TWO ISSUES IN THE PROSODY OF STANDARD DANISH:
THE LACK OF SENTENCE ACCENT AND THE REPRESENTATION OF SENTENCE INTONATION*

NINA THORSSEN

I. INTRODUCTION

As indicated by the title, this paper falls in two parts: I wish to discuss the lack of an obligatory sentence accent because, to judge from descriptions of more or less closely related languages, Standard Danish is somewhat exceptional in this respect, and I wish to discuss the tonal sequence representation suggested by Pierrehumbert (1980) and by Ladd in his contribution to this volume.

II. SENTENCE ACCENT

Sentence accent, primary accent, sentence stress, nuclear stress, tonic, focal accent, Satzakzent, etc. are more or less synonymous terms which designate the one stressed syllable in a stretch of speech which is more prominent than other stressed syllables.

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 framework 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. 259).

Halliday (1967) rather categorically associates the tonic with 'contrastive' or 'new' in the information structure. Experiments reported in Currie (1980, 1981) and Brown et al. (1980) demonstrate that this claim is far too strong: listeners are not unanimous in their tonic assignment (neither in read nor in spontaneous speech) and (consequently) a perceived tonic does not necessarily fall on the 'new' information in the utterance. Brown et al. (1980, p. 157) suggest that "in sentences read aloud there will generally be some extension of pitch movement on the last stressed syllable. This cannot be taken as marking new information, since all items in the sentence must be new. What it does mark, we suggest, is 'sentence final'", which ties in well with the quotation from Crystal (1969) above.

Bolinger's (1958, 1972) account of (American) English pitch accents does not include the concept of a sentence accent, and at first glance his analysis resembles that of Dutch (cf. below). However, Bolinger (1972) is perhaps rather more categorical in the association between information structure and pitch accents than 't Hart and Collier (1979) and Terken (1980). Furthermore, Bolinger (1958) clearly implies the possibility of more and less prominent pitch accents: "the stress ... that is the most prominent one in the utterances tested." (p. 113), "... the major stress .." (p. 116). At this point, though, I wish to point to a certain difficulty in comparing descriptions across languages and authors because the pragmatic constraints (if any) of the materials analyzed are not always explicated and in any case are rarely identical across investigations.

The existence of a Satzakzent in German, falling on the most important word in the utterance, seems to be uncontested, cf. von Essen (1956), Stock (1980, p. 79-80), and Klein (1980).

Sentence accent is also an undisputed reality in Standard Swedish, cf. Bruce (1977) and Gårding (1980). Bruce (1977) is not categorical in the association of sentence accent with 'new' information, but states (p. 21): "It seems to me that the Functional Sentence Perspective – the theme/rheme distinction – can account for a great deal of what has conventionally been attributed to emphasis and contrast." According to Carlson et al. (1974), prosodically neutral utterances, i.e. utterances with a non-contrastive (focus-free) stress pattern are possible (p. 212). (Note that they seemingly equate contrast and focus.)

The relation between sentence accent and information structure may not be straightforward and there may also be a marked difference between planned and spontaneous speech with respect to the number and placement of sentence accents. I do not think that it stretches the descriptions of Standard Swedish, English and German too far, however, to conclude that in pragmatically
neutral speech the last accented syllable in the phrase (or whichever the relevant chunk of speech may be) will normally be more prominent than preceding stressed syllables, a prominence which is realized by a greater duration of and a more elaborate fundamental frequency (F0) movement within or in the immediate environment of that syllable. Under different circumstances, the context may evoke an earlier location of this prominence, for instance under conditions of contrast.

In Standard Danish, in pragmatically and emotionally neutral speech (as for instance in the reading of context-free utterances), NONE of the stressed syllables is more prominent than the others. Such prominence is not evident in acoustic registrations (in duration and/or F0 contour) nor is it present auditorily. There is nothing incomplete about such prosodically neutral utterances and although they may not be very frequent in spontaneous speech, they certainly do occur, they are not unnatural, and they are very easy to elicit from speakers in a reading situation. I should add that this is a completely uncontroversial fact among Danish scholars. It appears, then, that three degrees will suffice to account adequately for the distribution of stresses in Standard Danish: main stress (assigned to the lexically stressed syllables in most non-function words), secondary stress (assigned to the second - and following - lexically stressed syllables in compounds), and weak stress, see further Basbøll (1978), Heger (1981, p. 118-119, 122), and Rischel (1972, 1975, 1980, forthcoming). In similar utterances in Standard Swedish, English, and German a fourth degree is required: the sentence accent.

