

Hocus Focus – How prosodic profiles of contrastive focus emerge and change in different elicitation contexts

Karin Görs, Oliver Niebuhr

Department of General Linguistics, ISFAS, Christian-Albrecht-University of Kiel, Germany

kgoers@arcor.de, niebuhr@isfas.uni-kiel.de

Abstract

The results of a production experiment with 8 native speakers of German show that the acoustic profiles of contrastive focus differ between read monologues, read dialogues and spontaneous dialogues. The differences suggest that contrastive focus is only a matter of F₀, and that two F₀-related subtypes, comparative and corrective focus, may be distinguished. The additional changes in duration and intensity that were found for corrective focus are not due to establishing an information-structure contrast, but to signaling emphatic reinforcement.

Index Terms: focus, information structure, emphasis, German.

1. Introduction

Particularly in spoken language the signalling of information structure (as defined by [1]) is an important component of a successful communication. In fact, hardly any of those languages and dialects that were touched on in modern phonetic and linguistic research was not investigated with regard to the realization of broad vs narrow focus, including minority and endangered varieties like Irish Gaelic [2], Tibetan [3], Porteño [4], Chickasaw [5], Basque [6], and Yucatec Maya [7]. When narrow focus occurs in a context in which a speaker differentiates a piece of information from a previous one, the resulting subtype of narrow focus can be referred to as contrastive focus (in line with [1], we assume here that contrastive and non-contrastive focus are to be distinguished).

Many languages have morphological or syntactic means to distinguish between broad and contrastive focus, cf. [8,9] for French and Bantu languages. However, in all languages, as it seems, the distinction is also made on a prosodic basis. Particularly in the last decade a large number of studies scrutinized the prosodic patterns of broad and contrastive focus with regard to phonatory and articulatory aspects.

Cross-linguistic meta-analyses of these studies yield at first glance a clear picture of the prosodic differences between broad and contrastive focus. Relative to broad focus, contrastive focus is characterized by longer intonation rises (which are analyzed in terms of different phonological pitch-accent categories), greater intensity levels, and by lengthening and hyperarticulating those syllables to which the focus is linked.

On closer inspection, however, quite a lot of variation emerges; and it is too simplistic to dismiss this variation as originating from language-specific frameworks, although these frameworks are surely part of the explanation. Crucially, language-specific frameworks do not account for the fact that a major part of the variation is caused by considerable inter- and intra-individual differences that were found within each language. For example, in German, which is the subject of the present paper, the studies of [10,11,12,13] showed that some speakers realized the longer intonation rises under contrastive focus in terms of F₀-peak alignment differences, whereas for other speakers ‘longer’ primarily meant higher and steeper. Nevertheless, intonational changes are at least consistently involved in distinguishing broad and contrastive focus. The same is not true for intensity changes, lengthening and local

hyperarticulation. These three focus constituents can be absent or present; and in the latter case they vary in the degree to which they are exploited by speakers. Moreover, they can concern different temporal domains like the accented vowel, the accented syllable, the foot, or the pre-accented syllable.

Given these inconsistencies in duration, intensity and articulatory effort, could it be that focus marking is in the end essentially a matter of intonation? If so, what causes the additional occurrences of enhanced intensity changes, lengthening and local hyperarticulation under contrastive focus? And, can the same cause also be responsible for the observed variation in the intonation patterns of contrastive focus?

In developing a system of forms and functions of emphatic accentuation in German, [14] found a type of accent whose function is to reinforce the truth value (the trustworthiness) of the accented information. A typical example in which this accent type occurs is the following. A man stated in a recent TV model show “*Das ist eine sehr schöne Jacke*” (this is a very beautiful coat) with reinforcement on “*sehr*” (very) in order to convince his wife to finish her extended shopping tour.

The prosodic profile of the reinforcement type of emphasis is shown in Figure 1. As can be seen, reinforcement is created by lengthening the accented syllable (especially the syllable-initial consonant), by increasing height and sharpness of the intonation peak (particularly by means of a steeper rise) and by hyperarticulating the contrast between the accented vowel and its adjacent consonants, which yields at the same time abrupt intensity changes that reach a higher level on the accented vowel. Moreover, reinforcement is introduced by a “ritardando”, i.e. by an articulatory deceleration that affects the pre-accented syllable (audio examples, cf. www.isfas.uni-kiel.de).

