303. On Langue française 5. *IRAL* 9: 180-181. 1971. [Review *Langue française* 5. Paris, 1970.].

304. On Linguistique et enseignement des langues étrangères. *Studia Linguistica* 25: 61-63. 1971. [Review de Grève, M., van Passel, F. *Linguistique et enseignement des langues étrangères*. Paris, 1968.].

305. On Littérature et signification. *Studia Linguistica* 25: 59-61. 1971. [Review Todorov, T. *Littérature et signification*. Paris, 1967.].

306. On Substance and structure of language. *Studia Linguistica* 25: 54-59. 1971. [Review Puhvel, J.(ed.) *Substance and structure of language*. Berkeley, 1969.].

307. Relatum. Analyse et critique de *Message et phonétique*. In *Bulletin d'audiophonologie* 2: 127-137. 1972. [Review Lafon, J.-C. *Eléments de message et phonétique*. Besançon, 1972.].

308. On Les langues romanes. *Le français moderne*. Octobre (4), 1975. 353-356. [Review Camproux, C. *Les langues romanes*. Paris, 1974.].

309. Signes et la théorie linguistique. *Semiotica* 17: 153-163. 1976. [Review Sebeok, T.A. (Ed.) *The tell-tale sign*. A survey of semiotics.].

310. On Semiotic perspectives. *Studia linguistica* 37: 193-199. 1983. [Review Hervey, S. *Semiotic perspectives*. London, 1982.].

311. Théories du signe – un aperçu sémiotique. *Semiotica* 44: 363-369. 1983. [Review Bailey, R.W.; Matejka, L.; Steiner, P. (eds.) *The sign: Semiotics around the world*. Ann Arbor, 1978.].

312. On Sprachwissenschaftliche Germanistik. *Moderna Språk* 81: 156-162. 1987. [Review Bahner, W. & Neumann, W. (eds.) *Sprachwissenschaftliche Germansistik. Ihre Herausbildung und Begründung*. Berlin, 1985.].

313. On El problema de las lenguas en contacto. *Studia Linguistica* 44: 82-85. 1990. [Review Sala, M. *El problema de las lenguas en contacto*. Mexico, 1988.].

314. On From Latin to Spanish. *Studia Linguistica* 44: 75-81. 1990. [Review Lloyd, P.M. *From Latin to Spanish; Vol 1: Historical phonology and morphology of the Spanish language*. Philadelphia, 1987.].

315. Les théories linguistiques de Brøndal. *Semiotica* 91: 341-358. 1992. [Review Brandt, P.A. (ed.) *Linguistique et sémiotique: Actualité de Viggo Brøndal*. Copenhagen, 1989.].

The Bertil Malmberg Bibliography was compiled by Marianne Gullberg

# On Parameters and Principles in Intonation Analysis

Eva Gårding

Introduction

This paper is an elaboration of two preprints (Gårding 1991a and 1991b) and oral contributions made at the conference *Fonetik 1991* in Stockholm and at the symposium *Intonation: Models and Parameters*, held at the International Congress of Phonetic Sciences in Aix en Provence the same year. Both preprints discuss the perceptual relevance of the production-based parameters of an intonation model that I have worked with. To give coherence to this presentation I shall start with a brief review of the production-based model and the parameters of my latest version of it. This review serves as a background for the main part of the paper which is a discussion of the perceptual aspects of the model parameters. In the last section I will comment on some principles used in recent intonation analyses.

The model was originally designed for Swedish dialects (Bruce & Gårding 1978). Our ambition was to create a unified model for Swedish intonation, which can generate dialectal pitch contours for the same temporally structured input sentences using identical phonological components and phonetic parameters. The theoretical background was simple: to separate functionally defined form (phonology) and substance (phonetics) in our analysis.[1]

The phonological components are the word accents (A1 and A2), accents on phrase and sentence level (Focus) and boundaries. Sentence intonation is represented by a set of auxiliary lines (later to be called tonal grid) on which accents and boundaries are inserted as points whose position relative

---

[1] The separation of form and substance at all levels in speech analysis was one of the key notions taught by BERTIL MALMBERG. His phonetics seminar was truly 'integrative' in nature, since, in keeping with his own wide interests, it was a forum for all those interested in LANGUAGE: phonology and phonetics, theory and application. I dedicate my paper to Bertil, my teacher and friend, with respect and affection.
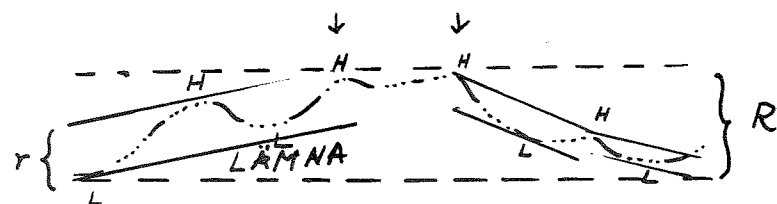
Figure 1. Parameters in the production-based model. Fo curve of *Man vill LÄMNA (A2) nåra långa (A2) nunnor (A2)* 'One wants to deliver some long baguettes' in declarative intonation. Stockholm speaker. A2 is manifested as HLH turning points in focus, after focus as HL (Bruce 1977). Parallel lines represent the grid. 'R' denotes global range, 'r' is the vertical distance between the grid lines. Pivots are marked by arrows. After focus the grid is compressed.

to the segments are dialect dependent and given by rule. The rules were established in analyses of an extensive collection of data produced by speakers representing different dialect areas.

In the generative program the word accents are subordinated to sentence intonation: The grid is generated first and constrains the pitch values of the target points of the accents. Interpolation is used to fill in the blanks between the generated points. In his overview of present intonation research, Hirst called it a target-transition model (1991).[2]

The model has been revised in various ways in work with more complex sentences (e.g. Bruce 1982, 1984, 1986). In my own work with different languages, I have developed the notions of tonal grid and pivot to account for prosodic constituents and the hierarchical structure of intonation (Gårding 1981, 1983, 1987a,b).

Production-based parameters

Figure 1 shows the parameters of the revised model which are all visual correlates of acoustic events. *Turning points* are the local maxima and minima of the Fo curve which are related to an accented syllable in a systematic way. The timing of these points in relation to the accented syllables is crucial for the distinction between the two accents of Swedish (A1 and A2) and typical for a particular dialect. In the generative program,

[2]An overview of present and past intonation models as well as visions of future ones was given by Hirst 1991.

the dialectal variation in intonation appears as different timing rules for word accent and sentence accent (Bruce & Gårding 1978).

The *tonal grid* is the auxiliary parallel lines enclosing curve pieces with a common general direction and range. The upper line of the interior grid (see Figure 1) is supported by local maxima of the curve and the lower line by local minima. In practice not all the the local maxima and minima reach the grid lines. This is a tonal reflex of the varying accentuation levels carried by the syllables which are enclosed by the same grid lines. The grid can be shifted up and down to fit changes in range, which are characteristic of spontaneous speech and discourse.[3]

There are two kinds of *range*, small 'r' and capital 'R'. Small 'r' corresponds to the width of the interior grid, capital 'R' denotes the exterior grid, i.e. the range between the lowest and highest point of the curve.