In Thorsen (1980a) I report the results of acoustic analyses of emphasis for contrast in Standard Danish, which may be summarized briefly as follows: When emphasis for contrast is introduced in a sentence, the stressed syllable of the emphasized word will stand out clearly from the surroundings; this is brought about by a raising of F0 (except in initial position), an elaborate F0 rise within that syllable and a deletion of the F0 deflections (the rises from the stressed syllables) 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, see figure 1. (Statements and syntactically unmarked questions differ with respect to the extent of the influence of emphasis on the prosodic patterns: F0 patterns two stress groups away from the emphasized word retain - at least partly - their rises to the post-tonic in statements but not in questions. It may 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, i.e. terminal declarative sentences.) For the sake of clarity I will disregard the limited influence of emphasis in terminal declarative sentences and conclude: Short utterances with emphasis for contrast reduce tonally to one prosodic stress group in the sense that only one LOW + HIGH-FALLING pattern occurs in them. The difference between statements and questions with contrast emphasis
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. Zero on the logarithmic frequency scale corresponds to 100 Hz. See further Thorsen (1980a).
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.

Clearly, then, Danish contrast emphasis shares some features with the acoustic manifestation of other languages' sentence accent. It is difficult to carry this comparison much further on a general basis, because there are differences between English, German and Standard Swedish: Satzakzent in German is not signalled by any special Fo movement within the stressed syllable (except if it is the last one in the utterance), but rather by what happens after it (cf. von Essen 1956, Isačenko and Schädlich 1970); word accent differences are not deleted in non-focal position in Swedish (Bruce, 1977). Furthermore, as I hinted above, some instances of sentence accent presumably fall in the category 'contrast', but the semantic/pragmatic distinction (to the extent that it can be made) and the phonetic differences (if any) between contrast and other types of prominence are rarely made explicit, with one notable exception: From Jones (1960, §§ 1049-1059) we learn that when contrast emphasis is superposed upon a sentence accent in utterance-final position, 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, his readers 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, according to Jones' description, we can distinguish phonetically between a neutral sentence accent and contrast emphasis in utterance-final position but not elsewhere.

The fact that non-final contrastive and non-contrastive sentence accents may not be distinguished phonetically and the fact that the manifestation of emphasis for contrast in Standard Danish has certain features in common with such non-final sentence accents does not of course affect my argument that Standard Danish lacks an obligatory sentence accent, because the test case remains: in pragmatically neutral utterances no special prominence is attached to the last (or any other) stressed syllable.

According to Bruce (personal communication) Southern Swedish also lacks an obligatory sentence accent, so Standard Danish is not a completely isolated case among the Nordic languages. Furthermore, Dutch also seems to defy the sentence accent analysis, but in a different manner: One or more of the lexical stresses (but not necessarily all of them) will be manifested as pitch accents, and these pitch accents are apparently all equally prominent, cf. for instance 't Hart and Collier (1979). The pragmatic rules which govern the assignment of pitch accents in Dutch are the object of a study by Terken (1980).
I am of course not blind to the fact that in spontaneous speech in Danish we certainly get uneven stress distributions, i.e. varying degrees of prominence among the stressed syllables of an utterance, without necessarily evoking the impression of emphasis for contrast. There are several reasons why I do not think this kind of emphasis is a sentence accent phenomenon: (1) There may be any number of emphasized words within one prosodic phrase. (2) If I can trust my impression from just listening, such non-contrastive emphases do not lead to a deletion of the neighbouring Fo deflections, so (3) their prominence derives from increased duration and probably intensity rather than from a stress reduction in the surroundings. Furthermore, a distinction between sentence accent and other types of prominence (emphasis) is probably also required in spontaneous speech in sentence accent languages, cf. Brown et al. (1980).

I do not know what the general psycholinguistic significance of the missing sentence accent in Danish would be, but I am certain that it is - or should be - an important feature in language teaching, both ways: Danes who omit the sentence accent in, say, English - in styles of speech where it would have been appropriate, for instance reading a scientific paper to an audience - are reported by speakers of English to sound dull, dead-pan and uninterested in their own subject. On the other hand, a Swede speaking Danish is likely to be labelled "too emphatic", "affected", etc.

