ARIPUC VOL. 17.

ISSN 0589-6681

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

COPENHAGEN 1983

ARIPUC VOL. 17.

ISSN 0589-6681

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

96, Njalsgade
DK 2300 Copenhagen

COPENHAGEN 1983

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

KØBENHAVNS UNIVERSITET
REPRO. AFDELINGEN

CONTENTS

Editorial note

Due to a change in grant policy, future ARIPUCs will be published in the spring. The next issue (vol. 18) is thus due already in the spring of 1984.

ARIPUC 17, 1983 p. 1-20

TWO EXPERIMENTS IN THE PERCEPTION OF FØ TIMING IN DANISH

NINA THORSEN

The perceptual relevance of rather finely timed differences in the occurrence of fundamental frequency (Fø) rises as observed in natural speech is established in listening experiments with synthetic speech stimulus material. (1) When vowel and consonant duration cues are ambiguous between /ku·lə/ and /kulə/ an early Fø rise, relative to the vowel-consonant boundary, will tend to favour identification of stimuli as /ku·lə/, while a later Fø rise increases the number of /kulə/ judgments somewhat. (2) When vowel duration cues are ambiguous between /'bilisd/ and /bi'lisd/ an Fø rise before the intervocalic consonant will significantly increase the number of /'bilisd/ identifications, whereas an Fø rise after the consonant yields more judgments of /bi'lisd/. The latter result corroborates the acoustic observation that initial voiced consonants in stressed syllables dissociate tonally from the vowel and join up with the preceding material, if any, to the effect that tonally the stress seems to begin with the vowel.

I. INTRODUCTION

A. VOWEL LENGTH AND FØ TIMING

On previous occasions (e.g. Thorsen 1982a and 1982b) I have suggested a description of the Fø pattern characterizing prosodic stress groups in Danish in terms of an essentially invariant Fø wave upon which the segments and syllables are superposed. This will make intrasyllabic Fø movements predictable from the shape of the wave where the syllables hit it (falling, rising, etc.). Any Fø movement may then of course be modified

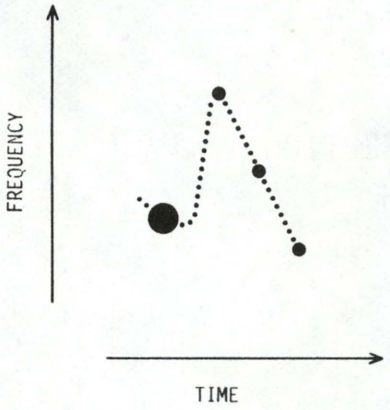


Figure 1

Stylized fundamental frequency pattern associated with the prosodic stress group in Standard Danish. The heavy dot denotes the stressed vowel, the light dots the unstressed syllables. The slope of the fall through the posttonics is subject to individual variation.

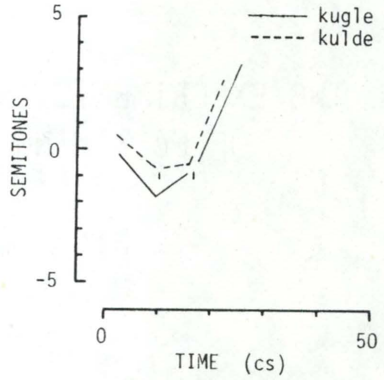


Figure 2

Fundamental frequency tracings (averages over 5 recordings) of two utterance medial words kugle and kulde [g̊ʰu·l̇ g̊ʰull̩]. *The vowel-consonant boundary is indicated with a vertical stroke. Zero on the logarithmic frequency scale corresponds to 100 Hz. Male speaker.*

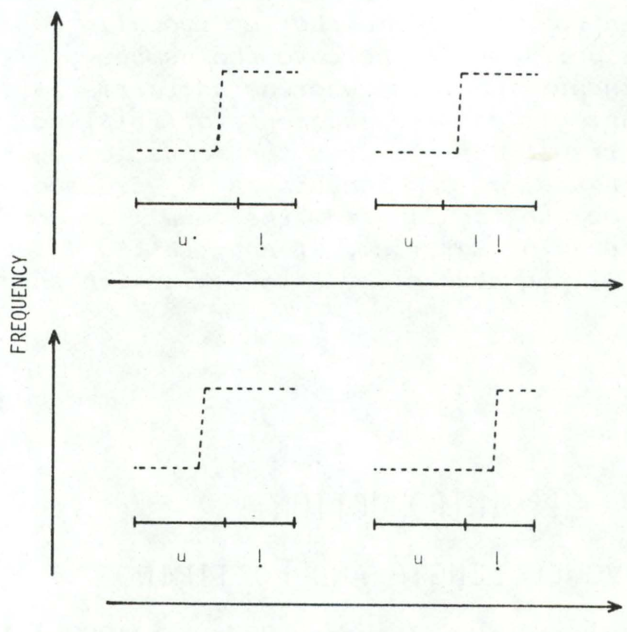


Figure 3

Stylized tracings of the normal alignment of segments with fundamental frequency in /-'V·C-/ and /-'VC-/ words (top) and of two different alignments of fundamental frequency with identical sequences of segments (bottom).

by microprosodic (segmental) phenomena. The shape of the Fø wave and the timing of stressed and unstressed syllables with respect to its trough and peak vary across dialects, cf. Thorsen and Nielsen 1981. In Standard Danish the stressed vowel hits the Fø wave in the trough before the fairly steep rise to the peak, so that generally short vowels have falling movements, whereas long stressed vowels and diphthongs will be falling-rising, cf. figure 1. With this account of Fø patterns in Standard Danish, a word pair like *kugle-kulde* (/ku·lə - kulə/ 'ball - cold') with assimilated schwa ([g̊ʰu·l̩ - g̊ʰull̩][1]) should be distinguished acoustically, inter alia, by an earlier Fø rise relative to the vowel/consonant boundary in the word with long vowel (where Fø rises before the boundary) than in the word with short vowel (where Fø rises at or after the boundary), and so is generally also the case, cf. figure 2. This account tallies well with Brink and Lund's (1975, p. 197) description of the consequences of schwa assimilation in these word types. Thus, if in figure 3 the two upper tracings are stylized illustrations of the normal alignment of segments with the Fø pattern, you would expect the two words in the lower tracings (where vowel and consonant durations are identical and presumably ambiguous) to be perceived as having a long vowel (left) and short vowel (right), respectively, if Fø rise timing has any perceptual cue value.

B. STRESS LOCATION AND Fø TIMING

Under certain circumstances in Standard Danish, a difference in stress location in otherwise comparable words can be seen, inter alia, as a rather finely timed difference in Fø movement. This is not only a question of timing the turning point in bidirectional Fø patterns (as in /ku·lə - kulə/ above) - quite the contrary. Take minimal stress pairs like *bílligst*[2] - *bilíst; Pállas - paláds; Nánna - Naná* (/'bilisd - bi'lisd -- 'palas - pa'las -- 'nana - na'na/ ['bilisd̥ - bi'lisd̥ -- 'bʰalas -'bʰa'las -- 'nana - na'na] 'cheapest - motorist'; the Greek goddess - 'palace'; Danish and French girl's name). When they are preceded exclusively by unstressed material they will all have an Fø rise from the first to the second syllable. *bílligst*, *Pállas* and *Nánna* have an Fø rise because this is the normal Fø pattern on a succession of a stressed plus unstressed syllable; *bilíst*, *paláds*, *Naná* have an Fø rise because utterance initial unstressed syllables are often lower than the first stressed syllable. The three leftmost columns in figure 4 depict Fø tracings of these words from utterance initial position by four speakers, averaged over 6 recordings. The difference in placement in a subject's Fø range (higher or lower) between the members of a pair is to be expected if the stressed syllables are to have approximately the same pitch. The difference in intrinsic Fø level between the high vowels and the [l] in *bílligst* - *bilíst* is barely noticeable with the males (NRP and JBC) but very apparent with the females (BH and NT), even though the scale is logarithmic. (I do not know why this should be so - such a difference between males and females is not to be expected from Reinholt Petersen's (1976) investi-

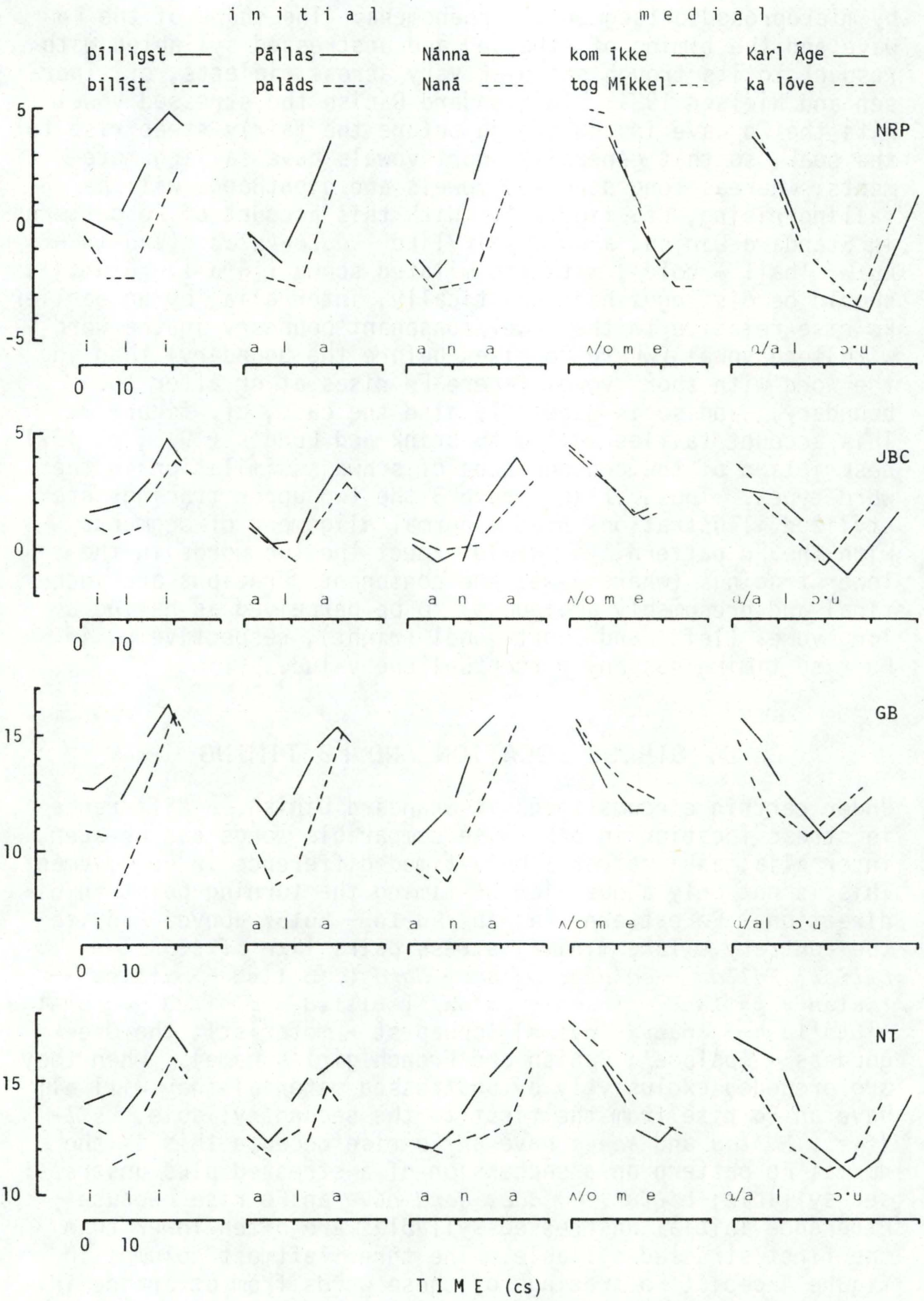


Figure 4

Fundamental frequency tracings (averages over 6 recordings) of three utterance initial word pairs (left) and two utterance medial sequences (right). The three initial word pairs, identified at the top of the figure, differ only by their stress placement. The two medial sequences differ by the location of the word boundary in a /-VCV-/ structure. Two male speakers (top) and two female speakers (bottom). Zero on the logarithmic frequency scale corresponds to 100 Hz.

gation.) Apart from differences in range and intrinsic Fø, there is a systematic difference in the relation between the vowels and the intervocalic consonant in words with first syllable stress (full line) versus words with second syllable stress (dashed line). The major Fø rise or step up begins after the consonant in the words with second syllable stress (dashed lines) but before the consonant in words with stress on the first syllable (full lines); or, in other words, the intervocalic consonant remains low and more or less at a level with the first, pretonic vowel in words with second syllable stress, whereas it rises away from the first, stressed vowel in words with initial stress. - In a pilot experiment (Thorsen 1982a) on the perception of Fø rise time differences in otherwise identical [bilisd̥]-stimuli, the location of a 2 semitone rise proved sufficient to shift listeners' location of the stress in the absence of other (e.g. durational) cues.

I take the difference in consonantal Fø movement to be an expression of a tonal dissociation between the syllable initial consonant and the stressed vowel in *bilíst, paláds, Naná*. (This interpretation will be supported by further examples below.) The rise during the consonant away from the stressed vowel in *bílligst, Pállas, Nánna* is probably not to be seen in a similar manner as a tonal dissociation from the stressed vowel but rather as a consequence of the way Fø is patterned in the prosodic stress group in Standard Danish, being low on the stressed syllable and high-falling on the post-tonics, cf. figure 1. Furthermore, the rise is not always coincident with the vowel-consonant boundary, cf. section A and figure 2 above.

Tonal dissociation between an initial consonant and stressed vowel is not confined to utterance initial position. The two rightmost columns in figure 4 show Fø tracings (averages) of the underlined medial sequences *Fár kom ikke méd til fésten. Mór tog Mikkel méd til fésten.* ('Father did not make it to the party. Mother brought Mikkel along to the party.', i.e. [ʌm ˈe]/[o ˈme]) and *Jeg trór Karl Åge får góde karaktérer. Jeg trór jeg kan lóve du får góde karaktérer.* ('I think Karl Åge will get good grades. I think I can promise you will get good grades.', i.e. [ɑ·l ˈɔ·u]/[a ˈlɔ·u]). The word boundary lies after and before the [m] and [l], respectively, but there are no apparent systematic differences in the course of Fø within each pair of tracings. (The sharp discontinuity in NRP's --*Karl Åge*-- can be ascribed to a brief glottal stop before the stressed vowel. Such boundary signalling is optional, but not very frequent in pragmatically neutral speech styles.) The tracings confirm a previous observation (e.g. Thorsen 1980) that word boundaries as a rule leave no trace in the course of Fø, and in this particular case the sonorant consonants smoothly connect the unstressed with the succeeding stressed vowel in /-- VC+V́--/ and /-- V+CV́--/ alike. However, these examples are not conclusive as to a tonal dissociation between initial consonant and stressed vowel (except maybe negatively, because there is no dissociation between the previous unstressed vowel and the initial consonant either). Such an interpretation only lends itself clearly from consideration of sequences of stressed vowels with less intervening segmental material. Figure 5 depicts the course of Fø in the underlined sequence of *Dén øl ér lúnken*. ('That beer is tepid.', i.e. ['øl ?'æʌ 'lɔ].) The initial [l] in *lúnken* behaves very much as if it were part of the preceding prosodic stress group; together with the preceding stressed vowel it performs the characteristic low + high-falling pattern. (The high consonant is not due to intrinsic Fø level differences between the [l] and the vowels because, if anything, the consonant has an intrinsically lower Fø.) In fact, the initial [l] in *lúnken* behaves qualitatively exactly as the final [l] in *øl*. Its more ample rising movement

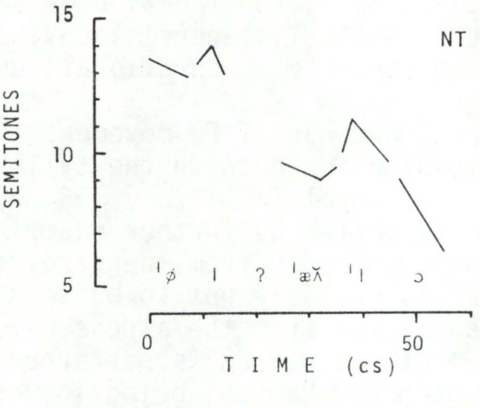


Figure 5

Fundamental frequency tracing of the underlined sequence in the utterance Dén øl ér lúnken. Zero on the logarithmic frequency scale corresponds to 100 Hz. Female speaker.

is due to the longer total duration of the sequence ['æʌl] versus ['øl], which allows for a more complete gesture before the downward course towards the next stressed vowel. In other words, the "tonal syllabification" seems to be V̆C-V́ rather than V̆-CV́, irrespective of phonological and morphological/syntactic boundaries. Note, however, that tonal dissociation between initial consonant and stressed vowel does not always make itself noticed. A consonant will not go out of its way to dissociate from the succeeding homosyllabic stressed vowel, cf. the rightmost examples in figure 4, but when two stressed vowels are sufficiently spaced in frequency (as they will be in short terminal declarative utterances), and when there is not sufficient segmental material to connect them smoothly, an initial consonant can be seen to team up with the preceding prosodic stress group.

Whether a stressed vowel can repel more than one initial consonant is hard to say, due to the phonetics and phonotactics of Danish. The sonorant consonants do not combine initially. The obstruents are all unvoiced, except [v]. [v] combines initially only with [ʁ] as in *vride* ('to wring') and with [l] in foreign names like *Vladimir* , so the material for investigation is limited, and it may be difficult to interpret due to intrinsic and coarticulatory effects of the [vʁ] and [vl] combinations.

To summarize the acoustic observations: An initial voiced consonant in a stressed syllable will be dissociated tonally from the stressed vowel and join up with the preceding material, if any. Postvocalic consonants after a stressed vowel behave tonally in an unmarked fashion, and their relation to the preceding vowel is determined by the Fø pattern which characterizes a prosodic stress group in Standard Danish. - Under certain circumstances, the relation between an intervocalic voiced consonant and the surrounding vowels may be the only tonal difference between words with different stress locations.

To what extent tonal dissociation between initial consonants and stressed vowels is manifest in other varieties of Danish, I cannot say, but I doubt that the phenomenon should be an exclusively Standard Danish one, or even a Danish phenomenon, for that matter. It matches the observation from some tone languages that postvocalic consonants may carry the final part of the distinctive tonal course but prevocalic consonants do not seem to be included (Pike, 1948, p. 10, 30; Selmer, 1928).

II. PROCEDURES

On the hardware parallel synthesizer (Rischel and Lystlund 1977) at the Institute of Phonetics, I synthesized two sets of "words", [g̊ʰul] and [b̥ilisd̥]. In figure 6, top left, the four variables in the [g̊ʰul̥] stimuli are shown schematically. (1) Vowel duration varies in three steps: 100, 140, and 180 ms;[3] (2) consonant duration varies in three steps: 100, 120, and 140 ms; (3) the timing of the Fø rise occurs at the vowel-consonant boundary or 60 ms before the end of the consonant, i.e. it is constant with respect to the end of the consonant; (4) the rise is 3 or 4 semitones high, and is accomplished in 5 ms. Only 26 (of a possible 36 total) stimuli were included in the test.

Five parameters varied in [b̥ilisd̥], cf. figure 6, top right. (1) the timing of the Fø rise, which occurred immediately after the first vowel or immediately after the [l]. That is to say, either the [l] was low - at the level of the first vowel - or it was high - at the level of the second vowel. Both vowels were level pitched and had identical intensities; (2) the magnitude of the Fø rise, which was 0, 2, 3 or 4 semitones high; (3) placement in the Fø range of the whole word, that is to say, the first vowel was at 89 Hz or at 100 Hz; (4) duration of the first vowel varied in three steps: 60, 80 and 100 ms; (5) duration of the second vowel varied in three steps: 80, 100, and 120 ms. These were values which I myself found would create definite /ˈbilisd/ and /biˈlisd/ stimuli, as well

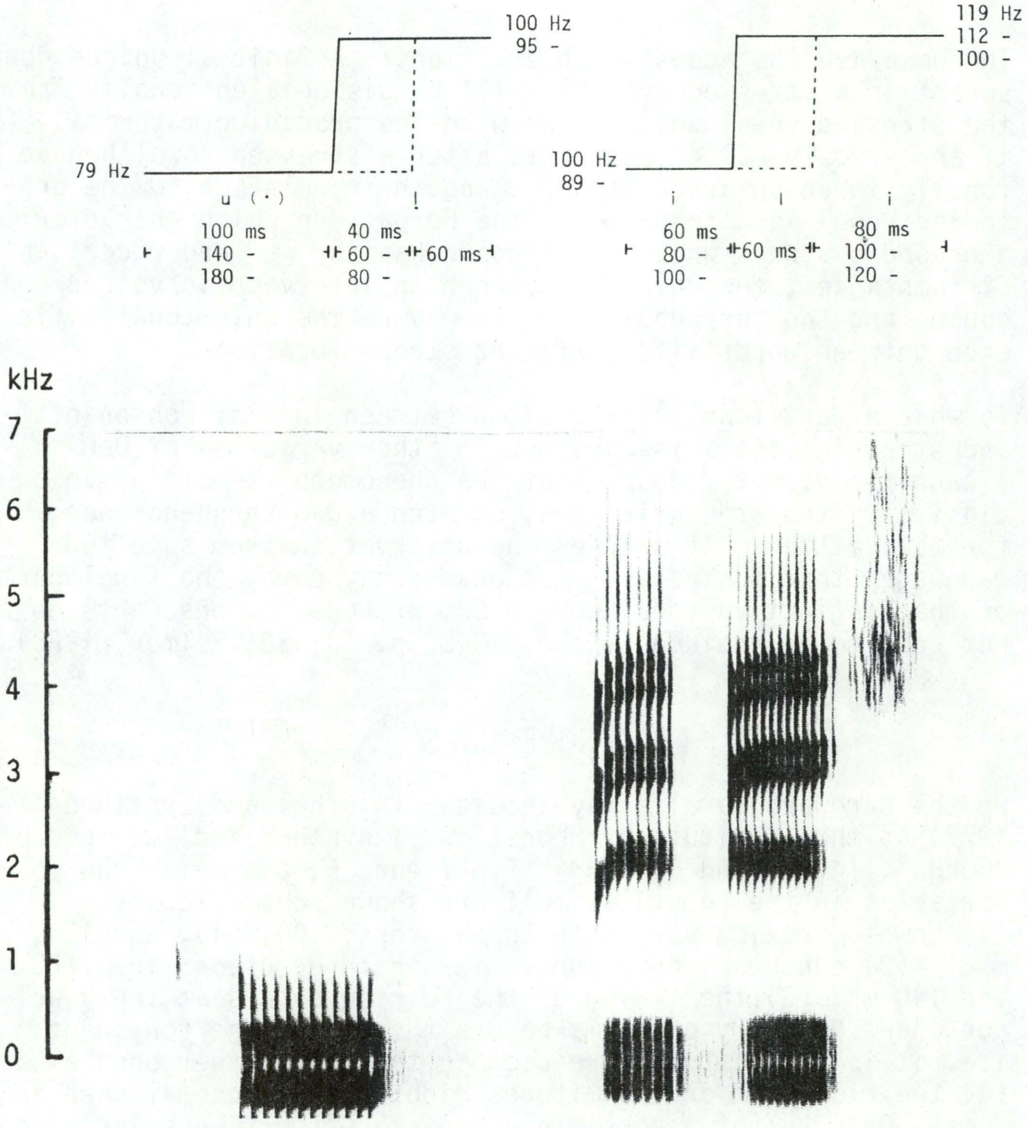


Figure 6

Scheme of the variables in two synthesized stimuli [g̊ʰuḷ] and [bilisd̥] (top), and spectrograms of two sample stimuli (bottom). See further the text.

as ambiguous ones. From a possible 126 different stimuli, 63 were selected for the test. (Actually, 12 more words were included, where [l]'s Fø was lower than the first vowel. Since these stimuli neither add to nor subtract anything from the conclusions to be reached from stimuli where [l]'s Fø is identical with the first vowel, they are disregarded in the following.)

The total of 101 (26 + 63 + 12) different words were randomized five times, each randomization on a separate tape. Every tape was introduced by the same two repetitions of three [b̥ilisd̥] and three [g̊ʰul] which swept the whole range of variation and were intended for practice and not to be responded to. The 101 test items were led by five and succeeded by four dummies, giving a total of 110 items to be identified. Each stimulus occurred once only, with 4 sec intervals between stimuli. Leader tape after every ten words automatically stopped the tape recorder, to help subjects keep track of their progression. One run of a test tape lasted 11 minutes.

Subjects were instructed (in writing) about the composition of the tapes. They were asked to identify each word as either *kugle* or *kulde* or as *billigst* or *bilist* (forced choice) by ticking off the appropriate box on prepared answer sheets. The tapes were presented to subjects individually, over headphones. Some listened to one tape a day, five days running, others distributed their trials over a couple of weeks. One subject heroically worked his way through all five tapes in one session. Nine colleagues and students at the Institute of Phonetics, representing different age groups and dialectal background, took this test.

III. RESULTS

A. STATISTICAL TREATMENT

The only suitable test of independence between response distributions such as depicted in figure 7 and 9 is the rather conservative, non-parametrical χ^2. χ^2 values are given in the lower right of each rubric and the confidence level, p, is indicated if it is 0.05 or better. (In a one-tailed test it takes a χ^2 above 13.36 to be significant at the 0.05 level at 8 degrees of freedom and above 10.64 at df = 6.) N is the total number of responses behind each data point.

The significance of the difference between scores on any pair of stimuli (or collapsed stimuli) may be tested in a student's t-test, from the score variances, according to

$$\sigma^2 = n \cdot p \cdot q$$

where σ^2 is the score variance, n is the total number of presentations of the stimulus (or collapsed stimuli), p is the proportion of responses in one of the two categories, i.e. p = score/n, and q = (1-p). When a score difference is statistically significant at the 0.05 level, or better, this is indicated with a star just below the zero line in figure 7 and 9.

For example: the two response distributions in figure 7 are not significantly different (χ^2 = 5.73, df = 8), but two stimulus pairs did receive significantly different scores.

Thus, the stimuli in the fifth pair received 13 and 19 *kugle* identifications, respectively. The score variances are 6.24 and 4.56, and t = 1.83. At df = (2·25-2) = 48 this value is significant beyond the 0.05 level (one-tailed test).

B. *kugle-kulde*

To four of the nine subjects vowel duration seems to have been the only perceptual cue to the identity of the test words. At the shortest vowel duration (100 ms) only *kulde* is heard, and at the two longer vowel durations (140 and 180 ms) only *kugle* is identified, irrespective of the duration of the consonant and of the timing of the Fø rise with respect to vowel-consonant boundary. In other words, the range of vowel durations is ill suited to these subjects, and they are therefore disregarded here.

Responses from the remaining five subjects to 9 stimulus pairs (those with a 3 semitone Fø rise) are shown in figure 7 where the percentage *kugle* responses (of a possible 25 total) is depicted. There is an effect from all three parameters, vowel duration, consonant duration, and Fø rise time, but the latter two are of little consequence only when the vowel is short (100 ms - no. 1-3) and long (180 ms - no. 7-9). At the intermediate vowel duration (140 ms - no. 4-6) there are more *kugle* identifications when the consonant is short and with an early Fø rise relative to the vowel-consonant boundary. The overall effect of Fø rise time is weak and not statistically significant, and in natural speech there is every reason to believe that vowel (and consonant) duration alone decide(s) the issue.

C. *bílligst-bilíst*

1. RESPONSES

Two subjects consistently identified very nearly 100% of the [b̥ilisd̥] stimuli as *bílligst*, and their data are disregarded in the following. However, their behaviour illustrates how different subjects may choose different (internal) judgment criteria, and how the choice of response category label therefore determines identification. Both subjects had participated in the pilot test six months previously (Thorsen, 1982a). One had then given balanced and consistent responses in both categories, and the other had surprisingly given almost exclusively *bilíst* responses. The latter subject had explained her reaction to the pilot test in the following manner: she had heard some adequate *bilíst*s and some which perhaps were not good but which she would not identify as *bílligst* at any rate. In other words, *bílligst* seems to have been the marked member of the pair to this subject in the pilot test, since her requirements for *bílligst* identification did not allow her to put doubtful stimuli in that slot. In the present experiment she turned the tables (but not consciously, because she

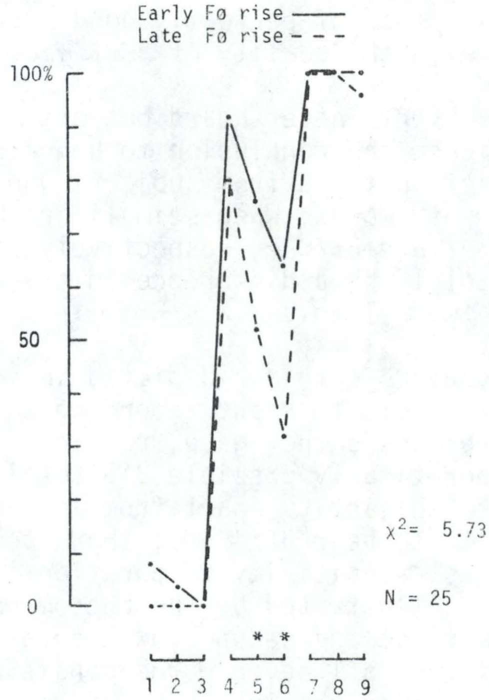


Figure 7

Identification functions (proportion of long vowel judgments, i.e. kugle) of two sets of synthetic [g̊ʰu(·)l̩] stimuli which differ with respect to the timing of the fundamental frequency rise from the stressed vowel to the post-tonic syllabic consonant. Vowel duration increases from left to right thus: 100 ms (no. 1-3), 140 ms (no. 4-6), 180 ms (no. 7-9). Within each triad consonant duration increases from left to right thus: 100 ms (no. 1,4,7), 120 ms (no. 2,5,8), 140 ms (no. 3,6,9). Fundamental frequency rises 3 semitones in 5 ms, either at the vowel-consonant boundary ("early") or 60 ms from the end of the consonant ("late"). χ^2 on the two response distributions is given in the lower right. The total number of responses behind each data point is 25 (5 subjects times 5 responses). A star beneath a pair of data points indicates that those two stimuli received significantly different response scores.

did not remember what her choice had been on the first occasion), and she now confessed to have heard some adequate *billigst*s and some which were really not very good, but none that could have passed as good *bilist*s. The other "uncooperative" subject described her reaction to this test in a similar manner. Now, clearly, if subjects had been asked to sort out the acceptable *billigst*s, i.e. if response categories had been labelled *billigst/not billigst*, these two subjects' response distributions might have been radically altered. But then, so would probably some of the other subjects' responses: at least two of them (ND and OT) declared to have been quite certain to have heard a number of perfectly good *bilists* but they were less satisfied with the quality of *billigst*.

The fact that some subjects never heard but one member of the pair does not invalidate the conclusion to be reached on the basis of response distributions from subjects who had identified occurrences of both words: When stimuli are heard predominantly as *billigst* and *bilist*, respectively, this difference may be caused solely by a difference in the timing of the Fø rise between the two syllables.

The results from seven subjects are depicted in figure 8. Subjects are ordered from left to right according to the total number of *billigst* responses they gave, ranging from 154 (ND) to 252 (PM) of a theoretically possible 315 total (63 stimuli times 5 responses per subject). Apart from an overall smaller or greater inclination to hear *billigst*, there seems to be a difference in subjects' sensitivity to durational variation, which is most clearly illustrated by the bottom row, where Fø is constant and only first and second vowel duration vary: AM and PM have identified all seven words consistently as *billigst*, JB has identified the last five stimuli in the septet, i.e. those with medium and long first vowel as *billigst*, irrespective of the duration of the second vowel. With JR and OT the increase in the duration of the second vowel can be seen as a decrease in *billigst* responses at medium first vowel durations (no. 3, 4, and 5 in the septet).

Despite quantitative differences between subjects, and different sensitivity to some of the parameters under investigation, subjects do not exhibit systematic qualitative differences that prohibit consideration of their pooled data. Accordingly, figure 9A-D and F present the responses to individual stimuli as the percentage *billigst* identifications (of a possible 7x5 = 35 total). Figure 9E,G, and H present the data collapsed in various ways, see further below.

2. INTERPRETATION

Figure 9A-E shows clearly that, everything else being equal, an [l] at the same Fø level as the preceding vowel (broken lines) yields a smaller proportion of *billigst* answers than an [l] at the level of the succeeding vowel (full lines). In other words, when the consonant is more closely associated tonally with the first vowel, the second vowel tends to be perceived as stressed and vice versa, when the consonant is more closely associated tonally with the second vowel, the first vowel tends to be perceived as stressed. The effect is dependent on duration, being generally nearly nil at the longest first vowel duration (no. 6 and 7 in each septet). The cue inherent in duration alone is apparent from figure 9F which depicts responses to stimuli with a monotone Fø at 100 Hz. That figure resembles very closely the distribution of collapsed responses

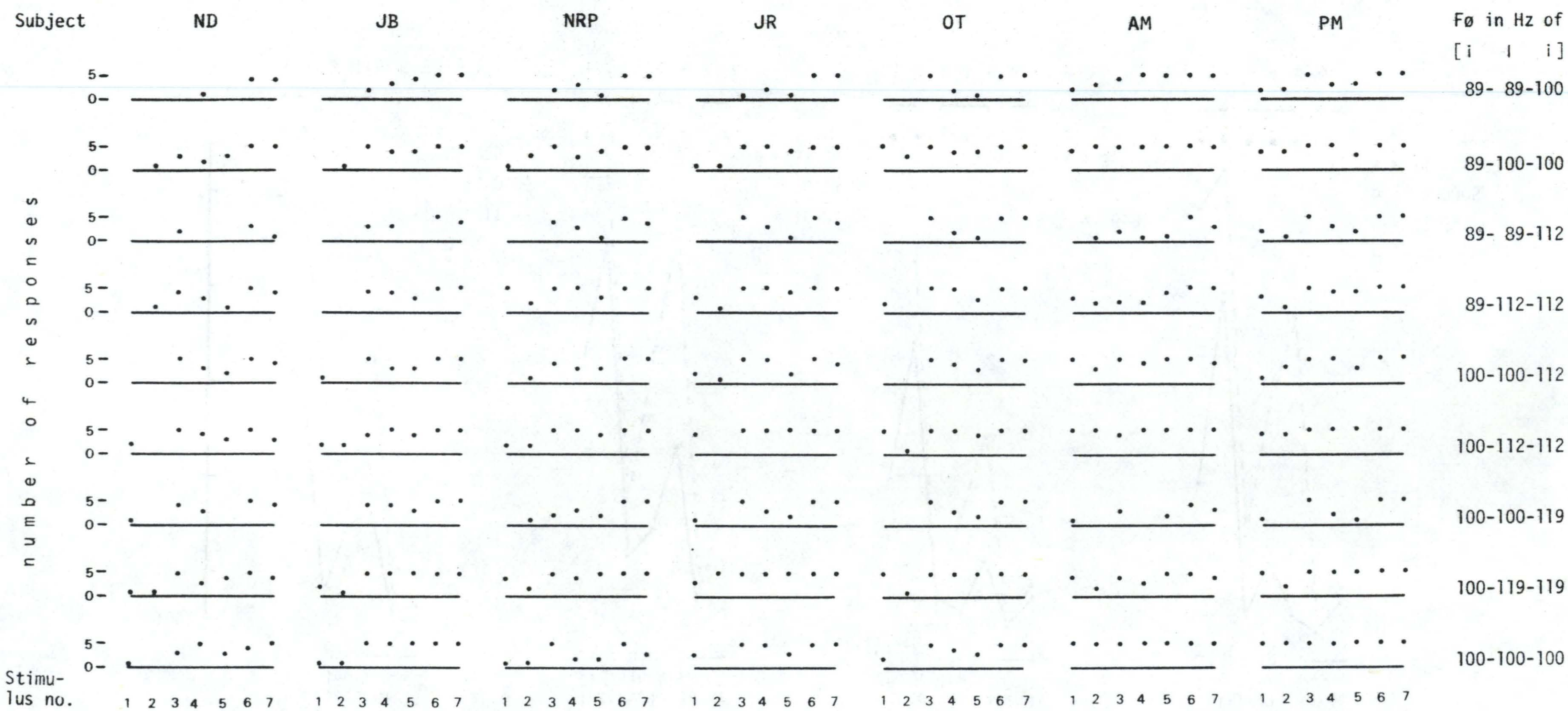
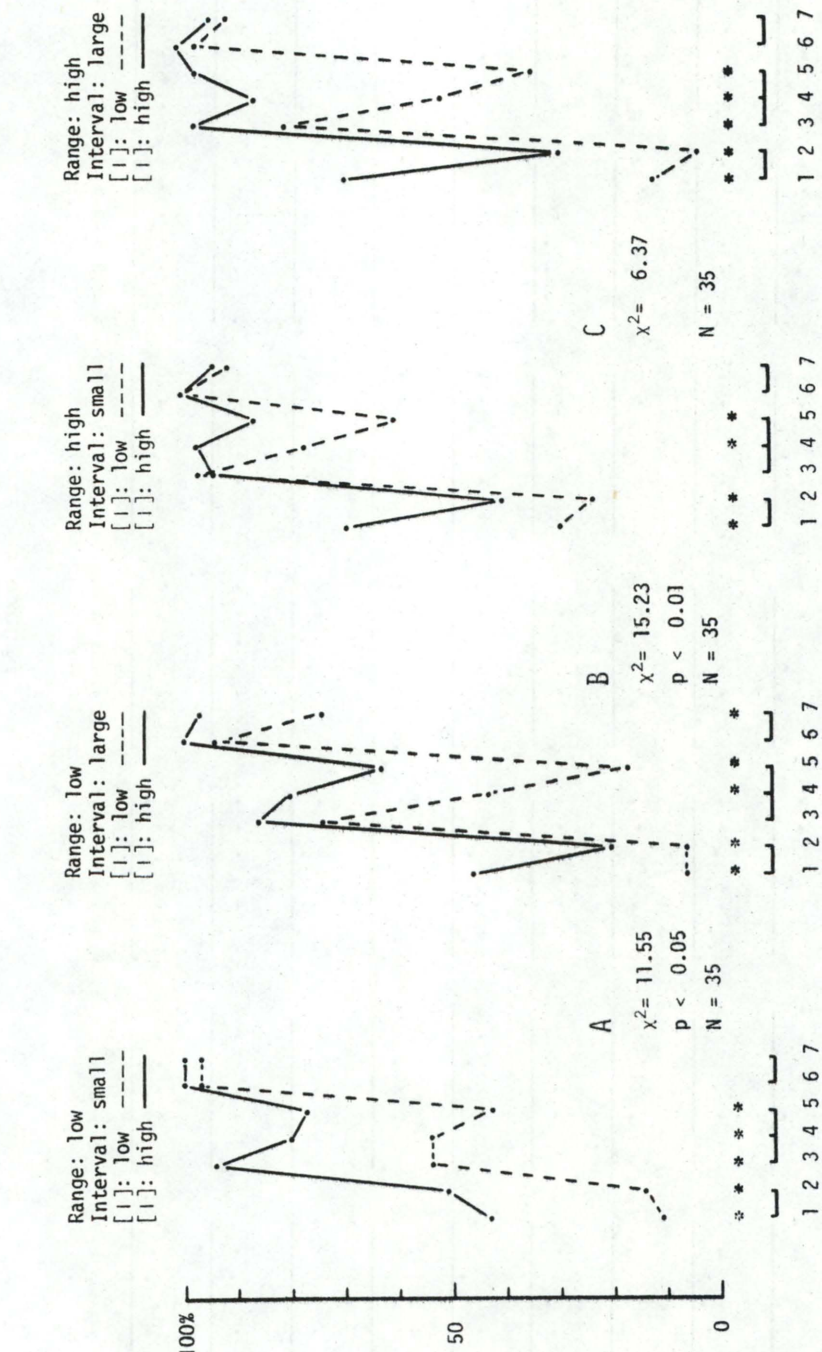


Figure 8

Seven subjects' judgments of initial stress (i.e. billigst) to 63 synthesized [b̥ilisd̥]-stimuli. There are 9 different fundamental frequency conditions, as indicated to the right in the figure. Within each septet of stimuli the duration of the first and second vowel varies thus: no. 1 - 60/100 ms; no. 2 - 60/120 ms; no. 3 - 80/80 ms; no. 4 - 80/100 ms; no. 5 - 80/120 ms; no. 6 - 100/80 ms; no. 7 - 100/100 ms. Subjects are ordered from left to right according to the smaller or larger total number of initial stress judgments they gave.



Range: low
Interval: small
[I]: low
[I]: high
A
χ^2= 11.55
p < 0.05
N = 35
Range: low
Interval: large
[I]: low
[I]: high
B
χ^2= 15.23
p < 0.01
N = 35
Range: high
Interval: small
[I]: low
[I]: high
C
χ^2= 6.37
N = 35
Range: high
Interval: large
[I]: low
[I]: high
D
χ^2= 21.34
p < 0.0005
N = 35
100%
50
0
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7

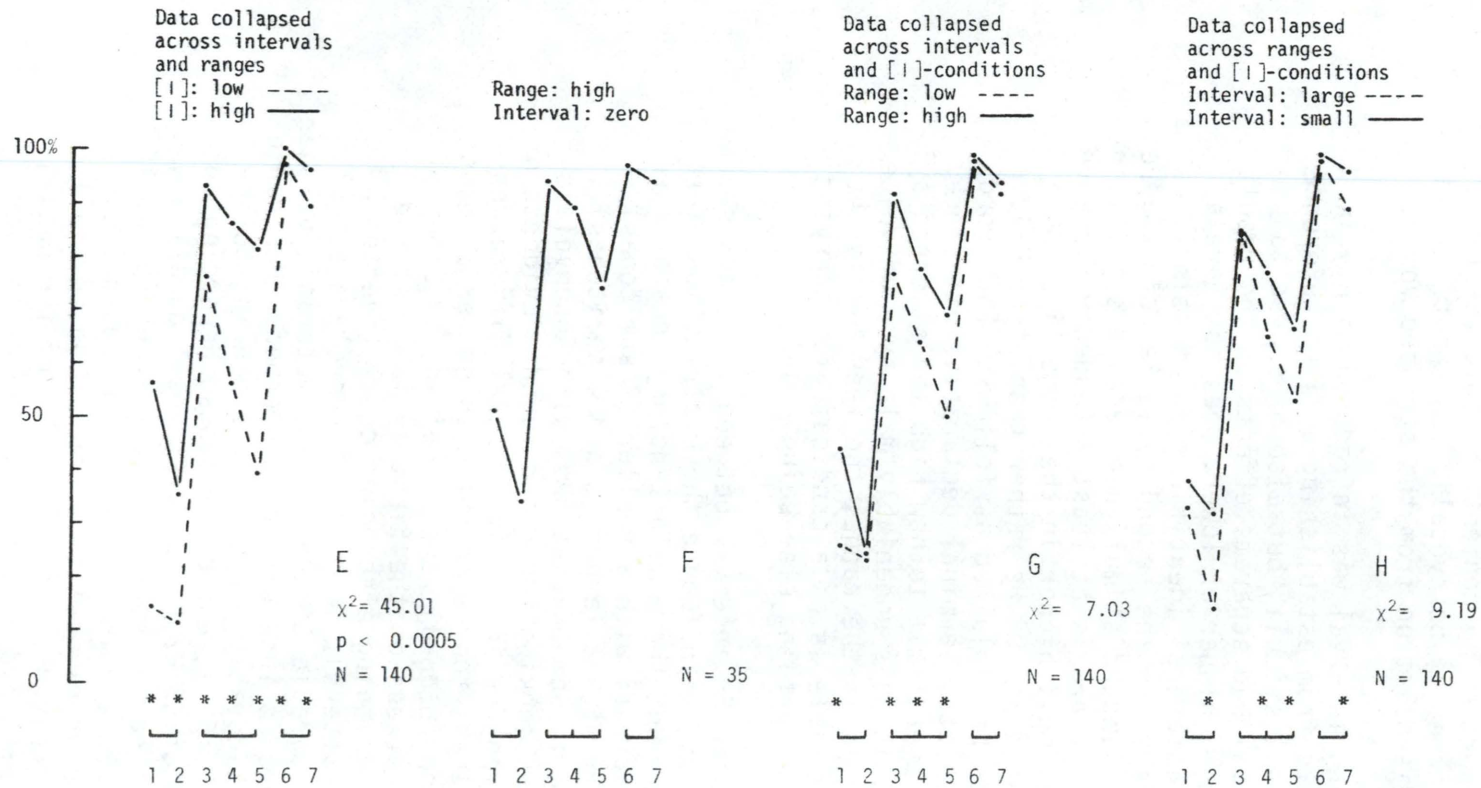


Figure 9

Identification functions (proportion of initial stress judgments, i.e. billigst) of 9 sets of synthesized [bilisd]-*stimuli, collapsed across seven subjects. See further the legend to figure 8. In rubrics E, G, and H the data are collapsed across intervals, ranges and* [l]-*conditions as indicated at the top.* χ^2 *on each pair of response distributions is given in the lower right, and p is indicated if 0.05 or better. N is the total number of responses behind each data point. A star beneath a pair of data points indicates that those two stimuli received significantly different response scores. "Range" refers to the frequency of the first vowel (89 Hz: low; 100 Hz: high). "Interval" refers to the frequency interval between the first and second vowel (2 semitones: small; 3 or 4 semitones: large). "low" and "high"* [l] *refers to the location of the consonant in the frequency range (on the same level as the first vowel: low; on the same level as the second vowel: high).*

obtained from stimuli with an [l]-condition which favours *bi̊lligst* responses (9E, full line). This can be taken to mean that an Fø rise timed immediately after the vowel is the "unmarked" Fø condition, since it does little to change listeners' identification of stimuli where Fø cues are totally absent, and it is a further support of my contention above (section I.B) that the rise performed by postvocalic consonants is not to be seen as a tonal dissociation from the stressed vowel.

