, ˈsiːtaː/ˈsiːttaː, ˈkɛːdaː/ˈkɛːddaː. In order to deny the existence of such pairs, one might propose three different arguments:

(1) It might be maintained that the vowels before geminated consonants are always central. This is what Bahri says on p. 93 but, as it was shown in 2.3, there is a clear distinction between central and peripheral vowels in pairs like ˈbəggaː/ˈbaːggɛː, ˈkɩllaː/ˈkeːllaː, etc.

(2) One might also maintain that the type ɟaːdiː, i.e. peripheral vowel plus short lenis consonant, does not exist. In the present material there are only few words of this type, but there are some obvious examples with voiced consonants. The words baːg(ə), laːg(ə), and káːḍ(ə) have short consonants in the pronunciation of the five main informants (RN, SD, SL, CS, VG) (see table 9). A comparison with table 7 shows that the duration of the final consonant corresponds to the average duration for single voiced consonants, and they agree with the duration of e.g. d̲ in ɟəd(ə) (general average 5.2 cs) as compared to ɟədd(ə) (11.7 cs). There is one example of final t̲ (ɟaːt(ə)) (6.7 cs), which shows the same (cp. table 7 and the general average duration of sət(ə) (7.9 cs) as compared to sətt(ə) (14.9cs).

Table 9

Duration of signle and geminated consonants in cs after peripheral vowels.

	RN	SL	SD	VG	CS	grand mean
bɑ:g(ə)	3.5	4.4	4.0	5.0	3.4	4.1
lɑ:g(ə)	4.6	5.6	5.2	5.6	4.3	5.1
kɑ́:ḍ(ə)	4.8	2.0	4.0	4.4	4.2	3.9
ɟɑ:t(ə)	6.4	5.0	5.0	9.6	7.6	6.7
ɟɑ:dɑ:	4.2	3.6	4.0	4.2	3.8	4.0
ɟɑ:ddi:	12.2	9.8	12.2	13.6	11.3	11.8
lɑ:gɑ:	9.4	8.0	11.8	8.2	8.4	9.2
lɑ:ggɛ:	10.8	9.9	12.6	12.4	10.6	11.3

Medially the pair ˈɟɑ:dɑ: (4.0 cs)/ˈɟɑ:ddi: (11.8 cs) shows a very clear contrast in consonant length, which may be compared to gɒˈḍɑ: (4.7 cs)/ˈgɒḍḍɑ: (12.7 cs) of table 9. The pair ˈlɑ:gɑ:/ˈlɑ:ggɛ: is, however, somewhat different. The g of ˈlɑ:gɑ: (9.2 cs) is closer to that of ˈlɑ:ggɛ: (11.3 cs) and ˈləggɑ: (11.3 cs) than to that of ˈləgɑ: (5.6 cs). Only CS and VG have no overlapping between ˈlɑ:gɑ and ˈlɑ:ggɛ:, and SL and SD have considerable overlapping. For this word pair, thus, there is a tendency to confusion.

The material does not contain any examples of unaspirated voiceless stops after peripheral vowels (it was recorded before I had read Bahri's paper), but there are a few examples of aspirated stops after peripheral vowels, viz. ˈtò:phɑ:, ˈtò:khɑ:, and ˈcɔ:thɑ:. The latter two were pronounced with geminated stop by all informants. As for tò:phɑ:, two informants had clearly a short consonant: SL 7.2 cs (his general average of aspirated single consonants is 8.1 cs) and VG 10.4 cs (general average 10.6 cs), but SD (15.4 cs) and CS (14.0 cs) have durations corresponding to their geminated aspirates, and RN's duration is in between his single and geminated aspirates, but he has an example of ph after ə which is almost as long (14.4 cs) and which overlaps extensively with ˈtò:phɑ:, but which is distinguished from ˈḍəpphɑ: (16.6 cs).

On the whole, the duration of the aspirated stops varies much more than the other consonants, and in this case there really seems to be a tendency to confusion in medial position after peripheral vowels. However, final aspirates as in cɔ:th, dɑ:kh, and tù:ph have, for all informants, the same short duration as those of e.g. tɷkh, but the absolute measures are somewhat uncertain here because of the difficulty of delimitation from the following s of the frame sentence.

As for the consonant l, there is a clear distinction of short and long l after peripheral vowels, cf. table 8.

Thus, single consonants, also stop consonants, are distinguished from geminates after peripheral vowels, although there seems to be some confusion for medial aspirates and in the case of g in ˈlɑ:gɑ:.

(3) One might maintain that the so-called geminated consonants after peripheral vowels are not really geminates but only fortis consonants, and this is what Bahri seems to ascertain on p. 92. That would probably mean that they should not be much longer than the single consonants but articulated with more force. Table 10 shows consonant durations for eight triplets (five of them with stops) of the type ˈbəgɑ:/ˈbəggɑ:/ˈbɑ:ggɛ:. All subjects have been combined since there is no consistent difference between them.

Table 10

Duration of geminated consonants in cs after short (central) and long (peripheral) vowels compared to single consonants in triplets. 5 subjects (RN, SL, SD, VG, CS) combined (for the first two triplets 6 subjects (+ AS)).

dəˈbɑ:	4.9	ˈdəbbɑ:	11.4	ˈdɑ:bbɑ:	10.8
ˈbəgɑ:	4.4	ˈbəggɑ:	12.9	ˈbɑ:ggɛ:	11.2
ˈɟɑ:dɑ:	4.0	ˈɟədd(ə)	11.7	ˈɟɑ:ddi:	11.8
gɷˈḍɑ:	4.7	ˈgɷḍḍɑ:	12.7	ˈgo:ḍḍɑ:	10.7
cəˈthɑ:	11.8	ˈkəttha:	17.8	ˈcɔ:tthɑ:	15.6
ˈkɩlɑ:	5.3	ˈkɩllɑ:	14.5	ˈke:llɑ:	12.5
ˈkɷli:	5.9	ˈkɷlli:	15.4	ˈko:lli:	13.3
ˈbo:li:	5.0	ˈpɷll(ə)	11.8	ˈbo:lli:	11.5

Table 10 shows that the geminated consonants are indeed slightly shorter, on the average, after peripheral than after central vowels (except ˈɟɑ:dɑ:/ɟədd(ə)). Of 42 pairs of averages (5 informants x 8 pairs + 2 for AS), 32 show a higher average value after a central vowel. This may not be due to pure chance; consonants often tend to be slightly shorter after long vowels than after short vowels, but this is a perfectly well-known compensation phenomenon which does not justify ascribing them to two different phonetic categories. Moreover, there is extensive overlapping between single tokens. This is shown in fig. 7 for some of the triplets, all subjects combined; the other triplets look the same (there is slightly less overlapping in kɩllɑ:-ke:llɑ:). The overlapping is not due to the combination of subjects. When the consonants are compared for each subject separately, there is overlapping in 38 of 42 pairs (the exceptions are VG kɩllɑ:/ke:llɑ: and kɵlli:/ko:lli: and CS kɩllɑ:/ke:llɑ:).

The pairs dɵbbɑ:/do:bbɑ:, kɵbbi:/bo:bbi:, sɵttɑ:/so:ttɑ:, gɵttɑ:/go:ttɑ: and tɩllɑ:/te:llɑ: show similar relations. On the other hand, it appears clearly both from table 9 and fig. 7 that geminated consonants after peripheral and central vowels are clearly distinguished from single consonants, even for all subjects combined.

3.6 The phonetic manifestation of consonant weakening

As mentioned in 3.1, there is a tendency to weakening of single intervocalic and final consonants in Dogri. This weakening consists mainly in (1) a shortening of the closure, (2) sometimes a slight voicing of unvoiced stops, and (3) fricativization.

Aspirated unvoiced stops and unaspirated voiced stops are more prone to this weakening than unaspirated unvoiced stops, which are relatively resistant. Retroflex consonants also seem to be less prone to weakening, but retroflex ḍ may become a retroflex flapped [ṛ], and dental n may become a retroflex flapped [ṇ]. The palatal affricate ch may be weakened to [ʃ].

In the following, the analysis will be restricted to labial, dental and velar stops, and only part of the material has been included. For each subject two examples of each consonant were

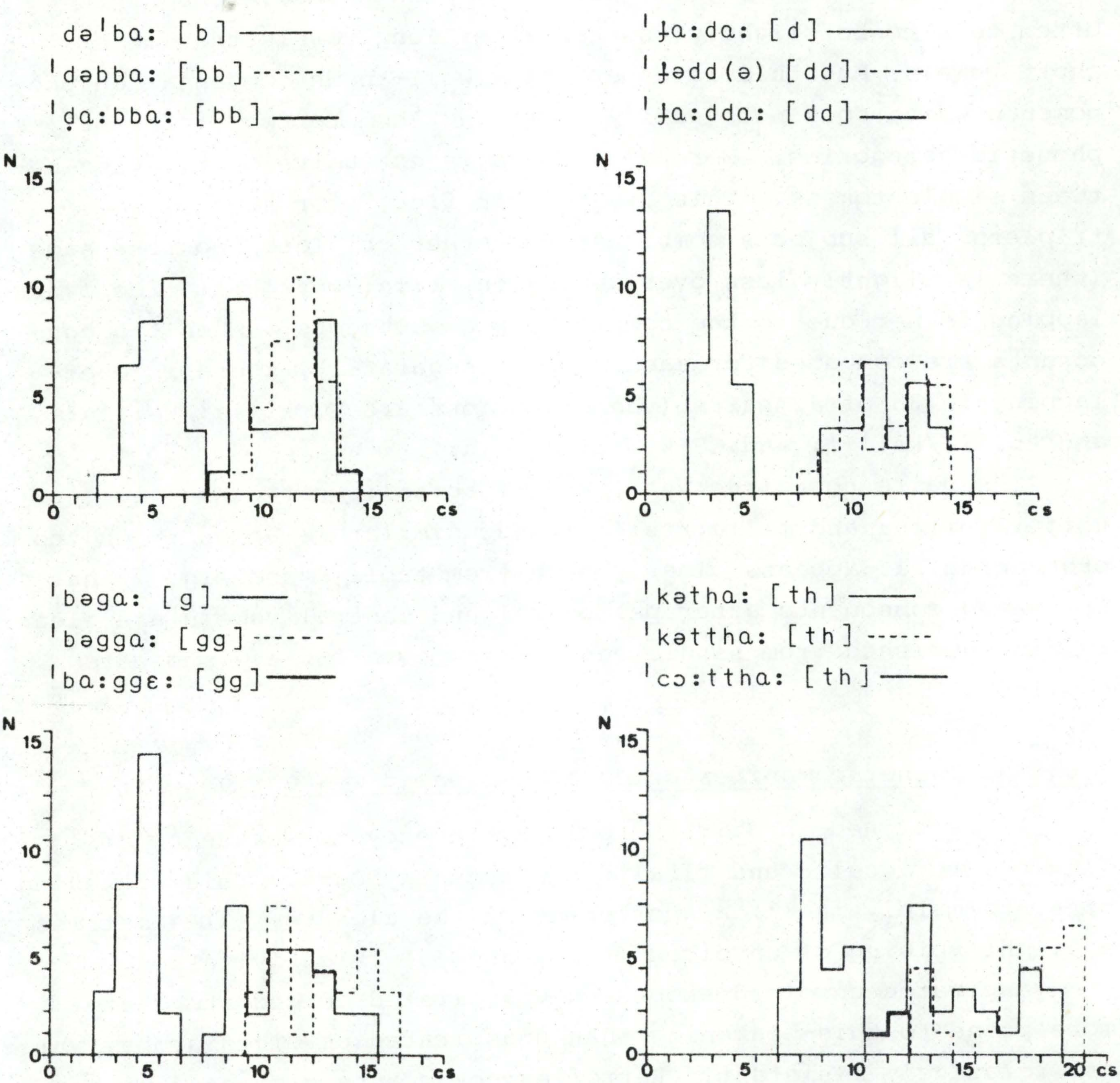


Figure 7

Duration of single consonants and geminated consonants after central and peripheral vowels. Six subjects combined (N = 23 - 28).

analysed in word initial stressed position and, as far as possible, one example for each of the following positions: word initial unstressed position, syllable initial stressed intervocalic position, syllable initial unstressed intervocalic position, and final position (in some cases no examples were available in the material). The results are given in table 11 for the informants RN, SL, SD and VG combined. The word phar was taken from a different list spoken in a somewhat slower tempo. For kh in stressed word initial position, only two informants had examples and only one example of each word. For the word tò:pha:, the duration measurements of SD and CS have been left out since they evidently pronounced it with geminated consonant. For voiced stops and unvoiced unaspirated stops only the total length, i.e. closure plus open interval, is given since the open interval is short and not influenced by the weakening. The individual averages of the open interval are 1.0 - 2.2 cs for the voiced stops (shortest for b, longest for g, but often it cannot be measured at all), and 1.0 - 3.2 cs for the unvoiced unaspirated stops (shortest for p, longest for k). In the case of the aspirated stops, both closure and aspiration are given, but in some cases they could not be distinguished so that the averages for total length may be based on more individual averages than those for closure and aspiration. In final position it was sometimes only possible to measure the closure for some informants, and only the total length for others (the number of individual averages included is in these cases indicated in parentheses, see e.g. co:th in table 11).

3.6.1 Shortening

It appears from table 11 that consonants are shortened in intervocalic and final positions. This is true of all subjects, but VG has relatively little shortening in unvoiced unaspirated stops. For unvoiced unaspirated stops and voiced stops it is the closure which is shortened as long as they remain stops, whereas the short open interval remains unchanged. This corresponds to what is found e.g. in many Germanic languages. But it is a characteristic feature of Dogri that the same is the case for aspirated consonants: the closure is shortened, but the aspiration remains almost the same. In Germanic languages, on the contrary, the aspiration is normally very much reduced in unstressed intervocalic

Table 11

Examples of shortening and fricativization of stops in medial and final position. Averages for five informants (RN, SL, SD CS, and VG) combined. The ciphers under fricativization indicate the number of cases (listening results for fricativization in parentheses). (N generally = 25.)

position	cons.	word	cs	fricativiz. –	?	+	cons.	word	cs	fricativiz. –	?	+
stress. word-in.	p	ˈpəttɑ: ˈpətɑ:	14.2 13.3	25 25			b	ˈbəġgɑ: ˈbəgɑ:	12.4 10.7	25 25		
unstr. word-in.	p	pəˈlɑ:	12.0	23			b	bəˈgɑ:	10.6	25		
stress. syll-in.	p	–					b	dəˈbɑ:	5.5	9 (8	2 1	14 16)
unstr. syll.-in.	p	ˈcəpəl(ə)	7.8	19 (23	5 2	1)	b	ˈdebək(ə)	4.7			25 (25)
stress. word-in.	t	ˈtə̀pphɑ: ˈtò:phɑ:	13.0 13.4	25 25			d	ˈdəbbɑ: ˈdɑ:kh(ə)	12.2 10.7	25 25		
unstr. word-in.	t	təˈkhɑ̀:	11.5	25			d	dəˈkhɑ:	11.0	25		
stress. syll.-in.	t	səˈtɑ:	8.2	23 (23	2	)	d	ləˈdɑ:	5.9	18 (23	5 1	2 1)
unstr. syll.-in.	t	ˈsəti:	9.2	22	3		d	ˈɟɑ:dɑ:	4.0	2 (13	7 1	16 11)
final	t	ɟɑ:t(ə)	7.3	16 (24	7 1	2)	d	ɟəd(ə)	5.1	8 (13	3 9	14 3)
stress. word-in.	k	ˈkəccɑ: ˈkəcɑ:	13.8 13.6	25 25			g	ˈgəɟɟe ˈgəlɑ:	12.0 10.8	25 23	 2	
unstr. word-in.	k	kəˈrɑ:	11.6	25			g	gəˈlɑ:	10.9	25		
stress. syll.-in.	k	ləˈkɑ:	8.6	25			g	bəˈgɑ:	5.4	12 (11	4 2	9 12)
unstr. syll.-in.	k	–					g	ˈbəgɑ:	4.5	1 (1	3 1	21 23)
final	k	ˈdebək(ə)	7.0	17	3	5	g	bɑ:g(ə)	4.0	7 (7	7 12	11 6)

Table 11 (continued)

posi- tion	cons.	word	cs clos.	cs open interv.	cs total	fricativization -	 ?	 +
stress. word-in.	ph	phɑ:r	13.4	7.5	(20.3)	19		
stress. syll.-in.	ph	ḍəˈphɑ:	6.6	5.2	11.4	14 (9	6 4	5 12)
unstr. syll.-in.	ph	ˈtò:phɑ:	5.7	7.0	12.8	16 (12	1 1	8 12)
final	ph	tù:ph(ə)	4.4 (2)		10.3 (3)	8 (5	7	10 20)
stress. word-in.	th	thò:	10.7	7.6	18.3	23		
stress. syll.-in.	th	cəˈthɑ:	4.9	6.9	11.8	22 (22	3 2	 1)
unstr. syll.-in.	th	ˈkəthɑ:	4.0	5.9	9.9	17 (19	5 2	3 4)
final	th	co:th(ə)	6.3 (3)	 (2)	8.8 (2)	22 (21	2 2	1 2)
stress. word-in.	kh	kəll(ə) ˈkhəbər	10.0 9.0	6.5 8.8	16.5 17.0	2 2		
unstr. word-in.	kh	khəˈbɑ:r	8.0	6.8	14.7	22	3	
stress. syll.-in	kh	dəˈkhɑ:	4.4	7.4	11.8	11 (11	3 1	11 13)
unstr. syll.-in.	kh	ˈpə̀khɑ:			8.8	1 (2	3 1	21 22)
final	kh	tə̀kh(ə)	9.9 (3)		10.6 (2)	12 (12	5 6	8 7)

position, so that they become practically unaspirated.

There is also a tendency to shorten consonants in unstressed position. In eight out of eleven comparable pairs, four or five subjects have shorter consonant in unstressed position, but the difference is smaller than that due to position.

3.6.2 Voicing

In medial and final position after a vowel the first 2-4 cs of a voiceless consonant are normally weakly voiced. That is, however, not specific for Dogri. It is rare that the whole consonant gets voiced, but it happens in many examples of dəbək (k̲) and in cəpel (p̲), particularly in the pronunciations of SD, CS and SL. It also happens in some cases of tù:ph(ə), tò:phɑ:, and ɟɑ:t(ə).

3.6.3 Fricativization

3.6.3.1 Acoustic investigation

The most characteristic feature of Dogri consonants is the tendency to fricativization in medial and final position, particularly of aspirated and voiced stops. The analysis of this phenomenon was concentrated on acoustic curves since, for all informants except VG, only tape recordings were available. Acoustically, fricativization should turn up as a weakening of the explosion of the stops and as fricative noise or voiced higher components during the closure of voiceless and voiced stops, respectively. As it is difficult to get a reliable measure of the intensity of the explosion which is very weak anyhow in voiced stops, particularly in b̲, the investigation was concentrated on the appearance of high noise or higher voiced components during the closure. For this analysis the mingographic recordings of all informants were used. They contained an intensity curve with high-pass filtering at 2000 Hz (with a steepness of 18 db/octave) added especially for this purpose. High components were expected to show up in this curve, and low frequency noise in the intensity curve high-pass filtered at 500 Hz. The analysis of the curves, however, turned out to be difficult. In the first place, the weakening of stops is a gradual phenomenon and one may find all intermediate steps between firm closure and full fricativization, not only in the time dimension (treated in the preceding paragraph as shortening of the closure), but also in the intensity of the noise. Thus, there will be cases of doubt, and no real quantization is possible. But the fricativization showed up very clearly in VG's curves. Fig. 8 contains some clear examples: there is noise during the whole medial consonant in /ˈpə̀khɑ:/[ˈpə̀xɑ:] (fig. 8a) but not in /ˈpə̀kkhɑ:/[ˈpə̀k:hɑ:] (fig. 8b), and there are high frequency components and high intensity in the /g/[ɣ] of /bəˈgɑ:/ but not in the /b/[b] (fig. 8c).

However, for the other informants there was noise or, more often, a mixture of noise and voiced consonants where it was not expected. To clear up this problem, a number of spectrograms were taken, and after a thorough comparison of mingograms and spectrograms, Eli Fischer-Jørgensen came to the conclusion that in most cases of noise during the closure in the intensity curve filtered at 2000 Hz, this apparent noise had nothing to do with the consonant but was due to a continuation of the vowel formants in the beginning of the consonant, caused by some phenomenon of resonance during the recording. This conclusion was based on the following facts:

(1) The noisy components were not only found in the consonants expected to be weakened but also in unvoiced stops and in voiced and voiceless geminates.

(2) The noise started immediately after the vowel and decreased in intensity during the following 5-10 cs.

(3) Its intensity depended on the relative intensity of the preceding vowel, and it was particularly strong for the subjects SD and CS, who have read the text in a relatively loud voice.

(4) The spectrograms showed that it was mostly found at the level of the (strong) first formant of the vowel.

(5) There was nothing to be seen at the frequencies of expected consonant noise, and it cannot be taken for preaspiration either, since aspiration noise is always weak at the frequency of the first formant (but strong in this case) and stronger at the frequency of higher formants. For an example, see fig. 9a ɟɑ:ttɑ: spoken by SD, and the corresponding spectrogram fig. 10a.

This type of noise should therefore be disregarded in the interpretation of the curves, and the preliminary results of the acoustic analysis were revised by Eli Fischer-Jørgensen. All noise following immediately after the vowel and showing decreasing in-

Figure 8

Mingograms of the words (a) ˈpə̀khɑː, (b) ˈpə̀kkhɑː, and (c) bəˈgɑː spoken by VG. 1: duplex oscillogram, 2: intensity curve HP filtered at 500 Hz, 3: intensity curve HP filtered at 2000 Hz, 4: intensity curve full frequency range, and 5: fundamental frequency curve.

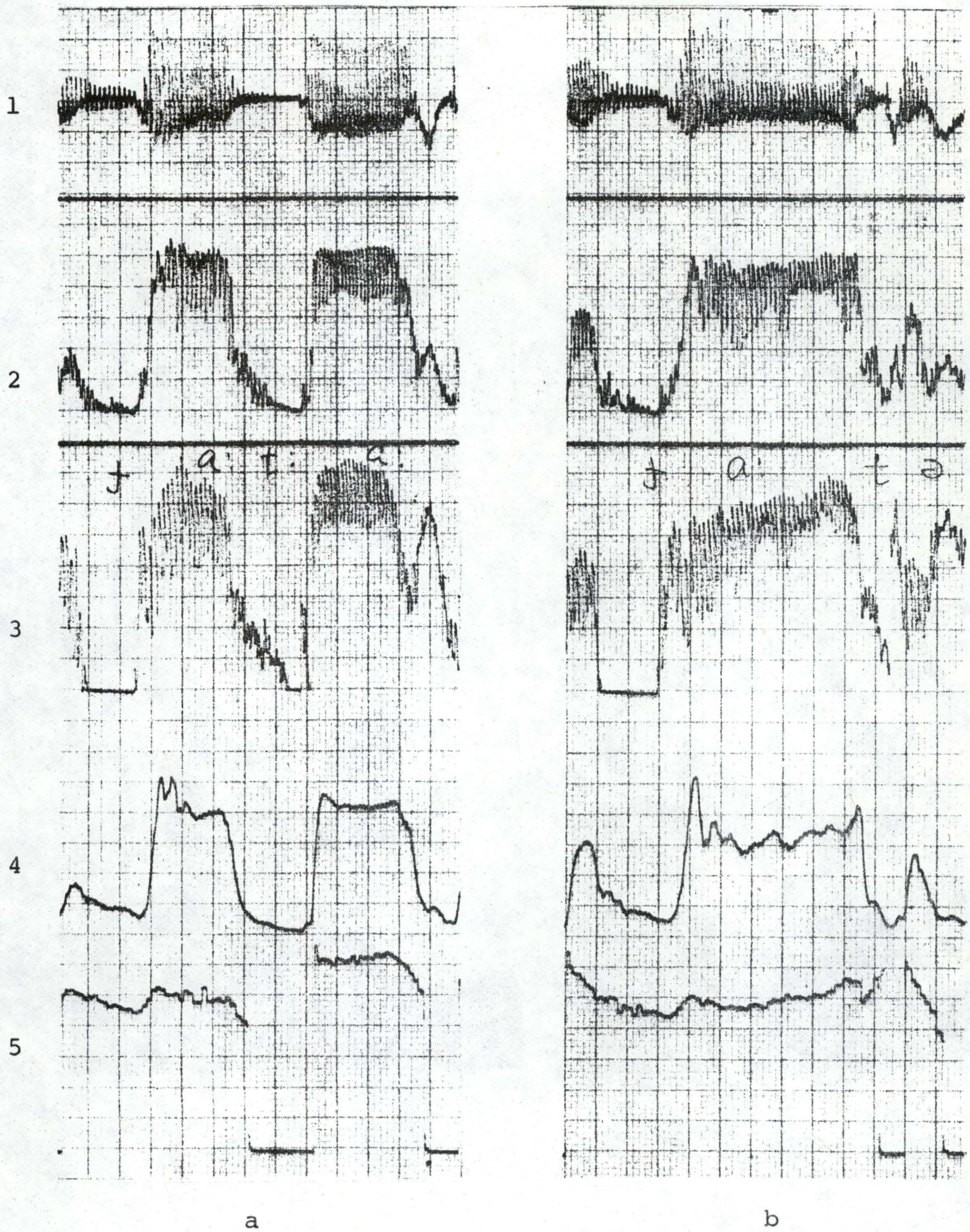
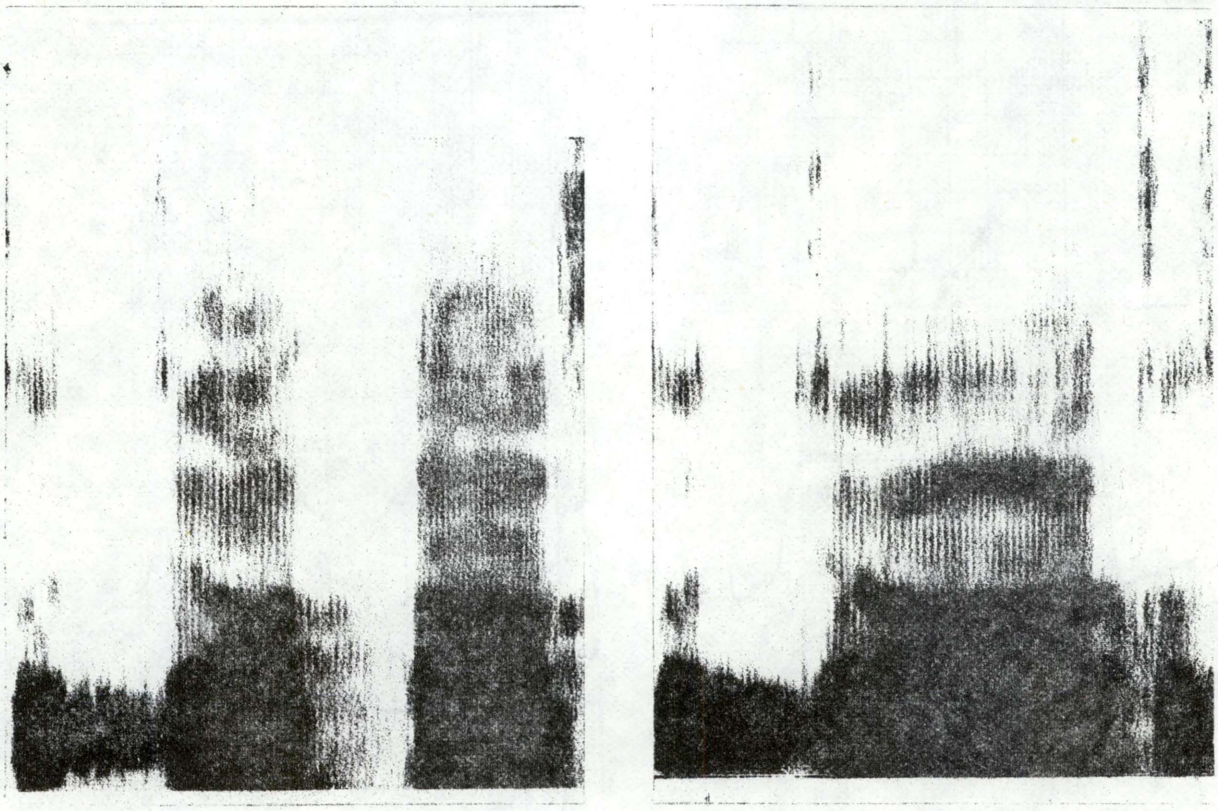


Figure 9

Mingograms of (a) ˈɟɑ:ttɑ: and (b) ɟɑ:t(ə) spoken by SD. 1: duplex oscillogram, 2: intensity curve HP filtered at 500 Hz, 3: intensity curve HP filtered at 2000 Hz, 4: intensity curve full frequency range, and 5: fundamental frequency curve.



ɟɑ:ttɑ:

a

ɟɑ:tə

b

Figure 10

Spectrograms of (a) ˈɟɑ:ttɑ: and (b) ɟɑ:t(ə) spoken by SD.

tensity, found in the intensity curve filtered at 500 Hz, which contained most of the resonance of the first formant, was disregarded.

With low formant 1 the resonance turns up in the curve high-pass filtered at 500 Hz only, but after ɑ: and ə (with relatively high first formant) and particularly after ɑ: (with a strong second formant around 1100 Hz), it is also found in the curve high-pass filtered at 2000 Hz, and it is difficult to distinguish between real consonant noise and the noise due to resonance, particularly in the cases of short closure.

The words utilized for the analysis of shortening were also utilized for the analysis of fricativization. In table 10, in which the durations are indicated, there is a column for fricativization. "-" means "no indication of fricativization", "+" means "clear signs of fricativization", whereas "?" means "problematic", i.e. some very weak noise is seen in the high-pass filtered curve, or there is strong noise, particularly in the start, and at the same time a clear explosion showing that it is not so probable that the consonant is weakened or, in the case of voiced stops, it may simply be an intermediate case.

A number of cases have been controlled by means of spectrograms, but it was not possible to make spectrograms of all examples. There may therefore be some cases of noise all through in the mingograms which are marked "+" but should have had a "?" or a "-" and, inversely, cases which have got a "-" instead of a "?" or "+", because the noise was so weak that it did not show up in the curves, or it was mixed up with the noise due to resonance. Therefore, the results of a listening test with one Danish phonetician, who listened to each word several times, has been added in parentheses. Here, "-" means "definitely stop", "+" means "definitely fricative", and "?" means "doubt".

It appears from table 11 that the agreement between the acoustic analysis and the listening results is sometimes very good, sometimes relatively poor. In the latter cases it is not always easy to say whether the curves or the listening is more reliable.

For word initial stops no weakening was found in the curves. There is only one word (khəˈbɑːr) which seemed to have some noise in the start of the initial kh in a few cases (by mistake, this word and the word ˈsətiː were not listened to). Moreover, the g in ˈgəlɑ: looked slightly fricativized in a few cases (it stood, exceptionally, after an l of the frame), but it was heard as a stop. On the whole, all initial stops were heard as stops, and no listening results are added for these consonants in table 10.

Medial voiceless unaspirated consonants are very rarely weakened, but the final stop in ɟɑ:t(ə) showed noise in the mingograms in various cases. It was, however, heard as a stop. The listening results turned out to be correct in a number of cases which were controlled by means of spectrograms showing that the noise was due to a continuation of the first and second formants (see fig. 9b and the corresponding spectrogram fig. 10b). On the other hand, the strong noise of the s of the frame (where there is no intervening ə), following relatively weak noise, may contribute to hearing a stop.

The aspirated consonants are often fricativized in medial position, particularly ph and kh. As for the labial consonant ph, a fricative (generally a bilabial [ɸ]) was heard in quite a number of cases, corresponding to dubious intensity curves or even to an apparent closure. This was particularly the case for SL. This may be due to the fact that the noise of a bilabial fricative may be rather weak, so that it does not show up in the curves, not even in the spectrograms. Where it does show up (see tò:phɑ, fig. 11a), it can be distinguished from aspiration by its higher frequency; moreover, the fact that it is strongest in the middle of the consonant distinguishes it both from the resonance phenomenon, which is strongest in the beginning of the consonant, and from aspiration, which shows up at the end.

The dental aspirate does not tend much to fricativization, and there is very little disagreement with the listening results for this consonant. There are several cases with strong but decreasing noise going all through the closure in the intensity curve high-pass filtered at 2000 Hz, but showing a clear release. They are heard as stops, and spectrograms show that the decreasing noise is due to a continuation of the vowel formants, cp. e.g. CS's cəˈthɑ: (fig. 11b and the corresponding spectrogram fig. 12b) versus SD's ˈkəthɑ: (figs. 11c and 12c), which has some real high frequency noise and is heard approximately as [ˈkəθtθɑ:].

The velar stop kh is the one that is most often fricativized, and for this consonant there is good agreement between listening

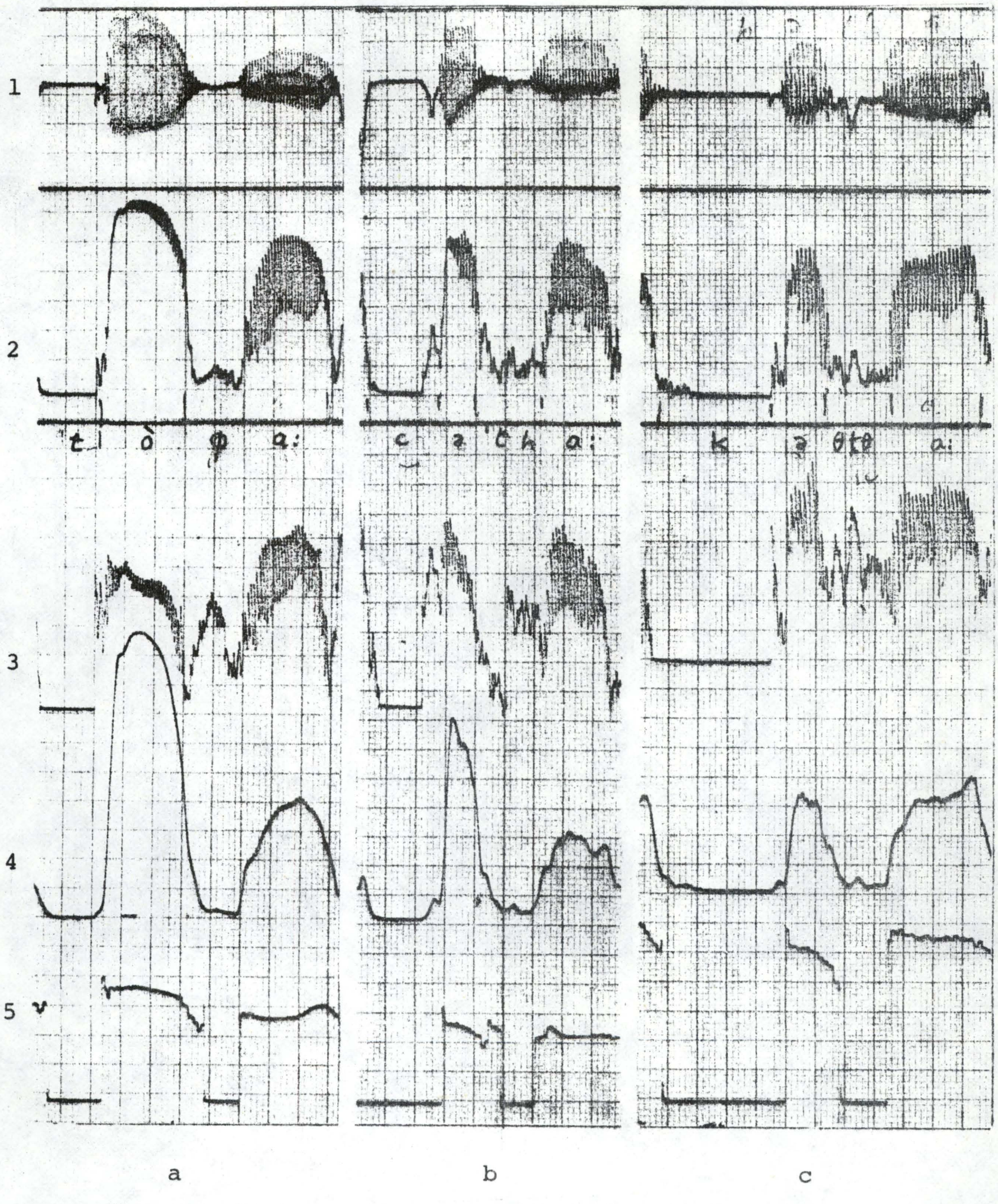
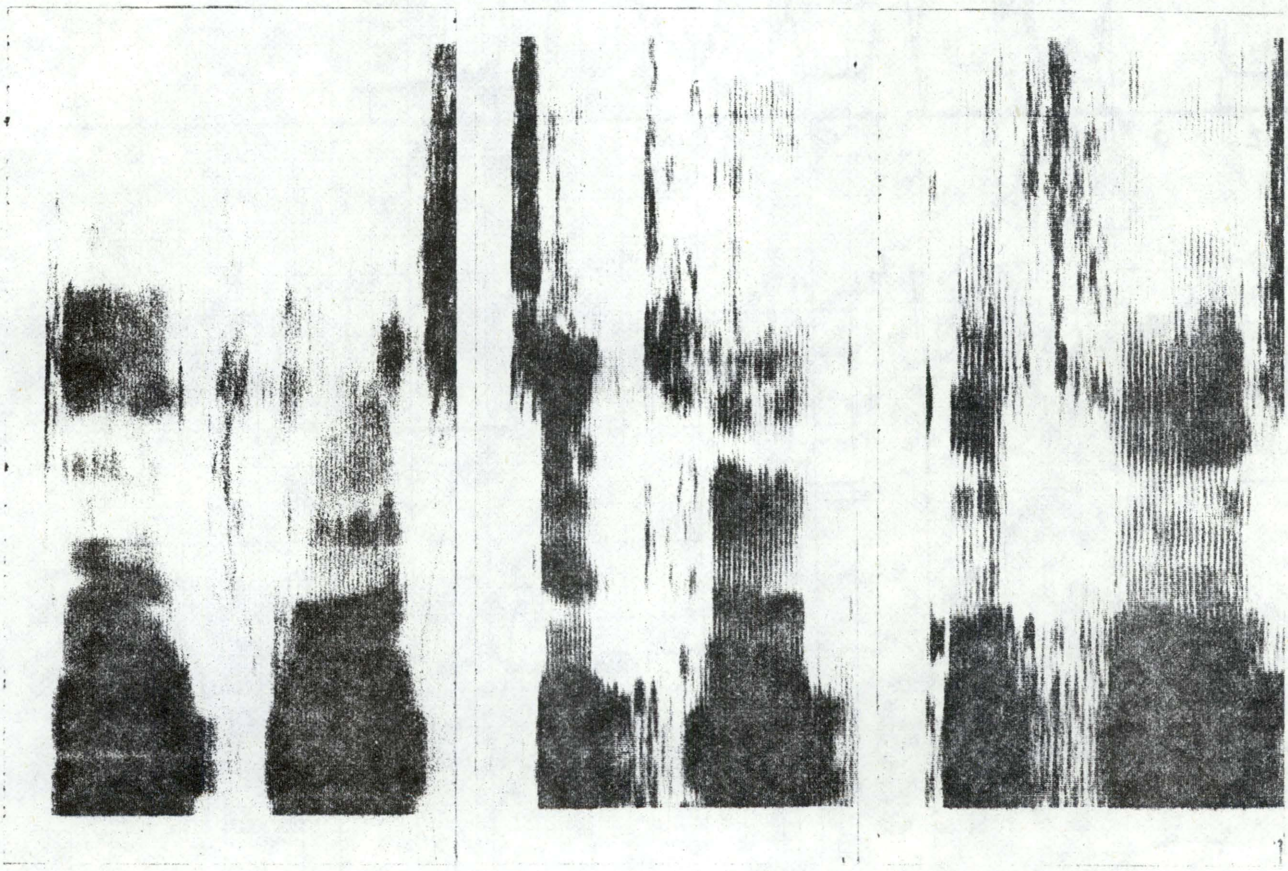


Figure 11

Mingograms of (a) ˈtò:pha: spoken by VG, (b) cəˈtha: spoken by CS, and (c) ˈkətha: spoken by SD. 1: duplex oscillogram, 2: intensity curve HP filtered at 500 Hz, 3: intensity curve HP filtered at 2000 Hz, 4: intensity curve full frequency range, and 5: fundamental frequency curve.



ˈtɔ̀:phɑ: cəˈthɑ: ˈkətha:

a b c

Figure 12

Spectrograms of (a) ˈtɔ̀:phɑ: spoken by VG, (b) cəˈthɑ: spoken by CS, and (c) ˈkəthɑ: spoken by SD.

results and acoustic analysis. The frication noise is generally rather strong. In some cases where a stop is listed, it was affricated rather than aspirated.

Voiced medial and final stops are often fricativized. The agreement between acoustic analysis and listening is very good for b and g (it was, however, very difficult to come to a decision about the final g in ba:g(ə)). In the case of ˈɟa:da: and ɟəd(ə) there is an obvious discrepancy, more stops being heard than expected on the basis of the curves (as on the whole in the case of dental consonants), but the stops heard seemed very short. In ˈɟa:da:, like in ɟa:t(ə), the consonant is preceded by a very sonorous vowel, and in some cases the transition from the consonant to the following vowel was very abrupt and the pitch curve had a dip like in stops. Thus, there were probably more stops than shown by the ciphers for the acoustic analysis, but a quick movement of the tip of the tongue may, on the other hand, give the impression of a stop even if there is not full closure.

In most cases there is more fricativization (and more shortening) in unstressed than in stressed syllables, both according to the acoustic analysis and to the listening results, cp. the medial consonants of ˈbəga: vs. bəˈga:, ˈɟa:da: vs. ləˈda:, ˈdebək vs. dəˈba:, (ˈkətha: vs. cəˈtha:), ˈpəkha: vs. dəˈkha:, but in tɔ̀:pha: vs. ḍəˈpha: it is rather the other way round.

The weakening of single consonants might be seen as a means to avoid confusion with the geminated consonants. However, as seen in table 7, the difference in duration is smallest for aspirated stops, and just in this case there is also less difference in fricativization: ˈtɔ̀:pha: is weakened both by those who pronounce the ph as a single and as a geminated consonant, and SD has complete fricativization of geminated pph in ˈḍəppha:, ˈḍəpphə, and ˈtə̀ppha: and an obvious affrication of kkh in ˈpəkkha:, and SL has a rather weak pph in some examples of ˈḍəppha: and ḍəpphə.

3.6.3.2 Prospects of a physiological analysis

It would have been of interest to undertake a physiological analysis of the weakening phenomenon. This was only possible in the case of VG, and such an analysis was planned but postponed till after the acoustic analysis, and then it was too late. The only investigations undertaken were a fiberoptic analysis of p, ph and b, and a registration of intra-oral pressure of a word list with labials and dentals spoken in the frame ku:sɛ: ... gəlɑ:yɑ:. A preliminary analysis of the fiberoptic investigation, undertaken by Birgit Hutters, showed that the glottis is wide open in ph, apparently still more open than in the strongly aspirated Danish stops, although the aspiration is not quite as long as in Danish (for VG around 6-8 cs in the present material). In the unaspirated p the glottis is much narrower, but somewhat more open than in Danish voiceless b. The wide opening of the aspirated consonant seems to be preserved in the positions of weakening d̥əˈphɑ:, which is in good agreement with the preservation of aspiration found in weak position for all informants.

A preliminary analysis of the intraoral pressure curves undertaken by Eli Fischer-Jørgensen showed that in initial position aspirated consonants had slightly shorter closure but a somewhat higher intraoral pressure than unaspirated consonants (which may be due to the wide open glottis). The voiced consonants b and d have a much slower rise of the curve, but it may attain the same level as in the voiceless unaspirated stops at the very end. Only labials were used in a list of words with stops in different positions. Neither stress nor position seemed to influence the intra-oral pressure of p, and stress did not influence the intra-oral pressure of initial ph or b either. Nor was there any consistent difference between single and geminated stops. But medial and final weakened ph (ˈtò:phɑ:, tù:ph(ə)) and b (ˈdəbɑ: and, partly, gɛ:b(ə)) had a lower intra-oral pressure than initial ph and b, and since the glottis opening was found to be the same in ˈphɑ:r̥ and d̥əˈphɑ: (and must be practically the same in initial and medial b), this points to a larger opening at the lips. The different degree of weakening also influenced the pressure in ˈdəbɑ vs. dəˈbɑ and ˈɟɑ:dɑ: vs. səˈdɑ:. The relations between stress and duration and between stress and pitch movement mentioned above were fully confirmed in this material.

A palatographic analysis might throw some light on the weakening of the closure of lingual consonants, registration of lip pressure might give some information for the labials, and air-flow curves might give good information about the degree of opening. But this must be done some time in the future.

References

Allen, W. S. 1953: Phonetics in Ancient India, London

Bahri, U. S. 1969: Phonology of Dogri. Pakha Sanjam I, Panjabi University, Patiala

Bloch, Jules 1965: Indo Aryan, English translation by Alfred Master, Paris

Burrow, T. 1937: The language of the Kharosthi documents from Chinese Turkestån, Cambridge

Chomsky, N. and M. Halle 1968: The Sound Pattern of English, New York

Delattre, P. 1968: "From acoustic cues to distinctive features", Phonetica 18, p. 198-230

Fant, G. 1959: Acoustic Analysis and Synthesis of Speech with Application to Swedish, Stockholm

Fischer-Jørgensen, Eli 1975: Trends in Phonological Theory, A historical Introduction, Copenhagen

Gauri Shankar 1931: "A short account of Dogri", Indian Linguistics, p. 1-83

Ghai, Ved Kumari 1968: "Word tones in Dogri", ARIPUC 2, p. 133-204

Ghai, Ved Kumari 1970: "Problems of stress and vowel quantity in Dogri", Proc. 1st All India Conference of Linguists, Poona, p. 158-160

Ghai, Ved Kumari 1979: "Nasal sounds in Dogri", Proc.Phon. 9, vol. I, p. 329 (summary)

Gill, H. S. and H. A. Gleason 1974: "Salient features of the Panjabi language", Pakha Sanjam, Patiala, p. 1-50

Gupta, Bansilal 1965: Dogri Bhaṣa aur Vyakaraṇa, Jammu

Jakobson, R., G. Fant and M. Halle 1965: "Preliminaries to speech analysis" (1952) 2nd edition, Cambridge

Jørgensen, Hans-Peter 1969: "Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantenfrequenzen", Phonetica 19, p. 217-245

Ladefoged, P. 1975: <u>A course in phonetics</u>, New York

Lehiste, Ilse 1970: <u>Suprasegmentals</u>, Cambridge, Mass.

Lehiste, Ilse, K. Morton and M.A.A. Tatham 1973: "Consonant gemination", <u>JPh</u> 1, p. 131-148

Lindau, Mona 1977: "Vowel features", <u>UCLA WPP</u>, p. 49-81

Ohala, Manjari 1979: "The phonological features of Hindi stops", <u>Proc.Phon.</u> 9 Vol. I, p. 309 (summary)

Pischel, R. 1965: <u>Comparative grammar of the Prakrit languages</u>, English translation by Subhadra Jha, Delhi

Sharma, D. D. 1971: "Syllabic structure of Hindi and Panjabi", Panjab University, Chandigarh

Verma Siddheshwar "Aspiration in North West Sub-Himalayan Indo-Aryan dialects", <u>Indian Linguistics</u> 26, Poona

(ARIPUC 14, 1980)

A PILOT INVESTIGATION OF THE F_0 PATTERN IN AMERICAN ENGLISH

Niels Dyhr

Abstract: Fourteen declarative sentences were recorded by one native speaker of American English. The F_0 pattern (the F_0 movement within the stress group) was investigated, and a model was proposed. This model was compared to one based on identical material, but with a British speaker.

1. Introduction

In British English and "East Coast American English", the stressed syllables have a higher fundamental frequency than the unstressed ones (Fry, 1958, Lieberman, 1960). Bolinger (1970) mentions, however, that in other types of American English a pattern may be found in which the fundamental frequency is lower in the stressed syllables than in the unstressed ones. The purpose of the pilot experiment reported below was to examine instrumentally this opposite F_0 pattern.

2. Procedure

The test material consisted of 14 declarative sentences, each consisting of a test word (either a natural word or a nonsense - but possible - word) embedded in a carrier phrase. To avoid the influence of intrinsic F_0 differences among vowels, the material was constructed in such a way that each test sentence contained either low or high vowels throughout.

The test material contained the following stress combinations: Stress on the first, second, third, and fourth syllable in the test word, followed by zero, one, and two unstressed syllables.

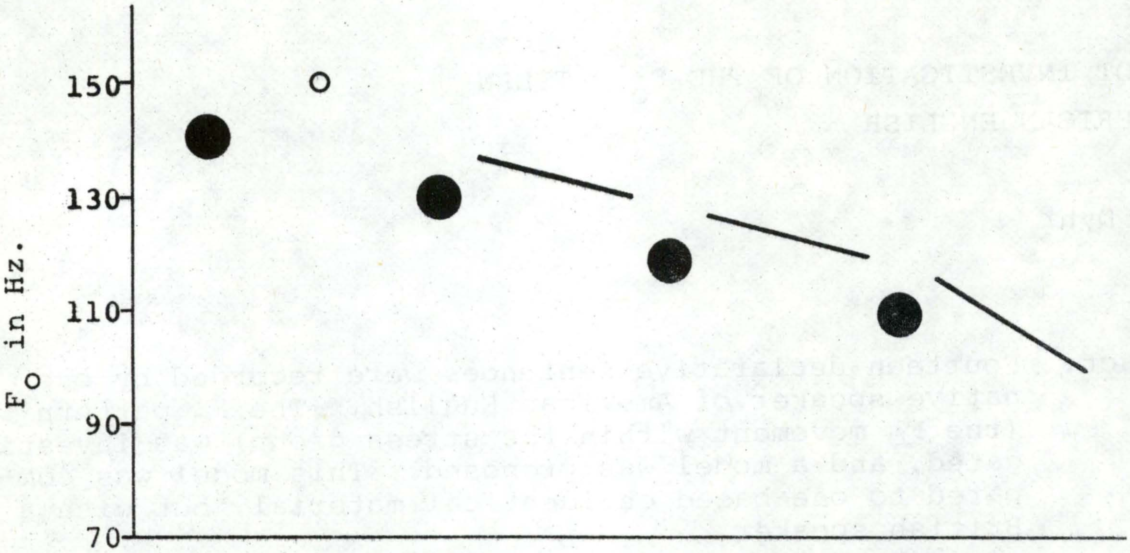


Figure 1

American model based on sentences with natural words. ● indicates stressed syllable, O unstressed syllalbe, —— a series of unstressed syllables.

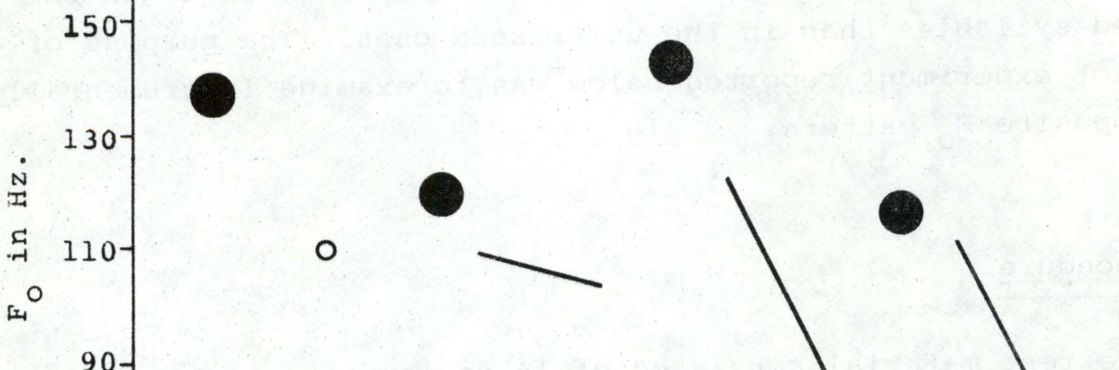


Figure 2

British model based on material identical with that of the American model. See further legend of figure 1.

The material was recorded ten times by a twenty-two year old American male speaker, born and raised in California.

Tracings of the F_o movements were made, and superimposed on one another. Average F_o curves were drawn by hand. A quantitative measure for the central tendency of the F_o level in the vowels was obtained by measuring the average F_o curves at a point two thirds from the vowel start, cf. Rossi (1971).

3. Results

It was found that in the natural words all the stressed syllables had a lower fundamental frequency than the unstressed ones. For the nonsense words the pattern was the opposite. This seemed to be due to the fact that the speaker, who had had no phonetic training, was unable to pronounce the nonsense words without emphasizing them. This change of pattern did not occur at the word boundary, but at the stress group[1] boundary (except when these coincide) - i.e., the F_o pattern did not change before the stressed syllable in the nonsense word, even when the stressed syllable was the last one in that word. This seems to indicate that the F_o pattern is controlled, not by the word, but by the stress group..

The results for the natural words can be described by the model shown in figure 1.

The model based on the data from the American speaker can be compared to one based on an identical test material, but with a British speaker, figure 2. It is evident that, apart from the opposite F_o pattern, a definite sentence nucleus (focus) is found only in the British model, which also has a greater F_o variation than the American model.

It must be pointed out that, considering the very restricted material, the present result should be viewed with care. A preliminary examination of data obtained for another American (Wisconsin) speaker, however, seems to be in good agreement with the tendencies outlined in the present paper.

---

1) The stress group is defined as a stressed syllable plus the following unstressed ones.

References

Bolinger, D.L. 1970: "Relative height", in: P. Léon, G. Faure, and A. Rigault (eds.) Prosodic Feature Analysis, Paris, p. 109-125

Fry, D.B. 1958: "Experiments in the perception of stress", LS 1, p. 126-151

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

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

(ARIPUC 14, 1980)

AN ACOUSTIC INVESTIGATION OF INTRINSIC VOWEL DURATION IN DANISH

Michael Bundgaard

Abstract: The duration of twelve short and long stressed vowels in disyllabic nonsense words, embedded in a carrier sentence, recorded 10 times by each of five speakers, was measured. The vowels can be grouped (roughly) in five tongue height categories, and duration was found to increase significantly from category 1 (highest) through 5 (lowest), in about 1 cs steps. Unstressed vowels show the same tendency but the increase in duration with lower tongue height is considerably smaller. Variation in speaking rate (even rather considerable) did not significantly influence the relationship between high and low vowels. In contradistinction to earlier investigations, it seems that the difference in duration between long and short vowels is constant over different tongue heights, i.e. $V_{short} = V_{long} - b$, where b is a constant (approximately 5 cs).

1. Introduction

Intrinsic vowel duration is the object of several investigations in several languages: Danish - Fischer-Jørgensen (1955 and 1964), German - Maack (1949), English - Peterson and Lehiste (1960), and Thai - Abramson (1962). All agree that there is a universal tendency for vowels with high tongue position to be of shorter duration than vowels with lower tongue position, everything else being equal.

Various explanations for this correlation between tongue height and vowel duration have been offered. Fischer-Jørgensen (1964) prefers the hypothesis, advanced by Jespersen (1926), that low vowels require more time for the speech organs to reach their target positions than do high vowels. Lindblom (1967) adheres to the same theory, and in his model of lip and jaw co-ordination vowel duration increases as a function of increased jaw opening. The jaw is described as a damped spring-mass system, where the differences in duration are due to inertia in the system, but on a higher level of production control, different vowel qualities are presumably normally programmed with equal durations.

The purpose of the present investigation is to examine the relationship between tongue height and vowel duration in Danish vowels in stressed (and unstressed) syllables, and to investigate the relation between long and short vowels. Danish vowels have been accounted for in these respects previously by Fischer-Jørgensen (1955, 1964) and by Holtse (1977), and my experiments are intended as a supplement to their investigations: The speakers in this study are all fairly young, and they speak Advanced Standard Copenhagen Danish (ASC - see Basbøll 1968). Fischer-Jørgensen's subjects all belong to an older generation, and both her and Holtse's subjects represent different dialects. Holtse's study examines only 4 vowel qualities, Fischer-Jørgensen's includes 10, whereas 12 vowel qualities are involved in this material.

2. Experimental procedure

2.1 Material

The material consists of 12 long and 12 short vowels:

[iː e̞ː ɛ̞ː æː ɑː yː øː œ̞ː uː o̞ː ɔ̞+ː ɒ̞ː]
[i e̞ ɛ̞ a ɑ y ø œ̞ u o̞ ɔ̞+ ʌ̝+][1]

According to Thorsen and Thorsen (1978), these vowels can be placed in the Cardinal Vowel Diagram as seen in fig. 1.

Long and short vowels in ASC Danish have not been subjected to acoustic analysis, but the quality differences between long and short is generally considered to be very small in ten of the twelve pairs above. The difference is greater between [æː] and [a], and between [ɒː] and [ʌ], but with regard to tongue height I shall assume that they are close enough to be treated as pairs. The listing and pairing is made on purely phonetic grounds. Thus, [æː, ɑː] and [a, ɑ] are variants of /aː/ and /a/, respectively,

---

1) The diacritics (except the length mark) are omitted in the following.

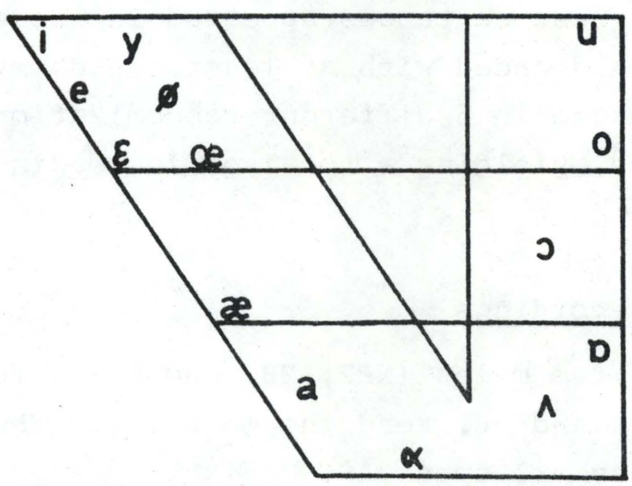


Figure 1

The positions in the Cardinal Vowel diagram of the vowels used in the experiment.

[o, ɔ] are variants of /o/, [ɔː, ɒː] are variants of /ɔː/, and [ʌ] is the phonetic realization of /ɔ/.

The 24 vowels occurred in nonsense words of the type: ˈbVbə and were embedded in a frame sentence:

'Trykket i bVbə ligger på første stavelse.'
(The stress in ˈbVbə is on the first syllable.)

In a small supplementary material one of the subjects read the words biˈbibə and bɑˈbɑbə - in the same frame sentence - to examine the relationship between tongue height and duration in unstressed vowels. To avoid a rhythmic reading of the text, 14 dummy sentences were mixed with the test sentences, in such a way that no more than 5 test sentences appeared in succession, and every page started and ended with at least one dummy. The 24 + 14 sentences were arranged in 5 different randomizations in a list which was read twice, yielding a total of 10 readings by each subject.

2.2 Subjects and recordings

Five subjects, two males (NRP, MB) and three females (KM, ER, NT), aged between 24 and 38, read the material. They are all phoneticians and they all speak ASC Danish.

The recordings took place in a sound treated room at the Institute of Phonetics with professional equipment (Revox A700 tape-recorder, 7½ i.p.s., Sennheiser MD21 microphone, Agfa PE36 tape). The ten readings were obtained in one recording session which lasted from 20 to 30 minutes. This may be an upper limit before fatigue effects occur, but none of the subjects complained and they all declared that it was an easy task.

3. Registration and measurements

The same tape recorder was used for recording and replay. The tapes were processed by intensity and pitch meters (F-J Electronics) and registered on an Elema 800 mingograph at a paper speed of 100 mm/sec.[1], cf. fig. 2. One intensity curve is high-pass filtered at 500 Hz, integration time 2.5 ms, logarithmic display, the other has full frequency range, integration time 2.5 ms, and linear display. A short integration time facilitates an accurate segmentation, which is crucial for duration measurements.

Segmentation, however, is not the only problem in vowel duration investigations. One must also decide on a definition of vowel beginning and end: Peterson and Lehiste (1960) regard the aspiration of preceding stop consonants as being part of the vowel. Fischer-Jørgensen (1964) considers the vowel to start where the higher formants appear in the spectrum which is also the procedure adopted by Holtse (1977), and I have used the same criterion so that the previous and present results for Danish may be directly compared. Accordingly, the vowels are considered to begin where the high-pass filtered intensity curve rises sharply. This is approximately 1 cs after the oscillations start - and it corresponds to the point in time on sonagrams where the higher formants appear.

The vowels terminate where the intensity curves start to drop sharply. There were, however, several cases where this point was difficult to establish: when the following b was voiced in the beginning of the closure. I had expected that the high-pass filtered intensity curve would eliminate this voicing, but it did not do so completely, and since this curve did show normal decreases for nasals (in the dummy sentences), this energy cannot be due to deficient filtering (but it can be due, at least partly, to the logarithmic display): Fig. 3 shows mingograms and sonagrams of ˈby:bə and biˈbibə, respectively, which clearly demonstrate that e.g. the interval with lower, but level, intensity in i in fig. 3c must be due to the energy in the 'voice bar'.

Measurements were made in whole millimeters (centiseconds), i.e. the measuring accuracy is ±0.5 cs.

---

1) The paper speed was checked intermittently throughout registration, and turned out to be constant, with distances between the 1 Hz pulses of exactly 100 mm.

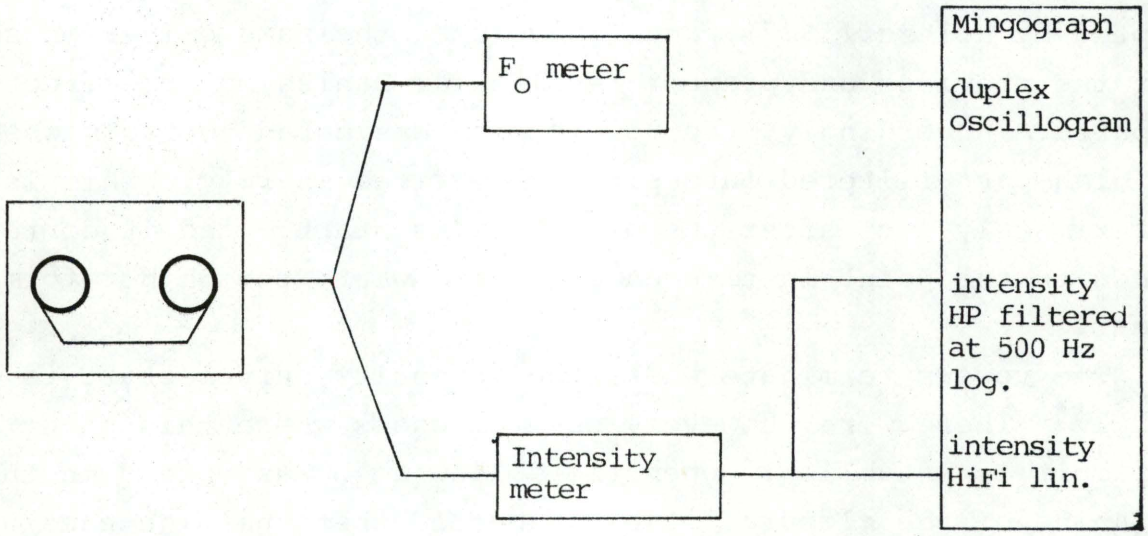


Figure 2
Block diagram of the experimental set up.

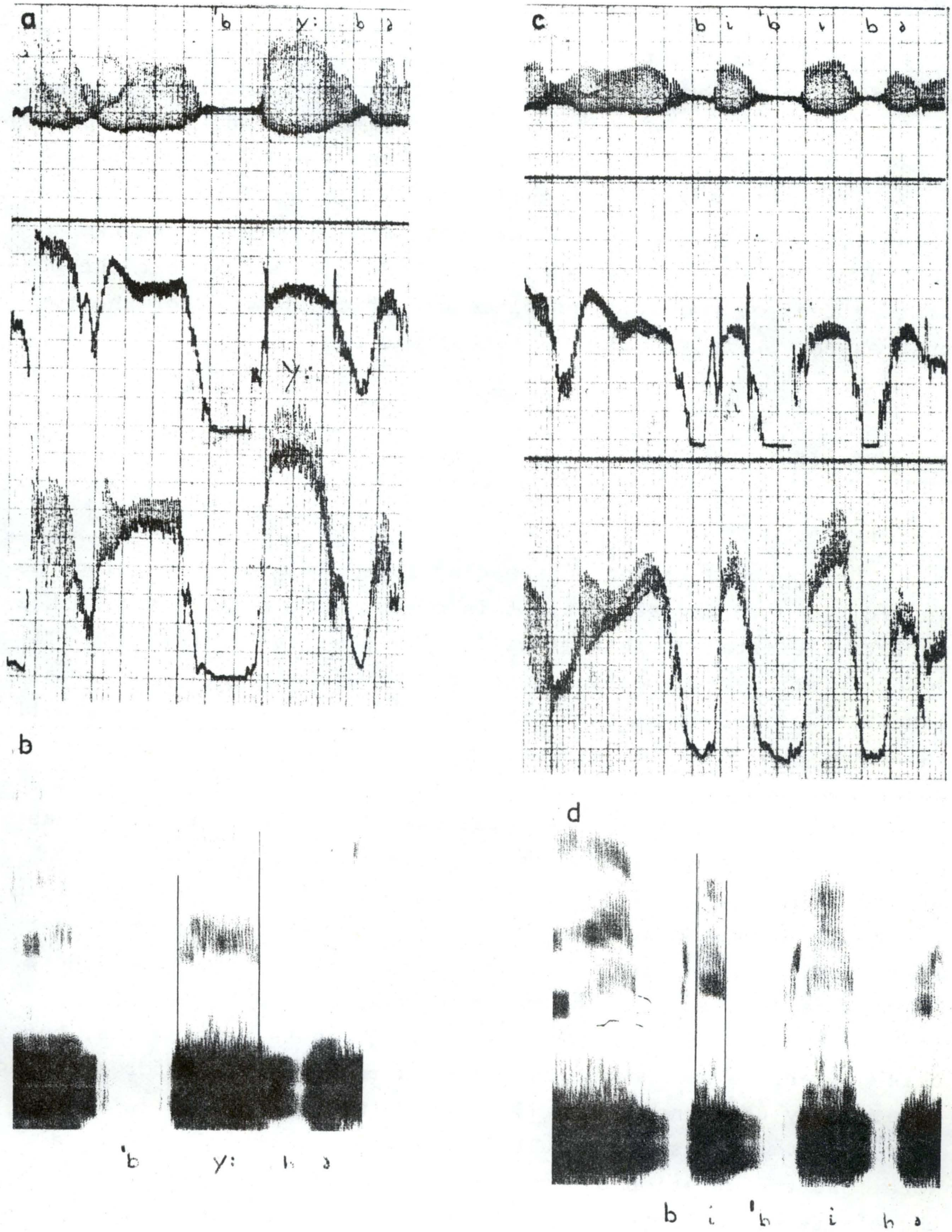


Figure 3

Mingogram (a) and spectrogram (b) of the word ˈby:bə, and mingogram (c) and spectrogram (d) of the word biˈbibə.

4. Statistical treatment

Each subject is treated separately. A one-way analysis of variance was performed to find out whether the durations obtained for the different vowels belong to the same or to different populations. Further, a multiple comparison procedure - Sheffé's method - (Ferguson 1971) was employed in which each vowel quality's mean value is compared to every other mean (within each of the two length categories). This was done in order to single out the possible effect of rounding and place of articulation. For the two unstressed vowels a Student's t-test was run.

5. Results

5.1 Speaking rate

Though no variation in speaking rate was observed during the recordings with any subject, it is of course possible that such a variation exists and that it will influence vowel durations. If vowel duration and speaking rate are correlated, and if it is a linear relationship, then homogeneity of the data (within as well as across subjects) can be achieved by applying a simple scaling procedure to the data.

Rather than measure the total duration of the utterances as an expression of speaking rate, the duration from the explosion of the utterance-initial t to the beginning of the first vowel in "første" was measured, in order to eliminate, to the extent that it is possible, an intra-utterance rate variation effect. There was a clear tendency in the material that the last word varied more in duration than did other words in the test sentences, but this difference has been eliminated in the expression for speaking rate. Only utterances with i, ɑ, and u were measured.

Fig. 4 depicts the results for each subject and each vowel. Note that the duration of the test vowel has been subtracted from the entity which expresses speaking rate in order to eliminate a "double" effect from any correlation between vowel duration and speaking rate. Apparently, duration and rate are not correlated, neither within, nor across subjects. E.g. NRP has a rather great variation in speaking rate, but his vowel durations are nearly

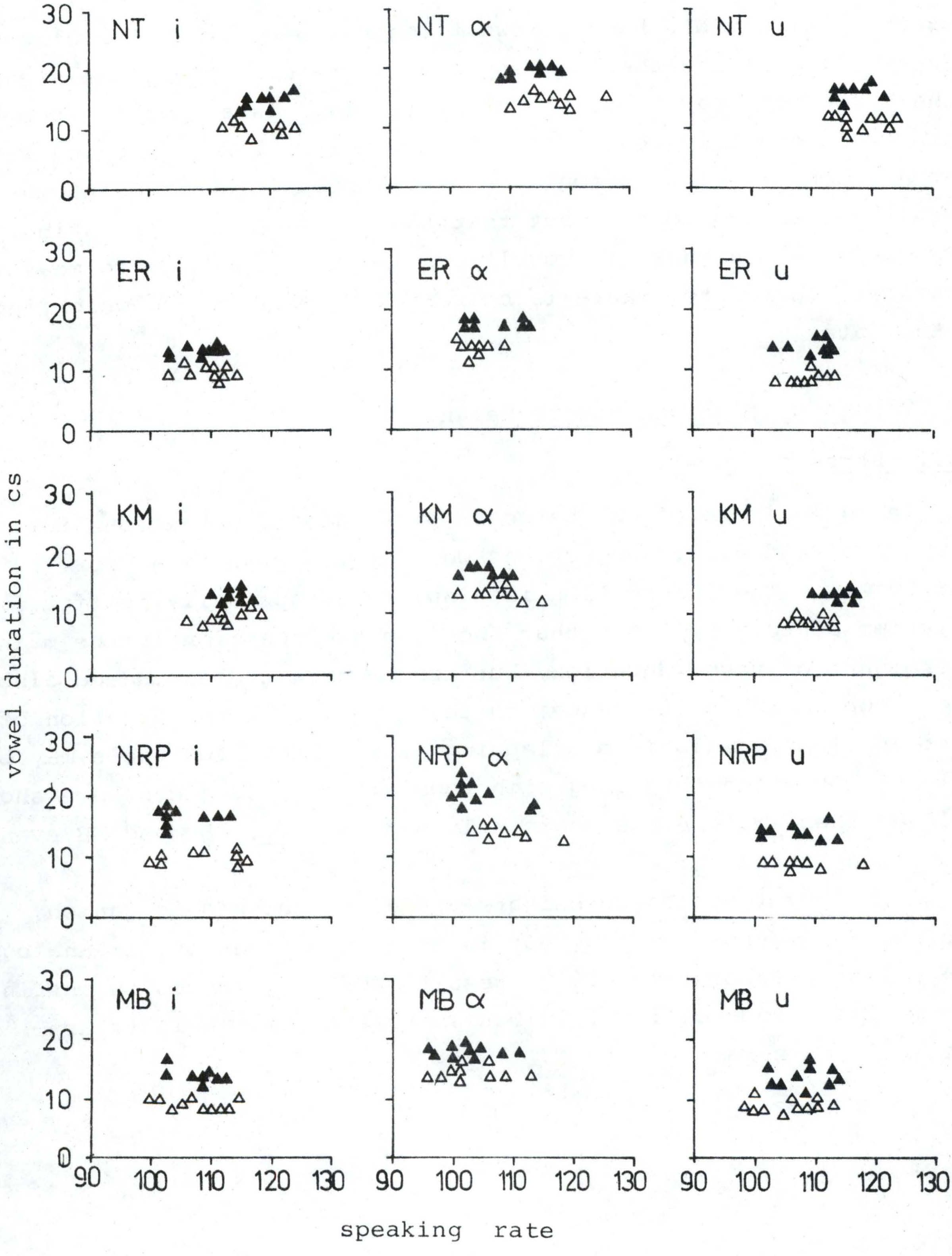


Figure 4

Vowel duration plotted as a function of speaking rate for the vowels i, α, and u. Filled and open triangles indicate long and short vowels, respectively. The derivation of the expression for speaking rate is explained in the text.

constant. NT and NRP have somewhat longer vowels than the other subjects, but NT has the highest speaking rate. KM has the second highest speaking rate, but the shortest vowel durations. Of course, it is true generally that vowels are longer at lower speaking rates. That no such correlation could be established in the present material must be due to the fact that the differences in speaking rate (and thus in the test vowel) are small in relation to measuring accuracy. Therefore, there is no reason to perform any weighting of the data.

5.2 Vowel duration and tongue height

5.2.1 Stressed vowels

In table 1 are given the means and standard deviations for every vowel and every subject, as well as the grand mean for every vowel. The information is displayed graphically in fig. 5. It is immediately apparent that the length opposition is the main determinant of vowel duration, but tongue height does indeed also have a considerable influence, to the effect that the duration of short and long vowels is overlapping, i.e. short low vowels may be of longer duration than long high vowels: Four subjects have short α longer than i:, one has short ʌ longer than u:. ER and MB even have α longer than o: and ø:.