Pivots are those places in the curve, marked by arrows, where the grid changes direction or range. They usually occur in connection with focus and demarcate constituents in the global contour.

The model with its parameters, the turning points, the tonal grid, the pivots and the ranges can be used in analysis as well as synthesis. It has been used to capture the gross features of Swedish intonation in synthesis. In analysis, the grid and the pivots make it possible to segment the curve into parts related to prosodic phrases or subphrases. The speaker's choice of range and register for such pieces indicates how phrases are weighted and related to each other.

In this way the model is fit to handle the hierarchical structure that is regarded as a main characteristic of language. For example, the model can coordinate, superordinate and subordinate elements of speech and thereby express one of the most important functions of syntax.

[3]At the present stage a strict algorithm for the construction of a tonal grid cannot be given; as many other terms in intonation analysis, e.g. upstep and downstep, it is still on the descriptive side. A demonstration of its utility for cross-language comparisons was given in Gårding 1984. It was used by Touati for his comparison of tonal features in French and Swedish (1987). A tonal grid construction based on regression lines for successive Fo maxima and minima was developed by Mountford for Bambara (1983).This method was used by Susan Herz in a fundamental frequency program, called Delta (1990) and by Huber for the analysis of intonation units in read Swedish texts (1988). Attempts to quantify the accentuation levels relative to a tonal grid were made by Merle Horne for English (1987).
  The perception-oriented analysis of accentuation patterns brings out invariant relations between the ranges of the grid enclosing different parts of the contours. For the intonation of these sentences at least, the pivot range (R) and place acted as a scale setter for the rest of the contour. If such relations hold in a larger material they may help in constructing an acceptable algorithm.
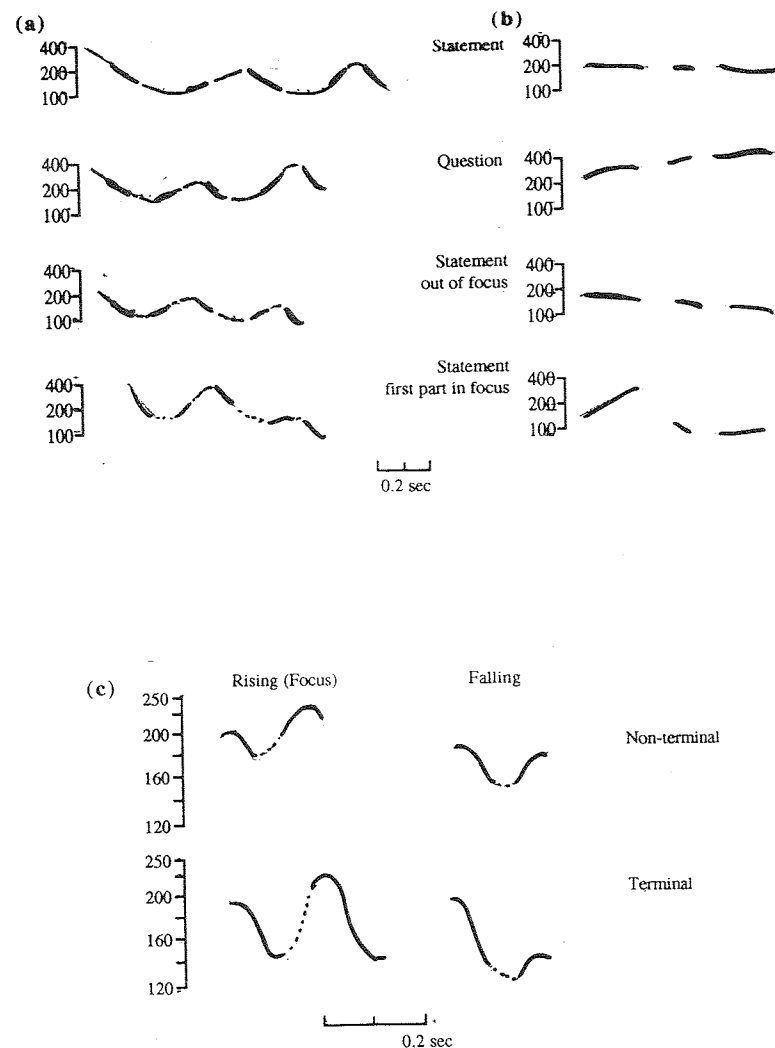
With support from the superposition principle and the grid, which can take various directions and ranges, it is possible to show how the surface shapes of the local tones and accents are determined by global intonation (see Figure 2 and arguments in the last section).

Perceptual aspects

Turning points

The timing of pitch peaks in relation to the accented syllable was used for my categorization of Swedish prosodic dialects (1973). This categorization, mainly based on Meyer's data (1937, 1954), was also inspired by the results of peak-shift experiments conducted by Malmberg which showed that the location of the peak in the accented syllable controlled the A1/A2 distinction (1967).The timing characteristics of the accents were essential in the generative program suggested by Gårding 1976, and by Bruce 1977 and they prevailed in the dialectal typology proposed by Bruce & Gårding 1978. Kohler 1987 and Ladd 1983 adopted the notion of peak location in their analyses of accents with different semantic connotations in German and English. Ladd even suggested 'delayed peak' as a unit in the phonologic inventory (1983). It seems clear that this kind of analysis is based on the visual correlates of accents in the acoustic record and that the aural impressions may call for a different phonological notation, as I will show in two examples.

Highs and Lows are often used to show for a given Fo curve how visual local maxima and minima relate to the text and to the surrounding highs and lows. For Swedish this kind of analysis applied to the Stockholm dialect has yielded High for A2 and Low for A1, based on the early high or low turning point after the C/V boundary of the accented syllable (Bruce 1977, 1986). In an intonation analysis of Hausa (traditionally and aurally analysed as having two tones, High and Low), Lindau showed that a 'high' tone is manifested as a rise to a high point at the end of the syllable and a 'low' tone as a corresponding fall to low (Lindau 1986). This finding suggested that it is the latter part of the syllable that has perceptual salience rather than the earlier one. In addition, a quick fall from an early peak may give the auditive impression of LOW whereas a fall from a late peak position is perhaps best described as HIGH. With the present interest in a general descriptive framework for tone languages and accent languages it would be desirable to keep apart 'visual' and 'auditive' parameters.

Figure 2. Output forms of tones and accents due to (semi)global intonation (the grid) can be explained by the superposition principle.
(a) Chinese *Sòng Yán mài niúròu* 'Song Yan sells beef' with alternating falling and rising tones.
(b) *Wāng Yī chōu xiāngyān* 'Wang Yi smokes cigarettes' with high tones only. (Gårding 1987b.)
(c) Swedish *Bara* (A2) ('only' and placename) in rising (focus) and falling intonation. (Gårding 1967.)

