

#### 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:

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.

Download date: 17. Dec. 2025

ICPhS 95 Stockholm

Session 81.10

# COMPUTATIONAL MODELLING AND GENERATION OF PROSODIC STRUCTURE IN SWEDISH

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

250

ğ

### ABSTRACT

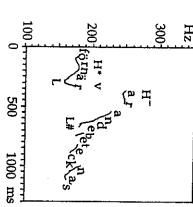
are needed for their recognition in texts linguistic and discourse parameters that prosodic hierarchy for Swedish and the various levels of structure assumed in a are presented A summary of the motivation for the

### INTRODUCTION

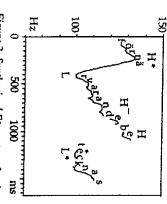
of the following word, betecknas 'is characterized', since the underlying end of the focussed expression for narvarande presently coincides with a low F0 point (L#) in the speech of the cannot reproduce this pattern since no generated using the current rule system Word. In Figure 2, it is observed that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

> Figure premainstress syllable be- and a L\* on the syllable -teck- [1]. Thus, the L# at automatically generated. the end of for narvarande such as in cannot currently

present the trend is characterized as (1) För närvarande betecknas very weak



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



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

Hz

100+

8

000

1500

2000

2500

3000

3500

E

procent il medan sexmanadersvaxiar gan upp 5 punkter till 10,50 procent| gått tillbaka 1 räntepunkt || till 10,58 (2) \langle Tolvmdnaders statsskuldväxlar hade

strength as that after procent 'percent'.

where the internal boundary after associated with the sentence in (2), exemplified, for example, in Figure 3, which presents part of the F0 contour majority of clauses/sentences. This is

runtepunkt 'interest point' has the same

gone back 1 point to 10.58 percent while six-month bonds had gone up 5 represents a Prosodic Phrase boundary) points to 10.50 percent. (welve-month state-debt bonds had (Where

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

#### STRUCTURE SWEDISH PROSODIC

Prosodic Word (PW)

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

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

procent || medan) with a clause-internal PP boundary after 'en ribitepunkt

Figure 3. Fo contour for a fragment of the sentence in (2) (1 rantepunks || iill 10.58

synthesis is that one has not been able to

Another problem with current

those which occur at the end of internal boundaries that are as strong as predict the location of clause-internal Prosodic Phrase boundaries, i.e.

ş

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

### Prosodic Phrase (PPh)

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

ICPhS 95 Stockholm

Session 81.10

Phrase' [4] and Lieberman's 'breath group [5]. Factors which determine the and a certain degree of both Pierrehumbert's 'Intonation houndary tone accent, a following pause following: location of PPh boundaries include the Lengthening ([3]). It corresponds to

associated with a PPh at some level of analysis. Furthermore, 337 of these clauses corresponded to sentence 93% of the sentence houndaries were PPh on some level of analysis. assigned a prosodic boundary equal to a were 362 sentences. This means that boundaries. In the whole corpus, there of the clauses ended in a boundary constituents, this means that the end of a Hypothesis in the hierarchy of prosodic Since we assume the Strict Layer PU is also the end of a PPh; thus, 69% Prosodic Utterance (PU) boundaries). classified as PPh boundaries and 95 as strong or stronger than a PPh (404 were ending in a houndary which was as 69% of these were characterized as containing 724 clauses, where clauses also included elliptical clauses, 499 or of a corpus of 36 radio broadcasts the end of a PPh. In an auditory analysis boundary corresponds in many cases to a) clause/sentence houndary: A clause

up. gd ner go down , falla 'fall', stiga tise IFF these complements were preceded by another focussed verb complement. Thus a PPh boundary (II) complements (beginning with till 'to' or conclusions can be drawn from the covironment for the insertion of these generalizations concerning extremely small, we decided, could occur before till in (4a) but not in sedan 'since') to the verbs gd upp 'go houndaries occurred before focussed we investigated, 12 clause-internal PPh data-base dealing with the stock-market investigation: in the domain-specific internal PPh boundaries. The following determine whether one could make any syntactic structure of the data to nevertheless to examine the lexico-(2%). Although the number was in clause-internal position. In our data of 724 clauses, we detected only 17 cases position, they may also occur optionally majority of cases in clause-final Although PPh houndaries occur in the b) clause-internal PPh boundaries the

> (relevant focussed expressions are (4b) where the first complement following the verb is not focussed written in bold script):

hade då fallit 4 punkter II till en ränta på (4) (a) IIF yra-åriga standardobligationen 'IIThe four-year standard bond had fallen 10,27 procently