Normally, four tongue heights are distinguished in Danish, high/semi-high/semi-low/low, but in this investigation, in analogy with Reinholt Petersen's (1977) results for differences in intrinsic fundamental frequency level in Danish vowels, five heights are distinguished here:

1.	i	-	y	-	u
2.	e	-	ø	-	o
3.	ɛ	-	œ	-	ɔ
4.	a/æ:		-		ʌ/ɒ:
5.			α		

The vowels have been grouped, in table 1 and fig. 5, according to length (short/long) and in three series of unrounded front, rounded front, and back vowels, and within each series according to tongue height. The tendency for duration to increase with increased

Table 1

Means and standard deviations for short and long vowels for the individual subjects and for all subjects pooled. The leftmost column gives the F-values obtained in the one-way analyses of variance applied to the data of each subject.

			i	e	ɛ	a/æ	ɑ	y	ø	œ	u	o	ɔ	ʌ/ɒ	F
NT	short	$\overline{X}$	9.9	11.1	12.3	13.8	14.6	10.5	12.2	12.8	10.7	11.7	12.6	13.7	30.64
		s	0.74	0.74	0.82	0.63	1.08	0.71	1.40	0.63	0.95	0.68	0.70	0.68	
	long	$\overline{X}$	14.6	15.9	17.4	17.8	19.3	15.4	16.3	17.4	15.7	16.2	18.1	18.5	12.60
		s	o.97	0.88	0.70	1.14	0.82	1.90	1.25	0.70	0.82	1.32	0.99	0.85	
NRP	short	$\overline{X}$	9.6	10.5	12.6	13.6	14.0	10.0	10.6	12.4	9.3	9.9	12.1	14.1	28.75
		s	0.70	0.53	1.27	0.52	1.05	0	0.70	0.52	1.25	0.32	0.57	1.10	
	long	$\overline{X}$	16.1	16.1	17.1	18.3	19.6	15.7	16.5	17.3	14.0	15.5	17.8	17.6	13.17
		s	1.10	0.88	0.88	1.16	1.57	0.68	1.18	0.68	0.94	1.43	1.40	0.97	
KM	short	$\overline{X}$	9.4	10.5	10.9	12.6	13.1	9.4	9.9	10.8	8.9	9.7	9.9	12.6	32.75
		s	0.97	0.71	0.32	0.52	0.86	0.70	0.32	0.63	0.74	0.68	0.74	0.52	
	long	$\overline{X}$	12.7	13.2	14.3	15.2	16.3	12.5	13.4	14.9	12.8	14.3	13.9	14.5	17.12
		s	0.95	0.92	0.48	0.78	0.52	1.08	0.97	0.57	0.63	0.48	0.57	1.43	
ER	short	$\overline{X}$	9.6	9.8	11.3	12.7	13.8	9.2	9.5	11.8	9.2	8.5	10.9	12.7	30.51
		s	0.70	0.42	0.48	0.48	0.79	0.63	0.97	0.79	0.92	1.18	0.74	0.68	
	long	$\overline{X}$	13.0	14.0	14.9	15.7	17.3	13.3	14.4	15.0	13.5	13.5	15.3	15.2	10.26
		s	0.67	0.94	0.99	0.68	0.48	1.16	1.43	1.25	1.08	1.18	1.83	1.03	
MB	short	$\overline{X}$	9.0	1o.0	11.6	13.5	14.3	8.8	9.5	11.0	9.1	10.7	11.6	12.9	26.64
		s	0.94	1.25	0.52	0.85	0.82	0.79	1.08	0.82	0.88	0.68	1.35	0.74	
	long	$\overline{X}$	13.8	14.5	15.6	17.4	17.6	12.8	13.7	16.7	13.9	14.8	16.1	17.4	13.47
		s	1.o3	1.52	1.35	1.07	1.o8	0.42	0.67	0.68	1.45	1.14	1.60	0.84	
Grand mean	short	$\overline{X}$	9.5	10.4	11.7	13.2	14.0	9.6	10.3	11.8	9.4	10.1	11.4	13.2	
		s	0.33	0.51	0.70	0.55	0.57	0.67	1.13	0.87	0.72	1.14	1.06	0.63	
	long	$\overline{X}$	14.4	14.7	15.9	16.9	18.0	13.9	14.9	16.3	14.0	14.9	16.2	16.6	
		s	1.37	1.24	1.35	1.35	1.37	1.50	1.45	1.23	1.07	1.05	1.75	1.70	

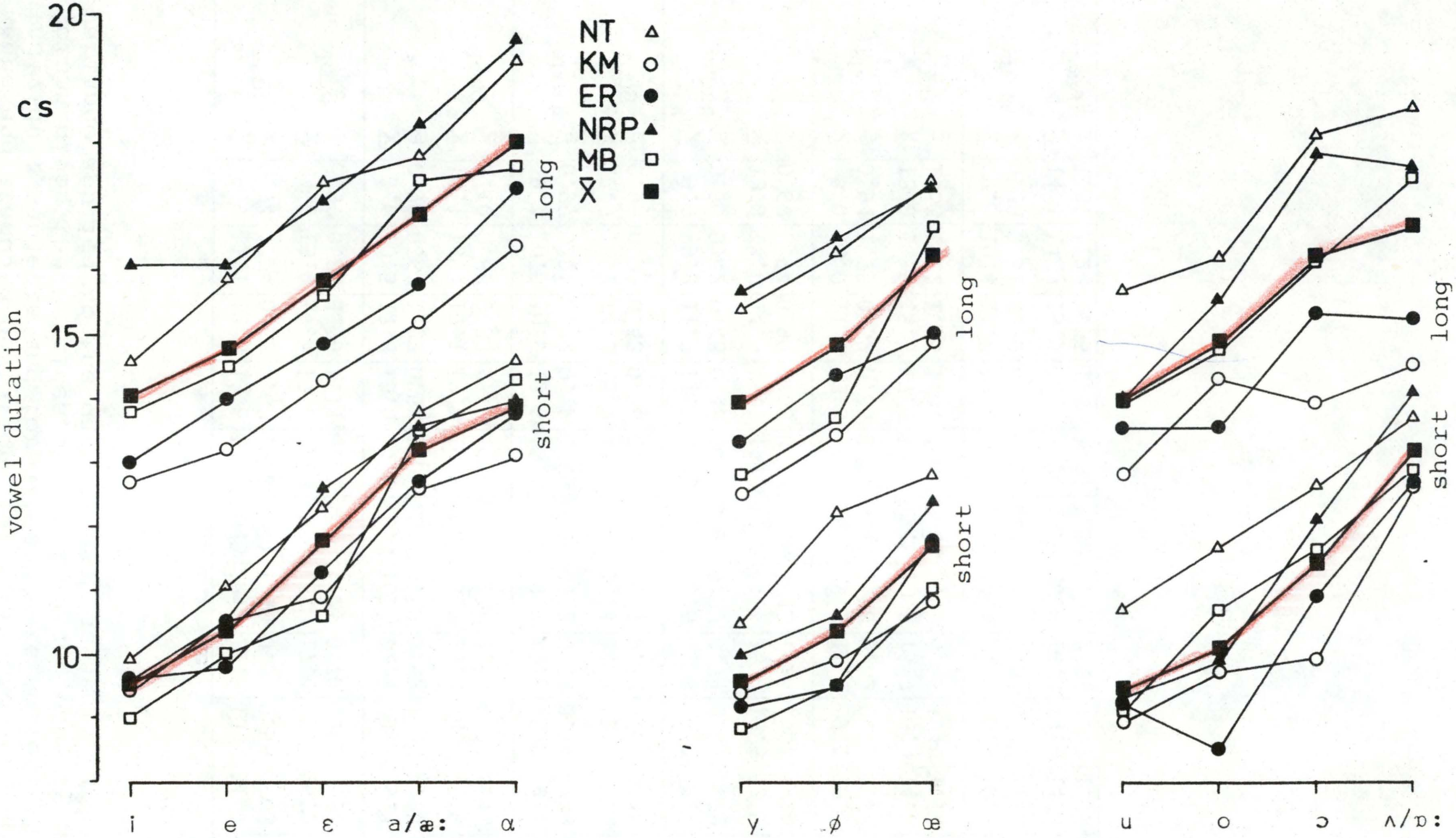


Figure 5

Mean durations for long and short vowels plotted for the individual subjects and for all subjects pooled (termed $\overline{X}$ and indicated by filled squares in the graph).

tongue height number is clear within each series, - there are only 6 mean values out of 120 that oppose this tendency, and no exceptions are found in the grand mean.

The linear dependency of vowel duration upon tongue height may be expressed as follows:

$$\text{Duration}_V = b + a \cdot \text{(tongue height)}$$

where a and b are constants, and tongue height is an integer ranging from 1 to 5. To simplify matters, the grand mean of long and short vowels, respectively, with the same tongue height was calculated, cf. table 2 and fig. 6. The slopes of the regression lines (least squares method) in fig. 6 are 1.0 and 1.2 for long and short vowels, respectively, and the scatter is small (the correlation coefficients are 0.995 and 0.992 for long and short vowels, respectively. Thus, the increase in duration is constant, about 1 cs per tongue height step, which is in agreement with Holtse's (1977) results, and he concludes that (my translation): "Roughly speaking one could say that the vowels are lengthened by 1 cs for every degree they are lowered". Note, however, that this formulation and conclusion imply that the jaw is lowered in equidistant steps from e.g. i through α, which we do not really know to be true, but we may, of course, accept the durational data as an indication of such equidistant tongue heights.

Holtse (1977), Fischer-Jørgensen (1964), and Reinholt Petersen (1974) find that the duration of long vowels increases more with lower tongue height than short vowel duration does. If anything, the tendency is in the opposite direction in this material. This question will be treated in further detail in section 5.3.

It appears from the multiple comparison test, cf. fig. 7, that place of articulation and rounding do not contribute to the durational differences between the vowels. This is in agreement with Fischer-Jørgensen (1964), who found no significant differences according to place of articulation or rounding in the position before b, but she did find such significant differences within vowels with the same tongue height in the position before d. The number of significant differences ($P < 0.05$) with the five subjects between every vowel compared to every other vowel within the same length category (short and long) are listed. Then the figures for

Table 2

Mean durations (in cs) for vowels having the same tongue height. All subjects pooled.

	tongue height number				
	1	2	3	4	5
Short vowels	9.51	10.27	11.64	13.22	13.96
Long vowels	13.99	14.82	16.12	16.76	18.04

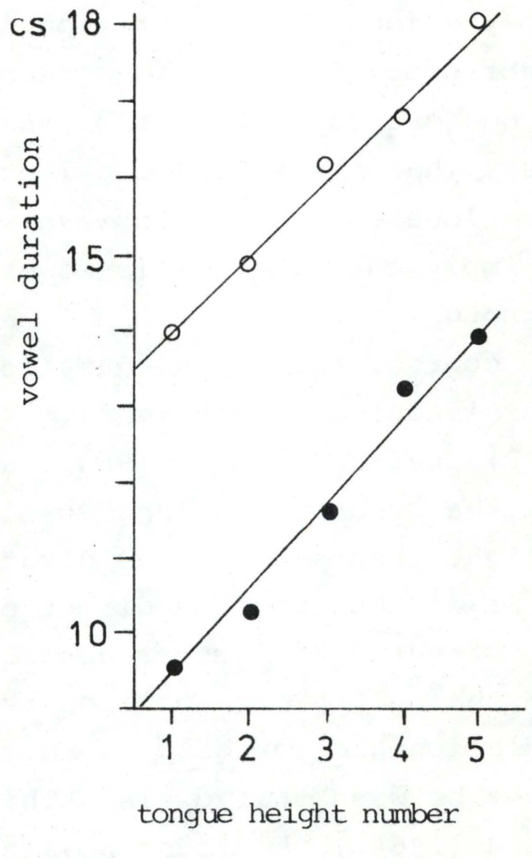


Figure 6

Vowel duration as a function of tongue height. Each data point represents an average over all subjects and all vowels having the same tongue height. Open and filled circles indicate long and short vowels, respectively. The straight lines are regression lines fitted to the data points.

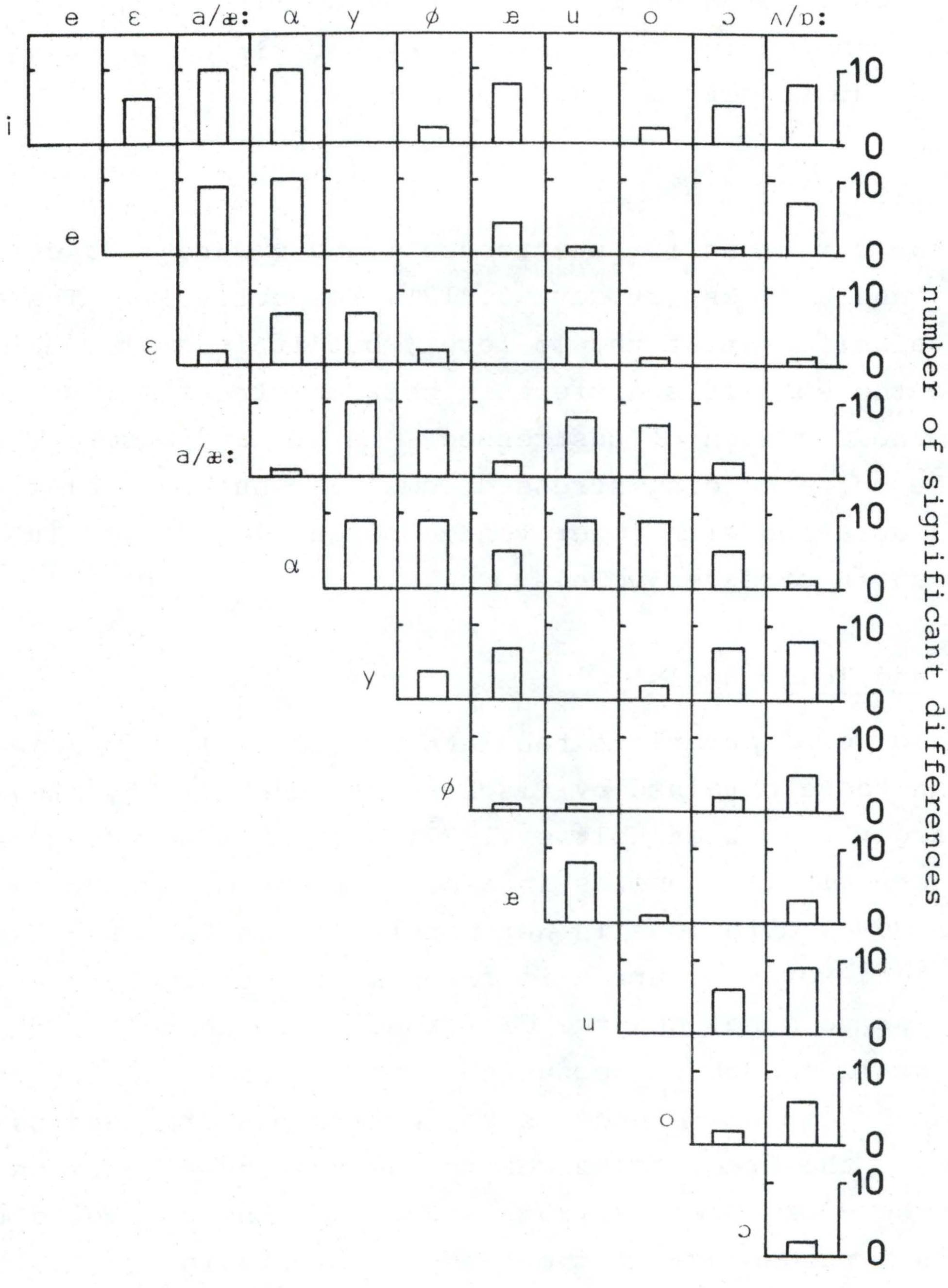


Figure 7

Number of significant differences ($p<0.05$) obtained by the multiple comparison procedure. For further explanation, see text.

long and short vowels are added, i.e. the maximum possible number of significant differences for each vowel quality is 10. Sheffé's method is a conservative one, which does not yield many significant differences, but it serves its purpose here.

5.2.2 Unstressed vowels

The mean durations for unstressed i and ɑ were 5.25 cs (st. dev. 0.635) and 6.35 cs (st.dev. 0.412), respectively. The difference is significant at the 1% level, but it is small (1.1 cs) compared to the subject's difference between stressed i and ɑ (4.7 cs). The duration of unstressed i is 53% and unstressed ɑ 43.6% of the corresponding stressed vowels. Thus the relative increase in duration with lower tongue height is smaller in unstressed than in stressed vowels.

5.3 Short and long vowels

The mean vowel durations tabulated in table 1, are somewhat smaller than those obtained by Fischer-Jørgensen (1955, 1964) and slightly larger than what Holtse (1977) got. Fischer-Jørgensen's test words were isolated words in a list, which may account for the greater vowel durations in her material. In Holtse's experiments, the test words occurred in frame sentences under conditions rather similar to mine, and the difference between Holtse's and my results is small enough to be due to differences between individuals and/or to a difference in the stress patterns and segmental composition of the frame sentences in the two investigations.

In Fischer-Jørgensen's (1955) study, the short vowel duration expressed as a percentage of the long vowel duration is 50.6%, in her 1964-study it is 57%; in Holtse's (1977) investigation the figure is 67%, and in the present study it is 71%. Holtse investigated vowels of height number 1 and 3 only, which occurred in words with a varying number of succeeding unstressed syllables and in varying consonant environments. The percentage for the corresponding vowels in my material is 69.5%, so Holtse's and the present results are in good agreement. It would have been tempting to explain the difference between Holtse's and my results as an indication of a relative lengthening of short vowels in ASC Danish, if it were not for the fact that Holtse's subjects were <u>not</u> ASC speakers, and the difference between his short:long fraction and the results that I have obtained (2.5%) is too small to be indicative of a relative lengthening of short vowels in ASC Danish. Fischer-Jørgensen's small values (as compared to Holtse's and mine) are due to the fact that long vowels are considerably longer in isolated words than in sentences, whereas short vowels are only slightly lengthened, cf. Fischer-Jørgensen's values (average over all vowels) of 13.9 cs and 26.9 cs (short and long vowels, respectively) as against 11.9 cs and 15.9 cs in the present material: the difference is evidently in the long vowels.

The relation between short and long vowels may be illustrated graphically as in fig. 8, where duration of the short vowels is depicted as a function of the corresponding long vowels. The regression line (least squares method) has the equation:

$$\text{short vowel} = 1.19 \cdot (\text{long vowel}) - 7.3$$

The slope of 1.19 deviates considerably from Holtse (1977) and Lindblom (1967), who both arrived at a slope of 0.75. Lindblom had only one subject, Holtse employed three, but Holtse calls attention to inter-individual differences and assigns the identity with Lindblom's result to a coincidence.

Even though I have treated the a/æ: and ʌ/ɒ: as short long pairs with the same tongue height, the short member of each pair does have a lower tongue position than the long one (cf. fig. 1), and thus presumably a longer intrinsic duration, which makes for relatively larger y-values in fig. 8. In fig. 9 those two vowel pairs have been excluded, and the regression line equation is now:

$$\text{short vowel} = 1.075 \cdot (\text{long vowel}) - 5.6$$

The slope is so close to unity that a simplification as follows seems justified:

$$V_{short} = V_{long} - b$$

where b is a constant, approximately 5 cs in this case. Thus, the relation between short and long vowels is constant over different vowel pairs. In fig. 10, the short:long difference of the 12 pairs, for individual subjects as well as the grand mean, is de-

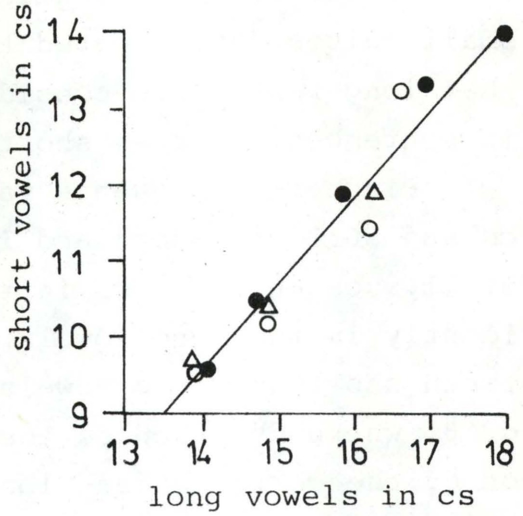


Figure 8

Short vowels plotted as a function of long vowels (all subjects pooled). Unrounded front vowels are indicated by filled circles, rounded front vowels by triangles, and back vowels by open circles.

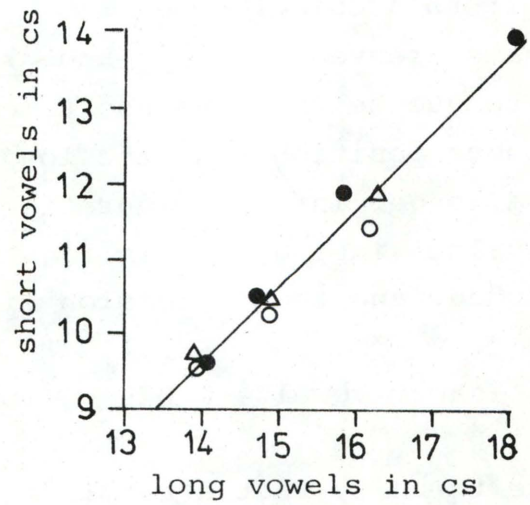


Figure 9

Same as fig. 8, but with the pairs a/æ: and ʌ/ɒ: left out.

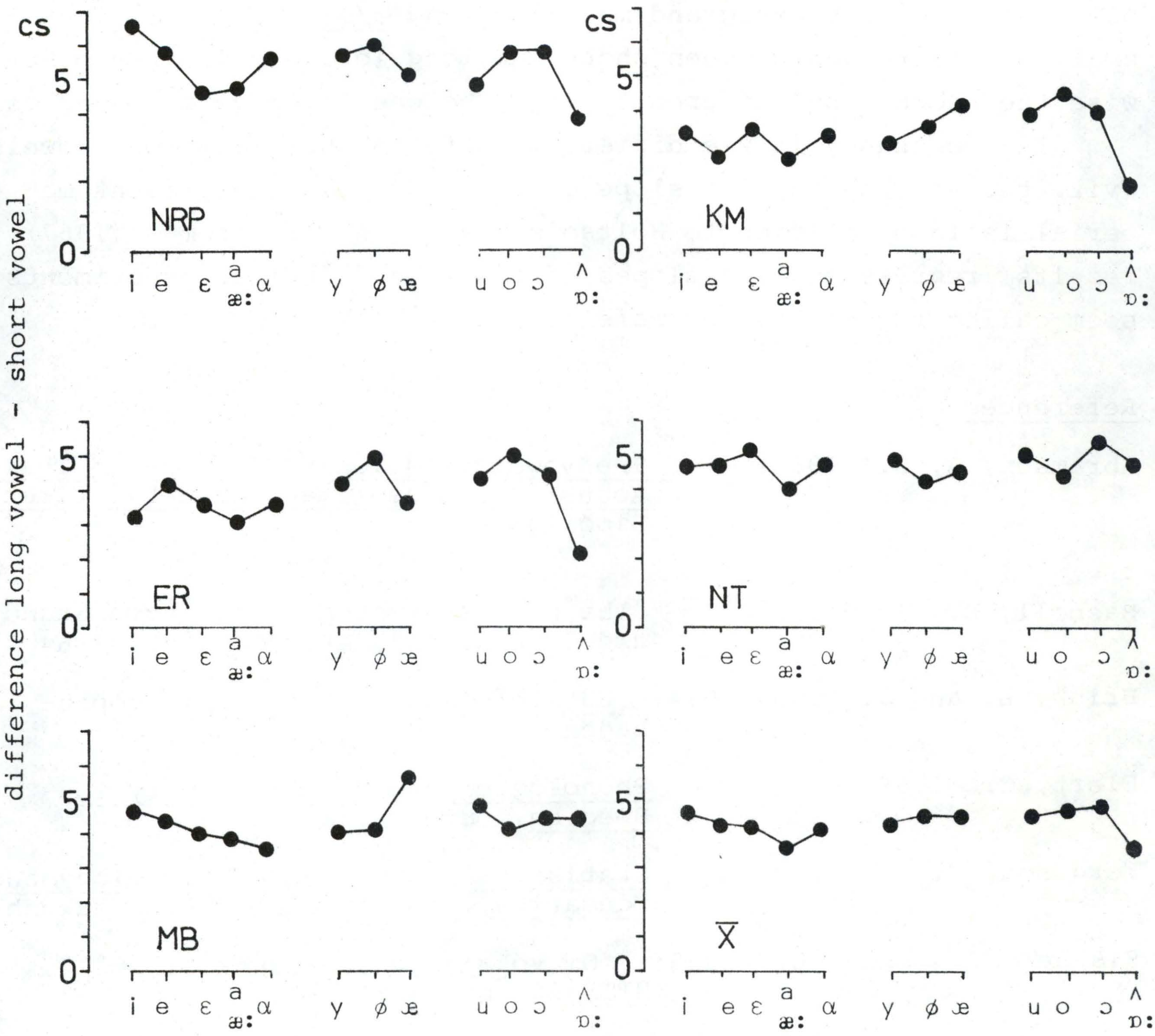


Figure 10

Differences between long and short vowels for the individual subjects, and averaged over all subjects.

picted. The constancy is not very apparent with individual subjects, but on the other hand, the variation seems to be random, and it is possible that it would disappear with a larger number of readings. With the grand mean a/æ: and ʌ/ɑ: stick out with smaller differences between short and long (cf. above), but otherwise the short/long difference ranges between 4.08 cs and 4.82 cs.

The constancy in the difference between short and long vowels (viz. the regression line slope of 1.0) found in the present material is in opposition to Holtse's (1977) and Lindblom's (1967) results (regression line slopes of 0.75), and further experiments seem called for before any safe conclusions can be drawn.

References

Abramson, A. S. 1962: The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments, Bloomington, Ind.

Basbøll, H. 1968: "The phoneme system of Advanced Standard Copenhagen", ARIPUC 3, p. 33-54

Brink, L. and J. Lund 1974: Udtaleforskelle i Danmark, Copenhagen

Elert, C. 1964: Phonologic Studies of Quantity in Swedish, Stockholm

Ferguson, G. 1976: Statistical Analysis in Psychology and Education, (fourth edition), Kogakusha

Fischer-Jørgensen, E. 1955: "Om vokallængde i dansk rigsmål", NTTS 15, p. 33-56

Fischer-Jørgensen, E. 1964: "Sound duration and place of articulation", Zs.f.Ph. 17, p. 175-207

Fischer-Jørgensen, E. 1972: "Formant frequencies of long and short Danish vowels", ARIPUC 6, p. 189-213

Holtse, P. 1977: "Variationer i vokallængde på dansk", (unpublished manuscript)

Jespersen, O. 1926: Lehrbuch der Phonetik, Leipzig

Lindblom, B. 1967: "Vowel duration and a model of lip mandible coordination", STL-QPSR 4, p. 1-29

Lindblom, B. 1976: "Durational pattern of Swedish phonology" (mimeographed), Stockholm

Maack, A. 1949: "Die spezifische Lautdauer deutscher Sonanten", Zs.f.Ph. 3, p. 190-232

Peterson, G. and I. Lehiste 1960: "Duration of syllable nuclei in English", JASA 32, p. 693-703

Reinholt Petersen, N. 1974: "The influence of tongue height on the perception of vowel duration in Danish", ARIPUC 8, p. 1-10

Reinholt Petersen, N. 1977: "Intrinsic fundamental frequency of Danish vowels", JPh 6, p. 177-190

Thorsen, N. and O. Thorsen 1978: Fonetik for sprogstuderende, Copenhagen.

(ARIPUC 14, 1980)

NEUTRAL STRESS, EMPHATIC STRESS, AND SENTENCE INTONATION IN ADVANCED STANDARD COPENHAGEN DANISH[1]

Nina Thorsen

Abstract: The relationship between stress and fundamental frequency in short declarative and syntactically unmarked interrogative sentences, both with emphasis for contrast in various positions, is investigated and compared to prosodically neutral statements and questions on the basis of recordings by four speakers. Emphasis for contrast has a radical influence on the course of fundamental frequency, to the extent that the stress group which contains the stressed syllable of the emphasized word and its neighbours tonally reduce to one stress group. The influence from emphasis seems to reach farther on marked (non-declarative) than on unmarked (declarative) contours. Durational differences between utterances with and without emphasis for contrast are small, consisting mainly in a slight lengthening of the emphasized stress group, a lengthening which is to some extent counterbalanced by a shortening of the preceding stress group, if any.

1. Introduction

Intonation in short sentences in Advanced Standard Copenhagen (ASC) Danish may be accounted for as in fig. 1, which is only a model - with the advantages and shortcomings that modeling almost always entails in terms of simplicity and inaccuracy, respectively.[2] For a detailed account of the material and procedure that led to the formulation of this model, see Thorsen (1978, 1979).

The complex course of fundamental frequency (Fo) in an utterance is assumed to be the outcome of a superposition of several components. (1) A sentence component which supplies the INTONATION

---

1) Revised and expanded version of a paper published in the Proceedings of the Ninth International Congress of Phonetic Sciences, Copenhagen, 6-11 August 1979, vol. II, p. 417-423.

2) This model is presented here for the umpteenth time, with more or less the same comments as in previous papers. I hope the reader will bear with me, since it does constitute the point of departure for the main part of the paper.

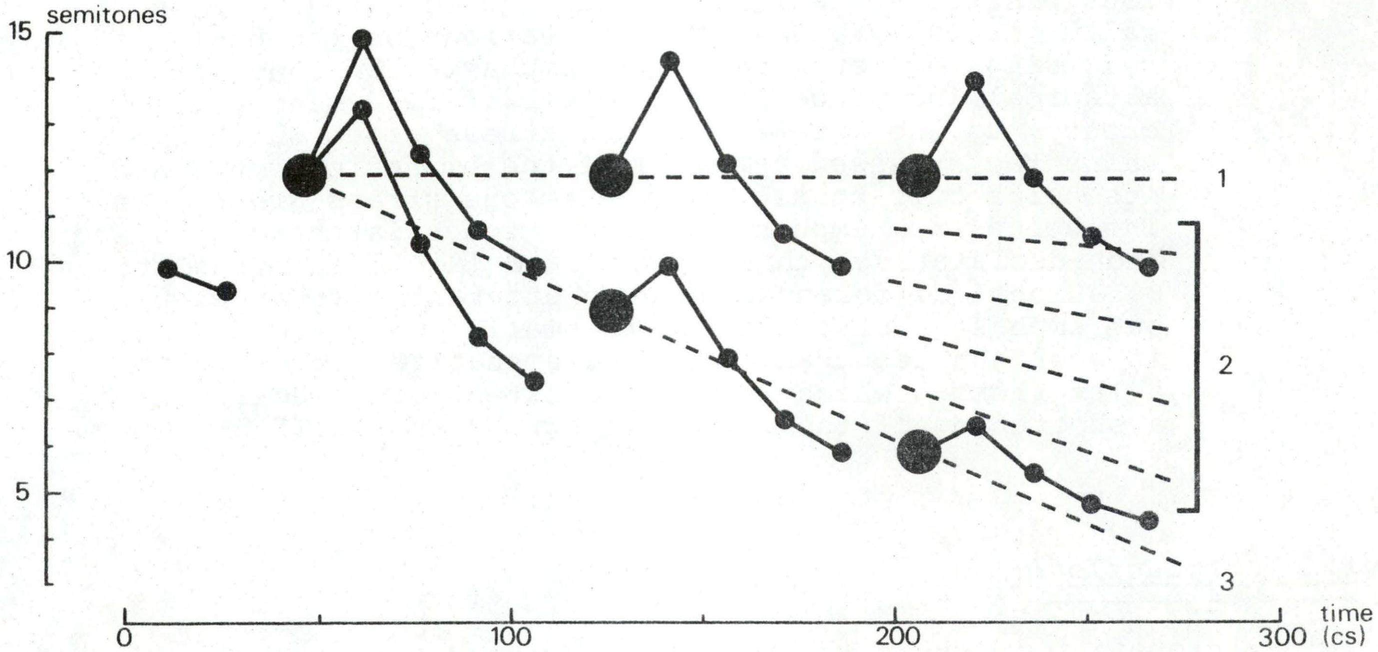


Figure 1

A model for the course of fundamental frequency in short prosodically neutral 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. The full lines represent the Fo pattern associated with stress groups, and the broken lines denote the intonation contours. Zero on the logarithmic frequency scale corresponds to 100Hz.

CONTOUR (broken lines in fig. 1). (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,[1] a stød component, rendering STØD MOVEMENTS (not included in the model). 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 can be ascribed to inherent properties of the speech production apparatus and which therefore is superfluous in the model from the point of view of the human speaker. (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 1977.) Similar points of view about components or "layers" in intonation have been expressed by several authors, e.g. Bolinger (1970), Bruce (1977), Carlson et al. (1974), Cohen and 't Hart (1967), Collier and 't Hart (1975), Fujisaki et al. (1979), Gårding and Lindblad (1973), 't Hart (1966), 't Hart and Cohen (1973), Lehiste and Peterson (1961), Öhman (1968); for a more detailed account, see Thorsen (1979).

2. Stress group patterns

What is said in this section about stress in ASC Danish applies to <u>prosodically neutral</u> utterances, i.e. utterances devoid of emphasis of any kind where all the stressed syllables are equally prominent.[2]

---

1) The Danish stød/non-stød distinction may be said to correspond to the Swedish and Norwegian Accent I/Accent II distinction (see further Basbøll 1972 and Gårding 1977), but its manifestation, which exhibits a good deal of dialectal and individual variation (Riber Petersen 1973), is not generally considered to be primarily tonal. In ASC Danish it may be described as a kind of creaky voice which attacks the final part of a long vowel or the succeeding voiced consonant (if the preceding vowel is short).

2) This definition of (prosodic) neutrality applies throughout the paper and is not to be confounded with "unmarked", cf. later sections.

In many languages linguistic stress and Fo (pitch) are interrelated, e.g. in Dutch ('t Hart and Cohen 1973), in English (Fry 1958, Lieberman 1960), in Swedish (Carlson et al. 1974, Bruce 1977), and also in Danish. 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. Thus Berinstein (1979), on the basis of acoustic analyses and perceptual experiments on English, Spanish, K'ekchi, and Cakchiquel, finds support for a hypothesis that "Change in Fo, increased duration, and increased intensity, in that order, constitute the unmarked universal hierarchy for perception of stress in languages with no phonetic contrasts in tone or vowel length; in languages with such contrasts the perceptual cue correlated with that contrast (i.e. Fo with tone and duration with length) will be superseded by the other cues in the hierarchy." (p. 2).

Danish has a phonemic contrast in vowel length, so according to Berinstein the hierarchy for the perception of stress would be: Fo, intensity, duration. I do not wish to dispute the primacy of Fo but I doubt whether intensity really comes second; besides, vowel quality is a factor not to be dismissed in a language like Danish where the system of vowels in unstressed syllables is reduced, cf. Basbøll (1968) and Rischel (1968), and where, further, excessive schwa-assimilation takes place. Thus, I would hypothesize a hierarchy for ASC Danish as follows: Fo, vowel duration and quality, intensity; but of course this is still an area for experimentation.

The entity which governs the patterning of fundamental frequency may vary between languages. Esser (1978) hypothesizes that in German the word is the unit which governs Fo, whereas in English it is the foot (whose definition resembles that of the stress group in Danish, cf. below: a foot consists of an ictus (salient syllable) - which may be silent - and a (non-obligatory) remiss (weak syllable(s)), in that order, cf. Abercrombie 1964, Halliday 1967, p. 12). Carlson et al. (1974) and Bruce (1977) both imply that in Swedish the word does not constitute the basis for stress/Fo patterning and this is true of Danish too: The unit which carries the Fo patterns (full lines in fig. 1) consists of a stressed syllable plus all succeeding secondary- and unstressed syllables (within the boundaries of the same intonation contour), irrespective of intervening syntactic boundaries (see Thorsen 1978, 1980a, 1980c for documentation); this unit is termed a stress group.

2.1 Stress groups with more than one post-tonic syllable

The stress group patterns have the same basic shape, a (relatively) low stressed syllable followed by a high(-falling tail of) post-tonic syllable(s),[1] but the model predicts that the magnitude of the rise from stressed to post-tonic will vary with time and with intonation contour. The decrease with time may be a consequence of either of two distinct processes or of a combination of them: It may be a signal of finality and/or it may be a physiological phenomenon, i.e. the nearer the end of the contour, the less physiological energy is expended and the less complete (or distinct) the gestures will be. The variation in Fo patterns with intonation contour is likely a consequence of the difference in the level of the following stressed syllable, i.e. the lower the succeeding stressed syllable (as in statements compared to syntactically unmarked questions) the less of a rise is performed, and vice versa, an assimilation which may also be conceived of as induced by a physiological constraint. (This explanation would be supported by data on the behaviour of stress group patterns before clause-internal prosodic phrase group boundaries, cf. Thorsen 1980a, section 5, but see also section 7.1 below.)

2.2 Stress groups with only one post-tonic syllable

Stress groups with one post-tonic syllable will of course be shorter than those with several, a feature which is not reflected in fig. 1. Further, a single post-tonic does not always rise as high as does the first of several, and under some circumstances (see further section 7.1.1 below) a single post-tonic can even fall at or slightly below the level of the preceding stressed syllable. A full rise may be intended by the speaker and the undershoot (or assimilation) be due to shortcomings in the peripheral speech production mechanism - but whatever the reason, we

---

1) Note that the post-tonic syllables may include secondary stresses (in compounds), which tonally behave like unstressed syllables, but which in all other respects (vowel quality, vowel quantity, and stød) resemble stressed syllables.

have here an indication that time supersedes Fo when the two are in conflict: the duration of the post-tonic may be insufficient to perform a full rise when this rise is to be succeeded by a large fall to the next stressed syllable and, rather than stretch the duration of the post-tonic, the Fo deflection shrinks.[1] This inflexibility of the time structure vis-a-vis Fo patterns in Danish may or may not be generalizable to other conditions - it would be premature to say. (Lyberg's (1979) experiments on Swedish - where vowel duration and more elaborate Fo movements are correlated - induce him to suggest the inverse priority: "The variation of the segment duration in different places in a sentence is then only a secondary effect of the Fo events which carry the major burden of signalling the prosodic contents and may be executed by a regulation system (feedback or feedforward) at a fairly peripheral level."; i.e. vowels will stretch in order to make room for a required Fo movement.)

2.3 Stress groups with no post-tonic syllables

In stress groups consisting only of a stressed vowel, two possibilities present themselves: either the characteristic Fo pattern is absent, or it is (partially) compressed to be contained within the stressed vowel (and/or succeeding sonorant consonants, if any). The second solution would be parallel to the situation in Swedish (cf. Lyberg 1979). Although this point warrants a separate investigation, based on a carefully controlled and diversified material, I am - on the basis of materials analysed so far - rather in favour of the first solution, because short utterance final stressed vowels (followed by unvoiced consonants) in statements do not show any trace of a compensatory rise (see Thorsen 1978 and section 7.2.2 below). Furthermore, although long vowels often end in a rise (a rise which in statements is typically smaller than the preceding fall, however, - see fig. 8, left), this rise is not confined to stress groups without post-tonics, - see further section 5 below, about intravocalic Fo movements.

---

1) However, there is reason to believe that a speaker may overcome the 'sluggishness' of the peripheral speech organs to produce what may be termed "maximally distinct" Fo patterns, - see further section 7.1.1.

We may ask: What is the perceptual cue to stress when the characteristic Fo rise is missing from the stress group? - First of all, this may be a situation where other, linguistic and acoustic, cues step in (vowel quantity/quality and stød;[1] vowel duration and possibly intensity). Secondly, the first stressed syllable in an utterance cannot be perceived as unstressed by virtue of its position rather high in the frequency range. Thirdly, a succession of two stressed syllables - on any kind of intonation contour - will rarely resemble a combination of stressed and post-tonic syllables, see fig. 1. And lastly, the rhythm of the utterance probably also plays an important role: In 'stress-timed' languages, the stressed syllables tend to be perceived as occurring with equal time intervals, cf. Lehiste (1977) and Donovan and Darwin (1979), and therefore a syllable which fills an otherwise empty slot in the rhythmic structure is likely to be identified as stressed. (Of course, this argument easily risks to become circular - I am thinking here of a situation where the rhythmic pattern has been or can be established by surrounding stressed syllables.)

2.4 Stress group patterns - conclusion

If the variation in Fo patterns with time, intonation contour, and stress group composition (one vs. several post-tonics) is physiologically conditioned (in a not too strict sense of the term), the speaker may be unconscious of it, and the listener may neglect or compensate for it. Accordingly, we are left with a stress group component which is invariant from a productional (and perceptual) point of view. (For speech synthesis specific rules must of course be included to take care of the context dependent modification in stress group patterns.) I shall return to the tonal properties of the stress group and look at supplementary data in section 7 below.

---

1) Unstressed syllables can have neither long vowel nor stød, cf. Basbøll (1968).

3. Intonation contours

The recurrency and predictability of the stress group patterns gave rise to the definition of the intonation contour as the course described by the stressed syllables alone (dotted lines in fig. 1). This does not mean that the course of the post-tonic syllables is irrelevant, e.g. for the identification of intonation contours, but it is strictly speaking redundant (see further Thorsen 1980b). A similar concept of intonation contour is found in Bolinger's (1958, 1970) treatment of American English and in Carlson et al. (1974). For a discussion of this definition vs. the current 'topline' and 'baseline' concepts see Thorsen (1980a, section 1).

In short utterances (containing no more than three or four stress groups) the intonation contours approach straight lines whose slopes tend to vary systematically with sentence type, as suggested by fig. 1: Declarative sentences have the most steeply falling (unmarked) contours at one extreme, syntactically unmarked questions have horizontal contours at the other extreme. In between are found other question types and non-final periods. Further, there seems to be a certain trade-off between syntax[1] and intonation contour: The more syntactic information is contained in the sentence about its interrogative or non-final function, the more declarative-like, i.e. the more steeply declining, is its intonation contour, and vice versa (a tendency also noted by Bo 1933, p. 82-83, and Jespersen 1897-99, p. 592). A similar trade-off has also been observed for other languages, see e.g. Bolinger (1962), Cohen and 't Hart (1967), Daneš (1960), von Essen (1956), Hadding-Koch (1961), and Mikoš (1976). Subjects may differ slightly in the way they spread out their contours: some tend to cluster all of them, except the syntactically unmarked question, rather low and close to the declarative contour, others show the opposite tendency, i.e. all but the declarative contour are close to horizontal, but even so an ordering among the contours can be observed and the intersubject agreement on the order

---

1) I do not wish to exclude the possibility of treating intonation as an integral part of the syntax of the language, and "syntax" and "syntactic" should just be regarded as convenient abbreviations for "other signals, such as word order, interrogative particles, and the like".

of the contours from horizontal to steepest is surprising: In two different sets of sentences, each consisting of five different question types, three non-final periods and a declarative utterance (see Thorsen 1978), the level of the last stressed vowel in each utterance (which serves well as an indication of the slope of the contour, cf. Thorsen 1980b) was measured and averaged over six readings by each of four subjects. The degree of concordance across subjects in the rank ordering of these vowel levels in the 9 sentence types was: W = .94 and .95, respectively, for the two sets of sentences (Kendall's coefficient of concordance).

In utterances of four or more stress groups the intonation contour breaks up into a succession of gradients with (slight) partial resettings between them, though preserving an overall downdrift (see Thorsen 1980a for a further account) but it still seems that the stress group patterns can be regarded as simple superpositions on the intonation contours, i.e. the behaviour of the post-tonic syllables can still be accounted for in terms of the stressed syllables.

4. Stress groups and intonation contours - conclusion

Evidently, if we combine all possible variations in stress group composition with variations in number of stress groups in the utterance (and, further, if we manipulate the syntactic structure to make the intonation contour discontinuities occur in different places, cf. Thorsen 1980a), we are left with an infinity of physically different Fo courses where stress group patterns and intonation contours are highly interactive. (This richness and in particular the many unilluminated points make me hesitate to propose a proper "generative" model for Danish intonation; further data seem called for.) However, as long as the post-tonic syllables are predictable in terms of the stressed syllables (i.e. in terms of the intonation contour which, in its turn, is determined by the syntax and function of the utterance and by its length), we may still conceive of a level of production where stress group patterns and intonation contours are independent entities which are simply superposed one upon the other, and we may also hypothesize a level in the perceptual process where a lot of the physical variation is obliterated, recreating the simple configurations depicted in fig. 1.

5. Intrasyllabic Fo movements

The account of Fo movements within the syllable (vowel) is much less complicated because Fo movements seem to be predictable from the stress group pattern and (indirectly) the intonation contour, and these movements are therefore not specified in fig. 1: An account of vowel movement will suffice because the Fo movement in voiced consonants seems always to be a smooth interpolation between the vowels, modified by specific intrinsic Fo properties of the consonant, as the case may be (see further Thorsen 1979).

The post-tonic vowels are generally falling. The first post-tonic, however, may be rising-falling or even purely rising, depending on the exact timing of the vowel with respect to the Fo maximum. Short stressed vowels are also generally falling on unmarked intonation contours, but may be rising in the first stress group, i.e. at the start of the intonation contour, which is always high, cf. fig. 1. On marked (less falling) contours short vowels are less falling. Long stressed vowels on falling intonation contours are falling-rising, but the fall is normally equal to or greater than the rise, except initially in the utterance. On marked intonation contours the rise may take the upper hand (see further section 7.2.1). - We might ask: Why are stressed vowels generally and mainly <u>falling</u> (on falling intonation contours), 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 (1) we also observe a tendency for a positive correlation between short stressed vowel movement and intonation contour slope (the less steep the contour, the less falling the vowel, although the variation in vowel movement is slight, see Thorsen 1980b, p. 1020) and (2) the rise in long vowels is greater on horizontal than on falling contours. However, I think the explanation for the mainly falling stressed vowels (on falling contours) is to be found in the way stressed vowels are signalled tonally in ASC Danish, being ones that are jumped (or glided) <u>up</u> from: a fall in the stressed vowel will maximize the perceptual distance to the succeeding higher post-tonic. The less steep falls/greater rises in stressed vowels on marked intonation contours may be accounted for in terms of the greater rise to the post-tonic (i.e. as an assimilation), and thus only indirectly ascribed to the intonation contour proper. - This interpretation is corroborated by a pilot study on some Danish dialects where rising stressed vowels seem to be the rule, also on falling intonation contours: These are dialects where the stress group pattern is one of high stressed plus low (or falling) post-tonic syllable(s), so the vowel movement may be seen again, not as influenced by the intonation contour, but as determined by the stress group pattern with its stipulation that the perceptual distance between (high) stressed and (low) post-tonic syllable(s) be at a maximum.

6. Emphasis for contrast

6.1 Types of emphasis

Mangold (1975) in his very broad, cross-linguistic survey of phonetic emphasis distinguishes between 'emphasis' which "signifies meaning" and 'phonetic emphasis' which is "emphasis as expressed phonetically" (p. 5). Needless to say, I am concerned here with phonetic emphasis - more specifically, with emphasis for contrast as expressed with fundamental frequency. Any further explicitation is made difficult by the lack of a universally agreed upon definition and sub-categorization of 'emphasis'.

To Jespersen (1897-99, chapter XXV) stress is synonymous with voice effort, 'expiratory accent', not to be confounded with pitch (p. 555).[1] Three factors determine, according to Jespersen, the way people choose to stress one syllable rather than another:

---

1) This is what Jespersen says in the chapter on stress. In the chapter on pitch he is a little more explicit (p. 583-84, my translation): "As the first of those factors which influence pitch, we shall treat stress. The stronger we pronounce a syllable or a sound, the higher we tend to make it (...). However, stress and pitch are not to be seen as two things that necessarily accompany each other or even as one and the same thing - a view which has led to many unfortunate and confusing terms like "hochton" and the like and especially "betoning" (i.e. accentuation, NT) and "betonet" (i.e. accentuated, NT), where only stress should be (continued)

tradition (e.g. some languages have free, others have fixed stress placement), psychological factors, and physical/physiological factors (sonority and rhythm). The psychologically conditioned stress is termed _'value stress'_, which is assigned to a word or words to which the speaker attaches special importance. Value stress can be assigned to a whole utterance (the speaker is then said to speak 'emphatically', 'with emphasis') or to a single word. Jespersen distinguishes two kinds of value stress: _novelty stress_ and _contrast stress_ but grants that they are not easily separated. Mangold (1975, p. 16) objects to Jespersen's 'value stress' because "... the meaning 'value' as described by Jespersen may be phonetically expressed through stress only, and not through other phonetic devices such as high pitch, length. Value emphasis thus defined seems to include any type of emphasis which will be treated here, maybe with the exception of predicate emphasis."

I have quoted Jespersen because his classification and description, although they are applied to several languages, draw heavily on examples from Danish. There are, of course, several other important 'schools' and more sub-divisions of emphasis possible, surveyed in Mangold (1975, p. 16-28). - A common distinction is one between explicit and implicit contrast: If both contrasted terms are contained in the same sentence, clause, or phrase, the contrast is _explicit_. If only one of the terms is mentioned, the contrast is _implicit_, although the contrasted term may be contained in a preceding sentence. According to this dif-

---

(continued from the preceding page)
talked about. Just as in music a low note can be forte or fortissimo and a high note piano or pianissimo, in language a strong (accentuated) syllable can have a low pitch and a weak syllable a high pitch (...). The common tendency to let an increase in stress be accompanied by a pitch heightening can often be neutralized by other circumstances, which involves a low pitch on the strong syllable." - In Modersmålets Fonetik Jespersen notes (p. 141, my translation): "Pitch movements are greatest in stressed syllables (...) but we must not from that be led astray and believe that pitch and stress are identical", and then he goes on in much the same fashion as in the first quotation. It thus seems that the fact that stressed and unstressed syllables do not always and invariably have the same pitch relation to each other (high vs. low, or low vs. high) excludes to Jespersen the possibility of seeing pitch as a perceptual cue to stress. This situation may be due to the fact that Jespersen does not clearly distinguish the linguistic function of and the phonetic cues to stress.

ferentiation, I have investigated implicit contrasts, see further the material below, section 6.4.1.

6.2 Phonetic correlates to emphasis

According to Mangold's (1975) survey of the literature, the phonetic devices which may signal emphasis are stress, pitch, length, pause, tamber. "Most authors mention either stress or pitch or both. Pitch and stress are often not separated, or it is difficult to ascertain how far stress includes also pitch." (p. 29). As we have seen above, Jespersen denies any direct and simple relation between pitch and stress, i.e. neutral as well as emphatic stress. Heger (1975) makes two interesting observations (my translation, NT):[1] "A special kind of stress is called emphatic stress or emphasis. It is the special weight which is attached to a syllable when you wish to emphasize it. Whereas the three aforementioned degrees of stress, main stress, secondary stress, and weak stress, are necessary components in normal neutral speech, emphasis is present only when the speaker wishes to emphasize a syllable or word perceptibly. With emphasis, all means are employed: great voice effort, violent pitch excursions, energetic articulation and lengthening of the sounds (...)" (p. 51-52). Emphasis to Heger is the emotional 'value stress' or 'intensity emphasis' as is clear from his examples and the following quotation (p. 54 - my translation, NT): "However, there is one way in which to employ the three degrees of stress (main, secondary, and weak, NT) which is hard to deduce from the written text ... the so-called contrast stress. Contrast stress consists in emphasizing a word by replacing the main stress on surrounding (chiefly succeeding) words with a secondary stress (...). As shown by the examples, 'stress for comparison' would be an appropriate term for this use of stress; in every case the emphasized word is compared to another (possibly imagined) word. Contrast stress may be accompanied by emphasis, but need not be so. On the other hand, emphasis is often accompanied by that which characterizes contrast stress, so that surrounding main stresses are replaced by secondary stresses." To sum up Heger's statements: Emotional emphasis involves greater voice effort, Fo movements, and duration; contrast may be achieved simply by a reduction of surround-

---

1) Broken underlining mine.

ing stresses; - this last contention corresponds to my own interpretation of the results of the experiments reported below, although I shall make a reservation concerning the distance over which a contrast emphasis will result in stress reduction.

6.3 Emphasis for contrast vs. sentence accent

Heger's first claim (which I believe would be corroborated by any linguist or phonetician who has worked with Danish) that stress degrees above main stress are not characteristic of normal neutral speech reflects a fact about Danish which presumably makes Danish rather different from the languages it is most closely related to: namely that Danish lacks what has been termed primary accent, sentence accent, sentence stress, nuclear accent, focal accent, Satzakzent, etc., in descriptions of English, Swedish, and German. In these languages, one of the stressed syllables will always have slightly greater prominence than the others, realized - very roughly speaking - as a more elaborate Fo movement within or in the environment of that syllable; if nothing else is specified by the context, this accent will be located on the last stressed syllable of the utterance.

The works which, implicitly or explicitly, assume the existence of a nuclear stress in British and American English are too numerous to permit an exhaustive listing. Crystal (1969) reviews past work on prosodic features and says himself (p. 207): "There is general agreement about the internal structure of the tone-unit in English. Minimally, a tone-unit must consist of a syllable, and this syllable must carry a glide of a particular kind. This is the obligatory element, and is usually referred to (in the British tradition) as the _nucleus_ of the tone-unit. (...) The presence of a nucleus is what accounts for our intuition of 'completeness' at the end of the unit: if it is omitted, the auditory effect is one of 'being cut short'." - Liberman and Prince (1977) - in a theoretical frame-work which differs from most previous descriptions of English stress - also assume a "main stress" which is the most prominent terminal element of a given constituent and is termed a _designated terminal element_ (p. 257). Bolinger (1958) seemingly takes a slightly different approach, in limiting the term 'stress' to the domain of the word, and talks of 'pitch accents' at the sentence level. However, I think that the difference is terminological rather than conceptual, inasmuch as Bolinger does not deny the existence of stresses that are more prominent (by their pitch movements) than others.

German too has a Satzakzent, cf. von Essen (1956) and Stock (1980, p. 79-80).

Focal accent is undoubtedly a reality in Swedish, see e.g. Bruce (1977) and Gårding (1980, p. 285) but according to Carlson et al. (1974) prosodically neutral utterances, i.e. utterances with a "non-contrastive (focus-free) stress pattern" (p. 212) are possible. (Note that Carlson et al. seemingly equate contrast and focus, see further below.)

In Dutch a special sentence accent is not distinguished from other accents (see e.g. 't Hart and Collier 1979). One or more of the lexical accents in an utterance will be manifested as pitch accents which apparently are all equally prominent. The rules which govern the assignment of these pitch accents are the object of a study by Terken (1980).

Now, what is it about Danish that makes it deviant from English, German, and Swedish (we will leave Dutch aside)? First of all, there is nothing "incomplete" about an utterance with a neutral prosodic pattern, i.e. one where none of the stressed syllables is more prominent than the others. Such neutral utterances may not be very common in spontaneous speech but they do occur, they are not unnatural, and they are very easy to elicit from speakers in a reading situation. Secondly, and this is the crucial point: In English, German, and Swedish the sentence accent does <u>not</u> imply, or result in, a deletion of the surrounding stress group patterns (i.e. a reduction of main to secondary stress) the way that emphasis for contrast does in Danish (cf. the quotation from Heger 1975 above and see the results below). Thus, to sum up: in neutral utterances in English, Swedish and German four degrees of stress are necessary to account adequately for the distribution of and relation between stresses, whereas three relational degrees are sufficient in Danish: Main stress (assigned to the lexically stressed syllable in most non-function, or non-empty, words, secondary stress (assigned to the second (and following) lexically stressed syllable in compounds), and weak stress, see further Basbøll (1978) and Rischel (1972, 1975). If we accept stress as a hierarchical feature which describes <u>relations between</u> syllables rather than a (absolute) <u>property of</u> syllables, or vowels (as suggested by Fischer-Jørgensen 1948, Rischel 1964 and 1972, and later by Liberman and Prince 1977, and Selkirk (forthcoming)), then the same three degrees will suffice in short[1] utterances with contrast: The stressed syllable of the contrasted word will retain its main stress, and the surrounding stressed syllables will be reduced to secondary stresses, i.e. they will retain all of their stressed syllable properties (vowel quantity, quality and stød, if any) except the tonal one. Precisely because the difference between main and secondary stresses in prosodically neutral utterances is primarily tonal and because the effect of a contrast emphasis is mainly tonal (cf. below), I see no reason why a distinction between secondary stresses which are the result of a reduction due to a contrast emphasis and 'original' secondary stresses should be called for - presumably both types will be secondary stresses on a par with each other.

Presumably, in languages with sentence accent, emphasis for contrast exists independently. From Jones (1960, §§ 1049-1059) we learn that when a sentence accent (finally in the utterance) is superposed by a contrast emphasis its Fo movement is even more elaborate. With contrast emphasis on some earlier word this Fo movement is moved back to that word's stressed syllable and succeeding syllables "have the intonation of unstressed syllables" (§1050). Now, we can conclude either that the sentence accent is deleted due to the contrast emphasis earlier in the utterance or that it is moved back to coincide with the contrast emphasis. The decision is an arbitrary one, because whichever we choose, the fact remains that we can distinguish between a neutral sentence accent and contrast emphasis finally in the utterance but not in other positions where sentence accents do not occur in neutral speech. The implication of this is not, of course, that all sentence accents are really contrastive emphases in disguise: As long as, e.g., an English utterance, in order to be natural and complete, must have a special prominence attached to its final content word, which can be distinguished from a contrast emphasis, it is justified to make a distinction between the two (which does not preclude that they are related, semantically and phonetically).

Gösta Bruce, in a personal communication has informed me that the Skåne dialect of Southern Swedish is not characterized by having a sentence accent, so ASC Danish is not the only Nordic

---

1) It does seem, however, that the length of the utterance is decisive, see further section 6.4.2, in particular 6.4.2.4.

language/dialect without it. (Whether other Danish dialects will turn up with a focal accent is still an open question.) - As far as I read the numerous works of Cohen, Collier, and 't Hart, sentence accent is not a very useful concept in the description of Dutch either.

Before proceeding to an account of the results of the experiments I wish to stress the fact that the material analysed is emotionally and attitudinally neutral, i.e. the contrast emphasis elicited by the speakers does not, and was not intended to, contain any 'intensity emphasis' or the like. - Thus, I do not claim that this is the only type of phonetic emphasis possible, or even that contrast emphasis may not be realized phonetically in a different manner - but I do claim that what I describe here is fairly normal speech.

6.4 Acoustic analysis of emphasis for contrast in ASC Danish

6.4.1 Material, subjects, and procedures

The same type of material as was used for the analysis which led to the postulation of the model in fig. 1 served here, i.e. nonsense words embedded in initial, medial, and final position in short carrier sentences, which are as much alike semantically, syntactically, and rhythmically as possible. (For a discussion of the choice of words, see Thorsen 1978.) The sentences are listed below (they translate as follows: '_______ has shorter syllables', 'The syllables of _______ are shortened', 'There are shorter syllables in _______', respectively):

1	'pipi giver kortere stavelser.	S1
2	'pipipi giver kortere stavelser.	
3	pi'pipi giver kortere stavelser.	
4	pipi'pi giver kortere stavelser.	
5	'pipi giver kortere stavelser?	Q1
6	'pipipi giver kortere stavelser?	
7	pi'pipi giver kortere stavelser?	

8	*'pipipi* giver kortere stavelser.	
9	*pi'pipi* giver kortere stavelser.	S1A
10	*pipi'pi* giver kortere stavelser.	
11	'pipipi giver *kortere* stavelser.	S1B
12	pi'pipi giver *kortere* stavelser.	
13	pi'pipi giver kortere *stavelser*.	S1C
14	*pi'pipi* giver kortere stavelser?	Q1A
15	pi'pipi giver *kortere* stavelser?	Q1B
16	pi'pipi giver kortere *stavelser*?	Q1C
17	Stavelserne i 'pipi forkortes.	
18	Stavelserne i pi'pi forkortes.	
19	Stavelserne i 'pipipi forkortes.	S2
20	Stavelserne i pi'pipi forkortes.	
21	Stavelserne i pipi'pi forkortes.	
22	Stavelserne i 'pipi forkortes?	
23	Stavelserne i pi'pi forkortes?	Q2
24	Stavelserne i pipi'pi forkortes?	
25	Stavelserne i *'pipipi* forkortes.	
26	Stavelserne i *pi'pipi* forkortes.	S2B
27	Stavelserne i *pipi'pi* forkortes.	
28	*Stavelserne* i pi'pipi forkortes.	S2A
29	Stavelserne i pi'pipi *forkortes*.	S2C
30	Stavelserne i *pi'pipi* forkortes?	Q2B
31	*Stavelserne* i pi'pipi forkortes?	Q2A
32	Stavelserne i pi'pipi *forkortes*?	Q2C

33	Det giver kortere stavelser med 'pipi.	S3
34	Det giver kortere stavelser med 'pipipi.	
35	Det giver kortere stavelser med pi'pipi.	
36	Det giver kortere stavelser med pipi'pi.	
37	Det giver kortere stavelser med 'pipi?	Q3
38	Det giver kortere stavelser med 'pipipi?	
39	Det giver kortere stavelser med pi'pipi?	
40	Det giver kortere stavelser med 'pipipi.	S3C
41	Det giver kortere stavelser med pi'pipi.	
42	Det giver kortere stavelser med pipi'pi	
43	Det giver kortere stavelser med pi'pipi.	S3A
44	Det giver kortere stavelser med pi'pipi.	S3B
45	Det giver kortere stavelser med 'pipipi?	Q3C
46	Det giver kortere stavelser med pi'pipi?	
47	Det giver kortere stavelser med pi'pipi?	Q3A
48	Det giver kortere stavelser med pi'pipi?	Q3B

(An indication of stress placement is unavoidable in the nonsense words and might have created a problem with naive subjects but for the four phoneticians who read the material this does not seem to be an obstacle for a natural rendering of the utterances. In the reading lists the emphasized words were underlined in one part of the material, italicized in another part - again this turned out to be a straightforward procedure and did not make for exaggerations of any kind.)

The code for the different sentences reads as follows:

S = statement[1]
Q = question
1 = the nonsense word is initial in the sentence
2 = the nonsense word is medial in the sentence
3 = the nonsense word is final in the sentence
A = emphasis is on the first content word in the sentence
B = emphasis is on the second content word in the sentence
C = emphasis is on the third content word in the sentence

Thus, "Q2A" is a question with the nonsense word in medial position where emphasis is on the first content word.

---

1) In all of the following account, when I talk about questions it is always this particular kind of question, i.e. syntactically unmarked (but prosodically marked) questions.

The prosodically neutral statements and questions were intended for a control and quantification of the stress group patterns (section 7) but they of course also serve as a frame of reference for the utterances with contrast.

Note that the number of different nonsense words is not the same in all sentence types: The problems that I wanted the material to solve were weighed carefully against economy, i.e. the amount of material that subjects were to record and, sure enough, in a few instances I was too stingy - on the other hand, a couple of omissions would not have invalidated the results but that really only became apparent when the data were analysed, see further section 7.

Although experiments by Liberman and Streeter (1978) and Nakatani and Schaffer (1978) indicate that nonsense words retain the prosodic features of natural words, we cannot entirely exclude the possibility that they will exhibit exaggerated prosodic patterns when subjected to emphasis for contrast, so the following sentences were also recorded ('There are many buses out of Tiflis.'):

49 Der går mange busser fra Tiflis.

50 Der går mange busser fra Tiflis.

51 Der går mange busser fra Tiflis.

52 Der går mange busser fra Tiflis.

The prosodically neutral utterances (S1,2,3 - Q1,2,3) were "naked", i.e. they were not embedded in a context but they were mixed with sentences where the tonal assimilation in schwa-syllables was at stake (22 pipi-sentences interspersed in 75 schwa-sentences). All the other (30) sentences were embedded in small dialogues, like the following:

A: Sorry, what did you say? - Do the buses leave from Tiflis or from Grosny?

B: There are many buses out of Tiflis. - As far as I can see there is no connection from Grosny at all.

A: pi'pipi giver kortere stavelser.

B: pi'pipi giver kortere stavelser?

A: Yes - that's reasonable enough. It is a three syllable word.

A: Does pi'pipi have shorter syllables or just shorter vowels?

B: pi'pipi has shorter syllables.

(Each subject would simultaneously take the role of both A and B.) One part of these dialogues was mixed with the neutral pipi- and schwa-sentences (yielding a total of six pages of reading material), another part was recorded at a later stage, mixed with yet another material, yielding two pages of reading material, in three different randomizations.

Four phoneticians, two males (NRP and JR) and two females (BH and NT) read the material. NRP, BH, and NT speak Advanced Standard Copenhagen, JR speaks a slightly more conservative variant. NRP, BH, and JR recorded the material six times each, NT (the author) recorded it ten times.

The recordings were made with semi-professional equipment (Revox A-77 tape recorder, Sennheiser MD21 microphone, larynx microphone) in a quasi-damped room at the Institute of Phonetics on Agfa PE36 tape at 7½ i.p.s. The tapes were processed by hardware intensity and pitch meters (F-J Electronics) and registered on a mingograph (Elema 800) at a paper speed of 100 mm/s. 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 allowing for a measuring accuracy of 1 Hz for males and 2 Hz for females.

In unidirectional Fo courses with constant slope only the beginning and end points were measured, according to a procedure outlined in Thorsen (1979, p. 63-66). 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 distance in time of each measuring point from the first one was measured, and so was the duration of each segment, to the extent that reliable segmentation could be made. Fo and time measurements were averaged over the six (ten) recordings by each subject: the average Fo values were converted to semitones (re 100 Hz) and average tracings drawn. No correction was made for intrinsic Fo level differences between vowels.

The standard deviations on the average Fo and duration values are generally small, about 3% to 5% of the mean and very rarely exceeding 10% (they are generally smaller for stressed than post-tonic vowels/syllabic consonants), so production stability across different readings is rather great and the figures to follow must be fairly reliable indications of subjects' behaviour.

6.4.2 Results

In figs. 2-5 average tracings of every type of sentence ({S,Q}{1,2,3}{A,B,C}) are shown for each subject. The nonsense word in the sentences is [b̥ʰiˈb̥ʰib̥ʰi] (which is the only one that occurs in all positions and conditions), except in final position in S3, S3C, Q3, Q3C where it is [ˈb̥ʰib̥ʰib̥ʰi] (in order that the final stress group be comparable to the initial and medial ones, that both have two post-tonic syllables, the second post-tonic being 'giver' [g̊iʌ̆] and 'for-' [fʌ], respectively). The neutral edition (full line) is compared to every emphatic condition (dotted line) in statements (left) and questions (right), for each position of the nonsense word (1: initial, 2: medial, 3: final). A large part of the material by NRP had to be discarded (all the emphatic questions and those emphatic statements where the emphasis occurs on the surroundings to the nonsense word. - This was the part that was recorded in the second round, and apparently NRP had become too conscious of the purpose of the investigation and no longer kept up the emotional neutrality of the first set of recordings). Thus, some of the averages across subjects in the figures to follow are averages over three subjects only (this will be clearly stated in the legends).

Apparently, there are minor differences between subjects (voice range is an obvious one). However, it would be - if not exactly impossible - tedious to give a full verbal description of each of the four subjects' tracings. Table 1 gives a simplified account in terms of plusses and minusses which indicate whether a

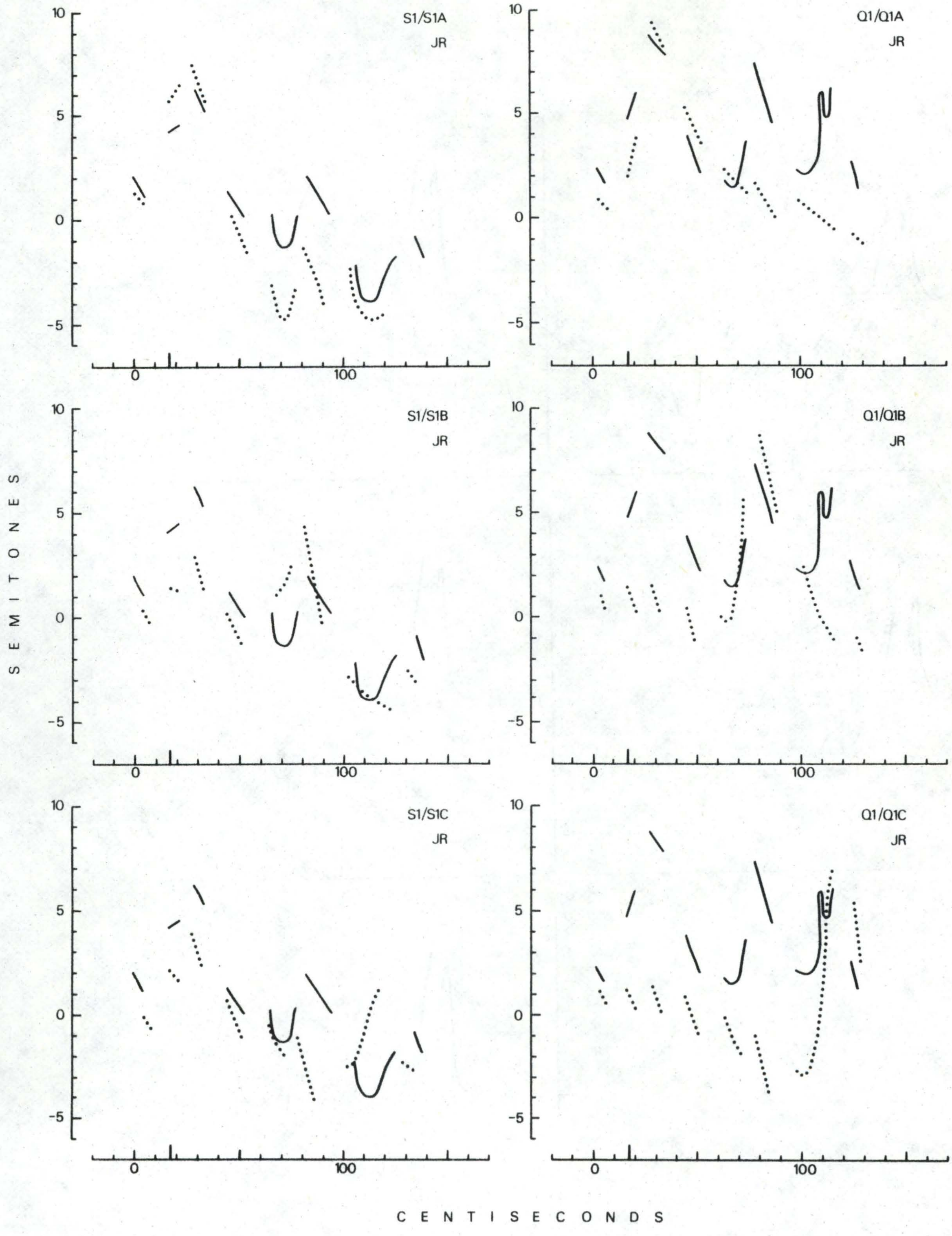


Figure 2-1

Fundamental frequency tracings of "pi'pipi giver kortere stavelser" (average of six readings): statements (S - left) and questions (Q - right), prosodically neutral (full lines) and with emphasis for contrast (dotted lines) in the first stress group (A - top), in the second stress group (B - mid), and in the third stress group (C - bottom) in the utterance. The tracings have been lined up (heavy stroke on the time scale) according to the beginning of the stressed vowel in the nonsense word. Zero on the logarithmic frequency scale corresponds to 100 Hz. - Subject JR.

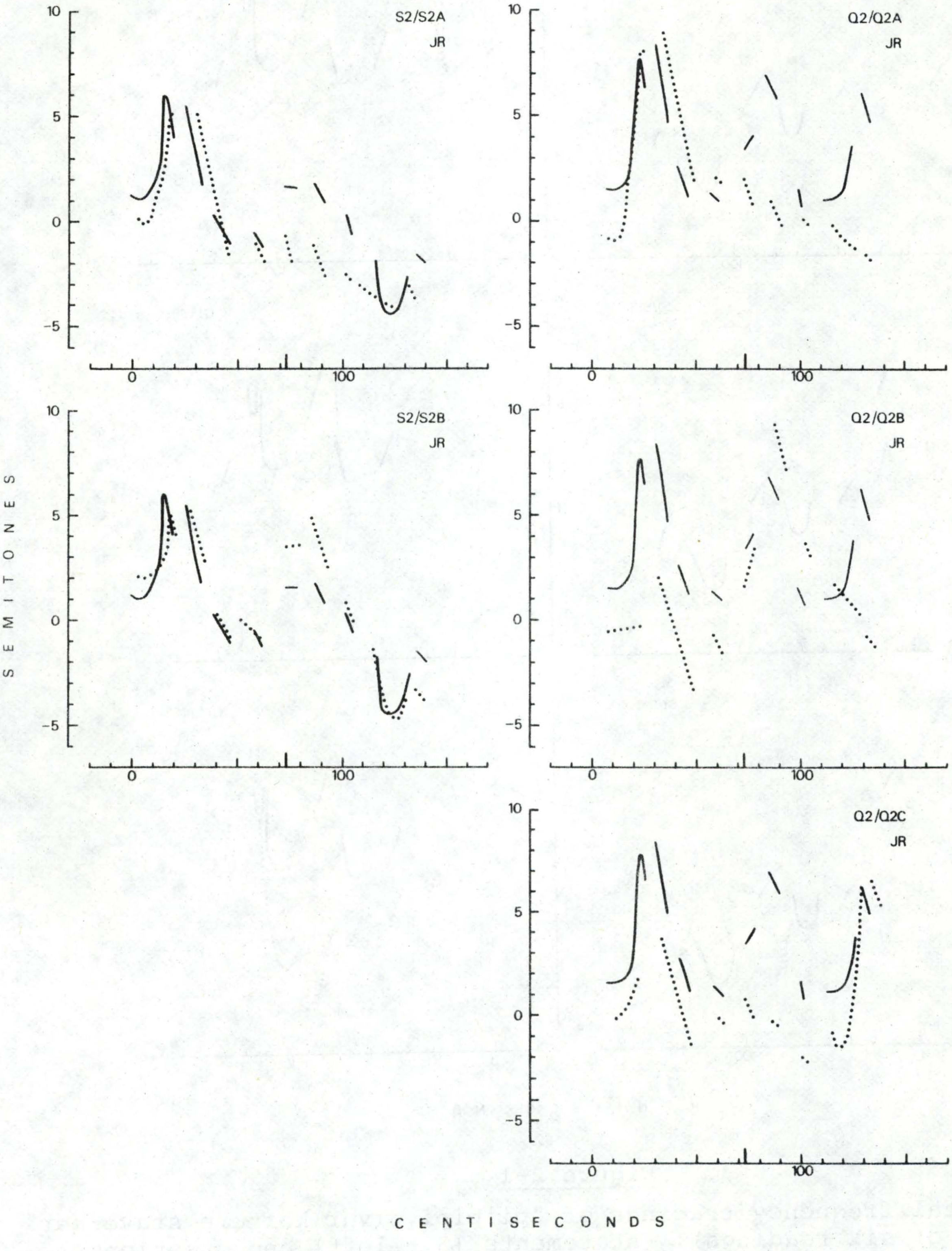


Figure 2-2

"Stavelserne i pi'pipi forkortes" Subject JR. See further the legend to fig. 2-1.

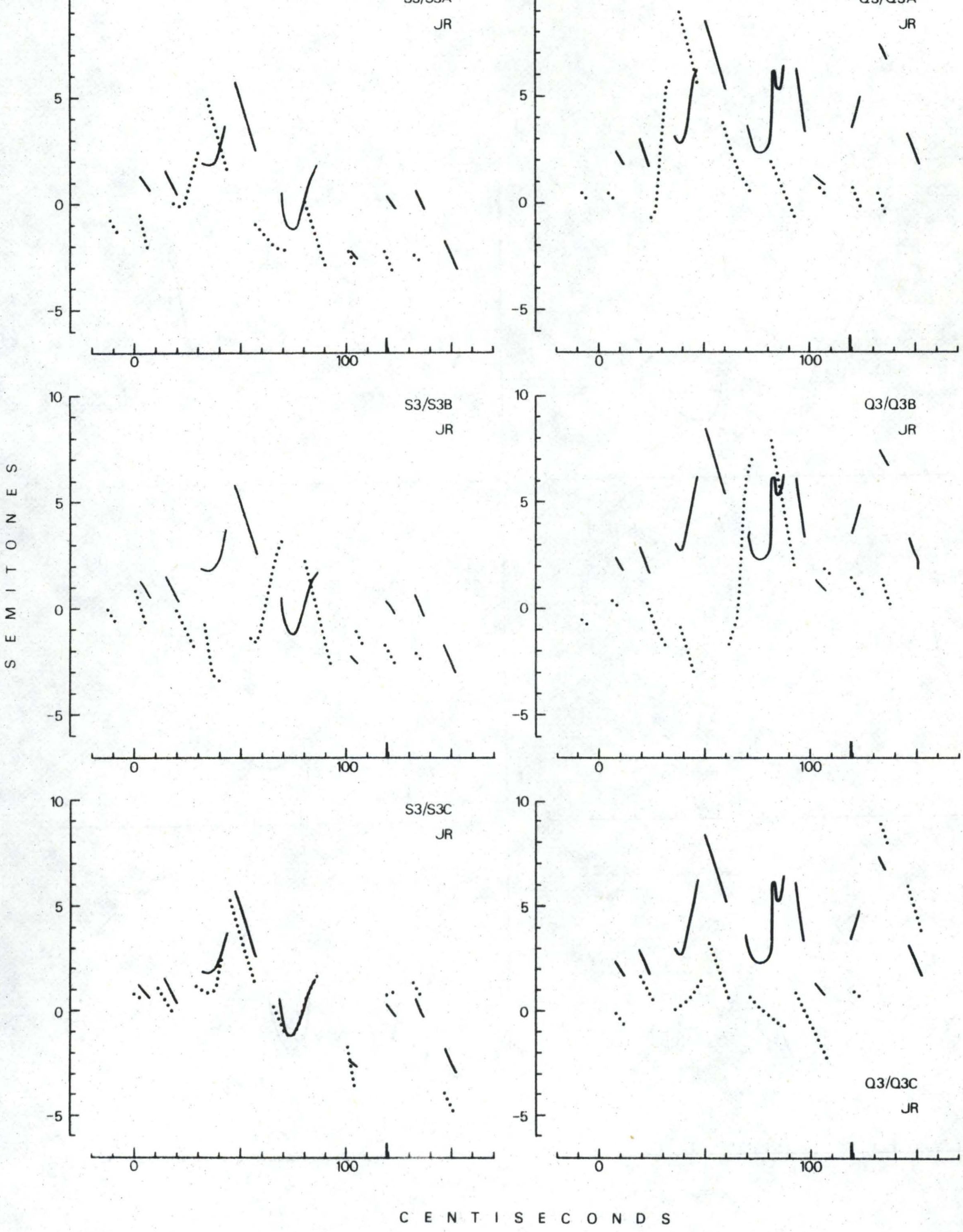


Figure 2-3

"Det giver kortere stavelser med 'pipipi" (S3, Q3, S3C, Q3C)
"Det giver kortere stavelser med pi'pipi" (S3A, S3B, Q3A, Q3B)
Subject JR. See further the legend to fig. 2-1.