Variation in Fø range and interval was introduced mainly in order to prevent subjects from establishing a fixed reference against which to judge all stimuli, but also of course to see whether range and interval had separate effects on the results. Placement in the Fø range, higher or lower, does not have a significant effect on listeners' identification of stimuli as bearing stress on the first or the second syllable, cf. figure 9G, but four pairs of collapsed stimuli did receive significantly different scores, and there is at least a tendency for more *bi̊lligst* responses to stimuli higher in the range (i.e. when the first vowel is located at 100 Hz rather than at 89 Hz). This trend might perhaps be explained as follows. Each word is simultaneously a complete, terminal declarative utterance. There is therefore an upper limit to how high the stressed syllable can plausibly be, in this undeniably male type of voice, and the higher [b̥ilisd̥] is in the range, the less likely the second, and highest, syllable is as a candidate for carrying (non-emphatic) stress, everything else being equal.

Likewise, magnitude of the Fø interval between first and second vowel does not significantly influence the distribution of responses, cf. figure 9H. But there is a tendency towards more *bi̊lligst* responses to stimuli with a smaller (2 semitones) interval than to stimuli with a larger (3 and 4 semitones) interval. This finding is somewhat at odds with Rosenvold's (1981) results. She was looking for - and found - evidence of a perceptual compensation for intrinsic Fø level differences between vowels of different tongue height. In her experiment synthesized [b̥id̥b̥iɡ̊] and [b̥id̥b̥aɡ̊] were perceived, by 9 out of 10 subjects, as having stress on the first syllable at the higher Fø intervals, everything else being equal (the tenth subject reversed this pattern). However, Rosenvold's stimuli are not immediately comparable to the present material, neither where duration nor where Fø interval are concerned. Her range of Fø intervals is considerably larger, corresponding to approximately 0, 2, 3½, 5, 6, 7, and 8 semitones, departing from a constant first vowel of 85 Hz. The first vowel duration was constant at 100 ms, the second vowel varied between 80 and 150 ms in 10 ms steps. If Rosenvold's data are collapsed across different durations and vowel qualities (figure 4a, p. 157), the second vowel is perceived as stressed (85% of the presentations) at intervals of 0 and 2 semitones, the cross-over point is at 3½ semitones, and intervals of 5 and more semitones are perceived as having stress on the first syllable (90% of the presentations). The fact that at the larger intervals, the first vowel is perceived as stressed may be explained along the same tentative lines as the effect of placement higher in the range in this experiment, cf. above: the second syllable may be too high to be a likely stressed syllable in a one word non-emphatic terminal declarative utterance. The fact that the second syllable is perceived as bearing stress when the interval is 0 or 2 semitones may be due to duration, rather than to Fø interval as such: a first vowel at 100 ms and a second vowel varying between 80 and 150 ms could be expected to provoke a predominance of second vowel stress judgements in the absence of other, overriding, cues (such as Fø).

No doubt Fø rise timing is the single most effective stress location cue in this experiment, disregarding vowel duration. The only instance where the difference between response distributions from the two different Fø rise time conditions is not statistically significant is when the stimuli were high in the range, and the Fø interval small, cf. figure 9C, a fact that may be explained as the combined effect of the trend towards a general preference for initial stress high in the range and with small intervals, cf. above. In other words, this is a condition where the tendency is towards perceived stress on the first syllable anyway, and differences in Fø rise timing will be less evident.

3. CONCLUSION

When everything else is equal, differences in the timing of an Fø rise between utterance initial syllables may influence the identification of the location of stress, but vowel duration - particularly the first vowel - may easily override the cue inherent in Fø rise timing. There is hardly any effect from Fø when the first vowel is sufficiently long. In view of the identification of stimuli with a monotone Fø, the proper way to interpret the results is probably to say that when the intervocalic consonant is dissociated tonally from the second vowel, then that vowel is heard as stressed. If not, then the first vowel is stressed.

Naturally, the timing of Fø events at this, segmental, level will hardly ever be the only acoustic cue to stress location. Mainly because in utterances with more than one prosodic stress group in them, the tonal relations to the surrounding syllables will decide the issue. If the stressed syllable in *bĭlligst* and *bilĭst* is, e.g., the second one in the utterance, then the stress cuing Fø rise will occur after *bĭl-* in *bĭlligst*, but after *-lĭst* in *bilĭst*, cf. figure 10. Furthermore, segment duration, vowel quality, degree of aspiration, etc. will help single out the stressed syllables of an utterance.

Tonal segregation of initial consonants may never serve any direct communicative function and may at best be said to be a redundant cue to stress location. But since it occurs and <u>can</u> be perceived - at least in laboratory conditions - there <u>must</u> be an explanation for it. It seems that as far as its tonal manifestation goes, <u>stress begins with the vowel</u>, because - contrary to expectations, perhaps - a stressed vowel

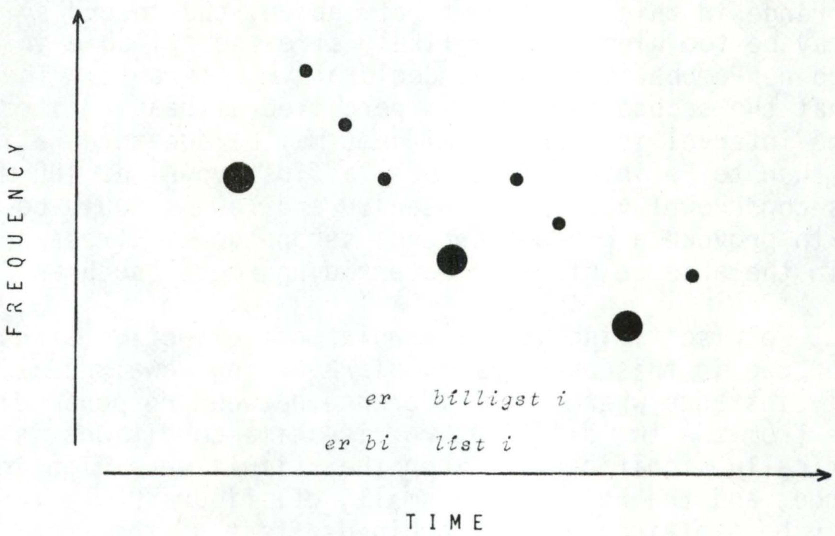


Figure 10

Stylized course of fundamental frequency in a terminal declarative utterance where the second stressed word is bi̋lligst and bili̋st, respectively. The heavy dots denote the stressed syllables, the light dots the unstressed ones. The utterances could be, e.g. Várerne er bi̋lligst i Ísland. and Ánni er bili̋st i Gali̋cien. ('The goods are cheapest in Iceland. and 'Anni is the motorist in Galicia.)

repels rather than attracts preceding homosyllabic consonants. One side-effect of this is that word and syllable boundaries are deleted in the course of Fø.

Prosodic segregation of initial consonants may not be a purely tonal phenomenon. Fischer-Jørgensen (1982, p. 159) presents evidence that under certain circumstances the pattern of segmental duration indicates a boundary between a pre-vocalic consonant and a stressed vowel. Recent experiments on Swedish indicate that the most reasonable account of rhythmic phenomena is achieved if the onset of the rhythmical unit is taken to be the onset of the stressed vowel rather than, say, the onset of the first pre-vocalic consonant. This was one of the conclusions drawn by Lubker et al. (1983) from data on articulatory compensation in bite-block experiments, and Strangert (1983) likewise found that the most appropriate segmentation for an account of inter-stress intervals is the onset of the stressed vowel.

IV. NOTES

1. A dot under a consonant symbol signifies syllabicity here.

2. The stressed vowels are indicated orthographically with acute accents in the following.

3. As on most speech synthesizers, actual segment duration may vary by as much as a whole period both at the onset and offset of voicing, because the settings pertain to the voice source amplitude gate control, and this control is independent of the repetition rate of the voice source pulses. Several repetitions of a given stimulus were recorded on master tape, and from oscillogram and intensity curves, those items were selected which best corresponded to the intended durations. However, measuring accuracy is hardly better than ± 5 msec, and minor deviations from intended vowel duration values may accordingly still occur.

ACKNOWLEDGEMENT

Thanks to Jørgen Rischel for valuable suggestions for improvements in the manuscript.

REFERENCES

Brink, L. and Lund, J. 1975: *Dansk Rigsmål 1-2* (Gyldendal, Copenhagen)

Fischer-Jørgensen, E. 1982: "Segment duration in Danish words and its dependency on higher level phonological units", *Ann. Rep. Inst. Phon., Univ. Copenhagen 16*, p. 137-189

Lubker, J., Lindblom, B., Gay, T., Lyberg, B., Branderud, P. and Holmgren, K. 1983: "Prosody and compensatory motor behavior", paper read at a symposium on *Prosody - Normal and Abnormal*, Zürich, April 6-8

Pike, K. L. 1948: *Tone Languages*, (University of Michigan Press, Ann Arbor)

Reinholt Petersen, N. 1976: "Intrinsic fundamental frequency of Danish vowels", *Ann. Rep. Inst. Phon., Univ. Copenhagen 10*, p. 1-29

Rischel, J. and Lystlund, S.-E. 1977: "The IPUC speech synthesizer", *Hamb. phon. Beitr., Misz IV* (= Band 22), p. 143-160

Rosenvold, E. 1981: "The role of intrinsic Fo and duration in the perception of stress", *Ann. Rep. Inst. Phon., Univ. Copenhagen 15*, p. 147-166

Selmer, E. W. 1928: "Noen bemerkninger om den musikalske aksent i dens forhold til den sterkt og svakt skårne aksent", in *Festskrift til Rektor J. Qvigstad. Tromsø Museums Skrifter 2*, p. 250-262

Strangert, E. 1983: "Temporal characteristics and rhythmic units in Swedish", paper read at the *Third Symposium on the Prosody of Nordic Languages*, Umeå, June 3-4

Thorsen, N. 1980: "Word boundaries and Fo patterns in Advanced Standard Copenhagen Danish", *Phonetica 37*, p. 121-133

Thorsen, N. 1982a: "On the variability in Fo patterning and the function of Fo timing in languages where pitch cues stress", *Phonetica 39*, p. 302-316

Thorsen, N. 1982b: "Selected problems in the tonal manifestation of words containing assimilated or elided schwa", *Ann. Rep. Inst. Phon., Univ. Copenhagen 16*, p. 37-100

Thorsen, N. and Nielsen, B. Jul 1981: "Fundamental frequency patterning and sentence intonation in two Jutlandic dialects - A first approach", *Ann. Rep. Inst. Phon., Univ. Copenhagen 15*, p. 1-12.

ARIPUC 17, 1983, p. 21-53

SUPPLEMENTARY NOTES ON VOWEL LENGTHENING IN DANISH

ELI FISCHER-JØRGENSEN

The recordings which formed the basis of Fischer-Jørgensen "Segmental duration of Danish words in dependency on higher level phonological units", ARIPUC 16, 1982, were supplemented by a number of new recordings containing mostly actual disyllables (i.e. words without apocope). The four West Jutlandish speakers read a supplementary list, and recordings were made of a number of new informants: seven dialect speakers from East Jutland, five from North Zealand, and two younger Copenhagen informants. It turned out that both West- and East Jutlandish speakers have vowel lengthening in actual disyllables but not in apocopated words (the same was true of the North Zealandish speakers who had apocope). There is thus more lengthening in the dialects than assumed in my 1982 report. Very conspicuous lengthenings were found in three North Zealandish coastal dialects, but somewhat less lengthening than expected in the speech of the two Copenhagen informants. The rejection of the hypothesis of a connection with apocope or with a special pitch contour was confirmed by the new recordings. The assumption of a connection with the weakness of the following consonant is upheld.

I. INTRODUCTION

In the preceding volume of this series I published a report (Fischer-Jørgensen 1982) showing that the tendency, found in many languages, to shorten the segments of disyllables compared to monosyllables was counteracted in Danish by a tendency to lengthen the short stressed vowel of the structure CV́CV(C) in disyllables (and also in words with more syllables). Thus in

Standard Danish as spoken in Copenhagen, represented by seven informants, the stressed vowel of *danne* ['d̥anə][1] is significantly longer than the vowel of *Dan* [d̥an] and, similarly, the stressed vowel of *natte* ['nad̥ə] is significantly longer than the vowel of *nat* [nad̥], the average difference of all pairs investigated being 3.2 cs. The stressed vowel of trisyllables of the type *dannede* ['d̥anəðə] is slightly shorter than that of ['d̥anə] but longer than in [d̥an]. The same difference is found when the structures CVC and CV̆CV are members of compounds, e.g. ['d̥anə,vɑŋʔ] vs. ['d̥an,fɔs]. The tendency seemed to be stronger in words with the ending *-e* [ə] than with the ending *-er* (pronounced [ɒ] or [ɔ]), and stronger before stops and nasals than before *s*. Before consonant clusters (*bast/ baste*) the difference is small and in most cases non-significant. The words were placed in frames of the type *Han sagde* ..., or *Han sagde ... fem (to) gange* 'he said ... five (two) times', a few also in small sentences.

In the lower sociolect of Copenhagen there is an audible lengthening of short vowels, which has been noticed by various observers, and according to the general impression it is more pronounced in disyllables than in monosyllables. This impression was supported by the recordings of a speaker of this sociolect (ST).

All the main informants had lived in Copenhagen for many years, but some of them had a still perceptible dialectal background (Jutlandish, and in one case Funish). The question was now whether their vowel lengthening was solely due to Copenhagen influence, or whether similar tendencies could be found in the dialects.

In order to throw some light on this problem I recorded samples of some Danish dialects and regional standards (RSD). There were three Zealandish, seven Funish, and six Jutlandish speakers. The test words were placed in small sentences.

Of the three Zealandish speakers, one (EP from Dragør on Amager near Copenhagen) had a considerable vowel lengthening in disyllables before single consonant. She had preserved the final vowel. The other two had apocope of final *-e*. One (from Bjeverskov near Køge) had few significant lengthenings, the other had a significant lengthening before nasals and in *takker* but no lengthening in other cases. Listening to older tape recordings of North Zealandish dialects revealed audible lengthenings in two cases, but there were also informants without audible lengthenings. The conclusion was that there are clear lengthenings in at least some Zealandish dialects, which may be connected with the lengthenings found in Copenhagen.

For the Funish speakers the results were very heterogeneous. Only one informant (IP from Western Funen) had significant vowel lengthening in most words (there were two exceptions). One informant (from Eastern Funen) had consistent lack of lengthening except in the actual disyllabic *bassen* [b̥asn̩]. (Both had apocope.) For the others there was no consistent pattern, and the examples of significant differences were found in different word pairs for different persons. In table II of the 1982-report the comparison between *bas* and *bassen* was left out, and since the table moreover contained a number of misprints (mostly lacking minus-signs, but fortunately only for non-significant differences), it is given here again in a corrected form. The definite forms of *bas* and *basse* (in both cases: *bassen*) were originally included in the list in order to see whether there would be any difference between old accent 1 and accent 2 words, which was not the case; but a comparison with *bas* is also of interest.

The recordings of Jutlandish informants gave much more consistent results. A typical speaker of the Århus RSD had lengthenings in exactly the same cases and of the same magnitude as the Copenhagen Standard speakers, and the same was, on the whole, the case for the RSD of five dialect speakers, one from Northern Jutland (Vendsyssel) and four from various parts of the West Jutlandish dialect area (with a number of exceptions for the speaker from Thy, TA).

In the recordings of the dialect forms of these five informants the lengthenings were quite different. All Jutlandish dialects have obligatory apocope of final [-ə]. The informant from Vendsyssel had, as should be expected for his dialect, a considerable lengthening of the vowel in apocopated forms, also before consonant clusters (there may be coalescence with phonologically long vowels). He had no lengthening in the actual disyllable *sønner* but some lengthening of *takker* and *danne*. Only for the latter two words was there agreement between his dialect and his RSD. The four West Jutlandish speakers had no lengthening in apocopated forms before stop consonant (in this case there is West Jutlandish stød), nor (for the two speakers who read these words) before *s*, but they had lengthening of the vowel and (with one exception) also of the consonant in *komme* and in the actual disyllable *takker*, two also in *sønner*. These were the only words in which all four informants had approximately the same lengthening in their dialect and their RSD. For other words there were only some scattered cases.

From these facts I drew the conclusion that the vowel lengthening in the RSD of the North- and West Jutlandish speakers is not based on their dialects, but is a different norm, which is in accordance with the Århus RSD. Further I concluded that this latter norm cannot be based on East Jutlandish dialects since it is well known that in East Jutlandish there is complete phonological merger between old disyllables with apocope and monosyllables without stød, also before sonorants, and in this case an allophonic lengthening of the old disyllables would be very unlikely. The lengthening in the Jutlandish RSD must thus be due to influence from the Copenhagen standard, which is now spreading rapidly.

Table I

Funish

Differences in vowel duration (in cs) between disyllables and monosyllables with phonologically short vowels for Funish speakers. Speakers are indicated by initials. - indicates shorter duration in disyllables, a star indicates that the difference is significant at the 1% level, a star in parentheses that it is significant at the 5% level (each average covers six tokens). ap. means apocope. D means dialect, RSD "Regional Standard Danish". A point after the words indicates final position.

	MA	LA	HV	IP	EK	EH	HC
	Frørup	Frørup	Haastrup	Vissen-bjerg	Odense	Odense	Odense/ Svendborg
	D +ap.	D +ap.	D -ap.	D +ap.	RSD -ap.	RSD -ap.	RSD -ap.
mad̥(ə). / mad̥.	-0.1	0.8	1.7*	0.5	1.1	2.5*	1.9*
snɑg̊(ə). / snɑg̊.	0.3	1.6(*)	1.9*	1.5*	0.4	-0.8	
kʰɛb̥(ə). / kʰɛb̥.	0.2	3.0*		2.7*			
tʰɑg̊(ə). / tʰɑg̊.				-0.4			1.1
sœnɒ / sœn	-0.1	-0.7	1.0	1.8*	-0.3	0.8	0.5
kʰɔm(ə). / kʰɔm.	-2.0*	-1.7	1.4	4.1*	2.8*	0.8	
b̥as(ə). / b̥as.	0.9	4.3*	0.0	2.4*	2.3*	-0.9	-0.6
lasd̥(ə). / lasd̥.				2.0*			
b̥asn̩.[I] / b̥as.	1.3*	-0.2	0.6	3.3*	2.2*	-0.5	
b̥asn̩.[II] / b̥as.	2.2*	0.3	-0.8	2.8*	1.8(*)	-0.8	

However, after a while, I was no longer so sure of the validity of my conclusions, particularly after having discussed the problems with Peter Molbæk Hansen, who told me that his East Jutlandish dialect in Himmerland had vowel lengthening in actual disyllables with short vowel + single consonant. If this should turn out to be a general phenomenon in East Jutlandish dialects, also in those closer to Århus, it would be conceivable that the lengthening of the vowels in the disyllables of the RSD, also in those disyllables which corresponded to dialectal monosyllables without lengthening (e.g. RSD [tʰɑg̊ə] vs. dialectal [thɑʔg̊] with apocope), was due to a generalization from the dialectal actual disyllables.

I therefore found it necessary to undertake recordings of Eastern Jutlandish dialects, and also some extra recordings of actual disyllables in Western Jutlandish. Moreover, I found that more recordings of North Zealandish dialects were required to find the basis of the Copenhagen lengthening.

II. NEW RECORDINGS OF JUTLANDISH DIALECTS

A. INFORMANTS AND MATERIAL

The informants for West Jutlandish were the same as in the preceding investigation, i.e. BT, TA, EA, and JD.

For the recordings of East Jutlandish the following informants were used:

PM, born 1946 in Himmerland, East Jutland, about 40 km south of Aalborg and about 70 km northwest of Århus. He spoke dialect as a child and still speaks it when visiting his family. His Standard Danish has a clear Jutlandish rhythm, but he has adopted the Copenhagen pitch contour. He has lived in Copenhagen since 1960.

NE, born 1906 in Hadbjerg, about 20 km north of Århus, still living in the same area (Halling), well preserved dialect.

NG, born 1917 in Røgen, about 35 km east of Århus, still living there, speaks genuine dialect.

SJN, born 1902 in Gl. Rye about 30 km southwest of Århus and still living there, has preserved his dialect, but there was some interference from the standard language when he translated from the written text.

MK, born 1903 in Tiset about 15 km southwest of Århus, still living in the same area (Astrup), speaks dialect, but somewhat closer to the standard language.

MN, born 1899 in Hundslund about 30 km south of Århus, where she also lives now. She has spent some years near Århus as a teacher of domestic science, and speaks both dialect and standard language, but keeps them apart.

IR, born 1930 in Malling, about 15 km south of Århus. She now lives in Århus and speaks the Århus RSD. Her dialect does not seem quite genuine. It has some West Jutlandish features, and in some cases she did not have apocope of [-ə].

The last three informants were female, the first four male. The places recorded are shown in the map, figure 1.

The West Jutlandish speakers read sentences containing the following supplementary words: *læs*, *læsse*, *læsset*, *søn*, *sønner*, *ven*, *vende*, *venner*, *tal*, *kalde*, *kalder*, *Kalle*.

The East Jutlandish speakers read the same words and, moreover, their list contained the following words: *tak*, *takke*, *takker*, *tæt*, *tætte*, *tættet*, *sæt*, *sætte*, *sætter*, *danse*, *danser*, *last*, *laster*. PM read more words than the other East Jutlandish informants. His recordings were used as a basis for the choice of a more restricted number of words for the other informants. Besides the words listed above his lists contained the words: *mat*, *matte*, *tække*, *tækker*, *tækket*, *tætter*, *bas*, *basse*, *basser*, *læsser*, *kande*, *kander*, *kom*, *komme*, *dans*, *danse*, *lasten*. Moreover, some of the words in his list were also placed finally in the sentence, but as the difference between monosyllables and disyllables was almost the same in medial and final position, only medial position was used for the other speakers.

The words ending in *-e* have apocope in the dialects, with the exception of the name *Kalle*. (In the tables (ə) indicates old disyllables, not facultative [ə], the apocope being obligatory in Jutlandish.) The ending *-er* is weak and comes very close to [ə] in most Jutlandish dialects. But since at least some of the informants for the present investigation do make a distinction between e.g. *kalder* and *Kalle*, the ending *-er* being slightly lower, I have transcribed *-er* by [ʌ].

No words with stop consonants were included in the West Jutlandish list because they would have West Jutlandish stød, which in the preceding recording had turned out to make difficulties for the segmentation, and which also made the comparison with the monosyllables somewhat problematic.

A few words had to be left out for some speakers because they had special forms, and the words *sæt*, *sætte*, *sætter* were not read by PM and IR.

The words were placed in small sentences, generally consisting of three stress groups with the test word placed in the second group, e.g. *'Else vil 'kalde på'Søren* 'Else will call Søren', *'Drengen hed 'Kalle til'fornavn* 'The boy's first name was Kalle'. In some cases, however, the first potentially stressed word had syntactically reduced stress (but it was less reduced

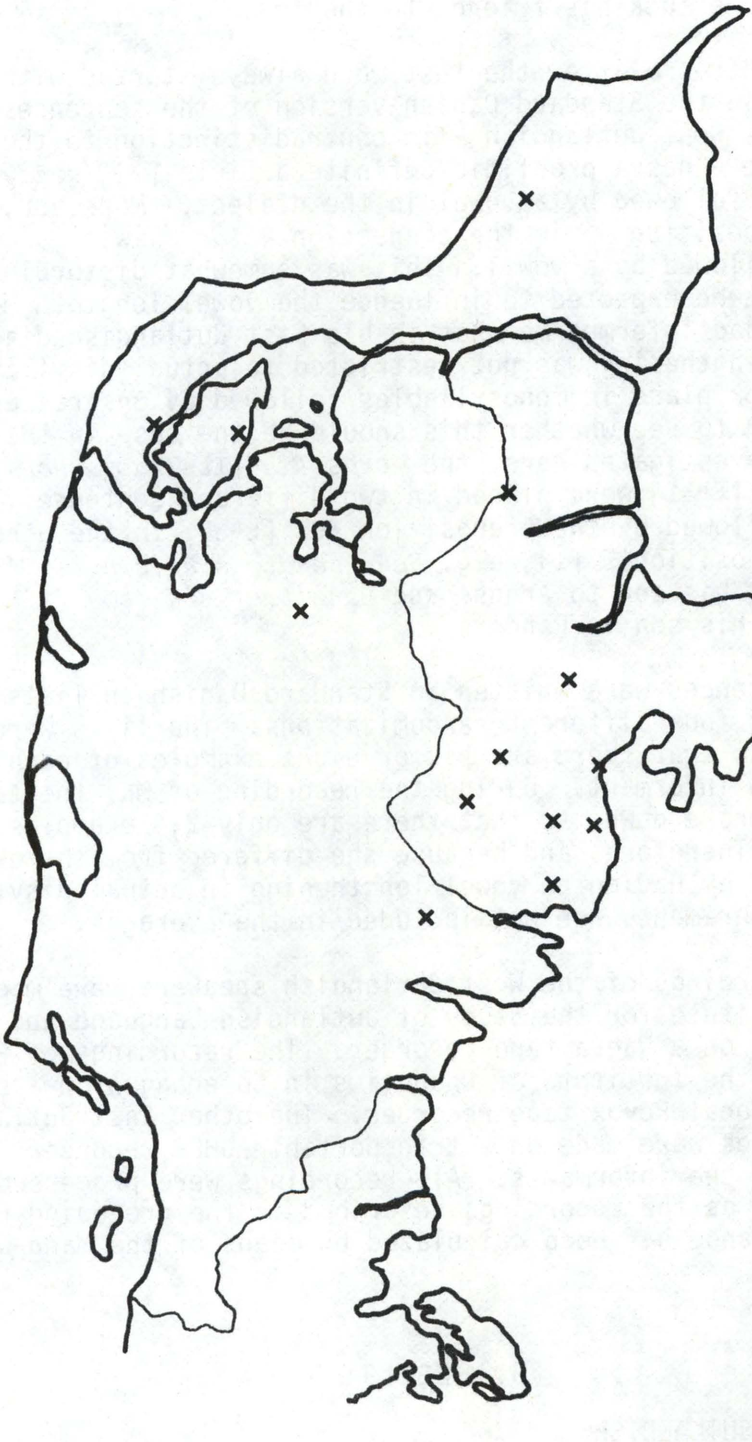


Figure 1

Map showing the Jutlandish recordings. The vertical line separates West and East Jutlandish.

in the pronunciation of the dialect speakers than in the standard language), e.g. *Han ,kørte sin 'ven til statiónen* 'He took his friend to the station' and *Han ,kørte sine 'venner til 'toget* 'He took his friends to the train'.

The first word after the test word always started with a consonant in the Standard Danish version of the sentences. However, as West Jutlandish - in contradistinction to the standard language - has a proclitic definite article [ɛ], *vende* and *læsse* were followed by a vowel in the dialect. Moreover, since the preposition *på* in the connection *kalde på* was *o*, *kalde* was also followed by a vowel. This was somewhat disturbing because it might be expected to influence the vowel length. Molbæk Hansen had informed me that in his East Jutlandish dialect the vowel lengthening was not restricted to actual disyllables but also took place in monosyllables followed by unstressed vowels. In order to see whether this should be the case in the dialects investigated here, the words *læs* [lɛs/las], *søn* [sœn], and *tal* [tʰal] were placed in two different sentences, in one case followed by the preposition *til* [tʰe], in the other by the preposition *i* [i], e.g. *Han ,sendte sin 'søn til 'Århus* 'He sent his son to Århus' and *Han 'traf sin 'søn i 'Randers* 'He met his son in Randers'.

The sentences were written in Standard Danish in lists with three or four different randomizations. The lists were read twice, so that there are six or eight examples of each word for each informant. During the recording of MK, the tape recorder broke down, so that there are only 2-4 examples of some words. Therefore, and because she differed from the other informants by having no vowel lengthening in actual disyllables, her measurements are not included in the averages.

The recordings of the West Jutlandish speakers were made at the Institute for the Study of Jutlandish Language and Culture in Århus on a Nagra tape recorder. The recordings of PM were made at the Institute of Phonetics in Copenhagen on a semiprofessional Revox tape recorder. The other East Jutlandish recordings were made on a transportable UHER recorder in the homes of the informants. All recordings were processed in the same way as the recordings described in the preceding report. Significance has been calculated by means of the Mann-Whitney U-test.

B. RESULTS

1. WEST JUTLANDISH

The results for the West Jutlandish speakers are given in tables II-IV and in figure 2A. The measurements from the two recordings (1982 and 1983) have been combined. The new measurements are indicated by a cross in the tables.

Table II

West Jutlandish

Difference in vowel duration (in cs) between old disyllables with apocope and old monosyllables with phonologically short vowels. Each individual average is based on 6 or 8 tokens. x refers to measurements from the supplementary list. (ə) in parentheses indicates that the word is an old disyllable; it is not a facultative [ə]. Apocope is obligatory. See further the caption to table I.

		BT	TA	EA	JD	average
	maʔd̥(ə)/mad̥	0.4	0.1	-2.4*	0.3	
	tʰɑʔg̊(ə)/tʰɑg̊	1.9*	-0.2	-0.8	-0.2	
	av.	1.2	0	-1.6	0.1	-0.1
	b̥as(ə)/b̥as	0.5				
x	lɛs(ə)/lɛs	0.6		0.4		
	av.	0.6		0.4		0.5
	kʰɔm(ə)/kʰɔm	1.5*	2.1*	1.6*	2.3*	1.9
x	wɛn(ə)/wɛn	0.5	-1.3(*)	0.2	-0.6	
x	kʰal(ə)/tʰal	-1.8*	-0.8	-2.5*	-1.1(*)	
	av.	-0.7	-1.1	-1.2	-0.9	-1.0
	d̥ans(ə)/d̥ans	0.6	-0.1	0.8	-0.2	
	lasd̥(ə)/lasd̥	-0.6	(6.5*)	0.3	0.2	
	av.	0	-0.1	0.6	0	0.1

Table II gives the differences in cs between old disyllables with apocope and monosyllables. It is obvious that there is no difference before stops (average -0.1 cs), nor before *s* (average 0.5 cs). (There are very few examples before *s* because the words used were inadequate (see section IIA).) Nor is there any difference before consonant clusters (average 0.1 cs). (TA is not included in the average for [lasd̥] because he has phonologically long vowel in this position.) The postvocalic consonant did not show any clear difference either.

The results before sonorants are confusing. All four informants have a significant lengthening of the vowel in the apocopated form of *komme* compared to *kom*, and three of the four informants also have a significant lengthening of the postvocalic consonant. However, in *vende* [wɛn] compared to *ven* and in *kalde* [kʰal] compared to *tal* there is no lengthening, neither of the vowel nor of the consonant, rather a tendency to shorten both vowel and consonant ([kʰal] was compared to [tʰal i], not to [tʰal tʰe], because it was followed by a vowel). According to Ringgaard (1959 and 1963) old West Jutlandish disyllables with short vowel followed by sonorant consonant are distinguished from monosyllables by having "dynamic circumflex". This is a somewhat problematic concept. In 1959 Ringgaard describes it as an extra intensity peak. But no measurements are mentioned. What is meant is probably the auditory impression of an extra stress peak or a syllable peak, but stress and syllabicity are more often related to duration and to pitch than to intensity. As mentioned all informants have a lengthening in *komme*, and JD, who is the one who does not lengthen the consonant, has a stronger fall in Fo and in intensity in *komme* than in *kom* (but not two peaks). The others do not have any difference, neither in Fo nor in intensity in any of the word pairs. But the speaker from Vendsyssel (PA) has a second intensity peak in *komme*. - The reason why *komme*, but not *vende* and *kalde* have lengthening should probably be looked for in the segmental and prosodic features of the environment. *komme* and *kom* were followed by a stressed word with initial consonant (*de vil 'komme 'snart, han sagde: 'kom 'så!*), whereas *vende/ven* and *kalde/tal* were followed by an unstressed word which (except for *ven*) had initial vowel. Now, according to oral information from Bent Jul Nielsen and Peter Molbæk Hansen the difference between words of the type *kom/komme* in West Jutlandish is only clearly audible in utterance-final position or before a syntactic break, and at any rate not before unstressed syllables starting with a vowel, unless there is a very clear syntactic break. I did not know this when I constructed the sentences. This may explain the different treatment of the word pairs, but moreover, the imperative *kom!* seems to have an extra short vowel.

Table III gives the differences in vowel duration between actual disyllables and monosyllables. The (very few) examples before *s* do not show any lengthening, but in other cases there is a tendency to a certain lengthening in disyllables, which is very clear and significant in *takker* vs. *tak* (average: 2.4 cs), but which is also - though less consistently - found before sonorants (here the average is 1.2 cs, and the difference is positive in 87% of the cases, though only significant in 56%). The tendency seems to be stronger before *n* (1.6 cs) than before *l* (0.4 cs).

A graphic display of the differences before stops and sonorants is found in figure 2A.

There is a clear tendency to shorten the postvocalic consonant in disyllables. This is true of 19 (or 83%) of the 23 word pairs, the general average being -1.0 cs, but the difference is only significant in six pairs (26% of the cases).

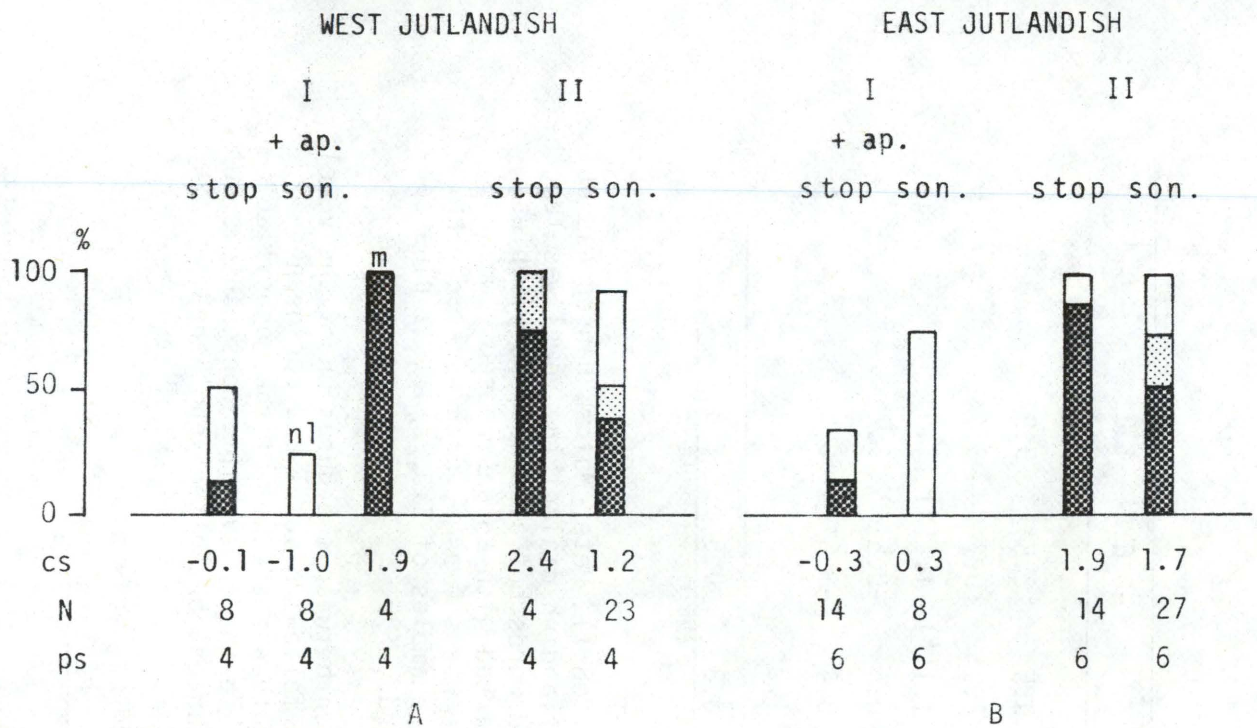


Figure 2

Vowel lengthening in CV́CV(C) words in West and East Jutlandish.
I: Old disyllables with apocope of -e (+ap.). II: Actual disyllables. stop = position before stops, son. = position before sonorants. cs = vowel lengthening in cs compared to corresponding monosyllables. N = number of individual word pairs, each comprising 6-8 tokens. PS = number of involved speakers. The height of the column gives the percentage of positive averages.

Percentage of significant differences at the 1% level.

Percentage of significant differences at the 5% level.

Table III

West Jutlandish

Difference in vowel duration (in cs) between actual disyllables and (old) monosyllables (see further the caption to table I).

		BT	TA	EA	JD	average
	tʰɑʔg̊ʌ/tʰɑg̊	3.5*	2.2*	2.3*	1.5(*)	2.4
	b̥asʌ/b̥as	1.3				
x	kʰasʌ/kʰas(ə)				1.0(*)	
x	lɛsʌ/lɛs(las)	0		-0.6		
	av.	0.7		-0.6	1.0	0.4
	sœnʌ/sœn	0.6	0.5	2.8*	3.9*	
x	sœnʌ/sœn	1.7*	2.0*	0.7	3.4*	
x	wɛnʌ/wɛn	1.5*	1.1(*)	1.1(*)	2.5*	
	d̥anə-/d̥an-	0.8	-0.4	-0.5	3.2*	
x	kʰalʌ/tʰal	0.7	0.4	-1.6*	0.5	
x	kʰalə/tʰal	0.7	1.2*		1.2*	
	av.	1.0	0.8	0.5	2.5	1.2

As mentioned above (in section I), Peter Molbæk Hansen has observed that in his East Jutlandish dialect the vowel was lengthened not only in disyllabic words but also in monosyllables, when the following word started with an unstressed vowel. Therefore, in Table III the words *læsser*, *sønner*, *kalder* and *Kalle* were compared to the examples of *læs*, *søn*, *tal* plus consonant in the following word.

Table IV brings a comparison between the duration of the vowels of *læs*, *søn* and *tal* before following vowel and consonant. There is a tendency to lengthen the vowel of the monosyllable before a vowel. In this case the postvocalic consonant is not shortened as in the disyllables, but often lengthened.

Table IV

West Jutlandish

Difference in word duration (in cs) between old monosyllables before unstressed vowel and unstressed consonant in the following word (see further the caption to Table I).

		BT	TA	EA	JD	average
x	las i / las[1] tʰe	0.6	0.9	0.6	0.4	0.6
x	sœn i / sœn tʰe	1.0(*)	1.3(*)	1.1(*)	1.2*	1.2
x	tʰal i / tʰal tʰe	2.9*	1.2(*)	-0.3	-0.2	0.9

1) BT lɛs

2. EAST JUTLANDISH

The most comprehensive material for East Jutlandish was read by the informant PM, and he also read some of the sentences in Standard Danish. His results are therefore given in a separate table (table V). It appears from the table (V,A) that PM has no lengthening in apocopated disyllables in his dialect but significant lengthening of the corresponding words (with preserved[ə] in his Standard Danish except before consonant clusters, as was also the case with the other speakers of Standard Danish. Before *s* he has a significant shortening of the vowel in his dialect recording. The unexpected shortening of *basse* [b̥as] compared to *bas* [b̥as] can be explained by the fact that *bas* stood before an unstressed vowel (*hans bas er for dyb*), whereas *basse* stood before a consonant. In this case the vowels of *bas* and *basser* have the same length. In his SD version, where both *basse* and *bas* have a vowel after *s*, the vowel of the disyllable is longer than that of the monosyllable. In contradistinction to the West Jutlandish speakers PM has no lengthening of the vowel in *komme*, but in his case the following word was unstressed (*de vil 'komme til 'aften*). The same was the case for *kande* (but here the *n* was lengthened).

In actual disyllables (table V,B) PM has significant lengthening of the vowel both in his dialect and in his SD before stops and sonorants, and the lengthening is of approximately the same magnitude. Before *s* the length is more irregular, like in the standard language; and before consonant clusters there is no significant lengthening. (Since he had no difference between old monosyllables and apocopated disyllables, this latter type has in some cases been used for comparison with the disyllables.)

Table V

PM, East Jutlandish (Himmerland)

Difference in vowel duration (in cs) between (A) old disyllables with apocope and old monosyllables, (B) actual disyllables and monosyllables, (C) monosyllables before vowel and before consonant in the informant's dialect (D) and in standard language (SD). - indicates shorter duration in the disyllable. Stars indicate significance at the 1% level, stars in parentheses at the 5% level. Each individual average is based on 8 tokens. (ə) in parentheses indicates the old disyllables. PM has obligatory apocope in the dialect, but no apocope in the SD words.

A. old disyllables

	D	SD		D	SD
mad̥(ə)/mad̥	-0.2	1.2*	kʰan(ə)/kʰan	0.3	2.0*
mad̥(ə)/mad̊	-0.5	1.9*	kʰɔm(ə)/kʰɔm	0.4	4.0*
tʰɑ̊g̊(ə)/tʰɑg̊	-0.5	4.8*	av.	0.4	3.0
tʰɛd̥(ə)/tʰɛd̥	0.2		lasd̥(ə)/lasd̥	0.6	-0.4
av.	-0.3	2.6	d̥ans(ə)/d̥ans	-0.3	-0.2
lɛs(ə)/lɛs	-1.5*				
b̥as(ə)/b̥as	-2.3*	1.5*			

Table V (cont.)

B. actual disyllables

	D	SD
tʰɑg̊ʌ/tʰɑg̊	4.3*	4.5*
tʰɛg̊ʌ/tʰɛg̊(ə)	2.4*	
tʰɛg̊əd̥/tʰɛg̊(ə)	3.2*	
tʰɛd̥ʌ/tʰɛd̥	1.8*	
tʰɛd̥əd̥/tʰɛd̥	1.6*	
tʰɛd̥əd̥./tʰɛd̥.	1.8*	
av.	2.5	
lɛsʌ/lɛs	-0.3	
lɛsəd̥/lɛs	1.5*	
b̥asʌ/b̥as(ə)	2.3*	
av.	1.2	
b̥asʌ/b̥as(+ə)	0.1	0.9

	D	SD
sœnʌ/sœn	1.9*	1.0(*)
sœnʌ/sœn	2.0*	
sœnʌ./sœn.	1.5*	
kʰanʌ/kʰan	2.1*	2.1*
d̥anə-/d̥an-	1.1(*)	0.7(*)
kʰalʌ/kʰal(ə)	1.5*	
kʰalə/tʰal(ə)	2.6*	
av.	1.8	1.3
d̥ansʌ/d̥ans(ə)	-0.1	
l̊asd̥n̩/l̊asd̥	0.1	
lasd̥ʌ/lasd̥	0.6	
av.	0.2	

C. monosyllables +V / +C

b̥as ə / b̥as(ə) tʰe	2.3*
sœn i / sœn pʰɔ	1.3*
d̥ans(ə) i / d̥ans(ə) pʰɔ	0.1
lasd̥ i / lasd̥ mɛ	0.3

The results for the other East Jutlandish speakers are given in tables VI-VIII. Parentheses indicate averages of less than five tokens. They are not included in the general averages. MK has been left out altogether in the general averages because she differs from the other informants in the duration of actual disyllables.

Table VI shows that there is no consistent difference between monosyllables reflecting respectively old disyllables with apocope and old monosyllables (just as for the other Jutlandish speakers). However, three informants have a significant lengthening of the vowel in *læsse* compared to *læs*. A possible explanation is that for two of these informants the word *vognen*, following *læsse*, started with a vowel (the diphthong [uə]), whereas for the others it started with a [v], and the third (IR), who was influenced by West Jutlandish, had a proclitic article [ɛ] before the [v].

Table VI

East Jutlandish

Difference in vowel duration (in cs) between old disyllables with apocope and old monosyllables. - indicates shorter duration in old disyllables. Stars indicate significance at the 1% level, stars in parentheses at the 5% level. Speakers are indicated by initials. Each individual average is based on 6 tokens. Numbers in parentheses are based on 2-4 tokens and are not included in the general averages. (ə) in parentheses indicates the old disyllable. There is obligatory apocope.

	NE Hadbjerg	NG Røgen	SJN Gl. Rye	IR Malling	MN Hunds-lund	MK Tiset	av.
sɛd̥(ə) / sɛd̥	-1.4(*)	0.1	-1.0(*)		0.2	(0.8)	
tʰɛd̥(ə) / tʰɛd̥			-1.0	(2.1)	1.3*	(1.2)	
tʰɑg̊(ə) / tʰɑg̊	-0.9	0.2	[1]	0	-0.1	-0.3	
	-1.2	0.2	-1.0	0	0.5		-0.3
lɛs(ə) / lɛs	-0.2	2.0*	0	1.4(*)	1.2(*)	1.6	0.9
vɛn(ə) / vɛn			(0.1)	0.3	(2.1)	(1.7)	
kʰal(ə) / tʰal	0.8	-0.5	0.1	-0.4	0.8	(-0.8)	
	0.8	-0.5	0.1	-0.1	0.8		0.2

[1] SJN's pair *tak/takke* has been left out, because translating from the written list he had pronounced final ə in *takke* in five of the six cases. And in all these cases he had a lengthening of the *a* (2.3 cs), whereas there was no lengthening in the example pronounced with apocope.

