Aspects of Fo Declination in Swedish

Gösta Bruce

INTRODUCTION

The topic of the present study is declination (downdrift) - the often observed tendency in many languages for comparable F_0 values to be lower later in an utterance than earlier in the same unit of speech.

Some issues of declination have recently turned out to be controversial (for a review see Cohen et al. 1982).

One such issue is whether declination is to be regarded as a direct consequence of some physiological mechanism, e.g. a property of the respiratory system (Lieberman 1967, Atkinson 1973, Collier 1975, Maeda 1976), or whether it is basically learned behavior and linguistically purposeful (Ohala 1978, Pierrehumbert 1979).

Another issue of debate is how the overall course of a declination is to be interpreted: as a globally specified declining intonation on which local excursions for accentuation are superimposed (see e.g. Öhman 1967, Cohen and 't Hart 1967, Maeda 1976, Fujisaki et al. 1979, Bruce and Gårding 1978, Gårding 1982, Thorsen 1980a and recent contributions in Cutler and Ladd 1983 by Gårding, Thorsen and Vaissière) or as formed by the pitch relations between successive, local excursions for accentuation (Pierrehumbert 1980, Ladd 1983). This issue also involves the question of the possible use of look ahead and look back strategies for the planning and execution of a declination.

The purpose of the present paper is to draw the reader's attention to certain facts about the production of declination (with exemplification from Swedish) that I think are important for a proper understanding of this particular aspect of intonation.

I also intend to let these facts about declination in my interpretation shed light on the debated issues above, in particular the one concerning the implementation of the overall course of a declination.

DATA BASE

The considerations about F_0 declination to be presented in this paper are based on two independent studies of Swedish intonation.

One consists of a series of experiments in Standard Swedish intonation (Bruce 1982a), and the other is a study of text intonation in South Swedish (Bruce 1982b).

In the present paper I will draw upon the results mainly from two of the experiments of the first study. One concerns the relationship between declination and utterance length. In this experiment the number of stress groups (from two to five) is systematically varied as is placement of focus (neutral focus assignment or focus on the first or the last stress groups).

In the other experiment the number of stress groups is varied in the same manner, but at the same time there is also a systematic variation of overall emphasis manifested as two degrees of involvement (detached - involved). The same test material containing meaningful Swedish sentences was used in both experiments, but the recordings took place on separate occasions.

Each test sentence was elicited as an answer to a question. The increase in the number of stress groups was achieved by a syntactic expansion to the right.

The test material was phonetically balanced so as to minimize known microprosodic effects on the F_0 contour. In each test sentence there is an onset of two unstressed syllables and an offset of two (in one case three) unstressed syllables. Between the stresses there are three unstressed syllables. Each stressed syllable carries accent 2, which is analyzed phonologically in Standard Swedish phrased in autosegmental terms as (L) $\overset{*}{H}$ L . Even if no special grouping or focusing is elicited, the first and the last stress group of a sentence will (normally) have a phrase accent in this dialect. It is analyzed phonologically as (L)H immediately following the $\overset{*}{H}$ L of accent 2.

The study of text intonation in South Swedish concerns the relationship between declination and text unit length. In the variation of text unit length three different meaningful Swedish sentences, each containing two stress groups, were used. Each of the sentences can form a complete text unit, but they can also be combined to form two-sentence and three-sentence text units. For these larger units the order of the component sentences has been systematically shifted. The coupling between the sentences in two and three-sentence units is temporal, although this has not been expressed directly by means of temporal adverbs. The recordings of the multi-sentence units contain pauses between the constituent sentences.

The same requirements for the phonetic composition of the test material as in the first study were also complied with in this study. Each constituent sentence has an onset and offset of two unstressed syllables respectively, and there are three unstressed syllables between the stresses.

Each stressed syllable carries accent 2, which can be characterized as $\overset{*}{L}$ H (L) in South Swedish. In each text unit equal prominence of successive accents was elicited. Unlike Standard Swedish there is usually no addition of a phrase accent to the first or last stress group for a neutral focus assignment.

