# INTONATION PARAMETERS IN PRODUCTION AND PERCEPTION

**Eva Gårding**

**Lund University, Sweden**

## ABSTRACT

My contribution is a discussion of perceptual aspects of the acoustic parameters of a Swedish intonation model which has been applied to several prosodic systems [5,7,20]. I emphasize the importance of perceptual analysis and close with examples showing how a phonological analysis may depend on the descriptive level used by the analyst.1)

## 1. A PRODUCTION-ORIENTED INTONATION MODEL

The model for Swedish intonation that I have worked with is oriented towards production. One goal is to reproduce the intonation of an utterance in the form of an Fo curve for a temporally structured string of words with labels representing the lexically distinctive accents (Accent 1 and Accent 2), phrase and sentence level accents, and labels indicating phrasing, sentence type and dialect. The generative scheme and its rules are based on analyses of materials chosen to explore the effect of the main communicative functions of intonation on the Fo curve.2) Superposition is a guiding principle in the analysis which is reflected in the generative scheme. This means that a local movement pertaining to an accent or tone is seen as added to a slow movement (global or semiglobal) representing sentence or phrase intonation [3,7].

The model makes use of the following parameters, which are all visual correlates of acoustic events. Fig.1.

1. *Turning points* are the local maxima or minima of Fo which are related to an accented syllable in a systematic way. The timing of these points, which is largely 'fixed' to specified syllables in different intonations, is crucial for the distinction between the accents and typical for a particular dialect.

2. The *tonal grid* is the auxiliary parallel lines enclosing curve pieces with a uniform general direction (see figure). Drawn in this way the upper line of the grid passes through local maxima of the curve (marked by H(igh) in the figure) and the lower line through local minima (marked by L(ow). In practice not all the local maxima or minima reach the grid lines, which is a tonal reflex of the varying accentuation levels carried by the syllables enclosed by the same grid. In synthesis the grid is generated before the accent target points.

3. *Pivots* are those places in the curve where the grid changes direction or range. They usually occur in connection with focus and demark constituents in the global contour.

*Interpolation* is used in the generative program to fill in the blanks of the Fo pattern between the target turning points.

This model, which captures the gross features of Swedish intonation, makes it possible to describe the accent contrast as a matter of *timing* the Fo target points in relation to to the accented syllable. The dialectal variation in intonation becomes a result of different timing rules for word accent as well as sentence accent (nuclear stress). Modality and weighting are expressed by the tonal grid.

With this description the common feature of the relation between the accents in various dialects is that the accentual movement of A2 comes later than that of A1. Interpreted at a

physiological level, the underlying accentual gesture is similar but timed differently in relation to the accented syllable. This interpretation has support in EMG data obtained from pitch controlling muscles [10].

The feature 'delayed peak' for accentual timing differences in English has been suggested by Ladd [17] and plays a role in the intonation analysis presented for German by Kohler [16].

## 2. PERCEPTION, A MISSING LINK

The explanatory value of the peak analysis is limited to the production side of speech. For Swedish it implies that speakers use their laryngeal muscles in a similar way to raise and lower pitch in order to produce a given tonal pattern and that the Swedish accents can be produced by timing a similar underlying laryngeal gesture in different ways in relation to the accent carrying syllable. The speaker's intention, on the other hand, is to reproduce tonal and rhythmical patterns and these patterns should also be described at a level that relates to perception. Such a description is important for the understanding of perception mechanisms in general, for the understanding of how prosodic patterns are acquired, learned and transferred to a foreign dialect or language, and for a comprehensive theory of intonation.

The following discussion of the possible perceptual relevance of the acoustic parameters described above are inspired by ongoing research, planned to explore perceptual correlates of intonation and accentuation.

1. *Turning points.* A given pitch curve can be efficiently reconstructed either from information about the position in time and frequency of the turning points or from suitable information about the falls and rises that make up the curve. Some of the turning points have very stable locations relative to the C/V boundary of the accented syllable which means that also the following falls and rises are relatively stable. Physiologically speaking a turning point indicates an onset-offset change in a nerve signal and like other transient signals this ought to have a strong attention-sharpening effect. It may correspond to what a neurologist would call a feature-trigger [e.g. Granit 13].

If we aim at a perceptual description of intonation, the falls and rises (or contrasts between two frequency levels) are more important than the turning points. Only a phonetician who has worked long enough with acoustic records can perceive the position of a turning point but almost anybody can hear a rise or fall if it is given the proper duration.

An experiment in which Fo peaks of different shapes were shifted in small steps in time over constant carriers has shown that a prosodic phrase pattern can be recognized when the position of the peak has produced patterns in which pitch movements over neighbouring vowels are prototypical. Our experiment suggests that an Fo peak is only indirectly important. It is the adjoining ramps over the vowels which have perceptual reality [8, 9].