It is perhaps self-evident that a given Fo curve can be reconstructed either from the position in time and frequency of the turning points or from suitable information about the falls and rises that make up the curve.[4] From a perceptual point of view, however, the pitch movement or the contrast between two frequency levels seems more relevant than the turning point even if the turning point is bound to have some important function in the auditory process. Following neurological basics about the importance of the onset and offset of a nerve signal, we may regard the turning points as attention-sharpeners. David House 1990 concludes from perceptual experiments with manipulated Fo patterns in Swedish nonsense syllables that it is during time intervals with spectral stability, hence the latter part of the vocalic segment, that the pitch configuration is interpreted as movement. In a letter Xiao-nan Susan Shen reports on experiments indicating that the turning points in connection with the vowels may act as boundaries for the tonal percept (1992).

The perceptual importance of the pitch movement over the vocalic part of an accented syllable for the identification of a prosodic pattern has been demonstrated by many researchers in experiments with several languages. Responses to stimuli with a time-shifted Fo peak in Chinese material and various other manipulations showed that for the recognition of T3 and T4, the peak position was only indirectly important (Gårding et al. 1986).

For a language with contrastive accents it seems that not only the accented syllable is significant but also the post-accented part. This may be explained not only by the recency effect (later cues having predominance over earlier cues) but by the fact that the post-accented part often has a substantial duration and a strong functional burden, namely to enclose the accented and the following unaccented and deaccented syllables in a common domain, which may carry information about focus and modality. The importance of the post-accented part is also emphasized by Lehiste and Ivić in their investigations of Serbocroatian accents (1986). Based on production data and perceptual experiments they conclude that a specification of the Fo peak location within a syllable is not a sufficient basis for identifying an accented syllable (p. 168). Primary importance

---

[4]Any reasonable curve can be well approximated by a piecewise linear curve. This property has been used by 't Hart, Collier & Cohen to analyse and synthesize intonation curves. One of their results is that rather rough approximations of this kind cannot be perceptually distinguished from the originals. Perception-orientation is a principal characteristic of their model. For a comprehensive account, see 't Hart, Collier & Cohen 1990.

belongs to the relationship between the two syllables constituting the accented disyllabic sequence (p. 170).

The perceptual salience of the post-accented part explains the classic 'ear-phoneticians' prosodic characterization of Swedish dialects into dialects with rising accents (e.g. Central Swedish) and dialects with falling ones (e.g. South Swedish). A similar categorization has been used in Norwegian dialectology (Alnæs 1916, see also Fretheim 1984:69) and has been applied to other Germanic languages without any accent distinction (e.g. for German, Sievers 1912).

Grids and pivots
I shall start with a summary of some general ideas of the perceptual aspects of grids and pivots, already published in the preprints, and continue with a brief account of three sets of experiments, conducted in collaboration with David House and Lars Eriksson.

The tonal grid which may be level or exhibit various degrees of fall and rise, encloses a pattern made up of syllables with often varying degrees of accentuation. The fact that the Fo movements relative to the grid are retained in different intonations is seen as a consequence of the superposition principle, as exemplified in Figure 2. The communicative value of the direction of the grid may be explained in the following way: The auditory system makes a running average of the Fo values of the phrase and these values are retained in memory to be matched with stored prototypes having semantic connotations.

The perceptual importance of *the pivots* stems from a discontinuity in direction or range of the global Fo pattern. Just like the turning points, the pivots are not directly perceptual units but may serve as attention sharpeners (and boundaries?). The new direction and range signal a changed communicative value for the carrying syllables.

Three experiments
In the first experiment to be reported, phrasing is varied and accentuation and intonation are constant. In the second, accentuation levels are varied and the intonation and phrasing are constant. In the third, which is still preliminary, intonation is varied and accentuation patterns are kept constant.

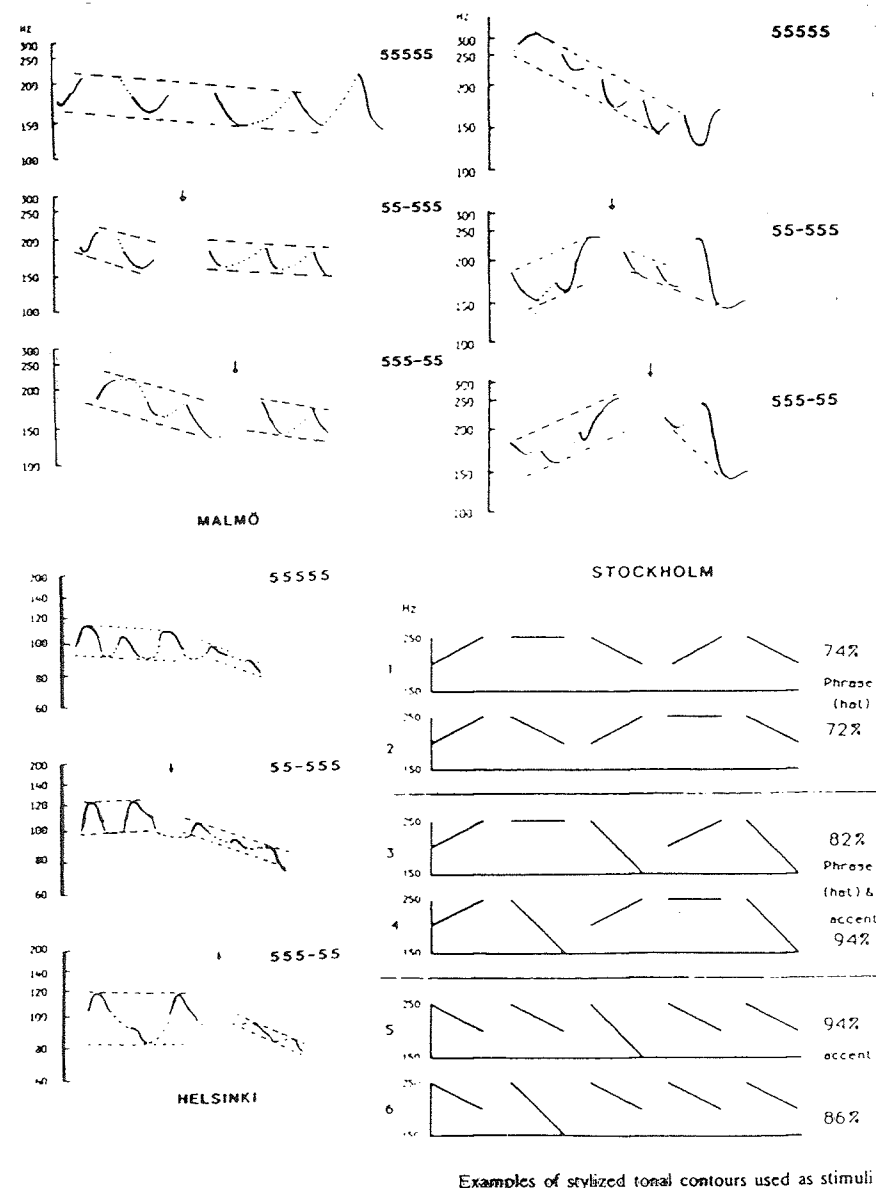### 1. *Phrasing varied. Accentuation and intonation constant*

These experiments, conducted with David House, used telephone numbers as material to avoid syntactic and semantic influence on the production. Our speakers represented different dialects and languages. Figure 3 exemplifies how speakers of Swedish dialects and Helsinki Finnish produce different phrasings.

