

TONAL AND ARTICULATORY MARKING OF FOCUS IN GERMAN

Stefan Baumann, Johannes Becker, Martine Grice & Doris Mücke

IfL Phonetik, University of Cologne

{stefan.baumann; becker.johannes; martine.grice; doris.muecke}@uni-koeln.de

ABSTRACT

A production study on read German speech shows an increase in prominence-lending parameters as the focus domain is narrowed, or when contrast is expressed over non-contrast. Prominence-lending parameters included tonal and durational cues (e.g. nuclear pitch accent type and scaling, duration of segments and syllables) as well as those relating to supralaryngeal articulation (e.g. formant values of vowels in nuclear syllables). These cues are exploited by different speakers to different degrees and in different combinations, but are all used for signalling aspects of focus.

Keywords: focus, contrast, pitch accent, downstep, localised hyperarticulation

1. INTRODUCTION

In studies examining the prosodic marking of focus in different languages, the investigation is often restricted to single items or constituents which are either structurally focussed, or part of the background, depending on their presence or absence in the immediately preceding context. Thus, usually only a specific kind of focus is investigated, and the question of the focussed or presupposed item's prosodic marking is generally limited to its (de)accentuation.

However, it is commonly accepted that different *types* of focus (e.g. contrast vs. non-contrast) as well as different *sizes* of focus domain (ranging from broad to narrow, see e.g. [9]), can and should be distinguished. Furthermore, a number of discrete as well as gradient prosodic means have been found to be used for marking focus, and to contribute to the degree of perceived prominence of the focus exponent. Some of these means are pitch accent type [1], accent peak height and timing [3], [10], pitch excursion [6], and duration of focussed constituents [8].

The role of articulation in focus marking has so far received only limited attention. Most of the studies on hyperarticulation restrict their analysis to accents and boundaries (e.g. [4] and [7] for English). However, [5] investigated lip movements in contrastive focus in French and found lips more protruded (hyperarticulated) under focus. In the present study, which is a refinement of [1], we investigate tonal, durational and articulatory ways of marking different

focus structures (broad, narrow, and contrastive – in fact, corrective – focus).

We expect an increase in prominence-lending cues ranging from broad to narrow to contrastive focus as a linguistic manifestation of the Effort Code [6] indicating increased emphasis or importance, expressed on three levels:

- *tonal* fewer prenuclear accents, fewer downsteps, higher, later and steeper nuclear accents
- *durational* increased duration of segmental domains (syllable, foot)
- *articulatory* increase in articulatory effort in vowels (localised hyperarticulation).

2. METHOD

2.1. Reading material

For each of eight target sentences we composed three question-answer pairs eliciting the three focus structures. The focus exponents of the target sentences consisted of either two or three syllables, with the stressed syllable containing one of the four long vowels /i:/, /a:/, /o:/ or /u:/. Each target sentence contained a proper noun in topic position and a postfocal verb. An example of a question-answer set is given below.

Questions:

1. Was gibt's Neues? *What's new?*
2. Was will Marlene schälen? *What does Marlene want to peel?*
3. Will Marlene eine Kartoffel schälen? *Does Marlene want to peel a potato?*

Answers:

- Marlene will eine Banane schälen.
1. [_____]_{focus} broad
 2. [_____]_{focus} narrow
 3. [_____]_{focus} contrastive

lit.: *Marlene wants a banana to-peel*

2.2. Speakers and recordings

Six native speakers of standard German (three female, three male) took part in the experiment. They were aged between 20 and 28 and originated from the low Franconian speech area near to Düsseldorf.

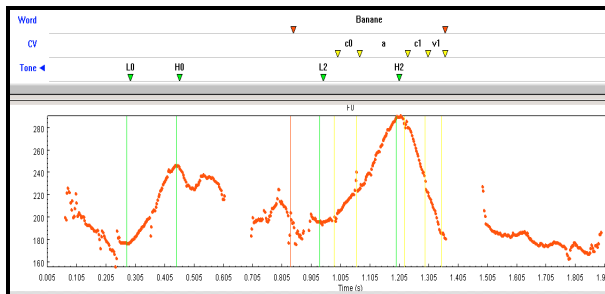
The subjects were asked to read aloud the answers to the questions, which were presented to them both

visually and auditorily (pre-recorded sound files read by the instructor), in a contextually appropriate manner. After a test block of five question-answer pairs each subject read out the target sentences (eight sentences, three focus conditions) three times in pseudo-randomised orders, leading to 72 tokens per speaker.

2.3. Analysis

Segmental and tonal labels were assigned using the speech database system EMU. Labels included the target word, phoneme-sized segments, and the L and H tones associated with pitch accents. An example screen shot is given in Fig.1. In addition, we extracted formant frequencies of the first two formants (F1 and F2) in the steady state of all target vowels.

