AN ACOUSTICAL INVESTIGATION OF DANISH INTONATION: PRELIMINARY RESULTS

Nina Thorsen

Abstract: The relationship between linguistic stress and fundamental frequency in short declarative sentences and the intonation contours of short declarative sentences, interrogative sentences, and non-terminal clauses in Advanced Standard Copenhagen Danish are investigated. A tentative model for F_o patterns and -contours in such short simple (i.e. non-compound) sentences is proposed. The model predicts the following (which is to a large extent confirmed by the data):
A stressed syllable and all succeeding unstressed syllables within the same simple sentence constitute a stress group, irrespective of intervening morphological and syntactical boundaries.
Sentence initial unstressed syllables constitute a separate entity and are lower than the first stressed syllable which is always relatively high.
A stressed syllable is lower than the immediately succeeding unstressed syllable (in the same stress group). The unstressed syllables in the stress group describe a falling F_o course. The more unstressed syllables there are, the more likely the last ones are to be lower than the stressed syllable in the succeeding stress group.
The stressed syllables describe a falling contour in declarative sentences and an almost level one in interrogative sentences where the only cue distinguishing the interrogative from the declarative sentence is the intonation. The rest of the interrogative sentences and the non-terminal clauses are situated between these two extremes.

1. Introduction

Danish intonation is an under-developed province in the world of phonetics. Research in this area will, therefore, increase our knowledge of Danish and it may also contribute to our knowledge of intonation as a universal phenomenon. Furthermore, the results

can be applied in the teaching of Danish to foreigners, and - by constituting a basis for comparative studies - in the teaching of foreign languages to Danes. Finally, the results will be useful in rule-synthesis of Danish.

1.1 Delimitation of the subject

An exhaustive description of Danish intonation is, of course, a formidable task, which could keep several people busy for years.

This paper deals with Advanced Standard Copenhagen Danish (see Basbøll 1968) and focusses on two phenomena:

1) the relationship between linguistic stress (in the following referred to simply as: stress), and fundamental frequency (abbreviated: F_o) in short[1] declarative sentences (statements), and

2) the intonation in short

a) declarative sentences,

b) interrogative sentences (questions), and

c) non-terminal clauses.

1.2 Terminology

The terms will be employed in accordance with the following definitions:

To a single syllable (stressed as well as unstressed) is attached a pitch or a pitch movement which is realized as an F_o or an F_o-movement, modified by intrinsic properties of the segment or chain of segments. - A stressed syllable and all succeeding unstressed syllables within the same short simple (i.e. non-compound) sentence constitute a stress group, regardless of morphological and syntactical boundaries. (This is not a definition in the strict sense of the word but, rather, a rationalization of the results, cf. section 3.1.3.) - To a stress group is attached a pitch pattern,

---

1) 'short' meaning: containing three or four stress groups.

realized as a (modified) F_o pattern. One or more stress groups may form a simple sentence which may be led by one or more unstressed syllables. To a simple sentence is attached an intonation with a specific pitch contour, realized as an F_o contour, which is that figure, or shape, or line one can 'draw' through the stressed syllables and upon which the unstressed syllables are superimposed. (This is a rather narrow definition of sentence intonation but it appears to be justified by the results, cf. section 3.2.3.1.) - Presumably, there is a rather close connection between the actual, physical F_o course and the perceived pitch course ('course' is employed indiscriminately about 'movement', 'pattern', and 'contour'), especially as far as the slowly varying pitch/F_o contours are concerned, wherefore the term intonation contour is employed in the following for both pitch- and F_o contour. When no distinction is made between physical (F_o) and auditory (pitch) phenomena, the term 'tonal' is employed. Unstressed syllables before/after the stressed syllables are occasionally, for the sake of brevity, called pretonic and posttonic, respectively.

1.3 The relationship between stress and F_o

Pitch is one (the most important, perhaps) of several auditory cues for the perception of stress in Danish, or at least in Standard Copenhagen Danish. To the foreigner (or synthesizer) who desires to speak intelligible and authentic Danish it is thus important, among other things, to attach the right pitch patterns to the stress groups. Those pitch patterns will therefore have to be mapped - at first by describing the F_o patterns of the stress groups and then, later, by establishing their perceptual correlates.

This will permit us, in a given signal (an F_o-tracing), to separate such variations which are due to the distribution of stressed and unstressed syllables from the intonation contour proper, which will then appear more clearly and simpler, to obvious descriptive and pedagogical advantages.

In the teaching of Danish and in connection with synthesis by rule it works the other way round: Given a certain (simple) sentence with a specific intonation contour the student or the synthesizer (the computer) will, by way of rather simple rules, add the pitch/F_o patterns proper to the stress groups in the sentence.

It seems reasonable, in the light of the above remarks, to include an investigation of intrinsic F_o (F_o differences) pertaining to single segments and of those variations that arise from the interaction between neighbouring segments. The procedure would then be as follows: A given signal is cleansed or corrected for intrinsic and coarticulatory F_o variations, and next for variations due to the distribution of stressed and unstressed syllables, thus producing a clean intonation contour.

And the other way round: to the intonation contour are added first the pitch/F_o patterns pertaining to the stress groups, next the intrinsic and coarticulatory variations at the segment level. The last step is, of course, unnecessary in the class-room situation, F_o variations at the segment level being universal phenomena. However, an investigation of F_o at the segment level is a rather extensive project, and it was decided to omit this aspect from the investigation. (Niels Reinholt Petersen has just completed an investigation of intrinsic F_o in Danish vowels (this issue); his results are turned to account in the treatment of the present material, cf. section 3.1.2.) Besides, it is possible to some extent to dodge the problems introduced by intrinsic and coarticulatory F_o variations, cf. below, section 2.1.

1.4 Advanced Standard Copenhagen Danish

This dialect/sociolect was selected partly because it is the most thoroughly described variety of Danish in recent literature partly because it is that type of Danish which most foreigners will probably need to learn. (Besides, it happens to be the author's own language and one for which subjects are easy to get hold of.)

2. Method

2.1 Material

One corpus is aimed at the stress/F_o relationship, and one at intonation contours in short simple sentences/clauses.

2.1.1 Stress and F_o

Pitch/F_o movements, -patterns, and -contours are probably exceedingly sensitive to the speaker's attitude to what he is saying, which is one reason why nonsense words are of dubious value in an investigation of this kind but they are unavoidable if one wants to introduce changes in one and only one parameter, namely the position of the main stress, in words (otherwise identical) of two and three syllables. There are a few word pairs in Danish, distinguished solely by stress placement, but we do not have five words as below (they are, however, possible words in Danish, - the acute accents mark the stressed syllables):

pípi
pipí
pípipi
pipípi
pipipí

The vowel being i ([i]) in all cases and the consonant being p ([p^{h}]), one can disregard intrinsic and coarticulatory F_o variations, so that observed differences in F_o (F_o movements) in the words must be ascribed to different stress distributions. (Still: the aspiration of p is stronger (of longer duration) before a stressed vowel than before an unstressed vowel which may well influence F_o, at least in the beginning of the following vowel.)

An aspirated consonant was chosen (rather than an unaspirated one or a sonorant) in order to obtain as many clearly discernible line-up points as possible for the subsequent averaging

procedure. It was also thought that the shortest possible voiced stretch within each syllable would be the closest approximation to the perceptual model, i.e. to the perceived pitch pattern, thus, at least to some extent, eliminating the difficult question of what part of the voiced stretch in a syllable is perceptually relevant.

One cannot, beforehand, assume that a given word will produce identical F_o patterns in different positions in a sentence. Therefore, each of the five nonsense words were placed initially, medially, and finally in short declarative sentences, which are as much alike semantically and syntactically as possible:

'_____ giver kórtere stávelser.' (_____ has shorter syllables.)

'Stávelserne i _____ forkórtes.' (The syllables of _____ are shortened.)

'Det giver kórtere stávelser med _____.' (There are shorter syllables in _____.)

2.1.2 Intonation contours

This corpus consists of two principal statements:

'Der går mánge bússer fra Tíflis.' (There are many buses out of Tiflis.)

and

'Der er for mánge tímer i statistík.' (There are too many classes in statistics.)

From these two main themes were made up two times five different types of questions and two times three different types of non-terminal clauses, and in such a way that the underlined sequences in the statements above were unchanged in all 8 sentences/clauses of each set. In this fashion it will be possible, later on, to perform identification tests on the intonation contours of those sequences:

(A1) 'Der går mánge bússer fra Tíflis?' (There are many buses out of Tiflis?)

Interrogative sentence: no word order inversion, no interrogative particle, the intonation alone signals the question. (This is a rather unusual type of question. It is more in the nature of a polite expression of doubt about somebody else's statement, or a way to seek confirmation of an opinion of one's own than it is a proper information seeking question.) It may be termed an intonation question.

(A2) 'Mon der går mánge bússer fra Tíflis?' (Would there be many buses out of Tiflis?)

Interrogative sentence: no word order inversion, interrogative particle 'mon'.

(A3) 'Går der mánge bússer fra Tíflis?' (Are there many buses out of Tiflis?)

Interrogative sentence: word order inversion, no interrogative particle.

(A4) 'Hvórfor går der mánge bússer fra Tíflis?' (Why are there many buses out of Tiflis?)

Interrogative sentence: word order inversion, interrogative particle 'hvórfor'. Note that this sentence contains four stress groups.

(A5) 'Går der mánge bússer fra Tíflis, eller går der kún få?' (Are there many buses out of Tiflis or are there only a few?)

Interrogative clause: word order inversion, no interrogative particle, - followed by a coordinate interrogative clause.

(A6) 'Der går mánge bússer fra Tíflis, så vi kan gódt lade bílen stå.' (There are many buses out of Tiflis, so we may well leave the car.)

Non-terminal main clause, - followed by coordinate main clause with 'så'.

(A7) 'Hvis der går mánge bússer fra Tíflis, kan vi gódt lade bílen stå.' (If there are many buses out of Tiflis, we may well leave the car.)

Non-terminal conditional clause with conjunction 'hvis', - followed by main clause.

(A8) 'Går der mánge bússer fra Tíflis, kan vi gódt lade bílen stå.' (The translation would be identical to A7.)

Non-terminal conditional main clause with word order inversion, - followed by main clause.

(B1) 'Der er for mánge tímer i statistík?' (There are too many classes in statistics?)

(See A1)

(B2) 'Mon der er for mánge tímer i statistík?' (Would there be too many classes in statistics?)

(See A2)

(B3) 'Er der for mánge tímer i statistík?' (Are there too many classes in statistics?)

(See A3)

(B4) 'Hvórfor er der for mánge tímer i statistík?' (Why are there too many classes in statistics?)

(See A4)

(B5) 'Er der for mánge tímer i statistík, eller er der for få?' (Are there too many classes in statistics, or are there too few?)

(See A5)

(B6) 'Der er for mánge tímer i statistík, så vi bliver nødt til at strýge et par stýkker.' (There are too many classes in statistics, so we will have to leave out a couple.)

(See A6)

(B7) 'Hvis der er for mânge tîmer i statistîk, bliver vi nødt til at strŷge et par stŷkker.'

(If there are too many classes in statistics, we will have to leave out a couple.)

(See A7)

(B8) 'Er der for mânge tîmer i statistîk, bliver vi nødt til at strŷge et par stŷkker.'

(The translation would be identical to B7.)

(See A8)

Furthermore, there were two times two statements where the key-words Tîflis and statistîk occur sentence initially and -medially, in order to permit a comparison with the nonsense words pîpi and pipipî in three positions in the sentence:

'Tîflis må lîgge i Kâukasus.' (Tiflis should be situated in the Caucasus.)

'Bûsserne fra Tîflis var forsînkede.' (The buses from Tiflis were behind schedule.)

'Statistîk er i første seméster.' (Statistics is in the first term.)

'Eksâmen i statistîk er frîvillig.' (The examination in statistics is non-compulsory.)

Tîflis and statistîk also occur in isolation as declarative and interrogative utterances, and as members of a reeling off, i.e. non-terminal words. This was thought to make differences in intonation contours easy to ascertain and interpret.

In order to ensure a natural performance, the sentences were incorporated in small dialogues. In the reading list these dialogues were interrupted by the statements containing the nonsense words.

The reading list was recorded six times in all, but only five recordings were included in the subsequent treatment of the material.

2.2 Subjects, recordings, and registration

The subjects were four phoneticians or students of phonetics between 20 and 35 years of age, two males, NRP and SH, and two females, BH and JG, who all speak Advanced Standard Copenhagen Danish.

The recordings took place at the Institute of Phonetics in a quasi-damped room via a condensator microphone (B&K 4131) and an amplifier (B&K 2607) on a professional tape recorder (Lyrec TR47-2) at a speed of 7½ ips on Agfa HiFi Low Noise tape (PE 36).

The interval between subsequent readings of the list ranged between a couple of hours (in rare cases) and several days. No attempt was made to control the rate of speech. Each subject was allowed to read at the speed most comfortable to her/him.

The tapes were processed by an intensity meter (F-J Electronics) and a pitch meter (F-J Electronics), the outputs of which were registered on a mingograph (Elema 803) at a paper speed of 100mm/sec. Fig. 1 is an example of the tracings. Note that there are two editions of the F_o registration: one ordinary and one amplified, which will permit an accuracy of measurement of about 1 Hz for the male voices and 2 Hz for the female voices (corresponding to about ½mm on the scale to the left) in the central portion of their respective F_o ranges.

2.3 Processing of the tracings

The beginning of every stressed vowel was marked, and also the beginnings of all unstressed vowels after unvoiced consonants. The F_o course was then touched up with a solid line, but only in the vowels and the (voiced) sonorants where these could be delimited safely from the surroundings (which was not always the case), assuming that the F_o course in voiced obstruents is irrelevant for the perception of pitch patterns and -contours.

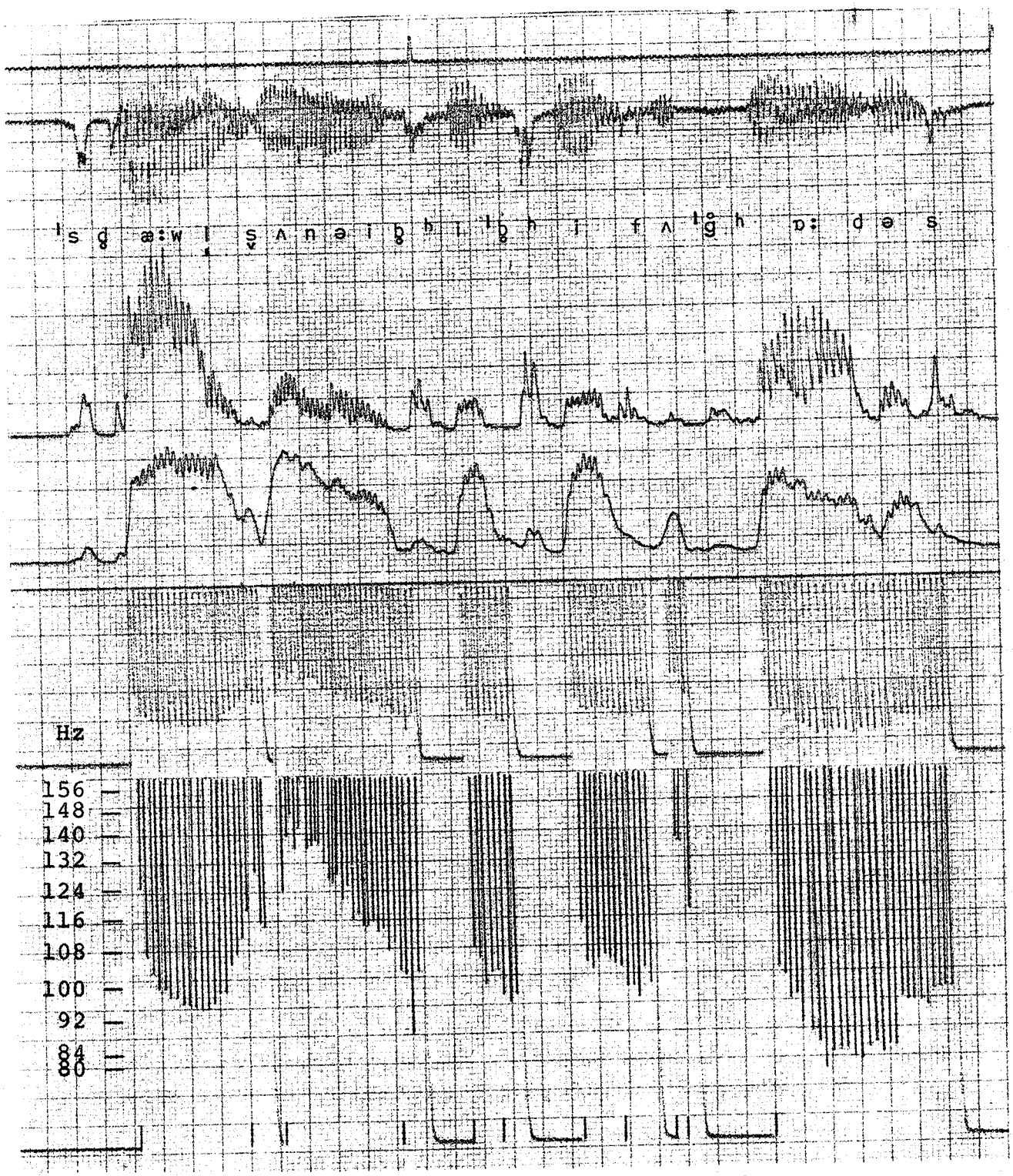


Figure 1

Mingogram of the sentence 'Stávelserne i pipí forkórtes.' (subject NRP). The traces are, from top to bottom: 1) duplex oscillogram, 2) intensity, HP 500Hz, linear display, integration time 5ms, 3) intensity, HiFi, linear display, integration time 10ms, 4) 'ordinary' F_o, 5) amplified F_o (the amplification causes the ink traces to be almost invisible - they have been touched up in crayon).