Table VII shows that also these East Jutlandish speakers have lengthening of the vowel in actual disyllables with a single consonant, with the exception of MK (who is not included in the general averages). Although in many cases MK had few examples, it was evident that she did not make any distinction.

For *læs*, *søn* and *tal* the examples before following consonant have been used for the comparison with disyllables.

Table VII

East Jutlandish

Difference in vowel duration (in cs) between actual disyllables and monosyllables (see further the caption to table VI).

	NE	NG	SJN	IR	MN	MK	av.
tʰɑg̊ʌ tʰɑg̊	2.7*	1.4*	2.7*	2.1*	0.9	(0.5)	
tʰɛd̥əd̥ tʰɛd̥			0.1	2.3*	1.9*	(0.5)	
sɛd̥ʌ[1]) sɛd̥	(1.1)	(2.9)(*)	(-0.9)		(1.1*)	(0.5)	
av.	2.7	1.4	1.4	2.2	1.4		1.8
lɛsʌ lɛs	0	1.7(*)	0.3	1.3*	1.4*	0.6	0.9
sœnʌ sœn	1.4*	0.9(*)	2.1*	1.9*	3.8*	(0.3)	
vɛnʌ vɛn	0.8	2.1*	1.3	2.4*	1.2	(-0.3)	
kʰalʌ tʰal	1.6(*)	0.5	1.6	-0.1	2.4*	(-2.5)	
kʰalˊə tʰal	0.8	1.2	2.6*	1.6(*)	2.3*	(0.2)	
av.	1.2	1.2	1.9	1.5	2.4		1.6
dansʌ d̥ans(ə)			0.5		0.4	(-0.8)	
lasd̥ʌ lasd̥	0.4	0	0.5	-1.2(*)	0.4	(0.7)	
av.	0.4	0	0.5	-1.2	0.4		0

[1]) The pair *sætter/sæt* has been excluded from the averages, although there were sufficient examples, because *sæt* was followed by a vowel. The pair is, however, not consistently different from *tak/takker*.

The lengthening of the vowel before stop consonant is of somewhat smaller magnitude than for the informant PM and, on the whole, there is less significance. Whereas PM had significant lengthening in all 13 pairs before stop and sonorant consonant, the lengthening for the other East Jutlandish speakers - though positive in 27 of the 28 pairs - is only significant in 18 (or 64%) of the pairs.

PM and the other East Jutlandish speakers have been combined in the graphic illustration of vowel lengthening before stops and sonorants in figure 2B. Both PM and the other East Jutlandish speakers have a clear tendency to shortening of the following postvocalic consonant in disyllables. It is shortened in 91% of the cases (in 61% significantly), the general average being -1.8 cs.

Table VIII shows that these speakers also have a tendency to lengthening in monosyllables before a following unstressed vowel, though not of the same magnitude as within a disyllabic word. The following consonant is shortened in *søn* and mostly in *tak* before vowel.

Table VIII

East Jutlandish

Difference in vowel duration (in cs) between monosyllables before vowel and before consonant in the following word (see further the caption to table VI).

	NE	NG	SJN	IR	MN	MK	av.
t^hɑɡ̊ i t^hɑɡ fɒ	1.5*	0.5	0.9(*)	0.5	1.1	(-0.3)	0.9
lɛs i lɛs t^he	-1.1(*)	0.5	-0.2	1.2*	1.2*	(0.2)	0.7
sœn i sœn t^he	0.4	0.9(*)	0.6	2.8*	2.6*	(-0.2)	1.4

There is thus more vowel lengthening in Jutlandish dialects than I assumed in my previous report, but only in actual disyllables (except for some cases before sonorants in West Jutlandish).

III. NEW RECORDINGS OF NORTH-ZEALANDISH DIALECTS

A. INFORMANTS AND MATERIAL

The following new informants were used:

PA, born 1898 in Hove about 20 km west-northwest of Copenhagen and still living there. She has consistent apocope, and her dialect seems genuine.

LJS, born 1900 in Snostrup about 30 km northwest of Copenhagen. He has now for some years lived in Slagslunde, 10 km closer to Copenhagen. His dialect is also well preserved. He does not have apocope except in *takke* and *vende*.

OJ, born 1902 in Gundsømagle about 25 km west-northwest of Copenhagen; he has now for some years lived in Frederiksværk about 55 km north-west of Copenhagen. He is somewhat closer to the standard language.

LJF, born 1897 in Auderød near Frederiksværk, now living in Frederiksværk; also somewhat closer to the standard language.

KJ, born 1890 in Gilleleje on the north coast of Zealand, about 60 km north-northwest of Copenhagen. He speaks a more genuine dialect.

LJS, OJ, LJF and KJ did not have apocope in the recordings; this has also been controlled in free conversations for LJF and OJ. Here, too, apocope was rare except after sonorants.

Moreover, a new recording was made of EP, Dragør, because the first recording was technically bad.

The text used for the five new informants was the same as that used for the East Jutlandish speakers, i.e. it contained the words *tak*, *takke*, *takker*, *sæt*, *sætte*, *sætter*, *tæt*, *tætte*, *tættet*, *læs*, *læsse*, *læsser*, *søn*, *sønner*, *ven*, *vende*, *venner*, *tal*, *kalde*, *Kalle*, *kalder*, *danse*, *danser*, *laste*, *laster*, placed medially in short sentences. The words *læs*, *søn* and *tak* were found in two positions: before an unstressed word beginning with a vowel, or with a consonant. LJF did not read the sentences with the words *sæt*, *sætte*, *sætter*. The new recording of EP comprised the words: *tak*, *takke*, *takker*, *væk*, *vække*, *vægge* [vɛ·gə], *mat*, *matte*, *søn*, *sønner*, *kom*, *komme*, *ven*, *penne*, *pæne*, *vask*, *vaske*. The sentences were read six times by all speakers except by KJ, who got tired and stopped at the third reading.

The recordings were made in the homes of the speakers, using a Nagra tape recorder. The material was processed in the same way as in the other cases.

Since the recordings covered only parts of North Zealand (particularly the dialects directly north of Copenhagen were missing), I also listened to a number of older tapes of connected texts, recorded by the Institute of Danish Dialectology. Those mentioned in my 1982 report (from Stenløse, Uvelse, Karlebo, Græse and Tisvilde) were supplemented by recordings from Farum, Blovstrød, Alsønderup, Tårbæk, Skovshoved, Esbønderup and Hornbæk, all north of Copenhagen, and Greve, south-west of Copenhagen. The places are indicated on the map, figure 3.

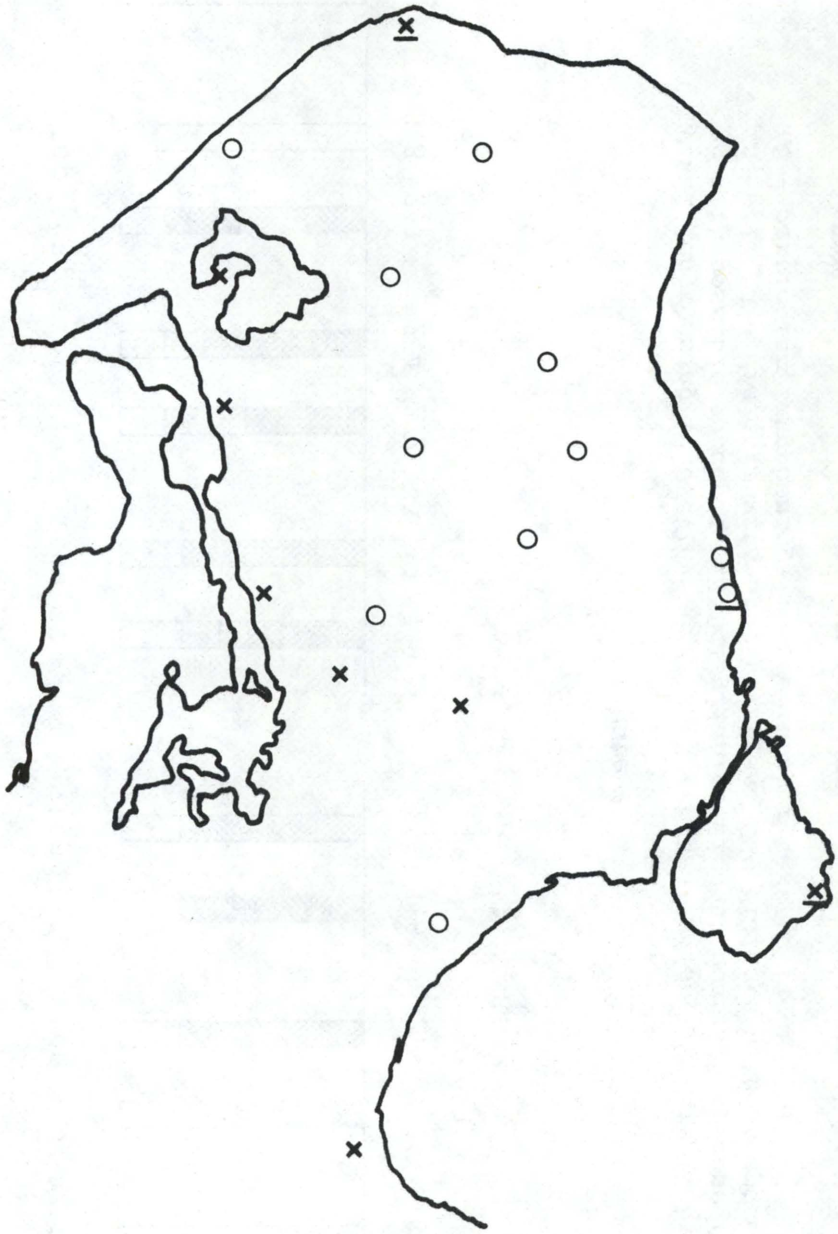


Figure 3

Map showing the recordings of test words (crosses) and of connected texts (small circles). The places with particular lengthening (group C) are underlined.

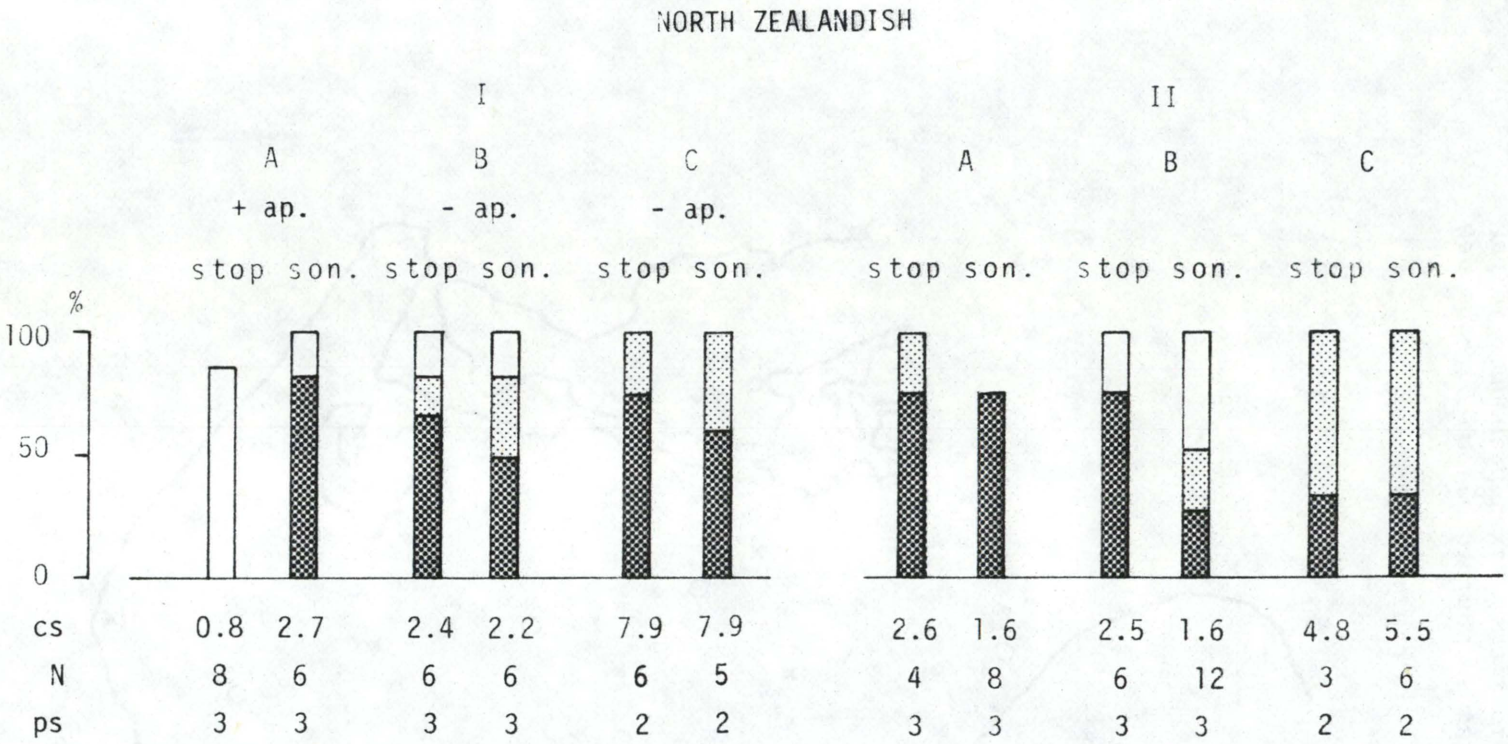


Figure 4

Vowel lengthening in CV́CV(C) words in North Zealandish. I: disyllables in -e with possible apocope. II: disyllables without possible apocope (mostly in -er). A = dialects with apocope, B and C = dialects without apocope (C = particular lengthenings). See further text to figure 2.

Percentage of significant differences at the 1% level.

Percentage of significant differences at the 5% level.

B. RESULTS

The results of the measurements are given in tables IX and X and (for stops and sonorants) in graphical form in figure 4. The two earlier informants (KH and OJ) have been included, and the two recordings of EP have been combined (the tempo was the same). The 8 informants are divided into three groups: A with apocope, B without apocope, and C without apocope and with very conspicuous vowel lengthening. The speakers with apocope do not belong to a geographically coherent area (PA 20 km west-northwest of Copenhagen, KH 40 km northwest, and OJ 40 km south of Copenhagen). It is rather a question of more or less influence from the standard language. The two with particular lengthening (EP and KJ, group C) are not close to each other either, but this strong lengthening may be characteristic of the coastal dialects. Similar lengthenings were found in a connected text read by an old speaker in Skovshoved at the coast about ten km north of Copenhagen.

It appears from tables IX-X that the speakers who have apocope (group A) are in agreement with the Jutlandish speakers (who all had apocope) in so far as they do not generally have any significant vowel lengthening in apocopated words with obstruents (table IX), whereas they have a significant lengthening in actual disyllables (table X).[2] There is, however, disagreement between the Jutlandish and the North Zealandish speakers in the case of old disyllables with sonorant consonants. Here the East Jutlandish speakers did not have any lengthening (there was merger between old disyllables with apocope and old monosyllables in all cases), and the West Jutlandish speakers had only lengthening in some cases, but the North Zealandish speakers have a clear and significant lengthening of the vowel before nasals and before *l*. With the exception of PA's *kalde* there is also lengthening of the following sonorant consonant, so that the sequence vowel + consonant is 5-11 cs longer than in the monosyllable, i.e. there is not merger between monosyllables and apocopated disyllables with sonorant consonant.

Group B (without apocope) has vowel lengthening in all cases before stops and sonorants (but not before *-s* in *læsse*). Sonorant consonants are slightly longer in disyllables with [-ə], whereas other consonants are shorter (2.0 cs, on the average). In disyllables with the endings *-er* [ɒ] and *-et* [əð], both group A and B have shortening of a following obstruent (2.8 cs, on the average), significant in 14 out of 18 averages (78%), whereas sonorants are variable (average 0.2 cs).

Table IX

North Zealandish

Differences in the duration of vowels (in cs) between (old) disyllables and monosyllables. (ə) indicates the old disyllable; ap. means apocope; - indicates shorter duration in old disyllables; a point after the word indicates final position; stars indicate significance at the 1% level, stars in parentheses at the 5% level.

	A (+ap.)				B (-ap.)				C (-ap.)		
	PA (Hove)	KH (St. Havelse)	OR (Bjeverskov)	av.	LJF (Auderød)	OJ (Gundsømagle)	LJS (Snostrup)	av.	KJ (Gilleleje)	EP (Dragør)	av.
mad(ə)/mad̥.	-	1.1	2.8		-	-	-		-	7.9*	
mad̥(ə)/mad̥	-	0.5	0.8		-	-	-		-	4.8*	
sɛd(ə)/sɛd̥	(-1.2)	-	-			(0.6)	(1.3*)		(3.2)	-	
tʰɛd(ə)/tʰɛd	0.7	-	-		1.5	2.0*	4.2*		7.0(*)	-	
tʰɑg̊(ə)/tʰɑg̊	0.9	-0.7	0.6		2.7*	1.3*	2.8(*)		10.9(*)	6.0*	
vɛg̊(ə)/vɛg̊.										8.5*	
av.	0.8	0.3	1.4	0.8	2.1	1.7	3.6	2.4	9.0	6.8	7.9
lɛs(ə)/lɛs	-0.4	-	-		0.1	-0.2	0.6		9.5(*)		
b̥as(ə)./b̥as.		-0.1	2.0				-			2.7*	
m̥is(ə)./m̥is.		-0.5	1.8*		-		-			2.2*	
av.	-0.4	-0.3	1.9	0.4	0.1	-0.2	0.6	0.2	9.5	2.5	
vɛn(ə)/vɛn	2.9*	-	-		1.1	2.8*	1.6(*)		9.9(*)	8.6*	
lan(ə)-/lan-		3.7*	3.4*		-	-	-		-	6.4*	
kʰɔm(ə)/kʰɔm	-	2.3*	1.9			-	-			4.9*	
kʰal(ə)/tʰal	2.2*		-		1.4(*)	2.2*	4.1*		8.5(*)	-	
av.	2.6	3.0	2.7	2.7	1.3	2.5	2.9	2.2	9.2	6.6	7.9
lasd̥(ə)/lasd̥		0.5								1.2	

Table X

North Zealandish

Differences in vowel duration (in cs) between actual disyllables and monosyllables (see further the caption to table IX).

	A				B				C		
	PA (Hove)	KH (St. Havelse)	OB (Bjeverskov)	av.	LJF (Auderød)	OJ (Gundsømagle)	LJS (Snostrup)	av.	KJ (Gilleleje)	EP (Dragør)	av.
sɛd̥ɒ/sɛd̥[1])	(0.7)	-	-			(0.8)	(0.8)		(1.7)	-	
tʰɛd̥əð/tʰɛd̥	1.4(*)		-		1.4	2.5*	4.2*		5.3(*)	-	
tʰɑg̊ɒ/tʰɑg̊	2.5*	2.8*	2.9*		2.8*	0.7	3.3*		7.2(*)	3.4*	
av.	2.0	2.8	2.9	2.6	2.1	1.6	3.8	2.5	6.3	3.4	4.8
sœnɒ./sœn.		3.7*	-1.9(*)						4.5(*)	3.9*	
sœnɒ/sœn	2.6*	2.2*	-0.2		1.7(*)	0.7	1.8(*)				
vɛnɒ/vɛn	2.7*	-	-		0.5	1.6(*)	2.5*		4.4(*)	6.2*	
kʰalɒ/tʰal	2.3*	-	-		1.6	0.3	1.1		4.5(*)		
kʰalə/tʰal[2])	3.8*	-	-		0.2	2.4*	4.2*		10.1(*)		
av.	2.9	3.0	-1.1	1.6	1.0	1.3	2.4	1.6	5.9	5.1	5.5
lasd̥ə/lasd̥	-1.3	-			0.4	0.8	-2.1	-0.3	-2.0	4.3	

[1]) The words *sæt*, *sætte*, *sætter* have not been included in the averages because *sæt* was followed by a vowel.

[2]) The name *Kalle* has been listed together with the words in *-er* and *-et* because (just as in Jutlandish) it does not have apocope.

Group C comprises the two speakers KJ from Gilleleje and EP from Dragør. EP's two recordings (from 1982 and 1983) have been combined in the tables. The pairs *tak/takker*, *mat/matte* and *søn/sønner* were found in both lists. The averages for these pairs are therefore based on 12 tokens. For the first two pairs the difference was exactly the same in the two recordings, but for *søn/sønner* the difference is considerably larger in the second recording. Both speakers have a conspicuous lengthening in disyllables. For KJ the difference beweeen monosyllables and disyllables in [-ə] is 9.2 cs on the average, and for words with other endings 6.1 cs. The distinction is thus very clear, and there are no overlappings, but as he has read only three examples of each word the significance for the individual pairs is only 5% (Mann-Whitney test). His "short" vowels are so long (21.4 cs on the average in words in [-ə]) that one may wonder whether he makes any distinction between short and long vowels. Unfortunately, this cannot be tested since he was not interested in further recordings.

For EP the differences between monosyllables and disyllables is 5.2 cs for disyllables with [-ə] and 4.3 cs for words with other endings. Her short vowels are not as long as KJ's vowels, except in final position. The second recording contained two short/long pairs in this position: [vɛg̊ə/vɛ·g̊ə]: 17.9 and 18.2 cs, and [pɛnə/pɛ·nə]: 19.4 and 24.0 cs. In the first pair the duration of the short vowel is 98% of the long vowel, and there is complete overlapping; in the second pair the percentage is 82, and there is no overlapping. In both cases the short vowel sounds as a long vowel to a speaker of the standard language.

KJ has generally shortening of the following consonant in disyllables (2.1 cs). For EP this could not be measured, since the delimitation of the consonants was uncertain in the first recording and, moreover, most words were found in final position.

As can be seen in table XI, all three groups have longer vowels in monosyllables before an unstressed vowel in the following word when the consonant is a stop or *-n*, though the lengthening is slightly less than in disyllables (before *-s* most speakers have no difference, but in this case there is hardly any lengthening in disyllables either).

For KJ and EP the lengthening is significantly less than in disyllables (only 2.2 cs on the average).

As for the speakers of connected texts to which I listened at the Institute of Danish Dialectology, I did not - with one exception - hear any particularly lengthened vowels. This does not mean that they did not have any vowel lengthening before single consonant, only that this lengthening did not exceed what I expected from the standard language. I had listened beforehand to three of the speakers of the supplementary list, PA, LJS and KJ. In the case of PA I did not notice any lengthenings, although her lengthening before sonorant is approximately as in the standard language (before stops (in non-

Table XI

North Zealandish

Differences in vowel duration (in cs) between monosyllables before unstressed vowel and before unstressed consonant in the following word (see further the caption to table IX).

	A		B		C	
	PA	LJF	OJ	LJS	KJ	EP
tʰɑg̊ i/ tʰɑg̊ fɒ	0.9(*)	2.6*	0.5	1.8	2.7	3.0*
sœn i/ sœn pʰɔ	2.1*	1.5*	0.8	1.9*	-0.5	3.0*
lɛs i/ lɛs tʰe	1.2*	-0.9	-0.3	-0.1	3.3(*)	

apocopated words) it is not quite of the same extent). I had heard some slightly lengthened vowels in the case of LJS, who also has somewhat more lengthening, his vowels in disyllables often being 2-3 cs longer than those of PA. In the case of KJ there was, of course, no doubt. Thus the speakers of connected texts, who generally did not have apocope, may have lengthenings of the same extent as other speakers of group B.

The exception was EM from Skovshoved (born 1890), who had clearly audible lengthenings, particularly before a pause (she does not have apocope). Recordings of selected examples confirmed this impression, e.g. [jɔlɒ] 23 cs, [sɔmɒ] 20 cs, [tʰʁab̥ə] 20 cs.

IV. NEW RECORDINGS OF COPENHAGEN INFORMANTS

A. INFORMANTS AND MATERIAL

The 1982 recordings contained only one speaker of the lower sociolect (ST), and she was not an extreme case. I did not succeed in finding any typical speakers of the lower sociolect who could also read naturally and fluently, but I recorded two young Copenhageners who did not really speak the lower sociolect.

JBC, born 1953 in Copenhagen (Christianshavn). He has spoken the lower sociolect until recently, but it turned out that, as a student of phonetics, he had consciously changed his speech somewhat, not specifically the vowel duration but the quality of the *a*-sounds, which might have involved some other changes, so that he was not the best choice.

ODL, born 1960 in Copenhagen (Gladsaxe). He was one of Molbæk's subjects (see my 1982 report), and one who had overlapping between short and long vowels. But his speech did not sound very pronouncedly "low sociolect".

The words read by the two new informants were: *mat*, *matte*, *tak*, *takke*, *takker*, *væk*, *vække*, *vægge*, *bas*, *basse*, *base*, *søn*, *sønner*, *ven*, *venner*, *kom*, *komme*, *tal*, *kalde*, *last*, *laste*. *mat/matte* were found both finally and medially in the sentence, *væk*, *vække*, *vægge* [vɛ·g̊ə] only finally, the others only medially.

B. RESULTS

The results are given in tables XII and XIII and, for vowels before stops and sonorants, in abbreviated form in figure 5C. (The results for the Standard Copenhagen and for the informant ST are given in figure 5A and 5B for comparison.)

The two new informants have less vowel lengthening than ST, and before nasals + *-er* they have no significant lengthening at all, thus also less than the Standard Copenhagen speakers. They have apocope of *-e* after sonorants, and here they have a significant lengthening of the vowel (see table XII) (it is 1.7 cs on the average for JBC and ODL) and of the consonant (3.9 cs). But the vowel lengthening is of slightly smaller magnitude than in Standard Copenhagen (and the small difference is not due to a longer duration in monosyllables).

Before stops ODL has the same difference between mono- and disyllables as Standard Copenhagen, whereas JBC has a larger difference before [g̊], and particularly finally. In the pair [vɛg̊ə/vɛg̊] in final position both ST and JBC have a considerable difference (8.1 and 5.7 cs, respectively). Before [s] they have more lengthening than in the standard language, and JBC has a rather short and weakly voiced [s] in *basse*. Both JBC and ODL also have a longer vowel in monosyllables before vowel, like the other informants.

The difference between long and short vowels is relatively small for these speakers. The percentual duration of the short vowel is for [vɛg̊ə/vɛ·g̊ə]: ST 82%, JBC 82%, ODL 88%, and for [b̥asɒ/b̥a·sə]: ST 90%, JBC 89%, and ODL 86%. There is overlapping for ODL [vɛg̊ə/vɛ·g̊ə] and ST [b̥asə/b̥a·sə]. Bundgaard (1980) found for 5 Copenhagen speakers 74% for [ɛ] and 78% for [a] before [b̥]. But there is also a difference in absolute duration. For [ɛ] before [b̥] Bundgaard gives the averages: [ɛ] 11.7 cs, [ɛ·] 15.9 cs. But in the present data short [ɛ] before [g̊] is longer than the long [ɛ·] in Bundgaard's material (ST 21.6, JBC 17.0, and ODL 17.7 cs), and JBC and ODL even read the sentences quickly. The difference from a somewhat more conservative norm is still larger. Personally, I read the examples with [vɛg̊ə/vɛ·g̊ə] with the durations 12.2 cs and 19.3 cs (63%).

Table XII

Copenhagen

Difference in vowel duration (in cs) between disyllables in [-ə] and monosyllables (see further the caption to table IX).

	ST	JBC	ODL	av.
mad̥ə./mad̥.	-	4.1*	3.5*	
mad̥ə/mad̥	2.9*	1.2(*)	3.3*	
vɛg̊ə./vɛg̊.	8.1*	5.7*	4.0*	
tʰɑg̊ə/tʰɑg̊	-	5.1*	3.6*	
av.	5.5	4.0	3.6	4.4
b̥asə./b̥as.	4.0*			
b̥asə/b̥as		3.8*	3.7*	3.8
vɛnɒ/vɛn		1.2*		
lanə-/lan-	3.2*			
kʰɔmə/kʰɔm		2.1*	1.8*	
kʰalə/tʰal		2.5*	0.9	
av.	3.2	1.9	1.4	2.1
lasd̥ə/lasd̥	-	0.1	-0.6	

Table XIII

Copenhagen

Difference in vowel duration (in cs) between disyllables in -er [-ɒ] and -en [n] and monosyllables (see further the caption to table IX).

	ST	JBC	ODL	av.
mɔd̥n̩/smɔd̥	5.9*			
sd̥ɔg̊n̩./ sd̥ɔg̊.	4.3*			
tʰɑg̊ɒ/tʰɑg̊	-	3.2*	3.3*	
av.	5.1	3.2	3.3	3.9
sœnɒ/sœn	2.0*	-0.4	1.1(*)	
vɛnɒ/vɛn	-	0.2	0.7	
av.	2.0	-0.1	0.9	0.9

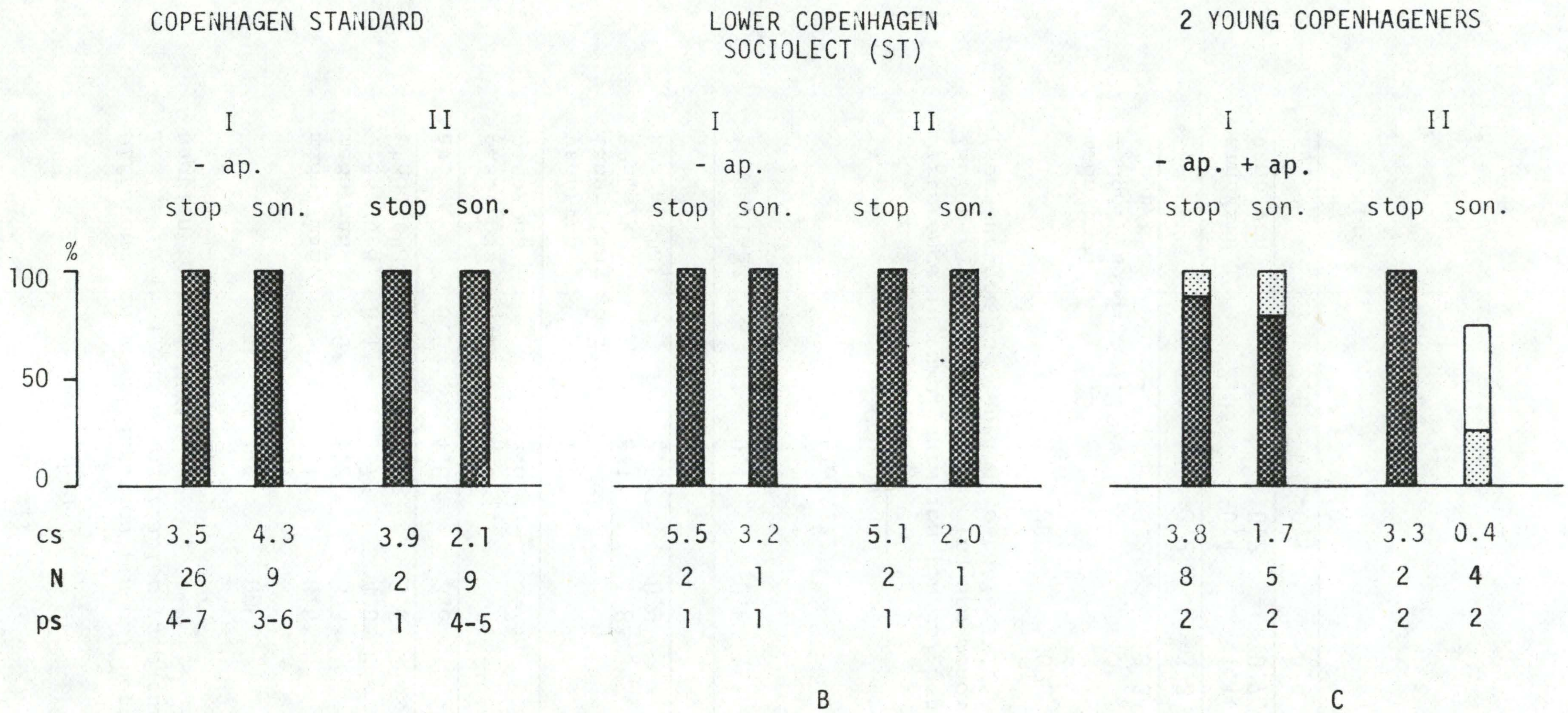


Figure 5

Vowel lengthening in CV́CVC words in Standard Copenhagen (A), the lower Copenhagen sociolect (B), and in the speech of two young Copenhageners (C). See further text to figure 2.

Percentage of significant differences at the 1% level.

Percentage of significant differences at the 5% level.

A feature which seems to be common to the present three speakers compared to the somewhat older speakers of Standard Copenhagen is a stronger tendency to final lengthening. For the Standard speakers the difference in vowel duration between examples with the test word in medial and final position was small - for short vowels on the average 1.5 cs for monosyllables, and 1.6 cs for disyllables (see my 1982 report), whereas JBC and ODL had clearly larger differences (3.0 cs for monosyllables and 3.8 cs for disyllables). For ST the difference is still larger: 5.0 cs and 6.8 cs, respectively. This latter difference is, however, combined with a difference in the following consonant [d̥ -g̊], which can, however, hardly be responsible for more than about 1 cs of the difference.

V. CONCLUSIONS

It has turned out that there is considerably more vowel lengthening in the Jutlandish dialects than it appeared from my first report, but the lengthening is restricted to actual disyllables and is not found in apocopated words except in some cases before sonorants in West-Jutlandish. It is thus not unlikely that the dialect speakers when speaking the standard language generalize this lengthening to all disyllables in the standard language, also to those which correspond to apocopated words without lengthening in their dialect. But this generalization is probably supported by the lengthenings in Standard Copenhagen. This assumption is supported by the fact that the speaker from Vendsyssel, who in his dialect had phonologically long vowels in apocopated words, also before consonant clusters, did not transfer these lengthenings to his RSD, where the lengthening before single consonant did not exceed the standard magnitude, and where there was no lengthening before clusters.

Old informants from Dragør, Skovshoved and Gilleleje at the coast of North Zealand had very conspicuous lengthenings. It is a possibility that lengthening in the old coastal dialect has influenced the Copenhagen sociolect (but the pitch contours of these speakers differed from the Copenhagen norm). The two new younger Copenhagen speakers did not have as much lengthening as expected, particularly not before sonorants, but one of them had more lengthening before [g̊] and both more before [s] than the speakers of Standard Copenhagen. It would be desirable to record more Copenhagen speakers of different generations to find out in which direction the development goes.

As for the possible explanation of the vowel lengthening, there were some relevant findings. In my 1982 report I rejected the hypothesis that the lengthening should have anything to do with the apocope. This rejection was corroborated by the new recordings, in particular the finding that neither the Jutlandish nor the North Zealandish dialects with apocope have any lengthening in apocopated words with obstruents. There is thus no functional reason for the lengthening (e.g. a tendency to preserve the difference between old mono- and disyllables).

One might use this fact in an attempt to date the development, assuming that the vowel lengthening must be later than the apocope, which is old in Jutlandish, but a much more recent phenomenon in Zealandish, probably starting in the last century. However, Molbæk Hansen has found a certain lengthening of the vowel in apocopated words with sonorant in the speech of the older generation in his East Jutlandish dialect, a lengthening which has disappeared in the speech of the younger generation, whereas they have kept the lengthening in actual disyllables. A similar development before obstruents cannot be excluded, and it is thus dubious to date the development on this basis.

The later recordings also supported the rejection of the hypothesis that the lengthening has something to do with the specific Copenhagen pitch contour, which is characterized by a quick rise or jump up from the stressed syllable to the first posttonic syllable, a jump which might be expected to require some preparation and thus a longer vowel in the stressed syllable (although the truncations found by Thorsen (1982) do not make this very probable). It is true that in PM's East Jutlandish dialect and in the Århus RSD (and also in some North Zealandish dialects) the first posttonic syllable is high after single consonant, i.e. just in the case where there is vowel lengthening, but low after clusters. But this apparently striking parallelism cannot be used to support the assumption of an influence from the pitch contour. For in reality the pitch contour is exactly the same in both cases. The higher posttonic in the first case is simply due to the fact that the pitch contour has not come so far down after a short consonant as after a cluster. This appears very clearly from a superposition of the curves with the vowel start as line-up point (see my paper on the manifestation of Danish stress in the next volume of this series). Moreover, there is also lengthening in the dialects which have a lower posttonic syllable both after single consonant and after clusters, e.g. in West- and North Jutland; finally, those speakers who had a particular lengthening did not have particularly high posttonics; one, EM, did not have any rise at all.

What remains is thus the connection with a relatively short and weak following consonant. As an example of the difference between Danish and some other languages in this respect, table XIV gives some measures of comparable Danish, German, and Swedish words. The Danish measurements are averages of 5-7 speakers of Standard Copenhagen, each reading the word 8 times. The Swedish examples are taken from Elert (1964) and represent the average of 8-9 speakers. The German examples (from Fischer-Jørgensen 1969) were read by a typical North German speaker (three times each) in a natural tempo (two other speakers had longer medial consonants). The Swedish examples, of course, contain geminates, and thus are not directly comparable to the Danish words, but the German examples with Vt also show very obvious differences from the Danish examples with V$\mathring{d}$. According to Kohler (1979) the relation of duration between vowel and following closure constitutes the most important cue for German fortis vs. lenis stop. He indicates

Table XIV

Examples of duration of vowels and following consonants (in cs) in Danish compared to Swedish and German.

		V	C	C/V		V	C	C/V
Dan.	*nat*	11.9	6.2	0.57	*natte*	13.2	4.2	0.32
Germ.	*Bett*	9.5	12.0	1.26	*Betten*	8.3	13.5	1.63
Swed.	*tätt*	11.2	17.8	1.59	*sätta*	8.2	10.4	1.27
Dan.	*danne*	13.5	5.9	0.44	*basse*	15.1	8.1	0.54
Germ.	*Kanne*	12.0	10.8	0.90	*Basse*	10.3	19.0	1.84
Germ.	*Kladde*	8.0	10.2	1.16				

the relations by the ratio V/V+C (whereas I have used C/V), and states that for long vowels a ratio higher than 0.70 is a cue for lenis stop, whereas a ratio lower than 0.60 is a cue for fortis stop. The ratios of our subject KV after long [a·] are in accordance with these indications. Kohler does not mention short vowels, which are very rare before lenis stops in German. Here the ratios must, of course, be lower. KV also read the word *Kladde*, and it is conspicuous that his C/V ratio here is also much higher than in Danish *natte*. His V/V+C ratio in this word is 0.46, whereas it is 0.76 in Danish *natte*, which thus has a ratio corresponding to German lenis stops after long vowels, and must be characterized as extremely lenis. - The examples with [n] and [s] show that there are corresponding differences between North German and Danish for other types of consonants as well.

In the Danish dialects the C/V ratio for disyllables is somewhat higher for stops than in the standard language (50% vs. 35%). Before nasals it is the same (around 50%), whereas before [s] it is somewhat higher (around 80%) but still considerably lower than in German. The young Copenhagen speakers have approximately the same C/V ratio as the Copenhagen Standard speakers, but they tend to have somewhat more voicing in the consonants. ODL has in most cases fully voiced medial [g] and [d], and even voiced *f* and *p* in the phrases *tákker for gáven* and *sǿnner på gáden*.

Danish consonants are also heard as having a looser "contact" (German "Anschluss" or "Silbenschnitt") than e.g. German, but this is a rather dubious concept, and since it has been shown (Fischer-Jørgensen 1969) that Danish phoneticians' judgements about "contact" depend mainly on the duration of the vowel, it would be a vicious circle to explain the duration by the loose contact.

Anyhow, it is probable that there is a connection between the weak and short consonants and the lengthening of the vowels, but exactly how and when this development has taken place remains uncertain.

VI. NOTES

1. I have transcribed *b d g* as [b̥ d̥ g̊] to remind the reader of the fact that they are voiceless initially and finally and may be voiceless or weakly voiced medially. Since the only difference between initial *p t k* and *b d g* is the aspiration of *p t k*, it would have been consistent to transcribe *p t k* as [b̥h, d̥h, g̊h]. However, I have preferred to use the more familiar notation [pʰ tʰ kʰ].

2. OJ has a significant difference in *mis/misse*, but in most tokens of this word he had preserved the [-ə]. However, the words without [-ə] do not have a shorter vowel. On the whole, OJ has a much more variable pronunciation than the other informants.

ACKNOWLEDGEMENTS

I am very grateful to the Institute of Phonetics at the University of Copenhagen for the use of their laboratory, and particularly to Svend-Erik Lystlund for competent technical assistance. I also want to thank Lilian Andersen for reliable help with measurements and computations, Niels-Jørn Dyhr for drawing of graphs, and Else Parkmann for competent typing of a difficult manuscript.

I am indebted to the Institute of Danish Dialectology for allowing me to listen to a number of tapes, and particularly to Finn Köster and Magda Nyberg for finding informants on Zealand and in Jutland, and for useful information.

My sincere thanks are also due to the staff at the Institute for the Study of the Jutlandish Language and Culture in Århus for help with the recordings of Jutlandish dialects.

I have profited very much from discussions with Peter Molbæk Hansen; Nina Thorsen and Jørgen Rischel have read the manuscript and given valuable suggestions for improvements.

REFERENCES

Bundgaard, M. 1980: "An acoustic investigation of intrinsic vowel duration in Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 14*, p. 99-120.

Elert, C.-C. 1964: *Phonologic Studies of Quantity in Swedish* (Stockholm: Almqvist & Wiksell).

Fischer-Jørgensen, E. 1969: "Untersuchungen zum sogenannten festen und losen Anschluss", *Kopenhagener germanistische Studien 1*, p. 138-164.

Fischer-Jørgensen, E. 1982: "Segment duration in Danish words in dependency on higher level phonological units", *Ann. Rep. Inst. Phon. Univ. Cph. 16*, p. 137-189.

Hansen, P. Molbæk 1980: "Længdeforhold i himmerlandske en- og tostavelsesord" (unpublished manuscript).

Kohler, K. 1979: "Dimensions in the perception of fortis and lenis plosives", *Phonetica 36*, p. 332-343.

Ringgaard, K. 1959: "Når tostavelsesord bliver enstavelses", *Sprog og Kultur 21*, p. 39-51.

Ringgaard, K. 1963: "The apocope of disyllabics", *Phonetica 10*, p. 222-230.

Thorsen, N. 1982: "Selected problems in the tonal manifestation of words containing assimilated or elided schwa", *Ann. Rep. Inst. Phon. Univ. Cph. 16*, p. 37-100.

(For further references, see my report in ARIPUC 16, 1982.)

ARIPUC 17, 1983, p. 55-86

THE EFFECT OF CONSONANT TYPE ON FUNDAMENTAL FREQUENCY AND LARYNX HEIGHT IN DANISH

NIELS REINHOLT PETERSEN

Fundamental frequency (Fo) and larynx height were measured in words of the type 'CV:fi *where the initial consonant varied in both manner and place of articulation. The main findings were a clear effect of voicing in the obstruents* [f] *and* [v] *on both Fo and larynx height in the following vowel, a small but consistent effect of aspiration on Fo but not on larynx height in the following vowel, and an effect of place of articulation on larynx height which was not, however, reflected in Fo.*

I. INTRODUCTION

In recent years the vertical movements of the larynx during speech have been playing an increasing role in the discussion of possible explanations for the segmentally conditioned fundamental frequency (Fo) variation, particularly the Fo perturbation in vowels after different consonant types. The basis for taking the vertical larynx movements into consideration in this discussion is the observation that - everything else being equal - Fo and larynx height are positively correlated (see e.g. Shipp 1975, Ohala 1973, Ewan 1979).

Greatest attention has perhaps been given to the effect of the voicing distinction in obstruents on the fundamental frequency in the following vowel. It is commonly observed that due to the reduction of the pressure differential across the glottis during the closure/constriction of voiced obstruents Fo is lowered locally in voiced obstruents, and is lower in the following vowel in comparison with Fo after voiceless obstruents (Di Cristo et Chafkouloff 1977, Fischer-Jørgensen 1972, House and Fairbanks 1953, Johansson 1976, Lea 1973, Lehiste and Peterson 1961, Löfqvist 1975, Mohr 1971, Jeel 1975, Hombert, Ohala and Ewan 1976).

The difference is largest at the beginning of the vowel but can also be observed even at the end of the vowel (Löfqvist 1975). It is also a common finding that the larynx is lower in voiced than in voiceless obstruents (Lindqvist, Sawashima and Hirose 1973, Ewan and Krones 1974, Riordan 1980, Westbury 1983). The low larynx position can be assumed to be the result of an active lowering (Kent and Moll 1969, Bell-Berti 1975, Westbury 1983), which contributes to an increase of the volume of the oral cavities, and consequently to the preservation of a sufficient pressure drop across the glottis for voicing to be maintained during the closure/constriction. The low larynx - like the low fundamental frequency - is preserved well into the following vowel. On the basis of this finding in conjunction with the commonly observed tendency for Fo and larynx height to be positively correlated, Hombert, Ohala and Ewan (1976) and Ewan (1979) suggest that the difference in larynx height after voiced versus voiceless obstruents causes (or is an indicator of) a difference in the vertical tension of the vocal chords, which in turn is responsible for the observed fundamental frequency difference.