Figure 1: Target sentence *Marlene will eine Banane schälen* (contrastive focus, speaker J) with labels and F0 trace in EMU.



3. RESULTS AND DISCUSSION

3.1. Tonal marking of focus

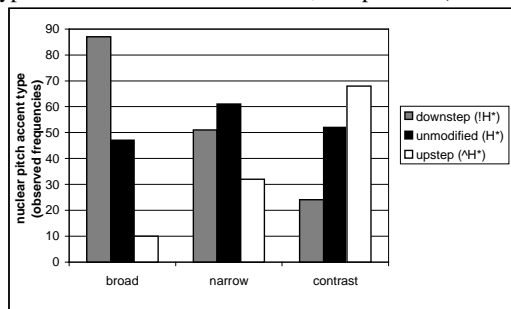
Is there a prenuclear accent on the subject of the sentence? Across all speakers, we found more prenuclear accents in broad (98%) and narrow focus (94%) than in contrastive focus (86%). While the results for broad focus as well as the general decrease of prenuclear accents from broad to contrastive focus were expected, the high percentage of prenuclear accents in narrow and contrastive focus contexts come as a surprise, since the sentence topics carrying the accent were already given in the question. Obviously, an accent on the subject argument is highly preferred for structural or rhythmic reasons. In some cases, we even observe a minor phrase break after the subject noun, strongly favouring accentuation of the subject. Nevertheless, two speakers (C,O) use fewer prenuclear accents than the others, especially for indicating contrastive focus.

Is the scaling of the nuclear accent peak lower, higher or the same in relation to the prenuclear peak, if present? We calculated speaker-specific upstep and downstep factors (see [2]), allowing us to ‘translate’ the observed scaling differences (in semitones)

between the prenuclear peaks (H0) and the nuclear peaks (H2) into the phonological categories !H* (downstep), ^H* (upstep) and H* (neither downstep nor upstep).

We found more downsteps in broad focus and more upsteps in contrastive focus (across all speakers), as shown in Fig.2 (highly significant differences in choice of accent type: $\chi^2(4) = 85.651$; $p < 0.001$). This result is in line with our assumption that a higher accent (relative to a preceding one) sounds more prominent than a lower accent and is thus more appropriate for marking a semantic contrast.

Figure 2: Observed frequencies of nuclear pitch accent type in relation to focus structure, all speakers (N=432).



Speakers vary as to their preferred choice of contours, though. One speaker (O) has a strong preference for (non-downstepped) peak accents in all conditions while another (B) favours downstepped nuclear accents in all three focus structures. Yet another speaker (C) restricts his usage of downsteps to broad focus.

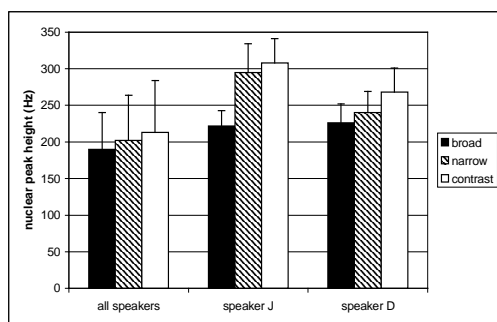
We observe alternation between two different but equally frequent downstepping contours: a) a plain downstepped accent (!H*) and b) an early peak accent with the H tone on the pretonic syllable (H+!H*). The distribution of both types of contour amongst the three focus structures investigated is very similar across all speakers (plain downstep: broad 53%, narrow 26%, contrast 21%; early peak: broad 55%, narrow 37%, contrast 8%). However, single speakers show considerable differences as to their preferred type of downstep: speaker J, e.g., exclusively uses early peaks (47% of all utterances), whereas speaker B does not use a single early peak contour, although 65% of his utterances include a downstep. Two speakers (O,A) unexpectedly use early peaks in particular for marking narrow focus.

What is the actual timing and height of the nuclear accent peak? In terms of nuclear peak timing (in relation to the onset of the accented syllable) our data for all speakers show a significantly later peak in contrastive than in non-contrastive focus (narrow~contrast $p < 0.01$; broad~contrast $p < 0.001$; ANOVA: $F(2) = 10.993$; $p < 0.001$). The same order, namely an increasing delay from sentence type 1 (broad) to 3 (contrast), can be found in five out of six

speakers, with contrast showing the latest peaks for all six speakers. Significance is reached in the data of speakers C (broad~narrow $p=0.001$; broad~contrast $p<0.001$) and A (narrow~contrast $p<0.05$; broad~contrast $p<0.05$), underlining the relevance of nuclear peak timing as one strategy for marking degrees of prominence indicating different focus structures.