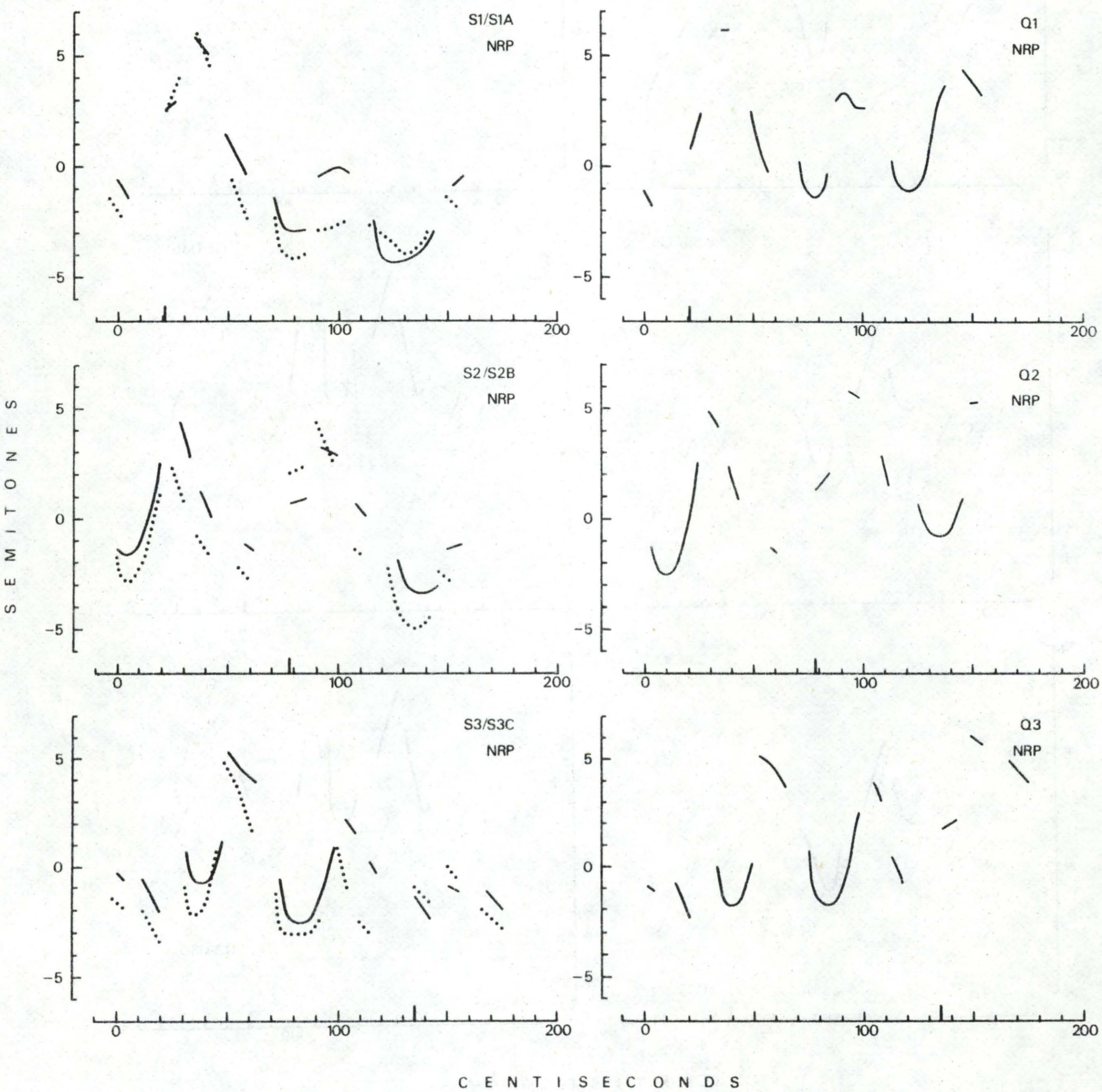
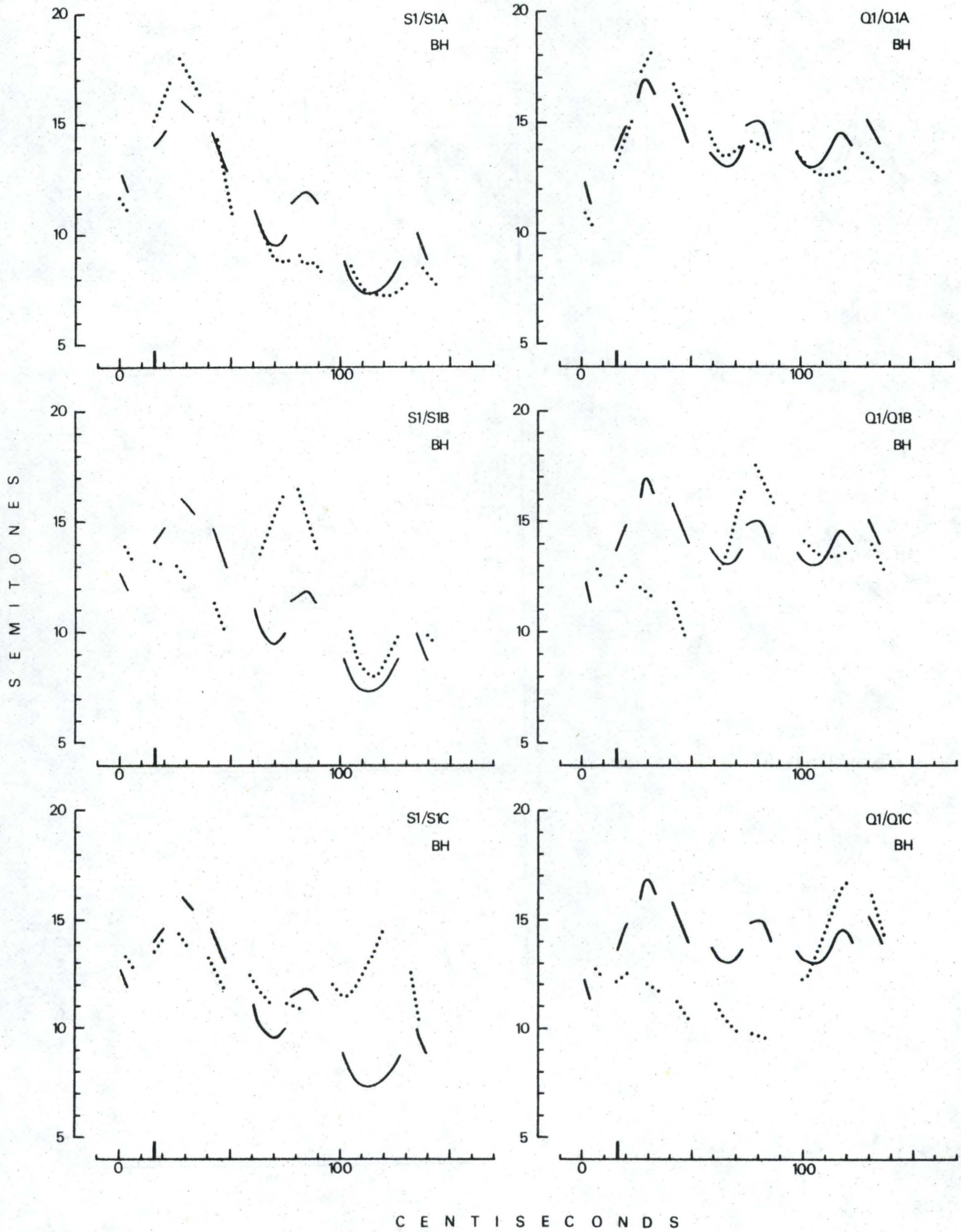


Figure 3

"pi'pipi giver kortere stavelser" (top)
"Stavelserne i pi'pipi forkortes" (mid)
"Det giver kortere stavelser med 'pipipi (bottom)
Subject NRP. See further the legend to fig. 2-1.



<u>Figure 4-1</u>

"pi'pipi giver kortere stavelser" Subject BH. See further the legend to fig. 2-1.

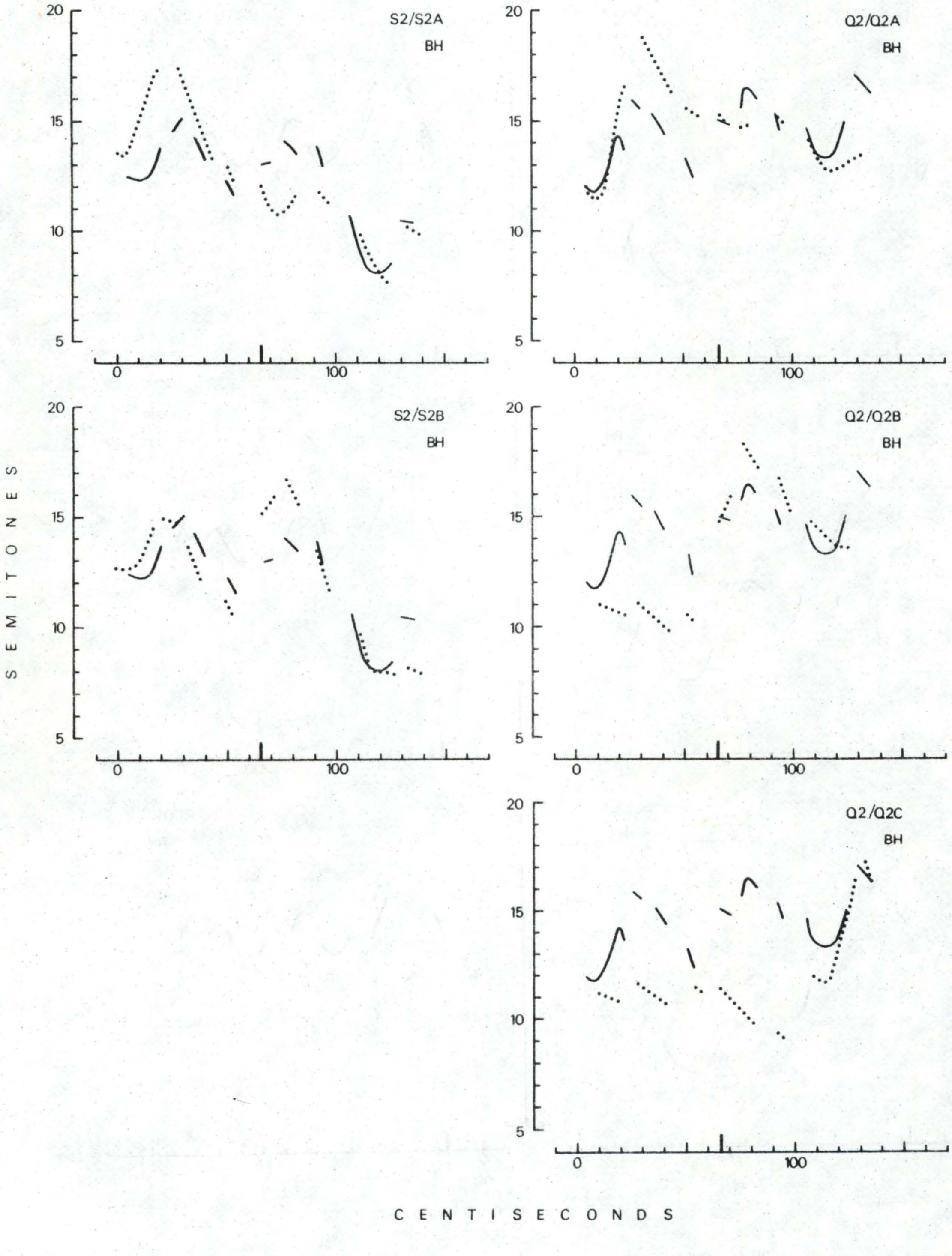


Figure 4-2

"Stavelserne i pi'pipi forkortes" Subject BH. See further the legend to fig. 2-1.

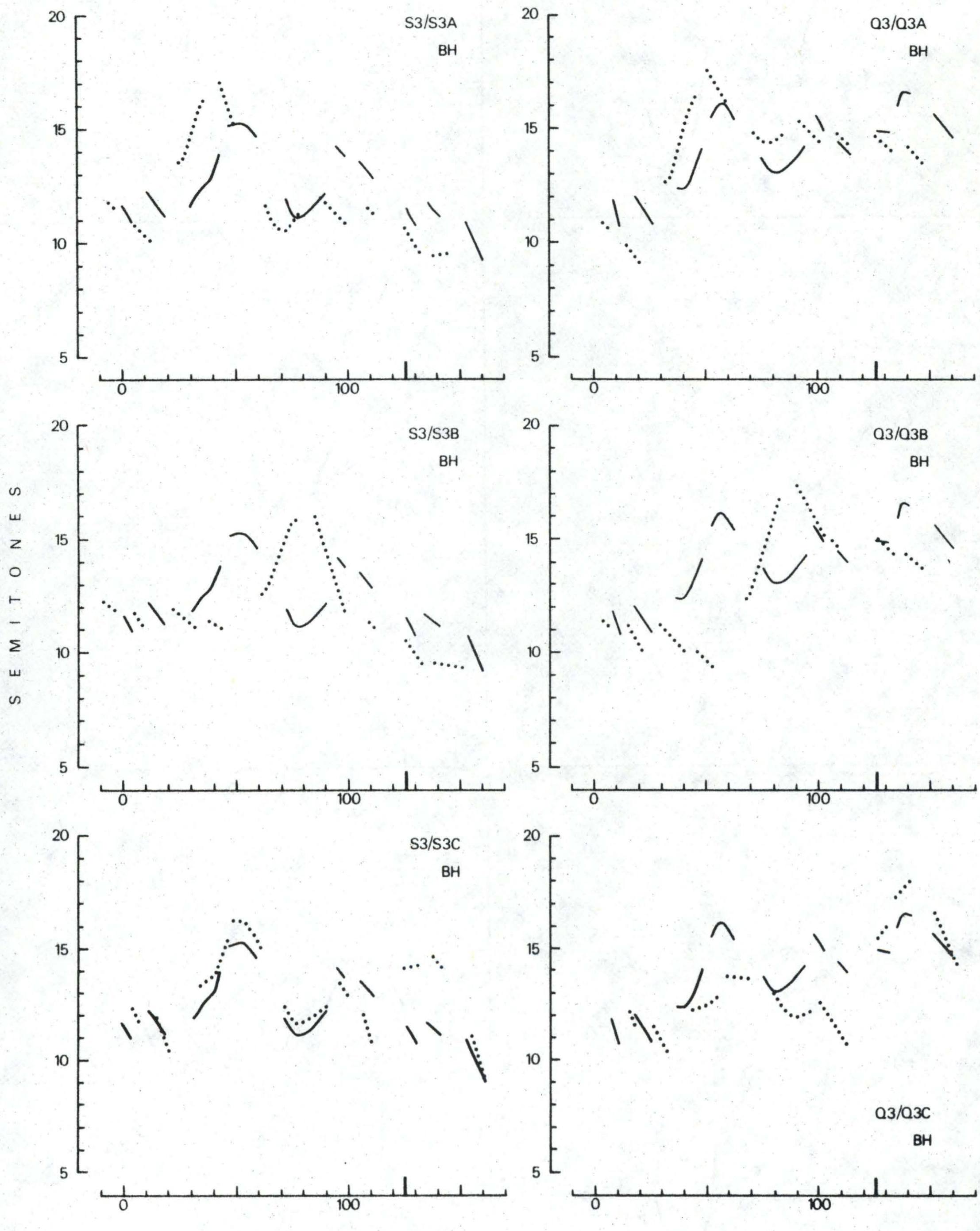


Figure 4-3

"Det giver kortere stavelser med 'pipipi" (S3, Q3, S3C, Q3C)
"Det giver kortere stavelser med pi'pipi" (S3A, S3B, Q3A, Q3B)
Subject BH. See further the legend to fig. 2-1.

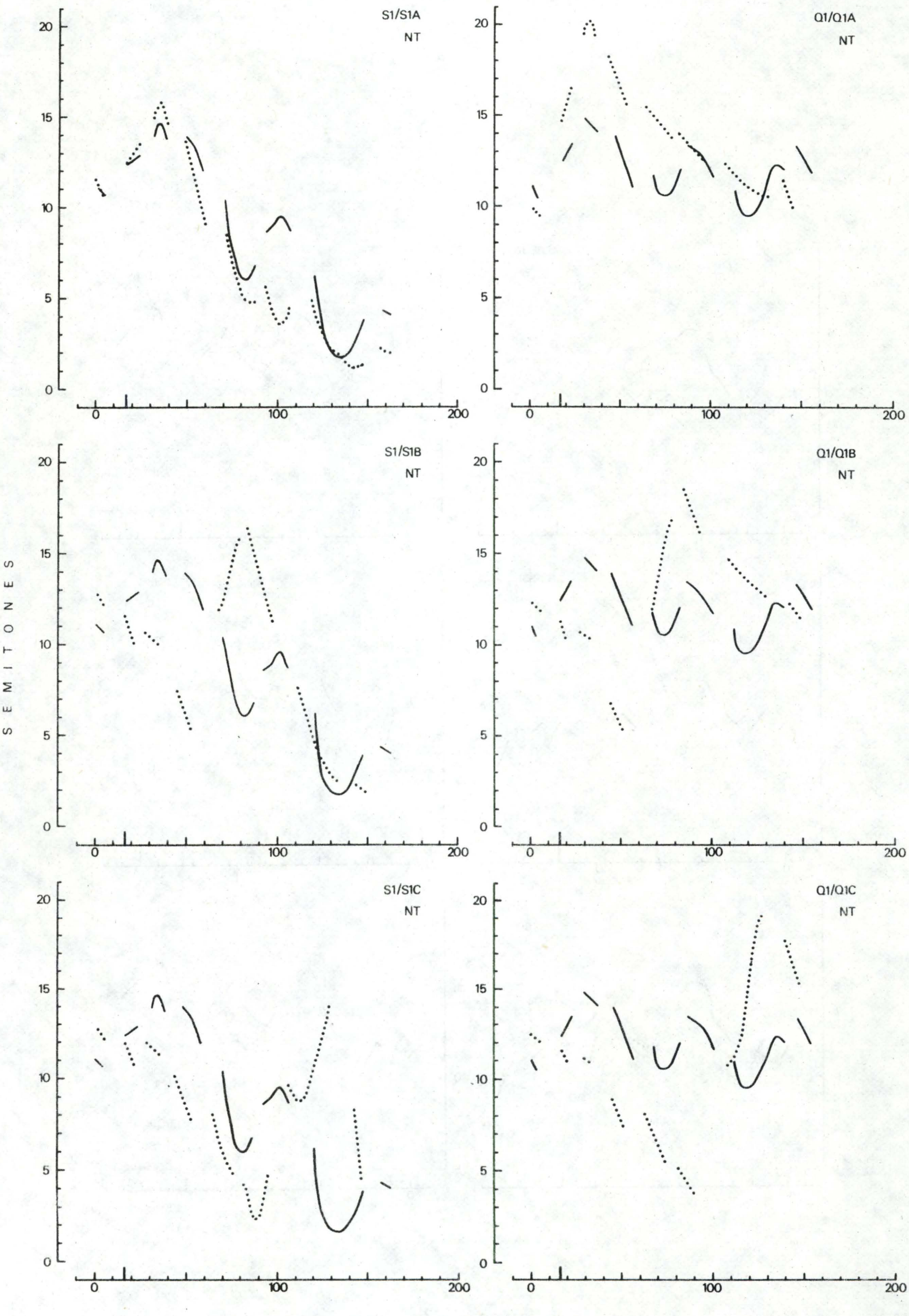


Figure 5-1

"pi'pipi giver kortere stavelser" Average of ten readings. Subject NT. See further the legend to fig. 2-1.

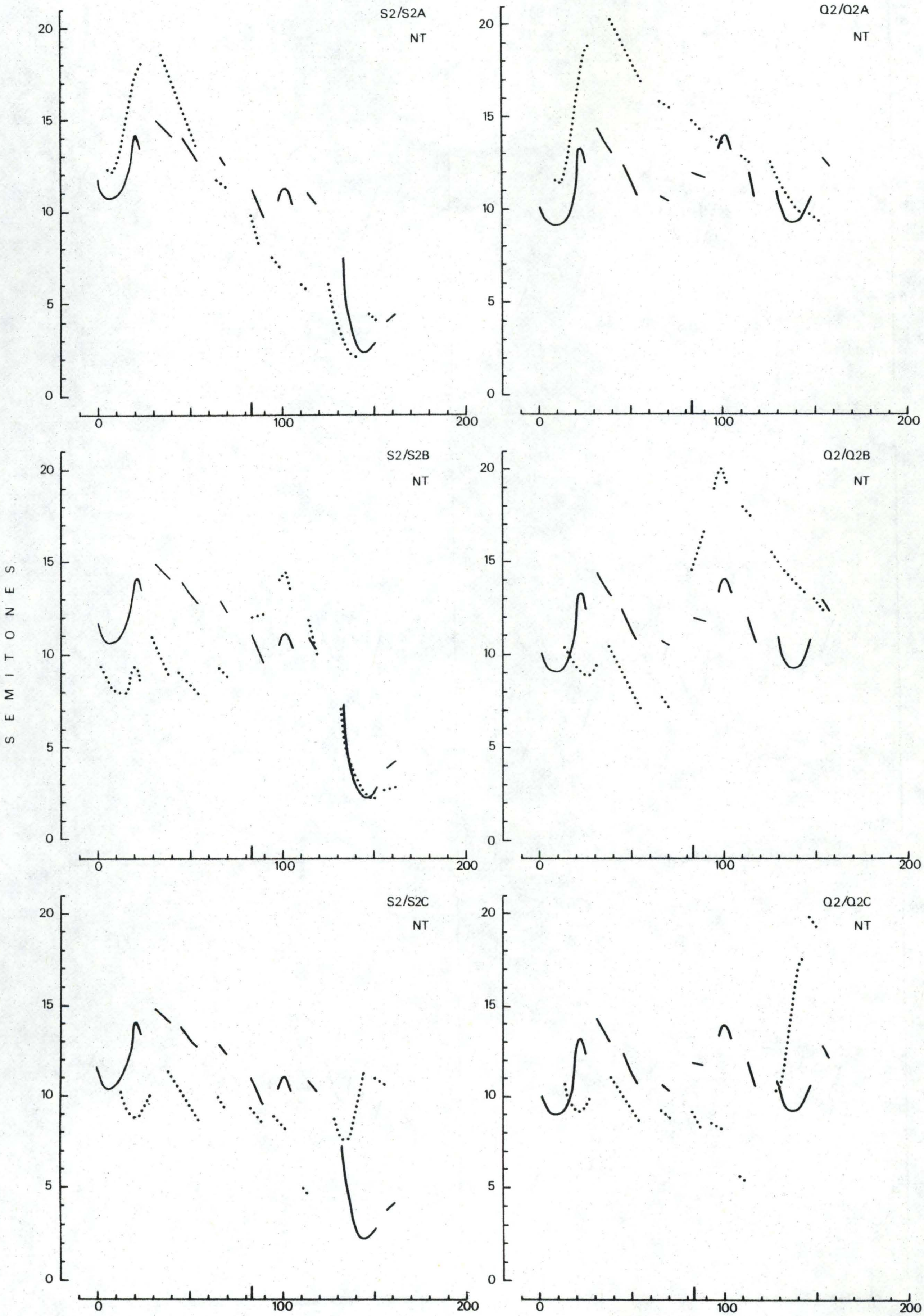


Figure 5-2
"Stavelserne i pi'pipi forkortes" Average of ten readings. Subject NT. See further the legend to fig. 2-1.

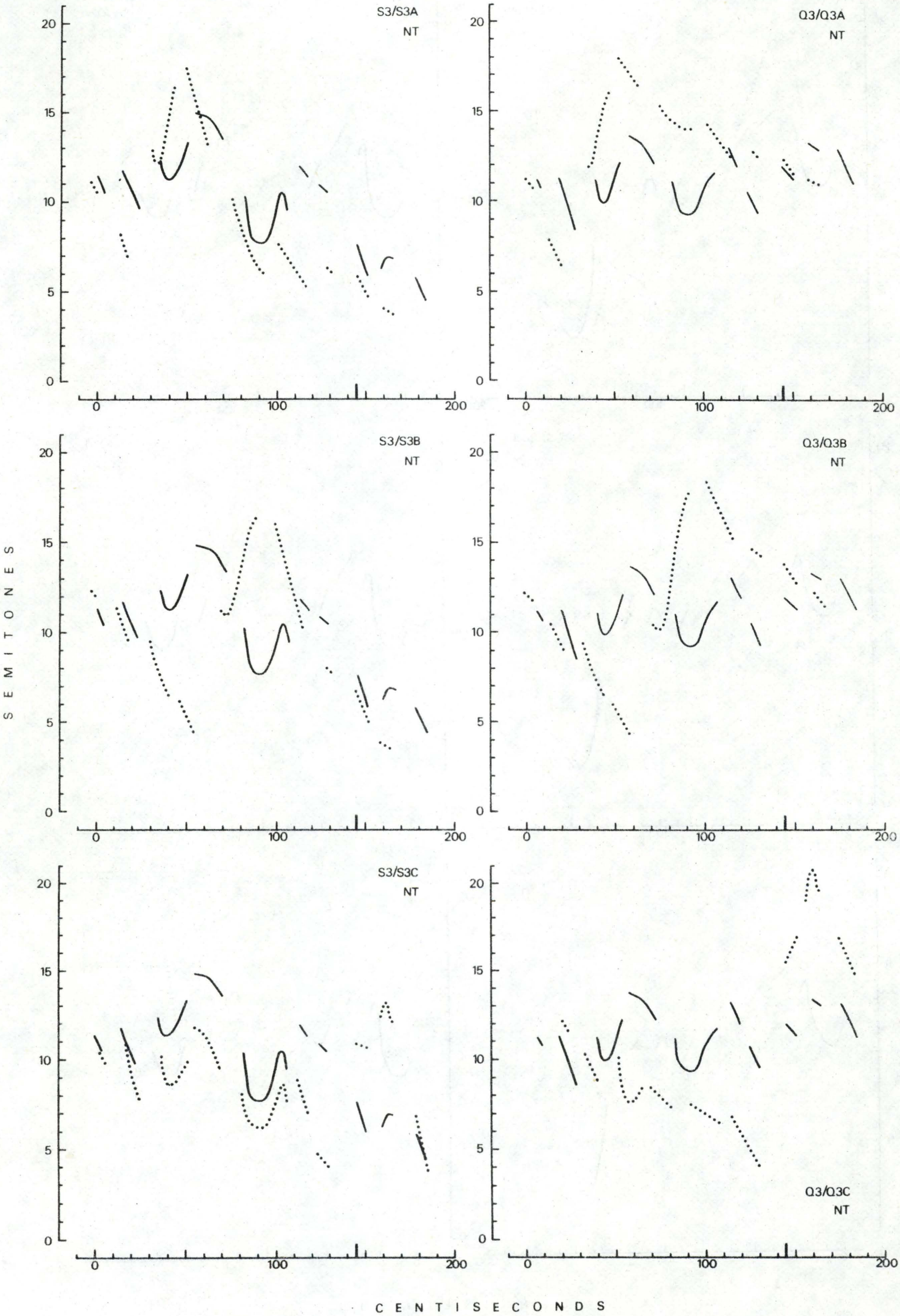


Figure 5-3

"Det giver kortere stavelser med 'pipipi" (S3, Q3, S3C, Q3C)
"Det giver kortere stavelser med pi'pipi" (S3A, S3B, Q3A, Q3B)
Average of ten readings. Subject NT. See further fig. 2-1.

given feature is present or not <u>relative to</u> the neutral edition - with each subject, in each of the sentences with emphasis for contrast. Of course, the assignment of plus or minus in some cases is a matter of opinion and in one type of feature a further differentiation seemed justified: When a stress group in the environment of the emphasized[1] one contains a partial rise (compared to the full rise in the neutral case) this is indicated by "½". Plusses and minusses pertaining to the nonsense words are framed in full lines, 'stavelser-' in broken lines, and 'kortere/-kortes' in dotted lines. Table 2 summarizes the information in table 1: It indicates the number of plusses for a given feature as a percentage of the total number of observations, for nonsense and real words separately, as well as for both types of words together. - On the whole, inter subject differences seem to be random and quantitative rather than qualitative and the generalizations in the following will be based on the average tracings across all subjects. Individuals' relation to these generalizations will be noted with reference to table 1.

6.4.2.1 Nonsense words

Fig. 6 is a highly stylized tracing, averaged over three subjects, where each syllable is depicted as a point only (to facilitate a comparison with fig. 1; intravocalic Fo movements in the nonsense words are accounted for in section 7.2). It is pieced together of the "nonsense stress groups" from different sentences, including the unstressed word/syllable 'giver' and 'for-' when necessary to obtain the required number of post-tonics. Thus, the tracings in fig. 6 (and 7 - which compares emphatic statements and questions directly) represent a long succession of pi-syllables from sentences as indicated below (note that a stress group commences with the stressed syllable and thus pretonic pi-syllables are <u>not</u> included):

---

1) When, in the following, I talk of "emphatic/emphasized syllables/vowels" it is shorthand for "the stressed syllable/vowel of the emphasized word", and "emphatic/emphasized stress group" is short for "the stress group which contains the stressed syllable of the emphasized word".

Table 1

Summary of the differences to be seen in figs. 2-5 in the course of fundamental frequency in statements (S) and questions (Q) with emphasis for contrast <u>relative to</u> prosodically neutral utterances. "1, 2, 3" indicate the position of the nonsense word in the sentences (initial, medial, final), "A, B, C" indicate the position of the emphasized word (initial, medial, final), see further section 6.4.1. "+" indicates that the difference is present, "-" that it is absent, "½" indicates a partial rise. Plusses and minusses pertaining to the nonsense words are framed in full lines, 'stavelser-' in broken lines, and 'kortere/-kortes' in dotted lines.

The abbreviations characterizing the various features read as follows:

Emph.syll.high.	=	The emphasized syllable is higher
Emph.syll.ris.	=	The emphasized syllable is more rising
Post-emph.ris.high.	=	The rise from the emphasized syllable to the post-tonic is higher
Post-emph.fall steep.	=	The fall through the post-tonic syllables after the emphasized one is steeper
Prec.st.gr.low.	=	The stress group immediately preceding the emphasized syllable is lower
Prec.post-ton.fall steep.	=	The post-tonics in the immediately preceding stress group fall more steeply
Prec.st.gr.ris.	=	The stress group immediately preceding the emphasized syllable has a rise
2.prec.st.gr.low.	=	The second stress group before the emphasized syllable is lower
2.prec.post-ton.fall steep.	=	The post-tonics in the second stress group before the emphasized syllable fall more steeply
2.prec.post-ton.ris.	=	The second stress group before the emphasized syllable has a rise
Succ.st.gr.low.	=	The stress group immediately succeeding the emphasized syllable is lower
Succ.st.gr.ris.	=	The stress group immediately succeeding the emphasized syllable has a rise
2.succ.st.gr.low.	=	The second stress group after the emphasized syllable is lower
2.succ.st.gr.ris.	=	The second stress group after the emphasized syllable has a rise

	S1A				S1B			S1C			S2A			S2B				S2C	S3A			S3B			S3C				no. of +/½ with the	total in parenthesis
	NRP	JR	BH	NT	JR	BH	NT	JR	BH	NT	JR	BH	NT	NRP	JR	BH	NT	NT	JR	BH	NT	JR	BH	NT	NRP	JR	BH	NT		
Emph.syll.high.	−	+	+	−	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−	+	+	−	+	+	−	+	+	+	22	(28)
Emph.syll.ris.[1]	+	+	+	+	+	+	+	+	+	+	−	+	+	−	−	+	+	+	+	−	+	+	+	+	−	−	+	+	22	(28)
Post-emph.ris. high.[1]	−	−	−	−	−	−	−	+	+	+	−	+	+	−	−	−	+	−	−	−	−	+	+	+	−	−	−	+	10	(28)
Post-emph.fall steep.	+	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+		−	−	+	−	+	+	+	+	+	+	23	(27)
Prec.st.gr.low.					+	+	+	−	−	+				+	−	−	+	+				+	+	+	+	−	−	+	12	(18)
Prec.post-ton.fall steep.[2]					−	+	+	+	−	−				−	−	+	−	+				−	−	−	+	+	+	+	9	(18)
Prec.st.gr.ris.					½	−	−	−	−	−				+	+	+	½	−				−	−	−	+	+	½	½	5/4	(18)
2.prec.st.gr.low.								+	−	+								+							+	+	−	+	6	(8)
2.prec.post-ton.fall steep.[2]								−	−	+								−							+	−	−	−	2	(8)
2.prec.post-ton.ris.								+	−	−								½							+	+	+	+	5/1	(8)
Succ.st.gr.low.	+	+	+	+	−	+	−				+	+	+	+	−	−	−		+	+	+	+	+	+					15	(20)
Succ.st.gr.ris.	½	+	−	−	+	½	−				−	−	−	+	½	−	−		½	−	−	−	−	−					3/4	(20)
2.succ.st.gr.low.	+	+	−	−							−	−	−						+	+	+								5	(10)
2.succ.st.gr.ris.	+		½	½							½	+	+						−	−	−								3/3	(9)

1) In the case of 'stavelser-' it is not possible to determine whether it is [ˈæ:] or [w|] which is the more rising.

2) A minus may indicate either that all syllables are low and level or that they continue a slight fall in the preceding "stressed" vowel.

(continued)

(Table 1 continued)

	Q1A J B N R H T	Q1B J B N R H T	Q1C J B N R H T	Q2A J B N R H T	Q2B J B N R H T	Q2C J B N R H T	Q3A J B N R H T	Q3B J B N R H T	Q3C J B N R H T	no. of +/½ with the total in parenthesis	
Emph.syll.high.	− − +	− − +	− − +	− − +	− − +	− − +	− − +	− − +	− + +	10	(27)
Emph.syll.ris.[1]	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	− + +	26	(27)
Post-emph.ris.high.[1]	+ + +	− − −	+ + +	+ + +	+ + +	− − −	− − −	+ + +	+ + +	18	(27)
Post-emph.fall steep.	− − −	− − −	− − +	− − −	− − −		− − −	+ + −	− + +	5	(24)
Prec.st.gr.low.		+ + +	+ + +		+ + −	+ + +		+ + +	+ + +	17	(18)
Prec.post-ton.fall steep.[2]		− − +	− − −		− − −	− − +		− − −	− − −	2	(18)
Prec.st.gr.ris.		− − −	− − −		− − −	− − −		− − −	− − −	0	(18)
2.prec.st.gr.low.			+ + +			+ + −			+ − +	7	(9)
2.prec.post-ton.fall steep.[2]			− − −			− − −			− − −	0	(9)
2.prec.post-ton.ris.			− − −			½ − ½			½ ½ −	0/4	(9)
Succ.st.gr.low.	− + −	+ − −		+ − −	− − −		+ − −	+ − −		5	(18)
Succ.st.gr.ris.	− ½ −	− − −		− − −	− − −		− − −	− − −		0/1	(18)
2.succ.st.gr.low.	+ − −			+ − −			+ + −			4	(9)
2.succ.st.gr.ris.	− − −			− − −			− − −			0	(9)

1) In the case of 'stavelser-' it is not possible to determine whether it is [ˈæ:] or [w̩] which is the more rising.

2) A minus may indicate either that all syllables are low and level or that they continue a slight fall in the preceding "stressed" vowel.

Table 2

Summary of the differences to be seen in figs. 2-5 in the course of fundamental frequency in all statements and all questions with emphasis for contrast relative to the prosodically neutral utterances, across all subjects. The number of times a given feature is present is given as a percentage of the total number of observations, for nonsense and real words separately, as well as for both types together, cf. table 1. Rises from a stressed syllable to its post-tonic may be full or only partial, indicated to the left and right of the dash, respectively. For an account of the feature abbreviations, see the caption to table 1.

	STATEMENTS			QUESTIONS		
	nonsense words	real words	total	nonsense words	real words	total
Emph.syll.high.	75	81	79	44	33	37
Emph.syll.ris.[1]	67	88[3]	79	89	100[3]	96
Post-emph.ris.high.[1]	17	50[4]	36	100	50[4]	67
Post-emph.fall steep.	100	73	85	22	20	21
Prec.st.gr.low.	100	57	67	100	92	94
Prec.post-ton.fall steep.[2]	75	43	50	33	0	11
Prec.st.gr.ris.	0/25	36/21	28/22	0/0	0/0	0/0
2.prec.st.gr.low.	67	80	75	100	67	78
2.prec.post-ton.fall steep.[2]	33	20	25	0	0	0
2.prec.post-ton.ris.	33/0	80/20	63/13	0/0	0/67	0/44
Succ.st.gr.low.	100	64	75	33	25	28
Succ.st.gr.ris.	0/0	21/29	15/20	0/0	0/8	0/6
2.succ.st.gr.low.	100	29	50	67	33	44
2.succ.st.gr.ris.	0/0	50/50	33/33	0/0	0/0	0/0

1) In the case of 'stavelser-' it is not possible to determine whether it is [ˈæ:] or [w̩] which is the more rising.

2) A minus (i.e. a low percentage) may indicate that all syllables are low and level or that they continue a slight fall in the preceding "stressed" vowel.

3) All long vowels.

4) All syllabic laterals.

sentence no.	sentence initial syllable	first stress group	second stress group	third stress group
S	3	2	19	34
S-A	9	8	28	43
S-B	12	11	25	44
S-C	13	13	29	40
Q	7	7	22	38
Q-A	14	14	31	47
Q-B	15	15	30	48
Q-C	16	16	32	45

The philosophy behind this piecemeal construction of the course of Fo in fig. 6 (and 7-9) is of course that this is what the utterances would look like had they actually consisted of a long succession of pi-syllables (or, in the case of figs. 8 and 9: syllables with long vowels and sonorant consonants). In reply to an objection that this is a dubious procedure I would like to call attention to the rather amazing regularities exhibited by the prosodically neutral statements and questions, in particular to the "intonation contours" which come very close to being straight lines. Thus, I think the procedure may be justified not only by its purpose but also by its results.

(a) Statements

When emphasis for contrast is on the first word (fig. 6, top left) the stressed syllable of that word is higher (and has a more elaborate rise, cf. figs. 2-5 and the full frame in table 1, S1A); the rise to the first post-tonic resembles the neutral stress group but the fall through the succeeding post-tonics is considerably steeper. The level of the second and third stressed syllables is lowered considerably and the rises to the post-tonics are deleted. Tonally, the second and third stress groups thus resemble a conti-

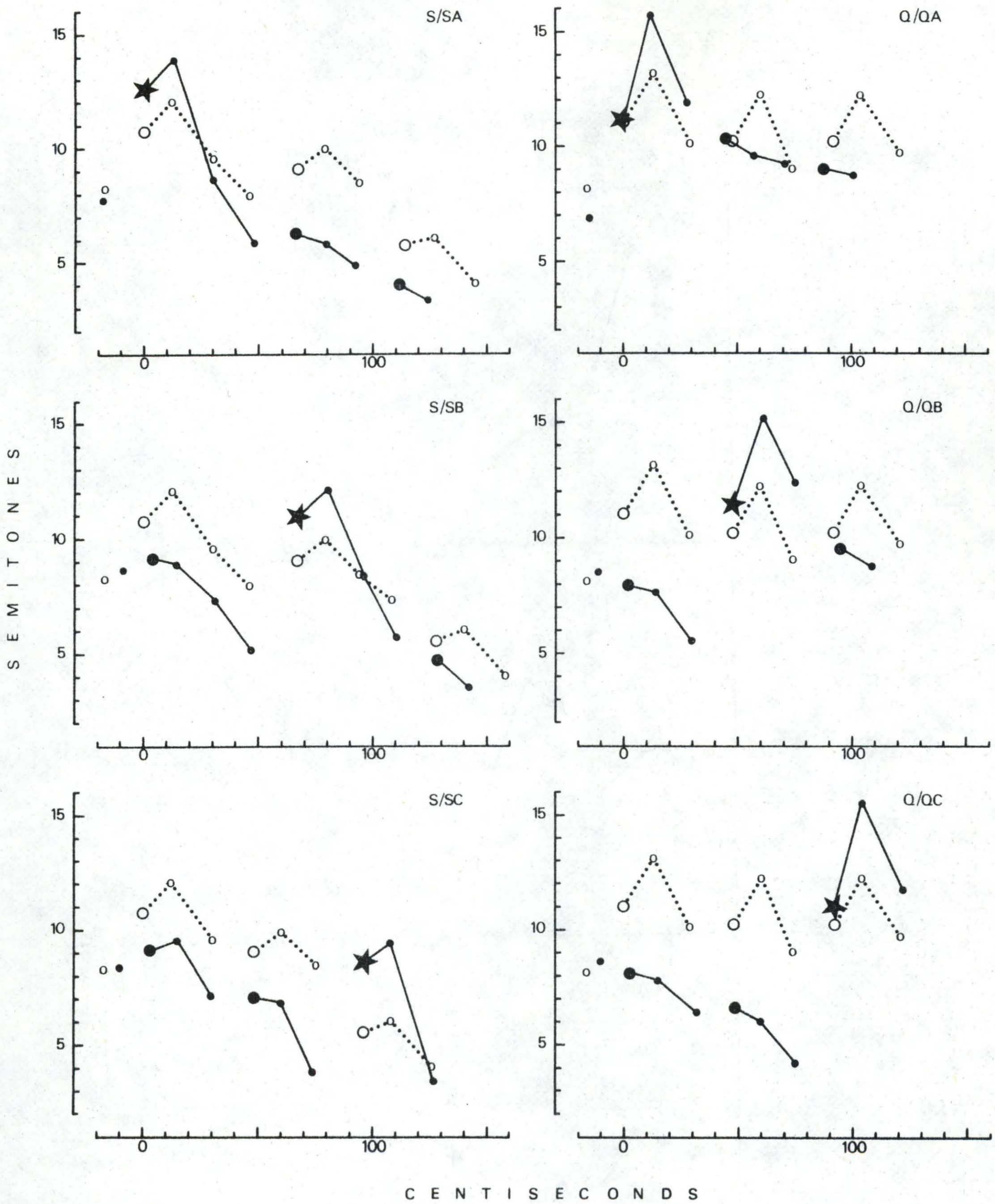


Figure 6

Stylized tracings of the course of fundamental frequency (mean of means over six readings by each of two subjects and ten readings by one subject) in statements (S - left) and questions (Q - right): prosodically neutral (empty circles - dotted lines) and with emphasis for contrast (stars and full circles - full lines) initially (A - top), medially (B - mid), and finally (C - bottom) in the utterance. Stars denote the emphasized syllable, large circles the stressed syllables, and small circles unstressed syllables. The tracings are composed of nonsense stress groups from different sentences, see further section 6.4.2.1. They have been lined up according to the emphasized syllable. Zero on the logarithmic frequency scale corresponds to 100 Hz.

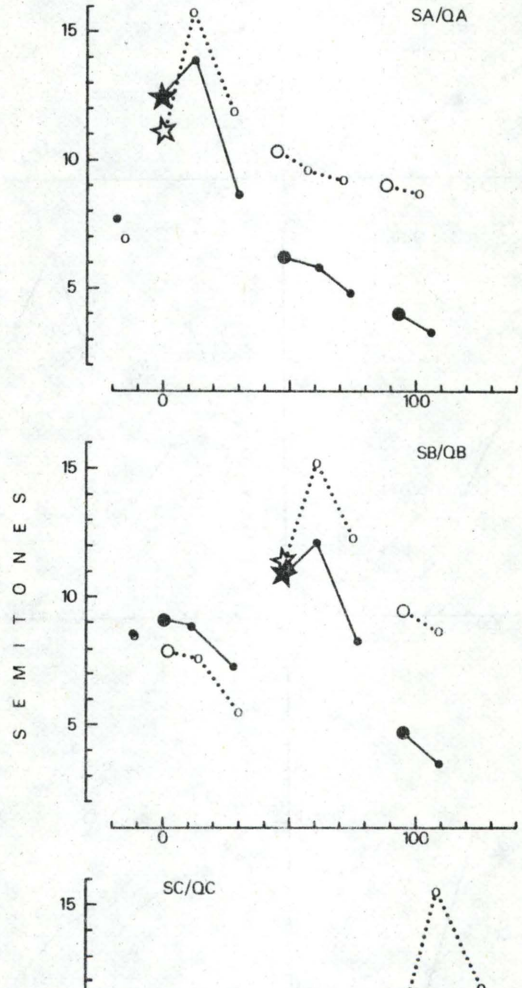


Figure 7

Statements and questions with emphasis for contrast from fig. 6 (statements: S - full stars and circles, full lines; questions: Q - empty stars and circles, dotted lines; emphasis initially (A - top), medially (B - mid), and finally (C - bottom)).

nuation of the high-falling tail of post-tonic syllables in the first stress group. Individual subjects conform well with this description, cf. the full frames in table 1, S1A, S2A, S3A, though only JR and BH have higher emphasized syllables.

Emphasis on the second content word[1] (fig. 6, mid left) repeats the pattern described above and in addition the preceding stress group is lowered, it has no rise to the first post-tonic, and the unstressed syllables describe a rather steep fall. Again individuals agree well, cf. the full frames in table 1, S2B, S1B, S3B, though only two subjects have more elaborate rises in the emphasized syllable.

With emphasis on the third content word (fig. 6, bottom left) we get a very considerable raising of the stressed syllable of the emphasized word, a rise to the post-tonic which is only very slightly larger than in the neutral stress group and a drastic fall to the second post-tonic. Both preceding stress groups are lowered,[2] the first one (farthest away from the emphasized stress group) retains a slight rise to the post-tonic which may be seen as a hint that the effect of an emphasis wears off at a distance. Subjects agree well to the description of the emphasized stress group (full frame in S3C in table 1). With JR and NT the second preceding stress group is lower, with NT its post-tonics describe a steeper fall, with BH and NT it has no rise (see the full frame in S1C in table 1).

(The behaviour of stress groups with only one post-tonic as well as intravocalic movements will be treated below and in section 7.)

The feature common to the three cases is that the stressed syllable of the emphasized word stands out clearly from the surroundings, which is brought about by a raising of that stressed syllable as well as by a lowering and shrinking of the Fo deflections in the surrounding stress groups, in a manner so that the immediate surroundings - except the first post-tonic - fall away

---

1) 'content word' is to be understood as opposed to 'function word' or 'empty word', not in contradistinction to 'nonsense word'.

2) Note that there is only one subject behind the second stress group in SC in fig. 6 (and S2C in table 1).

sharply from the emphatic syllable. If we consider percentages at or above 75% to be high and percentages at or below 25% as low, then table 2, statements - nonsense words, of course confirms that the emphasized syllable is higher (75%), the fall through the post-tonics is steeper (100%); the preceding stress group is lower (100%), its post-tonics more steeply falling (75%), it has no rise (25% partial rise); succeeding stress groups are both lower (100%) and have no rise (0%). - An account of word boundaries is given in section 7.3 - suffice it here to say that word boundaries leave as little trace in the emphatic Fo courses as they do in neutral stress group patterns.

A comment on the prosodically neutral statement and question is called for: A connection of the large empty circles (the stressed syllables) in the statement in fig. 6, top and mid left, would not result in a straight line. This seems to be due to the longer first stress group, compare fig. 6, bottom left, where the "intonation contour" is indeed almost straight (see further Thorsen 1980a). The rise to the first post-tonic decreases with time, as predicted by the model (fig. 1).

(b) Questions

The patterns observed in the statements reappear in the questions in fig. 6 (and table 1), with some modifications: The stressed syllable of the emphasized word is not raised very pronouncedly (the rises in fig. 6 seem mainly to be due to NT, cf. the full frames in Q1A, Q2B, Q3C in table 1) but the rise to the post-emphatic unstressed syllable is considerably higher (and is more elaborately rising) than in the neutral stress group, and on this point subjects are unanimous, cf. table 1. Stressed syllables succeeding the emphasized one are not very remarkably lowered but the rise to the post-tonic is clearly deleted (cf. the full frames in Q2A, Q3B, and Q3A in table 1), and the impression of all of the post-emphatic syllables is still that they tonally resemble one tail of post-tonic syllables.

Stress groups preceding the emphasized syllable in the questions behave qualitatively and quantitatively like pre-emphatic syllables in the statements, to the effect that the relative lowering of preceding stressed syllables (fig. 6, mid and bottom right) is greater in questions than in statements. Subjects are unanimous on the lowering and lack of rise in preceding stress groups, cf. the full frames in table 1, Q1B, Q2C, and Q1C. - Thus, the questions do not show any signs of a wear-off of the effect of emphasis two stress groups away, as did the statements, compare fig. 6 bottom left and right.

Table 2, questions - nonsense words, reflects these observations: The emphasized syllable is more elaborately rising (89%), its post-tonic rise is higher (100%); both preceding stress groups are lower (100%) and have no rise (0%); succeeding stress groups have no rise (0%).

In the prosodically neutral questions the three stressed syllables are not perfectly horizontal, but very nearly so. Furthermore, the rise to the post-tonic does not decrease linearly with time, and in section 7.1.2 we shall look at more data which have the same feature.

(c) Comparing statements and questions

The significant difference between statements and questions (fig. 7) lies in the course of the post-emphatic syllables, which run higher in questions than in statements. The preceding Fo courses do not differ markedly. With emphasis in the first position (fig. 7, top) the emphasized syllable is higher in the statement, in the second position (fig. 7, mid) statement and question are about equally high, and in the third position (fig. 7, bottom) the emphasized syllable is higher in the question. In other words: In statements the emphasized syllable lowers progressively with time, in questions it remains at the same altitude, a pattern which is reminiscent of the gradients described by the stressed syllables in neutral statements and questions, cf. fig. 1.

From fig. 6 it appears that in statements the changes in succeeding stress groups are greater than in preceding ones (the stressed syllables after the emphasized syllable sink relatively lower than those before it), whereas in questions the opposite relation holds (preceding stressed syllables sink lower than succeeding ones). - However, this is hardly the salient point: Some very preliminary experiments with synthetic speech suggest that the minimal requirement for perceiving a contrast emphasis is the shrinking of the surroundings, i.e. the deletion of the rise in surrounding stress group patterns and not the raising and lowering, respectively, of the emphatic syllable and its surroundings. During a visit to the Institute of Linguistics at Uppsala

University I had occasion to use the ILS-system for analysis and synthesis. I recorded the sentence 'Det er sidste bus til Tiflis.' (It is the last bus for Tiflis.) and re-synthesized it with a large number of different Fo courses. It turned out that (at least to my ear) in order to create the impression of a contrastive stress on the second stressed word 'bus', it was sufficient to delete the rise in the preceding word ('sidste'), i.e. to just lower its post-tonic syllable, whereas raising 'bus' alone would not do the trick. Likewise, I could get emphasis on the word 'sidste' by deleting the rise in the succeeding stress group, i.e. lower the unstressed word 'til' to the level of 'bus', and again, raising the stressed syllable of 'sidste' will not alone give the impression that that word is contrasted to some other word.

<u>6.4.2.2 Natural words</u>

In figs. 8 and 9 the procedure employed for the construction of fig. 6 is repeated for the natural words in the pipipi-sentences.[1] Fig. 8 depicts the course of the words 'kortere' ([ˈg̊ʰɒːd̥ʌːʌ]) (initially and medially) and '-kortes' ([ˈg̊ʰɒːd̥əs]) (finally), and fig. 9 shows tracings of 'stavelserne i' ([ˈsd̥æːw̩sʌnə i]), 'stavelser med' ([ˈsd̥æːw̩sʌ mɛ]), and 'stavelser' ([ˈsd̥æːw̩sʌ]), respectively, extracted from different sentences in a manner so as to reproduce, to the extent that it is possible by such a procedure, the Fo course in utterances where the syllables contain long vowels and sonorant consonants. (E.g. the emphatic question with emphasis on the second element, QB, in fig. 8 is pieced together of the 'kortere' words from sentence 48 (initial position), sentence 15 (medial position), and sentence 30 (final position).)

<u>(a) 'kortere'</u>

Note, first of all, that the stressed syllable [ˈg̊ʰɒː] in the neutral edition confirms the observations about vowel Fo movement (section 5): In initial position in statements the vowel is mainly rising but medially and finally it is falling-rising, and the fall is of greater extent than the rise. In questions the rise is greater than the fall. Note also that with these words

---

1) The words are all extracted from sentences that have the nonsense word "pi'pipi" in them, except that 'stavelserne i' initially and '-kortes' finally in the neutral question come from sentences 23 and 22, respectively.

the straight line intonation contours (cf. fig. 1) appear very clearly, also in the questions, cf. above about fig. 6) but, again, the magnitude of the rises to the post-tonic in the questions do not conform to the model.

On a few points fig. 8 deviates from fig. 6:[1] With emphasis on the first stressed vowel in the statement (fig. 8, top left) the stress group does not alter its level or shape appreciably, compared to the neutral stress group, and the third stress group retains most of its rise. In fig. 8, mid left, the final stress group does not lower much but the rise to the post-tonic is deleted. With emphasis on the final element (fig. 8, bottom left) there is no lowering of preceding stressed syllables and hardly any change in the first one at all but the rise in the second stress group is clearly deleted. Thus, the suspicion grows that emphasis for contrast in the statements exerts a weaker influence on stress groups "twice removed". - The questions (fig 8, right) compare very well indeed with the nonsense words (fig. 6, right), except that the relative lowering of preceding stress groups is not so great in fig. 8, bottom right.

The emphasized syllable has rising vowel movements in all positions and generally more elaborate rises in the questions.

(b) 'stavelser --'

The first post-tonic syllable in 'stavelser' is carried by a syllabic [l̩] due to assimilation of /ə/, which is why the 'stavelser' tracings (fig. 9) look slightly different from the 'kortere' ones: The maximum Fo value is generally not reached until the beginning of the vowel of the second post-tonic syllable ([ʌ:]). This is not a unique or exceptional case but to what extent schwa-syllables deviate from post-tonic full vowels, I hesitate to say at present (Thorsen, forthcoming, treats this problem in detail).

Apart from that, fig. 9 does not add much to what has been said already about fig. 8, i.e.: With the statements, the most obvious changes introduced by emphasis for contrast are the shrinking of immediately surrounding Fo patterns as well as a raising of the stressed syllable in medial and final positions but with the questions we also get a lowering of preceding stress groups, as with the nonsense words - though not so great.

---

1) Reference to tables 1 and 2 are assembled at the end of this section.

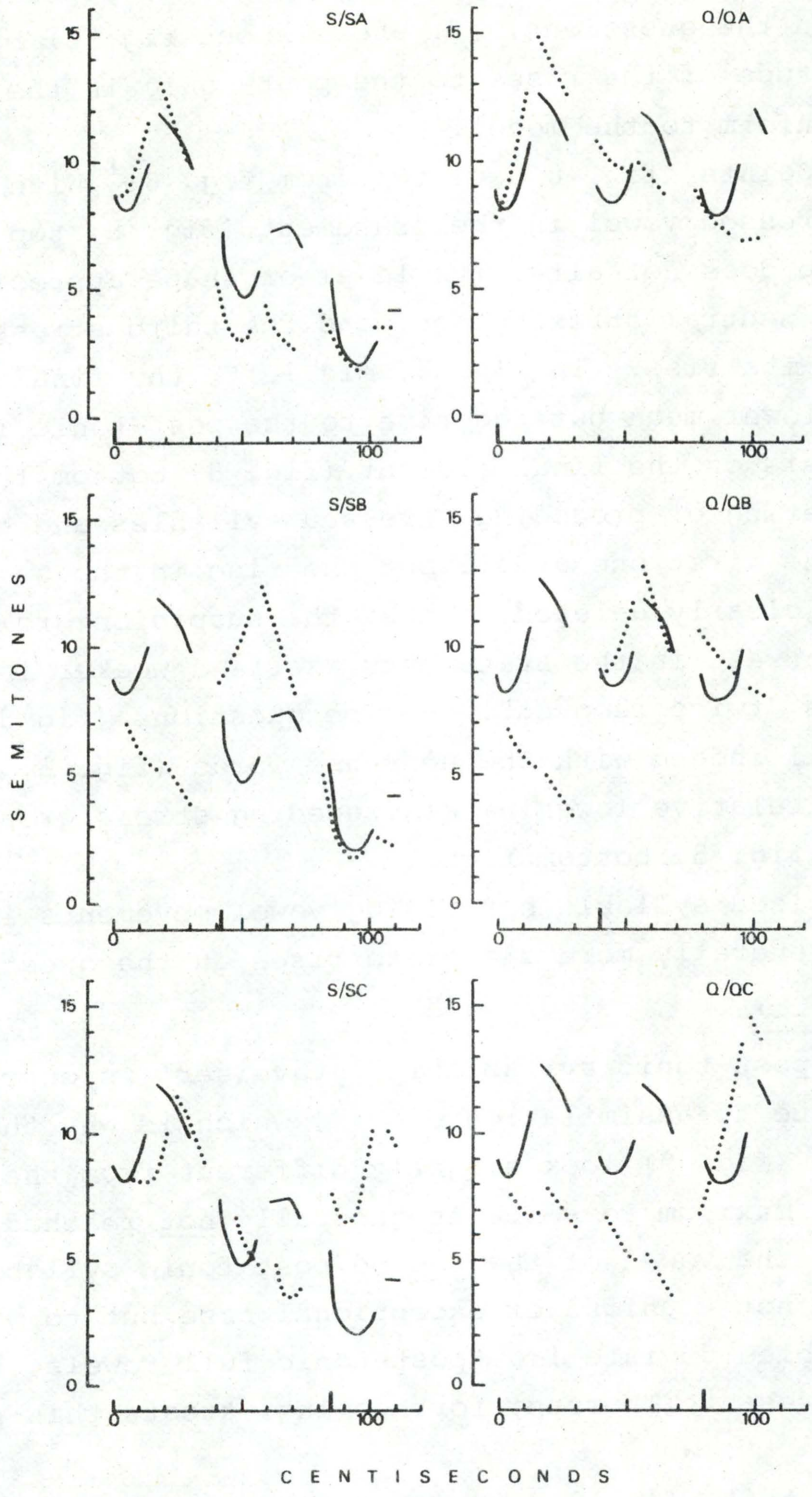


Figure 8

Fundamental frequency tracings (mean of means over six readings by each of two subjects and ten readings by one subject) in statements (S - left) and questions (Q - right): prosodically neutral (full lines) and with emphasis for contrast (dotted lines) initially (A - top), medially (B - mid), and finally (C - bottom) in the utterance. The tracings are composed of stress groups from different sentences, see further section 6.4.2.2. They have been lined up according to the beginning of the emphasized vowel (heavy stroke on the time scale). Zero on the logarithmic frequency scale corresponds to 100 Hz. - The stress groups depicted initially and medially consist of the word 'kortere' [ˈg̊ʰɒːd̥ʌːʌ], the final one is '(for)kortes' [ˈg̊ʰɒːd̥əs].

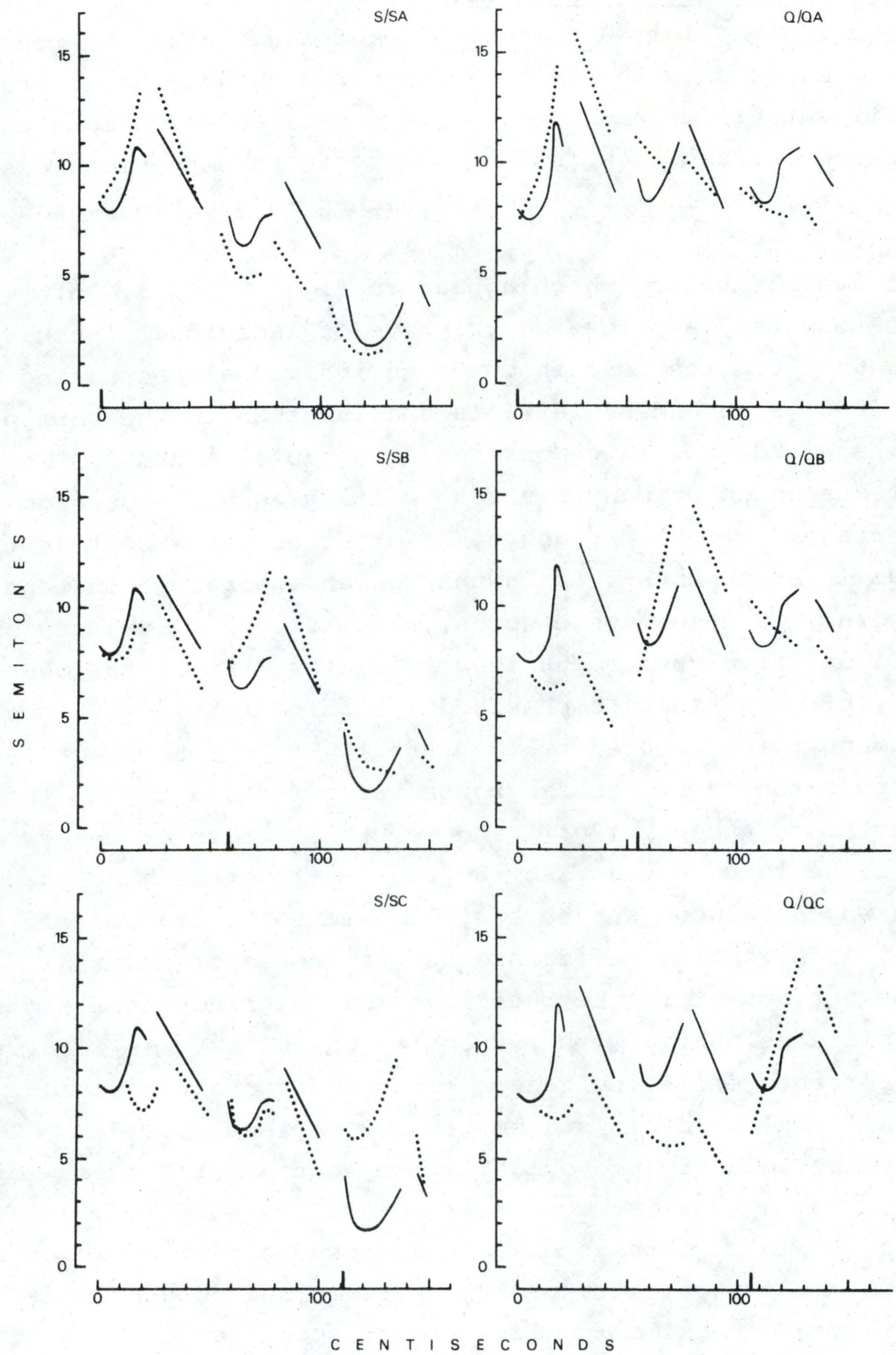


Figure 9

Fundamental frequency tracings of 'stavelserne i' [ˈsd̥æːw̩sʌnə i] (initially), 'stavelser med' [ˈsd̥æːw̩sʌ mɛ] (medially), and 'stavelser' [ˈsd̥æːw̩sʌ] (finally). See further the legend to fig. 8.

(c) Comparing statements and questions

Comparing the left and right side of figs. 8 and 9, respectively, it seems - as with the nonsense words - that the general and significant difference between statements and questions with emphasis for contrast lies in the course of the emphatic and succeeding syllables which rise and run higher in questions than in statements.

The average trends which appear in figs. 8 and 9, left (i.e. in the statements), are less descriptive of individuals' behaviour than fig. 6. I.e. the likelihood of an individual recreating the patterns in figs. 8 and 9, left, is smaller than in the case of the nonsense words, as is demonstrated by tables 1 and 2: There are fewer instances of many plusses or minusses and thus high and low percentages (i.e. at or above 75% or at or below 25%) in the real words in statements than in the nonsense words (5 instances in real words, 10 in nonsense words, cf. table 2).[1] Subjects only agree well on (a) higher emphasized syllables (81% of the observations), (b) more rising emphasized syllables (88%), (c) lower stress groups twice removed to the left (80%), (d) the fact that the post-tonics in stress groups twice removed to the left do not fall more steeply (20%), and (e) that stress groups twice removed to the left have a rise (80/20%). - On the other hand, the differences between subjects still seem to be random (cf. table 1). - A similar difference between nonsense and real words does not appear in the questions (although nonsense and real words differ slightly with respect to which features receive high/low percentages). The reason for the different behaviour of nonsense and real words in statements may reflect a tendency to treat nonsense words in prosodically non-neutral utterances more carefully, i.e. a nonsense word will more often (and to a greater extent) and in more features exhibit a prosodic pattern which deviates from the prosodically neutral pattern and produce what might be termed a "maximally distinct" prosodic pattern. - The reason for the different behaviour of statements and questions in this respect may be attributed to a difference between marked (question) intonation contours and unmarked (declarative) contours: On a marked intonation contour, prosodic non-neutrality is signalled more often, more clearly, and in more features than on unmarked intonation contours. Thus, the vacillation in the non-

---

1) This is partly, but not exclusively, due to the fact that in some respects the two real words behave differently.

neutral prosodic patterns of real words in statements disappears in (syntactically unmarked) questions where these patterns are "maximally distinct". (Note that the vacillation in emphatic statements is an inter subject affair - not a reflection of intra-subject variation because the standard deviation on Fo means is not greater in emphatic utterances than in prosodically neutral ones, cf. the end of section 6.4.1.)

6.4.2.3 Natural sentences

In fig. 10 the prosodically neutral edition of 'Der går mange busser fra Tiflis.' is compared to one with emphasis on 'mange' (top), 'busser' (mid), and 'Tiflis' (bottom).[1] Note that no correction for intrinsic Fo level differences between the stressed vowels has been performed (a correction which would have raised the low stressed vowel of 'mange' in relation to the succeeding stressed, high vowels) and thus the gently slanting slope of the intonation contour is not immediately apparent.

Emphasis on 'mange' does not raise the stressed syllable but the post-tonic syllabic nasal has a higher rise in the emphasized condition. (This resembles the situation in figs. 8 and 9, top left). Nor are the succeeding stressed syllables lowered but the rise to the post-tonic is deleted in the second stress group. No change is apparent in the third stress group.

Emphasis on 'busser' repeats the patterns in figs. 8 and 9, mid left, i.e. the stressed syllable of 'busser' is raised considerably and, further, the fall from the first to the second post-tonic is rather sharp. The preceding stressed vowel is not lowered (or only slightly so), but the rise to the post-tonic is deleted (the slightly higher position of the syllabic nasal in relation to the preceding low vowel may well be due to intrinsic Fo level differences). There is also a slight comparative lowering of the post-tonic syllable in the final stress group.

---

1) Fig. 10 is an average over all four subjects, which is why the tracings are all lower on the frequency scale than in figs. 6-8. It would of course have been possible to compensate for the variation introduced in the figures by the alternate inclusion and exclusion of NRP if the data had been zero-offset adjusted for each subject prior to averaging. I did not do this because it is really immaterial for an account of the results - the same relations would hold between utterances, which is one advantage of a logarithmic frequency scale.

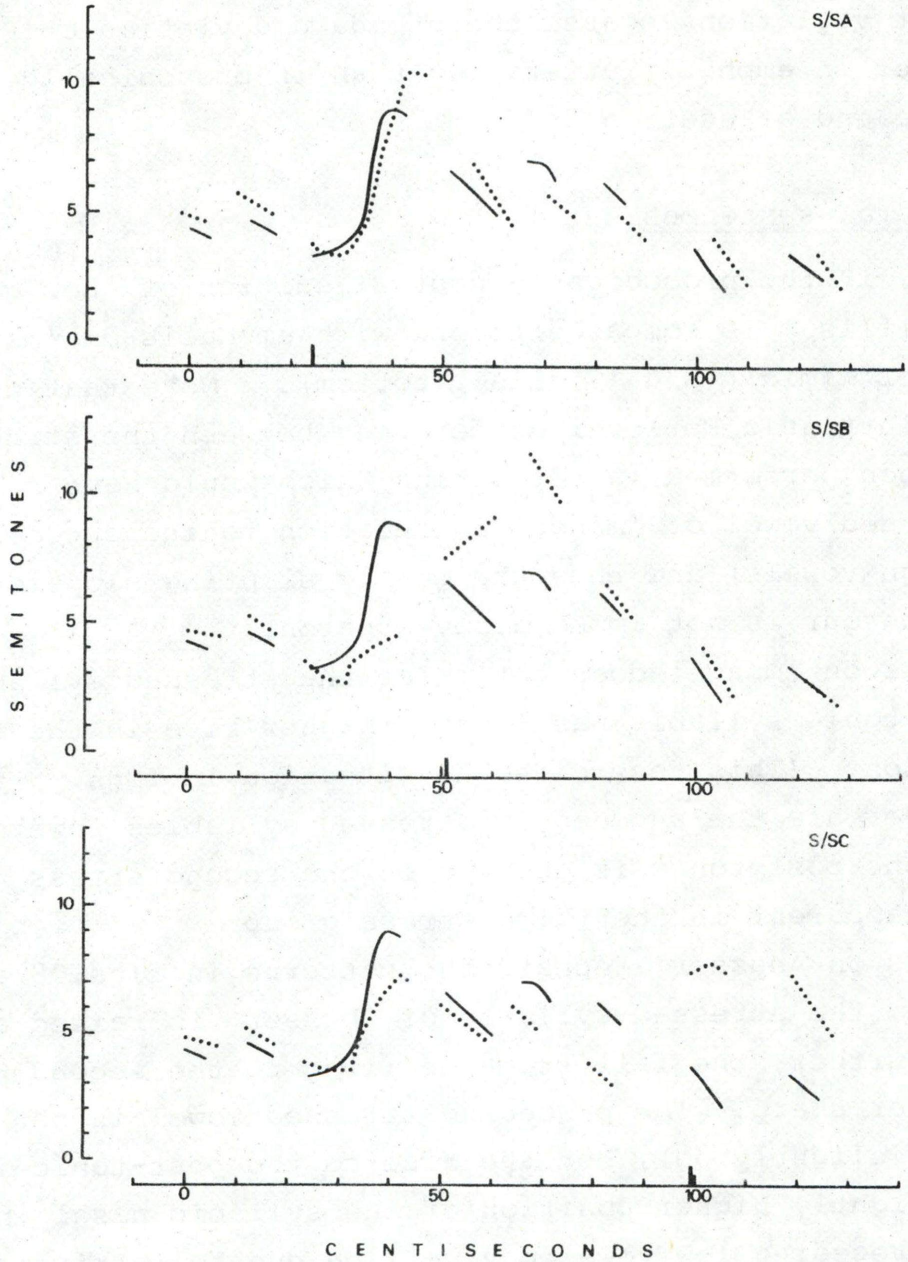


Figure 10

Fundamental frequency tracings (mean of means over six readings by each of three subjects and ten readings by one subject) of 'Der går mange busser fra Tiflis.' [d̥ɑ g̊ɒ ˈmɑŋ̊ ˈb̥usʌ fʁ̥ɑ ˈd̥ˢifl̥is]: prosodically neutral (S - full lines) and with emphasis for contrast (dotted lines) on 'mange' (SA - top), on 'busser' (SB - mid), and on 'Tiflis' (SC - bottom). Zero on the logarithmic frequency scale corresponds to 100 Hz. The tracings have been lined up according to the beginning of the emphasized vowel (heavy stroke on the time scale).

Emphasis on 'Tiflis' makes for a raising of 'Tif-' but no further rise to the post-tonic (which seems to be a general feature of emphasized stress groups with only one post-tonic, at least if that post-tonic is not carried by a syllabic nasal, cf. section 7 below). The preceding stressed vowels are again not lowered, and the shrinking of the Fo pattern in the first stress group is modest, though clear in the second one.

(The inter subject agreement is good in the natural sentences and this account would adequately describe any one of them, except that NT does have a lowering of stress groups preceding the emphasized one as well.)

At this point the suspicion becomes a near certainty that - at least in statements (the natural sentences did not include questions) - only in the immediately neighbouring stress groups are the Fo rises to the post-tonic completely deleted. Stress groups further away retain (a smaller or larger part of) their rises.

6.4.2.4 Fo and emphasis for contrast - conclusion

Slight differences between different parts of the material aside, a general conclusion may be formulated: When emphasis for contrast is introduced in a sentence, the stressed syllable of the emphasized word will stand out clearly from the surroundings which is brought about by a raising of (except in initial position) and an elaborate rise within that syllable and by a deletion of the Fo deflections in neighbouring stress groups, to the effect that the immediate surroundings, except the first post-tonic, fall away sharply from the stressed syllable of the emphasized word.

In this material statements and questions differ with respect to the extent of the influence of emphasis on the prosodic patterns: Fo patterns two stress groups away from the emphasized word retain (at least partly) their rises to the post-tonic in the statements but not in the questions. (In table 2 we observe a full or partial rise in the second preceding and succeeding stress groups in 76% (63%+13%) and 66% (33%+33%) of the observations, respectively, in statements but only 44% partial rises in the second preceding stress group and no rises at all in the second succeeding stress group in questions.) To the extent that this difference is a stable one, there must be a reason for it (akin to the explanation at the end of section 6.4.2.2(c) for the difference in the treatment of natural words vs. nonsense words): It could be that more of a change is invoked in the prosodic patterns on marked intonation contours (which accompany syntactically unmarked questions) than on unmarked contours (declarative sentences). - The material is not comprehensive enough to establish just how far emphasis "reaches" on marked contours.

For the sake of clarity of the exposition I will temporarily disregard the limited influence of emphasis in declarative sentences and state that: 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. The difference between statements and questions with emphasis for contrast is mainly located in the movement within the emphatic syllable, which rises higher in questions, and in the post-emphatic syllables, which run higher in questions than in statements. Thus, to a certain extent, emphatic utterances (with more than one stress group) resemble prosodically neutral (real) one word utterances of which it has previously been shown (Thorsen 1978) that the difference between statement and question manifests itself in the [level (higher in questions) and] movement (rising in questions) of the stressed syllable as well as in higher post-tonics in questions. This is not to say, of course, that one word utterances and utterances with contrast emphasis must necessarily be identical phonetically (as they certainly are not semantically). A real one stress utterance and a longer one with emphasis will usually be different tonally in the height of and movement within the emphasized syllable.

With this generalization of the results a notable difference between short prosodically neutral utterances and utterances with emphasis for contrast exists: In prosodically neutral sentences the information about sentence type (e.g. statement vs. question) is contained in the ensemble of stressed syllables, whereas in emphatic utterances the one "stressed" syllable and its "post-tonics" (which may or may not include (reduced) main, i.e. secondary stresses) will carry the burden alone. But, precisely, if we regard utterances with contrast emphasis as one word utterances, this does not jeopardize the definition of intonation contour as given in the introduction, nor the contention that intonation contour and stress group pattern essentially are two invariant entities which interact on the "physical" level only.

If we accept stress as a relational property and accept the reduction of Fo patterns surrounding a stress group which contains the stressed syllable of an emphasized word as evidence that the stressed syllables of those stress groups are reduced from main to secondary stresses, then it seems that three degrees of stress will suffice to account adequately also for short utterances with emphasis for contrast, - at least as long as these utterances are emotionally neutral, cf. sections 6.2 and 6.3. Note, again, that this is not necessarily a statement to the effect that phonologically, neutral main stress and emphasis for contrast are identical: they certainly have different effects on the surroundings and intuitively it seems wrong to treat them as inherently equal (and, as mentioned above, they may also well be different phonetically, although this difference - which consists in a higher and more elaborately moving emphasized syllable - may not be a necessary requirement for the perception of contrast, cf. section 6.4.2.1(c)). But it seems that two stress degrees, 'secondary' and 'weak', below 'main' and 'emphatic' are sufficient in both instances, i.e. three degrees will suffice. If it were not for the fact that neutral main and emphatic stresses are not <u>always</u> very different phonetically, and that such a difference does not seem to be a necessary requirement, then we might of course resolve the issue as follows: weak syllables are assigned a stress degree "0", secondary stresses "1", neutral main stresses "2", and contrast emphases "3" and then note that the occurrence of "3" precludes the occurrence of "2" (or, in other words: reduces "2" to "1") (at least in the immediate environment, cf. below). As the matter stands, however, I think that it is more appealing to account for the difference between neutral main stresses and emphases for contrast as one of kind rather than one of degree, i.e. to see emphasis for contrast as a phenomenon which is not fully integrated in the ordinary stress hierarchy.

A complication is introduced by the fact that (at least in declarative sentences, cf. above) stress groups further away from the emphatic one are unaffected, i.e. utterance length (in terms of number of stress groups) cannot be disregarded, and we cannot always state that the whole utterance reduces tonally to one stress group when emphasis for contrast occurs. - Instead we shall probably have to state that the stress group containing the stressed syllable of the emphasized word together with the immediately neighbouring stress groups reduce to one stress group. (In questions the emphasis may be more far reaching, but just how far I cannot say at present.) - We can presumably still make do with three relational stress degrees to describe such longer utterances with emphasis for contrast, but the stress relations that we describe must then be confined within the boundaries of each stress group (whether prosodically neutral or emphasized) and cannot extend over a whole utterance. (This does not obliterate the difference between Danish contrast emphasis and e.g. English nucleus, because in English, too, the <u>four</u> stress degrees describe relations within the tone group, of which there may be several in a sentence.)

I hesitate to say at present what constitutes the acoustic and perceptual cues to sentence intonation in long utterances with emphasis for contrast where the smoothly slanting slopes of the prosodically neutral intonation contour will of course be broken, due to the raising of the emphasized syllable - is it the overall downdrift, or is it e.g. the "post-emphatic" downdrift only? Further investigations are called for.

If the interplay on the physical level between stress groups and sentence intonation components were less complicated, i.e. less dependent on stress group composition and utterance length, a predictive model could be proposed that also took emphasis for contrast into account. And it is indeed possible to lay out upper and lower limits for (emphatically) stressed syllables, and for the first post-tonic in each stress group, - but I do not feel that such a graph satisfies the demands on a model that should generate the Fo course of all parts of the utterance. - Instead I think that one must, at least for the present, resort to fig. 6, which gives a rough indication of what to expect from short utterances with contrast emphasis in different positions.

Finally, I wish to stress once more the fact that what I have investigated are utterances that are emotionally neutral. - I do not wish to claim that emphasis in general must necessarily be signalled in the same way as emphasis for contrast, - in fact I am certain that this is not the case: A type of emphasis where the stressed syllable is low in relation to its surroundings is fairly common; it invokes the impression of the speaker being patronizing or indignant. (Further, a low emphasized syllable is the rule in the second contrasted term in explicit emphasis.) - I cannot even claim that emotionally neutral emphasis is inevitably realized as I have described above: I think it is possible (though maybe not common in spontaneous speech) to introduce non-emotional emphasis (be it for contrast or some other purpose) without necessarily reducing the surrounding stress groups - it seems to be fairly common with e.g. news readers on radio and TV. I do not know what constitutes the acoustic characteristics of such utterances but I think it likely that the emphasized stress group, besides the higher and more elaborate Fo movement of the stressed syllable, may be considerably lengthened (a feature which is not characteristic of the present material, cf. below).

6.4.2.5 Emphasis for contrast and duration

Durations have been measured in all of the material. However, only the results for the natural ('Tiflis') sentences are presented here. The nonsense word material conforms well with the conclusions to be drawn below.

In table 3 the duration of each segment, each stress group, and the whole sentence is given (mean of means from four subjects), and in fig. 11 this information is displayed graphically. Differences in total duration between the prosodically neutral edition (S) and the ones with emphasis for contrast (SA-SC) are fairly small, and seem mainly to be ascribable to the longer emphasized stress group. On this point there is overall agreement between subjects: In table 4 the durations are compared (qualitatively) for each subject. A plus indicates that the segment, stress group, or sentence, respectively, was longer in the neutral edition, and a minus that it was longer in the emphasis condition. (The asterisks indicate that the difference is statistically significant (one-tailed t-test) at the 5% level or better.) When all subjects agree as to the direction of a durational difference (which is comparatively rare), this is indicated by a plus or minus in the "overall agreement" row. Note that the pre-vocalic consonant in the emphasized stress group is significantly longer than in the neutral case with all subjects (except JR: [d̥^{s}]).