This explanation, however, meets with difficulties (which the authors are fully aware of) when the relation between Fo and larynx height in and after nasal consonants is taken into consideration. Hombert, Ohala and Ewan (1976) report data indicating that the fundamental frequency course after [m] corresponds very closely to that found after voiced obstruents. This is surprising, since the free passage through the nose in nasal consonants should not neccessitate a larynx lowering and, therefore, a higher Fo should be expected after nasal consonants than after voiced obstruents. Similarly, Lea (1973) reports that while Fo in nasals does not display the local Fo drop characteristic of the voiced obstruents there is no essential Fo difference in the following vowel after the two consonant types. As for the vertical position of the larynx, Ewan (1979) reports that the larynx height of nasals corresponds more closely to that of voiceless obstruents than to that of voiced obstruents; this is true in the consonant as well as in the following vowel (unfortunately the larynx height data are not accompanied by simultaneous Fo data). Thus, there seems to be a discrepancy between the effect of nasal consonants on Fo and their effect on larynx height in the following vowel: Fo behaves like in voiced obstruents, and larynx height behaves like in voiceless obstruents. There are, however, some divergencies among the - rather few - investigations of larynx height in nasal consonants. Lindqvist, Sawashima and Hirose (1973) found like Ewan (1979) a high larynx position in nasals, while Perkell (1969), Bothorel (1979), Riordan (1980) and Westbury (1983) have found larynx heights in nasals similar to those of voiced obstruents. This last is what should be expected on the basis of the low fundamental frequency in and after nasals, but then, of course, the problem is to explain the low larynx position in the nasals. In view of the importance of the Fo- and larynx height conditions associated with nasals for the evaluation of the vertical tension hypothesis outlined above, and considering the divergencies of previous investigations, one purpose of the experiments reported below was to obtain additional data on Fo and larynx height in nasal consonants.

A point, where data from Danish will be of particular interest, is the relation between Fo and larynx height after aspirated versus unaspirated stops. In Danish the two series of stops, *ptk* and *bdg*, are both voiceless, and are differentiated only by the aspirated/unaspirated distinction (the pair *t-d* also by the presence/absence of affrication). This distinction seems to give rise to a fundamental frequency difference in the following vowel, which approaches the magnitude of the voiced/voiceless Fo difference (Jeel 1975, Reinholt Petersen 1978). The effect of aspiration, however, seems not to be universally agreed upon to the same degree as does the effect of voicing. Fischer-Jørgensen (1968) found no Fo difference after Danish aspirated versus unaspirated stops, and data from other languages having an aspirated/unaspirated distinction show no consistent trend for the effect of aspiration on Fo. Now, in the cases where Fo is higher after aspirated than after unaspirated stops, the difference can hardly be accounted for by Hombert, Ohala and Ewan's (1976) vertical tension hypothesis, and it will therefore be of interest to obtain Fo and larynx height data for these types of stops, and consider them together with data for other consonant types in the discussion of the hypothesis.

Although the present study will focus upon the effect of manner of articulation in consonants on fundamental frequency and larynx height, and the relation between the two, the role of place of articulation will also be taken into consideration. The reason for including this feature is that there seems to be a tendency for larynx height to be influenced by place of articulation. Ewan and Krones (1974) and Ewan (1979) have found the larynx to be higher in dental stops than in labial stops, Westbury (1983) reports a high larynx position in velars, lower in dentals and still lower in labials, and Lindqvist, Sawashima and Hirose (1973) have found that the larynx is higher in velars than in dentals. This pattern of variation (viz. labial<dental<velar) does not seem to be reflected in the fundamental frequency of the following vowel.

II. METHOD

A. MATERIAL

The material consisted of test words of the type [CV:fi], where C represents the consonants [b̥ʰ, d̥ˢʰ, g̊ʰ, b̥, d̥, g̊, f, v, m, n] and V: the long vowels [i:], [u:], and [ɑ:][1]. The test words were inserted in the frame sentence *"Vokalen i ... forkortes"*. [vo'g̊ʰæˀln̩ i .. fʌ'g̊ʰɒ:d̥əs] "The vowel of ... is shortened". The 30 sentences were arranged in four different randomizations in a reading list.

B. RECORDING

The recording equipment consisted of a television camera (Sony AVC-3250 CES) and a video-recorder (Sony U-Matic type 2630). The frame frequency of the equipment was the normal 50 frames per second. The speech signal was recorded on the sound track of the video-tape via a Sennheiser MD 21 microphone placed about 15 cm from the subject's mouth. In order to synchronize speech and video signals a timer signal was recorded on the video-tape using a timing device (FOR-A CO. type VTG 33). On playing back the tape, the timer signal was displayed on the monitor screen in minutes, seconds and centiseconds, and could be registered together with the speech signal on an ink writer as pulses for seconds and centiseconds. In this manner it was possible to relate each TV-frame to the speech signal.

During recording the subject was seated in a dentist's chair with a fixed head-rest. The camera was placed at the level of the subject's thyroid prominence and at right angles to his mid-sagittal plane at a distance which allowed the area between the subject's chin and sternum to be covered by the field of vision. For calibration purposes each recording of the speech material was immediately preceded by a short recording of a millimeter scale placed in front of the subject's thyroid prominence in his mid-sagittal plane. Each subject produced 2 recordings of the list, so that altogether 8 repetitions of each test word were obtained.

C. SUBJECTS

Recordings were made of three male speakers, PD, PM, and NR (the author). PM and NR are phoneticians, and PD is an engineer and member of the technical staff of the Institute. They are all speakers of Standard Danish, although PM, who has grown up in Jutland, has some dialectal influence.

D. REGISTRATION AND MEASUREMENTS

The following acoustic curves were made: duplex oscillogram, two intensity curves, and an Fo curve. The timer signal was also registered on the mingograph. On the basis of the acoustic curves the following reference points were determined: 1) the midpoint in time of the pretonic [i] immediately preceding the test word, 2) the beginning of the test consonant, 3) the beginning of the stressed vowel, 4) the midpoint of the stressed vowel, 5) the end of the stressed vowel, and 5) the midpoint of the first posttonic [i].

On the basis of the timer signal and the acoustic curves the locations on the video-tape of the frames or sequences of frames to be measured were determined. Since the interval between frames was 2 cs (the frame frequency being 50 Hz), the temporal inaccuracy of a frame in relation to the corresponding point of reference was $\pm$ 1 cs. The video recorder was equipped with a step function which made it possible during playback to "freeze" the picture and step forward frame by frame and read off larynx height in the frames selected for measurement. Larynx height was determined from a scale drawn on the monitor screen on the basis of the millimeter scale which had been recorded on the tape prior to the reading of the material. The vertical position of the larynx could be measured with an accuracy of $\pm$ 0.5 mm. Larynx height measurements were made at the first reference point, in all frames between the beginning of the test consonant and the end of the stressed vowel, and at the sixth reference point. Similarly, Fo was measured at the first and last reference points and at 2 cs intervals during the voiced portion of the [ˈCV:] sequence.

III. RESULTS

A. STATISTICAL TREATMENT

In order to eliminate a systematic influence on the larynx height measurements, deriving from the fact that it had not been possible to place the millimeter scale in exactly the same position in the two recordings, the larynx height measurements in each of the 8 randomizations were converted into deviations from the mean of all measurements in that randomization, before being submitted to further statistical treatment. After this normalization procedure, means and standard deviations for each test word were computed at reference points 1, 2, 5, and 6. In the [ˈCV:] sequence means and standard deviations were computed continuously throughout the sequence (i.e. at 2 cs intervals) using the beginning of the vowel as a line-up point.

For the further statistical analysis a series of one-way analyses of variance was carried out at each of the 6 reference points.

Figure 1 displays the between group variance estimates (s_b^2) at the reference points. It is seen that the maximum consonantal influence on both Fo and larynx height occurs at the onset of the vowel following the test consonant, and although it decays through the vowel the effect is still significant at the 5 per cent level or better at the end of the vowel in about 50 per cent of the cases.

Fo shows only few cases of a significant effect at points earlier than stressed vowel onset, whereas the significant effects of consonants on the vertical position of the larynx tend to be dispersed over all reference points in the measured sequence, although they occur most frequently at the beginning and at the middle of the stressed vowel.

The analyses of variance, of course, give no detailed information on the pattern of influence of the various consonant types on Fo and larynx height. Therefore an a posteriori multiple comparison procedure had to be applied in order to detect which consonant distinctions were responsible for the observed effects

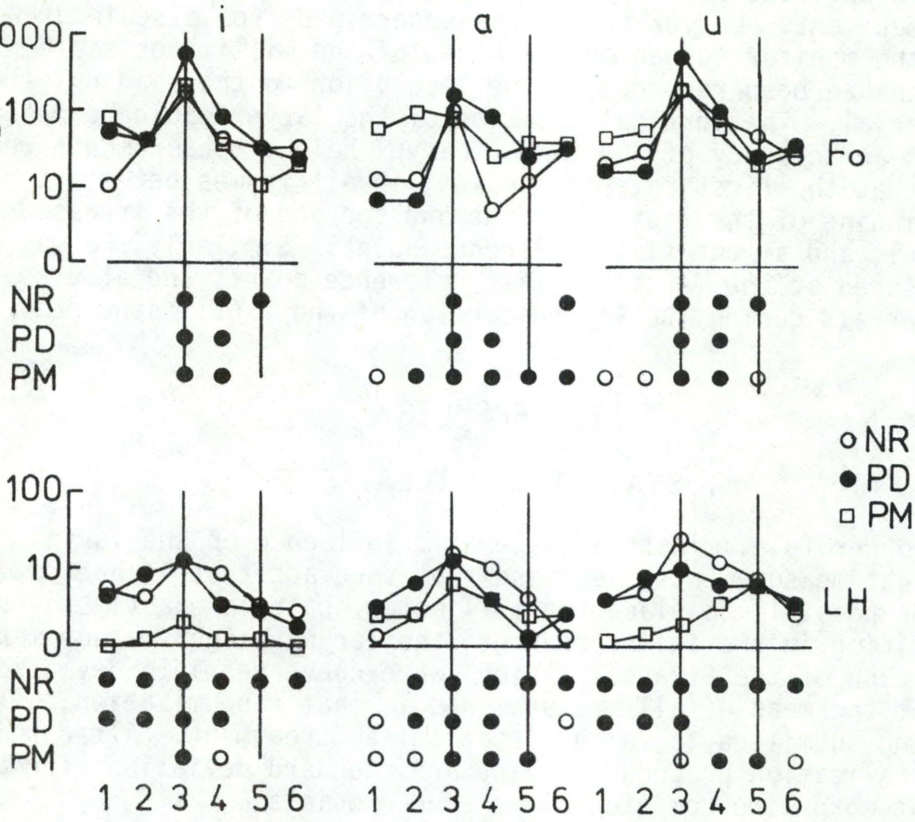


Figure 1

The effect of consonants (expressed in terms of between-group variance estimates) on fundamental frequency and larynx height at the six reference points in [i]-, [ɑ]-, *and* [u]-*words. (Note that a logarithmic scale is used.) Data from three subjects. The circles under the graphs indicate that the effect is statistically significant at the 5 per cent level (open circles), or at the 1 per cent level (filled circles).*

and which were not. The procedure employed in the present study was the Scheffé procedure, which has the advantage that it does not require an equal number of observations in the groups to be compared as do other methods like the Newman-Keuls and the Duncan methods (Ferguson 1976). The demand for an equal number in the groups could not be met in all cases since a number of measurements had to be discarded for various reasons (but no mean represents less than 5 observations and in the vast majority of cases the number of observations is 8). The drawback of the Scheffé procedure is, however, that it focuses on minimizing the probability of Type One errors, i.e. the risk of accepting H_1 (there is a difference) when H_0 (there is no difference) is in fact true. This means that it will lead to fewer significant differences than the other methods. To overcome this bias Ferguson (1976) suggests that a lower level of significance be used, e.g. the 10 per cent instead of the 5 or 1 per cent level. In the treatment of the present data the 10 per cent level was used, but still very few differences were shown to be significant. This does not, per se, discredit the method, but it means that differences which, although they were small, were consistent throughout the material, were obscured by the multiple comparison procedure. An example will illustrate this point: At the midpoint of the vowel the fundamental frequency is higher after aspirated than after unaspirated stops. The difference is quite small, 3 Hz on the average and varying between 0 and 8 Hz, and can be proved statistically significant in only 4 cases out of 27 (3 places of articulation x 3 vowels x 3 subjects). On the other hand, the difference is consistent in the sense that Fo is higher after aspirated stop in 21 cases, in 6 cases there is no difference, and in no cases is Fo lower after aspirated than after unaspirated stop. So, in this instance the multiple comparison procedure has failed to reveal a general and consistent tendency in the data. Because of this, for each pair of consonantal conditions compared, a count was made at each reference point over subjects and vowels (3x3) of the number and directions of the differences between the means of the members in the pair. The results of these counts were taken into consideration in the evaluation of the tendencies in the material.

B. THE INFLUENCE OF MANNER OF ARTICULATION ON THE FUNDAMENTAL FREQUENCY AND ON THE VERTICAL POSITION OF THE LARYNX

Figures 2, 3, and 4 show mean fundamental frequencies, and figures 5, 6, and 7 mean larynx heights at the 6 reference points, cf. above. The data are plotted in such a manner that the effect of the manner of articulation of the test consonants may be readily evaluated. The results of the counts of differences are given in tables 1 and 2.

1. FUNDAMENTAL FREQUENCY

Apart from the local fundamental frequency drop during the [v] which is not found in the nasal consonants (see figure 8), it seems that the effect on Fo exerted by the manner of articulation is confined to the following vowel. There are, of course, differences at the midpoint of the first pre-tonic, at the onset of the consonant, and at the first post-tonic. But these differences do not appear to constitute any systematic pattern.

The consonants can be very clearly divided into two groups according to their influence on Fo at the onset of the following vowel. One group includes the stops and [f], after which Fo is high, and the other consists of the nasal consonants and [v], after which it is low. Within each of the groups, however,

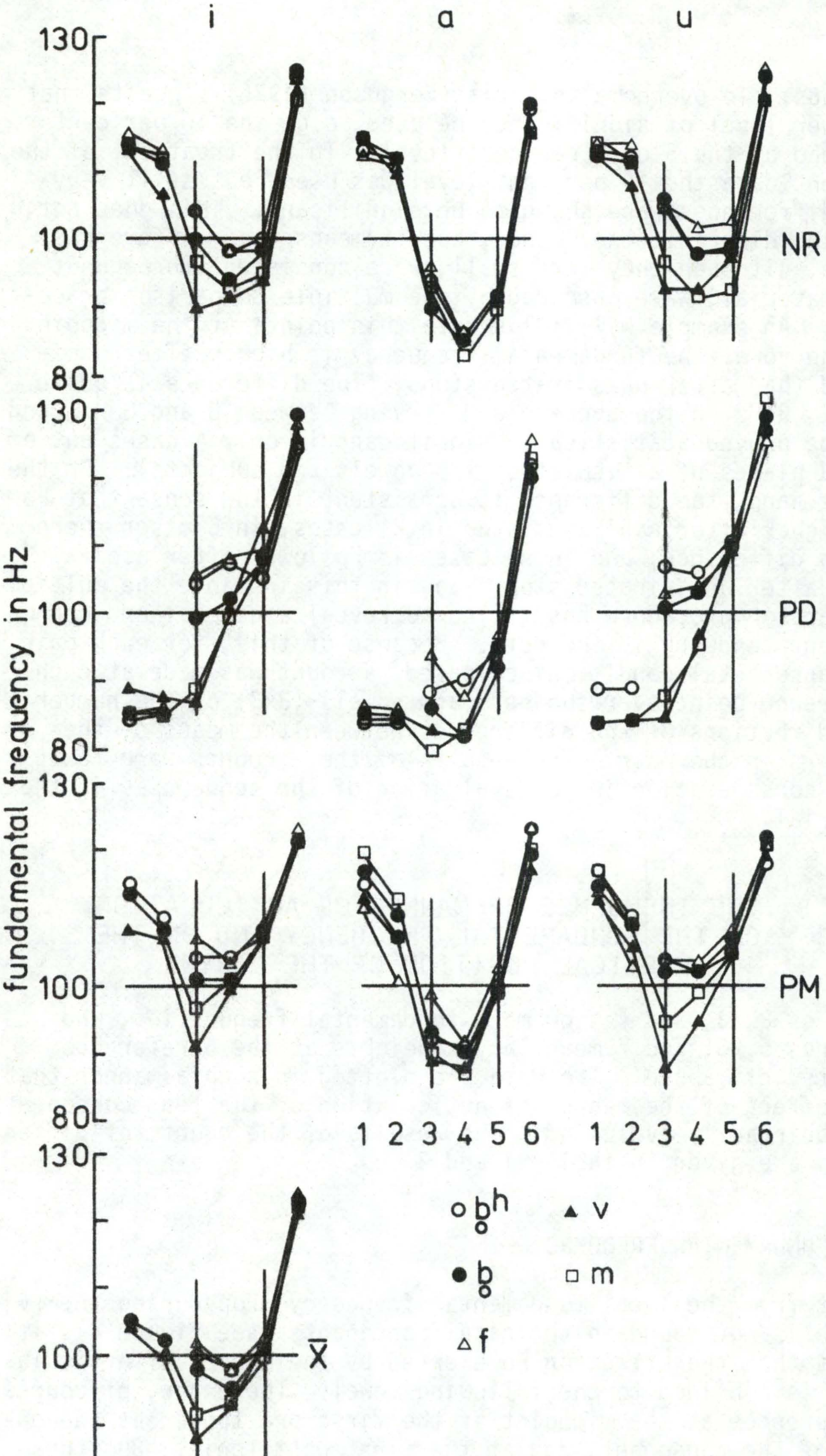


Figure 2

Mean fundamental frequency at six reference points in ['CV:fi] *words with initial labial consonants varying in manner of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

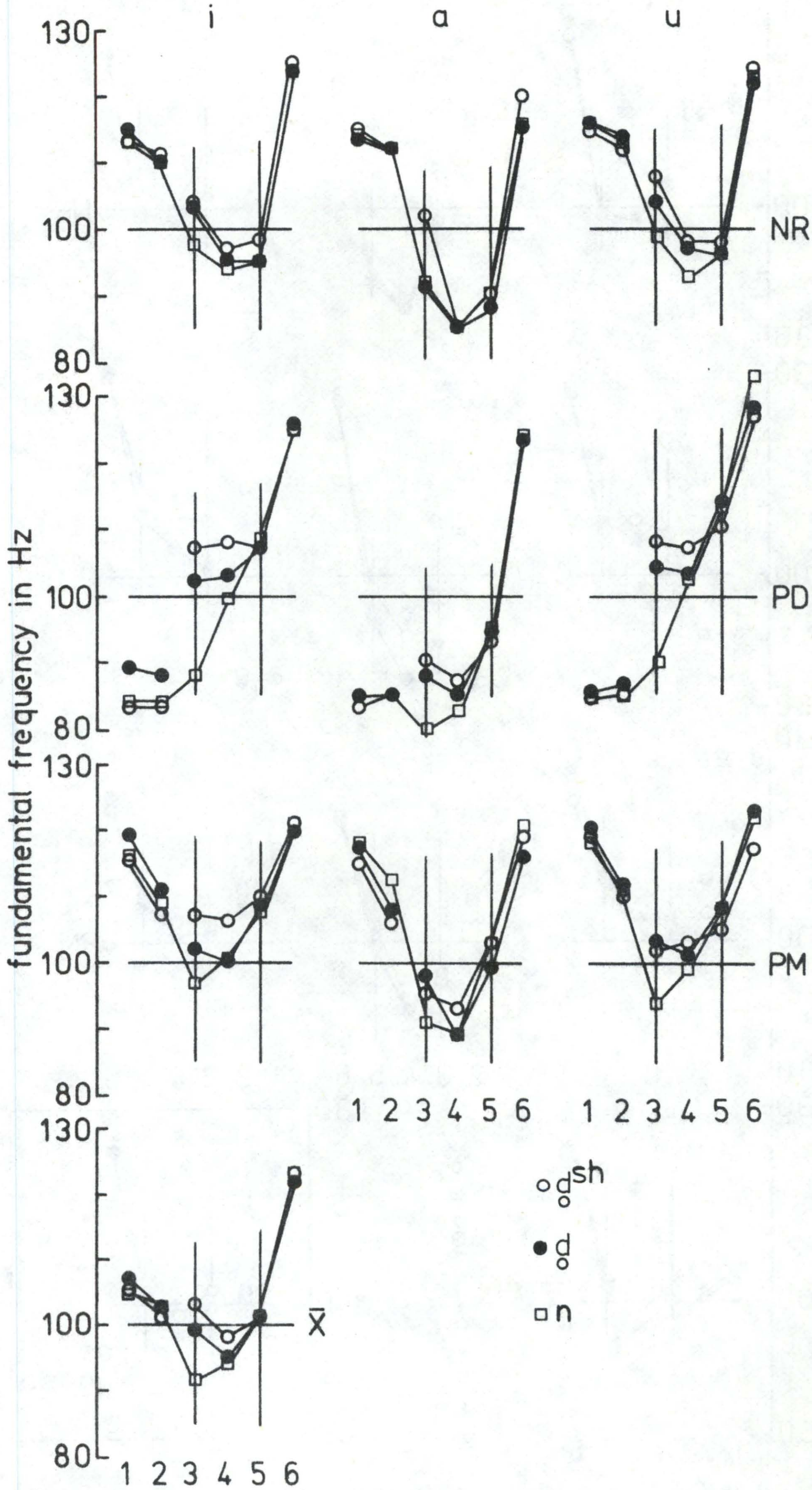


Figure 3

Mean fundamental frequency at six reference points in ['CV:fi] *words with initial alveolar consonants varying in manner of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

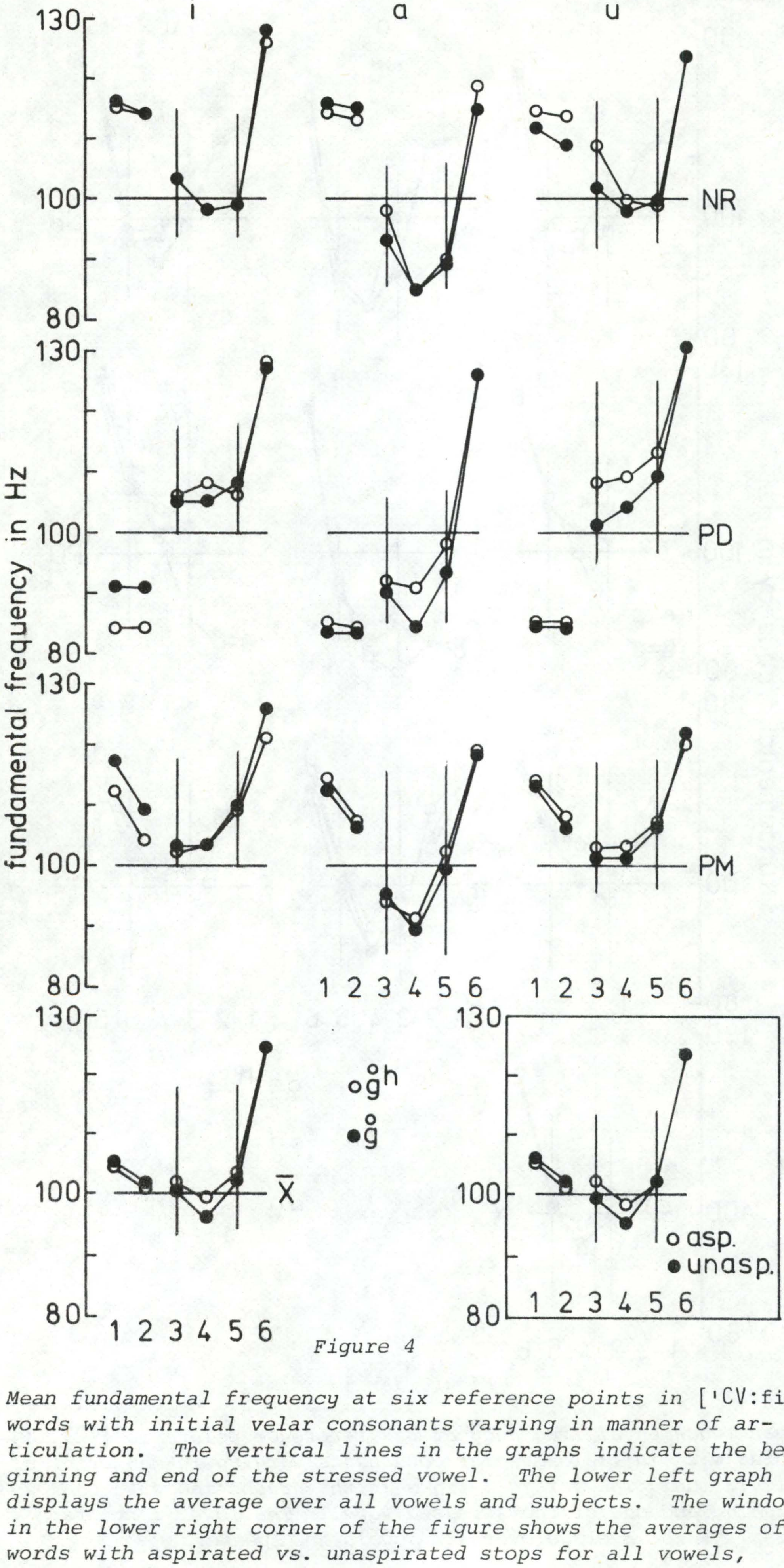


Figure 4

Mean fundamental frequency at six reference points in ['CV:fi] *words with initial velar consonants varying in manner of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and subjects. The window in the lower right corner of the figure shows the averages of words with aspirated vs. unaspirated stops for all vowels, speakers, and places of articulation pooled.*

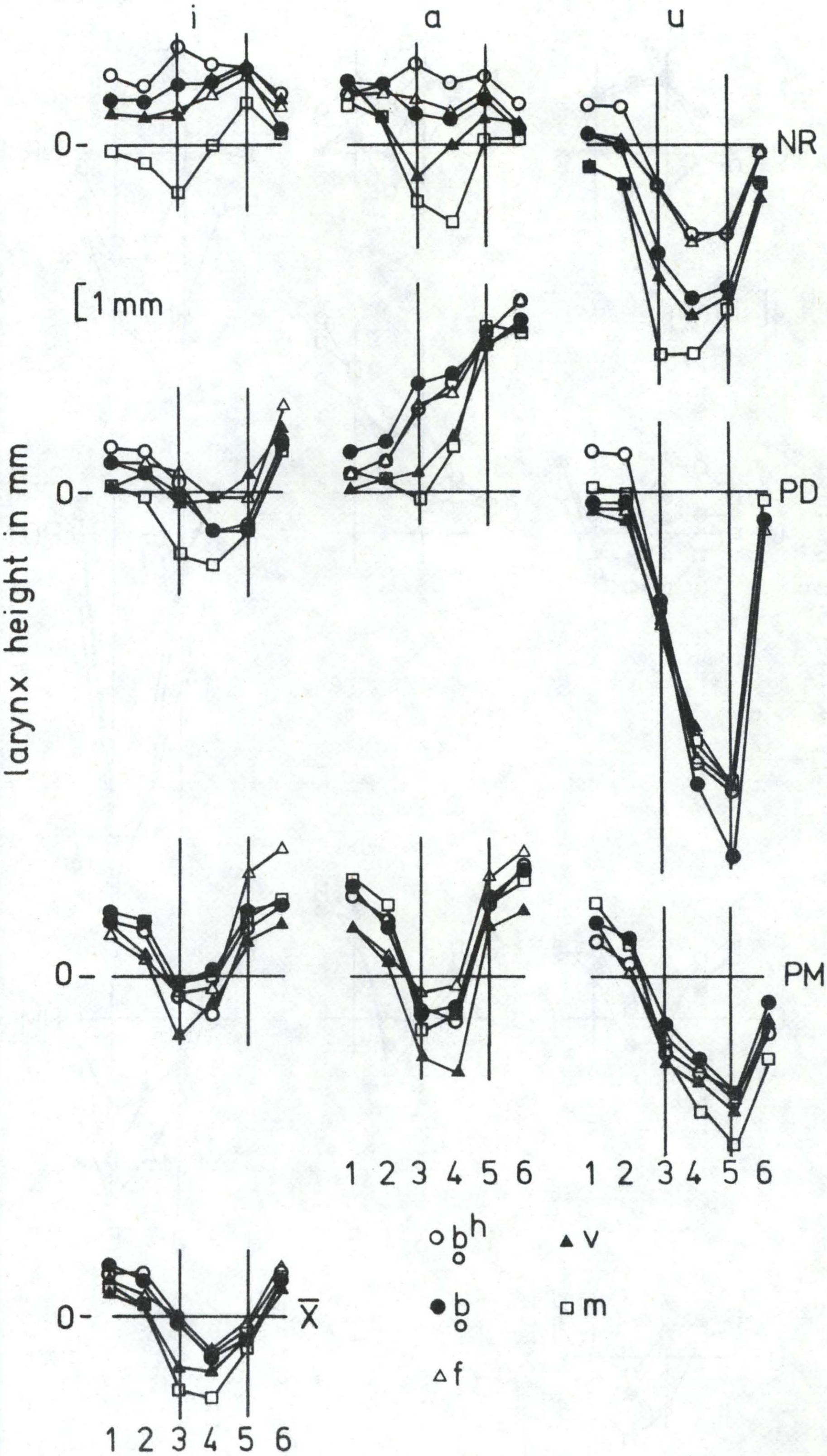


Figure 5

Mean larynx height at six reference points in ['CV:fi]-*words with initial labial consonants varying in manner of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

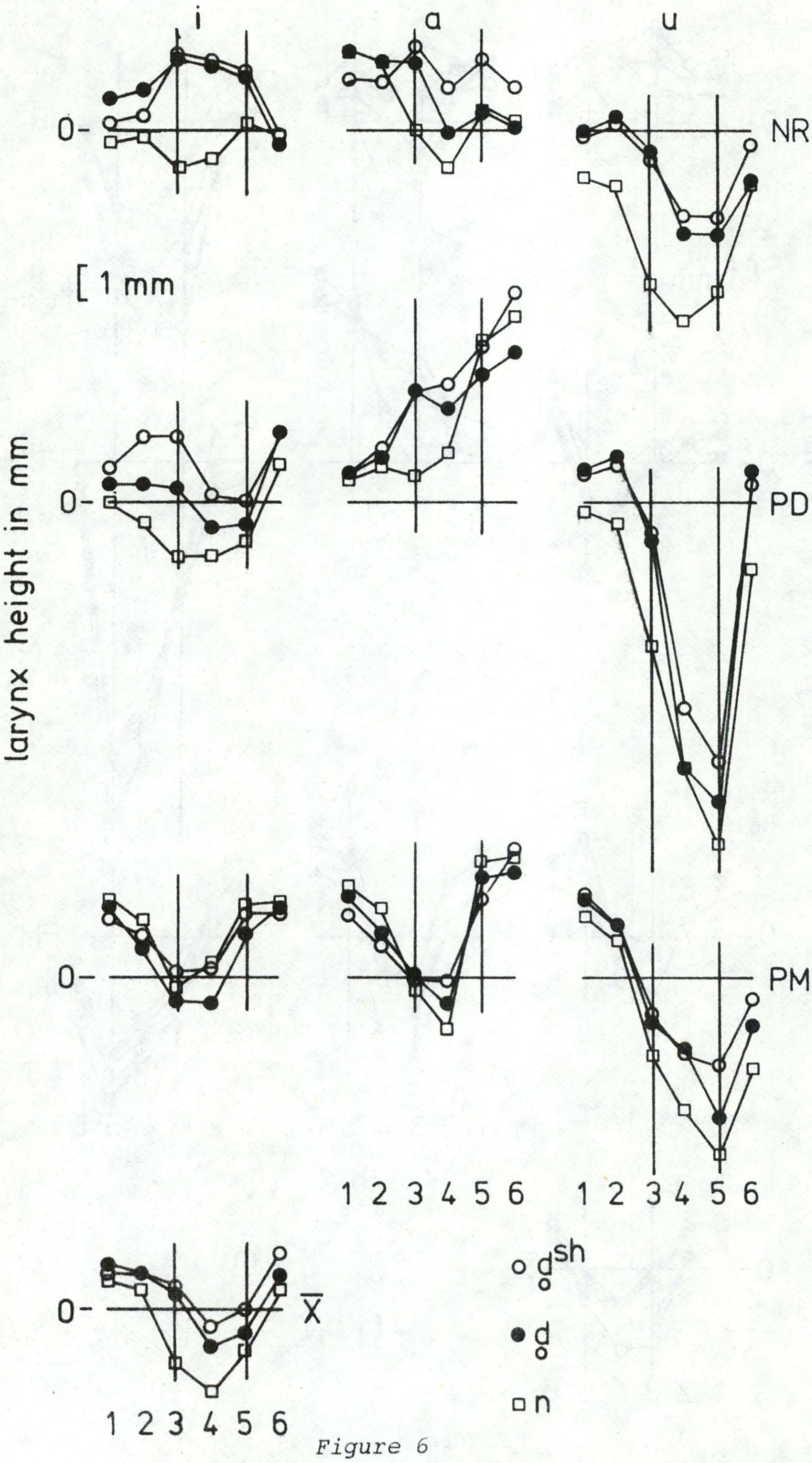


Figure 6

Mean larynx height at six reference points in ['CV:fi]-words with initial alveolar consonants varying in manner of articulation. the vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.

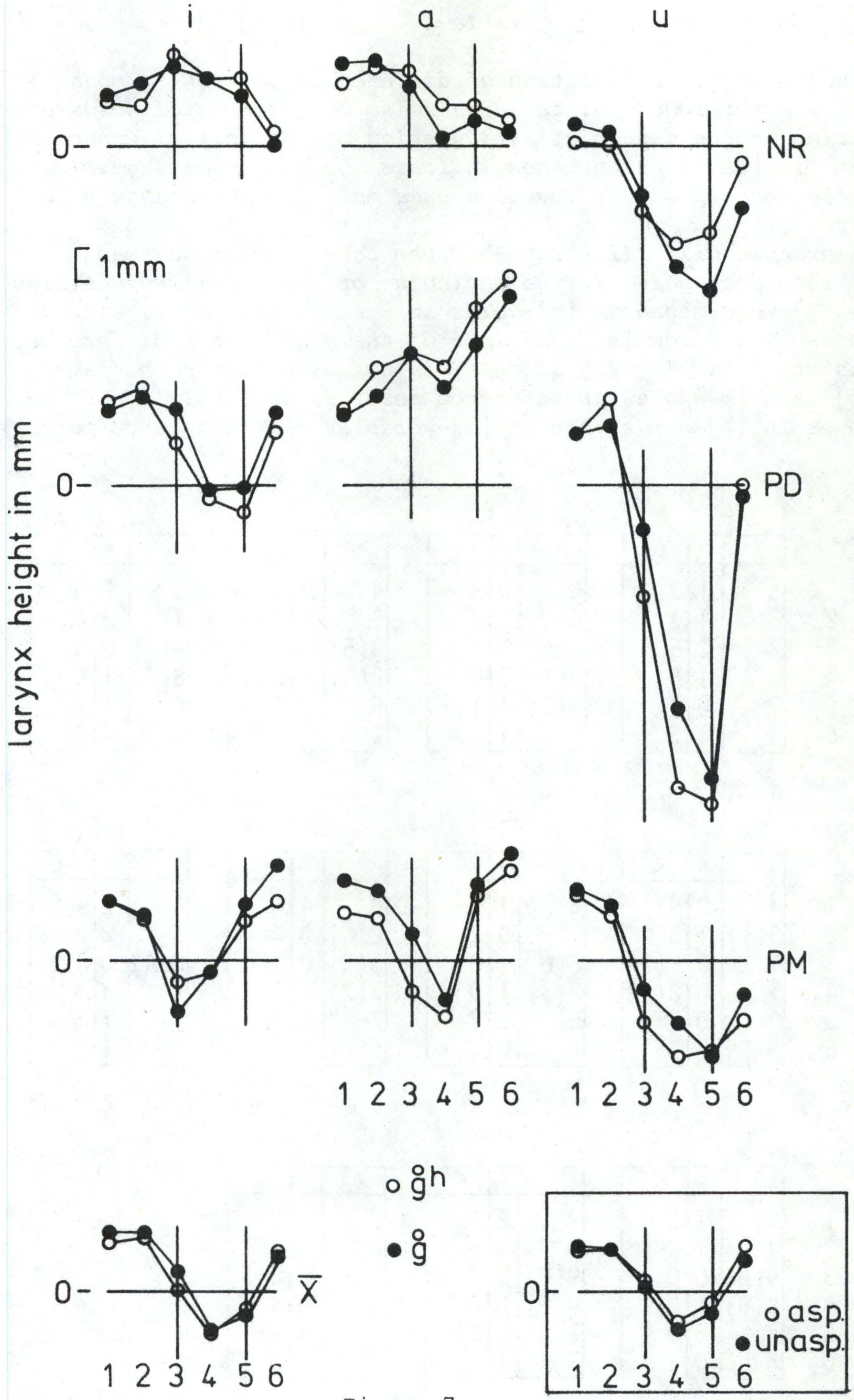


Figure 7

Mean larynx height at six reference points in ['CV:fi]-*words with initial velar consonants varying in manner of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and subjects. The window in the lower right corner of the figure shows the averages of words with aspirated vs. unaspirated stops for all vowels, speakers, and places of articulation pooled.*

Table I

Results of the enumeration of differences between Fo means at the six reference points in pairwise comparisons of words differing in the manner of articulation of the initial consonant. The figures in parentheses indicate the number of significant difference ($p < 0.1$) revealed by a multiple comparison procedure.
An example will illustrate how the table should be read: at reference point 3 (the beginning of the vowel) Fo was higher in [b̥]-words than in [m]-words in 8 cases (out of 9, viz. 3 speakers x 3 vowels), and in 5 of the 8 cases the difference was statistically significant; in no cases Fo was the same in [b̥]- and [m]-words at reference point 3, and in 1 case Fo was lower in [b̥]-words than in [m]-words at that reference point.

b̥^{h}-f

	>	=	<
1	5	1	3
2	5	1	3
3	2	2	5
4	3	1	5
5	2	0	7
6	3	1	5

b̥^{h}-v

	>	=	<
1	6	0	3
2	6	1	2
3	9(6)	0	0
4	9(5)	0	0
5	4(1)	4	1
6	5	1	3

b̥^{h}-m

	>	=	<
1	4	0	5
2	3	4	2
3	9(5)	0	0
4	9(5)	0	0
5	6	2	1
6	4	1	4

d̥^{sh}-n

	>	=	<
1	2	2	5
2	1	2	6
3	9(7)	0	0
4	8(3)	1	0
5	4	1	4
6	3	2	4

b̥-f

	>	=	<
1	4	2	3
2	5(1)	1	3
3	1	0	8(2)
4	1	0	8(1)
5	2	0	7
6	5	0	4

b̥-v

	>	=	<
1	6	1	2
2	7	0	2
3	8(6)	1	0
4	8(3)	1	0
5	5	2	2
6	7	0	2

b̥-m

	>	=	<
1	0	4	5
2	2	3	4
3	8(5)	0	1
4	7(2)	1	1
5	7	1	1
6	5	1	3

d̥-n

	>	=	<
1	6	2	1
2	5	2	2
3	8(4)	0	1
4	5	3	1
5	3	1	5
6	2	1	6

f-v

	>	=	<
1	5	1	3
2	4	1	4
3	9(8)	0	0
4	8(6)	1	0
5	7	0	2
6	8	0	1

f-m

	>	=	<
1	2	2	5
2	3	0	6
3	9(8)	0	0
4	9(6)	0	0
5	7	0	2
6	7	1	1

v-m

	>	=	<
1	2	1	6
2	2	1	6
3	1	0	8(3)
4	5	0	4
5	1	4	4
6	3	1	5

b̥^{h}-b̥

	>	=	<
1	4	1	4
2	5	2	2
3	7	0	2
4	6(1)	3	0
5	4	1	4
6	2	3	4

d̥^{sh}-d̥

	>	=	<
1	1	1	7
2	1	3	6
3	7(1)	0	2
4	8(2)	1	0
5	5	0	4
6	5	1	3

g̊^{h}-g̊

	>	=	<
1	5	1	3
2	5	1	3
3	6(1)	1	2
4	7(1)	2	0
5	5	2	2
6	3	3	3

Table II

Results of the enumeration of differences between larynx height means at the six reference points in pairwise comparisons of words differing in manner of articulation of the initial consonant. See further the caption to table I.

b̥^{h}-f

	>	=	<
1	6	1	2
2	8	1	0
3	4	0	5
4	4	0	5
5	5	0	4
6	3	2	4

b̥^{h}-v

	>	=	<
1	6(1)	1	2
2	8(1)	0	1
3	9(2)	0	0
4	6(1)	0	3
5	6	0	3
6	7	0	2

b̥^{h}-m

	>	=	<
1	6(2)	0	3
2	6(2)	0	3
3	7(3)	0	2
4	6(3)	0	3
5	5(1)	0	4
6	7	0	2

d̥^{sh}-n

	>	=	<
1	6	0	3
2	6(2)	0	3
3	9(5)	0	0
4	8(3)	0	1
5	6(3)	0	3
6	7	1	1

b̥-f

	>	=	<
1	5	3	1
2	6	0	3
3	5	0	4
4	4	0	5
5	3	0	6
6	2	0	7

b̥-v

	>	=	<
1	8	0	1
2	8	1	0
3	9(2)	0	0
4	6	0	3
5	5	0	4
6	6	0	3

b̥-m

	>	=	<
1	5	1	3
2	7	0	2
3	8(2)	1	0
4	8(2)	0	1
5	6	0	3
6	5	1	3

d̥-n

	>	=	<
1	6	1	2
2	6(1)	1	2
3	8(3)	0	1
4	8(2)	0	1
5	5	0	4
6	4	0	5

f-v

	>	=	<
1	4	1	4
2	7	0	2
3	8(3)	0	1
4	6(2)	1	2
5	6	1	2
6	7	1	1

f-m

	>	=	<
1	5	0	4
2	5(1)	0	4
3	7(4)	0	2
4	7(2)	0	2
5	6(1)	0	3
6	7	0	2

v-m

	>	=	<
1	4	0	5
2	2	4	3
3	5(1)	0	4
4	7(1)	0	2
5	5	0	4
6	4	1	4

b̥^{h}-b̥

	>	=	<
1	4	0	5
2	5	1	3
3	5(1)	0	4
4	4(1)	1	4
5	5	2	2
6	6	2	1

d̥^{sh}-d̥

	>	=	<
1	2	0	7
2	3	0	6
3	7	0	2
4	8	0	1
5	8	0	1
6	6	1	2

g̊^{h}-g̊

	>	=	<
1	2	2	5
2	3	0	6
3	3	0	6
4	3	2	4
5	5	0	4
6	5	0	4

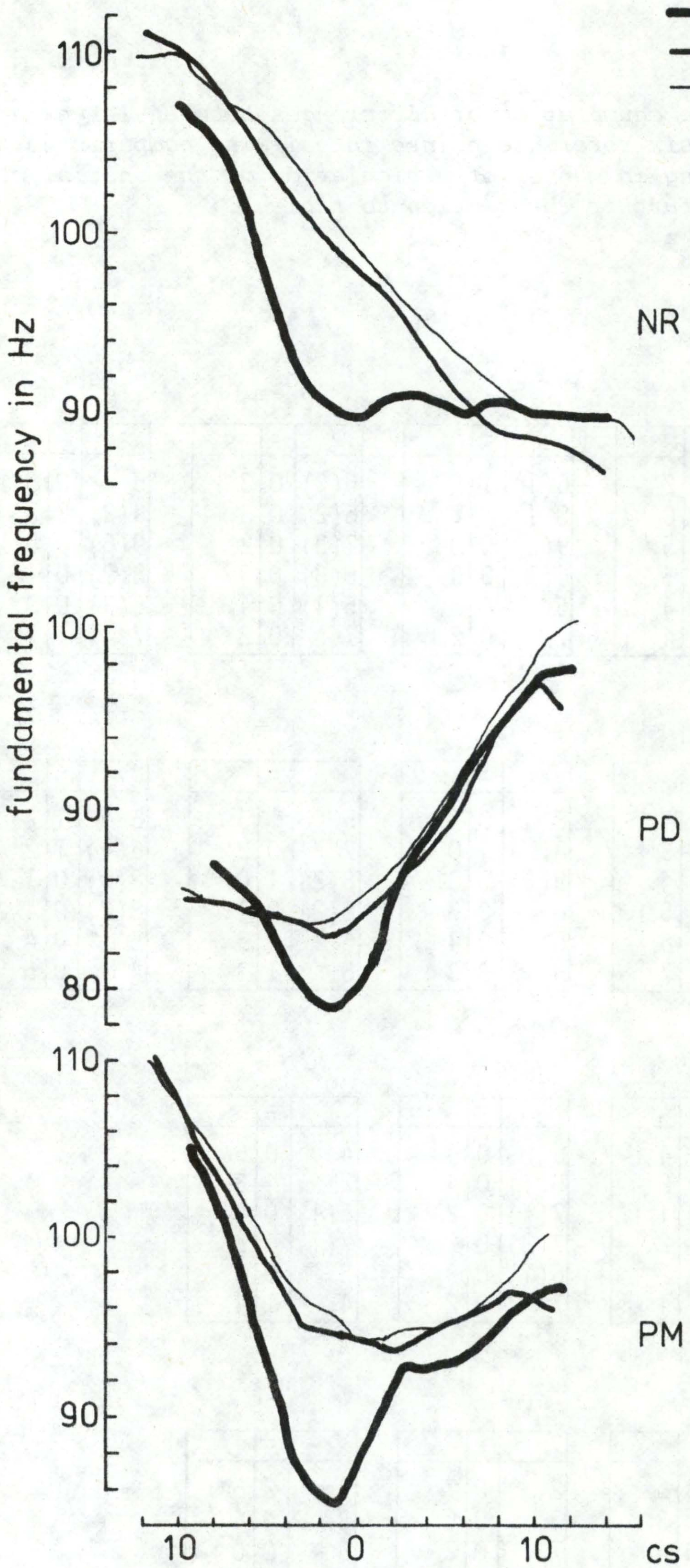


Figure 8

Mean fundamental frequency tracings in sequences of [v], [m], and [n] plus stressed vowel (lined up around the onset of the vowel). [i]-, [ɑ]-, and [u]-sequences are pooled.

there are fairly consistent differences to be found. There is a general tendency for Fo to be higher after aspirated than after unaspirated stops. The difference is small, 2 Hz on the average, but the relation holds in 20 out of 27 comparisons (3 vowels x 3 speakers x 3 places of articulation).

