

## Prosodic Effects of Focus in Dutch Declaratives

Judith Hanssen, Jörg Peters and Carlos Gussenhoven

Department of Linguistics  
Radboud University Nijmegen

{j.hanssen; j.peters; c.gussenhoven}@let.ru.nl

### Abstract

Seventeen speakers of Standard Dutch participated in a production experiment that investigated the effects of focus condition (broad, narrow, corrective focus) on the prosodic realization of nuclear accented words in declarative sentences. It was found that focus has an effect on the duration of onset and coda of the nuclear accented syllable as well as on the timing, scaling, and slope of  $f_0$  targets and movements. A significant effect exists mainly between broad focus accents on the one hand, and narrow and corrective focus accents on the other. The results are interpreted in terms of phonetic variation, but such that phonetic differences between these two conditions are to a large extent conventionalized.

### 1. Introduction

From a prosodic perspective, the concept of focus can be subdivided into several categories on two levels. On one level, a distinction can be made between broad and narrow focus, a difference in the size of the focus constituent [1]. In the broad focus sentence (1a), all information is new, and the focus constituent covers the entire sentence. In a narrow focus reading (1b), only part contains new information, the focus constituent *London*. On the second level, focus is divided into types of focus. One distinction that has been made is that between 'informational' and 'corrective' focus (cf. [2], and references therein). In both types, the size of the focus constituent can vary, but corrective focus is typically narrow. In narrow informational focus (1b), the focused element is a piece of new information that was requested by the hearer. In narrow corrective focus (1c), the hearer has presented a choice from some set of alternatives, which is rejected by the speaker. The focus constituent is indicated by square brackets.

- |                              |                             |     |
|------------------------------|-----------------------------|-----|
| a. Broad                     | <i>What happened?</i>       | (1) |
|                              | [We went to London].        |     |
| b. Narrow<br>(informational) | <i>Where did you go to?</i> |     |
|                              | We went to [London].        |     |
| c. Narrow<br>(corrective)    | <i>Did you go to Paris?</i> |     |
|                              | No, we went to [London].    |     |

Earlier studies report a number of different strategies for expressing semantic differences of the sort shown in (1). Restricting ourselves to prosodic distinctions, speakers can use different pitch accents [3], [4], vary the distribution of (prenuclear and nuclear) pitch accents [3], or use different phrasings [5]. Second, speakers can make differences in the phonetic realization. Strategies that have been reported include variation in peak scaling [3], [6]; variation in pitch range of both the focused and the nonfocused components [6]; differences in segmental duration of the focused element [3], [6], [7]; varying the slope of the pitch accent [6]; and finally, [3] report differences in timing of tonal targets between focus domains of various sizes.

A combination of phonological and phonetic correlates is shown for German by [3], who describe phonological differences between focus conditions in terms of variation in the distribution and frequency of occurrence of pitch accents. From broad to narrow to corrective focus, the authors observe a decrease in the occurrence of prenuclear accents, downstepped nuclear accents (!H\*), and early peak accents (H+!H\*). Phonetically, tonal, durational, and articulatory prominence-lending cues are exploited to distinguish between focus categories. Pitch accent peaks are later and higher, pitch excursion of the rise towards the peak is larger, and durations of the accented foot are longer in corrective focus than in narrow and broad focus, respectively. Articulatory effort in the production of the accented vowel is greater in narrow and corrective than in broad focus. Speakers employ both the phonological and phonetic cues in different ways and to different extents, though.

A study on American English [6] also reports the use of several prosodic prominence cues. An expansion of the pitch range is present on the focused syllable, accompanied by a suppression of the pitch range of syllables in postfocal position. In addition, focused elements show higher  $f_0$  peaks, an increase in the speed of the rise towards the peak, and longer duration of the stressed syllable. A comparison of [6] with a study of British English [7] suggests that the use of prosodic cues to focus varies cross-linguistically. [7] found only one prosodic cue that speakers employ to distinguish between focus categories. Their findings suggest no differences in peak height or peak timing that are directly attributable to focus condition. Speakers only showed increased duration of the element in narrow or corrective focus when compared to broad focus.