III. THE REPRESENTATION OF SENTENCE INTONATION

A. THE TONE SEQUENCE APPROACH

I have previously presented data and arguments to support a view of the composition of Standard Danish intonation in terms of a layered system of simultaneous, non-categorical components, see further Thorsen (1979, 1980a, 1980b, and forthcoming). I have also argued that although the stress group and sentence intonation components are highly interactive on the concrete articulatory and physical level (where the stress group pattern is subject to quantitative variation), there may be a level in the speech production and perception processes where the two can be viewed as invariant entities, see further Thorsen (1980a). Figure 2 illustrates three stages in the superposition process, and in Thorsen (1979) an example is given of the inverse process, i.e. the decomposition of the fundamental frequency course in an utterance into its constituent components. Figure 3 summarizes the results of analyses of stress group patterns and intonation contours in short utterances.

Pierrehumbert (1980) takes a radically different approach. (Her thesis abstract is quoted verbatim in the appendix.) She suggests that Standard Danish be re-analyzed in a fashion similar to her analysis of English - in terms of a sequence of pitch accents, consisting either of mono-tonal H or L accents or bi-tonal combinations of the two (connected with "+"
Illustration of the superposition principle in the production of fundamental frequency in an utterance *Ammerne i Alabama var i strejke*. (The nurses in Alabama were on strike.): Three stress group patterns of different length, superposed upon a declining sentence intonation contour, will - by way of adjustment rules and microprosodic rules - yield the Fo contour shown (spoken here by a female speaker).
A model for the course of fundamental frequency in short sentences in Standard Copenhagen Danish. 1: syntactically unmarked questions, 2: interrogative utterances with word order inversion and/or interrogative particle; non-terminal declarative and interrogative sentences (variable), 3: terminal declarative utterances. 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.

in the transcriptions). In bi-tonal pitch accents, either of the tones may be the stronger one, lining up with the stressed syllable, leading or lagging behind a weaker tone. Tones aligned with stressed syllables are denoted with a star. Figure 4 is an example (from a personal communication) of the tonal representation of an utterance with different prominence relations between the two stressed syllables.

Pierrehumbert observes that in a number of terminal declarative sentences, the degree of overall downdrift is too large to be accounted for by the slight, physiologically determined, baseline declination. In order to keep within a framework of categorically different intonational units which are non-interacting and simply follow each other in time, and so maintain a description of intonation in terms of tonal sequences which together make up or, rather, which ARE the intonation of the utterance, Pierrehumbert introduces a RULE OF DOWNSTEP which will lower a H tone in the context H+L_ and H L+, independently of the relative prominence (starring) of the tones (compare the top of figure 5 to the lower tracings). Overall downdrift is a result, then, of contextually determined downsteps which are completely locally governed. The actual scaling of a pitch accent requires no look-ahead and is totally
Figure 4

Fundamental frequency tracings and tonal representation of an utterance Anna came with Manny with different prominence relations between the two stressed syllables: Anna is 'backgrounded' in the upper tracing, Manny is 'backgrounded' in the lower one. From Pierrehumbert (personal communication).
Three examples of fundamental frequency contours and their tonal representation: declination of H* pitch accents (top), downstep of H+L* pitch accents (mid), and downstep of L*+H pitch accents (bottom). From Pierrehumbert (personal communication).
specified by its tonal type, its prominence, the baseline value, and by the immediately preceding tone. Further, accents are downstepped through the utterance by a constant factor, smaller than unity, relative to the preceding H, which creates asymptotically declining contours.

![Semitones Graph](chart)

**Figure 6**

Model for the course of fundamental frequency in a short terminal declarative utterance in Standard Danish and its tonal representation, as suggested by Pierrehumbert (1980). See further the legend to figure 3.

Confronted with the Danish data, particularly the low-plus-high-falling stress group pattern, Pierrehumbert faces what looks like an obstacle on the surface. In her system, Standard Danish has a bi-tonal L*+H accent, cf. figure 6. If trional accents are to be avoided, there is no room for a L after the H in the Standard Danish pitch accent, i.e., there is no context for the downstep rule. Pierrehumbert circumvents this problem by positing a downstep rule which applies in declaratives in Danish and which downsteps L* in relation to the preceding H. She bases this on an observation from my model (figure 3 here) that the L*+H interval is constant, whereas the H L* interval varies with degree of overall downshift; in other words, the L*+H interval is implemented in a constant way, whereas a L* dissipates from the preceding H to varying degrees in different types of utterances. With a rule which downsteps L* in relation to the preceding H by a factor, k, which varies according to the status of the sentence, all Danish sentence intonation contours are the product of locally determined pitch accent scaling. (Pierrehumbert's first observation, that the L*+H interval is constant, is not ac-
curate. It may be true at a more abstract level of production and perception, but not at the level where the actual physical scaling of Fo is taking place, and that is the level Pierre-humbert is dealing with where the downstream rule is concerned.)

