

# ACCENTUATION OF DOMAIN-RELATED INFORMATION IN SWEDISH DIALOGUES

Merle Horne, Petra Hansson, Gösta Bruce, Johan Frid and \*Arne Jönsson

Dept. of Linguistics and Phonetics, Lund University, Lund, Sweden  
\*Dept. of Computer Science, Linköping University, Linköping, Sweden

## ABSTRACT

The issue of the focal accentuation of contextually ‘given’ travel domain concepts in task-related Swedish dialogues is examined. An explanation for this focal accentuation is proposed using concepts within Centering theory, i.e. it is assumed that it is domain-related ‘Backward Centers’ that are associated with the prominent focal accents. The variation in the timing of the focal  $H^-$  is further accounted for by relating it to different orderings of the Optimality theory **Align XP** and **Wrap XP** constraints with respect to **Align Focus**. Support for the theoretical assumptions is obtained from a perception test using constructed man-machine Initiative-Response dialogues where the ‘machine’ Responses vary as to the type of F0 contour on the Backward Centers (focal vs non-focal and early vs. late timing of the focal  $H^-$ ).

## 1. INTRODUCTION

One issue that has intrigued researchers in prosody for many years is the relationship between accentuation and information structure. A lot of effort has been spent for example on trying to come to a better understanding of the relationship between the ‘new/given’ dichotomy and the ‘accented/non-accented’ distinction (see e.g. [9, 16]). It has now become clear, however, that one cannot draw a simple binary distinction between ‘new’ and ‘given’ information. In particular, it has been convincingly argued by a number of researchers [1, 8, 14] that there are different kinds of what one can term ‘givenness’ in relation to the discourse context. What is more likely involved is rather a hierarchical relationship of ‘givenness’ or ‘accessibility’ or ‘centrality’ between referents which finds expression in terms of different constellations of morphological, syntactic and prosodic form.

Ariel [1] for example, claims that the morphological form used to express a referent (e.g. definite NP, indefinite NP, demonstrative NP) is governed by how ‘accessible’ the referent is assumed to be for the hearer. Terken and Hirschberg [16] also see a correlation between a concept’s new-given status and a change in grammatical function. Using the theory of ‘centering’ [8], Kameyama [12] proposed a hierarchy of grammatical relations that express different degrees of ‘centering’ of discourse referents in a speaker’s ‘attentional state’. In centering theory, referents are classified as either Forward Centers ( $C_f$ ) or Backward Centers ( $C_b$ ), where ( $C_f$ ) constitute ‘new’ information and

( $C_b$ ) constitute ‘given’ information. Each utterance can have one or more ranked  $C_f$  but only one  $C_b$ . The problem of course is in coming to grips with the factors that determine the relative ordering of ‘centrality’ of referents in a particular discourse.

## 2. CENTERING AND DOMAIN-SPECIFIC DIALOGUES

We find the notion of ‘centering’ particularly well-suited to tackling the issue of information structuring in domain-specific dialogues. One of the factors that most surely determines to a large extent the centrality of referents in this type of discourse is whether or not they occur in the lexicon and in other domain-specific knowledge structures associated with the lexicon. That is to say, highly-ranked centers are referents and attributes that figure in these domain-related knowledge sources.

In the LINLIN dialogue management system [11] being used in the Swedish Dialogue Systems-project (<http://www.ida.liu.se/~nlplab/sds/>), the domain-specific lexicon for each dialogue task is associated with the background knowledge sources, i.e. the information the dialogue system can provide. One of these knowledge sources associated with the lexicon is the Domain model which describes lexical semantic relations holding between domain-specific concepts as well as inference relationships between these concepts. For example, as regards the Travel domain, it contains information such as the fact that *Star hotel* is in *Heraklion* which is a *city on Crete*, and that *Crete* is an *island in the Greek archipelago* or the fact that a *ticket* can be either *one-way* or *return*.

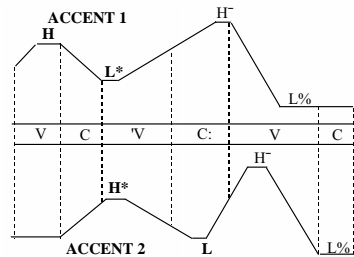
These domain-specific concepts that occur in the domain knowledge sources are those, then, that would be expected to constitute highly-ranked Centers in task-related dialogues between travel agents and clients and the ones that tend to be focally accented when functioning as  $C_b$ .

