(ARIPUC 14, 1980)

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

Nina Thorsen

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

1. Introduction

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

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

---

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

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

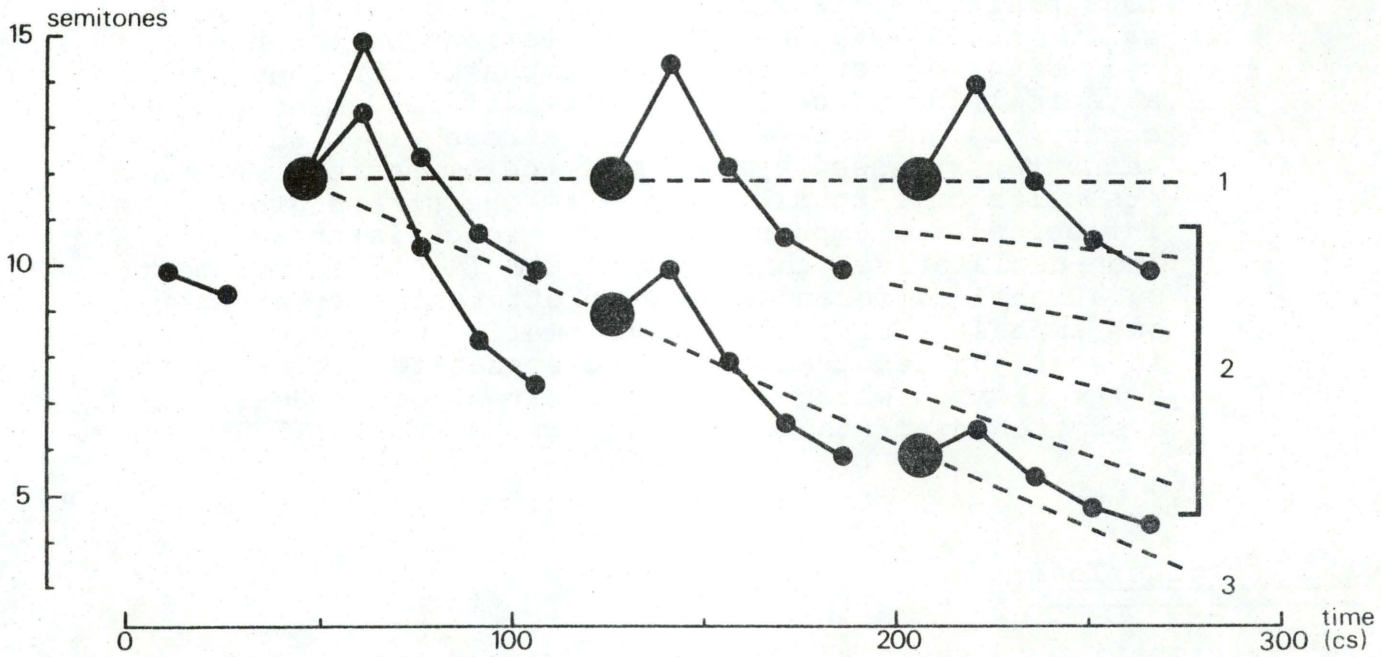


Figure 1

A model for the course of fundamental frequency in short prosodically neutral sentences in ASC Danish. 1: syntactically unmarked questions; 2: interrogative sentences with word order inversion and/or interrogative particle and non-final periods (variable); 3: declarative sentences. The large dots represent stressed syllables, the small dots unstressed ones. The full lines represent the Fo pattern associated with stress groups, and the broken lines denote the intonation contours. Zero on the logarithmic frequency scale corresponds to 100Hz.

CONTOUR (broken lines in fig. 1). (2) On the contour is superposed a stress group component which furnishes the STRESS GROUP PATTERNS (full lines). (3) To the resultant of those two components is added, in words containing stød,[1] a stød component, rendering STØD MOVEMENTS (not included in the model). These first three components are language specific and thus "speaker controlled".
(4) Finally, intrinsic Fo level differences between segments, and coarticulatory variations at segment boundaries supply a microprosodic component, which - at least in non-tonal languages - is not consciously controlled by the speaker, but can be ascribed to inherent properties of the speech production apparatus and which therefore is superfluous in the model from the point of view of the human speaker. (It seems as though in tone languages those perturbations are actively brought within time limits where they will not interfere with the perception of the tonal distinctions in the language, cf. Hombert 1977.) Similar points of view about components or "layers" in intonation have been expressed by several authors, e.g. Bolinger (1970), Bruce (1977), Carlson et al. (1974), Cohen and 't Hart (1967), Collier and 't Hart (1975), Fujisaki et al. (1979), Gårding and Lindblad (1973), 't Hart (1966), 't Hart and Cohen (1973), Lehiste and Peterson (1961), Öhman (1968); for a more detailed account, see Thorsen (1979).

2. Stress group patterns

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

---

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

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

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

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

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

the boundaries of the same intonation contour), irrespective of intervening syntactic boundaries (see Thorsen 1978, 1980a, 1980c for documentation); this unit is termed a stress group.

2.1 Stress groups with more than one post-tonic syllable

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

2.2 Stress groups with only one post-tonic syllable

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

---

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

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

2.3 Stress groups with no post-tonic syllables

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

---

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

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

2.4 Stress group patterns - conclusion

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

---

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

3. Intonation contours

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

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

---

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

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

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

4. Stress groups and intonation contours - conclusion

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

the physical variation is obliterated, recreating the simple configurations depicted in fig. 1.

5. Intrasyllabic Fo movements

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

The post-tonic vowels are generally falling. The first post-tonic, however, may be rising-falling or even purely rising, depending on the exact timing of the vowel with respect to the Fo maximum. Short stressed vowels are also generally falling on unmarked intonation contours, but may be rising in the first stress group, i.e. at the start of the intonation contour, which is always high, cf. fig. 1. On marked (less falling) contours short vowels are less falling. Long stressed vowels on falling intonation contours are falling-rising, but the fall is normally equal to or greater than the rise, except initially in the utterance. On marked intonation contours the rise may take the upper hand (see further section 7.2.1). - We might ask: Why are stressed vowels generally and mainly falling (on falling intonation contours), even though the Fo pattern is on the way up for the post-tonic syllable? and we might seek the answer in an influence from the intonation contour, since (1) we also observe a tendency for a positive correlation between short stressed vowel movement and intonation contour slope (the less steep the contour, the less falling the vowel, although the variation in vowel movement is slight, see Thorsen 1980b, p. 1020) and (2) the rise in long vowels is greater on horizontal than on falling contours. However, I think the explanation for the mainly falling stressed vowels (on falling contours) is to be found in the way stressed vowels are signalled tonally in ASC Danish, being ones that are jumped (or glided) up from: a fall in the stressed vowel will

maximize the perceptual distance to the succeeding higher post-tonic. The less steep falls/greater rises in stressed vowels on marked intonation contours may be accounted for in terms of the greater rise to the post-tonic (i.e. as an assimilation), and thus only indirectly ascribed to the intonation contour proper. - This interpretation is corroborated by a pilot study on some Danish dialects where rising stressed vowels seem to be the rule, also on falling intonation contours: These are dialects where the stress group pattern is one of high stressed plus low (or falling) post-tonic syllable(s), so the vowel movement may be seen again, not as influenced by the intonation contour, but as determined by the stress group pattern with its stipulation that the perceptual distance between (high) stressed and (low) post-tonic syllable(s) be at a maximum.

6. Emphasis for contrast

6.1 Types of emphasis

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

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

---

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

(continued)

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

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

---

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

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

6.2 Phonetic correlates to emphasis

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

---

1) Broken underlining mine.

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

6.3 Emphasis for contrast vs. sentence accent

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

The works which, implicitly or explicitly, assume the existence of a nuclear stress in British and American English are too numerous to permit an exhaustive listing. Crystal (1969) reviews past work on prosodic features and says himself (p. 207): "There is general agreement about the internal structure of the tone-unit in English. Minimally, a tone-unit must consist of a syllable, and this syllable must carry a glide of a particular kind. This is the obligatory element, and is usually referred to (in the British tradition) as the nucleus of the tone-unit. (...) The presence of a nucleus is what accounts for our intuition of 'completeness' at the end of the unit: if it is omitted, the auditory effect is one of 'being cut short'." - Liberman and Prince (1977) - in a theoretical frame-work which differs from most previous descriptions of English stress - also assume a "main stress" which is the most prominent terminal element of a given constituent and is termed a designated terminal element (p. 257). Bolinger (1958) seemingly takes a slightly different approach, in limiting the term 'stress' to the domain of the word, and talks of 'pitch accents' at the sentence level. However, I think that the difference is terminological rather than conceptual, inasmuch

as Bolinger does not deny the existence of stresses that are more prominent (by their pitch movements) than others.

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

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

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

Now, what is it about Danish that makes it deviant from English, German, and Swedish (we will leave Dutch aside)? First of all, there is nothing "incomplete" about an utterance with a neutral prosodic pattern, i.e. one where none of the stressed syllables is more prominent than the others. Such neutral utterances may not be very common in spontaneous speech but they do occur, they are not unnatural, and they are very easy to elicit from speakers in a reading situation. Secondly, and this is the crucial point: In English, German, and Swedish the sentence accent does <u>not</u> imply, or result in, a deletion of the surrounding stress group patterns (i.e. a reduction of main to secondary stress) the way that emphasis for contrast does in Danish (cf. the quotation from Heger 1975 above and see the results below). Thus, to sum up: in neutral utterances in English, Swedish and German four degrees of stress are necessary to account adequately for the distribution of and relation between stresses, whereas three relational degrees are sufficient in Danish: Main stress (assigned to the lexically stressed syllable in most non-function, or non-empty, words, secondary stress (assigned to the second (and following) lexically stressed syllable in compounds), and weak stress, see further Basbøll (1978) and Rischel (1972, 1975). If we accept stress as a hierarchical feature which describes <u>relations between</u> syllables rather than a (absolute) <u>property of</u> syllables, or vow-

els (as suggested by Fischer-Jørgensen 1948, Rischel 1964 and 1972, and later by Liberman and Prince 1977, and Selkirk (forthcoming)), then the same three degrees will suffice in short[1] utterances with contrast: The stressed syllable of the contrasted word will retain its main stress, and the surrounding stressed syllables will be reduced to secondary stresses, i.e. they will retain all of their stressed syllable properties (vowel quantity, quality and stød, if any) except the tonal one. Precisely because the difference between main and secondary stresses in prosodically neutral utterances is primarily tonal and because the effect of a contrast emphasis is mainly tonal (cf. below), I see no reason why a distinction between secondary stresses which are the result of a reduction due to a contrast emphasis and 'original' secondary stresses should be called for - presumably both types will be secondary stresses on a par with each other.

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

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

---

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

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

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

6.4 Acoustic analysis of emphasis for contrast in ASC Danish

6.4.1 Material, subjects, and procedures

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

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

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

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

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

The code for the different sentences reads as follows:

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

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

---

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

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

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

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

49 Der går mange busser fra Tiflis.
50 Der går mange busser fra Tiflis.
51 Der går mange busser fra Tiflis.
52 Der går mange busser fra Tiflis.

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

A: Sorry, what did you say? - Do the buses leave from Tiflis or from Grosny?
B: There are many buses out of Tiflis. - As far as I can see there is no connection from Grosny at all.

A: pi'pipi giver kortere stavelser.
B: pi'pipi giver kortere stavelser?
A: Yes - that's reasonable enough. It is a three syllable word.

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

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

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

The recordings were made with semi-professional equipment (Revox A-77 tape recorder, Sennheiser MD21 microphone, larynx microphone) in a quasi-damped room at the Institute of Phonetics on Agfa PE36 tape at 7½ i.p.s. The tapes were processed by hardware intensity and pitch meters (F-J Electronics) and registered on a mingograph (Elema 800) at a paper speed of 100 mm/s. The signal from the larynx microphone was processed in the hold mode. This, in combination with adjustment of the zero-line to the lower limit of the subject's voice range and full exploitation of the record space of the mingograph galvanometer, yields a good solution of the frequency scale, generally allowing for a measuring accuracy of 1 Hz for males and 2 Hz for females.

In unidirectional Fo courses with constant slope only the beginning and end points were measured, according to a procedure outlined in Thorsen (1979, p. 63-66). In more complex Fo courses three to six points were measured, in a manner so that the traces could be accurately reconstructed by smooth interpolation through the measuring points. The distance in time of each measuring

point from the first one was measured, and so was the duration of each segment, to the extent that reliable segmentation could be made. Fo and time measurements were averaged over the six (ten) recordings by each subject: the average Fo values were converted to semitones (re 100 Hz) and average tracings drawn. No correction was made for intrinsic Fo level differences between vowels.