CHARACTERISTICS OF DECLINATION

Facts concerning declination obtained from the two separate studies of Swedish intonation will be summarized below. For a more detailed account see Bruce (1982a, 1982b). The similarities between the two studies seem to be great enough to allow a joint presentation.

It is clear that there are certain linguistic factors that will affect the course of F_0 declination. I have argued elsewhere that placement of focus is one such factor (Bruce 1982a). Up to the focus of an utterance declination appears to be absent or gentle, while after focus there is a conspicuous, stepwise declination. The same declination course can also be found for certain syntactic groupings (Bruce 1982a, section 4). Sentence type (question/statement) will also tend to affect declination. For question intonation (Gårding 1979) the declination is largely suspended. This also holds for a related

language, namely Danish (Thorsen 1980a).

In the material chosen for the present paper these linguistic factors have not been varied. Therefore I will concentrate on units of speech that are statements (answers to questions), that have neutral focus assignment, no special syntactic groupings and thus equal prominence of succesive accents.

Declination and unit length

The boundary conditions for F_0 declination are that F_0 starts Low and ends Low, and that the initial Low is higher in frequency than the final Low. The final Low (offset) is usually the lowest F_0 point of an utterance and also the least variable F_0 point. It is constant across variation in unit length, i.e. length of an utterance or a text unit in terms of number of stress groups (see Figs 1 and 2). But interestingly enough the same Low value will also be reached before a pause in non-final

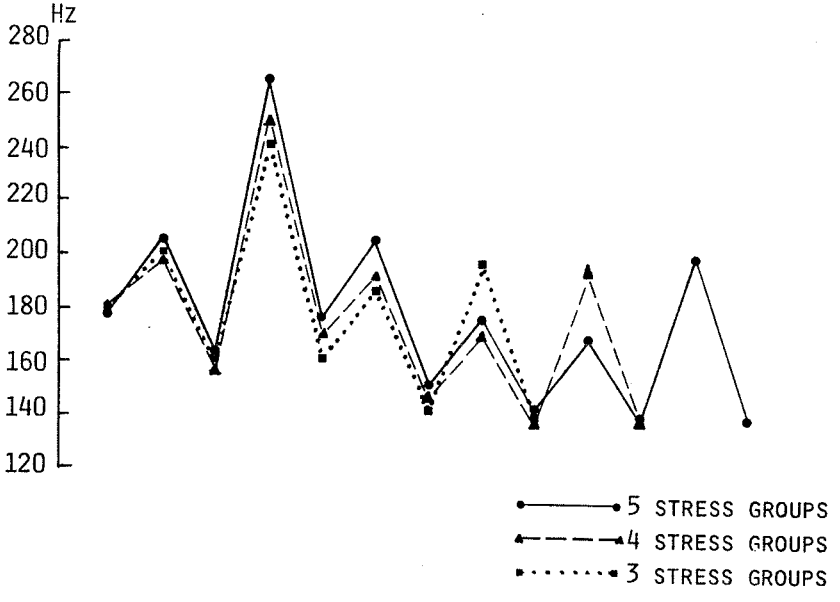

Figure 1. F_0 declination and utterance length. Stylized F_0 contours - means in Hz (9-11 repetitions) of successive F_0 minima and maxima - of utterances of varying length (3-5 stress groups) lined up from the beginning of the utterance. Neutral focus version for a female speaker of Standard Swedish.