## 3. ACCENTUATION AND DOMAIN-SPECIFIC LEXICAL ITEMS

In a previous study on accentuation and dialogue structure [4], it was observed how accentual patterns were related to lexical semantic structuring. In a dialogue dealing with a recipe for making a warm tuna-fish sandwich, for example, basic-level domain-specific concepts such as *tuna-fish*, *bread*, and *majonnaise* were observed to be prominently accented, whereas superordinate concepts such as *mishmash*

as well as overspecific descriptions such as *light majonnaise* (as opposed to just *majonnaise*) and *tuna-fish in water* (as opposed to just *tuna-fish*) were not accented even when the attributes *light* and *in water* were contextually ‘new’. This difference was there assumed to be related to the concept’s centrality in the discourse, i.e. their importance with respect to the development of the discourse topic. With respect to the discussion above in §2, we would say that *light* in the context of *majonnaise* and *in water* in the context of *tuna-fish* are concepts which do not occur in the Domain model of the tuna-fish sandwich-making task.<sup>1</sup>

In our current investigations of data from the travel domain, we are again observing similar tendencies, i.e. that domain-related, ‘central’ referents are prominently (focally) accented, both when they constitute a highly-ranked  $C_f$  or a  $C_b$  in the discourse. However, their linear position in an utterance as well as their prosodic form are seen to vary depending on whether the referent is a  $C_f$  or a  $C_b$ .



**Figure 1:** Schematic representation of the association of Swedish word-accent (Accent 1 and Accent 2), focal  $H^*$ , and phrase accent ( $L\%$ ) with the segmental string in phrase-final position (following [2]). The two word-accent are distinguished from each other as to their timing relative to the segmental string. Accent 1 is represented in the standard (Stockholm) dialect as  $HL^*$  whereas Accent 2 is represented as  $H^*L$ , where the starred tone is critically aligned with the beginning of the vowel in the stressed syllable. When a word is made prominent (focussed), it is also assigned a further tonal gesture consisting of a high tone ( $H^-$ ) which, in phrase-final position, is followed by a phrasal tone ( $L$ ).

In phrase-initial position, the accentual pattern on a prominent  $C_b$ , is not the same as that occurring on  $C_f$  (≈new information). Specifically, if the word constitutes new, central domain-related information (i.e. if it is a highly ranked  $C_b$ , the focal  $H^-$  peak occurs within the word and is also followed by a phrase tone when occurring in final position (see Figure 1). On the other hand, if the phrase-initial domain-related word constitutes given information but is nevertheless the  $C_b$ , the word is also assigned a focal  $H^-$  peak, but this  $H^-$  occurs in the following word. This use of

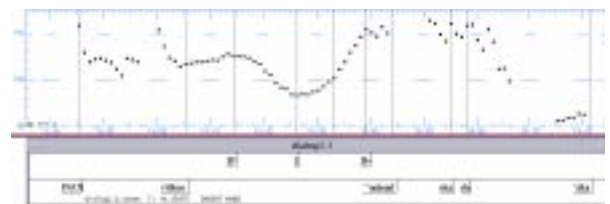
an accented nonpronominal form to refer back in the discourse can, as Grosz et al. [8] suggest be thought to ‘convey some additional information, i.e. lead the hearer or reader to draw additional inferences. The hearer or reader not only infers that the  $C_b$  has not changed even though no pronoun has been used, but also recognizes that the description holds of the old  $C_b$ ’.

## 4. TIMING OF FOCAL $H^-$ IN SWEDISH

In a study on the timing of the focal  $H^-$  in Accent 2 words [3], Bruce made the claim that the only critically timed tone is the word-accent  $H^*$  (see Figure 1) and that the association of the other word-related tones occurs relatively unconstrained within the domain of the ‘foot’. In that study, Bruce did not, however, make any claims as to the exact nature of the timing of the focal accent rise ( $H^-$ ). As noted above, however, we have observed in our studies of dialogue data, is that the timing of the focal  $H^-$  peak seems to be constrained by a number of pragmatic and discourse factors.

