The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/159834

Please be advised that this information was generated on 2019-01-17 and may be subject to change.
Phonetic Effects of Focus in Five Varieties of Dutch

Judith Hanssen\textsuperscript{1}, Jörg Peters\textsuperscript{2}, Carlos Gussenhoven\textsuperscript{3}

\textsuperscript{1,3} Radboud University, Nijmegen, the Netherlands
\textsuperscript{2} Carl von Ossietzky University, Oldenburg, Germany

judithhanssen@gmail.com, joerg.peters@uni-oldenburg.de, c.gussenhoven@let.ru.nl

Abstract

This study examined the effects of focus on the realization of non-final nuclear falls in five varieties along the Dutch North-Sea coast. While phonetic effects surfaced more clearly in some varieties than others, we found no dialect-specific responses to the focus manipulation. In line with the findings for Standard Dutch reported in \cite{1}, focus overall affected variables associated with the falling part of the nuclear contour. The results are interpreted in terms of hyper-articulation to express differences in communicative urgency. For sentences with higher degrees of urgency, speakers sought to maximize the pronunciation of the fo fall inside the accented word, leading to shorter and steeper falls, which went down lower and sometimes started a little earlier. By lowering fo in the postnuclear stretch even further, speakers added to the communicative effect of signaling greater urgency or importance in sentences with narrow or corrective focus, compared to broad focus.

Index Terms: intonation, varieties of Dutch, focus, tonal timing, fundamental frequency, hyperarticulation

1. Introduction

A focus constituent in West Germanic languages can be larger than the word carrying the nuclear pitch accent that signals the focus (e.g., \cite{2,3}). While a distinction is traditionally made between broad focus (sentence-wide) and narrow focus (applying to constituents smaller than the sentence), to use the terms used in \cite{4}, a focus constituent can have any size, including constituents smaller than the syllable when referred to metalinguistically (cf. \cite{5}). In (1a), the focus is ‘broad’, while in (1b) and (1c) the object NP is in focus. In addition to size, different ‘focus types’ have been distinguished. The focus meaning of (1a) and (1b) is ‘informational’ \cite{6}, while (1c) is ‘corrective’ (e.g. \cite{7}). This systematic focus ambiguity in size and type is illustrated in (1), where the focus constituent is indicated by square brackets.

\begin{enumerate}
  \item a. Broad (informational)
    \begin{enumerate}
      \item What’s happening?\footnote{\texttt{H^H*L}}  \hspace{1cm} (1)
      \begin{enumerate}
        \item [They’re drinking COFFEE].
      \end{enumerate}
    \end{enumerate}
  \item b. Narrow (informational)
    \begin{enumerate}
      \item What are they drinking?\footnote{\texttt{H^H*L}}
      \begin{enumerate}
        \item [They’re drinking COFFEE].
      \end{enumerate}
    \end{enumerate}
  \item c. Narrow (corrective)
    \begin{enumerate}
      \item Are they drinking milk?\footnote{\texttt{H^H*L}}
      \begin{enumerate}
        \item [No.] They’re drinking [COFFEE].
      \end{enumerate}
    \end{enumerate}
\end{enumerate}

While (1a,b,c) are generally analyzed as having the same phonological form, the difference in focus constituent size and focus type may have phonetic effects. Cross-linguistically, higher degrees of urgency or significance are associated with prominence-increasing properties, such as higher and later or earlier peaks; larger, steeper and longer fo excursions, and longer segmental durations (e.g., 8,9,10,11,12). Within West-
to the baseline (wide information focus), however, corrective focus pitch accents were realized with increased segmental durations, higher and later peaks, lower preceding valleys, and larger rise and fall excursions. The specific enhancement strategies varied per dialect. Contrary to our pilot results, [17] did not find peak retraction, nor lowering of postfocal fo.

The purpose of this contribution is to see if our findings for Standard Dutch can be replicated in a number of dialects of Dutch. We test the hypothesis that the falling part of rising-falling nuclear pitch accents is hyperarticulated when it is communicatively more important. To this end, we designed a reading task with syntactically and lexically identical sentences that were phonologically ambiguous with respect to the size of the focus constituent (broad vs narrow) and focus meaning (i.e., informational vs. contrastive, more specifically corrective). The declarative sentences favored a rising-falling nuclear pitch accent on a non-final syllable.