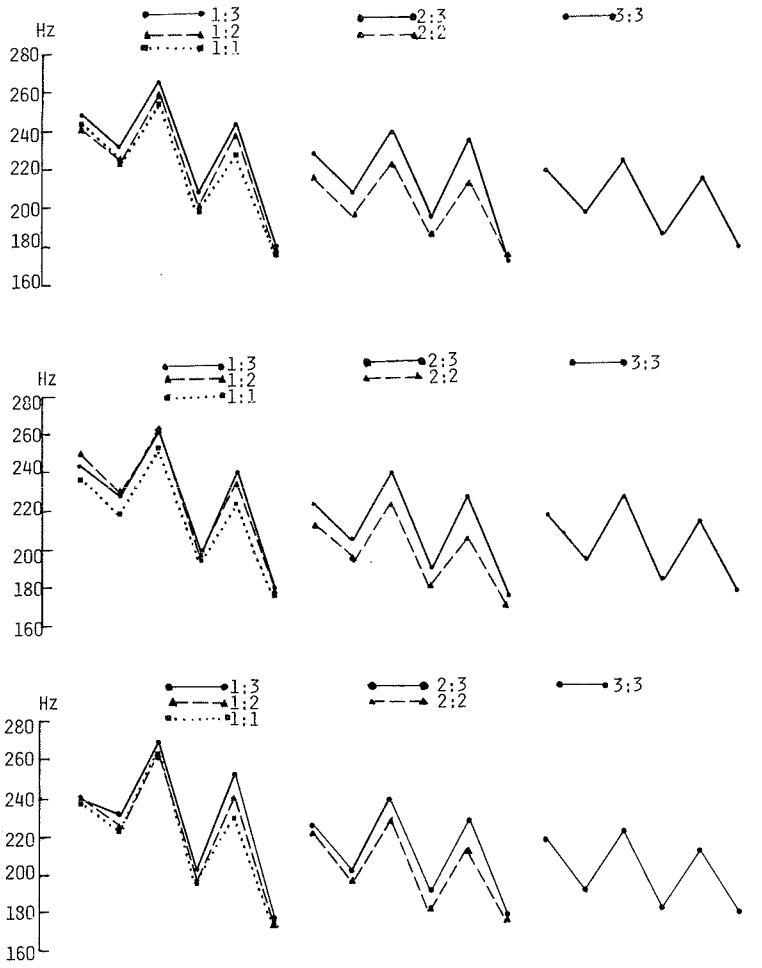

Figure 2. F_0 declination and text unit length. Stylized F_0 contours for three different sentences (upper, middle and lower part) - means in Hz (7 repetitions) of successive F_0 minima and maxima - in different positions of text units consisting of one, two or three sentences (two, four or six accents) lined up from the beginning of the text unit for a female speaker of South Swedish.

position within a text unit (see Fig 2). The final Low is also constant across variation in degree of involvement (see Fig 3). The constancy of this F_0 point is well attested (see e.g. Maeda 1976, Fujisaki et al. 1979, Pierrehumbert 1980).

The initial Low (onset) is constant over variation in unit length (see Figs 1-3), but varies (decreases) with position in a text unit (see Fig 2). It also varies (increases) with degree of involvement (Fig 3), as do all F_0 points except the final Low (see below).

This means that the initial Low of an utterance is a variable distance in frequency above the final Low of the preceding utterance. This is assumed to be the boundary signal, regardless whether or not there is a physical pause in between. The relative degree of frequency shift is an indication of the coupling between the actual utterances.

Another relevant F_0 point in the description of declination is the peak, the highest F_0 point of a unit of speech. It should be noted that in the study of Standard Swedish the peak is equal to the H of the first phrase accent which comes after the first accent, while in the study of South Swedish the peak is the H of the first accent of the unit.

From the peak there is a self-evident, successive lowering of F_0 values of both accent maxima and minima (see Figs 1-3). This is true, if we add the following qualification. In a text unit consisting of two or three sentences the decrease in F_0 values through the unit is not truly successive. The declination is arrested at each new start of a second or third sentence of a unit. This can be regarded as a kind of tonal coupling or adaptation of the earlier part of the F_0 contour of a later sentence to the later part of the F_0 contour of an earlier sentence (Fig 2).

The decrease in F_0 values is steeper in the beginning than later in a declination (see Figs 1-3). For similar results see Fujisaki et al. 1979 for Japanese, Thorsen 1980b, 1981 for Danish, Pierrehumbert 1980 and Sorenson and Cooper 1980 for American English. Ideally the declination appears to be exponential and asymptotic to the F_0 bottom of the speaker's voice range (the final Low)(cf. Pierrehumbert 1980).

There is no apparent difference in slope (rate of declination) between a longer and a shorter unit, lined up from the peak.