### 4.1 Constraints on Phrasing. Optimality Theory

As far as the phonology of Swedish is concerned, it is interesting to attempt to specify how the variation in the timing of the focal  $H^-$  peak can be related to differences in focus assignment and phrasing. According to ideas within Optimality Theory (see [15]), there are a number of constraints that interact to determine how an utterance’s prosodic form is associated with its syntactic and information structure. Among these are constraints that both assign and align focus with prosodic or grammatical constituents and constraints that align prosodic and syntactic structure. For Swedish, one can assume a constraint **Assign Focus** (cf.[15]) which assigns focus to a particular linguistic unit and another one, **Align Focus**, which has the function of aligning a focussed constituent with the right edge of a prosodic constituent. This can be thought of as a constraint that would line up the focal  $H^-$  peak with the right edge of the word. This is illustrated by the example in Figure 2, where the word *månad* ‘month’ constitutes a highly-ranked  $C_f$  (new information) and where the focal  $H^-$  occurs at the right edge

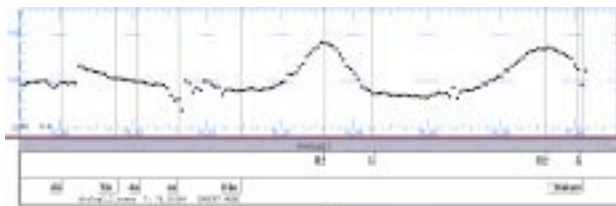


**Figure 2:** F0 curve showing the Accent 2 word *månad* ‘month’ in the utterance *Vilken månad vill du åka?* ‘What month do you want to travel?’ where it constitutes a highly-ranked  $C_f$  (new information). Notice the Focal  $H^-$  which occurs at the right edge of the word.

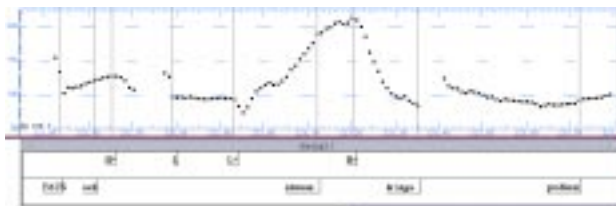
<sup>1</sup> Within the LINLIN model, the concept *light* would instead be associated with what is termed the ‘Conceptual model’ [6] which contains more generic knowledge than the ‘Domain model’.

of the word. **Align XP** on the other hand, is a constraint which aligns syntactic and prosodic boundaries [15]. The **Align Focus** constraint does not take into consideration syntactic boundaries. The **Align XP** constraint has the potential of making each maximal phrase category a prosodic phrase. This constraint is the one which would lead to the association of a phrase accent (L), in addition to the focal H<sup>-</sup> at the right edge of the focussed constituent. Compare Figure 2 with Figure 3 where the word *Malmö* is associated with both a focal H<sup>-</sup> and a phrasal L tone.

In initial position, **Align XP** can come into conflict with a third constraint **Wrap XP**, also a syntactically-based constraint [17]. This latter constraint has the function of ‘wrapping’ an entire utterance into one prosodic phrase whereas **Align XP**, unconstrained, would lead to an utterance having as many prosodic phrases as there are maximal phrase categories. In speech production, however, the **Wrap XP** constraint has the effect of weakening or even deleting the phrase boundaries after each maximal XP. Figure 4 presents an example of a focal accent on a C<sub>b</sub> where the H<sup>-</sup>, unlike in Figure 2, occurs to the right of the C<sub>b</sub> (=utresan), i.e. in the following word (*är* ‘is’). Our hypothesis is then that the **Wrap XP** constraint takes precedence over the **Align Focus** and the **Align XP** constraints as regards the accenting and phrasing behaviour of phrase-initial C<sub>b</sub>. We expect, furthermore, that the C<sub>b</sub> is optimally assigned a prominent focal accent. This prominence, together with a phrasing that groups the C<sub>b</sub> together with the rest of the utterance creates the impression of a coherent, well-planned speech unit.



**Figure 3:** F0 curve on the Accent 2 word *Malmö* in the utterance *Då får du en från Malmö* ‘You can get one [i.e. flight] then from Malmö’ which is associated with both a focal H<sup>-</sup> and a following L phrasal tone.