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

6.4.2 Results

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

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

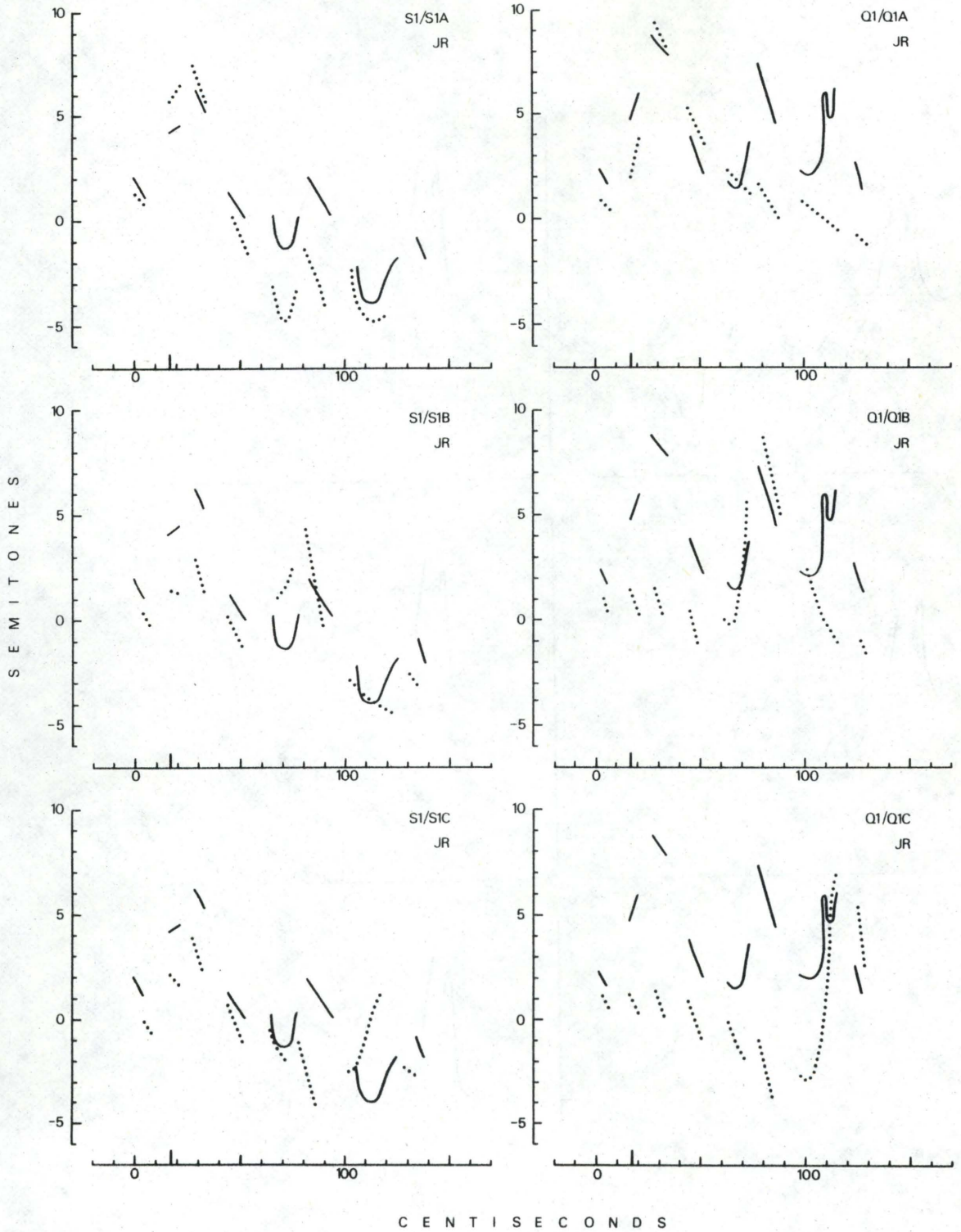


Figure 2-1

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

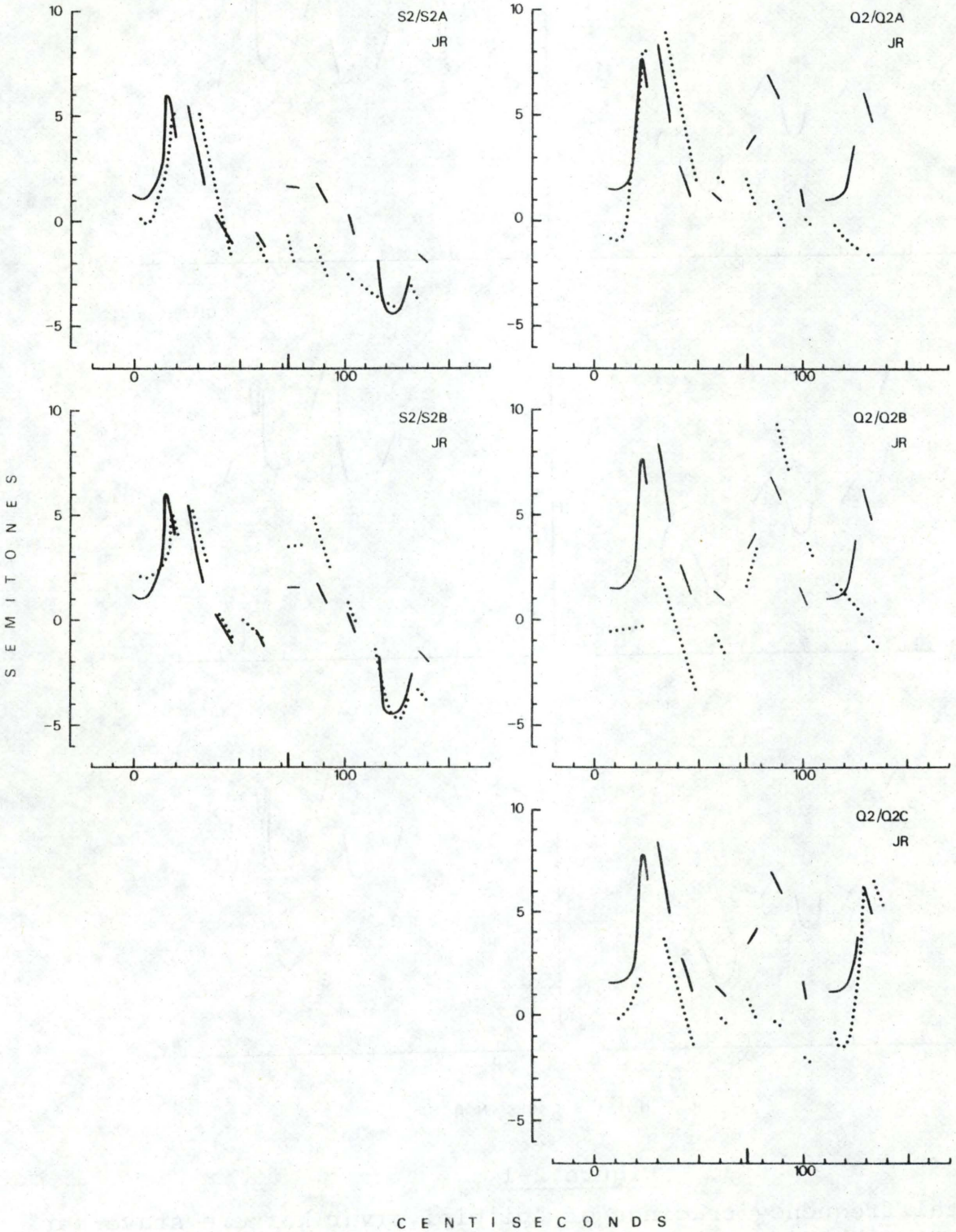


Figure 2-2

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

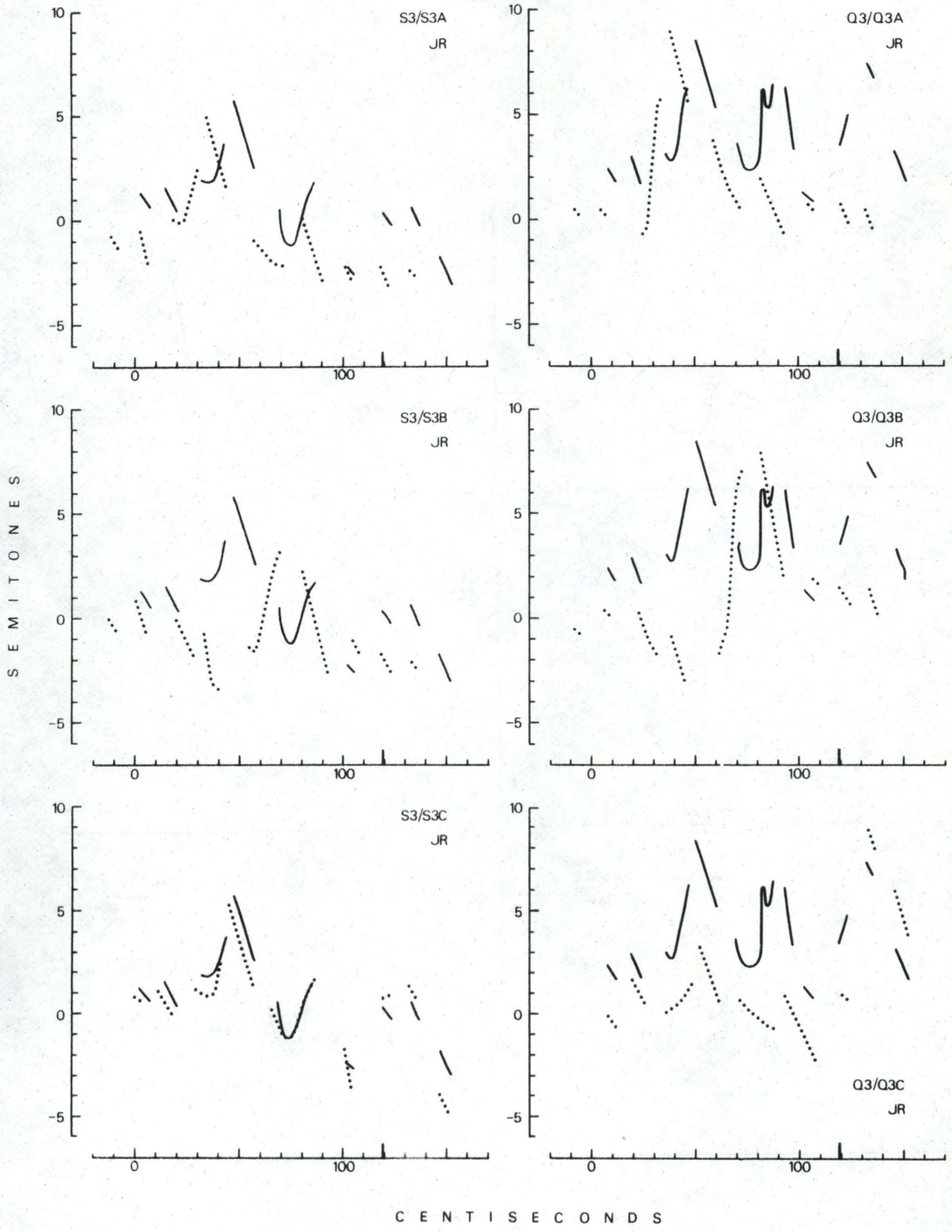


Figure 2-3

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

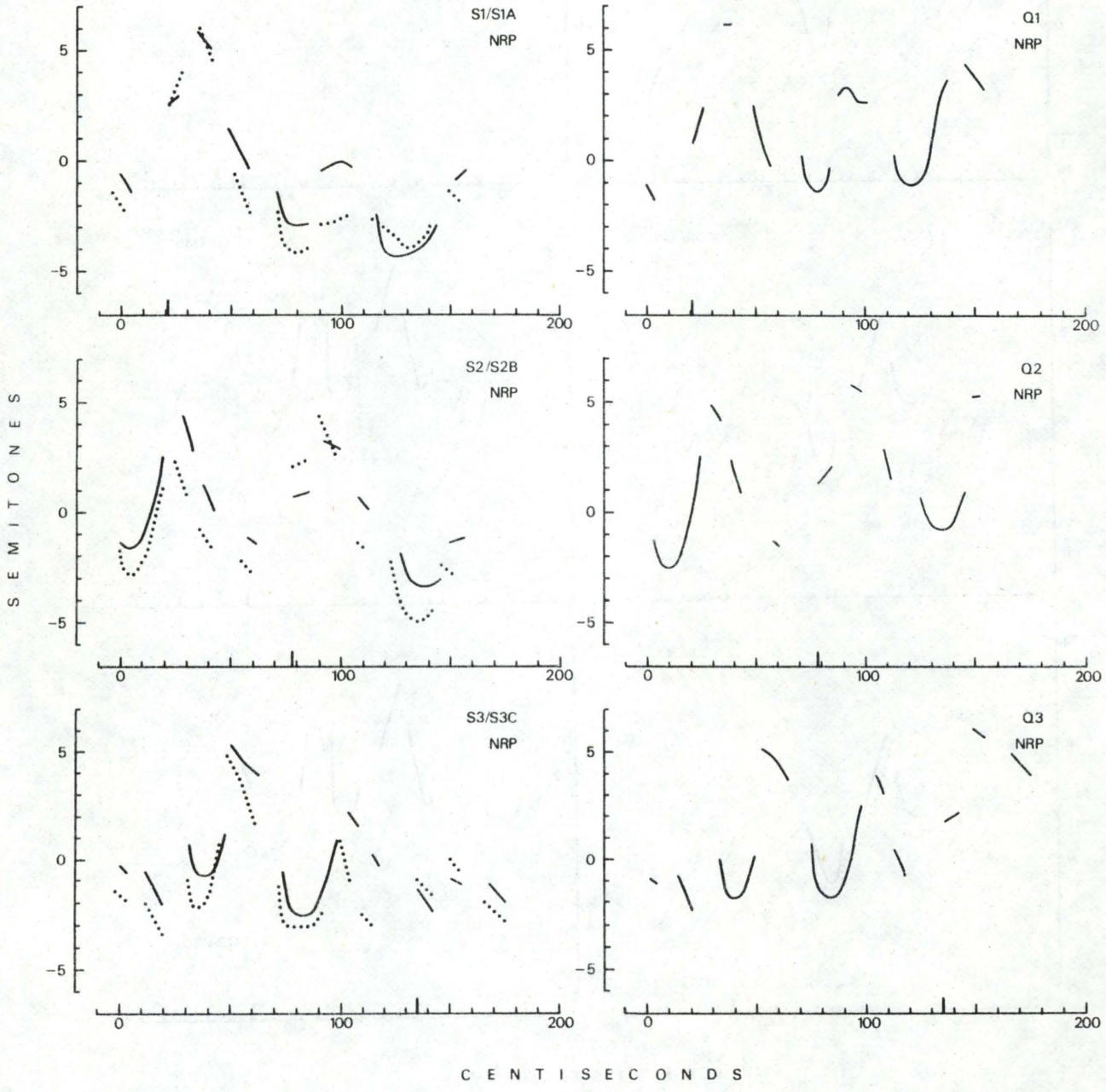


Figure 3

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

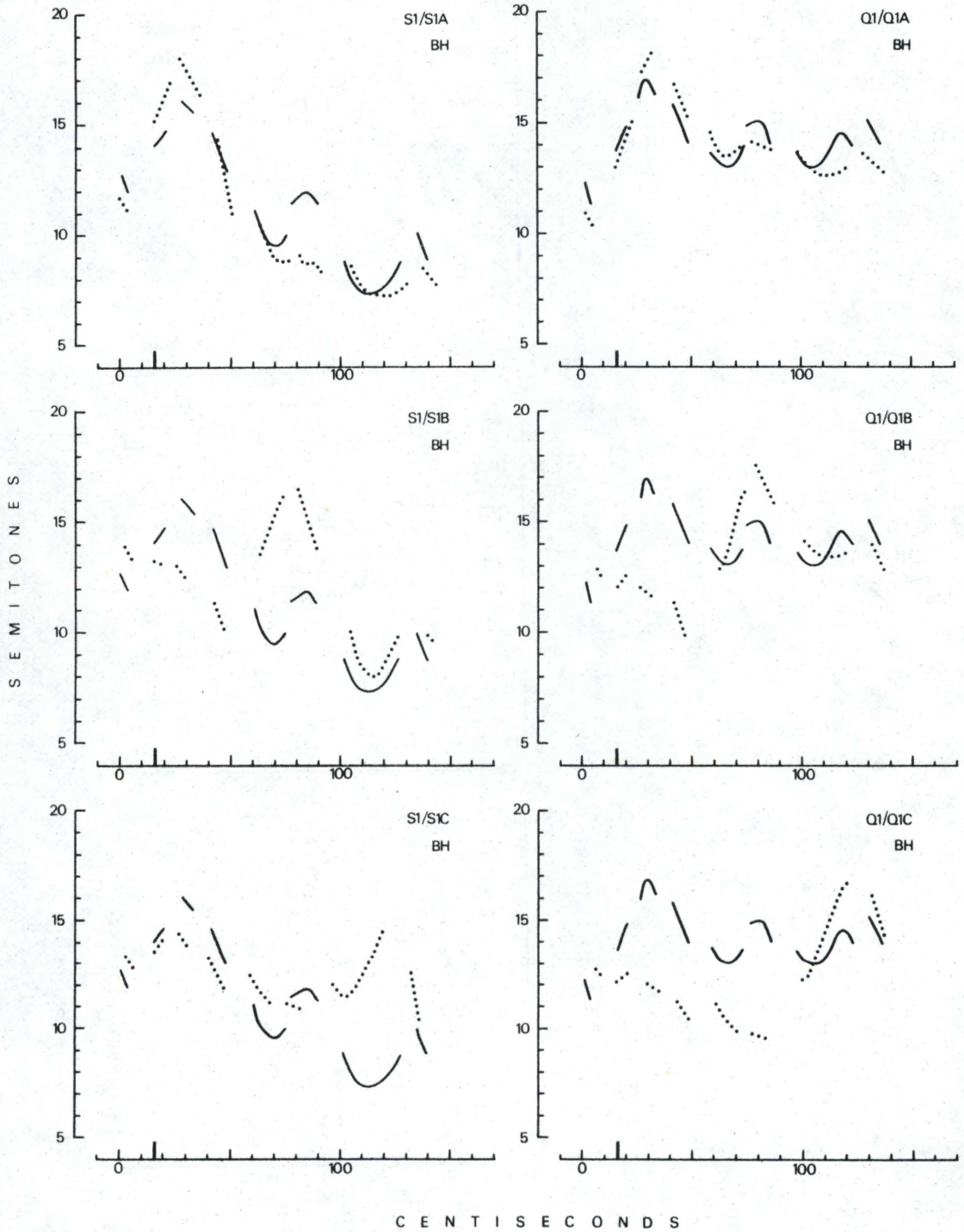


Figure 4-1