2. *The tonal grid,* which may be level or exhibit various degrees of fall and rise, encloses a phrase pattern made up of syllables with 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 exemplfied in Fig.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 and matched with stored prototypes.

Since a listener is capable of identifying an accentuation pattern in different intonations, it is bound to have some invariant features. Our recent work with contrastive Swedish accentuation patterns confirms that their relative patterns of duration, intensity and Fo are retained in different intonations. In

particular, pitch relations between neighbouring vowels are perceptually important. In a given intonation an accentuation pattern is characterized by its typical pitch relations over neighbouring vowels which sets it off from other patterns. (Fig.3.)

From a perceptual point of view, then, intonation, represented by the grid, carries relative invariant features of an accentuation pattern which recur in a global rising, falling or level Fo pattern with different communicative values.

3. The perceptual importance of the *pivots* stems from a discontinuity of 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 getters. The new direction or range is the important signal. In Swedish the position of the main pivotal point of an Fo curve in relation to the text is dialect dependent. In South Swedish declaratives the main pivot occurs on the accented syllable of the focussed word and in Central Swedish dialects it has a position after the accented word, described as 'floating' by Bruce [2].

The *interpolation* follows a mechanical rule which has the purpose of connecting two target points with each other. It has been interesting to note in ILS experiments that this interpolation can neglect small-scale pitch obtrusions caused by the accents without any noticeable effect on perception [12]. In this way, a description of the curve in terms of perceived local Highs and Lows may differ from a description relying rigorously on production data. Our phonological interpretation of the ILS experiments is that the accent domain may encompass not only the following unaccented syllables up to the next accent (which is a current definition) but also the deaccented ones.

## 3. REPRESENTATION

Highs and Lows may be used to show for a given Fo curve how 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 may yield High for A2 and Low for A1, marking a high or low turning point near the beginning of the vocalic segment as the distinctive feature [2]. A phonetic rule would generate the following pitch movement. The representation makes no claim to perceptual validity.

In an intonation analysis of Hausa (traditionally 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 [18]. This finding suggested that it is the latter part of the syllable that has perceptual salience rather than the earlier part. Further support comes from perceptual experiments with Chinese [11] and with Swedish nonsense syllables [15].

## 4. CONCLUDING REMARK AND QUESTION

Our discussion leads to the following conclusion:

Parameters and their representations differ depending on the orientation of the model and the descriptive level chosen by the analyst.

It also leads to an interesting question: What side of speech should the phonological units represent, production, perception, both or neither?

1) The necessity of distinguishing between levels of description in speech analysis, has been strongly argued for by Repp [19].

2) In classical tradition we have considered the lexical-distinctive function,the grouping function with demarcative and connective features and the weighting, modal and expressive functions.

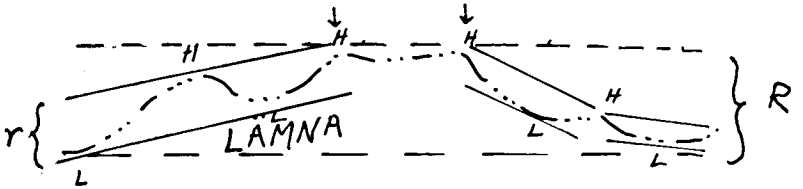


Fig.1. **Parameters in the production-oriented model.** Fo curve of *Man vill LÄMNA* (A2) *nåra långa* (A2) *nunnor* (A2) in declarative intonation (One wants to deliver some long nunnies). A2 is manifested as HLH turning points in focus, after focus as HL. Parallel lines represent the grid. 'R' denotes global range, 'r' is the vertical distance between the grid lines. The curve is from Bruce [1].

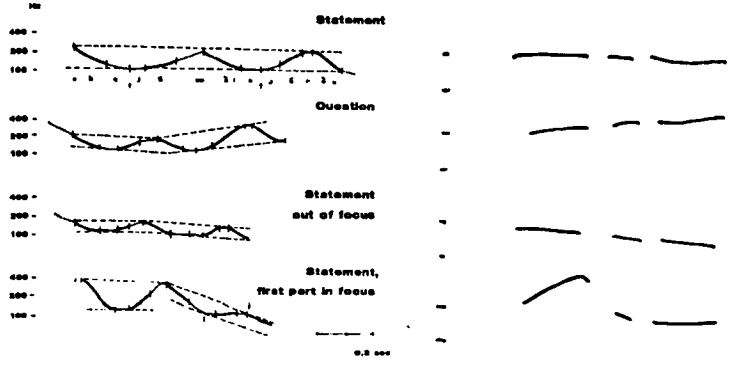


Fig.2. **Deformation of tone shapes** due to (semi)global intonation can be explained by the superposition principle. (a) Chinese *Sùn Yán mài niúròu* (Sun Yan sells beef) with alternating falling and rising tones, (b) *Wāng Yī chōu xiāngyān* (Wang Yi smokes cigarettes) with high tones only [6].