2. Method

2.1. Varieties and subjects

We made recordings in five Dutch locations: Zeelandic in Zuid-Beveland (ZB), Hollandic in Rotterdam (RO) and Amsterdam (AM) (all Low Franconian), West Frisian in Grou (GR) and Low Saxon in Winschoten (WI). Data from 95 speakers were selected for analysis (17-23 speakers per variety, aged between 14-49, 40 male speakers). ZB, GR and WI speakers were bilingual with Dutch and their local language. All regional speakers and their one or both parents were raised in the selected place and spoke the indigenous variety fluently. SD recordings were included if the geographical origin of the participants could not be determined by their accent.

2.2. Materials and procedure

We used twelve declarative sentences of the type We willen in Manderen blijven wonen (‘We want to stay in Manderen’) as answers to a preceding question. Four questions elicited an answer with sentence-wide informational focus (henceforth broad focus, BF), four with narrow informational focus (NF) and four with narrow corrective focus (CF), assuming CF to have a higher ‘information weight’ than NF, and NF than BF. Example Q/A pairs are listed in (2) for each condition.

Broad focus

A. Wat is er met jullie (What’s the matter?)
B. [We willen in Manderen blijven wonen.] (2)

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.

A non-final falling nuclear pitch accent was expected to occur on the target word Manderen, a fictitious place name. Each of the target words had the metrical pattern sww (Momberen, Memberen, Manderen and Munderen) and was followed by two verbs with the pattern sw. The onset consonant was kept constant to be able to detect durational effects. We chose/m/ to limit interruptions or perturbations of the fo signal.

The Dutch set of test sentences was used for ZB, RO and AM. West Frisian (GR) and Low Saxon (WI) have their own standardized spelling systems and we therefore translated the Dutch materials into their local varieties, keeping the rhythmic, lexical, and segmental context as comparable as possible. Both language varieties reverse the order of the modal verb and the full verb, which means that the word following the test word is variable (e.g., fyse ‘to cycle’, ite ‘to eat’, ride ‘to drive’) for West Frisian.

Speakers were recorded in pairs and read each part of the mini-dialogue once. The recordings were made in a quiet room either in the homes of our speakers or in a public building, using a portable digital recorder (Zoom H4) with a 48kHz sampling rate, 16 bit resolution and stereo format. The mini-dialogues were interspersed with 61 filler sentences (used for other experiments) and presented in a booklet in pseudo-randomized order, which was reversed for half of the subjects per variety.

To ensure that our subjects interpreted the information in the broad focus condition as all-new, these four sentences all appeared in the first block of approximately twenty sentences. NF sentences appeared in the second block and CF in the last block. In a control experiment with eight speakers of Standard Dutch, we statistically tested whether the order of presentation in block 1, 2 or 3 (ORDER) affected the phonetic realization of the nuclear pitch accent in terms of duration and fo. We found that ORDER did not affect the timing and scaling of tonal targets, but shortened the segmental duration of words that were realized later in the reading task. In what follows, focus effects on segmental duration should therefore be interpreted with care if they are in the same direction as the effects of ORDER. In other words, if we find that CF or NF shorten segmental durations relative to NF or BF, respectively, this may be a consequence of ORDER instead of focus condition. Duration effects in the opposite direction (CF/NF showing longer durations compared to NF/BF) must be attributed to the factor FOCUS, although there is no way of establishing the exact degree to which differences are obscured by our confounding factor order of presentation.

2.3. Data analysis

Using the speech processing software Praat [23], we inserted the labels listed in (3), and stored their fo value (f) and time (t) to compute the dependent variables in (4). To neutralize gender differences in fo excursion, fo levels were converted to semitones re 100 Hz. Segmental labels were placed manually at segment boundaries on the basis of visual and auditory inspection of waveform and broadband spectrogram. Tonal labels were either low (L) or high (H). L1 and H were determined semi-automatically using a Praat function that traces the location of the highest or lowest fo value in a selected interval. L2 was determined visually by selecting the location of the highest change in the speed of the fo movement near the bottom line of the nuclear contour (cf. [24]). If two elbows were visible in the low-pitched section after the peak, we selected the first one.