Table 3

Duration (in centiseconds) of the segments and stress groups in, as well as the total utterance: 'Der går mange busser fra Tiflis.', prosodically neutral (S), with emphasis for contrast on 'mange' (SA), on 'busser' (SB), and on 'Tiflis' (SC). Average over four subjects (mean of means).

	ɑ	g̊	ɒ	-	m	ɑ	ŋ	-	b̥	u	s	ʌ	fʁ̥	ɑ	-	d̥ˢ	i	fl̥	i	s	-	total
S	5.1	6.8	5.5	17.4	6.4	9.1	9.6	25.1	8.8	8.4	5.9	5.3	9.7	4.1	42.2	14.3	5.9	13.4	6.2	14.3	54.1	138.8
SA	4.9	6.2	6.1	17.3	8.7	10.8	11.5	31.0	8.2	8.4	6.9	5.1	10.3	3.8	42.7	14.1	6.1	13.4	5.3	14.2	53.1	144.1
SB	5.3	6.3	5.1	16.7	6.2	8.5	9.2	23.9	10.4	9.2	6.6	6.1	9.6	4.6	46.5	14.1	6.2	13.5	5.5	14.3	53.6	140.7
SC	5.4	6.1	5.7	17.2	6.2	8.8	9.7	24.7	8.0	8.0	5.9	5.3	9.7	4.5	41.4	16.1	7.0	12.7	6.8	15.5	58.1	141.5

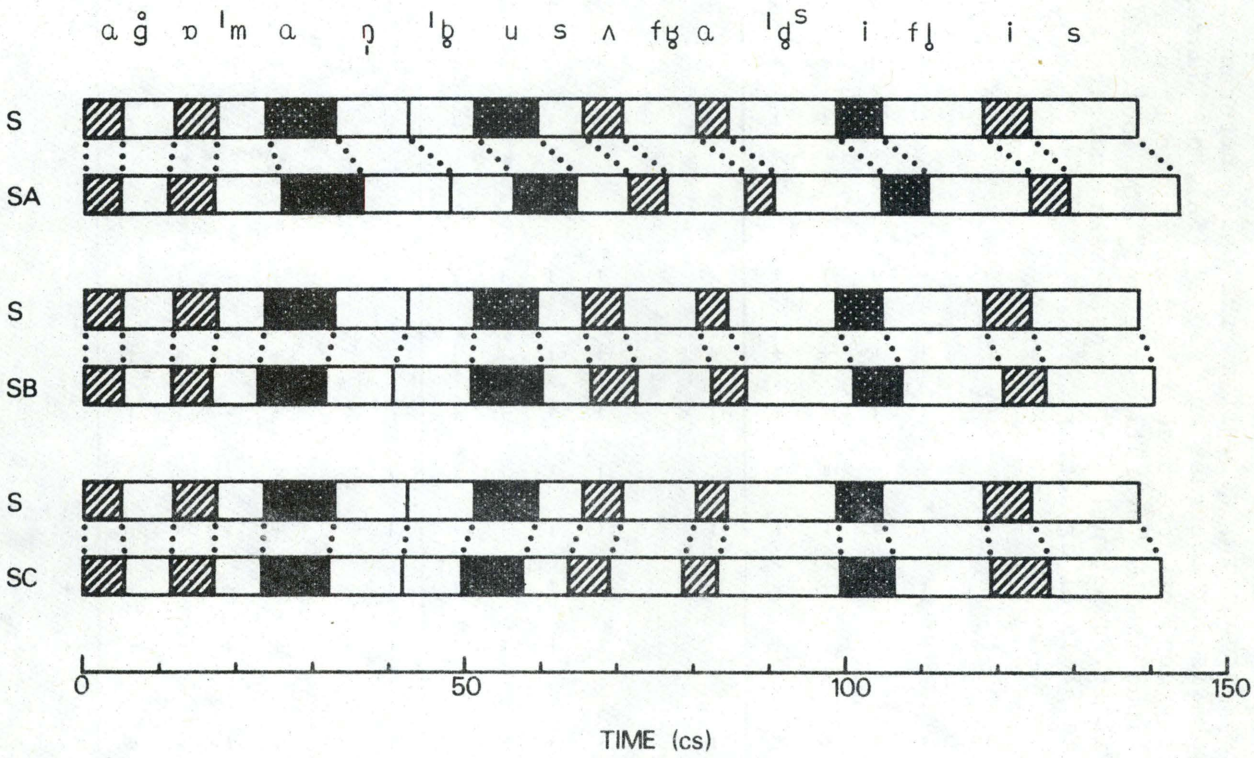


Figure 11

Graphical display of durations (mean of means over six readings by each of three subjects and ten readings by one subject) in 'Der går mange busser fra Tiflis.': prosodically neutral (S), with emphasis for contrast on 'mange'(SA - top), on 'busser' (SB - mid), and on 'Tiflis' (SC - bottom). Vowel segments are hatched, stressed vowels cross-hatched.

Table 4

Durational differences between segments and stress groups in, as well as the whole utterance: 'Der går mange busser fra Tiflis.', prosodically neutral (S), with emphasis for contrast on 'mange' (SA), on 'busser' (SB), and on 'Tiflis' (SC). Plusses indicate that the segment, stress group, or utterance is longer in the prosodically neutral edition (S) than in the emphatic condition. A plus or minus in the "overall agreement" row indicates that all four subjects agree as to the "direction" of the difference, if it is not nil ("="). Asterisks indicate that a difference is statistically significant (one-tailed t-test) at the 5% level or better.

	ɑ	g̊	ɒ	-	m	ɑ	ŋ	-	b̥	u	s	ʌ	fʁ̥	ɑ	-	d̥ˢ	i	fl̥	i	s	-	total
S/SA: NRP	+	+	−	+*	−*	−*	−*	−*	+*	−	−*	+	−	+*	+*	+	+	−*	+*	−	−	−*
JR	+	−	−	−	−*	−*	−	−*	=	−	−*	+	−	=	−*	+	−	−	+	−	−	−*
BH	−	+	−	−	−*	−*	−*	−*	−	+	−	=	−	+	−	−*	+	+	+	+	+	−*
NT	+	+	−	+	−*	−*	−*	−*	+	+	−	−	−	+	−	+	−	+	+	+	+*	−
overall agreement			−		−	−	−	−			−		−	+					+			−
S/SB: NRP	−	+*	+	+*	+	+*	−*	+	−*	−	−	−	−	+*	−*	−	+	−*	+*	−	−	−
JR	−	−*	+	−	−*	+	+*	+	−*	−	−	−	+	−	−*	+*	−*	+*	+	−	+	−
BH	−	−	+	−	+	+	−	+	−*	−*	−*	−	+*	−*	−*	−*	+	+	+	+	+*	−
NT	=	+*	+	+*	+	+	+*	+*	−*	−	−	−*	+	−*	−*	+	−	+	+	−	−	−
overall agreement	−		+			+		+	−	−	−	−			−				+			−
S/SC: NRP	=	+*	−	+*	+	+*	+	+	+	+	+	+	−	+	+*	−*	−	+*	−	−*	−*	+
JR	−	−	+	−	−	−	+	−	+	+	−	−	+	−*	+	−	−	+	+	−	−	−
BH	=	+	−	−	+	+	−*	−	+	+	=	=	−	=	−	−*	−*	=	−	+	−*	−
NT	−	+	+	=	+	−	+*	+*	+	+	−	−	−	−	−	−*	−*	−*	−*	−*	−*	−*
overall agreement	−								+	+						−	−				−	

The emphatic vowel is longer too, but not significantly so in all positions (initial - SA, medial - SB, final - SC) and with all subjects. There is a tendency for the stressed vowel preceding the emphasis to be shortened but no such tendency appears in succeeding stressed vowels. Thus, to a certain extent, the lengthening of the emphasized stress group is counterbalanced by a shortening of the preceding one - which is probably the reason why emphasis in the first stress group makes for a comparatively longer total duration than emphasis on the second and third stress groups (there is no preceding stressed vowel to shorten).

On the whole, the changes induced by contrast emphasis on the durations and durational relations within the sentence can be said to be less drastic than the changes induced on fundamental frequency. - No doubt, there are types of emphasis which will change duration more radically, but emotionally neutral emphasis (except maybe the "news reader emphasis", cf. above) does not interfere markedly with the rhythmic structure of the utterance.

7. A closer look at the tonal properties of the stress group

Fundamental frequency tracings of all the stress groups involving nonsense words in the material, except those that surrounded an emphasized stress group, are depicted in figs. 12-15 for individual subjects. (The second half of the Fo course of [g̊iʌ̆] has been deleted in figs. 12-16 when it is the first, i.e. only, post-tonic, to make it more directly comparable to post-tonic [b̥ʰi]. Note that the stress group commences with the stressed syllable, i.e. pre-tonic pi-syllables are <u>not</u> included.) Subjects differ among themselves but not in a systematic and qualitative respect, and a grand mean (mean of means, fig. 16) will not obscure differences of a fundamental kind between the four speakers. (Note that NRP is not included in the questions with emphasis, and see the footnote in section 6.4.2.3.)

Apparently, in many instances the course of fundamental frequency, i.e. the shape of the Fo pattern, is insensitive to stress group composition: In some positions and conditions the Fo patterns are concurrent regardless of the number of unstressed syllables in the stress group (except of course that the tail is longer with more post-tonics); in other positions and conditions

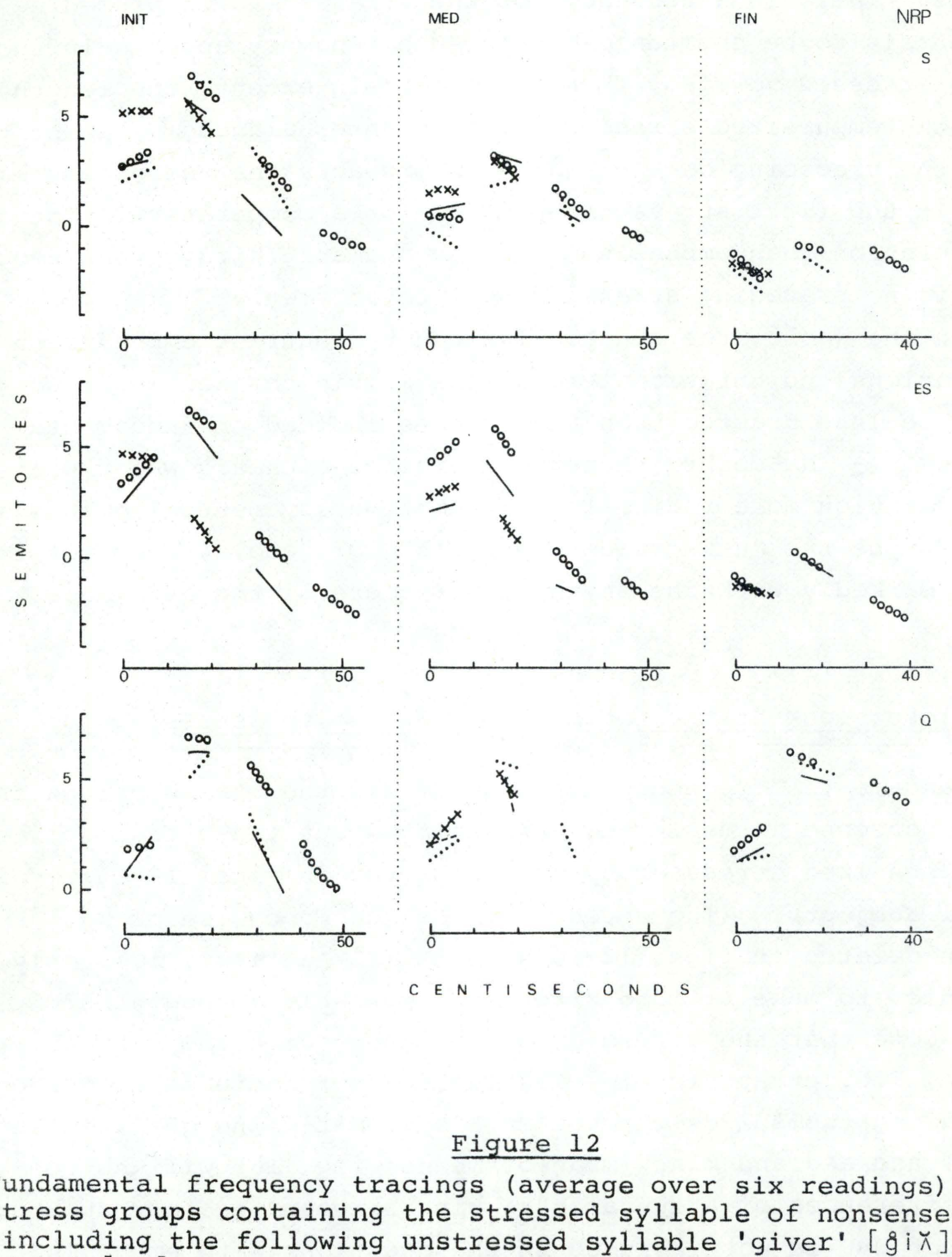


Figure 12

Fundamental frequency tracings (average over six readings) of stress groups containing the stressed syllable of nonsense words (including the following unstressed syllable 'giver' [ĝiʌ̆] and 'for-' [fʌ], respectively) initially (left), medially (mid), and finally (right) in prosodically neutral statements (S - top), statements with emphasis for contrast on the nonsense word (ES - higher mid), prosodically neutral questions (Q - lower mid), and questions with emphasis for contrast on the nonsense word (EQ - bottom). See further section 6.4.1. The nonsense words are designated as follows: 'pipi ·········· - (pi)'pi ----- - 'pipipi **ooooooo** - (pi)'pipi ——— - (pipi)'pi **xxxxxx** Zero on the logarithmic frequency scale corresponds to 100 Hz. - Subject NRP (no questions with emphasis for contrast).

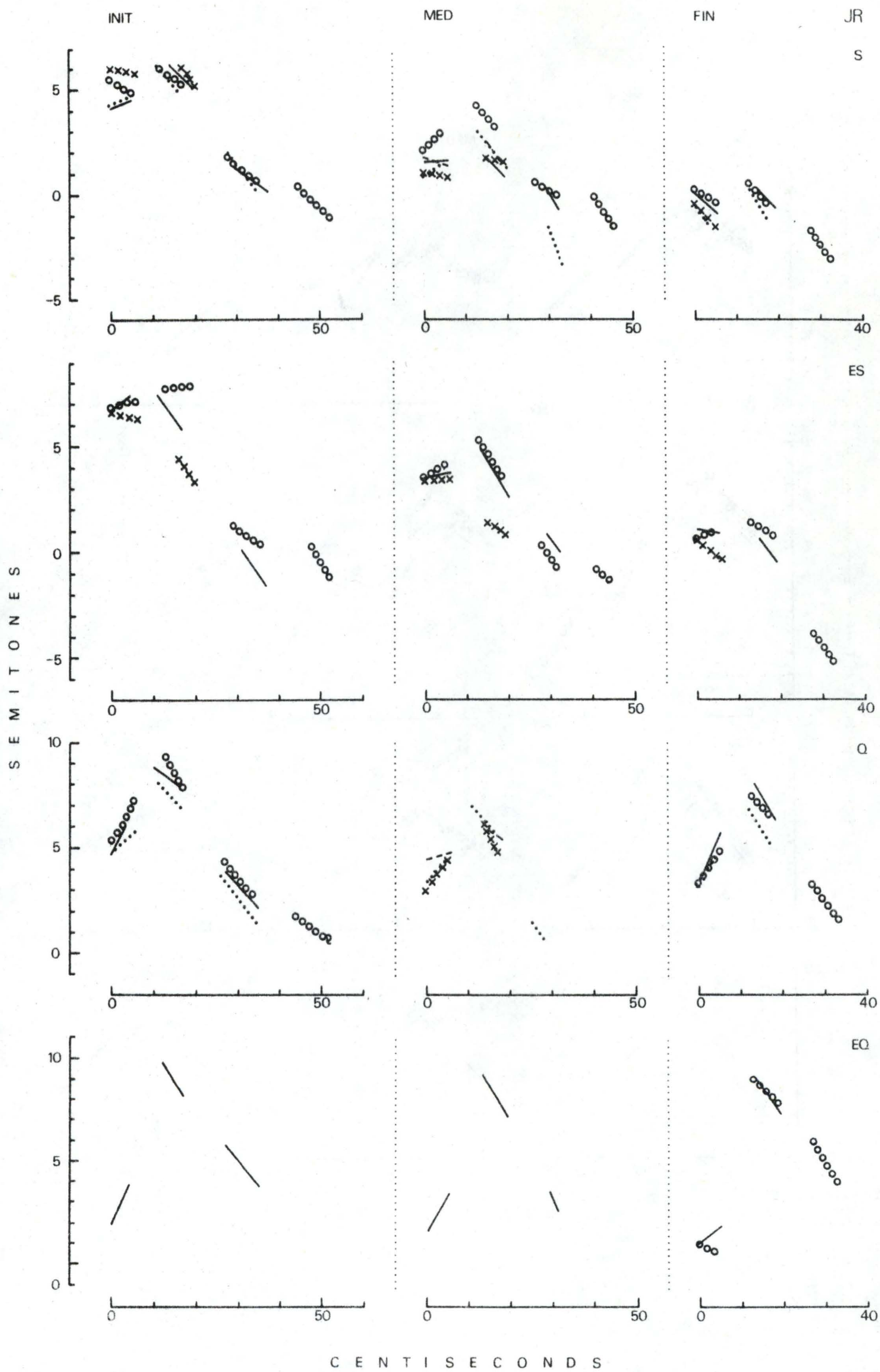


Figure 13
Subject JR. See further the legend to fig. 12.

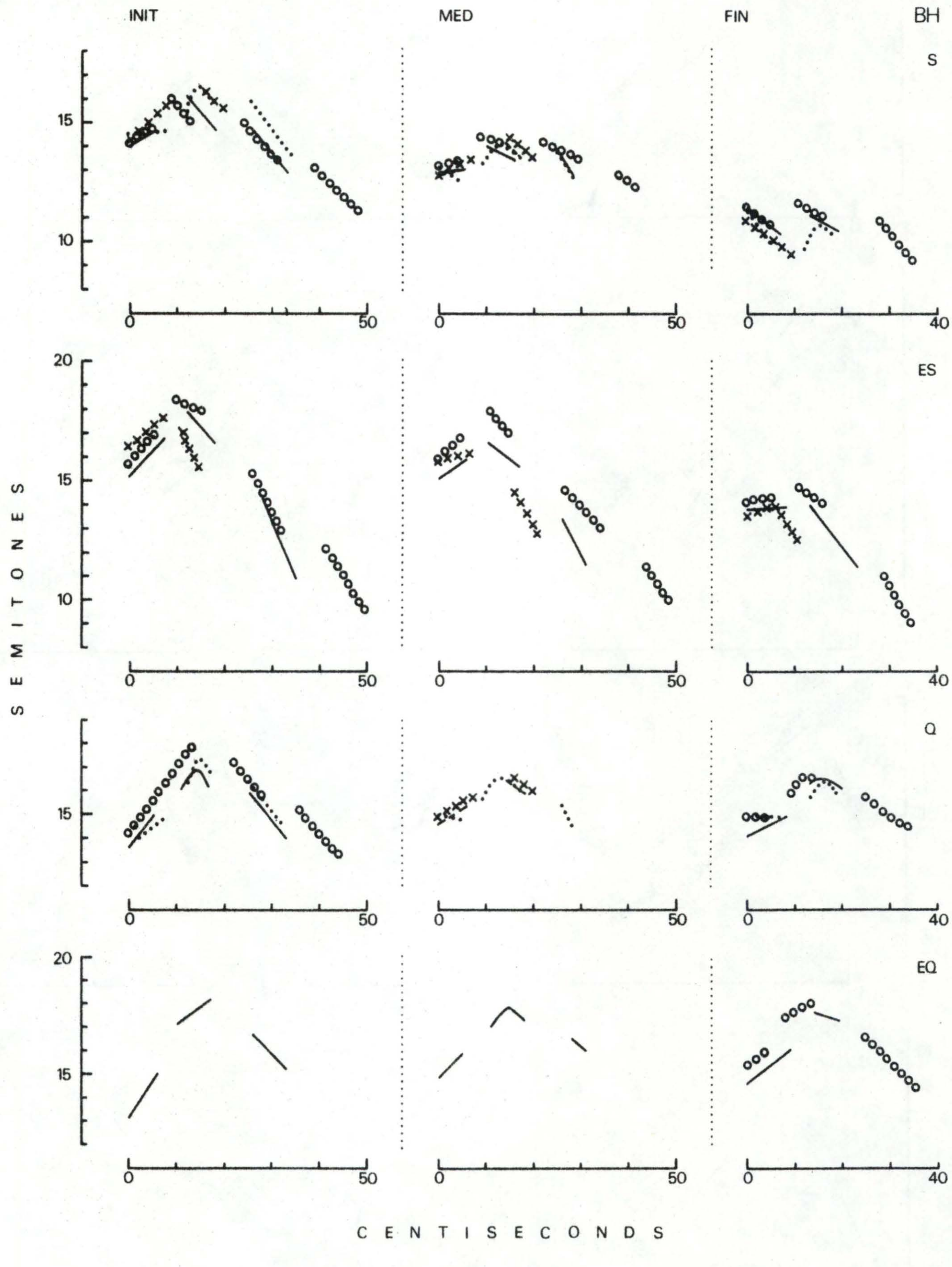


Figure 14

Subject BH. See further the legend to fig. 12.

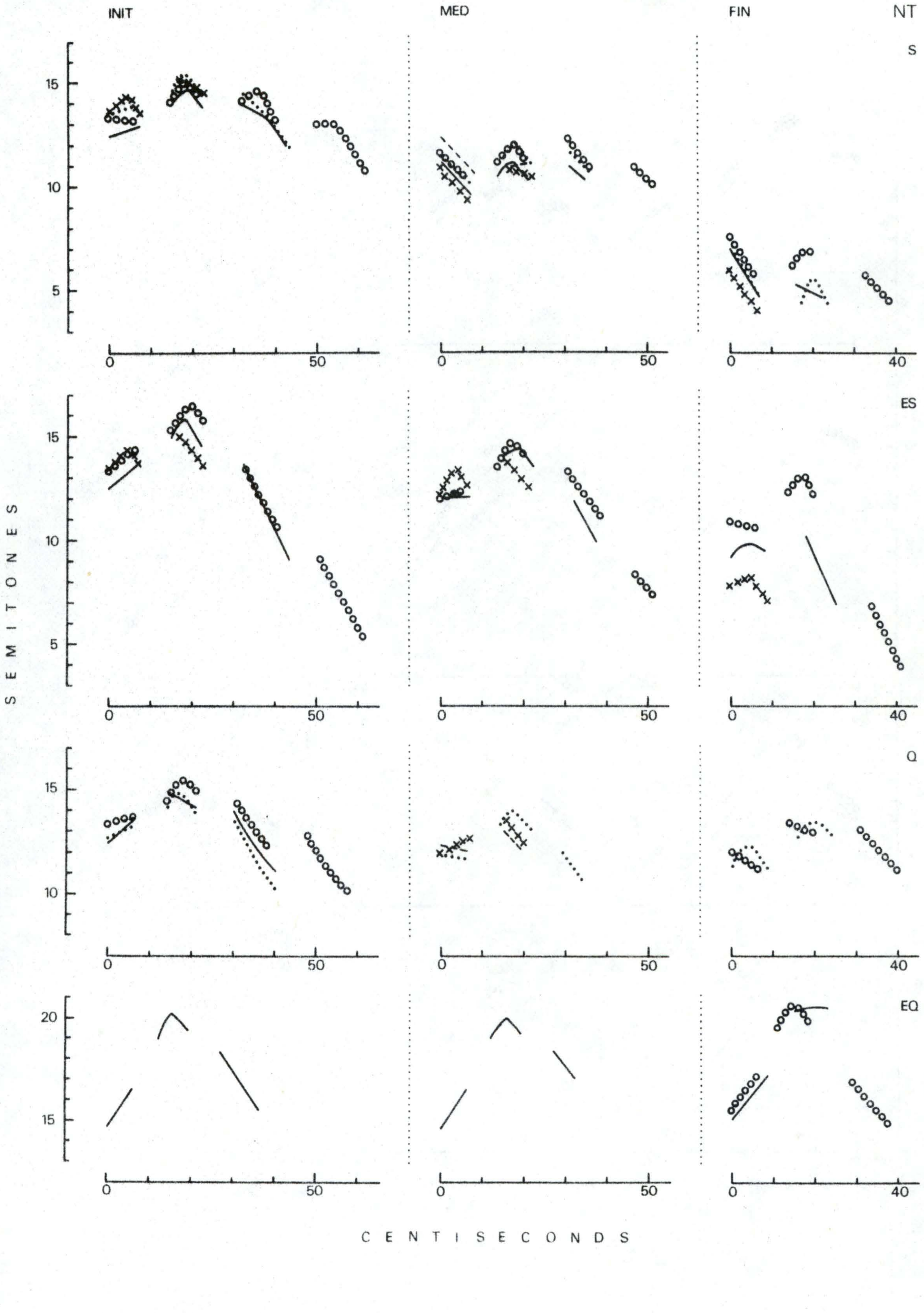


Figure 15

Subject NT. Average over ten readings.
See further the legend to fig. 12.

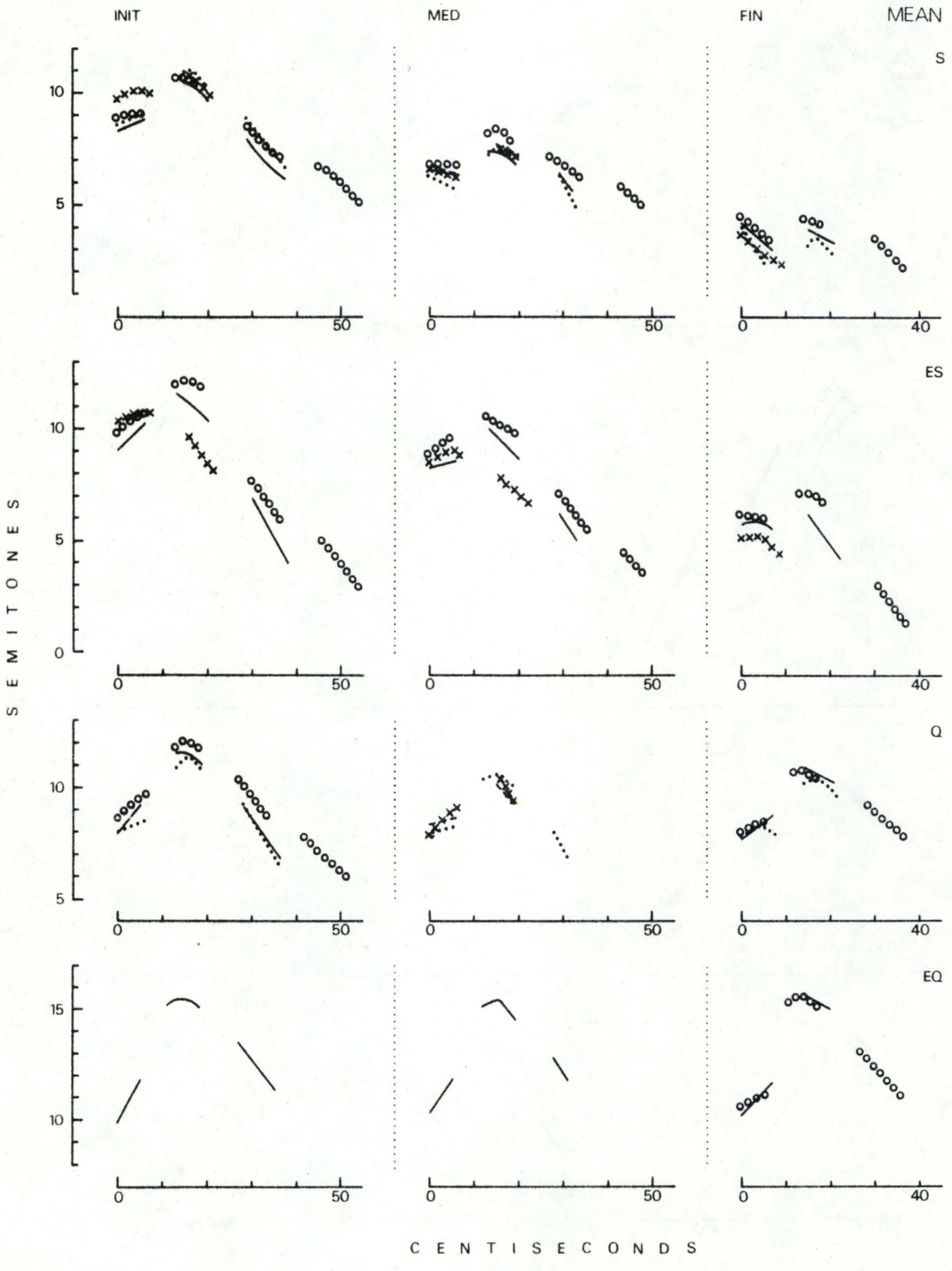


Figure 16

Average over four subjects (NS, ES, NQ) and three subjects (EQ), mean of means. See further the legend to fig. 12.

it appears that we must distinguish stress groups with one from stress groups with several post-tonics. - Accordingly, in some cases a common mean[1] over all words (and subjects) is possible; the result can be seen in fig. 17. Fig. 18 is a stylized version of fig. 17 where each vowel is depicted as a point (corresponding to a point in time at 2/3 of the distance from vowel onset or to the Fo maximum in rising-falling movements).

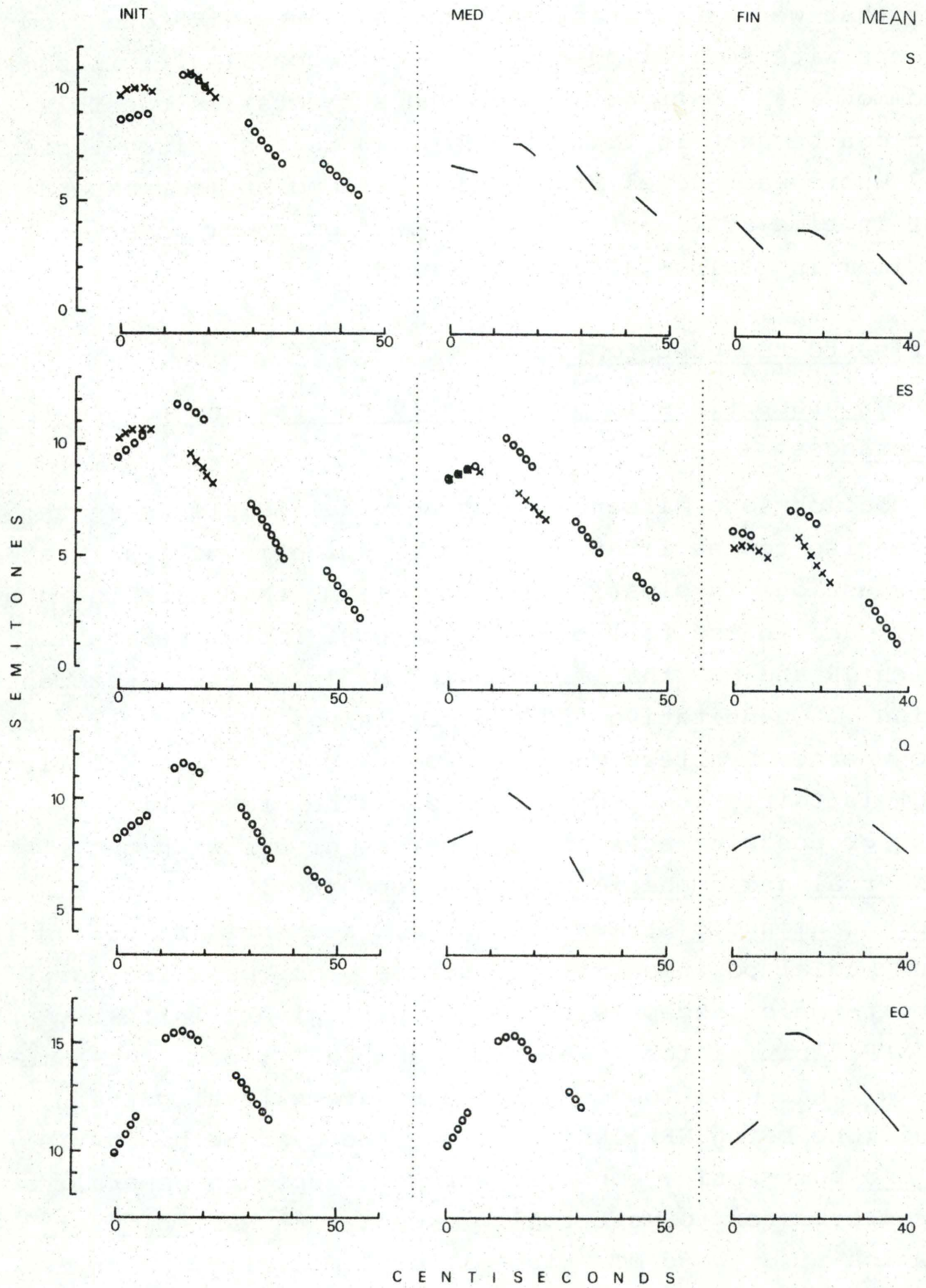
7.1 The rise to the post-tonic

7.1.1 Stress group rises in prosodically neutral and non-neutral utterances

The most obvious difference between individuals is in the size of the rise to the first post-tonic, which in neutral statements and questions is clearly larger with NRP than with the other subjects, and in the fall through the post-tonics, which is steeper with JR and NRP than with BH and NT in neutral statements and questions. Quantitative differences aside, the overall agreement as to whether a stress group rises or not (i.e. to the first post-tonic) is fairly good, cf. table 5. (The broken lines separate stress groups with several from stress groups with one post-tonic). Neutral and emphatic stress groups are invariably rising in questions. Emphatic stress groups in statements are rising in initial and medial position, if the stress group contains more than one post-tonic (except with JR "pi'pipi giver" initially and medially, NRP "'pipipi for-" medially, and BH "pi'pipi for-" medially). In final position two subjects have a level pattern, two have a rise (one being NRP, which is why the average has a rise). With only one post-tonic, emphatic stress groups are generally falling from stressed to post-tonic (except with NRP "pi'pipi" finally, which rises, and NT "pipi'pi for-" medially, which is level). The variation in neutral statements is slightly larger (cf. the speculations in sections 6.4.2.2 and 6.4.2.4 about distinctness of prosodic patterning on marked vs. unmarked intonation contours) but it does not seem an oversimplification to state that all stress groups are level in final position (i.e. a rise or fall is not larger than half a semitone), they are all

---

1) These means are calculated from the data behind figs. 12-15, not fig. 16, i.e. if e.g. three words are involved, then N=12 (four subjects times three words).



<u>Figure 17</u>

Fundamental frequency tracings of the nonsense words in fig. 16 where several words have been pooled, - see further section 7. The means are calculated from the data behind figs. 12-15, i.e. if e.g. three words are pooled, then N = 12 (four subjects times three words). <u>Full lines</u> indicate stress groups with only <u>one post-tonic</u> syllable. <u>Broken lines</u> indicate stress groups with <u>more than one</u> post-tonic. <u>Dotted lines</u> indicate that both stress groups with <u>only one as well as</u> stress groups with <u>several post-tonics</u> are involved, and <u>crosses</u> indicate that both stress groups with <u>no post-tonic as well as</u> stress groups with <u>one post-tonic</u> are involved. Zero on the logarithmic frequency scale corresponds to 100 Hz.

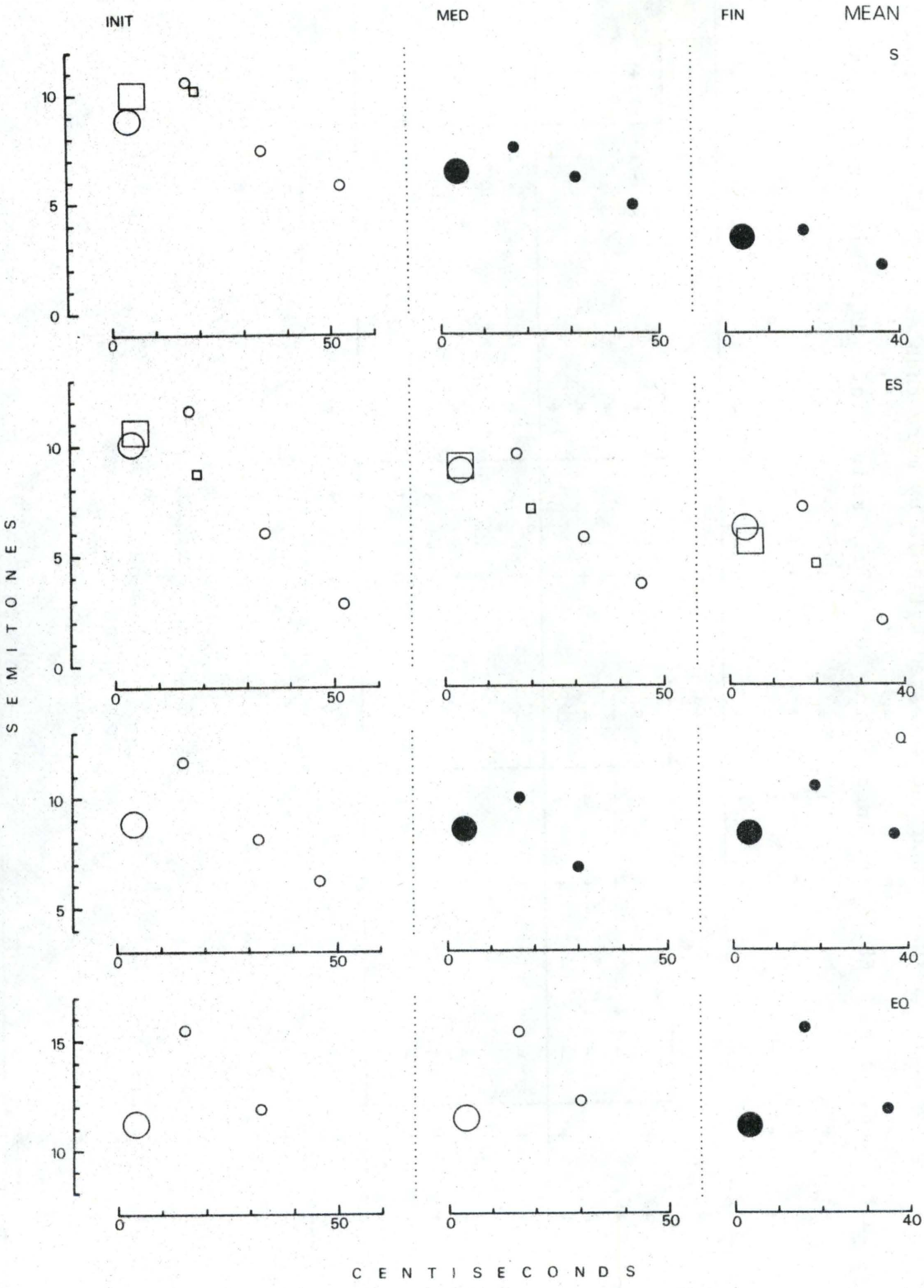


Figure 18

Stylized tracings of the data presented in fig. 17. Each vowel is depicted as a point in time at 2/3 of the distance from vowel onset, or at the maximum in rising-falling movements. Large circles and squares pertain to stressed syllables, small ones to unstressed syllables. Filled squares indicate stress groups with only one post-tonic syllable. Filled circles pertain to stress groups with more than one post-tonic. Empty circles indicate that both stress groups with only one as well as stress groups with several post-tonics are involved. Empty squares indicate that both stress groups with no post-tonic as well as stress groups with one post-tonic are involved.

Table 5

Qualitative account of the degree of rise from stressed to post-tonic syllable in the stress group patterns depicted in figs. 12-15 for individual subjects and in fig. 16 for the average, in neutral statements (S) and questions (Q) and in statements and questions with emphasis for contrast (ES, EQ). "+" indicates the presence of a rise of more than about half a semitone; "=" indicates that the stressed and post-tonic syllables are on about the same level; "-" indicates that the post-tonic is lower than the stressed syllable by more than about half a semitone.

		S: init.	med.	fin.	ES: init.	med.	fin.	Q: init.	med.	fin.	EQ: init.	med.	fin.
'pipipi	JR	=	+	=	+	+	=	+		+			+
(giver/for-)	NRP	+	+	+	+	=	+	+		+			
	BH	+	+	=	+	+	=	+		+			+
	NT	+	+	=	+	+	+	+		+			+
	average	+	+	=	+	+	+	+		+			+
(pi)'pipi	JR	+	=	=	=	=	-	+		+	+	+	+
(giver/for-)	NRP	+	+	+	+	+	+	+		+			
	BH	+	+	=	+	=	-	+		+	+	+	+
	NT	+	+	-	+	+	-	+		+	+	+	+
	average	+	+	=	+	+	-	+		+	+	+	+
'pipi	JR	+	+	=				+	+	+			
(giver/for-)	NRP	+	+	+				+	+	+			
	BH	+	+	=				+	+	+			
	NT	+	+	=				+	+	+			
	average	+	+	=				+	+	+			
(pipi)'pi	JR	-	+		-	-			+				
(giver/for-)	NRP	-	+		-	-			+				
	BH	-	+		-	-			+				
	NT	=	+		-	=			+				
	average	=	+		-	-			+				
(pi)'pi	JR		=						+				
(giver/for-)	NRP		+						+				
	BH		+						+				
	NT		=						+				
	average		+						+				

ARIPUC 14

e r r a t u m

On page 189, between lines 19 and 20, insert:

"rise was highest, when the post-tonic constituted a separate word the" rising in medial position, and rising in initial position except for stress groups with only one post-tonic, which are level initially. - Fig. 18 illustrates these observations. (The fact that the stressed syllable of the stress group with only one post-tonic is clearly higher initially in the neutral statement than the stressed syllable of stress groups with several may be accidental (there is only one word behind the squares in fig. 18, top left, whereas there are three behind the circles) or it may be an indication of "stress group elasticity", i.e. a compensation for the lack of rise to the post-tonic.)

The lack of rise to a single post-tonic in initial stress groups in prosodically neutral statements is not a must, however: The material analysed for Thorsen (1980c) was three sentence initial stress groups (segmentally identical, with only one post-tonic) that differed among themselves only in terms of the syntactic boundaries surrounding the post-tonic. The same subjects served for that investigation: NRP, BH, and NT invariably had rising Fo patterns; JR exhibited a syntactic boundary dependency: when the stressed and post-tonic syllables belonged to the same word, the rise was smaller and when the post-tonic belonged to the same word as the succeeding stressed syllable, stressed and post-tonic described a falling pattern. (JR does not exhibit a similar syntactic boundary dependency in this material, see further section 7.3.) I cannot say what triggers this difference in initial stress groups with only one post-tonic - the utterances of both materials sound alike, i.e. there are no apparent differences in their stress patterns in terms of slightly heavier stress (relative to the succeeding stressed syllables) on the first stressed syllable when it has a rising Fo pattern. It could be, of course, that since a level stressed and post-tonic initially in a statement can never resemble a succession of two unstressed or two stressed syllables - by virtue of the high position of the stressed syllable and the lack of a substantial fall (which characterizes a succession of two stressed syllables in short statement) - it is immaterial to the perception of stress distribution in such a stress group whether the post-tonic actually rises - and when it does it may be a sign of a more careful, i.e. maximally distinct, prosodic patterning (without invoking the impression of stronger stress). Such a distinctness may have been applied to the 1980c material, where the three utterances were mixed with utterances that served a different purpose and which were semantically and rhythmically very different, i.e. the three utterances stuck out clearly from the rest of the material. (If the lack of rise in initial stress groups with only one post-tonic were a constant phenomenon I would be hard put to explain why medial stress groups of a similar composition always come up with a rise - at least in materials analysed so far.)

The lack of rise, i.e. the falling Fo pattern in emphasized stress groups in statements supports the contention that time supersedes Fo when the two are in conflict (cf. section 2.2): In emphasized stress groups in statements the post-tonics must fall away rapidly from the emphatic syllable. It seems as though this fall must be accomplished, at least to a large extent, within the stress group itself which is why a stress group with only one post-tonic does not exhibit a rise, when emphasized: there is not time enough. (If there was no constraint on the domain of the post-emphatic fall there would be no reason why a single post-tonic (i.e. single in the prosodically neutral edition) should not rise, at least initially and medially, where the succeeding stress group(s) reduce tonally to a tail of unstressed syllables.) Whether a rise to a single post-tonic (a maximally distinct pattern) could be provoked with more favourable segmental conditions (long vowels and/or sonorant consonants) I cannot say. Fig. 10, top, which depicts a clear rise in the syllabic nasal of emphasized 'mange' cannot be considered conclusive evidence because 'mange' is not raised nor is the succeeding "stressed" syllable ('bus-') lowered relative to the prosodically neutral edition the way that is characteristic of utterances with nonsense words, i.e. the interval between the emphatic and succeeding "stressed" syllable is not particularly great.

Further speculations about the relationship between Fo patterning and rhythmic structure must await a separate investigation which should include careful manipulation of the duration of stress groups initially in neutral statements (partly by varying speaking rate, partly by varying the segmental composition: short/long stressed vowel, unvoiced/voiced post-vocalic consonant(s) in the stressed as well as post-tonic syllable).

7.1.2 Stress group rises as a function of intonation contour

Figs. 12-18 confirm the model in fig. 1 as far as higher rises to the post-tonic in neutral questions than in neutral statements go. - In section 2.1 I speculated that this variation in stress group pattern with intonation contour might be a physiologically conditioned assimilation: the lower the succeeding stressed syllable, the less of a rise is performed from stressed to post-tonic, and vice versa. However, if we are dealing with a physiological conditioned constraint, it is _not_ a strong one: The tendency towards higher rises on less falling contours is true only when averages of rather large data bases are considered. I noted at the end of section 6.4.1 that standard deviations on the means of Fo measurements are small, generally less than 5% of the mean (and smaller in stressed than in unstressed vowels). However, when the magnitude of the rise from stressed to post-tonic is calculated and averaged we get standard deviations which are often as great as or greater than the mean. There are two reasons for this: (a) with a dispersion of the same order of magnitude on Fo measurements in the stressed vowel, the post-tonic vowel, _and_ the rise from stressed to post-tonic, we will naturally get larger standard deviations on the mean of the rise, which is numerically much smaller than the means of stressed and post-tonic vowels. (b) Fo of the stressed and post-tonic (and succeeding stressed) vowel vary within rather narrow ranges, but they are not otherwise strongly correlated, i.e. a stressed vowel in the lower end of the range may be succeeded by a post-tonic in the upper end of the range (which may be succeeded by a stressed vowel in the lower end of the range). Documentation for this "production instability" (in sentence final stress groups) can be found in Thorsen (1980b, p. 1019-1021) and has been confirmed by spot checks in the present material. Thus, in individual recordings we may come across rises in statements which are higher than some rises in questions, and vice versa. Another point in disfavour of a strictly physiologically determined variation in stress group rises on falling and horizontal intonation contours is the fact that the _final_ stress group rise is also higher in questions than in statements, although there is no succeeding stressed syllable that can be made responsible for the variation. - On the other hand, to see higher rises on question contours as a controlled signal (added to the information contained in the course of the stressed syllables) of interrogative sentence intonation is not very appealing either, because you would then expect a greater production stability in these rises. - A third possibility presents itself, which ties up with the speculations in sections 6.4.2.2(c) and 6.4.2.4, namely that stress group patterns tend to be more distinct (i.e. have higher rises) on marked intonation contours, so in a sense this variation is speaker controlled but the demands for production stability are not severe. (If we assume that prosodic phrase group boundaries - which entail partial resettings in the course of long falling intonation contours (Thorsen 1980a) - receive special attention on the part of the speaker (which is not an unreasonable assumption) then the rises from stressed to post-tonic in stress groups preceding such boundaries, which are comparatively higher than if no boundary succeeds, can be explained along the same lines, i.e. as maximally distinct prosodic patterning.)

The successively lower rises from first through last stress group implied in fig. 1 (cf. section 2.1) is confirmed for the neutral statement - but not for the question, cf. fig. 18, top, and see also figs. 6, 8, and 9: the final rise is too high or, alternatively, the medial rise is too low. The lower medial rise compared to the final rise cannot be due to an artefact of the material because both stress groups in fig. 18 are the average of one stress group with one post-tonic and two stress groups with two post-tonics (times <u>four</u> subjects in both cases as well). Nor can it be due to deviant behaviour on the part of one particular subject because all subjects share this feature, cf. figs. 12-15. - Again, to see the comparatively higher rises <u>finally</u> in questions as an intonation contour signalling, i.e. as an acoustic cue to interrogative intonation contours, is not appealing due to the rather large dispersion in the magnitude of this rise, and because perception experiments show that the terminal rise is poorly correlated with listeners' identification of utterances as declarative, non-final, or interrogative, see further Thorsen (1980b, p. 1019-1021). (This rejection may be supported by the fact that the emphatic questions are not characterized by having comparatively higher rises when the emphasis is in the final stress group. Such a comparatively higher rise would not have been incongruous, because the difference between emphatic statements and questions lies mainly in the higher course of the "post-tonics" in questions, and with emphasis in initial or medial position there is a rather long tail to carry this high Fo course, but with emphasis in final position the question signalling must be carried by a few syllables only, which might have provoked a rise from emphatic to post-tonic greater than in initial and medial position.) If the extent of the final rise does not correlate with listeners' classifications of utterances as more or less interrogative, it cannot even be regarded as a redundant cue, i.e. it would have no perceptual purpose. So to account for the shape of the "topline" (the line connecting the first post-tonic in each stress group) in prosodically neutral questions it might be more profitable to not regard the final (or medial) stress group as being "aberrant" but to wonder whether "toplines" on marked contours may not be characterized by an asymptotic decline as compared to the linear decline on statement contours, i.e. we get a fall from first to second post-tonic, but after that the declination levels out. - More material, i.e. longer questions, is called for to further illuminate this point.

7.2 Intravocalic Fo movements

7.2.1 Long and short vowels

Long stressed vowel movements have been accounted for in section 6.4.2.2(a), which confirmed the observations in section 5 that on prosodically neutral falling intonation contours, long stressed vowels are falling-rising, but the fall is generally of greater extent than the rise, except initially in the utterance where the rise may be greater, probably due to the high start of the intonation contour. On horizontal intonation contours, long stressed vowels are also falling-rising, but the rise is slightly greater than the fall, which may be seen as an anticipation of, or assimilation to, the comparatively higher post-tonic syllable in questions. When subjected to emphasis for contrast, long stressed vowels are invariably rising, and more elaborately so in questions than in statements, cf. fig. 8. When subjected to stress reduction due to contrast emphasis in the environment, long stressed vowels exhibit a unidirectional movement, generally level or falling, regardless of the statement/question dichotomy, cf. figs. 2-5.

Short unstressed vowels are falling under all circumstances, except in the first post-tonic where it may be rising-falling, or even (but rarely) purely rising, cf. figs. 12-15 (NT is the only subject who has a majority of rising-falling post-tonic vowels).

Short stressed vowels were said in section 5 to be falling on falling contours, except initially, and to be less falling on less falling intonation contours. However, from figs. 12-15 it appears that initially in prosodically neutral statements, the majority of the nonsense words have rising movements, medially they are rising or falling, about fifty-fifty, finally they are all falling, but in the other conditions the majority of the words have rising stressed vowels, except finally in emphatic statements where the distribution is fifty-fifty. Thus, this is yet another instance of differences between different speech materials, because in the material analysed for the 1978 investigation, which led to the model in fig. 1, short stressed vowels (when succeeded by a post-tonic, see further below) were falling in the overwhelming majority of instances: The material consisted of nine different sentence types (five different questions, including a syntactically unmarked one, three different non-final periods, and a statement) all variations on the theme '(Der går) mange busser fra Tiflis' (which turned out with intonation contour slopes varying between the two extremes set by the statement and syntactically unmarked question, respectively). Four subjects (among them NRP and BH) recorded these sentences 6 times each: out of a total of 216 items, only four had non-falling stressed vowels in the final word ('Tiflis') (three of these occurred in the syntactically unmarked question). The stressed vowel of the medial word ('busser') was also invariably falling. The difference between the earlier and present materials cannot be ascribed to differences in segmental composition of the words, i.e. to different environments for the vowels, since they are unvoiced obstruents in all cases. The 1978 material also contained nine different sentences: '(Der er for) mange timer i statistik', and here there were more instances where the final, stressed vowel of 'statistik' was rising, invariably so in the syntactically unmarked question, and with a tendency for a correlation with intonation contour, i.e. more instances of and greater rises on less falling contours, and vice versa, i.e. more falling movements on more falling contours, and invariably falling in the statement (see further Thorsen 1980b, p. 1020-1021).

I do not think that the explanation for the greater number of instances with short stressed rising vowels in the present material should be sought in a difference between nonsense vs. real words. (The rationale behind such a difference would be a more distinct prosodic patterning of nonsense words to ensure the correct perception of stressed vs. unstressed syllables, and <u>if</u> we assume - as I did in section 5 - that the perceptual distance to the higher post-tonic is enhanced by a falling stressed vowel, then rising stressed syllables in nonsense words contradict such a hypothesis of more distinct prosodic patterning of nonsense vs. real words.) Rather, the reason for the difference in vowel movements in the two materials could be that in the present material the stressed syllables medially and finally were preceded by longer tails of unstressed syllables, to the effect that the end of these tails are lower on the frequency scale than the stressed syllable, which means that a rise is performed between the preceding unstressed and the stressed syllable, and since, further, the stressed syllable is succeeded by a rise to the post-tonic, it is not surprising, after all, that its own movement is not falling. In the 1978 material the stress groups were shorter and accordingly, more often than not, a fall was performed from the last unstressed syllable in the preceding stress group to the stressed syllable, which explains the generally falling movements. Note that this does not invalidate the contention in section 5 that vowel movements are predictable in terms of the Fo pattern, only surrounding (i.e. preceding) Fo patterns must be taken into account as well.

<u>7.2.2 Compression or truncation of Fo movements and -patterns</u>

There is ample evidence that the Fo <u>movement</u> in short vowels can be regarded as a truncation of the movement in long vowels, i.e. we never get in short vowels a compression of the falling-rising movement characteristic of long vowels.[1] This resembles the situation as described by Öhman (1965) for Swedish where "... the articulatory segments ... float on the tonal contour in certain ways." (p. 19), whereas Erikson (1973) says (also about Swedish) that "... the underlying rule is to synchronize the supraglottal events of a given syllable with its Fo variation." (p.28).

---

1) I.e. not under otherwise identical circumstances.

The results for vowel movements in the stressed (final) vowel of utterance final 'statistik' (cf. above) led naturally to considerations about (partial) compression or truncation of stress group patterns in stress groups with no post-tonic syllable(s), cf. section 2.3. I do not have documentation for stress groups without post-tonics in sentence initial and medial positions, but it seems reasonable to assume that if e.g. no compression takes place in sentence final position, the same will be valid in other positions in the sentence. - The fact that vowel movements in utterance final 'statistik' in statements (cf. above) is invariably falling, not falling-rising or plain rising, is one piece of evidence in favour of truncation of Fo patterns; secondly, the utterance final stressed vowels in figs. 12-15 do not seem to be systematically different from stressed vowels succeeded by a post-tonic, i.e. "pipi'pi" in the neutral statements is falling as in "'pipi, 'pipipi, pi'pipi", and in the emphatic statement it is clearly not more rising (if at all) than "'pipipi, pi'pipi". Unfortunately, there is a gap in the material, i.e. the questions do not encompass sentence final words with final stressed syllable, but the stressed vowel of "'pipi, 'pipipi, pi'pipi" in neutral and emphatic questions is rising anyway, and I do not think that utterance final stressed vowels will come up with any clear signs of a compressed stress group pattern in the shape of an even higher rise in the stressed vowel.

The fact that the final stressed vowel of 'statistik' is rising on intonation contours that are not steeply falling, whereas the stressed vowel of 'Tiflis' is falling under the same conditions (cf. above) can be viewed as an instance of Fo pattern compression, conditioned by the utterance final position in utterances with marked intonation contours, but it can also be viewed as a kind of assimilation due to the intonation contour, because it probably serves as a (redundant) cue to the perception of sentence intonation, cf. Thorsen (1980b, p. 1021). Thus, stress group pattern compression is not a general phenomenon, but restricted to utterance final stressed syllables on marked intonation contours (if we attribute the rise in such syllables to a compression at all).

7.3 Word boundaries in utterances with emphasis for contrast

An account of word boundaries in prosodically neutral statements has been given in Thorsen (1980c and 1980a). In the 1980c investigation (which was conducted prior to the 1980a one) JR deviated from the three ASC-speaking subjects (NRP, BH, and NT) by exhibiting signalling of syntactic boundaries (cf. section 7.1.1 above) but in the 1980a material the same four subjects all agreed on the non-existence of word (and syntactic) boundary signalling in the course of fundamental frequency. And likewise, in the present material no subject shows any signs of word boundaries in their tracings: If a boundary is signalled we should expect 'giver' ([g̊iʌ̯]) and 'for-' ([fʌ]) in "pi'pipi giver/pi'pipi for-" to look consistently different from, i.e. be either clearly higher or lower than, the last syllable of the nonsense word in "'pipipi giver/'pipipi for-" and if one looks at the third syllable from the left in figs. 12-15 initially and medially in neutral statements (S), emphatic statements (ES), and neutral questions (Q), no such consistent difference can be detected - on the contrary: 'giver' and 'for-' seem generally to be concurrent with pi-syllables in the same position in the stress group (especially if the stressed syllables are zero offset adjusted prior to comparison), i.e. the position on the frequency scale of 'giver' and 'for-' seems to be determined solely by their distance from the stressed syllable: they are higher, the closer they are to the preceding stressed syllable, just like post-tonics belonging to the same word as the stressed syllable. Likewise, pre-tonic pi-syllables show no signs of being associated tonally with the stressed syllable: the pre-tonics are not depicted in figs. 12-16, but even though figs. 2-5 only contain one type of nonsense word, "pi'pipi", this will suffice to prove the point (which is of course supported by that part of the material which is not shown here): The pre-tonic smoothly continues the fall in the preceding stress group's post-tonic tail, and the dissociation with the stressed syllable is especially clear when the nonsense word is emphasized, and the interval between pre-tonic and stressed syllable is considerable.

The conclusion to be drawn is identical to the one in Thorsen (1980a - this volume, at the end of section 5), and it gains in strength, I think, from the results on utterances with emphasis for contrast: The fact that word boundaries (which are also sometimes noun phrase/verb phrase boundaries) leave no trace in the course of fundamental frequency presents an argument in favour of a theory expounded in Selkirk (forthcoming) that prosodic categories (in casu: stress groups) are distinct entities in the phonology that do not have an isomorphous relation to syntactic structure. Rischel (1972) argues in a similar fashion: Danish stress is best represented in a hierarchy (a tree structure) which is not necessarily congruent with the syntactic structure. - (The autonomy of prosodic structure does not, of course, deprive it of a relation to syntax, on the contrary, prosodic categories can be seen as reconciling the syntactic structure to the phonetic output - in casu: the course of fundamental frequency.)

7.4 Stress group patterns - conclusion

As it appears from this section and section 2 above, stress group patterning is governed by several prosodic and structural properties of the utterance: There is a dependency of intonation contour, of time (or placement in the utterance), of prosodic neutrality or non-neutrality, of stress group composition (one vs. several post-tonics), all in a rather intricate interplay. However, regularities can still be formulated and stress group patterns predicted from the other properties of the utterance. Stress group patterns of stress groups with only one post-tonic initially on prosodically neutral unmarked intonation contours seem to be subject to assimilation or "slurring" under circumstances which might be termed "not maximally distinct patterning". - On one point the model requires a modification, namely where the magnitude of the rise to the post-tonic in stress groups on prosodically neutral question contours are concerned. But since this calls for a supplementary investigation, I shall leave the model in fig. 1 as it is for the time being and remind the reader that it covers short prosodically neutral utterances only - for an accout of utterances with emphasis for contrast, one must resort to figs. 6-9 and 18.

Finally, I should add that I do not consider the stress group and its properties a closed chapter - evidently its temporal properties need to be investigated.

8. Postscriptum about the variety of Danish investigated

Three subjects in the investigation speak the Copenhagen variety for which Basbøll (1968) coined the term 'Advanced Standard Copenhagen': "ASC ... is the language spoken by a large group of the younger generation in Copenhagen, whose language is normally considered to be a variety of Standard Danish. In many respects it differs from what might to-day be called Conservative Standard Danish ..., the language described by Jespersen [1934, NT] ... ASC is clearly different from both the Copenhagen dialect (sometimes termed "vulgar") and the language (sometimes termed "affected") spoken by the upper class in the northern parts of Copenhagen." (p. 33). The fourth subject (JR) speaks a slightly more conservative variant (described in Rischel 1968).

The varieties of Danish treated in Basbøll (1968) and Rischel (1968) differ with regard to both vowel and consonant systems (and stød), but if JR can be considered representative of the slightly conservative Copenhagen norm, then it seems that as far as intonation contours and stress group patterns (in prosodically neutral utterances as well as in utterances with emphasis for contrast) are concerned, the two varities of Copenhagen Danish do not differ in any systematic respect from each other. (In Thorsen 1980a, section 5, I argue that the word boundary signalling that JR exhibited in the 1980c material does not constitute an example of a difference between ASC and more conservative norms. Rather, it demonstrates that it is <u>possible</u> for a speaker to signal word boundaries, also with fundamental frequency, if he so desires, a possibility which is presumably also open to ASC speakers.) - Thus, I think that the limitation implied by the term 'advanced' may not be necessary, and the results of the various intonation investigations that have been conducted can probably be taken to extend to Copenhagen Danish in general. - Furthermore, the term 'advanced' may be inappropriate to-day for the simple reason that the younger generation in 1968 are now approaching their forties, and there are certainly more "advanced" norms being spoken in Copenhagen now.

9. Postscriptum about reduced main stresses

In sections 6.2, 6.3, and 6.4.2.4 an argument was presented that main stresses in prosodically neutral utterances are reduced to secondary stresses when a contrast emphasis occurs in the immediate surroundings, i.e. the rise from stressed to post-tonic, characteristic of prosodically neutral utterances, is deleted but otherwise the "stressed" syllable retains all of its stressed syllable properties (vowel quality, vowel quantity, and stød, if any). It occurs to me, after the bulk of the manuscript has been typed out, that one might object that the kind of nonsense material on which the main part of this investigation is based, poses special problems because pre-tonic, stressed, and post-tonic syllables in the nonsense words are exactly identical, and a distinction between secondary and weak stress cannot be made on any of the criteria mentioned above. So how does a listener distinguish secondary from weak stress? - The material is ill suited to make systematic comparisons between nonsense words with different stress placement in reduced stress groups: there are really only two such words, in sentences 11 and 12. Inspection of the traces of these two sentences by JR, BH, and NT (NRP's recordings were discarded) reveal no consistent Fo differences (and nor should we expect any) between the two words, but durations come out with clear differences in the pre-vocalic consonants, differences which recur in all the rest of the material: the prevocalic [b̥ʰi] is significantly longer in the reduced main stress than in pre- and post-tonic syllables. Differences in vowel duration are generally small and show no similar clear trend, although there is a tendency for the pre-tonic vowel to be shorter, and the post-tonic vowel to be longer than the reduced main stress. Presumably, the longer pre-vocalic consonant is sufficient to identify the reduced main stress. (There is no doubt, to me, that the reduced main stresses are identifiable, and very clearly so - whether preception experiments would come out with unambiguous results I do not know; it makes little sense to run such experiments on the present material, since stress placement is not systematically varied in the reduced stress groups.)

Acknowledgements

Thanks to Hans Basbøll, Eli Fischer-Jørgensen, and Oluf Thorsen who made valuable suggestions for improvements in the manuscript. Special thanks to Jeanette Holtse who competently performed all of the tedious measuring and calculating as well as the graphics work.

References

Abercrombie, D. 1964: "Syllable quantity and enclitics in English", in In Honour of Daniel Jones (eds.: D. Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott, and J.L.M. Trim), Longmans, London p. 216-222

Basbøll, H. 1968: "The phoneme system of Advanced Standard Copenhagen", ARIPUC 3, p. 33-54

Basbøll, H. 1972: "Some remarks concerning the stød in a generative grammar of Danish", in Derivational Processes, Proceedings of the KVAL Spring Seminar April 9-10, 1972.

Basbøll, H. 1978: "A note on boundaries and stress rules in Danish phonology", in Nordic Prosody (eds.: E. Gårding, G. Bruce and R. Bannert), Department of Linguistics, Lund University, p. 65-71

Berinstein, A.E. 1979: "Stress", UCLA WPP 47, p. i-59

Bo, A. 1933: Tonegangen i Dansk Rigsmål, Studier fra Sprog- og Oldtidsforskning, 164, Branner, København

Bolinger, D.L. 1958: "A theory of pitch accent in English", Word 14, p. 109-149

Bolinger, D.L. 1962: "Intonation as a universal", Proc. Ling. 9, p. 833-848

Bolinger, D.L. 1970: "Relative height", in Prosodic Feature Analysis (eds.: P. Léon, G. Faure, and A. Rigault), Marcel Didier, Paris, p. 109-125

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

Carlson, R., Y. Erikson, B. Granström, B. Lindblom and K. Rapp 1974: "Neutral and emphatic stress patterns in Swedish", Preprints from Speech Communication Seminar, Stockholm, Aug. 1-3, vol. II, p. 209-218

Cohen, A. and J. 't Hart 1967: "On the anatomy of intonation", Lingua 19, p. 177-192

Collier, R. and J. 't Hart 1975: "The role of intonation in speech perception", in Structure and Process in Speech Perception (eds.: A. Cohen and S.G. Nooteboom), Springer-Verlag, Berlin, p. 107-121

Crystal, D. 1969: Prosodic Systems and Intonation in English, Cambridge University Press, London

Daneš, F. 1960: "Sentence intonation from a functional point of view", Word 16, p. 34-54

Donovan, A. and C.J. Darwin 1979: "The perceived rhythm of speech", Proc.Phon. 9, vol. II, p. 268-274

Erikson, Y. 1973: "Preliminary evidence of syllable-locked temporal control of Fo", STL-QPSR 2-3, p. 23-30

von Essen, O. 1956: "Hochdeutsche Satzmelodie", Zs.f.Ph. 9, p. 75-85

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

Fischer-Jørgensen, E. 1948: "Some remarks on the function of stress with special reference to the Germanic languages", Congr. Intern. Sc. Anthropol. and Ethnol., Comptes-Rendus, IIIe session, Bruxelles 1948, p. 86-88 - also in: 25 Years' Phonological Comments, Wilhelm Fink Verlag, München, 1979, p. 55-57

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

Fujisaki, H., K. Hirose, and K. Ohta 1979: "Acoustic features of the fundamental frequency contours of declarative sentences in Japanese", Annual Bulletin, Research Institute of Logopedics and Phoniatrics 13, p. 163-173

Gårding, E. 1977: "The Scandinavian Word Accents", Travaux de l'Institut de Linguistique de Lund XI, p. 1-116

Gårding, E. 1980: Introduction to the discussion in the symposium on 'The relation between sentence prosody and word prosody', Proc.Phon. 9, vol. III, p. 283-287

Gårding, E. and P. Lindblad 1973: "Constancy and variation in Swedish word accent patterns", Phonetics Laboratory, Lund University, Working Papers 7, p. 36-110

Hadding-Koch, K. 1961: "Acoustico-phonetic studies in the intonation of southern Swedish", Travaux de l'Institut de Linguistique de Lund 3, p. 1-211

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

't Hart, J. 1966: "Perceptual analysis of Dutch intonation features", IPO APR 1, p. 47-51

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

't Hart, J. and R. Collier 1979: "On the interaction of accentuation and intonation in Dutch", Proc.Phon. 9, vol. II, p. 395-402

Heger, S. 1975: Tale og tegn, Elementær Dansk Fonetik 2, Gjellerup, København

Hombert, J.-M. 1977: "Development of tones from vowel height?", JPh 5, p. 9-16

Jespersen, O. 1897-99: Fonetik, Det Schubotheske Forlag, København

Jespersen, O. 1934: Modersmålets Fonetik, 3rd edition, 5th reprinting 1966, Gyldendal, København

Jones, D. 1960: An Outline of English Phonetics, 9th edition (1967 reprint), Heffer & Sons Ltd., Cambridge

Lehiste, I. 1977: "Isochrony reconsidered", JPh 5, p. 253-263

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

Liberman, M. and A. Prince 1977: "On stress and linguistic rhythm", Linguistic Inquiry 8, p. 249-336

Liberman, M.Y. and L.A. Streeter 1978: "Use of nonsense-syllable mimicry in the study of prosodic phenomena", JASA 63, p. 231-233

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

Lyberg, B. 1979: "Final lengthening - partly a consequence of restrictions on the speed of fundamental frequency change?", JPh 7, p. 187-196

Mangold, M. 1975: "Phonetic emphasis, a study in language universals", Forum Phoneticum 10, p. 1-161

Mikoš, M.J. 1976: "Intonation of questions in Polish", JPh 4, p. 247-253

Nakatani, L.H. and J.A. Schaffer 1978: "Hearing 'words' without words: prosodic cues for word perception", JASA 63, p. 234-245

Öhman, S.E.G. 1965: "On the coordination of articulatory and phonatory activity in the production of Swedish tonal accents", STL-QPSR 2, p. 14-19

Öhman, S.E.G. 1968: "A model of word and sentence intonation", STL-QPSR 2-3, p. 6-11

Riber Petersen, P. 1973: "An instrumental investigation of Danish 'stød'", ARIPUC 7, p. 195-234

Rischel, J. 1964: "Stress, juncture, and syllabification in phonemic description", Proc.Ling. 9, p. 85-93

Rischel, J. 1968: "Notes on the Danish vowel pattern", ARIPUC 3, p. 177-205

Rischel, J. 1972: "Compound stress in Danish without a cycle", ARIPUC 6, p. 211-228

Rischel, J. 1975: "Problemer ved en generativ beskrivelse af dansk tryk", Selskab for Nordisk Filologi. Årsberetning for 1971-73, p. 22-32

Selkirk, E.O. forthcoming: "On the nature of phonological representation", in The Cognitive Representation of Speech (eds.: J. Anderson, J. Laver, and T. Myers), North Holland Publishing Company

Stock, E. 1980: Untersuchungen zu Form, Bedeutung und Funktion der Intonation im Deutschen, Akademie-Verlag, Berlin

Terken, J.M.B. 1980: "The relation between information structure and accentuation: some exploratory investigations", IPO Rep. no. 381, p. 1-40

Thorsen, N. 1978: "An acoustical analysis of Danish intonation", JPh 6, p. 151-175 (a more extensive version is published in ARIPUC 10, 1976, p. 85-147)

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

Thorsen, N. 1980a: "Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish", ARIPUC 14, p. 1-29 (this vol.)

Thorsen, N. 1980b: "A study of the perception of sentence intonation - Evidence from Danish", JASA 67, p. 1014-1030

Thorsen, N. 1980c: "Word boundaries and Fo patterns in Advanced Standard Copenhagen Danish", Phonetica 37, p. 121-133 (also in ARIPUC 13, 1979, p. 121-134)

Thorsen, N. (forthcoming): Selected problems in the tonal manifestation of words containing assimilated or elided /ə/ (in preparation)

Williamson, K. 1979: "Sentence tone in some Southern Nigerian languages", Proc.Phon. 9, vol. II, p. 424-430

(ARIPUC 14, 1980)

TEMPORAL RELATIONS IN DANISH TAUTOSYLLABIC CV SEQUENCES WITH STOP CONSONANTS[1]

Eli Fischer-Jørgensen

Abstract: This paper presents some facts about the duration of closure and open interval in Danish stop consonants (5.1 - 5.2). 5.3 deals specifically with the relation between durations of closure, open interval and following vowel (1) in ptk vs. bdg, (2) depending on place of articulation of the consonant, (3) in connection with long and short vowels, (4) with high and low vowels, and (5) with rounded and unrounded vowels. In section 6 it is discussed whether the relations found are due to compensation within the CV sequence or to specific production mechanisms. It is suggested that only vowel length in (1) may be due to compensation.

1. Purpose

The purpose of this paper is to present some data on the temporal relations in CV sequences with Danish stops, and to discuss the possible explanations of these relations, in particular whether they can be assumed to be due to compensation phenomena or to specific physiological conditions.

2. Material

The Danish stop consonants are /ptk/ and /bdg/. These two classes are distinguished phonologically in syllable initial posi-

---

1) This is an enlarged and revised version of Eli Fischer-Jørgensen 1979. It contains a much more extensive documentation, and some mistakes have been corrected. In the last section references to findings in other languages have been added, and more arguments and points of view are included in the discussion. I am very grateful to Birgit Hutters for a number of critical remarks which have resulted in considerable improvements of the paper. -

This investigation has been supported by the Danish Research Council for the Humanities.

sition before (sonorant consonant +) full vowels only. /ptk/ are voiceless aspirated stops [ph th kh]. /bdg/ are voiceless unaspirated stops [b̥ d̥ g̊]. Medially before [ə] and the weak endings <u>ig</u> [i] and <u>ing</u> [eŋ], as well as finally, only one set of stops is found, which is almost always pronounced as (very) weakly voiced [bdg] medially, whereas finally there is free variation between [ph th kh] and [b̥ d̥ g̊], [ph th kh] being used mainly in absolute final position. This paper deals exclusively with the position in which /ptk/ and /bdg/ are distinguished, and only the position before stressed vowels is taken into account.

Some of the material presented is not of recent date. But these old recordings have either not been utilized at all, or only partly so, or for other purposes. They have been supplemented with one rather extensive more recent recording, which is part of an EMG-investigation. The recordings are either spectrographic (indicated by the letter S) or mingographic (indicated by the letter M). The latter recordings comprise duplex oscillograms, Fo curves, and intensity curves. In cases of doubt they have been supplemented by spectrograms.

S1 (1952) contains 17-45 words spoken by 9 different speakers. As the occurrence of different stops under different conditions is not very regular, this list is only referred to very briefly. (S2 is a very short list not utilized in this paper.)

S3 (1954) comprises 60 words of the type <u>CV:nə</u>, where C represents /p t k b d g/ and V the 10 long Danish vowels /i: e: ɛ: a: y: ø: œ: u: o: ɔ:/. It has been spoken twice by two subjects (No. 5 and No. 9) and once by one subject (No. 6).

S4 (1954) comprises 30 words of the type CVC, where C- and -C are phonetically identical [ph th kh], and V represents the ten short Danish vowels. This list was spoken twice by two subjects (No. 5 and No. 9) and once by one subject (No. 6).[1]

S5A and B (1956) comprise 18 different words of the type <u>CV:lə</u> and <u>CVlə</u>, where C represents /p t k/. In S5A the vowels are /a(:) i(:) u(:)/, in S5B only /a(:)/ occurs. The lists have been spoken by 5 subjects (Nos. 1, 5, 6, 8, and 9), S5A 6 times and

---

1) The recordings of lists S3 and S4 were made in Stockholm through the courtesy of Gunnar Fant at a time when we did not yet have any sonagraph in Copenhagen.

S5B 12 times.

S6 (1957) comprises 60 words of the type CV:lə, where C represents /p t k/ and /b d g/ and V the ten long Danish vowels. This list has been spoken once by 3 subjects (Nos. 5, 6, and 9).

M1 (1964). M1A comprises 9 words of the type CVCə, where C- and -C are phonetically identical and represent [bdg] and V is short /i y u/. M1B comprises 12 words of the type CVCə, where C- is /p t k b d g/ and -C always [d]; V is short /u/ and /i/. M1A has been spoken 12 times and M1B 6 times by 7 subjects (Nos. 1, 2, 5, 7, 10, 11, and 12).

M2 (1974) comprises 12 words of the type CV:lə or CVlə, where C represents /p t k/ and /b d g/ and V /a(:)/ and /i(:)/ (for Nos. 7 and 17 also /u(:)/. This list has been spoken 10 times by 6 subjects (3, 4, 7, 13, 15, 18), 6 times by 2 subjects (14 and 16), and twice by 2 subjects (2 and 17).[1]

The lists contain a mixture of existing words and nonsense words, except for the last list (M2), which contains existing words only. In list M2 the phonological vowel length is different for /i/ and /a/ except where velars are involved. This has given some restrictions on the comparisons of vowels.

On the whole, the purpose has been somewhat different for different lists. S4, S6 and M1 consisted of isolated words which do not permit the measurement of the closure duration. In S3 the words were preceded by [ɔ] ('and') (but subject No. 6 made a small pause after it, so that his closures cannot be measured). The words of S5 and M2 were said in a frame: [di sæ:] —— or [han sæ:] —— or [de vɑ —— di sæ:] or [de heðɒ] —— ("they said ——", "he said ——", "it was —— they said", and "it is called ——").

A restricted number of words with initial consonants other than stops have been measured for comparison, namely in list S3 30 different words (120 in all) with initial /f s h/, and in list M2 8 different words with /f v s h/ (672 in all) and 4 different words with /m n l/ only spoken by some subjects (136 in all).

The whole material comprises about 4000 tokens.

---

1) This list was made for the purpose of an extensive EMG investigation, and the recordings and measurements have been made in cooperation with Jørgen Rischel, Birgit Hutters and Anders Löfqvist.

