

# Modelling compound intonation in Dala and Gotland Swedish

Susanne Schötz<sup>1</sup>, Gösta Bruce<sup>1</sup>, Björn Granström<sup>2</sup>

<sup>1</sup>Department of Linguistics & Phonetics, Centre for Languages & Literature, Lund University

<sup>2</sup>Department of Speech, Music & Hearing, School of Computer Science & Communication, KTH

## Abstract

As part of our work within the SIMULEKT project, we are modelling compound word intonation in regional varieties of Swedish. The focus of this paper is on the Gotland and Dala varieties of Swedish and the current version of SWING (SWedish INtonation Generator). We examined productions by 75 speakers of the compound mobiltelefonen 'the mobile phone'. Based on our findings for pitch patterns in compounds we argue for a possible division into three dialect regions: 1) Gotland: a high upstepped pitch plateau, 2) Dala-Bergslagen: a high regular pitch plateau, and 3) Upper Dalarna: a single pitch peak in connection with the primary stress of the compound. The SWING tool was used to model and simulate compounds in the three intonational varieties. Future work includes perceptual testing to see if listeners are able to identify a speaker as belonging to the Gotland, Dala-Bergslagen or Upper Dalarna regions, depending on the pitch shape of the compound.

## Introduction

Within the SIMULEKT project (Simulating Intonational Varieties of Swedish) (Bruce et al., 2007), we are studying the prosodic variation characteristic of different regions of the Swedish-speaking area. Figure 1 shows a map of these regions, corresponding to our present dialect classification scheme.

In our work, various forms of speech synthesis and the Swedish prosody model (Bruce & Gårding, 1978; Bruce & Granström, 1993; Bruce, 2007) play prominent roles. To facilitate our work with testing and further developing the model, we have designed a tool for analysis and modelling of Swedish intonation by resynthesis: SWING. The aim of the present paper is two-fold: to explore variation in compound word intonation specifically in two regions, namely Dala and Gotland Swedish, and to describe the current version of SWING and how it is used to model intonation in regional varieties of Swedish. We will exemplify this with the modelling of pitch

modelling of pitch patterns of compounds in Dala and Gotland Swedish.

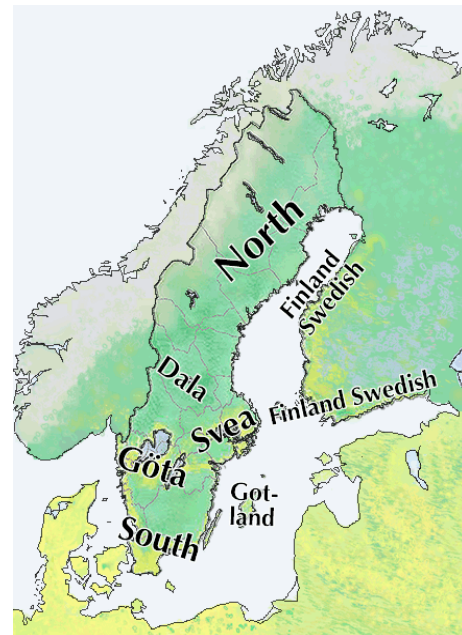


Figure 1. Approximate geographical distribution of the seven main regional varieties of Swedish.

## Compound word intonation

The classical accent typology by Gårding (1977) is based on Meyer's (1937, 1954) pitch curves in disyllabic simplex words with initial stress having either accent I or accent II. It makes a first major division of Swedish intonation into single-peaked (1) and double-peaked (2) types, based on the number of pitch peaks for a word with accent II. According to this typology, the double-peaked type is found in Central Swedish both to the West (Götå) and to the East (Svea) as well as in North Swedish. The single-peaked accent type is characteristic of South Swedish, Dala and Gotland regional varieties. Within this accent type there is a further division into the two subtypes 1A and 1B with some difference in pitch peak timing – earlier-later – relative to the stressed syllable. It has been shown that the pitch patterns of compound words can be used as an even better diagnostic than simplex words for distinguishing

