

TEXT AND INTONATION - A CASE STUDY*

NINA GRØNNUM THORSEN

This is a report about a study of intonation in a text which was not designed for intonation analysis, and which exists in one edition only. The text is the tape-recorded instruction to a language laboratory exercise which, incidentally, rehearses tag-questions in British English. It is a sample of self-sufficient, self-explanatory, non-fictive prose, unique in its kind in terms of human communication, though very familiar. It was recorded once, by the author of the programme, MG.

This study represents my first, cautious step away from the laboratory on the way to spontaneous speech. The need for investigations of intonation in spontaneous speech is obvious. There are basic research problems involved, and the results may be applied in teaching, as well as in synthesis and automatic recognition of speech.

Appendix I is a copy of the first page (of a total of five) of my transcript of the recorded text. I have added stress-marks to the orthographic representation and indicated the duration of pauses as well. Occasional glottal attacks are also marked.

The list of questions which you may reasonably expect the analysis of this and similar texts to answer is fairly long. For reasons of time and space I shall limit myself here to answer two of them, which both test the validity of aspects of an intonation model for isolated, read sentences (see Thorsen 1978, 1979, 1980a, 1980b, 1982a, 1982b, 1983a, 1983b, 1984, 1985, forthcoming):

- 1) Will the tonal pattern associated with the prosodic stress group occur as regularly and in the same shape? I.e. will

*) *Paper read at the IVth Nordic Prosody Meeting in Middelfart, June 6-8, 1986. To appear in the proceedings.*

