LUND UNIVERSITY

Constraints on prosodic phrasing in spontaneous speech

Hansson, Petra

2001

Link to publication

Citation for published version (APA):
Hansson, P. (2001). *Constraints on prosodic phrasing in spontaneous speech*. (Working Papers, Lund University, Dept. of Linguistics; Vol. 48). http://www.ling.lu.se/disseminations/pdf/48/Hansson.pdf

Total number of authors:
1

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

Download date: 11. Sep. 2026

Lund University, Dept. of Linguistics
Working Papers 48 (2001), 87–99

Constraints on prosodic phrasing in spontaneous speech

Petra Hansson

1. Introduction

1.1 Purpose

The present paper makes some preliminary hypotheses about the constraints on prosodic phrasing in spontaneous Swedish speech. Within Optimality Theory, linguistic phenomena such as prosodic phrasing are explained through the interaction of constraints (McCarthy & Prince 1993). One may assume that there are a number of constraints that interact to determine what prosodic form an utterance is assigned given its syntactic and information structure. Among these are constraints that both align prosodic phrase structure with syntactic and information structure, as well as constraints that assign word and focal accents and delimit the prosodic phrases' size and accentual content.

The constraints on output representations are hypothesized to be universal. Languages differ only in the ranking order of the constraints. Therefore, it is interesting to investigate in what way constraints previously claimed for other languages can be assumed to interact in Swedish.

In the present paper, results from a production data study will be presented. Taking a number of constraints suggested to exist in other languages as a starting point, a possible constraint hierarchy for Swedish is discussed.

1.2 Optimality theory

Within optimality theory, linguistic phenomena are explained through the interaction of universal constraints. An optimal output form for a given input is selected from among a number of competing surface forms. The form that best satisfies the highest-ranking constraint, on which the candidates conflict, is considered to be optimal, as shown in Tableau 1. Constraints can be violated, but only minimally, i.e. only in order to satisfy a higher ranked constraint. The constraints are universal, but languages may differ with respect to each other in the ranking order of the constraints.

Tableau 1. An example of a constraint tableau[1]. Constraint A is ranked higher than constraint B (Constraint A >> Constraint B).

Candidates	Constraint A	Constraint B
☞ $Cand_1$		*
$Cand_2$	*!	

1.3 Previous studies

The constraints on prosodic phrasing have previously been discussed for a number of different languages, see for example Delais-Roussarie 1996, Selkirk (in press) and Truckenbrodt 1999. In the present study, we take the constraints claimed by Selkirk as a starting point in our discussion of the constraints on prosodic phrasing in spontaneous Swedish speech.

Selkirk has made some preliminary hypotheses about the hierarchy of constraints on prosodic phrasing in English. The constraints suggested to play a role are constraints on the syntax-phonology interface (Align-XP,R and Wrap-XP), constraints on the focus-phonology interface (Align Focus, R) and constraints on the size and accentual content of prosodic phrases (Bin(MaP) and MiPAccent).

The constraints Wrap-XP and Align-XP,R are claimed to occupy the same rank in the English constraint hierarchy. Wrap-XP calls for the elements of an input morpho-syntactic constituent of type XP to be contained within a prosodic constituent of type MaP (major phrase) in output representation. Align-XP,R, on the other hand, calls for the right edge of any XP (maximal projection) in syntactic structure to be aligned with the right edge of a phrase in prosodic structure. Wrap-XP comes into conflict with most (but not all) boundaries demanded by Align-XP,R.

The Align Focus, R constraint is ranked above the Wrap-XP and Align-XP,R constraints. Align Focus, R calls for the right edge of a focus constituent in information structure to be aligned with the right edge of a prosodic phrase. Ranked above the Align Focus, R constraint, is the constraint MiPAccent[2]. MiPAccent calls for a minor phonological phrase to contain at least one accent. Finally, below Wrap-XP and Align-XP,R in the constraint

[1]The optimal candidate is indicated with a '☞'. A candidate wrongly considered to be optimal is indicated with a '✋'. The leftmost constraint in the tableau is the highest-ranking constraint and the rightmost constraint is the lowest-ranking. The cells of constraints that are irrelevant for the choice of optimal candidate are shaded. '*' marks a violation of a constraint and '*!' a fatal violation.