The task was to group a sequence of five digits in 3 different ways (5 in a series, and series of 3+2 and 2+3), to give the same weight to every digit and to perform the grouping without pauses in declarative intonation. The acoustic records suggested that the utterances contained both general and language or dialect-specific demarcative and connective signals. We could define 'tonal grid' as the set of frames enclosing prosodic groups and 'pivot' as the point where the grid changes direction or range. Synthetic stimuli which simulated the natural productions with configurations of equal duration (see Figure 3) were tested with Swedish listeners. The results showed that large-size Fo movements were efficient boundary signals corresponding to pivots. Important connective signals were patterns which in the visual analysis could be regarded as enclosed by a common frame. Within these frames the digits had either similar pitch movements or special dialect-dependent shapes in the form of troughs or hats (Gårding & House 1987).

2. *Accentuation varied. Intonation and phrasing constant*

In this set of experiments we used the phrase *dom e långa män* 'they are tall men', produced by a Stockholm speaker. The sequence *långa män* may be tied to different meanings on account of its accentuation pattern, namely *långa män* (focus free), *LÅNGA män* (focus left), *långa MÄN* (focus right), and *långa-män* (compound) (Gårding & Eriksson 1991). Some stimuli were produced by shifting Fo peaks, more precisely open triangles with bases of different size, in small steps in time over constant carriers. Listeners' reactions showed that a prosodic phrase pattern can be recognized when the position of the peak had produced a pattern in which pitch movements over neighbouring vowels agreed with the original. Our experiment suggests that the Fo peak position is only indirectly important. It is the adjoining ramps over the vowels which have perceptual reality.

Other synthetic stimuli were produced by stripping off acoustic properties from the carrier signals layer by layer. From the results of categorization tests with these manipulated stimuli we could confirm that

Figure 3. Phrasing varied, constant accentuation and intonation. The Swedish speakers have used the digit *fem* and the Finnish speaker the colloquial form *viis*. Examples of stylized tonal contours used as stimuli in the right-hand corner. From Gårding & House 1987.

the vocalic part is the most important carrier of pitch, the temporal pattern is very important and the intensity pattern is by no means unimportant for a phrase to be easily recognized.[5] These results guided our choice of descriptive parameters which are demonstrated in Figure 3, namely Fo, impulse area and duration relative to the whole phrase. (For impulse area and its use in speech analysis, see Fant 1954.)

Figure 4 presents stylized versions of our original phrases. I have enclosed all the sentences with exterior grid lines, expressing the overall trend of the contours. Within the exterior global lines there are interior grid lines which enclose the rising parts of the Fo curves. The figure leads to the following comments.

Duration and intensity. The temporal pattern is represented by durations measured from vowel onset to vowel onset and from vowel onset to the end of the phrase. It is expressed as eighths of the total duration. Intensity is represented by the impulse area under the decibel envelope over the vowels, also expressed as eighths of the total energy. The durational pattern and the intensity pattern go in the same direction and I have collapsed them into one parameter that I call rhythm. This kind of measurement gives a rather good correspondence to aural impressions of the rhythm of the phrases. I have transcribed the rhythm according to classical tradition with macrons and microns. The first and the last phrase are similar in rhythm. The focus patterns differ: focus left is a dactyl, focus right is an anapest.

The tonal patterns have been stylized based on our experience from earlier ILS experiments which showed that curved lines could be replaced by straight ones without causing any loss of prosodic information (compare also Note 4). The frequency scale is in semitones. The position of the pivot in relation to the text is determined by the accentuation pattern. The fall from the beginning point to the lowest end point is 3 semitones for all the patterns, which may be an effect of declination or an expression of their common declarative sentence intonation (see discussion in the last section). Another expression of their common intonation is the fall from the pivot, i.e. the last turning point of the global contour. This pitch interval is largest in the sentences with focus, and late focus has a smaller interval than the

---

[5]We were surprised to find that Fo did not play the predominant role that we expected. Instead, different acoustic correlates turned out to have different importance for different prosodic patterns (Gårding & Eriksson 1991).
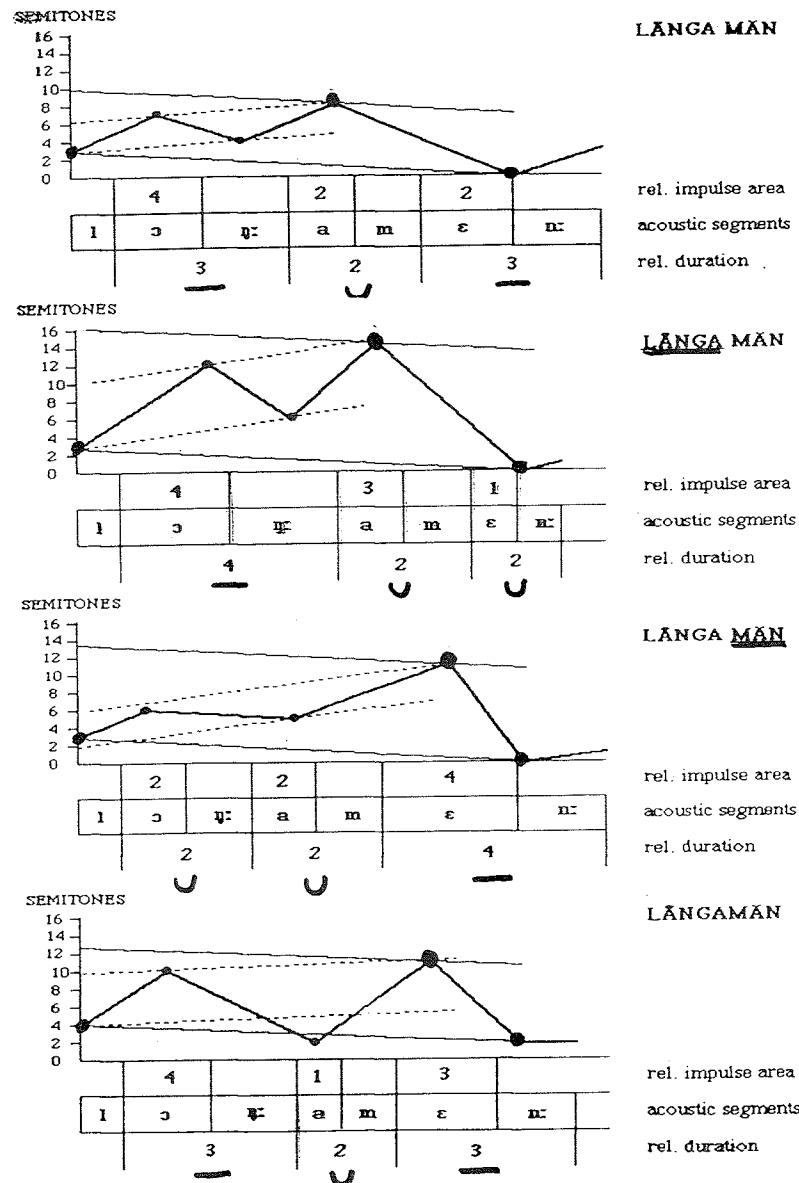