This assumption will have to be tested in perceptual experiments but, so far, it seems justified in the light of a case like the one below (which is but one out of many):
The word ståvelserne may appear in any of the three shapes depicted in fig. 2:

(a): a rise in the first two syllables, continuing through the [s̬], followed by a fall in the last two syllables,[1]

(b): a rise in the first two syllables, a fall and a rise in the [s̬], and a fall in the last two syllables, which corresponds to the fall in case (a), and

(c): a rise in the first two syllables, 'silence' during the [s], a jump upwards to the fall in the last two syllables, which corresponds to the falls in cases (a) and (b).

It seems reasonable to assume that the 'basic production model' looks somewhat like case (a) and it may be realised as such (in rapid speech) when the glottis remains 'closed' and the vocal cords vibrate during the [s], and, probably, when at the same time the constriction at the alveolar ridge is fairly loose. It may be modified as in case (b) when the vocal chords vibrate around a somewhat more open position, a modification which is not voluntary but due to mechanico-acoustico-aerodynamical factors. Finally, the model may be realised (in fairly slow speech) as in case (c) if the glottis is wide open during the [s], impeding vocal chord vibrations. Unless one's attention is specifically drawn to the detection of voicing in an /s/, the three editions are probably perceived in the same fashion, i.e. there is as much information for the listener in case (c) as in case (a), and the F_o course in the obstruent is thus irrelevant in the sense that it passes unnoticed.

---

1) Danish does not have an opposition between unvoiced and voiced alveolar fricative, but /s/ may be voiced between voiced sounds, especially in rapid speech.

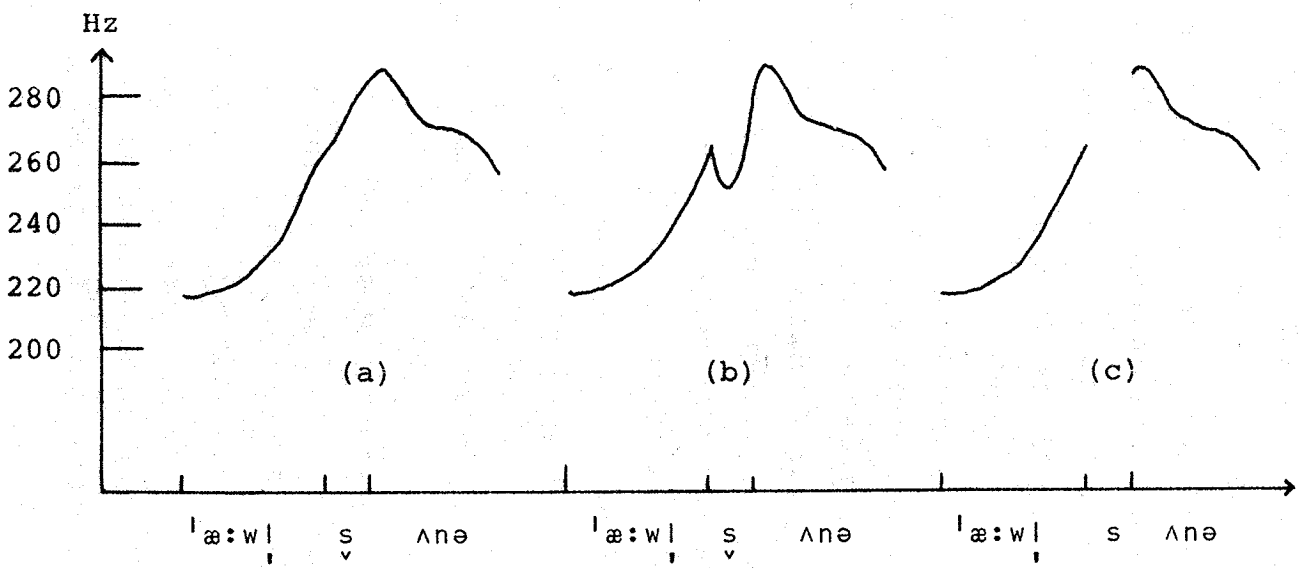


Figure 2

F_o tracings of three recordings of the word 'Stávelserne ..' (subject JG).

In vowels one very often finds, after an aspirated consonant or an (unvoiced) /s/ or /f/ (i.e. consonants produced with an open glottis and with forceful airflow), one or two vibrations which are between 10 and 40 Hz above the succeeding ones. Such vibrations are left out, cf. the line of argument above for /s/.

The pruning of the F_o course is, regrettably, not uncontroversial. The question soon arises: how much of a movement (a rise or a fall) in a vowel can be ascribed to influence from neighbouring consonants? Or, in other words, what does the intended F_o course - the production model - look like before it is superimposed by modifications introduced by mechanico-acoustico-aerodynamical factors, resulting in the rather complicated F_o course observed in the tracings?

One can also pose the question this way: what is perceptually relevant? Ideally, one might trim the traces in accordance with their perceptual correlates (which are not necessarily identical to the production model). But before one can do that, a good deal more must be known about pitch perception in connected speech, i.e. signals which are far more complicated and contain far more information than do e.g. sinusoidal waves.

One example will suffice to illustrate the schism: The first syllable in the word pipí in the sentence 'pipí giver kǿrtere stǻvelser.' produces, with all subjects, a continuous but rather steep fall (cf. figs. 4-7). The beginning of this falling movement may be ascribed to an influence from the preceding aspirated consonant and the end of it may be ascribed to the articulatory transition to the following stop consonant. The intended F_o (movement) may be a level F_o at a frequency corresponding to the central part of the fall, but we cannot be sure (see also section 3.1.1.2). How is the actual movement perceived? As a fall? And in that case: of which extent? As a level pitch corresponding to the beginning? or to the middle? or to the end of the fall?

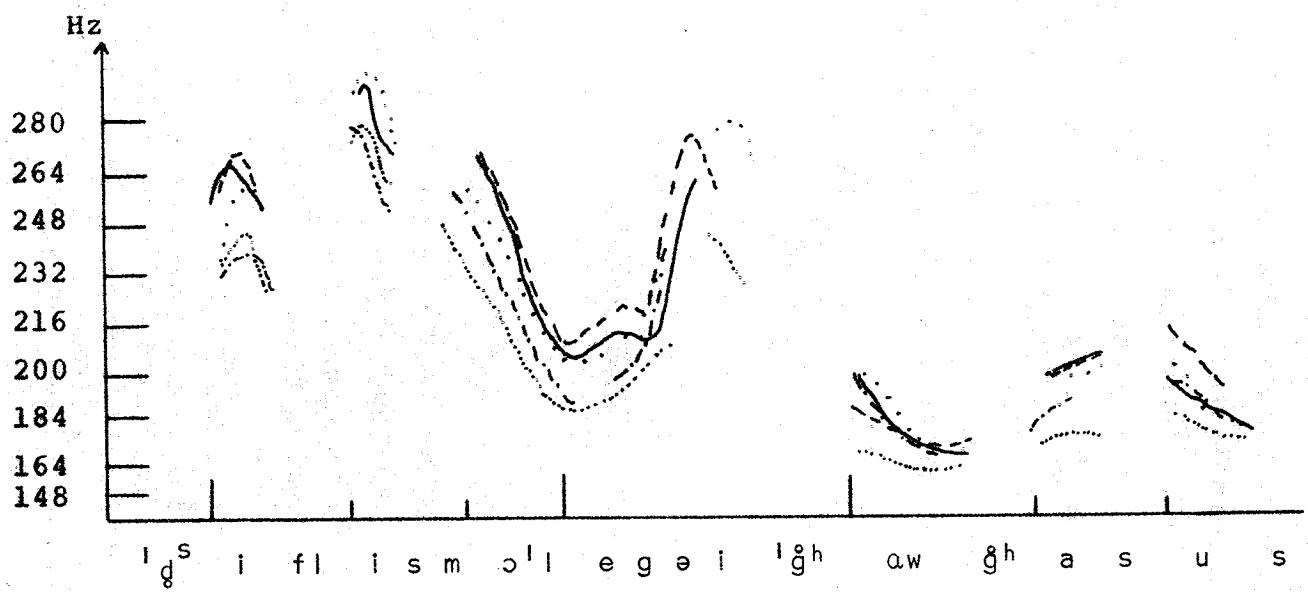


Figure 3

F_o tracings of five recordings of the sentence 'Tíflis må lígge̊ i Káukasus.' (Subject BH).

The various types of lines indicate the recordings thus:

1st recording: ———
2nd - : - - - - - -
3rd - : . . . . .
4th - : ……
5th - : -·-·-·-

Because of this uncertainty the touched-up tracings have been transferred to tracing paper and have been evaluated qualitatively, rather than quantitatively. I.e. no measurements have been performed at this stage. A further justification for this procedure is the fact that a visual display does contain more information than one can immediately convert to numericals. All the line-up points (7-10, depending on the sentence, cf. section 2.1.1) were turned to account in order to obtain the least possible dispersion over the five recordings of each item. In this way differences in duration are, of course, conceiled, and the 'time-axis' is distorted. Fig. 3 is an example of tracings of five recordings of the sentence 'Tíflis må lígge i Káukasus.' by subject BH.

3. Results

3.1 The relationship between stress and F_o

3.1.1 Nonsense words

Figs. 4-7 are averages of five recordings of each of the five words in each of the three carrier sentences for each subject. (The averaging was done by eye and hand.) Note that the time-axis is twice distorted, so to speak: Apart from the distortion

---

Figures 4-7

F_o tracings (averages of five recordings) of five nonsense words in three different carrier sentences by four subjects. The tracings have been moved along the 'time-axis' in order that the line-up points that mark the beginnings of the stressed syllables in the nonsense words merge. The five types of thin lines indicate the words thus:

pípi ——————
pipí . . . . .
pípipi - - - - -
pipípi —·—·— -
pipipí ·····

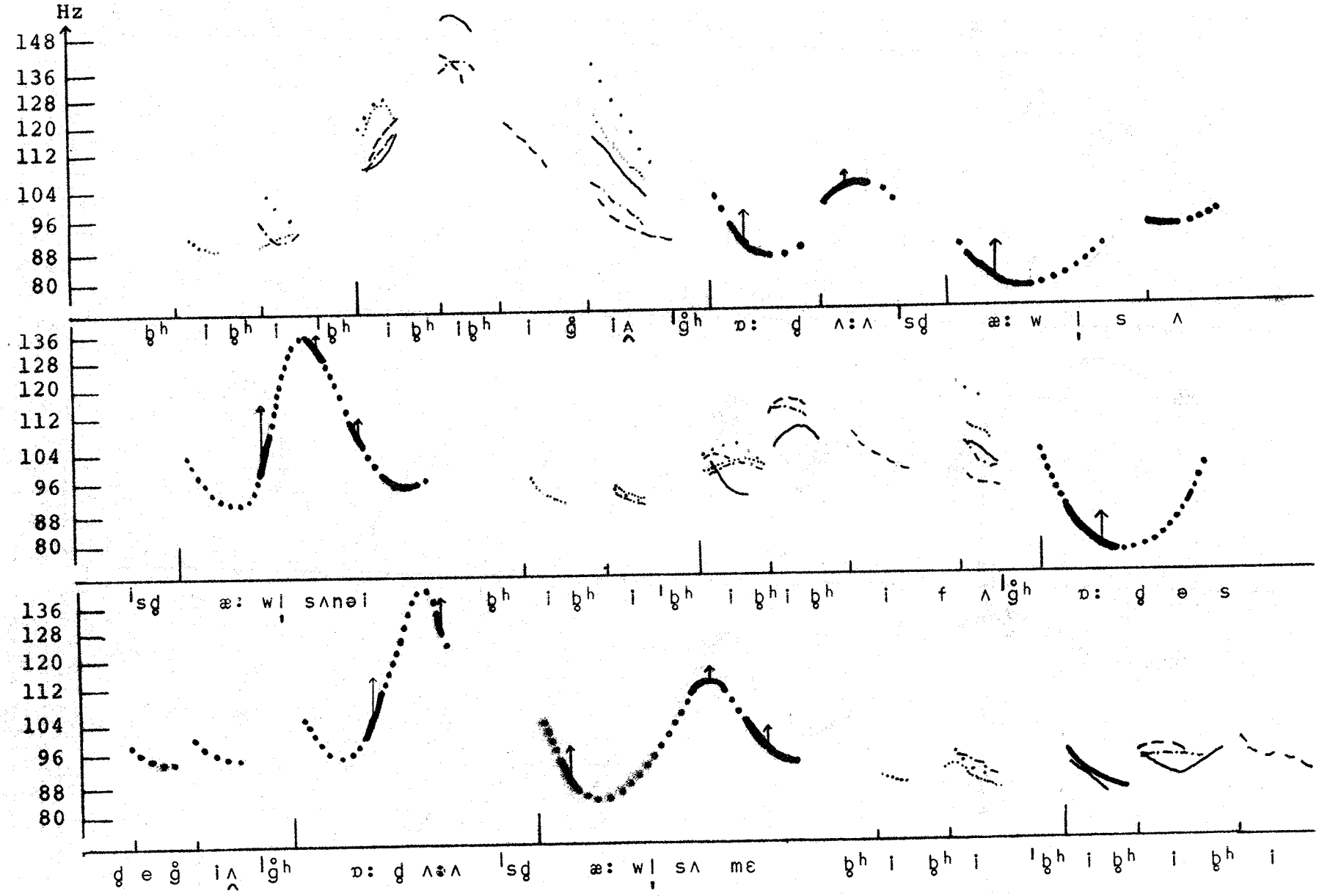

Hz
148
136
128
120
112
104
96
88
80
136
128
120
112
104
96
88
80
136
128
120
112
104
96
88
80

Figure 4 (MRP)

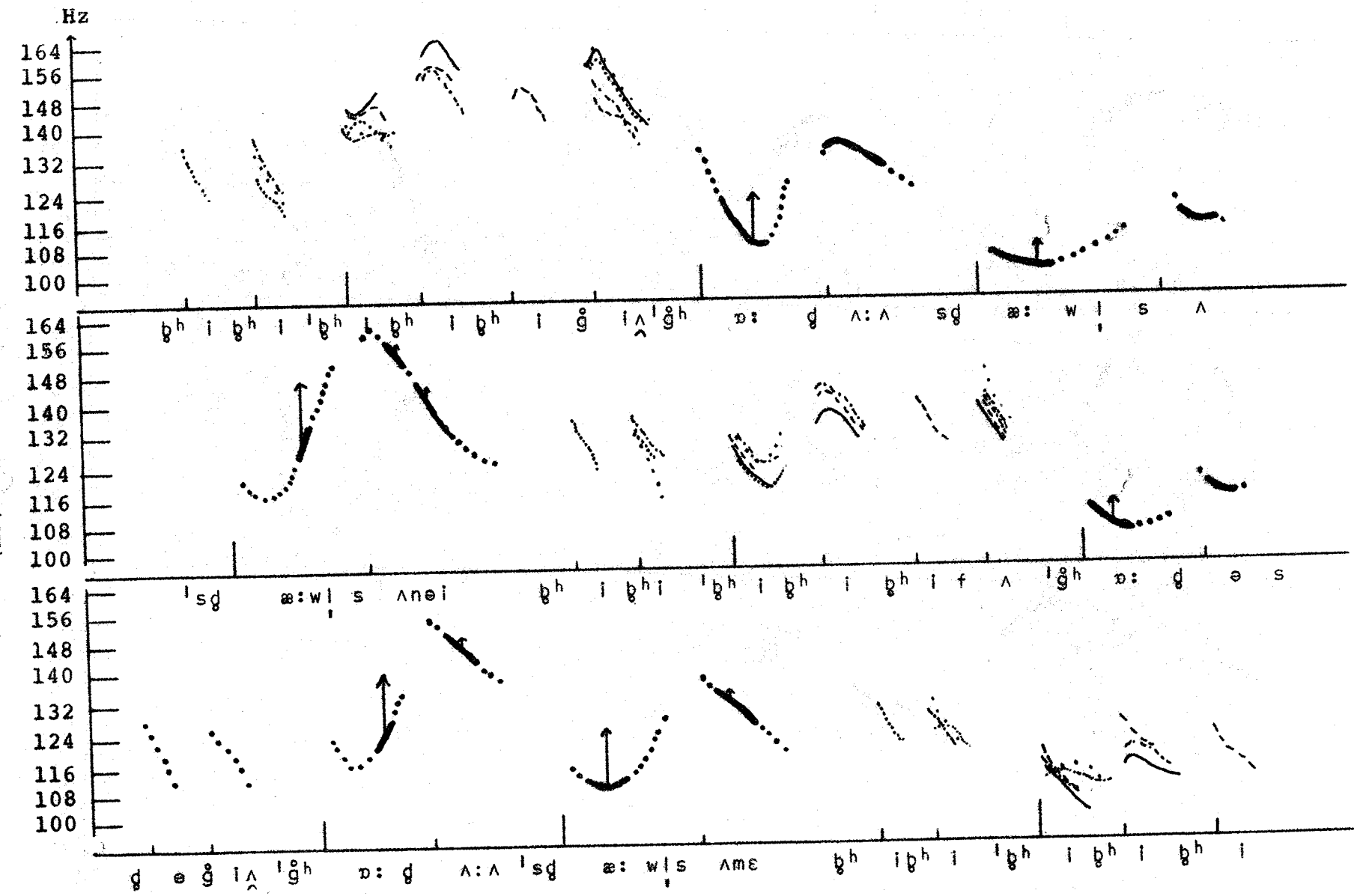

Hz
164
156
148
140
132
124
116
108
100
pʰ i pʰ i ˈpʰ i pʰ i pʰ i g̊ iˈg̊ʰ ɒ: ɡ ʌ:ʌ sɡ æ: w ˌl s ʌ
ˈsɡ æ:wˌl s ʌnəi pʰ i pʰi ˈpʰ i pʰ i pʰ i f ʌ ˈg̊ʰ ɒ: ɡ ə s
ɡ ə g̊ iˈg̊ʰ ɒ: ɡ ʌ:ʌ ˈsɡ æ: wˌl s ʌmɛ pʰ iˈpʰ i ˈpʰ i pʰ i pʰ i