[2]A DEPAccent constraint is also assumed, with the same ranking as MiPAccent, which assures that an accent in the output representation have a corresponding accent in the interface representation.

hierarchy, the constraint Binary(MaP) which calls for a major phrase to consist of just two minor phrases is found.

In Selkirk's study, examples are given to support the hypothesizing that the constraints Align-XP,R, Wrap-XP and Align Focus, R, which interact with the phonological constraints on the prosodic phrases' size and tonal content, form a full account of English phrasing. However, she remarks that the results "need to be solidified on the basis of non-intuition-based investigation".

Whereas most previous studies on constraints on prosodic phrasing have been limited to read stimuli, we have chosen to use spontaneous dialogues as data in the present study. Although we are aware of the fact that prosodic structure rarely is fully determinative and that several alternative prosodic realizations of a given syntactic structure often are possible (see Cutler et al. 1997 for a discussion), we assume that the prosodic phrasing produced by the recorded speakers is optimal.

2 Method

2.1 Speech material

The examined speech material consists of four spontaneous, natural dialogues from the research project Swedish Dialogue Systems' database[3], recorded at travel agencies in Lund, Sweden. Four male and three female speakers' speech was analyzed[4].

2.2 Prosodic transcription

The recordings were stored as digitized audio files. Subsequently, they were segmented and transcribed orthographically using the speech wave and spectrograms. Finally, after listening to the dialogues and visually inspecting F0 contours of the material, the transcription was completed with information about the distribution of prosodic phrase boundaries. Weak boundaries were indicated with '|'and strong boundaries with '||'. Although a distinction was made between strong and weak boundaries in the transcription of the material, both boundary types will be dealt with in the same manner in the analysis below.

In two of the dialogues, the prosodic transcription was further completed with information about the positions of word accents ('') and focal accents ('').

[3]Information about the research project *Swedish Dialogue Systems* can be found at the following address: http://www.ida.liu.se/~nlplab/sds/

[4]One of the speakers appeared in more than one dialogue.

det kostar 'nittio 'kronor om du inte "har det |
'it's ninety crowns if you don't have one |'

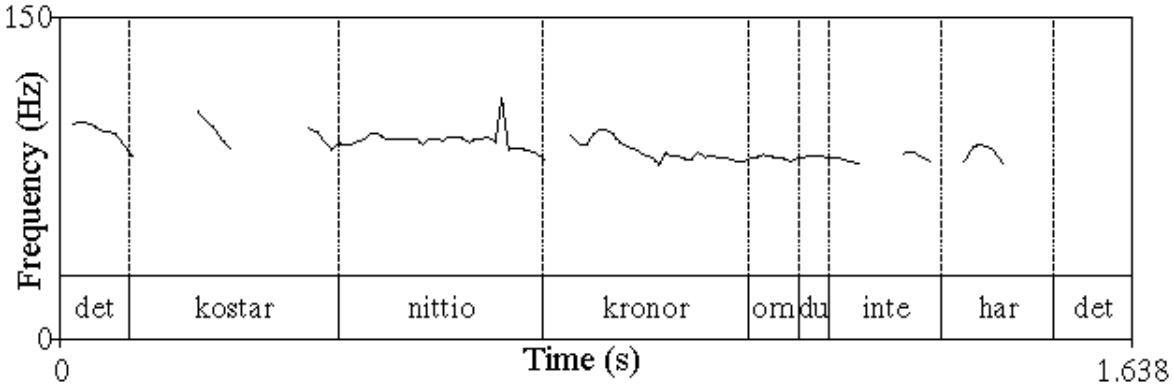

Figure 1. F0 contour and orthographic transcription of the utterance *det kostar nittio kronor om du inte har det* 'it's ninety crowns if you don't have one'.

3 Empirical analysis

3.1 Speech repairs, Align-XP,R and Wrap-XP

Spontaneous speech is typically less structured syntactically than written discourse. It contains fragments and there is little subordination. A first observation about the prosodic phrasing in our data is that very few utterances have internal prosodic phrase boundaries, suggesting that Align-XP,R be ranked below Wrap-XP in spontaneous Swedish. See for example the prosodic transcription and F0 contour in Figure 1.

Utterances containing certain kinds of speech repairs constitute an exception. In spontaneous speech, speakers do not produce perfect utterances. Speakers interrupt themselves, repeat and modify what they say. Heeman 1997 and Schriberg 1994 have shown that 'speech repairs' or 'disfluencies' are normal and frequent in spontaneous speech.