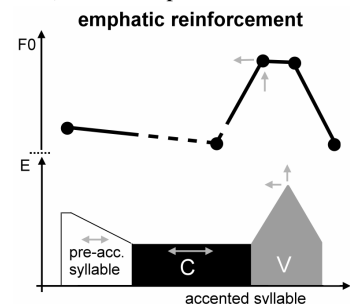


Figure 1: Prosodic profile of reinforcement in German, cf. [6]

The prosodic profile of reinforcement is similar to that of contrastive focus; and even the function of reinforcement is compatible with that contrastive focus, albeit reinforcement does not require expressing a contrast and is hence applicable to more communicative contexts. Against this backdrop, one may wonder whether the prosodic profiles of German contrastive focus that were set up in studies like [10,11,12,13] are actually bipartite and result from *contrasting and reinforcing* the corresponding information. Under these circumstances, the variation that was found in the realizations of contrastive focus may be explained by the presence, absence or the degree of additional reinforcement. If reinforcement is the source of variation, then it is likely that the variation is facilitated by the

fact that a number of elicitation procedures were used across studies. Some studies asked the subjects to read text passages; others involved the subjects in more or less spontaneous dialogues with a machine, a virtual listener, or a real dialogue partner. Each procedure is claimed to elicit contrastive focus, although they create very different communicative contexts.

These elicitation differences and their different potentials to encourage reinforcement marked the point of departure of our present study. We performed a production experiment against the backdrop that reinforcement is used in order to convince a dialogue partner of the speaker's opinion. We elicited focussed target words in three different contexts, i.e. read monologues, read dialogues, and spontaneous dialogues. In addition to broad focus we included two potentially different types of contrastive focus. In the first type the speaker compares two pieces of information with each other, like in the following example (target word underlined): "*Das war vor der Katastrophe in Japan so und ist nach dem Unglück noch verstärkt worden*" (It was like this before the disaster in Japan; and after the accident it was urged even more). We will refer to this first type as comparative focus. The second type is the prototypical type of contrastive focus in which a speaker corrects the dialogue partner: "*Er ist nicht nach links, sondern nach rechts abgebogen*" (He did not turn left, he turned right). This second type will be called corrective focus.

On this basis we hypothesized that, if there is an influence of reinforcement on the focussed target words, this influence will become stronger the more the communicative situation requires arguing with a dialogue partner. More specifically, we expected that reinforcement will not occur in read monologues as well as under broad and comparative focus, whereas corrective focus will show additional reinforcement, more clearly so in the spontaneous than in the read dialogues.

2. Method

We pursued our hypothesis by means of a multi-parametric acoustic analysis. The corresponding production experiment was based on 7 monosyllabic target words. Each of them had a CVC(CC) structure and specified temporal or spatial properties of the syntactically associated noun or verb. Three target words had short-vowel nuclei: "*rechts*", "*links*", "*weg*" (left, right, away). In the other four target words the nucleus was a long vowel or a diphthong: "*vor*", "*nach*", "*rein*", "*raus*" (before/in front of, after/behind, into, out of). The group of subjects included three male and five female native speakers of German, aged 20-30 years. Compared with the heterogeneous methods of previous production experiments on focus and the related limitations of previous meta-analyses, the crucial advantage of our method is that the same group of 8 speakers produced the same set of 7 target words in 3 different and increasingly interactional elicitation contexts.

Read monologues: In the monologue context the target words were embedded in two short texts that were to be read by the subjects in a newsreader style. One text dealt with the recent turnaround in the German nuclear policy. Its semantic-pragmatic make-up was designed to elicit the target words with comparative focus. The other text summarized different contradictory witness statements on a series of famous robberies in Kiel. It served to elicit corrective focus on the target words. Broad-focus productions of the target words were induced by both texts. Four subjects started with reading the robbery text; the other four subjects read the text on the German nuclear policy first. All subjects were given 15 min at the beginning of the recording session in order to familiarize themselves with the texts and the recording situation.