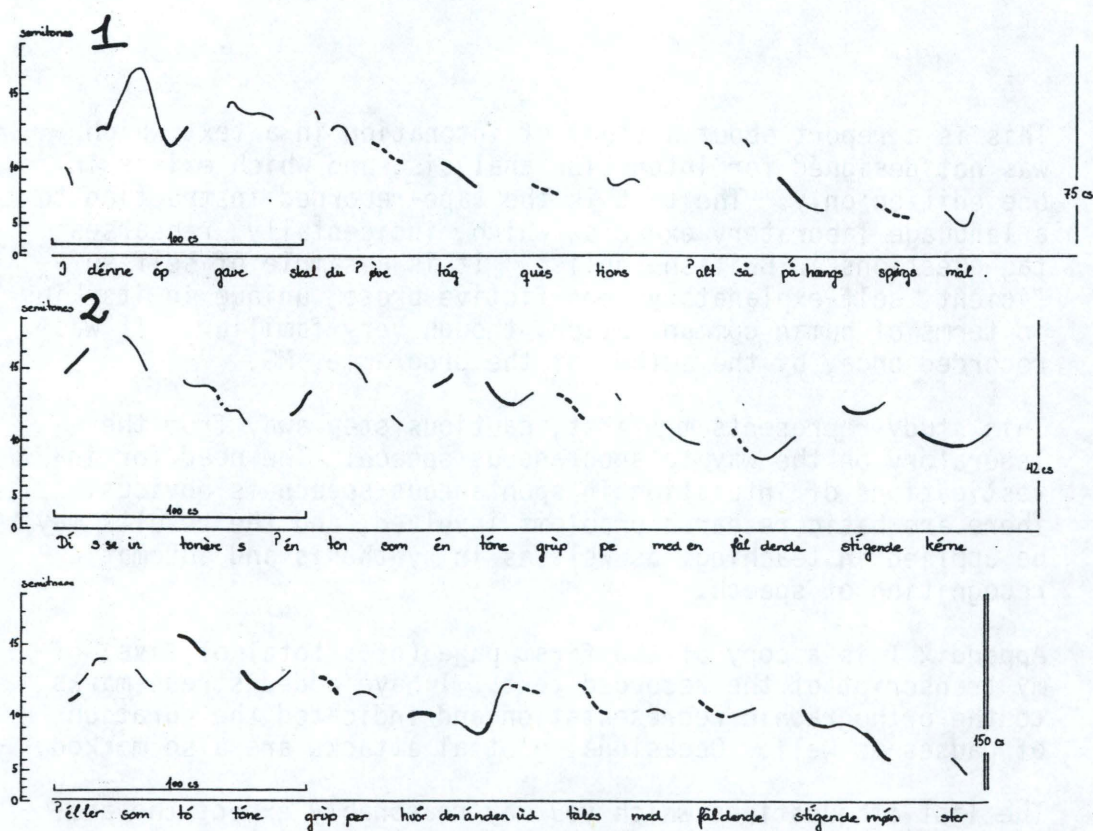


Figure 1

Fo tracings of the first two sentences in the text. The stressed vowels have been touched up with thicker lines. Syllables with secondary stress are indicated with broken lines. Vertical bars indicate pauses, with pause duration in centiseconds added. Zero on the logarithmic frequency scale corresponds to 100 Hz.

any succession of a stressed plus all following unstressed syllables describe a low plus high-falling fundamental frequency (F_0) pattern?

- 2) Does the magnitude of the rise from low stressed to high post-tonic decrease gradually through the sentence or phrase, as it does in pragmatically neutral, read speech?

Before I proceed to answer these two questions I wish to say that I have found this a rather difficult project, because it is hard to separate chance variability from systematic variation, and it is difficult to know when there are no more interesting observations to be wrung from the data.

The tape was processed with hard-ware intensity and F_0 meters (F-J Electronics) and registered on an Elema 800 mingograph, at a paper speed of 100 mm/sec, to yield the standard oscillogram, intensity and pitch curves. The F_0 tracings were transferred to tracing paper without any further processing. Figure 1 is an example - of the first two sentences in the text.

First a little statistical information to characterize the text. There are 60 sentences in the material. I have defined a sentence here operationally, as any passage which appeared between periods in MG's own text from which she read aloud and recorded the tape. These 60 sentences consist of 111 clauses. On the basis of visual inspection of the curves AND the auditory impression I have counted 92 intonational phrases. An intonational phrase does not necessarily coincide with a sentence or a clause, as can be seen from these figures.

Sentence boundaries are always accompanied by pauses whose duration average 104 centiseconds (with a very large standard deviation of 34.3 cs). There are 49 sentence internal clause boundaries, 26 of which are accompanied by a pause, with a mean duration of 46 cs (SD = 19.7 cs). Despite the large standard deviations, the difference in sentence and clause boundary induced pauses is statistically significant beyond the 0.0005 level. There are 14 pauses which are not associated with any major syntactic boundary. Their average duration is 21 cs (SD = 2.3 cs). The difference between these pauses and clause boundary pauses is statistically significant at the 0.025 level. - There is thus a clear tendency that heavier boundaries are associated with longer pauses.

There are 421 stressed syllables in the text, which I have determined by listening. That makes 7 stressed syllables per sentence, on the average (with a minimum of 2 and a maximum of 26). There are 111 clauses, i.e. 3.8 stressed syllables per clause (with a minimum of 1 and a maximum of 10). There are 92 intonational phrases, i.e. 4.6 stressed syllables per phrase (minimum: 1, maximum: 11). All this is probably characteristic of the text as a sample of speech read from a manuscript. Spontaneous speech will hardly turn up with as many complete sentences, nor with so many long clauses, sentences, and intonational phrases.

Now to the question of stress group patterns. This question derives from my previous analyses of Danish, based on fairly short sentences, read out of context - typical laboratory material and speech. A prosodic stress group is associated with a pitch pattern which in Standard Danish can be characterized as low plus high-falling. The stressed syllable is succeeded by a rise in F_0 which normally reaches its peak during the first post-tonic syllable, see the full lines in figure 2.

There are 421 prosodic stress groups in the text. 80 of those are monosyllabic, i.e. they consist only of a stressed syllable with no post-tonics. Those are not expected to have any rising-falling F_0 pattern (cf. Thorsen 1984). The remaining 341 stress groups are polysyllabic. Only 32 of those, or 9%, do NOT have a rise in F_0 after the stressed syllable. It appears justified to conclude that a sequence of stressed plus unstressed syllables is normally - in this type of text - accompanied by a rise in F_0 .