between intonational varieties of Swedish (Riad 1998, Bruce 2001, 2007). A compound word in Swedish contains two stresses, primary stress (ˈ) on the first element and secondary stress (ˌ) on the final element. In most varieties of Swedish a compound takes accent II. The exception is South Swedish where both accents can occur (Bruce, 2007). A critical issue is whether the secondary stress of a compound is a relevant synchronisation point for a pitch gesture or not. Figure 2 shows stylised pitch patterns of accent II compounds identifying four different shapes characteristic of distinct regional varieties of Swedish (Bruce 2001). The target patterns for our discussion in this paper are the two types having either a single peak in connection with the primary stress of the compound or a high plateau between the primary and the secondary stresses of the word. These two accentual types are found mainly in South Swedish, and in the Dala region and on the isle of Gotland respectively. It has been suggested that the pitch pattern of an accent II compound in the Gotland and Dala dialect types has basically the same shape with the high pitch plateau extending roughly from the primary to the secondary stress of the word. The specific point of interest of our contribution is to examine the idea about the similarity of pitch patterns of compound words particularly in Gotland and Dala Swedish.

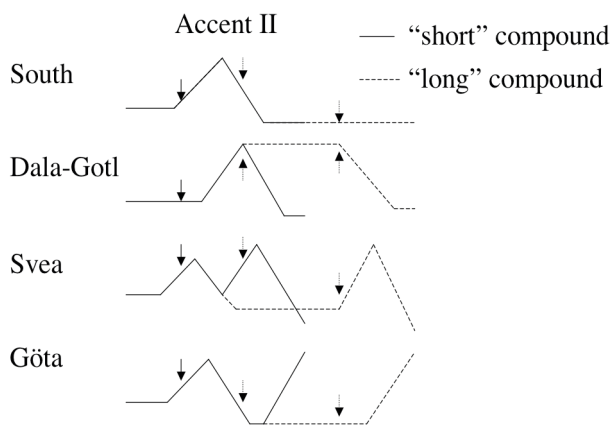


Figure 2. Schematic pitch patterns of accent II compound words in four main intonational varieties of Swedish (after Bruce, 2001). The first arrow marks the CV-boundary of the primary stress, and the second/third arrow marks the CV-boundary of the secondary stress. In a “short” compound the two stresses are adjacent, while in a “long” compound the stresses are not directly adjacent.

## Speech material and method

The speech material was taken from the Swedish SpeechDat (Elenius, 1999), a database containing read telephone speech. It contains speech of 5000 speakers registered by age, gender, current location and self-labelled dialect type according to Elert’s suggested 18 Swedish dialectal regions (Elert, 1994). As target word of our examination, we selected the initial “long” compound /moˈbilteˌfoˌnen/ from the sentence *Mobiltelefonen är nittiotalets stora fluga, både bland företagare och privatpersoner*. ‘The mobile phone is the big hit of the nineties, both among business people and private persons’. Following Elert’s classification, we selected 75 productions of *mobiltelefonen* from the three dialect regions which can be labelled roughly as Gotland, Dala-Bergslagen and Upper Dalarna Swedish (25 speakers of each dialect).

$F_0$  contours of all productions were extracted, normalised for time (expressed as the percentage of the word duration) and plotted on a semitone scale in three separate graphs: one for each dialectal region.

## Tentative findings

Figure 3 shows the  $F_0$  contours of the speakers from the three dialectal regions examined. Even if there is variation to be seen among  $F_0$  contours of each of these graphs, there is also some constancy to be detected. For both Gotland and Dala-Bergslagen a high pitch plateau, i.e. early rise + high level pitch + late fall, for the compound can be traced. A possible difference between the two dialect types may be that, while speakers representing Dala-Bergslagen Swedish have a regular high pitch plateau, Gotland speakers tend to have more of an upstepped pitch pattern for the plateau, i.e. early rise + high level pitch + late rise & fall.

Among the speakers classified as representing Upper Dalarna Swedish there is more internal variation of the  $F_0$  contours to be seen. However, this variation can be dissolved as consisting of two basic distinct pitch patterns, either a single pitch peak in connection with the primary stress of the compound or a high pitch plateau. These two patterns would appear to have a geographical distribution within the area, so that the high pitch plateau is more likely to occur towards South-East, i.e. in places closer to the Dala-Bergslagen dialect region.