Read dialogues: 63 short A-B dialogues were created in which the target words were elicited with broad, comparative,

and contrastive focus in B's response to A. In the comparative focus condition A asked a question and B answered in such a way that s/he reached a different conclusion, compared to what was suggested by A. For example, A: "*Treffen wir uns vor dem Film?*" – B: "*Wir treffen uns nach dem Film.*" (A: Do we meet before the movie? – B: We meet after the movie.). The corrective focus condition was similar to that of comparative focus, except that B's answer explicitly corrects a categorical, but wrong statement of A. For example, A: "*Wir müssen hier links abbiegen.*" – B: "*Nein, hier müssen wir rechts abbiegen.*" (A: We must turn left here. – B: No, we must turn right). The subjects read B's parts in the dialogues, whereas A was always the experimenter (KG). The recording session for each subject started with the verbal instruction to read B's responses in a natural everyday fashion. The instruction was followed 10 A-B dummy dialogues that served to practise the reading procedure and to familiarize the subjects with the recording situation. The 63 A-B dialogues were produced in each session in a differently randomized order. The subject and the experimenter sat next to each other, which precluded direct eye contact.

Spontaneous dialogues: The spontaneous dialogues evolved in the established 'maptask' procedure, cf. [15]. In individual recording sessions each subject was to explain the plotted route to the experimenter (ON) so that he could draw the described route on his map. Due to this navigation task broad focus productions of the target words were elicited automatically. Comparative focus productions were evoked by means of the spatial arrangement of the landmarks on the map and the resulting route along or through them. In order to provoke corrective focus productions of the target words, the experimenter pretended misunderstandings. Subject and experimenter sat face-to-face to each other at separate tables.

The order of elicitation conditions was permuted across the 8 subjects. All recordings were made in a sound-treated booth at the University of Kiel with a 96 kHz sampling rate and a 24-bit quantization. Each target word in the recordings was acoustically analyzed with PRAAT, using the default settings. The durational measurements (in ms) concerned the pre-accented and accented syllables, and within the accented syllables the onset, vowel, and coda durations. As for F0, we determined the F0 minimum and maximum of the pitch-accent rise (in st) as well as the point in time of the rise offset (in ms). The intensity measurements included the peak levels in the pre-accented and accented syllables (in dB) as well as their points in time. Based on these multi-parametric measurements, the following 7 variables were calculated with reference to the acoustic profile of reinforcement (cf. [14], Fig.1) and the state-of-the-art on German broad vs contrastive focus, cf. [12].

- (a) F0-peak maximum alignment after the accented-vowel onset,
- (b) F0 range of the pitch-accent rise,
- (c) Duration of the pitch-accent rise,
- (d) Intensity-peak alignment after vowel onset,
- (e) Increase in intensity from the peak level of the pre-accented to the peak level of the accented syllable,
- (f) Duration ratio of pre-accented to accented syllable,
- (g) Duration ratio of onset to onset plus nucleus (i.e. C/C+V) in the accented syllable.

As the three elicitation contexts yielded different numbers of target words, we reduced the number of target words to 42 tokens per subject and elicitation context for the purposes of the statistical analyses. The tokens for these sub-samples were selected at random. We further created separate sub-samples for the long-vowel and short-vowel target words.

3. Results

The data of one of our female speakers had to be excluded from the analysis. For the remaining 7 speakers the acoustic measurements of the long-vowel and short-vowel target words were analyzed in separate repeated-measures MANOVAs, each with the fixed factors focus (broad vs comparative vs corrective) and elicitation context (read monologues vs read dialogues vs spontaneous dialogues). The dependent variables were the 7 calculated parameters (a)-(g).

The results of the long-vowel and short-vowel MANOVAs were overall similar. In the long-vowel MANOVA the two fixed factors focus type and elicitation context became highly significant ($F[14,154]=113.028, p<0.001, \eta_p^2=0.877$; $F[14,154]=33.966, p<0.001, \eta_p^2=0.693$); and there was a relatively weak, but significant interaction between them ($F[28,644]=26.780, p<0.001, \eta_p^2=0.396$). Pairwise post-hoc comparisons (with Bonferroni corrections) that were performed between the 3x3 factor levels of each dependent variable provided in-depth insights into the origins of the global significances. All differences in the descriptions below refer to mean values and are based on $p<0.01$; the lack of differences means at least $p>0.1$.

