The prosody of contrastive topics in Southern Swedish

Ambrazaitis, Gilbert; Frid, Johan

Published in:
Proceedings FONETIK 2012 : The XXVth Swedish Phonetics Conference, May 30–June 1, 2012

2012

The prosody of contrastive topics in Southern Swedish

Gilbert Ambrazaitis and Johan Frid
Centre for languages and literature, Lund University

Abstract

This paper presents a pilot study on the prosodic marking of a contrastive topic in Southern Swedish. A test sentence was elicited in three experimental conditions: initial focus; final focus; contrastive topic (initial word) plus focus (final word). F0 patterns were analysed in recordings of 10 speakers. A majority of the speakers distinguished clearly between the conditions, but speakers employed different strategies. The most common one involved modifying the range and/or level of the F0 patterns given by the lexical pitch accents. Another frequent strategy involved a rising pitch accent on the initial word, where the lexical pitch accent would normally stipulate a fall. This result was unexpected and indicates that Southern Swedish sentence intonation might be more complex than typically assumed.

Introduction

In many languages such as German, English, or Central (e.g., Stockholm) Swedish, the focus of an utterance is signalled by means of a sentence accent (or focal accent). Also in languages that lack sentence accents – at least in the sense of *additional* post-lexical pitch gestures – such as Southern Swedish or Mandarin Chinese, focus can be signalled prosodically by means of F0-modifications (e.g., increasing the pitch range) of the lexical tone or accent (see, e.g., Bruce and Gårding, 1978, for Southern Swedish, Wang and Xu, 2011, for Mandarin).

Beyond the marking of focus, prosody can fulfil a couple of further functions related to information structure, which are unexplored as far as Southern Swedish is concerned. This paper presents a pilot study on the prosodic marking of a *contrastive topic* (e.g., Molnár, 2006), as exemplified in (1).

(1) A: What do wild animals eat?
B: [Bears]$_{CT}$ eat [berries]$_{Focus}$.

A contrastive topic (CT) can be conceived of as a ”topic in focus” (Molnár, 2006), a characterisation which suggests that a CT is marked prosodically in a similar – but not necessarily identical – way as a focus. There are a couple of differences between a CT and a focus. First, focus does not always imply contrast, and when it does (in contrastive focus), this contrast is of another kind as in a CT (*exhaustive* contrast in focus; *non-exhaustive* in CT, cf. Molnár, 2006).

Second, a CT is bound to the sentence-initial position and always followed by a focus later in the sentence (as in (1) above), while a focus can, and often does, occur without a preceding CT. Therefore, if a CT is highlighted prosodically, then a sentence containing a CT can be expected to contain two sentence-level prominences (one for the CT, one for the focus). However, also in sentences lacking a CT – containing a focus only – a sentence-initial prominence (also referred to as *pre-nuclear accent*) can often be observed, as for instance in German (e.g., Ambrazaitis, 2009), or in Central Swedish (Bruce, 1982).

In order to investigate the prosodic effects of CT, we thus need to compare sentences with a contrastive topic and a focus (henceforth, CT-F) like B’s reply in (1) with corresponding sentences (a) with initial focus (henceforth, F-IN) – a somewhat marked, but fully possible situation – and (b) with final focus, but without preceding CT (henceforth, F-FI).

In German, a sentence with CT can be clearly distinguished from a sentence without CT by means of prosody: A CT is typically highlighted by a late-rising pitch accent (L*+H), which differs clearly from the default accent used for – initial or final – focus (H*), as well as from the pre-nuclear accent found in sentences with final focus but without CT. In addition, in a sentence with CT, the two accents for CT and focus are often connected by a high plateau, resulting in a so-called hat pattern (see, e.g., Ambrazaitis, 2009).

Ambrazaitis (2009) studied the prosodic marking of CT in Central Swedish using the same laboratory material as in this study (see Method). He found that the initial word in all three conditions mentioned above (CT-F, F-IN, F-FI) was realised with a rising sentence accent (H-); this rise had a larger F0 range and reached a higher maximum when the initial word was in focus (F-IN) than when not (F-FI); in a CT, the accent had approximately the same F0 range as in initial focus, but it was produced at a some-

what higher F0 level. These results suggest that Central Swedish distinguishes prosodically between CT and focus, but in a less categorical way than German.