**Figure 4:** F0 curve on the Accent 2 word *utresan* ‘trip out’ in the utterance *Utresan är inga problem* ‘The trip out is no problem’ associated with a focal but late-timed H<sup>-</sup>.

## 5. FOCAL VS. NON-FOCAL, EARLY VS. LATE-TIMED FOCAL H<sup>-</sup> : WHAT DO LISTENERS PREFER?

Since the number of occurrences of the late-timed focal H in our current database is limited, it is not possible to say with any certainty using just the production data if the hypotheses we are making are statistically valid. It is important, however, for the development of the man-machine dialogue system to know, for example, if a particular intonation contour is more natural or preferred by users. In our case, we are interested in getting some initial answers to two questions:

- i) Do speakers prefer focal accents instead of non-focal accents on C<sub>b</sub>?
- ii) If the answer to i) is affirmative, do speakers prefer a focal accent with a late timing of the H<sup>-</sup> (i.e. outside the focussed word) or do they prefer the H<sup>-</sup> to be located in the focussed word?

In an attempt to get some insight into the issue, we decided to conduct a listener test where we could present listeners with pairs of short ‘man-machine’ Initiative-Response dialogues where the ‘machine’ Responses have a C<sub>b</sub> associated with different F0 contours (and where the C<sub>b</sub> is the same as the C<sub>f</sub> in the preceding ‘human’ Initiative). We hypothesized that the focal contour with the late timing of the H<sup>-</sup> (**Wrap XP** >> **Align Focus**) would be the preferred contour since it gives the intuitive impression of a more engaged and interactive dialogue participant who picks up on the preceding utterance’s most highly-ranked center and comments on it.

## 6. PERCEPTION TEST

### 6.1. Construction of Stimuli

As stimuli, we used short Initiative/Response dialogues where the Initiative was produced by a human and the Response was a synthetic voice. The Initiative utterance contained a domain-specific C<sub>f</sub> which also constituted the C<sub>b</sub> of the Response. The Responses are synthetic versions of actually occurring spontaneous travel-agent utterances from our data-base. Following in (1) are the dialogues used. The C<sub>b</sub> words whose accentual patterns are manipulated in the experiment are written in bold:

- (1) **A: Accent 1 words:**
  - i) I: Kan du se vad ni har på charter där?  
‘Can you see what you have in charter there?’  
R: **Charter** är ju helt fullbokat.  
‘Charter is totally booked up.’
  - ii) I: Men då vet man inte vilket pris det blir?  
‘But then you don’t know what the price will be?’  
R: Nej **priset** är inte klart ännu.  
‘No, the price is not fixed yet’.
  - iii) I: Flyg till Paris den tolfte, trettonde april.  
‘Flight to Paris the twelfth, thirteenth of April.’

R: **Tolfte** ser jag att det är fullt.  
'Twelfth is full I see'

**B) Accent 2 words**

iv) I: Jag skulle vilja flyga till Paris så billigt som möjligt.

'I would like to fly to Paris as cheap as possible'

R: Det **billigaste** vi har är Air France.

'The cheapest we have is Air France'

v) I: Hur är det med bokningen då?

'What's the situation with the booking?'

R: **Bokningen** är alltså hundra procent klar.

'The booking is one hundred percent ready'

vi) I: Är utresan lika jobbig?

'Is the trip out just as hard?'

R: **Utresan** är inga problem. Då byter man bara plan.

'The trip out is no problem. You just have to change planes'

## 6.2. Generation of Synthetic Responses

On the basis of the transcription of the dialogue data, the travel agent utterances were generated using the LUKAS concatenation synthesis [7]. Each phone's duration was initially adjusted to correspond to the original duration (see, however, (3)-(4) below).

Three prosodically different versions of the synthesized Response utterances were generated with different accentual patterns on the initial  $C_b$ :

**(2) F0 Patterns:**

- Version 1 (**Non-Focal**): non-focal accent generated on  $C_b$ ;
- Version 2 (**Early Focal  $H^-$** ): focal accent with  $H^-$  and boundary L inside  $C_b$ ,
- Version 3 (**Late Focal  $H^-$** ): focal accent with  $H^-$  outside  $C_b$  (i.e. in following word)

The F0 generation rules in LUKAS were used to first generate an intonation contour that corresponded as close as possible to the original (using F0-rules associating labels for word accents and focal accents to F0 values).