Figure 5 (SH)

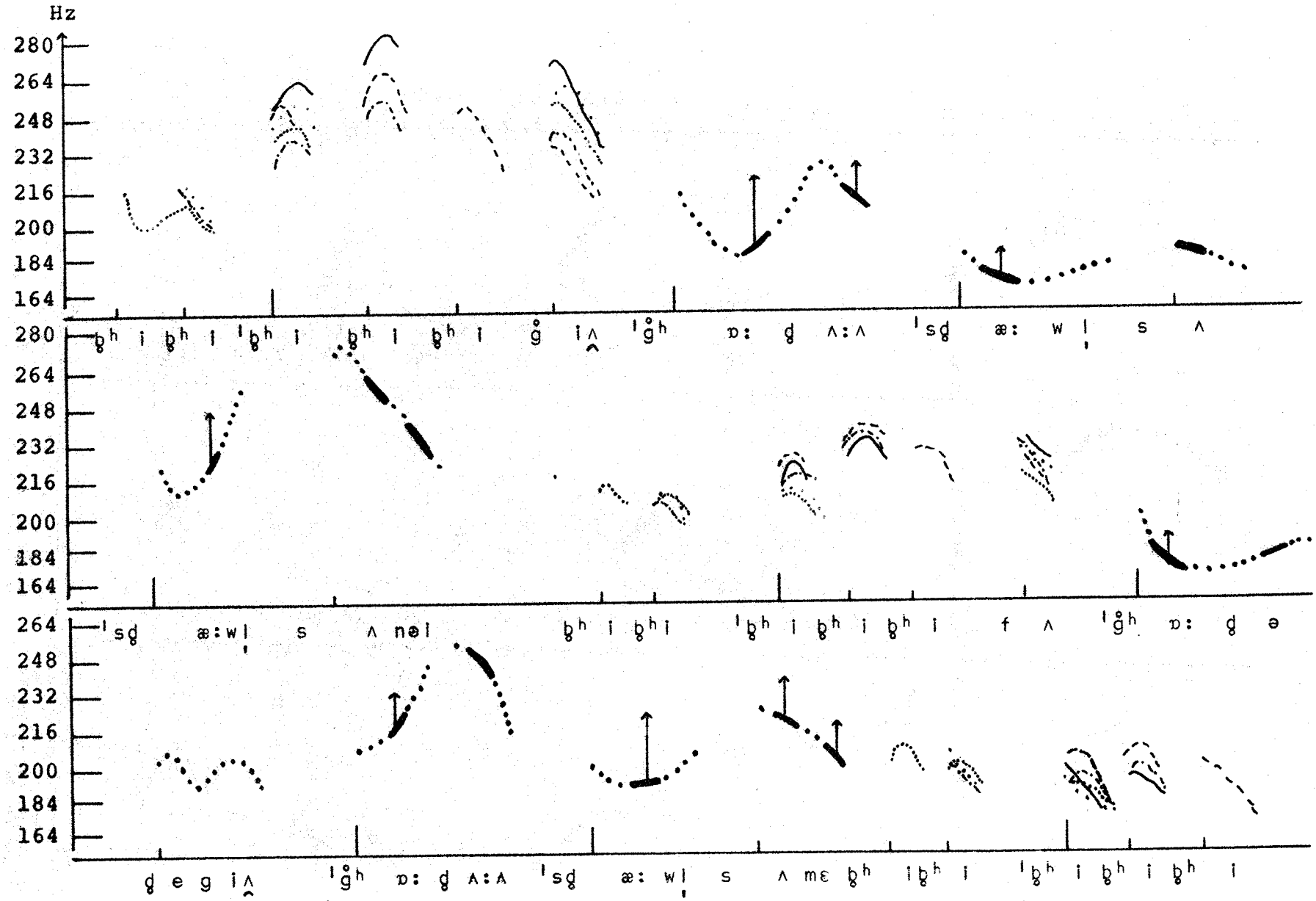

Hz
280
264
248
232
216
200
184
164
ɓʰ i ɓʰ i ˈɓʰ i ɓʰ i ɓʰ i g̊ i^ ˈg̊ʰ ɒ: ɡ ʌ:ʌ ˈsɡ æ: w ˌ s ʌ
ˈsɡ æ:wˌ s ʌ nəi ɓʰ i ɓʰi ˈɓʰ i ɓʰ i ɓʰ i f ʌ ˈg̊ʰ ɒ: ɡ ə s
ɡ e g i^ ˈg̊ʰ ɒ: ɡ ʌ:ʌ ˈsɡ æ: wˌ s ʌ mɛ ɓʰ ˈɓʰ i ˈɓʰ i ɓʰ i ɓʰ i

Figure 6 (BH)

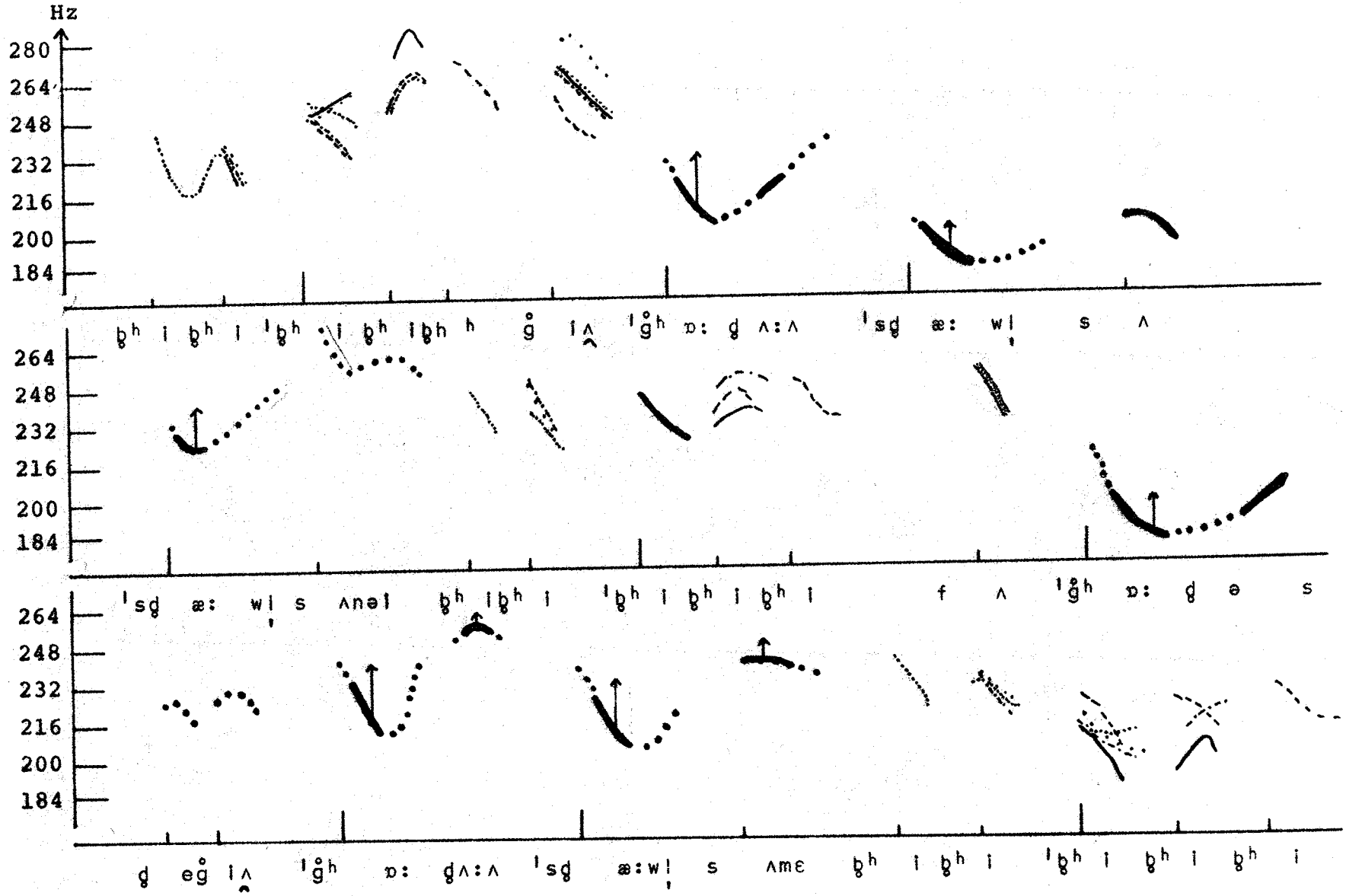

Hz
280
264
248
232
216
200
184
pʰ i pʰ i ˈpʰ i pʰ ˈpʰ h g̊ ˈʌ ˈgʰ ɒ: g ʌ:ʌ ˈsg æ: wˌ s ʌ
264
248
232
216
200
184
ˈsg æ: wˌ s ʌnəl pʰ ˈpʰ i ˈpʰ i pʰ i f ʌ ˈgʰ ɒ: g ə s
264
248
232
216
200
184
g eg̊ ˈʌ ˈgʰ ɒ: gʌ:ʌ ˈsg æ:wˌ s ʌmɛ pʰ i pʰ i ˈpʰ i pʰ i pʰ i

Figure 7 (JG)

introduced by the adjustment to each line-up point through the sentences (cf. above), the five nonsense words were also moved along the axis in such a way that the line-up points marking the beginnings of the stressed vowels merged, which is why there seem to be five pi-syllables in the sentences. - The heavy dotted/full lines represent the averages of 25 recordings (five recordings times five nonsense words) of the carrier sentences. (The explanation for the use of dotted/full lines and for the arrows will be given later, in section 3.1.2.) (There are occasional heavy full lines in the nonsense words (NRP and JG), representing a complete merging of 20 and 25 recordings, respectively.) The dispersion in the carrier sentences over the 25 recordings was no greater than the dispersion within each set of five recordings, and the carrier sentences are thus a good frame of reference for the nonsense words.

3.1.1.1 Stressed followed by unstressed syllables

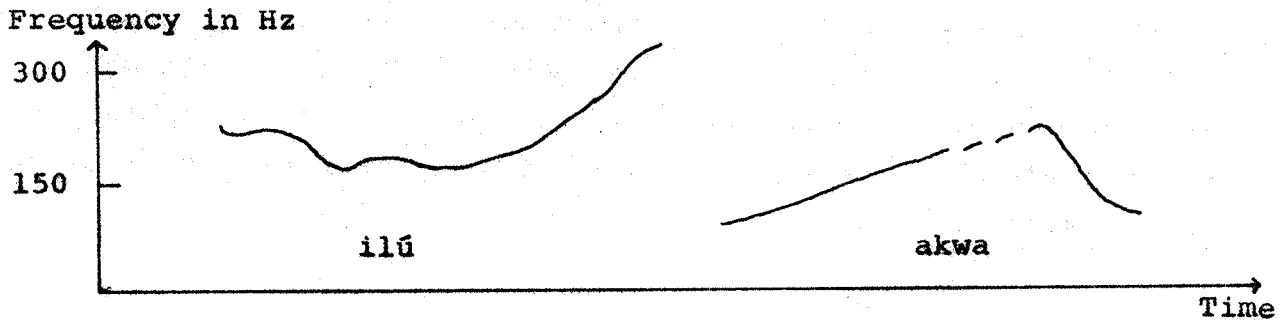
A general tendency shows up in the figures:
the stressed syllable in the nonsense words has a lower F_o than the immediately succeeding unstressed syllable (which may belong to the next word, as in the case of pipí and pipipí in sentence initial and -medial positions).

SH (fig. 5) has not a single exception to this rule. But the remaining subjects do (although not in all recordings of a given word):

JG (fig. 7): two words in two positions, namely medial and -final pípi and pípipi, are exceptions to the rule in 17 (out of 20) recordings, but then in both words the stressed syllable is clearly lower than the immediately preceding unstressed syllable, and thus the stressed syllable is probably marked already as being stressed. Further, in the case of the sentence medial words, and sentence final pípi, one may wonder whether these words are exceptions perceptually. We are dealing with movements of the falling (in the stressed vowel) - rising (in

the unstressed vowel) kind. It is possible that such a movement is perceived as a sequence 'low-high'.

(Bertil Malmberg (1961) has investigated the relation between F_o movement and pitch perception in various words in Yoruba and Ibo and found that he perceived movements like the ones below:



as sequences of 'low-high' (ilú) and 'high-low' (akwa), respectively.)

BH (fig. 6): BH is deviant in sentence initial pipí (2 recordings), sentence initial pipipí (3 recordings), sentence final pípi (4 recordings), and sentence final pipípi (3 recordings). In the first three cases the tracings are open to the same interpretation perceptually as in JG's case. The sentence initial words, however, also hold another possibility, see below about NRP.

NRP (fig. 4): NRP deviates in one (maybe two) case(s), namely sentence initial pipipí giver[1] (5 recordings). Most of the F_o movement in the stressed syllable lies above the succeeding unstressed syllables, and it is presumably also perceived as being higher. In sentence initial pipí giver the same may be true.

---

1) giver is pronounced in one syllable ([g̊i$_{\wedge}$]).

These are instances where also BH deviates, and an explanation may be sought in the fact that
sentence initial unstressed pi-syllables are (considerably) lower than the first stressed syllable, which is always relatively high.
The stressed syllable in the sentence initial words is thus always marked, namely in relation to the preceding unstressed syllable(s). Therefore, it may not be an absolute necessity to continue further up the scale on the succeeding unstressed syllable, especially when, as in these instances, there is only one unstressed syllable before the next stressed syllable in the sentence, a stressed syllable which has to be lower than the first stressed syllable, since
F_o of the stressed syllables is falling from the first to the last stressed syllable in short statements (cf. sections 3.1.2, 3.1.4, and 3.2.2).
In cases like these, where there is only one unstressed syllable between two stressed syllables, the choice of strategy is seemingly free:
NRP has giver lower than pipipí (and maybe also pipí).
BH has giver higher in two recordings, lower in 3 recordings than pipí and pipipí. JG and SH have giver higher than pipipí . and pipí in all instances.
(Incidentally, there are other factors besides pitch which contribute to the correct identification of stress: the unstressed word giver is pronounced (in one syllable) with a short stødless vowel, and p of the stressed syllable in (pi)pipí has a stronger aspiration than p in the preceding unstressed syllable(s).)

3.1.1.2 Stressed preceded by unstressed syllables

The F_o relations between a stressed syllable and the preceding unstressed syllable(s) are less clear than the relations to succeeding unstressed syllable(s).

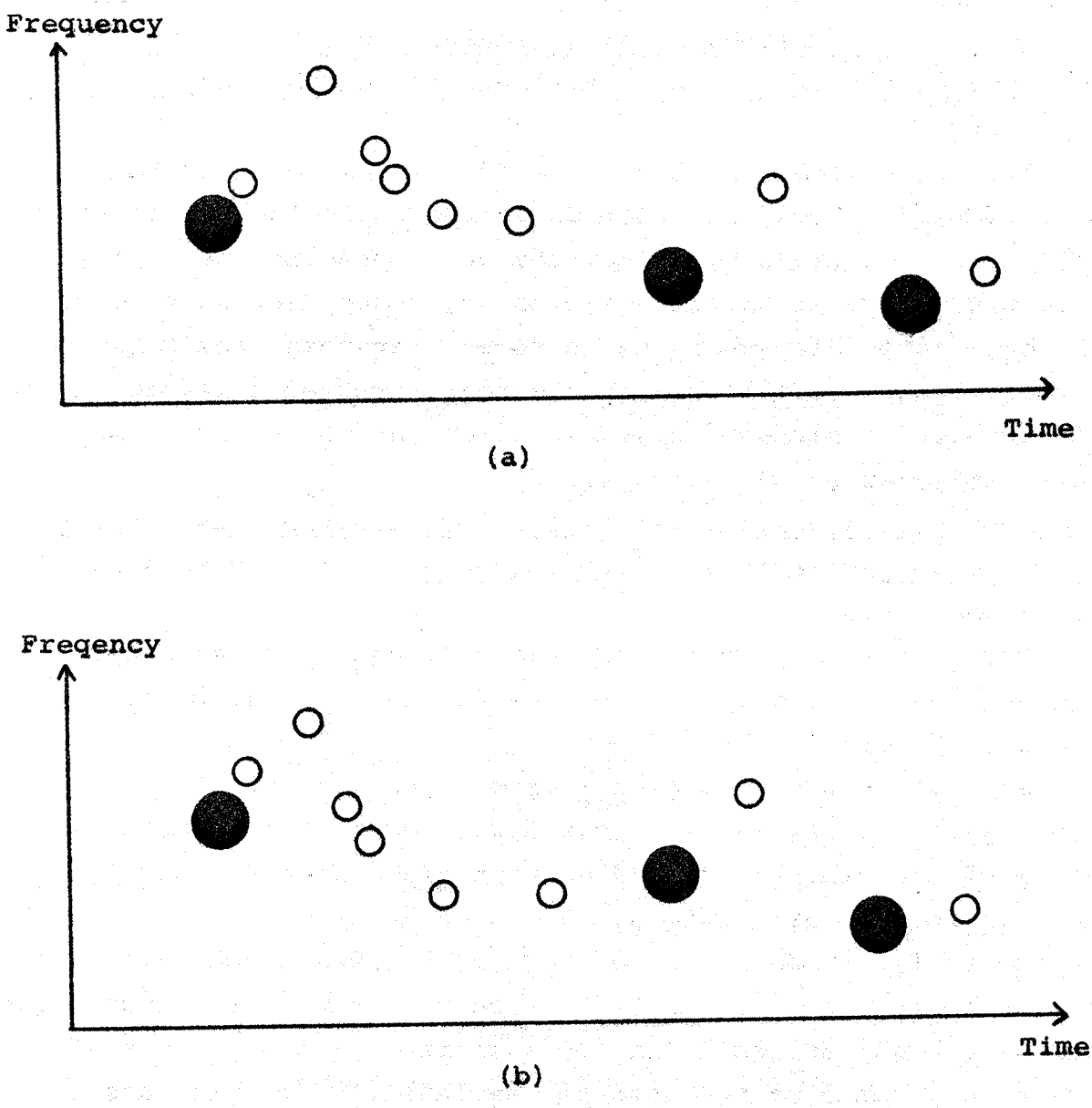


Figure 8

F_o relations between stressed (large filled circles) and unstressed (small empty circles) syllables in the sentence 'Stá-vel-ser-ne i pi-pi-pí for-kór-tes.' (a): subject SH, (b): subjects NRP, BH, and JG.