What is the optimal way of phrasing an utterance which contains a speech repair? How does the speaker let the hearer know that a part or all of what has just been said should be disregarded?

Heeman has shown for English that the end of a disfluency reparandum (the interruption point) is often accompanied by a disruption in the intonation contour. This indicates that speech repair phenomena need to be taken into consideration when formulating the constraints on prosodic phrasing in spontaneous speech.

Heeman divides speech repairs into three different categories: 'fresh starts', 'modification repairs' and 'abridged repairs'. 'Fresh starts' occur where the speaker abandons the current utterance and starts again. Heeman claims that "they are defined in terms of a strong acoustic signal marking the abandon-

om jag 'titt(ar) | "fjortonde har jag 'plats ||
'if I loo(k) | on the fourteenth I have seats ||'

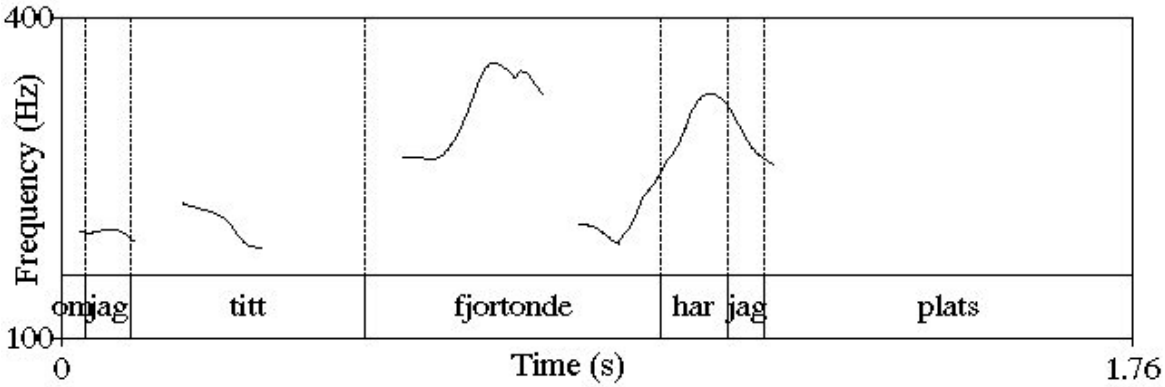

Figure 2. F0 contour and orthographic transcription of the utterance *om jag titt(ar) fjortonde har jag plats* 'if I loo(k) on the fourteenth I have seats' (female speaker of South Swedish)

vi 'har ju Air F vi 'har ju "SAS 'också
reparandum *alternation*
'we have Air F we have SAS as well'

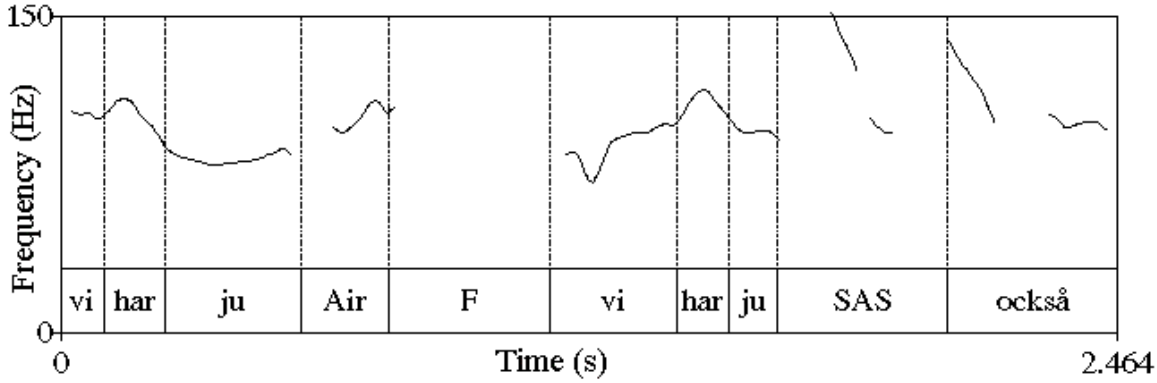

Figure 3. F0 contour and orthographic transcription of the utterance *vi har ju Air F vi har ju SAS också* 'we have Air F we have SAS as well' (male speaker of Standard Swedish)

ment of the current utterance" (1997:12) and that the interruption point "often is accompanied by a disruption in the intonational contour" (1997: 9). Intonational disruptions are also found in fresh starts in Swedish, see e.g. Figure 2. Therefore, we may choose to allow the morpho-syntactic constraints to call also for the interruption points of uncompleted XPs to be aligned with prosodic phrase boundaries, at least the interruption points of fresh starts.