The elicitation contexts had qualitative and quantitative effects on the acoustic profiles of the three focus conditions. In the read monologues, the three focus conditions only differed in terms of F0 parameters. Compared with broad focus productions, both comparative and corrective focus productions showed 30-80 ms longer and about 1.5-2.0 semitones higher pitch-accent rises that peaked 20-50 ms later relative to the accented-vowel onset (which means in absolute terms that the F0 maximum was reached close to or clearly after the accented-vowel offset). These F0 differences in relation to broad focus were significantly more pronounced for the corrective than for the comparative focus productions.

While in the read monologues comparative and corrective focus differed from broad focus only quantitatively – i.e. in the degree of increase in rise duration, alignment and scaling – the difference became a qualitative one in the two dialogue contexts. Firstly, the acoustic differences between corrective focus on the one hand and comparative and broad focus on the other spread from the F0 parameters to the duration and intensity parameters; secondly, the F0 differences between comparative and corrective focus developed in opposite directions. In the case of comparative focus, changing the elicitation context from read monologues to read dialogues and from read dialogues to spontaneous dialogues only led to successively longer and later aligned pitch-accent rises. Relative to broad focus, whose acoustic profile remained statistically constant across all three elicitation contexts, the pitch-accent rise of comparative focus was still 60 ms longer and ended still 20-50 ms further away from the accented-vowel onset than in the read monologues. The range of the pitch-accent rise did not increase. It remained on average about 1.5 semitones higher under comparative than under broad focus in all three elicitation contexts.

In contrast, from the read monologues to the read dialogues and further to the spontaneous dialogues the pitch-accent rises of the corrective focus productions became successively *shorter* and higher (and hence *steeper*), and they peaked successively *earlier* relative to the accented-vowel onset. So, in terms of F0 the corrective focus productions in the two dialogue contexts diverged from the comparative focus productions and turned back to the broad focus productions, except for the F0 range of the rise which was 5-6 semitones higher under corrective than under broad focus. For example, in the spontaneous dialogues the pitch-accent rises of corrective focus were on average 45 ms shorter than those of

comparative focus, but still 35 ms longer than those of broad focus. The pitch-accent peaks of corrective focus were aligned on average almost 40 ms earlier after the accented-vowel onset than those of comparative focus and were in this respect statistically identical to the pitch-accent peaks of broad focus.

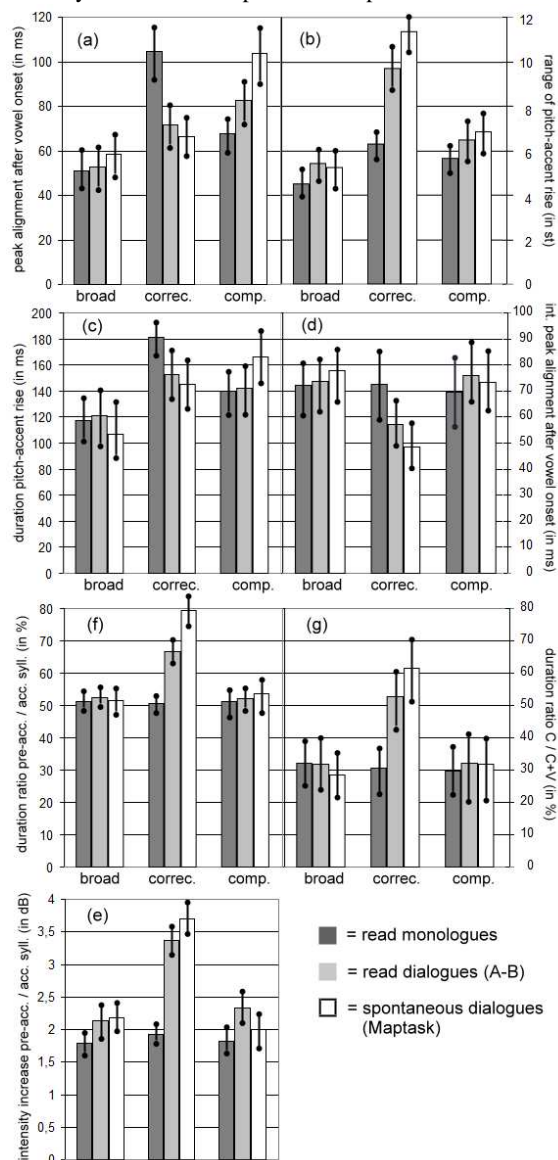


Figure 2: Mean values for the 7 dependent variables (a)-(g) of the short-vowel target words produced in read monologues (dark gray), read dialogues (light gray) and spontaneous dialogues (white bars). Each bar $n=42$.