Labels

L1 elbow before nuclear peak
H maximum fo of pitch accent
L2 elbow after nuclear peak
O1 beginning of nuclear onset
V1 beginning of nuclear vowel
C1 beginning of nuclear coda
O2 beginning of onset first postnuc. unstressed syllable
O4 beginning of onset first postnuc. stressed syllable
V4 beginning of vowel of first postnuclear stressed syllable
Variables

ONSETDURATION of accented syllable (ms)  O1 to V1
Accented RIME DURATION (ms)  V1 to O2
Accented SYLLABLE DURATION (ms)  O1 to O2
Accented WORD DURATION (ms)  O1 to O4
RISE DURATION (ms), RISE EXCURSION (ST) and RISE SLOPE (STs)  L1 to V1
FALL DURATION (ms), FALL EXCURSION (ST) and FALL SLOPE (STs)  H to L2
POSTNUCLEAR EXCURSION (ST)  H to V4
L1 TIMING from onset (ms)  L1 to O1
H TIME from nuclear vowel (ms)  V1 to H
L1 SCALING, H SCALING, L2 SCALING and V4 SCALING (ST)

We analyzed the data using the Linear Mixed Effect Model procedure in SPSS, including SPEAKER and WORD as random factors, and FOCUS (BF/CF/NF/CF) as fixed factor. Pairwise comparisons between the three levels of the fixed factor were carried out using the Bonferroni correction. To estimate the additional amount of variance explained by adding the fixed factor FOCUS to the model, as opposed to a model that only includes the random factors, we used $\Omega^2$, following [25]. The formula is

$$\Omega^2 = 1 - \frac{\text{variance residuals model random & fixed}}{\text{variance residuals model random}}$$

3. Results

3.1. Segmental timing

We observed a significant lengthening effect of FOCUS in Winschoten for ONSET DURATION: $F(2,159) = 3.70, p<.05, \Omega^2 = .0445$, with BF>CF, $p<.05$ in posthoc comparisons, and CODA DURATION: $F(2,159) = 3.94, p<.05, \Omega^2 = .0468$, with NF>BF, $p<.05$. Focus did not significantly affect segmental duration in any of the other varieties.

3.2. Scaling of tonal targets

Table 1. Estimated means of tonal scaling per variety.

<table>
<thead>
<tr>
<th></th>
<th>L1</th>
<th>L2</th>
<th>H</th>
<th>ZB</th>
<th>BF</th>
<th>NF</th>
<th>CF</th>
<th>BF</th>
<th>NF</th>
<th>CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>5.9</td>
<td>6.6</td>
<td>6.5</td>
<td>6.55</td>
<td>6.49</td>
<td>6.49</td>
<td>6.5</td>
<td>6.54</td>
<td>6.52</td>
<td>6.5</td>
</tr>
<tr>
<td>V2</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
</tr>
</tbody>
</table>

As table 1 shows, FOCUS generally had a lowering effect on low targets in all varieties and did not raise the high target of the nuclear peak, which was in fact lowered by FOCUS in ZB and AM.

The scaling of the elbow leading up to the nuclear peak (L1 SCALING) was significantly affected in ZB, RO, and WI, with BF higher than either CF or NF. (ZB $F(2,166) = 5.39, p<.01, \Omega^2 = .0669$) with BF>CF, $p<.01$; RO $F(2,175) = 3.14, p<.05, \Omega^2 = .0329$ with BF>NF, $p<.05$; WI $F(2,159) = 4.12, p<.05, \Omega^2 = .0525$ with BF>CF, $p<.05$.

FOCUS had a significant lowering effect on peak height (H SCALING) in ZB $F(2,163) = 3.49, p<.05, \Omega^2 = .0438$, with BF>CF, $p<.05$ and in AM $F(2,180) = 9.00, p<.001, \Omega^2 = .0910$, with BF>CF, $p<.001$ and NF>CF, $p<.01$.

We also found a lowering effect on the elbow after the peak (L2 SCALING) in ZB $F(2,163) = 3.89, p<.05, \Omega^2 = .0486$, with BF higher than CF, $p<.05$.

Finally, the clearest effect of FOCUS on scaling was found when we looked at V4 SCALING (to measure at the first postnuclear stressed vowel). FOCUS lowered postfocal material in all varieties, with BF higher than CF and/or NF. (ZB $F(2,164) = 3.26, p<.05, \Omega^2 = .0383$, with BF>CF, $p<.05$; RO $F(2,178) = 5.97, p<.01, \Omega^2 = .0632$, with BF>NF, $p<.01$ and BF>CF, $p<.01$; AM $F(2,180) = 8.92, p<.001, \Omega^2 = .0902$, with BF>CF, $p<.001$; GR $F(2,222) = 19.91, p<.001, \Omega^2 = .1517$, with BF>CF, $p<.001$ and BF>CF, $p<.001$; WI $F(2,159) = 5.18, p<.01, \Omega^2 = .0633$, with BF>CF, $p<.01$.)