The fricative [f] tends to be followed by a higher Fo at the onset of the vowel than does the corresponding stop [b̥ʰ]. The mean difference is 2 Hz, which is the same as the aspirated-unaspirated difference, but referring to table I it is evident that the [b̥ʰ-f] difference is less stable than the [b̥ʰ-b̥] difference. Thus, on the basis of their effect on the fundamental frequency at the onset of the following vowel it seems reasonable to subdivide the voiceless consonants into two groups, [b̥ʰ d̥ˢʰ g̊ʰ f] versus [b̥ d̥ g̊].

Within the group of voiced consonants the fundamental is in most cases lower after [v] than after [m] at the start of the vowel. This difference must be ascribed to the aerodynamic conditions during the [v]-constriction, which are still effective at or immediately after the release.

At the midpoint of the vowel, which lies between 6 and 10 centiseconds after its onset, the effect on Fo exerted by the preceding consonant is considerably smaller than at the beginning of the vowel, but two groups of consonants can still be clearly discerned, namely the aspirated stops and [f], that are accompanied by a relatively high Fo and the nasal consonants and [v], after which Fo is low. Further, the difference between [m] and [v], which was found at the beginning of the vowel, is eliminated at the midpoint of the vowel. The relation of the unaspirated stops to the two groups is less clear. As is seen from table II Fo after [b̥] is higher than after [v] and [m] in 8 and 7 (out of 9 cases), respectively, and after [b̥ʰ] Fo is higher in 6 cases (out of 9) than after [b̥], i.e. the [b̥ʰ-b̥] difference is slightly less consistent than the [b̥-m] and [b̥-v] differences. On the other hand, the [b̥-v] and [b̥-m] differences are slightly smaller (2 Hz on the average) than the [b̥ʰ-b̥] differences (3 Hz on the average). This tendency comes out more markedly in the alveolars, where Fo after [d̥] is only 1 Hz higher than after [n] but is 3 Hz lower than after [d̥ˢʰ]. Furthermore, the [d̥ˢʰ-d̥] difference is consistent to a high degree, whereas the [d̥-n] difference is not. In the velars, no comparison can be made with voiced consonants, of course, but it should be noted that Fo after [g̊ʰ] is 3 Hz higher than after [g̊] which is identical to the corresponding differences in labials and alveolars. Thus, judging from their influence on the fundamental frequency at the midpoint of the following vowel, the unaspirated stops tend to be more similar to the voiced than to the other unvoiced consonants, where, at the beginning of the vowel, they were clearly separated from the voiced consonants.

At the end of the vowel the effect of the manner of articulation of the preceding consonant is very small, and follows no consistent pattern.

2. LARYNX HEIGHT

As mentioned above the consonantal influence on the vertical position of the larynx seems to be more evenly dispersed over the measured sequence than does the influence on the fundamental frequency.

This tendency can also be read from table II where it appears that the degree of consistency of the outcome of a given comparison remains fairly constant over all points of reference. But apart from that the pattern of consonantal influence on larynx height is subject to gross variation between speakers, and also - to a lesser degree - to variation between the vowels following the consonant.

The most consistent feature of the data seems to be the tendency for nasals to be associated with a lower larynx position than the other consonants, particularly at the onset and the midpoint of the following vowel. The tendency is very clear for subject NR, less so for PD, and it is only found for PM in the [u]-words and at the midpoint of [ɑ] after [n].

In the first pretonic vowel, at the beginning of the consonant, and at the beginning of the following vowel the larynx is generally lower in [v]-words than in [b^h]- [b]- and [f]-words. At the beginning of the vowel this is true in all cases except one (PD's [u]-words). The tendency still remains at the midpoint of the vowel, but the differences are considerably smaller and less consistent than at the beginning. Further, at the midpoint of the vowel the vertical position of the larynx in [v]-words is closer to the position in [b]-, [b^h]- and [f]-words than to the position in [m]-words, whereas the opposite was true at the beginning of the vowel.

The tendency for Fo to be higher after aspirated than after unaspirated stops is hardly - if at all - reflected in the larynx height data. Although the larynx tends to be higher after aspirated labial and alveolar stops, the difference is consistent for the alveolars only.

C. PLACE OF ARTICULATION

In figures 9 to 14 Fo means and larynx height means at the 6 reference points are plotted in such a manner as to display the effect of place of articulation. The results of the counts of differences and their directions are given in tables III and IV.

1. FUNDAMENTAL FREQUENCY

An effect of place of articulation of the test consonants is only found at the beginning of the following vowel. Here Fo is slightly higher, 3 Hz on the average, after [d^{sh}] than after [b̥^{h}] in 7 out of 9 cases. A similar degree of consistency

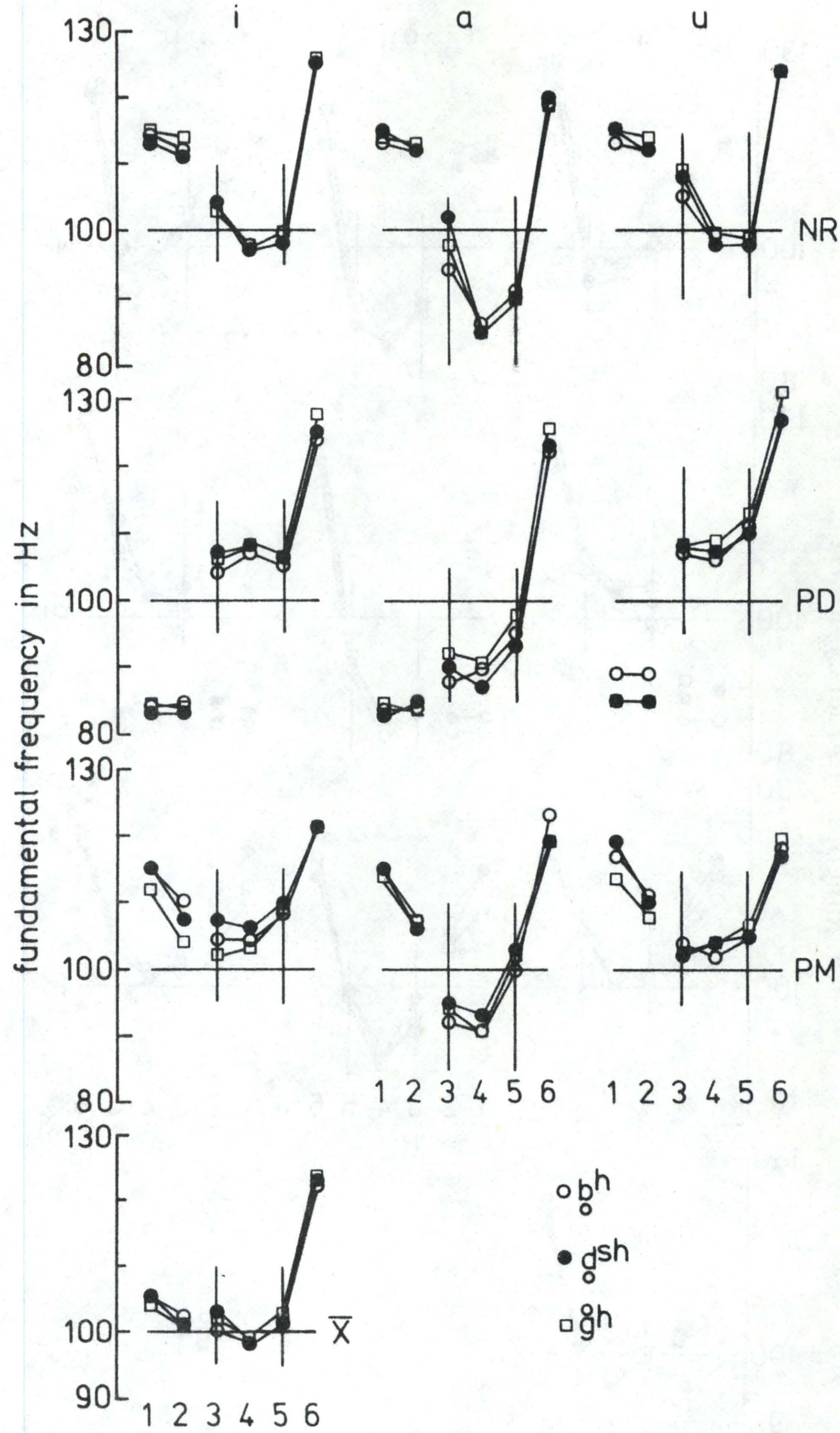


Figure 9

Mean fundamental frequency at the six reference points in [ˈCV:fi]*-words with initial aspirated stops varying in place of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

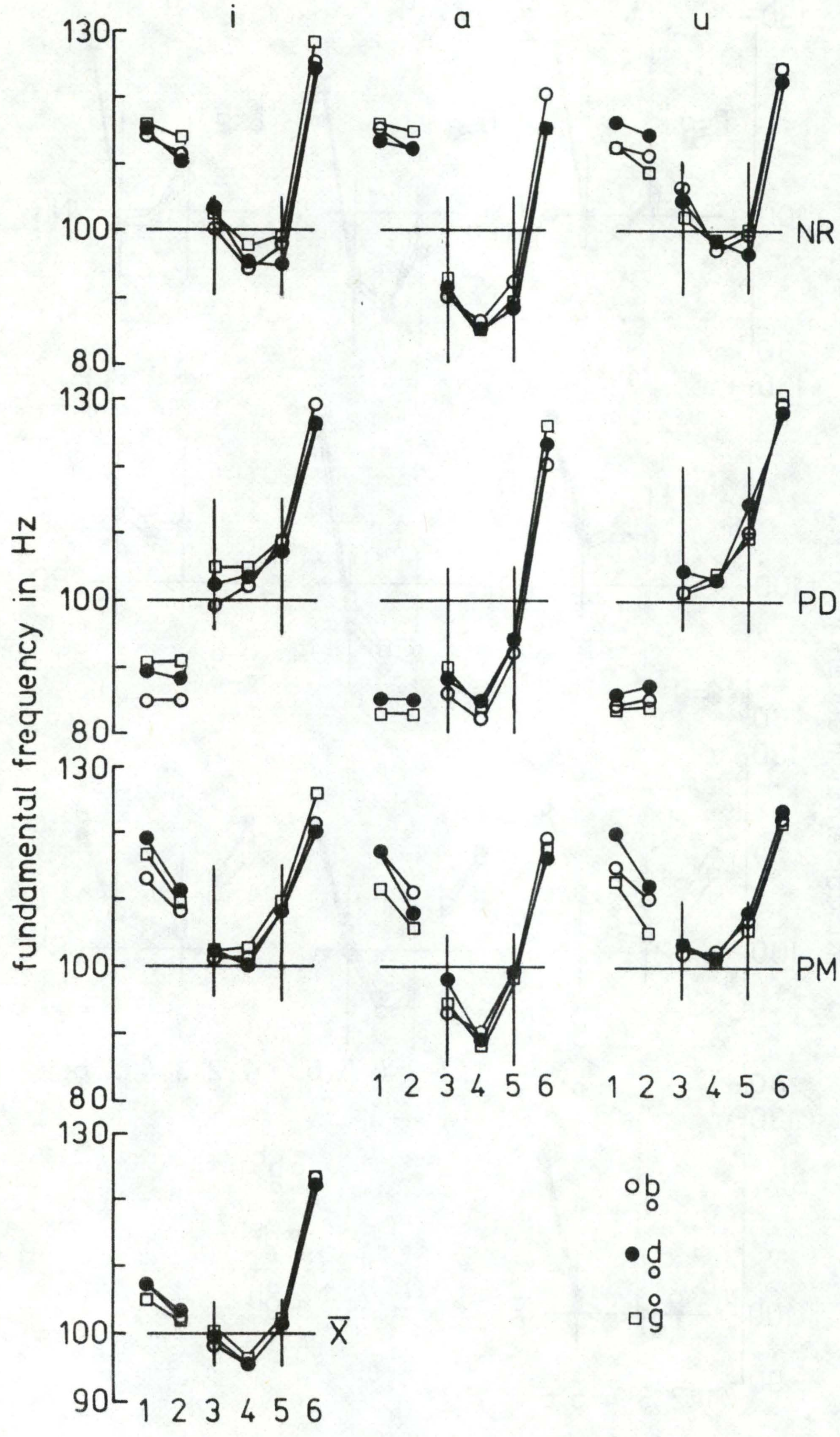


Figure 10

Mean fundamental frequency at the six reference points in ['CV:fi]*-words with initial unaspirated stops varying in place of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

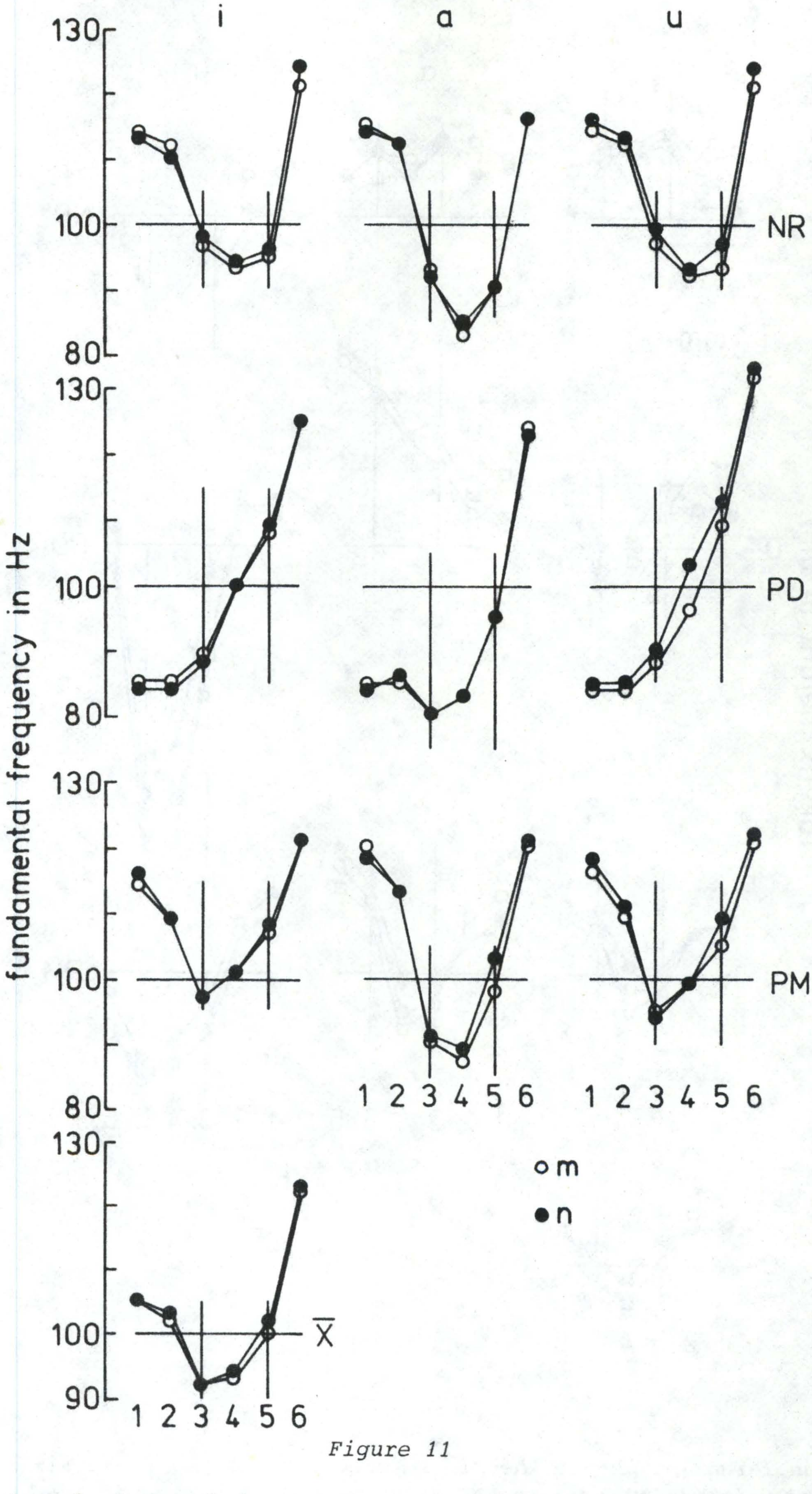


Figure 11

Mean fundamental frequency at the six reference points in ['CV:fi]*-words with initial nasal consonants varying in place of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

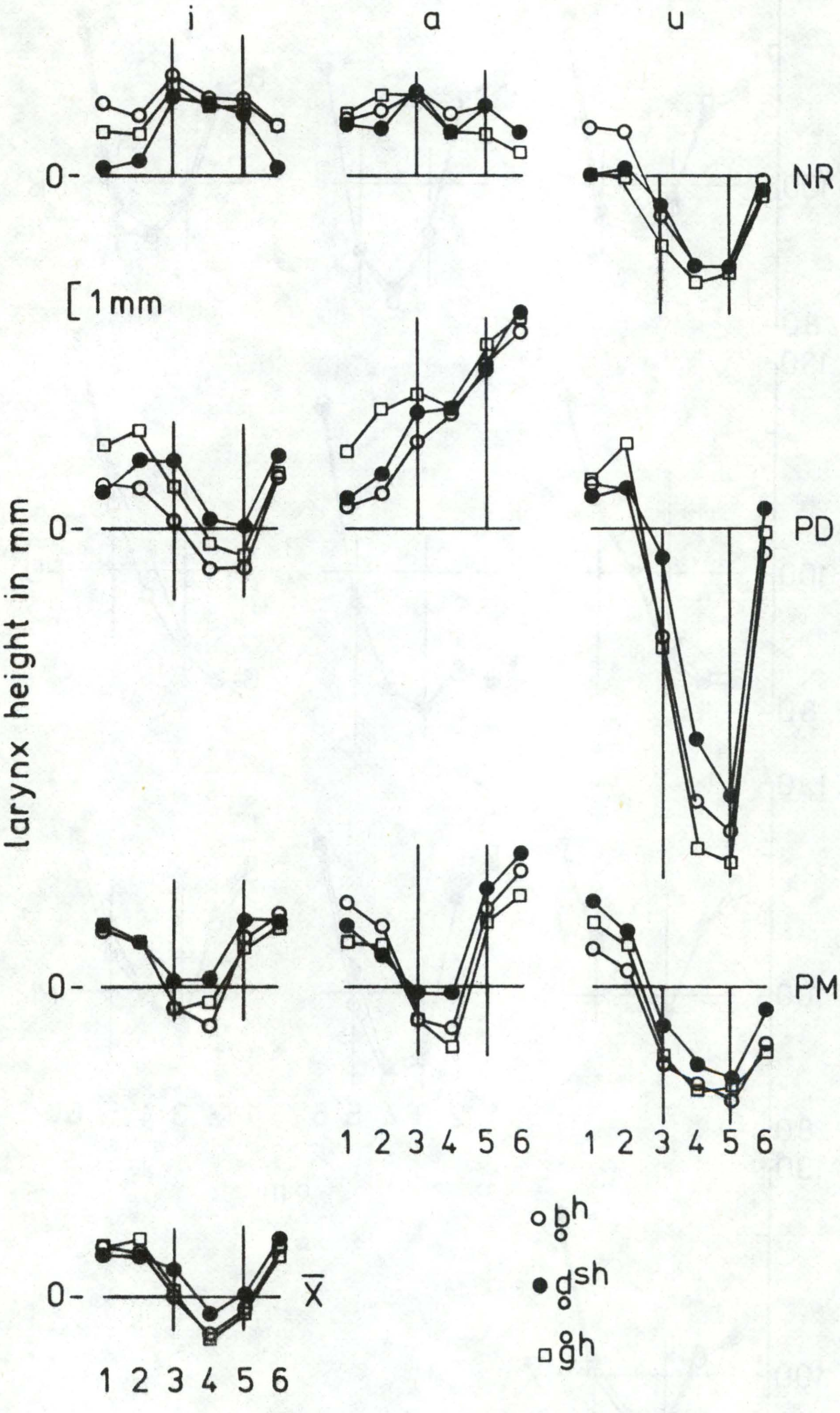


Figure 12

Mean larynx height at the six reference points in [ˈCV:fi]-*words with initial aspirated stops varying in place of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

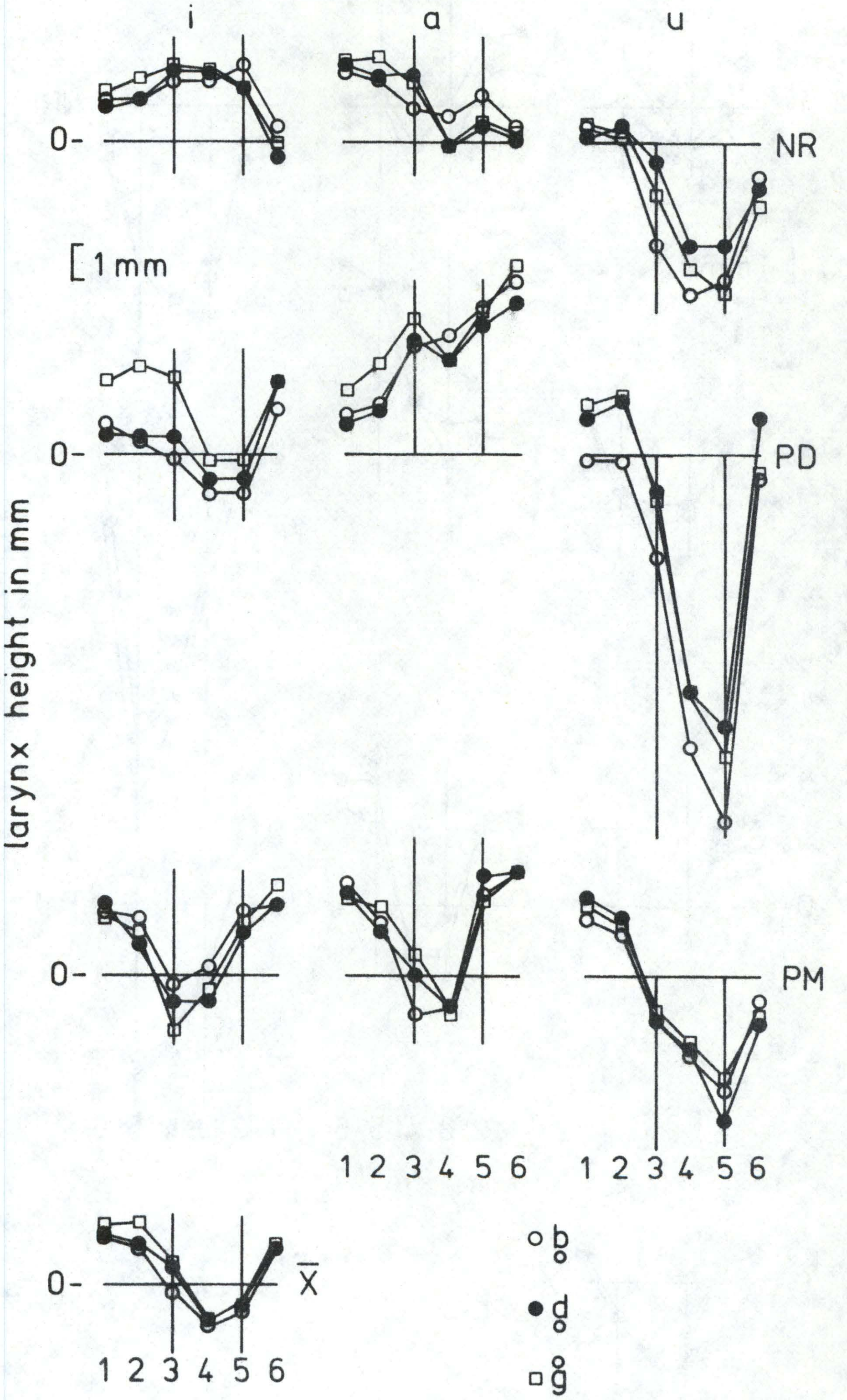


Figure 13

Mean larynx height at the six reference points in [ˈCV:fi]-*words with initial unaspirated stops varying in place of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

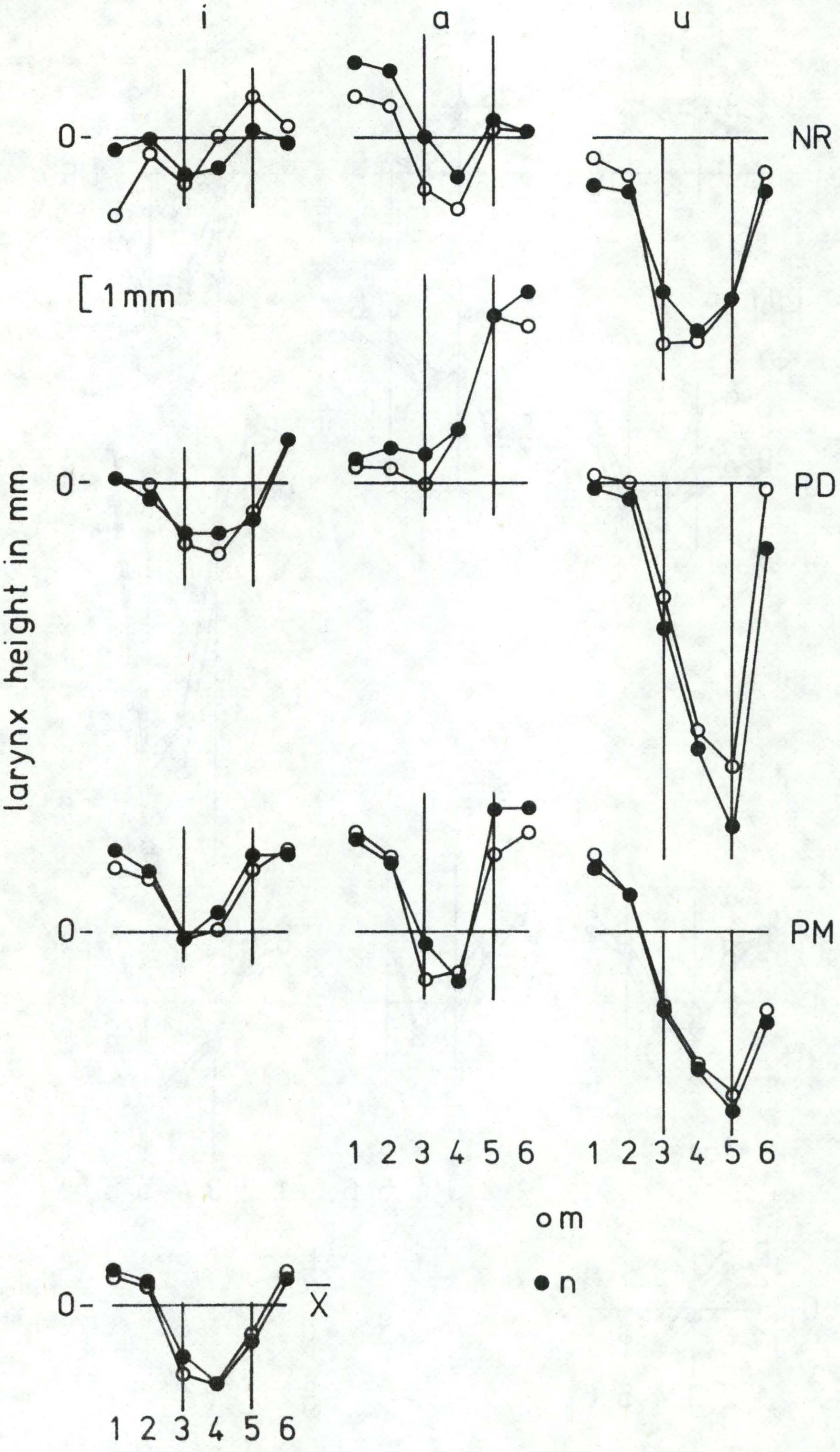


Figure 14

Mean larynx height at the six reference points in [ˈCV:fi]-*words with initial nasal consonants varying in place of articulation. The vertical lines in the graphs indicate the beginning and end of the stressed vowel. The lower left graph displays the average over all vowels and speakers.*

Table III

Results of the enumeration of differences between Fo means at the six reference points in pairwise comparisons of words differing in place of articulation of the initial consonant. See further the caption to table I.

b̥^{h}-d̥^{sh}

	>	=	<
1	4	2	3
2	6	2	1
3	1	1	7(1)
4	3	1	5
5	4	2	3
6	2	4	3

b̥^{h}-g̊

	>	=	<
1	4	1	4
2	4	2	3
3	3	0	6
4	2	2	5
5	2	0	7
6	1	3	5

d̥^{sh}-g̊

	>	=	<
1	3	1	5
2	3	1	5
3	5	1	3
4	2	3	4
5	2	2	5
6	1	3	5

b̥-d̥

	>	=	<
1	1	3	5
2	2	2	5
3	2	0	7
4	5	1	3
5	2	2	5
6	7	0	2

b̥-g̊

	>	=	<
1	4	1	4
2	5	0	4
3	2	1	6
4	3	1	5
5	3	2	4
6	3	2	4

d̥-g̊

	>	=	<
1	6	0	3
2	6	0	3
3	4	1	5
4	1	3	5
5	4	2	3
6	1	1	7

m-n

	>	=	<
1	4	1	4
2	2	3	4
3	3	2	4
4	1	3	5(1)
5	1	1	7
6	0	3	6

Table IV

Results of the enumeration of differences between larynx height means at the six reference points in pairwise comparisons of words differing in place of articulation of the initial consonant. See further the caption to table I.

b̥^{h}-d̥^{sh}

	>	=	<
1	6(1)	0	3
2	4	2	3
3	1	1	7
4	2	1	6
5	2	2	5
6	3	1	5

b̥^{h}-g̊^{h}

	>	=	<
1	3	0	6
2	3	1	5(1)
3	4	2	3
4	6	0	3
5	6	0	3
6	6	0	3

d̥^{sh}-g̊^{h}

	>	=	<
1	2	0	7
2	2	1	6
3	7	0	2
4	6	3	0
5	7	0	2
6	7	1	1

b̥-d̥

	>	=	<
1	4	0	5
2	3	1	5
3	1	1	7
4	3	1	5
5	5	0	4
6	6	1	2

b̥-g̊

	>	=	<
1	2	0	7
2	1	0	8(1)
3	1	0	8(2)
4	4	0	5
5	6	0	3
6	4	1	4

d̥-g̊

	>	=	<
1	2	1	6
2	2	0	7
3	3	0	6
4	2	0	7
5	2	1	6
6	2	0	7

m-n

	>	=	<
1	4	1	4
2	4	0	5
3	2	1	6
4	4	1	4
5	4	2	3
6	5	2	2

is found for the pair [b̥-d̥], but the average difference is only 1 Hz. Fo after [ɡ̊ʰ] and [ɡ̊] are also slightly higher than after [b̥ʰ] and [b̥] but not very consistently so. After nasals there is no effect of place of articulation on Fo.

2. LARYNX HEIGHT

The larynx tends to be higher for alveolars than for labials, especially at the beginning of the following vowel. This applies to aspirated and unaspirated stops and nasals alike, and the tendency is fairly consistent. At the midpoint of the first pre-tonic and at the beginning of the consonant the velars give rise to a higher larynx position than do labials and alveolars, most markedly so in PD's [i]- and [ɑ]-words. In the following vowel [ɡ̊ʰ] seems to associate with [b̥ʰ] (i.e. to be lower than [d̥ʰ]), whereas [ɡ̊] associates with [d̥] (i.e. is higher than [b̥]).

D. THE ASSOCIATION BETWEEN LARYNX HEIGHT AND FUNDAMENTAL FREQUENCY

In order to assess quantitatively the overall degree of association between larynx height and fundamental frequency, a correlation analysis (Pearson's r) was made. Since Fo and larynx height vary not only as a function of the type of consonant, which is the variation in focus here, but also as a function of time over the reference points and as a function of vowel quality, it was necessary to eliminate these effects by expressing any mean Fo and larynx height as the deviation from the grand mean over the 10 consonants at each reference point and in each vowel. In figure 15 the normalized means are plotted, Fo against larynx height. It is seen that there is some degree of association between the two variables. The correlation coefficient is 0.409. This is not very high, but because of the great number of data points (540 = 3 subjects x 3 vowels x 10 consonants x 6 reference points) it is highly significant, $p < 0.001$. By squaring the correlation coefficient a measure can be derived of that portion of the entire variation in one variable which can be accounted for by the variation in the other ($r^2 = s^2_{y'}/s^2_y$). Thus, in the present material, 17 per cent of the entire Fo variation can be explained by the variation in larynx height.

IV. DISCUSSION

The results of the present study are in good agreement with those of previous investigations in the field. This applies not only to the magnitudes and directions of the main effects of consonants on larynx height and fundamental frequency in the following vowel, but also - and to a higher degree, perhaps - to the tendency for the pattern of influence of consonants on larynx height to be subject to variation between subjects and vowel qualities, whereas Fo shows a considerably more consistent pattern of variation over the entire material.

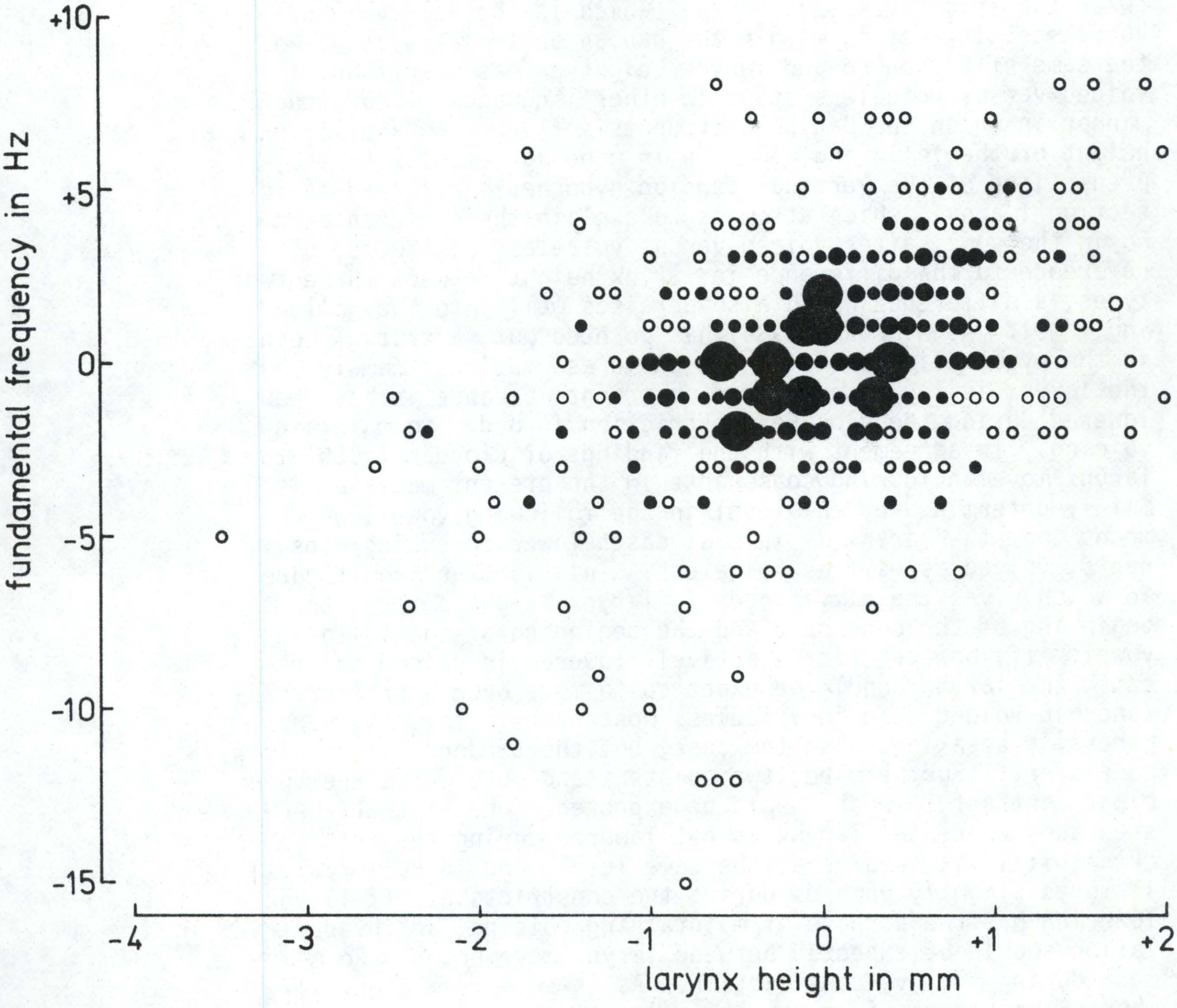


Figure 15

Fo means versus larynx height means. The number of data points at each pair of coordinates is indicated as follows: ○ = 1, ● = 2-3, ● = 4-6, and ● = 7 or more data points. The total number of data points is 540, viz. 3 speakers x 3 vowels x 10 consonants x 6 reference points.

One of the main findings of the present experiment is that both Fo and larynx height tend to be low following the voiced obstruent [v] and high following the voiceless obstruents. In this connection it is worth noting that the distinction between the fricatives [f] and [v] (which is the only voiced/voiceless distinction within the Danish obstruent system) has the same effect on Fo and larynx height as has been found for voiced versus voiceless stops in other languages. Thus, the manner in which the Danish obstruents influence Fo and larynx height of the following vowel is in good agreement with the predictions of the vertical tension hypothesis referred to in section I above, which attempts to explain the difference in Fo in the vowel after voiced versus voiceless obstruents by reference to the difference in larynx height between these two types, a difference which also persists well into the following vowel. As Riordan (1980) has pointed out, a crucial point in the hypothesis is one of its basic assumptions, namely that the larynx is low after voiced obstruents because it has been lowered during the closure/constriction in order to maintain voicing. In agreement with the findings of Riordan (1980) the larynx movement during consonants in the present material is mainly determined by the level in the following vowel, which means that the larynx is in most cases lowering during consonants, voiced as well as voiceless. This is seen from figure 16 which gives the differences in larynx height between the beginning of the consonant and the beginning of the following vowel. If, however, it is actively lowered in voiced obstruents, the larynx should be expected to move over a greater distance in voiced than in voiceless obstruents. This is, very generally speaking, also the case, but the tendency is far from consistent. Furthermore, two points stand out, which are in clear contrast to what should be expected. One is that there are cases where the larynx is not lowered during the [v]; in NR's [vi:fi] it remains at the same level, and in PD's [vɑ:fi] it moves slightly upwards during the constriction. If larynx lowering has the purpose of maintaining voicing, an inverse relation should be expected between larynx movement and Fo movement during the [v]-constriction. As is seen from figure 17, where Fo movement (from the beginning of the consonant to the Fo minimum at or near the end of the consonant) has been plotted against larynx movement, the extent and direction of larynx displacement has no influence on Fo, not even in the cases of no or upward displacement; if anything, the correlation is positive.

The other point which speaks against the assumption that the larynx is lowered in order to maintain voicing is the behaviour of the nasal consonants. Not only is the larynx lower in the vowel after nasals than after [v] but the larynx is also displaced farther downwards during nasals than during [v], although the purpose of the larynx lowering can hardly be to preserve a sufficient pressure drop across the glottis for voicing to be maintained during the nasal consonants. The nasal consonant

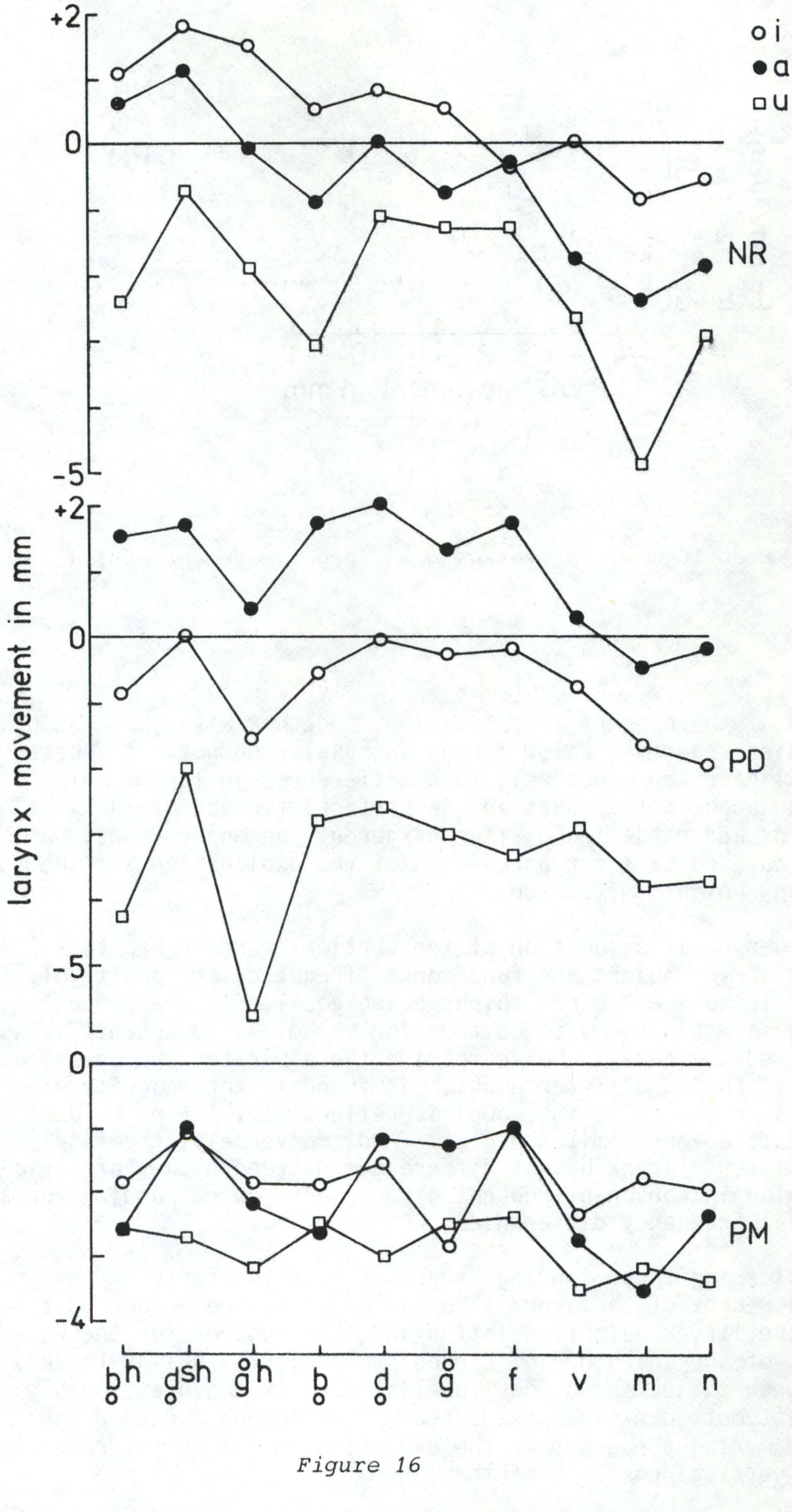


Figure 16

Mean larynx movement in mm during the test consonant. A positive value indicates an elevation of the larynx and a negative value a lowering of the larynx.

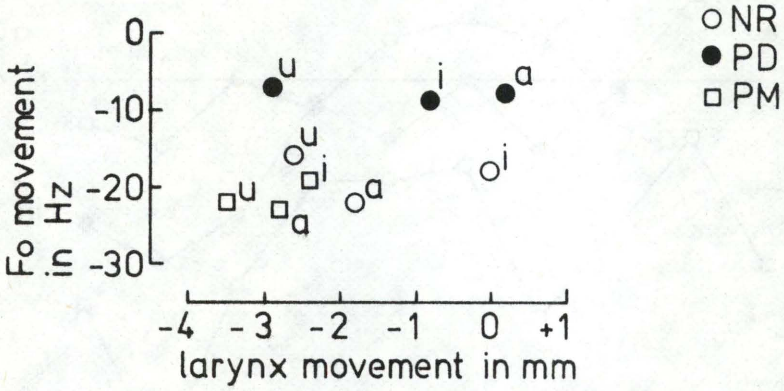


Figure 17

Mean Fo movement versus mean larynx movement in [v].

data of the present investigation, together with those of other studies reporting a low larynx in nasals, do not, of course, invalidate the hypothesis that differences in larynx height bring about differences in the vertical tension of the vocal chords and hence in Fo after different consonant types, but the data do call for a revision of the explanation for the larynx height differences.