In order to create Version 1 (Non-Focal  $H^-$ ) from this base version, the focal accents on the  $C_b$  were associated with non-focal word accents (see Figure 5, top F0 curve) generated without the  $H^-$  label after the word accent.

In order to generate Version 2 (Early Focal  $H^-$ ), a L phrase tone was generated at the end of the  $C_b$  and the focal  $H^-$  was moved to the left in order to accommodate the fall from the  $H^-$  (see Figure 5, bottom F0 curve). Moreover, a certain degree of final lengthening was in some cases added to the final syllable rime of the  $C_b$  in order to make the synthesis sound reasonably natural with the more complex F0 pattern ( $H^-$  L). The segments affected and the degree of lengthening were as follows:

**(3) Final Lengthening:**

- **Bokningen** - /e/ 30%, /n/ 50%
- **Charter** - /r/ 185%
- **Priset** - /e/ 30%,
- **Tolfte** - /t/ 30%, /e/ 50%
- **Utresan** - /a/ 30%, /n/ 50%

In order to generate Version 3 (Late focal  $H^-$ ) with F0 outside the  $C_b$  (see Figure 5, middle F0 curve) the Praat speech analysis package was used to place the F0 peaks in the following word. (The LUKAS rules in their present form place the  $H^-$  midway between the word accent H and a following L% or word accent H). In cases where the following word began with a consonant, the  $H^-$  peak was placed at the beginning of the following word's first vowel; in cases where the following word began with a vowel, the peak was placed at the end of the vowel. The only exception was the utterance containing the  $C_b$  **tolfte**, where, because of the following word accent on *ser*, the F0 peak could not be placed further to the right than the end of the final vowel in *tolfte*.

In all versions, a L phrase tone was generated at the end of the last voiced segment in the utterance. Finally, in order to create approximately the same rate of speech in all the synthetic Responses, a number of further duration manipulations were carried out. These were:

**(4) Speech-Rate manipulations:**

Response (ii): - whole utterance lengthened by 20%

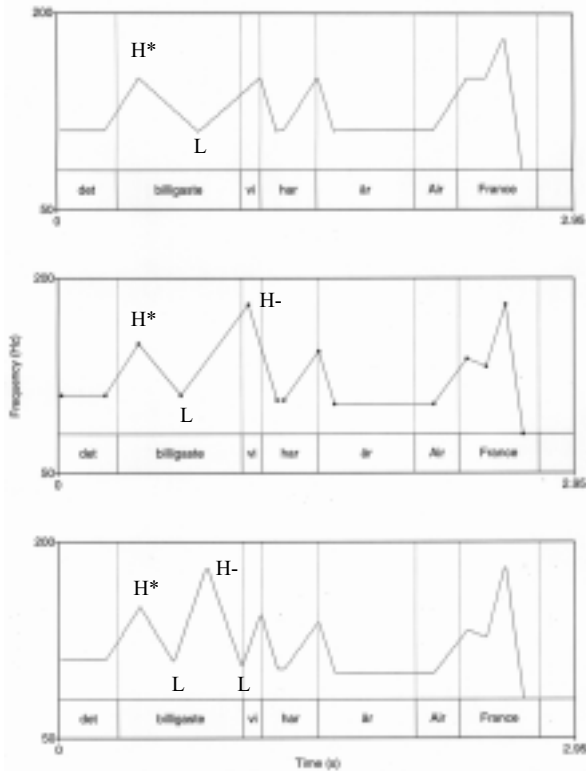
Response (v): - whole utterance first lengthened by 20%, then the rime in final word *klar*: /a/ lengthened by 30%, /r/ by 50%

Response (vi): - whole utterance lengthened by 10%

The dialogues in (1) were then put together where the Initiatives, uttered by a male speaker of Southern Swedish, were grouped together with the synthesized Responses. The Version 3 Responses (Late focal  $H^-$ ), which we hypothesized would be the preferred (more *engaged*) ones were thus compared with either Version 1 (with non-focal accent) or Version 2 (Early focal  $H^-$  and boundary L inside  $C_b$ ) in each pair of dialogues. The dialogue pairs were thus organized in the following way: Version 1 compared to Version 3, Version 3 compared to Version 1, Version 3 compared to Version 2 and Version 2 compared to Version 3. The order of presentation of the dialogue pairs was done in a random manner. All in all there were thus 24 dialogue pairs for the listeners to judge.

