

# Prosodic Marking of Focus Domains - Categorical or Gradient?

Stefan Baumann, Martine Grice & Susanne Steindamm

I/L - Phonetik

University of Cologne, Germany

{stefan.baumann; martine.grice}@uni-koeln.de

## Abstract

This paper reports on a production experiment in German eliciting focus domains of various sizes, ranging from broad to narrow focus, as well as contrastive focus. Results show that speakers use categorical as well as gradient prosodic means to indicate different focus structures, with an increase of prominence-lending cues as the focus domain narrows. Contrast is shown to enhance certain differences between narrow and broad focus. There is a clear indication that speakers differ considerably as to the combination of strategies they employ for marking focus structure.

## 1. Theoretical background

In most approaches to information structure, whether a constituent is focussed or in the background represents a binary choice. This is also the case for contrast: a constituent is either contrasted with another one or not. The present study investigates the prosodic marking of these two aspects of information structure, and in particular whether they are expressed by categorical or gradient means.

### 1.1. Focus and accentuation

Focus is conceived of here as a genuinely semantic-pragmatic notion. It denotes the part of an utterance which the speaker presents as being important and/or which the speaker assumes to be most informative for the listener (see e.g. [13], [17]). The uninformative part of an utterance is often referred to as background.

In controlled experimental set-ups (like the one used here), focus is determined on the basis of the immediately preceding context. That is, the focus and the background part of a target utterance is defined in terms of Question-Answer Congruence (see [4]), as in (1).

(1) Q: Who did you call?

A: [ I called ]<sub>background</sub> [ MArY ]<sub>focus</sub>

Since the verb *call* has been mentioned in the question, and since the addressee is both textually (*you*) and situationally given, the constituents *I* and *called* form the background of the answer, whereas the newly introduced *Mary* is the focus. Note, however, that focussed information is not necessarily new in the sense of 'not known to the listener' (see [6]).

The structure in (1) represents a so-called *narrow focus*. In West-Germanic languages like English, German and Dutch, narrow focus is marked by a pitch accent on the focussed constituent (indicated by capitalisation of the first syllable of *Mary* in (1)). A special type of narrow focus is *contrastive focus*, which in this study involves a correction of what has previously been said. An example of contrastive focus used in a correction is in (2).

(2) Q: Did you call John?

A: No, [ I called ]<sub>background</sub> [ MArY ]<sub>focus</sub>

In *broad focus* structures, the focus is not restricted to a single constituent. The relation between focus and accent is no longer straightforward. Here, a pitch accent on one word marks a larger focus domain, a phenomenon called *focus projection* ([4],[21]).

(3) Q: What happened?

A: [ I called MArY ]<sub>focus</sub>

In (3), the complete answer is focussed. Thus, the word carrying the pitch accent, the *focus exponent*, is not coextensive with the focus domain, as in (1) and (2). It is claimed for West-Germanic languages that – in the unmarked case – arguments rather than predicates tend to qualify as focus exponents ([15]).

### 1.2. Categorical and gradient prosody

It is clear that the differences between the answers in (1), (2) and (3) are discrete: it is either [Mary] or [I called Mary] which is in focus, and [Mary] is either contrasted with another specific person, or is singled out from a larger set. The question we ask here is whether these differences are marked prosodically, and if so, whether the prosodic marking involves *discrete* means, i.e. *phonological categories* such as pitch accent type, or *gradient* means, such as duration, or F0 timing and scaling differences, which do not lead to a difference in phonological categories.

In an earlier study on the prosodic marking of broad versus narrow focus in German, [9] looked for categorical distinctions and did not find any. The results of a production experiment with nuclear accents occurring early in the target sentence revealed that speakers used the same nuclear pitch accent type (namely H\*L) across the two different focus domains, that is, in both broad and narrow focus, as in (4) and (5) below ([9]: 71, 188).

(4) Q: Was ist los? *What's wrong?*

A: [ ANna ist weggelaufen ]<sub>focus</sub>

*Anna has run away.*

(5) Q: Wer ist weggelaufen? *Who has run away?*

A: [ ANna ]<sub>focus</sub> [ ist weggelaufen ]<sub>background</sub>

*Anna has run away.*

It has recently been shown that categories in information structure can be prosodically marked using gradient means. Contrastive topics, or themes (occurring in prenuclear

position), in German have later and higher peak placement than non-contrastive ones [3].