Wang and Xu (2011) found for Mandarin Chinese what seems to involve the opposite of Ambrazaitis' finding for Central Swedish: a higher F0 in initial focus than in CT. However, their experimental condition for CT is not entirely comparable to ours.

Based on this background, we hypothesise that Southern Swedish might distinguish prosodically between the three information-structural conditions CT-F, F-IN, and F-FI, basically by means of F0-modifications of the lexically-conditioned pitch accents.

Method

The test sentence *Wallander förlänger till november* ('Wallander is continuing until November') was embedded in three dialogue contexts, in order to elicit focus on the initial word *Wallander* (F-IN), focus on the final word *november* (F-FI), or a contrastive topic on the initial word followed by focus on the final word (CT-F). Note that only words with Swedish Accent I were used in the material.

Each test context consisted of the situational frame context displayed in (2) and a unique context question shown in (3-5), where the test sentence is represented by three dots (. . .).

(2) *Du är polis och träffar en gammal kollega. Ni pratar om pensioneringen och om möjligheten att förlänga sin tjänst.*
'You are a police officer meeting a former colleague. You are talking about retirement and the possibility to continue working.'

(3) **F-IN: initial focus:**
Och vem förlänger till november? – **. . .**
'And who is continuing until November?'

(4) **F-FI: final focus:**
Och fram till när färlänger Wallander? – **. . .**
'And until when is Wallander continuing?'

For condition CT-F, the test sentence was embedded in an utterance frame:

(5) **CT-F: contrastive topic plus final focus:**
Och vet du redan fram till när dina kollegor förlänger? – ***Alltså , . . . Det är allt jag vet.***
'And do know already how long your colleagues will continue working? – Well, ... That's all I know.'

Speakers and recording procedure

The recordings of ten speakers of Southern Swedish were analysed in this study: five female (aged 26-47), five male (30-62).

Data were collected in an experimental studio at the Humanities Laboratory of Lund University, using the *BAS SpeechRecorder* software (Draxler and Jänsch), as follows: First, only the frame context (2) was displayed on a computer screen. After the speaker clicked a button, the (pre-recorded) context question (3-5) was played to her via headphones and simultaneously, the test sentence appeared on the screen (embedded in an utterance frame in the case of CT-F, see (5) above). The subject's task was to say the test sentence (in its utterance frame) as naturally and appropriately to the provided context as possible. (For further details, see Ambrazaitis et al., forthcoming.)

The material was randomised and mixed with similar materials used for other studies; each test context occurred five times in the recording session, resulting in 15 recorded utterances per speaker, or 150 in total.

The context questions were pre-recorded in the anechoic chamber at the Humanities Laboratory by a 39-year-old male speaker of Southern Swedish (from Malmö).

Data analysis

Data analysis in this pilot study was limited to an auditory and visual inspection of F0 patterns with the goal to look for general trends in the data. The F0 data were normalised for time and speaker, in order to support visual presentation and comparison of the data and to make it possible to calculate mean F0 contours across several repetitions of the same intonation patterns, as produced by different speakers.

For the time normalisation, we segmented the recordings into syllables and took five temporally equidistant F0 measurements for each syllable (the function word *till* and the voiceless fricative [f] in *förlänger* were excluded). We used the PRAAT script *Prosody Pro* (Xu, 2005-2011) for this purpose.

Our speaker normalisation involved converting the F0 measurements (in Hz) into semitones, at the same time rescaling them so that a speaker's base F0 value (*Fb*) would roughly correspond to 0 on the semitone scale. We estimated *Fb* using the formula proposed by Traunmüller and Eriksson (1995), in a modified version. (For further details, see Ambrazaitis et al., forthcoming.)

Results and discussion

In accordance with the hypothesis, the final word in the test sentence (*november*) was always – by all speakers and in all conditions – produced with a falling pitch accent (henceforth, *fall*) typical for Accent I in Southern Swedish (see Fig. 1).