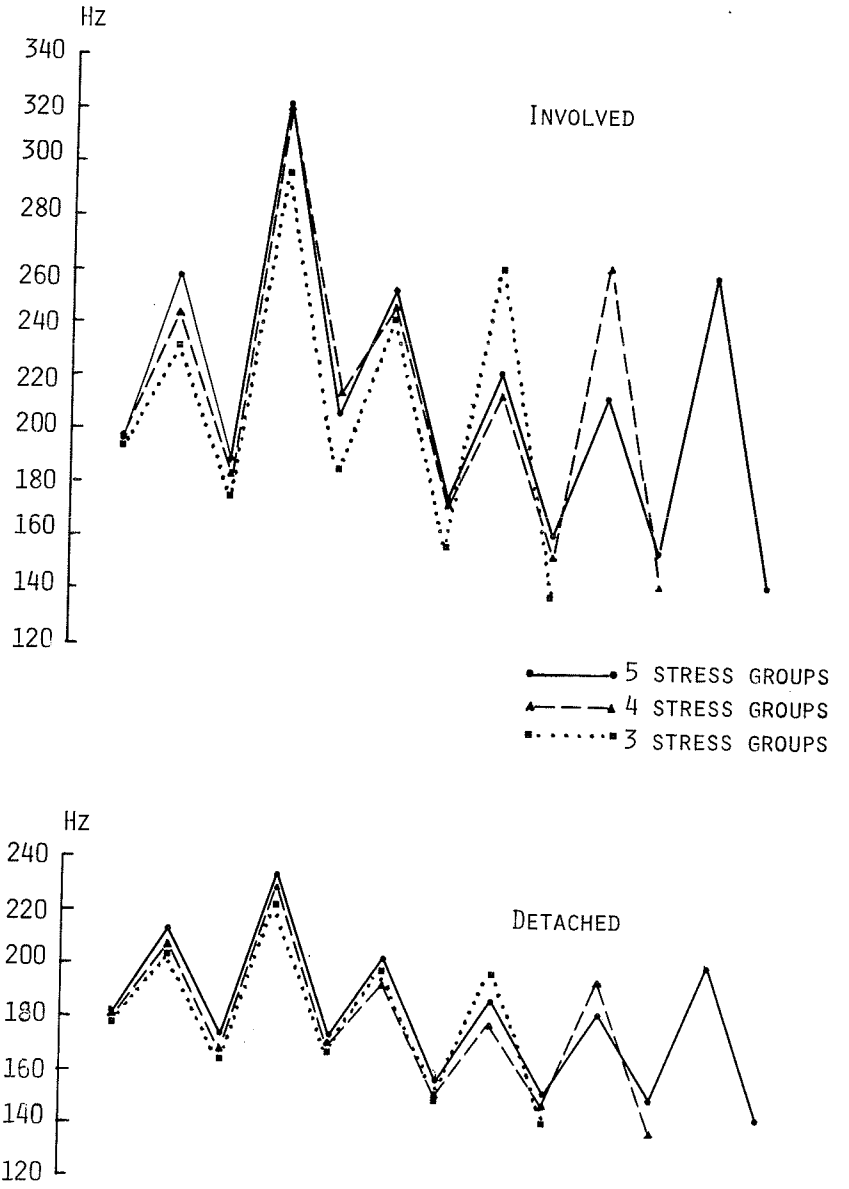

Figure 3. F_0 declination and involvement. Stylized F_0 contours - means in Hz (6 repetitions) of successive F_0 minima and maxima - of utterances of varying length (3-5 stress groups) lined up from the beginning of the utterance in an involved version (above) and a detached version (below) for a female speaker of Standard Swedish.

There is some adjustment of F_0 values to the length of the actual unit expressed as the number of stress groups (accents) contained. This means that lined up from the peak each successive maximum and minimum is usually higher in frequency in a longer utterance except for final values (see Figs 1-3).

I think it is important to note that the decrease in F_0 on a plateau consisting of several unaccented syllables between two accents, if any, is negligible compared with the decrease in F_0 taking place from the minimum preceding the accent excursion to the one following it. According to my data, it is only in utterances containing a single (or two) accent(s) with several unstressed syllables initally or finally that there is an obvious F_0 drop on such a plateau. Therefore, to account for the declination in Swedish described here as primarily one of accentual downstepping it seems possible to assume that the frequency value of each accent L (or H) is a constant ratio of the immediately preceding accent L (or H), as was suggested for American English by Pierrehumbert (1980). The overall declining course of an utterance can be seen as the result of this kind of local rule together with a sensitivity of F_0 values to the length of the unit of speech (For a preliminary test see Bruce 1982a).