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

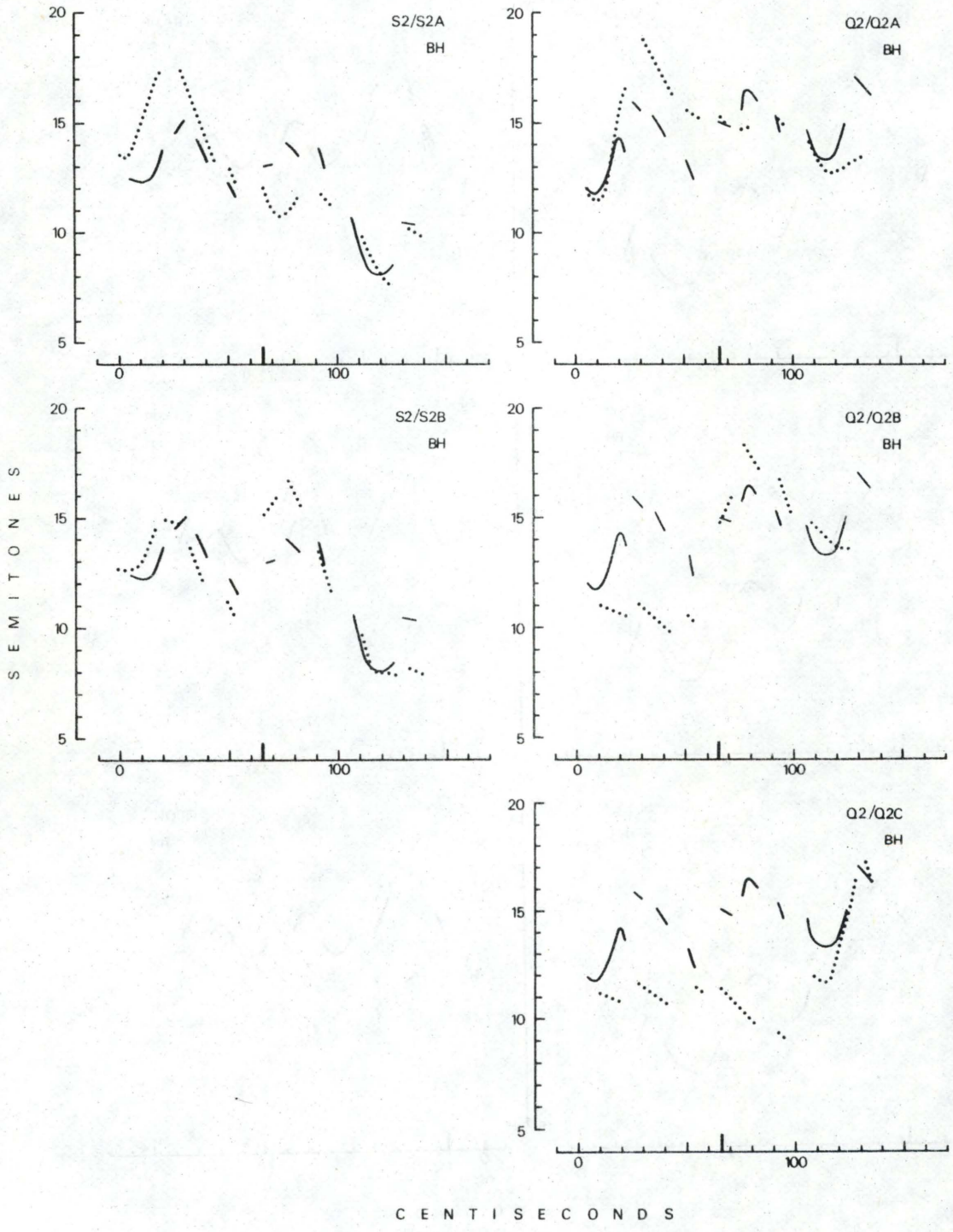


Figure 4-2

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

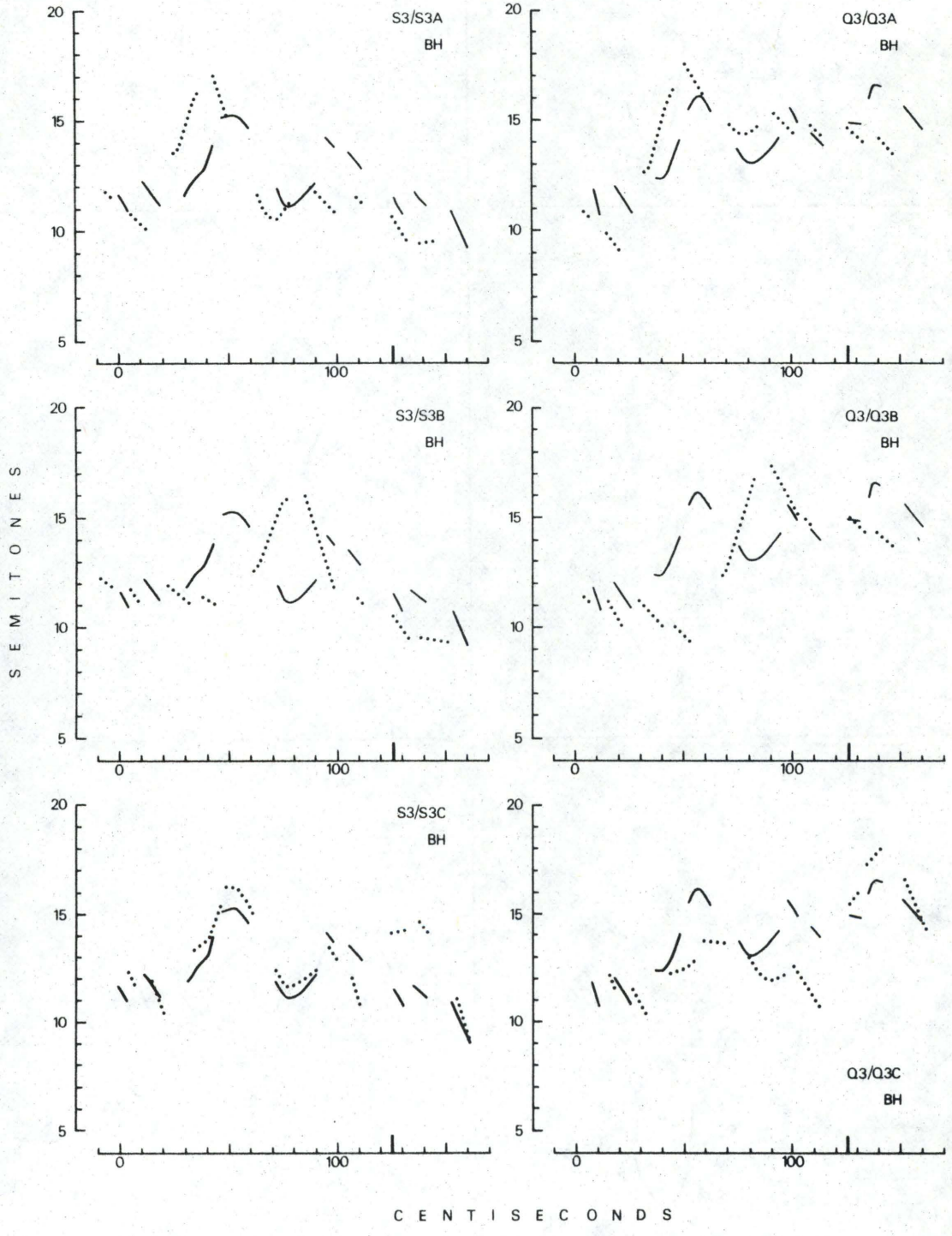


Figure 4-3

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

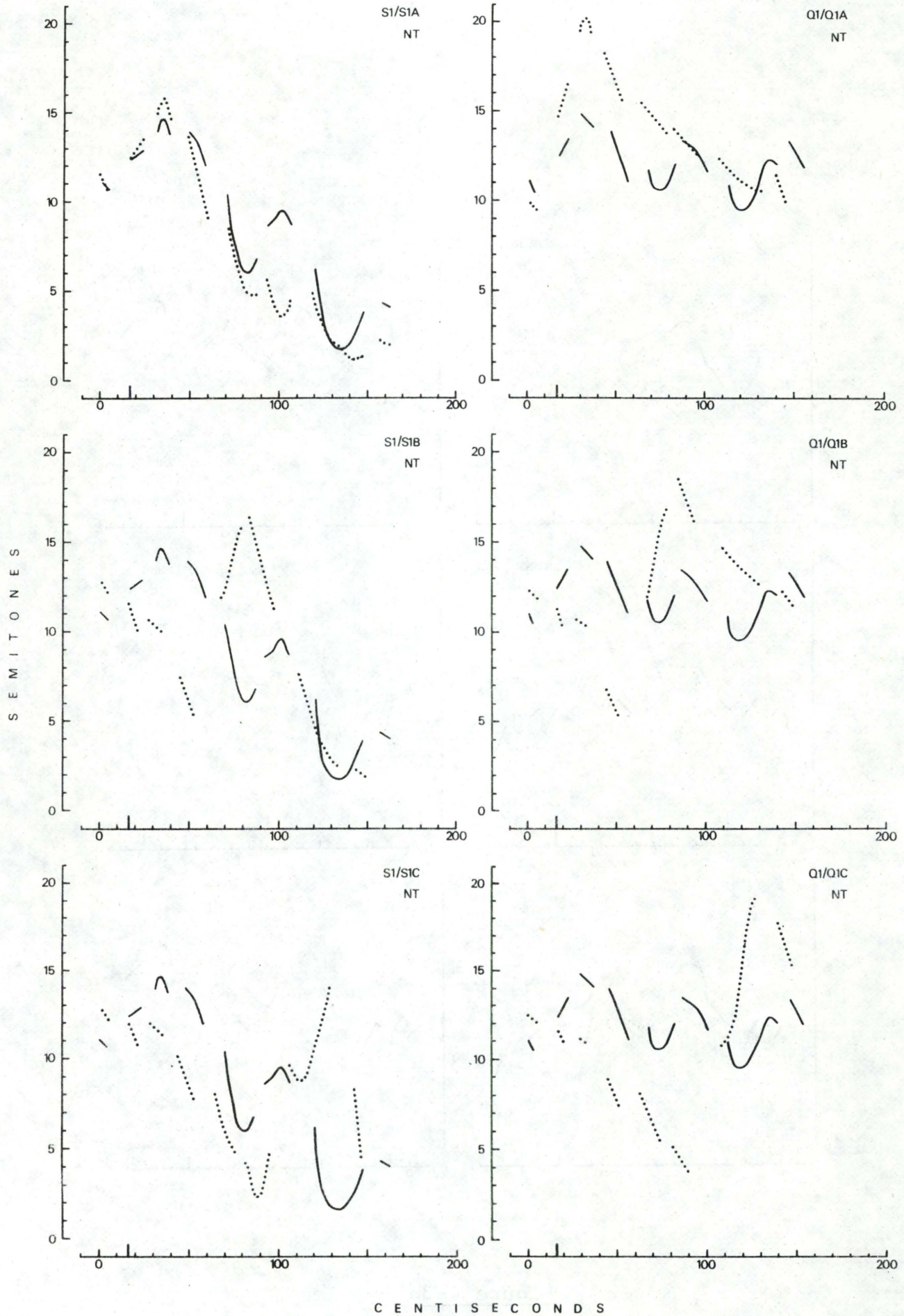


Figure 5-1

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

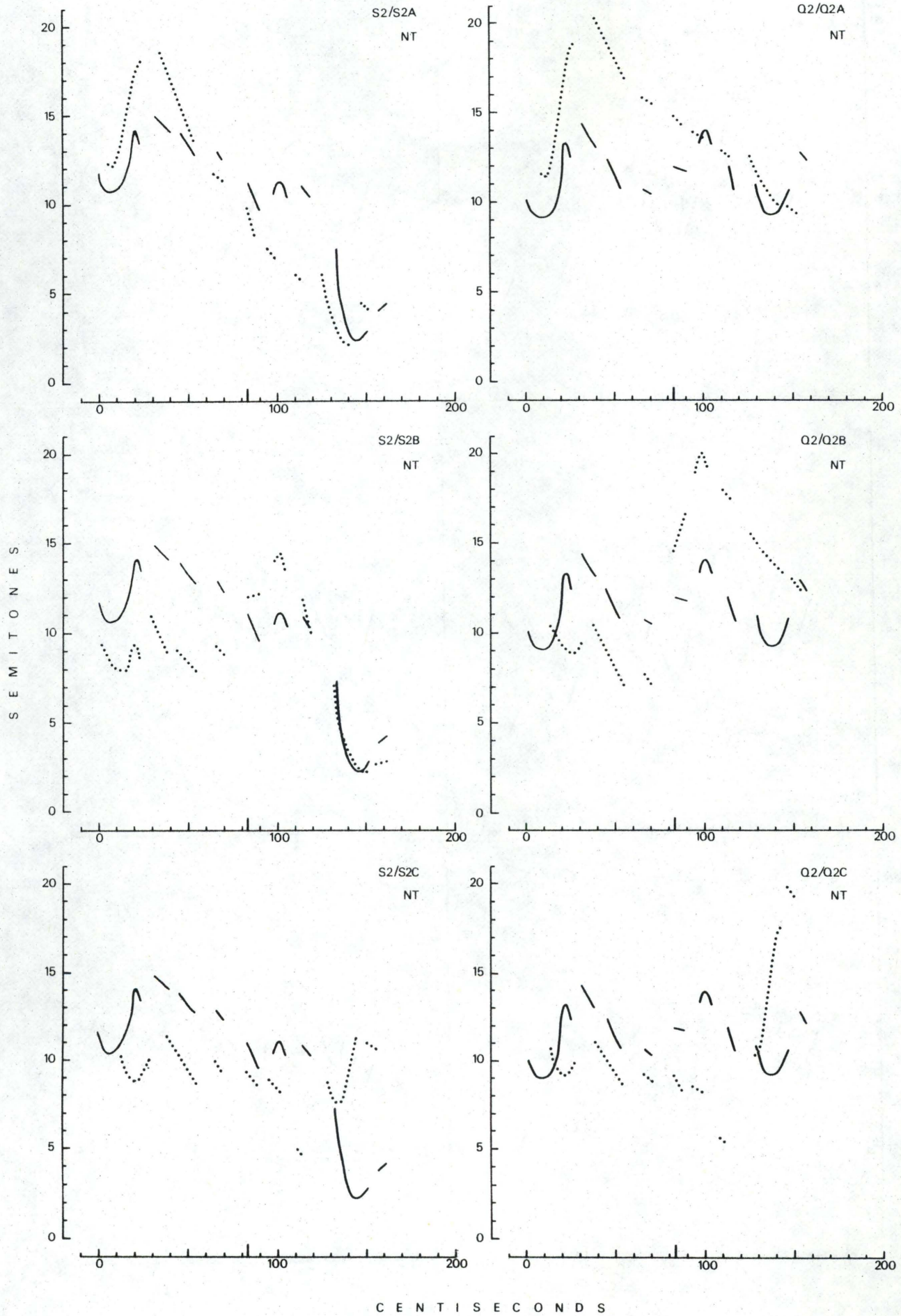


Figure 5-2

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

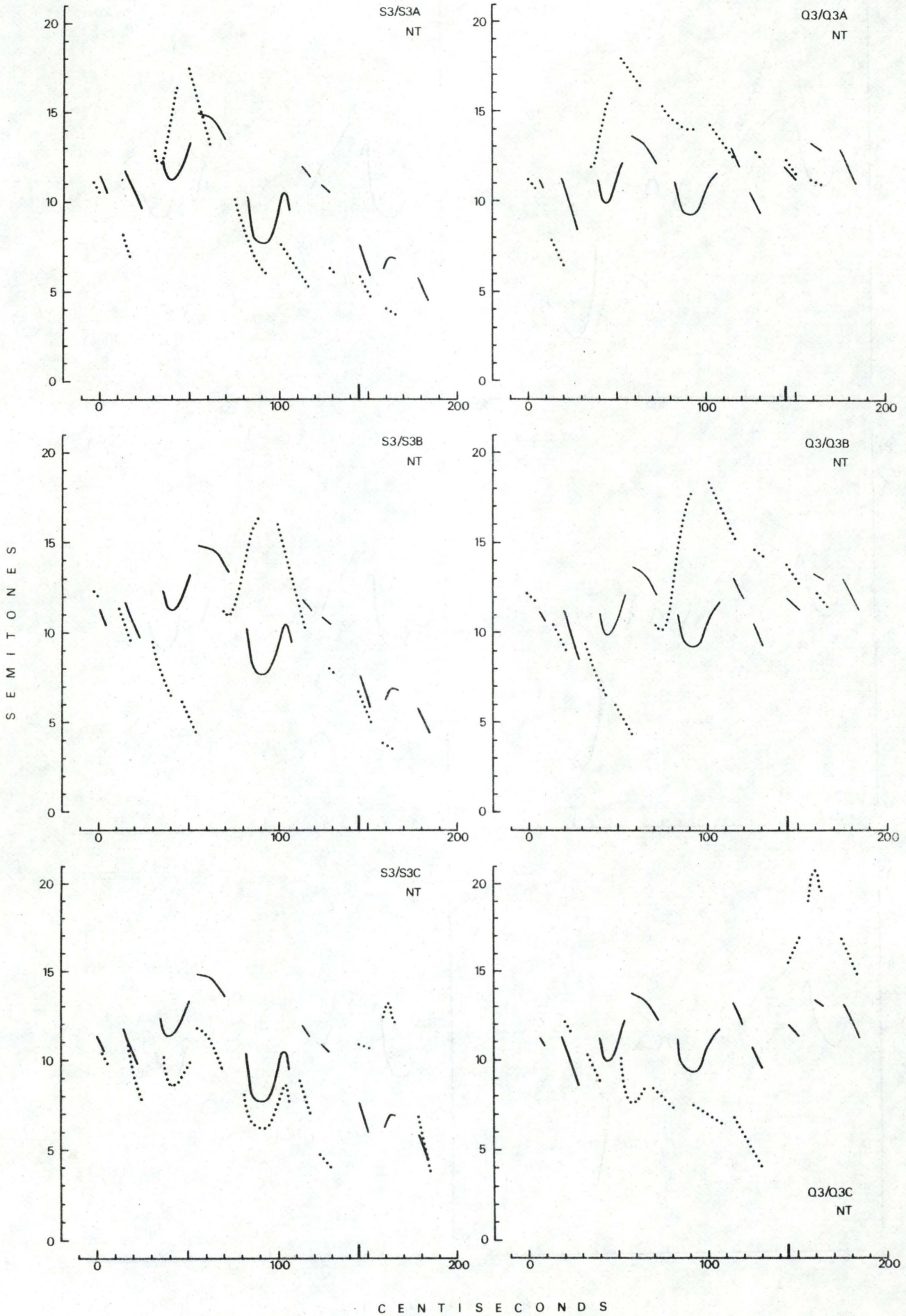


Figure 5-3

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

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

6.4.2.1 Nonsense words

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