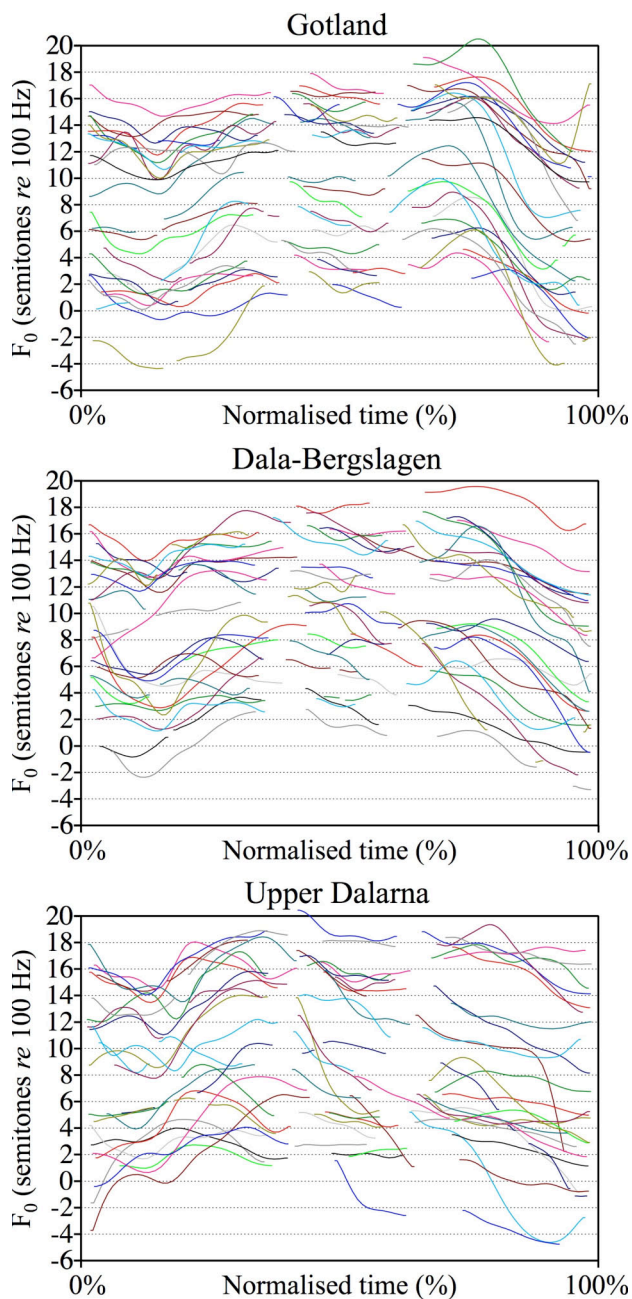


Figure 3. Variation in compound word intonation. F<sub>0</sub> contours of the compound word *mobiltelefonen* (accent II) produced by 25 speakers each of Gotland, Dala-Bergslagen and Upper Dalarna Swedish.

Figure 4 shows examples of typical F<sub>0</sub> contours of the compound word *mobiltelefonen* produced by one speaker from each of the three dialect regions discussed. For the speaker from Gotland the high plateau is further boosted by a second rise before the final fall creating an up-stepped pitch pattern through the compound word. The example compound by the speaker from Dala-Bergslagen is characterised by a high pitch plateau instead, i.e. a high pitch level extending between the rise synchronised with the primary stress and the fall aligned with the secondary stress. A single pitch peak in connec-

tion with the primary stress of the compound followed by a fall and a low pitch level in connection with secondary stress of the word is characteristic of the speaker representing Upper Dalarna Swedish.

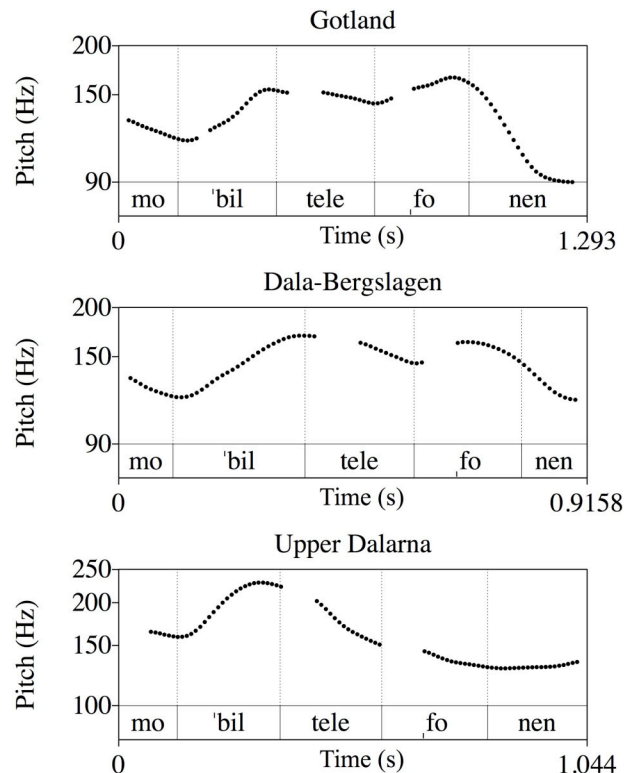


Figure 4. Compound word intonation. Typical example F<sub>0</sub> contours of *mobiltelefonen* produced by a speaker from each of the three dialect regions Gotland, Dala-Bergslagen and Upper Dalarna Swedish.