However, the initial word (*Wallander*) was either produced with a *fall* or with a *rise*, i.e. a pattern rising from the pre-stress to the stressed syllable and ending in a high plateau on the final syllable (see Fig. 2).

Taking into account the experimental conditions, we observed that speakers made use of theses two basic prosodic patterns found in the initial word in three different ways:

- **Group A:** five speakers (three female, two male) produced a *fall* (on the initial word) in all conditions;
- **Group B:** four speakers (one f, three m) used either an initial *fall* or a *rise* depending on the condition, even though the distribution of patterns across conditions varied in this group;
- **Group C:** one female speaker mostly used an initial *rise* in all conditions.

We classified a speaker as belonging to a specific groups if s/he had produced the corresponding pattern(s) in at least three out of five repetitions; occasional deviating patterns occurred in all groups. For both groups A and B, we found that a majority of the speakers distinguished clearly between the experimental conditions.

Speakers of group A

Figure 1 displays normalised F0 curves averaged across all cases with initial *falls* in the recordings of the group A speakers. As the figure shows, all three content words of the test sentence were produced with the typical Accent I pattern of Southern Swedish – a sharp F0 fall through the stressed syllable (which is always the second syllable in these words). Furthermore, the the speakers of group A varied F0 parameters in order to distinguish between the conditions.

As a general trend, the F0 range of the accent on the final word was larger when this word was in focus (F-FI and CT-F) than when it was out of focus (F-IN). For the initial word, an increased F0 range can be observed in the two conditions F-IN and CT-F, compared to F-FI; in addition, the F0 pattern is shifted to a somewhat higher level in CT. This upwards shift for CT is also observed in the medial word – even though this word is irrelevant for the definition of the test conditions – and is hence reminiscent of the high-pitch concatenation found in German (see Introduction).

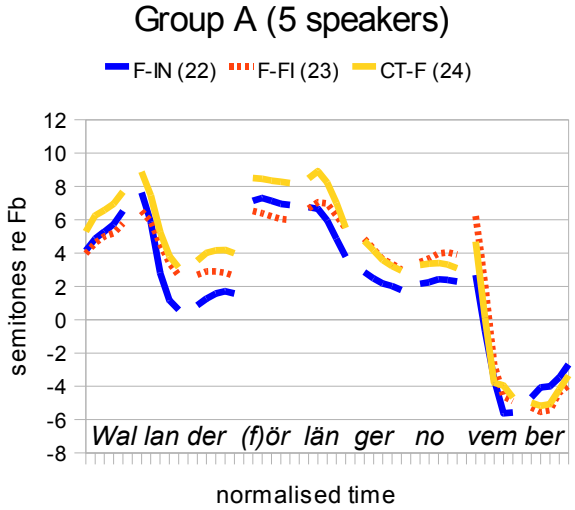

Figure 1. Mean F0 contours (*N* in parentheses) of cases with initial *fall* produced by speakers of group A in the three conditions F-IN, F-FI, and CT-F. The normalised time scale indicates the number of measurements; gaps in the curves indicate syllable boundaries (the function word *till* is excluded from this graph; there are hence nine syllables). The extremely low F0 values (< 0 st) reached in the last two syllables are due to creaky voice.

For at least three of the five speakers included in Figure 1, we can perceive rather clear differences in prominence relations between the conditions, despite seemingly subtle differences in F0 (strong prominence on the initial word in F-IN, on the final word in F-FI, and on both words in CT-F). However, further prominence cues beyond F0 may be involved here. One speaker, for instance, consistently produced a salient phrase boundary after the initial word in F-IN, and a strong hesitation before the final word in F-FI. We have also informally observed salient duration effects.

Speakers of groups B and C

Speakers of group B distinguished between the conditions partly by means of F0-adjustments like the speakers of group A, and partly by means of different types of F0 patterns. This is exemplified in Figure 2, which summarises the data of two of the four speakers of group B (one female, one male). These two speakers preferred an initial *fall* in conditions F-IN and F-FI, while they used an initial *rise* for the contrastive topic (CT-F). The conditions were thus much more clearly distinguished by these

speakers compared to speakers of group A.