## 6.3 Subjects and Task

Twenty listeners (colleagues and friends) agreed to participate in the listening test. They were asked to listen to the dialogue pairs one at a time and to choose the one in which they deemed the 'machine' to sound more 'engaged' in the dialogue. By clicking on a number on a computer screen corresponding to one of the 24 dialogue pairs they could listen to a dialogue pair as many times as they wished before making a 'forced' choice.



**Figure 5:** An example of the different F0 contours generated on the  $C_b$  *billigaste* in the machine Response in (1)(iv). From top to bottom these are: Non-Focal contour, Focal contour with late  $H^-$  (resulting from **Wrap XP** >> **Align Focus**), and Focal contour with early  $H^-$  and phrasal L (resulting from **Align Focus** >> **Align XP**).

## 6.4 Results

We will discuss the results in relation to the two questions that we posed above, namely:

- i) Do speakers prefer focal accents instead of non-focal accents on domain-related  $C_b$ ?
- ii) If the answer to i) is affirmative, do speakers prefer a focal accent with a late timing (i.e. outside the focussed word) or do they prefer the  $H^-$  to be located in the focussed word?

As regards i), results from the dialogues contrasting Version 1 (Non-focal) and Version 3 (Late Focal  $H^-$ ) show (with all results pooled together) that listeners choose a focal accent (Version 3 (Late Focal  $H^-$ )) over a non-focal accent (Version 1) 70,4% of the time on a  $C_b$ . Thus the results lend support to the hypothesis that given information that functions as a  $C_b$  is preferably focally accented. A  $\chi^2$  test comparing the results for each individual dialogue pair shows a statistically reliable preference ( $p < 0.05$ ) for Version 3 (Late Focal  $H^-$ ) in 6 of the 12 dialogue pairs (See Table 1). The results show further no statistically significant preference for the non-focal version. For the  $C_b$  = **charter** (1-i) and  $C_b$  = **utresan** (1-vi) Responses,

the preference for the Late-focal  $H^-$  was statistically significant for both presentation orders of the Response version. One of the dialogues ( $C_b$  = **bokningen**) appeared to be particularly problematic for the listeners. Reactions from listeners after the test indicated that all three Response versions seemed to be “unengaged” in comparison with the human Initiative. This impression could be very well due to a lack of F0 register ‘coherence’ or ‘synchrony’ between the Initiative and the Response utterances, i.e. the  $C_b$  has a narrower range than the  $C_f$ . In other words, it quite likely the case that the register width of the focal accent on the  $C_b$  should not be narrower than that on the user’s Initiative in order for the Response to sound *engaged* (see [5]).

As regards ii), results from the the dialogues contrasting Version 2 (Early Focal  $H^-$ ) with Version 3 (Late Focal  $H^-$ ), show (with all results pooled together) that listeners preferred Version 3 in 64.2% of the cases. Thus, the results point to a certain preference for the focal pattern with the late timing of the focal  $H^-$  as hypothesized. A  $\chi^2$  test comparing the results for each individual dialogue pair shows a reliable preference for this version (Version 3) in 4 of the 12 dialogue pairs, whereas the results show no significant preference for Version 2. For the  $C_b$  = **tolfte** dialogue, the preference for Version 3 (Late Focal  $H^-$ ) was significant regardless of the order of the Response versions. Again, the same dialogue ( $C_b$  = **bokningen**) appeared to be particularly problematic for the listeners.

## 7. CONCLUSION

Although the results are not overwhelmingly conclusive, they do provide support for the hypotheses made at the outset. The focal accenting of  $C_b$  does appear to be a preferred strategy in Swedish as does the late timing of the focal  $H^-$  on  $C_b$ . That is to say, the **Wrap XP** constraint appears to be more highly ranked than **Align Focus** in utterances containing an initial domain-related focal  $C_b$ .