Figure 4. Accentuation varied, intonation constant. Stylized Fo curves of one of several productions of each of the distinctive accentuation patterns. Stockholm speaker. For measurement see text.

early one (due to declination?).The slope of the fall varies depending on the focus position and the time interval from the pivot to the end point. This agrees with results obtained with English material as shown convincingly by Cooper & Sorensen 1981:37ff.). The pitch interval to the pivot from the nearest turning point to the left is the same for both phrases with focus and possibly position independent.

In the figure the grid lines appear as an aid in establishing interesting perceptual relations within the contours. Some examples:

The range of the interior grid has a special relation to the global range for each pattern. It is the narrowest when it contains only unaccented syllables and may be neglected in synthesis. It is about half the global range in the first two patterns and in the compound it is the largest, almost equal to the global range.

With a basis of such relative measures from a larger material and several speakers it would perhaps be possible to construct a quantitative model similar to the one proposed for Chinese to account for variations in tempo (Gårding & Zhang 1987).

Perceptual analyses and considerations may even lead to a tenable definition of intonation which is usually missing or evaded in production based studies and models. For instance, intonation is the global melody of a phrase or sentence to which the local pitch movements are subordinated. In Fig. 3 the pertinent points of the global melodies have been marked by large dots.

3. *Intonation varied. Accentuation and phrasing constant*

In the third set of experiments we want to know how different accentuation patterns behave in different kinds of intonation and speech effort and how they are recognized.

To the left in Figure 5 we have displayed our *långa män* sentences as they have been uttered in declarative rising-falling intonation in a normal degree of speech effort. To the right the same accentuations appear in a falling compressed version that can be expected after a strongly focussed word, in this case *JA* 'yes'. (Part of our Chinese material was designed in the same way. See Figure 2.) The compressed versions are about two thirds of the normal ones in duration. With the same relative measures as we have used before we find that the rhythmic pattern of a given intonation is largely retained when we go from left to right in the figure. We may hypothesize that listeners in their identifications are guided by the relational invariance of the rhythmic pattern.
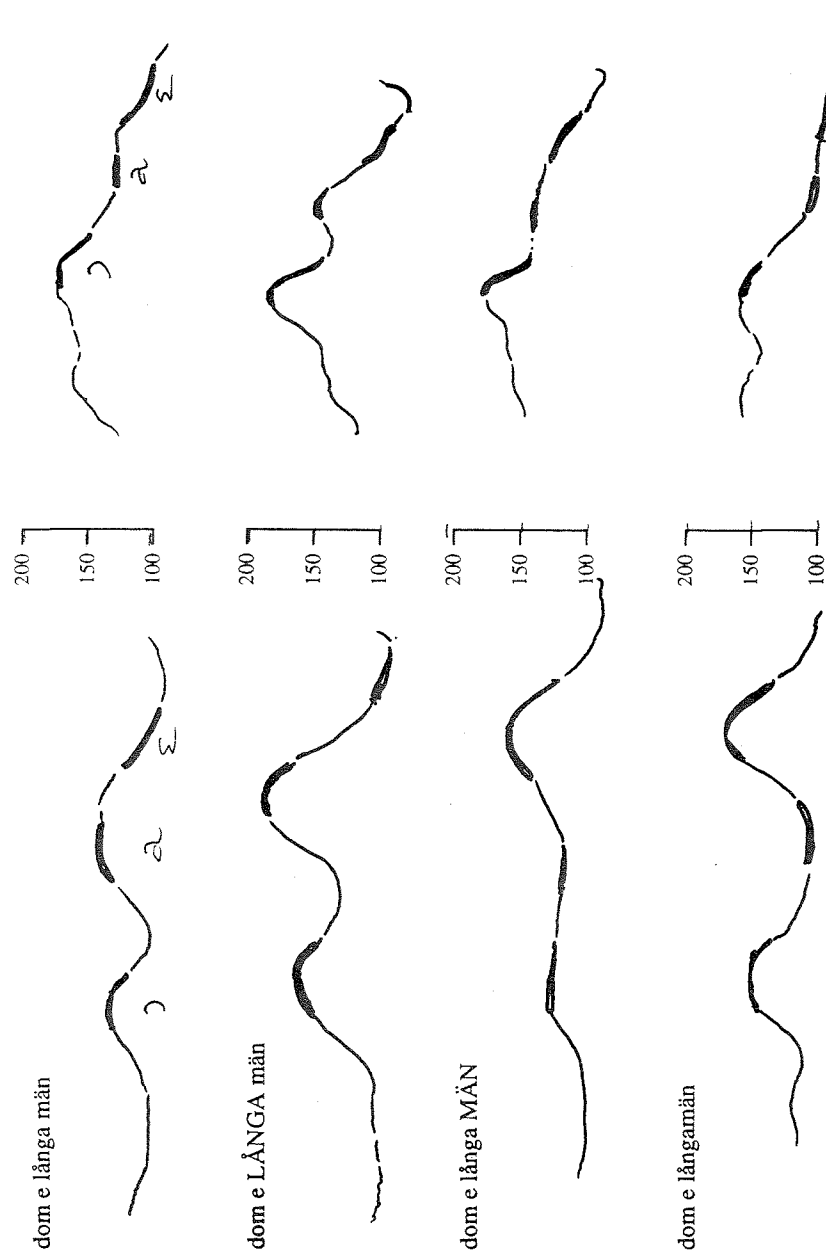
Figure 5. Intonation varied, accentuations constant. To the left the four accentuation patterns of *dom e långa män* in a rising-falling intonation. To the right the same patterns compressed in a falling intonation after a strongly focussed *JA* 'yes', not included in the curve.

Even the tonal patterns have invariant features. We find that the average fundamental frequencies for the vowels in our paradigm occur in the same mutual order on the frequency scale for both intonations. This is immediately noticeable for the vowel [a]. It has its lowest average value in the compound and its highest in focussed *långa*. In between we find the [a] of the other two accentuation patterns. In the compressed patterns the highest and lowest vowels are at a distance of about 6 semitones from each other, in the patterns to the left the distance is twice as large.

