



LUND UNIVERSITY

Computational modelling and generation of prosodic structure in Swedish

Horne, Merle; Filipsson, Marcus

Published in:

Proceedings of the XIIIth International Congress of Phonetic Sciences

1995

[Link to publication](#)

Citation for published version (APA):

Horne, M., & Filipsson, M. (1995). Computational modelling and generation of prosodic structure in Swedish. In K. Elenius, & P. Branderud (Eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences* (Vol. 4, pp. 364-367)

Total number of authors:

2

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

COMPUTATIONAL MODELING AND GENERATION OF PROSODIC STRUCTURE IN SWEDISH

Merle Home and Marcus Filipsson
Department of Linguistics and Phonetics, University of Lund

ABSTRACT

A summary of the motivation for the various levels of structure assumed in a prosodic hierarchy for Swedish and the linguistic and discourse parameters that are needed for their recognition in texts are presented.

INTRODUCTION

Current speech synthesis systems, which lack detailed prosodic structure cannot generate many of the intonational patterns that one observes in natural speech. Prosodic phenomena associated with the boundaries of clause-internal word groups constitute one problem area. More specifically, the transitions that the current Swedish text-to-speech system generates between word accents do not always correspond to those one finds in naturally occurring speech. As the F0 curve in Figure 1 (corresponding to part of the sentence in (1)) shows, the end of the focussed expression *för närvarande* 'presently' coincides with a low F0 point (L#) in the speech of the radio commentator we are modelling. This L#, we claim below, corresponds to the end of the prosodic constituent which we will define as a {+focal} Prosodic Word. In Figure 2, it is observed that the corresponding synthetic F0 curve generated using the current rule system cannot reproduce this pattern since no low point after the focal high is predicted in clause-internal position. The F0 transitions are only triggered by the positions of the word accents which can be either focal (i.e. followed by a H*) or nonfocal (i.e. without an additional following H*). Thus, after the H*L (Accent 2) word accent on the syllable *-ndr-*, there is a rise throughout the remainder of the word (due to an associated focal H*) and the first syllable of the following word, *betecknas* 'is characterized', since the underlying accent pattern of an Accent 1 word like *betecknas* is HL*, with a H on the

premainstress syllable *be-* and a L* on the syllable *-ick-* [1]. Thus, the L# at the end of *för närvarande* such as in Figure 1 cannot currently be automatically generated.

(1) *För närvarande betecknas tendensen som mycket svag* 'At present the trend is characterized as very weak'

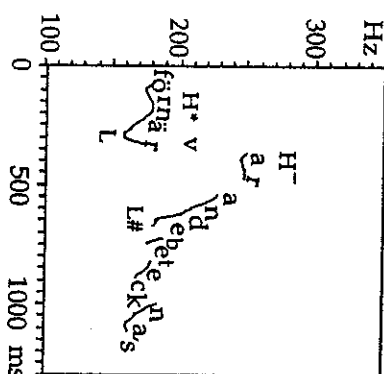


Figure 1. A partial F0 contour for the sentence in (1) uttered by a professional radio commentator.

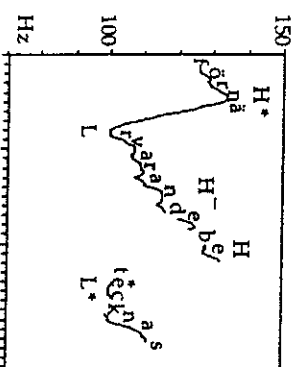


Figure 2. Synthesized F0 contour for the same sentence fragment as in Figure 1.