Furthermore, stressed vowels have a significantly longer duration in contrastive themes. Similarly, Second Occurrence Focus (SOF; [10]) involves a longer duration of the target word. SOF is induced by a focus operator such as *even* after the main focus of the phrase, that is, in the unaccented stretch following the nuclear accent.

In both of the above cases gradient means are used to express a binary opposition: contrastive or non-contrastive themes; SOF or no focus. In our own study, the size of focus domains can be seen as discrete but not necessarily binary - as the size of focus domains can be extended step by step to include more and more constituents, bounded only by sentence length. Contrastive or non-contrastive narrow focus, on the other hand, can be considered a binary distinction, although not all theories distinguish narrow focus from contrast, since narrow focus is also contrastive in some way [20].

## 2. Production experiment

### 2.1. Recordings

A production experiment was designed to investigate whether prosodic means are used in German to differentiate between three different sizes of focus domain involving focus projection, and between these and narrow focus, and, within the narrow focus category, contrastive focus. Our hypotheses are based on the fact that gradient variation has been found to express other differences in information structure (see 1.2). However, this variation does not preclude a categorical distinction e.g. in pitch accent type.

#### 2.1.1. Reading material

Reading material consisted of five question-answer pairs with the answer 'Manuela will Blumen malen.' (*Manuela wants to paint flowers.*). The main criterion the target sentence had to fulfill was its continuous voicing, so as to be able to accurately measure exact peaks and valleys in the F0 contour. The questions are listed below, followed by the focus domains according to question-answer congruence.

Questions:

1. Was gibt's Neues? *What's new?*
2. Was gibt's Neues von Manuela? *What about Manuela?*
3. Was will Manuela? *What does Manuela want?*
4. Was will Manuela malen? *What does Manuela want to paint?*
5. Manuela will Gesichter malen? *Manuela wants to paint faces?*

Answers:

- |          |                            |             |
|----------|----------------------------|-------------|
|          | Manuela will Blumen malen. |             |
| 1.       | [ _____ ] focus            | broad       |
| 2.       | [ _____ ] focus            | ↓           |
| 3.       | [ _____ ] focus            | narrow      |
| 4.       | [ _____ ] focus            | contrastive |
| 5. Nein, | [ _____ ] focus            |             |

lit.: *Manuela wants flowers paint*

### 2.1.2. Speakers and recording procedure

Six speakers (three female, three male) between the ages of 23 and 27 took part in the experiment. All of them were students at the University of Cologne. Four speakers originated from the north-west of Germany, one from the west (just below the Benrath isogloss), and one from the north of Bavaria.

The recordings were carried out in a soundproof room, with the instructor reading out the questions, and the subjects giving the answers. The five sentences were interspersed with fillers and read aloud four times in randomised orders by each speaker, leading to 20 tokens per speaker. Thus, 120 utterances in total entered the analysis.

### 2.2. Analysis

Using the speech analysis tool EMU [5], we labelled the onset and the end of the nuclear word (which was the word *Blumen* in all cases), and the start and end of each segment. The prenuclear and nuclear pitch accents were transcribed in GToBI [11] with an additional label for the beginning of the nuclear rise. Example contours are given in Fig. 1.

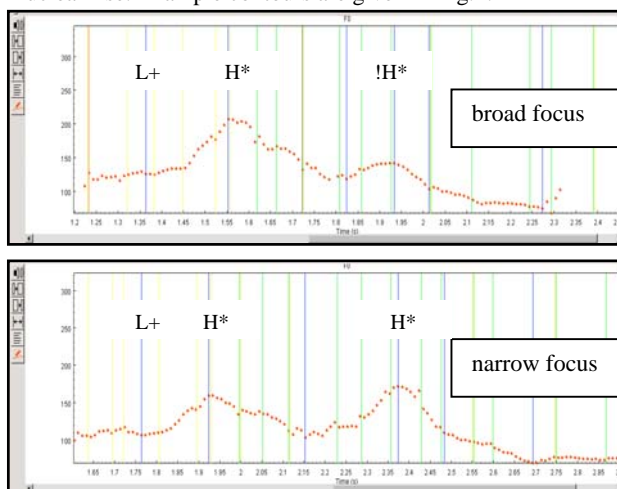


Figure 1: Example F0 contours for broad and narrow focus (answers 1 and 4, speaker CB)

## 3. Results and discussion

### 3.1. Categorical means

As a first result, contrary to predictions in the literature, both the size of the focus domain and type of focus affect the choice of accent *type* on the focus exponent: in broad(er) focus structures (sentences 1 and 2) a downstepped nuclear accent was produced in 42% of all cases, while in narrower focus domains (sentences 3 and 4) fewer downsteps occurred (25% and 17%, respectively). In contrastively focussed utterances no downstep was produced at all (Fig.2).