B. PROBLEMS WITH THE TONE SEQUENCE APPROACH IN STANDARD DANISH

(i) A general point of difficulty with the Tonal Sequence approach is that it seems to operate with an open scale of degree of prominence (cf. Pierrehumbert 1980, Chapter 3 in particular). It demands of the speaker and listener (and analyst) that they can reliably and consistently produce and identify a large number of prominence degrees. If such an ability to identify highly varied prominences cannot be assumed with all speakers and hearers, the Tone Sequence theory faces a serious obstacle.

(ii) Pierrehumbert's downstream rule cannot account for the way a sequence of two successive stressed syllables (i.e. with no intervening unstressed syllables) are scaled in, say, a terminal declarative, let alone a whole utterance like Per så Lis [ˈbɛə ˈsɔː ˈlis] (Per saw Lis), which is just as downdrifting/downstepping as utterances with unstressed material between the stressed syllables. I do not think that introducing an abstract or underlying H after the L* can solve the problem, since I do not see how an abstract tone can be a factor in the actual computation of the scaling of a succeeding tone. This difficulty with the downstream rule would disappear if L* could be downstepped in relation to the previous L*. That would be much better motivated phonetically and would also remove the difficulty that the variation in the magnitude of the rise from stressed to first post-tonic (the L*+H interval) creates. However, I suppose that since the previous L* is not always IMMEDIATELY preceding, it is unacceptable as a general context for the downstream rule to apply in, because it will disrupt the strict locality principle. This illuminates a fundamental problem with the Tone Sequence approach: Unstressed syllables after the first post-tonic (H) must get their Fo course by a phonetic fill-in rule. That is in itself reasonable enough, also in the light of my own speculations that stress group patterns may be invariant entities at a higher level in production than the concrete articulate one. However, in order to scale the succeeding L* stressed syllable, these intermediate unstressed syllables must be disregarded; in other words, the speaker must employ some kind of phonetic 'look-back' mechanism to locate the H which - together with other factors - determines the actual frequency value of the next L*. The only reason why this phonetic mechanism might not look back one more syllable, to the preceding stressed (i.e. L*) one, is the constraint that the tonal representation and the locality principle impose. That is to say, the distinction between those features of a speaker's output which are relevant to his ongoing production and those which are not is NOT independently established but is a consequence of the theoretical framework itself. In this particular case it will prevent a phonetically
more plausible downstep rule (L* is downstepped after and scaled in relation to L* rather than H), and one which would also reflect the fact that the stressed syllables are the relevant ones in a listener's identification of intonation contours, cf. Thorsen (1980b).

Ladd (in press, and this volume) modifies Pierrehumbert's theory while staying within its general framework. He finds her account of (Danish) sentence intonation contours unconvincing and introduces a downstep FEATURE, which "gives separate representation to the overall downward slope, treating it as an independently selectable phonological phenomenon without giving up the advantages of TS [tonal sequence] phonetic specification." - It seems to me that this has obvious advantages over the downstep rule, and it also solves the problem of environment mentioned above. Ladd suggests that the downstep feature be carried over to the analysis of English and together with some other modifications he achieves a rather less abstract tonal representation.

Independently of the choice of either a downstep rule and its environment or a downstep feature, there are, however, still certain points which make the Tone Sequence representation descriptively inadequate for Standard Danish:

(iii) If an intonation contour (or phrase contour) is NOT asymptotically declining, and if this is NOT due to uneven prominence of the stressed syllables, then the downstep factor or feature must of necessity take different values as it creeps along the utterance. Figure 7 offers evidence for non-asymptotically declining phrase and intonation contours; see further Thorsen (forthcoming).

(iv) Consider utterances of varying length: In the conceptually simplest case where range is constant over utterances of different length and where consequently the slope of the intonation contour (the downstep factor) is inversely proportional to the length of the utterance it spans, the downstep must be factored differently in short and long utterances, something which presumably demands a certain amount of look-ahead for its computation. In more complicated cases where range increases with utterance length, but not linearly so, the demands on the computation of the manifestation of the downstep are still harder; see Thorsen (forthcoming).