3. Subjects

There are 18 subjects; they were all students or teachers of phonetics (with one exception, No. 13 (MF), who is a phoniatrist). Phoneticians have the advantage that they are able to speak lists containing nonsense words more naturally and freely than phonetically naïve people. Four of the subjects grew up in Jutland: No. 1 OT, born in 1928, No. 2 BF, born in 1935, No. 3 PM, born in 1946, and No. 4 BM, born in 1947 (BM, however, moved to Copenhagen at the age of seven). Three grew up in Funen: No. 5 FJ (the author, born in 1911), No. 6 NK, born in 1915, and No. 7 JR, born in 1934, but only No. 6 is influenced by the dialect. Nos. 5 and 7 spoke Standard Danish at home, No. 5 also with playmates. Eleven subjects grew up in Copenhagen or in North Zealand: No. 8 KS, born in 1928, No. 9 BL, born in 1930, No. 10 KB, born in 1936, No. 11 BS, born in 1922, No. 12 HP, born in 1938, No. 13 MF, born in 1939, No. 14 HU, born in 1945, No. 15 JJ, born in 1949, No. 16 BH, born in 1945, No. 17 NR, born in 1942, and No. 18 LG, born in 1946).

4. Measurements and delimitation

The following measurements were made: (1) duration of the closure, (2) duration of the open interval, (3) duration of the following vowel. The beginning of the closure has been determined as the point where the vowel formants stop on the sonagrams, and the intensity and frequency curves go abruptly down on the mingograms. This may be about 1 cs before the real occlusion, but this point can be identified in a more consistent way on the curves.

The term "open interval" (which has been taken over from an early paper by Gunnar Fant) indicates the distance from the start of the release to the start of the vowel. It thus includes (a) the transient phase, (b) the frication phase, and (c) the aspiration phase (Fant 1960). It would have been preferable to make these further delimitations, but very often they could not be made with certainty (in S3 frication and aspiration have been delimited for t). Generally, "aspiration" is used in the sense of "open interval" for ptk, but it is not a very appropriate term for bdg. Fant (1970) uses "burst", but this may also indicate the release transient only. More recently, (positive) VOT duration is used in the same sense. But, as the start of the vowel and not the start of the vibrations is chosen as limit in this investigation, "voice onset time" would be a slightly misleading term. In most cases start of vibrations and start of vowel coincide. But when p and k are followed by an open vowel, particularly a, some speakers start the voicing before the vocal cords have come together, which gives low frequency voicing with gradual rise of the intensity curve and low rising Fo before the formants are fully developed (the vibrations are often very clearly seen on the oscillogram). In the present investigation pa and ka were spoken in this way by subjects No. 1, 4, 5, 13, 14, 16, and 18 and sometimes by Nos. 8 and 9.

Statistical significance has been tested by means of a simple pair test, based on averages for specific consonants, contexts, and subjects, and on a Poisson test.

The consonant following the vowel is not taken into account in this investigation. In disyllabics with long vowel (S3, S6 and parts of S5 and M2) the consonant belongs to the following syllable. In monosyllabics of the CVC-type (S4) it evidently belongs to that syllable, and in disyllabics with short vowel (M1 and parts of S5 and M2) it should probably be considered ambisyllabic. (Accordingly, the term "CV-syllables" used in my 1979 paper has in this paper been changed to the more correct "CV-sequences".) Since compensations are normally supposed to take place within the syllable, one might argue that the consonant should have been included. However, some earlier measurements seem to indicate that this measure would be irrelevant, at least for Danish. In Fischer-Jørgensen 1955 the duration of the consonant l was measured after phonologically long and short vowels and no significant difference found, and in Fischer-Jørgensen 1964 bdg were measured after long and short vowels (in list M1B) and no consistent difference found. Moreover, in a sample of words from list S3 (type: CV:nə) the consonant n has been measured, and no dependence on the initial consonant, nor on the degree of opening of the vowel was found. It therefore seemed legitimate to examine CV-sequences separately.

5. Results

5.1 Individual averages of the duration of closure and open interval

The average durations of closure and open interval for stop consonants in the various lists are shown in tables 1 - 6. Only in the cases where the words were spoken in a frame has it been possible to measure the closure (tables 1, 3, 4, and 6).

Table 1

List S3. Average durations of closure in cs and open interval in cs and in % of closure + open interval (ptk and bdg + ten long vowels)

			p			t			k		
S		N	clos.	op.i.	%	clos.	op.i.	%	clos.	op.i.	%
5	FJ	20	8.8	5.2	37	6.9	7.1	51	8.6	6.1	42
6	NK	10	(14.2)	5.4	-	(11.5)	7.1	-	(12.4)	6.6	-
9	BL	20	11.1	6.2	36	8.8	9.2	51	9.7	6.4	40
			b			d			g		
5	FJ	20	10.4	1.2	10	10.3	1.2	10	10.6	1.6	13
6	NK	10	(17.0)	1.5	-	(15.9)	1.7		(15.2)	2.7	-
9	BL	20	14.5	1.5	9	14.2	2.0	12	12.6	2.7	18

Table 2

List S4. Average durations of open interval in cs for ptk + ten short vowels + ptk.

S		N	p	t	k
5	FJ	20	5.3	6.3	6.2
6	NK	10	3.7	6.2	6.5
9	BL	20	4.1	6.9	5.6

Table 3

List S5. Average durations of closure in cs, and open interval in cs and in % of closure + open interval (ptk + a(:) u(:) i(:)).

			p			t			k		
S		N	clos.	op.i.	%	clos.	op.i.	%	clos.	op.i.	%
1	OT	30	9.5	7.8	45	8.8	8.5	49	9.6	8.1	46
5	FJ	30	8.4	6.2	42	5.7	8.9	61	7.6	7.0	48
6	NK	30	11.9	4.8	29	9.4	8.2	47	10.0	7.7	44
8	KS	30	8.8	5.3	38	7.0	7.5	52	8.0	6.4	44
9	BL	30	10.0	7.7	44	7.1	9.3	57	8.3	7.8	49

Table 4

List S6. Average durations of open interval of ptk and bdg in cs (ptk and bdg + ten long vowels).

			p	t	k	b	d	g
5	FJ	10	6.3	6.5	6.3	1.7	1.8	2.4
6	NK	10	6.0	8.5	7.9	2.2	2.9	3.9
9	BL	10	8.2	9.4	8.4	2.4	3.0	4.2

Table 5

List M1. Average durations of open interval of ptk and bdg in cs (M1A: bdg + short i y u M1B: ptk and bdg + short i u).

	M1A				M1B						
S	N	b	d	g	N	p	t	k	b	d	g
1 OT	36	3.7	3.2	3.9	12	5.7	6.8	7.0	2.2	2.6	4.0
2 BF	36	1.8	1.9	2.5	12	5.3	5.4	6.5	1.9	2.3	2.9
5 FJ	36	1.5	2.4	2.2	12	5.0	6.4	6.0	1.4	1.6	2.5
7 JR	36	1.8	2.2	2.7	12	5.6	5.7	6.8	2.0	2.2	2.9
10 KB	36	2.2	1.9	3.0	12	7.3	7.8	8.3	2.0	2.0	3.3
11 BS	36	2.5	2.1	3.1	12	6.7	8.4	9.1	2.8	2.6	3.4
12 HP	36	2.2	2.2	3.2	12	7.6	9.2	9.4	1.9	2.2	3.4

Table 6

List M2. Average durations of closure in cs and open interval in cs and in % of closure + open interval (ptk and bdg + a i· (u)).

S		N	p clos.	p op.i.	p %	t clos.	t op.i.	t %	k clos.	k op.i.	k %
2	BF	4	12.0	6.3	34	8.5	9.5	53	9.0	9.3	51
3	PM	20	8.6	6.0	41	7.1	7.7	52	7.9	7.0	47
4	BM	12	9.8	7.6	44	8.5	8.9	51	9.7	9.6	50
7	JR	30	8.1	8.1	50	7.0	8.5	55	7.8	7.9	50
13	MF	20	13.4	9.2	41	14.2	10.7	43	15.3	9.6	39
14	HU	20	9.6	7.8	45	8.7	8.8	50	9.4	9.4	50
15	JJ	20	13.2	10.1	43	9.2	12.8	57	10.5	11.5	52
16	BH	12	10.2	9.5	48	8.1	11.8	59	9.3	10.4	53
17	NR	6	13.0	10.3	44	9.0	13.5	60	10.5	11.8	53
18	LG	20	7.9	8.8	53	5.5	12.2	69	7.6	10.8	59

S		N	b clos.	b op.i.	b %	d clos.	d op.i.	d %	g clos.	g op.i.	g %
2	BF	4	13.8	0.8	5	12.5	2.0	14	10.3	3.0	23
3	PM	20	9.7	1.7	15	9.3	3.0	24	8.4	3.4	29
4	BM	20	13.2	1.1	8	10.6	2.0	16	11.5	2.8	20
7	JR	30	9.7	1.2	10	9.8	1.9	16	9.3	2.7	23
13	MF	20	17.7	1.2	6	16.1	2.1	12	17.9	2.7	13
14	HU	12	11.0	1.5	12	11.0	2.2	17	11.0	2.7	20
15	JJ	20	15.0	1.1	7	14.2	2.4	14	13.1	3.8	22
16	BH	12	13.3	1.1	8	12.1	2.9	19	11.6	3.4	23
17	NR	6	18.3	0.9	5	16.5	1.9	10	16.3	2.2	12
18	LG	20	11.8	1.1	9	11.1	3.0	21	11.6	3.1	23

5.2 Range of open interval in ptk and bdg

The difference between Danish /ptk/ and /bdg/ is one of aspiration. Both sets are voiceless, and /ptk/ cannot be considered to be more fortis than /bdg/ since /bdg/ have a longer closure and a tendency towards stronger organic pressure. As shown in table 7, there is a clear difference in the average duration of the open interval for /ptk/ vs. /bdg/ for all subjects, and no subject shows any overlapping of single tokens in this material. In connected texts and particularly in unstressed position, cases of overlapping may, however, occur (this was found in an older kymographic material).

Table 7

Differences in cs between the open intervals of ptk and bdg. Averages for the individual subjects on the basis of lists S3, S6, M1B and M2.

S	1	2	3	4	5	6	7	8	9	10
p > b	3.5	4.5	4.3	6.5	4.1	4.4	5.3	-	5.3	5.3
t > d	4.2	5.3	4.7	6.9	5.5	5.5	5.1	-	7.0	5.8
k > g	3.0	6.0	3.6	6.8	4.0	4.0	4.6	-	4.6	5.0

S	11	12	13	14	15	16	17	18	average
p > b	3.9	5.7	8.0	6.3	9.0	8.4	9.4	7.7	6.0
t > d	5.8	7.0	8.6	6.6	9.8	8.9	11.6	9.2	6.9
k > g	5.7	6.0	6.9	6.7	7.7	7.0	9.6	7.7	5.8

Moreover, the open interval of /ptk/ contains aspiration noise and /t/ is affricated, most strongly before high vowels (see section 5.3.4). The open interval of /bdg/ may, however, also contain some noise, particularly in di and gi.

The duration of the open interval of ptk is very variable. The individual averages vary between 3.7 and 10.3 cs for p, 5.4 and 13.5 cs for t, and 5.6 and 11.8 cs for k, and in single tokens the aspiration may be as short as 3.0 cs (p for subject No. 6) and as long as 17 cs (t (+ i) for subjects No. 15, 16, and 17).

Variations between different lists spoken by the same subjects are mostly due to the following vowel, the aspiration being longer before high vowels (see section 5.3.4). That may be the reason why list S5, which contains two high vowels and one low vowel, shows relatively longer values for subjects No. 5, 6, and 9 than list S3. However, M1 with high vowels only (u and i) has relatively short aspirations compared to S5 (subject No. 1) and M1 (subjects No. 2 and 7). The type of word may play a role here. Words of the type pidə, tidə, kidə, bidə, etc. containing short vowels surrounded by stop consonants may invite to rattling off the list.

The same is still more tempting for list S4: pip, pep, pɛp, pap, etc. do not sound like real words, the occurrence of the same consonant before and after the vowel being relatively rare except in onomatopoeia. Particularly subjects 6 and 9 have extraordinarily short aspirations for p (3-4 cs), and No. 6 has hardly any aspiration noise. This list has therefore not been included in the general means for the open interval in the following. There is no general tendency to shorten the aspiration before short vowels. List S5 was intended to test this assumption. It contained words with long and short vowels of the type talə, ta:lə, tulə, tu:lə, etc., but of 45 comparable pairs of averages (3 consonants x 3 vowels x 5 subjects), 21 had a longer and 21 a shorter aspiration before long vowels, and 3 had the same duration.

The differences between individual subjects reveal a tendency due to dialectal background and age. In table 8 the individual duration averages of the open interval of ptk were combined into three groups: A Non-Copenhageners, B Copenhageners born before 1939, and C Copenhageners (and subjects grown up in North Zealand) born after 1939. There is a tendency to increasing duration of aspiration from group A through B to C. The difference between groups A and C is significant at the 0.1% level.

No. 4 (BM) has been left out in the means because he moved from Jutland to Copenhagen at the age of 7, and he evidently does not belong in the group of non-Copenhageners.

No. 8 (KS) has shorter aspirations than the other Copenhageners. His ptk do in fact sound almost unaspirated in some cases, and in a listening test some of his intended p's and k's were heard as b's and g's (the same was true of some of NK's p's).

The year 1939 as the limit between group B and C has nothing magical in it. It is an artefact of the material. With more subjects, another year might have been chosen as an appropriate dividing line. Relatively long aspirations may also be found with subjects born earlier. E.A. Meyer (1904) has measured the duration of the aspiration in a few words spoken by a Copenhagener, and he found the values 7.8 and 10.0 cs for p and 11.5 cs for t. Moreover, H. Abrahams, who has undertaken a kymographic investigation of his own speech (1949), gives the range for p as 7-9 cs and mentions that the aspiration of t and k often goes beyond 10 cs (he is born in Copenhagen 1907). In the list S1, which is

Table 8

Individual average durations of the open interval of ptk, grouped according to dialectal background and age.
A Non-Copenhageners
B Copenhageners born before 1939
C Copenhageners born 1939 and later.[1]

				p	t	k
A.						
Non-Copenh.	1	OT	(1928)	6.8	7.7	7.6
	2	BF	(1935)	5.8	7.5	7.9
	3	PM	(1946)	6.0	7.7	7.0
	[4	BM	(1947)	7.6	8.9	9.6]
	5	FJ	(1911)	5.7	7.2	6.4
	6	NK	(1915)	5.1	7.7	7.2
	7	JR	(1934)	6.6	7.1	7.4
	average (N = 322)			6.0	7.5	7.3
B.						
Copenh.	8	KS	(1928)	5.4	7.5	6.4
	9	BL	(1930)	7.4	9.3	7.5
	10	KB	(1936)	7.3	7.8	8.3
	11	BS	(1922)	6.7	8.4	9.1
	12	HP	(1938)	7.6	9.2	9.4
	average (N = 150)			6.9	8.4	8.1
C.						
Copenh.	13	MF	(1939)	9.2	10.7	9.6
	14	HU	(1945)	7.8	8.8	9.4
	15	JJ	(1949)	10.1	12.2	11.5
	16	BH	(1945)	9.5	11.8	10.4
	17	NR	(1942)	10.3	13.5	11.8
	18	LG	(1946)	8.8	12.2	10.8
	average (N = 98)			9.3	11.5	10.6

---

1) 1-4 (born in Jutland) and 5-7 (born in Funen) are ordered chronologically. In groups B and C the ordering is not strictly chronological, because I did not want to change the numbering of my 1979 paper.

not utilized in full here, two Copenhageners (born 1899 and 1909) have several examples with aspirations of more than 10 cs. It should also be kept in mind that subjects 10, 11 and 12 have only spoken list M1, which has relatively short aspirations for the other subjects. On the other hand, No. 13 (HU), born 1945, has relatively short aspirations. This is not an accident due to the list in question. In the fiberoptic investigation undertaken by Birgit Hutters (1979), utilizing some of the same subjects, No. 13 (HU) also has shorter aspirations than Nos. 13 (MF) and 18 (LG). She has also recorded subject No. 5 (FJ) of group A, who shows considerably longer aspirations in this recording (1979) than in those used in the present investigation (recorded 1954-1963), viz. p 7.0 cs, t 9.8 cs, and k 9.0 cs. This may be due to a recent influence from Advanced Copenhagen speech, but it may also be due to a slower overall tempo in the fiberoptic recording, since the relative values (p 39%, t 56%, and k 51%) are hardly different from the earlier recordings.

In contradistinction to the absolute durations of the open interval, the relative durations do not show any difference between groups A and B (but it can only be calculated for two subjects of group B, 8 and 9, and, as mentioned above, 8 is not typical of group B).

The averages for the three groups are given in table 9.

Table 9

Average relative durations of the aspiration of ptk in groups A, B and C (see table 8).

	p	t	k
A (subjects 1,2,3,5,6,7)	40%	52%	47%
B (subjects 8 and 9)	39%	53%	45%
C (subjects 14 - 18)	47%	59%	53%

Subject 13 (MF) has not been included in the average because she has spoken extremely slowly and particularly with very long closures; her relative values (p 41%, t 43%, k 39%) are therefore very low compared to the other subjects of group C. That this is in fact due to her slow speed of delivery can be seen by comparing these values with those obtained in the investigation made by Birgit Hutters, where the absolute duration of MF's aspiration was

approximately the same as in the present investigation, but where the relative durations were 54, 61, and 48%.

On the whole, it must be kept in mind that the different lists are not exactly comparable, and that tempo may play a role. In order to get a clear picture of the differences, more subjects should be included and they should read the same lists. But the tendency found in the three groups is in complete agreement with the general auditory impression of an increasing aspiration in modern Copenhagen pronunciation of ptk.[1]

The open interval of bdg shows much less variation than that of ptk. The range of the average durations is 1.1 to 2.2 cs for b, 1.5 to 3.0 cs for d, and 1.8 to 4.0 cs for g, and there is no consistent difference between the three groups of subjects, A, B, and C, which were set up for ptk. The average durations for these three groups are given in table 10.

Table 10

Average duration of the open interval of bdg in cs for the groups A (Non-Copenhageners), B (Copenhageners born before 1939) and C (Copenhageners born in 1939 and later), based on lists S3, S6, M1B and M2.

A				B				C			
N	b	d	g	N	b	d	g	N	b	d	g
160	1.7	2.3	3.0	66	2.2	2.3	3.4	90	1.2	2.4	3.0

This means that the larger difference between the open intervals of ptk and bdg found in table 7 for subjects 13-18 is almost exclusively due to a longer aspiration of ptk.

One might have expected those subjects who have very long open intervals in ptk to have a corresponding lengthening of the open intervals of bdg, but this is only true to a certain extent. It is true of d and g for subjects No. 15, 16 and 18 but not for subjects No. 14 and 17 (see table 6) and, on the other hand, No. 1 and No. 3, who have rather short aspirations of ptk, have rather long open intervals of d and g (they have almost the same dialectal background in Northern Jutland, which might perhaps play a role). The lengthening for these 5 subjects is above all due to an affrication of di and gi. The averages for di and gi for these subjects are given in table 11.

---

1) Abrahams 1949 sees this development as a continuation of the Germanic and Old High German consonant shifts.

Table 11

Average duration of the open interval of di and gi in cs for subjects 1, 3, 15, 16, 18, based on lists M1B and M2.

S	1	3	15	16	18
N	6	10	10	6	10
di	2.9	4.0	3.4	4.0	4.4
gi	4.2	4.2	5.4	5.1	3.3

Single tokens may have as long an open interval as 6.0 cs (di) and 7.0 cs (gi). As ti and ki have still longer open intervals, this does not lead to any overlapping, but it means that the VOT-boundary may be around 7.0 cs before the vowel i for some subjects, which is higher than what is generally found in other languages (e.g. Lisker and Abramson 1964). Zlatin (1974) has, however, found a high perceptual VOT-boundary for g/k in English (6.5 cs).

5.3 Relations between closure, open interval and following vowel

5.3.1 Differences between /ptk/ and /bdg/

The relatively long open interval of ptk involves a shortening both of the closure and of the following vowel compared to bdg. Only lists S3 and M2 permit a comparison of both closure and vowel shortening. The averages for the subjects of these two lists are given in table 12, divided into two groups: A (Non-Copenhageners + BL from group B, whose durations in S3 are of the same order as those of group A) and C (Copenhageners born in 1939 and later).

Table 12

Differences between ptk- and bdg-sequences: open interval, closure, and following vowel, in cs, and differences in total CV length in cs. A(+B): subjects 2, 3, 4, 5, 7, 9; C: subjects 13-18 (lists S3 and M2).

	A(+B)			C		
	N=87 p>b	t>d	k>g	N=90 p>b	t>d	k>g
open int.	+5.3	+6.5	+5.1	+8.2	+9.0	+7.6
closure	-2.2	-3.3	-1.6	-3.5	-4.6	-3.0
total cons.	+3.1	+3.2	+3.5	+4.7	+4.4	+4.6
vowel	-1.4	-2.4	-1.5	-2.4	-1.9	-2.1
CV	+1.7	+0.8	+2.0	+2.3	+2.5	+2.5

In lists S6 and M1 and for subject 6 in S3, only aspiration and vowel duration can be compared. The averages are given in table 13.

Table 13

Differences between ptk- and bdg-sequences: open interval and following vowel in cs for subjects 6 (list S3), 5, 6, and 9 (list S6) and 1, 2, 5, 7, 10, 11, 12 (list M1B).

N		p>b	t>d	k>g
124	open int.	+4.3	+5.1	+4.3
	vowel	-1.3	-1.3	-1.1

The difference in closure between ptk and bdg is very stable and significant at the 0.1% level. A comparison of the individual averages (with separate means for consonants before i and a in list M2 and before high, mid, and low vowels in list S3) gives 78 comparable pairs. Of these, 75 have a longer closure in bdg than in ptk, 2 have a slightly longer closure in ptk (0.5 and 0.6 cs), and one has the same duration. On the average, the difference between t and d is somewhat larger than that between p and b and between k and g. The shortening of the vowel after ptk compared to bdg is also consistent and significant at the 0.1% level. Out of 113 comparable pairs of averages in lists S3, S6, M1B and M2, 102 have a longer vowel after bdg, 9 have a slightly longer vowel after ptk (0.1 - 1.1 cs), and 2 have the same duration.[1] On the average, the vowel is less shortened than the closure, but this is not consistent, and there are individual differences. No. 16 shortens the vowel more in all cases, whereas No. 17 and No. 18 show definitely more shortening of the closure in all cases.

In figs. 1-3 the relations are shown graphically for list M2. The numbers at the bottom indicate the subjects. The values for the individual subjects are combined by lines in order to make the consistency stand out more clearly. It should be noticed that

---

1) If the delimitation between aspiration and vowel is made at voicing start, the difference in vowel duration after ptk and bdg will be somewhat less regular.

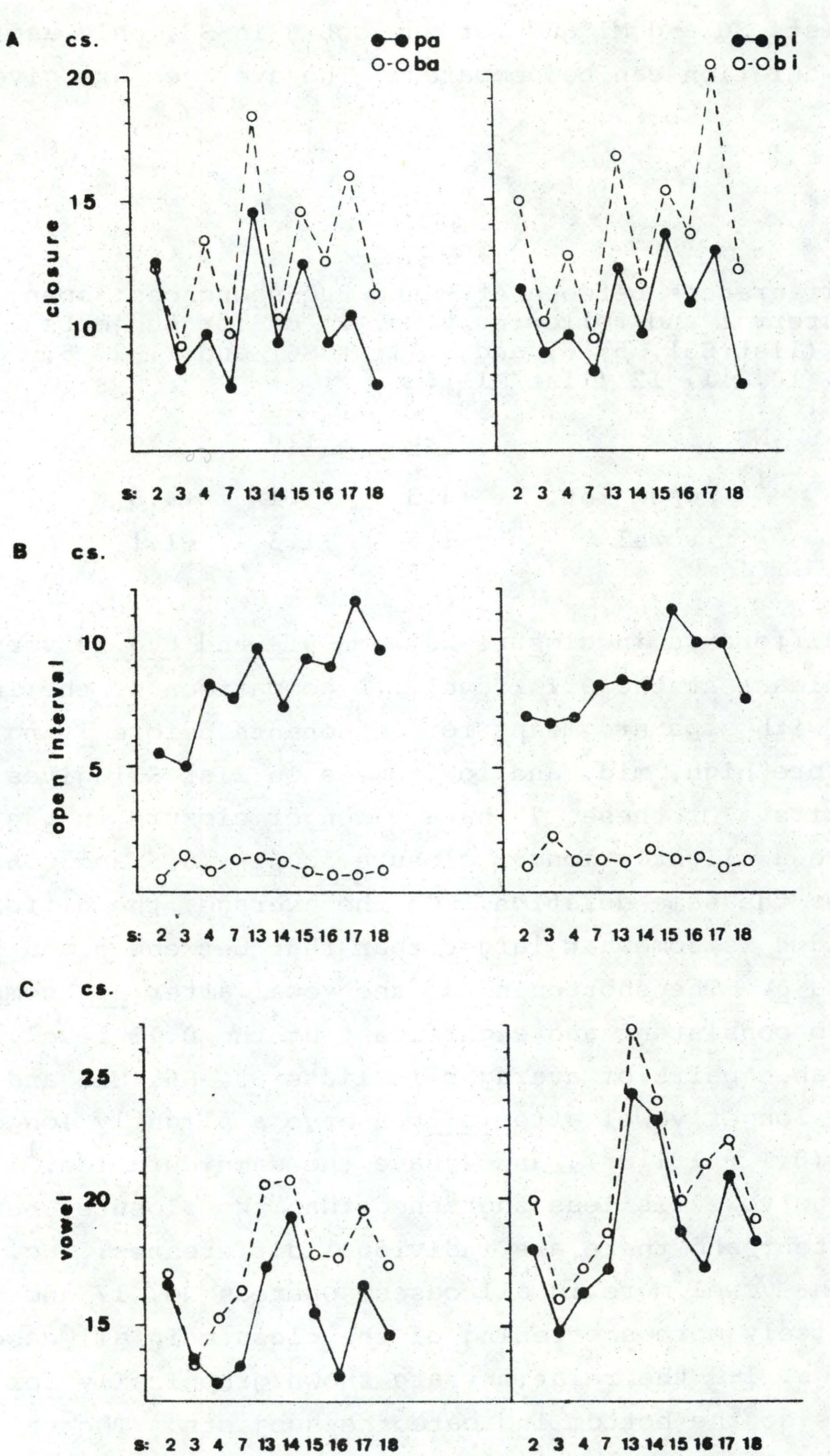


Figure 1

Duration of closure, open interval and vowel
List M2 (S = subject number)

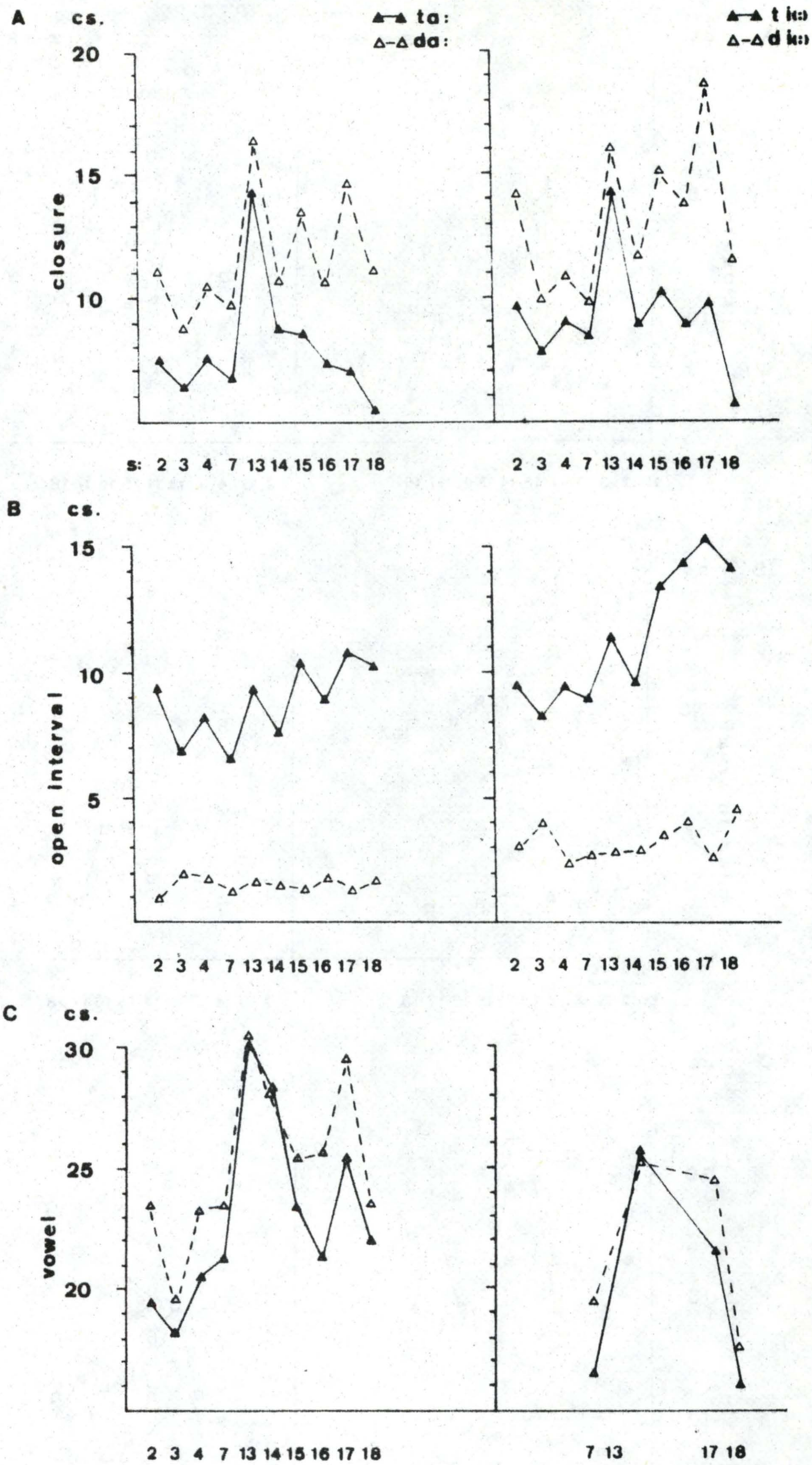


Figure 2

Duration of closure, open interval and vowel.
List M2 (S = subject number)

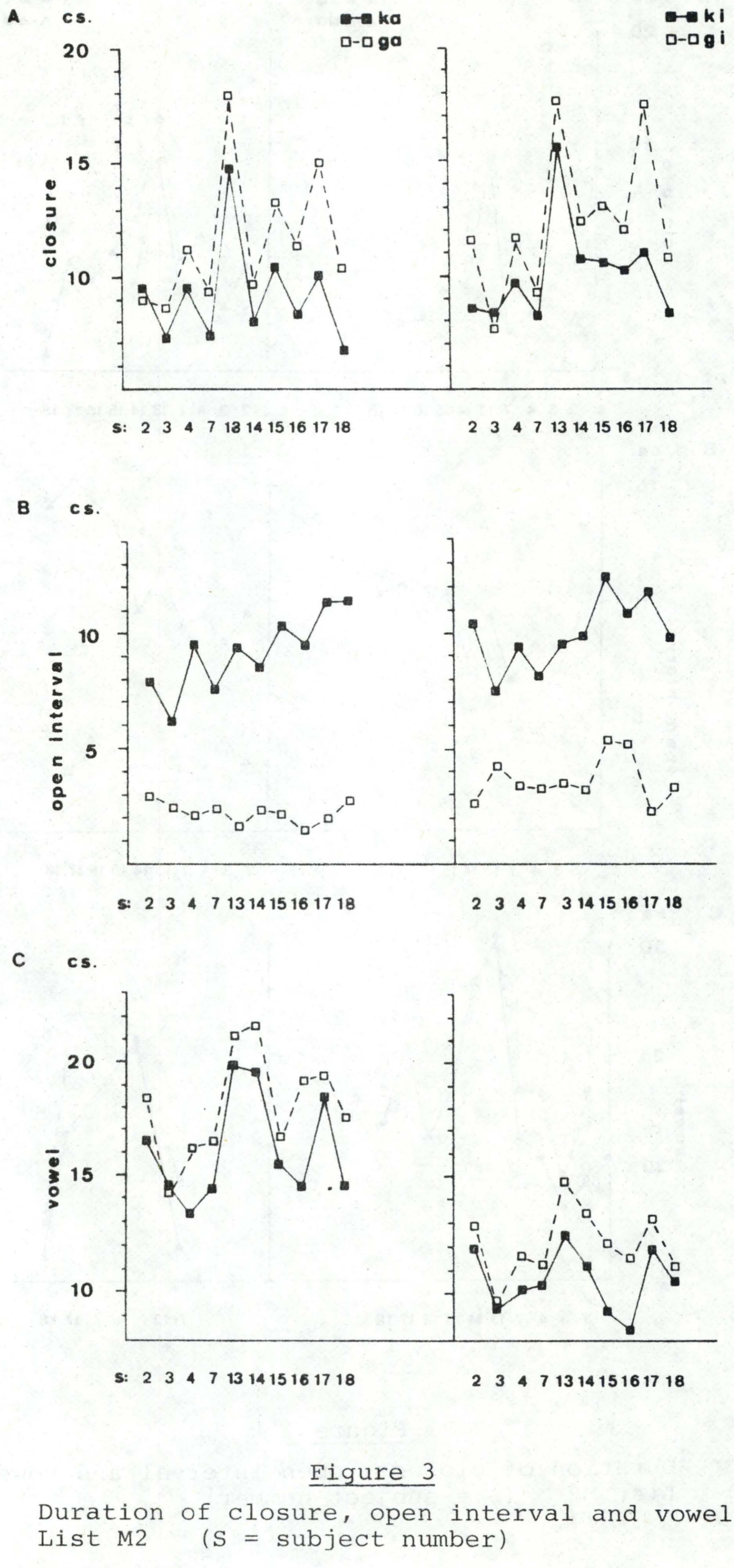


Figure 3

Duration of closure, open interval and vowel.
List M2 (S = subject number)

not all the absolute vowel durations are comparable. The words are: palə, balə, ta:lə, da:lə, kalə, galə, pi:lə, bi:lə, ti(:)ə, di(:)ə, kilə, gilə. In ti(:)ə, di(:)ə the vowel [i] is of variable length since there is no phonological distinction of duration before [ə]. The segmentation of [i] and [ə] was only safe for four subjects.

The high peaks in the closures and vowel durations of subjects 13 and 17 are due to slow speech. No. 14 did not speak slowly, but she has relatively long vowels. The rising slope of the graphs for the open interval reflects the tendency towards longer open intervals in modern Copenhagen speech. In ta the closure shows an opposite downdrift, and a comparison between groups A and B in table 12 seems to indicate a general tendency for those subjects who have long aspirations in ptk to shorten the closure more (in relation to bdg), although the shortening does not fully make up for the longer aspiration, the differences in total consonant length and in CV length between ptk- and bdg-sequences being somewhat larger in group C. It is, however, not a very regular phenomenon in individual averages. Subject 17 has, e.g., very large closure differences, No. 15 and No. 18 only in t-d, No. 13 not in t-d.

There is no tendency to shorten long vowels more than short vowels.

The difference in closure duration between ptk and bdg never reaches the difference in open interval, except for subject 9 in the case of k and g before open vowels. This means that the total duration of the consonants ptk is consistently longer than the total duration of bdg. When the shortening of the vowel is also taken into account, it turns out that the shortening of closure and vowel may make up for the lengthening of the open interval for some subjects (5, 9, and 16). But for most subjects the ptk sequences are clearly longer. The individual differences for total consonant length and CV-length are shown in table 14.

When averages of individual consonants before a and i in list M2 and before high, mid, and low vowels in list S3 are taken separately, the difference between CV-sequences with ptk and with bdg can be compared in 72 pairs. The ptk-sequences are longer in 62 pairs. This means that the difference is significant, although it is not very large (it is less than 1.5 cs in 35 out of the 72 cases).

Table 14

Differences in total consonant duration (closure + open interval) and in CV duration between ptk and bdg for individual subjects (cs).

S	2	3	4	5	7	9
tot.cons. ptk > bdg	+4.1	+2.9	+4.3	+2.6	+4.2	+1.3
CV ptk > bdg	+2.5	+2.6	+2.0	+0.7	+2.8	0

S	13	14	15	16	17	18
tot.cons. ptk > bdg	+4.9	+4.6	+5.8	+5.0	+3.6	+4.0
CV ptk > bdg	+3.3	+3.5	+3.9	+0.6	+1.1	+2.3

The longer closure in bdg than in ptk and the longer overall length of ptk compared to bdg is also found as consistent features in the material investigated by Hutters (1979), and I have found the same relation in an older kymographic material and in a material used for air pressure measurements.

The shorter vowel after ptk was also found for 7 out of 8 subjects in list S1. The same relation for the vowels was found by Hutters (1979) and by Holtse (1977) (two of Holtse's three subjects and four of Hutters' five subjects are identical to subjects used in the present investigation).

It may be of interest to compare the duration of vowels after ptk and bdg with vowel duration after other consonants. This has been done for list M1B (see Fischer-Jørgensen 1964, p. 187). The averages for seven subjects (6 examples of each) are for the short vowels i u after ptk 9.5 cs, after fsh 9.8, after bdg 10.7, and after mnlv 10.0 cs. In list M2 it is possible to compare palə, balə, falə, valə and pi:lə, bi:lə, fi:lə, vi:lə and ta:lə, da:lə, sa:lə. The total consonant durations and the vowel durations are shown in table 15.

It appears from table 15 that f is shorter than p but longer than b, and that the vowel after f is longer than after p and shorter than after b, i.e. there seems to be compensation. But there is only very slight compensation in the vowel for the difference of duration between f and v, and none between the vowels

Table 15

Averages of total consonant duration and duration of following vowel in cs for labial and alveolar consonants in list M2 (10 subjects).

N = 76	p	f	b	v	f>p	f>b	f>v	b>v
cons.	18.5	15.8	13.8	10.7	-2.7	+2.0	+5.1	+3.1
vowel a	15.2	17.2	17.6	17.5	+2.0	-0.4	-0.3	+0.1
cons.	19.1	17.9	15.1	11.6	-1.2	+2.8	+6.3	+3.5
vowel i:	18.9	19.8	20.6	20.6	+0.9	-0.8	-0.8	0
N = 38	t	s	d		s>t	s>d		
cons.	17.0	16.1	13.1		-0.9	+3.0		
vowel a:	23.0	25.3	24.3		+2.3	+1.0		

after b and v. The inverse relation between consonant and vowel duration holds for 14 out of 20 individual averages for f-p and for 13 out of 20 for f-b; for s-t and s-d it holds for 9 and 6 averages, respectively, out of 10. This is not statistically significant, but the tendency is clear.

One of the two subjects in list S3 shows the same relations for p-f-b, but none of them for t-s-d, both consonant and vowel being longer in s + V than in t + V.

Hutters (1979) has found the same inverse relations between consonant and vowel duration for f-p and f-b before i for four out of five subjects, i.e. pi>fi>bi for the consonant and bi>fi>pi for the vowel. For the alveolars the order of the consonant durations is correspondingly ti>si>di, but the order for the vowels is si>di>ti for four out of five subjects (the same subjects have a longer vowel after d than after s in the present material). The relations thus seem to be less stable for the alveolars than for the labials.

Holtse (1977) found vowels after f and s to be of approximately the same duration as after bdg. He did not measure the consonants.

In list M2 it is also possible to compare f and h + vowel for three subjects and in list S5 f s h + vowel for two subjects.

In list M2 both consonant and vowel are longer in fa than in ha. In S3 the consonants f and s are considerably longer than h (by 4.8 and 6.4 cs, respectively), but the following vowels are only very slightly (and insignificantly) shorter (by 0.5 and 1.1 cs, respectively). There is thus hardly any compensation in this case.

5.3.2 Differences between labials, alveolars, and velars

The duration of both closure and open interval differs according to place of articulation.

For bdg the closure durations can be compared on the basis of list S3 and list M2 (subjects 2-5, 7, 9, and 13-18), the open interval, moreover, on the basis of S6, M1B and M1A (all subjects). For ptk the closure durations can be compared on the basis of lists S3, M2, and S5 (subjects 1-9 and 13-18), and the open interval, moreover, on the basis of lists S4, S6, and M1B (all subjects). The durations have been compared by means of individual averages for different subjects and consonants before i a and u, taken separately for lists S5, M1B and M2, and consonants before high, mid, and low vowels, taken separately for lists S3, S4, and S6. As for M1A, bdg before i, y, and u have been combined. Table 16 contains a survey of the results.

Table 16

Differences between labials, alveolars, and velars in the duration of closure and open interval, indicated as the absolute and relative number of averages showing the difference.

closure (28 aver.)	b > d		b > g		d > g	
	23/28	82%	25/28	89%	15/28	54%
open interval (61 aver.)	d > b		g > b		g > d	
	47/61	77%	60/61	98%	57/61	93%
closure (58 aver.)	p > t		p > k		k > t	
	57/58	98%	47/58	81%	53/58	91%
open interval (93 aver.)	t > p		k > p		t > k	
	86/93	92%	75/93	81%	62/93	67%

Except for d > g the differences are all significant at the 0.1 or 1% level. It appears from table 16 that the order for the closure of bdg is b > (d > g) (where the order d > g is unstable) and the order for the open interval is the opposite: g > d > b. Here, d > b has more exceptions than g > d and g > b, but a closer inspection of the material shows that in almost all cases where b has a longer interval than d, the following vowel is rounded (see section 5.3.5). The order for ptk is different. For the closure it is p > k > t, and for the open interval the opposite: (t > k) > p. t > k is only valid in 67% of the averages and, whereas the other cases do not show any differences for lists or subjects, there is a difference between lists for t > k. There is majority for t > k in all lists except list M1B, which has only one case of t > k out of 14. Without list M1B the percentage would have been 77%. As mentioned in section 5.1, the open intervals of ptk are relatively short in this list, and a comparison of list M1B with the other lists for those subjects who have spoken other lists as well (i.e. Nos. 1, 2, 5, 7) shows that in six out of seven cases the shortening of the open interval of t in list M1B is more pronounced than that of k, and in five cases more pronounced than that of p. This may perhaps be due to the fact that in list M1B the consonant following the vowel was also an alveolar (the word type is: pudə, tudə); in list S4, where the same consonant is found before and after the vowel in all cases, there was a general shortening of the open interval. Or the quick tempo of list M1 might shorten the affrication of t. This is, however, pure guesswork.

If the subjects are considered separately (cp. table 6), 11 out of 18 (or 61%) have a longer open interval in t than in k, but if M1B is left out, it will be 13 out of 15 (or 87%).

It may be of interest to see whether the inverse relation between closure and open interval in the differences between labials, alveolars and velars holds in the cases where closure and open interval can be compared directly within individual averages, - not only when they are examined separately on the basis of the whole material as in table 16 (in many cases only the open interval, not the closure, could be measured). This can be investigated for ptk in lists M2, S3 and S5 (i.e. for subjects 1-9 and 13-18) and for bdg in lists M2 and S3 (i.e. for subjects 2-5, 7, 9, and 13-18), where closure and open interval for the same examples can be measured. The results are shown in table 17. For bdg there is good agreement between tables 17 and 16. The relation between d and g does not show any inverse relation, since the closure difference between d and g is unstable as shown in table 16, whereas the relations between g and b and between b and d do show this inverse relation.

For ptk the inverse relation of t-p is very stable, as should be expected from table 16. But t-k should not be expected to show a higher percentage of inverse relation than p-k.

Table 17

Occurrence of inverse relation between differences of closure and open interval within the same pairs of averages.

ptk - 58 averages

closure	p > k, open interv.	k > p	40/58	69%
-	p > t, -	t > p	50/58	86%
-	k > t, -	t > k	42/58	72%

bdg - 28 averages

closure	b > g, open interv.	g > b	27/28	96%
-	d > b, -	b > d	24/28	86%
-	d > g, -	g > d	15/28	53%

The reason is that the relatively low percentage for the open interval (t > k 67%) in table 16 was due mainly to list M1B, which is not included in table 17, because the closure cannot be measured in this list.[1]

The absolute magnitude of the differences is shown in table 18 for the lists in which both closure and open interval could be measured, i.e. S3 and M2 for bdg and S3, S5 and M2 for ptk.

---

1) If the delimitation of the open interval had been determined by the voicing start instead of the vowel start, the open interval of k would have been shorter in some cases, and the percentage t > k would have been still higher. In Fischer-Jørgensen 1979 it was said that there was no consistent relation for t-k and that the relation d-b was not significant. This was not correct. correct.

Table 18

Differences in cs between consonants of different places of articulation (closure, open interval, and total duration), based on lists S3 and M2 (subjects 2-5, 7, 9, 13-18) for bdg and S3, S5 and M2 (subjects 1-5, 7-9 and 13-18) for ptk.

bdg N = 204	b > d		b > g		d > g	
	S3	M2	S3	M2	S3	M2
clos.	+0.2	+1.1	+0.9	+1.4	+0.8	+0.3
op.int.	-0.3	-1.1	-0.8	-1.7	-0.6	-0.6
tot.dur.	-0.1	0	+0.1	-0.3	+0.2	-0.3

ptk N = 354	p > k			p > t			t > k		
	S3	S5	M2	S3	S5	M2	S3	S5	M2
clos.	+0.7	+1.0	+0.9	+2.1	+2.1	+2.0	+1.3	+1.1	+1.1
op.int.	-0.6	-1.0	-1.4	-2.5	-2.1	-2.1	-1.9	-1.1	-0.7
tot.dur.	+0.1	0	-0.5	-0.4	0	-0.1	-0.6	0	+0.4

Table 18 shows that shortenings and lengthenings counterbalance each other, so that the total duration of the consonants is almost identical. A graphic display of the differences t-p and g-b is given in figs. 4-5. The durations of p and b are given the value zero. Overlapping of single measurements has been examined for list S5. Correlation diagrams were made for the single values of aspiration versus closure for consonant + a (18 examples of each) and consonant + i and u (12 examples) in pairs. In fig. 6 examples are given of ta/pa for subject 5 (hardly any overlapping), ta/ka for subject 9 (somewhat more overlapping), and ka/pa for subject 1 (almost complete overlapping). On the whole, No. 1 shows almost complete overlapping except for ta/pa, whereas Nos. 5, 8, and 9 have very good separation for t-p, somewhat less for t-k (but still very good for No. 5), and rather bad separation for k-p. No. 6 has a small pause before each word, and this pause cannot be delimited from the closure. With the pause included he shows very good separation for t-p and p-k but not for t-k.

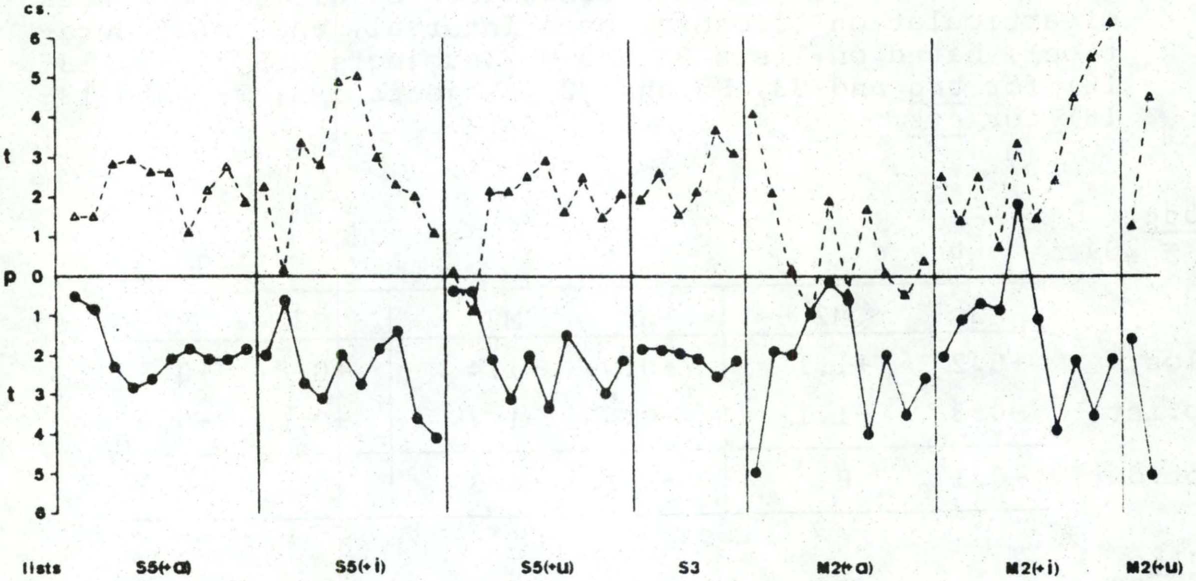


Figure 4

Differences between the durations of the closures (●——●) and open intervals (△----△) for t - p. p is in both cases given the value zero.

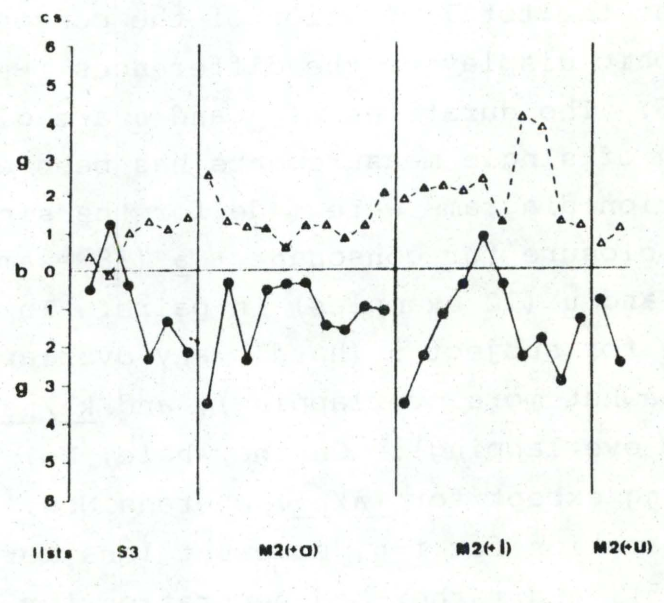


Figure 5

Differences between the durations of the closures (●——●) and open intervals (△---△) for g - b. b is in both cases given the value zero.

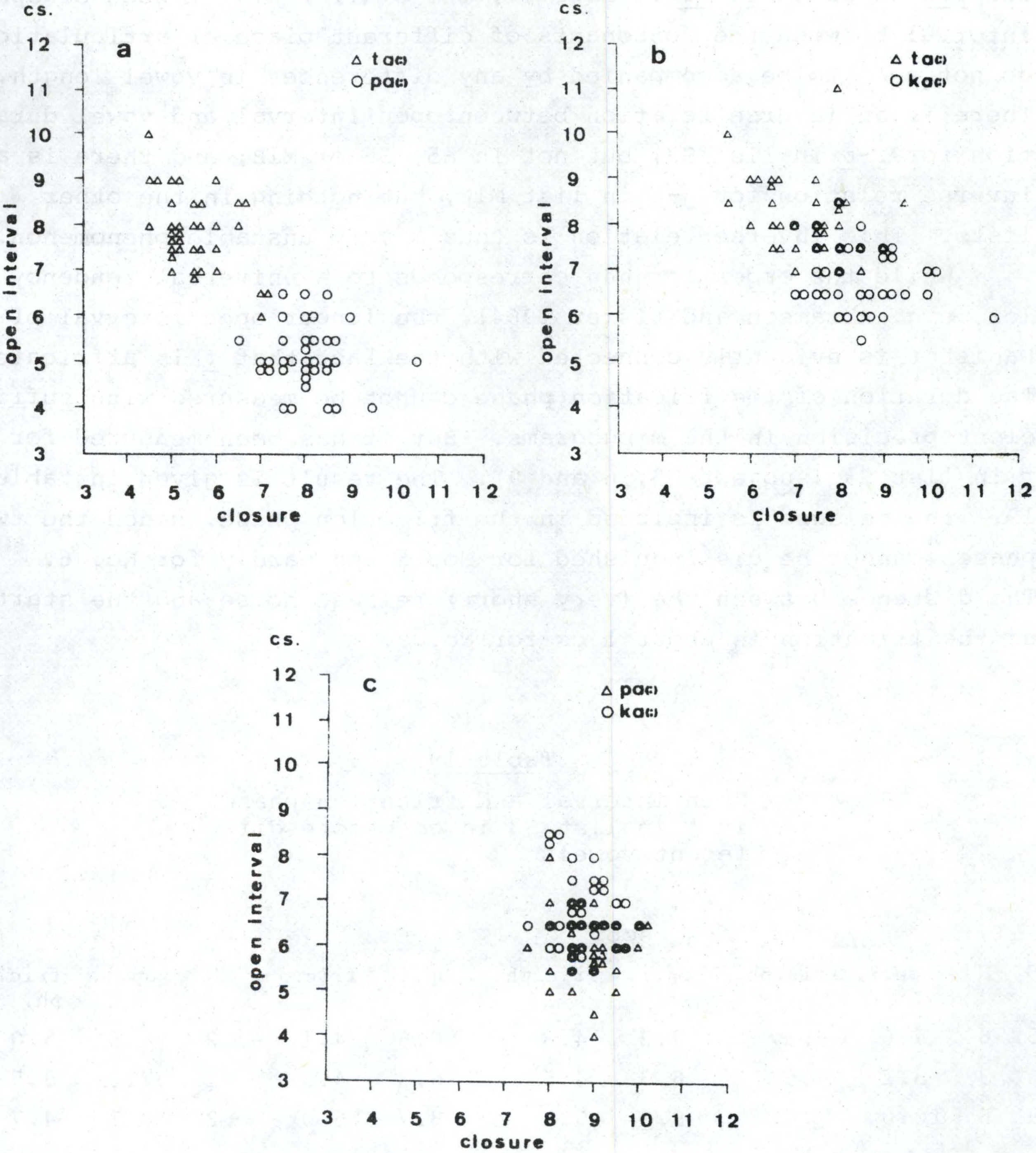


Figure 6

Examples of correlation diagrams of closure and open interval for single tokens (list S5):
(a) tɑ/pɑ (subject 5), (b) tɑ/kɑ (subject 9),
(c) kɑ/pɑ (subject 1).

The material has also been examined in view of a possible compensation between the duration of the open interval and the duration of the following vowel, since a shortening of the vowel was found for ptk compared to bdg. However, the smaller differences of open interval between the consonants of different place of articulation do not seem to be accompanied by any differences in vowel length. There is an inverse relation between open interval and vowel duration for t-p in list S3, but not in S5, S6 or M1B; and there is an inverse relation for g-b in list M1B, but nothing in the other lists. This inverse relation is thus a very unstable phenomenon.

While the order for bdg corresponds to a universal tendency (cf. e.g. Abramson and Lisker 1964), the longer open interval of Danish t is evidently connected with the fact that t is affricated. The duration of the frication phase cannot be measured with sufficient precision in the mingograms. But it has been measured for t in list S3 (subjects 5, 6 and 9). The result is given in table 19. The release is included in the frication phase, since the two phases cannot be distinguished for No. 5 and hardly for No. 6. The distance between the (very short) release noise and the start of the frication is about 1 cs for No. 9.

Table 19

Open interval and frication phase in t in list S3 in cs before different vowels

		iyu		eøo		ɛœɔ			a	
S	N	op.i.	fric.ph.	op.i.	fric.ph.	op.i.	fric.ph.	N	op.i.	fric. ph.
5	6	7.6	6.3	7.3	4.8	6.5	4.1	2	6.5	5.0
6	3	8.2	6.5	6.3	4.5	6.3	4.3	1	7.0	3.5
9	6	10.0	7.0	9.0	5.1	8.7	5.0	2	8.3	4.7

The frication phase is longer before high vowels than before low vowels, and the differences are somewhat larger than for the open interval, so that the frication phase takes up more of the open interval in high vowels. Spectrograms were also taken of some of the examples of S5. The frication phase is approximately of the same length here as in S3 for subjects 5, 6 and 9, but No. 8, who did not read list S3, has relatively short frications (before i 5.0 cs, before u 4.0 cs, and before a 3.2 cs). Mingograms of the other subjects seem to indicate that the frication phase generally takes up most of the open interval before high vowels, also when the open interval is very long, whereas this is generally not the case before a, and spectrograms of subject No. 16 show frication phases of 9-10 cs before i, but only 3-4 cs before a. Thus, it cannot be maintained for Danish (as it can for English, according to Klatt (1975) and Zue (1976) that when the frication phase is subtracted, the aspiration is approximately of the same length for p, t and k.

5.3.3 Influence of phonological vowel length

The main purpose of list S5 was to investigate the influence of long and short vowels on preceding /ptk/. In table 20 the five subjects are listed separately.

Table 20

Relations of duration between CV- and CV:-sequences (list S5).

CV: > CV

Subjects	1	5	6	8	9
vowel	+8.4	+4.8	+6.1	+7.6	+9.8
clos.	-0.1	0	-0.6	-0.4	-0.7
op. int.	-0.3	-0.3	+0.4	-0.2	+0.2

It appears from the table that subjects 6 and 9 have no inverse relation between open interval and vowel, and subject 5 no inverse relation between closure and vowel, and for the other subjects the differences are very small compared to the large vowel differences; moreover, they are based on varying single averages and are not statistically significant.

The following consonant has not been measured, but earlier investigations (Fischer-Jørgensen 1955 and 1964) have shown that there is no significant difference in consonant duration after long and short vowels.

5.3.4 Influence of vowel height

It is well known that there is a general tendency for high vowels to be shorter than low vowels, and this relation has also been found for Danish (Fischer-Jørgensen 1955 and Holtse 1977). The material of my 1955 paper was identical with S3 and S4 of the present investigation; but the same relations are found in S5, S6 and M2, which were recorded later. In the whole material high vowels are shorter than low vowels in all 123 comparable pairs of averages. This difference has relation to differences in the preceding consonants.

Table 21 shows the differences between high and low vowels and the corresponding differences in the duration of closure and open interval, based on the lists which permit a measurement of the closure, i.e. lists S3, S5 and M2. Table 22 shows the difference between high and low vowels and the corresponding differences in open interval of the preceding consonants for lists S4 and S6. S3, S4 and S6 contain examples of all Danish vowels (long in S3 and S6, short in S4), but here only the high vowels iyu and the low vowels ɛœɔ are compared. In S5 i and u are compared to a, separately for long and short vowels, and in M2 i is compared to a. In M2 only the vowels i-a after velars have the same phonological length. After alveolars the length is undetermined (ti(:)ə di(:)ə tu(:)ə du(:)ə), but the speakers have all pronounced them as long vowels, and the same is the case for /u(:)/ after labials. /i:-a/ after labials and /u:-a/ after velars have different phonological length. In the cases where the vowels differ in phonological or intended length, the vowel durations have been put in parentheses. (i after alveolars could only be measured for four subjects; u: after velars was only spoken by two subjects.) A graphical display is given in figs. 7-9.

Tables 21 and 22 and figs. 7-9 show that there is a consistent inverse relation between vowel duration and duration of the open interval of the preceding consonant, but the difference in open interval is considerably smaller than the difference in vowel duration, so that sequences with low vowels remain longer. The difference in open interval is, however, highly significant. For ptk it is valid in 92 out of 103 individual averages with comparable vowels (89%), and for bdg in 32 out of 34 averages (94%). It appears from table 21 that the consonants are also lengthened in

Table 21

Differences (in cs) of vowel, closure and open interval between CV-sequences with high and low vowels (subjects 1-9, 13-18).

S3 /i: y: u:/ > /ɛ: œ: ɔ: a:/

N = 15/20	p	t	k	b	d	g
vowel	-2.9	-3.2	-3.6	-3.0	-3.3	-3.5
closure	0	+1.0	+0.3	-0.2	-0.3	-0.7
open int.	+1.1	+1.3	+0.6	+0.7	+0.7	+0.6
CV	-1.8	-0.9	-2.7	-2.5	-2.9	-3.6

S5 /i u/ > /a/ (short and long)

	short			long		
N = 120	p	t	k	p	t	k
vowel	-5.7	-5.1	-5.0	-4.7	-4.0	-6.8
closure	+1.1	+0.4	+0.2	+0.5	+0.6	+0.5
open int.	+1.3	+0.7	+1.3	+1.4	+1.4	+1.6
CV	-3.3	-4.0	-3.5	-2.8	-2.0	-4.7

M2	p	t	k	b	d	g
	i:>a	i:>a:	i>a	i:>a	i:>a:	i>a
	N=76	N=32	N=76	N=76	N=32	N=76
vowel	(+3.5)	-4.7	-5.6	(+3.7)	-4.8	-5.9
closure	+0.4	+0.8	+1.0	+0.9	+1.0	+0.8
open int.	+0.2	+3.2	+0.8	+0.4	+1.6	+1.4
	(+4.1)	-0.7	-3.8	(+5.0)	-2.2	-3.7

N = 12	p	t	k	b	d	g
	u:>a	u:>a:	u:>a	u:>a	u:>a:	u:>a
vowel	(+6.6)	-2.6	(+3.2)	(+4.3)	-4.4	(+4.4)
closure	+2.9	+1.8	+0.5	+2.3	+1.1	+1.4
open int.	-1.2	+2.8	+0.1	+0.7	+0.7	+0.3
CV	(+8.3)	+2.0	(+3.8)	(+7.3)	-2.6	(+6.1)

Table 22

Differences (in cs) of vowel and open interval between CV-sequences with high and low vowels (subjects 1, 5, 9).

S4 /i u y/ > /ɛ œ ɔ a/

N = 15/20	p	t	k
vowel	-2.3	-2.9	-1.8
open int.	+0.8	+0.6	+0.9

S6 /i: y: u:/ > /ɛ: œ: ɔ: a:/

N = 0/12	p	t	k	b	d	g
vowel	-2.1	-1.9	-2.6	-3.9	-2.4	-2.9
open int.	+1.4	+0.6	+0.7	+0.7	+0.9	+0.9

cases where the vowels are of phonologically different length (we shall come back to this in the discussion), so that these cases can be included, but the percentages will be only slightly changed (85 and 94%, respectively). Most exceptions are due to 9 cases of pi-pa in M2 and S5 in which pa has the longer aspiration.[1] The percentage for pi-pa is 75%.

As for the lengthening of the closure before high vowels, table 21 shows no consistent difference in S3, but a consistent difference in S5 and M2. Comparison of individual averages shows that the difference is significant for ptk (52 out of 75 averages or 69%) but not for bdg (27 out of 45 averages or 60%). There are, however, individual differences among the subjects. Some have no exceptions, and two have the opposite relation, i.e. shorter closure before high vowels, viz. No. 9 and No. 13. Apparently, younger Copenhageners tend to lengthen the closure more before high vowels than do older subjects. At any rate, there is a striking difference between the subjects of list S3 (Nos. 5, 6, 9), list S5 (Nos. 1, 5, 6, 8, 9) and list M2 (Nos. 2, 3, 4, 7, 13-18). The percentage of averages showing a longer closure before high vowels than before low vowels is for S3 33%, for S5 66%, and for M2 80% (and for the five youngest Copenhageners of M2 it is 92%).

---

1) These exceptions would be reduced if voicing start were taken as marking the end of the open interval.

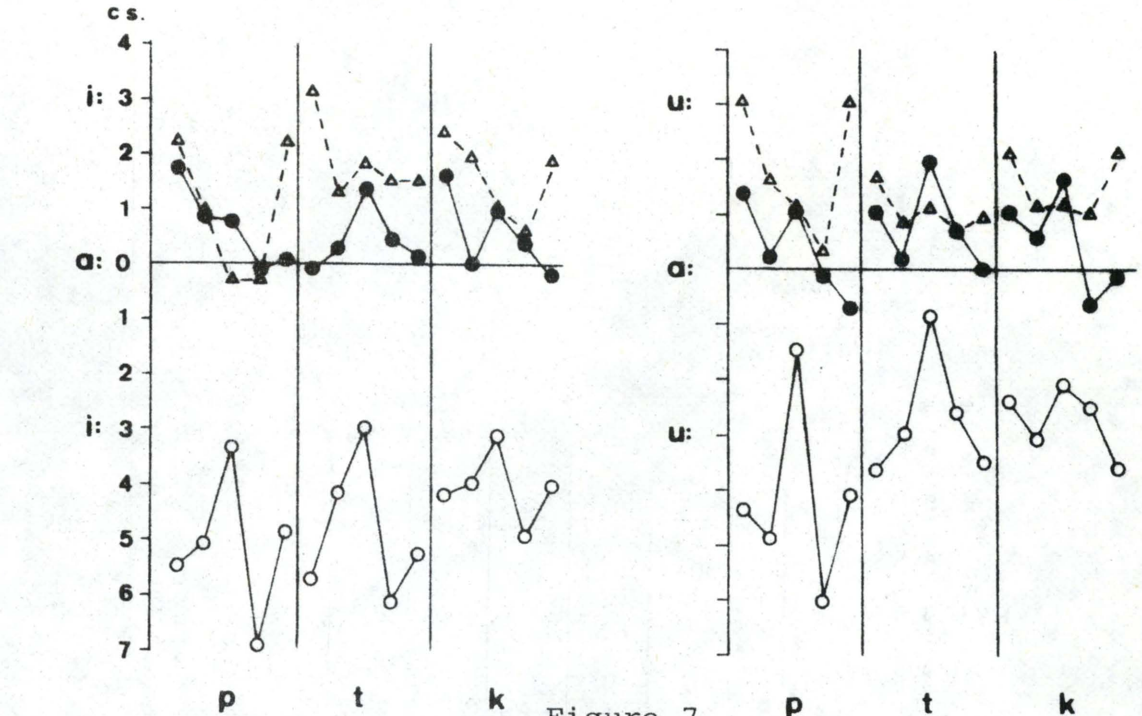


Figure 7

Differences between the durations of the closures (●——●), open intervals (△---△) and following vowels (○——○) for ptk + i: u: - ptk + ɑ: (list S5). The durations of ptk + ɑ: have been given the value zero.

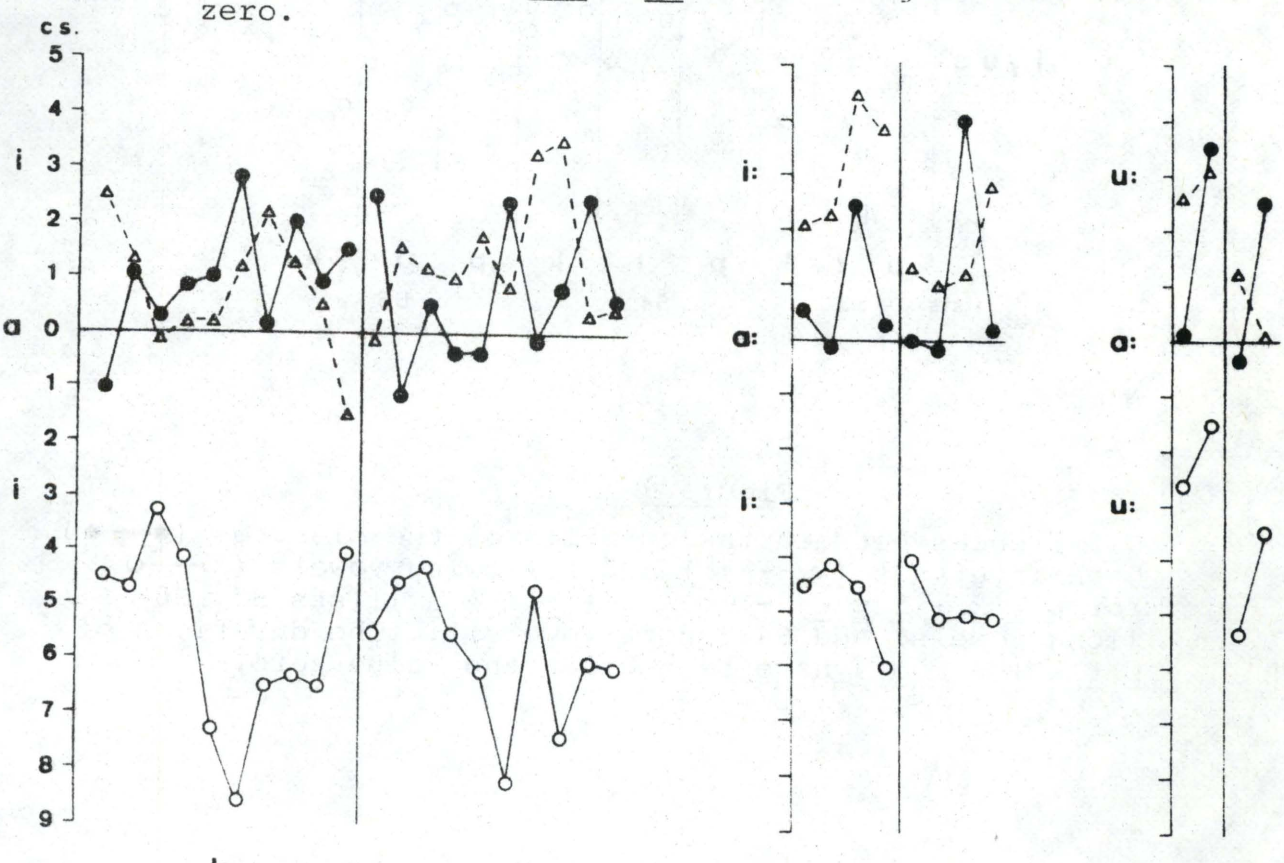


Figure 8

Differences between the durations of the closures (●——●), open intervals (△---△) and following vowels (○——○) for k g t d + i i: u: - kg td + a i: (list M2). The durations of k g t d + a a: have been given the value zero.

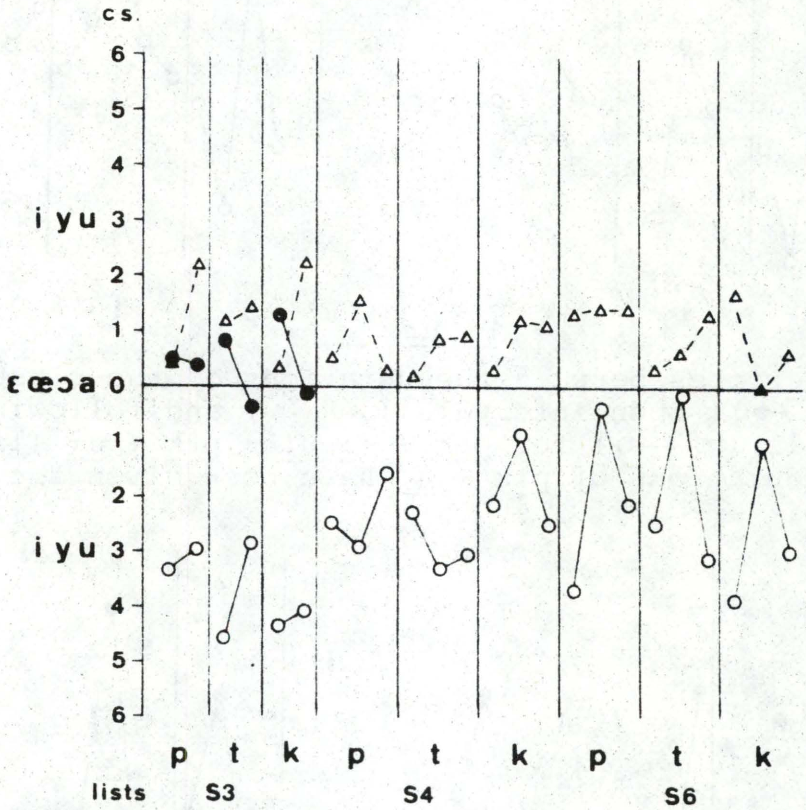


Figure 9

Differences between the duration of the closures (●——●) open intervals (△----△) and following vowels (○——○) for p t k + i y u - p t k + ɛ œ ɔ a (lists S3, S6 (long vowels) and S4 (short vowels)). The durations of p t k + ɛ œ ɔ a have been given the value zero.

And whereas in S5 there are only 5 cases out of 30 (17%) where the lengthening of the closure of ptk exceeds the lengthening of the open interval, the corresponding number in M2 is 16 out of 36 (44%). There is a slight inverse tendency for the open interval, i.e. more lengthening in S3 and S5 than in M2, but this may be due to inconsistencies in the delimitation of open interval and vowel, whereas the delimitation of the closure is pretty safe.

In Hutters 1979 the lengthening of the closure before i compared to a is found in 11 out of 15 averages (73%), whereas her subjects lengthen the open interval before i in only 53% of the averages.

In lists S3 and M2 there are also examples of other consonants. The fricatives f s v show the same lengthening before i and u compared to a as the stops, as seen in table 23.

Table 23

Difference in cs between CV-sequences with f,s,v before high and low vowels (Subjects 2 - 7, 9, 13 - 18).

	S3		M2					
	N = 15/20		N = 76					N = 12
	/i: y: u: > ɛ: œ: ɔ: a:/		/i:>a/.		/i:>a:/	/u:>a/		/u:>a:/
	f	s	f	v	s	f	v	s
vowel	-3.7	-3.4	(+2.7)	(+3.1)	-5.0	(+4.2)	(+3.7)	-3.0
cons.	-0.3	+0.2	+2.0	+0.9	+3.0	+1.9	+2.2	+2.3

The negative difference for f in S3 is due to subject 6 (5 and 9 have a positive difference). The consonant is longer before high vowels in 33 out of 36 individual averages (91%). There is also a small difference for h in S3, but the delimitation is not too certain, and in M2 h is variable. The same is true of the examples of m and l (spoken by 5 subjects of list M2).

Again it is found that the consonant lengthening is independent of the phonological length of the following vowel (the vowels of M2 have the same length only after s). In Hutters 1979 all five subjects have longer f and s before i than before ɑ.

5.3.5 Influence of vowel rounding

The relation between vowel rounding and the open interval of the preceding stop is different for labials, alveolars, and velars, and it is only significant for labials. Table 24 shows the differences in open interval and vowel duration for labials.

Table 24

Differences in cs between CV-sequences with labial stop + rounded and unrounded vowel (subjects 1-2, 5-12, 17). S3, S6, M2 long vowels, S4, M1 short vowels, S5 short + long vowels).

	S3 N=15	S4 N=15	S6 N=6	S3 N=15	S6 N=6	M1A N=84
p	yøœ > ieɛ			b yøœ > ieɛ		y > i
op.int.	+1.5	+0.3	+1.5	+0.5	+1.4	+0.6
vowel	+1.1	-0.2	+1.1	+1.8	+0.8	-0.5
p	uoɔ > ieɛ			b uoɔ > ieɛ		
op.int.	+1.5	+0.3	+1.0	+0.4	+0.9	
vowel	+1.4	+0.6	+1.2	+2.4	+0.4	
	M2 N=12	**M1B N=42**	**S5 N=60**	**M2 N=12**	**M1B N=42**	**M1A N=84**
p	u > i			b u > i		
op.int.	-0.3	+0.9	+0.7	+0.5	+0.5	+0.5
vowel		+0.6	+1.6		+1.4	-0.3

It appears from table 24 that the open interval of labials is longer before rounded than before unrounded vowels. The difference is significant for both p and b (for p the open interval is longer in 31 out of 37 individual averages (84%) and for b in 30 out of 35 (86%). Rounded vowels are also generally longer than unrounded vowels, but this depends in part on the following consonant. Rounded vowels are longer than unrounded ones in lists S3, S5, S6, M2 and M1B, where the following consonant is alveolar (n, l or d), but u and y are shorter than i in list M1A and variable in list S4. In these two lists the following consonant was b or p. (This influence has been shown in detail in Fischer-Jørgensen 1964.) The open interval is longer, irrespective of this vowel difference. The closure is not longer before rounded vowels.

As for alveolars, there is an opposite but not significant tendency to have a longer open interval before unrounded vowels. As for velars, the relations are irregular. This explains that although b has generally a shorter open interval than d, this is not always the case in list M1A, because the vowels are i, y and u. It also explains why the open interval was found to be in decreasing order in g>d>b except before rounded vowels, where the order is g>b>d.

The tendency to longer duration of the consonants f s h before rounded than before unrounded vowels mentioned in my 1979 paper is weak and not significant.

6. Summary and discussion

6.1 Summary

The results of the present investigation of Danish CV-sequences with stop consonants can be summarized as follows:

6.1.1 p t k vs. b d g

CV-sequences with ptk have a much longer open interval than those with bdg, and at the same time a shorter closure and following vowel, but the shortenings of closure and vowel do not generally counterbalance the lengthening of the open interval, so that for almost all subjects, CV-sequences with ptk are longer than those with bdg (grand mean 2.1 cs) (see section 5.3.1 and figs. 1-3).

6.1.2 Place of articulation

Labials, alveolars, and velars differ both in closure and in open interval, the order for the closure being b>d>g and p>k>t and for the open interval g>d>b and t>k>p. The differences are significant with exception of the closure d>g and, consequently, the inverse relation between closure and open interval for d-g. There is thus an inverse relation between closure and open interval, and it is of almost the same magnitude so that CV-sequences with consonants of different place of articulation are approximately of the same length. There is generally no difference in the following vowel (see section 5.3.2 and figs. 4-5).

6.1.3 Phonological vowel length

There is no consistent compensation in consonant duration in relation to the phonological length of the following vowel (see section 5.3.3).

6.1.4 Vowel height

The open interval - and often the closure - is longer before high vowels than before low vowels. But the differences in vowel duration (the high vowels being longer than the low vowels) are in almost all cases larger than the differences in open interval and closure, so that the CV-sequences with low vowels are normally longer than those with high vowels (grand mean +2.5 cs). (See section 5.3.4 and figs. 7-9). The fricatives f and s are also longer before high vowels, but there is no clear tendency in the case of h m l.

6.1.5 Vowel rounding

Labial stops have a longer open interval before rounded vowels than before unrounded vowels. Since rounded vowels are, on the average, longer than unrounded vowels, the CV-sequences with labials + rounded vowels will be longer than those with unrounded vowels

6.2 Discussion

The problem is now whether the relations which have been found should be explained as compensation, in the sense that the inverse relations are due to a tendency to maintain a constant duration of the CV-sequences, as it has often been proposed for VC-sequences and for larger units, or whether they are due to more specific physiological mechanisms.