Our purpose here is to give a description of the prosodic correlates of broad focus (BF), narrow informational focus (NF), and narrow corrective focus (CF) in Standard Dutch declarative sentences, and classify any differences as either phonological or phonetic. A production experiment with read speech was set up with identical sentences in three focus conditions. We expected a falling nuclear contour, consisting of a rising and a falling pitch movement, to be used as prominence marker on the focused word. The experiment was designed to investigate whether focus condition affects (a) the segmental durations in the nuclear syllable, (b) the scaling of pitch targets in the nuclear contour, (c) the slope of the nuclear rise and fall, and (d) the timing of pitch targets relative to segmental boundaries.

Our hypothesis, based on earlier findings for Germanic languages, was that Dutch shows (from BF to NF to CF) (i) longer segmental durations in the nuclear syllable, (ii) lower scaling of the valleys, and higher scaling of the nuclear peak, (iii) steeper slopes of the rising and falling pitch movements, and (iv) later peaks.

## 2. Method

### 2.1. Materials

We used short declarative sentences of the type *We willen in Manderen blijven wonen* 'We want to stay in Manderen' in three different contexts to elicit one broad, one narrow, and one corrective reading of the test sentence. An example set of one test sentence plus three different context questions is given in (2); focus constituents are indicated by square brackets; the nuclear accented word is underlined.

#### Broad focus (2)

- A. Wat is er met jullie? (*What's the matter?*)  
 B. [We willen in Manderen blijven wonen.]

#### Narrow focus

- A. Waar willen jullie blijven wonen? (*Where do you want to stay?*)  
 B. We willen in [Manderen] blijven wonen.

#### Corrective focus

- A. Willen jullie in Montfort blijven wonen? (*Do you want to stay in Montfort?*)  
 B. Nee, we willen in [Manderen] blijven wonen.

The nuclear pitch accent in all sentences is expected to occur on the target word *Manderen*, a place name. We created three sets of sentences, resulting in nine test sentences. Target words varied in segmental composition, although we used sonorants for the stressed syllable (CVC) where possible. The stress patterns of all target words and test sentences were identical.

### 2.2. Subjects and procedure

We recorded six male and eleven female speakers aged between 18 and 30. All subjects were judged to be speakers of Standard Dutch by the first author; subjects with a marked regional accent were excluded from participation.

The mini-dialogues were presented on cards in pseudo-randomized order, which was reversed for half of the speakers. The test sentences were interspersed with 93 filler sentences from other experiments. Subjects were recorded in pairs; they read each part of the mini-dialogue once. The recordings were made in a sound-treated booth at Radboud University Nijmegen using a portable DAT recorder (TASCAM DA-P1). The data were digitally transferred to a computer and downsampled to a sampling rate of 16 kHz (mono).

### 2.3. Acoustical analysis

Using the *Praat* speech analysis software [8], we inserted the labels listed in (3) below. Labels on the segmental tier were placed through visual and auditory inspection of the waveform and spectrogram; labels on the tone tier were inserted automatically when possible.

#### Labels (3)

- L1 elbow before the nuclear peak  
 H maximum  $f_0$  of the pitch accent  
 L2 elbow after the nuclear peak  
 O1 beginning of nuclear onset  
 V1 beginning of nuclear vowel  
 C1 beginning of nuclear coda  
 O2 beginning of first postnuclear onset  
 V4 beginning of vowel of first postnucl. stressed syllable

Of all labels, timing and  $f_0$  level were computed and saved.  $F_0$  levels were later converted to semitones re 100 Hz. Using the inserted labels, we calculated the variables in (4).