(v) Apart from their descriptive adequacy, the two representations may be evaluated in terms of their ability to reflect speech production and perception processes, to the extent that such processes are known, or at least hypothesized. The Tone Sequence theory explicitly requires no look-ahead mechanism, but we know that speakers do look ahead (see e.g. Perkell, 1980) also in the planning and execution of intonational phenomena, cf. Bruce (1981), Lehiste (1975). This fact is mirrored in the global representation of sentence intonation in the layered system. - The stressed syllables (the starred tones) in the Tone Sequence theory have no special role or
Figure 7

Intonation contours (full/broken lines) and stress group patterns (dotted lines) in two sets of terminal declarative sentences containing from one to eight stress groups. Average over four subjects (mean of means over six recordings). Large dots represent stressed syllables, small dots unstressed ones. The broken lines indicate boundaries between prosodic phrases, see further Thorsen (forthcoming). Zero on the logarithmic frequency scale corresponds to 100 Hz.
status in connection with the determination of overall down-
drift (downstep), cf. the formulation of the downstep rule
above - but at least in Standard Danish, the stressed syl-
lables seem to serve as anchor-points in the identification
of intonation contours and this fact is also reflected in the
layered system.

To conclude: If the downstep rule or feature is to capture
the amount of variability and complexity in intonation con-
tours like those presented in figure 7, it must be a scalar
and continuous feature which will often have to take differ-
et values within one and the same phrase or sentence. I.e.
it must have at least some of the properties it presumably
was intended to rid intonation analyses of. - It seems to me
that the Standard Danish data are more easily accommodated
by a descriptive system or theory that requires a look-ahead
mechanism and which allows its components to be parametric,
simultaneous and physically interacting than by a theory whose
components are sequential and categorical and generated without
look-ahead from left-to-right.

REFERENCES

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

Word 14, p. 109-149.

Bolinger, D.L. 1972: "Accent is predictable if you're a mind-

Intonation, (Croom-Helm, London).

Bruce, G. 1977: Swedish word accents in sentence perspective,
Travaux de l'Institut de Linguistique de Lund 12,
(Gleerup, Lund).

Bruce, G. 1981: "Om prosodisk koping i svenska", Svenskans
Beskrivning 12, Acta Universitatis Upsensis, Umeå Studies
in the Humanities 37, p. 64-78.

Carlson, R., Erikson, Y., Granström, B., Lindblom, B. and
Rapp, K. 1974: "Neutral and emphatic stress patterns in
Swedish", Preprints from Speech Communication Seminar,

Crystal, D. 1969: Prosodic Systems and Intonation in English,

Currie, K.L. 1980: "An initial "search for tonics"", Language


Ladd, D.R. in press: "Phonological features of intonational features".


Rischel, J. forthcoming: "On unit accentuation - and the distinction between deep and surface phonology", Folia Linguistica 17.


APPENDIX

Below is quoted verbatim the abstract from Janet Breckenridge Pierrehumbert's doctoral thesis The Phonology and Phonetics of English Intonation:

"This thesis develops a system of underlying representation for English intonation. It gives an account of what different tunes are possible and how they are aligned with different texts. It characterizes the rules which map the underlying representations into phonetic realizations.

The different tunes are described as structured strings of L and H tones generated by a finite-state grammar. The strings consist of one or more pitch accents, which are aligned with stressed syllables on the basis of the metrical pattern of the text, plus two additional tones which characterize the intonation of the end of the phrase. The pitch accents are either a tone, or a pair of tones on which a strength relation is defined. The two additional tones are the boundary tone, found at the end of the phrase regardless of the metrical structure of the text, and the phrase accent, which follows immediately after the pitch accent on the main phrase stress and controls the intonation from there to the boundary.

Local context-sensitive rules map the string of tones into the quantitative values which determine the fundamental frequency contour. These rules apply left to right, and include downstep and upstep rules resembling those which have been studied in African tone languages. A transform of the fundamental frequency domain which makes these rules linear is proposed on the basis of experimental data. Evidence is presented that superficially nonlocal intonational characteristics, such as the overall trend of the contour, really arise from local rules. The thesis also reviews other experimental results and explains how they are accommodated within the framework proposed."