The magnitude of the F_0 rise from stressed to post-tonic syllable varies a good deal, and I cannot assert that it decreases in a monotonous and simple fashion through the phrase or utterance, as it does in short, context-free sentences (cf. figure 2). Nor can I find any other systematic trend in the variation of the low-to-high interval. I venture to guess that varying the magnitude of the pitch rise from the stressed syllable - as this speaker does - serves pragmatic purposes, in that it lends slightly varying degrees of prominence to the stressed syllables, without their falling out of the range of what may be termed neutral, non-emphatic, main stress. Thus, greater prominence would be associated with a higher rise after the stressed syllable.

It is important to note that rise in pitch from a stressed syllable to the post-tonic is not the only cue to its stressedness. (This is evident in the case of monosyllabic stress groups, where there are no post-tonics to carry such a rise.) Instead, I propose, the relation to surrounding stressed syllables may determine a syllable as stressed. That is, if a syllable falls into place in a rhythmic pattern and an intonational contour, established by preceding and/or succeeding stressed syllables, and if its duration, its vowel quality, and its lack of stød do not contradict it, such a syllable will be perceived as stressed. It may further warrant a characterization as ACCENTED if it is succeeded by one or more post-tonics and has an associated pitch rise; if there is no such pitch rise to post-tonics the stressed syllable is NON-ACCENTED. The distinction between accented and non-accented stressed syllables does not apply to monosyllabic stress groups.

The distinction between stress and accent, as outlined above, would bring Standard Danish on a par with German, as described by Bannert (1985), with British English, as described e.g. by O'Connor and Arnold (1961), and with Dutch, as described by 't Hart and Collier (1979), where a similar distinction can be made between stressed syllables which are tonally prominent (accented) and stressed syllables which are not (non-accented).

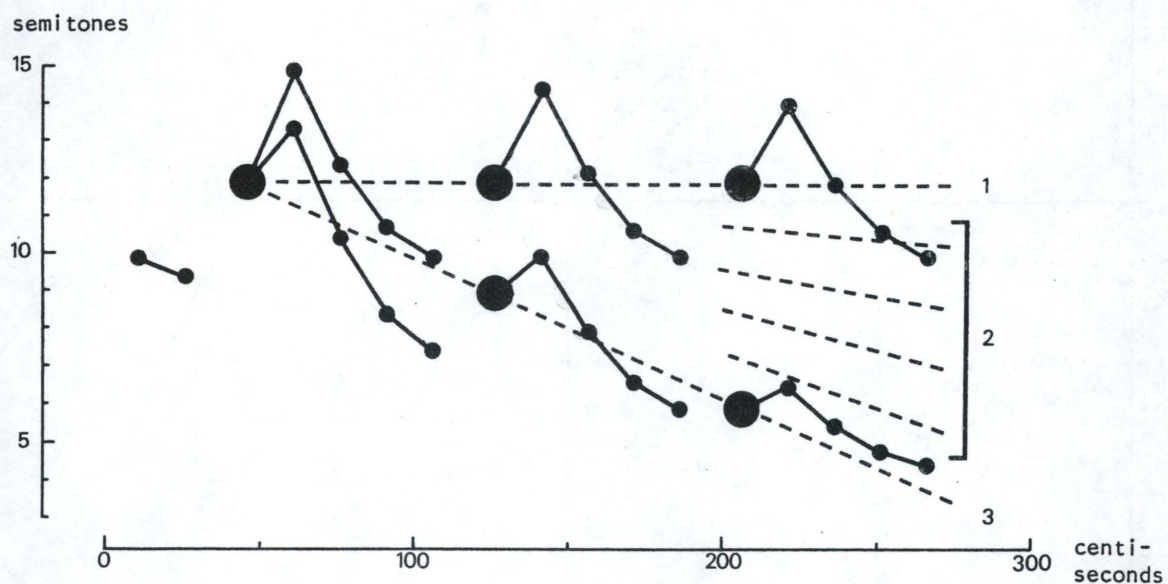


Figure 2

A model for the course of fundamental frequency in short sentences in Standard Danish. (1) Syntactically unmarked questions. (2) Questions with word order inversion and/or interrogative particle; non-final declarative and interrogative clauses. (3) Terminal declarative statements. Large points denote stressed syllables, small points depict unstressed syllables. Full lines represent the F_0 pattern of prosodic stress groups; broken lines indicate the sentence intonation contour.