<b>Variables</b>	(4)
Onset duration of accented syllable	O1 to V1
Nucleus duration of accented syllable	V1 to C1
Coda duration of accented syllable	C1 to O2
Height of elbow before nuclear peak	L1 height (st)
Height of nuclear peak	H height (st)
Height of elbow after nuclear peak	L2 height (st)
Rise slope	(H(st) – L1(st))/(H(sec)-L1(sec))
Fall slope	(L2(st) – H(st))/(H(sec)-L2(sec))
Alignment of L1 relative to onset	L1 to O1
Alignment of H relative to nucleus	H to V1
Alignment of L2 relative to postnucl. stress	L2 to V4

## 3. Results

### 3.1. Focus effects on segmental duration

Figure 1 shows mean durations of the onset, nucleus, and coda of the nuclear syllable in broad, narrow, and corrective focus. The nuclear syllable in broad focus is somewhat shorter than the nuclear syllables in narrow and corrective focus. Apparently, this difference can be attributed to a shorter onset and, to a lesser degree, to a shorter coda.

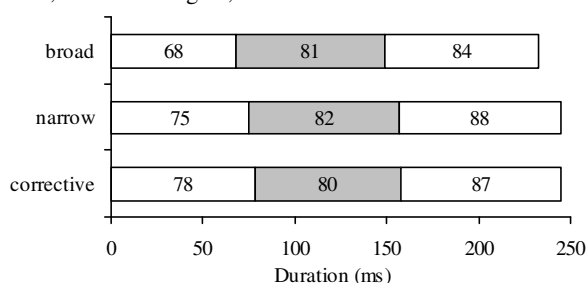


Figure 1: Mean durations (in ms) of onset, nucleus and coda of the nuclear syllable in broad focus, narrow focus and corrective focus (for each segment type,  $N = 143$ ).

To examine the null hypothesis that focus does not affect segmental durations of the nuclear syllable, we carried out three repeated measures ANOVAs using onset duration, nucleus duration, and coda duration as dependent variables and FOCUS (BF, NF, CF) as a fixed within-subjects factor. For onset duration, Mauchly's test indicated that the assumption of sphericity has been violated,  $\chi^2 = 6.27$ ,  $p < .05$ . Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ( $\epsilon = .871$ ). The results show that focus condition affected significantly onset duration ( $F [1.74, 69.65] = 6.591$ ,  $p < .01$ ) and coda duration ( $F [2, 80] = 3.396$ ,  $p < .05$ ), but not nucleus duration ( $F [2, 80] = .257$ ,  $p > .05$ ). Pairwise comparisons revealed that the onset in both NF and CF is longer than in BF ( $p < .05$  and  $p < .001$ , respectively), whereas no significant difference was found between the onset duration of NF and CF. Coda duration was found to differ significantly between BF and NF only ( $p < .05$ ).

### 3.2. Focus effects on $f_0$ contours

Figure 2 shows averaged  $f_0$  contours in the three focus conditions. The contours were created with the help of a Praat script by Yi Xu [9] using 100 measurements per interval. They span the nuclear foot (Xxx) and the first postnuclear foot (Xx). The  $f_0$  contours are aligned with the beginning of the onset of the nuclear syllable.

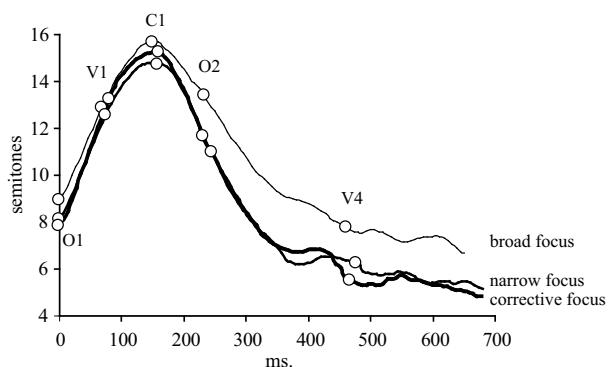


Figure 2: Averaged contours for BF (thin line;  $N = 44$ ), NF (medium line;  $N = 51$ ), and CF (thick line;  $N = 51$ ). Circles mark the beginnings of the nuclear onset (O1), the nuclear vowel (V1), the nuclear coda (C1), the first postnuclear onset (O2), and the first postnuclear stressed vowel (V4).

The  $f_0$  contours in both NF and CF show a steeper fall and reach a lower level than the  $f_0$  contour in BF. On the other hand, the  $f_0$  peak in BF is higher rather than lower than in NF and CF. In addition, we found that the elbow (L2) of all three contours occurs about 150 ms before the beginning of the vowel of the following stressed syllable (V4).