Since the average frequency value of the first vowel of the accentuation patterns [ɔ] is very much the same for all the accentuations it cannot serve as a distinctive mark. A particular cue value must be attributed to the difference in frequency between the first and the second vowel. If the difference is large and positive, then the compound is likely to occur, if it is small and negative, the focussed *långa* pattern is the probable alternative etc. This points to a relational invariance in the tonal part of the patterns in our paradigm which is retained independently of the intonation. At least two syllables are needed for the listener to make a successful guess and this points to the *foot,* as defined in classical metrics, as a possible significant unit in accent analysis. The significance of the foot for accent and intonation analysis was observed by Grønnum for Danish (Thorsen 1984, for a summary of her works, see Grønnum 1990), by Fretheim for Norwegian (1978) and by Bruce for Swedish (1987). The importance of the intersyllabic relations also recalls Lehiste's and Ivić's conclusion (mentioned above) that it is the bisyllabic sequence which is the carrier of the accent contrast in Neoštokavian, not the peak position of the first accented syllable.

Another relational invariance concerns the entire pattern, namely the position of the three vowels relative to the general direction of the falling intonation. In all cases, the frequency averages of the mid vowels are either above (the first two sentences) or below (the last two) the mean of the average frequencies of the other two.

The intersyllabic relational invariance in the paradigm of patterns explains how a listener may be able to identify an accentuation in a rising intonation with the same pattern in a falling one. What is absolutely a rise in one intonation may be a fall in another. (See e.g. Figure 5.) We may conclude that listeners in their identifications make corrections not only for the tempo but also for the global amplitude and intonation of the phrase.[6] In

---

[6]Results of experiments conducted by Pierrehumbert 1979 and Terken 1989 indicate that listeners make correction for intonation in their judgements of accentual prominence: for two

this way invariant mental representations of tonal patterns as well as rhythmic patterns may come about. Transfer phenomena in a foreign accent give evidence for the existence of such dissociated patterns. (For examples, see Touati 1987.)

I believe that the tonal invariance becomes natural if we regard the local tonal characteristics as subordinate to and constrained by the phrasal intonation. If we disregard the focus manifestation in the last sentence to the left it is easy to visualize the compressed patterns as compressed versions of the left ones superimposed on a falling intonation.

Some principles in intonation analysis

In this section I shall relate my own view of intonation to that of other researchers in a discussion of some principles which concern intonation analysis in general.

Coordination of local and global features

Grid and superposition. In the model that I have proposed the local accents or tones are coordinated with global intonation by being superposed on a global expression of intonation, the tonal grid.

Superposition is not a phonological concept. It is a technical term which is used for the addition of local movements to a global one. It can be used successfully when the slow movement is approximately linear as in regularly rising, falling or level stretches of speech. Together with the grid it gives a realistic picture of the interaction of local and global features in intonation and at the same time it reflects a simple mechanism that speakers may use in the production of their prosodic patterns. In perception a listener may recognize such superposed patterns by their constant rhythmical characteristics, aided by tonal features which are constant in relation to phrase intonation.

Superposition and grid range make it possible to represent the basic elements of an intonation model in a parsimonious way. In the generative program proposed in our earlier work (e.g. Gårding 1984) the tonal grid, expressing phrasing and speech act, was generated before the local tone marks which in a language with distinctive lexical accents or tones have rather fixed positions not only in relation to the vocalic segments but also to the grid.

---

accented syllables to be heard as equally prominent, the second should have lower pitch in falling intonation.

I think this order of events is well motivated for several reasons: (1) It is natural to see a small movement as added to the large movement. (2) The modified form of a tone or accent shape in connected speech will be practically 'gratis' without much need of e.g. tone sandhi rules or accent sandhi rules (see e.g. Shi 1990). In this way we can go straight from the phonological primitive (HL or LH, etc) and another primitive, the tonal grid, to a good approximation of the surface form. (3) Psycholinguistic considerations: Speech is always some kind of speech act. An accent or tone cannot be uttered without modality, not even a citation form. A child signals modality very early and only later learns to insert lexical items in prosodic modality patterns (Waterson, e.g. 1991).

Superposition models are not unusual in intonation research. They go back to the physiology-based model proposed by Öhman 1968 and his basic idea that sentence and word intonation should be seen as separate commands to a laryngeal component and that these commands, if added in different ways, could generate the variety of Fo output required by the Swedish dialects. Öhman's model has inspired not only my own work but also that of Fujisaki (e.g.1991) and Grønnum (for an overview of her work see Grønnum 1992).

Tonal grid, register, tonal frame
Clements, in his discussion of two contributions to a phonology meeting, one by Inkelas & Leben and the other by Ladd (see Kingston & Beckman 1990) welcomes the notion of *register* as it is represented in the phonological system of Hausa (Inkelas & Leben) and English (Ladd). Register (not to be confused with its use in Asian tone languages) is defined by Clements (1990:59) as a smaller interval or frequency band internal to the speaker's *range* which determines the highest or lowest frequency within which tones can be realized. Registers can be shifted upwards and downwards. The shift may affect all subsequent tones in the new register. By this definition, registers are represented by piecewise constant levels which change from time to time (see Ladd 1990:40).

In his discussion Clements introduces his own new development of the concept (p. 66), which shows that not only piecewise constant tone levels are taken into consideration as in the *tonal frame,* proposed in his earlier work (Clements 1979) but also *register ramps* with varying directions. In mathematical terms he replaces the piecewise constant structure with a piecewise linear one.

Figure 6 shows the difference between piecewise constant and piecewise linear configurations. According to Clements piecewise linear configurations would make it possible to include rising and falling register ramps of certain well defined domains such as the tone-group into the phonology. In this way the register ramp becomes very similar to a tonal grid.

*Downstepping, declination and falling intonation*
It has been generally recognized that there are at least two factors involved in a falling intonation (downtrends), a phonetic factor due to a physiological process resulting in a gradual lowering of pitch over the phrase and a phonological factor, controlled by the speaker to express speech act, e.g. declarative intonation, which in African tone languages is often associated with *downstepping*. Downstepping was introduced by analysts of African tone languages to describe the declining pitches of the tones that they heard towards the end of a phrase. In general these scholars focussed on the tones and were not overly concerned with global intonation (with some notable exceptions, see references in Clements).

A new term, *catathesis*, was introduced by Poser 1984 who recognized three factors in downtrends, apart from the two mentioned above, also final lowering. Pierrehumbert & Beckman 1988 in their analysis of Japanese tone structure leave the term downstep to scholars of African languages and distinguish between *catathesis* (apparently synonymous with downstep and falling intonation) for the phonologically controlled pitch lowering of successive accents after a 'triggering' high tone and *declination* for the gradual lowering of pitch at the end of an utterance due to uncontrolled phonetic factors.