I think that a description of the pitch relations between successive, local excursions for accentuation only in terms of F_0 points (minima and maxima) tends to be incomplete and to conceal interesting facts about declination. It seems wise to include also an account of the F_0 changes (rises and falls) contained in these excursions as a complement for an adequate description of declination (For a similar view see Ohala 1982).

There is an apparent asymmetry between rises and falls of the local excursions for accentuation. As a natural component of a declination an F_0 fall generally covers a wider range than the preceding F_0 rise of the same excursion. More interesting is perhaps the fact that the range of successive accent falls decreases gradually within a unit, while the accent rises preceding the falls tend to have a rather constant range independent of their position in a declination (see Figs 4 and 5). This means that the difference in frequency range between fall and rise is greater higher up in a declination. Towards the end of a declination F_0 falls have almost the same narrow range as do

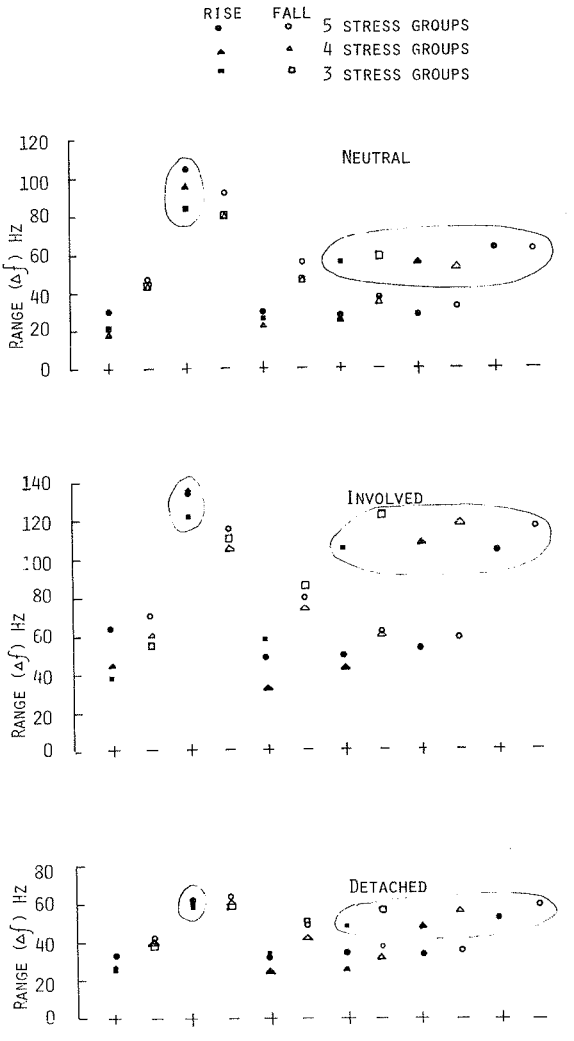

Figure 4. Rise - fall asymmetry. F_0 ranges - means in Hz of successive rises (+) and falls (-) - of utterances of varying length (3-5 stress groups) lined up from the beginning of the utterance. Neutral focus version (9-11 repetitions, upper part), involved version (6 repetitions, middle part), detached version (6 repetitions, lower part) for a female speaker of Standard Swedish. Encircled symbols are manifestations of phrase accent.