First, we examined the null hypothesis that focus condition does not affect the relative  $f_0$  level of the nuclear contours (see table 1). We carried out four repeated measures ANOVAs using the  $f_0$  levels (in semitones) of L1, H, L2 and V4 as dependent variables and FOCUS (BF, NF, CF) as a fixed within-subjects factor. For H and V4, Mauchly's test indicated that the assumption of sphericity has been violated,  $\chi^2 = 7.32$ ,  $p < .05$  and  $\chi^2 = 7.29$ ,  $p < .05$ , respectively. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ( $\epsilon = .854$  in both cases). The results show that focus condition affected significantly the  $f_0$  level of H ( $F [2, 80] = 4.190$ ,  $p < .05$ ), L2 ( $F [2, 78] = 13.966$ ,  $p < .001$ ), and V4 ( $F [1.71, 68.35] = 14.127$ ,  $p < .001$ ), but not of L1 ( $F [2, 80] = 1.699$ ,  $p > .05$ ). Pairwise comparisons revealed that the  $f_0$  peak (H) in both BF and CF is significantly higher than in NF ( $p < .01$  and  $p < .05$ , respectively), whereas the difference of the  $f_0$  levels in BF and CF was not significant. L2 was found to be significantly higher in BF than in NF and CF, whereas L2 in NF and CF do not differ significantly. Similarly, V4 was found to be significantly higher in BF than in NF and CF, whereas no significant difference of V4 height was found for NF and CF.

Table 1: Mean  $f_0$  levels and standard deviations at L1, H, L2, and V4 in semitones (re 100 Hz).

	<i>N</i>	Mean (st)	<i>SD</i>	
L1	Broad focus	41	8.5	4.6
	Narrow focus	41	8.1	4.3
	Corrective focus	41	8.1	4.6
H	Broad focus	41	16.1	5.2
	Narrow focus	41	15.4	4.9
	Corrective focus	41	15.8	5.2
L2	Broad focus	40	13.2	5.8
	Narrow focus	40	11.9	5.0
	Corrective focus	40	11.4	5.3
V4	Broad focus	41	8.4	5.1
	Narrow focus	41	6.7	5.0
	Corrective focus	41	6.6	5.0

Next, we examined the null hypothesis that focus condition has no effect on the rate of relative  $f_0$  change, or  $f_0$  slope, of the rising and the falling  $f_0$  movement (see table 2). We carried out four repeated measures ANOVAs using slope of rise and slope of fall as dependent variables and FOCUS (BF, NF, CF) as a fixed within-subjects factor. For the slope of rise, Mauchly's test indicated that the assumption of sphericity has been violated,  $\chi^2 = 11.03$ ,  $p < .01$ . Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ( $\epsilon = .802$ ). The results show no significant effect of focus on the slope of rise ( $F [1.61, 64.19] = .582$ ,  $p > .05$ ). The difference in the slope of fall observed in figure 2, however, was found to be significant ( $F [2, 80] = 7.733$ ,  $p < .01$ ). Pairwise comparisons revealed that in both NF and CF the slope of fall is significantly larger than in BF ( $p < .01$  and  $p < .001$ , respectively), whereas the slopes of the falls in NF and CF do not differ significantly.

Table 2:  $F_0$  slopes of the nuclear rise and fall (in semitones re 100 Hz/s).

	<i>N</i>	Mean (st)	<i>SD</i>	
Rise	Broad focus	41	49.0	17.3
	Narrow focus	41	49.4	17.5
	Corrective focus	41	51.2	18.9
Fall	Broad focus	41	40.4	13.7
	Narrow focus	41	47.8	15.1
	Corrective focus	41	49.2	15.6