---

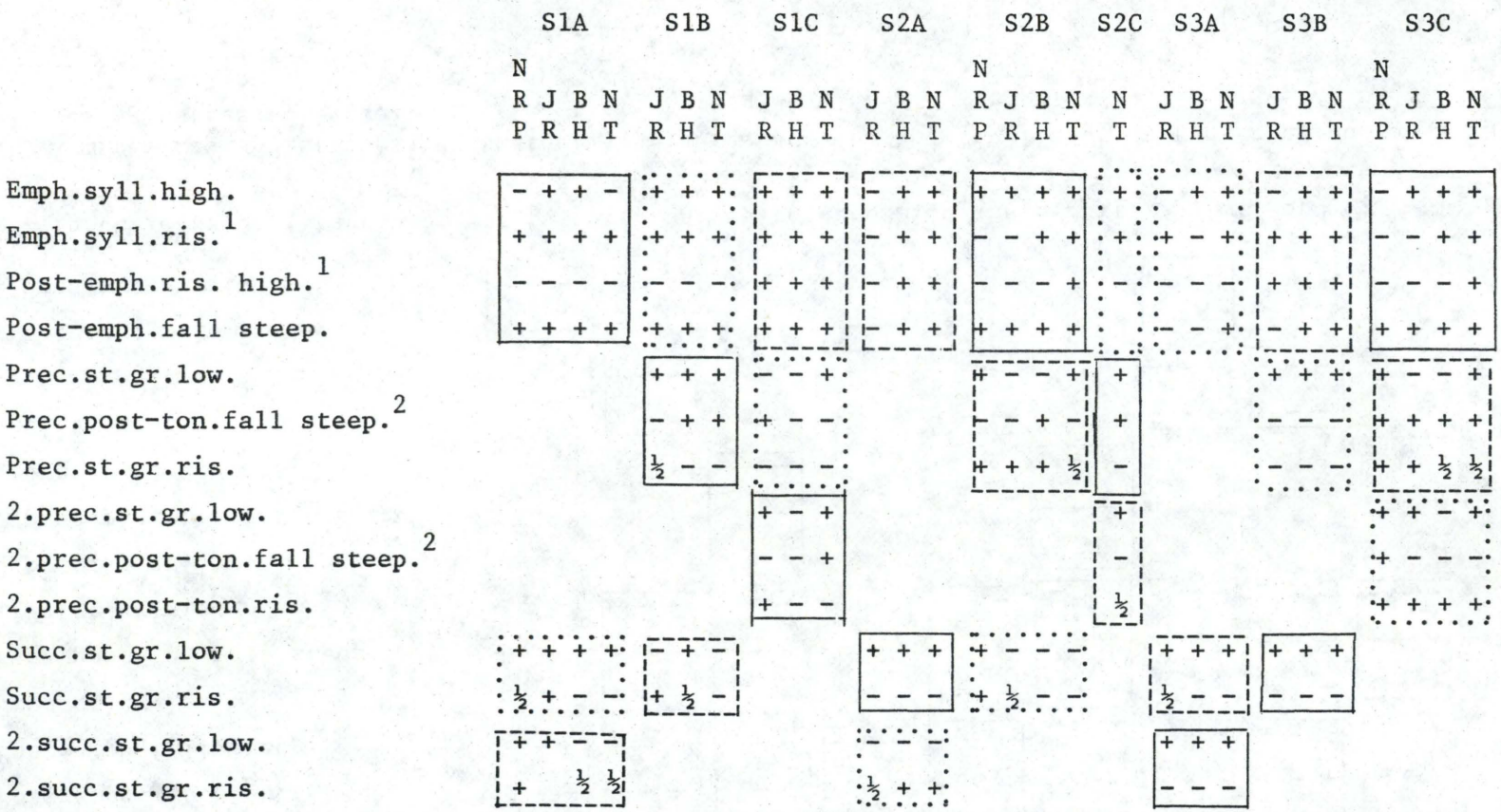
1) When, in the following, I talk of "emphatic/emphasized syllables/vowels" it is shorthand for "the stressed syllable/vowel of the emphasized word", and "emphatic/emphasized stress group" is short for "the stress group which contains the stressed syllable of the emphasized word".

Table 1

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

The abbreviations characterizing the various features read as follows:

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

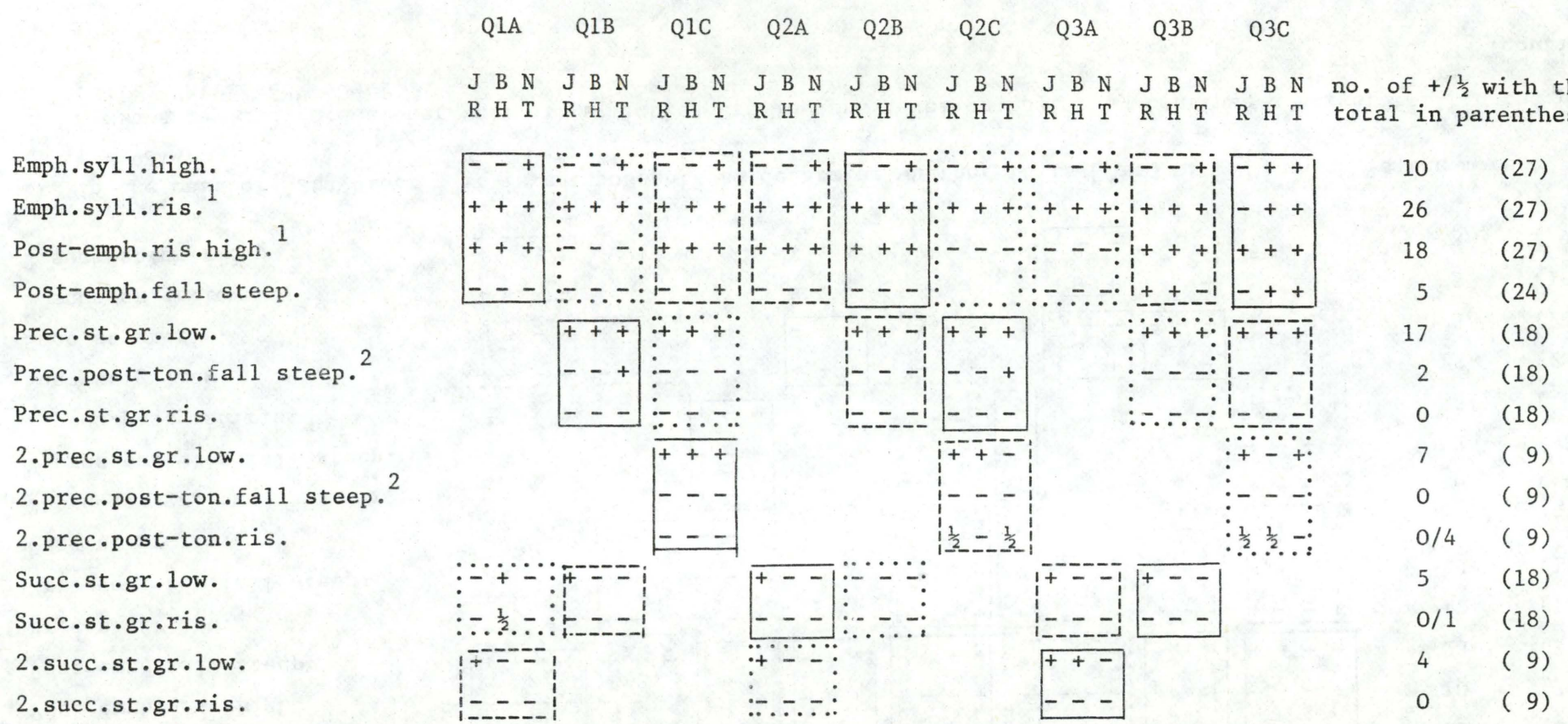


1) In the case of 'stavelser-' it is not possible to determine whether it is [ˈæː] or [wˌ] which is the more rising.