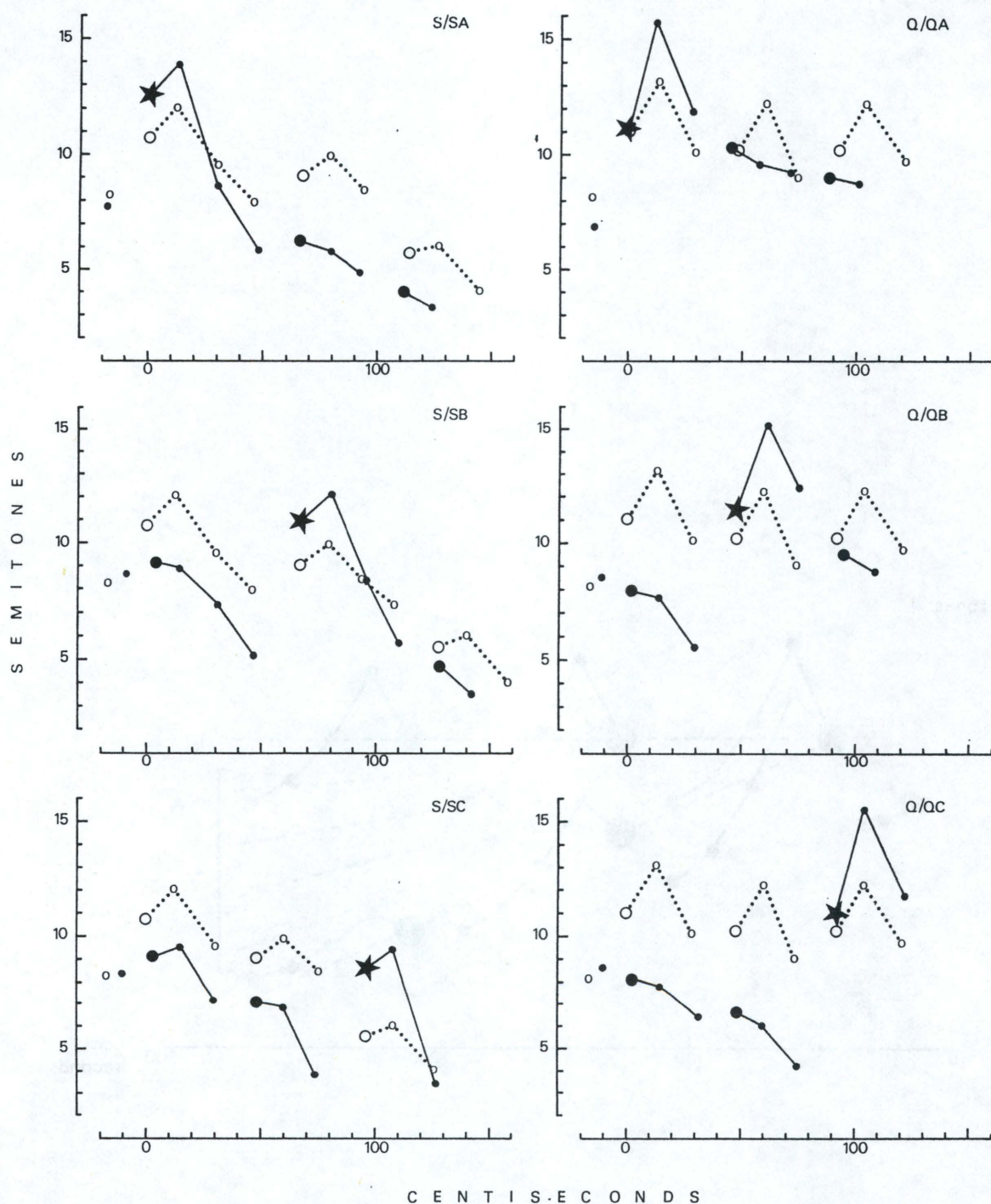


Figure 3

Stylized tracings of the course of fundamental frequency in statements (S - left) and questions (Q - right), which are prosodically neutral (open circles and dotted lines) or have emphasis for contrast (stars, points and full lines) in initial position (top), medial position (mid), and final position (bottom). Stars denote the emphatically stressed syllables, large points denote other stressed syllables, and small points depict unstressed syllables.

It appears, then, that in read speech in Danish, at least the speech I have analysed so far, stressed syllables are normally also accented. In the materials I have looked at before the present one, this has been true without exception, which is why I have not thought to introduce a distinction between stress and accent before. In retrospect, the behaviour of stressed syllables surrounding an emphasis for contrast could have provoked a discussion of accented versus non-accented: Stressed syllables which are neighbours to an emphasis for contrast suffer a reduction or deletion of their F_0 pattern, cf. figure 3. However, I am not sure whether this reduction is not also a de-stressing (that is how it has been treated previously). This is clearly a point where empirical research is called for, both in terms of acoustic and perceptual studies. - It is not my intention, and I am not able here, to do more than suggest the utility of a distinction between stressed syllables which are associated with a tonal change, and those which are not. I also want to suggest that stress is binary: either a syllable is stressed or it is not (whatever the criteria may be that determine this), but accentuation is a continuous function. Syllables are more or less accented in proportion to the amount of pitch change they are associated with, everything else being equal.