Another basic assumption of the vertical tension hypothesis is that larynx height and fundamental frequency are positively correlated - all other things being equal. In the present material this seems to be true for voiceless consonants versus [v] and the nasals, but comparing the aspirated and unaspirated stops, the Fo difference which is found in the majority of cases in the following vowel is reflected by larynx height only to a very limited degree. And, conversely, the fairly consistent larynx height differences between places of articulation in consonants do not give rise to corresponding fundamental frequency differences.

Considering these findings together with the fact that only 17 per cent of the overall Fo variation can be accounted for by the larynx height variation and the tendency for the pattern of the influence of consonants on larynx height to vary between subjects and vowel qualities, it seems questionable to attempt to explain the effect of consonant types on the fundamental frequency in the present material by reference to the vertical position of the larynx.

V. NOTE

1. The material was designed so as to investigate the effect of vowel quality on Fo and larynx height, also. The analysis of the results from that point of view will be dealt with in a later paper.

REFERENCES

Bell-Berti, F. 1975: "Control of pharyngeal cavity size for English voiced and voiceless stops", *J. Acoust. Soc. Am. 57*, p. 456-451

Bothorel, A. 1979: "Déplacement de l'os hyoïde", *Séminaire Larynx & Parole, Institut de Phonétique, Grenoble - 8-9 fév. 1979*, p. 183-196

Di Cristo, A. et Chafkouloff, M. 1977: "Les faits microprosodiques du français: voyelles, consonnes, articulation", *8. Journees d'études sur la Parole, Aix-en-Provence, 25-27 mai, 1*, p. 147-158

Ewan, W.G. 1979: "Laryngeal behaviour in speech", *Report of the Phonology Laboratory 3, University of California, Berkeley*, p. 1-93

Ewan, W.G. and Krones, R. 1974: "Measuring larynx height using the thyroumbrometer", *J. Phonetics 2*, p. 327-335

Ferguson, G.A. 1976: *Statistical Analysis in Psychology and Education*, (McGraw-Hill, Tokyo)

Fischer-Jørgensen, E. 1968: "Voicing, tenseness, and aspiration in stop consonants, with special reference to French and Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 3*, p. 63-114

Fischer-Jørgensen, E. 1972: "PTK et BDG français en position intervocalique accentuée", *Papers in Linguistics and Phonetics in the Memory of Pierre Delattre*, (The Hague) p. 144-200

Gandour, J. and Maddieson, I. 1976: "Measuring larynx movement in Standard Thai using the cricothyrometer", *Working Papers in Phonetics, UCLA, 33*, p. 160-190

Hombert, J.-M., Ohala, J.J., and Ewan, W.G. 1976: "Tonogenesis: Theories and queries", *Report of the Phonology Laboratory 1, University of California, Berkeley*, p. 48-92

House, A.S. and Fairbanks, G. 1953: "The influence of consonantal environment upon the secondary acoustical characteristics of vowels", *J. Acoust. Soc. Am. 25*, p. 105-113

Jeel, V. 1975: "An investigation of the fundamental frequency of vowels after various consonants, in particular stop consonants", *Ann. Rep. Inst. Phon. Univ. Cph. 9*, p. 191-211

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

Kent, R.D. and Moll, K.L. 1969: "Vocal tract characteristics of the stop cognates", *J. Acoust. Soc. Am. 46*, p. 1549-1555

Lea, W.A. 1973: "Segmental and suprasegmental influences on fundamental frequency contours", in *Consonant Types and Tone, Southern California Occasional Papers in Linguistics 1*, (Hyman, L.M. ed.), p. 15-70

Lehiste, I. and Petersen, G.E. 1961: "Some basic considerations in the analysis of intonation", *J. Acoust. Soc. Am. 33*, p. 419-425

Lindqvist, J., Sawashima, M., and Hirose, H. 1973: "An investigation of the vertical movement of the larynx in a Swedish speaker", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics, Tokyo, 7*, p. 27-34

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

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

Ohala, J.J. 1973: "Explanations for the intrinsic pitch of vowels", *Monthly Internal Memorandum, Phonology Laboratory, University of California, Berkeley*, p. 9-26

Perkell, J.S. 1969: *Physiology of Speech Production* (MIT Press, Cambridge, Mass.)

Petersen, N. Reinholt 1978: "The influence of aspiration on the fundamental frequency of the following vowel in Danish: Some preliminary observations", *Ann. Rep. Inst. Phon. Univ. Cph. 12*, p. 91-112

Riordan, C.J. 1980: "Larynx height during English stop consonants", *J. Phonetics 8*, p. 353-360

Shipp, T. 1975: "Vertical laryngeal position during continuous and discrete vocal frequency change", *J. Speech and Hearing Research 18*, p. 707-718

Westbury, J.R. 1983: "Enlargement of the supraglottal cavity and its relation to stop consonant voicing", *J. Acoust. Soc. Am. 73*, p. 1322-1336

ARIPUC 17, 1983, p. 87- 109

AN ORTHOGRAPHY NORMALIZING PROGRAM FOR DANISH*

PETER MOLBÆK HANSEN

The paper is a description of a program which transforms Danish orthographic texts into so-called normalized notation, i.e. a format which can serve as input to a text-to-speech-by-rule algorithm which is under development at the institute. The main features of the program are described, and the peculiarities of Danish orthography which motivate these features are presented. The capabilities and the limitations of the program are illustrated, and the possibilities of future improvements and extensions of the program are outlined.

I. INTRODUCTION

In Molbæk Hansen (1982) I reported on the construction of a grapheme-to-phone algorithm for Danish. That paper reflects a first approximation to a solution of the main problems connected with the design of a computer program capable of converting unrestricted Danish text to a normalized notation which is consistent with the pronunciation of the text and hence acceptable as input to a letter-to-sound-by-rule program.

During the past year the main features of the text normalizing program (henceforth TNP) were developed, and the basic code has been written. The present paper presents the basic design of the TNP and the factors (including some important characteristics of Danish orthography) which have been decisive in selecting that particular design.[1]

In section II some general features of text-to-speech-systems are touched upon, and a rudimentary typology of such systems is suggested. This furnishes the background for a characterization of the text-to-speech algorithm that is being developed for Danish.

---

*) *This work is carried out for the Telecommunications Research Laboratory within the synthesis-by-rule project.*

In section III the features of Danish orthography which motivate the choice of TNP-type are discussed, and the main features of the TNP and of its output, i.e. the normalized notation, are presented.

Section IV contains a description of the organization of the data part and of the processor part of the TNP. Some examples are given to illustrate what the program actually does to unrestricted Danish text.

In section V, finally, the capabilities as well as the limitations of the TNP are commented upon, and some perspectives of overlaying the TNP with facilities for operating with syntactic information are presented.

II. THE CHOICE OF A TEXT-TO-SPEECH SYSTEM

A. TYPES OF SYSTEMS

As is well known, alphabetic writing systems exhibit various degrees of consistency in the sense of Molbæk Hansen (1982, p. 128), ranging from near one-to-one correspondences between letters and phonemes, as in Finnish, over various approximations to more or less complex, yet contextually definable many-to-one correspondences between letters (letter combinations) and phonemes (phoneme combinations) as in French, to the abundance of many-to-one and, more seriously, one-to-many correspondences dominating such orthographies as Danish and English.

In some orthographies there are many irregularities, but such that each of them affects only a few words. In other orthographies there are only one or a few irregularities or other shortcomings, but such that a major part of the lexicon is affected by them. This is by and large true of the Russian and the Italian orthographic systems in which the most serious problem is that distinctive word stress is not indicated, cf. Sherwood (1978) and Lesmo et al. (1978). Another example of how a major part of the lexicon may be crucially dependent - as far as the predictability of the pronunciation is concerned - on an ambiguous spelling is the word final sequence *-ent* in French, which is pronounced [ã] in nouns and adjectives but is mute in 3rd person plurals of verbs.

Thus, although any text-to-speech algorithm has to perform at least some normalizing, it is not surprising that the various existing text-to-speech algorithms differ widely both in general build-up and in complexity. Among the existing text-to-speech systems two main approaches seem to be dominant:

Some systems are exclusively or predominantly rule oriented, i.e., except for a small preprocessor identifying a handful of words with extremely exceptional spellings or with a special phonetic behaviour, they consist of one monolithic rule component which must take care of both letter-to-sound rules

- including "rules" changing exceptional spellings - and phonological and phonetic rules - including rules for manipulating acoustic parameters. In such systems exceptional spellings are changed by "rules" of the same formal status as the rules referring exclusively to context of the expression plane (whether graphemic, phonemic or phonetic). This can be done because, except for cases of homography at word or morpheme level, the spelling of a word or of a morpheme is a reliable cue to its identity. This is the approach of Maggs et Trescases (1980) who take care of the above mentioned problem with French *-ent* by including in their rule system a large number of rules of the type *souv[ent]* $\rightarrow$ *ã* (in their notational format meaning "*'ent'* becomes [ã] between *'souv'* and space"). The Swedish system developed by Carlson and Granström also belongs here, being based on the explicit desire *"to minimize the lexicon and make the inventory of rules more exhaustive"* (Carlson and Granström (1975, p. 22)).

Other systems consist of several components of comparable complexity, typically including at least (1) a component taking care of deviating spellings by (hopefully) finding them in a more or less comprehensive morph lexicon which contains entries for a large number of morphs, each entry being linked with information on the pronunciation of the morph in question and on its syntactic and other grammatical behaviour, (2) a letter-to-sound rule component, applying to the spellings which are not found in the lexicon and hence considered normal, (3) one or more components taking care of prosodic rules, and (4) a component computing the acoustic parameters. The MITALK text-to-speech system for American English is of this type, cf. e.g. Allen (1976, 1981).

It is probably no coincidence that the lexicon minimizing systems were developed for French and Swedish: the orthographic irregularities of these languages are fewer and simpler than the irregularities of the English orthographic system. However, the choice of system also depends on the applications aimed at, on demands for speech quality, and on demands for real time performance. As a rule, the quality is high in the lexicon type system. The inclusion of a comprehensive lexicon, on the other hand, is to some extent incompatible with demands for (limited memory and) real time performance, cf. Sherwood (1978, p. 670).

B. THE DANISH SYSTEM

The text-to-speech system being developed for Danish is essentially a combination of the two types: we try to combine the advantage of having access to a morph lexicon including grammatical information with the advantage of being able to change, improve, and experiment with the rule system in the easy and flexible way this can be done if the rules are written in a special, linguistically oriented high level programming language based on the ideas of Carlson and Granström (1975).
For more general information on the system, in particular the

rule component and the construction of a rule language compiler, see Holtse (1982).

This particular combination also allows us to divide the text-to-speech algorithm into two natural parts: the TNP and the rule component. Obviously, it is conceptually more satisfactory not to have to include "rules" like the French one mentioned above.

Since we aim at high quality output, and since Danish orthography is of the highly irregular type, it became clear at a relatively early stage that the TNP must include a morph lexicon that is basically of the same type as the one described by Allen (1976 and 1981) and developed and improved for English during the last fifteen years. That it actually differs in many details is due primarily to the special features of Danish orthography, as we shall see below.

So far, considerations of memory requirements and program efficiency play only a minor role in our scheme.

III. GENERAL CONDITIONS AND INDICATIONS FOR A TNP

The ideal TNP should, of course, be capable of converting the orthographic representation of any Danish sentence to a normalized notation from which the pronunciation of the sentence can be derived in an unambiguous way.

This ideal goal implies that such important phenomena as sentence stress and intonation patterns should be derivable by rule from the normalized notation. In view of the fact that such phenomena still await thorough research it seems a good policy to start with a TNP which (1) is capable of normalizing single words correctly in the overwhelming majority of cases where this can be done in a principled way, and (2) is designed in such a way that future extensions taking care of sentence phonetic phenomena can be easily and harmonically added to it. This has been the leading principle in the design of the TNP. The more specific requirements were mainly dictated by the specific properties of Danish orthography.

In this section those features of Danish orthography which have been decisive for the design of the TNP will be outlined. In Molbæk Hansen (1982) I mentioned some of these features in a rather unsystematic fashion. At that time I did not fully realize the implications of some of the problems, and I overestimated the importance of others. In the following subsections I shall describe and illustrate those features of Danish orthography which I now consider the most significant ones - both quantitatively and qualitatively - for any attempt to write a TNP for Danish.

A. THE REGULAR SPELLING OF FOREIGN WORDS

For some reason one is apt to believe that irregular spellings occur more often in foreign words than in "genuine" Danish words, i.e. roughly words inherited from Old Danish and medieval loan words from Low German. This is true only if foreign words are defined narrowly as recent loan words, say, words borrowed in the 20th century. However, apart from these there is a large number of older loan words of Latin or Greek origin which are now more or less assimilated to the Danish sound system, and for these the spelling is impressively regular.

One only has to read a few pages in any dictionary (ordinary or retrograde, depending on what one is looking for) to become convinced that if a word is found to contain a morpheme belonging to the group of such classical loans, then, in the overwhelming majority of cases, the pronunciation of the whole word is straightforwardly derivable by rule. This is particularly obvious in the case of words containing obligatorily stressed suffixes like e.g. the verbalizing suffix *-er-*, borrowed (via French *-er* or German *-ier-en*) from the Latin 1st conjugation infinitive ending *-are*, yielding Danish infinitives in *-ere* with stress and stød on the penultimate syllable. Since information on stress and stød is crucial for eventually determining the pronunciation, the identification of such stressed suffixes is, of course, essential, but such an identification would be of less interest if the stems preceding such suffixes were typically characterized by irregular letter-sound correspondences. Fortunately, that is not the case. In Holmboe's dictionary (Holmboe (1978)) there are roughly 1400 verbs of this type, and the overwhelming majority of these exhibit a regular spelling, irrespective of whether or not the *-er-* is considered part of the root. Similar conditions prevail in other suffixes belonging to this category, e.g. *-itet*, *-ens*, *-isme*, *-at* as in *identitet* 'identity', *tendens* 'tendency', *kommunisme* 'communism', and *professorat* 'professorship', respectively.

Foreign words - in particular the numerous ones of Latin origin - which do not contain a stressed suffix also typically take stress according to a general rule, viz. the last vowel followed by a consonant is stressed, cf. *kommando* 'command', *kaskade* 'cascade', *melon* 'melon', *analyse* 'analysis', although this rule is far from being without exceptions, cf. Rischel (1970).

Obviously, these regularities should be exploited somehow in the TNP. We shall see below how this can be done.

B. THE IRREGULAR SPELLING OF GENUINE DANISH WORDS

The major part of the irregular spellings occur in the bulk of genuine Danish words which by and large can be divided into monosyllables and initially stressed disyllables with schwa, orthographically *e*, as the second vowel.

Three main features of the spelling of such words are important for the design of a TNP:

(1) certain vowel letters, notably *i*, *y*, *u*, and *o* are ambiguous, cf. *sidst* /sisd/ 'last' vs. *vidst* /vesd/ 'known'; *tykke* /tygə/ 'thick (plur.)' vs. *stykke* /sdøgə/ 'piece'; *pund* /punʔ/ 'pound' vs. *bund* /bonʔ/ 'bottom'; *ost* /osd/ 'cheese' vs. *post* /pɔsd/ 'mail'; *slog* /slo:ʔg/ 'hit' vs. *klog* /klɔ:ʔg/ 'clever';

(2) vowel length is unpredictable from the spelling in many monosyllables, cf. *sal* /sa:ʔl/ 'hall' vs. *bal* /balʔ/ 'ball'; *ben* /be:ʔn/ 'leg' vs. *pen* /pɛnʔ/ 'pen';

(3) the occurrence of stød is unpredictable from spelling in many words, cf. *solen* /so:ʔlən/ 'the sun' vs. *skolen* /sgo:lən/ 'the school'; *skov* /sgɔvʔ/ 'wood, forest' vs. *tov* /tɔv/ 'rope, wire'.

The following facts should also be taken into consideration:

(4) the bulk of homographs occurs in genuine Danish words, primarily due to one or more of the factors 1-3, cf. *lod* /lo:ʔd/ 'let (vb. pret.)' vs. *lod* /lɔd/ 'destiny'; *læs!* /lɛ:ʔs/ 'read!' vs. *læs* /lɛs/ 'load'; *føl* /fø:ʔl/ 'feel!' vs. *føl* /føl/ 'colt';

(5) the majority of common and frequent compounds are composed of elements belonging to the genuine vocabulary, cf. such common compounds as *ægteskab* 'marriage', *sagfører* 'lawyer', *lommetyv* 'pickpocket', *brombær* 'blackberry' (in which *brom-* is a "quasi-genuine" element which does not occur in isolation, cf. the English word *cranberry*), and/or their initial or final elements belong to this group: *overassistent* 'senior clerk' (in which *over* is the genuine element, and *assistent* is a regularly spelled loan word of the classical type mentioned above), and *regentskifte* 'exchange of sovereign' (in which *skifte* is the genuine element);

(6) double consonant letter before a schwa alternating with word final single consonant letter is typical of such words, viz. in case the vowel is short, cf. *kys* 'kiss' - *kysset* 'the kiss' (vs. *lys* 'light' - *lyset* 'the light' with long vowel), *hjem* 'home' - *hjemmet* 'the home', etc.

C. THE LOCAL INFORMATION IN DANISH WORDS

The rightmost letters of a Danish word contain most of the information needed in order to eventually determine its pronunciation. The leftmost letters contain less information than the rightmost letters, but more information than the medial letters.

The importance of the rightmost letters of a word form is primarily due to two factors: (1) in the majority of cases such important phenomena as word stress, vowel length in stressed syllables and stød in stressed syllables can only be determined from the particular combinations of roots with suffixes and/or endings, (2) a handful of very frequent final letter combinations symbolize either an ending or the final part of a stem, and the interpretation of such final sequences as one or the other is crucial for the pronunciation.

This may be illustrated by the word final letter combination *en*. In *arsen* /ar'se:ʔn/ 'arsene', *pen* /pɛnʔ/ 'pen', and *igen* /i'gɛn/ 'again' the final sequence *en* belongs to the root. In *arsen* the vowel is long /e:/, and the word has stød according to the general rule that word final stressed syllables have stød if their vowel is long. In *pen* and *igen*,*en* represents /ɛn/, but *pen* has stød, whereas *igen* is stødless (stød is not predictable from the surface segmental structure of a stressed word final syllable with stød basis and short vowel). In *solen* /so:ʔlən/ 'the sun' *en* represents the allomorph /ən/ of the definite singular common gender ending, which has the property of retaining the regular stød in roots like *sol* (with long vowel, cf. *arsen* above) and of adding stød to roots like *søn*, which are stødless in other forms. In *talen* /ta:lən/ 'the speech' *e* represents the final schwa of the regular stødless disyllabic word *tale*, whereas the *n* represents the allomorph /n/ of the same ending as in *solen* and *sønnen*.

Similar one-to-many correspondences between spelling and morphological/phonological structure are characteristic of practically all word final letter combinations which can appear as (parts of) endings and also of a few suffixes consisting of one or more schwa syllables, cf. such frequent final letter sequences as *ens*, *et*, *ets*, *er*, *ers*, *erne*, *ernes*, *ene*, *enes*, *e*, *s*, *t*, *ede*, *edes*, *ende*, *ere*, *eres*, *else*, *elses*. A frequent word final letter combination like *el* behaves exactly like e.g. *en*, although it does not ever represent an ending: it represents a schwa syllable in words like *middel* 'means', *adel* 'nobility' (both with stød, cp. *sønnen* and *solen* above), and *sadel* 'saddle' (stødless, cp. *talen* above); it represents root final /e:ʔl/ in *kamel* 'camel' (cp. *arsen* above); it represents root final /ɛl/ in *farvel* 'good bye', (cp. *igen* above), and it represents root final /ɛlʔ/ in *model* 'model' (cp. *pen* above). Final *e* often represents root final /ə/ of many disyllables belonging to the genuine vocabulary - among Danish phonologists often referred to as β-words - such as *tale* 'speech', cf. above, *kone* 'wife', and many others; in the following sections such letters or letter sequences will be referred to as quasi-endings when they are root final, and the remaining parts of such roots, e.g. the *kon* of *kone*, will be referred to as quasi-roots.

Other final letter sequences represent structures that obligatorily entail stress on the preceding syllable. This is true e.g. of "suffixes" like *ium* as in *radium* 'radium', etc. Others represent, more or less uniquely, suffixes which obligatorily take stress themselves if not followed by other suffixes, cf. *ik* as in *musik* 'music', etc.

That the leftmost letters of a word form are also important for the phonological behaviour of the whole word is mainly due to the fact that stød on the root is obligatory (if it is segmentally permitted) after certain prefixes, cf. cases like *skue* 'see, view' vs. *beskue* 'take a view of', where the occurrence of stød on the latter form is entirely predictable from the presence of the prefix *be-*.

In this case, as well as in all the other cases involving stress and stød, the heart of the matter is - as is well known among Danish phonologists - that the occurrence and placement of stress and stød in Danish word forms is by and large predictable from the combined information provided by (a) the identity of certain bound morphemes, (b) the segmental structure of root morphemes, (c) the morphological type of the word form, provided that the segmental structure of roots is considered at a rather abstract level, see e.g. Rischel (1970) and Basbøll (1972).

What all this amounts to is that, once certain initial and final conditions are determined, the phonological structure of the whole word form is also basically determined, and that the medial part of a long word form typically exhibits regularity or at most minor segmental irregularities in its orthographic shape.

Apart from cases like *ium* where the spelling uniquely defines the suffix, it is a problem with most of the above mentioned affixes that their spelling will not *per se* identify them. In many cases this is due to the fact that the spelling might as well be part of a larger structure, as was illustrated by the examples with endings. Since the morphological status of such final sequences is so crucial for the determination of stress and stød, it becomes an important task for a Danish TNP to identify these sequences correctly.

D. INDICATIONS OF DANISH ORTHOGRAPHY FOR A TNP

The above mentioned facts of Danish orthography point to a TNP with the following characteristics:

(1) It should have access to a lexicon of morphs and other structures containing crucial information of the kind needed to derive the phonetic shape of a word form by rule.

(2) The morphs to be placed in the lexicon should be (a) all or almost all genuine roots - since these contain the major irregularities; (b) all or almost all affixes - since these contain crucial information and help to delimit the roots. The other entries to be placed in the lexicon should be quasi-morphs of the sort which phonologically act in a morph-like way, cf. e.g. the quasi-ending *-el* which was mentioned above, and a quasi-root like *ad* in *adel* 'nobility'. (This word may thus be said to consist of a quasi-root and a quasi-ending.)

(3) It should scan word forms maximalistically, i.e. in a longest-match-first fashion, from the right end of the word as well as from the left end of the word, to check the input word form against the lexicon. In most cases in running text all morphs should be identified correctly by this approach, the normalized shape of the word forms should be retrievable, and phonologically relevant grammatical information pertaining to the morphs should be accessible. In cases where the word or part of it is not identified - typically when there is an unidentifiable medial residue such as the many foreign roots which are not in the lexicon - that part (or that whole word) should be considered regular.

E. NORMALIZED TEXT

When it is decided that the items to be placed in the lexicon should be all affixes and all roots belonging to the core of genuine items, cf. section II D, much of the design of the normalized notation which is to be the output of the TNP is also determined: If, e.g., a word like *officer* /ɔfi'se:ˀr/ 'officer' is taken to be regular and hence not to be placed in the lexicon, its spelling must be considered normal, and this means that the endings, suffixes, and genuine roots in which final *er* represents unstressed /ər/ must be normalized in such a way that they will not come out ending in *er* in the same way as *officer*.

The normalized notation is comparable to the so-called World English Spelling (WES), see Dewey (1971), although its motivation is of quite another kind. At the moment it is not defined rigidly (partly because its final design may have to depend on technical considerations and possibilities), but an operational version of it which I use in the daily testing work, cf. section IV, has the following main properties:

Word final occurrences of the letters *t* and *s* are separated from roots by a + if they represent endings, cf. e.g. *spis+t* 'eaten' vs. *gnist* 'spark', and *fod+s* 'of a foot' vs. *klods* 'block'.

Word final occurrences of the letter *e* are separated from preceding roots and quasi-roots by a + if they appear as endings or quasi-endings (/ə/) in cases where the preceding syllable has (stress and) stød, cf. e.g. *extern+e* 'external (definite or plural)', *Johann+e* (a personal name), *radikal+e* 'radical (definite or plural)', etc.

In word final mono- or disyllabic sequences with *e* (representing schwa) as vowel letter(s), the first or only *e* is considered initial in an ending or quasi-ending, if the word form in question is stødless, otherwise the (first or only) *e* is dropped. In both cases the letter preceding the *e* is taken to be root final, and the root is separated from the ending or quasi-ending by a hyphen, cf. e.g. the following normalized notations: *mand* 'man', *mand-n* (spelled *manden*) 'the man', *mand-ns* (spelled *mandens*) 'the man's', *kon-e* 'wife', *kon-er* 'wives', *sad-el* 'saddle', *sadl-er* 'saddles', *hus* 'house', *hus-t* (spelled *huset*) 'the house', *hus-e* 'houses', *hus-ene* 'the houses', *mål-ne* (spelled *målene*) 'the measures', *rat* 'steering-wheel', *ratt-t* (spelled *rattet*) 'the steering-wheel', etc. In a similar way, *-e* or, in some cases, just - is inserted before certain endings and quasi-endings in order for the word forms in question to conform to the same structural features as the ones just mentioned, cf. *spis-ete* (spelled *spiste*) 'ate' vs. *spis-te* (spelled *spiste*) 'eaten', *kis-ete* (spelled *kiste*) 'coffin', and *pås-eke* (spelled *påske*) 'Easter'.

The vowel letters *i*, *y*, *u*, are replaced by the corresponding upper case letters in cases where they symbolize mid or high-mid as opposed to high vowels, e.g. in *fUnd* 'finding, discovery, find' vs. *pund* 'pound', *lIste* 'list' vs. *pisk* 'whip', *lYgte* 'lantern' vs. *skygge* 'shadow'.

The letter *o* is replaced by (upper case) *O* in cases where it represents /o/ as opposed to the more frequent reflex /ɔ/, and in cases where it represents /ɔ:/ as opposed to the more frequent reflex /o:/, cf. *trOmme* 'drum' vs. *fromm-e* 'pious (pl.)' and *bOg* 'book' vs. *slog* 'hit'.

Boundaries between parts of a compound are marked by =, e.g. *lUft=skib*, 'airship'.

Boundaries between certain prefixes and roots are marked by +, e.g. *be+tragte-lig* 'considerable'.

Final stressed root syllables are changed in many cases where the spelling of such syllables exhibits less than two postvocalic consonants: if the roots in question have stød-basis and stød in uninflected forms, their final letter is rewritten in geminate form, cf. e.g. *gåå* 'go, walk', *ball* 'ball'. The final consonant letter is also rewritten in geminate form if it represents an obstruent and the vowel is short, cf. *kYss* 'kiss'. If the syllable is closed and the vowel is long, their spelling is considered normal, e.g. *hospital* 'hospital'; otherwise an *h* is added to the final letter, e.g. *metalh* 'metal', *nuh* 'now'.

As can be seen from these examples, the normalized notation is neither phonemic nor morphophonemic in any reasonable sense: many-to-one correspondences between letter and phoneme are allowed; the endings *s* and *t* are not separated from preceding unstressed endings, cf. e.g. *mand-ns* 'the man's'. But, ideally at least, an autonomous phonemic transcription can be derived from it.

IV. THE TNP

The TNP consists of two parts: a data part, i.e. the lexicon, and a program part, i.e. the text processor. In the following the organization of these two parts will be outlined. It must be mentioned that what has actually been implemented is a TNP skeleton consisting of a small lexicon of test items representing the main types of entries and an *ad hoc* version of the processor. There is still much testing and experimenting to be done before a reasonably well defined system which will normalize a large part of the Danish vocabulary correctly, can be implemented. The following description is based on the well established main features of the TNP and on experiences with current *ad hoc* implementations with a limited lexicon.

No attempt will be made to describe concrete, more or less trivial technical details, since such details are dependent on current configurations of hardware and software. Suffice it to mention that the TNP is being developed and implemented under the UNIX operating system running on the PDP-11/60 computer assigned to the text-to-speech project, and that the basic code is written in the C programming language. These conditions guarantee a maximum of portability and flexibility of the basic software, so that it can be moved and adapted to various concrete configurations and applications without much trouble.

A. THE LEXICON

1. THE ENTRIES OF THE LEXICON

The lexicon is to consist of a large number of entries. Each entry is a data structure with two elements:

(1) a character string which has two functions: it permits an identification with a letter combination which actually exists as a Danish morph or is a recurrent part - a quasi-morph - of one or more Danish words; and it contains information on the normalized shape of that morph or quasi-morph. We shall see below how a letter sequence can simultaneously contain two sorts of information.

(2) a bit-pattern containing either/or information on grammatical and other characteristics of the entry. (In the current implementation the size of the bit pattern is 16 bits, which is the size of a PDP-11 machine word, but this is only a practical compromise between the amount of information needed for an entry and the hardware at hand and has nothing to do with the basic software design.)

The following subsections describe the general organization of the lexicon, some important details of the components of the entries, and some of the types of morphs and quasi-morphs the lexicon should contain.

2. THE ORGANIZATION OF THE LEXICON

The entries are arranged in several groups according to the length in letters of the item (morph or quasi-morph) they represent; thus all four letter items are in the same group, etc. Within each group the entries appear in the order that the items they represent would have in a retrograde dictionary, i.e., they are sorted alphabetically from the right end of the letter string, whereby upper case letters are evaluated as equivalent with the corresponding lower case letters.

The grouping according to length is motivated by the search algorithm of the text processor which will search for the longest possible match first.

The dictionary order of the items is motivated by the binary search algorithm that the processor invokes once the group length has been selected.

The retrograde principle is motivated by the fact that both right-to-left and left-to-right scanning technically proceeds letterwise right to left.

In addition to the entries, the lexicon contains some auxiliary information required by the processor in order to determine the beginning and the end of each length group.

3. THE COMPONENTS OF THE ENTRIES

As was mentioned above, each entry consists of a character string and a bit pattern.

The character strings are basically reverted copies of the morphs they represent; for instance, the character string part of the entry corresponding to the morph *lov* 'law' appears as "vol". The character strings in the entries contain both upper case and lower case letters. This is one of the ways in which identificational and normalizational information is condensed; for instance, the character string corresponding to *mund* 'mouth' appears in the lexicon as "dnUm", i.e. with upper case "U", but its place among the four letter words it is grouped with is the one it would have had if it had contained a lower case "u" instead. This structuring of data permits the processor to identify the entry string "dnUm" with the input string *mund* and to replace the latter with the string *mUnd*, which is the correct normalized notation of that morph when it occurs as an isolated word form.

The bit-pattern component contains a number of static one-bit flags which are set or cleared from the outset according to the behaviour of the entry.

For instance, in the current implementation each entry has a NOUN flag (set if the entry is a noun or may be part of a noun, cleared otherwise), a COMMON flag (set if the entry is (part of) a noun of the common gender), a VERB flag, and ADJECTIVE flag, a PLUR-R flag (set for a root entry if it takes *-r* in the plural, and set for the entry, whose orthographic form is *er* (since it can appear as a plural ending; the latter entry also has its VERB flag set, since it can appear as the present tense ending in verbs)). There is also a BETA flag (set for those entries which can appear as endings or quasi-endings in stødless word forms, and set for the quasi-roots which are of the regular stødless disyllabic type ending in schwa - among Danish phonologists often referred to as β-words, cf. section III E). Thus a word like *gade* 'street' is split up into a quasi-root *gad* and a quasi-ending *e* (the splitting up of words in quasi-morphs will be further commented on below).

Other important single flags are the ADD-H flag and the DOUBLE-CONSONANT flag, one of which must be set for those morphs and quasi-morphs which end in a stressed short vowel syllable and are spelled with less than two consonants after the stressed vowel: the ADD-H flag must be set in roots not taking stød when not followed by an ending, e.g. *tal* 'number, figure', and *nu* 'now'; the DOUBLE-CONSONANT flag must be set in roots that take stød when not followed by an ending, and in roots without stød-basis, e.g. *lem* 'limb', and *kys* 'kiss', cf. section III E.

The function of these flags is to supply the processor with the information needed to eventually select the correct normalized notation of an input word.

4. THE TYPES OF ENTRIES

In section III D the basic types of entries to be put into the lexicon were mentioned. Three other important features of the contents of the lexicon must be mentioned.

The idea of including in the lexicon the bulk of old mono- and disyllabic roots was motivated by the fact that, in addition to containing crucial information for the determination of stress and stød, this class of words contains the majority of irregular spellings of single segments. However, the decision that practically all such words should be accessible in the lexicon does not necessarily mean that they all have to be entered as such. These words have a property which makes it possible to reduce the size of the lexicon considerably:

A comparison of words like *mod* 'courage', *moder* 'mother', *modet* 'the courage', *moden* 'mature', *moderne* 'modern (unfortunately homographic with *moderne* 'the fashions'; but that is irrelevant to the phenomenon under consideration), *model* 'model', *modellen* 'the model', *models* 'of (a) model', *modellens* 'of the model', *modeller* 'models', *modellers* 'of models', *modellerne* 'the models', *modellernes* 'of the models', *modellere* 'to model', etc. suggests that we need only have entries for the quasi-root *mod-* and for the (quasi)-suffixes and/or (quasi)-endings *-en*, *-el(l)*, *-et*, *-er*, *-erne*, *-ernes*, etc. in order to generate these forms. Since the quasi-suffixes (cf. the remarks on foreign words with stressed (Latin) suffixes in section III A) and the inflexional endings are needed anyway, there is no need to list entries such as *model*, *moden*, *moder* separately. The forms may be generated by adding a few extra flags to the bit-pattern of the entries (such as a STRESSED-SUFFIX flag to be set for quasi-roots like *mod* and for the entry *el*, so that the processor can select *modell* as the correct normalization on finding that the input word contains matches for the entries *mod* and *el* in that order, but *mod-er* as the correct normalization on finding that the input word contains matches for the entries *mod* and *er* in that order.

Since potential endings like *er*, *erne*, *ens*, *et*, *ets*, etc. must be identified anyway, there is a problem with roots ending in the corresponding letter sequences: In a word like *lanterne* the *erne* sequence is identical to the final sequence in *planterne* 'the plants' where *erne* represents the definite plural ending of a noun. Since *erne* must appear in the lexicon anyway, there is no need to list *lanterne* in its entirety; indeed it would be erroneous to do so, for thanks to the maximalistic way of searching, cf. section III D and IV B, *lanterne* would be found before *erne*, and this would prohibit *planterne* from being correctly identified, unless, of course, still longer word forms, containing *lant* as a substring were to be listed too, i.e. inflected forms like *planterne*, *planternes*, *slanterne*, *slanternes*, etc., which is the very thing to be avoided, since stuffing the lexicon with numerous entries containing identical information would actually be to abandon the quasi-morph policy. What this means is that lexical entries which cannot appear as endings or other important right-end material are permissible as entries only if they do not contain a rightmost letter sequence which is (orthographically) identical with such an item. Thus a word like *stjerne* 'star' must be split up in the lexicon as a quasi-root *stj* and the quasi-ending *erne* which is needed anyway.

On the other hand, there are cases of homography between certain suffixes and inflected forms of other suffixes or endings which necessitate the inclusion in the lexicon of certain quasi-roots which do not belong to the bulk of old words. Thus, since word final *ens* can be (1) a definite genitive of a common gender noun, as in *gadens* 'of the street', (2) a homophonous quasi-ending as in *Assens* (the name of a town), (3) the stressed quasi-suffix *ens* (with a short vowel) as in *tendens* 'tendency', or (4) the genitive of a root ending in stressed /e:ˀn/ as in *arsens* 'of arsene', the two latter types must have entries for the corresponding quasi-roots (the *tend* of *tendens* and *ars* of *arsen*), since it seems safest to reserve the normalization *-ns* (definite singular genitive of a finally stressed noun) for foreign roots not found in the lexicon, so that e.g. a word like *mikadoens* 'the Mikado's' will be normalized as *mikado-ns*, consistent with its phonemic structure

/mi'ka:ˌdo:ʔəns/, and not as *mikadoen+s* (whitch the rule system would transform into /mikado'e:ʔns/), or as *mikadoens* (which the rule system would transform into /mikado'ɛnʔs/).

B. THE PROCESSOR

1. THE INPUT AND SCANNING ROUTINES

The processor operates in the following fashion:

Text is input one sentence at a time, in which process upper case letters are converted to lower case. For each word the length in letters, the number of syllables, and the position in the sentence is memorized.

Each word is subject to one or both of the following routines:

(1) The right-to-left scanning routine: The rightmost letters of the word are compared to lexical entries in the following fashion: if the word is no shorter than the longest group of entries in the lexicon, a match is looked for in that group, otherwise a match is looked for in the group whose length corresponds to the length of the word. If a match is found, the search is stopped, and the address of the matching entry is memorized. If no match is found in that group, a match is looked for in the group with entries shorter by one, and so on, until either a match is found or the whole lexicon has been scanned (thanks to the structuring of the lexicon and the design of the search algorithm, this does not, of course, mean that every entry in the lexicon has to be compared with the input). If no match is found in any of the groups, the right-end scanning is stopped, and program control transferred to the left-to-right scanning routine, cf. below. If a match is found, the whole right-to-left scanning routine is repeated, starting from the position in the word to the right of which the match was found. This routine is repeated until either the left end of the word is reached or no match is found. The left end of the word being reached thus means that the whole word was matched morphwise or quasi-morphwise, and program control is transferred to the interpreter routine, cf. below.

A non-match being recognized before the left end of the word is reached means that only the rightmost portion of the word was matched, viz. the portion from the position reached after the last match found to the right end of the word. In this case program control is transferred to the

(2) left-to-right scanning routine, which is (conceptually, at least) a mirror image of the right-to-left scanning routine, i.e. it scans the word from left to right up to, but not including the position reached by the right-to-left scanning routine, and memorizes the addresses of any matching entries.

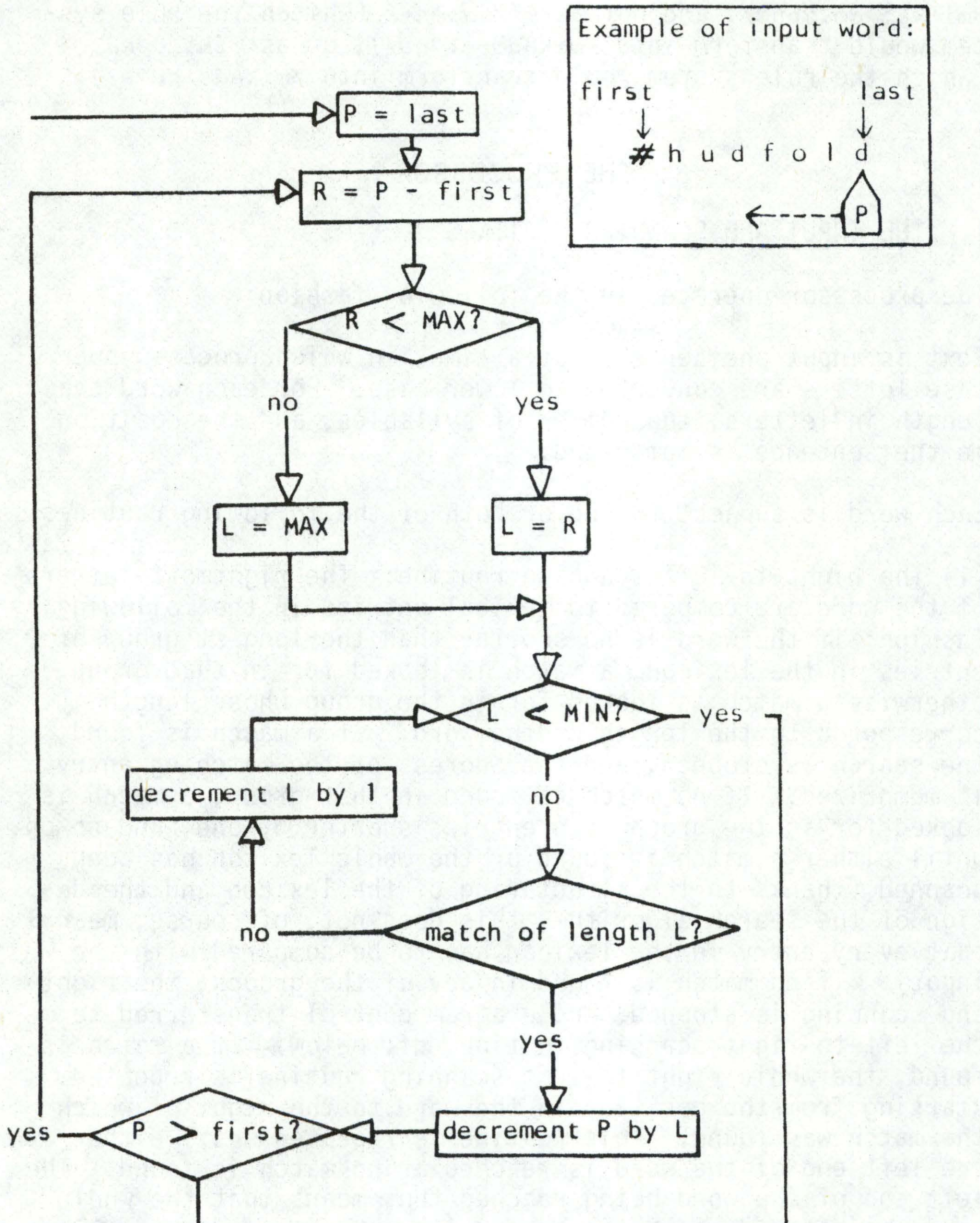


Figure 1

Flow chart of the right-to-left scanning routine. P *= current position in the word;* last *= rightmost position in the word;* first *= leftmost position in the word - 1;* R *= length of the current left remainder of the word;* L *= current length of match to be looked for;* MAX *= length of longest entries in the lexicon;* MIN *= length of shortest entries in the lexicon.*

What all this means is that the processor carves its way, as it were, from both ends of the word until either the word is matched in its entirety, or a left-end, right-end, or - more frequently - medial residue is left unidentified.

Figure 1 is a flow chart of the right-to-left scanning routine.

2. THE INTERPRETER ROUTINE

This routine is the intelligent part of the processor. It does the important job of combining and interpreting the information handed over from the input and scanning routines, and selects a normalization procedure according to its interpretation. The information at its disposal is (1) its knowledge of found matches and access to their entries (which, as we have seen, consist of two components, each with its own type of information), (2) its knowledge of unidentified residues, and (3) its knowledge of the position in the word of morphs, quasi-morphs, and the residue (if any).

The following example will give the reader an idea of the basic philosophy underlying the interpreter routine:

In a word like *afreagerer* 'abreacts, works off' the scanning routines will first identify the two rightmost letters as the lexical entry *er*. Then the next two letters are recognized as the same entry *er*. Then the next four letters are recognized as the entry *reag* (although this entry is neither a genuine root (or quasi-root) nor an affix, it should appear in the lexicon because of its combinability with the stressed suffix *ens*, cf. *reagens* 'reagent'; see further section IV A 4). Finally, the remaining two letters are recognized as the entry *aF* (see further below). In other words, the interpreter routine is presented with one of the easier cases, viz. that of the whole word being matched. It now begins to weigh the evidence, as it were, and its way of doing this can be described, somewhat metaphorically, as follows:

The rightmost item, *er*, considered in isolation, reduces the possible analyses of the word to be one of the following:

(a) a word in which *er* belongs to the root and is phonologically /ɛr/ (occurs probably only in two monosyllables, viz. *er* 'is', and the pronoun *jer* 'you (plur.)';

(b) one of a few monosyllables ending in /e:ˀr/, cf. *sner* 'snows' (where *e* belongs to the root and *r* is the present ending), and *fjer* 'feather' (where *er* belongs to the root);

(c) a polysyllable where *er* is a stem final syllable, in which case *er* is phonologically /e:ˀr/, cf. *klaver* 'piano';

(d) a polysyllable where *er* is unstressed and phonologically is /ər/, and where the preceding syllable takes stress and stød, cf. *mager* 'lean' (where *er* belongs to the root), *bøger* 'books' (where *er* is a plural ending), *kommer* 'comes' (where *er* is the present ending of a verb as in the actual example);

(e) the same as d), except that the preceding syllable does not take stød, cf. *bager(1)* 'baker' (where *er* is a nominalizing suffix) and *bager(2)* 'bakes' (where *er* is the present tense ending);

(f) the same as e), except that the preceding syllable is unstressed, cf. *magiker* 'magician' (where *er* is a nominalizing suffix), *hoveder* 'heads (noun plur.)' (where *er* is a plural ending).

Consider now how the identification of the next item from the right will reduce the possibilities: At this stage, thanks to the combined information of the flags of the two rightmost items - which happen to be identical - and of the positions of these items, possibilities a, b, and e can be immediately discarded, i.e. the remaining phonological representations of the last two syllables could be (i) /eˈre:ʔr/ as in *gerer dig!* 'behave yourself!', (ii) /əre:ʔr/ as in *moderer!* 'moderate! (imp.)', (iii) /ərər/ (which does not occur in Danish words), or (iv) /e:ʔrər/ as in the actual example.