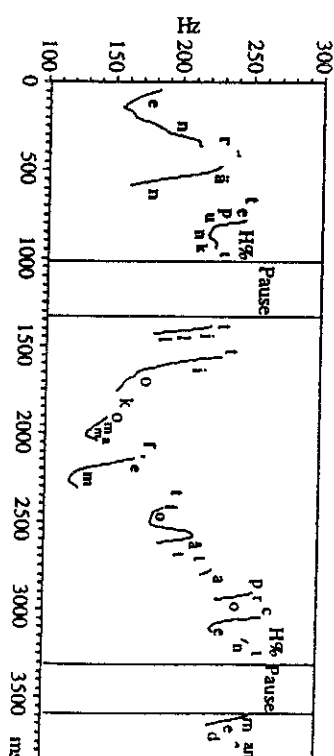


Figure 3. F0 contour for a fragment of the sentence in (2) (1 rättepunkt || till 10,58 procent || median) with a clause-internal PP boundary after 'en rättepunkt'.

Another problem with current synthesis is that one has not been able to predict the location of clause-internal Prosodic Phrase boundaries, i.e. internal boundaries that are as strong as those which occur at the end of the majority of clauses/sentences. This is exemplified, for example, in Figure 3, which presents part of the F0 contour associated with the sentence in (2), where the internal boundary after *rättepunkt* 'interest point' has the same strength as that after *procent* 'percent'.

(2) *1 rättepunkt || till 10,58 procent || median sexmånadersväxlar går upp 5 punkter till 10,50 procent* 'Twelve-month state-debt bonds had gone back 1 point to 10,58 percent while six-month bonds had gone up 5 points to 10,50 percent.' (where || represents a Prosodic Phrase boundary)

In order to be able to recognize such internal Prosodic Phrase boundaries, one must have access to more lexico-grammatical information than is currently available in text-to-speech systems.

SWEDISH PROSODIC STRUCTURE

Prosodic Word (PW)

Three levels of prosodic structure are being assumed over the level of the syllable [2]. The smallest of these is the Prosodic Word (PW) which is defined as corresponding to a content word and

any following function words up to the next content word within a given Prosodic Phrase (PPh). At the beginning of a PPh, the PW can also begin with one or more function words.

The PW is characterized by a word accent. It is also marked by a boundary tone which is realized by a final rise in the case where the content word is not focussed (i.e. contextually given) (H#) or a fall when the content word is focussed (L#). These boundary tones, we claim, play an important role in creating the transitions between consecutive PW's in a larger PPh. The unit does not necessarily correspond to a syntactic constituent; the grouping is, however, characteristic of well-planned speech. It is a rhythmic grouping with a left-headed character, where a content word can be grouped together prosodically with following function words in a manner analogous to the way the definite article and other morphological endings are attached and prosodically clefted to the right of a lexical stem in Swedish (e.g. *bil+ar+na* 'car+pl+the'). Thus, a PW can consist of a content word and a following preposition (e.g. *köpt över*, 'bought over') where the preposition is syntactically a member of a constituent that does not include the content word, as in *Lars har köpt över 100 skivor*. Lars has bought over 100 records'.

Prosodic Phrase (PPh)

One or more PW's make up a PPh which is marked by a L% or H%

- [13] Horne, M., Strangert, E. and Heidner, M. (1993), 'Prosodic boundary strength in Swedish: Final Lengthening and Silent Interval duration', *Proc. XIII ICPhS*, Stockholm.
- [14] Pierrehumbert, J. (1980), *The phonetics and phonology of English intonation*, Ph.D. Diss., M.I.T.
- [15] Lieberman, P. 1967, *Intonation, perception and language*, Cambridge, Mass.: MIT Press.
- [16] Swerts, M. (1993), 'On the prosodic prediction of discourse finality', *Proc. 4th SCSA Workshop on Prosody, Working Papers* (Dept. Ling., U. of Lund) 41, pp. 96-99.
- [17] Grosz, B. and Hirschberg, J. (1992), 'Some intonational characteristics of discourse structure', *Proc. ICSLP 92*, Banff, pp. 429-432.
- [18] Lindström, A., Horne, M., Svensson, L., Lingqvist, M. and Filipsson, M. (1995), 'Generalizing prosodic structure for restricted and 'unrestricted' texts', *Proc. XIII ICPhS*, Stockholm.
- [19] Horne, M. and Filipsson, M. (1994), 'Generalizing prosodic structure for Swedish text-to-speech', *Proc. ICSLP 94*, Yokohama, pp. 711-714.