There is one snag about the dichotomy accented/non-accented, if it is to apply to stressed syllables only. There are 104 syllables with secondary stress in the text. (The decision about syllables carrying secondary stress has been reached on purely formal grounds: secondary stress occurs when a main stress gets reduced in composite words, and in phrasal units (cf. Rischel 1980). This is an emergency procedure, since I do not know how to localize, consistently, secondary stresses in running speech by prosodic criteria. Even so, a certain arbitrariness invariably persists in the assignment of secondary stress.) 95 of the 104 secondary stresses are succeeded by unstressed syllables. 21 of those 95, i.e. 22%, are associated with a rise in F_0 . Does this mean that secondary stresses may also be accented? Or does it mean that the stress reduction is only partial? Previously, secondary stresses have been characterized prosodically as lacking the pitch rise typical of main stress, but retaining all other stress cues (duration, quality, stød). - If accented/non-accented is a relevant distinction in stressed syllables, then lack of pitch rise is not a feature confined to secondary stresses. Conversely, if there are secondary stresses with associated pitch rises, then accentuation is not confined to stressed syllables.

There is a vast amount of research to be done on stress in Danish, not least perceptual experiments - and it will also be evident from the discussion in the afternoon that stress and accent are among the most elusive phenomena in speech production and perception.

ACKNOWLEDGEMENT

I am grateful to Mette Greiffenberg for letting me use her tape.