Finally, we examined whether there are differences in the timing of the pitch gesture relative to segmental boundaries and the position of the postnuclear stress (see table 3). We carried out four repeated measures ANOVAs using the distance of L1 from the nuclear onset, the distance of H from the beginning of the nuclear vowel, and the distance of L2 from the beginning of the first postnuclear stressed vowel as dependent variables and FOCUS (BF, NF, CF) as a fixed within-subjects factor. The results show that focus condition affected significantly the distance of H from the beginning of the nuclear vowel ( $F [2, 80] = 4.592$ ,  $p < .05$ ) and the distance of L2 from the beginning of the first postnuclear stressed vowel ( $F [2, 80] = 15.336$ ,  $p < .001$ ), but not the distance of L1 from the beginning of the onset ( $F [2, 80] = .841$ ,  $p > .05$ ). Pairwise comparisons revealed that the  $f_0$  peak occurred later relative to the beginning of the nuclear vowel in BF than in NF and CF ( $p < .05$  in both cases), whereas the alignment of the  $f_0$  peak in NF and CF does not differ significantly. In addition, L2 was found to be aligned closer to the first postnuclear stressed vowel in BF than in NF and CF ( $p < .001$  in both cases), whereas the alignment of L2 was not found to differ significantly in NF and CF.

Table 3: Mean distances of L1 to onset, H to nucleus, and L2 to stress (V4).

	<i>N</i>	Mean (ms)	<i>SD</i>	
L1 to onset	Broad focus	41	17.5	33.7
	Narrow focus	41	24.9	38.8
	Corrective focus	41	23.3	37.8
H to nucleus	Broad focus	41	109.8	27.3
	Narrow focus	41	101.1	23.8
	Corrective focus	41	96.8	24.2
L2 to stress	Broad focus	40	77.1	83.1
	Narrow focus	40	125.5	67.2
	Corrective focus	40	122.8	64.1

## 4. Discussion

The results of our experiment suggest that focus condition affects nuclear contours of Dutch declaratives in a number of ways. (1) *Segmental duration*. The onset of the nuclear syllable is lengthened when using either NF or CF. In addition, the coda was found to be lengthened in NF where the lengthening in CF was too small to reach statistical significance. (2) *Pitch scaling*. Focus condition was found to affect the scaling of the pitch contour. Both the elbow (L2) and the beginning of the first postnuclear stressed vowel (V4) were found to be scaled higher in BF than in NF and CF. These findings show that both NF and CF lower the postnuclear pitch, as can be seen in figure 2. This finding is in line with [6]. Surprisingly, the  $f_0$  peak was found to be lower in NF than in BF as well. Even if this difference is small, we note that neither NF nor CF raise the  $f_0$  peak as may be expected from [3], [6], and [10]. (3) *Slope*. The fall of the accentual gesture (but not the rise, contrary to [3] and [6]) was found to be affected by focus condition. Both in NF and CF, the falling movement was found to be steeper than in BF. (4) *Alignment*. The  $f_0$  peak (H) was found to be aligned earlier relative to the beginning of the nuclear vowel in NF and CF than in BF. In addition, the elbow (L2) was found to be aligned earlier relative to the following stressed vowel in NF and CF than in BF. Thus, in NF and CF, the whole accentual pitch gesture is somewhat time-compressed when compared to the pitch gesture in BF.

Three conclusions can be drawn from these findings. First, the data give us no ground for assuming variation in the marking of focus types by accentual peaks. In all cases, we appear to be dealing with realizations of some such nuclear contour as %L H\*L L%, or alternatively L+H\* L- L%. More generally, since neither in BF nor in NF or CF did L2 occur in or close to the first postnuclear stress, the L2 does not represent the target of an associating tone in any of the three conditions. On the basis of theories representing L2 as a phrase accent [11], the target of L might have been expected to align with the first post-nuclear stress. Earlier, [12] showed that the alignment of the beginning of the rise for the boundary H% in Dutch is not affected by any stressed syllables in the post-nuclear section of the utterance. We conclude that Standard Dutch lacks evidence for a 'phrase accent'.