The 'modification repairs' have a strong word correspondence between reparandum and alternation, see e.g. Figure 3. Heeman suggests that the word correspondence can help the hearer to determine the extent of the reparandum as well as function to signal that a modification repair has

nej för att jag hade 'bytt äh "datum 'nämligen ||
'because I had changed uh the date you see ||'

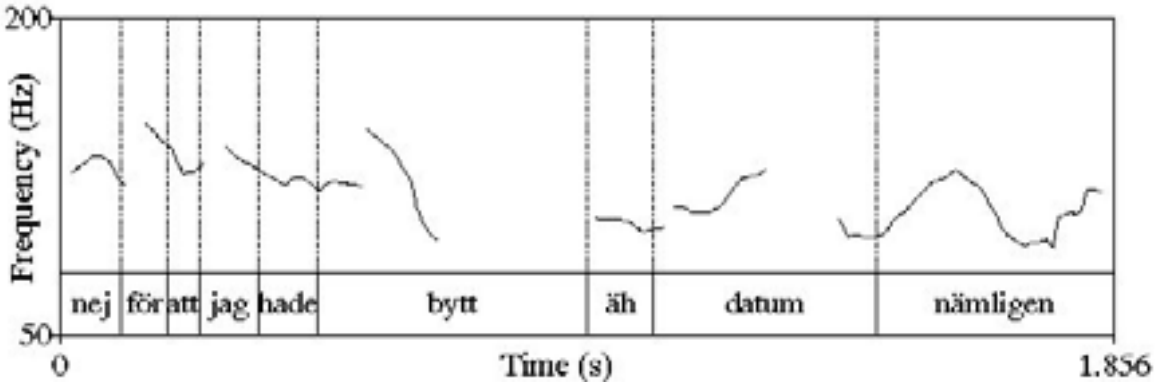

Figure 4. F0 contour and orthographic transcription of the utterance *nej för att jag hade bytt äh datum nämligen* 'because I had changed uh the date you see' (male speaker of South Swedish)

occurred. Therefore, an acoustic marking of the interruption point of a modification repair, e.g. a boundary tone or a F0 reset, does not appear to be necessary, as confirmed by the examined examples in the present study.

'Abridged repairs' consist of an editing term, e.g. a hesitation sound / filled pause, but of no reparandum, see Figure 4. In spontaneous speech, filled pauses are frequent e.g. after conjunctions and before important content words. However, as the phrase is continued after the repair, we do not expect the Align-XP,R constraint to call for an alignment of the abridged repair with a prosodic phrase boundary. Rather, a cohesional strategy is expected.

In conclusion, it is necessary to allow the morpho-syntactic constraints to also apply to incompleted XP in a constraint hierarchy for prosodic phrasing in spontaneous speech. Whereas abridged repairs do not appear to need an acoustic marking of their interruption point, fresh starts do. Modification repairs are not necessarily acoustically marked since they appear to rely on a strong word correspondence to delimit the extent of the reparandum.

When examining the prosodic correlates of speech repairs, it appears that it may be necessary to divide the Align-XP,R constraint into two distinct constraints: Align-XP,R and Align-XP,L. The Align-XP,R constraint would be responsible for associating the right edges of a XP with a boundary tone and final lengthening and the Align-XP,L constraint for associating the left edge of a XP with a F0 reset. One would then assume that the Align-XP,L constraint applies to each XP or part of XP, while the Align-XP,R only applies to complete XPs. The interruption points of speech repairs do not always show final lengthening and a boundary tone (see e.g. Figure 2 above).

och det är "bara 'här jag kan 'göra det då | 'köpa den hos 'er ||
[och det är bara här jag kan göra det då]$_{ROOT}$ [köpa den hos er]$_{E.C.}$
'and it's only here I can do that then | buy it here with you ||'