A Spearman's Rho correlation analysis showed a significant interaction between nuclear pitch accent type and sentence type ( $p < 0.001$ ).

Although it has been argued for English [8] that downstep cannot be reliably transcribed, implying that H\* and !H\* should not be treated as separate categories, a one-way ANOVA with "peak2 (Hz) minus peak1 (Hz)" as independent variable and accent type as factor revealed a highly significant difference between H\* and !H\* ( $p < 0.001$ ). These results

support the finding of [1] for English that !H\* was perceived as significantly less prominent than L+H\* or H\*.

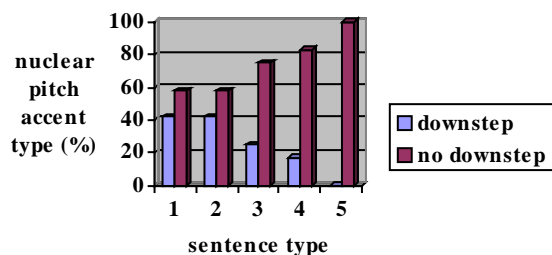


Figure 2: Differences in nuclear pitch accent type in relation to sentence type, all speakers (N=120)

In 20% of all cases there was no prenuclear H tone, since the nuclear accent was the only accent in the phrase (78% of the prenuclear accents were of the type (L+)H\* and 2% of the type L\*+H). Half of the single-accent phrases occurred in contrastive utterances, which is in line with the observation of [7] that the prominence of an accent can be increased by deaccenting other words in the phrase.

The observation already made by [3] in their investigation for contrastive and non-contrastive themes in German that speakers vary considerably as to the (combination of) means they employ for signalling aspects of information structure, is supported by our data.

As for the use of different accent types, for example, four out of six speakers use downstepped nuclear accents for marking broad focus and non-downstepped peak accents for marking narrow and, in particular, contrastive focus. The other two speakers do not use downstepping contours at all, i.e. all prenuclear and nuclear accents were of the type (L+)H\*.

### 3.2. Gradient means

As the focus domain narrows, we also observe the use of the following gradient means:

- increased duration of the focus exponent
- higher peak on the nuclear accent (marking the focus exponent)
- greater pitch excursion to the peak of the nuclear accent
- delay in the nuclear accent peak.

Across all speakers, duration varied consistently with the size of focus domain but it did not distinguish between contrast and non-contrast (Fig.3). In a one-way ANOVA with sentence type as independent factor, the focus domain had a highly significant effect on the duration of the focus exponent ( $p < 0.001$ ). Scheffé posthoc tests revealed significant differences between the two narrow focus sentence types (4 and 5) on the one hand, and the case with broad focus over the whole sentence (1) on the other.

In fact, all but one speaker ('SI') produced a more or less gradual increase of the focus exponent's duration from sentence 1 (broad focus) to 5 (contrastive focus). The values of two speakers ('NP' and 'CS') showed a highly significant interaction between these two variables. Posthoc tests revealed significant differences between broad focus and contrast for 'NP' and even between broad focus, narrow focus and contrast for speaker 'CS'. The gradual increase in duration is exemplified with the data of speaker 'NP' in Fig.3.

For one of the two speakers who did not use downstepped nuclear accents ('MG') we found a gradual increase from

sentence type 1 through 5 in the difference between the prenuclear and nuclear accent peak height, although there was no consistent pattern as to whether it was a lowering of the first peak or a raising of the second which was responsible for

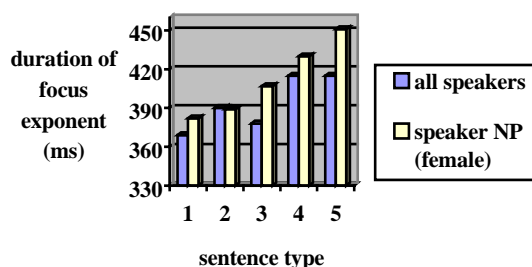


Figure 3: Differences in duration of focus exponent (mean values) in relation to sentence type, speaker NP

the difference. It is important here to note that an investigation of the prosodic marking of focus structures should not be restricted to nuclear accents alone but should include an examination of the relation between prenuclear and nuclear accents. This claim does not question the predominant role of nuclear accents in focus marking, though (see [21]).