Sentence initial unstressed pi-syllables are, as mentioned above, always lower than the succeeding stressed syllable.

With SH the remaining stressed syllables are always lower than the preceding (as well as succeeding) unstressed syllables, and his F_o patterns are thus the simplest to describe (cf. below). With NRP, BH, and JG the situation is slightly more complicated: When the number of unstressed syllables between two stressed syllables is three, or less than three,[1] the unstressed syllables are higher than the succeeding stressed syllable (i.e. the pattern that is general for SH). But when there are more than three[1] unstressed syllables between the stressed ones, as in the sentence 'Stå-vel-ser-ne i (pi)-pi-pí forkórtes.', number four and following unstressed syllables are as low as or lower than the succeeding stressed syllable. The difference between SH on one hand and NRP, BH, and JG on the other may be depicted graphically as in fig. 8. (The somewhat peculiar position of '-vel-' will be dealt with in section 3.1.2.)

SH's statements did sound a little more insisting than did the statements of NRP, BH, and JG, and this impression could be attributed to the fact that his stressed syllables are always lower than the unstressed syllables (excluding the sentence initial unstressed syllables). This is corroborated by informal listening to Copenhagen Danish in various situations: the lower the stressed syllables in relation to the unstressed syllables, the more obstinate a speaker sounds.

---

1) The limit may not be three exactly. It may also, it seems, be two or four.

3.1.1.3 F_o movement in individual syllables in the nonsense words

There are variations in F_o movement in individual syllables, but the picture is complicated by the fact that there are inter- and intra subject variations as well as variations seemingly due to position in the sentence and to the difference between stressed and unstressed syllables. However, the following statements hold true, not only of the averages but also of individual recordings.

(i) Sentence initial position

The pretonic syllables are generally falling.

The stressed syllable is rising with NRP, rising-falling with BH, slightly rising or rising-falling with SH, and rising slightly or falling with JG.

The (first) posttonic syllable (of two) is very slightly rising-falling with NRP, and rising-falling with SH, BH, and JG. However, with SH the fall is greater than the rise, with BH the rise and fall are of almost equal extent, and with JG the rise is greater than the fall.

The second of two posttonic syllables as well as the following word giver are falling with all subjects.

The very rough general picture is one of falling unstressed syllables and rising stressed syllables.

(ii) Sentence medial position

The pretonic syllables are falling with all subjects.

The stressed syllables are very slightly rising-falling with NRP, falling or falling-rising with SH, rising-falling with BH, and falling with JG.

The description of the posttonic syllables in medial position is exactly the same as in sentence initial position.

The rough general picture is one of falling unstressed and stressed syllables alike.

(iii) Sentence final position

Stressed and unstressed syllables alike are generally falling, but only slightly so with NRP, and with exceptions in the posttonic syllable for NRP (pípi) and JG (pípi and pipípi).

The very crude (too crude, perhaps) picture is thus one of falling unstressed syllables in all positions in the sentence and of rising stressed syllables in sentence initial position, falling stressed syllables in sentence medial and -final positions.

It should be remembered, however, that much of the falling movement in stressed and unstressed syllables may be due to influence from preceding aspiration and following closure, - and it is hard to say how much of the falling movement is inherent in the production of [i]. The rise in the sentence initial stressed syllables can hardly be ascribed to surrounding sounds, and it is not likely to be an inherent feature of [i]-production either (since this is the only position where stressed [i] is rising, and the rise is not found in all words with all subjects, cf. above). - This rise, when it occurs, may simply be a glide towards the high start of the intonation contour proper.

In order to get a clearer picture of F_o movements in individual syllables one must look at other vowels besides [i] and other surrounding sounds besides [b̥ʰ]. Niels Reinholt Petersen (this issue) investigated eleven long stressed vowels in words of the type [b̥ʰV$_1$ˈb̥ʰV$_1$:b̥ʰV$_1$] in sentence medial position. He found that the F_o movement was almost identical in all vowels, being generally falling-rising (the fall was of greater extent than the rise). Three vowels, [i:, ɑ:, u:] were also investigated between [b̥], [f], and [m] where the same falling-rising movement showed up but with a (not very pronounced) tendency for the extent of the fall to be greater after [b̥ʰ] and [f]. Thus, it does not, at present, seem as if different vowel qualities carry inherently different F_o movements in Danish, and different surrounding consonants do not seem to influence the direction of the movement,

only the extent. The falling-rising movement in Reinholt Petersen's material as opposed to the falling movement in the present material may reflect a difference between long and short vowels, and both movements may reflect a property specific of Danish vowels.

3.1.2 The natural words in the sentences

The F_o relations in the nonsense words reappear in the real, or natural, words in the sentences (stávelser(ne), kórtere, forkórtes), but it is not quite so easy to demonstrate from the tracings. In most cases the F_o course is greatly varying: which portion in a given course is then to be compared to what portion in another course?

If we assume that the tonal patterns in the (pi)pipi words shall recur in the other words (an assumption which can be verified or falsified in perception experiments), the state of affairs is greatly simplified (thereby acting as an indirect support for the assumption). One can compare the courses in the underlined sequences in the six pairs of sentences below, where the nonsense words have been chosen to give the best fit to the natural words of interest in each pair:

(a) 'pípipi giver kórtere stávelser.'
'Stávelserne i pípi forkórtes.'

(b) 'pipípi giver kórtere stávelser.'
'Det giver kórtere stávelser med pípi.'

(c) 'Stávelserne i pípi forkórtes.'
'pípipi giver kórtere stávelser.'

(d) 'Stávelserne i pípipi forkórtes.'
'Det giver kórtere stávelser med pípi.'

(e) 'Det giver kórtere stávelser med pípi.'
'pipípi giver kórtere stávelser.'

(f) 'Det giver kórtere stávelser med pípi.'
'Stávelserne i pípi forkórtes.'

Note that the best fit to kőrtere in (a) and (e) is (pi)pípi rather than pípipi, although kőrtere is a three syllable word, and, likewise, the best fit to stávelser in (b) and (f) is pípi rather than pípipi.

Phonemically, kőrtere and stávelser are /ˈkɔʁdəʁə/ and /ˈsda:vəlsəʁ/, respectively, phonetically they are [ˈg̊ʰɒ:ɡʌ:ʌ] and [ˈsɡæ:wˌsʌ], respectively, i.e. the third syllable in kőrtere is assimilated to the second syllable (but it still counts as a syllable), and the second syllable in stávelser has /ə/ elision, resulting in a syllabic lateral, [ˌ]. It may be that sequences of two syllables, where the second is weakened through /ə/ assimilation or -elision and separated from the preceding syllable by zero or a sonorant consonant, behave tonally as one syllable, i.e. they show what may be termed tonal assimilation. - In fig. 8 was seen that the -vel- ([wˌ]) of stávelserne occupied a position on the frequency axis between the preceding stressed syllable and the succeeding unstressed syllable.[1] - This is somewhat unusual since the first posttonic syllable throughout the rest of the material is the highest one in the stress group, which is a further indication of a peculiarity with such types of unstressed syllables. - More research is needed in order to establish this point with any certainty and to formulate the exact conditions (segmental and otherwise) under which the tonal assimilation takes place.

If the assumption is valid, namely that the same tonal patterns are intended in the nonsense and natural words (in these particular sentences which are of approximately the same semantic and rhythmical structure), one can prune the F_o course in the natural words if, at the same time, one takes intrinsic F_o differences between different vowel qualities into account. This procedure was facilitated by the investigation of Niels Reinholt Petersen who used the same subjects. The relevant figures from his table 1 are repeated below in table 1, expressed as differences in Hz between [i:] and the other 10 vowels of his material.

---

1) The exact location is difficult to establish on the mingograms but it is clearly below the level of the vowel in -ser- ([ʌ]).

Table 1

Differences in Hz (rounded off to the nearest integral) between the (long) vowel [i:] and various other (stressed) vowels (in sentence medial position). The difference is positive if [i:] has the highest frequency of the pair, negative otherwise. For subjects NRP, BH, and JG the figures represent the mean between two values, as observed in two different recordings in Reinholt Petersen's material. [æ:] is identical to [a:] in his notation.

Subject	NRP	SH	BH	JG
Vowel pair				
[i:] - [e:]	3	7	10	9
[i:] - [ɛ:]	6	12	23	16
[i:] - [æ:]	8	16	24	17
[i:] - [ɑ:]	9	15	34	25
[i:] - [y:]	-2	2	0	-8
[i:] - [ø:]	3	7	6	6
[i:] - [œ:]	7	12	17	15
[i:] - [u:]	-2	-1	-3	-16
[i:] - [o:]	2	5	1	5
[i:] - [ɔ:]	4	10	18	15

(NRP did not measure [ɒ:], but presumably it behaves like [ɑ:] (cf. that there is only a negligible difference between [i:] and [y:]).)

The heavy full lines in figs. 4-7 represent those parts of the F_o courses in the natural words which merge with the courses in the nonsense words when the tracings are moved up the frequency axis to an extent indicated by the arrows (which tally reasonably well with Reinholt Petersen's results, cf. table 1). The heavy dotted lines represent the remainder of the F_o courses in the natural words. - In consequence of the remarks above about the [wi] of stávelserne, a fit to this syllable has not been attempted and, likewise, not to the third syllable of kórtere either. - Fittings are missing in some instances where it proved impossible to perform them. - In several instances it seemed unnecessary to move the unstressed syllables at all to obtain a fit.

Reinholt Petersen measured only long stressed sentence medial vowels. Presumably, his results (or at least the tendencies, if not the exact figures) can be generalized to short vowels as well (the differences may be smaller), and to sentence initial and -final positions. However, in order to obtain the best possible fit the intrinsic differences will have to be slightly greater in sentence initial position, and slightly smaller in sentence final position than in medial position, and although this point needs to be verified, it was deemed justifiable to adjust the procedure accordingly.

Reinholt Petersen did not measure the unstressed vowels, but it seems as though the intrinsic F_o differences are smaller and the dispersion of measurements greater in unstressed than in stressed position, cf. Reinholt Petersen's section 3.2.4. The moving of the unstressed syllables ([ʌ, wˌ, ə, i, ɛ]) along the frequency axis was performed simply till the best fit was obtained, which proved always to be a smaller distance than for the stressed vowels nearest to the unstressed ones in quality, which would be: [ʌ∼ɔ-ɑ, wˌ∼u-i, ə∼ɛ-œ, i∼i, ɛ∼ɛ].

It appears that there are several factors involved which may vitiate the procedure with possible errors, and the outcome of the trimming must be regarded with some reservation.

Keeping those reservations in mind, however, it seems nonetheless as if the former statements about F_o relations in the nonsense words can be extended to be valid in short statements in general:

<u>Sentence initial unstressed syllables are always lower than the first stressed syllable, which is always relatively high</u>.
(The difference between JG's sentence initial unstressed syllables and the first stressed syllable is smaller than for NRP, SH, and BH, however.)
<u>A stressed syllable is lower than the immediately succeeding unstressed syllable</u>.

One unstressed syllable between two stressed ones, however, may be lower than the preceding stressed syllable.
The unstressed syllables in the stress group describe a falling F_o course. The more unstressed syllables there are, the more likely the last ones are to be lower than the stressed syllable in the succeeding stress group (the limit seems to be two to four unstressed syllables).
Furthermore, it seems as if
the difference in Hz between a stressed syllable and the immediately succeeding unstressed syllable is greatest in sentence initial and smallest in sentence final position.
F_o of the stressed syllables (the intonation contour) is falling from the first to the last stressed syllable.

3.1.3 The segmental domain of F_o patterns

An interesting problem is that of the structure of the F_o patterns in relation to the chain of segments. Are the F_o patterns related to word boundaries? Or to the boundaries between larger entities within the simple sentence? Or do the stressed syllables alone determine the F_o patterns? (The answer to the last question seems to be affirmative, a fact that has already been turned to account in the preceding pages.) In this case, as in numerous cases mentioned above, a special material should be prepared, aimed specifically at this problem, and the following is more in the nature of speculations than of results proper:

One can search for those portions of the F_o course that vary and look at the extent of the variation in each of the three sentence types where the nonsense words occur initially, medially, and finally:

In sentence initial position in '_____ giver kórtere stávelser.' it is evident that the variation in F_o is not limited to the (pi)pipi words but extends over the following unstressed giver ([g̊iˆ]), in such a fashion that giver is higher the closer it is to the preceding stressed syllable in time, i.e. the fewer

unstressed pi-syllables there are: giver is lower after pípipi than after any of the other nonsense words, and it is highest (not quite so clearly so) after pipí and pipipí. The same tendency appears medially, but only clearly so for NRP, and not at all for JG. - This points in the direction of word boundaries being non-decisive for the F_o patterning. However, the boundary between (pi)pipi and giver and between (pi)pipi and forkórtes is not just a word boundary but also a noun phrase + verb phrase boundary, i.e. a substantial syntactic boundary within the simple sentence. Thus it seems as if F_o patterns and their recurrence are governed by the stressed syllables alone. A further, but rather weak, support for this claim is found in the pretonic pi-syllables in sentence medial and -final positions which show less F_o variation than do the stressed and posttonic pi-syllables. These pretonic pi-syllables may be part of the preceding (invariant) course, thus: 'Stávelserne i (pi)pi ...'/'(Det giver kórtere) stávelser med (pi)pi ...'. - This means that a stressed syllable and all succeeding unstressed syllables within the same simple sentence constitute the unit for the F_o patterning, irrespective of intervening morphological and syntactical boundaries. This unit may be termed a stress group. Sentence initial unstressed syllables constitute a separate entity.

The restriction 'in the same simple sentence' above is inserted because then one can be reasonably sure, on the basis of the present material, that the statement holds true. (But the material is, admittedly, limited and future investigations may, of course, show that the definition of 'stress group' should be changed or modified.)

There are indications, however, that a stress group may also sometimes extend across the boundary between two simple sentences (clauses) in a compound sentence: There are two times four compound sentences in the material, namely A5-8 and B5-8 (cf. section 2.1.2). Only the initial clauses of these sentences have been subjected to processing because they constituted the focal points of interest. But when one looks at the mingograms the following facts appear:

In the majority of a total of 160 recordings (8 clauses times 4 subjects times 5 recordings) the initial unstressed syllables in the terminal clauses are higher than the succeeding

stressed syllable, contrary to the low position of sentence initial unstressed syllables throughout the rest of the material. This is universally true of the terminal clause in sentences A7,8 and B7,8, i.e. main clauses preceded by conditional clauses. Whether the high position of the unstressed syllables is a property of this particular combination of clauses or whether it should be ascribed to an 'across clause boundary stress group effect' is debatable. - In favour of the first solution speak the facts that in 8 instances (A7: 2, A8: 4, B7: 1, and B8: 1) a pause intervenes between the two clauses, thus breaking up what seems to be the normal temporal relations within the stress group. And in 29 out of 40 instances in B7 and B8, where the last syllable of the initial clause is stressed (statistík) the unstressed syllables (two in both cases) in the succeeding terminal clauses are lower than the preceding stressed syllable and thus do not follow the normal F_0 pattern of a stress group. In A7 and A8 there are 6 instances where the first unstressed syllable of the terminal clause is higher than the preceding unstressed syllable (Tíflis), thus breaking the normal stress group pattern. But the second solution would probably be correct for 11 out of 40 instances in B7 and B8 where there is no intervening pause and where the clause final stressed syllable is lower than the clause initial unstressed syllable and for 22 out of 40 instances in A7 and A8 where there is no intervening pause and where the 2 clause initial unstressed syllables together with the clause final unstressed syllable form the smooth falling course characteristic of the unstressed syllables of a stress group.

In A6 and B6 (two coordinate main clauses) both high and low (in relation to the succeeding stressed syllable) clause initial unstressed syllables (three in both cases) are found, and in almost equal numbers. A pause may intervene before high and low unstressed syllables alike. There are only 3 instances where the final syllable(s) of the initial clause and the initial syllables of the terminal clause may be said to constitute a stress group, according to the criteria mentioned above.

In A5 (two coordinate interrogative clauses) subjects unanimously produce the four clause initial syllables on the same low level in relation to the succeeding stressed syllable, and in 13 out of 20 instances without an intervening pause. - The second clause in A5 contains two stressed syllables. The second clause in B5 is extremely short, containing only one stressed syllable, which may be the reason why it is treated differently from A5. SH and BH in 9 out of 10 instances produce what looks like an across clause boundary stress group, but NRP and JG do not.

The conclusion seems to be that the initial unstressed syllables in the terminal clause in a compound sentence may be high or low in relation to the succeeding stressed syllable, whereas sentence initial unstressed syllables always are low. In compounds of conditional and main clauses the initial unstressed syllables of the terminal clause are always high, which

may be a signal for a close link between such clauses. In compounds of coordinate (interrogative) clauses both high and low unstressed syllables are found initially in the terminal clause, but the low ones take the upper hand, which could be an expression of the fact that coordinate clauses are not as closely linked as are compounds of conditional and main clauses. If a stress group is a temporal and tonal entity it may (in about 40% of the instances) extend over the conditional + main clause boundary, whereas this happens only rarely (in 15% of the instances) across the boundary between two coordinate clauses.