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

(continued)

(Table 1 continued)



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

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

Table 2

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

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

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

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

3) All long vowels.

4) All syllabic laterals.

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

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

(a) Statements

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

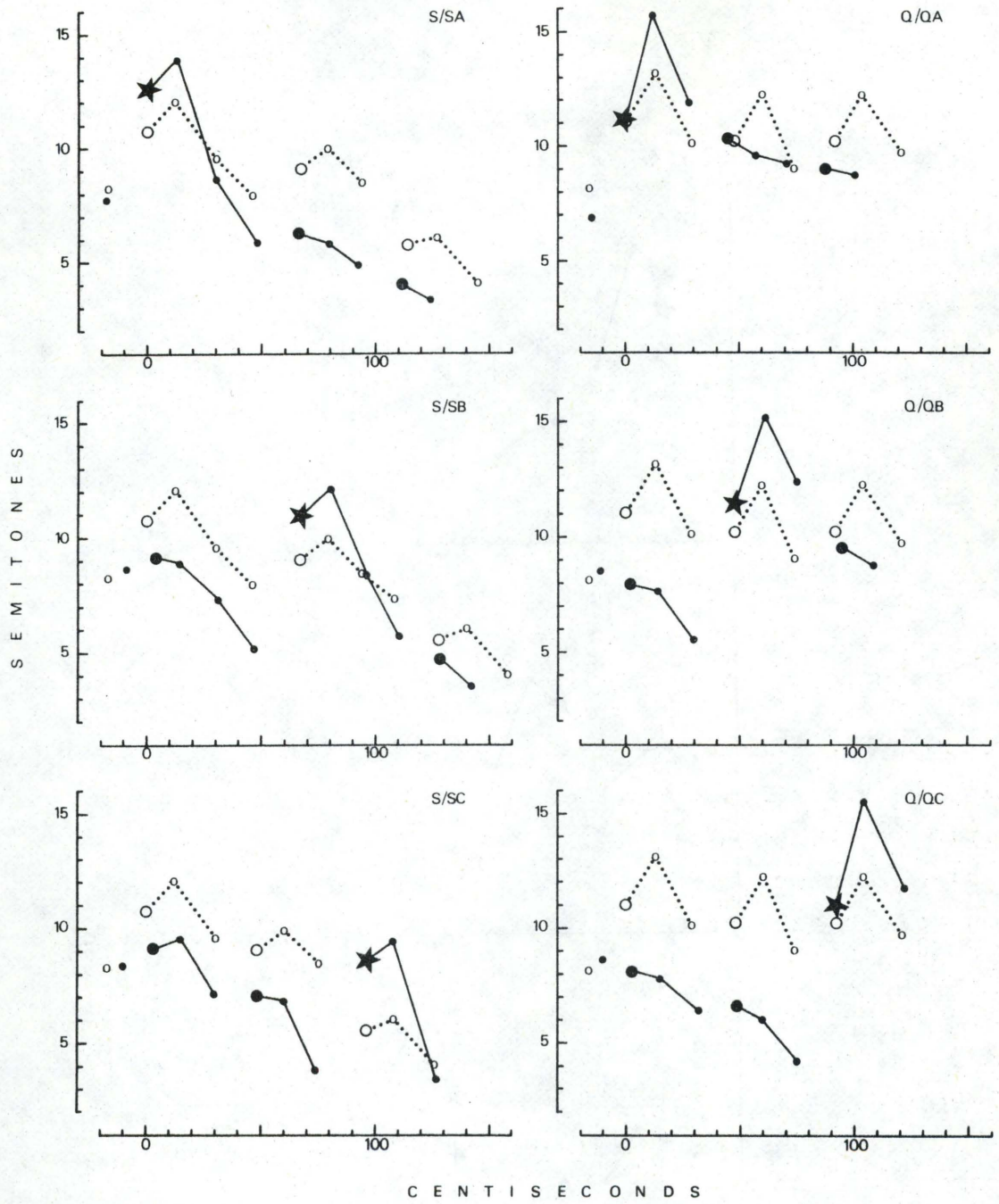


Figure 6

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

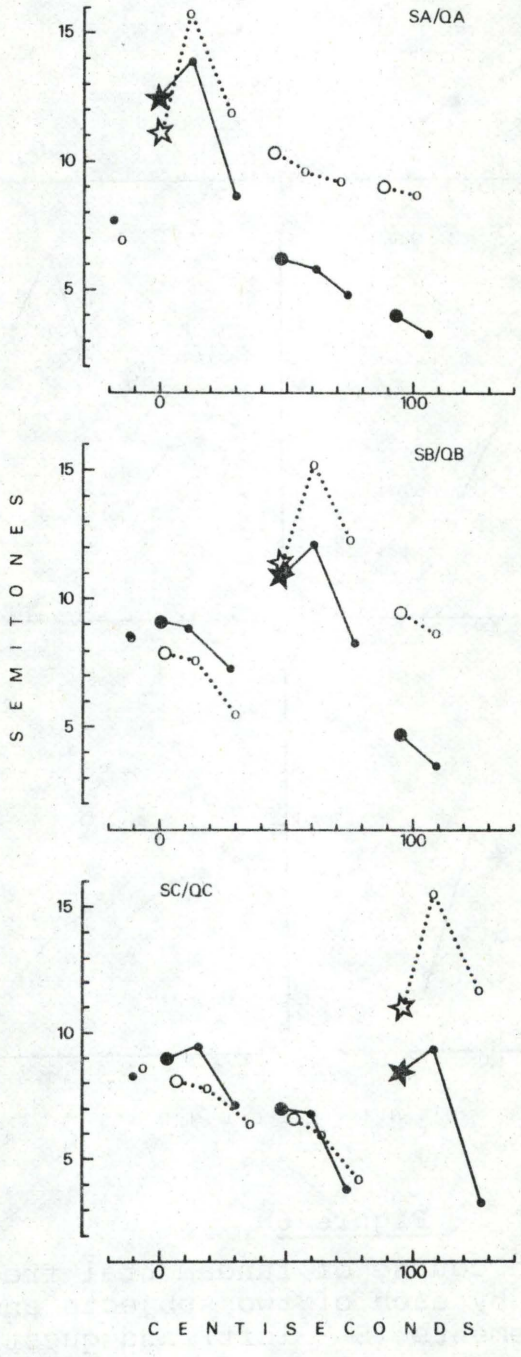


Figure 7

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

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

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

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

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

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

---

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

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

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

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

(b) Questions

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

Stress groups preceding the emphasized syllable in the questions behave qualitatively and quantitatively like pre-emphatic syllables in the statements, to the effect that the relative lowering of preceding stressed syllables (fig. 6, mid and bottom right) is greater in questions than in statements. Subjects are unanimous on the lowering and lack of rise in preceding stress

groups, cf. the full frames in table 1, Q1B, Q2C, and Q1C. - Thus, the questions do not show any signs of a wear-off of the effect of emphasis two stress groups away, as did the statements, compare fig. 6 bottom left and right.

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

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

(c) Comparing statements and questions

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

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

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

6.4.2.2 Natural words

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

(a) 'kortere'

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

---

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

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

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

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

(b) 'stavelser --'

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

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

---

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

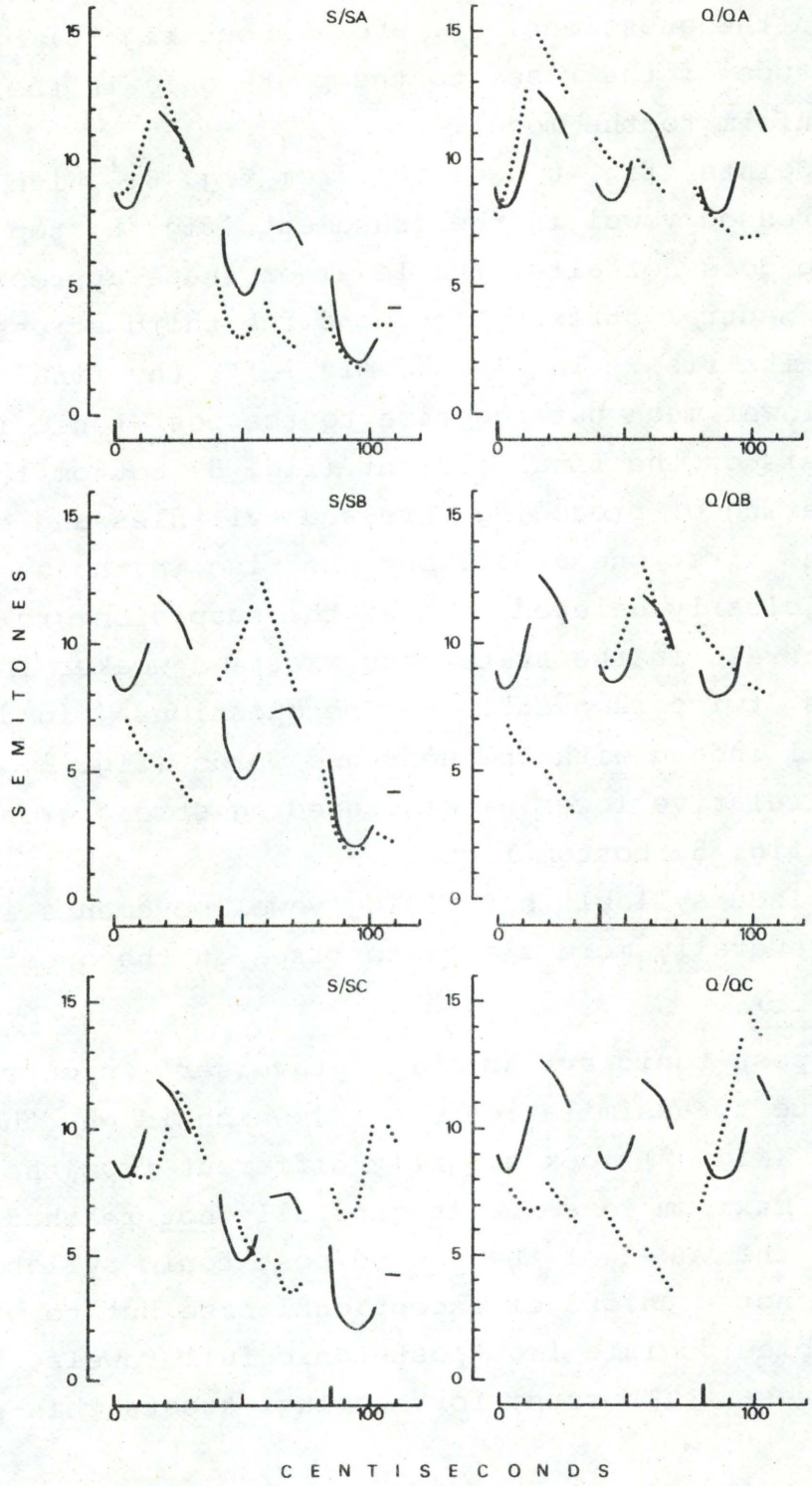


Figure 8

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

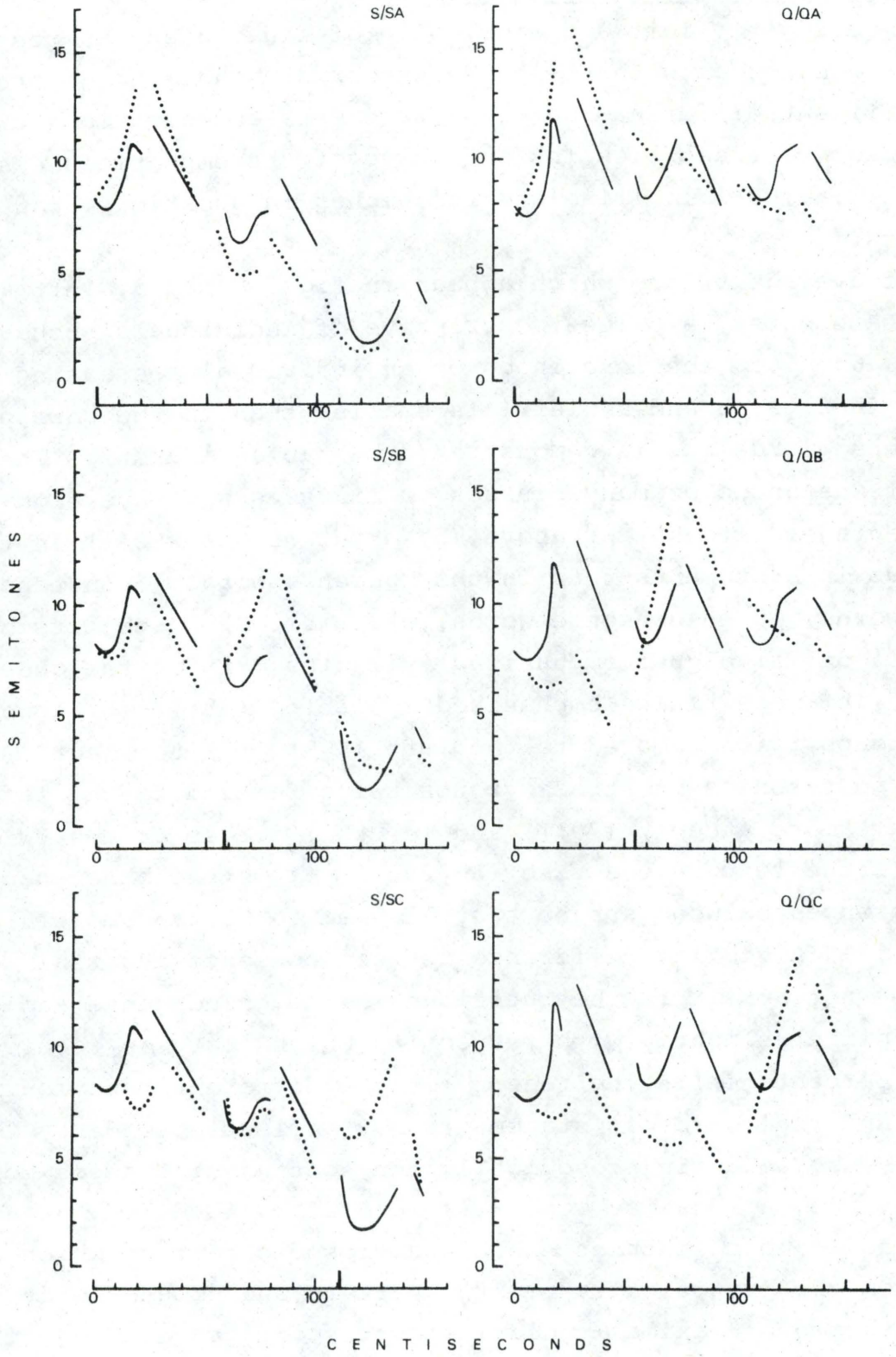


Figure 9

Fundamental frequency tracings of 'stavelserne i' [ˈsɡæ:wˌsʌnə iˌ] (initially), 'stavelser med' [ˈsɡæ:wˌsʌ mɛ] (medially), and 'stavelser' [ˈsɡæ:wˌsʌ] (finally). See further the legend to fig. 8.