Since the relations between ptk and bdg (section 6.1.1 above) and between different places of articulation (section 6.1.2) are the most problematic, I shall treat them at the end and start with a discussion of phonological vowel length, vowel height, and vowel rounding (sections 6.1.3, 6.1.4, and 6.1.5, respectively).

6.2.1 Phonological vowel length and rounding

It is obvious that no compensation takes place between phonological vowel length and preceding consonant (6.1.3). One might set up the hypothesis that programmed phonological differences do not cause compensation within the syllable. This seems to be true of phonological vowel length in Danish, also in relation to the following consonant, but it is not a universally valid rule; it is, e.g., not valid for languages like Swedish and Italian, where long vowels are followed by short consonants and short vowels by long consonants. One might also try the hypothesis that there is no compensation between vowel and preceding consonant. This is valid for phonological vowel length and also for the phonetic difference due to rounding of the following vowel (6.1.5). (Labials before u have longer open interval than before i, although the vowel u is also generally longer than i.)

6.2.2 Vowel height

It might be objected that the hypothesis of no compensation is not valid for differences in consonant duration due to vowel height. Here an inverse relation between vowel and preceding consonant was found. But this may not be a case of compensation but a coincidence due to specific conditions of sound production.

Since, as seen in table 21, the lengthening of the open interval before i and u compared to a is quite independent of the phonological length of the vowel, i.e. the consonant is lengthened before long i: compared to short a, although i: is in fact longer than a, one might even take this as proof that the duration of the consonant has nothing to do with the duration of the following vowel but only with its quality (cf. Klatt, 1975). This is, however, not sufficient proof; it only shows that it is not crude phonetic duration which is at stake. It might be so that phonological length does not cause compensation, as already shown above, but within long and short vowels there could be a compensation for differences of duration due to vowel height.

(Nooteboom (1972) found a difference of duration between words with long and short vowels but not between words with high and low vowels.) There is, however, a further and, as far as I can see, more convincing argument against considering the longer open interval before high vowels as due to compensation, viz. that it is possible to give independent physiological explanations of the durations of consonants and vowels and a common explanation for rounding and vowel height.

The longer duration of the open interval of labials before rounded vowels can be explained by the fact that, particularly in high vowels, where the longer open interval is very clear, the lips are only opened very little from consonant to vowel, consequently the airflow is relatively slow, and the supraglottal pressure goes down slowly, which delays the start of voicing. Similarly, in the case of high vowels, the opening from consonant to vowel is of relatively small extent, and the airflow is delayed. This is particularly evident for the Danish affricated t, but it is also evident for the velars. As for the labials, the degree to which the airflow is delayed may depend on the degree of coarticulation and the relation is, in fact, also found to be more variable, also individually, in labials. In the present material pi has in most cases a longer open interval than pa, but some subjects have the opposite relation.

The difference in duration between high and low vowels has been explained by the extent of the articulatory movement. Jespersen (1932) seems to have been the first to give this explanation (a fact which I had overlooked in earlier papers (1955, 1964), cf. also Maack 1953, Fischer-Jørgensen 1955 and 1972, Kohler and Künzel 1978). The longer distance is partly, but not fully, compensated for by the speed of the movement (Lindblom 1967). The slightly longer duration of rounded vowels can, similarly, be explained by the extra articulation. This means that both the differences in vowel duration and the difference in open interval can be physiologically explained, and their interrelation in the time domain may be accidental and has nothing to do with compensation.

This explanation is corroborated by the fact that neither the difference in vowel length nor in length of open interval are specific Danish phenomena but well documented in a large number of languages (see, for the literature on vowel height, Lehiste 1970

p. 18 and Lindblom 1967, and for the open interval, e.g. for French Fischer-Jørgensen (1968 and 1972), Serniclaes (1974), Wajskop (1979), for German Fischer-Jørgensen (1976) (ti > ta and ki > ka, but pi-pa variable), and Kohler (1978), for English Klatt (1975), Summerfield (1975) (ki > ka, but pa > pi). I have found similar relations in a Dutch material. The dependency of the open interval on vowel height thus seems to be a general, physiologically conditioned tendency.

The lengthening of the fricatives f, v and s before high vowels can be explained in the same way, viz. as a delay of the airflow, and the fact that this lengthening is not consistent for h, m and l, which have a freer air escape, corroborates this assumption.

The physiological explanation is also corroborated by actual measurements of air pressure decay and airflow rise. Fischer-Jørgensen (1968) gives an analysis of the stop consonants of a French-Danish bilingual subject, which shows complete agreement between decay of intra-oral air pressure after the release, rise of airflow and length of open interval before the vowels i, a, u for bdg and ptk in French and Danish (with one exception). There is also agreement between decay of air pressure and length of open interval before i a u for pt (three German subjects) and k (only one subject measured) in Fischer-Jørgensen (1976), again with only one exception (only the air pressure data are given in the paper). Danish material will be published later; it shows a clear correlation between decay of intra-oral air pressure and length of open interval for ptk. The explanation is also supported by the fact that the most consistent lengthening of the open interval before i compared to a is found for the consonant t (also in Hutters 1979), and t is affricated and has a very slow escape of air after the release, particularly before i.

The most serious counter-argument against a purely physiological explanation is the fact that in many cases the closure of the stop consonants is also lengthened before high vowels, although in my material (but not in Hutters 1979) less consistently than the open interval. The German material shows the same very clearly, whereas only a weak tendency is found for the French-Danish bilingual subject. The problem is whether this forces us to recognize a real case of compensation, i.e. to assume that the lengthening of the closure is due to a tendency to keep Ci- and Ca-sequences at approximately the same length (as mentioned in 5.3.4, the compensation is never complete, so that the goal is not reached).

I do not think that the possibility of compensation can be excluded, but, on the other hand, it is also possible to find a physiological explanation for the closure. Lindblom (1967) has shown that the opening movement of the jaw for the vowel begins during the preceding consonant and that it begins earlier for low than for high vowels. This would favour an earlier release of the consonant and thus shorten it. Lindblom finds that this effect may be prevented by a later opening activity of the lips for low vowels. This turns out to be the case for his subjects A and C, but not for B. Thus, there seems to be different individual strategies. Lindblom has only investigated labials, where a further complication is the possibility of more or less coarticulation with the tongue movements. It might be assumed that in general the jaw movement is dominating, and that the shortening of the consonant is thus not prevented by other factors, but the irregularity of the shortening shows that more factors may be at work.

6.2.3 Place of articulation

As concerns the difference between labials, alveolars and velars, it may be useful first to examine what is universal in the relations found in Danish and how the deviation from the universal tendencies can be explained. As mentioned above (5.3.2), the order in Danish stops is b>(d>g) and p>k>t for the closure, and the opposite order g>d>b and t>k>p for the open interval.

The relations for bdg correspond to a universal tendency, see Lehiste 1970, p. 27-28, Fischer-Jørgensen (French) 1968 and 1972 (in the latter case only as an average over different speakers). The same relations were found in a Hindi material. For the closure the tendency is rather weak. Kohler (1978) and Mansell (1979) have found the order b>g>d for German. This is also common for Danish subjects, and two of six German subjects in my investigation (1976) deviated in different ways from the normal order b>d>g. For the open interval the order g>d>b is, however, very stable, cf. Lisker and Abramson 1964 and 1965 (various languages), Bothorel-Witz and Pétursson 1972 (Icelandic), Klatt 1975 and Zue 1976 (English), Löfqvist 1976 (Swedish), and Wajskop 1979 (French).

As concerns ptk, the relations p>k>t for the closure and t>k>p for the open interval are, however, specific for Danish, the normal relations being the same as for bdg, i.e. p>t>k for the closure and k>t>p for the open interval (see the references for bdg above). It is interesting that the deviation in Danish concerns both the closure and the open interval, which means that there is an inverse relation between closure and open interval in all cases.

It was mentioned in 5.3.2 that the reason for the deviating relations for Danish ptk was the strong affrication of t, involving a longer open interval and a shorter closure. French speakers, who have affrication of t before i also have a longer open interval in t than in k in this case, and German affricates compared to plosives are characterized by a shorter closure and a longer open interval and have a slower opening of the articulators, evidenced by a slower decay of the intra-oral pressure and a slower rise of the airflow after the release (Fischer-Jørgensen 1976).

On this background it seems promising to look for similar explanations of the universal relations. The normal short closure and long open interval of velars may be explained by the slow movement of the body of the tongue compared to the tongue tip and the lips. As the closing and opening movement of a stop is regarded as being part of the preceding and following vowel (or the open interval), respectively, velars will have a shorter closure and a longer open interval. In labials, on the other hand, the movement of the lips is independent of the vowel articulation, the closure can therefore start earlier and last longer without hampering the articulation of the vowels, the lips can also move more quickly than the body of the tongue. This explanation has the advantage of being of the same type as that offered for the relations due to vowel height and rounding, and it has the advantage of seeking a common cause for the durations of closure and open interval and their inverse relation, so that it is not necessary to invoke the concept of compensation. The correlation diagrams in fig. 6 also show an inverse relation between the consonants (e.g. t and p), but no inverse relation within the same consonant. It has also been argued that there will be a higher pressure behind the point of articulation the farther back it is, and that it will take a longer time for this pressure to decrease to a level which permits voicing. But this higher pressure might also be expected to lead to a stronger airflow instead of a slower decay of the pressure. It might, however, contribute to the shortening of the closure.

There may, however, be other factors involved. Hutters (1978 and 1979) has found differences in the glottal activity in Danish stops according to place of articulation, the maximum glottal opening being larger in k than in p, whereas t varies. Moreover, the oral release comes earlier in relation to the glottal movement in the order t>k>p, at least before the vowel i. Similarly, g has a larger maximum glottal opening and an earlier oral release relative to the glottal closing movement than b. A larger glottal opening in k than in p has also been found for other languages, e.g. Japanese (Sawashima and Niimi 1974) and Icelandic (Pétursson 1976). The problem is whether these differences are artefacts of the method, perhaps due to a raising of the larynx in velars and, if they are real, whether the differences are programmed or whether they are due to some sort of reflex mechanism conditioned by aerodynamic factors, e.g. the pharyngeal pressure (Hutters 1979). This needs further investigation.

Klatt (1975) has a different explanation of the longer open interval in k. For bdg he assumes physiological constraints, but for p-k he supposes that there are perceptual reasons. He thinks that a longer VOT is needed in a voiceless plosive with slower formant transitions, such as k, to prevent listeners from hearing the low frequency energy cue that would indicate a g. It does not seem very plausible to explain the parallel order k>t>p and g>d>b in quite different ways. The slower transitions in velars should rather be considered as the acoustic result of the same slower movement which hampers the airflow and consequently the voicing. If these causal relations were disrupted, it would of course cause confusion, but I cannot see that perceptual factors should be more important in this specific case than in other cases.

6.2.4 ptk vs. bdg

Finally, there is the problem of the differences between the CV-sequences with ptk vs. bdg, where it was found that in the ptk -sequences the open interval is longer, but both the closure and the following vowel shorter than in the bdg-sequences. Can these differences be explained by the physiological mechanisms or is it a case of compensation?

It should first be stated that the relations found in Danish are not universal. The symbols "ptk" and "bdg", when used in broad transcription of different languages, cover differences which may have some auditory similarities (although there are overlappings, cf. that the unaspirated voiceless stops in Chinese, Icelandic and Greenlandic are sometimes transcribed as b d g, sometimes as p t k), but which must be described by means of three different oppositions: voicing, aspiration or fortis-lenis, or by various combinations of these oppositions. In Danish the main difference is, as stated earlier, one of aspiration, but there may be a concomitant difference of fortis-lenis, so that the aspirated is more lenis (see below).

The mechanism of aspiration is fairly well understood. In Danish (Frøkjær-Jensen, Ludvigsen and Rischel 1971, Hutters 1979) as well as in Hindi (Kagaya and Hirose 1975) and Icelandic (Pétursson 1976) the difference between aspirated and unaspirated stops is evidently due to a different programming of the glottal gesture: aspirated stops have a wider glottal opening and an approximately symmetric closing-opening movement with its maximum close to the oral release, whereas unaspirated stops have a smaller opening with the maximum opening in the beginning of the closure period so that the glottis is practically closed at the release. These different glottal mechanisms explain the differences in aspiration (cf. also Kim 1970), but the cause of the shortening of the closure is not quite clear. It might be assumed that a stronger intra-oral pressure in the aspirated stops would contribute to weaken and open the oral closure. Two of my Indian subjects have a definitely higher intra-oral pressure in aspirated stops. But one did not have any evident difference. In Danish the difference is very small, and in Icelandic the two types do not show any difference in intra-oral pressure. So this is not a sufficient explanation. It is more probable that a separate timing command of the release is part of the aspiration feature. A very long closure would prevent aspiration. The French-Danish bilingual subject whose stops I investigated in my 1968-paper had the same duration of French and Danish bdg, but her French ptk had a longer closure and her Danish a shorter closure. Löfqvist 1976 has also drawn attention to the very long closures of the Swedish medial unaspirated stops, which may have a relatively large glottal opening. Generally, however, the differences in the magnitude and particularly the differences in the position of the maximum of the glottal opening are so large that they give room for quite a variation in oral closure duration with preservation of the difference aspirated - unaspirated; and the different timing does not explain the weaker closure in aspirated stops which has been found in several languages (see below), and which must also be the cause of the further development to affricates in Danish t (and in the Old High German consonant shift). It may therefore be assumed that there is a concommittant fortis - lenis feature.

The fortis - lenis opposition, manifested initially by lengthening of the closure, a stronger organic pressure, and, according to Debrock (1977), by a quicker rise of the intensity of the following vowel, only rarely occurs alone. It generally accompanies the voiced - voiceless opposition, so that the voiceless member is more fortis (e.g. in French and Dutch). It may also accompany the opposition of aspiration, but in two different ways:

(1) In languages in which ptk and bdg are distinguished initially by aspiration and voicing (like Swedish) or by aspiration and optional voicing (like English and German), and medially before unstressed vowel by voicing, ptk are often assumed to be more fortis than bdg. This seems to be true for the medial position (cf. Löfqvist 1976 for Swedish and Kohler 1977 for German), but for the initial position one generally finds rather inconsistent indications of duration, organic pressure and EMG-activity (cf. e.g. for English Lisker 1966, Harris, Lysaught and Schvey 1965, Kent and Moll 1969, and Lubker and Parris 1970). Thus, the fortis feature is only evident when combined with obligatory voicing (i.e. medially in German and English).

(2) In languages which have a pure opposition of aspiration without a voicing opposition (like Danish and Icelandic) or with the two oppositions clearly independent of each other (like many Indian languages), it is the unaspirated member which seems to be more fortis (thus in Danish and Icelandic the stops that are written bdg in initial position). Danish bdg do not sound as fortis as e.g. French or Indian unaspirated ptk, but they have a longer closure than Danish aspirated ptk and a tendency to higher organic pressure and to stronger EMG activity, at least at the release of b, for some also at the implosion (Fischer-Jørgensen and Hirose 1974). Icelandic bdg sound almost like Danish, but Pétursson transcribes them as [ptk] and considers them as fortes (Bothorel-Witz and Pétursson 1972, Pétursson 1976 and Löfqvist and Pétursson 1976). They normally have a longer closure than the aspirated stops, and more contact on palatograms and a closer jaw distance. As for Indian unaspirated ptk, they have been found to have a longer closure than the aspirated stops (Kagaya and Hirose 1975, Dixit 1975 and Benguerel and Bhatia 1980 (Hindi), and Senn 1935 (Bengali)). I have found the same relations in Hindi, Gujarati and Dogri (unpublished material), and for a Gujarati speaker a lower organic lip pressure in ph than in p. Rousselot (1897-1908) also found a lower lip pressure in Armenian p^h than in p.

I therefore assume that the shorter duration of the closure in Danish aspirated stops is mainly due to weakness of articulation. In any case I do not think it is simply a question of compensation.

As for the shortening of the following vowel after aspirated stops, it is a consistent phenomenon in Danish. I have not found any indications for Icelandic, but I have found the same shortening in Gujarati and Dogri. Peterson and Lehiste (1960) found a slight shortening in American English. As for Swedish, I mentioned in my 1979-paper that Fant (1970) had found a very drastic shortening of the vowel after aspirated stops (6 cs). This quotation was, however, misleading. This large difference was only found in the word type pa'pa:pa vs. ba'ba:ba, where the following consonant is evidently mainly responsible, as Fant also remarks himself. When the following consonant is identical, the shortening is much smaller.

It is evident that the start of voicing is delayed by aspiration, so that the vowel starts later, but it would be possible to delay the end of the vowel correspondingly and thus get the same duration after aspirated and unaspirated stops. The fact that this is not done, is probably due to a tendency to maintain a relatively equal duration of the whole sequence, i.e. to a factor of compensation. This assumption is corroborated by the observation that vowels may also differ in duration after ptk and bdg in languages with unaspirated ptk. Thus in French vowels are shorter after ptk than after bdg, see Wajskop (1979) and Fischer-Jørgensen (1972). In the latter paper it is also shown that the lengthening of vowels after bdg exceeds the small difference in open interval after ptk and bdg in French. In 1972 I proposed that it might be due to the fortis character of ptk, but that is not a good explanation. It is more probable that it is a compensation for the longer closure of ptk. I have found a similar lengthening of vowels after bdg compared to unaspirated ptk in Hindi, Dogri and Gujarati. (But Kozhevnikov and Chistovich (1966) did not find any difference of vowel duration after t and d in Russian.) Moreover, measurements of stops in Gujarati have shown that the affricates č and čh have a longer open interval than k and kh, respectively, but also a longer following vowel, the difference being quite consistent. At the same time, however, the affricates have a shorter closure, so that the total duration of the affricates is shorter than that of the plosives. The difference of the vowels must therefore be explained as a compensation phenomenon. This is not, however, a quite general phenomenon for CV-sequences since the vowel does not seem to be lengthened after h (see 5.3.1 above); but perhaps, as Reinholt Petersen has suggested to me, h may have a special status since it has no supra-laryngeal articulation. The fact that Lehiste (1971) did not find any consistent compensation in CV-sequences does not directly invalidate the assumption made here, since what she measured was the variation of the same word, repeated a large number of times, thus the variation of the same sounds in a single token, whereas what has been discussed here is a problem of inverse relation between sequences containing different consonants or different vowels with different intrinsic duration. Fig. 4a also shows that there may be a clear inverse relation between closure and aspiration in ta and pa (whether due to real compensation or to physiological factors), whereas there is no inverse relation between the durations of closure and aspiration in single tokens of the two individual words.

However, in Danish it is not so that the vowel stops at the same distance from the release of ptk and bdg, nor is it so in the Indian languages. That would give an acoustic vowel duration after ptk which would probably be below an acceptable minimum. After long aspirations it might even disappear completely. One seems to strike a compromise between equal duration of vowel gesture, and equal duration of perceptible (voiced) vowel, but a compromise which is considerably closer to the acoustic vowel duration after bdg (cf. table 12 in 5.3.1, showing that the vowel is shortened much less than the open interval after ptk is lengthened).

Fant (1969) also emphasized that the temporal organization is not simply a matter of delay of voicing in ptk vs. bdg at the expense of vowel duration, the differences in the (voiced) vowel durations after ptk and bdg being much smaller than the differences in open interval. Moreover, there may be differences in the vowel gesture according to the preceding consonant. Fant quotes Öhman (1965), who has shown that when curves of Swedish words of e.g. the types kà:da and gà:da are lined up in such a way that the tone contours cover each other as exactly as possible, then the start of the vowel after g will not coincide with the start of the vowel after k, but will correspond to a point in the middle of the aspiration interval of k, and the release of k will lie about 4 cs earlier than the release of g. Öhman mentions that exactly the same result will be obtained if the words are lined up according to formant transitions instead of tone contours. Fant has found this assumption corroborated in a number of Swedish spectrograms. He found a displacement of the release of ptk vs. the release of bdg of about 3 cs when the formant transitions are lined up, and he adds that this involves either that the formants start at different frequencies at the release of ptk compared to bdg, i.e. closer to the consonant target, which means that there is less coarticulation with ptk, or they start at the same point but move more slowly in the beginning after ptk. He finds that the former explanation is valid for labials, and probably the latter for palatals and alveolars. Gay (1979) has supported this assumption by means of EMG measurements showing that the muscle activity for the following vowel starts earlier in bip and bap than in pip and pap, whereas this is not the case for t-d. Fant suggests some rules for the lengthening of the vowel compared to the differences in open interval and in release, but this will probably require further investigations. A preliminary examination of a number of Danish spectrograms has shown a pronounced displacement for t-d, which may be explained by the affrication of Danish t, involving a later start of the vowel movement, and in labials before back vowels, but no displacement of labials before i and hardly any difference in velars. These relations thus seem to be very complicated. A very preliminary examination of some French spectrograms did not show any clear displacements, and I would venture the hypothesis that in so far as the differences are due to a slower vowel gesture after ptk, this may be part of the lenis feature of aspirated stops. It must, however, be mentioned that Kozhevnikov and Chistovich (1965) have found a slower opening also of unaspirated p than of b in Russian.

Since the vowel movement takes place partly during the aspiration after ptk, one might think of considering the aspiration as part of the vowel and not as part of the consonant. This would, however, give absurd vowel durations in Danish. A third possibility is to consider it as a separate element. This would be in agreement with the phonological interpretation of Danish ptk as bdg + h, first proposed by Uldall, and accepted by Hjelmslev (1951) and others. Frøkjær-Jensen, Ludvigsen and Rischel (1971) have also found that glottograms of Danish aspirated p are very similar to glottograms of final b + initial h, and they suggest an interpretation of p as a summation of a b- and an h-gesture. A very similar description is given by Kagaya and Hirose (1975) of Indian aspirated stops. Holtse (1977) has measured the distance from the release of the stop to the end of the vowel in syllables with p and b + V, rV, lV and jV for two subjects. He finds that the duration of pV is approximately the same as that of bjV, brV and blV, whereas pjV, prV and plV have still longer durations. This is in agreement with Lehiste's measurements of English clusters with stop consonants (1972). She found that the duration of aspiration following initial voiceless stops is almost the same as that of resonants following voiced stops. Holtse draws the conclusion that p might be interpreted as b+h, where h is a separate unit in the programming. The shortening of the vowel (and probably of the closure, which Holtse has not measured) could then be seen as a compensatory shortening due to the larger number of segments in the syllable. This is an interesting suggestion. I am, however, inclined to think that more plausible explanations are reached if the aspiration is considered to be part of the stops, as I have done above.

My main endeavour in the preceding pages has been to try to see how far it is possible to find physiological explanations of the different temporal relations without taking recourse to the concept of compensation, a concept which is certainly important, but which is sometimes too easily advocated and which may sometimes prevent an analysis of the production mechanism. The main result was that physiological production mechanisms may explain the temporal relations in most cases: the longer open interval in labials before rounded vowels, the longer open interval (and perhaps also the closure) before high vowels, the differences according to place of articulation, and the longer open interval and the shorter closure in ptk vs. bdg. - But I found indications of a compensation mechanism for the different vowel durations after ptk vs. bdg.

References

Abrahams, H. 1949: Etudes phonétiques sur les tendances évolutives des occlusives germaniques Aarhus.

Benguerel, A.P. and T.K. Bhatia 1980: "Hindi stop consonants, an acoustic and fiberoptic study", Phonetica 37, p. 134-148

Bothorel-Witz, A. and M. Pétursson 1972: "La nature des traits de sonorité et d'aspiration dans le système des occlusives de l'allemand et de l'islandais", Travaux de l'Institut de Phonétique de Strasbourg 4, p. 277-356

Debrock, M. 1977: "An acoustic correlate of the force of articulation", JPh 5, p. 61-80

Dixit, R.P. 1975: Neuromuscular aspects of laryngeal control with special reference to Hindi, Ph.D. thesis, Austin

Fant, G. 1960: Acoustic Theory of Speech Production, The Hague

Fant, G. 1970: "Stops in CV-syllables", STL-QPSR 4/1969, and Speech Sounds and Features 1973, p. 110-139

Fischer-Jørgensen, Eli 1955: "Om vokallængde i dansk rigsmål", NTTS 15, p. 33-56

Fischer-Jørgensen, Eli 1964: "Sound duration and place of articulation", Zs.f.Ph. 17, p. 175-207

Fischer-Jørgensen, Eli 1968: "Les occlusives françaises et danoises d'un sujet bilingue", Word 24, p. 112-153

Fischer-Jørgensen, Eli 1972: "ptk et bdg français en position intervocalique accentuée", Papers in Linguistics and Phonetics in Memory of Pierre Delattre (The Hague), p. 143-200

Fischer-Jørgensen, Eli 1976: "Some data on North German stops and affricates", ARIPUC 10, p. 149-200

Fischer-Jørgensen, Eli and H. Hirose 1974: "A preliminary electromyographic study of labial and laryngeal muscles in Danish stop consonant production", Haskins SR 39/40, p. 231-254

Frøkjær-Jensen, B., C. Ludvigsen and J. Rischel 1971: "A glottographic study of some Danish consonants", F&S (Hammerich, Jakobson and Zwirner eds.), p. 123-140, Copenhagen

Gay, T. 1979: "Coarticulation in some consonant-vowel and consonant cluster-vowel syllables", Frontiers of Speech Communication Research, Festschrift for Gunnar Fant, eds. Lindblom and Öhman, (London), p. 69-76

Harris, K.S., H. Lysaught and M.M. Schvey 1965: "Some aspects of the production of oral and nasal labial stops", LS 8, p. 135-148

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

Holtse, P. 1977: Variationer i vokallængde på dansk (unpublished prize essay)

Hutters, Birgit 1978: "The glottal gestures in some Danish consonants", ARIPUC 12, p. S87-S101

Hutters, Birgit 1979: Glottisfunktionen ved produktion af ustemte obstruenter i dansk, en glottografisk og fiberoptisk undersøgelse, (unpublished manuscript (127 pp))

Jespersen, O. 1932: Lehrbuch der Phonetik (1912) [5]1932, Leipzig, p. 181

Kagaya, R.A. and H. Hirose 1975: "Fiberoptic, electromyographic and acoustic analyses of Hindi stop consonants", Ann. Bull. Univ. Tokyo, Logopedics and Phoniatrics 9, p. 27-46

Kent, R.D. and K.L. Moll 1969: "Vocal tract characteristics of stop cognates", JASA 46, p. 1549-1555

Kim, C.-W. 1970: "A theory of aspiration", Phonetica 21, p. 107-116

Klatt, D.H. 1975: "Voice Onset Time, frication and aspiration in word-initial consonant clusters", JSHR 18, p. 686-706

Kohler, K. 1977: "The production of plosives", Institut für Phonetik, Kiel, Arbeitsberichte 8, p. 30-131

Kohler, K.J. and J. Künzel 1978: "The temporal organization for sequences of vowels and plosives in German", Institut für Phonetik, Kiel, Arbeitsberichte 10, p. 117-167

Kozhevnikov, V.A. and L.A. Chistovich 1966: Speech, Articulation and Perception, translated by Joint Publications Research Service, Washington D.C. No. IPRS 30.543

Lehiste, Ilse 1970: Suprasegmentals, Cambridge, Mass.

Lehiste, Ilse 1971: "Temporal organization of spoken languages", F&S (Hammerich, Jakobson and Zwirner eds.), p. 159-169

Lehiste, Ilse 1972: "The syllable nucleus as a unit of timing", Proc.Ling. 11, II, p. 929-933

Lindblom, B. 1967: "Vowel duration and a model of lip mandible coordination", STL-QPSR 4, p. 1-29

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

Lisker, L. and A. Abramson 1965: "Voice onset time in the production and perception of English stops", Haskins SR-3, 3, p. 1.1 - 1.7

Lisker, L. 1966: "Measuring stop closure duration from intra-oral pressure records", Haskins SR-7/8, p. 5.1 - 5.6

Lisker, L. and A. Abramson 1971: "Distinctive features and laryngeal control", Language 47, p. 767-785

Lubker, J. and P.J. Parris 1970: "Simultaneous measurements of intra-oral air pressure, force of labial contact, and labial electromyographic activity during the production of the stop consonant cognates p and b", JASA 47, p. 625-633

Löfqvist, A. 1975: "On the control of aspiration in Swedish", Working Papers, Lund, 12, p. 99-106

Löfqvist, A. 1976: "Closure duration and aspiration for Swedish stops", Working Papers, Lund, 13, p. 1-40

Löfqvist, A. and M. Pétursson 1978: "Swedish and Icelandic stops: a glottographic study", The Nordic Languages and Modern Linguistics, p. 454-461

Maack, A. 1953: "Die Beeinflussung der Sonantendauer durch die Nachbarkonsonanten", Zs.f. Ph. 7, p. 104-128

Mansell, Ph. 1979: "The articulation of German plosives", Forschungsberichte 11, Institut für Phonetik, München, p. 1-207

Meyer, E.A. 1904: "Zur Vokaldauer im Deutschen", Nordiska Studier tillägnade A. Noreen, p. 347-356

Nooteboom, S.G. 1972: Production and perception of vowel duration (Thesis, Utrecht)

Öhman, S.E.G. 1965: "On the coordination of articulatory and phonatory activity in the production of Swedish tonal accents", STL-QPSR 2, p. 14-19

Peterson, G. and Ilse Lehiste 1960: "Duration of syllable nuclei in English", JASA 32, p. 693-703

Pétursson, M. 1976: "Aspiration et activité glottale, examen expérimental à partir de consonnes islandaises", Phonetica 33, p. 169-198

Rousselot, P.-J. 1897-1908: Principes de phonétique expérimentale, Paris

Sawashima, M. and S. Niimi 1974: "Laryngeal condition in articulation of Japanese voiceless consonants", Ann. Bull. Univ. Tokyo, Logopedics and Phoniatrics 8, p. 13-17

Senn, A.C. 1935: "An experimental study of Bengali occlusives", Proc.Phon. 2, p. 184-193

Serniclaes, W. 1974: "Influence du contexte vocalique sur la perception du voisement des occlusives", Rapport d'activité, Inst. Phon. Univ. Libre de Bruxelles 8, p. 101-108

Slis, I.H. 1967: "What causes the voiced-voiceless distinction?", IPO APR 2, p. 71-76

Summerfield, Q. 1975: "Aerodynamic versus mechanics in the control of voicing onset in consonant/vowel syllables", Speech Perception, Belfast, 4, p. 61-72

Wajskop, M. 1979: "Segmental durations of French intervocalic plosives", Frontiers of Speech Communication Research, Festschrift for Gunnar Fant (Lindblom and Öhman eds.), London, p. 119-123

Zlatin, M.A. 1974: "Voicing contrast: Perceptual and productive voice onset time characteristics in adults", JASA 56, p. 981-994

Zue, V.W. 1976: Acoustic characteristics of stop consonants, M.I.T. Lincoln Laboratory, Technical Report 523

(ARIPUC 14, 1980)

PHONOLOGICAL SYLLABIFICATION IN DANISH ONCE MORE: A PROPOS MOLBÆK'S PAPER

Hans Basbøll[1]

Abstract: This paper contains an evaluation of Molbæk Hansen's critical discussion of my earlier syllabification principles of Danish phonology. A more recent version of the principles, which has been tested within the DANFON-project, is presented - although very sketchily - and Molbæk Hansen's counter-examples to the earlier principles are evaluated with respect to the DANFON-version. Particular attention is paid to the manifestation of short /a/, synchronically as well as diachronically.

1. Introduction[2]

Peter Molbæk Hansen's recent paper (1979) contains an interesting discussion of my principles of phonological syllabification in Danish published earlier (see below), and a most challenging set of apparent and real counter-examples to them. The editors of ARIPUC have been so kind as to allot me in this volume the space necessary for a reaction to some of Molbæk's criticism. I shall try to clarify (in section 2) my position on some issues where I am not sure I agree with Molbæk's interpretation, and, in particular, I shall present (in section 3) an outline of the phonological syllabification principles which have in fact been used in the DANFON-project, which is a computer testing of a generative phonology of Danish, conducted by Kjeld Kristensen and myself (see

---

1) Scandinavian Department, Odense University.

2) Most of the contents of the paper were presented - in a rather condensed form - at the 4th International Conference of Nordic and General Linguistics, Oslo, June 23-27, 1980. I am indebted to John Dienhart for stylistic suggestions, and to Peter Molbæk Hansen for many valuable comments on the manuscript.

Basbøll and Kristensen 1975), and evaluate Molbæk's counter-examples with respect to these latter principles. (Particular attention is paid to the manifestation of short /a/ as [a] or [ɑ].) It is hoped that this evaluation will clarify to what extent the shortcomings of the earlier principles are due to the tentative nature of their formulation and/or to the inherent failure of my whole strategy of syllabification. It is shown that the DANFON syllabification principles are observationally more adequate than those of the earlier version, but that certain classes of counter-examples nevertheless seem to remain. Lastly (in section 4) I discuss Molbæk's suggestions in favour of a unit larger than the syllable but smaller than the word, and an alternative account is sketched. (I should like to add that section 2 contains very little that is new, but that I have considered it as a prerequisite for the remainder of the paper.)

2. General remarks on my syllabification principles

2.1 Preliminaries

My point of departure is the following hypothesis (see e.g. Basbøll 1978a): The function of boundaries, which are linearly ordered, is to delimit "domains" (like "syllable", "phonological word" and "stress group"). The primary function of those domains is to define the "universe of application" of phonological rules and phonotactic constraints; but they also occur as units in such rules (e.g. syllables in tone- or stress-rules). Boundaries should not occur properly included in the Structural Description of a phonological rule, i.e. they may occur at the very beginning or at the very end of the Structural Description, but not within it.

This framework considers the function of syllabic boundaries and phonologically relevant grammatical boundaries to be phonologically alike, and thus the syllable should be a possible domain for phonological rules, just as it is for phonotactic constraints. This was also one of the main points of my treatment of the short vowels in Danish (Basbøll 1972), even though the whole junctural framework was only worked out later. The domains of the present model in fact represent a simple type of hierarchically structured phonological organization (cf. Basbøll (forthcoming) for a discussion of recent versions of metrical or hierarchical phonology which seem generally over-structured to me).

As already stated, it has been claimed (Basbøll 1972) that a number of important phonological processes in Danish have the syllable as their domain (such syllable-dependent rules are the adjustment of short /o/ and /a/ in closed syllables and a whole range of consonant gradation-phenomena which are so characteristic of Danish (see Rischel (1970)), and which constitute a main reason why spoken Danish is so difficult to understand for other Scandinavians). It is then necessary, of course, to propose some principles of syllabification, since these are presupposed by such an account. In different papers (e.g. 1972 and 1974) I have discussed, in general terms, such principles (for instance that certain grammatical boundaries function as syllable boundaries too; that a stressed vowel attracts neighbouring consonants; that a full vowel attracts more consonants than a schwa; and that certain consonant clusters, like obstruent plus liquid in some languages, or /s/ plus certain consonants in others, function as single consonants with regard to syllabification). I proposed that such general principles (including the so-called "Hjelmslev's law", according to which a medial cluster should be split up into a possible final cluster plus a possible initial cluster) interact in different ways in different languages, and quite tentatively I suggested that the resultant principles of syllabification for Danish should be something like the following:

(1) The grammatical boundaries preceding pauses, words, stems and (primarily or secondarily) stressed native suffixes function as syllable boundaries too.

(2) Medial clusters are split up into a possible word-final cluster plus a possible word-initial cluster.

(3) Before syllables with full vowels, the syllable boundary goes as far to the left as permitted by the preceding principles, whereas before schwa-syllables it goes as far to the right as possible, with the exception that the syllable boundary goes between a sonorant consonant and a stop other than /g/.

2.2 Molbæk's critique and "Hjelmslev's law"

Quite recently, Peter Molbæk Hansen (1979) has critically examined the above principles and their empirical consequences in an interesting paper. He agrees with my principles concerning grammatical boundaries, which I shall therefore not consider any further here. But then he points rightly to the following difficulty with regard to my principle (2) applying "Hjelmslev's law": I consider this principle to mean that split up medial clusters should not be in conflict with any general phonotactic restrictions of final, respectively initial, consonant clusters. I have stated these restrictions myself, however, on a phonological level which is more concrete (closer to phonetics, if you like), than the phonological rules which presuppose syllabification, in particular on a level at which consonant gradation has already applied (Basbøll 1973a). I should like to emphasize (and Molbæk has not claimed otherwise) that there is nothing circular in such a procedure, of course: the restrictions can be defined, in a non-circular fashion, to apply at a more abstract level; but the principles will then be less general, and what is worse, they will lose their otherwise convincing phonetic motivation (which lies in the sonority hierarchy).

Of course, Molbæk's criticism of my use of "Hjelmslev's law"[1] also applies to other approaches to phonological syllabification which make crucial use of the notions "possible initial cluster" and "possible final cluster", e.g. Anderson and Jones (1974). I should like to make clear (cf. Molbæk 1979, p. 96) that Kahn's strategy (e.g. 1976, p. 22) treats "permissible initial/final cluster" as an important theoretical primitive, or at least as a notion that is crucially presupposed by the syllabification within his framework. This is one of the facts about Kahn which make me feel somewhat uneasy about Molbæk's classification of him within

---

1) I should add that I entirely agree when Molbæk points to the important distinction between strong and weak syllables with respect to phonotactics (1979, p. 98). Although I have criticized Haugen's (1956) definition of the syllable and other definitions in Fischer-Jørgensen (1952) using just this argument (1974, p. 94-95), I must admit that my early treatment of these matters (1972, p. 194) is in fact objectionable in exactly the way Molbæk says (ibid.).

"natural generative phonology", a trend which (in Molbæk's words) should have syllable boundaries inserted "according to phonetically and typologically based hypotheses of natural syllabification of sequences of segments, leaving relatively little room for language specific deviations" (ibid.). I also feel that Molbæk's characterization of my own use of syllable boundaries as reminiscent of e.g. their use within glossematics, "in that relatively large freedom is allowed in connection with the placement of syllable boundaries" (ibid.), can easily be misunderstood by others, so I want to make the following point clear: One basic idea of my procedure is that syllable boundaries are inserted by rule; there is thus <u>no</u> freedom once the rules have been settled. This procedure differs markedly from that of Hjelmslev (1951), who not only provided no rules, but even located the syllable boundaries in places which were crucially different in words of exactly the same phonological structure, just to account for the distinction in manifestation (cf. Basbøll 1971, p. 207-211).

In fact, I have stated explicitly (e.g. 1974, p. 83) that this Hjelmslevian principle (for lack of a better term) might well be dispensable in the final analysis, if we presuppose a certain elaboration of the third (and last) principle. In the DANFON-project we have operated with a set of syllabification rules in Danish which <u>only</u> depend on grammatical boundaries and on the sequence of segments. This set of rules has been used since 1975 with only minor modifications as far as syllabification principles are concerned. The system has given rise to very few ill-formed constructions where the placement of the syllable boundary is a cause of the failure, and it may thus be taken to represent at least some degree of observational adequacy. It must be emphasized, however, that the project has, for various reasons, not yet been concluded, and that there has been no really systematic testing of the syllabification rules in particular (cf. note 2, p. 269). The syllabification rules in DANFON have never been published, so Molbæk can of course not be blamed for not having considered them. I shall, however, briefly present their contents here (as of late 1975), in section 3.1 below, so that they can be used in the evaluation of Molbæk Hansen's counter-examples to my proposals (they remain, of course, counter-examples to the older and published proposals).

2.3 Phonological syllable boundaries as descriptive devices

Before we turn to the specific rules with their examples and possible counter-examples, two remarks of a preliminary nature may be in order. First and most important, the phonological syllable boundaries, as I have used them, are descriptive devices to account for[1] a number of phonological phenomena. They are subject to certain general restrictions on any type of syllable, e.g. that they conform to some sort of sonority hierarchy, and that a stressed syllable should be a possible word (disregarding prosody, e.g. the stød); but they are not claimed to have any phonetically or psychologically demonstrable existence at all.

Phonetic syllable boundaries should be phonetically demonstrable, on the other hand, e.g. in terms of duration of initial vs. final allophones. I suggested (1974, p. 72) that the universally unmarked way to syllabify a given sound chain is the phonetic (as opposed to phonological, viz. as concrete vs. abstract) syllabification, which also depends, of course, on language-specific (abstract-)phonological factors. Quite naturally, phonological syllable boundaries most often coincide with phonetic syllable boundaries, e.g. in the German example ein Esel (cf. Basbøll 1974, p. 74), where I consider the prevocalic glottal attacks to be manifestations of both phonetic and phonological syllable boundaries. That the different types of syllable boundaries within my framework coincide, is the unmarked (or natural) case. I think the apparent disagreement between Molbæk (1979, p. 95) and myself on this issue is purely terminological.

Secondly, it follows from this conception of phonological syllabification that such boundaries may be partly indeterminate. For example, when the only distinction that matters phonologically is one between open and closed syllables, then it will, of course,

---

1) Of course, such an account in no way qualifies as an explanation - which must ipso facto involve (well-known) explanantia external to the explicandum in order to avoid circularity - but is an instance of a scientific generalization (hopefully a linguistically significant one). The latter point presupposes, naturally, that several different phonological phenomena are captured under one description, which is the case here. Notice that within this conception of reality, there is no reason to prefer /sal$mə/ to /salm$ə/ - or the other way round - except for what Molbæk Hansen calls "economy of formulation" (1979, p. 100).

be quite empty to insist on one unique inter-segmental boundary in all cases when its placement cannot be tested phonetically or psychologically, and when different placement of the syllable boundary has no phonological consequences.

What is essential in an evaluation of my proposed principles of phonological syllabification in Danish is, in my opinion, the following: Can a set of (preferably not too unnatural) syllabification rules be given from which one can derive - from grammatical boundaries and the sequences of segments, with or without the inclusion of prosodic information - in a non-circular manner, the correct output forms as far as syllable-dependent phenomena like short vowel adjustment and consonant gradation are concerned? Or, more specifically: can the correct allophones [o, ɔ], [a, ɑ], [d, ð], [g, ɣ], and so on be predicted from underlying forms with invariant /o, a, d, g/ etc., by means of automatically inserted syllable boundaries and some simple syllable-dependent rules?[1] The tentative answer delivered by the DANFON-project is in the affirmative, but it should be borne in mind that so far only rather limited sets of data have been tested.

3. Molbæk's counter-examples and the DANFON syllabification principles

3.1 The net effect of the DANFON syllabification principles

The syllabification principles of DANFON are approximately as follows ($ indicates a syllable boundary). These principles (4-10) replace 1-3 above (2.1). Remember that this is only an informal statement of the net effect of the syllabification rules taken together.[2] (5) through (10) only apply if no syllable boundary has been placed in accordance with (4).

---

1) These manifestation principles may be stated informally like this: /o/ → [ɔ] in a closed syllable; /a/ → [+grave] (i.e. [ɑ]) before a tautosyllabic grave consonant; /d,g/ → [ð,ɣ] in the final part of the syllable (the last rule is somewhat simplified).

2) It should be emphasized that the rules (4) through (10) below cannot be determined from the program of the DANFON-project, but that I have considered their over-all effect to be a reasonable approximation to the over-all effect of the DANFON syllabi-

(4) Certain grammatical boundaries are syllable boundaries too (unchanged, see (1) above): e.g. mados, dumhed (the examples are rendered in the orthographic form, except for $).

(5) A single intervocalic consonant belongs with the preceding vowel if the following vowel is schwa, otherwise with the following vowel (unchanged): e.g. bade, nota, O$da.

(6) $ goes immediately to the right of /g/ preceded by a voiced consonant if the vowel of the following syllable is schwa (that /g/ is the weakest plosive is in agreement with the hierarchies discussed e.g. in Foley 1977): e.g. alg$e, ærg$re.[1]

(7) $ goes immediately to the left of a plosive followed by a voiced continuant followed by a full vowel (exception: /tl/ (and possibly /dl/ too) is hetero-syllabic): e.g. hydra, Nigra.

(8) /s/ plus a plosive function as a plosive (different from /g/, see point (6) above), cf. (7): e.g. bi$skop, ek$stra.

(9) $ goes between a two consonant group and a nasal: e.g. ast$ma, øks$ne.

(10) Otherwise the syllabification of a medial consonant cluster is the unmarked one, namely as equal as possible, but with a preference to the left in the case of an odd number of consonants, informally speaking, viz. CC, CCC, CC$CC: e.g. sal$me, æn$dre, fæng$sle, tundra, alfa.

Notice that the principles do not appear completely ad hoc: the clusters plosive plus liquid (or glides, etc.) and /s/ plus plosive are well attested as close-knit units in several other languages, and so is the hetero-syllabicity of /tl/ (and possibly

---

2) (cont.) fication principles. It is clear that a reliable evaluation of those rules can only be given after a more definitive report on the DANFON-rules has been presented. Unfortunately, Kjeld Kristensen and I have not been able to work on the DANFON-project for the purposes of the present paper. I should also like to mention that a late DANFON-rule optionally deaspirates /p t k/ before schwa. The variables entering into such an optional rule are not encoded within the DANFON-project (it is the other way round: such variable rules should be investigated departing from the output of the DANFON-project, see Basbøll and Kristensen 1975).

1) Molbæk Hansen (1979, p. 101) apparently finds it in some way objectionable that I have treated otherwise similar clusters containing /g/ and /k/ according to different criteria. I fail to see why, given the unique behaviour of /g/ (as opposed to the other plosives) in this respect.

/dl/ too). There is also a high degree of similarity between the principles accounting for groups of one, two, three and four intervocalic consonants. Finally, the significance of the distinction between full vowels and schwa should come as no surprise either.

3.2 Molbæk's counter-examples to the earlier syllabification principles, evaluated with respect to the DANFON-version

Molbæk Hansen (1979) classifies his examples (which are all problematic with respect to my tentative rules corresponding to (1) - (3) in section 2.1 above) into seven groups which will be exemplified below:

I. Cases like manøvrere, aula, which are correct according to the DANFON-rules.

II. Cases like Sigrid, Børglum, prognose, of which the names Børglum, Sigrid, Sigvald, and Sigvard, where /g/ is not pronounced as a plosive, are the only exceptions to the DANFON-rules. The latter two examples can be correctly generated by restricting principle (7) to apply only to plosive-liquid-sequences (and possibly plosive-glide sequences as well), not to all clusters of a plosive followed by a voiced continuant (cf. Muta cum Liquida as a category in other languages); or the last three of them by recognizing a strong word-internal grammatical boundary between Sig- and -rid, -vard, -vald (cf. Ingrid, Edvard, Thorvald, and other names mentioned by Molbæk Hansen). The latter alternative is of the hocus-pocus-type (see below), and would be even more so if a strong word-internal boundary was postulated also in a case like Børglum (between Børg- and -lum).

III. Examples like adjudant, advokat, Gudrun (with /d/ pronounced [ð]) are counter-examples to both sets of my proposed syllabification principles. According to the first alternative just mentioned (under II), the cluster /dv/ (and possibly /dj/ too, cf. above) will be split up, correctly; and according to the second alternative there, Gudrun will be accounted for by means of a grammatical boundary (cf. Gudmund and other such names). Notice, however, that this account in terms of boundaries is only to be considered a lexical shorthand device, so to speak, since this boundary is of course semantically quite unpredictable in such proper names.

IV. Cases like pingvin, jonglør, Ingrid, which when pronounced without [g] are exceptional also with regard to the DANFON-rules. These forms would be correctly generated if the velar nasal were taken as one underlying segment. One can find independent arguments in favour of this ("concrete-phonological") solution, but the adoption of it would, admittedly, diminish the number of cases accounted for by my syllabification rules a little. I shall not pursue this issue any further here, and the forms must at present be considered exceptions, although only marginal ones. Another problem with /g/-words is a name like Hauge, which was formerly correctly generated with a pronounced stop, but which is malgenerated by the DANFON-rules. These may be changed without complications, as far as I can see, so that /g/ follows the unmarked C$C-pattern between /j, v/ and schwa (such forms are only rare names and the like, where one is on uncertain ground anyhow).

The preceding groups of problematic cases seem to me to be rather marginal when related to the DANFON-rules, which also by and large agree with the result of Molbæk Hansen's useful overview of certain $V_1C_1C_2V_2$-sequences. Some cases in the remaining three groups appear to be less marginal, however.

V. Cases like gamma, Abba, Bacchus, gummi, Gunna, in which underlying /a, o/ are pronounced as if the following consonant closes the preceding syllable, although it is followed by a full vowel.[1] Incidentally, as Molbæk remarks, I in fact mentioned the vacillating pronunciation [a/ɑ] in words like papir, akademiker (1974, p. 67), where the /p/ or /k/ which follows /a/ is aspirated and thus clearly syllable-initial. I suggested that this might be the symptom of a phonological change in progress whereby the rule assimilating an /a/ to a following grave consonant enlarges its domain (it is, for instance, my impression that older people, "socially higher" people and people from Jutland use the [a]-pronunciations in such words more than other people, provided that

---

1) Gummi and Gunna are Molbæk's only examples with /o/; the relation between short [o], [ɔ] and [ʌ] is extremely complex when viewed in its entirety, and I shall not discuss these examples further here, but only refer to Brink and Lund (1975, p. 180-183). These authors recognize that phenomenological syllable boundaries ("oplevede stavelsesgrænser") play a role for the distribution of [o]: [ɔ] (their [å]) essentially like the one I have proposed since (1972), but without any reference to my work.

they have the syllable-dependent [a/ɑ]-alternation in other words).

When we consider this possibility of a phonological change in progress (the question of the nature of the change will be taken up below), it seems to me that there are a couple of important observations to be made with regard to the examples presented by Molbæk Hansen (1979, p. 109) and other similar examples:

(1) The preceding consonant: Most of Molbæk's problematic examples, i.e. those with obligatory or optional [ɑ] before a grave consonant followed by a full vowel, have a grave consonant before them, and there is only one example (the place name Malacca) which is preceded by an acute consonant.[1] The examples of vacillating pronunciation adduced by Brink and Lund (1975, p. 71-73) seem to agree well with my analysis.[2] (Since /a/ adjacent to /r/ is always pronounced [ɑ] (or [ɑ]) in the varieties of Standard Danish considered here, examples with /ar/ or /ra/ have been disregarded, of course, in the present context.) I conclude that the [ɑ]-pronunciation is favoured by a preceding grave consonant and impeded by a preceding acute consonant.[3]

As to other examples of a preceding consonant influencing vowel quality, consider r-colouring in Danish as well as many instances of vowel nasalization (where the influence of a preceding nasal consonant is clearly inferior to that of a following tautosyllabic one, but nevertheless not quite negligible.

---

1) The name Jakob is not a real counter-example, since the general manifestation of the first part of the /aj/-diphthong (as [ɑ] or something intermediate between [a] and [ɑ]) shows that synchronically, at least, /j/ functions differently from acute true consonants with respect to this rule. Diachronically, however, /j/ was on a par with the other acute non-syllabics when the change started, cf. Brink and Lund (1975, p. 67).

2) Brink and Lund do not consider this question, but it turns out that only a couple of the roughly thirty relevant examples which they give of the vacillating pronunciation have an acute consonant before the /a/ (the exact figures depend on how you count and would be insignificant anyhow).

3) Cf. the fact that none of the examples of vacillating pronunciation given by Molbæk Hansen have any consonant before the /a/ (words like akademi(ker) belong here too). This generalization does not hold when further material is adduced, however, but the order grave cons. - zero/h - acute cons. still seems to obtain.

(2) Stress: Only 2 out of Molbæk's 16 examples with obligatory [ɑ] (viz. Hammurabi and akkurat[1]) have unstressed /a/, whereas both examples with vacillation [a]/[ɑ] have unstressed /a/. Brink and Lund (1975, p. 73) give the rules that [ɑ] is more frequent in stressed than in unstressed position, and that [ɑ] is clearly more frequent in the first of several pretonic syllables than in the only pretonic syllable. All this strongly suggests the following principle: the more prominent the /a/-syllable with respect to the following syllable, the more likely /a/ is to be influenced by a following grave consonant. If, as seems intuitively evident, a schwa-syllable is considered to be minimally prominent, the distinction between full vowels and schwa with respect to syllabification appears to be just a special case of this general tendency.[2]

Synchronically, the following three aspects of a phonological rule can be distinguished: (a) the contents of the rule (traditionally in terms of Structural Description and Structural Change, but other (alternative or additional) structurings are certainly possible[3]), (b) its domain of application (like syllable, word, and so on, see Basbøll 1978a), and (c) its mode of application (in terms of obligatory vs. variable rules, cf. Labov 1970). These three aspects of the rule of /a/-manifestation will be briefly considered in turn.

(a) Contents of the rule: (1) above in this section seems to agree very well with my statement (1974, p. 66) that the rule is an auditory assimilation rule. That the acute vowel is considered to be synchronically basic is due to arguments of formal

---

1) Molbæk (1979, p. 110) gives the pronunciation [ɑkuˈʁɑˀd], but in fact pronunciations with stress on the first syllable of this word are also frequently heard, which makes the tendency even clearer.

2) This by no means implies that the syllabification effect of the distinction between full vowels and schwa can be predicted from the general tendency: The importance of this vowel distinction for syllabification, as compared to e.g. different degrees of stress, is an interesting fact about Danish which had to be discovered (and its discovery was gradual: cf. Martinet (1937), Andersen (1954), Rischel (1970) and Basbøll (1972)).

3) One could consider such questions as which segment is affected, which direction does the change take, and so on. The important question of a typology of phonological rules also belongs here, at least in part (see, e.g., Linell 1977 and Dressler 1980).

simplicity, but I think it is in accordance with the intuition of most present-day speakers of Standard Danish (cf. the fact that /a/ is manifested [a] before zero, i.e. in the neutral context). Diachronically, the grave vowel changed into an acute one in more and more contexts (see Brink and Lund (1975, p. 67-96)); this development has been explained by Davidsen-Nielsen and Ørum in terms of the acoustic-auditory feature 'gravity' (1978), and further discussed in such terms by Brink and Lund (1975, p. 81). I shall return briefly to the diachronic problems below.

(b) Domain of application: It still seems to me that syllable boundaries play a decisive role for the manifestation of short /a/, cf. the detailed discussion by Brink and Lund (1975, p. 71-73 and 730-734).[1] Most of Molbæk's counter-examples with stressed /a/ (like gamma and other examples with 'non-grading' consonants) may be reconcilable with a syllabic analysis, presupposing that the syllable boundary occurs to the right of the consonant. In cases where the consonant, obligatorily or optionally has its "initial" manifestation (like kappa, Bacchus, etc.), this analysis meets with difficulties. And forms like papir, fakultet - pronounced with [α] - where pretonic /a/ as a rule is followed by aspirated /p/ or /k/, apparently cannot be analysed in these terms at all. In some sense, the domain of the /a/-rule for such forms seems to be larger than it is for otherwise similar forms pronounced with [a], regardless of whether this is accounted for in terms of a difference in the location of the syllable boundary (so that it is intra-segmental in the forms pronounced with [α], see section 4 below), or in terms of a rule domain larger than the syllable. Now the important point is that pronunciations of these and similar words with [α] seem to be more recent than those with [a],[2] according-

---

1) As in the case of the auditory nature of the rule, Brink and Lund make no reference to my proposals concerning the relation between the /a/-manifestation rule and syllable boundaries.

2) When Molbæk (1979, p. 110) expresses his scepticism as to whether the pronunciation [pαˈpiɐ̯ʼ] "is a new phenomenon (to the extent that it occurs)", it should be said, first, that it occurs without any doubt (cf. Brink and Lund (1975, p. 72-73)), and, second, that the [α]-pronunciation, of course, is an old phenomenon in the sense that it is attested long before the [a]-pronunciation (viz. before the fronting of [α] started), but that my claim concerns something else, namely that the [α]-pronunciation has reappeared (at least as a possibility) after an interval of "pure" [a]-pronunciation, presupposed, of course, that the sociolinguistic variables are kept constant (this, however, I cannot prove).

ing to my impression (unfortunately, this question has never been investigated, and no conclusions can be drawn from the material presented by Brink and Lund (ibid.)). If this is so, the /a/-assimilation rule seems to be in the process of enlarging its domain in one of the senses just hinted at: either so that the syllable boundary seems to be located intra-segmentally in more and more cases like those just mentioned, or so that the blocking effect of syllable boundaries with respect to this rule seems to be diminishing.

(c) Mode of application: It follows from what has already been said that the rule is variable in the sociolinguistic sense of Labov (1970), both with respect to different speakers (classified according to sociological, geographical and chronological criteria) and with respect to phonology (cf. (1) and (2) above in this section). Within the syllable, the rule is obligatory (for those speakers of Standard Danish considered here), and it never applies across word boundaries; in between those two domains, it is variable (but cf. section 4 below).

What I conclude from all this, although quite tentatively, is the following diachronic picture: The original [ɑ] gradually was replaced by [a] in more and more contexts, with the proviso that a following tautosyllabic grave consonant impeded the change.[1] Towards the end of this process, a rule accounting for alternations and vacillations of /a/ would synchronically treat [a] as basic (since /a/ is manifested as [a] in the neutral context, viz. before zero). This latter rule (which essentially assimilates an /a/, with respect to the feature "gravity", to a following tautosyllabic grave consonant) then is applied in more and more contexts, [ɑ]-pronunciation being favoured by higher relative prominence of the syllable in question as compared to the following syllable, and favoured, respectively impeded, by a preceding grave, respectively acute, consonant. According to the present account, the expansion of [ɑ]-pronunciations in such words would thus be a symptom of a (variable, in Labov's (1970) sense) enlargement of the domain of the rule of auditory /a/-assimilation in one of the two respects

---

1) For certain speakers, the labials clearly were not impeding the change in the way velars were, cf. Brink and Lund (1975, p. 67 and 71).

just mentioned. All this needs further investigation, of course. The same applies to the influence of spelling: it is the impression of both Molbæk (personal communication) and myself that double ("grave") consonants in the orthography favour [ɑ]-pronunciation of a preceding short /a/. There are a number of methodological difficulties in investigating the character of this influence, however, and this issue will not be pursued any further here.

VI. Cases like Harry, paritet, terracotta, Karoline, in which /r/ is realised as a glide before an unstressed full vowel. The conclusion suggests itself that the realisation of /r/ is not always syllable-dependent (cf. Basbøll 1972, p. 196).

VII. Cases like Canada, Paludan, which have stød on a sonorant consonant followed by a weakly-stressed full vowel. Although there is no descriptive problem in first assigning stød to syllable peaks and then having the stød spelled out, late in the derivation, on a consonant which phonologically, at earlier stages of the derivation, belonged to the following syllable, Molbæk is certainly right that this description is at odds with my basic conception of stød as a syllabic prosody. Although I could still say that the stød-consonant occurs in the same phonetic syllable as the preceding (stressed) vowel, the description seems unsatisfactory.

The cases mentioned under VII, some of those under V, and possibly those under VI, suggest to me that in certain cases, a consonant occurring between a fully stressed short vowel and a weakly stressed short full vowel, may seem to close the preceding syllable (cf. Basbøll 1974, p. 88), see further below.[1]

4. Concluding remarks

Molbæk Hansen concludes with the suggestion (1979, p. 118) that an additional hierarchical unit in between the syllable and the word might be phonologically relevant. He gives it no name,

---

1) In the syllabification rules given in (1973b, p. 25), I treat certain instances of short /e/ and /i/ (viz. the vowel of the endings -ing, -ig, and certain -isk) on a par with schwa, cf. the fact that Martinet operates with "i de très faible intensité" as a phonological entity which conditions (just like schwa) a neutralization of the aspiration correlation in the preceding consonant (1937, § 3-5).

and does not refer to recent versions of hierarchical phonology, but he evidently has in mind some sort of <u>foot</u>, consisting of one salient syllable followed by zero, one or more subordinate syllables with a limited vowel repertoire.[1] The foot would be "internally consolidated by certain obligatory structural properties: /a/- and /o/-adjustment, the restricted occurrence of medial aspirated stops before sonorants,[2] the occurrence of at most one stød, and probably some more" (ibid.). He suggests that e.g. the pronunciations [ɑgvaˈvid/akvaˈvid] are due to different foot-formation: <u>akva-vit</u> vs. <u>a-kva-vit</u>, and similarly [sygloˈtʁoˀn/sykloˈtʁoˀn]: <u>cyklo-tron</u> vs. <u>cy-klo-tron</u>,[3] but he of course realizes that this would mean the introduction of a new unpredictable structure. I find this structural addition empirically ill-supported by the type of examples he gives. Notice that only a very small and specific part of the consonant gradation-phenomena can be accounted for in terms of feet, that different placement of the syllable boundary in cases like this will have the same effect as different foot-structure, and that the stød-restrictions offer no real arguments for the foot, either (since weak syllables generally do not have stød).[4]

One of the more challenging consequences of Molbæk Hansen's competent discussion is that it brings into the open certain incongruencies (within my framework) between 1) the concept of the syllable which is decisive for the manifestation of /a, o/, and

---

1) According to Molbæk (1979, p. 118), it should be "schwa or one of the full vowels /a o i y u/ but not /e ɛ ø æ ɔ/". This set must be erroneous anyhow, as shown by words like <u>Ammon</u> ([ɑmʌn], also cf. <u>madding</u> [maðeŋ], unless this ending is posited with an underlying schwa), but I shall not go into that problem here.

2) This restriction is not quite as strict as Molbæk seems to think (1979, p. 113), e.g. <u>cyklus</u> has an aspirated /k/ in my normal pronunciation. (I disagree with some other pronunciations given by Molbæk, but this is not important here.)

3) If such different pronunciations are tonally distinct, an important independent argument for foot-structure might be established from such a distinction (this point was suggested by Jan Katlev at the conference).

4) It should be observed, however, that stød-words of the type <u>Canada</u>, mentioned under VII in section 3.2 above, are unproblematic within Molbæk's account, whereas they seem to presuppose that the syllable boundary does not occur before /n/ within my framework (which is quite acceptable to me).

2) that which defines the domain of consonant gradation intervocalically. Since not the foot in Molbæk's sense, but something like the syllable in my sense, is decisive for the manifestation of a single intervocalic /d/ or /g/, one can in fact construct a better case for the foot, or at least a better counter-case to my analysis, than the one presented by Molbæk. Consider the following (constructed) examples (where all vowels are short):

(1) /ádə/, pronounced [áðə] (e.g. in (sn)adde)
(2) /ága/, pronounced [ɑ́ga] (e.g. spelled Agga)
(3) /agá/, pronounced [agá] (e.g. in (prop)aga(nda))

Molbæk would ascribe the following dual structure to these examples, if I have understood him correctly:

(1) syllables: \$ád\$ə\$, foot: -ádə-
(2) syllables: \$á\$ga\$, foot: -ága-
(3) syllables: \$a\$gá\$, feet: -a-gá-

This seems rather straightforward (presupposed that the manifestation of /d,g/ is determined with the syllable as its domain, and the manifestation of /a/ with the foot as its domain).

In my analysis, on the other hand, there would be trouble in ascribing a syllabic structure to (2), since consonant gradation would seem to presuppose \$ to the left of /g/, and /a/-manifestation to the right. In agreement with my strategy as applied to French (cf. Basbøll 1978b and section 2.3 above), I could define the notions "open and closed syllable" so that /g/ closes the preceding syllable in (2), due to the prominence relation of the vowels. The rules of /a/- (and /o/-) manifestation thus, naturally, would obtain in closed but not in open syllables. The consonant manifestation would be "initial", so to speak, due to the following full vowel. This proposal could be rendered in syllabic notation somewhat like this:[1]

(1) \$ád\$ə\$, pronounced [áðə]
(2) \$á$\overset{\$}{g}$a\$, pronounced [ɑ́ga]
(3) \$a\$gá\$, pronounced [agá]

---

1) The principles of syllabification lying behind this notational proposal might be rendered something like this: <u>In the case of a single intervocalic consonant, \$ occurs to the right of the consonant if and only if the following vowel is schwa.</u> Concerning the difference between intra-consonantal and pre-consonantal location of the syllable boundary, only some variable rules not reach-

This is, of course, a type of ambisyllabicity proposal (cf. Anderson and Jones (1974) and Kahn (1976)), which I should like to rephrase as follows: the intervocalic consonant at the same time closes the preceding syllable and begins the next one[1] (cf. the notion "close contact"). Viewed in this light, the difference between the treatment suggested here and Molbæk's account in terms of foot-structure, is perhaps not essential.[2] In addition to a possible (but by no means forcible) methodological reason for preferring my own account to Molbæk's (parsimony of levels), I want to briefly point out how the two sketchy proposals would account for a few complicated cases.

Consider pronunciations like agent, papir [ɑˈgɛnʼt, pɑˈpiɒ̯ʼ] (cf. note 1, p. 279), which seem to presuppose that /a/-adjustment optionally may apply across foot-boundaries (but still within words only). This optionality across foot-boundaries at the same time accounts for the pronunciation [ɑkvaˈvit], which is not mentioned by Molbæk Hansen (1979, p. 118): a-kva-vit. But now consider forms like kappa, Bacchus which have obligatory [ɑ], but where the stop may be pronounced with or without aspiration. Those words would consist of just one foot, according to Molbæk's analysis; and in order to account for the optional lack of aspiration,

---

1) (cont.) ing the level of complete predictability can be given, viz. that the syllable boundary most often occurs to the left of the consonant if the first vowel is less prominent than the second, and within the consonant if both vowels are full, and the first vowel is more prominent than the second. These variable rules should, of course, be made much more precise. The formulation is deliberately vague regarding a sequence of equally prominent full vowels, which is exactly the case where most [a/ɑ] vacillation occurs, but where the manifestation of the consonant is "initial" as a rule, e.g. in words like Agamemnon, fakultet, and so on. (The principles of syllabification just stated only apply within word boundaries, of course, in the usual fashion.)

1) It is by no means surprising that the initial manifestation of the consonant overrides the final one, so to speak, since initial is in many respects the stronger of these two positions.

2) Cf. Kiparsky's claim (forthcoming) that phonological phenomena which were earlier considered (in particular by Kahn (1976)) to be arguments for ambisyllabicity, can in general be accounted for as foot-bound phenomena. I do not subscribe to all of Kiparsky's claims concerning the foot, however (cf. Basbøll (forthcoming)).

an optional rule of de-aspiration must be postulated to apply within the foot.[1] Within Molbæk's analysis, the optional lack of aspiration of /k/ in Bacchus and akvavit is thus due to two unrelated structural properties, viz. the optional rule of de-aspiration within the foot in the former case, and the structural ambiguity between a two-feet and a three-feet analysis combined with "the restricted occurrence of medial aspirated stops before sonorants" (ibid.) within the foot, in the latter case. Whether this structural complexity can be substantiated by any independent evidence still remains to be shown.

Within my proposal, the three possible pronunciations of akvavit may be accounted for as follows: [akvaˈvit]: akvavit; [ɑgvaˈvit]: akvavit; [ɑkvaˈvit]: a$\overset{\$}{k}vavit. Notice that the pronunciation [agvaˈvit] is excluded within this notational system, as desired. It is my impression, however, that the three pronunciations given are very different with regard to distinctness: the [g]-form is clearly less distinct than the two others.[2] In view of this, I would prefer to limit the freedom of syllable boundary location in cases like the one at hand to $a$kvavit vs. a\overset{\$}{k}vavit$ (pronounced [akvaˈvit, ɑkvaˈvit], respectively), and to account for the pronunciation [ɑgvaˈvit] by means of an optional rule of de-aspiration applying to ambisyllabic /p, t, k/ before weak syllables (it is no surprise that ambisyllabic stops are more liable to de-aspiration than initial stops, cf. the fact that syllable-final stops are unaspirated in Danish except before pause; also cf. note 2, p. 280). This rule accounts at the same time for the optional de-aspiration in words like kappa and Bacchus.[3]

---

1) Notice that /d, g/ are not optionally pronounced as continuants in this position, which shows that the often claimed parallelism between /t, k/ and /d, g/ as instances of a common process of "weakening" in certain positions ('consonant gradation') is not complete.

2) I am not in a position to wholly exclude the possibility that even the pronunciation [agvaˈvit] can in fact be heard, as a very indistinct form, but I very much doubt that it will ever be encoded in serious communication, in contradistinction to [ɑgvaˈvit]. Investigations of such matters would be welcome.

3) Whether the normal pronunciation of apotek: [ɑboˈteˀk] can be accounted for in this way, or whether it has an underlying /b/, must be left entirely open here. The normal pronunciation of chokolade: [ʃogoˈlæːðə], in addition to [ʃɔgoˈlæːðə], might be interpreted as an instance of lexical restructuring (from /k/ to /g/), but other interpretations are possible too. All this is nothing but speculation, of course.

The sketchiness of this suggested proposal, just like Molbæk's on the importance of a unit like the foot, can hardly be overestimated. But it has at least become clear, I think, that Molbæk's detailed criticism of my own work on phonological syllabification as applied to Danish has been highly stimulating. I hope work in this area will be continued in the same spirit.

References

Anderson, J.M. and C. Jones 1974: "Three theses concerning phonological representations", JL 10, p. 1-26

Basbøll, H. 1971: "A commentary on Hjelmslev's Outline of the Danish Expression System", ALH 13, p. 173-211

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: Konsonanter I+II, Mimeographed notes

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

Basbøll, H. 1978a: "On the use of "domains" in phonology", Proc.Ling. 12, p. 763-766

Basbøll, H. 1978b: "Schwa, jonctures et syllabification dans les représentations phonologiques du français", ALH 16, p. 147-182

Basbøll, H. (forthcoming): "Metrical theory and the French foot", in: W.U. Dressler et al. (eds.), Phonologica 1980

Basbøll, H. and K. Kristensen 1975: "Further work on computer testing of a generative phonology of Danish", ARIPUC 9, p. 265-291

Brink, L. and J. Lund 1975: Dansk rigsmål. Lydudviklingen siden 1840 med særligt henblik på sociolekterne i København, vol. 1-2, Gyldendal, Copenhagen

Davidsen-Nielsen, N. and H. Ørum 1978: "The feature "gravity" in Old English and Danish phonology", ALH 16, p. 201-213

Dressler, W. U. 1980: Morphonology, Karoma Press

Fischer-Jørgensen, E. 1952: "On the definition of phoneme categories on a distributional basis", AL 7, p. 8-39

Foley, J. 1977: Foundations of theoretical phonology, Cambridge University Press, Cambridge

Haugen, E. 1956: "The syllable in linguistic description", FRJ, p. 213-221

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 [English translation: "Outline of the Danish expression system with special reference to the stød", in: Essais linguistiques II = TCLC 14, p. 247-266]

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

Kiparsky, P. (forthcoming): "Main report on rhythm and metrics in phonology", in: W.U. Dressler et al. (eds.), Phonologica 1980

Labov, W. 1970: "The study of language in its social context", Studium Generale 23, p. 30-87

Linell, P. 1977: "Morphophonology as part of morphology", in: W.U. Dressler and O.E. Pfeiffer (eds.), Phonologica 1976, Innsbrucker Beiträge zur Sprachwissenschaft 19, p. 9-20

Martinet, A. 1937: "La phonologie du mot en danois", Bulletin de la société linguistique de Paris 38, p. 169-266

Molbæk Hansen, P. 1979: "A comment on Basbøll's phonological syllabification as applied to Danish", ARIPUC 13, p. 87-120

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

(ARIPUC 14, 1980)

FILTERING OF EMG SIGNALS

Jørgen Rischel[1] and Birgit Hutters

Abstract: One of the ways to eliminate different kinds of disturbances in EMG signals is by filtering in the frequency domain. This procedure requires both a knowledge of the frequency content of the disturbing components and a knowledge of the frequency range of the relevant signal. Therefore, the frequency characteristics of an EMG material of various internal laryngeal muscles recorded by bipolar hooked-wire electrodes have been examined. We conclude that the highpass cutoff frequency normally proposed in the literature - i.e. 100 Hz or lower - is generally not sufficient to eliminate microphony and especially the movement artifacts. But the choice of filtering frequency will often be a compromise between the cutoff that is optimal with respect to removal of spurious components and the attenuation of the overall signal caused by the filtering. However, it is generally advisable to highpass filter the signal even at the expense of some overall attenuation, and the greatest improvement is often achieved by highpass filtering with cutoff frequencies well above the low range normally proposed.

1. The background of the present study

Over the past several years a very considerable amount of EMG recording has been made at the Institute of Phonetics. This research has been largely concentrated around a project involving several researchers and comprising investigations of the functioning of various internal larynx muscles in speech, particularly with a view to the production of various types of consonants and the production of Danish stød and word stress. These recordings from larynx muscles were all made with bipolar hooked-wire electrodes. Dr. Hajime Hirose[2] performed the insertions of the electrodes. (Information about the project as a whole will appear in a later report.)

---

1) Institute of Linguistics, University of Copenhagen.
2) Institute of Logopedics and Phoniatrics, University of Tokyo, Japan.

The bulk of recordings made in 1974 exhibit occasional or, in some recordings, frequent or constant disturbances due to various factors external to the motor unit potentials being picked up. Typical disturbances encountered were: (I) hum, (II) other background noise of various kinds, (i.e. probably a summation of amplifier noise and background noise caused by fluctuations in the physiological properties of the tissue, the so-called "tissue noise", see Hayes 1960), (III) microphony (the EMG electrode functioning as a transducer picking up the periodic vibrations of voiced portions of the speech signal), and (IV) occasional aberrant spikes in the signal which are likely to be caused by movements in the tissue jerking an electrode back and forth, or even causing two electrodes to touch intermittently, i.e. the "movement artifacts" normally referred to in EMG work.

The identification of artifacts of type (IV) was in all cases done on the basis of the raw signal, which contains information both concerning the frequency content and the amplitude of the EMG signal. The main criterion for the identification of such artifacts is the presence of a dominant low frequency component at some point in the signal. Strong artifacts of this kind are normally visible also in integrated EMG curves, but for the identification of weaker artifacts these curves are not in themselves sufficient, and if one has access to integrated curves only, spurious energy due to artifacts may be erroneously identified as being a part of the useful signal, cf. Fig. 3a. - In most cases it does not seem to be very difficult to identify artifacts on the basis of their frequency content, but it must be admitted that we have no principled criterion for the decision.

All of the four factors mentioned above are very well known sources of trouble in EMG work. In the recordings made at the Institute, great care was taken to minimize such effects, and in fact the general quality of our recordings is high. In the design of the instrumentation a very low level of hum and noise was achieved, and thus the two factors ((I) and (II) above) are generally negligible. (They come to play a role only in cases of extremely low signal level, which is indicative of too large a distance from the motor units recorded. Such recordings are of dubious value anyway since, with the very small muscular structures of the larynx, the pickup area must be very close to some active muscle in order to achieve a signal that can be uniquely associated with that particular muscle.)

However, disturbances of types (III) and (IV) occur frequently enough to pose a problem in the processing and evaluation of the signals. As for "voice bar" microphony (III), which in the case of inner larynx muscles sometimes occurs with a magnitude comparable to a level of moderate EMG activity, the obvious way to eliminate this distortion is by filtering in the frequency domain. This procedure, of course, requires both a knowledge of the frequency content of the microphonic component, and a knowledge of the frequency range of the useful EMG signal, in order to see what kind of filtering (if any), can be applied successfully, i.e. without seriously degrading the useful signal.

As for (apparent or real) artifacts of type (IV), there are various possible solutions. The treatment of the problem partly depends on whether the appearance of the artifacts is "systematic" or random, and partly on where in the test sentence the artifacts appear. In the case of random appearance the averaged signal will be less dominated by such artifacts if the number of tokens included in the average is not too restricted - and for that purpose the median mean is preferable since it is less sensitive to random artifacts than the arithmetic mean. But if the appearance of artifacts is systematic, i.e., if it comes up several times at the same place in the test sentence, averaging will not help much. One may choose to disregard this kind of distortion if the artifacts occur outside the passage of primary importance. But whenever it appears essential to avoid this kind of averaged signals we have the possibility of eliminating artifacts in the time domain, i.e. by setting a window in the raw signal at the appropriate width (duration) and suppressing the part of the signal that is contained within the window.[1] In this case, too, information on the frequency characteristics of the various signal components is desirable, as a basis for the interpretation of a suspicious component as a relevant (though perhaps unexpected) signal spike, or as an artifact. But if a "systematic" artifact appears in the important part of the test sentence the tokens in question must either be eliminated - which in certain cases will result in a considerable reduction in the number of tokens to be averaged -

---

1) An EDP programme for this kind of surgery is available at the Institute.

or it must be attenuated as much as possible by filtering, as in the case of "voice bar" microphony. Again, some kind of information on the frequency content of the various kinds of signals is a prerequisite.

For the reasons stated above we found it essential to examine the frequency characteristics of the various EMG recordings. This was done mainly by highpass, lowpass and (to a more limited extent) bandpass filtering, the cutoff frequencies being varied stepwise within a frequency range chosen in accordance with the properties of each individual recording, and the effect on the signal shape and level was examined by visual inspection of mingograph tracings.

Since, in general, information on the frequency characteristics of EMG signals is very scanty, to say the least, the overall results of our filtering experiments may be of some interest to others, specifically as information on the properties of signals from very small muscles. We must, however, emphasize the limitation of the present study, viz. that we have confined ourselves to a qualitative appraisal of the effects of filtering (rather than performing quantitative measurements on frequency spectra) in accordance with our immediate goal: that of evaluating how much frequency limitation - especially in the form of highpass filtering - various kinds of EMG signals can stand, and how effective such filtering is in removing spurious components of the signal.

2. Previous research on the frequency characteristics of EMG signals and its relevance to the present study

The literature on EMG techniques and EMG research being by now extensive, we have found surprisingly little space devoted to questions having to do with the frequency characteristics of EMG signals and especially with the artifact problem. Several researchers touch upon the question of frequency filtering of EMG signals, however, and we refer to some of the more important statements below.