## SWING

SWING (SWedish INtonation Generator) is a tool for analysis and modelling of Swedish intonation by resynthesis. It comprises several parts joined by the speech analysis software Praat (Boersma & Weenink, 2009), which also serves as graphical interface. Using an input annotated speech sample and an input rule file, SWING generates and plays PSOLA resynthesis – with rule-based and speaker-normalised intonation – of the input speech sample. Additional features include visual display of the output on the screen, and options for printing various kinds of information to the Praat console (Info window), e.g. rule names and values, the time and F<sub>0</sub> of generated pitch points etc. Figure 5 shows a schematic overview of the tool.

The input speech sample to be used with the tool is manually annotated. Stressed syllables are labelled prosodically and the corresponding vowels are transcribed orthographically. Figure

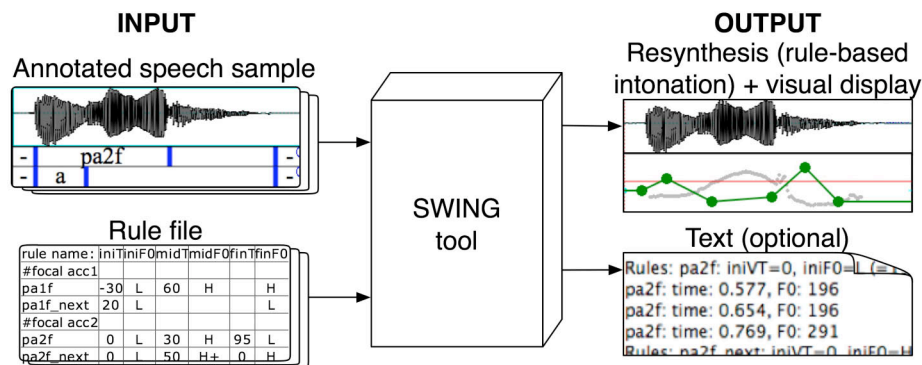


Figure 5. Schematic overview of the SWING tool.

6 displays an example compound word annotation, while Table 1 shows the prosodic labels that are handled by the current version of the tool.

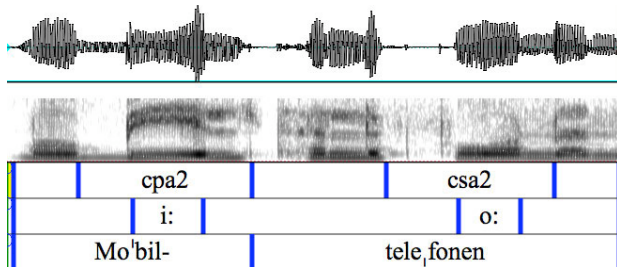


Figure 6. Example of an annotated input speech sample.

Table 1. Prosodic labels used for annotation of speech samples to be analysed by SWING.

Label	Description
pa1	primary stressed (non-focal) accent 1
pa2	primary stressed (non-focal) accent 2
pa1f	focal focal accent 1
pa2f	focal focal accent 2
cpa1	primary stressed accent 1 in compounds
cpa2	primary stressed accent 2 in compounds
csa1	secondary stressed accent 1 in compounds
csa2	secondary stressed accent 2 in compounds

### Rules

In SWING, the Swedish prosody model is implemented as a set of rule files – one for each regional variety of the model – with timing and  $F_0$  values for critical pitch points. These files are text files with a number of columns; the first contains the rule names, and the following comprise three pairs of values, corresponding to the timing and  $F_0$  of the critical pitch points of the rules. The three points are called *ini* (initial), *mid* (medial), and *fin* (final). Each point contains values for timing ( $T$ ) and  $F_0$  ( $F_0$ ). Timing is expressed as a percentage into the stressed syllable, starting from the onset of the

stressed vowel. Three values are used for  $F_0$ : L (low), H (high) and H+ (extra high, used in focal accents). The pitch points are optional; they can be left out if they are not needed by a rule. New rules can easily be added and existing ones adjusted by editing the rule file. Table 2 shows an example of the rules for compound words in Gotland, Dala-Bergslagen and Upper Dalarna Swedish. Several rules contain an extra pitch gesture in the following (unstressed) segment of the annotated input speech sample. This extra part has the word ‘next’ attached to its rule name; see e.g. *cpa2* in Table 2.