Finally, the hypothetical nature of all of the preceding section should be stressed. The properties of the stress group will have to be investigated in much further depth, taking also temporal relations into account.

3.1.4 Conclusion

To recapitulate the results, a model for the F_o course in short statements, which contain only stressed and unstressed syllables (i.e. no secondary stresses), is proposed, cf. fig. 9. It cannot be rejected on the basis of the present material, but its corroboration must await further experiments.

Any sentence initial unstressed syllables are relatively low. The first stressed syllable is relatively high. The first posttonic syllable of several constitutes the F_o maximum in the stress group and the sentence (with possible exceptions if this first posttonic syllable is 'weak' and separated from the preceding stressed syllable by zero or a sonorant consonant, cf. section 3.1.2). A single posttonic syllable may be lower than the preceding stressed syllable, but it is always higher than the succeeding stressed syllable. Three[1] or less posttonic syllables follow the course of the full line. Four or more posttonic syllables may follow the course of either the full line or the broken line, in which last case the first three syllables lie above the heavily dotted line connecting the stressed syllables. This dotted line constitutes the intonation contour proper. The posttonic syllables in the second and, probably, third stress group(s) describe the same figure as those of the

---

1) See footnote to section 3.1.1.2.

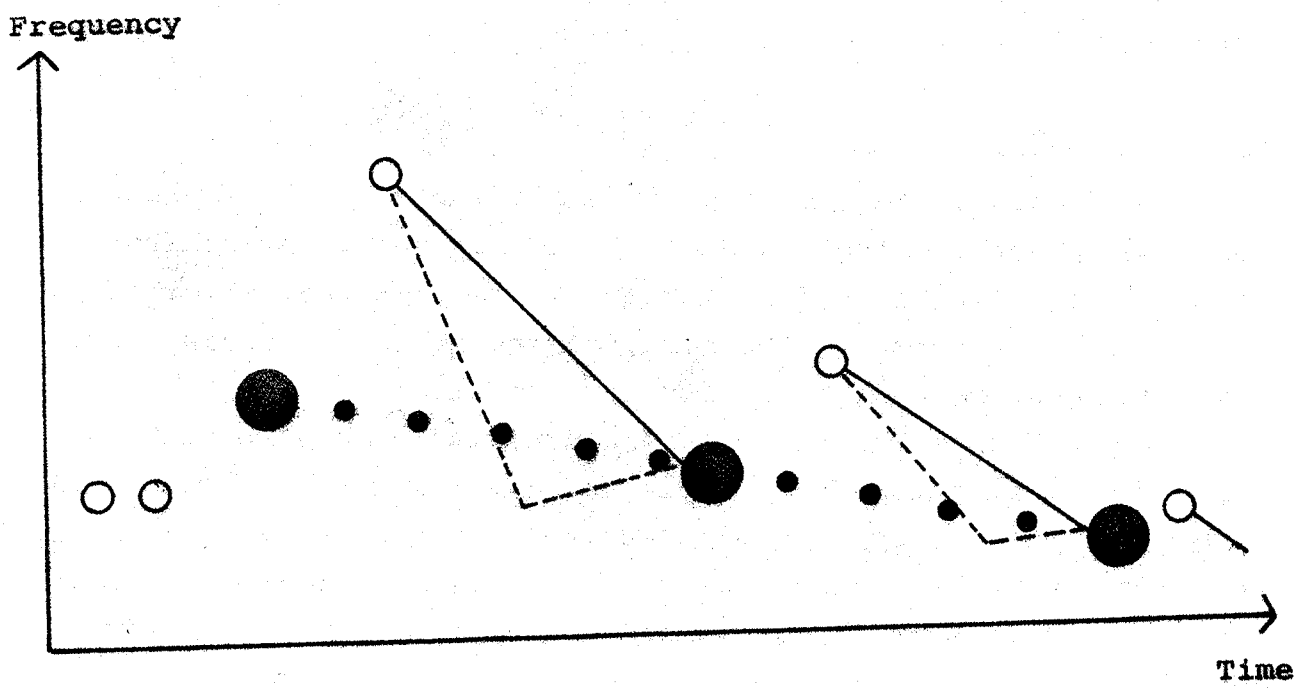


Figure 9

A model for F_0 in short declarative sentences. Large filled and small empty circles indicate stressed and unstressed syllables, respectively. - The small filled circles indicate the intonation contour. For an account of the thin full and broken lines, see the text.

first one, but the distance in Hz between the stressed and the first posttonic syllable decreases progressively. There seems to be the same freedom in choice of strategy, i.e. full or broken line, in the second stress group. (NRP is the only subject who actually follows the broken line in this case.)

The intonation contour describes a smooth fall from beginning (the first stressed syllable) to end (the last stressed syllable).

3.1.5 A postscriptum about normalizing/non-normalizing the tracings

The question whether a normalization (a zero offset adjustment, strictly speaking) of the traces from the outset would not have facilitated the subsequent processing is a very reasonable one. One might, for each subject, move the traces up or down the frequency scale, as the case might be, so that, for instance, F_o minima merged in all five recordings of a given sentence, and maybe also in all sentences of the same type. - However, Reinholt Petersen found in the course of his work that the dispersion does not reduce in this way. He attempted a normalization based on F_o minima in the stressed vowels in the pV'pV:pV words. - It is evident, also in this material, when one looks at individual recordings that such a normalization may well reduce the dispersion in one instance but increase it in another: In the statements the dispersion is generally greatest on the initial stress group, smaller on the medial and smallest on the final stress group, regardless of sentence type and nonsense vs. natural words (fig. 3 is an example). Thus, if one normalizes after the minima in the first stressed syllable, a dispersion is introduced on the last stressed syllable, which is not present in the recordings. If one normalizes after the last stressed syllable, the dispersion on the first stressed syllable is not substantially reduced. A normalization which will reduce the dispersion all over the sentence must therefore be performed for each individual stress

group, as it turns out. In future experiments one might try a normalization along these lines.

3.2 Intonation contours in short statements, questions, and non-terminal clauses

The material was accounted for in section 2.1.2.

3.2.1 Isolated words and reeling-off members

JG is left out in this section. Her recordings were not quite successful in that she did not treat the statements as separate utterances, and her questions sounded overly disbelieving or surprised.

Figs. 10 and 11 are averages of five recordings of the two words in each of the three functions by NRP, SH, and BH.

3.2.1.1 Stress and F_o in the words Tíflis and statistík

(i) Statements

(i, a) Tíflis

With NRP and SH the stressed (first) syllable of Tíflis is lower than the unstressed syllable, i.e. a relationship which is found with pípi in all positions in the sentence with these two subjects (cf. figs. 4 and 5).

With BH the two syllables show no appreciable F_o difference and this fact, together with their position on the frequency axis, makes them resemble her sentence final pípi closely.

Both syllables in Tíflis are falling with all three subjects.

(i, b) statistík

With NRP the two unstressed syllables are clearly lower than the stressed (third) syllable. The same is true only of the second unstressed syllable with SH and BH, but we may still conclude that the relationship between unstressed and stressed syllables resembles that of sentence initial pipipí, and since

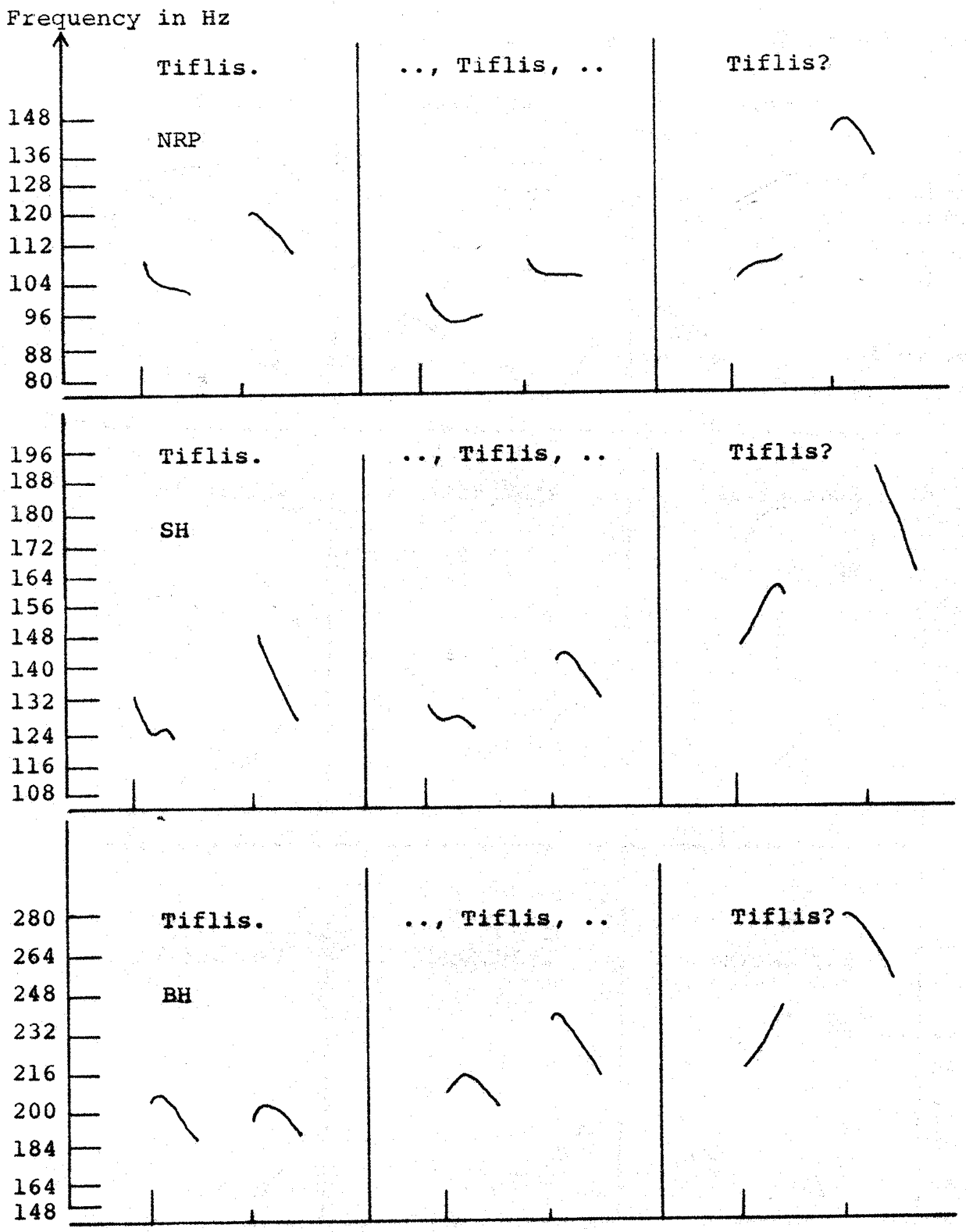


Figure 10

F_0 tracings (averages of five recordings) of the word Tíflis by three subjects, from top to bottom: NRP, SH, and BH. The word occurs 1) in isolation, as a declarative utterance (left column), 2) as a member of a reeling-off (middle column), and 3) in isolation, as an interrogative utterance (right column).

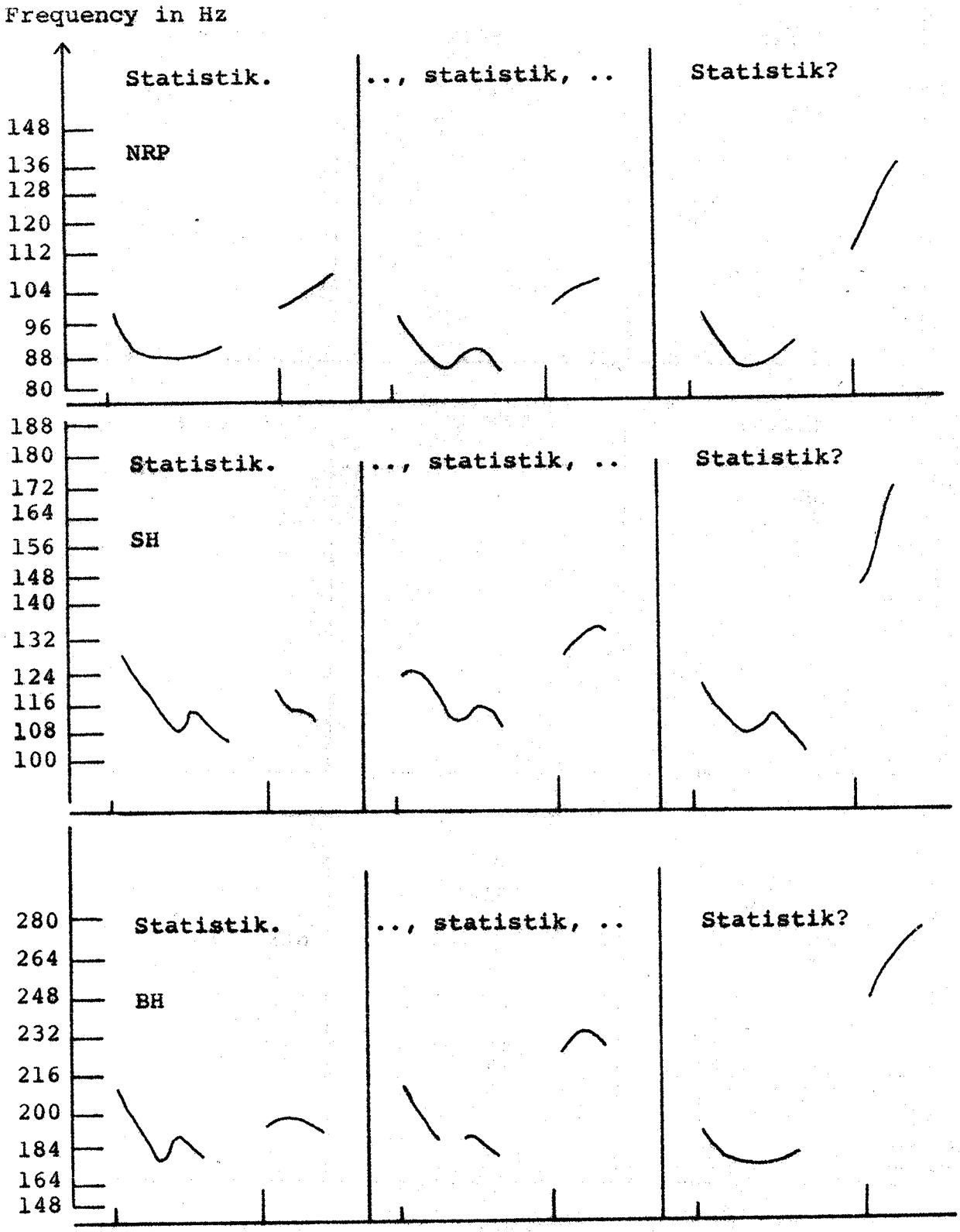


Figure 11

F_o tracings (averages of five recordings) of the word statistík by three subjects, from top to bottom: NRP, SH, and BH. The word occurs 1) in isolation, as a declarative utterance (left column), 2) as a member of a reeling-off (middle column), and 3) in isolation, as an interrogative utterance (right column).

that relationship may or may not be reversed in sentence medial and final positions (cf. section 3.1.1.2), a direct comparison with pipipí in those positions is not possible.

All three syllables are more or less falling with SH and BH. With NRP the stressed syllable is clearly rising (this is true of all five recordings). This corroborates a subjective impression of NRP's statement being slightly 'non-final' in character, i.e. more or less leading on to the following sentence in the dialogue.

(ii) Reeling-off members (i.e. non-terminal words)

(ii, a) Tíflis

With NRP and SH the description would be the same as for the statement, cf. above, except that the two syllables are about 8 Hz lower on the frequency axis with NRP.

With BH there is in this function a clear difference between the stressed and unstressed syllable.

The two syllables are, roughly, falling with all subjects.

(ii, b) statistík

With NRP the description would be the same as for the statement, cf. above.

With SH and BH there is a clear difference between both of the unstressed syllables on one hand and the stressed syllable on the other, which difference is brought about by the stressed syllable being (considerably) higher than in the statement, whereas the position of the unstressed syllables is the same.

Note that the stressed syllable is rising (or rising-falling) with all subjects.

(iii) Questions

(iii, a) Tíflis

There is a large and clear F_o difference between the (low) stressed and (high) unstressed syllable with all subjects.

The stressed syllable is rising, the unstressed one falling with all subjects.

(iii, b) statistík

The F_o difference between unstressed and stressed syllables is considerable. The unstressed syllables resemble those of the statement and reeling-off member with all subjects, i.e. they are low and falling (or falling and slightly rising).

The stressed syllable is steeply rising with all subjects, and the extent of the rise is greater than in sentence initial pipipí.

(iv) Conclusion

We may conclude that the stress/F_o relationships observed in the nonsense words recur in the natural words. Further, the general impression is one of falling unstressed syllables all over and of falling stressed syllables in the declarative, rising in the interrogative utterances. In the non-terminal words Tíflis is falling, statistík is rising.

3.2.1.2 Intonation contours in the words Tíflis and statistík

It may not be proper to speak of intonation contours in isolated words, intonation contour being defined as the figure described by a succession of stressed syllables in a (simple) sentence. However, in utterances containing only one stressed syllable the F_o movement in that syllable, together with its relative position on the frequency axis may be an adequate definition of intonation contour.

BH is very straightforward to describe: The stressed syllable is falling in the declarative, rising-falling in the non-terminal words, and steeply rising in the interrogative word. The position on the frequency axis becomes progressively higher as we go from declarative, through non-terminal, to interrogative words. The distance between unstressed and stressed syllables increases progressively in the same fashion.