Second, the findings on segmental lengthening, pitch scaling, slope, and the alignment of  $f_0$  turning points suggest that BF is pronounced less precisely than either NF or CF. However, in view of the earlier findings for West Germanic, we had not anticipated the specific phonetic differences we found. Speaker behavior appears to be conventionalized, not only to the extent that in our case NF and CF show few if any differences, but also that only specific aspects of the contour are enhanced. Surprisingly, peak height was not used in the expected way, and in fact we found a significant difference in the opposite direction, with the peak for BF being higher than in the NF and CF conditions. Since the L2 turned out to be lower in the NF and CF conditions than in the BF conditions, it would appear that instead of expanding the pitch range at the top end in the NF and CF conditions, the pitch range was expanded at the lower end.

Third, since the focus condition affects the  $f_0$  trajectory of the second half of the nuclear syllable rather than the first half, we assume that the second half of the peak is communicatively more important. First, the peak in the BF condition occurs somewhat later than in the NF and CF conditions. Although

this difference is small (approx. 11 ms), the later peak can be interpreted as a less controlled realization of the H\* in the BF condition (cf. [10:93]). More substantially, the alignment of L2 is considerably later (approx. 47 ms, taking the beginning of the following stressed syllable as the reference point) and the slope of the fall considerably shallower in the BF condition, which implies that the falling part of the accentual peak is hyperarticulated, while the rising part remains unaffected. This strongly suggests that an off-ramp analysis is superior to an on-ramp analysis. Off-ramp analyses, which take the fall to reflect the morphological and phonological element, were standardly assumed in the British tradition (e.g., 'the high fall' in [13]) and within the autosegmental tradition in ToDI [14], rather than an on-ramp analysis, as in ToBI [15]. This suggests that H\*L, preceded by a L-tone either from the boundary or a preceding pitch accent, is a better way of describing the peak than LH\*, with a following L-tone either from a following pitch accent or from a boundary.

## 5. References

- [1] Ladd, D.R., 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
- [2] Gussenhoven, C., 2007. Types of focus in English. In *Topic and focus: Cross-linguistic perspectives on meaning and intonation*, C.M. Lee; M. Gordon; D. Büring (eds.). Dordrecht: Springer, 83-100.
- [3] Baumann, S.; Becker, J.; Grice, M.; Mücke, D., 2007. Tonal and articulatory marking of focus in German. *Proc. of the 16<sup>th</sup> ICPHS, Saarbrücken*, 1029-1032.
- [4] Avesani, C.; Vayra, M., 2003. Broad, narrow and contrastive focus in Florentine Italian. *Proc. of the 15<sup>th</sup> ICPHS, Barcelona*, 1803-1806.
- [5] Jun, S.-A., 1996. *The phonetics and phonology of Korean Prosody*. New York: Garland Publishing.
- [6] Xu, Y.; Xu, C.X., 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33(2), 159-197.
- [7] Sityaev, D.; House, J., 2003. Phonetic and phonological correlates of broad, narrow and contrastive focus in English. *Proc. of the 15<sup>th</sup> ICPHS, Barcelona*, 1819-1822.
- [8] Boersma, P.; Weenink, D., 1994-2007. *Praat: doing phonetics by computer* (Vers. 4.5.14). Computer program.
- [9] Xu, Y. *A Praat script for F0 analysis* (version 2.6).
- [10] Gussenhoven, C., 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- [11] Grice, M.; Ladd, D.R.; Arvaniti, A., 2000. On the phrase accent in intonational phonology. *Phonology* 17(2), 143-185.
- [12] Peters, J.; Hanssen, J.; Gussenhoven, C., 2007. Alignment of the second low target in Dutch falling-rising pitch contours. *Proc. of Interspeech 2007, Antwerp*, 2641-2644.
- [13] O'Connor, J.D.; Arnold, G.F., 1973. *Intonation of colloquial English*. London: Longman.
- [14] Gussenhoven, C., 2005. Transcription of Dutch intonation. In *Prosodic Typology – The Phonology of Intonation and Phrasing*, S.-A. Jun (ed.). Oxford: Oxford University Press, 118-145.
- [15] Beckman, M.E.; Hirschberg, J.; Shattuck-Hufnagel, S., 2005. The original ToBI system and the evolution of the ToBI framework. In *Prosodic Typology – The Phonology of Intonation and Phrasing*, S.-A. Jun (ed.). Oxford: Oxford University Press, 9-54.