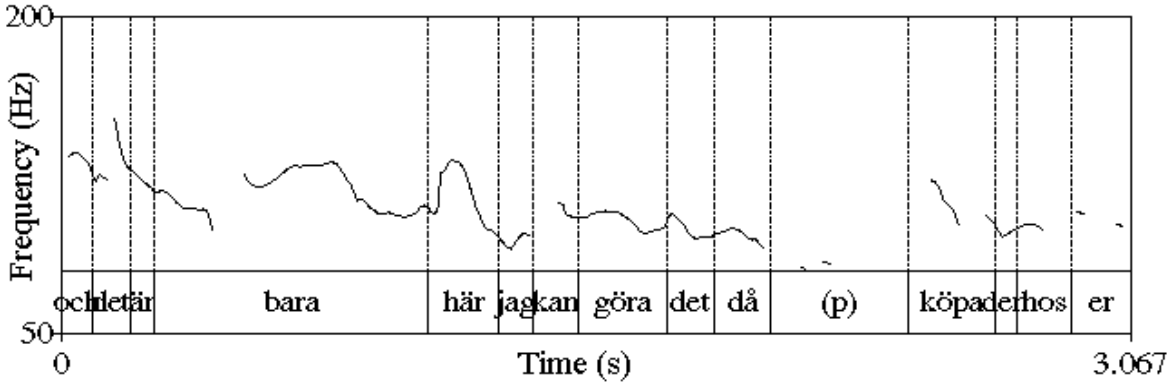

Figure 5. F0 contour and orthographic transcription of the utterance *och det är bara här jag kan göra det då köpa den hos er* 'and it's only here I can do that then, buy it here with you'.

3.2 Root sentences, Align-XP,R and Wrap-XP

Certain types of constructions, parenthetical expressions for example, introduce internal prosodic phrase boundaries by forming intonation domains on their own within the root sentence. These constructions are external to the root sentence with which they are associated (Nespor & Vogel 1986). In (1) the parenthetical, clarifying expression, a so-called right detachment or antitopic construction (Lambrecht 1994), is found at the end of the root sentence. Here we may assume that Wrap-XP wraps the root sentence and the external construction (E.C.) separately, as two distinct morpho-syntactic constituents.

(1) nej det är den 'billigaste jag 'har | Air "France 'där ||
[nej det är den billigaste jag har]$_{ROOT}$ [Air France där]$_{E.C.}$
'no it's the cheapest I have | Air France there ||'

The phrasing in Figure 5 can be described in the same manner, i.e. as the result of Wrap-XP wrapping the root sentence and the clarification separately, as two distinct morpho-syntactic constituents.

However, in (2) the external construction is internal to the root sentence, and a violation of Wrap-XP is inevitable. If the external construction itself is wrapped, then the larger XP is not, i.e. then the larger XP is not contained in a single prosodic phrase and vice versa. Note that with the violation of Wrap-XP, Align-XP,R right-aligns all XPs with the edges of prosodic phrases (except

Tableau 2. PwdAccent >> Wrap-XP >> Align-XP,R

jo du måste ha [ett sådant kort]$_{NP}$ [till]$_{IP}$ [ett sådant]$_{E.C.}$ [till bägge de här va]$_{PP}$ för det är [studentbiljett]$_{NP}$	PWd Accent	Wrap-XP	Align-XP,R
a. (jo du måste ha ett sådant kort till ett sådant till bägge de här va för det är studentbiljett)$_{PPh}$		*	**!**
b. (jo du måste ha ett sådant kort till)$_{PPh}$ (ett sådant till bägge de här va för det är studentbiljett)$_{PPh}$		*	**!*
c. (jo du måste ha ett sådant kort till)$_{PPh}$ (ett sådant)$_{PPh}$(till bägge de här va för det är studentbiljett)$_{PPh}$		*	**!
☞ d. (jo du måste ha ett sådant kort till)$_{PPh}$ (ett sådant)$_{PPh}$(till bägge de här va)$_{PPh}$ (för det är studentbiljett)$_{PPh}$		*	*
e. (jo du måste ha ett sådant kort)$_{PPh}$ (till)$_{PPh}$ (ett sådant)$_{PPh}$ (till bägge de här va)$_{PPh}$ (för det är studentbiljett)$_{PPh}$	*!	*	

the noun phrase *ett sådant kort* ‘one of those cards’ which will be discussed below).