The other two (male) speakers of this group behaved slightly differently: One also made a clear distinction between all three conditions, although he used the initial *fall* only in the final focus condition (F-FI), and the *rise* in both CT-F and F-IN. The last speaker in this group preferred the *rise* in CT-F and in F-FI, and the *fall* only in F-IN. He did not seem to make a clear distinction between CT-F and F-FI.

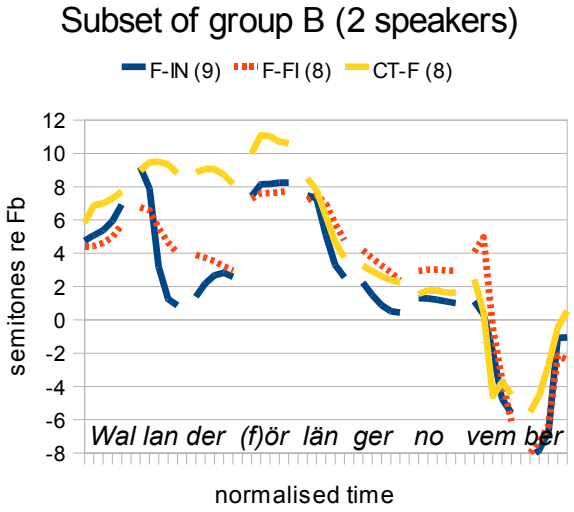

Figure 2. Mean F0 contours of cases with initial *rises* (observed in condition CT-F) and initial *falls* (F-IN and F-FI) produced by two speakers of group A. See Figure 1 for further explanations.

The single female speaker categorised as a separate group C preferred the initial *rise* in all conditions; she did not seem to make a clear distinction between CT-F and F-FI.

Conclusions

The major result of this pilot study was that a CT can be distinguished prosodically from focus in Southern Swedish more clearly than we have hypothesised. In particular, speakers employed two different strategies for signalling prominence on the initial word: Beyond the expected pattern – which involved modifying the range and/or level of the F0 patterns given by the lexical pitch accents – another frequent strategy involved a rising pitch accent on the initial word, where the lexical pitch accent would normally stipulate a fall. This result was unexpected and indicates that Southern Swedish sentence intonation might be more complex than typically assumed.

This study is part of the project *Function- and production-based modeling of Swedish prosody*, which aims at building a novel model of Southern and Central Swedish prosody, based on the idea of parallel encoding of communicative functions (e.g., Wang and Xu, 2011). In future research on CT, we will, among others, need to include materials with Accent II, investigate durational patterns, as well as investigate the perceptual relevance of the acoustic correlates of CT.

Acknowledgements

This study was supported by the Swedish Research Council (2009-1566).

References

Ambrazaitis G. (2009) Nuclear Intonation in Swedish – Evidence from Experimental-Phonetic Studies and a Comparison with German. Travaux de l'institut de linguistique de Lund 49. Lund: Lund University.

Ambrazaitis G., Frid J., and Bruce G. (forthcoming) Revisiting Southern and Central Swedish intonation from a comparative and functional perspective. In Niebuhr O. (ed) Prosodies – Context, Function, Communication, 135–158. Berlin: de Gruyter.

Bruce G. (1982) Developing the Swedish intonation model. Working Papers (Dept. of Linguistics, Lund University) 22, 51–116.

Bruce G. and Gårding E. (1978) A prosodic typology for Swedish dialects. In Nordic Prosody – Papers from a Symposium (Lund, Sweden, 1978), 219–228. Lund: Dept. of Linguistics, Lund University.

Draxler C. and Jänsch K. SpeechRecorder (Bavarian Archive for Speech Signals) [Computer program]. Available from: http://www.phonetik.unimuenchen.de/Bas/software/speechrecorder/

Molnár V. (2006) On different kinds of contrast. In Molnár V. and Winkler S. (eds) The architecture of focus, 197–233. Berlin: Mouton de Gruyter.

Traunmüller H. and Eriksson A. (1995) The frequency range of the voice fundamental in the speech of male and female adults. Manuscript.

Wang B. and Xu Y. (2011) Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. Journal of Phonetics 39, 565–611.

Xu Y. (2005-2011) ProsodyPro.praat. Available from: *http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/*