Table 2. Rules for compound words in Gotland, Dala-Bergslagen and Upper Dalarna Swedish with timing ( $T$ ) and  $F_0$  ( $F_0$ ) values for initial (*ini*), mid (*mid*) and final (*fin*) points (‘\_next’: extra gesture; see Table 1 for additional rule name descriptions).

	<i>iniT</i>	<i>iniF0</i>	<i>midT</i>	<i>midF0</i>	<i>finT</i>	<i>finF0</i>
<i>Gotland</i>						
cpa2	50	L				
cpa2_next	30	H				
csa2	30	H			70	H+
csa2_next	30	L				
<i>Dala-Bergslagen</i>						
cpa2	50	L				
cpa2_next	30	H				
csa2	30	H				
csa2_next	30	L				
<i>Upper Dalarna</i>						
cpa2	0	L			60	H
cpa2_next	30	L				
csa2						

### Procedure

Analysis with the SWING tool is fairly straightforward. The user selects one input speech sample and one rule file to use with the tool, and which (if any) text (rules, pitch points, debugging information) to print to the Praat console. A Praat script generates resynthesis of the

input speech sample with a rule-based output pitch contour based on 1) the pitch range of the input speech sample, used for speaker normalisation, 2) the annotation, used to identify the time and pitch gestures to be generated, and 3) the rule file, containing the values of the critical pitch points. The Praat graphical user interface provides immediate audio-visual feedback of how well the rules work, and also allows for easy additional manipulation of pitch points with the Praat built-in Manipulation feature.

## Modelling compounds with SWING

SWING is now being used in our work with testing and developing the Swedish prosody model for compound words. Testing is done by selecting an input sound sample and a rule file of the same intonational variety. If the model works adequately, there should be a close match between the  $F_0$  contour of the original version and the rule-based one generated by the tool. Figure 7 shows Praat Manipulation objects for the three dialect regions Gotland, Dala-Bergslagen and Upper Dalarna Swedish modelled with the corresponding rules for each dialect region.

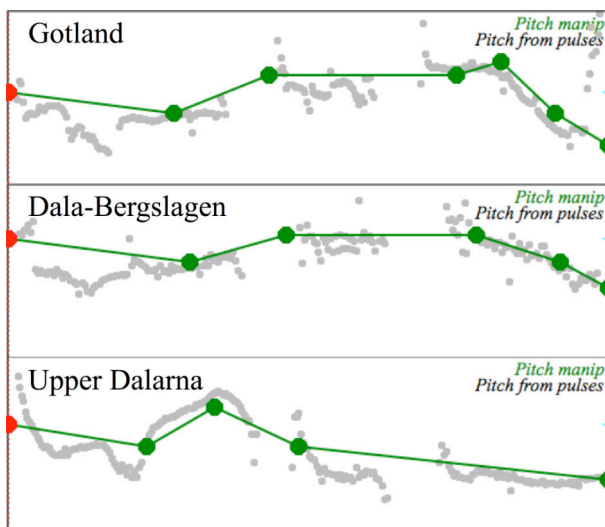


Figure 7. Simulation of compound words with SWING. Praat Manipulation displays of *mobiltelefonen* of the three dialect regions Gotland, Dala-Bergslagen and Upper Dalarna Swedish (simulation: circles connected by solid line; original pitch: light-grey line).

The light grey lines show the original pitch of each dialect region, while the circles connected with the solid lines represent the rule-generated output pitch contours.

As can be seen in Figure 7, the simulated (rule-based) pitch patterns clearly resemble the corresponding three typical compound word

intonation patterns shown in Figure 4. There is also a close match between the original pitch of the input speech samples and the simulated pitch contour in all three dialectal regions.

## Discussion and additional remarks

The present paper partly confirms earlier observations about pitch patterns of word accentuation in the regional varieties of Dala and Gotland Swedish, and partly adds new specific pieces of information, potentially extending our knowledge about compound word intonation of these varieties.