In order to pursue the investigation on the relationship between accentuation and information structure assumed here, we need to better control the stimuli. As we have observed, listeners had difficulty with stimuli where F0 range on the  $C_b$  did not synchronize with that on the  $C_f$ . Further, due to the fact that we used actually occurring Response utterances as test utterances, some of the differences between the two focal versions were extremely small due to the absence of unstressed syllables between the focal  $H^-$  and the following word accent. Even control of listener dialect can be thought to be important since the realization of word-accent and focal-accent patterns differs according to speaker dialect.

## 8. ACKNOWLEDGMENTS

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

Non-Focal + Late Focal H <sup>-</sup>	Late Focal H <sup>-</sup> + Non-Focal
<b>Billigaste:</b> Late-Focal H <sup>-</sup> : 10, Non-focal: 10, $\chi^2_{\text{obt}} = 0,00$	<b>Billigaste:</b> Late Focal H <sup>-</sup> : 15, Non-Focal: 5, $\chi^2_{\text{obt}} = 5,00^*$
<b>Bokningen:</b> Late-Focal H <sup>-</sup> : 13, Non-Focal: 7, $\chi^2_{\text{obt}} = 1,80$	<b>Bokningen:</b> Late Focal H <sup>-</sup> : 14, Non-Focal: 6, $\chi^2_{\text{obt}} = 3,20$
<b>Charter:</b> Late-Focal H <sup>-</sup> : 15, Non-Focal: 5, $\chi^2_{\text{obt}} = 5,00^*$	<b>Charter:</b> Late Focal H <sup>-</sup> : 15, Non-Focal: 5, $\chi^2_{\text{obt}} = 5,00^*$
<b>Priset:</b> Late Focal H <sup>-</sup> : 14, Non-Focal: 6, $\chi^2_{\text{obt}} = 3,20$	<b>Priset:</b> Late Focal H <sup>-</sup> : 15, Non-Focal: 5, $\chi^2_{\text{obt}} = 5,00^*$
<b>Utresan:</b> Late-Focal H <sup>-</sup> : 16, Non-Focal: 4, $\chi^2_{\text{obt}} = 7,20^*$	<b>Utresan:</b> Late-Focal H <sup>-</sup> : 17, Non-Focal: 3, $\chi^2_{\text{obt}} = 9,80^*$
<b>Tolfte:</b> Late-Focal H <sup>-</sup> : 13, Non-Focal: 6, $\chi^2_{\text{obt}} = 2,58$	<b>Tolfte:</b> Late-Focal H <sup>-</sup> : 14, Non-Focal: 6, $\chi^2_{\text{obt}} = 3,20$

Late-Focal H <sup>-</sup> + Early-Focal H <sup>-</sup>	Early-Focal H <sup>-</sup> + Late-Focal H <sup>-</sup>
<b>Billigaste:</b> Late-Focal H <sup>-</sup> : 14, Early-Focal H <sup>-</sup> : 5, $\chi^2_{\text{obt}} = 4,05^*$	<b>Billigaste:</b> Late-Focal H <sup>-</sup> : 9, Early-Focal H <sup>-</sup> : 10, $\chi^2_{\text{obt}} = 0,05$
<b>Bokningen:</b> Late-Focal H <sup>-</sup> : 10, Early-Focal H <sup>-</sup> : 10, $\chi^2_{\text{obt}} = 0,00$	<b>Bokningen:</b> Late-Focal H <sup>-</sup> : 10, Early-Focal H <sup>-</sup> : 10, $\chi^2_{\text{obt}} = 0,00$
<b>Charter:</b> Late-Focal H <sup>-</sup> : 11, Early-Focal H <sup>-</sup> : 9, $\chi^2_{\text{obt}} = 0,20$	<b>Charter:</b> Late-Focal H <sup>-</sup> : 11, Early-Focal H <sup>-</sup> : 9, $\chi^2_{\text{obt}} = 0,20$
<b>Priset:</b> Late-Focal H <sup>-</sup> : 16, Early-Focal H <sup>-</sup> : 4, $\chi^2_{\text{obt}} = 3,20$	<b>Priset:</b> Late-Focal H <sup>-</sup> : 13, Early-Focal H <sup>-</sup> : 7, $\chi^2_{\text{obt}} = 1,80$
<b>Tolfte:</b> Late-Focal H <sup>-</sup> : 15, Early-Focal H <sup>-</sup> : 5, $\chi^2_{\text{obt}} = 5,00^*$	<b>Tolfte:</b> Late-Focal H <sup>-</sup> : 15, Early-Focal H <sup>-</sup> : 5, $\chi^2_{\text{obt}} = 5,00^*$
<b>Utresan:</b> Late-Focal H <sup>-</sup> : 18, Early-Focal H <sup>-</sup> : 2, $\chi^2_{\text{obt}} = 12,80^*$	<b>Utresan:</b> Late-Focal H <sup>-</sup> : 13, Early-Focal H <sup>-</sup> : 7, $\chi^2_{\text{obt}} = 1,80$