As for the possibility of improving the signals by filtering, Hayes (1960) finds that tissue noise and movement artifacts tend to be concentrated at low frequencies, and he recommends the use of band limited amplifiers with a lower cutoff frequency at about 20 Hz and a higher cutoff frequency in the range 100-200 Hz. Such band limiting also reduces amplifier noise, and it is applicable, according to Hayes, because the muscle action potentials exhibit sharply peaked spectra (he reports that biceps voltage is reduced by only about 3 dB when passing through a 35-75 Hz bandpass filter). Hayes' reasoning presupposes that movement artifacts etc. are indeed characterized by a preponderance of very low frequencies (ideally below the cutoff of some 20-30 Hz!). This is in disagreement with Geddes (1972), who briefly states that movement artifacts caused by disturbance of the electrical "double layer" are in the frequency range of many of the bioelectric events, and that for this reason "filtering techniques can seldom be employed with success" (p. 211).

Hayes' findings are of little direct relevance to our project since they are based mainly on EMG recordings from a large muscle by means of surface electrodes. Scott (1967), however, calculated power density spectra for signals measured with subcutaneous wire electrodes inserted in four large muscles, and he found "significant energy only in the frequency range from approx. 30 Hz to approx. 200 Hz" (p. 303), i.e. a result similar to that of Hayes as regards the frequency spectrum of the useful signal. He concludes that "methods of reducing mains frequency interference by using high pass filters are relatively inefficient" (p. 303), and also that "amplification of frequencies in excess of 200 Hz is likely to degrade the signal-to-noise ratio of a myo-electric system" (p. 303).

Trimble et al. (1973) performed a frequency analysis of single motor unit potentials recorded from a very small muscle (the medial rectus muscle of the eye) by means of a concentric needle electrode. They found energy at a much higher frequency range than found for other muscles by most other investigators, the predominant frequency band extending from 470 Hz to 1400 Hz (-3 dB points). They explain this surprisingly high frequency band by the very fact that a single motor unit was involved, whereas in other cases entire electromyograms have been analysed, which might be expected to reflect algebraic summations of asynchronous signals from different motor units. Such summation causes a lowering of the predominant frequency. With signals in which a single motor unit may predominate, the amplifier bandwidth must of course correspond to the resulting high frequency emphasis in the power spectrum for faithful reproduction of the single motor unit potential. This situation may well obtain in the case of EMG from larynx muscles.

If we look at the EMG processing employed for speech research a highpass filtering at 100 Hz or lower is normally proposed in order to eliminate movement artifacts and hum.

Fromkin (1965), in her EMG-investigation of the facial muscles using surface electrodes, states that the lower roll-off frequency of 40 Hz introduced by the tape recorder "gets rid of artifacts due to movements of the electrode" (p. 66). For the upper frequency limit she suggests a cutoff frequency at 10.000 Hz. Approximately the same bandwidth is proposed for research on muscles in the lips, tongue, and larynx by Hubler (1967), who mentions that "movement artefacts are minimized by rolling off the preamplifier at 40 Hz" (p. 27).

A higher HP cutoff frequency is proposed by Hirano et al. (1967) in their presentation of the hooked-wire equipment. In some of their recordings of various speech muscles they have noted the occurrence of low frequency movement artifacts, and according to the authors, recordings containing these artifacts can be highpass filtered at about 80 to 100 Hz, "which will completely eliminate the movement artifacts without diminishing the overall signal strength by too great a factor" (p. 23). The same highpass frequency is employed by Hirano and Smith (1967), using bipolar hooked-wire electrodes inserted in external tongue muscles, in order to eliminate movement artifacts. Furthermore, they find it necessary to lowpass filter the signal at 1500 Hz in order to eliminate some extraneous high-frequency noise, and they find no noticeable loss of "strength" due to this bandpass filtering. Also at the Haskins Laboratories a highpass filtering at 80 Hz is employed in order to reject movement artifacts and hum (Kewley-Port 1977).

In order to find out what the contributions of various frequency components to the total amplitude of the EMG signals are for typical speech gestures, Mansell (1979) has made a frequency band analysis of EMG signals provided via a surface monopolar derivation from the lower lip of one subject. The frequency bands exmined were 56-560 Hz, 560-1000 Hz, and four bands in the higher frequency range comprising frequencies from 1000 Hz to 4700 Hz. He finds that although the signal is attenuated when rejecting the components of the EMG signal below 560 Hz in favour of the frequency band ranging from 560 Hz to 1000 Hz, the signal-to-noise ratio is maintained because of attenuation of a band of EMG activity in the region below 560 Hz which appears to be relatively undifferentiated with respect o the segmental gestures of the EMG signal. This "cleaning" of the signal results in a small drop in peak variance and a greater distinction between linguistic types. Furthermore, Mansell finds a much better alignment of individual traces and a better differentiation of multi-modal gestures, exemplified by the two-phase nature of the average signal for the labial gesture in the German word Pute.

Various authors have approached the possibility of using the frequency spectrum (profile) in the diagnosis of various aspects of the behaviour of the muscles involved. Thus it has been found that the frequency spectrum of the signal changes with the level of contraction and with fatigue (Kaiser and Petersén 1965; Kadefors, Kaiser and Petersén 1968; Kwatny, Thomas and Kwatny 1970). It has been repeatedly observed that "when the levels of contraction rise, the motor unit potentials become synchronized, which results in augmented potential duration and amplitude" (Herberts 1969, p. 21). Such synchronization, if causing broader activity spikes, necessarily affects the power spectrum, with the effect of boosting the lower frequency region (cf. above). However, there is probably no simple, uniform statement applying to all muscles under all conditions. Thus, according to Kaiser and Petersén (1965, p. 234), the spectral change accompanying increased muscle contraction intensity may take different courses depending on the intensity level in question. Kadefors, Kaiser and Petersén (1968) decompose the spectral change under fatigue into a smooth increase in low-frequency content, and a rapid decrease in high-frequency content, and they suggest that "the dropping-out of small motor-unit potentials as fatigue develops might cause part of the high-frequency decay" (p. 71). Kwatny, Thomas and Kwatny (1970) find that "fatigue tends to place more emphasis on lower frequency components, indicated by a rise in the mean frequency of the spectrum" (p. 311). In their investigation the effect of fatigue on the EMG spectrum seems to override that of increasing contraction level.

These findings, which relate to sustained contraction of muscles of a much larger size than those of the larynx, are in themselves hardly applicable to the study of speech production. However, it is noteworthy that the spectrum of the EMG signal may vary under varying conditions of muscular behaviour, and it would seem advisable not to base an appraisal of the effect of filtering on a single type of speech gesture.

As for the question of large versus small muscles, Kadefors and Petersén (1970) report on earlier studies according to which the decrease in high-frequency activity accompanying fatigue is "considerably different in different muscles, minor muscles showing steeper decrease (and lower recovery rates) than major ones" (p. 46). In the study under consideration they consider the frequency spectrum of two sphincter muscles as well as a levator muscle, and they find that "the static frequency spectrum of EMG from the two sphincter muscles studied contains much more high-frequency energy than muscles of the extremities in general..." (p. 63). The sphincter muscles in question agree in this respect with another sphincter muscle: the orbicularis oris. The sphincter muscles also generally exhibit less change in the spectra of EMG under maximal voluntary effort than do muscles of the extremities previously studied (p. 64). The authors attribute the greater amount of high-frequency energy of such muscles "to short duration of the motor-unit potentials and to small size motor-units" (p. 66), and they also establish some association between more high-frequency emphasis and faster muscle contraction.

All of this strongly suggests that one should not rely on generalized statements concerning the useful frequency range of EMG signals, and that one should be particularly aware of the possibility of a preponderance of higher frequency energy in the case of small and fast acting muscles such as the larynx muscles, as compared to the much larger muscles which have been most studied by physiologists.

As stated by Abbs and Watkin (1976), there appears to be little agreement by electromyographers as to the frequency band necessary for EMG signal recordings. Needless to say, comparisons of findings concerning different muscles make sense only if the same setup and exactly the same electrodes are used. Ideally, even the number of motor units involved and their distances from the electrode should be similar. This makes it very difficult to arrive at valid generalizations beyond the particular muscles studied and the particular technique employed. Generally speaking, surface electrodes pick up from a more diffuse area and at a greater distance from the motor units under investigation than do needle or hooked-wire electrodes, and they are expected to yield signals with more low-frequency emphasis than the other types of electrodes. As the larynx muscle potentials are picked up by hooked-wire electrodes which are placed (hopefully) closely adjacent to the muscle fibres, the output is a priori likely to have a relatively high amount of high-frequency energy, although this would depend very much on the distance from a motor unit. The presence of strong high-frequency components in the signal makes the possibility of signal improvement by frequency filtering more promising than with many other types of EMG signals. At the same time there is a particular need for such signal processing because the electrodes are placed in a small-size region of which all parts move almost continuously, so that the likelihood of movement artifacts here is extremely high.

3. Filtering approach in the present project

In the filtering experiments various types of filters were employed. The results reported on here were based almost exclusively on two types, viz. active RC filters (FJ-Electronics) coupled as lowpass, highpass or bandpass filters, and a passive LC filter for highpass filtering only. The former type of filter had an attenuation slope of 36 dB/octave; the latter type had a much steeper roll-off, the attenuation at 1/2 octave above the cutoff frequency being at least some 50 dB. The range of available cutoff frequencies for the active filter covered the entire frequency range of possible interest, whereas the passive filter could be set at cutoff frequencies only from 200 Hz upwards. The use of one type or the other was not found to make any essential difference in the effect of highpass filtering on the EMG signals (strictly speaking, the results were not exactly comparable because of differences in the nominal cutoff frequencies for the two filter types, but this was of no consequence because of the rather narrow spacing of the available cutoff frequencies for each type).

The EMG recordings were played back on an FM multi-channel tape recorder (Lyrec), viz. the one on which they had been recorded. The filters were connected to the output of the tape recorder via suitable high quality pre-amplifiers set to a gain minimizing the background noise caused by the instrumentation. The outputs from the filters were rectified and smoothed and recorded on a mingograph (with or without interstage amplifiers). Whenever filtering attenuated the signal so much that its characteristics were no longer clearly discernible, the gain of the interstage amplifiers was raised by a suitable amount (the gain being varied in calibrated steps, so that the real changes in signal level as a consequence of filtering can be observed by comparison of the various mingograms).

The total range of cutoff frequencies used was from 22 Hz to 2000 Hz in the case of highpass, and from some 4700 Hz to 22 Hz in the case of lowpass filtering. Bandpass filtering[1], if performed at all, was made within the frequency region 22 Hz - 12000 Hz. (The bandwidth of the amplifying and recording equipment greatly exceeded the range within which filtering was performed.)

We did not, however, perform filtering at closely spaced cutoff frequencies across these total ranges for all subjects. The filter settings were often varied in large steps at first, i.e. a recording specimen was filtered at various rather widely spaced frequencies, closer spacings (or an extension of the frequency range of the filter settings) being used especially if the signal turned out to be clearly sensitive to filtering in a certain frequency region.

---

1) The signal was bandpass filtered by setting a highpass filter and a lowpass filter (coupled in cascade) at the same nominal cutoff frequency, so that the pass band consisted of only a narrow peak with attenuation slopes of 36 dB/octave on either side of the center frequency, the point in bandpass filtering being to locate particularly prominent frequency components of artifacts. In this setup, an additional 30 dB amplification was used.

By this more or less selective use of filtering, a considerable amount of informative data concerning a variety of EMG signals could be produced. However, even with this attempted minimization of redundant and excessively detailed information, the filtering was a laborious task. The frequency range from about 30 Hz to some 500 Hz, in particular, was important, and in virtually all cases most of this range required detailed analysis for a full description of the relationship between the frequency contents of artifacts versus useful signals.

4. Material

The EMG signals studied comprised specimens from recordings of 12 subjects: 8 males (PM, BM, BF, LG, JJ, NR, JR, JPF) and 4 females (MF, BH, HU, MW). One of these subjects was French (JPF), one was German (MW), the rest were Danish. All subjects spoke test sentences, i.e. frame sentences plus test words, in their own language. In this paper we have not taken language differences into consideration.

The recordings (made with bipolar hooked-wire electrodes, as mentioned earlier) were from the following muscles of the larynx: the cricothyroid muscle (CT), the vocalis muscle (VOC), the posterior crico-arytenoid muscle (PCA), and the interarytenoid muscle (INT).

In the survey of findings below, we have found it most useful to group the results according to the muscle studied (rather than according to the subject). It may be doubtful whether one can arrive at clear differentiations between the performances of these muscles as regards the frequency characteristics of the EMG signals. It is, however, interesting to distinguish between the set of muscles (CT and VOC) which were reached by insertion of a needle through the neck, and the rest, which were reached via the oral and pharyngeal cavities, since the conditions for movement artifacts may be expected to be somewhat different in the two cases. Moreover, it may be interesting to distinguish between VOC and the other muscles, since VOC is obviously more vulnerable to microphony ("voice bar" interference) than the others. The clearest presentation, then, is obtained by simply treating each muscle separately. The specimens chosen for analysis include artifact types (III) - microphony - and (IV) - movement artifact - if they appear at all in the recordings.

5. Survey of results from filtering

5.1 The crico-arytenoid muscle (CT)

The frequency characteristics of recordings from CT were examined for subjects BF, BH, HU, BM, JR, and JJ. All subjects are Danish.

Subject BF: two specimens were analyzed. One of these was highpass filtered in steps from 47 to 1600 Hz, and lowpass filtered in steps from 3300 to 100 Hz. The other specimen was filtered within the same range except for the highpass filtering which started at 200 Hz.

On visual inspection of the raw curves these look "clean", except that there appears to be mains interference in one of the specimens.

Results: the hum is obviously 50 Hz mains interference, which is easily removed by HP filtering, e.g. at 68 Hz. Otherwise, it is a signal which is very resistive to filtering, the overall shape being essentially constant at least in the frequency region 390 - 1600 Hz, and with rather evenly distributed energy in the region 390 - 800 Hz. LP filtering at 100 Hz is more detrimental to the output than is HP filtering, even at 1600 Hz.

Subject BH: two specimens were analyzed. One was highpass filtered in steps from 200 to 1000 Hz and lowpass filtered in steps from 1000 to 100 Hz. The other specimen was highpass filtered in steps from 250 to 1600 Hz and lowpass filtered in steps from 470 to 47 Hz.

On visual inspection the raw curves exhibit some (occasional?) hum as well as some apparent movement artifacts. It is likely that the electrodes had been misplaced in this recording, so that the muscle involved is in fact not CT (or there may be contamination from another muscle).

Results: the hum is easily removable by HP filtering at 250 Hz (the lowest cutoff frequency employed) without affecting the useful signal. The spurious spikes, which we identify as movement artifacts, are suppressed by HP 400 Hz and more completely by HP 630 Hz. The useful signal seems to have rather evenly distributed energy in a wide frequency region from 100 Hz upwards, although with considerable attenuation at HP filtering above 400 Hz.

Subject HU: one specimen was highpass filtered in steps from 200 to 2000 Hz and lowpass filtered in steps from 4700 to 47 Hz.

On visual inspection the raw curves do not seem to be contaminated by disturbances.

Results: the signal information is fairly constant over the frequency range from some 470 Hz upwards, but increasingly changed at lower frequencies. The low frequency component is relatively weak. It seems clear that a more well-defined signal is obtained by HP filtering at 470 Hz or even higher.

Subject BM: one specimen was analyzed. It was highpass filtered in steps from 200 to 1600 Hz, lowpass filtered in steps from 1000 to 470 Hz, and bandpass filtered with (exceptionally) a pass-band comprising the range 1000 - 1500 Hz.

On visual inspection the raw curves do not seem to be disturbed by artifacts of any kind.

Results: the result of the filtering was that only very minor changes occurred in the shape of the curves, and the signal has a very broad frequency spectrum, ranging from below 470 Hz to above 1600 Hz.

Subject JR: two specimens were analyzed. One of these was highpass filtered in steps from 22 to 2000 Hz, lowpass filtered in steps from 4400 to 82 Hz, and bandpass filtered with a passband comprising the range 22 - 330 Hz. The other specimen was highpass filtered in steps from 200 to 2000 Hz and lowpass filtered in steps from 4700 to 82 Hz.

On visual inspection the curves exhibit numerous spurious spikes and occasional microphony.

Results: microphony disappears by HP filtering at e.g. 200 Hz. The artifacts have a frequency spectrum which is most prominent at frequencies below 200 Hz, whereas the useful signal has its dominant spectral components at higher frequencies, apparently with most energy in the region 300 - 600 Hz. But the signal is well preserved, even at frequencies well above 1000 Hz. Thus it seems straightforward to improve the signal by HP filtering somewhere above 200 Hz, and it is probably safe to choose a cutoff frequency somewhere in the region of most spectral prominence (i.e. somewhat above 300 Hz).

Subject JJ: one specimen was analyzed. It was highpass filtered in steps from 22 to 1600 Hz, lowpass filtered in steps from 1000 to 47 Hz, and bandpass filtered in steps from 22 to 3300 Hz.

On visual inspection the raw curves seem free from disturbances except that there is microphony.

Results: BP and HP filtering showed that the microphony contains both the first and the second harmonic of the voice fundamental frequency, although the second harmonic does not do any harm at HP 200 Hz, since the total signal energy above this frequency completely overrides the microphony. At higher HP cutoff frequencies there are various changes in the shape of the curves, and the signal level becomes considerably weaker with HP cutoff frequencies in the vicinity of cutoff 1000 Hz.

5.2 The vocalis muscle (VOC)

The frequency characteristics of recordings from VOC were examined for subjects BF, BH, HU, NR, JR, MW, JJ, and BM. All subjects are Danish except for MW, who is German.

Subject BF: five specimens were analyzed. These specimens were chosen from different parts of a long recording because the performance varied quite a lot during the recording. The signal level was good in the beginning and later, but low in an intermediate portion, the temporary weakening having to do with the fact that the electrode was touched and got into a less suitable position but apparently restored its original conditions at a later stage. One of the specimens was chosen to exemplify a passage with some strange spikes, viz. just before p and s.

Three of the specimens were highpass filtered in steps from 47 to 2000 Hz, lowpass filtered in steps from 3300 to 47 Hz, and two of these were bandpass filtered in steps from 47 to 1000 Hz. The other two specimens were highpass filtered in steps from 200 to 1000 Hz and lowpass filtered in steps from 3300 to 270 Hz.

On visual inspection the signals exhibit no disturbances, except for the specimen extracted from the intermediate portion of the recording, which is not only weak but also filled with spurious spikes and microphony.

Results: the noise and movement artifacts in the weak recording can be attenuated by HP filtering at 200 Hz, but the signal remains poor. The remaining specimens exhibit a very good signal with a relatively even distribution of energy in the region 100 - 1000 Hz, but with the most prominent energy in the middle of this range. It is hardly appropriate to process these different parts of the recording together in terms of filtering, since they differ widely in signal level and do not become similar in quality anyway.

Subject BH: two specimens were analyzed. One specimen was highpass filtered in steps from 200 to 1600 Hz, lowpass filtered in steps from 1000 to 47 Hz, and bandpass filtered in steps from 39 to 680 Hz. The other specimen was highpass filtered in steps from 200 to 630 Hz and lowpass filtered in steps from 1000 to 150 Hz.

On visual inspection the signal is characterized by spurious spikes, which are interpreted as movement artifacts.

Results: the frequency content of the spurious spikes covers a rather wide band with most prominence around 100 - 150 Hz. But as the wide frequency band of the spurious spikes overlaps with the prominent frequency band of the useful signal, they can hardly be eliminated completely by filtering, although HP 400 Hz seems a reasonable compromise between too little suppression of artifacts and too much signal attenuation. With high cutoff frequencies the curve becomes less smooth with rather prominent single spikes reflecting the action potentials.

Subject HU: two specimens were analyzed. Both specimens were highpass filtered in steps from 200 to 2000 Hz and lowpass filtered in steps from 4700 to 47 Hz. Moreover, one of the specimens was bandpass filtered in steps from 47 to 4700 Hz.

On visual inspection the curves exhibit mains interference, microphony and various spurious spikes.

Results: the mains interference is removed by HP filtering at 200 Hz (the lowest cutoff frequency employed). Removal of the microphony required a somewhat higher cutoff frequency. HP filtering at higher cutoff frequencies gradually changed the fine structure of the signal, and in some cases this change is rather considerable at high cutoff frequencies, while at the same time the overall signal gets weaker. However, the signal is in part well preserved even at HP 2000 Hz. The apparent artifacts are of varying frequency content, and they are not all easily removable by filtering.

Subject NR: one specimen was analyzed. This specimen was highpass filtered in steps from 22 to 2000 Hz and lowpass filtered in steps from 4700 to 47 Hz.

On visual inspection the curves appear to contain microphony and artifacts.

Results: microphony and artifacts can be at least partially eliminated by HP filtering in the vicinity of 150 Hz. Most of the artifacts have their strongest spectral energy below 100 Hz, but one (perhaps not a mere artifact?) has most of its energy at higher frequencies. Since the fine structure of the signal fluctuates quite a lot from one frequency band to another, it may not be advisable to eliminate too much of the signal by filtering, and hence a compromise must be found.

Subject JR: two specimens were analyzed. One of the specimens was highpass filtered in steps from 200 to 2000 Hz, lowpass filtered in steps from 4700 to 68 Hz, and bandpass filtered in steps from 22 to 330 Hz. The other specimen was highpass filtered in steps from 200 to 500 Hz, lowpass filtered in steps from 4700 to 68 Hz, and bandpass filtered in steps from 68 to 330 Hz.

On visual inspection this is a signal without much disturbance, so that there is not a priori a need for filtering.

Results: the most prominent frequency range is 100 - 500 Hz, but there is energy reflecting the action potentials even at much higher frequencies.

Subject MW: two specimens were analyzed. Both specimens were highpass filtered in steps from 200 to 2000 Hz, and lowpass filtered in steps from 560 to 27 Hz.

On visual inspection the signal seemed to be rather noisy and disturbed by movement artifacts.

Results: the movement artifacts are most prominent at some 220 Hz, but can be removed effectively only by HP filtering at a high cutoff frequency (some 800 Hz or more). At this frequency the useful signal is also attenuated a good deal, however. Still, HP filtering has the additional advantage of removing various noise components at relatively low frequencies. The "cleanest" looking signal is obtained by HP filtering above 1000 Hz.

Subject JJ: three specimens were analyzed. One of these was highpass filtered in steps from 22 to 2000 Hz, lowpass filtered in steps from 1000 to 33 Hz, and bandpass filtered in steps from 22 to 680 Hz; another was highpass filtered in steps from 47 to 1600 Hz, lowpass filtered in steps from 1000 to 69 Hz, and bandpass filtered in steps from 47 to 1000 Hz. The last specimen was only highpass and bandpass filtered from 200 to 2000 Hz and from 68 to 1200 Hz, respectively.

Inspection of the raw curves reveals occasional hum and spurious spikes, of which at least some are interpreted as movement artifacts.

Results: the filtering showed that the spurious spikes were not eliminated to the same degree by filtering, but in most cases they are eliminated by HP filtering with cutoff 150 Hz (which also eliminates occasional hum). The signal has components at much higher frequencies, so it seems perfectly possible to filter somewhere in the range 200 - 400 Hz. The signal level at HP 400 Hz is not much below that at HP 200 Hz, and the removal of spurious spikes is better.

Subject BM: one specimen was analyzed. It was highpass filtered in steps from 200 to 2000 Hz, lowpass filtered in steps from 1000 to 47 Hz, and bandpass filtered in steps from 22 to 12000 Hz.

Inspection of the raw curves reveals microphony and hum.

Results: microphony and hum disappear with the lowest HP filtering attempted, viz. HP 200 Hz. The filtering showed that the signal was very resistive to filtering, its overall shape being essentially constant, irrespective of the choice of cutoff frequency in the range HP 200 - 1600 Hz, and BP from 680 Hz and upwards, although the signal was weakened very much at very high BP frequencies.

5.3 The interarytenoid muscle (INT)

The frequency characteristics of recordings from INT were examined for subjects JPF, MF, LG, and PM. JPF is French, the others are Danish.

Subject JPF: four specimens were analyzed. One of the specimens exhibited an unstable signal since the pickup of the signal by the electrode suddenly became very weak during this period; after the failure of this electrode, another one was inserted, and one specimen represents the signal from the new electrode.

The first recording was highpass filtered in steps from 22 to 1600 Hz and lowpass filtered in steps from 560 to 56 Hz. The second recording was highpass filtered in steps from 200 to 2000 Hz and lowpass filtered in steps from 560 to 39 Hz.

On visual inspection both recordings (the specimens with the electrode inserted at the start of the recording session versus the specimen with a new electrode inserted) are characterized by spurious spikes, although these are not very prominent except in the specimen showing a very weak signal. Altogether, the signal is not very "clean".

Results: by HP filtering the signal is gradually weakened; this effect is noticeable from about HP 100 Hz upwards and very pronounced with HP 340 - 500 Hz. Above 500 Hz the signal is very weak with only occasional peaks of activity. Conversely, LP filtering affects the signal very little until rather low cutoff frequencies (below LP 200) are reached; it is only with cutoff frequencies below 100 Hz that the signal is considerably weakened. The apparent artifacts are essentially eliminated by HP 250 Hz. These characteristics are shared by both recordings.

Thus, both recordings exhibit a not very "clean" signal, dominated by the frequency region around and just above 100 Hz. Only in the beginning of the first recording is the signal so much better that HP filtering at higher frequencies (e.g. 630 Hz) can be successfully applied. The signal-to-noise ratio is such that one cannot generally achieve any considerable improvement by HP filtering (HP 250 seems to be the upper practical limit).

Subject MF: one specimen was analyzed. The signal was highpass filtered in steps from 150 to 2000 Hz and lowpass filtered in steps from 1000 to 22 Hz.

On visual inspection this signal contains only few artifacts (there is occasional hum in other parts of the recording, which has been found to be easily removable by HP filtering).

Results: HP filtering practically removed a spurious spike at cutoff 200 Hz. The useful signal was rather resistive to HP as well as to LP filtering, except with LP filtering with very low cutoff frequencies (100 Hz or lower). The above mentioned artifact appears to have energy in the range 40 - 200 Hz; the useful signal is well represented throughout the spectral range from some 100 Hz to at least 2000 Hz, but apparently with most energy in the range 300 - 500 Hz.

Subject LG: two different recordings (with different electrodes) were made. These are treated separately below.

(1) One specimen was analyzed. The signal was highpass filtered in steps from 82 to 2000 Hz, lowpass filtered in steps from 1500 to 47 Hz, and bandpass filtered in steps from 22 to 220 Hz.

On visual inspection the signal appears to contain some artifacts, but the distinction between artifacts and the useful signal is not altogether clear.

Results: HP filtering removes various spurious spikes, but they do not all disappear at the same cutoff frequency: some require only a cutoff frequency of 82 Hz to be sufficiently eliminated, others require HP 200 or 250 Hz. BP filtering shows that the former have their spectral energy below 100 Hz, but that the latter have most of their energy in the range 100 - 200 Hz. The useful signal is rather resistive to both HP and LP filtering: there is a discernible signal even at HP 2000 Hz and at LP 100 Hz (the latter strongly contaminated by the spurious spikes). However, most of the spectral energy is at relatively low frequencies: at HP 630 Hz the signal is very considerably weakened.

(2) One specimen representing the signal from the other electrode was analyzed. The signal was highpass filtered in steps from 200 to 2000 Hz, lowpass filtered in steps from 1000 to 33 Hz, and bandpass filtered in steps from 22 to 220 Hz.

On visual inspection this is a "clean" signal, though with apparent artifacts.

Results: the spurious spikes are at least in part artifacts; these are essentially eliminated by HP filtering with a cutoff of 200 Hz. The useful signal has a broad spectrum, ranging all the way from 27 Hz to above 2000 Hz, although it is considerably weakened towards higher frequencies.

Subject PM: one specimen was analyzed. The signal was highpass filtered in steps from 200 to 1000 Hz, and lowpass filtered in steps from 1000 to 100 Hz.

According to Dr. Hirose, this is not a very clear signal, but it must be INT, since there is not much possibility of contamination with other muscles (during insertion the needle pointed at the middle of the muscle). What we have here is an EMG interference pattern with different motor units involved, and it is hard to tell what information is present without averaging over several tokens.

Results: the signal is relatively constant over most of the frequency range studied, but the signal gets perceptibly weaker with LP 270 Hz (or lower), and with HP 630 Hz (or higher), i.e., the spectral energy is most prominent between these frequencies. There are artifact-like spikes which have most of their energy below 100 Hz, but the ratio between the remainder of the signal and these spurious spikes is improved by HP filtering at higher cutoff frequencies. The optimum is reached with HP 400 Hz, which does not weaken the useful signal very much.

5.4 The posterior crico-arytenoid muscle (PCA)

The frequency characteristics of recordings from PCA were examined for subjects MF, LG, and PM, all speakers of Danish.

Subject MF: two different specimens were analyzed. One of these was highpass filtered in steps from 47 to 2000 Hz, and lowpass filtered in steps from 1000 to 22 Hz; the other specimen was highpass filtered in steps from 200 to 1000 Hz and lowpass filtered in steps from 1000 to 100 Hz.

On visual inspection this signal is found to contain artifacts. There is occasional hum elsewhere in the recording, but not in these specimens.

Results: the signal is found to have a broad frequency spectrum but with particular prominence of a rather narrow region (the signal is considerably weakened both by LP filtering with cutoff 270 Hz and by HP filtering with cutoff 630 Hz). The artifacts are eliminated by HP filtering with cutoff 150 or 200 Hz (which also eliminates occasional hum), which can be done safely since the useful signal has its essential energy at higher frequencies.

Subject LG: one specimen was analyzed. It was highpass filtered in steps from 200 to 2000 Hz, lowpass filtered in steps from 1000 to 33 Hz, and bandpass filtered in steps from 47 to 2200 Hz.

On visual inspection this signal contains spurious spikes which are interpreted as artifacts, and it is also contaminated by microphony.

Results: around 47 Hz there is activity only in some of the spurious spikes. The microphony is narrow-banded: it is weakened by LP filtering with cutoff 100 Hz and eliminated by HP filtering with cutoff 200 Hz. The overall (useful) signal seems to have a broad spectrum ranging all the way from some 100 Hz to above 2000 Hz, but with a rather uneven distribution of energy in different frequency regions. As for the apparent artifacts, these are not all affected in the same way by filtering, but generally HP filtering improves the ratio between the useful signal and the spurious spikes. Some of the latter exhibit energy at relatively high frequencies, however. In this signal it is not very easily determined what is an artifact and whether it may be necessary to distinguish between different kinds of "artifacts" in a broad sense.

Subject PM: three different specimens were analyzed. These were highpass filtered within different ranges, together encompassing the range 22 - 1600 Hz, and (one specimen) lowpass filtered in steps within the range 1000 - 47 Hz.

On visual inspection the signal seems very weak, and it contains spurious spikes of the type interpreted as movement artifacts.

Results: suppression of various spurious spikes is obtained by HP filtering with cutoff 200 Hz, and in part even at 100 Hz. The signal itself seems to have a certain spectral dominance somewhere around 300 - 400 Hz: it is quite considerably attenuated by LP filtering with cutoff 270 Hz and likewise by HP filtering with cutoff frequencies above 400 Hz. HP filtering with a cutoff frequency in the range 100 - 200 Hz would seem advisable in such a case.

6. Discussion and conclusion

There is a great deal of variation within the material considered here. As regards the subjects for which two or more muscles have been recorded, i.e. most of the subjects, it is possible to see whether the frequency characteristics and general appearance of the EMG signals differ more characteristically from one person to another, or more characteristically from one muscle to another. However, on the basis of the results above, it seems hard to draw any such conclusions.

There are, however, some obvious recurrent characteristics of these EMG curves regarding the frequency content of (1) the useful signal (recording of action potentials), (2) apparent movement artifacts, (3) microphony, and (4) hum.

(1) The useful EMG signal, in the case of larynx muscles, mostly covers a frequency range from some 100 Hz upwards. There is often energy even at 2000 Hz or higher, but it is normally so that the spectral energy is most prominent below 1000 Hz and

especially in a region from some 200 Hz to some 600 Hz. The signals which seem generally least contaminated by noise and which present the clearest picture of the EMG activity typically exhibit energy also at high frequencies. In such cases, a faithful reproduction of the overall shape of the curves can be obtained even with HP filtering at very high frequencies (in the region 1000 - 2000 Hz), provided that adequate amplification is available (cp. figs. 1 and 2).

This result is clearly at variance with the general experience from large muscles.

(2) Spurious spikes which, according to their location in time etc., must be interpreted as movement artifacts, typically have their spectral center of gravity at low frequencies, i.e. essentially below the useful frequency range of the EMG signal or only slightly overlapping it. As a rule of thumb, these artifacts are situated in the frequency region up to 100 or 150 Hz (see fig. 3a and 3b). However, there are other spurious spikes which have a broader spectrum, or a spectrum shifted towards higher frequencies (see fig. 4). Some of these occur during the test words and typically in connection with features such as the Danish stød (see fig. 5). It should be investigated whether there is a difference in frequency content between <u>different kinds of movement artifacts</u> (e.g. between artifacts caused by external pull in the leads of the electrodes and artifacts caused by violent muscle contraction). This is important in order to know to what extent it is at all possible to discriminate between movement artifacts and suspicious spikes which are in fact part of the EMG signal (note that the interference pattern of contributions from different motor units may give quite a complex curve at times).

(3) Microphony is easily removed by HP filtering. It is mostly dominated by the first harmonic, so that HP filtering in the range 100 - 250 Hz should suffice (depending on the voice), see fig. 6. Contrary to expectation, microphony did not occur only in recordings from the vocalis muscle; we have found no explanation of why it sometimes occurs with one muscle, at other times with another, and why it sometimes occurs only intermittently.

(4) Hum is mains interference, probably picked up if the electric shielding somehow fails to be complete. In our recordings it is only occasionally a problem, and it is generally no problem to remove this interference by HP filtering (cutoff somewhere between 50 and 200 Hz).

Thus, for these EMG data from the internal laryngeal muscles recorded by bipolar hooked-wire electrodes, we conclude that the highpass cutoff frequency normally proposed in the literature - i.e. 100 Hz or lower - is generally not sufficient to eliminate microphony and movement artifacts. Especially in the case of movement artifacts a considerably higher cutoff frequency is fairly often required. The most suitable cutoff frequency seems to depend both on the general characteristics of the individual recording and on the position of the artifact in relation to the articulation involved. The choice of filtering frequency will often be a compromise between the cutoff that is optimal with respect to removal of spurious components and the attenuation of the overall signal caused by filtering. However, it is generally advisable to highpass filter the signal even at the expense of some overall attenuation, and the greatest improvement is often achieved by highpass filtering with cutoff frequencies well above the low range normally proposed.

References

Abbs, J.H. and K.L. Watkins 1976: "Instrumentation for the study of speech physiology", Contemporary Issues in Experimental Phonetics (ed. N.J. Lass), Academic Press

Fromkin, V.A. 1965: "Some phonetic specifications of linguistic units: An electromyographic investigation", UCLA WPP 9, p. 170-199

Geddes, L.A. 1972: Electrodes and the measurements of bioelectric events, New York: Wiley-Interscience

Hayes, K.J. 1960: "Wave analysis of tissue noise and muscle action potentials", Journal of Applied Physiology 15, p. 749-752

Herberts, P. 1969: "Myoelectric signals in control of prostheses", Acta Orthopaedica Scandinavia, suppl. 124

Hirano, M. and T. Smith 1967: "Electromyographic study of tongue function in speech: A preliminary report", UCLA WPP 7, p. 46-56

Hirano, M., J. Ohala, and T. Smith 1967: "Current techniques used in obtaining EMG data", UCLA WPP 7, p. 20-24

Hubler, S.A. 1967: "A high input impedance electromyography preamplifier", UCLA WPP 7, p. 25-34

Kadefors, R., E. Kaiser, and J. Petersén 1968: "Dynamic spectrum analysis of myopotentials and with special reference to muscle fatigue", Electromyography 8, p. 39-74

Kaiser, E. and J. Petersén 1965: "Muscle action potentials studied by frequency analysis and duration measurements", Acta neurol. Scandinavia, suppl. 13, p. 213-236

Kewley-Port, D. 1977: "EMG signal processing for speech research", Haskins SR 50, p. 123-146

Kwatny, E., D.H. Thomas, and H.G. Kwatny 1970: "An application of signal processing techniques to the study of myoelectric signals", IEEE trans. on Biomedical Engineering, BME-17, p. 303-313

Mansell, P. "The articulation of German plosives" Section IIID (Frequency analysis of EMG signals), Forschungsberichte 11 München, p. 102-118

Scott, R.N. 1967: "Myo-electric energy spectra", Med. & Biol. Engineering 5, p. 303-305

Trimble, J.L., L. Zuber, and S.N. Trimble 1973: "A spectral analysis of single motor unit potentials from human extraocular muscle", IEEE Trans. on Biomedical Engineering, BME-20, p. 148-151

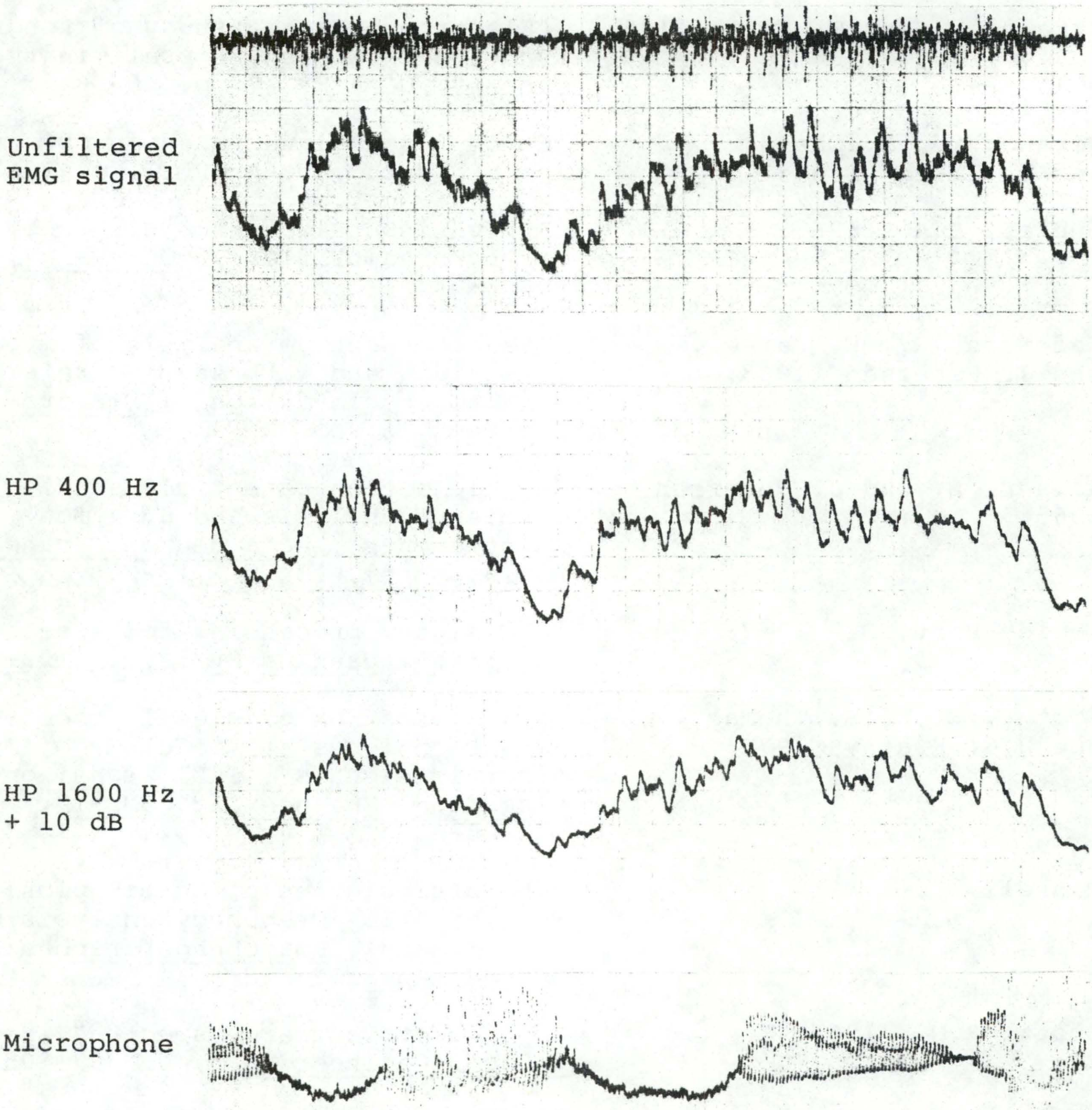


Figure 1

EMG signal of the interarytenoid muscle, highpass filtered at 400 Hz and 1600 Hz. In all figures the uppermost curve shows the unfiltered signal and the lowest one the microphone signal. Subject: MF, text: "(d)e sagde sile" (the text in brackets is not shown in the figure).

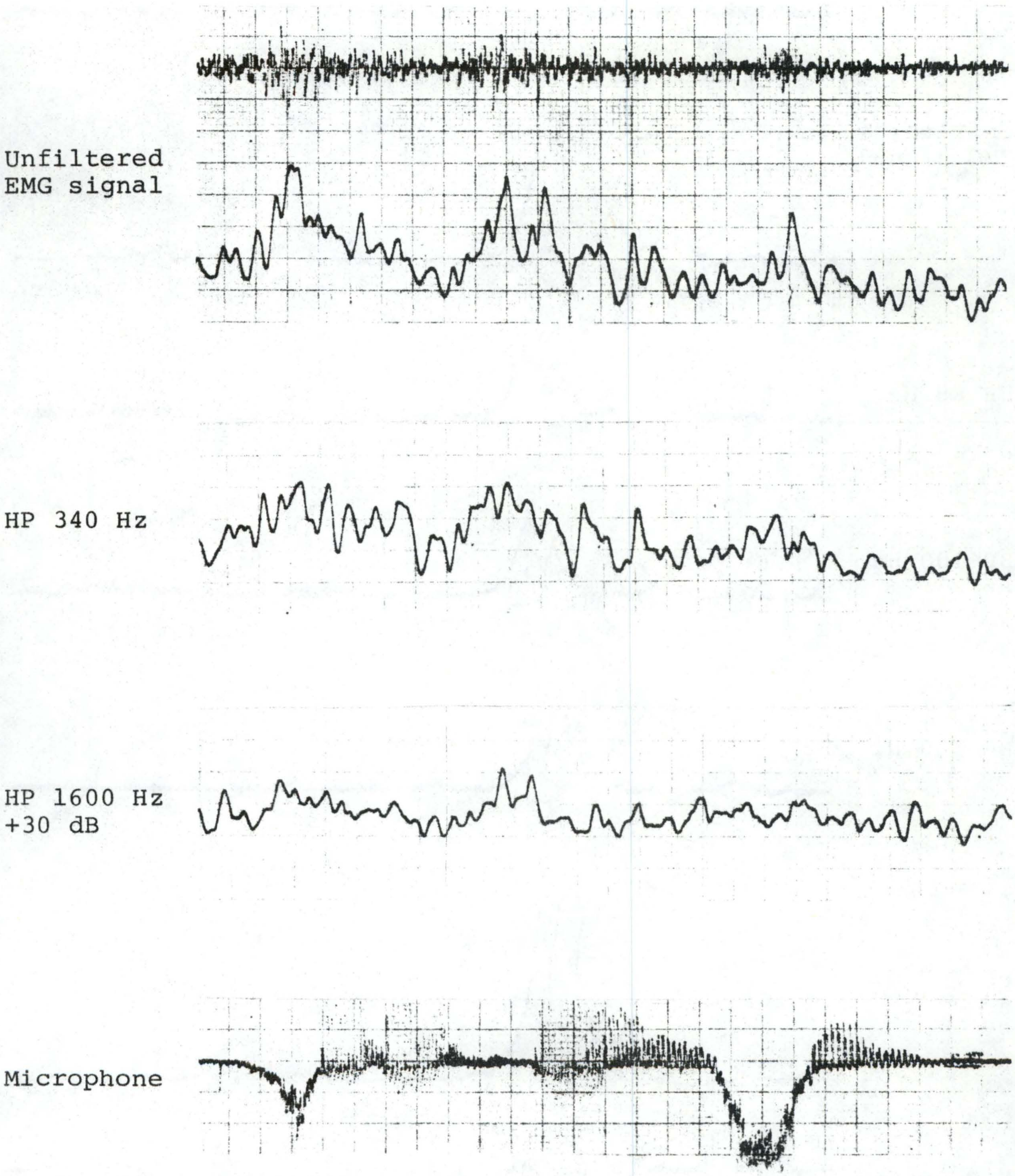


Figure 2

EMG signal of the interarytenoid muscle, highpass filtered at 340 Hz and 1600 Hz.
Subject: JPF, text: "c'est la panne ici".

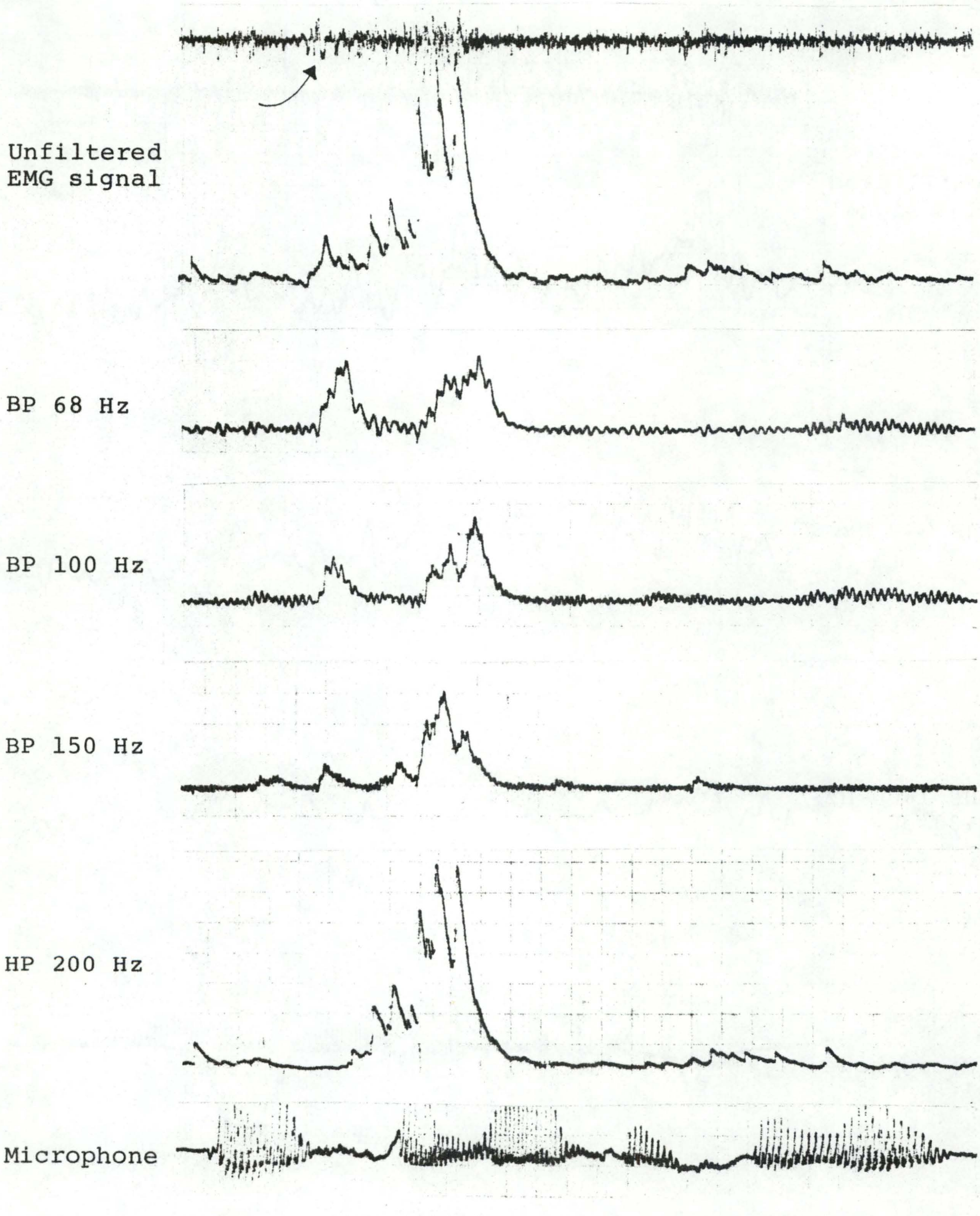
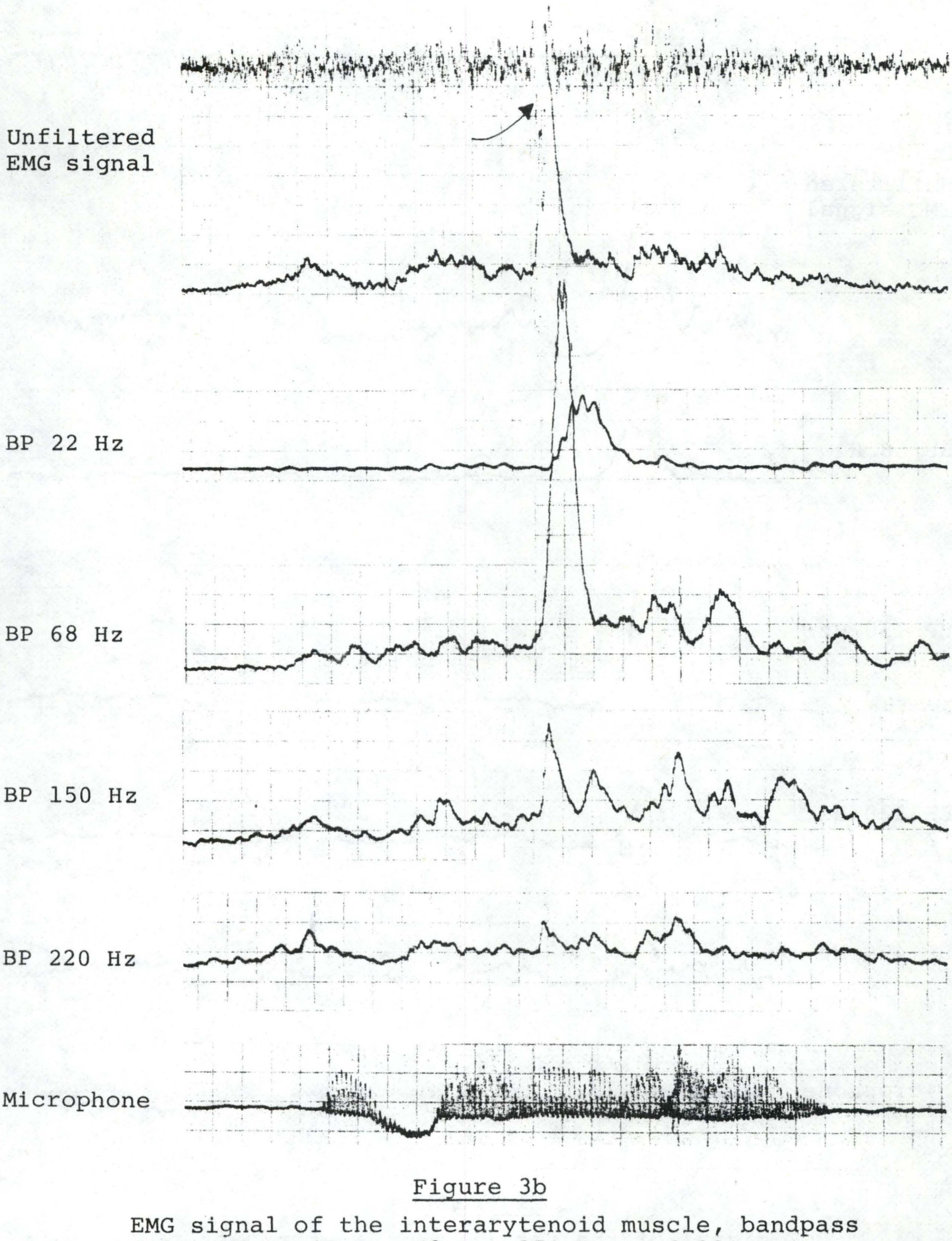


Figure 3a

EMG signal of the vocalis muscle, bandpass filtered at 68 Hz, 100 Hz, and 150 Hz. The lowest EMG curve shows the signal highpass filtered at 200 Hz. The arrow in the unfiltered curve points at the artifact. Subject: JJ, text: "det er hu'en de siger" (' indicates the Danish stød).



<u>Figure 3b</u>

EMG signal of the interarytenoid muscle, bandpass filtered at 22 Hz, 68 Hz, 150 Hz, and 220 Hz. The arrow in the unfiltered curve points at the artifact.
Subject: LG, text: "de sagde male".

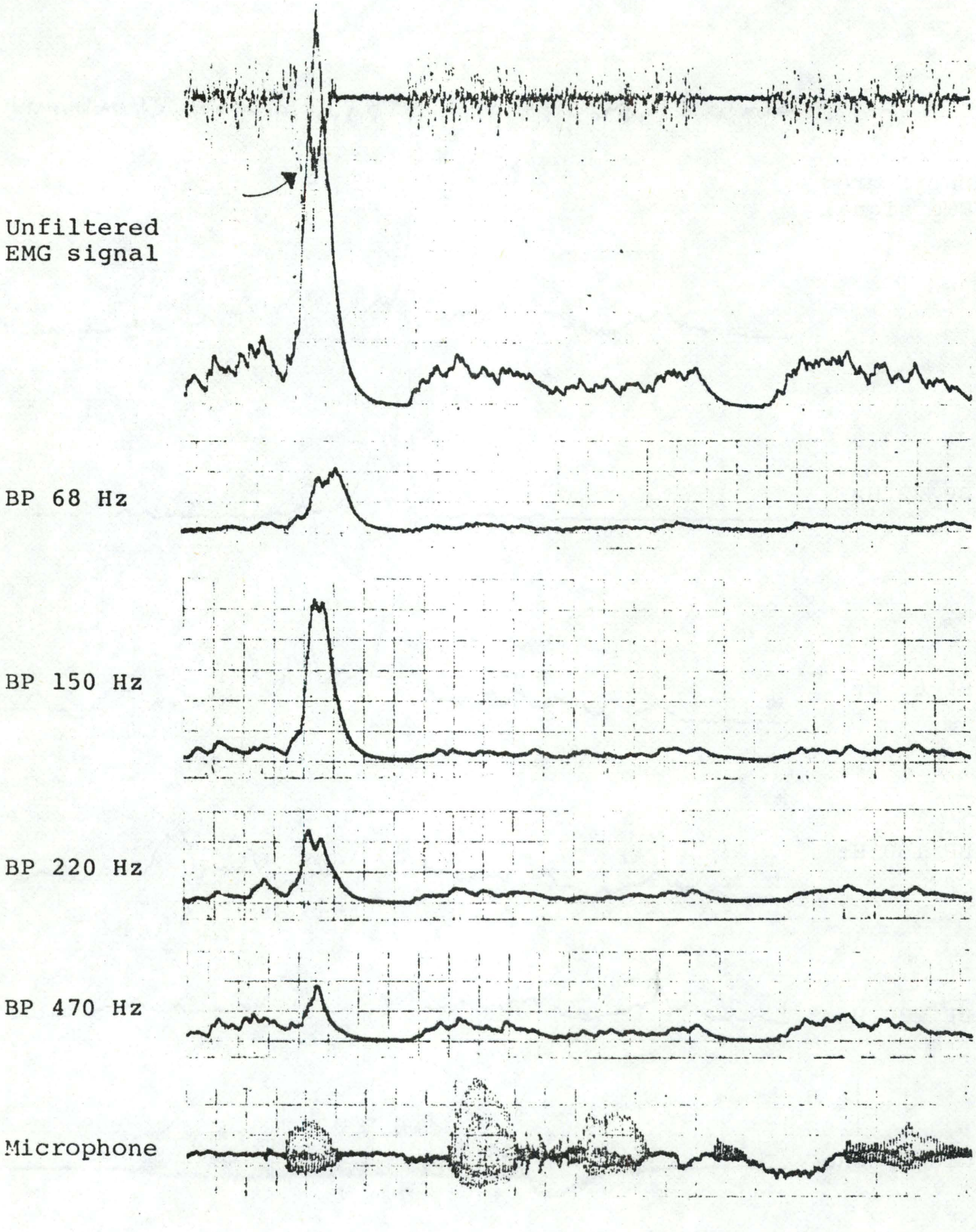


Figure 4

EMG signal of the vocalis muscle, bandpass filtered at 68 Hz, 150 Hz, 220 Hz, and 470 Hz. The arrow in the unfiltered curve points at the artifact. Subject: BH, text: "det er kæ'ler de siger".

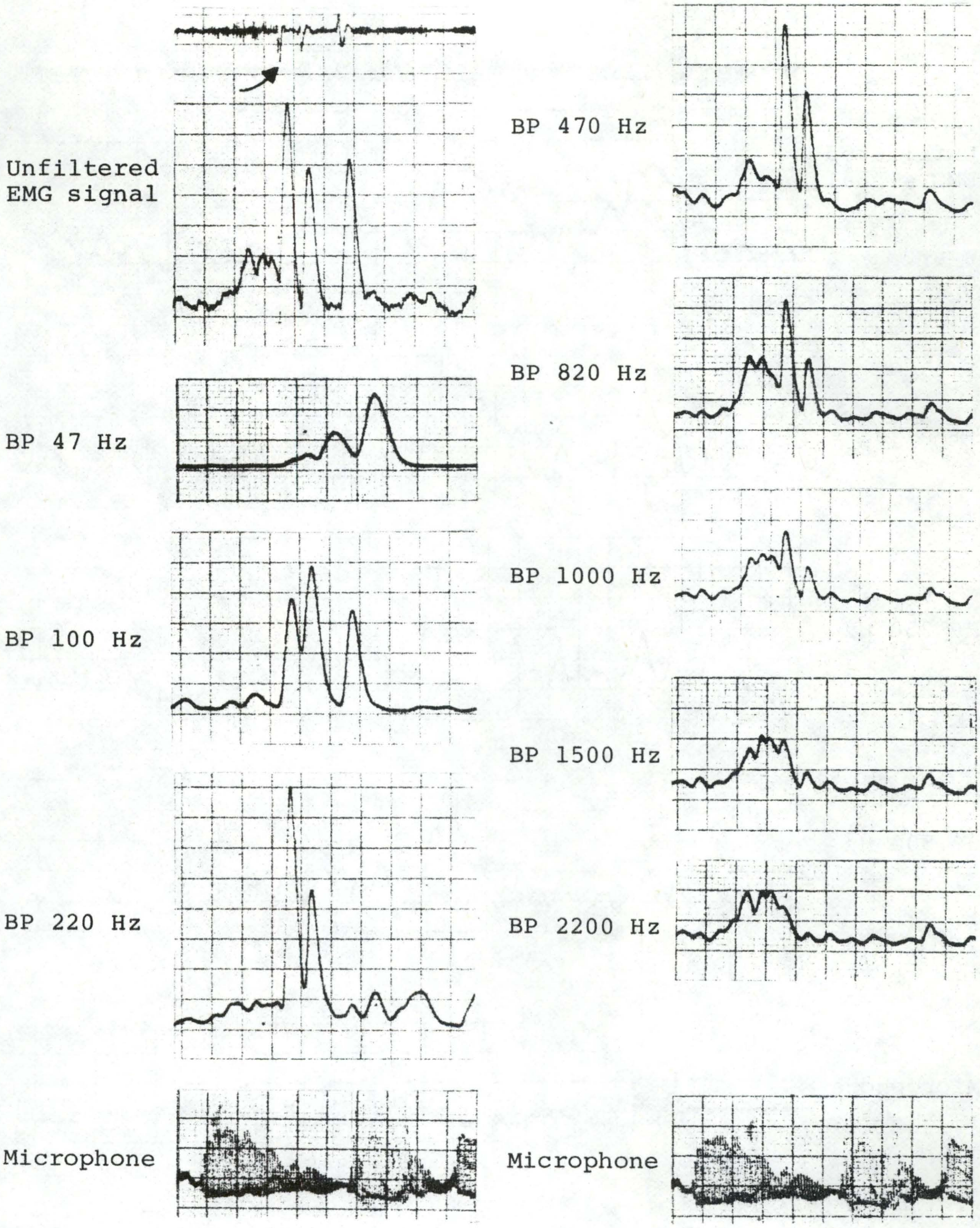


Figure 5

EMG signal of the vocalis muscle, bandpass filtered at different frequencies between 47 and 2200 Hz. The arrow in the unfiltered curve points at the artifact occurring in the stød-phase. Two other artifacts are seen.
Subject: HU, text: "(det er p)i'ber d(e siger)".

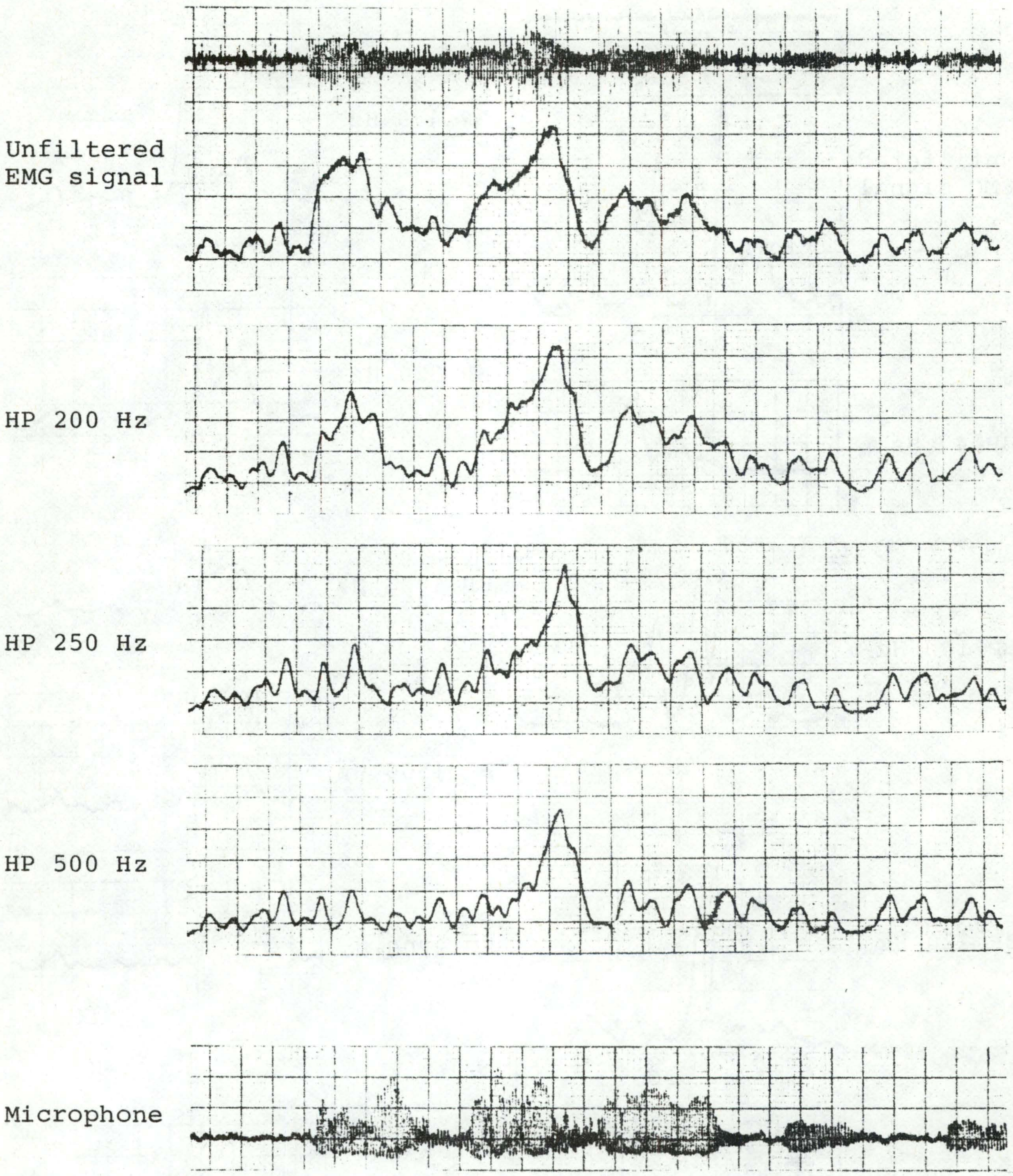


Figure 6

EMG signal of the vocalis muscle, showing microphony in a female voice. The signal is highpass filtered at 200 Hz, 250 Hz, and 500 Hz. Subject: HU, text: "det er ven'nen de siger".

(ARIPUC 14, 1980)

COARTICULATION OF INHERENT FUNDAMENTAL FREQUENCY LEVELS BETWEEN SYLLABLES

Niels Reinholt Petersen

Abstract: In Advanced Standard Copenhagen (ASC) Danish the inherent Fo level differences between high and low vowels are of an order of magnitude comparable to that of the Fo deflection in the stress groups. In theory, this could imply that the intended stress distribution in an utterance might be perceived incorrectly, since the inherent Fo level variation could distort the linguistically relevant Fo pattern in the stress groups. It was hypothesized that such distortion is compensated for by coarticulation of inherent Fo levels between syllables. Experiments carried out to test this hypothesis showed that Fo level in one vowel is consistently influenced by vowel height in the preceding syllable, being higher after a high than after a low vowel. There seems to be no consistent influence from the succeeding vowel. The greatest amount of compensation is found in the first posttonic vowel in the stress group, and it occurs whether the consonant between the two vowels is a sonorant (m) or a voiceless fricative (f). Formant measurements of first posttonic ɑ showed that the formant frequencies are influenced (although slightly) by both the preceding and the following vowels. An explanation of the discrepancy between the manner in which vowel quality (and hence tongue body position) and the manner in which Fo is influenced by adjacent vowels is attempted in terms of a spring-mass model describing the relation between the tongue body and the laryngeal structures.

1. Introduction

The observed course of fundamental frequency in an utterance can be viewed as the result of contributions from several simultaneous components. In a non-tonal language such as Advanced Standard Copenhagen (ASC) Danish four components have to be considered (Thorsen 1979): (1) a sentence component which gives the intonation contour of the sentence, (2) a stress group component which supplies the Fo patterns of the stress groups (a stress group in ASC Danish is constituted by a stressed syllable plus the following unstressed ones, irrespective of intervening syntactic boundaries on the same intonation contour, cf. Thorsen 1980b), (3) - in words with stød - a stød component which yields Fo movements characteristic of the stød, and (4) a segmental (or microprosodic) component which renders the Fo variation attributable to the segments constituting the utterance, such as intrinsic Fo level differences between segments and coarticulatory effects on Fo at and across segment boundaries. Since it is a consequence of inherent properties of the speech production system, and cannot be voluntarily controlled by the speaker[1], the segmentally determined Fo variation cannot - unlike the contributions from components 1, 2, and 3 - carry linguistically relevant prosodic information. On the contrary, if sufficiently large, the segmentally determined Fo variation could be expected to interfere with and possibly distort the contributions to the fundamental frequency course from the linguistically relevant components.

The present paper examines the interaction between the segmental component (with the emphasis on inherent Fo level differences between vowels) and one of the linguistically relevant components, viz. the stress group component.

The Fo pattern of the ASC stress group can be described as a relatively low stressed syllable followed by a high-falling tail of unstressed syllables (Thorsen 1979, 1980b). The Fo rise from the stressed to the first posttonic syllable varies (with position in time and in the intonation contour) between 3 and 0.5 semitones. These values are very similar to the inherent Fo level differences between high and low vowels, which have been found by Reinholt Petersen (1978, 1979) to vary between 1 and 3 semitones. In theory this might imply that the segmental composition of an utterance could distort the Fo pattern in such a manner that - assumed that Fo is a cue to the perception of stress - the perceived stress pattern of an utterance would differ from that intended by the speaker due to the qualities of the vowels in the utterance.

---

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

Now, apparently this does not happen. People normally perceive the stress patterns as intended by the speaker. One explanation for this could be that the listener perceptually compensates for inherent Fo differences, and thus reconstructs the intended Fo contour and patterns. A similar perceptual compensation seems to take place as far as intrinsic durational differences between high and low vowels are concerned (Reinholt Petersen 1974). However, the inherent Fo level differences are considerably larger in stressed than in unstressed syllables (Reinholt Petersen 1979), and this means that the perceptual system would have to know the stress distribution in order to be able to select the appropriate correction factor for the reconstruction of the intended stress group pattern. Another possibility, which is the one under investigation here, could be hypothesized, namely that the compensation takes place in the speech production system as coarticulation of inherent Fo levels between syllables. A coarticulatory effect of this kind would more or less eliminate the Fo distortion introduced by the segmental component, and thus preserve the intended fundamental frequency course. For example, in a sequence of high stressed vowel plus low first posttonic, where the distortion due to vowel height could reduce the intended fundamental frequency rise from the stressed to the first posttonic syllable, the inherently high Fo level of the stressed vowel could be hypothesized to be carried over to the first posttonic vowel, raising its Fo, or (less likely, perhaps) the inherently low Fo level in the first posttonic could be anticipated in the stressed vowel, lowering the Fo of that vowel. In both cases the distortion introduced by the vowels in the sequence would be smoothed out and the intended Fo pattern preserved (to a greater or lesser extent depending on the magnitude of the coarticulation effect).

2. Experiment I

Experiment I concentrated on the inherent Fo level differences between vowels and their influence on Fo in adjacent syllables. More specifically, the following questions were considered:
(1) Can Fo in one syllable be influenced by a high vs. low vowel (i.e. a vowel with high vs. low inherent Fo level) in adjacent

syllables? (2) Is the effect, if any, directional (i.e. is it the preceding or the following syllable which has the stronger effect)? (3) Are there differences between syllables in different positions in the stress group as to how strongly Fo is influenced?

2.1 Method

2.1.1 Material

The test material consisted of nonsense words of the structure mVmVmV with the vowels i, u and ɑ[1] alternating within the words as described below. The test words were embedded in carrier sentences where the word immediately following the test word could be either markeres [mɑˈg̊ʰeˀʌs] 'is/are marked', muteres [muˈd̥ˢeˀʌs] 'is/are mutated', or miteres [miˈd̥ˢeˀʌs] a non-existing, but possible word in Danish. By having i, u, or ɑ in the first syllable of the word following the test word, the sequence within which the vowels could be systematically varied could be extended from three to four syllables. This was considered justified, since word boundaries do not seem to have any influence on the Fo pattern in ASC Danish (Thorsen 1980a). The frame preceding the test word was Stavelserne i ... [ˈsd̥æwˌsʌnə i] 'The syllables of ...".

The vowels examined for effects on Fo from preceding and following high and low vowels occurred in stressed, first posttonic, and second posttonic position. In addition to this, vowels in first pretonic position (i.e. vowels in the last unstressed syllable in the preceding stress group) were examined, in order to see whether stressed vowels might produce a stronger effect on the preceding vowel than do unstressed vowels. (It should be noted that the effect (if any) of stressed vowels on first pretonic vowels operate across the stress group boundary, whereas effects of first and second posttonic vowels on preceding vowels operate within the stress group.) The material consisted of the following 46 test sequences (henceforth a test sequence, i.e. the three-syllable test word plus the first syllable of the following words, will also be referred to as a test word):

---

1) The symbol ɑ is used in this paper to denote a low, unrounded mid vowel [ɑ̄-] or [ɑ+].

1 ˈmimimi mi-
2 ˈmimimi ma-
3 ˈmimima ma-
4 ˈmimama ma-
5 ˈmamama ma-
6 ˈmamama mi-
7 ˈmamami mi-
8 ˈmamimi mi-
9 ˈmimima mi-
10 ˈmimami mi-
11 ˈmamami ma-
12 ˈmamima ma-

13 ˈmumumu mu-
14 ˈmumumu ma-
15 ˈmumuma ma-
16 ˈmumama ma-
17 ˈmamama mu-
18 ˈmamamu mu-
19 ˈmamumu mu-
20 ˈmumuma mu-
21 ˈmumamu mu-
22 ˈmamamu ma-
23 ˈmamuma ma-

24 mimiˈmi mi-
25 mimiˈmi ma-
26 mimiˈma ma-
27 mimaˈma ma-
28 mamaˈma ma-
29 mamaˈma mi-
30 mamaˈmi mi-
31 mamiˈmi mi-
32 mimiˈma mi-
33 mimaˈmi mi-
34 mamaˈmi ma-
35 mamiˈma ma-

36 mumuˈmu mu-
37 mumuˈmu ma-
38 mumuˈma ma-
39 mumaˈma ma-
40 mamaˈma mu-
41 mamaˈmu mu-
42 mamuˈmu mu-
43 mumuˈma mu-
44 mumaˈmu mu-
45 mamaˈmu ma-
46 mamuˈma ma-

The vowels underlined in the list are those which were to be examined for effects on the fundamental frequency from vowel height variation in adjacent syllables. In words 1-23 first and second posttonic vowels were examined and in words 24-46 first pretonic and stressed vowels were examined. To give an example: first posttonic i could be measured in i-i environments in word 1, in i-a environments in word 3, in a-i environments in word 8, and in a-a environments in word 12.

2.1.2 Recordings and speakers

The reading list consisted of 46 test sentences and 14 distractor sentences arranged in six different random orders. The recordings took place in a sound treated room in the Institute of Phonetics by means of a REVOX A700 professional tape recorder and a Sennheiser MD21 microphone.

There were four subjects, two females (KM and SI) and two males (PA and NR (the author)), who were all phoneticians and all speakers of ASC Danish. The subjects were instructed to read the list with a neutral declarative intonation at a comfortable speech rate.

As the list was long and proved rather difficult to read, the recordings were divided into readings of one or two randomizations at a time. As it turned out, this caused the overall Fo levels for the individual speakers to vary between readings (see further section 2.2 below). Altogether, six repetitions of each test word were obtained.

2.1.3 Registrations and measurements

The apparatus for registration was a REVOX A77 tape recorder, two intensity meters, a fundamental frequency meter (F-J Electronics), and a Mingograph (Elema 800). The following acoustic curves were made: duplex oscillogram, linear HiFi intensity curve (integration time 2.5 ms), two logarithmic intensity curves (integration time 2.5 ms) HP-filtered at 500 and 2000 Hz, respectively, and a fundamental frequency curve. The paper speed of the Mingograph was 100mm/sec. The duplex oscillogram and the intensity curves were used for segmentation purposes only. The accuracy of segmentation was ± 0.5 cs. The fundamental frequency of the vowels was measured at a point in time two thirds from the vowel onset (cf. Rossi 1976 and 1978). The scale of the Fo curve varied between 1 and 3 Hz/mm depending on the speaker, i.e. an accuracy of measurement of 1 Hz or less was achieved.

2.2 Results

As mentioned in section 2.1.2 above, the recording of the material was divided in several readings of one or two randomizations at a time. This meant a risk for differences in overall Fo level between readings to add unduly to the statistical variability of the measurements. A Friedman two-way analysis of variance by ranks (Siegel 1956) applied to the data showed a statistically significant effect of reading (randomization) on the overall Fo level for all subjects (KM: $\chi_r^2 = 12.39$, $p < 0.05$; SI: $\chi_r^2 = 112.21$, $p < 0.001$; PA: $\chi_r^2 = 63.88$, $p < 0.001$; NR: $\chi_r^2 = 121.06$, $p < 0.001$). Therefore, a normalization procedure was employed in which each of the measurements in one randomization was converted into the deviation from the mean of all measurements in that randomization. This procedure reduced the variation among the six repetitions of the test vowels, and the original means could be restored by adding the mean deviations over the six repetitions to the grand mean, i.e. the mean of all measurements in all randomizations.

The means and standard deviations computed on the basis of the normalized data are listed in tables 1 to 4 for ɪ and ɑ in ɪ/ɑ-words and u and ɑ in u/ɑ-words in each of the four stress group positions examined, viz. first pretonic, stressed, first posttonic, and second posttonic positions. In figs. 1 to 4 the mean Fo values for the individual subjects are plotted as a function of the tongue heights of the preceding and following vowels, respectively (i.e. the data points in the figures represent the row ($\overline{X}_{r.}$) and column ($\overline{X}_{c.}$) means given in tables 1 to 4). Fig. 5 shows the means over all subjects expressed in semitones derived by converting the normalized measurements for each subject into semitone deviations from that subject's grand mean, and then averaging the deviations over the subjects.

In the evaluation of the data as presented in tables 1 - 4 and figs. 1 - 5 it should be kept in mind that the material was organized in such a manner that first posttonic and second posttonic vowels (indicated by +1 and +2 in the figures) appeared in syllables number two and three (i.e. consecutively) in words with the stress on the first syllable (words 1-23), and that first pretonic and stressed vowels (indicated by -1 and 0 in the figures) appeared in syllables number two and three in words with the stress on the third syllable (words 24-46) (see section 2.1.1 above). This means that, apart from comparisons within stress group positions (which was the object of the design of the material), the Fo levels measured can only be compared between the first and second posttonic syllables, and between the first pretonic and stressed syllables. Comparing the Fo levels of stressed and first posttonic vowels, on the other hand, in order to derive the Fo rises from stressed to first posttonic syllables would lead to an overestimation (although not very great) of the rises, because the

Table 1

Mean fundamental frequencies (in Hz) and standard deviations for the vowels i, u, and ɑ in first pretonic, stressed, first posttonic, and second posttonic syllables before and after high vs. low vowels. Speaker KM.