One point of discussion is the internal variation within Dala Swedish with a differentiation of the pitch patterns of word accents into Upper Dalarna and Dala-Bergslagen intonational sub-varieties. This division has earlier been suggested by Engstrand and Nyström (2002) revisiting Meyer's pitch curves of the two word accents in simplex words, with speakers representing Upper Dalarna having a slightly earlier timing of the relevant pitch peak than speakers from Dala-Bergslagen. See also Olander's study (2001) of Orsa Swedish as a case in point concerning word intonation in a variety of Upper Dalarna.

Our study of compound word intonation clearly demonstrates the characteristic rising-falling pitch pattern in connection with the primary stress of a compound word in Upper Dalarna as opposed to the high pitch plateau between the primary and secondary stresses of the compound in Dala-Bergslagen, even if there is also some variability among the different speakers investigated here. We would even like to suggest that compound word intonation in Dala-Bergslagen and Upper Dalarna Swedish is potentially distinct. It would also appear to be true that a compound in Upper Dalarna has got the same basic pitch shape as that of South Swedish. Generally, word intonation in Upper Dalarna Swedish and South Swedish would even seem to be basically the same.

Another point of discussion is the suggested similarity of word intonation between Dala and Gotland Swedish. Even if the same basic pitch pattern of an accent II compound – the pitch plateau – can be found for speakers representing varieties of both Dala-Bergslagen and Gotland, there is also an interesting difference to be discerned and further examined. As has been shown above, Gotland Swedish speakers tend to display more of an upstepped pitch shape for

the compound word, while Dala-Berslagen speakers have a more regular high pitch plateau. Our preliminary simulation of compound word intonation for Dala and Gotland with the SWING tool is also encouraging. We are planning to run some perceptual testing to see whether listeners will be able to reliably identify a speaker as belonging to the Gotland, Dala-Bergslagen or Upper Dalarna regions depending on the specific pitch shape of the compound word.

## Acknowledgements

The work within the SIMULEKT project is funded by the Swedish Research Council 2007-2009.

## References

- Boersma P. and Weenink D. (2009) Praat: doing phonetics by computer (version 5.1) [computer program]. <http://www.praat.org/>, visited 30-Mar-09.
- Bruce G. (2001) Secondary stress and pitch accent synchronization in Swedish. In van Dommelen W. and Fretheim T. (eds) *Nordic Prosody VIII*, 33-44, Frankfurt am Main: Peter Lang.
- Bruce G. (2007) Components of a prosodic typology of Swedish intonation. In Riad T. and Gussenhoven C. (eds) *Tones and Tunes*, Volume 1, 113-146, Berlin: Mouton de Gruyter.
- 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. Lund: Department of Linguistics.
- Bruce G. and Granström B. (1993) Prosodic modelling in Swedish speech synthesis. *Speech Communication* 13, 63-73.
- Bruce G., Granström B., and Schötz S. (2007) Simulating Intonational Varieties of Swedish. *Proc. of ICPhS XVI*, Saarbrücken, Germany.
- Elenius K. (1999) Two Swedish SpeechDat databases - some experiences and results. *Proc. of Eurospeech 99*, 2243-2246.
- Elert C.-C. (1994) Indelning och gränser inom området för den nu talade svenskan - en aktuell dialektografi. In Edlund L.E. (ed) *Kulturgränser - myt eller verklighet*, Umeå, Sweden: Diabas, 215-228.
- Engstrand O. and Nyström G. (2002) Meyer's accent contours revisited. Proceedings from Fonetik 2002, the XVth Swedish Phonetics Conference, Speech, Music and Hearing, Quarterly Progress and Status Report 44, 17-20. KTH, Stockholm.
- Gårding E. (1977) The Scandinavian word accents. Lund, Gleerup.
- Gårding E. and Lindblad P. (1973) Constancy and variation in Swedish word accent patterns. Working Papers 7. Lund: Lund University, Phonetics Laboratory, 36-110.
- Meyer E. A. (1937) Die Intonation im Schwedischen, I: Die Sveamundarten. *Studies Scand. Philol.* Nr. 10. Stockholm University.
- Meyer E. A. (1954) Die Intonation im Schwedischen, II: Die Sveamundarten. *Studies Scand. Philol.* Nr. 11. Stockholm University.
- Olander E. (2001) Word accents in the Orsa dialect and in Orsa Swedish. In Fonetik 2001, 132-135. Working Papers 49, Linguistics, Lund University.
- Riad T. (1998) Towards a Scandinavian accent typology. In Kehrein W. and Wiese R. (eds) *Phonology and morphology of the Germanic languages*, 77-109. Tübingen: Max Niemeyer.