**Table 1:** Results from Perception test. The two columns present results for the two orders of presentation of the Response versions. The number of responses for each version are indicated following the accent type (Non-Focal, Late-Focal H<sup>-</sup>, and Early Focal H<sup>-</sup>). For the  $\chi^2$  test:  $\chi^2_{\text{crit}} = 3,841$  (df = 1,  $\alpha = 0.05$ ). Statistically significant results are marked with an asterisk (\*).

## 9. REFERENCES

- Ariel, M., "Referring and accessibility", *Journal of Linguistics* 24: 65-87.
- Bruce, G., *Swedish Word Accents in Sentence Perspective*. Gleerups, Lund, 1977.
- Bruce, G., "How floating is focal accent", *Nordic Prosody* 4, ed. by K. Gregersen & H. Basbøll, 41-49. Odense, Odense University Press, 1987.
- Bruce, G., Filipsson, M., Frid, J., Granström, B., Gustafson, K., Horne, M., House, D., Lastow, B. and Touati, P. "Developing the modelling of Swedish prosody in spontaneous dialogue." *Proc. ICSLP 96*: 370-373, 1996.
- Buder, E. and Eriksson, A., "Prosodic cycles and interpersonal synchrony in American English and Swedish". *Proc. Eurospeech 97*: 235-238, 1997.
- Dahlbäck, N. and Jönsson, A., "Integrating domain specific focusing in dialogue models", *Proc. Eurospeech 97*, 2215-2218, 1997.
- Filipsson, M. and Bruce, "LUKAS - a preliminary report on a new Swedish speech synthesis." *Working Papers* (Dept. of Ling., Lund Univ.) 46: 45-56. 1997.
- Grosz, B., Joshi, A., and Weinstein, S. "Centering: A framework for modelling the local coherence of discourse", *Computational Linguistics* 21: 204-225, 1995.
- Horne, M. "Why do Speakers Accent 'Given' Information?" *Proc. Eurospeech91*, Vol. 3: 1279-1282, 1991.
- Horne, M., Filipsson, M., Ljungqvist, M., and A. Lindström, "Referent Tracking in Restricted Texts using a Lemmatized Lexicon: Implications for Generation of Prosody". *Proc. Eurospeech '93* Vol. 3: 2011-2014, 1993.
- Jönsson, A., "A Model for Habitable and Efficient Dialogue Management for Natural Language Interaction". *Natural Language Engineering* 3(2/3): 103-122, 1997.
- Kameyama, M., "Stressed and unstressed pronouns: complementary preferences", *Working Papers of the Institute of Logics and Linguistics* (IBM Deutschland Informationssysteme GmbH, Heidelberg) 8, Vol. 3: 475-484, 1994.
- Kanerva, J. "Focusing on phonological phrases in Chichewa", in *The Phonology-Syntax Connection*, ed S. Inkelas & D. Zec. Chicago: University of Chicago Press, 145-162, 1990.
- Prince, E. "Towards a taxonomy of Given-New information", in *Radical Pragmatics*, ed. P. Cole. New York: Academic Press, 223-256, 1981.
- Selkirk, E., "The interaction of constraints on prosodic phrasing". In *Prosody: Theory and Experiment. Studies Presented to Gösta Bruce*, ed. M. Horne. Dordrecht: Kluwer, In press.
- Terken, J. and J. Hirschberg, "Deaccentuation of Words Representing 'Given' Information: Effects of Persistence of Grammatical Function and Surface Position". *Language and Speech* 37: 125-145, 1994.
- Truckenbrodt, H. "On the Relation between Syntactic Phrases and Phonological Phrases", *Linguistic Inquiry* 30: 219-255, 1999.