An additional argument for the relevance of prenuclear accents is the strategy of yet another speaker ('SI') whose values for prenuclear pitch peaks significantly decrease with narrowing focus domains ( $p < 0.015$ ). This increases the potential difference in F0 to the nuclear accent peak and might thus enhance its degree of prominence.

Greater prominence has often been claimed to be marked by higher accent peaks ([19], [16]), which in turn has been claimed to be the most reliable cue to contrastive focus [2]. In fact, two speakers ('NP', 'CB') show a highly significant effect of nuclear accent pitch height on sentence type (see the data for 'NP' in Fig.4). Interestingly, the statistically different subgroups for both speakers of the Scheffé posthoc tests are sentences 1 to 4 on the one hand and sentence 5 on the other, i.e. non-contrast versus contrast.

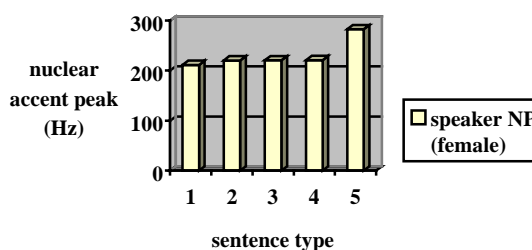


Figure 4: Differences in pitch height of nuclear accent peak (mean values) in relation to sentence type, speaker NP

As [12] points out, perceived prominence is not a correlate of pitch height, but of relative pitch excursion. Our production data show that for all speakers the nuclear rise excursion in sentence type 1 (broad focus) is significantly smaller than the rise in sentence type 5 (contrastive focus). Looking at specific speakers, we find similar pitch excursion effects for the two speakers which already showed significant effects for pitch height ('NP', 'CB'). When measuring the extent of the nuclear rise in semitones, only speaker 'NP's results reach significance (again with non-contrast versus contrast as subgroups), although the data for 'CS' and 'CB' also indicate tendencies towards a gradual increase of the pitch excursion as the focus domain narrows, plus a sharp increase from narrow to contrastive focus (Fig.5).

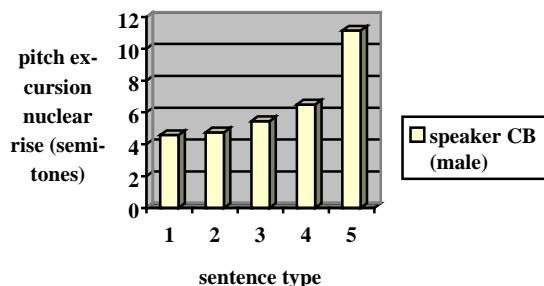


Figure 5: Differences in pitch excursion of nuclear accent rise (mean values) in relation to sentence type, speaker CB

Another indicator of prominence which might also be an indicator of narrow(er) focus or contrast, consists in late accent peaks. The effect of higher prominence of late peaks is attested in perception studies on English [16] and German [14]. There is also indirect evidence from production data in a variety of German. [18] found that in Hamburg German narrow focus expressions are marked by late peaks as opposed to broad focus expressions.

From our six speakers, two speakers showed such a trend ('MG' and 'NP'): the smaller the focus domain, the later the peak measured in ms from the accented syllable onset. This trend was highly significant for one speaker ('NP':  $p < 0.004$ ). Scheffé posthoc tests revealed that broad focus (1, 2, and 3) was significantly different from contrastive focus, although broad and narrow focus were not significantly different (Fig.6).