SH deviates slightly from the simple pattern of BH in that declarative and non-terminal Tíflis look very much alike, and the stressed syllable of non-terminal statistík is rising.

NRP deviates from BH in that declarative and non-terminal statistík look alike, and both have rising stressed syllables. Furthermore, the stressed syllable of non-terminal Tíflis is lower than that of declarative Tíflis, and the stressed syllable of interrogative Tíflis is not higher than that of declarative Tíflis, but it is rising, and the succeeding unstressed syllable is considerably higher than in the other two words.

Conclusively it can be said that declarative isolated words are characterized by low and falling stressed syllables. The distance to any preceding (lower) and succeeding (higher) unstressed syllables is rather small. Interrogative isolated words are characterized by high and rising stressed syllables. The distance to any preceding (lower) and succeeding (higher) unstressed syllables is rather great. Non-terminal words may resemble the statements or they may be intermediate between declarative and interrogative words.

3.2.2 Declarative sentences

There were six statements in the material (cf. section 2.1.2). Two of these have been accounted for already (cf. section 3.1.2 and fig. 9). The other four, where Tíflis and statistík occur sentence initially and -medially, only confirm the results and the model in fig. 9 with regard to stress/F_o relations as well as to intonation contour: The intonation contour describes a smooth fall from the first stressed to the last stressed syllable in the sentences.

3.2.3 Interrogative sentences

The intonation contours in the questions will be described in relation to the statement contour. Figs. 12-15 are tracings

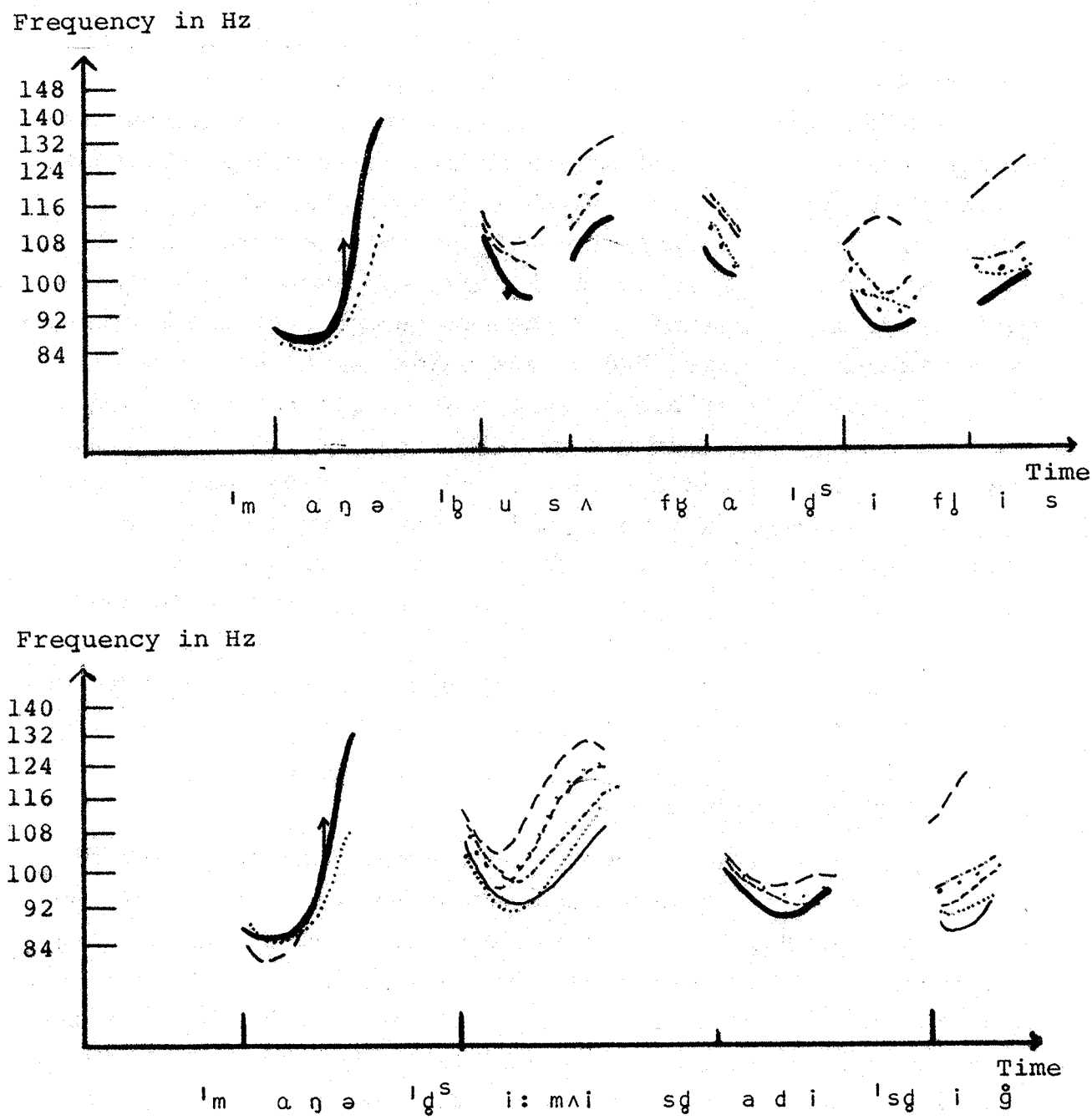
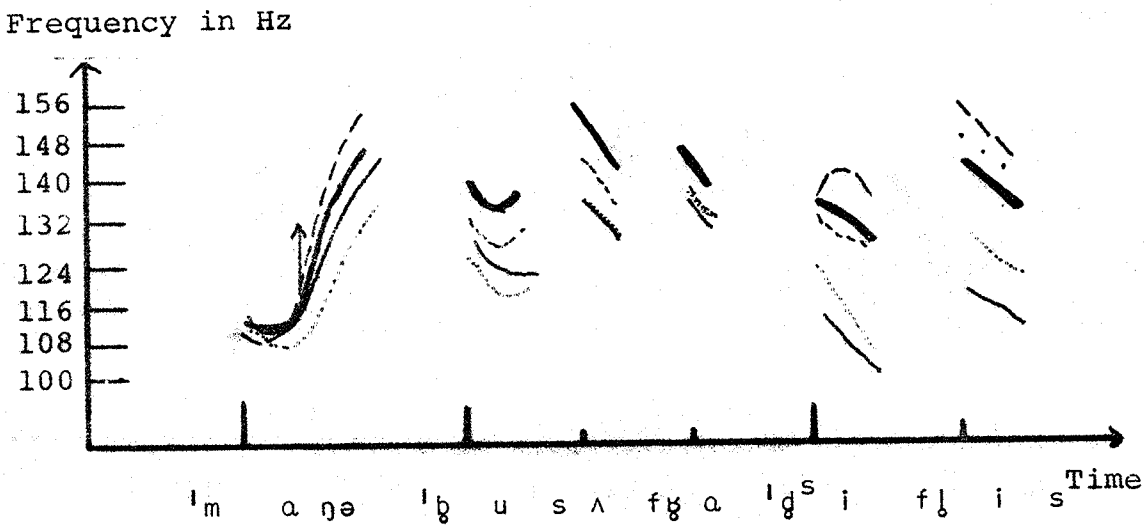


Figure 12

(Subject NRP)

F_0 tracings (averages of five recordings) of the sequence '.. mánge bússer fra Tíflis' (above) and '.. mánge tímer i statistík' (below) in 6 editions: one statement and five different types of questions. See the text for an account of the lines.



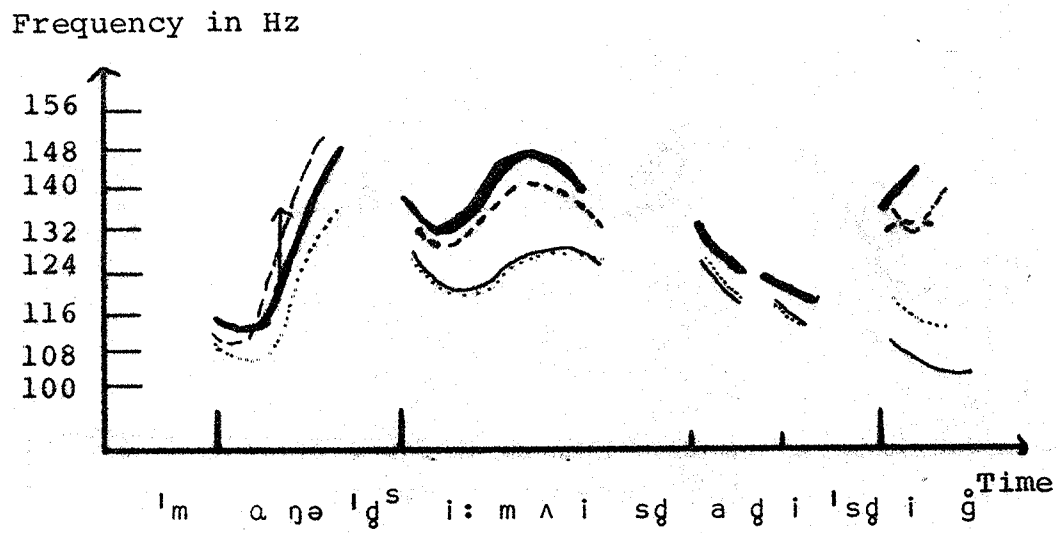
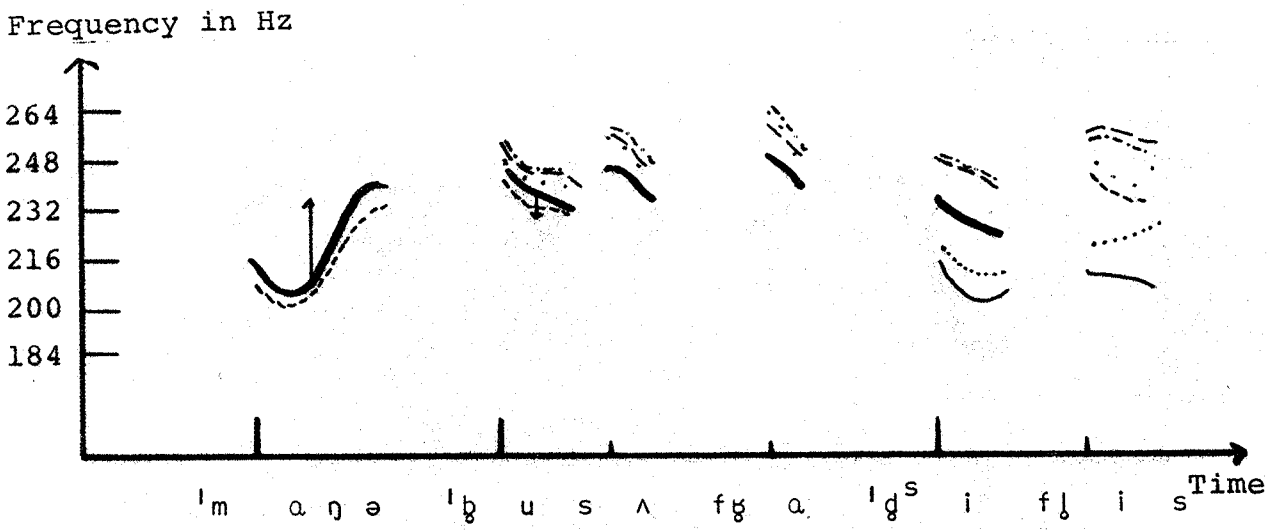


Figure 13

(Subject SH)

See the legend to figure 12.



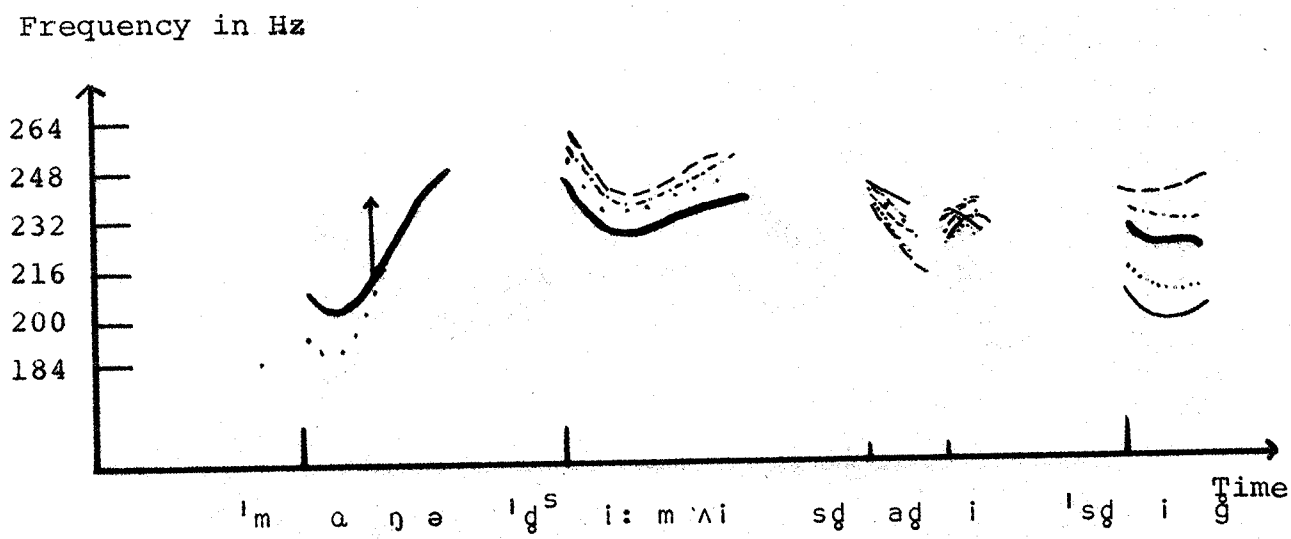
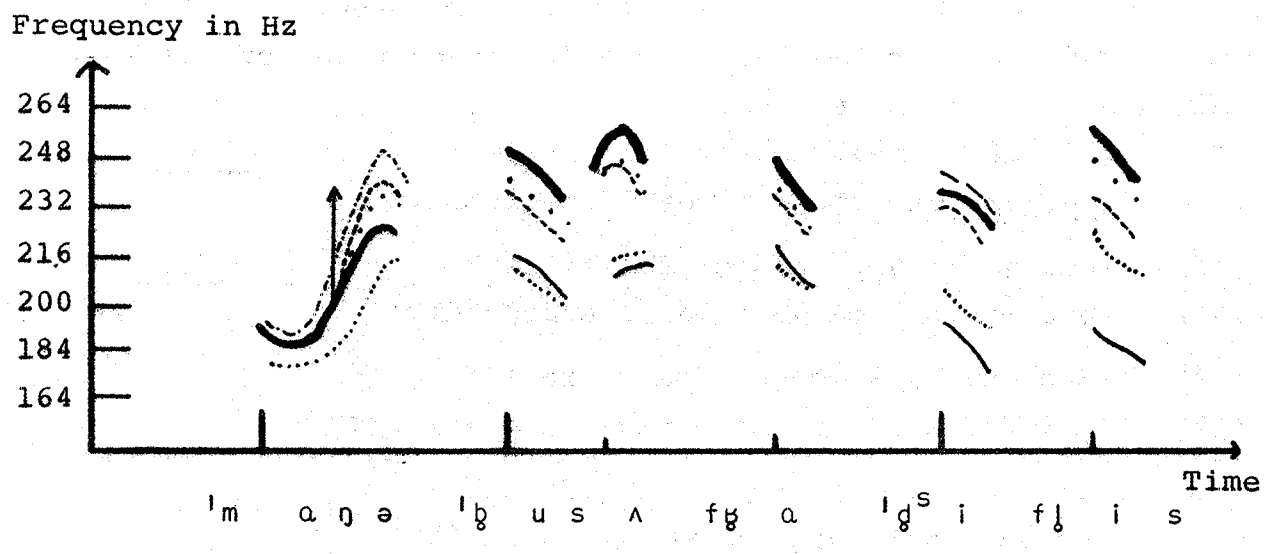


Figure 15

(Subject JG)

See the legend to figure 12.



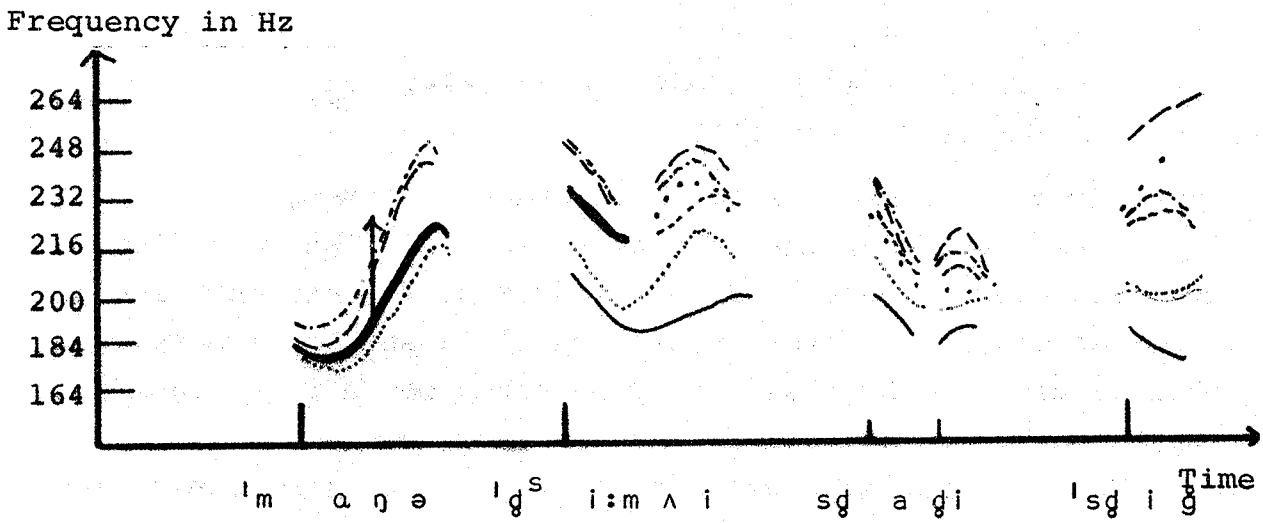


Figure 14

(Subject BH)

See the legend to figure 12.