3.3. Nuclear contour shape

This section looks at the shape (duration, excursion and slope) of the rise leading up to the nuclear peak, and the shape of the subsequent fall. We found no clear pattern across varieties for the L1H1 rise. The nuclear fall tended to be shorter in NF and CF compared to BF, with increasing excursions and steeper slopes. Shape differences between CF and NF were not always in the expected direction. We found significant effects of FOCUS on both rising and falling movements in RO, AM and WI. The effect is largest in AM, where FALL DURATION is considerably shorter, while excursion is somewhat smaller, in corrective focus than in broad and narrow focus.

RISE DURATION was affected in AM $F(2,184) = 4.37, p<.05, \Omega^2 = .0467$, with BF>CF, $p<.05$ and in WI $F(2,159) = 6.78, p<.001, \Omega^2 = .0786$, with posthoc tests showing that BF<CF, $p<.001$ and NF>CF, $p<.05$.

RISE EXCURSION was affected in RO $F(2,175) = 3.22, p<.05, \Omega^2 = .0346$, with BF<CF, $p<.05$ and in WI $F(2,161) = 3.40, p<.05, \Omega^2 = .0399$, with posthoc tests showing that BF<CF, $p<.05$.

FALL DURATION was affected in RO $F(2,173) = 5.65, \Omega^2 = .0686$, with posthoc tests showing BF>NP, $p<.05$ and BF>CF, $p<.01$; and AM $F(2,183) = 6.284, p<.01, \Omega^2 = .0652$, with BF>CF, $p<.01$ and NF>CF, $p<.05$.

FALL EXCURSION was affected in AM $F(2,183) = 6.23, p<.01, \Omega^2 = .0644$, with BF>CF, $p<.05$ and NF>CF, $p<.01$; and WI $F(2,159) = 3.08, p<.05, \Omega^2 = .0406$, with no significant variation in posthoc tests.

FALL SLOPE was affected in RO $(2,175) = 12.03, p<.001, \Omega^2 = .1255$, with BF<CF, $p<.05$ and BF<CF, $p<.01$.

Furthermore, the excursion from H* to the first postnuclear syllable (POSTNUCLEAR EXCURSION) was significantly smaller in BF than in NF and/or CF in RO $F(2,176) = 4.17, p<.05, \Omega^2 = .0468$, with BF<CF, $p<.05$ and BF<CF, $p<.05$; GR $F(2,223) = 11.11, p<.001, \Omega^2 = .0922$, with BF<CF, $p<.001$ and BF<CF, $p<.05$; and WI $F(2,161) = 4.59, p<.05, \Omega^2 = .0540$, with BF<CF, $p<.05$.

3.4. Tonal timing

FOCUS had a significant effect on tonal timing in GR and WI, with timing of both L1 and H earlier in NF and/or CF, compared to broad focus sentences.

L1 TIMING: GR $F(2,222) = 5.19, p<.01, \Omega^2 = .0431$, with BF>CF, $p<.05$ and NF>CF, $p<.05$; and WI $F(2,159) = 7.18, p<.001, \Omega^2 = .0872$, with BF>CF, $p<.01$ and NF>CF, $p<.01$. 

738
H₁: TIMING: GR \( F(2,222) = 8.71, p<.001, \Omega^2 = .0728 \), with
BF>NF, \( p<.01 \) and BF>CF, \( p<.001 \); and WI \( F(2,158) = 8.31, \)
\( p<.001, \Omega^2 = .0950 \), with NF>CF, \( p<.001 \).

The effect of focus condition on peak timing in WI is
different from the other varieties, because the peak is timed
later, not earlier, in NF compared to BF. This finding is in line
with the narrow focus results for H₁: SCALING (sec. 3.3), although
we have no explanation for it at present.

4. Discussion and conclusions

Our results show that in most of the varieties investigated, there
are small differences in phonetic realization of nuclear falling
contours as a function of focus condition. Segmental durations
in WI were longer in the NF and CF condition than in the BF
condition. Our ORDER confound may have obscured any other
durational effects. Secondly, low targets were realized lower in
all varieties in sentences with more intensified focus meanings.
The lowering effect was most obvious in the \( f_0 \) after the elbow,
which means that speakers use the postnuclear stretch to express
communicative differences. An additional lowering effect on
the nuclear peak could be observed for ZB and AM. ZB was not
otherwise affected by FOCUS. Timing effects were observed for
WI and GR. Finally, the shape (duration, excursion and slope of
\( f_0 \) movements) was most notably affected in RO, AM and WI.
In AM, the effect of FOCUS on FALLDURATION as well as
RISE-DURATION reported in sec. 3.3 may incidentally have a
phonological basis. The AM data includes both regular-peak
falls and (late) half-completed falls. The latter are associated not
only with later peaks, but also with shallower and longer falling
movements. The longer rise and fall durations in BF sentences
can be explained by a larger proportion of late-peak falls in this
condition.