(2) jo du "måste 'ha ett 'sådant "kort till | ett 'sådant | till "bägge de 'här va || för det 'är stu'dentbil'jett |
[jo du måste ha ett sådant kort till [ett sådant]$_{E.C.}$ till bägge de här va för det är studentbiljett]$_{ROOT}$
‘yes you have to have one of those cards for | one of these | for both of these right | because it’s a student ticket |’

3.3 Constraints on accentual content and size

The speakers’ utterances in the examined domain, i.e. the travel domain, are generally short and they are usually wrapped into a single prosodic phrase. This tendency provides support for a higher ranking of the Wrap-XP constraint than the Align-XP,R constraint. However, some evidence to support an interaction between the syntactic constraints and constraints on the size and tonal content of prosodic phrases has been found.

Since the defining feature of a prosodic word is that it contains an accent, we may formulate a constraint PWdAccent that sets the minimal tonal content of a prosodic word to one accent. Because a prosodic phrase consists of at least one prosodic word, PWdAccent can be assumed to be the constraint which prevents *till* ‘to/for’ from constituting a prosodic phrase on its own (i.e. the phrasing in candidate (e) in Tableau 2). In order to do so, the PwdAccent constraint is ranked above Wrap-XP and Align-XP,R (PwdAccent >> Wrap-XP >> Align-XP,R).

(3) jo du "måste 'ha ett 'sådant "kort till | ett 'sådant | till "bägge de 'härva || för det 'är stu'dentbilÆjett |
jo du måste ha [ett sådant kort]$_{NP}$ [till] [ett sådant]$_{E.C.}$ [till bägge de här va]$_{PP}$ för det 'är [stu'dentbilÆjett]$_{NP}$
'yes you have to have one of those cards for | one of these | for both of these right || because it's a student ticket |'

The extent of a prosodic phrase can be described in a number of ways. We can, for example, characterize a prosodic phrase as containing a certain number of syllables or feet, and can measure its duration in time or count the number of focal and/or non-focal accents it contains.

In Horne & Filipsson 1996, it is shown that syllable count plays a role in determining the position of prosodic phrase boundaries. In their material (read speech), where the speech rate was on the average of about 5 syllables per second, prosodic phrases contained between 7 and 63 syllables, with a mean at 24 syllables (standard deviation 10.3). In the spontaneous and unscripted speech examined in the present study, the prosodic phrases contain a much smaller number of syllables. However, because there are few long syntactic structures in the examined material, we are not, at this point, able to determine if the observed units are optimal chunks for linguistic processing in spontaneous speech, or if Wrap-XP would wrap even larger constituents.

It is necessary to assume that there is an interaction between the syntactic constraints and further constraints on the maximum and minimum size of prosodic phrases. However, the examined material does not allow us to draw any conclusions as to their nature.

3.4 Maximal focus projection and Align Focus, R

In English, the constraint Align Focus, R, ranked above Wrap-XP and Align-XP,R aligns the right edge of a focus constituent in information structure with the right edge of a prosodic phrase (Selkirk, In press).

The *focus* is the part of the utterance that makes it into a piece of information (Lambrecht 1994), and it is the focus part which is responsible for conveying the 'new' information (Molnár 1998). The focal information is 'new' "in the sense that the speaker presents it as not being recoverable from the preceding discourse" (Halliday 1967:204).

The focus may cover a larger or smaller part of the utterance. In an utterance with a broad focus, the focus may cover the entire utterance (a maximal focus projection).

[jag 'glömde 'plånboken 'hemma]$_{Focus}$ |
'[I forgot my wallet at home]$_{Focus}$ |

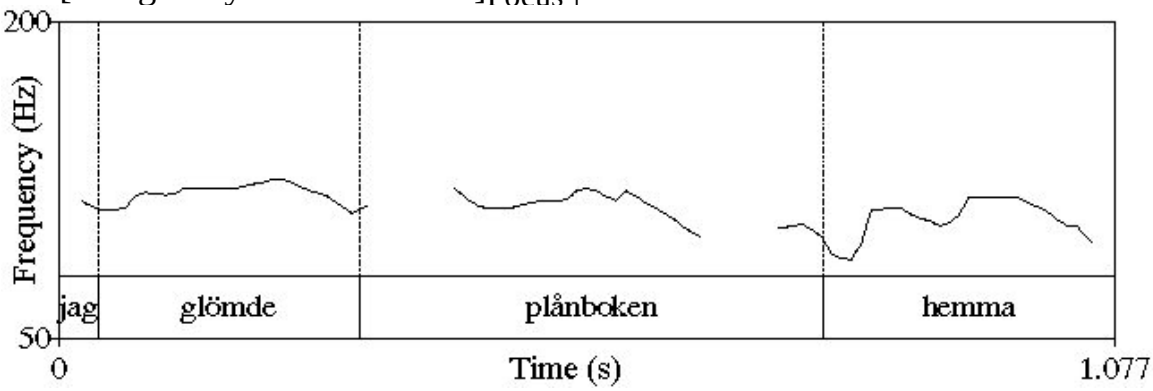

Figure 6. F0 contour and orthographic transcription of the utterance *jag glömde plånboken hemma* 'I forgot my wallet at home' (male speaker of South Swedish). The lack of a focal accent marking the focus is not unusual in South Swedish.