With phrase intonation as a general autonomous basic element it may not be necessary to distinguish between downstepping and other forms of falling intonation. Figure 2 shows how in Chinese the resulting combination of what I interpret as falling phrase intonation and superimposed compressed tones have a downstepped, stair-case like character in one tone sequence and a mere declining slope in another. (A similar staircase is present in the Swedish example in Figure 1.) In the quoted examples at least, the difference between falling intonation and downstepping seems to depend on the visual form of the lexical items which happen to be involved in the contour.

Normal   Expanded   Contracted
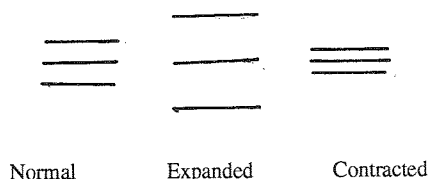
(a) Tonal frame: Piecewise constant tone levels in three different expansions. (Clements 1979:549.)

H H   H L   H   H L   L %
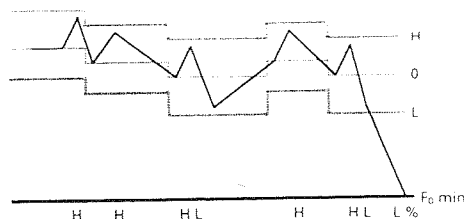
(b) Ladd's piecewise constant register. Model of idealized Fo data for the sentence *When Harry arrived at the airport he was arrested immediately.* (Ladd 1990:40.)

H H   H L   H   H L   L %
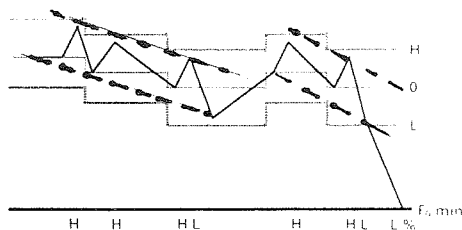
(c) Piecewise linear register applied to Ladd's idealized Fo data.

Normal   Expanded   Compressed
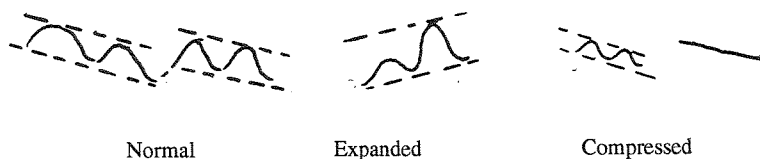
(d) Tonal grid: Piecewise linear pitch levels with idealized Fo data in four different expansions. (Gårding 1987.)

Figure 6. Piecewise constant and piecewise linear tone levels.

Linear and hierarchical structure

A transcription of an Fo contour in terms of a linear sequence of Highs and Lows or markings describing the auditive impression of the contour is valuable for many purposes. However, it does not by itself reflect the hierarchical structure of intonation. A representation which mirrors the hierarchy is desirable.

In autosegmental analyses hierarchy is very much in focus. In an autosegmental representation for Japanese (Pierrehumbert & Beckmann 1988:21) the local phonological entities (H, L, HL) from the tone tier are assigned to the units on the upper tiers which represent nodes in a tree structure. For Japanese six levels are relevant: mora, syllable, word, accentual phrase, intermediate phrase and utterance. The phonological input to the pitch algorithm is a phrasal sequence equipped with tone marks, whose assignment to the relevant segments and pitch values are determined by rules based on the results from a number of experiments. The algorithm used to simulate actual Fo contours (e.g. p. 177) builds up the contour from these local marks. It is only indirectly connected with the autosegmental analysis, since it does not show any hierarchy and the steps in the algorithm do not reflect the order established in the theoretical representation. The important steps (apart from smoothing and devoicing) are linear interpolation between the pitch marks and declination. In spirit the model is very similar to the one presented by Bruce 1977.

In his analysis of English intonation, Hirst proposes a simpler system with three levels. He believes that a phonetic model may be considerably sparser than a typical autosegmental representation would suggest (Hirst 1991). In this respect my own view of intonation is similar to that of Hirst.

My model mirrors a structure which can be informally described as follows: The basic string is an utterance represented by segments including marks for quantity, marks for accents on three levels, boundaries of different weight and modality. This string reflects the intention of the speaker and is the blueprint for the final contour. The rest of the program is the phonetic implementation. An important point is that, in contrast to other similar schemes, the global features are generated first: The highest level, representing speech act, controls the nature and assignment of the pivots and the pivot range (R) is the scale setter for the ranges of the subordinated phrases (grids). At the same time the pivot rearranges the rhythmical structure of the input. The position and nature of a pivot determine the direction of the grids. Finally the accentuation patterns

indicate where to insert the turning points between which the contour is completed by interpolation.

Conclusion

The main point of the early part of my paper was to emphasize the importance of perceptual analysis. The results of some experiments suggested that perceptual analyses of tonal and rhythmical properties, particularly on the phrase and sentence level, may contribute to a better understanding of the nature of intonation. In addition, if representations of the phonological units on the two sides of the speech process, production and perception, are kept apart, less confusion in cross-language comparison can be expected.

In the latter part of the paper, I have confined myself to only a few principles that I consider important in intonation analysis. Yet, I hope that this discussion conveys the general idea that intonation is more similar across languages than existing models, parameters, notational systems, and terminology would suggest.

References

Alnæs, I. 1916. *Norsk Sætningsmelodi.* Kristiania: Aschehoug.

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

Bruce, G. 1982. 'Developing the Swedish intonation model'. *Working Papers* 22, 51-117. Dept. of Linguistics, Lund University.

Bruce, G. 1984. 'Aspects of Fo declination in Swedish'. *Working Papers* 27, 51-63. Dept. of Linguistics. Lund University.

Bruce, G. 1986. 'How floating is floating accent?'. In K. Gregersen & H. Basbøll (eds.), *Nordic Prosody* IV, 41-49. Odense University Press.

Bruce, G. & E. Gårding. 1978. 'A prosodic typology for Swedish dialects'. In E. Gårding, G. Bruce & R. Bannert (eds.), *Nordic Prosody*, 219-228. Department of Linguistics. Lund University.

Clements, G.N. 1979. 'The description of terraced-level tone languages'. *Language* 55:3, 536-558.

Clements, G.N. 1990. 'The status of register in intonation theory: comments on the papers of Inkelas & Leben and Ladd'. In J. Kingston & M. Beckman (eds.), *Papers in laboratory phonology I,* 58-71. Cambridge University Press.

Cooper, W.E. & J.E. Sorensen. 1981. *Fundamental frequency in sentence production.* New York: Springer.

Fant, G. 1954. 'Phonetic and phonemic basis for the transcription of Swedish word material'. *Acta otolaryngologica,* Suppl. 116.

Fretheim, T. 1984. 'Discrete features of East Norwegian intonation contours'. In C.-C. Elert, I. Johansson & E. Strangert (eds.), *Nordic Prosody III,* 67-77. Almqvist & Wiksell.