(averages of five recordings) of the F_o courses in the declarative and five interrogative sentences. (The sentence initial unstressed syllables have been omitted. No pruning of the traces has been carried out.) The different types of lines indicate the sentences thus:

'(Der går) mánge bússer fra Tíflis.' _____
'(Der er for) mánge tímer i statistík.'

(A1) '(Der går) mánge bússer fra Tíflis?' _ _ _ _ _
(B1) '(Der er for) mánge tímer i statistík?'

(A2) '(Mon der går) mánge bússer fra Tíflis?' _ _ _ _ _ _ _ _
(B2) '(Mon der er for) mánge tímer i statistík?'

(A3) '(Går der) mánge bússer fra Tíflis?' . . . . .
(B3) '(Er der for) mánge tímer i statistík?'

(A4) '(Hvórfor går der) mánge bússer fra Tíflis?' .....
(B4) '(Hvórfor er der for) mánge tímer i statistík?'

(A5) '(Går der) mánge bússer fra Tíflis,
(eller går der kún få)?' - - - - - - - - -
(B5) '(Er der for) mánge tímer i statistík,
(eller er der for få)?'

Heavy lines occur when two or more traces are merging.

Note that the averages are averages of five recordings only, and since a certain dispersion exists within each set of five recordings, the overlap between the 6 sets of traces is considerable and the observed differences may not be significant in all cases.

The four subjects are much alike in some (important) respects, but they differ in others:

With NRP and JG the differences in the F_o courses appear clearly only in the second and third stress groups, and with JG they are even rather small in the second stress group. (In the sentences A4 and B4 ('Hvórfor går der mánge .../ Hvórfor er der for mánge ...' (closely dotted line)) mánge is the second stress

group, which is reflected, by the lower position, in NRP's tracings.)

SH and BH have small differences in the first stress group and rather large differences throughout the rest of the sentences. mánge of sentences A,B4 ('Hvôrfor ...') is also lower than in the remaining sentences with these two subjects.

The F_o courses are (remarkably) scattered in the figures, a scattering which must reflect several different interrogative intonation contours (related to the semantic/syntactic function of the sentences, cf. below). (NRP may be an exception, though! His intonation questions (A,B1) are rather clearly separated from the other questions and the statement, which in their turn are relatively low and close together, at least in the last stress group. Perception experiments will reveal whether the contours in these other questions are perceived as being interrogative at all. If not, then it seems as if NRP only utilizes an interrogative contour when absolutely necessary, i.e. when it is the only cue distinguishing the question from a statement. However, the four questions and the statement do not merge, and their order of succession on the frequency axis does follow the same pattern, as far as a pattern can be established, as the courses of the other subjects.)

Keeping the abovementioned differences between NRP and JG on one hand and SH and BH on the other in mind, the F_o courses (i.e. F_o in stressed as well as unstressed syllables) in the sentences can be accounted for in the following manner:

The lowest course is always the statement and the highest is always the intonation question (A,B1), provided they do not merge with some other sentence. The lowest but one is the 'hvôrfor' question (A,B4). - The remaining three questions do not form a clear pattern which recurs with all subjects. With SH they are all rather high and close together and with NRP they are, as mentioned above, all rather low and close together. There is a tendency, however, for the non-terminal clause (A,B5)

to be the highest of the three. This is so for NRP and JG (A5 is very slightly above and almost merges with the intonation question with JG). With SH and BH only B3 is higher than B5 (merging with the intonation question with SH), whereas A3 and A5 merge. There is also a tendency for A,B2 ('Mon' questions) to be the lowest of the three. This is so with NRP (A3 merges with the statement) and BH and also with SH and JG, insofar as it does not merge with one of the other questions.

Looking at the statements and the questions as singular events, their contours may be described thus: The intonation contour, i.e. F_o of the stressed syllables, describes a smooth slanting slope in statements. (The order of magnitude of this fall is about 30 Hz for NRP and SH, about 40 Hz for JG and 50 Hz for BH, which is seen most easily in figs. 4-7.) In intonation questions the contour seems to be almost horizontal. In the remaining types of questions the contours are more or less slanting slopes between these two extremes. The steepest fall is found in the 'hvǿrfor' questions. Cf. fig. 16. With some subjects some of the contours in the interrogative sentences may not be identified as interrogative contours at all.

Thus it seems as if there is a tendency for the intonation contour to be more steeply falling (i.e. more statement-like) the more syntactic information about the interrogative status is contained in the sentence. This presupposes that an interrogative particle overrides word order inversion with respect to 'interrogative information', which is not unreasonable in light of the fact that word order inversion is also found in non-interrogative sentences.

3.2.3.1 Stress and F_o in interrogative sentences

The different slopes of the intonation contours may influence the F_o relations between stressed and unstressed syllables, and any changes in relation to the statements would be most marked in the intonation questions which are treated in the following:

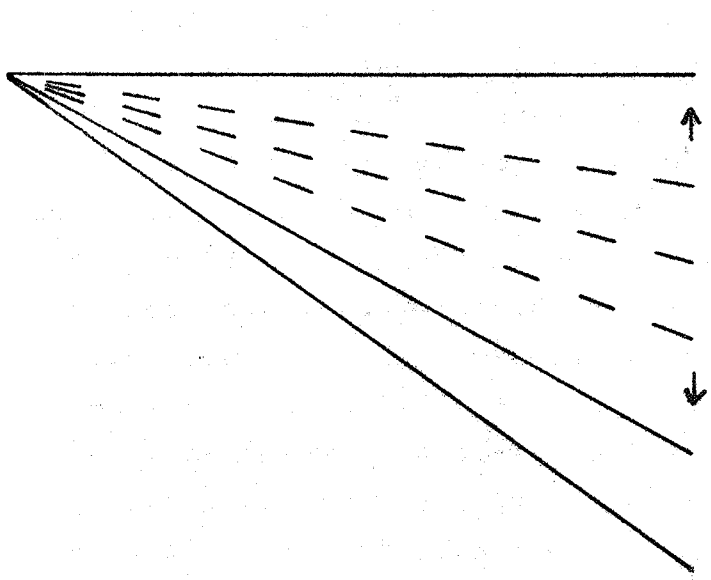


Figure 16

Stylized graph of the intonation contours in declarative sentences and in five different types of interrogative sentences.

The last two stress groups in each of the two intonation questions show a difference from the F_o course in the statements with all four subjects: In the sentence '... bûsser fra Tíflis?' there are only two unstressed syllables between the stressed ones, and they seem with all subjects to be higher than both surrounding stressed syllables if one introduces a correction (even if slight) for intrinsic F_o differences between [ɪ] and [ʊ, ʌ, ɑ]. This is a confirmation of the rule for stress/F_o relations in statements. Furthermore, it does not seem as if the distance in Hz between the stressed and the succeeding unstressed syllable is neither smaller nor greater than in the statement, except, perhaps, with BH. In the sequence '.. tímer i statistík?' there are four unstressed syllables between the stressed ones, and the last two of these (stati-) are (considerably) lower than the surrounding stressed syllables with all subjects. (The difference is rather small with JG.) A likely explanation is the following: The unstressed syllables in a stress group always perform a fall (which may straighten out, or maybe even perform a small rise, towards the end) after the initial rise from the preceding stressed syllable. The slope of this fall may vary (cf. the difference between SH as against NRP, BH, and JG in fig. 8, and the full versus broken line in fig. 9) and if there are enough unstressed syllables in a stress group the last ones will be below the succeeding stressed syllable. The higher this stressed syllable is in relation to the preceding stressed syllable, the 'sooner' the level of the preceding unstressed syllables will be brought below it. This is, in fact, corroborated by the traces in figs. 12-15 (and 17-20).

If the description of the stress/F_o relationship is correct, i.e. the same basic F_o pattern recurs with slight modifications which are predictable from the intonation contour, then the definition of intonation contour as the contour described by the stressed syllables alone seems justified. - This does not mean, of course, that the F_o course of the unstressed syllables is irrelevant for the perception of a given contour, but it is, strictly speaking, redundant.

3.2.4 Non-terminal clauses

The intonation contours in the non-terminal clauses will also be described in relation to the statements, cf. figs. 17-20. The different types of lines indicate the sentences thus:

'(Der går) mánge bússer fra Tíflis.' _____
'(Der er for) mánge tímer i statistík.'

(A6) '(Der går) mánge bússer fra Tíflis,
(så vi kan gódt lade bílen stå).' - - - - - - - -
(B6) '(Der er for) mánge tímer i statistík,
(så vi bliver nødt til at strýge et
par stýkker).'

(A7) '(Hvis der går) mánge bússer fra Tíflis,
(kan vi gódt lade bílen stå).' . . . . . .
(B7) '(Hvis der er for) mánge tímer i sta-
tistík, (bliver vi nødt til at strýge
et par stýkker).'

(A8) '(Går der) mánge bússer fra Tíflis,
(kan vi gódt lade bílen stå).' _ . _ . _ . _
(B8) '(Er der for) mánge tímer i statistík,
(bliver vi nødt til at strýge et par
stýkker).'

The reservation mentioned in section 3.2.3 about the averages must be kept in mind.

The four F_o courses are dispersed to a greater or lesser degree with each subject in much the same fashion as in interrogative sentences, except that the dispersion in the first stress group is small or non-existent with all subjects. The mutual relations between the four courses are almost exactly the same for all subjects (insofar as traces do not merge completely): The lowest F_o course is the statement, next (from bottom to top) sentences no. 6, 7, and 8 (cf. above).

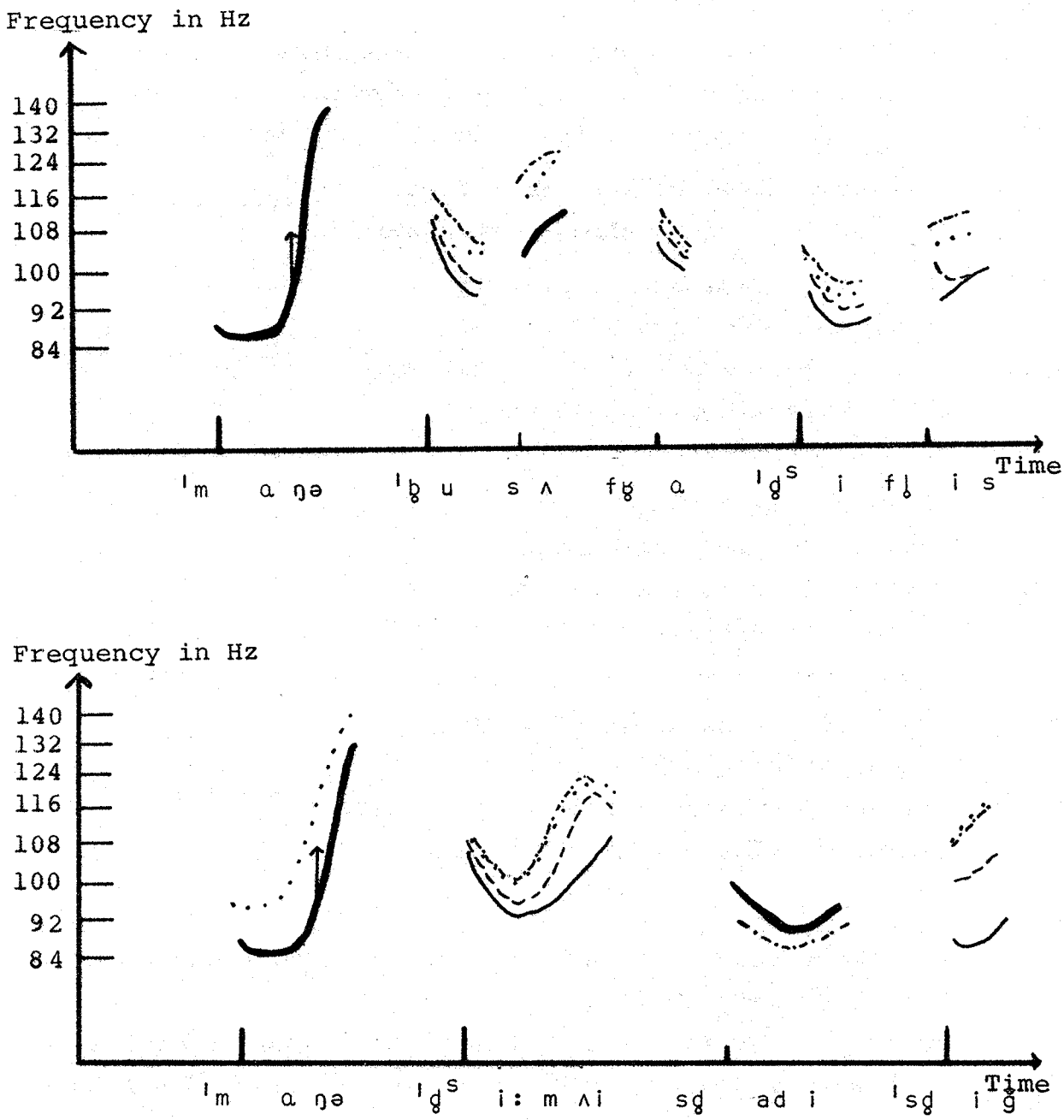
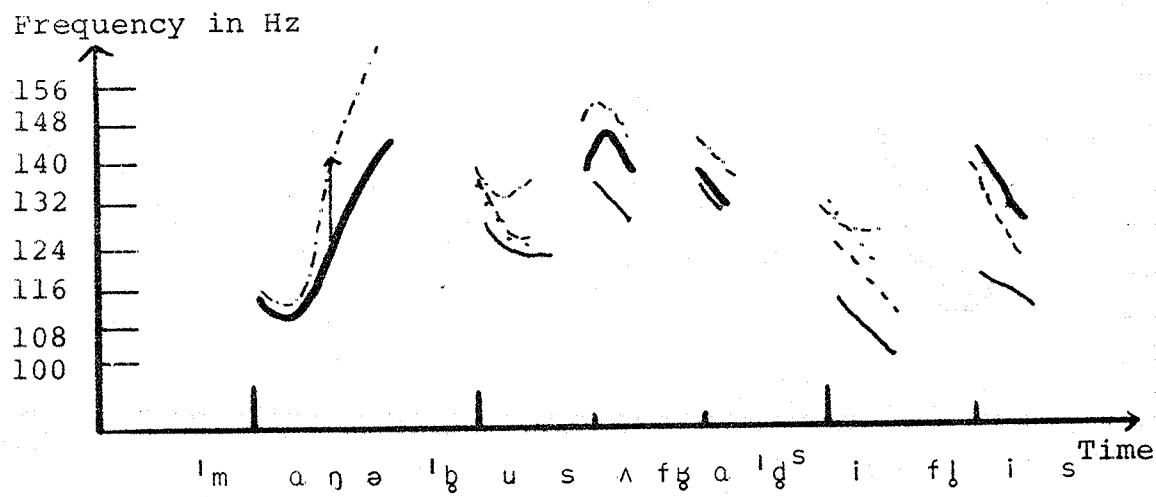


Figure 17

(Subject NRP)

F_0 tracings (averages of five recordings) of the sequence '°.. mánge bússer fra Tíflis' (above) and '.. mánge tímer i statistík' (below) in 4 editions: one statement and three different types of non-terminal clauses. See the text for an account of the lines.



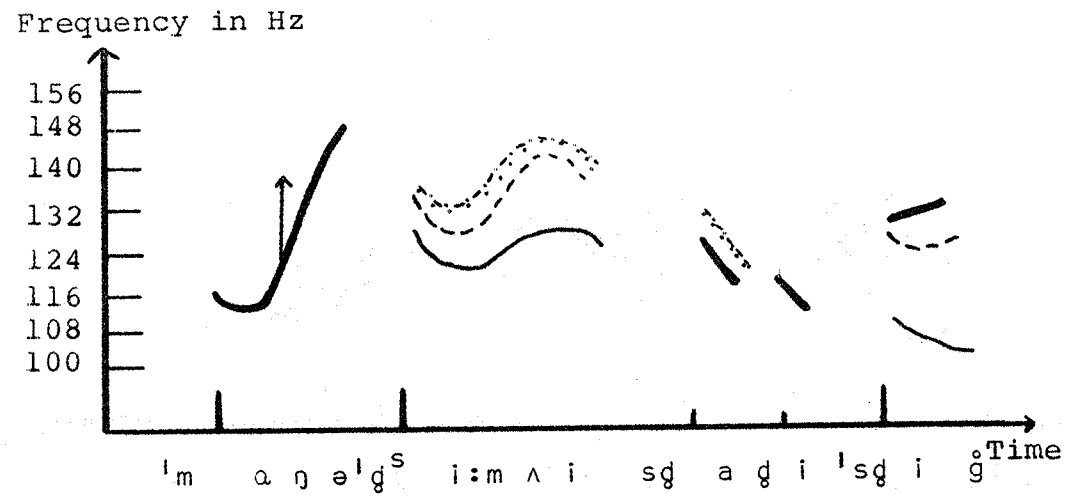


Figure 18

(Subject SH)

See the legend to figure 17.