(c) Comparing statements and questions

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

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

---

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

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

6.4.2.3 Natural sentences

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

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

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

---

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

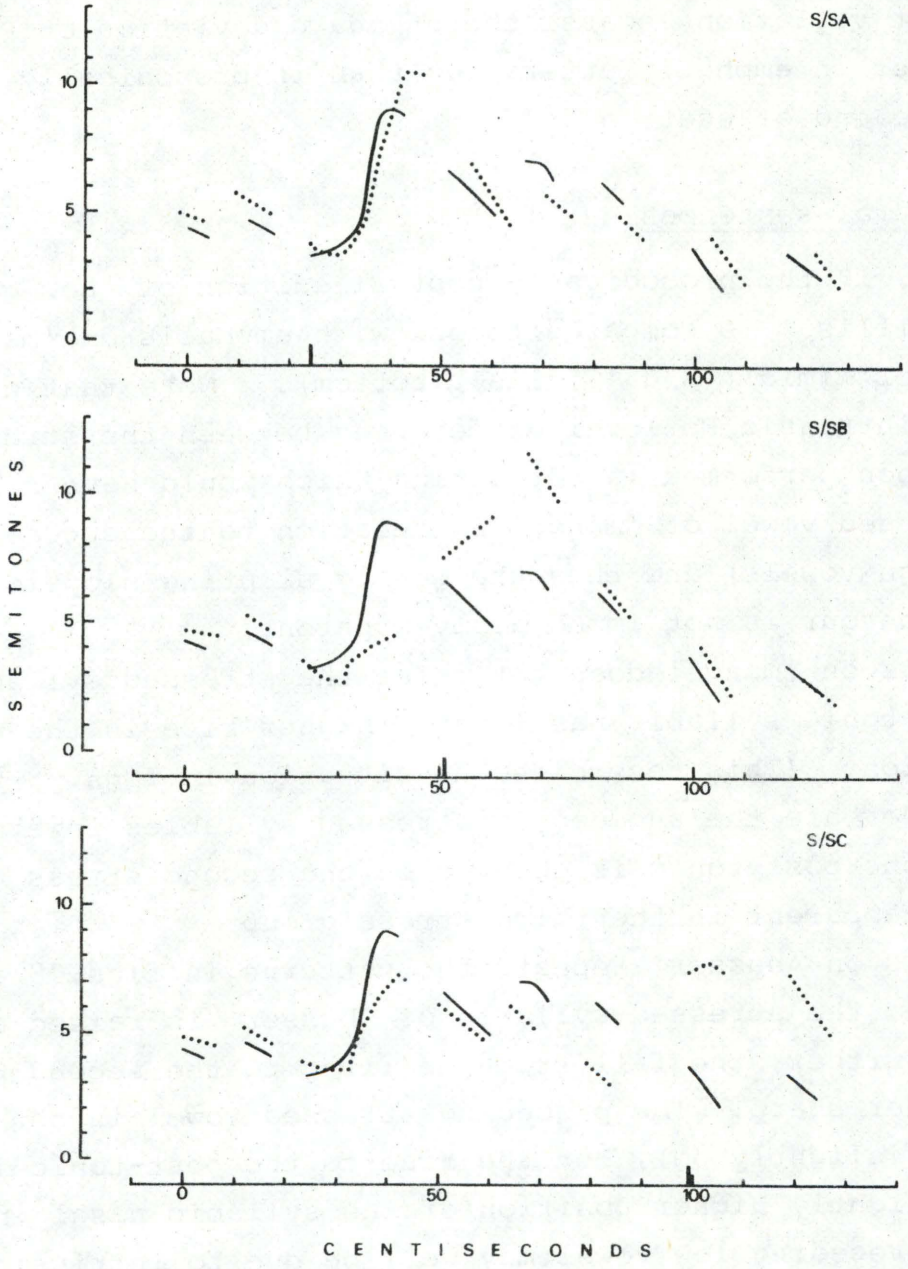


Figure 10

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

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

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

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

6.4.2.4 Fo and emphasis for contrast - conclusion

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

In this material statements and questions differ with respect to the extent of the influence of emphasis on the prosodic patterns: Fo patterns two stress groups away from the emphasized word retain (at least partly) their rises to the post-tonic in the statements but not in the questions. (In table 2 we observe a full or partial rise in the second preceding and succeeding stress groups in 76% (63%+13%) and 66% (33%+33%) of the observations, respectively, in statements but only 44% partial rises in the second preceding stress group and no rises at all in the second succeeding stress group in questions.) To the extent that this difference is a stable one, there must be a reason for it (akin to the explana-

tion at the end of section 6.4.2.2(c) for the difference in the treatment of natural words vs. nonsense words): It could be that more of a change is invoked in the prosodic patterns on marked intonation contours (which accompany syntactically unmarked questions) than on unmarked contours (declarative sentences). - The material is not comprehensive enough to establish just how far emphasis "reaches" on marked contours.

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

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

tour and stress group pattern essentially are two invariant entities which interact on the "physical" level only.

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

A complication is introduced by the fact that (at least in declarative sentences, cf. above) stress groups further away from the emphatic one are unaffected, i.e. utterance length (in terms of number of stress groups) cannot be disregarded, and we cannot always state that the whole utterance reduces tonally to one

stress group when emphasis for contrast occurs. - Instead we shall probably have to state that the stress group containing the stressed syllable of the emphasized word together with the immediately neighbouring stress groups reduce to one stress group. (In questions the emphasis may be more far reaching, but just how far I cannot say at present.) - We can presumably still make do with three relational stress degrees to describe such longer utterances with emphasis for contrast, but the stress relations that we describe must then be confined within the boundaries of each stress group (whether prosodically neutral or emphasized) and cannot extend over a whole utterance. (This does not obliterate the difference between Danish contrast emphasis and e.g. English nucleus, because in English, too, the four stress degrees describe relations within the tone group, of which there may be several in a sentence.)

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

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

Finally, I wish to stress once more the fact that what I have investigated are utterances that are emotionally neutral. - I do not wish to claim that emphasis in general must necessarily be signalled in the same way as emphasis for contrast, - in fact I am certain that this is not the case: A type of emphasis where