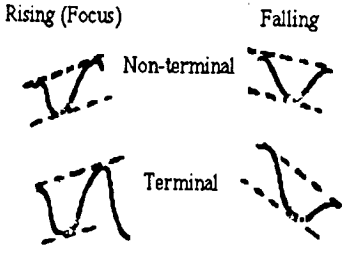


Fig.2(c).**Deformation of accent shapes.** Swedish *bara* A2 ('only' and placename) in rising and falling intonation. [4].

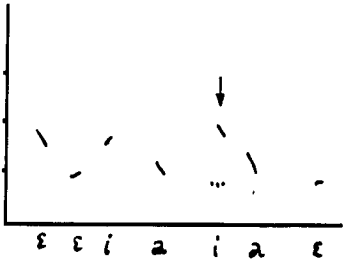


Fig.3. **Perceptual importance of intersyllabic relations.**Pitch curve in maximally unvoiced context.*Ester* (A1) *fick FATT(*A1) *i katten(*A1) (Ester got hold of the cat). Lowering the post-accented syllable (arrow) shifts categorization from A1 to an A2 compound (fatti(g)katten(A2) [11].

## A NOTE ON INTONATION MODELS

Intonation models differ in scope and orientation. Moreover, to list some of the qualifying labels that have been used, they may be qualitative or quantitative, descriptive or generative, abstract or concrete. Analyses may depart from different theoretical assumptions (sometimes biassed by the analyst's native prosody!) and rely on different kinds of observations and measurements, described and interpreted at different levels in speech production or speech perception.

There is general agreement on the close correlation between intonation and Fo but there is disagreement on the best decomposition of a pitch curve, horizontally as well as vertically.

Evaluation of models seems to be simpler, at least if we adhere to the following criteria: A good model should be descriptivly adequate, general, parsimonious and have explanatory power.

## 5. REFERENCES


[1] BRUCE, G. (1977), Swedish word accents in sentence perspective. *Travaux de l'Institut de Linguistique de Lund XII*. Lund:Gleerup

[2] BRUCE, G. (1986), How floating is focal accent? *Nordic Prosody IV*, eds K.Gregesen & H. Basböll, 41-49. Odense University press.

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

[4] GÅRDING, E. (1967), Prosodiska drag i spontant och uppläst tal, *Svenskt talspråk*, ed. G. Holm, 40-85. Stockholm: Almqvist & Wicksell.

[5] GÅRDING, E. (1983), A generative model of intonation, *Prosody: Models and measurements*, eds. A. Cutler & R. Ladd, 11-21. Springer.

[6] GÅRDING E. (1987), Speech act and tonal pattern in Standard Chinese: Constancy and variation. *Phonetica* 44:1, 13-29.

[7] GÅRDING, E. (1989), Intonation in Swedish, *Working Papers* 35, 63-88. Contribution to *Intonation Systems*, eds. D. Hirst & A. Di Cristo.

[8] GÅRDING, E. & ERIKSSON L. (1989), Perceptual cues to some Swedish prosodic phrase patterns - a peak shift experiment. Fonetik 89. *STL-QPSR* 1/1989, 13-16.

[9] GÅRDING, E. & ERIKSSON L. (1991), On the perception of prosodic phrase patterns. To be published.

[10] GÅRDING, E., FUJIMURA O., HIROSE H. & SIMADA Z. Laryngeal control of Swedish word accents. *Working Papers* 10, 53-82. Dept of Linguistics, Lund University.

[11] GÅRDING, E., KRATOCHVIL, P., SVANTESSON, J.O. & ZHANG J.(1986),Tone 4 and Tone 3 discrimination in modern standard Chinese. *Language and Speech* 29:3, 281-293.

[12] GÅRDING, E. & STENBERG, M. West Swedish and East Norwegian intonation. *Nordic Prosody V*, eds. K. Wiik & I. Raimu, 111-131, Turku.

[13] GRANIT, R. (1977) The purposive brain.The MIT Press.

[14] HIRST, D. & DI CRISTO, A. (1991), Intonation systems. To be published.

[15] HOUSE, D. (1990), Tonal perception in speech. *Travaux de l'Institut de Linguistique de Lund* 24. Lund University Press.

[16] KOHLER, K. (1987), Categorical pitch perception. *Proceedings of the 11th Congress of Phonetic Sciences*, 5:331-333, Tallinn: Academy of Sciences of the Estonian S.S.R.

[17] LADD, R. (1983), Phonological features of phonological peaks. *Language* 59, 721-59.

[18] LINDAU, M. (1986) Testing a model of intonation in a tone language. *J.A.S.A.* 80: 757-764

[19] REPP, B. (1981) On levels of description in speech research. *J.A.S.A.* 5: 1462-64.

[20] TOUATI, P. (1987), Structures prosodiques du suédois et du francais. *Travaux de l'Institut de Linguistique de Lund* 21. Lund University Press.


( For further references see [7] )