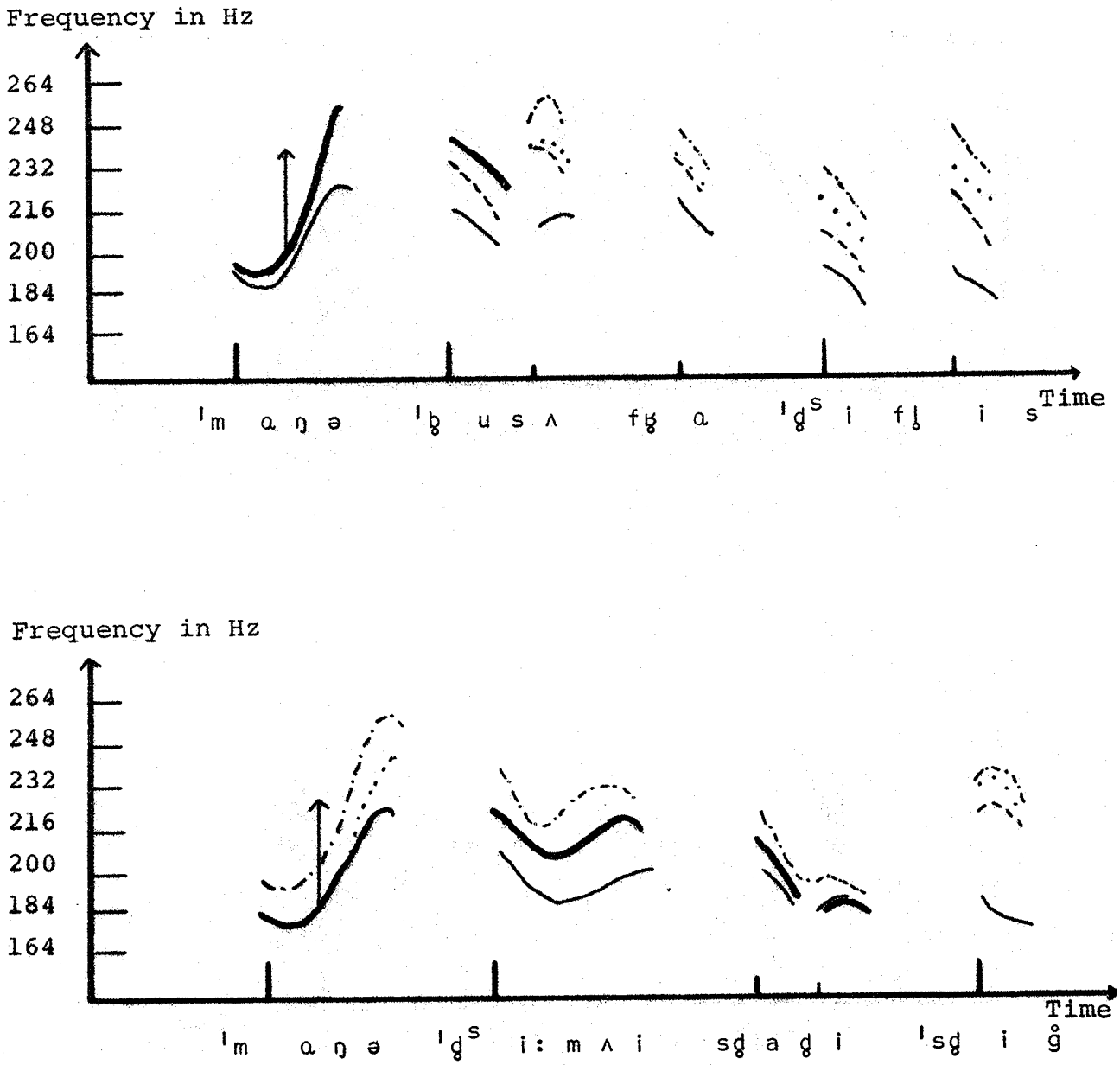
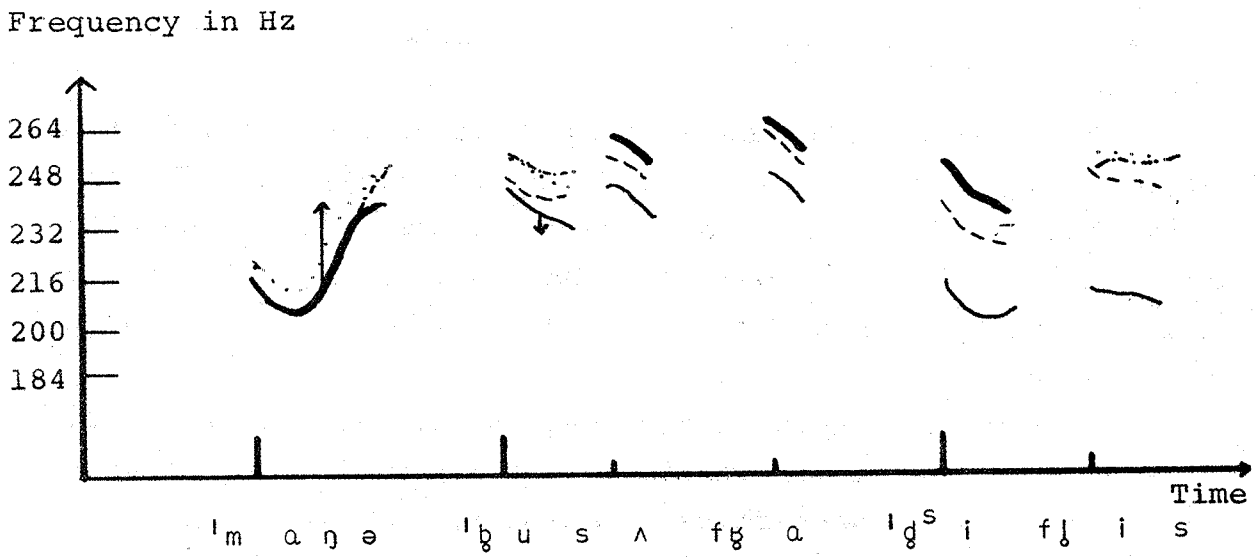


Figure 19

(Subject BH)

See the legend to figure 17.



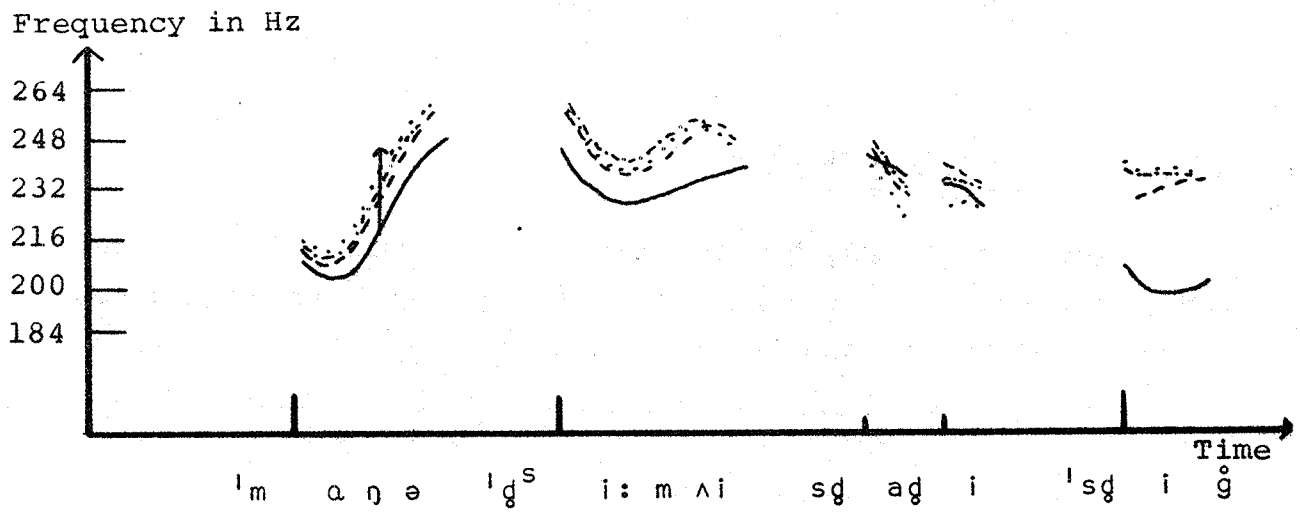


Figure 20

(Subject JG)

See the legend to figure 17.

With JG the three non-terminal courses are all relatively high and close together, - in fact, they are as high as her intonation questions, cf. fig. 15. The same is true of NRP, sentences B7 and B8.

It is evident that some of the intonation contours of the non-terminal clauses, also with SH and BH, merge with the contours in some of the interrogative sentences. I.e., the clear differences which were found with BH (and, partly, SH) in isolated declarative, non-terminal, and interrogative words do not generally exist (cf. section 3.2.1.2).

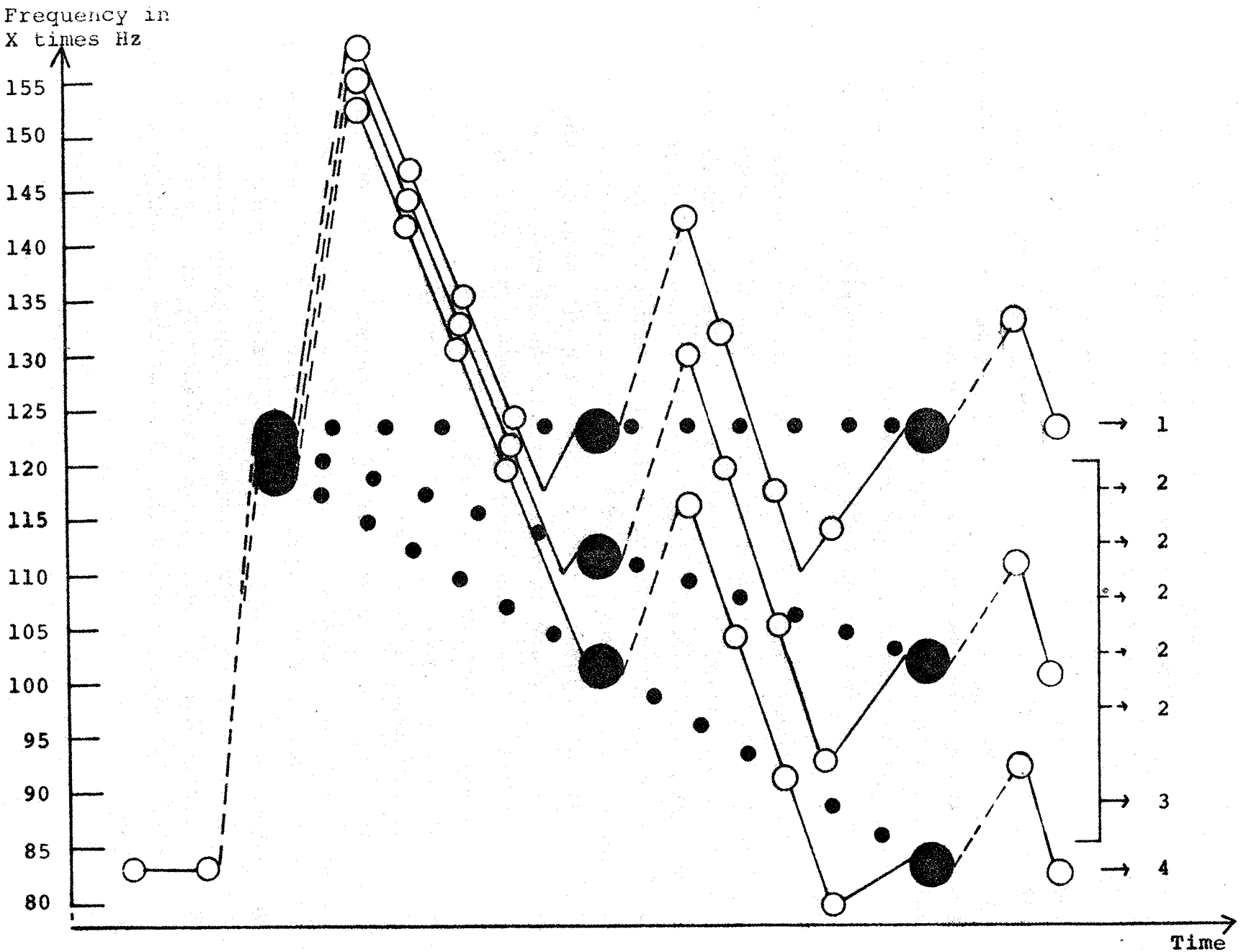
It seems more than likely that in identification experiments some of the 'interrogative' contours will be confused with some of the 'non-terminal' contours, and vice versa. It is possible, also, that the intonation contours will tend to be classified in two, rather than three, groups as terminal versus non-terminal, i.e. subjects may not be able to distinguish interrogative from non-terminal intonation (and such a distinction may not exist with the speaker, either).

4. Summary of the results

The results may be summarized graphically in a (rather tentative and hypothetical) model for F_o patterns and contours in short sentences and non-terminal clauses, cf. fig. 21.

Figure 21

A model for F_o in short sentences and non-terminal clauses: 1) Questions where the only cue distinguishing the sentence from a statement is intonation, 2) questions with word order inversion, non-terminal clauses in compound sentences, 3) questions with interrogative particle, 4) declarative sentences.


Frequency in
X times Hz
155
150
145
140
135
130
125
120
115
110
105
100
95
90
85
80
Time
1
2
2
2
2
2
3
4

This model is, of course, a greatly stylized and simplified version of the facts as they were found and it must be regarded as provisional. Future experiments will probably lead to adjustments and modifications.

A scale has been tentatively indicated. For X equal to one it would fit NRP. For other subjects X will have to be determined, which is probably not a very simple procedure. If one measures F_o of the second stressed syllable in statements for SH, BH, and JG and divides that figure by 100 (which is the value, for X equal to one, of that syllable in the model), then a coarse approximation to X is achieved, although it predicts too high values for the highest unstressed syllables of BH and JG.

The stressed syllables have been marked with large filled circles in order to avoid specifying the actual F_o movement within the syllable. The unstressed syllables are small circles on the thin, full lines. The number of unstressed syllables within a stress group may vary. If there are more than four they will probably crowd together on the lines connecting the first posttonic syllable to the succeeding stressed syllable, and if there are fewer they will spread out on those lines.

The black dots connecting the stressed syllables describe the intonation contours. The slope of the one in the middle may vary, partly according to sentence type, cf. the legend to fig. 21, partly with different subjects.

The difference between SH on one hand and NRP, BH, and JG on the other (cf. section 3.1.1.2 and fig. 8) is not accounted for in the model, which depicts the strategy of SH. By changing the slopes of the falling unstressed syllables in each stress group, i.e. by making them steeper, the appropriate model for NRP, BH, and JG is achieved, since relatively more unstressed syllables will thus be below the level of the succeeding stressed syllable.

The course of the unstressed syllables rests on three assumptions (which cannot be rejected on the basis of the present material), namely 1) that the difference in Hz between a given stressed syllable and the immediately succeeding unstressed syllable is, for a given subject, the same, irrespective of intonation contour; this difference decreases from the first to the last stress group. 2) The slope of the unstressed syllables is the same (for a given subject) in all stress groups, which is why the last (of several) unstressed syllables are below the level of the succeeding stressed syllable in non-statement contours, and most clearly so in the second stress group, and 3) unstressed syllables where /ə/ elision or -assimilation has taken place are assimilated tonally to the preceding syllable.

The model covers single word utterances (including reeling-offs) if the initial low unstressed syllables are connected to the second or third stress group (depending on the individual).

Presumably, the model would also cover slightly longer sentences, containing four and five stress groups, which would just cause a given falling contour to be slightly less steep, its upper and lower limits remaining the same.

5. Perspectives for the future

Some of the investigations which at present seem called for are listed below. Some concern universal phenomena, others are specific to Danish. When these investigations are undertaken, they will probably in their turn bring new problems to light and create a demand for still further research.

We need to know more about:

Intrinsic F_o in unstressed syllables

Intrinsic F_o in stressed and unstressed syllables in various positions in the sentence

F_o variations caused by surrounding sounds (there are various papers on this subject but it is far from exhausted yet (see Jeel 1975 and the references therein))

The F_o course of varying numbers of unstressed syllables within a stress group, in different positions in the sentence

The mechanism of tonal assimilation in words with unstressed /ə/ syllables

F_o in syllables with secondary stress, as opposed to main- and unstressed syllables

Intonation contours in compound sentences, and the dependence of stress groups upon clause boundaries

Intonation contours in free speech.

The perceptual relevance of F_o courses needs to be established. There are several, more or less sophisticated, lines of procedure:

Trained phoneticians could take down (i.e. transcribe) pitch/intonation as discrete phenomena (i.e. by means of lines and dots) and as continuous phenomena (i.e. in as good accordance with the perceived course as possible)

Trained phoneticians could beat rhythm and mimic intonation with a succession of identical CV syllables (i.e. they would hear, or produce themselves, a given sentence, whereafter they mimic it solely with CV syllables)

Listeners should adjust sinus- or triangular waves to the pitch of a given vowel (or consonant) in a given syllable

Listeners should identify the various F_o courses in the sequences '.. mánge bûsser fra Tíflis' and '.. mánge tîmer i statistîk.' These sequences can be mutilated in various ways: One can cut out words in the beginning of the sequence in order to find out how much information is necessary for the identification, and one can cut out words from the end to find out how early identifiable differences in the sequences occur

The difference limen for F_o in connected speech should be established. How well do listeners localize a given perceived difference in relation to the chain of segments? How are several simultaneous differences perceived?

Are F_o courses in voiced obstruents and (non-syllabic) sonorants of any perceptual relevance?

Finally, other varieties of Danish, besides Advanced Standard Copenhagen Danish, should be investigated.

References

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

Jeel, Vivi 1975: "An investigation of the fundamental frequency of vowels after various consonants, in particular stop consonants", ARIPUC 9, p. 191-212

Malmberg, Bertil 1961: "Haut-bas" ou "montant-descendant"? Language and Society, Essays presented to A.M. Jensen on his 70th birthday (Copenhagen)

Petersen, Niels Reinholt 1976: "Intrinsic fundamental frequency of Danish vowels", this issue.