4.1. Hyperarticulation

All effects of focus reported here are well-known from the
literature summarized in sec. 1 and can be interpreted in terms
of hyperarticulation [26]. Hyperarticulation can increase the
prominence of (parts of) an utterance, with the purpose of
increasing the distinctiveness of different levels of
communicative urgency [12]. More prominence, for example in
the form of larger pitch excursions, can be used to signal
emphasis, enthusiasm or increased importance. Conversely,
smaller pitch excursions are associated with less important
information or a lack of interest.

Larger pitch excursions have been reported to go hand-in-
hand with higher and later peaks. While this was confirmed in
[17], which is based on the same set of varieties and subjects as
ours, our results show earlier peaks and lowering of \( f_0 \) after the
peak. Nevertheless, our results as well as those in [17] can be
interpreted as hyperarticulation. As described in [27], there are
two ways in which a pitch peak can be enhanced. One is by
raising it, a strategy which may evolve into peak delay as a
substitute for raising, on the assumption that higher peaks are
reached later. The other strategy is to hyperarticulate the pitch
accent of which the peak is the realization. A more careful
pronunciation of a falling pitch accent in an accented syllable
may seek to maximize the pronunciation of the \( f_0 \) fall inside the
syllable rime, leading to a steeper fall that may begin earlier and
reach lower (cf. [11]). The literature on West Germanic
provides evidence for both these strategies. What unifies them
is that they both serve to signal the communicative importance
of the pitch accent’s focus constituent. We have referred to this
variation as communicative urgency, which has been
manipulated by changing the focus meaning (focus type) and the
size of the focus constituent, whereby smaller constituents are
assumed to signal greater communicative urgency.

The results of our investigation tend to confirm the strategy
used by speakers of Standard Dutch [1], whereby FOCUS overall
affected variables associated with the falling part of the nuclear
contour. Sentences with higher degrees of communicative
urgency are expressed by steeper falls, which go down lower
and may start a little earlier. The falls are also somewhat shorter
as a result of the increased steepness that is sought by the
speaker. The results reported in [17] point to the other strategy
of later and higher peaks to maximize pitch excursions. We
currently have no explanation for when which strategy is used.
We note that the corrective focus test sentences in [17]
contained three levels of urgency (CF on the nuclear accented
word, syllable or onset consonant), which may have triggered
speakers to use increasingly higher peaks.

Just like [15], who found considerable speaker variation in the
choice for particular strategies to mark focus structure, we also
have not been able to identify variety-specific preferences for
particular enhancement cues. Rather, the speaker’s goal to
hyperarticulate the fall can be attained by using a variety of
strategies. We therefore support their suggestion that speakers
can choose from different (phonological and phonetic) cues
within a functional cluster [28] to mark focus structure.

4.2. Focus size vs type

Whereas the results for Standard Dutch reported in [1]
suggested that speaker enhance the pronunciation of the fall as
a function of focus size rather than focus type, the current study
of dialects fails to show that particular distinction. We found 30
significantly different BF-CF pairs (= variation in size and
type), 13 significantly different BF-NF pairs (=size), and 12
significantly different NF-CF pairs (=type).

4.3. Contextual clues

It is likely that our speakers used, or even preferred, other cues
besides differences in the realization of intonation structures to
express or interpret focus structures. The literature reports, e.g.,
Downstep, deaccentuation, the use of a special pitch accent for
focus, or the choice for and number of prenuclear accents. Other
possibilities include visual cues [29], body language, eye
contact, or the shared context between discourse partners.

One unintended contextual clue to information structure in
our test materials was the presence of the focus marker ‘no’ at
the start of the corrective focus sentence. This disambiguating
morpheme may have had an effect on the necessity for speakers
to express the focus structure phonetically. The fact that [14],
whose test material also included the contrastive focus marker
nein ‘no’, didn’t find a durational difference between
contrastive and non-contrastive focus, supports this possibility.

5. Acknowledgements

As part of the project Intonation in Varieties of Dutch, this
research is funded by the Netherlands Organization of Scientific
Research (NWO). We thank Marrom C. Fort and Pieter Duijf for
translating the materials, Rachel Fournier and Joop Kerkhoff for
technical support, and Renske Teeuw and Lian van Hoof for
data collection and annotation. We are grateful to our speakers
and informants for giving us their time and best efforts.
6. References