Tableau 3. Align Focus, R >> Wrap-XP

[jag glömde [plånboken]$_{NP}$ hemma]$_{Focus}$	Align Focus, R	Wrap-XP
☞ a. (jag glömde plånboken hemma)$_{PPh}$		
b. (jag glömde plånboken)$_{PPh}$ (hemma)$_{PPh}$		*

According to the 'Focus-to-Accent' (FTA) approach, the focus of an utterance is marked by a focal accent (Gussenhoven 1983; Ladd 1996). If the Align Focus, R constraint is satisfied, the focus constituent's right edge is also aligned with a prosodic phrase edge.

When examining utterances with foci covering the entire utterance and foci in phrase-final position, a high ranking of the Align Focus, R constraint seems to also be relevant for Swedish, see Figure 6 and Tableau 3. A structure such as the one in Figure 6 does not lead to a conflict between Wrap-XP and Align Focus, R. Here a single prosodic phrase around the entire structure allows both the focus constituent to be right-aligned with a prosodic phrase and XP to be contained within a single prosodic phrase.

However, when focus is in non-final position, it is apparent from our data that the right edge of a focus constituent is not necessarily aligned with the right edge of a prosodic phrase in spontaneous Swedish speech, see e.g. Figure 7 and Tableau 4.

In Figure 7 we can also observe a late timing of the focal H¯. The focally accented Accent II-word (H*L) *månad* 'month' is associated with a focal gesture consisting of a high tone (H¯) which occurs in the word following *månad*. This 'delay' of the H¯, a possible coherence cue, has been found both

[vilken "månad]$_{Focus}$ ska du 'åka i |
'[what month]$_{Focus}$ are you leaving |'

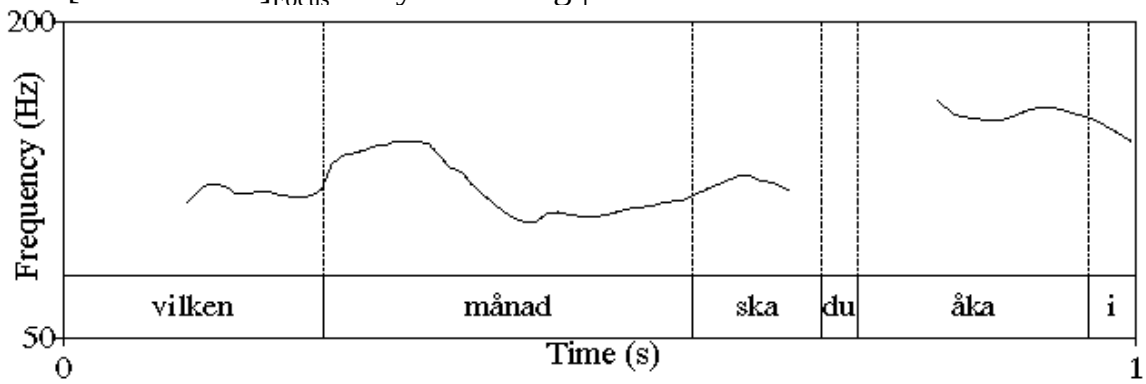

Figure 7. F0 contour and orthographic transcription of the utterance *vilken månad ska du åka i* 'what month are you leaving' (male speaker of Standard Swedish).

Tableau 4. Align Focus, R >> Wrap-XP

[[vilken månad]$_{Focus}$ ska du åka i]$_{VP}$	Align Focus, R	Wrap-XP
a. (vilken månad ska du åka i)$_{PPh}$	*!	
☞ b. (vilken månad)$_{PPh}$ (ska du åka i)$_{PPh}$		*

on words in utterance-initial foci and on so-called 'contrastive topics' in Swedish (Hansson 2000).