The investigation of absolute pitch height of the nuclear accent reveals a highly significant difference between all three focus structures across all speakers (ANOVA: $F(2) = 27.012$; $p<0.01$; separate subgroups in a Scheffé posthoc test). As expected, the mean Hz values are lowest in broad focus and are increasing over narrow to contrastive focus, thus enhancing the nuclear accent's prominence. This increase can be found in three out of the six speakers. Interestingly, the statistically significant step between the subgroups varies. Speaker J, for instance, makes a clear distinction between the two *sizes of focus domain* (i.e. broad focus on the one hand and narrow/contrastive on the other; $p<0.001$), while speaker D more clearly differentiates between the two *types of focus* (i.e. non-contrastive and contrastive focus; $p<0.01$)(Fig.3).

Figure 3: Differences in pitch height of nuclear accent peak (mean and s.d.) in relation to focus structure.



What is the degree of pitch excursion in rising nuclear accents, if present? It has been claimed [6] that pitch excursion is a more relevant cue to (perceived) prominence than absolute pitch height. We find significantly more rising nuclear accents in contrastive focus (74%) than in narrow (50%) and broad focus (42%)(ANOVA: $F(2) = 5.071$; $p<0.01$). Also, the pitch excursion (calculated as the difference between the nuclear L and H targets) is significantly higher in contrastive focus than in broad focus ($p<0.01$). In general, we observe an increase in the extent of the rise from broad focus through narrow focus to contrast. Like other cues discussed so far, some speakers exploit pitch excursion as a strategy for focus marking, such as D (significant difference between broad and contrastive focus; $p<0.05$), whereas others do not, like speaker B, who makes ample use of rising nuclear accents which have nearly identical values for pitch excursion.

3.2. Durational marking of focus

We examined the duration of all nuclear syllables and feet. In bisyllabic target words, the domain 'foot' is coextensive with 'word' (e.g. *Nase* 'nose'). For the durations of both nuclear syllables and feet we observe an increase in the mean values from broad to narrow to contrastive focus across all speakers. Significant differences can be found between broad focus and contrast for both domains, with increased significance for nuclear feet ($F(2) = 6.028$; $p<0.01$) compared to nuclear syllables ($F(2) = 4.063$; $p<0.05$). The same order of durational differences (broad < narrow < contrastive) for feet and syllables is displayed in the data of three speakers, and broad focus is marked by the smallest durations by five speakers.

One speaker (C) uses significantly longer nuclear syllables ($p<0.05$), and two speakers (C,B) produce longer nuclear feet in contrastive versus broad focus (C: $p<0.01$, B: $p<0.05$). This result suggests that the (nuclear) foot could be a more relevant domain for marking different focus structures than the syllable.

3.3. Marking of focus on accented vowels

The formant charts in Fig.4 show mean values in Bark for each vowel quality separately. For each speaker we used their initial (B,C,O; all male) to plot means for the vowels in different focus conditions: broad (small letters), narrow (CAPITALS) and contrastive (**BOLD**). We observed the following tendencies, indicated by the arrows in the charts: vowels were more peripheral in narrow and contrastive focus conditions than in broad focus, both for F1 (e.g. lowering for /a:/ in speaker C, and raising for /u:/ in speakers B,C,O), and for F2 (e.g. fronting for /i:/ in speakers B,C, and backing for /o:/ in speakers B,C,O). However, there was no systematic difference between narrow and contrastive focus (e.g. /u:/ is more raised in contrastive than in narrow focus for speaker B, but not for speaker O).

Fig.5 shows a schematised vowel quadrilateral with the four target vowels marked. The arrows indicate the peripheral vowel articulation (which, we stress here, was measured in the acoustic dimension) in narrow and contrastive focus. We found clear tendencies for peripheral articulation in narrow and contrastive focus in all vowels: /u:/ is more raised and retracted (speakers B,C,O), /o:/ is more raised and/or retracted (speakers B,C,O), /i:/ is more raised and/or fronted (speakers B,C) and /a:/ is more lowered and fronted (but only for speaker C). We found similar hyperarticulation strategies for the females (speaker J /o:/, a:/, speaker D /o:/, u:/, and speaker A /o:/).

Figure 4: Effect of focus structures on F1 and F2 in the acoustic vowel space (three male speakers: B,C,O). Formant values in Bark.