The information obtainable from the next item from the right, *reag*, which is known (from the bit-pattern component of its entry) to be a verbal root of the type which is right-compatible with stressed Latin suffixes like *ens* (forming deverbative nouns), and *er* (forming verbal stems with such roots), will - when combined with the information accumulated so far - discard possibilities i, ii, and iii, and the processor can now be certain that it is dealing with a simple word or a compound the last part of which is the present tense of the verb *reagere*. The remaining, leftmost item *aF* is known to be a word which can appear as the first part of a compound verb like *afreagere*. As can be seen, the leftmost item appears in the lexicon in the form *aF*, where the *F* is a normalized segment equivalent with *v* when not word final. The procedure selected by the interpreter routine will now produce the correct normalized notation *aF=reager-r*. Incidentally, thanks to the maximalistic search algorithm, if the word had been *afficerer*, the third item from the right would have been recognized as *affic*, this item being listed in the lexicon because it belongs to a handful of quasi-morphs in which initial orthographic *af* is not the (genuine) word *af*, cf. Molbæk Hansen (1982, p. 132).

It is an important property of the interpreter routine that it is, at least in principle, designed to act like an average Danish reader. In the case of *afreagerer* it happened to recognize *reag*, just as a Danish reader would - *ceteris paribus* - recognize that part of the word as something which exists and which has, among other properties, the ability to fit in between the prefix /av/ and the /e:ʔrər/ part which would indicate that he was dealing with a finite verbal form. But the interpreter routine should - if properly designed - produce the same normalized notation of the word even if it had not found *reag* in the lexicon: suppose we did not know the *-reag-* (or, for the benefit of linguists and Latinists, *-re-ag-*) part, what would our guess have been? It would definitely have been that we were dealing with some foreign verb which we happen not to be familiar with, but which is of the *-ere* type, in this case prefixed with /av/; this is what any Dane would do if he were presented with a nonsense form like *afpumonerer*. (There would have been the technical difference that the right-to-left scanning routine, not finding a medial letter sequence (like *pumon*) in the lexicon, would have left it to the left-to-right scanning routine to identify *af*, but that is of minor importance here.)

Another illustration of the work of the interpreter routine is a comparison between its treatment of (1) certain forms of *damp* 'steam', (2) *drabant* 'halberdier, satellite', and (3) *urbant* 'urbane (adj. neut. or adverb)':

damp would be recognized immediately and, appearing without an ending, it would be normalized as *damp*. If it had been the definite form *dampen* 'the steam', the information that it is a finally stressed, common gender noun (its BETA flag is cleared, and its COMMON flag is set) combined with the entry *en*, which can appear as the definite ending for such nouns, would have yielded the normalized notation *damp-n*. If it had been *dampet* 'steamed', the information (thanks to another flag) that it is compatible with the entry *et*, which can appear as the past participle suffix, would have yielded *damp-et*; (NB: not *damp-t*, since the entry *damp* has its COMMON flag set so as to prohibit it from combining with the (normalized) ending *-t*, which belongs with finally stressed neuter nouns).

In *drabant* the *ant* part would first be recognized as either representing the adjectivizing or nominalizing suffix *ant* or the stressed, stem final syllable of a regular adjective in *an* with the neuter (or adverbalizing) ending *+t*. Next *drab* would be recognized as being compatible with certain suffixes like *ant*, and that would settle the matter, yielding the normalized notation *drabant*. Incidentally, this case illustrates the entry-saving policy mentioned in section IV A 4: the entry *drab* is, of course, needed as the neuter noun *drab* 'homicide'; thus, in addition to having its STRESSED-SUFFIX flag set - so as to make it compatible with the lexicon-suffix *ant* as in the case under consideration - this entry should have its NOUN flag set, and its COMMON flag cleared. Notice that, if a nonsense form like **drabens* were to appear as input, it would be normalized as *drabens* (which the rule system would transform into /dra'bɛnˀs/, cf. section IV A 4) if the following conditions were met: (a) *drab* was found in the lexicon, (b) *drab* had its STRESSED-SUFFIX flag set, and (c) the entry *ens* had its STRESSED-SUFFIX flag set like the entry *ant*. The entry saving is possible in such cases because, typically, only one of several possible combinations of morphs or quasi-morphs exists. Thus a word like *fodens* 'of the foot' would be normalized correctly as *fod-ns*, because it would have its COMMON flag set (like *ens* would), and this would be decisive even if it had its STRESSED-SUFFIX flag set too; but the latter flag would be responsible for the normalization *fodant* - quite similar to *drabant* - if the non-existent form **fodant* were to appear as input, cf. also below.

In *urbant*, finally, *ant* would be recognized as in *drabant*; then, supposing that *urb* is not found in the lexicon, the next item from the right would be recognized as the entry *b*, which has to be in the lexicon because it is the initial consonant in a word like *ben* 'leg' (where *en* represents a potential ending, cf. IV A 4). Finally, the remaining part of the word would be recognized as the entry *ur*. In this case the interpreter routine - if properly designed - ought to discard the possibilities *ur=b=ant* (since single consonants cannot be parts of a compound), and *ur=bant* (since (a) *bant* was not found in the lexicon, and (b) *ant* is probably either the monomorphemic stressed suffix *ant* or the bimorphemic structure *an+t*, cf. above). Of the remaining possibilities *urbant* (i.e. an unidentified root *urb* suffixed with *ant*) should be discarded since that suffix is supposed to have its (quasi)-roots - like *drab* and *fod*, cf. above - in the lexicon, and the default case *urban+t* should be chosen.

These examples ought to illustrate the most important properties of the interpreter routine and, in particular, of the combined function of that routine and the special structuring of lexical entries outlined in section IV A.

V. CAPABILITIES, DEFICIENCIES, AND PERSPECTIVES OF THE TNP

An implementation of the TNP with a comprehensive lexicon and with an optimal choice and distribution of bit-patterns ought to successfully normalize the following types of isolated words, provided that they have no homographs:

(1) Most simple words belonging to the genuine vocabulary, and all inflected forms of such words, cf. *mand* 'man', *lyve* 'lie', etc.

(2) The majority of those high frequency compounds which consist of material belonging to the genuine vocabulary, and all inflected forms of such words, cf. *søpindsvin* 'sea urchin', *springkniv* 'flick knife', etc.

(3) Most simplex foreign words - possibly with affixes - and inflected forms of such words, cf. *kalamitet* 'calamity', *vegetar* 'vegetarian', etc.

(4) Many compounds the initial and/or final parts of which belong to the genuine vocabulary, and inflected forms of such words, cf. *kvindeemancipation* 'emancipation of women', *pensionsalder* 'pensionable age', *efterrationalisere* 'rationalize (afterwards)', etc.

The main shortcomings of the TNP are due either to the maximalistic way of scanning the lexicon, or to the fact that the processor loses track of the morph boundaries in certain cases, notably in long compounds consisting exclusively of foreign material.

As Allen (1976) has rightly pointed out, the maximalistic principle is insufficient in cases where the morph composition is ambiguous as e.g. in *solur*; this form can be interpreted as either *so=lur*, literally: 'sow-lure', 'nap taken by a sow' (!) or *sol=ur* 'sundial', literally: 'sun-watch'. Of these interpretations the latter is, of course, superior, mainly for pragmatic reasons; but there is no morphological, syntactic, or phonotactic reason for not preferring the former interpretation. Because of the maximalistic, right-end-first policy of the processor, *so=lur* will be the interpretation chosen; but it is easy to adduce cases where the actual choice will, by chance, be pragmatically superior to the alternative, cf. e.g. *sømand* which could be either *sø=mand* 'sailor', literally: 'sea-man', or *søm=and*, literally: 'nail-duck', of which the former is a more or less lexicalized, common compound, whereas the latter is pragmatically dubious, to say the least, but nevertheless entirely acceptable as a Danish compound (note that the pronuncations would be different in the two cases).

As Allen has pointed out, an algorithm finding both (all) the combinations could in some cases evaluate them and "pick a winner". For an English example - *scarcity* - see Allen (1976 p. 436). It is unclear to me whether the extra software overhead needed to take care of such situations would be worth while for Danish. It is clear, however, that this difficulty can be easily overcome in applicational versions of a text-to-speech system by including in the lexicon the relatively few problematic cases any selected corpus of words (e.g. a frequency dictionary) would contain.

The other main deficiency is the fact that the processor will lose track of the morph composition of long foreign compounds like *basilarmembran* 'basilar membrane' (note that in this case, as in many others of the type in question, the word is matched by a noun phrase in English). This is, of course, serious in texts with a high frequency of such words. But such compounds are typically highly specialized words occurring in special (technical) texts for which special extensions of the core lexicon could be made.

So far, there remains two major categories of problems, viz. those of homography and of sentence phonological phenomena closely connected with syntactic structure. Note, however, that the policy of supplying each lexical entry with a bit pattern containing information on such lexical properties as word class etc. is compatible, almost *par excellence*, with overlaying the TNP with routines for exploiting such information. This perspective has been one of the main motivations for designing a TNP with properties as outlined in this paper.

The current work is concentrated on preparing facilities for testing the output of the TNP against well defined lexical material such as, e.g., the corpus of a frequency dictionary (Maegaard and Ruus (1979)). Although such tests will probably disclose numerous minor flaws and redundancies in the current implementation - in particular in the choice of grammatical and other flags in the bit-pattern component of lexical entries - the general strategy is believed to be sound.

VI. NOTE

1. In Molbæk Hansen (1982) the abbreviation XCO (exception removing component) designates what roughly corresponds to the TNP.

ACKNOWLEDGEMENTS

I am indebted to Hans Basbøll, Jørgen Rischel, and Nina Thorsen for valuable suggestions.

REFERENCES

Allen, J. 1976: "Synthesis of speech from unrestricted text", *Proceedings of the IEEE, vol. 64, no. 4*, p. 433-442

Allen, J. 1981: "Linguistic-based algorithms offer practical text-to-speech systems", *Speech Technology 1,1*, p. 12-16

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

Carlson, R. and Granström, B. 1975: "A text-to-speech system based on a phonetically oriented programming language", *Speech Transm. Lab., Quart. Prog. and Status Rep. 1/1975*, p. 17-26

Dewey, G. 1971: *English Spelling: Roadblock to Reading.* Columbia University, New York: Teachers College Press

Holmboe, H. 1978: *Dansk Retrogradordbog* (Copenhagen, Akademisk forlag)

Holtse, P. 1982: "Speech synthesis at the Institute of Phonetics", *Ann. Rep. Inst. Phon. Univ. Cph. 16*, p. 117-126

Lesmo, L., Mezzalama, M. and Torasso, P. 1978: "A text-to-speech translation system for Italian", *Int. J. Man-Machine Studies 10*, p. 569-591

Maegaard, B. and Ruus, H. 1979: *HYPPIGHEDSLISTEN DANwORD L1A*, Inst. f. anv. og mat. lingv. og Inst. f. nord. fil. University of Copenhagen

Maggs, P. et Trescases, P. 1980: "De l'écrit à l'oral. Un programme sur micro-ordinateur pour machine à parler à l'usage des aveugles francophones", *Le français moderne, revue de linguistique française no. 3*, p. 225-245

Molbæk Hansen, P. 1982: "The construction of a grapheme-to-phone algorithm for Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 16*, p. 127-136

Rischel, J. 1970: "Morpheme stress in Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 4*, p. 111-144

Sherwood, B. A. 1978: "Fast text-to-speech algorithms for Esperanto, Spanish, Italian, Russian, and English", *Int. J. Man-Machine Studies 10*, p. 669-692.

ARIPUC 17, 1983, p. 111-124

THE RELATION BETWEEN BLOWING PRESSURE AND BLOWING FREQUENCY IN CLARINET PLAYING

NIELS BAK AND PREBEN DØMLER

Different opinions on the relation between blowing pressure and blowing frequency in clarinet playing provide the background for the investigations presented in this paper. By employing mechanical blowing and a newly developed artificial embouchure, test results have been obtained which show that the correlation between blowing pressure and blowing frequency is positive for all notes in the low register of the clarinet (the chalumeau register) when all other parameters are kept constant.

I. INTRODUCTION

It is the clarinet player's unequivocal opinion that an increase in blowing pressure causes a fall in blowing frequency. From the literature on the clarinet by such authors as Benade, Worman, and Schumacher one receives the impression that variations in blowing pressure have but a slight effect on the blowing frequency. In the cases where changes do occur, these are ascribed to factors connected with the characteristics of the tube resonator, which are fixed once and for all when the tube - including the sound holes - is bored out (Benade 1976, p. 477, and Worman 1971, p. 101). These factors are then supposedly decisive for the positive or negative correlation between blowing pressure and blowing frequency.

This is contradicted by the test results obtained from experiments carried out in the Laboratory of the Institute of Phonetics at the University of Copenhagen. In fact, the results show that blowing pressure and blowing frequency are generally positively correlated. Actually, an indirect prediction of this may be deduced from the classic literature, viz. Weber's well-known lecture, reprinted in 1830.

That the conclusions of Benade (1976), Worman (1971), and Schumacher (1978) are at variance with the empirical findings presented here is probably due to their having underestimated the elasticity of the clarinet reed in their work with theoretical models of the clarinet. In their measurements of clarinet tube resonance the reed has been replaced by a compact, inelastic material.

Our experiments were arranged in such a way that the related values of blowing pressure and blowing frequency could be determined for a scale of notes in the clarinet's lowest register (the chalumeau register) as a function of time, partly by mechanical blowing, and partly with the assistance of classically trained professional clarinetists.

II. THE EXPERIMENTAL SETUP

A. EMBOUCHURE AND EQUIPMENT

An artificial embouchure has been developed in order to make possible measurements by mechanical blowing; it was mounted on a reversible wall (see figure 1). We have thus been able to collect reliable and reproduceable data and also to compare measurements for blowing and suction as well.

By removing the lead pellets and the attached bar from F (figure 1) wall, clarinet, and embouchure can - as a total system - be removed from the box, reversed and replaced. Suction is achieved when the hose (in figure 2 marked C) is at-

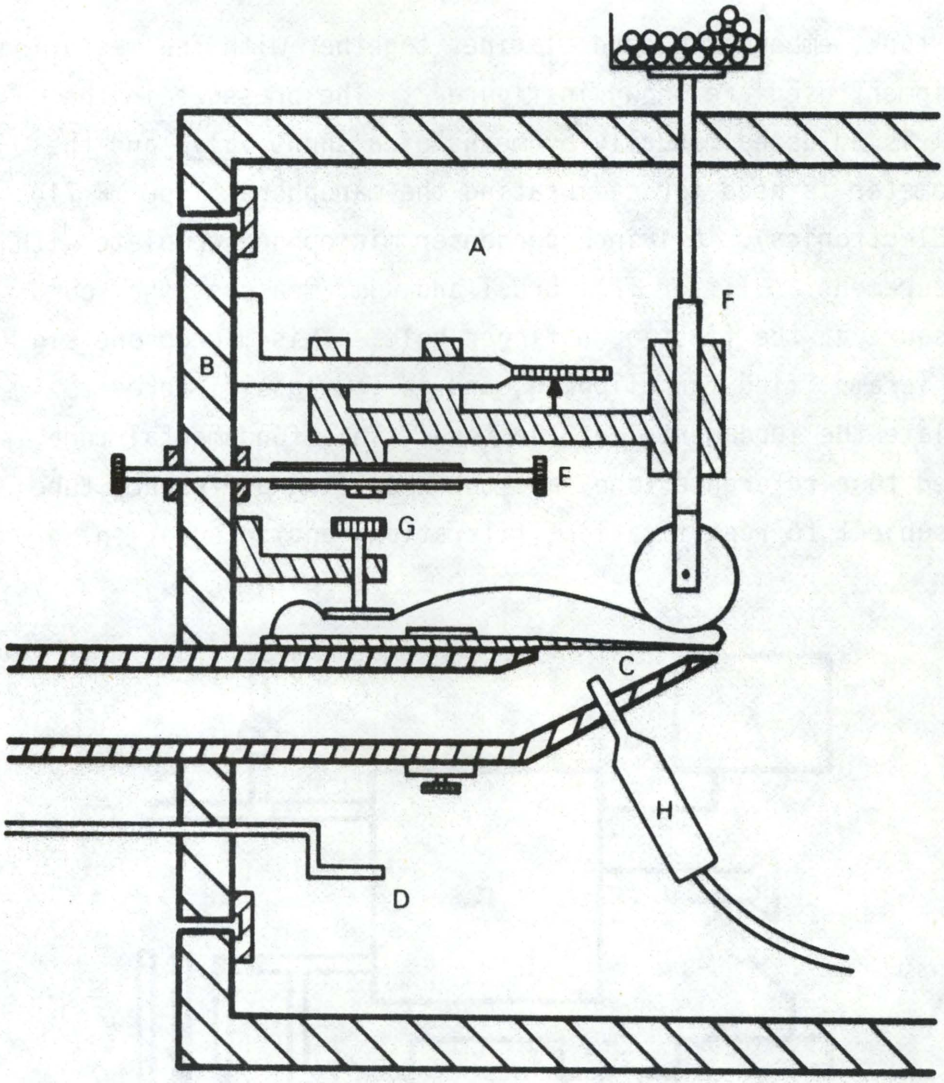


Figure 1

Schematic drawing of the artificial embouchure when mounted on the reversible wall of the air tank.
A) air-tight tank, B) reversible wall, C) mouth piece, D) remote control of speaker key used during suction, E) horizontally placed position device with adjustment screws and millimeter scale, F) lip pressure device; a container with lead pellets is placed on top of a rod which is supported by E, so that pressure is exerted on the nylon ball on top of the waterfilled balloon, G) lip damping device which regulates the water pressure in the balloon, H) quarter-inch Brüel and Kjær probe microphone.

tached to the suction unit of the vacuum cleaner.

Air tank, embouchure, and clarinet together with the measuring equipment used are shown in figure 2. The pressure in the tank is adjusted manually by means of a shunt valve and the manometer is used for calibrating the manophone (Type MF 710, FJ-Electronics). A 1-inch condenser microphone complete with measurement amplifier from Brüel and Kjær measures the sound pressure at the first open finger hole. This microphone signal is amplified for clipping, and is low-pass filtered to isolate the fundamental of the tone. This fundamental tone is added to a reference tone, and the resulting difference tone is subject to rectification, filtration, and integration.

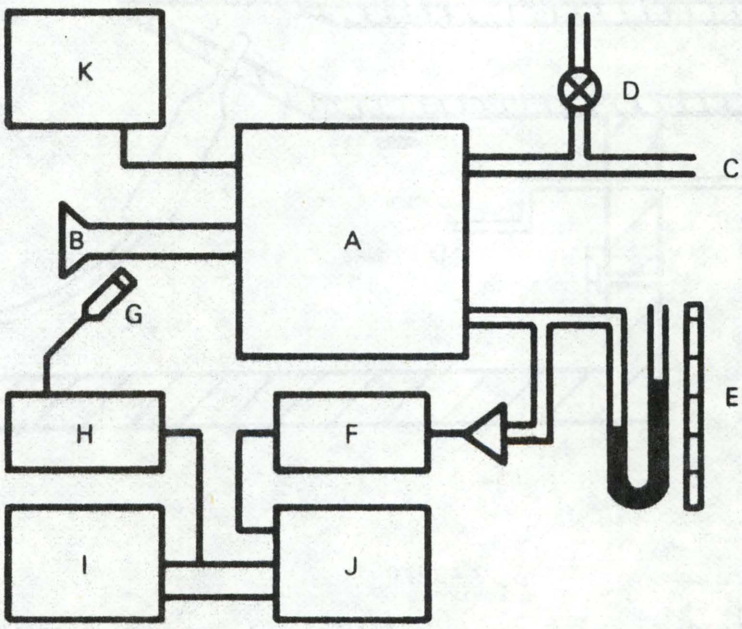


Figure 2

Schematic drawing of the experimental setup used by measuring sound pressure level (internal as well as external), blowing pressure, and blowing frequency.
A) air tank, B) clarinet, C) hose attached to an ordinary vacuum cleaner, D) shunt valve, E) Manometer, F) pressure transducer and amplifier, G) 1-inch condensator microphone, H) measuring amplifier, I) connected laboratory equipment for measuring blowing frequency, J) Mingograph (ink writer), K) spectrum analyzer.

In this way a DC-signal is obtained which - once necessary calibration has been carried out - is a measure of the deviation of the clarinet's tone from the equally tempered scale based on A_4 = 440 Hz.

The signal from the probe microphone (in figure 1 marked H) is analyzed by means of a spectrum analyzer, type 2033 from Brüel and Kjær. This apparatus makes it feasible to determine the position of the resonance peaks during blowing without excitation by means of the "running average" facility. In this way the internal noise signal of the mouth piece fills out the resonance peaks, and the resonance curve can be read as the resulting envelope curve.

Sound pressure from the external condenser microphone, as well as blowing pressure and blowing frequency were all registered as a function of time on a Siemens-Elema Mingograph. In experiments with "live" clarinetists, the blowing pressure was measured by inserting the manophone tube in the clarinetist's oral cavity at the corner of his mouth.

B. DEFINITION OF TERMS

A definition of key terms will facilitate the reading of the remaining pages:

P	blowing pressure; an independent variable.
P_{thr}	the threshold blowing pressure, i.e. the level where excitation begins.
P_{close}	the upper threshold blowing pressure, i.e. the level where excitation ceases.
f_{blow}	blowing frequency of the tone in question, a dependent variable.

Frequency deviations of the tempered scale's different tones are marked in cent, 100 cent equating one tempered semi-tone. The concert pitch is set to 440 Hz.

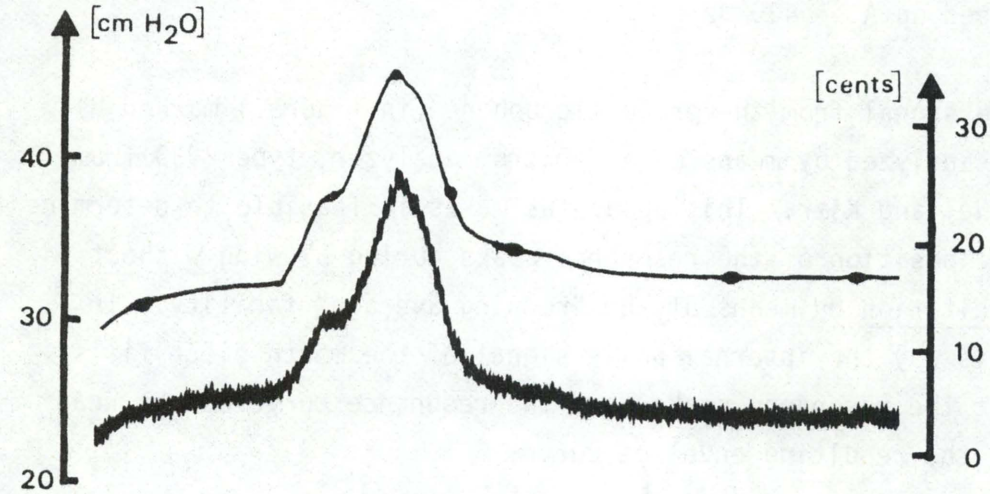


Figure 3

Mechanical blowing with artificial embouchure and constant lip pressure. The curves depict from top to bottom: blowing pressure and blowing frequency as functions of time.

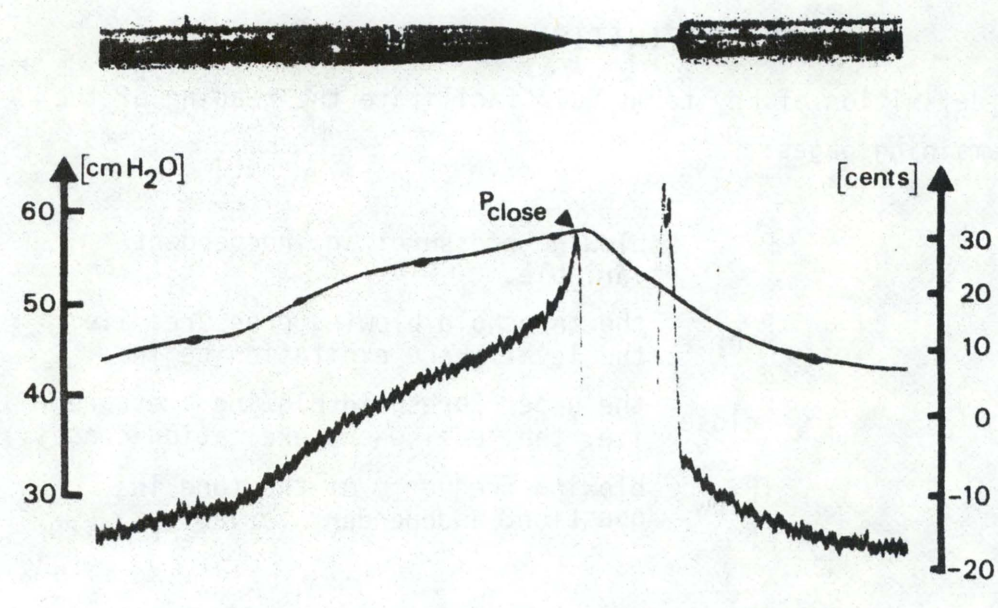


Figure 4

Mechanical blowing with artificial embouchure and constant lip pressure. The curves depict from top to bottom: sound pressure, blowing pressure, and blowing frequency as functions of time.

III. RESULTS

The characteristic correlation between P and f_{blow} appears from figures 3 and 4. Further, figure 4 provides an example of the cessation of excitation when $P = P_{close}$ and shows that oscillations are reestablished when the blowing pressure is reduced to a value somewhat below P_{close}. An equivalent hysteresis effect occurs at P_{thr}. This is a technical advantage as it makes it possible to measure f_{blow} and the first resonance frequency in turn at identical blowing pressures.

Natural clarinet playing is represented by the curves in figure 5, in which a professional performer of the classical clarinet carries out a register transition from F_3 to C_5, the only requirement being that this change be made without interruption

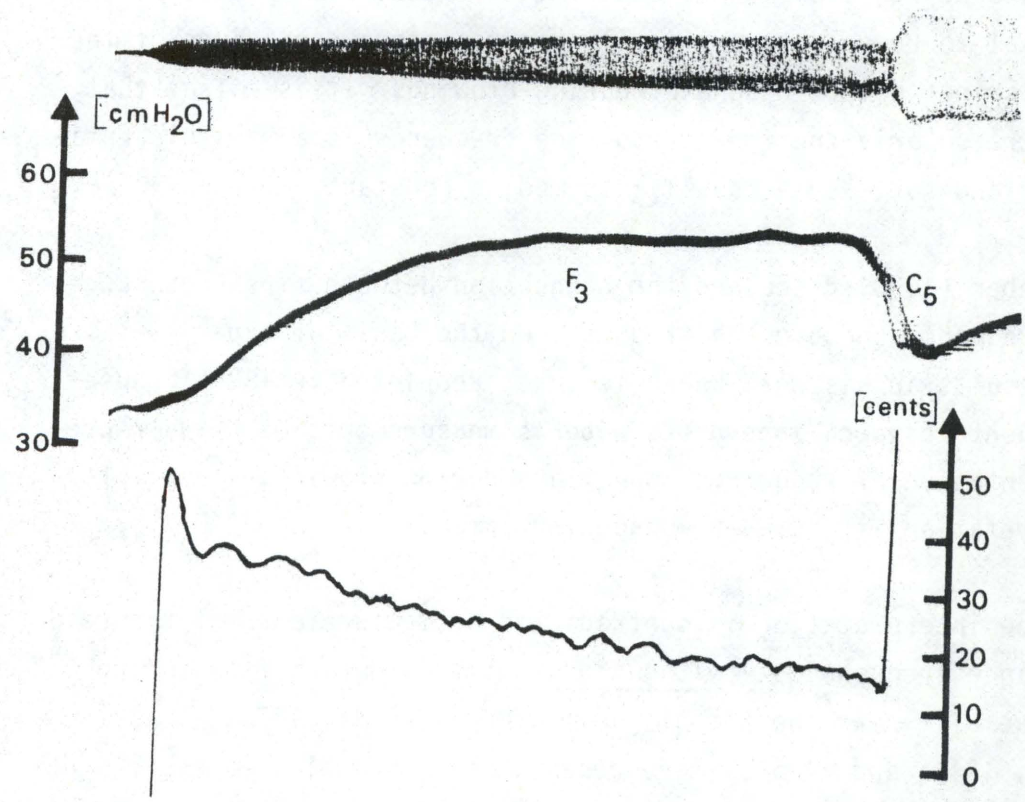


Figure 5

Professional performer of the classical clarinet. The curves depict from top to bottom: sound pressure, blowing pressure, and blowing frequency as functions of time.

of the tone production. It is interesting to note that the musician intonated a soft and relatively high tone during the initial intonation. From listening we can ascertain that the tone in question was a typical classical clarinet tone, which is characterized by a dominance of odd harmonics. During the maintenance of the tone F_3, P and f_{blow} are negatively correlated. Why do we find a negative correlation in this case, when mechanical blowing (figures 3 and 4) always shows a positive correlation? The explanation is simple: The musician may alter more than one of the embouchure parameters, whereas during the mechanical blowing only P was changed.

IV. DISCUSSION

A. THE REED CONDITIONING THE TUBE RESONANCE

In order to explain the experimental results it proves essential to reevaluate critically the resonance conditions of the clarinet's tube resonator during blowing. It is a fact that particularly the first resonance frequency is a variable entity and not, as is usually assumed, a constant.

Weber (1830) discussed the connection between first resonance frequency and blowing frequency on the basis of experimental results in his well-known lecture, reprinted in 1830. Subsequent research interprets Weber's measurements as being representative of resonance conditions during blowing. This interpretation is, however, unsubstantiated.

The interpretation is contradicted by measurements of the resonance frequencies - without excitation - during flow through the slit near the tip of the mouth-piece (Bak 1978, p. 132 and p. 134), and also by more recent, still unpublished experiments which prove that first resonance frequency - at constant lip pressure - drops drastically as pressure rises between 0 and P_{thr}. Around the threshold P_{thr}, first resonance frequency converges with the f_{blow} measured just when excitation begins.

In other words: within this pressure range the clarinet changes acoustically from a quasi-halfwave resonator to a quarter-wave resonator, and first resonance frequency is here negatively correlated to blowing pressure. This detail is vital for the understanding of the properties of the resonance tube.

In the pressure range P_{thr} to P_{close} measurements show that first resonance frequency and thus f_{blow} is positively correlated to P, and there is actually a natural explanation for this. What has till now been left out of consideration is the fact that an increasing blowing pressure forces a larger and larger part of the reed towards the curved part of the lay by the extreme tip of the mouth piece. Thus the flexibility of the reed decreases more and more just in the place where the internal sound pressure is at its highest. This has the effect of reducing the tube resonator's acoustic length, and the resonance frequencies - especially the first resonance frequency - rise. By constant blowing pressure an increase in lip pressure will, of course, have the same effect.

It is thus clear that at least two independent parameters can alter f_{blow} during the clarinetist's natural blowing. Changes in blowing pressure and lip pressure do, however, influence the dynamics and timbre of the tone in completely different ways:

1) At constant blowing pressure an increased lip pressure causes a reduction of the slit height whereby the volume flow decreases, with the result that sound pressure drops.
2) At constant lip pressure, a rise in blowing pressure produces a change of tone timbre and volume (Schumacher 1978, figs. 2 and 9), but whether this leads to a smaller or larger content of higher harmonics and whether the sound pressure rises or falls, depends entirely on the strength of the blowing pressure in relation to P_{thr} and P_{close}.

The experienced clarinet player is capable of adapting the variable parameters in such a way as to balance tone timbre and dynamics, and thus to maintain the required pitch. This is evident from figure 5, and we are now able to give a qualitative description of the missing parameter: the lip pressure.

Because the musician choses to blow softly in the beginning, lip pressure must be high if too low an intonation is to be avoided. Moreover, we know from experience that the chosen embouchure (high lip pressure and a relatively high P) will result in the typical classical clarinet tone in which the uneven harmonics - and in particular the first harmonic - are dominant. The coming change of register is determined by the first harmonic being decreased relatively to the third harmonic. To change the spectral composition of the tone, the performer has to decrease his lip pressure while at the same time increasing his blowing pressure. The change of register is now prepared and all that remains is to activate the speaker key.

B. THE CLARINETIST'S JUDGEMENT

Are clarinetists right then in maintaining that a fall in frequency - as shown in the bottom curve in figure 5 - occurs as a consequence of increased blowing pressure? The answer is no! All our conclusions so far point to the opposite. The lowering of the frequency is a result of the decreasing lip pressure. The fact is that not only a larger P but also an increase in the amount of air exhaled demands more of the performer's expiratory muscles. But whether the energy is used to increase pressure or to increase air flow he cannot possibly judge. His experience is the same. As shown by the empirical findings presented above, the blowing frequency for a given note is - contrary to common belief - positively correlated with blowing pressure. One might then ask what happens with the blowing pressure when an ascending scale of notes are played. Although this is slightly outside the scope of the present report, figure 6 is included to illustrate that on this point, too, musicians often misjudge what happens physiologically when playing the clarinet. It is a widespread assumption among clarinetists that the higher the note, the higher the blowing pressure. Figure 6 clearly shows that this is not the case.

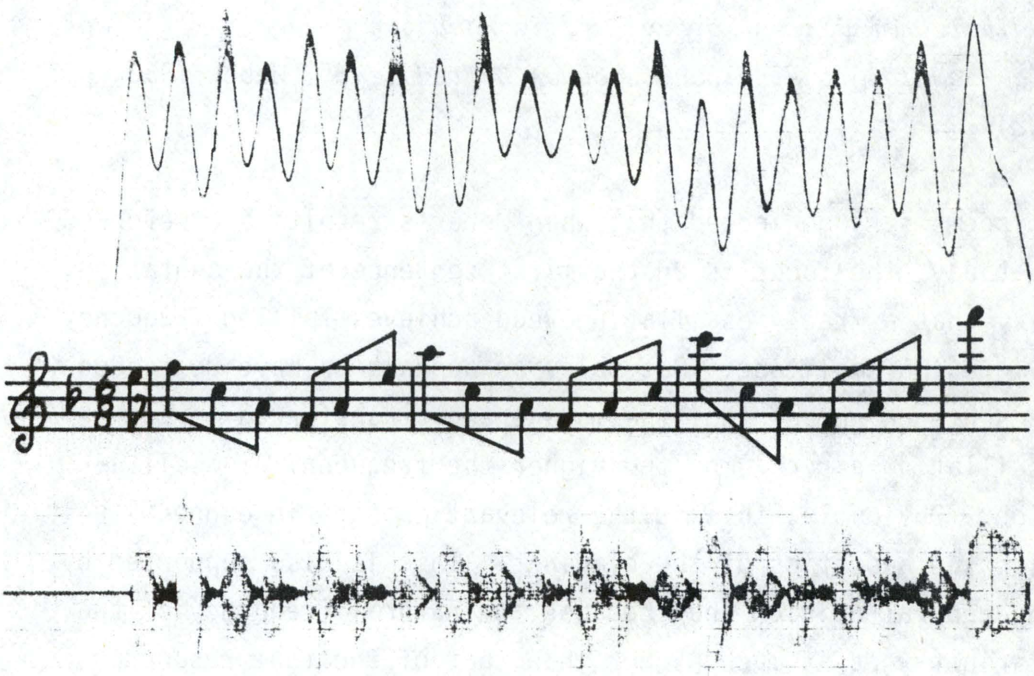


Figure 6

Fragment of Telemann's Triad in F-major played by a classically trained clarinetist. The top curve is the blowing pressure as a function of time, and the bottom curve is the sound pressure level as a function of time.

C. WEBER'S FINDINGS

Have no earlier researchers demonstrated the intimate relationship between P and f_{blow}? Apparently it has gone unnoticed that Weber is actually calling attention to this relation in his lecture: *"Ueberhaupt ist es eine Eigenthümlichkeit aller transversalschwingenden Körper, dass ihr Ton etwas tiefer bei stärkerer Schwingung, etwas höher bei schwächerer Schwingung ist. Die umgekehrte Eigenthümlichkeit haben aber alle longitudinalschwingenden Körper, und im höchsten Grade findet sie sich bei longitudinalschwingenden Luftsäulen;"* (Weber 1830, p. 402)

It seems strange indeed that when Weber's results are referred to today, the focus is on the first sentence of the quotation, viz. that a freely oscillating reed achieves falling frequency by growing amplitude. But oblivion appears to have descended on sentence number two: the more energy transferred to an oscillating air column, the higher the frequency of oscillation. Obviously, this is the relevant passage in connection with the acoustics of the clarinet. This is also supported by the general opinion that because the natural frequency of the clarinet reed is much higher than that of the tube resonator, and because much more oscillatory energy is accumulated in the air column than in the reed, it is the air column which is the driving oscillator and the reed which is the driven oscillator.

But Weber's experiments, which primarily served to clarify the blowing conditions of a freely swinging metal reed in combination with an organ pipe, cannot give an exhaustive explanation for every sequence of events during the blowing of the clarinet. For instance, the course of the frequency curve around P_{close} (figure 4) cannot be accounted for on the basis of Weber's findings.

V. FINAL COMMENTS

A modern symphony orchestra demands of a wind instrument the ability to vary a note greatly, both as concerns loudness and spectral composition, without the blowing frequency being affected.

That the clarinet is capable of meeting these demands is particularly due to the conditions investigated in this report: During playing, at least two parameters are positively correlated to f_{blow}, viz. blowing pressure and lip pressure. This provides the musician with almost inexhaustible possibilities for variation of tone and volume without getting off-key.

Similar options are not available when playing the traditional wood wind instruments from which the clarinet was originally developed. On these, only P may affect f_{blow}. They were, however, both useful and highly valued in older times when the esthetics of musical tradition differed from that of today. Furthermore, ensembles that cultivate music other than the symphonic may still employ these instruments where only one parameter determines the blowing frequency. Within this group of instruments the recorder in particular has been recognized as an excellent solo instrument under the right conditions.

However, a modern wind instrument must have at least two variable parameters which are positively correlated to the blowing frequency. The best way to realize this is perhaps to imagine a clarinet-like instrument whose blowing pressure is negatively correlated to f_{blow}, while its lip pressure is positively correlated to f_{blow} (Schumacher 1981, p. 81). It should be self-evident that such an instrument would not survive for very long, as its playing qualities would be very poor.

ACKNOWLEDGEMENT

The special instrumentation for this investigation was acquired on the basis of a grant from the Research Council of the Humanities.

We wish to thank the staff of the Laboratory of the Institute of Phonetics at the University of Copenhagen for their invaluable help in the setting up and carrying out the experiments upon which this paper is based.

REFERENCES

Bak, N. 1978: "Slit height and blowing pressure as parameters conditioning the function of the clarinet as a quasi-quarter-wave resonator", *Ann. Rep. Inst. Phon. Univ. Cph. 12*, p. 113-141

Benade, A. 1976: *Fundamentals of musical acoustics*, Oxford University Press, New York

Schumacher, J. 1978: "Self-sustained oscillations of the clarinet: An integral equation approach", *Acustica 40*, p. 298-309

Schumacher, R. 1981: "Ab initio calculations of the oscillations of a clarinet", *Acustica 48*, p. 71-85

Weber, W. 1830: "Compensation der Orgelpfeifen", *Annalen der Physik und Chemie, stk. 11*, p. 397-408

Worman, W. 1971: "Self-sustained non-linear oscillations of medium amplitude in clarinet-like systems", Ph.D. Thesis Case-Western Reserve University (unpublished).

ARIPUC 17, 1983, p. 125-144

LINGUISTICS, PHONETICS, AND FIELD-WORK

JØRGEN RISCHEL

The present, rather loosely structured and causerie-like paper is an attempt to summarize some of the reflections, expectations, and frustrations which have marred the author during several years of pendling between instrumental phonetics and theoretical and descriptive linguistics, and during many years of practical field-work. The main emphasis is on the - possibly futile - question: is field-work data likely ever to be of any interest to instrumental phonetics, or vice-versa? No attempt is made to deal systematically with general methodological aspects of field-work (for this reason also, references to the literature on field-work techniques are totally omitted). Moreover, what is said below about conditions to be met in connection with instrumental analysis, will be trivial to most readers. Nevertheless, it may be worth while spelling out <u>*what*</u> *it is that causes field-work activities and instrumental phonetic research to exhibit little or no mutual interaction in contemporary language study.*

I. LINGUISTICS VERSUS PHONETICS

What are the goals of linguistic research? What are the goals of phonetic research? At times the issue is stated as if those interested in language come in two totally different species: Linguists and Phoneticians, the assumption being that the research paradigms and interests of these two species of mankind overlap so little that there are limited possibilities of fruitful interaction. Admittedly, it is not difficult to establish a polarization between two extremes: the linguist who is devoted to abstract formalism and considers phonetic detail irrelevant, and the phonetician who is totally absorbed in physical or physiological measurements. Researchers in the two camps (henceforth referred to as "linguists" and "phoneticians", although the present writer, like many others, does not see this as a real dichotomy) may feel that their research interests overlap only (if at all) in the very general sense that it may be considered a common research objective of theoretical linguistics and instrumental phonetics to contribute to our understanding of the basic principles and constraints underlying human language. Some linguists and phoneticians may also agree on searching for explanations of why there are non-trivial universal constraints on what is a possible language, or a possible linguistic performance, but they may find that they differ profoundly in their assumptions about where to look for such explanations, and in their understanding of the very concept of Explanation. (The difference of opinion may extend also to the criteria to be met in order for a general feature of languages to count as a universal.)

The possible polarization referred to here has to do with a parameter of abstractness, this term being understood here in a very general sense. Phoneticians anchor their research in physically observable and hopefully quantifiable phenomena. Some linguists would claim that what counts in their analysis of raw data is discreteness from the point of view of semiotic function, viz. on the one hand syntagmatic discreteness (segmentability in one or several hierarchical layers) and, on the other hand, paradigmatic discreteness (distinctness and membership of paradigms). And then there are theoretical linguists whose real object of study is constituted by highly generalized and highly abstract relationships between abstract entities, and to whom the observable "surface" data are only marginally interesting in that they reflect these abstract structures. (This characterization of "abstractness" in linguistics is, of course, grossly over-simplified in that it refers solely to the relationship between linguistic form and phonetic substance and totally disregards related issues in semantics and pragmatics. Moreover, discreteness is only one among several possible components in the semiotic functioning of linguistic entities. Hopefully, this over-simplification may be tolerated for the sake of the argument.)

Needless to say, it is not generally true that phoneticians and linguists fall into totally separate categories in terms of "abstractness": there are phoneticians who take an interest in taxonomies and sometimes quite abstract formalizations of phonetic findings, and there are linguists who want to "hug the phonetic ground closely" (Ch.F.Hockett), also cf. the recent controversies about abstract versus "concrete" or "natural" trends in phonology. Nevertheless, it is widely held that there is a deep abstractness schisma between (much of) modern linguistics and phonetics.

There are other parameters than "abstractness" which are essential to a taxonomy of research activities in the universe of language study (parameters which are not necessarily orthogonal to each other or to that of "abstractness", but which nevertheless merit separate mentioning). One such parameter has to do with the question of language specific versus universal phenomena, and there is another, closely related distinction to be made between descriptive statement and general theory. It should be noted that on these issues there is no categorial difference between the full-fledged linguistic theoretician and the full-fledged phonetic experimenter: it is not possible to characterize the work of the latter as being generally less theoretically oriented than that of the former. The phonetician who works in the framework of certain (possibly rather speculative) assumptions about the temporal organization of speech in the central and peripheral nervous system of the speaker, or about the mechanisms underlying speech perception, is in a sense theorizing just as much as the linguist who posits universal principles of markedness, prosodic hierarchization, or rule application.

It is true, however, that the hypothesis-forming phonetician and the linguist may differ in the kind of corroborative evidence they look for. The experimenting phonetician shares the concepts of natural sciences and of experimental psychology as to what constitutes empirical evidence for or against a hypothesis. The theoretical linguist may take a stand anywhere between this strictly empiricist attitude and another extreme, viz. that of substituting labels and definitions for real evidence. In actual practice most linguists working with spoken language place themselves in between these extremes, drawing more or less on the speaker-listener's (possibly their own) intuition about the language as primary empirical data. Other linguists programmatically look for substantive evidence for universal and language specific structurings, searching in areas such as language acquisition, language change, slips of the tongue, pathological speech disturbances, and psychological testing of normal language behaviour. However, there is so far no generally recognized paradigm for integrating such studies into theoretical linguistics, and this is one reason why the study of language is as far from Unity of Science as ever.

Purely descriptive studies of languages by necessity furnish the main input to the formation of interesting hypotheses. This is equally true inside and outside of phonetics proper. Linguists may perhaps think that instrumentally oriented phoneticians tend to take delight in the very craftmanship of recording and processing physical or physiological data, and that they do not always bother enough about the relevance of their meticulous measurements in the wider context of language sciences. Phoneticians, on the other hand, may perhaps think that linguists prefer to recede into a sphere of abstraction in which the awkward and conflicting evidence furnished by real and unbiased data sometimes presents more of a nuisance than of a welcome challenge. Of course, neither of these two prejudices is fair (although they may reflect a real enough tendency toward a difference in scientific temperament). And more interestingly, the attitudes of linguists and phoneticians toward the use of real data from real language use is in fact not at all in line with the hypothetical contention one might be tempted to forward on the basis of the above-mentioned characterizations of linguists as abstractionalists and of phoneticians as empiricists, viz. the contention that phoneticians are necessarily more interested in such real data than are linguists.