In summary, the right edge of a focus constituent is not necessarily aligned with the right edge of a prosodic phrase in spontaneous Swedish speech. Quite the contrary, in terms of prosodic phrasing, it has been suggested that the observed 'delay' of the H¯ can be interpreted as a way of making the focussed word a part of the same prosodic phrase as the information that follows (Horne et al. 1999).

4 Discussion

In the present paper an attempt is made to determine whether empirical evidence can be found for theoretical claims made about a number of universal constraints on prosodic phrasing.

As pointed out by Selkirk, as far as the syntax-phonology interface is concerned, languages must opt for a dominant cohesional strategy or a dominant demarcative structure, as represented by the relative rankings of the morpho-syntactic constraints Wrap-XP and Align-XP,R. In the case of Swedish, the data analyzed indicates a cohesional strategy. However, the speech style influences the choice of phrasing strategy. Both rate and style of speech have an effect on the length of the prosodic phrases. It is plausible to

assume that the constraint ranking described in the present paper, rather than being a ranking specific for Swedish, is a more general ranking reflecting the prosodic phrasing of speech produced on-line.

Although further constraints on the prosodic phrases' maximum and minimal size need to be added to the hierarchy, the following preliminary ranking for Swedish spontaneous speech has been proposed: PWdAccent >> Wrap-XP >> Align-XP,R. Evidence against a high ranking of the constraint Align Focus, R in Swedish has been found.

Acknowledgements

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

References

Cutler, Anne, Delphine Dahan & Wilma van Donselaar. 1997. 'Prosody in the comprehension of spoken language: a literature review'. *Language and Speech* 40, 141-201.

Delais-Roussarie, Elisabeth. 1996. 'Phonological phrasing and accentuation in French'. In M. Nespor & N. Smith (eds), *Dam phonology: HIL phonology paper II*, 1-38. La Haye: Holland Academic Graphics.

Gussenhoven, Carlos. 1983. 'Focus, mode, and the nucleus'. *Journal of Linguistics* 19, 377-417.

Halliday, M. A. K. 1967. 'Notes on transitivity and theme in English 2'. *Journal of Linguistics* 3, 199-244.

Hansson, Petra. 2000. 'Focally accented topics'. *Proceedings FONETIK 2000. The Swedish phonetics conference. May 24-26, 2000*, 65-68. Department of Languages, Högskolan Skövde.

Heeman, Peter A. 1997. *Speech repairs, intonational boundaries and discourse markers: modeling speaker's utterances in spoken dialogue*. Ph.D. Dissertation, University of Rochester.

Horne, Merle & Marcus Filipsson. 1996. 'From prosodic structure to intonation contours'. *Nordic Prosody, Proceedings of the VII conference. Joensuu, Finland, 1996*, 127-139.

Horne, Merle, Petra Hansson, Gösta Bruce & Johan Frid. 1999. 'Prosodic correlates of information structure in Swedish human-human dialogues'. *Proceedings of EUROSPEECH'99, the 6th European conference on speech communication and technology, September 5-9, 1999, Budapest, Hungary*, Vol. 1, 29-32.

Ladd, Robert. 1996. *Intonational phonology* (*Cambridge studies in linguistics* 79). Cambridge: Cambridge University Press.

Lambrecht, Knud. 1994. *Information structure and sentence form. Topic, focus, and the mental representation of discourse referents* (*Cambridge studies in linguistics* 71). Cambridge: Cambridge University Press.

McCarthy, John J. & Alan Prince. 1993. 'Generalized alignment'. In G. Booij & J. van Marle (eds), *Yearbook of morphology 1993*, 79-153.

Molnár, Valéria. 1998. 'Topic in focus. On the syntax, phonology, semantics and pragmatics of the so-called "contrastive topic" in Hungarian and German'. *Acta Linguistica Hungarica* 45, 89-166.

Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris Publications.

Schriberg, Elisabeth. 1994. *Preliminaries to a theory of speech disfluencies*. Ph.D. Dissertation, University of California, Berkeley.

Selkirk, Elisabeth. in press. 'The interaction of constraints on prosodic phrasing'. In Merle Horne (ed.), *Prosody. Theory and experiment. Studies presented to Gösta Bruce.* Dordrecht: Kluwer.

Truckenbrodt, Hubert. 1999. 'On the relation between syntactic phrases and phonological phrases'. *Linguistic Inquiry* 30, 219-255.