REFERENCES

- Bannert, R. 1985: "Fokus, Kontrast, und Phrasenintonation im Deutschen", *Zeitschrift für Dialektologie und Linguistik* 52, p. 289-305
- 't Hart, J. and Collier, R. 1979: "On the interaction of accentuation and intonation in Dutch", *Proceedings of the Ninth International Congress of Phonetic Sciences*, vol. II, p. 395-402
- O'Connor, J.D. and Arnold, G.F. 1961: *Intonation of Colloquial English*, (London)
- Rischel, J. 1980: "Phrasenakzent als Signal des Objekts ohne 'determiner' im Dänischen", *Kopenhagener Beiträge zur germanistischen Linguistik*, Sonderbd. 1, p. 262-279
- Thorsen, N. 1978: "An acoustical analysis of Danish intonation", *J. Phonetics* 6, p. 151-175
- Thorsen, N. 1979: "Interpreting raw fundamental frequency tracings of Danish", *Phonetica* 36, p. 57-78
- Thorsen, N. 1980a: "A study of the perception of sentence intonation - Evidence from Danish", *J. Acoust. Soc. Am.* 67, p. 1014-1030
- Thorsen, N. 1980b: "Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish", *Ann. Rep. Inst. Phon. Univ. Copenhagen* 14, p. 121-205
- Thorsen, N. 1982a: "On the variability in F_0 patterning and the function of F_0 timing in languages where pitch cues stress", *Phonetica* 39, p. 302-316
- Thorsen, N. 1982b: "Sentence intonation in Danish", *Preprints for the Working Group on Intonation at the XIIIth International Congress of Linguists*, Tokyo, August 29 - September 4, 1982, p. 47-56
- Thorsen, N. 1983a: "Standard Danish sentence intonation - Phonetic data and their representation", *Folia Linguistica* 17, p. 187-220
- Thorsen, N. 1983b: "Two issues in the prosody of Standard Danish", In: *Prosody - Models and Measurements* (eds.: Cutler, A. and Ladd, D.R.), (Heidelberg), p. 27-38

- Thorsen, N. 1984: "Variability and invariance in Danish stress group patterns", *Phonetica* 41, p. 88-102
- Thorsen, N. 1985: "Intonation and text in Standard Danish", *J. Acoust. Soc. Am.* 77, p. 1205-1216
- Thorsen, N. (forthcoming): "Intonation and text in Standard Danish - with special reference to the abstract representation of intonation", In: *Phonologica* 1984.

APPENDIX I

Transcript (the first of five pages) of the analysed text. Stress marks have been added, to designate emphasis (ˈ), normal main stress (ˑ) and secondary stress (ˋ). Glottal attacks (ʔ) are also marked. Pause durations are likewise given in parentheses (in centiseconds).

I denne ópgave skal du øve tág-quèstions, altså páhængsspørgsmål.(75) Þíssa intonères ʔénten som én tónegrúppe med en faldende stígende kërne,(42) ʔéller som tó tónegrúpper, hvór den ánden údtáles med faldende stígende múnster.(150) Det første dú skal fóretáge dig i denne ópgave er at lýtte pá páhængsspørgsmål ʔúdtált som en énkelt tónegrúppe.(100) Betingelserne for at fá en ʔénkelt tónegrúppeúdtále ʔér,(60) ʔénten at páhænget fólger efter en impérativ,(68) ʔéller at báde páhænget ʔog den fóranstíllede deklárative konstruktión ʔer négative eller pós(i)tive,(55) ʔéller at táleren stíller den tíltálte persón frít (23) med hénsyn til at svàre já eller néj (16) pá det gívne spørsmål.

Øvelse ét.(120)

I denne øvelse skal du øve impérativer med et ʔéfterfólgende páhæng.(140) Du hórer først spéakerens ʔíndtáling af páhængsspørgsmålet,(60) og ʔúmídelbart héréfter kómmer sátningen ʔi en filtréret fórm,(52) sálédes at du hórer intonatiónen men ʔíkke segménterne.(150) Du skal forsóge at sýnge méd pá den filtrérede sátning(44) - du kan for eksémpel sýnge pá ʔím'(51) - ʔéller du kan vælge báre at lýtte til intonatiónen i den filtrérede sátning.(135) I den éfterfólgende páuse (16) er det dín túr til at índtále páhængsspørgsmålet med den rígtige intonatió.(182) Éfter dít forsóg hórer du spéakerens versión (28) ʔéndnu en gáng,(40) ʔog kan héréfter góre ʔéndnu et forsóg.(60) (Jævnfør modéllen.)(125) Húsk (at) sèt(te) tálleren pá núl ʔínden du stárter. (((exercise)))