4 points 1 to an interest-rate of 10,27

också gån tillbaka 4 punkter till 10,84 (h) HTolymanaders statsskuldväxlar hade procent

procentil also gone back 4 points to 10,84 'ITwelve-month national-debt bills had

relatively long subject (on the average of PPh boundaries occurred between a 15 syllables) and a focussed verb. The remaining 5 cases of clause-internal

segments and practically never between the second last and final clauses of a clauses together, did not contain (on an average of) more than 40 syllables. The segment boundaries. It was observed assigned a weaker, i.e. PW boundary. PW boundaries also replace PPh frequently at the beginning of discourse linking of clauses occurred most PPh, after the linking of two or more than 30 syllables and if the resulting only if the first clause contained less futhermore, that clause-linking occurred linking never occurred over discourse within a discourse segment. Such from the linking together of clauses instead of a PPh houndary all arose were associated with a PW boundary The 225 clause boundaries (31%) which elliptical clauses) in our database were containing less than 7 syllables (12 associated with a PPh boundary; clauses certain length in order for it to be sentence-internal clause must be of a syllables, with the mean at 24 syllables (SD=10.3). Our data also indicate that a the average of about 5 syllables/second, PPh's contained between 7 and 63 is often termed 'breath groups' [5]. In our material, where the speech rate is on given rate of speech since PPh's (as we or less fixed number of syllables at a c) length: A PPh will consist of a more have defined them) correspond to what

> considered as a cue to segment finality (see also [6] for other cues). that the non-linking of clauses can be discourse segment. Thus, it could

## Prosodic Utterance (PU)

place (i.e. the end of a 'discourse segment' [7]). In our data, 95 of the 727 clauses (13%) ended in a boundary which was classified as a PU boundary. In the texts which were originally read opening of a new paragraph immediately following the PU boundary (S. Haageon the radio, these correlate with the Paim, personal communication). with locations where a topic shift takes which is bounded by extended pauses [3]. These strong boundaries coincide One or more PPh's make up a PU

### STRUCTURE GENERATING PROSODIC

(JI, 140 subordinate conjunctions (e.g. att 'that' pronouns (e.g. som 'that') and after coordinate conjunctions and relative not occurring in lists of words having colon, semicolon, some commas (those the same word class), as well as before tain punctuation marks, e.g. full stop. defined. Clause boundaries occur at cerboundaries in a text since the clause is also crucial to be able to identify clause needed in order to define PW's. It is the basic domain over which PPh's are (c.g. prepositions, conjunctions) nouns, adjectives) and function words distinction between content words (e.g. kinds of information [8]. First of all, the able to recognize a number of different prosodic structure, it is important to be In order to automatically generate

prepositional complement to the verbs gd upp 'go up', gd ner 'go down', falla 'fall 'and stiga' rise'. specific as regards their lexical specification. This is not the case with boundaries before the second focussed module inserts clause-internal boundaries, which is domain-independent. Thus, the domain specific boundaries seemed to be so domainour algorithm. This was due to the fact included a domain-specific module in In order to generate the clause-internal (optional) PPh boundaries, we have the module that identifies clause that the locations of clause-internal PPh

> order to link two or more clauses into the prosodic parser, This information is currently being built its being linked to the following clause. number of syllables that will result after that a given clause consists of and the able to calculate the number of syliables logether into a PPh, it is necessary to be boundaries at the ends of clauses, i.e. in Moreover, in order to generate PW

Finally, in order to generate PU boundaries, one must be able to recognize discourse segment boundaries. triggered by paragraph boundaries. In the present algorithm [9] these are

## ACKNOWLEDGEMENT

This research has been supported by a grant from the Swedish HSFR/NUTEK Language Technology Programme.

### REFERENCES

to-speech system", STL-QPSR, pp. 31. 1) Bruce, G. and Granström, B. (1989). Modelling Swedish infonation in a lext-

prosodic structure for synthesis of Swedish intonation. Working Papers (Dept. Ling., Univ. of Lund), pp. 43, 72. [2] Horne, M. (1994), "Generating

[3] Horne, M., Strangert, E. and Heldner, M. (1995), "Prosodic boundary strength in Swedish: Final Lengthening and Silent Interval duration", Proc. XIIIth ICPhS. Stockholm.

[4] Pierrehumbert, J. (1980). The phonetics and phonology of English intonation. Ph.D. Diss., M.I.T.
[5] Lleberman, P. 1967. Intonation. perception and language. Cambridge. Mass.: MIT Press.

[6] Swerts, M. (1993), "On the prosudic prediction of discourse finality", Proc. ESCA Workshop on Prosudy, Working Papers (Dept. Ling., U. of Lund) 41, pp.

discourse structure", Proc. ICSLP 92. Banff, pp. 429-432. [7] Grosz, B. and Hirschherg, J. (1992). "Some Intonational characteristics of

[9] Horne, M. and Filipsson, M. (1994). "Generating prosodic structure for Swedish text-to-speech". *Proc.* ICSLP 94, Yokohama, pp. 711-714. [8] Lindström, A., Horne, M., Svensson, T., Ljungqvist, M. and Filipsson, M. (1995), "Generating prosodic structure for restricted and "unrestricted" texts." Proc. XIIIth ICPhS. Stockholm.