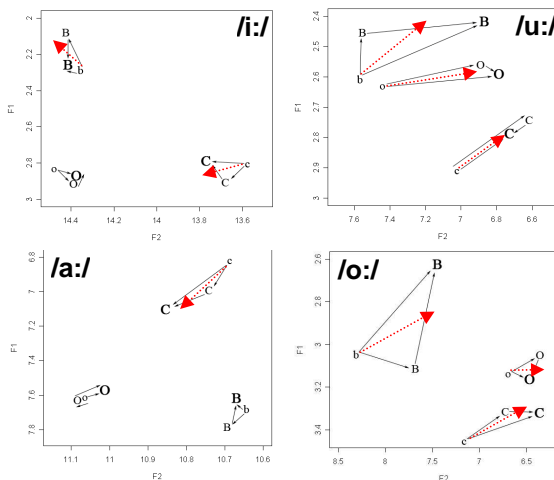
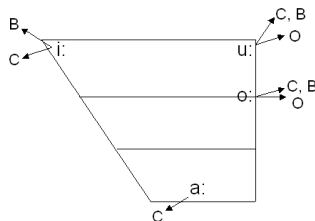


Figure 5: Vowel quadrilateral for vowel quality changes from broad to narrow/contrastive focus for /i:/, a:, o:, u:/ for three speakers (B,C,O).



We thus show that vowel articulation plays a role in marking different focus structures. To enhance focus prominence, speakers produce vowels more peripherally (localised hyperarticulation [4]). Systematic differences were found for broad vs. narrow focus and broad vs. contrastive focus, but not between narrow and contrastive focus.

It has been argued [4] that whereas hyperarticulation is related to tongue kinematics, and can therefore be expressed by F1/F2 vowel space expansion, this is less the case for sonority expansion, which is related to jaw and lip kinematics. We plan to include kinematic measurements in the next study, so as to explore the role of sonority expansion in focus marking.

4. CONCLUSIONS

Across all speakers, we observed a general increase in prominence-lending parameters in the tonal and durational domains from broad to narrow to contrastive focus. In the articulatory domain, we found an increase in localised hyperarticulation from broad to narrow focus, but not between narrow and contrastive focus. At the same time, the use of particular strategies is highly speaker-dependent (Table 1). For example, speaker D makes use of several tonal cues

such as higher and steeper nuclear accents, whereas speaker B prefers longer durations for signalling contrast. All speakers employ hyperarticulation, but not for all vowels in the same way.

Table 1: Single speakers' significantly preferred strategies for marking contrastive focus (over broad focus). Brackets indicate tendencies for tonal and durational strategies, and, for hyperarticulation, restriction to single vowels.

	J	D	C	O	B	A
no prenuclear accent			(+)	(+)		
upstep	+	+	+	(+)		(+)
higher peak	+	+				(+)
later peak	(+)	(+)	+		(+)	+
steeper rise		+	(+)			+
longer nuclear foot		(+)	+		+	(+)
hyperarticulation	(+)	(+)	+	(+)	(+)	(+)

As we can see from the table, both segmental and suprasegmental cues can be used to achieve the same interaction goal, i.e. to signal information structure. Along the lines of [11:117], we can thus identify a *functional cluster*, which is composed of cues from different phonetic and phonological systems.

5. REFERENCES

- [1] Baumann, S., Grice, M., Steindamm, S. 2006. Prosodic Marking of Focus Domains - Categorical or Gradient? *Proc. Speech Prosody 2006* Dresden, 301-304.
- [2] Berg, R. van den, Gussenhoven, C., Rietveld, T. 1992. Downstep in Dutch: Implications for a model. In: Docherty, G.J., Ladd, D.R. (eds), *Papers in laboratory phonology II*. Cambridge: CUP, 335-359.
- [3] Braun, B., Ladd, D.R. 2003. Prosodic correlates of contrastive and non-contrastive themes in German. *Proc. 8th European Conference on Speech Communication and Technology* Geneva, 789-792.
- [4] Cho, T. 2005. Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a, i/ in English. In: *JASA 117(6)*, 3867-3878.
- [5] Dohen M., Lævenbruck H., Hill H. 2006. Visual Correlates of Prosodic Contrastive Focus in French: Description and Inter-Speaker Variabilities, *Proc. Speech Prosody 2006* Dresden, 221-224.
- [6] Gussenhoven, C. 2004. *The Phonology of Tone and Intonation*. Cambridge: CUP.
- [7] Harrington, J., Fletcher, J., Beckman, M.E. 2000. Manner and place conflicts in the articulation of accent in Australian English. In: Broe, M., Pierrehumbert, J. (eds), *Papers in laboratory phonology V*. CUP, 40-51.
- [8] Ishihara, S., Féry, C. 2006. Phonetic Correlates of Second Occurrence Focus. *Proc. 36th North Eastern Linguistic Society (NESL36)*.
- [9] Ladd, D.R. 1996. *Intonational Phonology*. Cambridge: CUP.
- [10] Ladd, D.R., Morton, R. 1997. The Perception of Intonational Emphasis: Continuous or Categorical? *Journal of Phonetics* 25, 313-342.
- [11] Local, J. 2003. Phonetics and talk-in-interaction. *Proc. 15th ICPhS* Barcelona, 115-118.