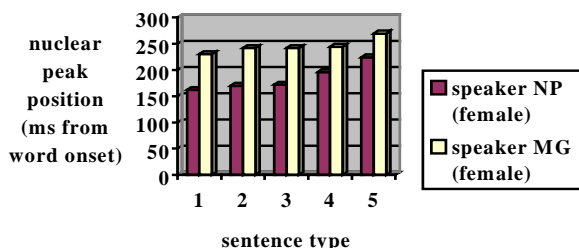


Figure 6: Differences in nuclear peak position (mean values) in relation to sentence type, speaker MG

In sum, speakers predominantly used duration to distinguish between broad and narrow focus, and differed more in their strategies for distinguishing between contrast and non-contrast.

#### 4. Conclusions

We have shown that speakers have a choice between a variety of different strategies for marking focus domains of differing sizes, and different types of focus. There is considerable variation as to what extent each strategy is used by any individual speaker. Only one speaker makes use of (nearly) all strategies examined, others restrict themselves to a specific strategy such as varying the duration of the focus exponent, the height of the nuclear peak (in relation to a potential prenuclear accent), the position of the nuclear peak, and the excursion size of the rise leading up to the accentual peak. In addition to these gradient prosodic means for the marking of broad and narrow focus, speakers use categorically different pitch accent types, in particular downstepping versus non-downstepping contours.

Contrast enhances some of the differences between narrow and broad focus, in that it is never marked by

downstep, and in that it reaches significantly higher values in comparison to the other focus structures in terms of peak height, rise excursion and peak delay. These cues undoubtedly play a role in increasing the prominence of contrasted constituents.

#### 5. References

- [1] Ayers, G., 1996. Nuclear accent types and prominence: Some psycholinguistic experiments. PhD Thesis, Ohio State University.
- [2] Bartels, C.; Kingston, J., 1994. Salient Pitch Cues in the Perception of Contrastive Focus. In *Focus and Natural Language Processing*, P. Bosch, R. van der Sandt (eds). IBM Working Papers 6.
- [3] Braun, B.; Ladd, D.R., 2003. Prosodic correlates of contrastive and non-contrastive themes in German. *8th European Conference on Speech Communication and Technology*, September 2003, Geneva, 789-792.
- [4] Büring, D., 2003. Focus Projection and Default Prominence. Ms. University of California, Los Angeles.
- [5] Cassidy, S.; Harrington, J., 2001. Multi-level annotation in the EMU speech database management system. *Speech Communication* 33 (1-2), 61-78.
- [6] Chafe, W., 1994. *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press.
- [7] Cruttenden, A., in press. The deaccenting of given information: a cognitive universal? In *The pragmatic organisation of discourse*, G. Bernini (ed.). The Hague: de Gruyter.
- [8] Dainora, A., 2001. Eliminating downstep in prosodic labeling of American English. *Workshop on Prosody in Speech Recognition and Understanding*, 41-46.
- [9] Féry, C., 1993. *German Intonational patterns*. Tübingen: Niemeyer.
- [10] Féry, C.; Ishihara, S., to appear. Interpreting Second Occurrence Focus. In *Methods in Empirical Prosody Research*, S. Sudhoff et al. (eds). Berlin: de Gruyter.
- [11] Grice, M.; Baumann, S.; Benzmüller, R., 2005. German Intonation in Autosegmental-Metrical Phonology. In *Prosodic Typology. The Phonology of Intonation and Phrasing*. Sun-Ah Jun (ed.). Oxford: Oxford University Press. 55-83.
- [12] Gussenhoven, C., 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- [13] Halliday, M.A.K., 1967. Notes on transitivity and theme in English, Part 2. *Journal of Linguistics* 3, 199-244.
- [14] Kohler, K., 1991. Terminal Intonation Patterns in Single-Accent Utterances of German: Phonetics, Phonology and Semantics. *AIPUK* 25, 115-185
- [15] Ladd, D.R., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- [16] Ladd, D.R.; Morton, R., 1997. The Perception of Intonational Emphasis: Continuous or Categorical? *Journal of Phonetics* 25, 313-342.
- [17] Lambrecht, K. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.
- [18] Peters, J., 2002. Intonation und Fokus im Hamburgischen. *Linguistische Berichte* 189, 27-57.
- [19] Rietveld, T.; Gussenhoven, C., 1985. On the Relation between Pitch Excursion Size and Prominence. *Journal of Phonetics* 13, 299-308.
- [20] Rooth, M., 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75-116.
- [21] Welby, P., 2003. Effects of Pitch Accent Position, Type, and Status on Focus Projection. *Language and Speech* 46(1), 53-81.