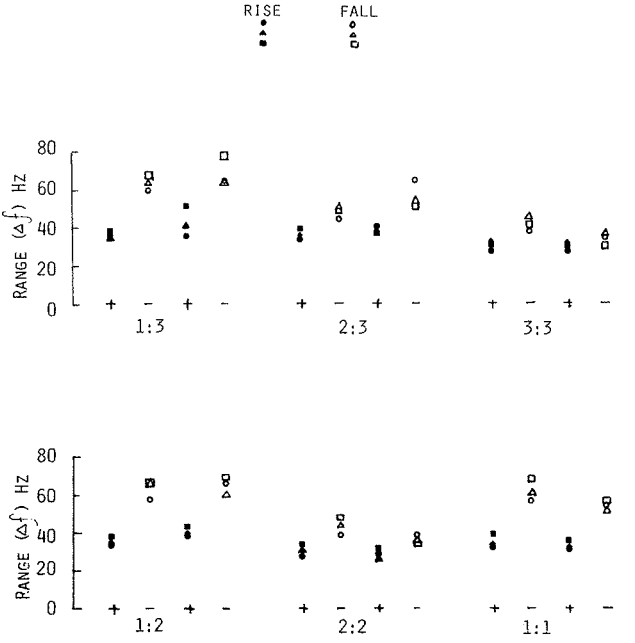

Figure 5. Rise - fall asymmetry. F_0 ranges for three sentences - means in Hz (7 repetitions) of successive rises (+) and falls (-) - in different positions of text units consisting of one, two or three sentences (two, four or six accents) for a female speaker of South Swedish.

corresponding rises. This is true for both studies of Swedish intonation in spite of of the apparent difference in timing of accent rises and falls relative to the stressed syllable between the two dialects described above.

While there is an adjustment to unit length of F_0 points, the range of each accent fall - in the same position lined up from the peak disregarding final characteristics - appears to be more constant (in Hz) across variations in unit length.

As for accent Hs and Ls the relation between successive accent falls of the same declination can profitably be described as one of a constant ratio, so that a particular fall is a

fraction of the immediately preceding one.

Finally, in the study of text intonation in South Swedish we observe that both the rise and fall of the second accent tend to be narrower in range than the first accent rise and fall in the same constituent sentence, provided it is text-final. If it is non-final, we find the opposite relation (see Fig 5). This means that there is a concomitant lowering of the text-final accent maximum (see Fig 2). This was not observed for Standard Swedish. Final lowering is a feature typical of American English (Liberman and Pierrehumbert 1982).

Declination and involvement

A change in involvement from detached to involved is clearly expressed as an increase in F_0 range (see Fig 3). Variation in F_0 range is achieved by a frequency expansion upwards; the lower limit of the range (final Low) is fixed, while the upper limit (peak) is highly flexible. F_0 maxima appear to be more affected than F_0 minima. While there is no slope difference between utterances differing in length, there is a superficially obvious slope difference between utterances differing in F_0 range. However, the same F_0 pattern is maintained with a change in F_0 range. Linear regression analysis shows a very high correlation between F_0 values (maxima and minima as well as rises and falls) of a detached and an involved attitude. The relation between successive F_0 values (maxima, minima, falls) is of approximately the same constant ratio in the two different studies. Very similar results are presented in Liberman and Pierrehumbert (1982).

It should be clear by now that a declination is not properly described as a chain of decreasing ranges simulating a successively decreasing degree of involvement. The range decrease in a declination concerns only the falls, while a shift from detached to involved or vice versa affects the range of both rises and falls.

CONCLUSION

I interpret the facts about declination in Swedish presented above as favoring the idea that in a declining intonation the overall course is essentially formed by the pitch relations between successive, local excursions for accentuation. The

possible, but not obligatory downtrend in F_0 on unaccented plateaus between accents can be conceived of as an extra means of amplifying a declination. Therefore, what is mainly responsible for the declination effect, in my opinion, is the fact that the fall of a local excursion for accentuation is controlled to cover a wider frequency range than the preceding rise of the same excursion and that this process repeats itself with a successive decrease of a constant ratio in the L value and in the range of the fall for each accent. This is the ideal way of expressing the combination of declination and equal weight of successive accents.

According to this view it is also more adequate to say that there is an adjustment of F_0 values to the number of upcoming accents rather than to utterance length per se.

A drop in F_0 (close) to the bottom of the speaker's voice range somewhere in the course and then a resetting to the previous course of declination will favor the perception of a boundary, e.g. a clause boundary. There is also preliminary evidence showing that a departure in the other direction from the ideal downstepping contour with instead no decrease in successive L values will be perceived as indicating less prominence and thus a grouping of the actual accents.