1st pretonic

prec. vowel	i in i/ɑ-words, foll. vowel i	foll. vowel ɑ	$\bar{X}_r$	ɑ in i/ɑ-words, foll. vowel i	foll. vowel ɑ	$\bar{X}_r$	prec. vowel	u in u/ɑ-words, foll. vowel u	foll. vowel ɑ	$\bar{X}_r$	ɑ in u/ɑ-words, foll. vowel u	foll. vowel ɑ	$\bar{X}_r$
i	190	197	194	181	186	184	u	195	190	193	185	187	186
	2.2	7.5		2.4	7.5			5.3	6.3		3.9	4.0	
ɑ	189	187	188	187	188	188	ɑ	189	187	188	185	188	187
	4.3	2.6		3.3	5.6			2.4	3.9		4.7	5.6	
$\bar{X}_c$	190	192		184	187		$\bar{X}_c$	192	189		185	188	

stressed

prec. vowel	i in i/ɑ-words, foll. vowel i	foll. vowel ɑ	$\bar{X}_r$	ɑ in i/ɑ-words, foll. vowel i	foll. vowel ɑ	$\bar{X}_r$	prec. vowel	u in u/ɑ-words, foll. vowel u	foll. vowel ɑ	$\bar{X}_r$	ɑ in u/ɑ-words, foll. vowel u	foll. vowel ɑ	$\bar{X}_r$
i	203	201	202	187	186	187	u	209	208	209	186	188	187
	5.8	4.4		5.4	5.9			4.6	6.2		5.6	5.3	
ɑ	204	203	204	185	190	188	ɑ	201	202	202	186	190	188
	6.8	3.5		6.6	5.2			4.8	7.4		4.7	5.2	
$\bar{X}_c$	204	202		186	188		$\bar{X}_c$	205	205		186	189	

1st posttonic

prec. vowel	i in i/ɑ-words, foll. vowel i	foll. vowel ɑ	$\bar{X}_r$	ɑ in i/ɑ-words, foll. vowel i	foll. vowel ɑ	$\bar{X}_r$	prec. vowel	u in u/ɑ-words, foll. vowel u	foll. vowel ɑ	$\bar{X}_r$	ɑ in u/ɑ-words, foll. vowel u	foll. vowel ɑ	$\bar{X}_r$
i	226	229	228	222	222	222	u	229	223	226	219	216	218
	5.1	3.5		5.5	7.1			9.0	2.7		3.0	2.2	
ɑ	220	221	221	215	214	215	ɑ	218	220	219	214	214	214
	5.2	5.5		2.5	5.7			2.9	5.2		5.1	5.7	
$\bar{X}_c$	223	225		219	218		$\bar{X}_c$	224	222		217	215	

2nd posttonic

prec. vowel	i in i/ɑ-words, foll. vowel i	foll. vowel ɑ	$\bar{X}_r$	ɑ in i/ɑ-words, foll. vowel i	foll. vowel ɑ	$\bar{X}_r$	prec. vowel	u in u/ɑ-words, foll. vowel u	foll. vowel ɑ	$\bar{X}_r$	ɑ in u/ɑ-words, foll. vowel u	foll. vowel ɑ	$\bar{X}_r$
i	207	206	207	193	192	193	u	212	212	212	185	182	184
	2.9	4.4		6.2	4.7			11.7	8.0		6.5	2.9	
ɑ	201	206	204	193	192	193	ɑ	200	203	202	191	192	192
	4.8	2.9		4.2	5.1			5.0	6.1		6.2	5.1	
$\bar{X}_c$	204	206		193	192		$\bar{X}_c$	206	209		188	187	

Table 2

Mean fundamental frequencies (in Hz) and standard deviations for the vowels i, u, and ɑ in first pretonic, stressed, first posttonic, and second posttonic syllables before and after high vs. low vowels. Speaker SI.

		i in i/ɑ-words			ɑ in i/ɑ-words		
		foll. vowel		$\bar{X}_{r.}$	foll. vowel		$\bar{X}_{r.}$
	prec. vowel	i	ɑ		i	ɑ	
1st pretonic	i	236 6.2	240 9.1	238	234 3.9	226 6.5	230
	ɑ	231 5.8	231 5.8	231	227 7.4	222 3.8	225
	$\bar{X}_{c.}$	234	236		231	224	
stressed	i	255 5.4	248 6.6	252	222 8.1	225 9.0	224
	ɑ	251 5.7	252 9.4	252	224 7.5	227 4.8	226
	$\bar{X}_{c.}$	253	250		223	226	
1st posttonic	i	278 4.8	276 7.6	277	276 18.0	277 11.9	277
	ɑ	271 9.4	274 6.3	273	264 11.3	262 4.4	263
	$\bar{X}_{c.}$	275	275		270	270	
2nd posttonic	i	260 4.0	260 3.7	260	243 4.5	244 5.6	244
	ɑ	253 8.1	258 4.5	256	241 3.9	240 3.2	241
	$\bar{X}_{c.}$	257	259		242	242	

		u in u/ɑ-words			ɑ in u/ɑ-words		
		foll. vowel		$\bar{X}_{r.}$	foll. vowel		$\bar{X}_{r.}$
	prec. vowel	u	ɑ		u	ɑ	
1st pretonic	u	241 7.9	240 3.7	241	238 5.0	226 2.7	232
	ɑ	231 8.6	231 6.1	231	222 7.2	222 3.8	222
	$\bar{X}_{c.}$	236	236		230	224	
stressed	u	258 7.4	256 4.1	257	224 4.5	226 5.6	225
	ɑ	243 4.2	250 6.6	247	222 7.5	227 4.8	225
	$\bar{X}_{c.}$	251	253		223	227	
1st posttonic	u	277 8.7	284 8.1	281	274 13.6	269 7.5	272
	ɑ	266 7.4	264 7.2	265	268 2.5	262 4.4	265
	$\bar{X}_{c.}$	272	274		271	266	
2nd posttonic	u	255 5.8	258 3.3	257	246 4.2	243 5.5	245
	ɑ	253 3.0	254 9.1	254	241 4.1	240 3.2	241
	$\bar{X}_{c.}$	254	256		244	242	

Table 3

Mean fundamental frequencies (in Hz) and standard deviations for the vowels i, u, and ɑ in first pretonic, stressed, first posttonic, and second posttonic syllables before and after high vs. low vowels. Speaker PA.

		i in i/ɑ-words			ɑ in i/ɑ-words			u in u/ɑ-words			ɑ in u/ɑ-words		
		foll. vowel		$\bar{X}_{r.}$	foll. vowel		$\bar{X}_{r.}$	foll. vowel		$\bar{X}_{r.}$	foll. vowel		$\bar{X}_{r.}$
	prec. vowel	i	ɑ		i	ɑ		u	ɑ		u	ɑ	
1st pretonic	i / u	109 3.1	109 1.1	109	105 1.5	104 5.2	105	113 1.4	110 3.3	112	107 2.5	108 2.4	108
	ɑ	111 2.6	110 3.7	111	105 1.6	103 1.4	104	115 4.1	112 2.3	114	109 3.2	103 1.4	106
	$\bar{X}_{c.}$	110	110		105	104		114	111		108	106	
stressed	i / u	121 5.9	117 3.3	119	109 3.2	106 3.1	108	122 3.9	123 5.0	123	109 3.1	107 4.5	108
	ɑ	113 3.2	120 3.7	117	107 7.1	104 4.9	106	116 4.6	121 3.2	119	108 3.8	104 4.9	106
	$\bar{X}_{c.}$	117	119		108	105		119	122		109	106	
1st posttonic	i / u	142 4.8	149 7.4	146	141 3.2	143 3.7	142	144 4.2	153 4.9	149	142 3.2	143 6.9	143
	ɑ	140 3.6	143 3.1	142	138 8.2	136 5.9	137	142 6.0	142 8.1	142	137 5.0	136 5.9	137
	$\bar{X}_{c.}$	141	146		140	140		143	148		140	140	
2nd posttonic	i / u	120 4.6	123 5.4	122	111 6.0	113 8.2	112	124 4.9	124 4.4	124	109 2.2	111 3.5	110
	ɑ	118 3.7	119 5.6	119	114 4.0	112 3.5	113	119 3.9	119 5.2	119	113 3.9	112 3.5	113
	$\bar{X}_{c.}$	119	121		113	113		122	122		111	112	

Table 4

Mean fundamental frequencies (in Hz) and standard deviations for the vowels i, u, and ɑ in first pretonic, stressed, first posttonic, and second posttonic syllables before and after high vs. low vowels. Speaker NR.

i in i/ɑ-words

Syllable	prec. vowel	foll. vowel i	foll. vowel ɑ	$\bar{X}_r$.
1st pretonic	i	89 1.5	87 1.8	88
	ɑ	83 4.3	85 1.6	84
	$\bar{X}_c$.	86	86	
stressed	i	91 1.4	91 2.5	91
	ɑ	89 3.4	88 1.6	89
	$\bar{X}_c$.	90	90	
1st posttonic	i	106 4.5	109 4.3	108
	ɑ	102 3.3	102 5.5	102
	$\bar{X}_c$.	104	106	
2nd posttonic	i	102 3.5	102 2.8	102
	ɑ	100 2.2	100 2.8	100
	$\bar{X}_c$.	101	101	

ɑ in i/ɑ-words

Syllable	prec. vowel	foll vowel i	foll vowel ɑ	$\bar{X}_r$.
1st pretonic	i	83 1.1	84 1.6	84
	ɑ	84 2.4	83 3.0	84
	$\bar{X}_c$.	84	84	
stressed	i	83 3.4	82 2.5	83
	ɑ	83 3.5	82 3.9	83
	$\bar{X}_c$.	83	82	
1st posttonic	i	110 2.8	106 4.0	108
	ɑ	102 2.6	99 3.2	101
	$\bar{X}_c$.	106	103	
2nd posttonic	i	96 3.0	95 6.8	96
	ɑ	96 3.1	96 2.3	96
	$\bar{X}_c$.	96	96	

u in u/ɑ-words

Syllable	prec. vowel	foll vowel u	foll vowel ɑ	$\bar{X}_r$.
1st pretonic	u	89 2.0	89 1.8	89
	ɑ	86 2.0	85 1.3	86
	$\bar{X}_c$.	88	87	
stressed	u	94 2.3	92 1.2	93
	ɑ	89 3.5	89 2.3	89
	$\bar{X}_c$.	92	91	
1st posttonic	u	108 3.8	111 6.0	110
	ɑ	105 5.3	106 6.3	106
	$\bar{X}_c$.	107	109	
2nd posttonic	u	103 4.6	103 6.0	103
	ɑ	98 3.0	99 2.4	99
	$\bar{X}_c$.	101	101	

ɑ in u/ɑ-words

Syllable	prec. vowel	foll vowel u	foll vowel ɑ	$\bar{X}_r$.
1st pretonic	u	86 1.4	85 2.4	86
	ɑ	84 0.6	83 3.0	84
	$\bar{X}_c$.	85	84	
stressed	u	84 3.7	84 3.2	84
	ɑ	83 4.1	82 3.9	83
	$\bar{X}_c$.	84	83	
1st posttonic	u	106 3.3	110 2.1	108
	ɑ	101 3.0	99 3.2	100
	$\bar{X}_c$.	104	105	
2nd posttonic	u	96 3.5	95 5.1	96
	ɑ	96 3.4	96 2.3	96
	$\bar{X}_c$.	96	96	

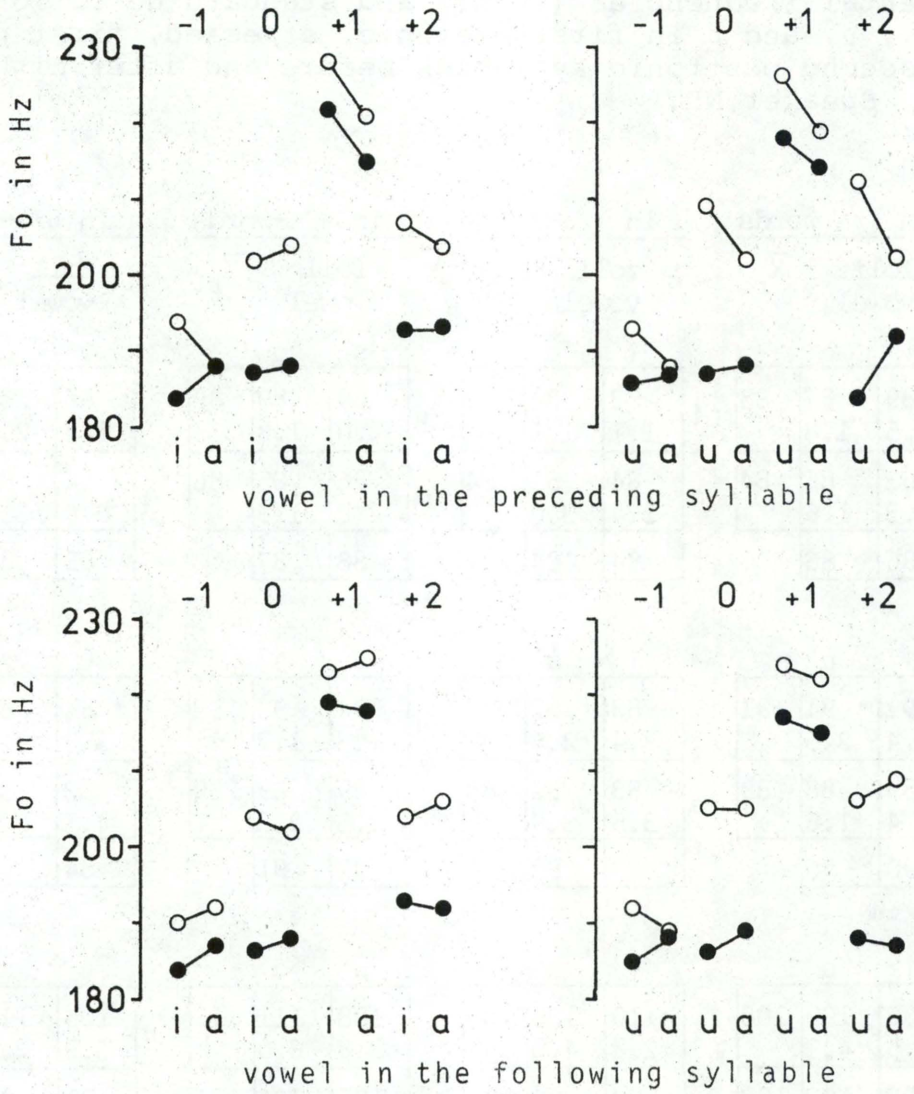


Figure 1

Mean fundamental frequencies in i (o), u (o), and a (•) in first pretonic (-1), stressed (0), first posttonic (+1), and second posttonic (+2) syllables as a function of high and low vowels in the preceding (upper graph) and following (lower graph) syllables. Left: i/a words; right: u/a words. Speaker KM.

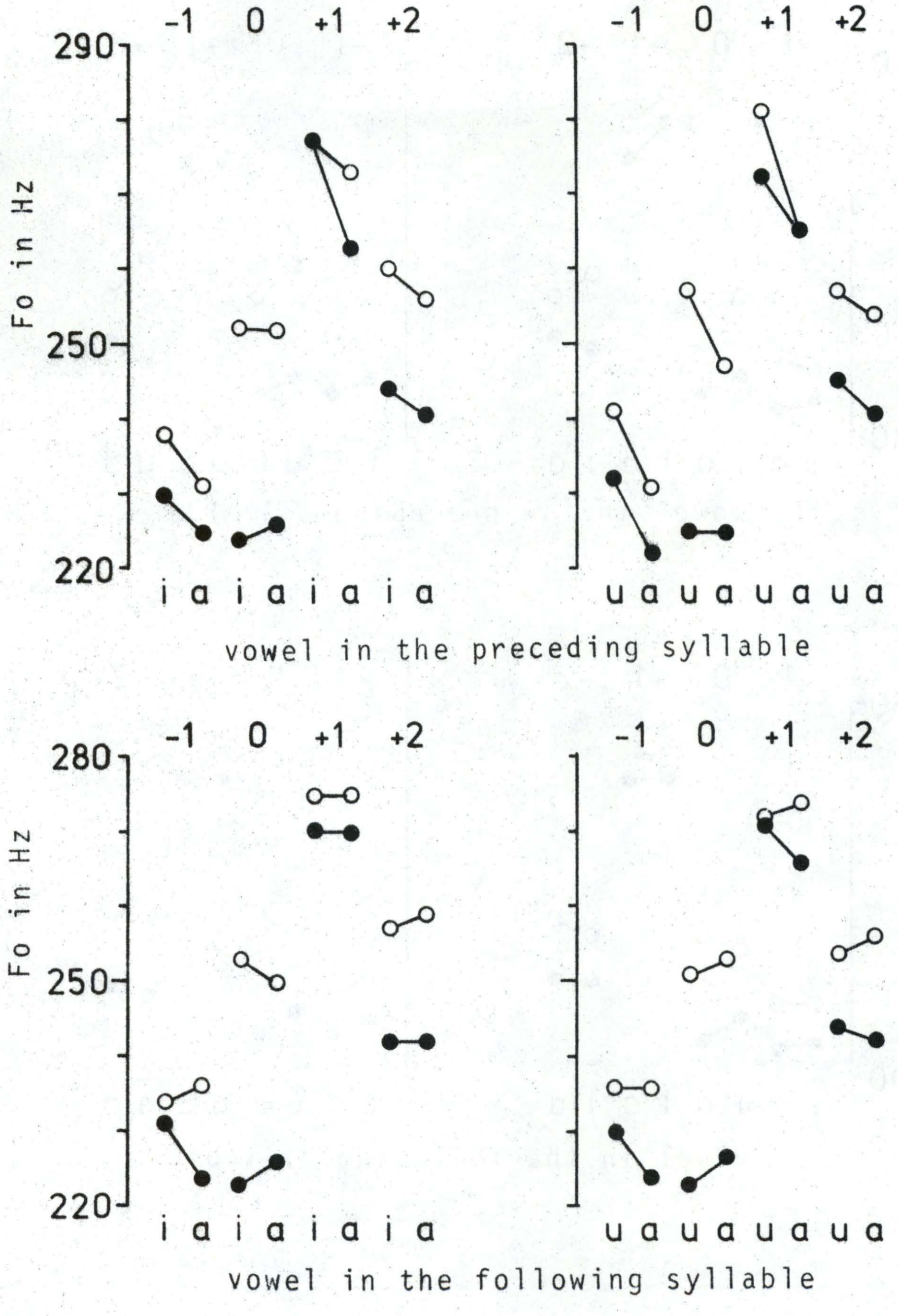


Figure 2

Mean fundamental frequencies in i (o), u (o), and a (•) in first pretonic (-1), stressed (0), first posttonic (+1), and second posttonic (+2) syllables as a function of high and low vowels in preceding (upper graph) and following (lower graph) syllables. Left: i/a words; right: u/a words. Speaker SI.

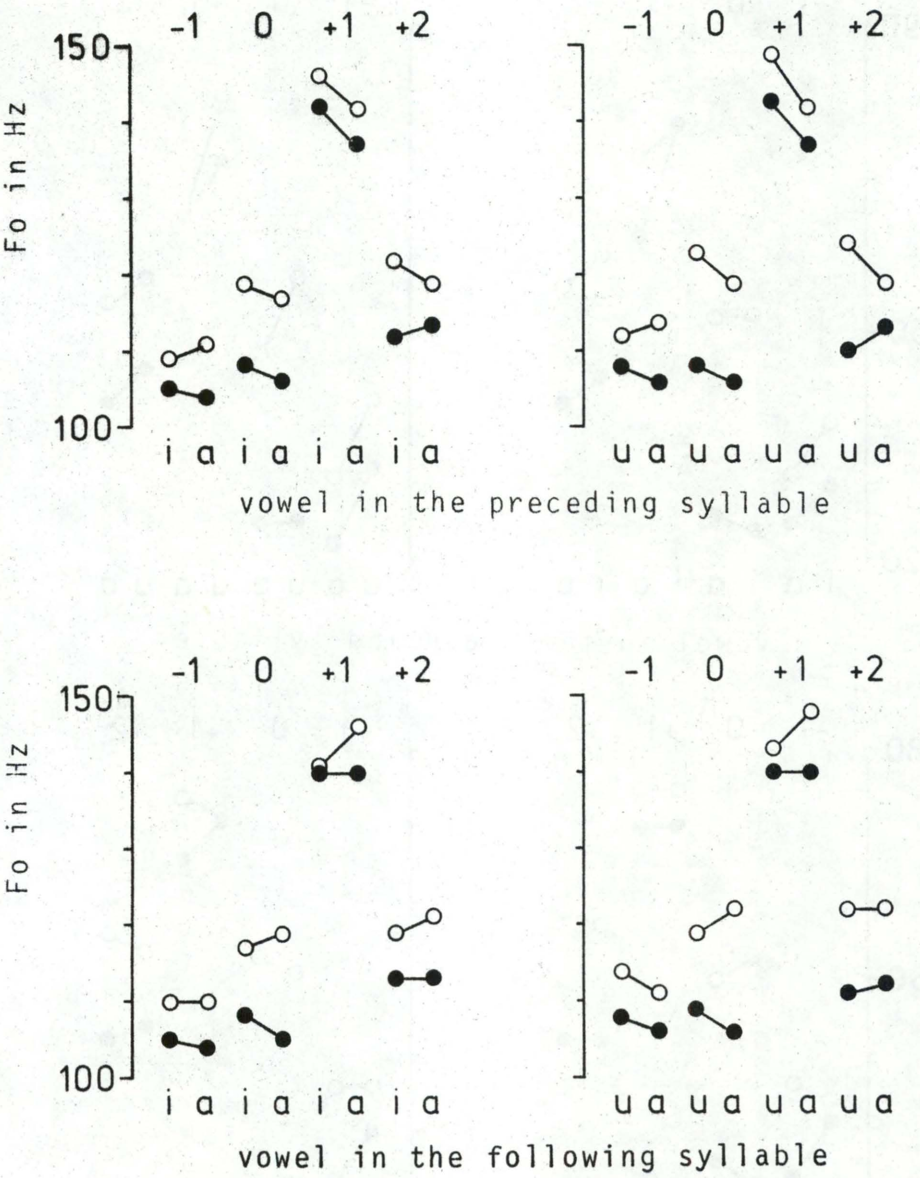


Figure 3

Mean fundamental frequencies in i (o), u (o), and ɑ (•) in first pretonic (-1), stressed (0), first posttonic (+1), and second posttonic (+2) syllables as a function of high and low vowels in the preceding (upper graph) and following (lower graph) syllables. t: Left: i/ɑ words; right: u/ɑ words. Speaker PA.

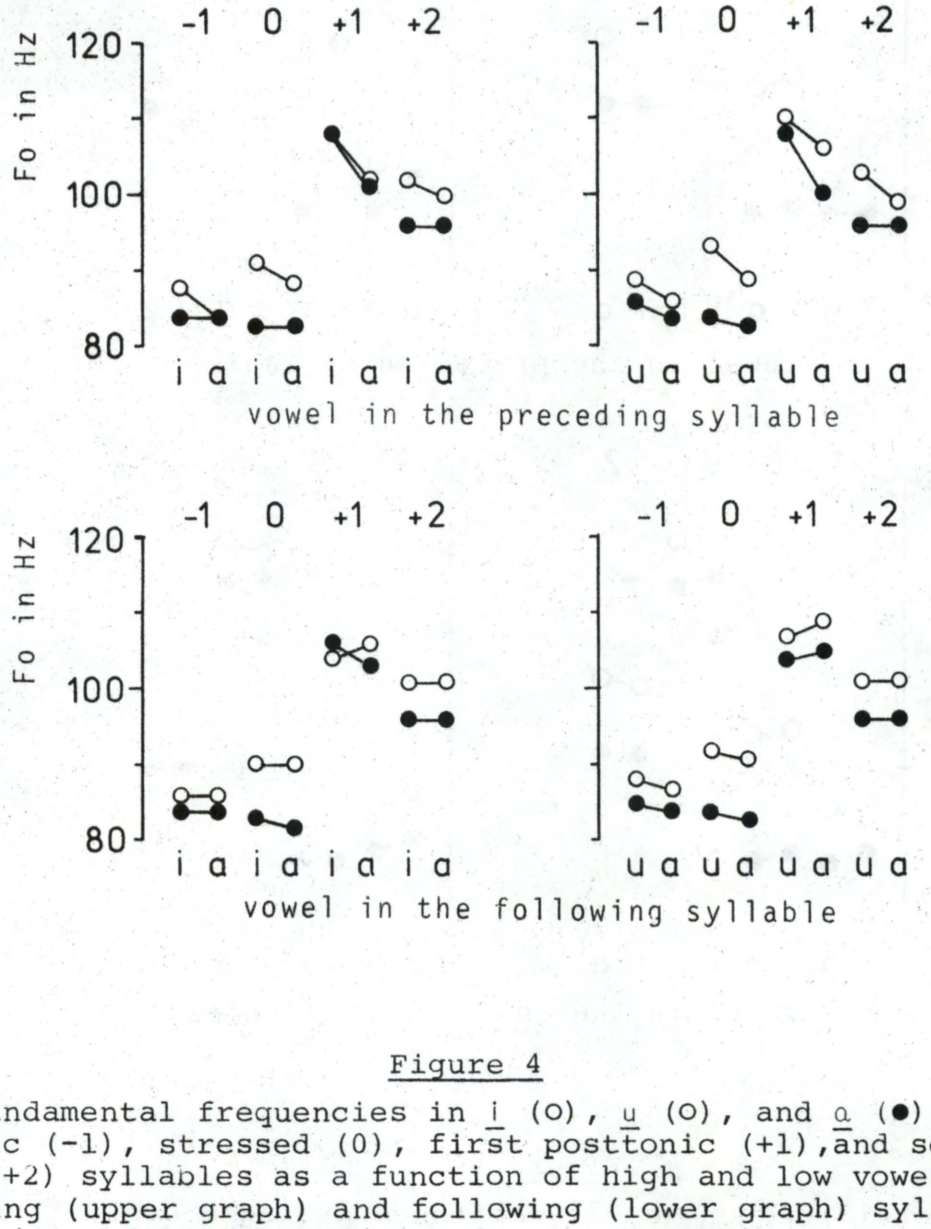


Figure 4

Mean fundamental frequencies in i (O), u (O), and a (●) in first pretonic (-1), stressed (0), first posttonic (+1), and second posttonic (+2) syllables as a function of high and low vowels in the preceding (upper graph) and following (lower graph) syllables. Left: i/a words; right: u/a words. Speaker NR.

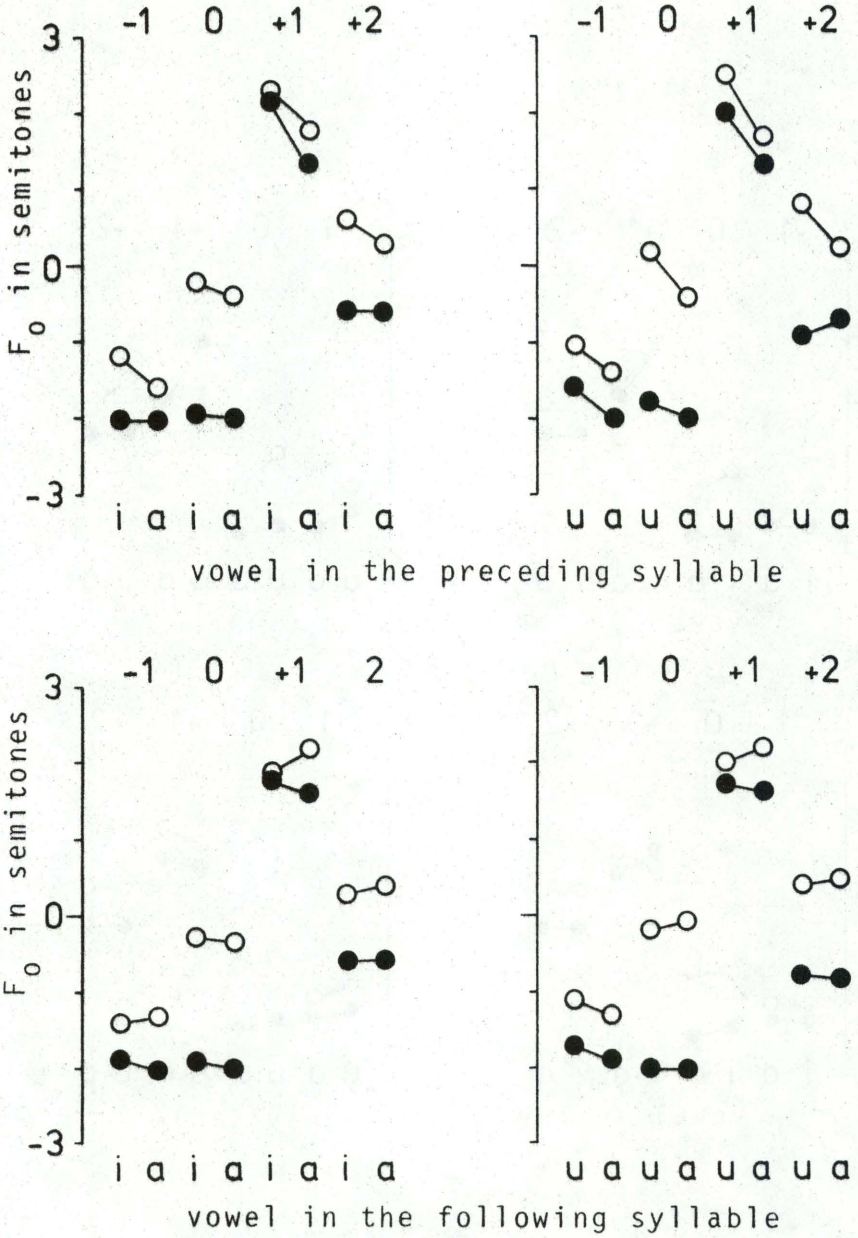


Figure 5

Mean fundamental frequencies in semitones averaged over all speakers in i (o), u (o), and ɑ (●) in first pretonic (-1), stressed (0), first posttonic (+1), and second posttonic (+2) syllables as a function of high and low vowels in the preceding (upper graph) and following (lower graph) syllables. Left: i/ɑ words; right: u/ɑ words.

first posttonic vowels measured occurred earlier in time on the falling intonation contour of the sentence than did the stressed vowels measured, viz. in syllables number two and three, respectively, in the two types of test words.

The data were submitted to a series of two-way analyses of variance (preceding vowel x following vowel). Each of the 2 x 2 matrices of tables 1 to 4 was analyzed separately. The results of the analyses are summarized in table 5.

Table 5

Results of the two-way analyses of variance applied to the data presented in tables 1 to 4. ++ indicates $p < 0.01$, and + indicates $p < 0.05$. Parentheses denote that the effect is significant, but goes in the "wrong" direction, i.e. Fo is higher before or after low vowels than before or after high vowels.

		i/ɑ-words				u/ɑ-words			
		i		ɑ		u		ɑ	
		prec.	foll.	prec.	foll.	prec.	foll.	prec.	foll.
1st pretonic	KM	+				+			
	SI	+		+	+	++		++	++
	PA						+		+
	NR	++				++		+	
stressed	KM					++			
	SI					++			
	PA					+			
	NR	+				++			
1st posttonic	KM	++		++		++			
	SI			+		++			
	PA		(+)	+		+		+	
	NR	++		++	+			++	
2nd posttonic	KM	+				++		(++)	
	SI							+	
	PA					+			
	NR					+			

The data demonstrate clearly that the fundamental frequency of one vowel can be influenced by the quality of neighbouring vowels, but it is equally clear that it is the <u>preceding</u> vowel that exerts the greater and more consistent influence, Fo being in a vast majority of the cases higher after a high vowel than after a low vowel.

The effect of the following vowel is generally considerably smaller and much less consistent. It may be of interest to note that four out of the five cases, where the effect of the following vowel was statistically significant, occurred in the first pretonic syllable, i.e. before a stressed vowel, and across the stress group boundary.

The fundamental frequency of vowels in different positions in the stress group is not influenced to equal degrees by the quality of the preceding vowel. The strongest effect is found in the first posttonic syllable as a result of vowel height variation in the stressed syllable. It is seen from fig. 5 that the effect in the first posttonic approaches one semitone on the average. In the other unstressed syllables the influence is weaker and less consistent. The stressed syllable seems also to be the most resistent with respect to the influence on Fo of the tongue height of the preceding vowel. The vowel u, however, is an exception to this general tendency, Fo in stressed u being strongly affected by the preceding vowel. The behaviour of stressed u and the finding that first pretonic vowels can be influenced by following stressed vowels show that coarticulatory effects on Fo between syllables can occur across stress group boundaries, a fact which weakens the case for the stress group as an autonomous unit in the production of the fundamental frequency course. On the other hand, the effect of stressed vowel on the first pretonic was only statistically significant in four cases out of 16 (cf. table 5 upper row), and - as pointed out above - apart from u, Fo in stressed syllables are not affected by the quality of the preceding vowel; this, of course, can be due to the resistance of the stressed vowel itself, but it may just as well be due to a blocking effect of the preceding stress group boundary.

2.2.1 Fo rises from stressed to first posttonic syllables

In the results given above one point deserves special attention, namely the finding that the strongest effect on the fundamental frequency is that of the stressed vowel on the first posttonic. This is interesting because it might be taken to suggest that the Fo rise from the stressed to the first posttonic syllable, which seems to be an important characteristic of the Fo pattern in the stress group in ASC Danish, tends to remain constant irrespective of vowel heights (and hence Fo levels) in these syllables.

But the above data also suggest that the tendency towards constancy is not complete, in the sense that the inherent fundamental frequency level variation in the stressed syllables is not fully compensated for in the first posttonics. Now, the data presented above do not provide direct information about the Fo rises, since - as pointed out - the organization of the material does not permit intersyllabic comparisons between stressed and first posttonic syllables. Therefore, in order to obtain such information two additional sets of Fo measurements were made, viz. in stressed vowels in words with the stress on the first syllable and in first posttonic vowels in words with the stress on the third syllable. Thus, on the basis of these measurements and the measurements already made in first posttonic vowels and stressed vowels, respectively, two sets of Fo rises from stressed to first posttonic syllables could be derived, both sets containing all possible combinations of i and ɑ and u and ɑ in the two syllables. The rises were computed from raw (i.e. non-normalized) measurements.

The mean Fo rises and standard deviations are given in table 6, and figs. 6 to 7 show the mean Fo values in stressed and first posttonic syllables under the various tongue height conditions. It is seen that Fo rises starting from low stressed vowels are considerably greater than those starting from high stressed vowels, but also that some compensation occurs in the first posttonic vowels, the fundamental frequency in these vowels being higher after high than after low stressed vowels. This is in line with the tendency suggested by the data dealt with above. It also appears from figs. 6 and 7 that the height of the preceding vowel contributes more to the Fo level variation in first posttonic vowels than do the inherent Fo level differences in these vowels; on an average, the effect of high vs. low vowel in the preceding syllable amounts to 4 Hz for KM, 10 Hz for SI, 5 Hz for PA, and 6 Hz for NR, which can be compared to inherent Fo level differences of 4 Hz, 5 Hz, 4 Hz, and 2 Hz, respectively.

Table 7 gives the amount of compensation in pairs of stressed-first posttonic sequences having high vs. low vowels in the stressed syllable and identical vowels in the first posttonic (i.e. ˈi-i/ˈɑ-i, ˈi-ɑ/ˈɑ-ɑ, etc.). The compensation index was derived by dividing the effect on Fo in the first posttonic (i.e. the Fo after high minus Fo after low stressed vowel) by the high-low difference in the stressed vowels.

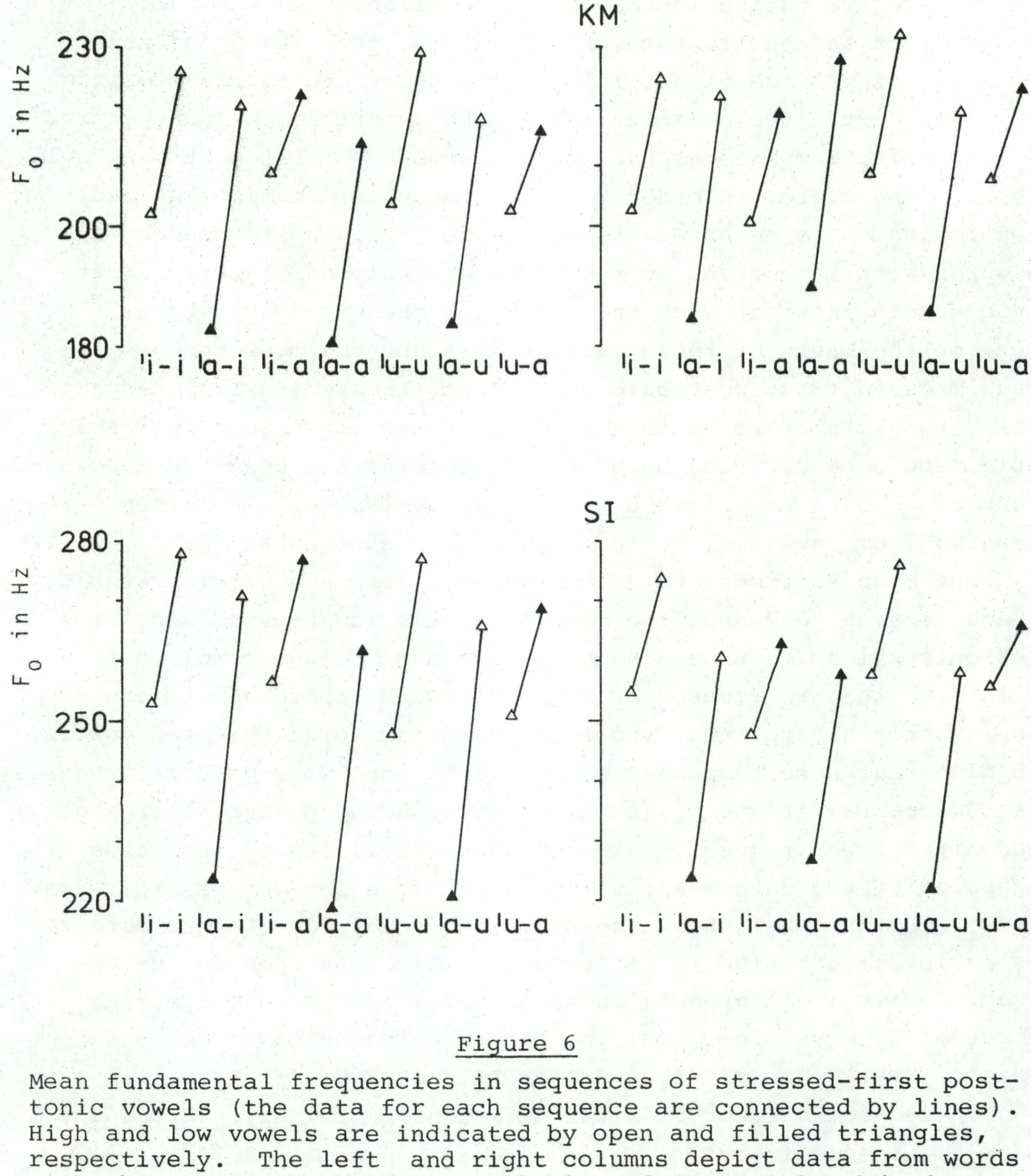


Figure 6

Mean fundamental frequencies in sequences of stressed-first post-tonic vowels (the data for each sequence are connected by lines). High and low vowels are indicated by open and filled triangles, respectively. The left and right columns depict data from words with the stress on the first syllable and from words with the stress on the third syllable, respectively. Speakers KM (upper graph) and SI (lower graph).

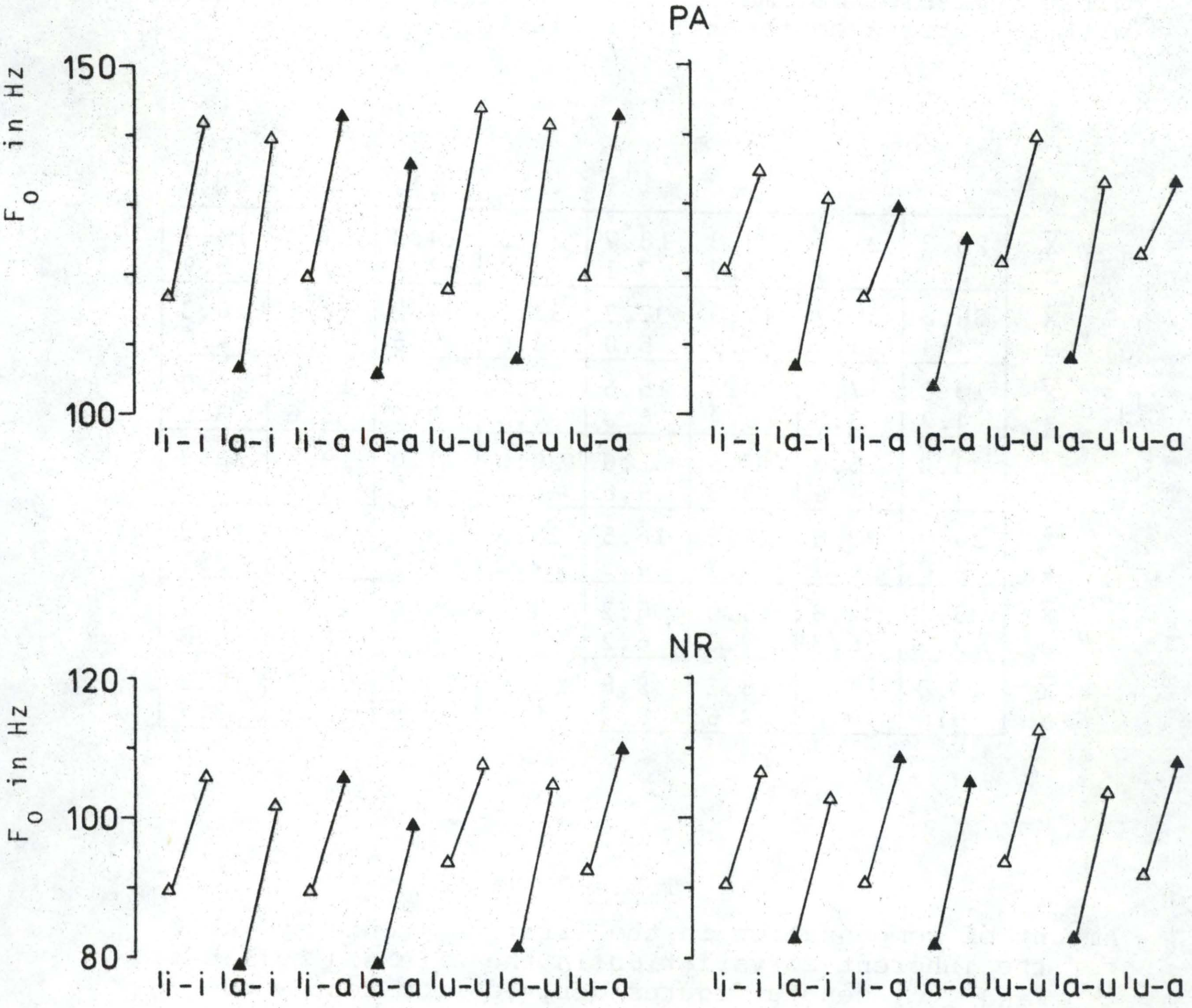


Figure 7

Mean fundamental frequencies in sequences of stressed-first posttonic vowels (the data for each sequence are connected by lines). High and low vowels are indicated by open and filled triangles, respectively. The left and right columns depict data from words with the stress on the first syllable and from words with the stress on the third syllable, respectively. Speakers PA (upper graph) and NR (lower graph).

Table 6

Mean fundamental frequency rises (and standard deviations) from stressed to first posttonic vowels. Columns marked ˈooo list data from words with the stress on the first syllable, columns marked ooˈo list data from words with the stress on the third syllable.

		KM		SI		PA		NR	
		ˈooo	ooˈo	ˈooo	ooˈo	ˈooo	ooˈo	ˈooo	ooˈo
ˈi-i	$\overline{X}$	23.3	21.8	25.8	18.7	24.2	14.3	14.3	15.7
	s	3.0	5.6	4.2	7.1	3.4	7.3	2.7	3.9
ˈa-i	$\overline{X}$	36.8	36.7	47.0	37.7	32.7	24.2	22.5	20.3
	s	4.4	5.5	4.9	5.0	3.0	5.5	5.0	4.5
ˈi-a	$\overline{X}$	13.3	17.5	20.5	15.5	23.2	13.2	15.5	18.0
	s	4.2	3.6	6.4	6.3	4.5	2.6	2.6	3.6
ˈa-a	$\overline{X}$	33.3	38.0	43.5	31.0	30.0	21.0	19.2	23.3
	s	4.0	6.6	3.3	9.6	5.7	3.7	5.1	4.0
ˈu-u	$\overline{X}$	25.0	22.8	29.2	16.5	25.5	18.0	14.2	19.2
	s	4.0	6.7	5.3	4.3	2.2	6.1	3.0	2.1
ˈa-u	$\overline{X}$	33.7	32.8	45.0	36.2	34.0	25.0	22.3	21.2
	s	2.7	6.3	9.3	6.2	3.9	4.3	6.2	5.8
ˈu-a	$\overline{X}$	13.3	14.5	18.2	9.8	23.8	10.8	16.7	16.2
	s	5.0	10.0	2.6	9.2	5.9	2.9	1.5	3.7

Table 7

Amount of compensation in the first posttonic syllable for the inherent Fo variation in the stressed syllable. ˈooo and ooˈo denote figures derived from test words with the stress on the first and third syllable, respectively. Full compensation = 1.0.

		ˈi-i/ˈa-i	ˈi-a/ˈa-a	ˈu-u/ˈa-u	ˈu-a/ˈa-a
KM	ˈooo	0.32	0.29	0.55	0.09
	ooˈo	0.17	-0.82	0.57	-0.28
SI	ˈooo	0.24	0.39	0.41	0.22
	ooˈo	0.42	0.24	0.50	0.28
PA	ˈooo	0.20	0.50	0.20	0.50
	ooˈo	0.29	0.29	0.50	0.42
NR	ˈooo	0.36	0.70	0.25	0.85
	ooˈo	0.50	0.44	0.82	0.30

The degree of compensation varies a great deal, but no systematic trend in the variation can be seen; on the average the compensation amounts to 0.34, i.e. the inherent fundamental frequency variation due to vowel height in stressed syllables is compensated for by approximately one third of the amount in the first posttonics.

Although not entirely within the scope of the present paper, one point concerning the Fo rises should be mentioned, namely the tendency which appears in table 6 for the magnitude of the Fo rises beginning with high vowels to be nearly the same for all speakers, irrespective of their overall fundamental frequency level. This is more clearly illustrated in fig. 8, where the Fo rises from high and low stressed vowels are plotted as a function of the Fo level of the stressed vowels. It is seen that the regression line fitted to the 'i- and 'u-rises is almost constant, at a level of about 18 Hz (slope = 0.0137), throughout the fundamental frequency range covered by the four speakers. The magnitude of the Fo rises beginning with the vowel ɑ, on the other hand, show an increasing tendency through the Fo range (slope = 0.1234). The difference between the slopes is, of course, due to the fact that the inherent fundamental frequency differences increase with the general Fo level (as was also found by Reinholt Petersen, 1978), and that these differences are only partly compensated for in first posttonic vowels. The vertical dispersion of the data points within each category (rises from i and u and rises from ɑ) can - at least in part - be ascribed to inherent Fo differences between vowels in first posttonic syllables, rises to high vowels tending to be greater than rises to low vowels. The tendencies lined out here are interesting and deserve further experimentation, intended to see whether they hold in a larger sample of speakers, and to examine the mechanisms underlying them.

2.3 Discussion

In section 1 above it was suggested that the inherent Fo level differences between vowels could distort the intended Fo pattern of stress groups in such a manner that the stress distribution perceived by the listener would differ from that intended by the speaker. Two possibilities were hypothesized which might ensure the correct transmission of stress distributions from speaker to listener, namely (1) that the perceptual system compensates for

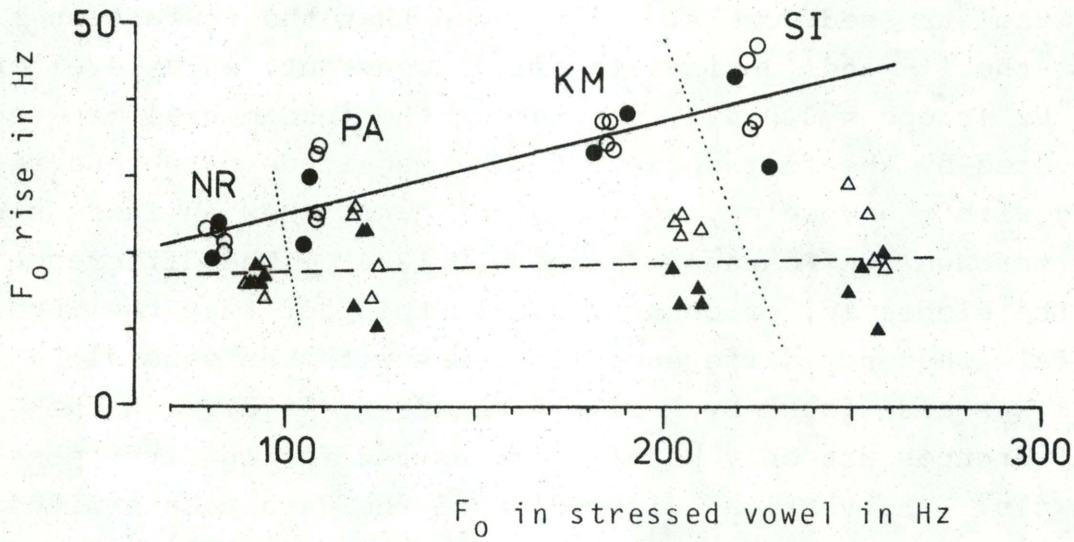


Figure 8

Mean F_0 rises from stressed to first posttonic vowels plotted as a function of F_0 in the stressed syllables. ˈi-i and ˈu-u rises are indicated by △, ˈi-ɑ and ˈu-ɑ rises by ▲, ˈɑ-i and ˈɑ-u rises by ○, and ˈɑ-ɑ rises by ●. The full and broken lines are regression lines fitted to the data points representing rises starting with low stressed and with high stressed vowels, respectively. The dotted lines separate the data for NR and PA and the data for KM and SI.

the inherent Fo variation and thus restores the intended Fo course, or (2) that the compensation takes place in the speech production system in the shape of coarticulation of inherent Fo level across syllable boundaries so that the linguistically irrelevant Fo variation is smoothed out and the intended Fo pattern preserved.

The above experiment has shown that the inherent Fo level differences are to some degree smoothed out as a result of carry-over coarticulation of Fo from one syllable to the next. This is particularly true of the effect of stressed vowels on Fo in first posttonic vowels. On the other hand, in the majority of cases in the present material the Fo rises from stressed to first posttonic syllables seem to be sufficiently great in comparison with the inherent Fo level differences between high and low vowels for them to be preserved even if no coarticulatory compensation for the inherent differences had occurred. But there are a few cases of stressed i or u followed by first posttonic ɑ, where the compensation effect is almost entirely responsible for the rise (see e.g. KM's ˈi-ɑ vs. ˈɑ-ɑ (left column of fig. 6), SI's ˈu-ɑ vs. ˈɑ-ɑ (right column of fig. 6), and PA's ˈu-ɑ vs. ˈɑ-ɑ (right column of fig. 7)). Even in such cases, however, it is a question whether the observed effect can be assumed to be a necessary prerequisite for - or to facilitate - the perception of stress. A further discussion of this matter will, however, be postponed till section 4 below, where the results of Experiment II will also be taken into consideration. The present discussion will focus upon the mechanisms responsible for the coarticulation of inherent Fo levels across syllable boundaries. More specifically the question will be considered whether it is tongue height that is coarticulated or it is Fo (or rather the laryngeal conditions which produce the inherent Fo level differences between vowels).

Before proceeding any further, it may be appropriate to review briefly the current hypotheses advanced to explain the inherent Fo level differences.[1] One hypothesis rests on the assumption that the acoustic impedance in the glottis is sufficiently low, relative

---

1) The following review is neither exhaustive nor does it attempt to evaluate the hypotheses sketched. Critical and more detailed reviews are given in e.g. Ewan (1979), Ohala (1973), and Reinholt Petersen (1978).

to the comparatively high input impedance to the vocal tract at low first formant frequencies, for some coupling to occur between the vocal tract and the glottis (e.g. Flanagan and Landgraf, 1968, Lieberman, 1970). If this is so, the first resonance of the vocal tract will cause a change, i.e. an increase of the fundamental frequency. Since such an effect becomes greater the lower the first formant, i.e. the closer the first resonance is to the Fo range, vowels with low first formants (e.g. i and u) will have higher fundamental frequency than vowels with high first formants (e.g. ɑ).

Another hypothesis suggests that the tongue, when elevated for the production of high vowels, pulls the hyoid bone and the larynx upwards (Ladefoged, 1964, Lehiste, 1970). This vertical pull is then thought to be translated into an increased longitudinal tension of the vocal cords, which in turn leads to a higher fundamental frequency.

A third hypothesis, advanced by Ohala (1973) and elaborated upon by Ewan (1975, 1979), is also based on the pull of the tongue on the laryngeal structures. But according to this hypothesis the increased tongue pull in high vowels gives rise to an increased vertical tension in the vocal cords, which is established through the mucous membrane and other soft tissues without involving the hyoid bone and the hard tissues of the larynx. Ewan (1975, 1979) has proposed a modification of the hypothesis, emphasizing not so much the pull as the retraction of the tongue and the constriction of the pharynx in low vowels, which increase the vibrating mass and decrease the vertical tension of the vocal cords, with a lower Fo as the result.

Now, in the present material the Fo variation in first posttonic syllables attributable to the variation in tongue height of the preceding (stressed) vowel approaches one semitone. If an effect of that magnitude were to be explained by coarticulatory assimilation of tongue height, a radical quality shift should be expected in first posttonic vowels. This applies no matter whether the inherent Fo level differences between vowels are to be accounted for by the acoustic source/tract coupling hypothesis or by the physiologically based tongue pull hypothesis. Listening to the recorded material did not, however, reveal any such changes of vowel quality. In order to obtain quantitative data to illuminate this point, the frequencies of the first and second formants in ɑ in first posttonic syllables were measured in the recordings of the two male speakers PA and NR. The mean formant frequencies are plotted in fig. 9 as a function of the vowels in the preceding and following syllables, respectively. The formant data were submitted to a series of two-way analyses of variance (preceding vowel x following vowel), the results of which are summarized in table 8.

Table 8

Results of the two-way analyses of variance applied to the formant data for the vowel ɑ depicted in fig. 9. ++ indicates $p < 0.01$, and + indicates $p < 0.05$.

	first formant				second formant			
	i/ɑ-words		u/ɑ-words		i/ɑ-words		u/ɑ-words	
	prec. vowel	foll. vowel	prec. vowel	foll. vowel	prec. vowel	foll. vowel	prec. vowel	foll. vowel
PA			+		++	++	++	
NR		+			++	+	++	++

It is seen (as could be expected on the basis of the auditory evaluation) that the formant frequencies of first posttonic syllables are only moderately influenced by the adjacent vowels; and the very small - and in most cases non-significant - effect observed in the first formant precludes the possibility of the influence on Fo from the preceding vowel to be accounted for by acoustic coupling between the vocal tract and the glottis. Furthermore, to the extent that the formant frequencies are influenced by the vowels in adjacent syllables (this applies to the second formant in particular), they are influenced by both preceding and following vowels, whereas Fo is influenced mainly by preceding vowels.

Thus, the formant data reveal a discrepancy between the effect of neighbouring vowels on the fundamental frequency and the effect of neighbouring vowels on the position of the tongue body, both as regards the magnitude and the direction of the effects. This interpretation of the formant data presupposes, of course, that the formant frequency changes observed in the present material can be ascribed to tongue body position changes. This can be assumed to be the case with reasonable certainty in the i/ɑ words, but not,

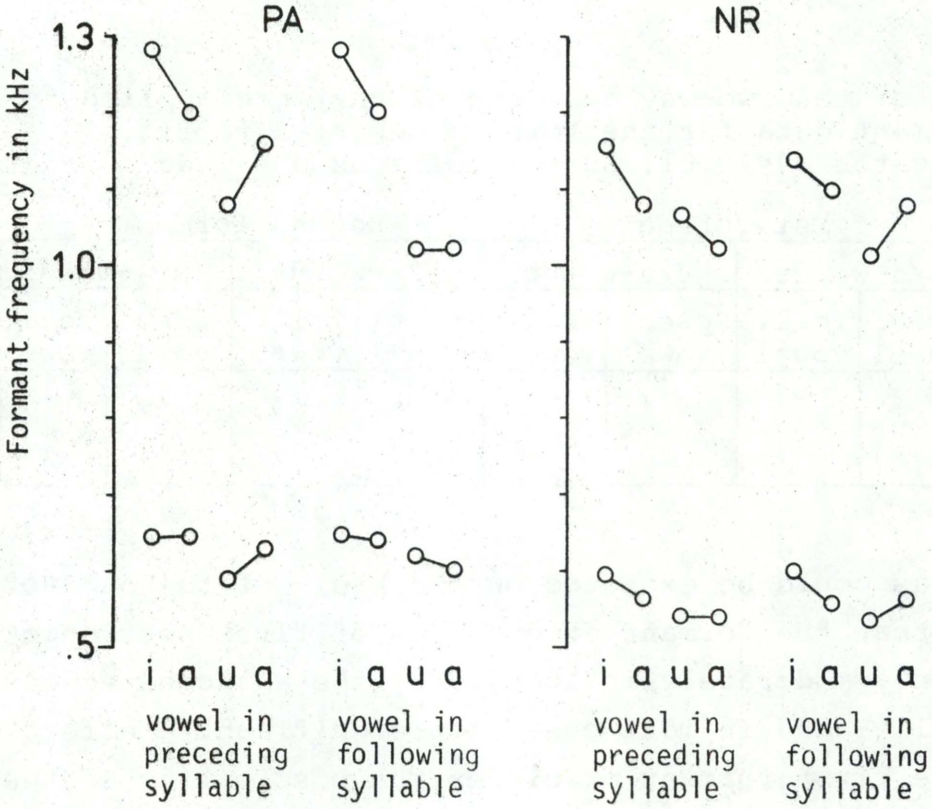


Figure 9

Mean frequencies of formant 1 and formant 2 in first posttonic a as a function of preceding and following high and low vowels. Speakers PA (left) and NR (right).

perhaps, in the u/ɑ words, where lip rounding is also involved.

The discrepancy can be explained under the tongue pull hypothesis - but not under the source/tract coupling hypothesis - if the relation between tongue height and the laryngeal conditions responsible for inherent Fo level differences between vowels can be described in terms of a damped spring-mass system, where the driving force is the tongue body, the spring represents the connection between the tongue body and the laryngeal structures, and the mass represents the tension and/or mass of the vocal cords. If this description is tenable, changes of the laryngeal conditions, and hence of the fundamental frequency, will occur with a delay relative to the changes in tongue height which produce them.

3. Experiment II

Experiment II was primarily intended to examine, as a supplement to Experiment I, whether the effect of a high vs. a low vowel in one syllable on the fundamental frequency in the following one would also show up if the consonant between the two syllables was voiceless instead of voiced as in Experiment I, but since it would imply only a limited extension of the material it was decided also to consider the question whether a segmentally determined Fo level variation in one syllable conditioned not by vowel height but by the initial consonant could produce an effect on Fo in the following syllable similar to that produced by a high vs. a low vowel.

3.1 Method

3.1.1 Material

Since the results of Experiment I indicated that Fo in one syllable is influenced almost exclusively by tongue height differences in the preceding syllable, and since the effect turned out most markedly in the first posttonic, it was thought appropriate to make up the test material so as to focus on the influence on Fo in the first posttonic from the segmental variation in the stressed syllable. The present material was limited to include only the vowels i and ɑ, and the consonants chosen were the labials m and f.

The test material consisted of all possible ˈCVCV combinations of the two consonants and the two vowels. The test words were embedded in carrier sentences similar to those of Experiment I.

3.1.2 Recordings and speakers

The sixteen test sentences and eight distractor sentences were arranged in six different random orders in a reading list, i.e. six repetitions of each word were obtained. The recordings were made under the same conditions as in Experiment I, and with the same subjects (see section 2.1.2 above). Since the list was comparatively easy to read all six randomizations were recorded in one session.

3.1.3 Registrations and measurements

The following acoustic curves were made: duplex oscillogram, two intensity curves (HiFi linear and HP-filtered at 500 Hz, logarithmic, integration time 2.5 ms for both), and a fundamental frequency curve. The fundamental frequency of the stressed and first posttonic vowels were measured at a point in time two thirds from vowel onset.

3.2 Results

Since all six randomizations of the test words were read in one session it was not considered necessary to carry out a normalization of the measurements as was done in Experiment I, and inspection of the variability of the present data did not reveal any systematic deviation from the variability of the normalized data of Experiment I.

3.2.1 Effect of initial f vs. m on Fo in the vowel

A prerequisite for considering at all whether f vs. m and i vs. ɑ in the stressed syllable have similar effects on the fundamental frequency in the first posttonic syllable was that the intrinsic Fo differences produced in the stressed syllable by the two different types of segmental variation (f vs. m and i vs. ɑ) were of the same order of magnitude. In order to see whether this was the case, means and standard deviations for Fo in the vowels of all mi, fi, mɑ, and fɑ syllables in stressed position were computed, and a series of t-tests were applied to test the statistical significance of the differences between means in the vowels following f and m. Table 9 gives the means, standard deviations, differences between means, and levels of significance achieved. The means are displayed graphically in fig. 10 in Hz for the individual subjects and in semitones averaged over the subjects.

It is seen that there is a clear tendency for Fo to be higher after f than after m, and the effect was significant at a very high level in all cases. Furthermore, the effects of initial consonant and of vowel height seem to be of comparable magnitudes, although the former effect is slightly smaller than the latter. On the average the Fo difference between i and ɑ amounts to 1.3 semitones compared to a difference of 0.9 semitones between vowels following f and m. The initial consonant also influences Fo in first posttonic vowels, but to a lesser degree than in stressed syllables. The difference in the posttonic vowels after m vs. f were 0.6 semitones on the average. The Fo intrinsic level difference between i and ɑ in first posttonic vowels was 0.5 semitones on the average.

Table 9

Mean fundamental frequencies and standard deviations in the stressed vowels i and ɑ and differences between means after initial f and m. The levels of significance of the differences between f and m are indicated by +++ for $p < 0.001$ and ++ for $p < 0.01$.

			f-	m-	f-m
KM	-i	$\overline{X}$	194.5	186.7	7.8+++
		s	5.595	5.247	
	-ɑ	$\overline{X}$	184.2	175.6	8.6+++
		s	3.978	3.717	
SI	-i	$\overline{X}$	229.8	218.0	11.8+++
		s	5.838	3.989	
	-ɑ	$\overline{X}$	210.5	203.0	7.5+++
		s	6.454	5.668	
PA	-i	$\overline{X}$	111.3	106.2	5.1++
		s	7.910	3.293	
	-ɑ	$\overline{X}$	105.0	100.4	4.6+++
		s	3.257	2.618	
NR	-i	$\overline{X}$	102.0	95.5	6.5+++
		s	2.136	1.351	
	-ɑ	$\overline{X}$	92.4	88.0	4.4+++
		s	1.974	2.828	

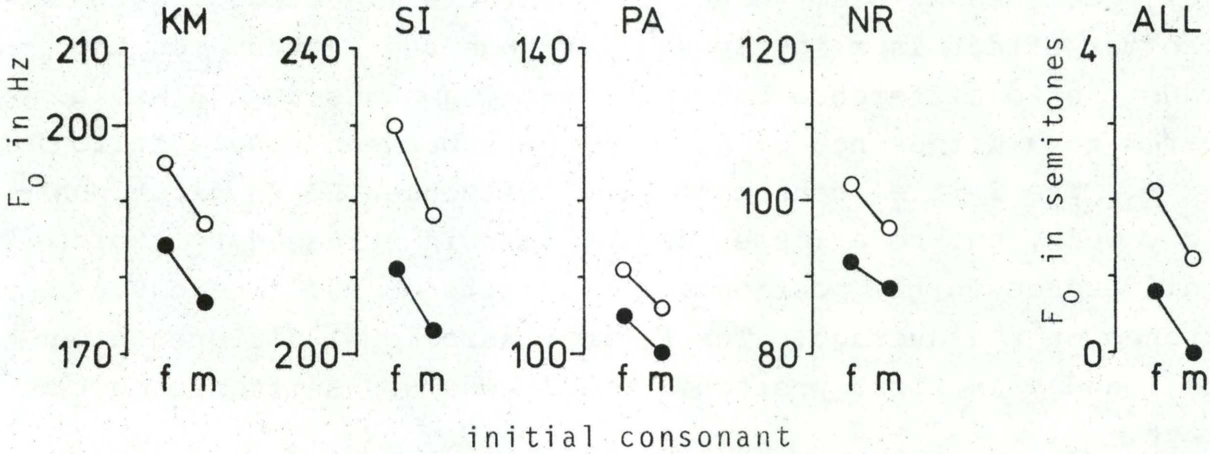


Figure 10

Mean fundamental frequencies in stressed i (o) and ɑ (●) as a function of initial f and m for the four subjects individually and averaged over all subjects (in semitones).

3.2.2 The effects of i vs. ɑ and f vs. m initially in the stressed syllable on Fo in the first posttonic

Table 10 presents the means and standard deviations of the Fo measurements in first posttonic fi, fɑ, mi, and mɑ syllables after i vs. ɑ and initial f vs. m in the preceding, stressed syllable. The means are displayed graphically in fig. 11 for the subjects individually in Hz, and averaged over the subjects in semitones. The data were submitted to a series of two-way analyses of variance (tongue height x consonant type in the stressed syllable). Each of the 2x2 matrices in table 10 were analysed separately. The results of the analyses of variance are given in table 11.

Before further considering the results, it should be noted that the data for one of the subjects, PA, seem to show a rather atypical pattern. From fig. 11 it is seen that in 3 out of the 4 posttonic syllables examined, PA has a higher Fo after ɑ than after i in the preceding syllable. This is in disagreement with what is found for the other subjects in the present experiment and also with PA's results in Experiment I (see fig. 3 above). In view of this, the evaluation of the data will be based mainly on subjects KM, SI, and NR.

The data show - like those of Experiment I - that the tongue height of the vowel in the stressed syllable influences the fundamental frequency of the first posttonic vowel. Furthermore, it is evident that this effect exists irrespective of whether the intervening consonant is voiced or voiceless.

In contradistinction to the effect of tongue height, Fo variation due to the initial consonant in the stressed syllable has no effect on the fundamental frequency of the first posttonic vowel. There is a slight, but not consistent, tendency for Fo in that vowel to be higher after an f than after an m initially in the stressed syllable, but the effect could be proved statistically significant only in one case (se table 11) out of 12 (leaving PA out), which can be compared to 9 out of 12 cases of significant tongue height effects.

Table 10

Mean fundamental frequencies (in Hz) and standard deviations in first posttonic -fi, -fa, -mi, and -ma syllables after i vs. a and f vs. m in the preceding (stressed) syllables.

		-fi: initial cons. in prec. syl.			-fa: initial cons. in prec. syl.			-mi: initial cons. in prec. syl.			-ma: initial cons. in prec. syl.		
		f	m	$\bar{X}_{r.}$	f	m	$\bar{X}_{r.}$	f	m	$\bar{X}_{r.}$	f	m	$\bar{X}_{r.}$
KM	prec. vowel i	202 / 6.2	203 / 2.0	203	201 / 6.0	198 / 4.7	200	205 / 3.5	200 / 4.3	203	199 / 3.1	199 / 2.3	199
	prec. vowel a	203 / 3.6	203 / 1.9	203	196 / 4.9	198 / 3.4	197	198 / 5.2	199 / 5.4	199	197 / 3.0	193 / 3.6	195
	$\bar{X}_{c.}$	203	203		199	198		202	200		198	196	
SI	prec. vowel i	243 / 3.8	239 / 5.2	241	239 / 4.8	231 / 5.3	235	234 / 4.8	234 / 7.5	234	230 / 5.2	230 / 4.3	230
	prec. vowel a	238 / 2.5	237 / 2.3	238	226 / 5.1	230 / 4.8	228	229 / 3.8	229 / 5.8	229	224 / 2.2	223 / 2.2	224
	$\bar{X}_{c.}$	241	238		233	231		232	231		227	227	
PA	prec. vowel i	133 / 5.9	125 / 3.3	129	129 / 15.3	123 / 3.4	126	123 / 6.9	124 / 4.3	124	117 / 6.0	124 / 7.9	121
	prec. vowel a	128 / 7.1	131 / 5.1	130	121 / 4.1	123 / 4.8	122	126 / 8.2	124 / 5.3	125	125 / 5.6	118 / 5.7	122
	$\bar{X}_{c.}$	131	128		125	123		125	124		121	121	
NR	prec. vowel i	111 / 2.2	114 / 2.4	113	112 / 2.3	111 / 2.2	112	106 / 2.5	106 / 2.5	106	104 / 3.4	104 / 2.8	104
	prec. vowel a	111 / 2.3	109 / 0.8	110	106 / 3.4	108 / 2.6	107	102 / 4.7	101 / 3.4	102	100 / 3.4	99 / 3.5	100
	$\bar{X}_{c.}$	111	112		109	110		104	104		102	102	

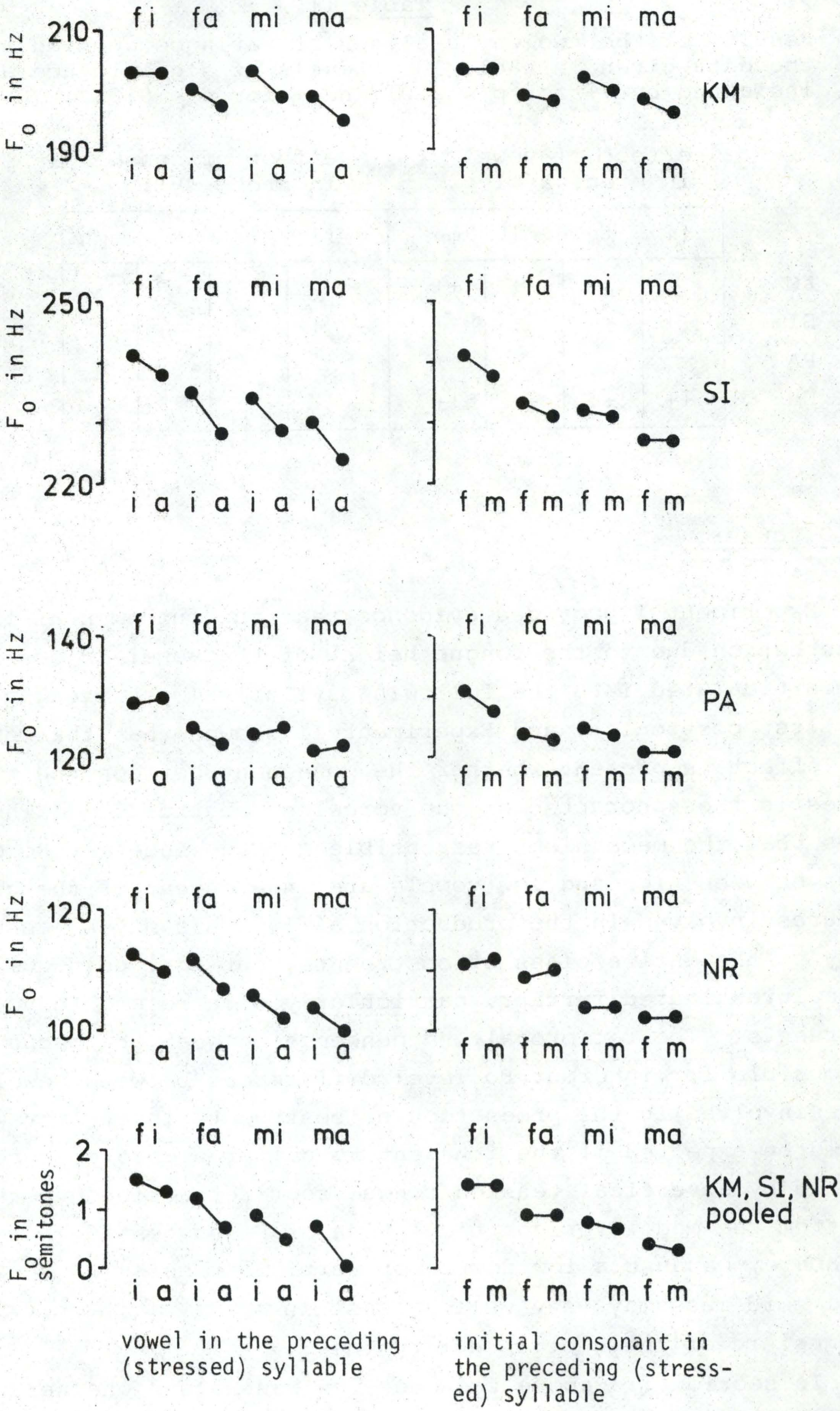


Figure 11

Mean fundamental frequencies in the vowels in first posttonic -fi, -fa, -mi, and -ma syllables as a function of i and a in the preceding syllable (left column) and initial f and m in the preceding syllable (right coulumn), for the speakers individually and averaged over the subjects KM, SI, and NR (in semitones).

Table 11

Results of the two-way analyses of variance applied to the data given in table 10. Levels of significance are indicated by ++ for $p < 0.01$ and + for $p < 0.05$.

	effect of vowel in prec. syll.				effect of cons. in prec. syll.			
	fi	fa	mi	ma	fi	fa	mi	ma
KM			+	++				
SI	+	++		++	+			
PA								
NR	++	++	++	++				

4. Discussion

Experiment I provided evidence that the fundamental frequency perturbation due to the tongue height of the vowel in one syllable is coarticulated into the following syllable (most clearly so in the first posttonic), and Experiment II showed that this carry-over effect is present whether the consonant between the two syllables is the sonorant m or the voiceless obstruent f. Thus, it seems that the mechanisms responsible for the inherent Fo differences between high and low vowels are independent of the glottal gestures involved in the production of f. This may be assumed to apply to the entire class of obstruents, but that point will have to be investigated further, particularly with regard to voiced obstruents. The reciprocal independence between the mechanisms responsible for inherent Fo level differences between vowels and those involved in the production of obstruents is in fact what should be expected if the inherent Fo differences are to be explained by a vertical tension/tongue root retraction hypothesis and from the model for coarticulation of Fo suggested in section 2.3 above, because a low degree of vertical tension or an increased vocal cord mass may very well persist in spite of the opening-closing gesture of the glottis during the f.

In section 2.3 above the question was raised whether the coarticulation of inherent Fo levels between syllables is a necessary prerequisite for the correct perception of stress patterns. The results of the experiments reported in the present paper can hardly be assumed to give an affirmative answer to that question. First, the compensation in one syllable for the Fo variation ascribable to tongue height found in Experiment I is only partial (even in first posttonic syllables, where it is greatest), and secondly - and more important, perhaps - the effect on Fo of the initial consonant in the stressed syllable is not (although of a magnitude comparable to the effect of tongue height in that syllable) carried over to the first posttonic. Thus, the perceptual system seems to be able to do without the smoothing out of (i.e. to compensate for) the segmentally determined fundamental frequency variation. This view is corroborated by the finding in Experiment II of cases in words of the type ˈfimɑ, where the rise from the stressed to the first posttonic vowel is extremely small or even negative. (The rises range from -2 to 8 Hz for KM, from -1 to 8 Hz for SI, from 4 to 16 Hz for PA, and from -3 to 4 Hz for NR in such words.) Furthermore, Thorsen (1980b) reports that in stress groups with only one unstressed syllable there may be (but is not always) no or a reduced Fo rise from the stressed to the unstressed syllable.

But, of course, this reasoning will have to be supplemented by perceptual experimentation before the question can be answered with reasonable certainty. In particular, experiments are called for which take into consideration the role of segment duration - and its interaction with the fundamental frequency - in the perception of stress patterns.

References

Ewan, W.G. 1975: "Explaining the intrinsic pitch of vowels", Paper presented at the fifth Linguistics Association conference, San José, May 4, 1975, p. 1-9

Ewan, W.G. 1979: "Laryngeal behavior in speech", Report of the Phonology Laboratory 3, University of California, Berkeley, p. 1-93

Flanagan, J.L. and L. Landgraf 1968: "Self-oscillating source for vocal tract synthesizers", IEEE Transactions on Audio and Electroacoustics AU-16, p. 57-64

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

Ladefoged, P. 1964: A Phonetic Study of West African Languages: an auditory-instrumental survey, Cambridge

Lehiste, I. 1970: Suprasegmentals, Cambridge, Ms.

Lieberman, P. 1970: "A study of prosodic features", Haskins SR 23, p. 179-208

Ohala, J.J. 1973: "Explanations for the intrinsic pitch of vowels", Monthly Internal Memorandum, Phonology Laboratory, University of California, Berkeley, p. 9-26

Petersen, N. Reinholt 1974: "The influence of tongue height on the perception of vowel duration in Danish", ARIPUC 8, p. 1-10

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

Petersen, N. Reinholt 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

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

Siegel, S. 1956: Nonparametric Statistics for the Behavioral Sciences, Tokyo

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

Thorsen, N. 1980a: "Word boundaries and Fo patterns in Advanced Standard Copenhagen Danish", Phonetica 37, p. 121-130

Thorsen, N. 1980b: "Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish", ARIPUC 14, p. 121-205