Unlike for broad and comparative focus, the step from monologues to dialogues caused additional significant changes in the *duration and intensity patterns* of the corrective focus productions. The increase in intensity from the pre-accented to the accented syllable grew by 2.0-2.5 dB, and the intensity maximum in the accented syllable was reached 20-25 ms earlier after the accented vowel than in the broad and comparative focus productions. The duration ratio of pre-accented to accented syllable doubled from 35 % to 70 % due to an increase in the duration of the pre-accented syllable. Likewise, the duration of the consonant in the accented syllable increased relative to the accented vowel so that the ratio of C to C+V grew by 15-25 %.

The results of the short-vowel target words differed only quantitatively from those of the long-vowel target words. That

is, the relative differences described above were similarly found for the short-vowel target words, but at higher or lower parameter levels due to the reduced vowel durations and their mediate effects on duration ratios, intensity changes and F0-peak alignments. Correspondingly, the repeated-measures MANOVA of the short-vowel target words showed significant main effects of focus ($F[14,154]= 63.857, p<0.001, \eta_p^2= 0.853$) and elicitation context ($F[14,154]= 14.827, p<0.001, \eta_p^2= 0.574$) as well as a significant interaction between the two fixed factors ($F[28,644]= 11.093, p<0.001, \eta_p^2= 0.325$). In terms of partial eta-squared, focus was again most powerful in explaining the variance of the measurements. Figures 2(a)-(g) summarize the results of the short-vowel target words in terms of the means received for the 7 dependent variables.

4. Discussion and Conclusion

The statistical analyses of our acoustic measurements yielded clearly significant main effects of focus and elicitation context. Post-hoc tests showed further that the significant interaction of the main effects had several reasons. The most important reasons were that the differences between broad focus on the one hand and comparative and corrective focus on the other were greater between the monologues and the dialogues than between the two dialogue contexts. Moreover, the phonetic profiles of comparative and corrective focus developed in diametrically opposed directions in the dialogue contexts.

In the read monologues the profile of corrective focus was distinguished from the profiles of broad focus and comparative focus merely in terms of F0. However, in the dialogue contexts, the corrective focus productions showed additional duration and intensity changes that were even stronger in the spontaneous than in the read dialogues. Among others, these changes included lengthening the pre-accented syllable as well as the subsequent onset consonant of the accented syllable. The resulting deceleration transitioned into an accelerated and higher intensity increase in the accented vowel. Likewise, the F0 peak on the accented syllable became higher and sharper and was reached earlier, whereas the peak of the comparative focus productions showed a longer, shallower and hence later aligned rise.

Thus we found typical characteristics of the emphasis type of reinforcement, cf. Figure 1, [14]; and we found them in those contexts in which the communicative situation – i.e. presence of a dialogue partner, need to correct her/him – required the speakers to reinforce the truth value (i.e. the trustworthiness) of their focussed information. In summary, the phonetic approximation of the reinforcement profile and the context-specific occurrence of this approximation clearly support our hypothesis. That is, in the dialogues our corrective focus productions were hybrid forms of contrastive focus and reinforcement. Our data suggest additionally that, when the reinforcement component is subtracted, signaling contrastive focus is a matter of F0 alone. Even though this claim must be cross-validated in perception experiments (which also include functional judgments), we think that our findings give sufficient rise to revisit the results of all those studies in which the supposed contrastive focus was tied to duration and intensity changes and/ or to local hyperarticulation, particularly if these changes extended beyond the accented syllable.