The hypothetical contention might run as follows: Phoneticians deal with language performance and should hence prefer raw data reflecting the unconstrained use (production and perception) of language. Linguists are in the first place concerned with linguistic competence (performance phenomena being a secondary object of study which is derivable from the study of competence), and hence are better off with idealized data without interference from communicative situation contexts or from variability in speaker (or listener) performance. - Off-hand this may perhaps seem a meaningful contention, and even a legitimate difference of research strategy, but clearly that is not the truth of the matter. Anyone working seriously and responsibly on syntax, for example, must realize that one does not get very far - if the statements are to represent a reasonable degree of adequacy - unless one looks (also) at genuine, non-monitored and non-idealized data (be it real texts in written language or real specimens of spoken language). Concoction of sample sentences serving to reassure the linguist of the correctness of his assumptions is convenient and useful for restricted purposes, but it is not enough, not even if the assumptions are derived from the most fashionable general theory of syntax. There may be linguists who prefer to remain in the belief that the language they have chosen to study (read: chosen to use for illustration of some general point) is the way it should be according to theory, rather than try to find out how it really is, but this cannot possibly constitute a legitimate research strategy, and probably all linguists would emphatically deny that they take such an attitude toward their data. And indeed, in descriptive syntax as much as in any other field of language study (and indeed much more than in phonology, for example) extensive study of real texts has always been the norm. Moreover, there has been a rapidly increasing interest in the study of texts as a component of communicative situations, the study of text coherence, and so on. Unfortunately, this trend has, so far, not manifested itself equally strongly in phonology, with the result that the study of prosodic phenomena is sadly lagging behind, because linguistic theory has not been developed sufficiently to cope with the intricate interplay between pragmatics, semantics, syntax, and phonology-phonetics which is so characteristic of prosodic phenomena at the sentence level.

So much in order to underline the inadequacy of idealized data as the only input to linguistic study. What then about the phonetician and his data? Phoneticians are of course basically interested in the issues mentioned above, and when dealing with prosodic phenomena, for instance, the instrumental researcher tries to take the complex scenario into consideration in the design of his techniques. However, there is a very serious difficulty which is rooted in the very nature of instrumental research: with the use of quantitative methods it is hard to come to grips with data varying in very complex fashions and involving factors which are in part inaccessible to quantitative treatment with present-day techniques. How is one to present quantitative data on sentence intonation and sentence rhythm, for instance, if the input sentences vary in segmental composition, in focal stress placements, in the use of intonation perturbations or hesitation phenomena to indicate speaker's attitude, and so on? Even at the level of segmental studies such factors as sentence length, sentence rhythm, and expressive lengthening may totally change the picture, and the same is true of differences in the distinctness level aimed at by the speaker (which may even be rapidly varying within an utterance, some items being enhanced and others slurred for the purpose of effective communication). On this background it is no wonder that, paradoxically, the phonetician - more than the explicitly data-oriented syntactician - tends to work on idealized data such as stereotype series of words, phrases or utterances spoken in suitable (invariant) contexts and designed to be as devoid af variable pragmatic factors as possible. The enormous popularity which nonsense words enjoy in phonetic research is witness of this trend. This restrictive research strategy is certainly legitimate as long as the issue is defined in terms of parameters (of sound production, acoustics, or perception) such as tongue root advancement, VOT, vowel duration, intrinsic pitch of vowels, or difference limens for formant frequencies, but obviously one has not said all there is to be said about the physical aspect of aspiration or vowel length in a language by measuring test words of certain types in stereotype utterances. Similarly, by making identification and discrimination tests on specially designed stimuli one certainly has not said all there is to be said about the perceptual categorization of consonants with varying amounts of aspiration or vowels with varying durations in real speech situations. One has, in fact, said rather little, and the experimenting phonetician knows this, but nevertheless he feels it is necessary to stick to the type of data he can handle with rigid methods, viz. data with few and well-known variables.

It should be emphasized also that the discovery and description of physiological and auditory mechanisms and constraints operating in speech production and perception, which is a major goal in general phonetics, can proceed quite far on limited types of data and rather require that the sampling of data be rich enough to permit satisfactory statistical treatment. It is, moreover, important to be able to distinguish between properties of speech production and perception which are strictly universal and organically conditioned, and properties which are sensitive to differences in linguistic patterning (some phoneticians tend to make too little of this distinction, perhaps). But in order to study the role of linguistic patterning it is often more immediately rewarding to take a narrowly defined type of data and study what happens with comparable data across a variety of languages than to study what happens across all kinds of contexts within one language.

One potential danger of the approach described above is that in narrowing his scope the phonetician may suppress his sense of proportions vis-à-vis natural language or, put differently, may define the goals of phonetics in a self-assuring way, so that phonetic research tends to circle around a limited set of standing issues forming their own closed universe without the researchers being sufficiently responsive to challenges from linguists who take a different approach and therefore raise quite different issues which also belong within the sphere of interest of phonetics in a broader sense. (It should not be overlooked, however, that this very tendency to keep circling around a set of highly specific, live issues is - at least equally if not more - characteristic of modern theoretical linguistics.) More generally speaking the phonetician has to face the fact that he is caught in a situation which is not very inviting from the point of view of sentence level phonetics: there is a kind of trading relationship so that one can either study some (artificially) limited aspects of linguistic performance by exact methods, or one can make relatively loose qualitative statements on the basis of observations of linguistic performance in a more realistic setting.

As stated earlier, linguistic research may suffer even more if its input is restricted to sentences or wordforms which have been filtered through the analyst's prejudices, but it is not all that simple to take in "unretouched" data in linguistics either. Non-monitored spoken texts are difficult to handle, especially if they stem from a complex communicative situation, and it is no wonder that the vast majority of syntactic studies have been done, and still are done, on written rather than spoken language. (Needless to say, both kinds of studies are necessary, and one cannot ever substitute totally for the other.)

II. FIELD-WORK DATA

Now, what kinds of input do linguists actually use? It is essential to distinguish here between first-hand data (data gathered by the researcher himself) and second-hand data (data gathered by somebody else). Along another parameter it is essential to distinguish between non-monitored and monitored data gathering, the former being what happens if, for example, the researcher taps telephone conversations or takes literary texts as data for a study of written language, and the latter being what happens if the researcher works with an informant. Along a related parameter there is a highly significant distinction between reproductive performance (e.g. when retelling a piece of oral tradition whose style is fixed by convention or, at the extreme, when reading a written text aloud) and creative performance (ranging from the restricted creativity involved in giving equivalents to sentences put in another language to the unlimited creativity involved in spontaneous speech). Monitored data, moreover, may represent non-interactive data gathering, e.g. if the informant is asked to render a narrative text or engage in a conversation with somebody else, or it may be a case of interactive data gathering, viz. if the researcher directly enters the process by which the linguistic material is generated, e.g. by putting questions to the informant, asking about his acceptance or non-acceptance of certain data, or having him record specimens worked out in advance. - Strictly speaking the user of field-work data should be conscious about the implications of using one or another kind of data.

Many linguists take much less interest in data gathering than in linguistic generalizations and, consequently, tend to use second-hand rather than first-hand data. In phonology this may be practically hopeless unless the source one consults already represents a systematic linguistic analysis using consistent criteria such as distinctness. In syntax it may be the other way round: a source rich in raw data with little or no systematization being forced upon the data may often be less misleading than a selective and fully formalized presentation, if the purpose is to make a "more insightful" restatement. Often enough, such analyses involving sweeping generalizations on the basis of very limited material may eventually turn out not to be warranted by the language as it really is, apparent new insights being artifacts of the accumulation of skewnesses caused by repeated restatement on the basis of the same primary material. Needless to say, the danger is imminent if the source consulted is itself a case of highly selective use of data stemming from interactive gathering. Such accumulation of errors may, of course, happen in all fields of linguistic research, and it certainly happens also in phonology (this is one reason why it is necessary to be somewhat cautious in using typological reference manuals, impressive and immediately useful as these may be).

It would help if it were always clearly stated in papers dealing with phenomena in specific languages on what kinds of data (in terms of parameters such as those mentioned above) the statements are based and, if the paper contains illustrative data, what is the status of the data included in the paper. Nevertheless, it cannot be refuted that the use of second-hand data is always somewhat problematic whenever the data must be sought in a presentation of the same kind as one's own study, as when one takes a specimen of phonological or syntactic analysis and restates it into a phonological or syntactic analysis of the second order. Second-hand data is less controversial (though certainly not unproblematic) in case it is drawn from a different kind of source (as when items from dictionaries are used in comparative studies), and indeed, real progress in linguistics would be impossible without such use of the work of one's predecessors or colleagues in other linguistic fields. It goes without saying that it is always preferable to have a first-hand knowledge of the language under study, and also the use of other linguists' materials is greatly facilitated if one has done field-work - including interactive gathering - oneself. The importance of this is sadly under-estimated by many an eminent representative of contemporary linguistics.

Linguistic field-work is typically predominantly interactive, the process of gathering going hand in hand with systematic analysis involving (preliminary) categorization and formulation of tentative generalizations. Linguists and phoneticians with little or no interest in field-work should realize that this is indeed the most crucial phase in linguistic research. This is where decisions are made which have consequences for all following steps in the analytical treatment of the data and for the generalizations based on the processed data. It is in the very first phase, when the field-worker sits face to face with his informant, that all kinds of errors may be made, and hopefully may be cleared up. Unnoticed errors made at this stage are the most fatal ones since the process of faulty decision-making at the stage of data gathering is irreversible if there is no independent evidence showing that one went wrong. This may not seem so serious if it is an easily accessible and perhaps well-known language, but with less easily accessible - but typologically or genetically highly interesting - languages there is a heavy burden of responsibility both on the field-worker and on the linguist who squeezes the available second-hand material in order to get answers to questions which the material may not really live up to. New insights may - and often do - emerge from reconsideration of earlier language descriptions, but basically, it is irresponsible to claim that one is making generalizations about a language if in fact one is making generalizations about some particular description of that language.

As emphasized already, the professional linguist doing field-work may go wrong, and he cannot be sure that he always discovers (even grave) errors. Neither can those who draw on his presentation. It is often very difficult to do field-work. There are many reasons for this, also reasons which are beyond the control of even the most well-trained linguist or phonetician. What then if anthropologists or missionaries with little or no linguistic training furnish the data? What can the linguist do if a field-worker with no linguistic background returns from visits to a remote and perhaps nearly inaccessible area, carrying with him unique specimens of a language which is of crucial interest to the linguist in question?

In a sense this situation is rather transparent, though certainly most frustrating: we know that from a linguistic point of view the data gathering was not done professionally (just as the anthropologist may find that linguists do not always ask the right questions in their data gathering!), we have no illusions as to the adequacy of phonetic transcriptions, and our attitude is to try to make the best of the data with all its shortcomings. And then there is one possible virtue of the material: since the field-worker did not start out with any theoretical linguistic bias (or at least not with the same biases as the linguists using his data), there is a genuine chance that his data may be a less deceptive specimen of the language than data gathered with specific purposes in mind by a trained linguist. The linguist's data may be incomparably better in all other respects, and provided that the linguistic field-worker is sensible enough to supplement his strongly monitored data (obtained through steady interaction) with specimens of narrative or fully spontaneous speech one is of course on immensely much safer ground with such data. But the basic problems of how to do proper field-work remain.

Whoever has done field-work himself is at least painfully aware of the shortcomings of his own data, and of the extent to which the interactive approach may distort the data. Probably many researchers have, on some occasion, suddenly come to realize that they were forcing the informant into a narrow scheme of linguistic performance which seemed so artificial to him that he either revolted or resigned, in the latter case suppressing his own intuitive judgment of appropriateness, acceptability or grammaticality (whatever that means) of sentences presented to him. Or - if the issue was a phonological one - the informant was perhaps losing interest in keeping track of what is "same" and "different", or getting confused as to what "same" and "different" means (sameness of sound?,of meaning?).

Interactive data gathering may at times be an extremely successful undertaking. At other times everything may go wrong: the researcher may be persistently on the wrong track on some vital point, he may overlook crucial bits of data, make errors out of sloppiness, and so on. - Many a breakdown in the researcher - informant interaction is caused by differences of expectation. The "naive" informant may expect the researcher to behave like any sensible person who is more interested in the information conveyed by speech than in the medium of communication as such. Some informants can be taught to take the professionalist's attitude toward their own language, others may come to find it funny that they speak the way they do, and take the whole thing as a game. Still others may not ever become "good" informants in any sense of the word.

The worst danger encountered in field-work is the generation of artifacts by informants being confused by leading and too persistent questions. It is, indeed, very unpleasant to discover, in the middle of a field-work session, that the informant is taking more efforts to please the linguist who is pestering him than to perform correctly and consistently in his own language, and that maybe some of the items one carefully recorded are gibberish to the informant when they are read (or played back) to him later. This is painful for the linguist, and it may be detrimental to his work if it happens too often and perhaps passes unnoticed at times. But it may be no less embarrassing to the informant.

The field-worker must be aware all the time that he is dealing with a human being, and he should not only do his best to make his informant "function" properly but certainly also pay due respect to the integrity of this other person. It must be remembered that the field-work situation may be utterly strange and perhaps a little unpleasant to the informant, and that in spite of this strangeness and unpleasantness the informant is

- at least part of the time - doing his best to assist in the documentation of his own language. He may not see the point of the task, but he may be very interested in his own language and conscious about his own linguistic usage. It may appeal to him that his language is important enough for people to come from far away to learn about it. Then again some informants are extremely aware of the fact that in getting involved in field-work they cause something to happen to their language. If the researcher is not careful and responsible enough in his attitude toward the target language the informant may feel compelled to revolt on behalf of his language ("We do not say such things to each other!", "You twist my language!", etc.). Many a syntactician suggesting phrases which are pragmatically silly or totally artificial must have experienced this. There most certainly is an ethical aspect of field-work, and it is not only a question of decent behaviour: it may also be crucial to ultimate success if one really wants genuine language data. In short: good field-work is a demanding task.

III. FIELD-WORK DATA AND PHONETIC ANALYSIS

Where does all this leave phonetics? The practical phonetician has an indisputable role in field-work involving phonetic transcription and preliminary phonological analysis on the basis of accurate impressionistic assessment of the sounds and sound sequences heard. So any field-worker should at least be well trained in practical phonetics. But what about the instrumental phonetic aspect?

It should appear from what has been said above that there is a kind of incompatibility between field-work and instrumental phonetics. Thus the issue may seem a pseudo-issue, the most reasonable answer being that field-work data should not be considered relevant input to instrumental study. To the linguist this may not be too much of a disappointment: he probably did not expect instrumental phonetics to facilitate or corroborate his phonological analysis anyway. The devoted laboratory phonetician may be equally at ease since he has interesting challenges enough and moreover feels at ease only with data generated and processed under ideal laboratory conditions. Why, indeed, bother about instrumental approaches to the analysis of field-work data?

Although there is much to be said in favour of this total segregation of field-work and instrumental phonetics one should not overlook the possibility of getting stimulating suggestions by combining the two approaches. In spite of the misfit between field-work data and research methods in instrumental phonetics it may well be the case that careful phonetic study of whatever data is available may help to sort out what is going on in areas which are not easily accessible to a consistent phonemic analysis, in particular certain aspects of prosodic patterning (such as complex interplay between tonemes and sentence intonation). Trivial types of phonetic specification, such as formant frequencies of vowels, or presence versus absence of vocal fold vibrations in a stop consonant, may also help to put the field-worker's qualitative description on safer ground, because of the experience we have as to the normal relationship between impressionistic and physical or physiological parameters. In particular, it should not be overlooked that an optimum of phonetic specification may be desirable if the linguistic data are used in the context of comparative studies and linguistic reconstruction.

In connection with hypotheses about sound change it is essential to know exactly what is the phonetic reality in the present state of one or several related languages; if such information is not available it seems rather futile to speculate about the phonetics of reconstructed phoneme systems. In the study of tonogenesis, for example, one should know what typically happens in the larynx, and how the fundamental frequency is typically perturbated under specific articulatory conditions, in thoroughly studied languages, and in addition one should know as much as possible about these physiological and physical aspects of the pertinent syllables in the particular languages under consideration. Otherwise it may be that crucial parts of the linguistic argumentation are based on pseudo-phonetic evidence, which is indeed worse than no evidence.

As for phonetic documentation the ideal situation is, of course, if informants can be brought to a place with laboratory facilities (be it a phonetic laboratory proper or a hospital with relevant equipment). To a limited extent, articulatory phenomena can also be studied instrumentally in the field if the phonetician is ingenious enough in selecting portable though adequate equipment. Altogether, however, it is not very common that phoneticians endeavour to create such a happy symbiosis of linguistic field-work and instrumental research as has been done by Peter Ladefoged in his study of African languages. It must be conceded that there are field-work situations which are not very favourable to instrumental research. In working with a tribe of very shy people living in a remote area one faces several difficulties in addition to the remoteness: the informants may be afraid of instruments, there may be severe difficulties in communicating, which adds to the awkwardness of the whole situation, and so on. Even taking a photograph may not be entirely unproblematic. Otherwise, it is self-evident that general visual inspection, corroborated by photography, must be the basic approach used to supplement the auditory impression (and its source of corroboration: the tape). Otherwise, instrumental techniques may come in later, viz. in making acoustic analyses of the tapes brought home from the field sessions. Strangely enough, this may turn out to be the most controversial issue.

To the phonetician such instrumental processing of material which is ill designed for the purpose and perhaps technically rather imperfect may turn out to be a most unrewarding task. However, it should not be overlooked that it may so happen that one can get a first notion of some highly interesting feature whose articulatory nature may be tentatively inferred from the acoustic signal. Limited and perhaps non-quantifiable observations on remote languages may well trigger off a burst of research activity on some hitherto neglected phonetic mechanism. And ideally, it would seem natural to strive for an all-round linguistic documentation of the language under study, comprising also phonetic detail to the extent that it can be at all specified. There are significant phonetic contributions of this kind for some language areas (such as Peter Ladefoged's study of African languages), but it is certainly not the norm in descriptive studies. More often than not the field-work is taken care of by non-phoneticians, and it is a source of enormous frustration to the phonetician who is called upon to assist in getting something out of bad tape recordings, let alone to provide data on physical parameters of sounds on the basis of such tapes, and who has to explain that it is really not worth while. This is so much more unfortunate since it would be desirable if the linguistic field-worker, even if he is not a phonetician, were stimulated to ensure good data for preliminary phonetic observations rather than getting the impression that this is not worth while anyhow. What can be done to make the a posteriori task of extracting phonetic information from tape recordings a less forbidding one?

IV. REQUIREMENTS ON FIELD-WORK RECORDINGS

Let us assume that the field-work is done under the most adverse conditions in some remote village. There is then little that can be done as far as primary observation of speech production is concerned. Perceptual studies, in turn, are largely confined to simple playback experiments serving as an aid to phonemic decisions. We are, then, left with impressionistic transcription and acoustic recording, the latter being always designed to make it possible to listen to the language on later occasions and sometimes also designed to permit some instrumental processing. Needless to say, it is essential that the recordings are both technically good and appropriately edited.

Before going into the subject-matter of good tape recording it must be emphasized that for the purpose of phonetic transcription even the best tape recordings can never wholly substitute for the first-hand impression which the field-worker gets out in the wilderness, or wherever the work takes place. This should be understood by field-workers such as anthropologists who perhaps feel that they have imperfect training in phonetic transcription and therefore tend to scrap their own notes as soon as something more authoritative is available. The field-work situation is unique in that it offers three enormous advantages, viz. (1) that the researcher can hear the speech sounds without any kind of distortion (except for extraneous noise, which, however, is normally much less disturbing in the field-work session than it is afterwards on the tape!), (2) that he can observe the person's mouth and whole appearance during speech, and (3) that he can ask the informant to repeat as much as necessary (the variability occurring in such repetitions being often more illuminating than the exact repetition of each token one gets from playing a tape over and over again). For such (and other) reasons transcriptions made during the field-work sessions should never be by-passed in favour of transcriptions made afterwards on the basis of tapes alone, no matter how careful the latter may be. The field-work notes may be faulty and should be used with much caution, but they often contain information (e.g. on the place of articulation of unreleased final stops) which may be virtually impossible to retrieve from a tape, especially if the researcher does not know exactly what to listen for.

In line with what has been said above it seems strongly advisable always to make dual documentation, viz. by supplementing transcriptions made in the field with (simultaneous or subsequent) tape recordings including the same types of data, which can then be transcribed at leisure afterwards. - In the most recent field-work project in which the present writer was engaged, we were three linguists/phoneticians working together, each taking down all data in phonetic transcription, and each monitoring a separate tape recorder (so as to make sure that at least one recording was reasonably satisfactory for each piece of data). We did not feel any undesirable redundancy in this approach, which was highly interactive in the sense that transcriptions could be compared throughout the sessions, and decisions could be checked over again and perhaps revised.

The ultimate question, then, is what is required in order for tape recordings from such "primitive" field-work to be reasonably satisfactory to the phonetician. Obviously there are requirements both on the quality of recording and the type of data recorded.

Quality of recording is a deplorable chapter in the history of field-work. First of all, much field-work has been done over the decades by researchers who did not realize how bad their recordings were, and how much better they might have been if certain precautions had been made. It is a sad experience when somebody proudly plays samples of a language he or she has recorded far from civilization, with people whose language is near extinction, and it then turns out to be very hard for others to hear what is going on on the tape (the situation may be reminiscent of that in which a phonetician demonstrates synthetic speech to his colleagues, who are utterly incapable of hearing what the machine is supposed to say).

There is an incredible amount of variation in the equipment used, ranging from professional reel recorders equipped with studio microphones down to low-grade cassette recorders equipped with the poorest of microphones. Then the equipment may be fine in terms of maintenance, or it may be on the point of breaking down after the field-worker has dropped it a couple of times when carrying his recorder across rocky paths or slippery ice, or after weeks in incredible humidity and extreme ambient temperatures. It may seem easy to warn the field-worker that he should always use decent equipment, since the technical expenses will probably be a modest entry in the total budget anyway. But it is not always that simple. The present author has been doing field-work with high-quality equipment in some cases but has been forced on other occasions to borrow or hire dubious equipment, sometimes with distinctly unsatisfactory results. And even the best and most well checked equipment may at times go on strike; there may be trivial problems such as running out of good quality tapes and having to buy the poor quality tapes available, running low on batteries way out in the jungle, or finding that contrary to expectations there is no electricity generator running to supply the bulky mains operated tape recorder which one has strived hard to bring along.

Then there is the acoustic environment! Maybe recordings take place inside a small room with disturbing reverberations or tinkling noise from the kitchen or the workshop. Maybe it takes place in open air under potentially ideal circumstances but with occasional or even constant disturbance from domestic animals, from cicadas or sea-gulls, or from the soft but incredibly penetrating hiss or drumming sound that accompanies field-work done in the rainy season. Such constant sound is apt to be ignored during the field-work session, but it has a terrible effect when mixing with speech on the tape. - Finally, in all likelihood there will be a lot of disturbance from curious spectators, or worse: a hopeless mixture of eager voices blending beyond recognition on the tape. This, of course, does not happen all the time (if it did, there would indeed be no point in recording), but it tends to occur at crucial points in the recordings, and it is only when listening to the tape afterwards that one realizes how bad the disturbance was.

Often the field-worker has to make an unpleasant compromise between getting two kinds of noise on the tape: he may be forced to place the microphone quite close to the informant's mouth in order to reduce the level of background noise proportionally, but then again it may cause severe distortion and spurious noise if the microphone is too close. Most microphones used for phonetic recording are <u>not</u> of the type used by pop-singers, who seem to almost <u>swallow</u> the microphone during their singing; on the contrary, the typical recording microphones react spuriously to such treatment. Occasionally a forcefully aspirated stop consonant may sound like a hurricane because of the airstream hitting the microphone grid, unless the field-worker has been ingenious enough to place a sock or the like around the microphone.

Finally, we have the informant himself. Often, what is linguistically most appealing is to work with an elderly informant who still remembers a good deal of the once so rich vocabulary of the language under study, and whose morphosyntax is less corrupt than that of his younger tribesmen (who perhaps have gone to school and learned about the higher prestige of other languages). The elderly informant may, however, have a very bad voice, and he may have very few - if any - teeth left (which would be just wonderful if it were a matter of doing X-ray of tongue positions but is most discouraging in the case of tape recording). In addition he may perform less satisfactory in the sense that he does not speak up clearly and produces too many hesitation sounds.

Unfortunately it is often so that the more relaxed and confident the informant is, the worse are the acoustic conditions for recording. The informant may be absorbed in smoking his pipe or eating candy sugar with his coffee, making terrible noises in between or, worse, on top of the words he is uttering, and the field-worker may hesitate to interfere in an otherwise perfect and genuine linguistic performance in order to deprive his informant of these symbols of ease of mind and self-assurance. Furthermore, the good informant may take such interest in the topic of his talking that he moves his head violently, while the field-worker is trying desperately to keep the microphone somewhere in the vicinity of his mouth.

Problems associated with speaker performance are so numerous that they cannot be listed here. In working with tribal informants who have had little or no contact with Westerners one may have a very hard time persuading the informants to speak up loud, and if they finally do, the result may be an unnatural and exaggerated performance. It goes without saying that rhythm (including absolute durations of vowels and consonants) and intonation, in particular, are vulnerable to such exaggerated speech.

On the other hand, field-work does <u>not</u> just consist in recording what informants choose to say in casual conversation. The existence of various levels of distinctness is a genuine feature of language, and so is the scale of possibilities of marking differences in syntactical structure or in the morphophonemic status of wordforms by various phonetic signals (boundary signals, presence or absence of coarticulation phenomena, etc.). The linguist never arrives at a deeper understanding of the functioning of the language without exploring this scale of possibilities. It is necessary to have words and phrases said at "normal" speed (whatever that means), but it is also necessary to have access to maximally distinct renderings of the wordforms to be sure what is going on. There is nothing wrong in asking the informant to repeat a wordform or phrase over and over again, perhaps slowly and loudly, if only one is sure to get <u>also</u> the more casual rendering occurring in normal conversation. One must pay attention to the possibly differential use of distinctness levels within one discourse unit, the speaker slowing down on focal points and slurring other parts of the utterance in order to ensure optimum communication. To the researcher who is striving with his pencil and eraser to get the transcription right this <u>variability</u> may be rather a nuisance, but it is indeed an extremely important characteristic of languages.

The problem is that the transcription of certain sounds or syllables may remain controversial (perhaps causing vehement disagreement among the field-workers, if there is more than one) even after numerous repetitions, and perhaps after the informant has been shouting right into one's ear. The informant may wonder about the field-worker's mental abilities, and the field-worker himself may also begin to wonder (if he has not wondered all along) about his own poor performance in transcribing (How do other people always arrive at consistent transcriptions so easily, if I have such difficulties?). Under such circumstances it is, of course, an enormous help to have high quality recordings which can be studied afterwards.

A digression on the processing of recordings may be in order here (although it takes up issues which have in part been dealt with already). Even with the best of equipment one should not ever permit oneself to be less attentive during the field-work session on the assumption that it is better and easier to listen carefully at home. Minute quality differences in consonants, in particular, are often heard more distinctly in the presence of the speaker than on the tape, and this may be true especially of consonant clusters. On the other hand, there are several kinds of phenomena which are more easily and safely assessed when listening repeatedly to a tape, and here the supplementary use of instrumental analysis immediately comes in very useful. One such feature is tonal manifestation; another feature is voicing. Pitch is perceived, of course, but by using pitch meter recordings one can get extremely detailed and accurate displays of even microprosodic phenomena; the difficulty is that it takes some phonetic expertise to interpret such raw curves in terms of what the listener actually may perceive. The analyst must know, for example, to what extent perception involves an integration of tonal movements over a certain span of time (physically changing Fo being perceived as steady pitch); to what extent the "tail" of decaying fundamental frequency vibrations tapering off at the end of an utterance (possibly an isolated wordform) should be included in the specification of pitch, and so on. The instrumental analysis may give extremely interesting information of tonal phenomena, and the like, but obviously the information is of a different kind than the impressionistic appreciation of tone, and even though the two descriptions complement each other significantly they should not be mixed up. As for voicing it is very easy to set up a filtering system and to produce an oscillogram enhancing the fundamental frequency oscillations so as to serve as a voicing indicator (or a more sophisticated data processing approach may be used, based on an analog-to-digital conversion of the raw speech signal). In this fashion one can get detailed information as to the actual presence or absence of periodicity in various phases of stop consonants, for example. The problem again is that one gets too much information for the correlation with one's auditory impression (such as weak oscillations which are in fact inaudible) and even spurious information (such as periodicity due to reverberation in the room in which the recording took place, or extraneous periodic noise occurring intermittently), disturbances which are perhaps easily "filtered out" by the ear but not equally easily detected in the oscillogram. The analyst must be able to see this in the oscillogram. Still, the present author has found that such voicing registration is useful not only in its own right, i.e. as a component of an acoustic analysis, but also sometimes as a practical help for the purpose of checking one's auditory impression. Needless to say, the transcription should not just be changed to agree with physical measurements, that would corrupt the whole idea of transcription on an impressionistic basis, but glaring discrepancies may perhaps stimulate the analyst to listen more carefully or to take in more or better material. - In passing, the present author wishes to mention that he has, on the whole, found the two approaches to be in reasonably good agreement. That is, if there is serious doubt (for the trained ear) about the transcription of a particular segment in a particular token as voiced or voiceless the curves typically show some relatively weak oscillations (often only during a minor part of the total segment duration). Thus the problem remains but perhaps its nature is better understood, which may be comforting or even useful to the person struggling with transcription. Above all, however, such sporadic information on the acoustics of speech in one's target language may be enormously suggestive of instrumental investigations which it may be worth while to perform.

V. CHOICE OF DATA TO BE RECORDED

Maybe the most important issue is: out of the mass of talking that occurs during a field-work session, what kinds of data should one record? It is not very practical just to let the tape recorder run all the time, since the subsequent job of sorting out what is going on on the tape is quite frustrating. On the other hand, if one does not ever press the switch until one has prepared a list of items to be recorded, or until something exciting is going on, there is a great risk of missing some of the most valuable instances of spontaneous linguistic performance, which it may be difficult or impossible to recreate afterwards. There is a bargain here. Often the most manageable approach to the recording of narrative texts is to work over the texts first and then make a recording, but the naturalness of the original version may to some extent be lost in the final version that is recorded. If the field-worker is interested in rhythm, and the like, he must calculate the risk that the version entering the tape is spoken in a pedantic, dictation-like fashion - or perhaps with a trace of boredom in the voice - which makes it less than ideal as a specimen of narrative style, although it may be very useful from the point of view of getting vowel and consonant qualities right in the phonetic transcription, and also useful for demonstration of how sentences are designed in this language. If the informant is at all talkative the obvious solution is to try to get both: monitored texts and totally spontaneous speech. As for the latter, one may then have to stop the informant if he starts to tell something interesting and to have him start over again while the tape recorder is running. This requires some tact, and it is successful only with some informants (and especially if the field-worker has cared to establish good enough social relations with his informant); others may lose all inspiration once stopped. And then there are potential informants whose linguistic performance is heavenly music to the field-worker, but who flatly deny to perform in a field-work session because they do not consider themselves bright enough to say something intelligent on tape, or consider their own speech too vulgar. In such a situation it is hard to face the fact that one may have to resign.

Now, to what extent are recordings of free narration and conversation useful to phonetic-phonemic analysis in particular? It may be very difficult to approach a systematic treatment of such data: its immediate value will be rather that there may be a random occurrence of crucially important sounds, sound combinations, sandhi phenomena, and the like, which perhaps escaped the analyst in the first systematic elicitation of information. And then, of course, natural connected speech is enormously suggestive in the context of rhythm and intonation studies. But altogether the immediate value of such data is that it is <u>suggestive</u> rather than being directly amenable to systematic treatment and generalized statements.

It is different with connected speech emerging from the <u>interactive</u> field-work situation as such. The linguist may prepare a text together with the informant and may see to it that the text is interesting not only in terms of its contents, lexical information, morphosyntax, etc., but also - and perhaps above all - in terms of the phonetic-phonemic phenomena it illustrates. In the early stages of field-work on a new language the first practical requirements are (1) to get a certain basic lexicon, and (2) to get a reasonably firm basis for consistent transcription, so for quite a while the phonetic-phonemic illustrativeness of the texts recorded must in fact be given first priority.

Often the analyst will find (or at least believe) that he is better off in his endeavours to get the patterns clear if he concentrates on short sentences and/or single-word utterances in which he has stripped off whatever can be dispensed with without totally losing the structural distinctions to be investigated. In the initial stages of the analysis it is indeed a delightful simplification, and at least for phonetics-phonemics it certainly holds true that one gets quite far with one-word utterances in the analysis of vowel and consonant systems, and with one-clause utterances in dealing with basic stress and tone patterns (whereas focal accent and most of the interesting aspects of sentence intonation require a larger domain of study, of course).

A too reductive approach may create difficulties, however. Let us assume that we are dealing with a tone language. In the initial stage of "discovery" of the tonal system it may seem expedient to concentrate on short single words in order to avoid complications stemming from tonal sandhi or uncertainty about the actual number of consecutive tones occurring on longer stretches of syllables. This is quite all-right, and similarly it may seem useful to take such isolated words forming minimal sets and record them for measurement and description of the acoustic correlates of tone in this language. However, it is strongly advisable to take also such words in a larger context in order to see what happens when the word is not bearer of the whole sentence intonation contour (on top of the toneme, as it were), and when it is not subject to final lengthening, as it probably is in isolation (words in isolation being a special case of utterance final position). One solution, then, is to have a stereotype frame containing material both before and after the test word and designed so that the word has full stress (impressionalistically speaking) and exhibits a minimum of tonal interference with the context.

For the study of vowels and consonants it is always crucial to have the items in several different positions. Often, it is most essential to look at word initial, word internal and word final position (there may also be more specific positions associated with consonant clusters, of course). Here again, one should be somewhat cautious with isolated words. Especially if the data may come to serve as input to instrumental investigation it must be realized that initial and final position posit severe problems of delimitation. If an isolated word starts with a voiceless stop it is totally impossible to measure where it begins, and if the final segment is a vowel it may be difficult to make a principled decision as to where the vowel ends (in some languages the oscillations may die out quite gradually). It is much better if the test word is placed in a neutral, stereotype frame (occurring before or after or on both sides of the word, depending on the items to be examined). The difficulty is that the segment of the frame which is immediately adjacent to the initial or final segment of the test word should be chosen so that there is a minimum of articulatory interaction (or possibly phonological alternation triggered by the frame), and also so that the boundary between frame and word comes out clearly in acoustic curves. This makes the choice of frame somewhat difficult since the differences within a list of test words may be such that there is no one frame which is equally suitable for all test words (a frame ending in a stop consonant is, for example, a bad choice if the test word begins with a stop consonant, whereas it may be an excellent choice if the test word begins with a continuant). Thus one may have to permit small variations in the segmental composition of the frame.

The above may seem a small technicality, but it must be emphasized that much material which was specifically arranged for the purpose of acoustic phonetic analysis turned out eventually to be very difficult or in part impossible to work with because the frame was not well chosen. It takes phonetic experience to foresee these difficulties, and even the experienced phonetician may make the silliest errors in the design of his data, which he then bitterly regrets afterwards.

VI. FINAL REMARKS

Hopefully, the random reflections communicated in this paper may serve two purposes: (1) To urge the field-work linguist to be as careful as possible with his tape recordings, both because of the inherent importance of phonetic documentation, and because his tapes may perhaps - with all their inescapable limitations - help phoneticians to get the scent of some interesting phenomenon, whose acoustic properties can at least be guessed about on the basis of acoustic analysis of selected items. (2) To urge phoneticians to take a wider interest in field-work and to put some of their practical and theoretical expertise to use in this context. But above all: too few young and healthy linguists <u>and</u> phoneticians (at least in these quarters of the world) seem to be willing to realize - or at least to take the consequences of realizing - that gathering information on neglected languages is an obligation for our generation, considering how rapidly languages are disappearing of which no satisfactory record is available. It should be realized also that field-work is an important entry to new insights, that it is indefensible to pursue typological studies if we do not do our best to widen the scope and create a safer basis for typological statements and sweeping generalizations about language universals. There is so much we (still) do not know about languages, and studying <u>Language</u> without bothering about the proliferance of virtually <u>unknown languages</u> (or poorly understood phenomena in otherwise well-known languages) surrounding us on all sides, cannot in the long run lead to real progress.

INSTITUTE OF PHONETICS
JULY 1, 1982 - JUNE 30, 1983

I. PERSONNEL OF THE INSTITUTE

PROFESSOR:
Jørgen Rischel, dr.phil.

ASSOCIATE PROFESSORS:
Børge Frøkjær-Jensen, cand.mag. (seconded to the Audiologopedic Research Group)
Peter Holtse, cand.phil.
Birgit Hutters, cand.mag.
Niels Reinholt Petersen, cand.phil.
Nina Thorsen, lic.phil.
Oluf Thorsen, cand.mag.

TEACHING ASSISTANTS:
Michael Bundgaard, stud.mag.
Peter Molbæk Hansen, cand.mag.
Eva Rosenvold, cand.mag.

ENGINEERS:
Otto Bertelsen, M.Sc.
Preben Dømler, B.Sc.

TECHNICIAN:
Svend-Erik Lystlund

SECRETARY:
Else Parkmann

TEACHERS FROM OTHER INSTITUTES:
Henning Andersen, Ph.D. (Institute of Linguistics)

GUEST RESEARCHERS:
Eli Fischer-Jørgensen, dr.phil.h.c.
Amon Thavisak, M.A., Mahidol University, Bangkok (October 1982 - June 1983)
Chalida Rojanawathanavuthi, M.A., Chulalongkorn University (October 1982 - June 1983)

II. PUBLICATIONS BY STAFF MEMBERS

Françoise Andersen in collaboration with Oluf Thorsen
La prononciation en chansons, Samfundslitteratur, Copenhagen 1982, 70 pp.

Birgit Hutters "Næseflow til vurdering af velofaryngal insufficiens - nogle foreløbige iagttagelser", Nordisk Tidsskrift for Logopædi og Foniatri (Scandinavian Journal of Logopedics and Phoniatrics), 7/2, 1982, pp. 135-146.

Jørgen Rischel "Review article: Quantity in historical phonology", Nordic Journal of Linguistics 5, 1982, pp. 163-171.

Jørgen Rischel Indføring i fonologiske grundbegreber, 185 pp.

Jørgen Rischel "Language policy and language survival in the North Atlantic parts of Denmark", Gegenwärtige Tendenzen der Kontakt-linguistik, ed. P. Nelde, 1983, pp. 203-212.

Nina Thorsen "On the variability in Fo patterning and the function of Fo timing in languages where pitch cues stress", Phonetica 39, 1982, pp. 302-316.

Nina Thorsen "Two issues in the prosody of Standard Danish", in A. Cutler and D.R. Ladd (eds.) Prosody: Models and Measurements, Springer-Verlag, Berlin 1983, pp. 27-38.

Nina Thorsen Review of M. Rossi, A. Di Cristo, D. Hirst, P. Martin, Y. Nishinuma: L'intonation, de l'acoustique ã la sêmantique, Phonetica 40, 1983, pp. 86-87.

Nina Thorsen "Beskrivelsen af dansk ytringsprosodi - en statusrapport" Dansk Audiologopædi 18, 1982, pp. 104-111.

III. LECTURES AND COURSES

1. ELEMENTARY COURSES IN GENERAL PHONETICS AND LINGUISTICS

One semester course (two hours a week) in elementary general phonetics (intended for students of phonetics and linguistics) were given by Niels Reinholt Petersen in the autumn semester 1982.

Courses in general and German phonetics were given by Michael Bundgaard in the autumn semester 1982 (one hour a week), and by Peter Holtse in the spring semester 1983 (two hours a week).

Two courses in general phonetics, for students of Modern Greek and for students of Italian (two hours a week), were given by Eva Rosenvold in the autumn semester of 1982.

Courses in general and Russian phonetics (two hours a week) were given by Peter Molbæk Hansen in the spring semester 1983.

An introductory course in general linguistics (two hours a week) was given in the autumn semester 1982 by Henning Andersen.

2. PRACTICAL EXERCISES IN EAR-TRAINING AND PHONETIC TRANSCRIPTION

Based on tape recordings as well as work with informants, courses of two hours a week were given as follows:

Consonants, in the autumn semester 1982 by Birgit Hutters.

Vowels and prosody, in the spring semester 1983 by Nina Thorsen.

Advanced students, in the autumn semester 1982 by Oluf Thorsen.

3. PHONOLOGY

Courses in phonology (two hours a week) were given by Peter Molbæk Hansen for more advanced students in the autumn semester 1982, and for beginners in the spring semester 1983 by Oluf Thorsen.

4. PHYSIOLOGY AND ACOUSTICS OF SPEECH

A course in the physiology of speech (two hours a week) was given in the autumn semester 1982 by Birgit Hutters.

A course in instrumental physiological phonetics (two hours a week plus individual exercises) was given by Birgit Hutters in the spring semester 1983.

An intensive course in instrumental experimental methods in phonetic research (two hours a week plus individual exercises) was given in the autumn semester 1982 by Peter Holtse, Niels Reinholt Petersen, and Jørgen Rischel.

5. OTHER COURSES

Jørgen Rischel presided over a bi-weekly seminar (graduate level) on Variation and norm in Danish sound history in the autumn semester 1982 (continued from the spring semester 1982).

Jørgen Rischel gave consultations concerning Danish phonology in the autumn semester 1982.

A course in the phonetics and phonology of French (two hours a week) was given by Oluf Thorsen in the autumn semester 1982.

A course in the theory and practice of the language laboratory (one hour a week) was given by Oluf Thorsen in the autumn semester 1982.

A course in statistics (two hours a week) was given by Niels Reinholt Petersen in the spring semester 1983.

Niels Reinholt Petersen presided at a series of weekly seminars for advanced students on topics in experimental phonetics (two hours a week) in the spring semester 1983.

Jørgen Rischel gave experimental phonetic seminars for scholarshipholders from Thailand in the spring semester 1983.

A course in computer science (two hours a week) was given by Peter Holtse in the spring semester 1983.

A course in the phonetics and the phonology of Danish (two hours a week) was given in the spring semester 1983 by Nina Thorsen.

Birgit Hutters presided at a series of weekly seminars for advanced students on topics in speech production (two hours a week) in the spring semester 1983.

Oluf Thorsen gave a 12 hour course in the didactics and methodology of pronunciation teaching at the Institute of applied and mathematical linguistics in the spring semester 1983.

6. SEMINARS

October 6th, 1982 - Jørgen Rischel: Impressions from the linguistic/phonetic institutes in Bangkok, from the XIIIth International Congress of Linguists in Tokyo, and from phonetic field-work in Thailand.

October 27th - Jens Bechsgaard Christensen: A VOT investigation with Danish and American listeners. Preliminary results.

November 10th 1982 - Jørgen Rischel: The Mlabri language in Northern Thailand. Impressions from the workshop on intonation at the XIIIth International Congress of Linguists in Tokyo.

November 12th 1982 - Amon Thavisak: Fieldword on ?Urak Lawoi?.

March 9th 1983 - Professor Klaus Kohler (Kiel): Investigating speech rate.

March 10th 1983 - Professor Bill Barry (Kiel): What is the basic encoding unit in speech production? Thoughts and observations.

April 20th 1983 - Dr. Robert Ladd (Giessen): Toward a model of pitch range and height.

May 11th 1983 - Dr. António Almeida (Cologne): Speakers and their formants.

IV. PARTICIPATION IN CONGRESSES ETC.

Jørgen Rischel participated in the XIIIth International Congress of Linguists, Tokyo, August 29 - September 4 1982 and gave a paper: "The abstractness paradox in Hjelmslevian linguistics".

Jørgen Rischel gave various courses and lectures in experimental phonetics at Chulalongkorn University, Bangkok, in August 1982.

Jørgen Rischel lectured at the University of Hamburg, November 11th 1982: "Lautsysteme und Lautwandel im grönländischen Dialektraum".

Oluf Thorsen visited the Speech Transmission Laboratory at the Royal Institute of Technology, Stockholm, April 25-30 1983, and assisted in the development of text-to-speech synthesis of French.

Peter Molbæk Hansen participated in a meeting about Speech synthesis and recognition in Slet by Århus, October 15th 1982, and gave a paper: "Text-to-speech, phonetics, semantics".

Nina Thorsen participated in Prosody - Normal and Abnormal, an Interdisciplinary Symposium, Zürich, April 6-8 1983, and gave a paper: "Two experiments in the perception of fundamental frequency timing".

Nina Thorsen participated in Nordic Prosody III, Umeå, June 3-4 1983, and gave a paper: "The tonal manifestation of words with assimilated /ə/ in Standard Danish".

V. VISITORS TO THE LABORATORY

Sten Larsen, The Royal Danish School of Educational Studies, worked with stimuli for a dichotic listening experiment.

Niels Bak, Ranum, discussed his work on the acoustics of clarinet playing with the technical staff.

Henrik Nielsen, Telecommunications Research Laboratory, worked on the transfer function of the vocal tract.

Svend Smith, Hellerup, did electroglottographic experiments.

Two groups of high school students were given an introduction to phonetics and methods of speech analysis and synthesis.

VI. INSTRUMENTAL EQUIPMENT OF THE LABORATORY

The following is a list of instruments that have been purchased or built during the period July 1, 1982 - June 30, 1983.

1. INSTRUMENTATION FOR SPEECH ANALYSIS

   1 pressure transducer, Gaeltec, type 8T-2.

2. INSTRUMENTATION FOR SPEECH SYNTHESIS

   1 microprocessor-controlled speech synthesizer, type PD.

3. TAPE RECORDERS

   1 semi-professional recorder, Revox, type B77.

4. LOUDSPEAKERS

   4 headphones, Sennheiser, type HD 430.

5. EQUIPMENT FOR EDP

   1 CTR terminal, Tele Video, type 925.