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

6.4.2.5 Emphasis for contrast and duration

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

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

Table 3

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

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

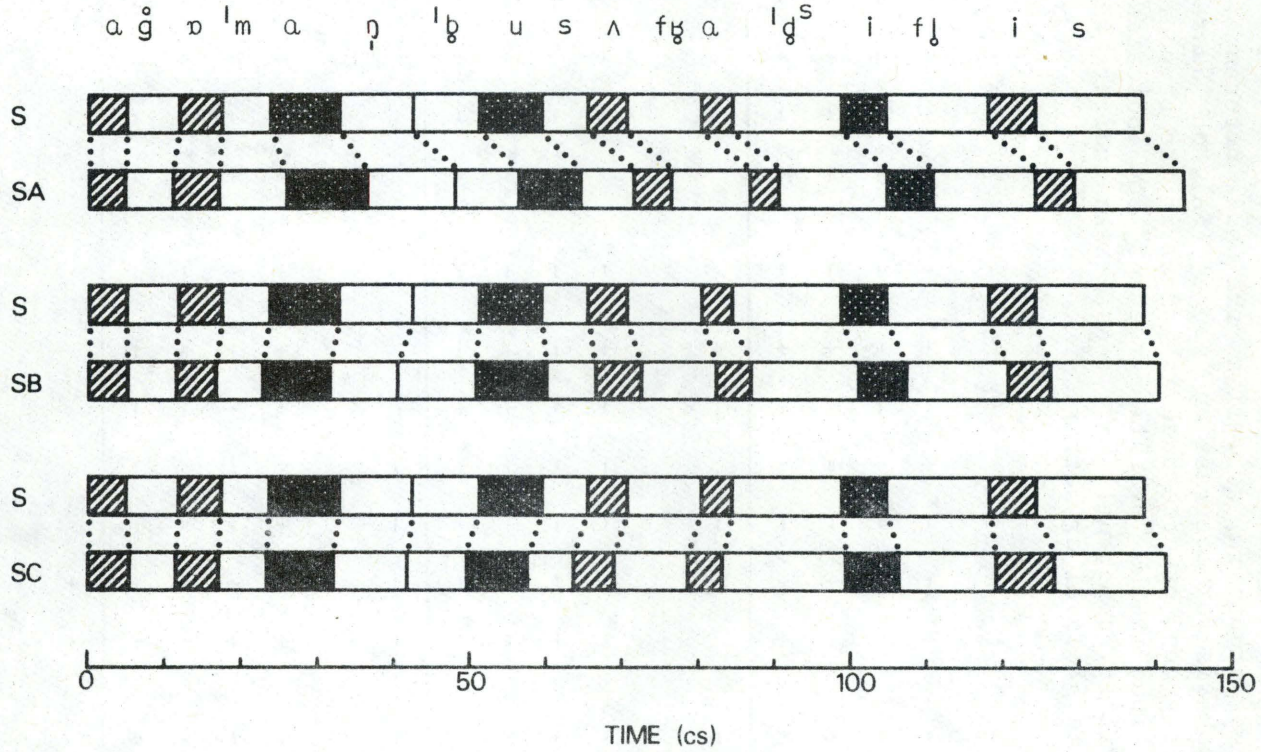


Figure 11

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

Table 4

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

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

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

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

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

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

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

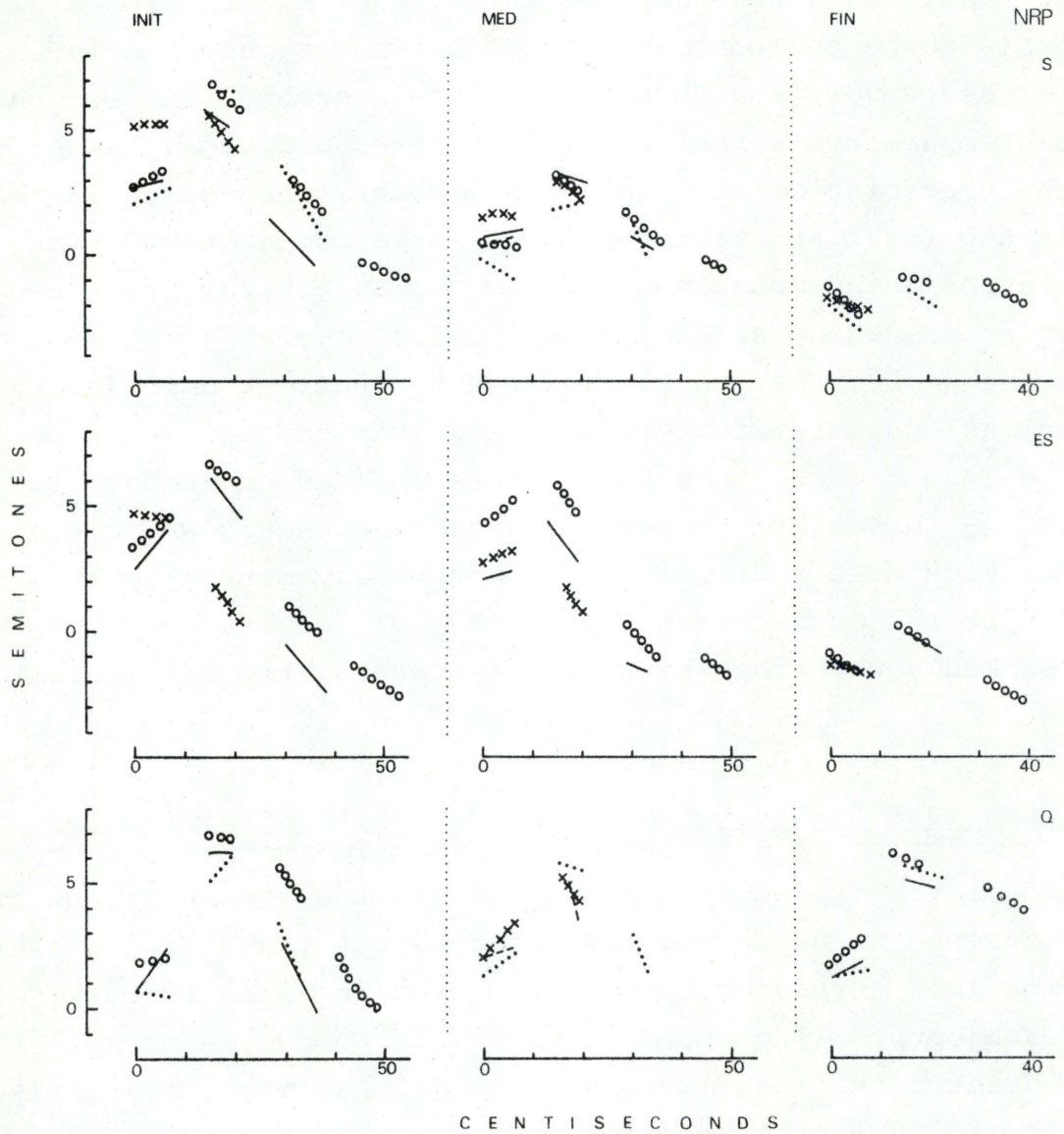


Figure 12

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

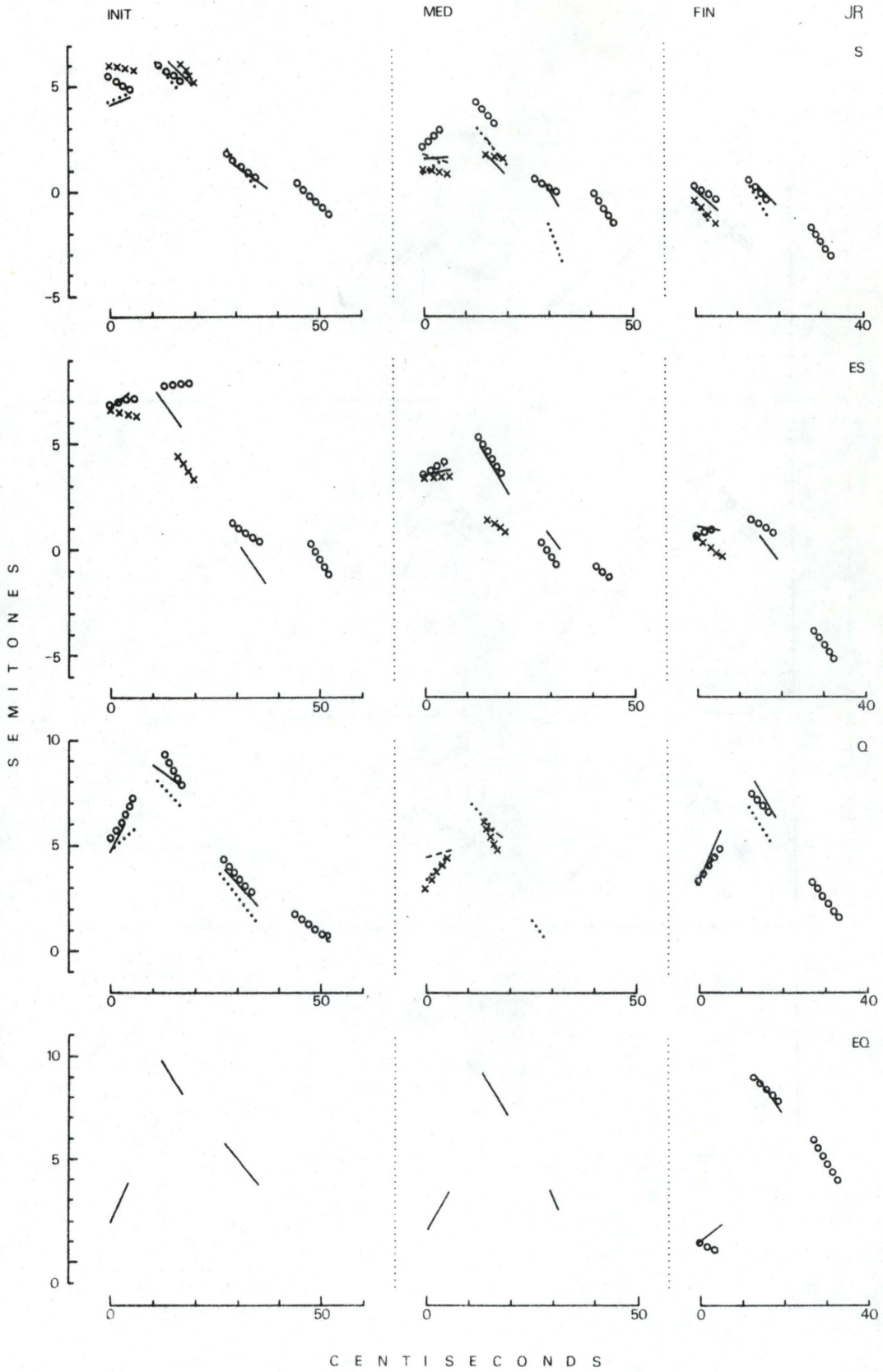


Figure 13

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

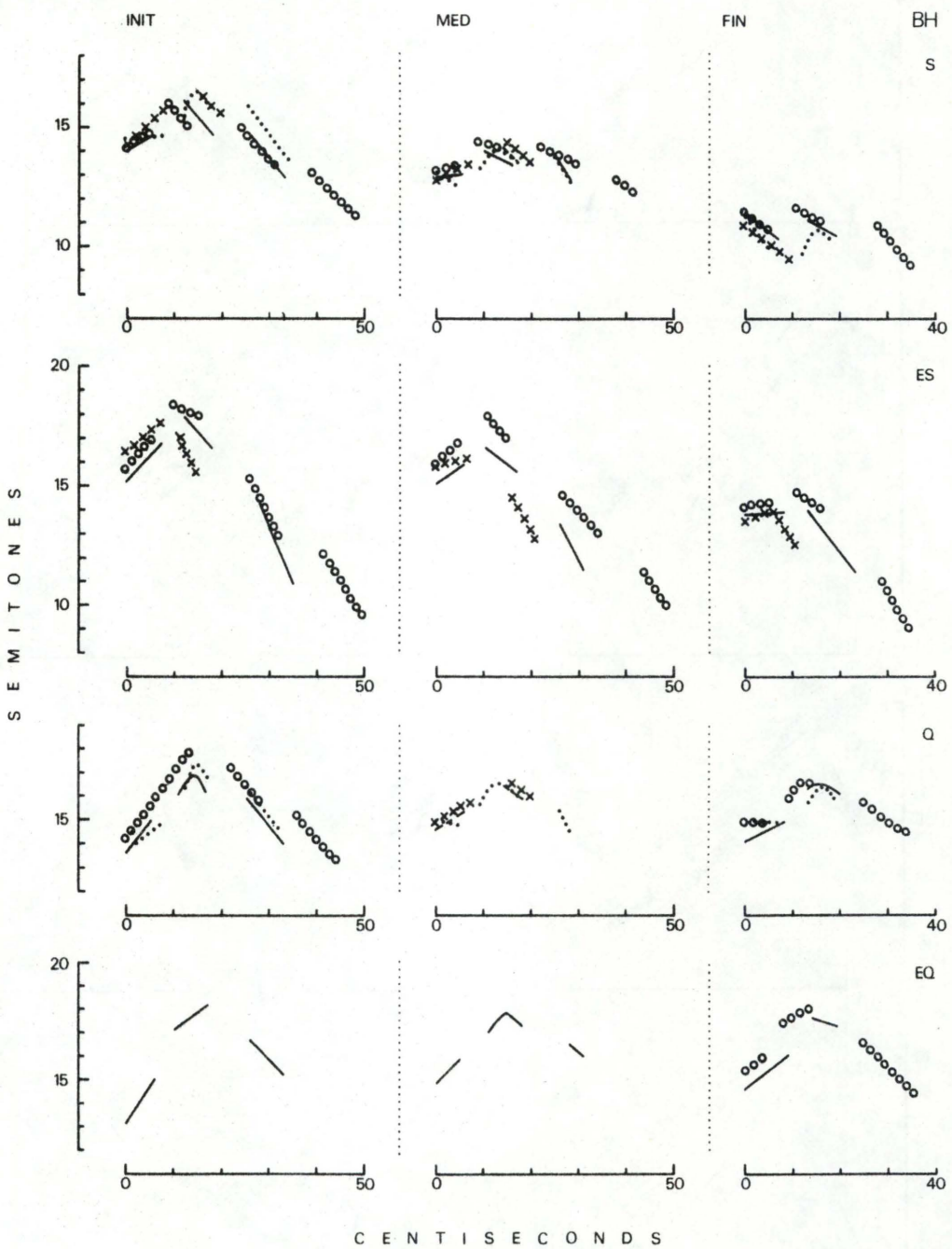


Figure 14

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

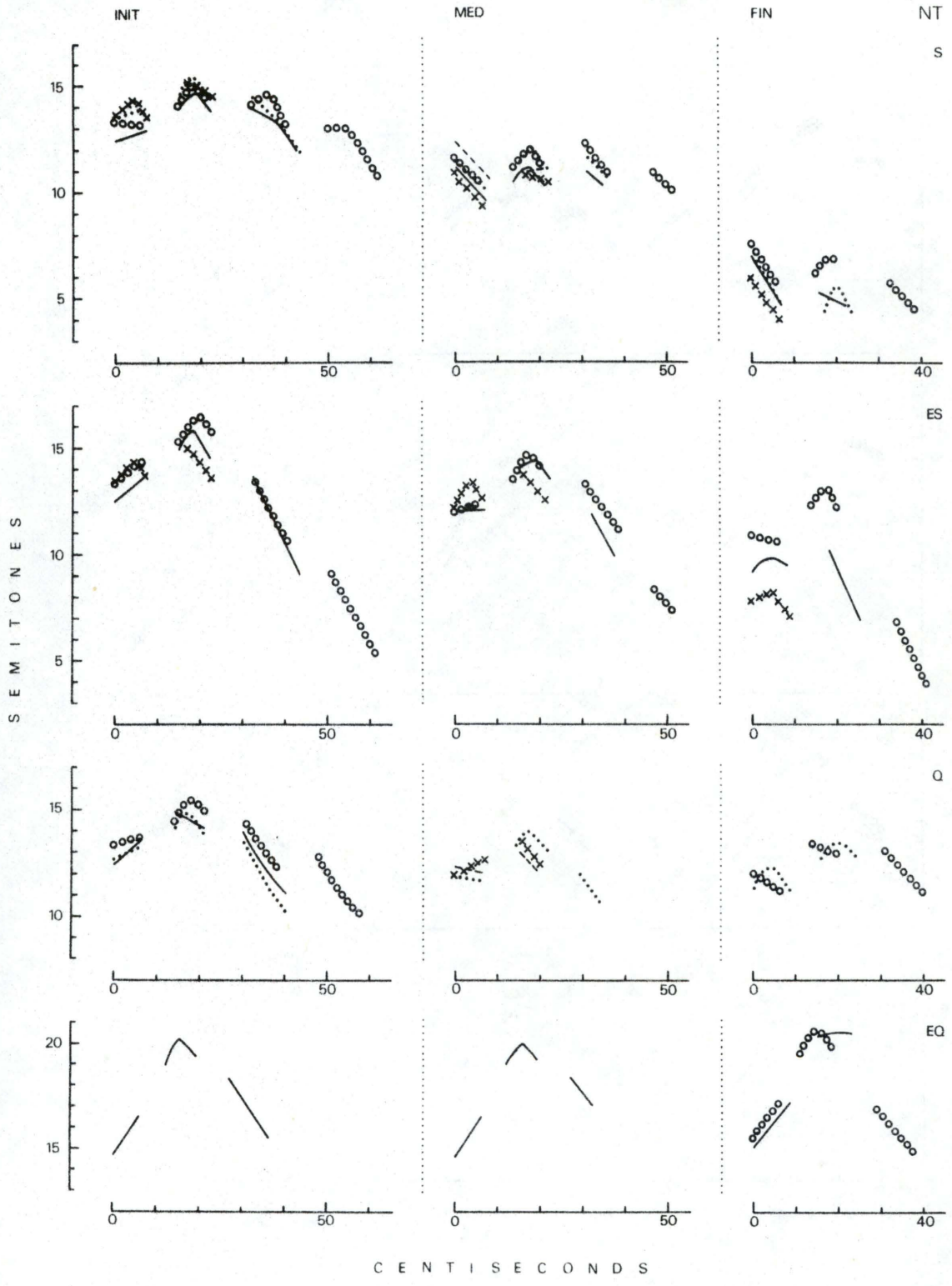


Figure 15

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

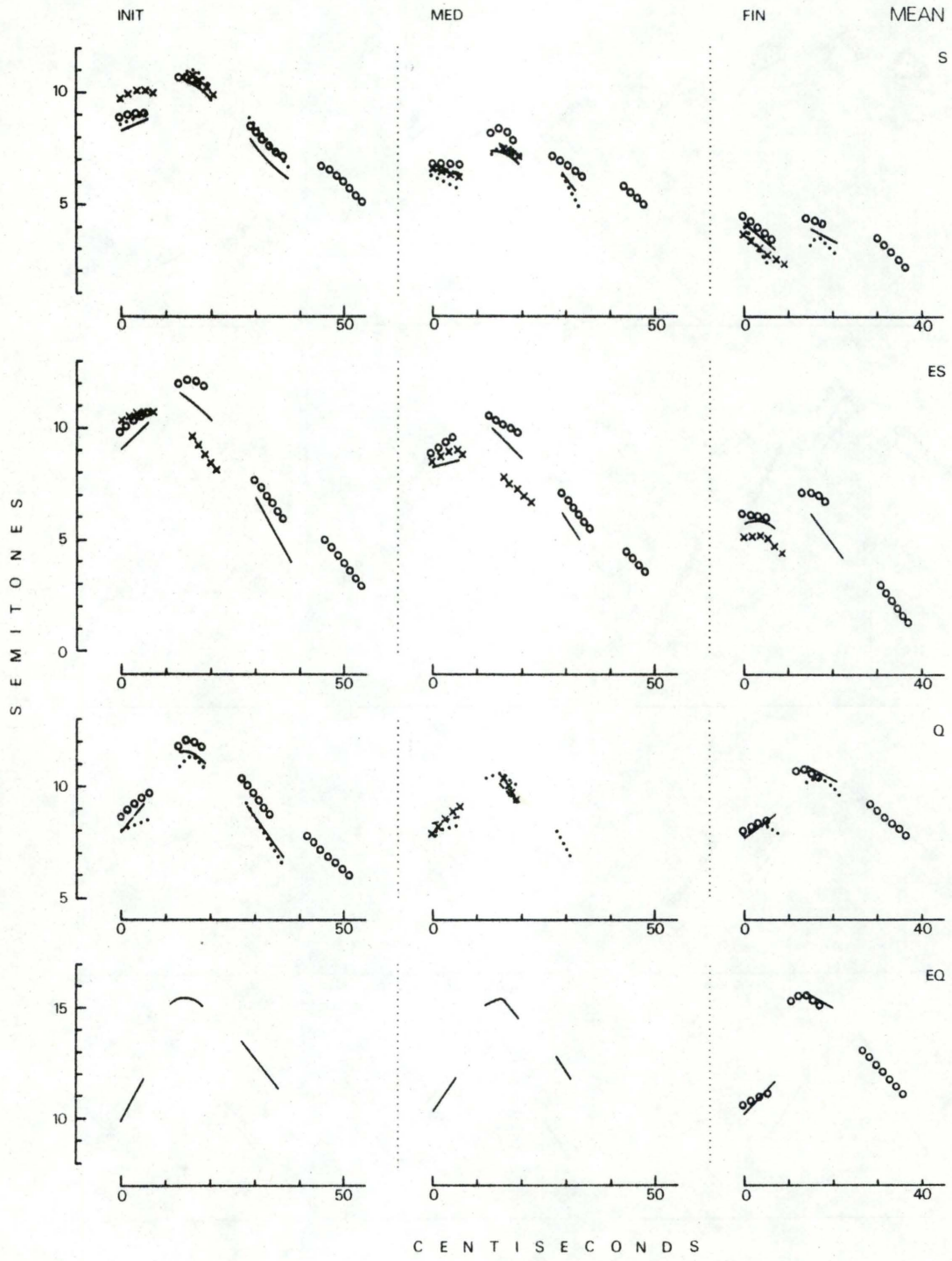


Figure 16

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

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

7.1 The rise to the post-tonic

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

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

---

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

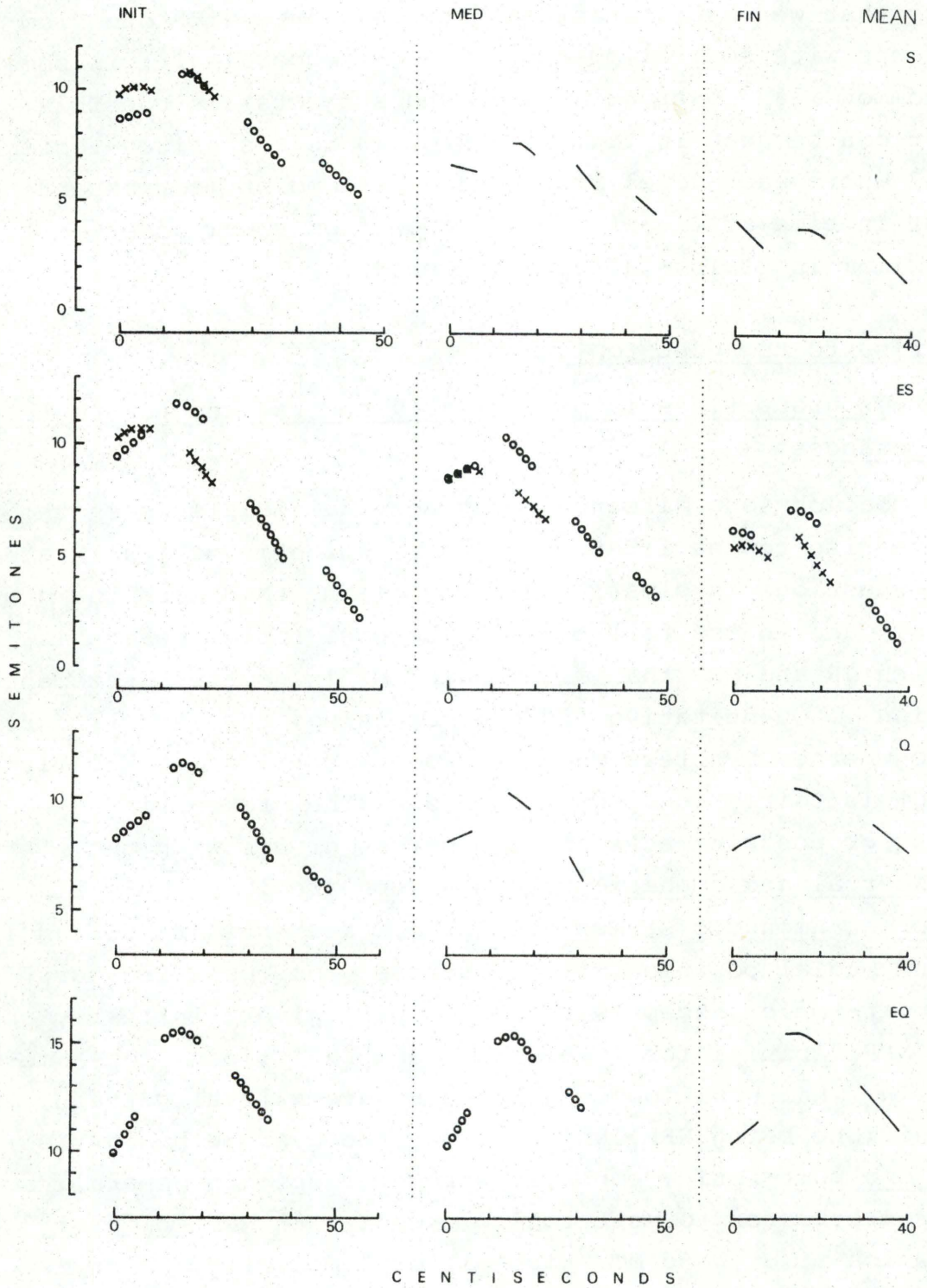


Figure 17

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

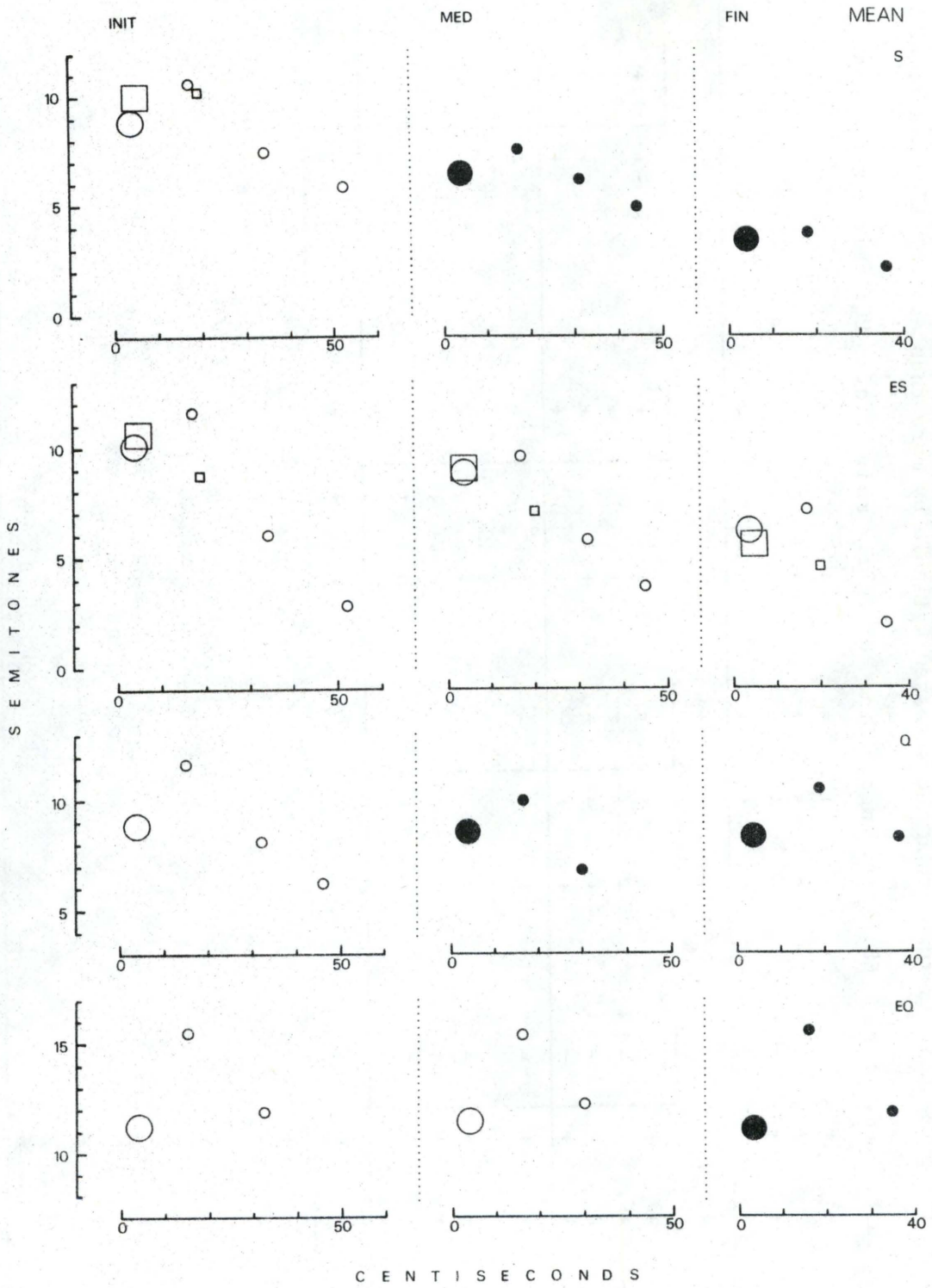


Figure 18

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

Table 5

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

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

ARIPUC 14

e r r a t u m

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

"rise was highest, when the post-tonic constituted a separate word the"

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

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

the three utterances were mixed with utterances that served a different purpose and which were semantically and rhythmically very different, i.e. the three utterances stuck out clearly from the rest of the material. (If the lack of rise in initial stress groups with only one post-tonic were a constant phenomenon I would be hard put to explain why medial stress groups of a similar composition always come up with a rise - at least in materials analysed so far.)

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

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

7.1.2 Stress group rises as a function of intonation contour

Figs. 12-18 confirm the model in fig. 1 as far as higher rises to the post-tonic in neutral questions than in neutral statements go. - In section 2.1 I speculated that this variation in stress group pattern with intonation contour might be a physiologically conditioned assimilation: the lower the succeeding stressed syllable, the less of a rise is performed from stressed to post-tonic, and vice versa. However, if we are dealing with a physiological conditioned constraint, it is not a strong one: The tendency towards higher rises on less falling contours is true only when averages of rather large data bases are considered. I noted at the end of section 6.4.1 that standard deviations on the means of Fo measurements are small, generally less than 5% of the mean (and smaller in stressed than in unstressed vowels). However, when the magnitude of the rise from stressed to post-tonic is calculated and averaged we get standard deviations which are often as great as or greater than the mean. There are two reasons for this: (a) with a dispersion of the same order of magnitude on Fo measurements in the stressed vowel, the post-tonic vowel, and the rise from stressed to post-tonic, we will naturally get larger standard deviations on the mean of the rise, which is numerically much smaller than the means of stressed and post-tonic vowels. (b) Fo of the stressed and post-tonic (and succeeding stressed) vowel vary within rather narrow ranges, but they are not otherwise strongly correlated, i.e. a stressed vowel in the lower end of the range may be succeeded by a post-tonic in the upper end of the range (which may be succeeded by a stressed vowel in the lower end of the range). Documentation for this "production instability" (in sentence final stress groups) can be found in Thorsen (1980b, p. 1019-1021) and has been confirmed by spot checks in the present material. Thus, in individual recordings we may come across rises in statements which are higher than some rises in questions, and vice versa. Another point in disfavour of a strictly physiologically determined variation in stress group rises on falling and horizontal intonation contours is the fact that the final stress group rise is also higher in questions than in statements, although there is no succeeding stressed syllable that can be made responsible for the variation. - On the other hand, to see higher rises on question contours as a controlled signal (added to the informa-