Fujisaki, H.1991. 'Modeling the generation process of Fo contours as manifestation of linguistic and paralinguistic information'. 12th International Congress of phonetic sciences. Aix-en-Provence.

Gårding, E. 1967. 'Prosodiska drag i spontant och uppläst tal'. (Prosodic features of spontaneous and read speech). In G. Holm (ed.), *Svenskt talspråk,* 40-85. Stockholm: Almqvist & Wiksell.

Gårding, E. 1973. 'Toward a prosodic typology for Swedish accents'. In K.H. Dahlstedt (ed.), *Proceedings of the Second International Conference of Nordic and General Linguistics,* 466-474. Umeå.

Gårding, E. 1976. 'The importance of turning points for the pitch patterns of Swedish accents.' In L.M. Hyman (ed.), *SCOPIL* 4, 27-35. Los Angeles.

Gårding, E. 1981. 'Contrastive prosody: a model and its application'. *Studia Linguistica* 35, 146-65.

Gårding, E. 1983. 'A generative model of intonation'. In A. Cutler & R. Ladd (eds.), *Prosody: Models and measurements,* 11-21. Springer.

Gårding, E. 1984. 'Comparing intonation'. *Working Papers* 27, 75-99. Dept. of Linguistics, Lund University.

Gårding, E. 1987a. 'How many intonation models are there in Lund?'. *RUUL* 17, 141-148. Dept. of Linguistics, Uppsala University.

Gårding E. 1987b. 'Speech act and tonal pattern in Standard Chinese: Constancy and variation'. *Phonetica* 44:1. 13-29.

Gårding, E. 1991a. 'Perceptual aspects of an intonation model'. *Perilus* XIII, 25-29. Institute of Linguistics, University of Stockholm.

Gårding, E. 1991b. 'Intonation parameters in production and perception'. *Actes du XIIeme congres international des sciences phonétiques* 1, 300-305.

Gårding, E. & L. Eriksson. 1991. 'On the perception of Swedish phrase patterns'. *Working Papers* 38, 45-70. Dept. of Linguistics, Lund University.

Gårding, E. & D. House. 1987. 'Production and perception of phrases in some Nordic dialects'. In P. Lilius & M. Saari (eds.), *The Nordic languages and modern linguistics* 6, 163-175. Helsinki University Press.

Gårding, E., P. Kratochvíl, J.-O. Svantesson & J. Zhang. 1986. 'Tone 4 and Tone 3 discrimination in modern standard Chinese'. *Language and Speech* 29:3, 281-293.

Gårding, E. & J. Zhang. 1987. 'Tempo and prosodic pattern in Chinese and Swedish'. *Working Papers* 31, 11-41. Dept. of Linguistics, Lund University.

Grønnum, N. 1992. *The groundworks of Danish intonation*. Museum Tusculanum Press.

Hart, J. 't, R. Collier & A. Cohen. 1990. *A perceptual study of intonation*. Cambridge: Cambridge University Press.

Herz, S. 1990. 'The Delta programming language: an integrated approach to nonlinear phonology, phonetics, and speech synthesis'. In J. Kingston & M. Beckman (eds.), *Papers in laboratory phonology* 1, 215-257, Cambridge University Press.

Hirst, D. 1991. Intonation models: towards a third generation. *Actes du XIIeme congres international des sciences phonétiques* 1, 305-310.

Horne, M. 1987. 'Towards a quantified, focus-based model for synthesizing sentence intonation in English'. *Working Papers* 31, 41-86. Dept. of Linguistics, Lund University.

House, D. 1990. *Tonal perception in speech*. Travaux de l'Institut de Linguistique de Lund 24. Lund University Press.

Huber, D. 1988. *Aspects of the communicative function of voice in text intonation*. Doctoral dissertation, Dept. of computational linguistics, University of Göteborg.

Inkelas, S. & W. Leben. 1990. 'Where phonology and phonetics intersect: the case of Hausa intonation'. In J. Kingston & M. Beckman (eds.), *Papers in laboratory phonology* I, 17-34. Cambridge University Press.

Kingston, J. & M. Beckman (eds.). 1990. *Papers in laboratory phonology 1. Between the grammar and physics of speech*. Cambridge University Press.

Kohler, K. 1987. 'Categorical pitch perception'. *Proc. of the 11th congress of Phonetic Sciences* 5, 331-333. Tallinn: Academy of Sciences of the Estonian S.S.R.

Ladd, R. 1983. 'Phonological features of intonational peaks'. *Language* 59, 721-59.

Ladd, R. 1990. 'Metrical representation of pitch register'. In J. Kingston and M. Beckman (eds.), *Papers in laboratory phonology* 1, 35-57. Cambridge University Press.

Lehiste, I. & P. Ivić. 1986. *Word and sentence prosody in Serbocroation*. Cambridge: MIT Press.

Lindau, M. 1986. 'Testing a model of intonation in a tone language'. *JASA* 80, 757-764.

Malmberg, B. 1967. *Structural linguistics and human communication*. Berlin: Springer.

Meyer, E.A. 1937. *Die Intonation im Schwedischen, I: Die Sveamundarten*. Studies Scand. Philol. 10. University of Stockholm.

Meyer, E.A. 1954. *Die Intonation im Schwedischen, II: Die norrländischen Mundarten*. Studies Scand. Philol. 11. University of Stockholm.

Mountford, K. 1983. *Bambara declarative sentence intonation*. Research in phonetics No 3. Indiana University, Bloomington.

Öhman, S.E.G. 1968. 'A model of word and sentence intonation'. *STL-QPSR* 22-3, 20-54.

Pierrehumbert, J.B. 1979. 'The perception of fundamental frequency declination'. *JASA* 66, 363-369.

Pierrehumbert, J. B. & M.E. Beckman. 1988. *Japanese tone structure*. Cambridge: The MIT Press.

Poser, W.J. 1984. *The phonetics and phonology of tone and intonation in Japanese*. Doctoral dissertation. MIT.

Shen, Xiao-nan Susan. 1992. Personal communication.

Shi, B. 1990. *A speech synthesis-by-rule system for Modern Standard Chinese*. Dept. of Phonetics & Linguistics. University College, London.

Sievers, E. 1912. *Rhytmisch-melodische Studien*. Heidelberg.

Terken, J.M.B. 1989. 'Fundamental frequency and perceived prominence of accented syllables'. *IPO Annual progress report* 24, 33-42.

Thorsen, N. 1984. 'Variability and invariance in Danish stress group patterns'. *Phonetica* 41, 88-102.

Touati, P. 1987. *Structures prosodiques du suédois et du français. Profils temporels et configurations tonales*. Travaux de l'institut de linguistique de Lund XXI. Lund University Press.

Waterson, N. 1991. *Nursery rhymes: Have they a role in language development?* Child language research institute. Paper 7. Lund University.