Over and above the issue of superimposed reinforcement, our study raised the question whether comparative and corrective focus represent two different subtypes of contrastive focus. In order to provide an initial answer to this question, we must concentrate on the read monologues, as we may assume that the target words in this context were consistently produced without reinforcement. Under these circumstances, the F0 rises of the pitch accents of corrective focus were on aver-

age about 40-50 ms longer than those of comparative focus. Together with the measurements for the rise-offset alignment, we can conclude that these 40-50 ms manifested themselves primarily in a delayed peak maximum. Crucially, this delay pushed the peak maximum beyond the accented syllable into the subsequent unaccented syllable (in both the long-vowel and the short-vowel conditions), while the rise onset remained more or less constantly aligned at about the accented-syllable onset. On this basis, we think that the differences in the durations and alignments of the pitch-accent rises would be associated with two different pitch-accent categories. Comparative focus was produced with L+H*, whereas the corrective focus productions showed L*+H accents. In previous studies, such phonological distinctions were exploited as an argument in favour of different focus types, cf. [10,16]. Following this line of argument, we suggest that comparative and corrective focus also represent two different subtypes of contrastive focus.

Apart from scrutinizing this suggestion, future studies must investigate focus multi-parametrically and with a well-conceived elicitation task that pays more attention to communicative functions. In this connection, it is also necessary to spend more effort on the cross-linguistic investigation of reinforcement and other types of emphasis.

5. References

- [1] Halliday, M., "Notes on Transitivity and Theme in English, Part 2", *Journal of Linguistics*, 3: 199-244, 1967.
- [2] Dorn, A. and Ní Chasaide, A., "Effects of focus on f0 and duration in Irish (Gaelic) declaratives", *Proc. 12th Interspeech Conference, Florence, Italy: 965-968, 2011.*
- [3] Wang, B., Wang, L., and Qadir, T., "Prosodic realization of focus in six languages/dialects in China", *Proc. 17th ICPhS, Hong Kong, China: 144-147, 2011.*
- [4] Feldhausen, I., Pešková, A., Kireva, E., and Gabriel, C., "Categorical perception of Porteoñol nuclear accents", *Proc. 17th ICPhS, Hong Kong, China: 128-131, 2011.*
- [5] Gordon, M., "The intonational realization of contrastive focus in Chickasaw", in C. Lee and M. Gordon [Eds], *Topic and focus: Cross-linguistic perspectives on meaning and intonation*, 65-78, Kluwer, 2004.
- [6] Elordieta, G., "A constraint-based analysis of the intonational realization of focus in Northern Bizkaian Basque", in T. Riad and C. Gussenhoven, *Tones & tunes, Vol. 1, Typological studies in word and sentence prosody*, 199-232, deGruyter, 2007.
- [7] Kügler, F. and Skopeteas, S., "On the universality of prosodic reflexes of contrast: The case of Yucatec Maya", *Proc. 16th ICPhS, Saarbrücken, Germany: 1025-1028, 2011.*
- [8] Dohen, M., and Loevenbruck, H., "Pre-focal rephrasing, focal enhancement and post-focal deaccentuation in French", *Proc. 8th ICSLP, Jeju Island, Korea: 313-316, 2004.*
- [9] Zerbian, S., "Intensity in narrow focus across varieties of South African English", *Proc. 17th ICPhS, Hong Kong, China: 2268-2271, 2011.*
- [10] Baumann, S., Grice, M., and Steindamm, S., "Prosodic marking of focus domains - categorical or gradient?" *Proc. 3rd Speech Prosody, Dresden, Germany: 301-304, 2006.*
- [11] Braun, B. and Ladd, D.R., "Prosodic correlates of contrastive and non-contrastive themes in German", *Proc. 8th Eurospeech Conference, Geneva, Switzerland: 789-792, 2003.*
- [12] Hermes, A., Becker, J., Mücke, D., Baumann, S., and Grice, M., "Articulatory gestures and focus marking in German", *Proc. 4th Speech Prosody, Campinas, Brazil: 457-460, 2008.*
- [13] Kügler, F., "The role of duration as a phonetic correlate of focus", *Proc. 4th Speech Prosody, Campinas, Br: 591-594, 2008.*
- [14] Niebuhr, O., "On the phonetics of intensifying emphasis in German", *Phonetica*, 67: 170-198, 2010.
- [15] Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S. and Weinert, R., "The HCRC Map Task Corpus", *Language and Speech*, 34: 351-366, 1991.
- [16] Kohler, K.J., "What is emphasis and how is it coded?", *Proc. 3rd Speech Prosody, Dresden, Germany: 748-751, 2006.*