tion contained in the course of the stressed syllables) of interrogative sentence intonation is not very appealing either, because you would then expect a greater production stability in these rises. - A third possibility presents itself, which ties up with the speculations in sections 6.4.2.2(c) and 6.4.2.4, namely that stress group patterns tend to be more distinct (i.e. have higher rises) on marked intonation contours, so in a sense this variation is speaker controlled but the demands for production stability are not severe. (If we assume that prosodic phrase group boundaries - which entail partial resettings in the course of long falling intonation contours (Thorsen 1980a) - receive special attention on the part of the speaker (which is not an unreasonable assumption) then the rises from stressed to post-tonic in stress groups preceding such boundaries, which are comparatively higher than if no boundary succeeds, can be explained along the same lines, i.e. as maximally distinct prosodic patterning.)

The successively lower rises from first through last stress group implied in fig. 1 (cf. section 2.1) is confirmed for the neutral statement - but not for the question, cf. fig. 18, top, and see also figs. 6, 8, and 9: the final rise is too high or, alternatively, the medial rise is too low. The lower medial rise compared to the final rise cannot be due to an artefact of the material because both stress groups in fig. 18 are the average of one stress group with one post-tonic and two stress groups with two post-tonics (times <u>four</u> subjects in both cases as well). Nor can it be due to deviant behaviour on the part of one particular subject because all subjects share this feature, cf. figs. 12-15. - Again, to see the comparatively higher rises <u>finally</u> in questions as an intonation contour signalling, i.e. as an acoustic cue to interrogative intonation contours, is not appealing due to the rather large dispersion in the magnitude of this rise, and because perception experiments show that the terminal rise is poorly correlated with listeners' identification of utterances as declarative, non-final, or interrogative, see further Thorsen (1980b, p. 1019-1021). (This rejection may be supported by the fact that the emphatic questions are not characterized by having comparatively higher rises when the emphasis is in the final stress group. Such a comparatively higher rise would not have been incongruous, because the difference between emphatic statements and questions lies

mainly in the higher course of the "post-tonics" in questions, and with emphasis in initial or medial position there is a rather long tail to carry this high Fo course, but with emphasis in final position the question signalling must be carried by a few syllables only, which might have provoked a rise from emphatic to post-tonic greater than in initial and medial position.) If the extent of the final rise does not correlate with listeners' classifications of utterances as more or less interrogative, it cannot even be regarded as a redundant cue, i.e. it would have no perceptual purpose. So to account for the shape of the "topline" (the line connecting the first post-tonic in each stress group) in prosodically neutral questions it might be more profitable to not regard the final (or medial) stress group as being "aberrant" but to wonder whether "toplines" on marked contours may not be characterized by an asymptotic decline as compared to the linear decline on statement contours, i.e. we get a fall from first to second post-tonic, but after that the declination levels out. - More material, i.e. longer questions, is called for to further illuminate this point.

7.2 Intravocalic Fo movements

7.2.1 Long and short vowels

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

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

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

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

7.2.2 Compression or truncation of Fo movements and -patterns

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

---

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

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

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

7.3 Word boundaries in utterances with emphasis for contrast

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

The conclusion to be drawn is identical to the one in Thorsen (1980a - this volume, at the end of section 5), and it gains in strength, I think, from the results on utterances with emphasis for contrast: The fact that word boundaries (which are also some-

times noun phrase/verb phrase boundaries) leave no trace in the course of fundamental frequency presents an argument in favour of a theory expounded in Selkirk (forthcoming) that prosodic categories (in casu: stress groups) are distinct entities in the phonology that do not have an isomorphous relation to syntactic structure. Rischel (1972) argues in a similar fashion: Danish stress is best represented in a hierarchy (a tree structure) which is not necessarily congruent with the syntactic structure. - (The autonomy of prosodic structure does not, of course, deprive it of a relation to syntax, on the contrary, prosodic categories can be seen as reconciling the syntactic structure to the phonetic output - in casu: the course of fundamental frequency.)

7.4 Stress group patterns - conclusion

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

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

8. Postscriptum about the variety of Danish investigated

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

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

9. Postscriptum about reduced main stresses

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

Acknowledgements

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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