I think it is natural to assume from the present data on declination in Swedish that both a look ahead and a look back are used in the control of F_0 declination. The look back to the preceding accent will ensure equal prominence of successive accents (or any prominence relations aimed at by the speaker) (cf. Pierrehumbert 1980). The look ahead will be involved in the adjustment of F_0 values to the number of upcoming accents in order to ensure that the F_0 bottom of the speaker's voice will not be reached until the end of the utterance.

I would be inclined to take the facts concerning declination presented here as favoring an interpretation of declination as learned behavior, actively controlled, linguistically purposeful and useful in the textual organization of speech.

The relative weight of what happens at the local excursions for accentuation compared to what happens in between, the rise-fall asymmetry of these excursions in combination with the effect of linguistic factors such as placement of focus, certain syntactic groupings and sentence type (question/statement) on

declination are all indicative of declination as a linguistically integrated phenomenon, which is a phonetically motivated and natural process that does not seem to be accounted for in any direct way by any specific physiological constraint.

REFERENCES

Atkinson, J. (1973) Aspects of intonation in speech: implications from an experimental study of fundamental frequency. Unpublished doctoral thesis. University of Connecticut

Bruce, G. (1982a) Developing the Swedish intonation model. Working Papers, Department of Linguistics, Lund University 22: 51-116

Bruce, G. (1982b) Textual aspects of prosody in Swedish. Phonetica 39: 274-287

Bruce, G. and Gårding, E. (1978) A prosodic typology for Swedish dialects. In Gårding, E., Bruce, G. and Bannert, R. (Eds.) Nordic prosody: 219-228. Department of Linguistics, Lund University

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

Cohen, A., Collier, R. and 't Hart J. (1982) Declination: Construct or intrinsic feature of speech pitch? Phonetica 39:254-273

Collier, R. (1975) Physiological correlates of intonation patterns. JASA 58: 249-255

Cutler, A. and Ladd, R. (Eds.) (1983) Prosody: models and measurements. Heidelberg: Springer

Fujisaki, H., Hirose, K. and Ohta, K. (1979) Acoustic features of the fundamental frequency contours of declarative sentences. Annual bulletin of the research institute of logopedics and phoniatrics, University of Tokyo 3: 163-172

Gårding, E. (1979) Sentence intonation in Swedish. Phonetica 36:207-215

Gårding, E. (1982) Contrastive prosody: a model and its applications. Studia Linguistica XXXVI: 146-165

Ladd, R. (1983) Peak features and overall slope. In Cutler, A. and Ladd, R. (Eds.) Prosody: models and measurements: 39-52. Heidelberg: Springer

Liberman, M. and Pierrehumbert, J. (1982) Intonational invariance under changes in pitch range and length. Manuscript, Bell Laboratories

Lieberman, P. (1967) Intonation, perception and language. Cambridge: MIT Press

Maeda, S. (1976) A characterization of American English intonation. Unpublished doctoral thesis. Massachusetts Institute of Technology

Ohala, J. (1978) Production of tone. In Fromkin, V. (Ed.) Tone: A linguistic survey: 5-39. New York: Academic Press

Ohala, J. (1982) Physiological mechanisms underlying tone and intonation. In Fujisaki, H. and Gårding, E. (Eds.) Working group on intonation, Preprints. The XIIIth International Congress of Linguists. Tokyo

Öhman, S. (1967) Word and sentence intonation: A quantitative model. STL-QPSR 2-3: 20-54

Pierrehumbert, J. (1979) The perception of fundamental frequency declination. JASA 66 (2): 363-369

Pierrehumbert, J. (1980) The phonology and phonetics of English intonation. Unpublished doctoral thesis. Massachusetts Institute of Technology

Sorenson, J. and Cooper, W. (1980) Syntactic coding of fundamental frequency in speech production. In Cole, R.A. (Ed.) Perception and production of fluent speech: 399-440. Hillsdale, New Jersey: Lawrence Erlbaum Associates

Thorsen, N. (1980a) A study of the perception of sentence intonation. JASA 67: 1014-1030

Thorsen, N. (1980b) Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish. ARIPUC 14: 1-29. University of Copenhagen

Thorsen, N. (1981) Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish: supplementary data. ARIPUC 15: 13-47. University of Copenhagen