The following full text is a publisher’s version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/151184

Please be advised that this information was generated on 2019-02-18 and may be subject to change.
Jörg Peters*, Judith Hanssen and Carlos Gussenhoven

The timing of nuclear falls: Evidence from Dutch, West Frisian, Dutch Low Saxon, German Low Saxon, and High German

Abstract: A reading experiment was designed to examine the effects of word boundaries and metrical structure on the temporal alignment of the accentual peak and the end of the fall in nuclear rising-falling accents. Participants were speakers of a number of closely related dialects and languages from the coastal area of the Netherlands and North-West Germany, covering Zeelandic Dutch, Hollandic Dutch, West Frisian, Dutch Low Saxon, German Low Saxon, and Northern High German. Our findings suggest that in no variety is the timing of the nuclear peak or the end of the fall systematically affected by the location of the final word boundary or that of the following stress. In most cases, the accentual peak was found to be stably aligned with the beginning of the nuclear accented syllable, while the end of the fall occurred at a fairly constant distance from the preceding $F_0$ peak. These findings do not support a representation of the nuclear fall by a sequence of a high accentual tone and a ‘phrase accent’ that is secondarily associated to a post-nuclear stress. In addition, we found substantial cross-linguistic variation in the overall timing of the beginning and end of the fall. One component in this variation is a geographically gradient shift in the alignment of the pitch gesture.

DOI 10.1515/lp-2015-0004

1 Introduction

1.1 Background

One issue that has arisen from two decades of research into the time alignment of tonal targets is the behaviour of post-nuclear tones. A number of cases have been
presented of such tones being aligned with stressed syllables, while in other cases
the targets of similar tones have been shown to have a timing relation either with
the targets of other tones or with the upcoming phrasal boundary. The ‘stress-
seeking’ kind was labelled a ‘phrase accent’ by Grice et al. (2000), meaning a non-
starred tone that associates with an unaccented stressed syllable. In at least three
cases they present, the effect of the location of the post-nuclear stress on the
alignment of the tone has been found to be considerable. First, their Eastern
European Question Tune, a L* H-L% melody which occurs in a number of
languages in southeastern Europe, shows H aligning with the stressed syllable
of the last unaccented word, in Athenian Greek creating peaks on the antepenult,
penult, or final syllable of the phrase, depending on the location of the word
stress of the final post-nuclear word (Arvaniti 2002). Similarly, the second target of
the Hι in the Roermond Dutch interrogative nuclear melody L* HιL, associates with
the last post-nuclear stressed syllable, which again is final, penultimate, or ante-
penultimate (Gussenhoven 2000; Hι and Lι are right-hand boundary tones of the
intonational phrase).¹ Third, the English vocative chant has a mid-pitched trailing
H-tone which associates with a post-nuclear stressed syllable (Liberman 1975;
Ladd 1978; Hayes and Lahiri 1992); a Dutch counterpart spreads the H-tone to
all post-nuclear stressed syllables (Gussenhoven 1993).

The status of ‘phrase accent’ has also been attributed to tones whose align-
ment varies less obviously with context. Barnes et al. (2006) found that the
alignment of the first target of L- in a H* L-L% contour varied with the distance
to the first post-nuclear stress in American English, suggesting that L- is stress-
seeking. An extended study showed, however, that this target was in fact aligned
at a more or less fixed distance from the target of H* (Barnes et al. 2010). This
latter finding is in line with Caspers and van Heuven (1993) for Dutch, who found
that increasing speech rate compressed the rising, but not the falling, movement
of nuclear rising-falling accents. Because the determination of the alignment of
the turning point in a concave pitch movement, a pitch ‘elbow’, may be ambigu-
ous, Benzmüller and Grice (1998) examined the timing of two alignments in
German. The first was defined as “the point where the slope of the fall turns from
convex to concave, usually well above the baseline”, while the second was “the
elbow, or point where the baseline level is reached [….]” (81). They observed that
L1 occurred at a fixed distance after the preceding F0 peak, while L2 was

¹ Gussenhoven (2000) used Hι and Lι rather than H% and L%, following the practice of
subscripting boundary tones for their prosodic constituent (e.g., Hayes and Lahiri 1992). The
iota here indicates the intonational phrase rather than the intermediate phrase, as in
Pierrehumbert and Beckman (1988: 135).
preferably aligned with the first postnuclear stressed syllable. The ambiguity in the timing of the low turning point is also illustrated by a preliminary conclusion drawn by Pierrehumbert (1980: 85), who suggested that the right edge of the accented word was a likely alignment point for her phrase accent L-. Van de Ven and Gussenhoven (2011) investigated the alignment of the end of the low plateau in nuclear rising-falling-rising contours in Dutch, varying both the interval from the last stress to (a) the phrase end and (b) the location of second-occurrence focus. They found that the alignment was largely determined by the distance to the final boundary, showing that the L-tone responsible for the beginning of the final rise is not stress-seeking. The evidence for the status of a post-H* ‘phrase accent’ in these less obvious cases therefore is ambiguous at best.

An understanding of the timing of the post-H* low target will require a contextualization within the entire rising-falling pitch accent. The beginning of rising pitch movements of rising and rising-falling pitch accents was found to be aligned with the beginning of the accented syllable or rime by Caspers and van Heuven (1993) and Ladd et al. (2000) for Dutch, Niebuhr and Ambrazaitis (2006) for German, and Ladd et al. (2009) for English. As for the $F_0$ peak, Silverman and Pierrehumbert (1990) report earlier prenuclear peaks in American English with smaller distances from the end of the accented word as well as from the first stressed syllable after the accented word. This finding was replicated in Greek for some speakers by Arvaniti et al. (1998). Earlier, Steele (1986) had found variation of the timing of the nuclear peak as a function of the distance to the upcoming IP boundary. For Dutch, Schepman et al. (2006) found no effect of word length on the timing of H* in nuclear rising-falling accents, while there was only a small effect of the distance to the IP-end.

For other languages, Dalton and Ni Chasaide (2007) report that the timing of the prenuclear and nuclear peak in Cois Fharraige Irish was unaffected by the number of adjacent unstressed syllables, while Prieto et al. (1995) report effects of the end of the accented word, the distance to the upcoming prosodic boundary, and the strength of that boundary on the timing of nuclear $F_0$ peaks in Mexican Spanish, with closer distances leading to earlier peaks. As for the end of the fall, Frota (2002) shows that the trailing L of the European Portuguese focal H*+L is timed with reference to H*, in contrast to the leading H of the H+L* pitch accent, which is rather timed independently of L*. For Pisa Italian, Gili Fivela (2002) found that the post-peak low target was later in words containing two or three postnuclear syllables than in words containing one postnuclear syllable, at least in sentences with broad focus, which in her data can be interpreted either as an effect of the distance to the final word boundary or the final IP boundary. Though methodologies vary in these investigations, the variation in the findings may in part reflect cross-linguistic differences.
1.2 Cross-linguistic variation

In the British School (e.g., O’Connor and Arnold 1973), the rising-falling contour is classified as the High Fall nuclear tune. The determination of any systematicity in the dependence of the alignment of the falling slope of the High Fall on landmarks to its left and right will require a large database, because effects may be small. Also, there is likely to be regional variation, as suggested by various findings that cross-linguistic variation in West Germanic contributes to variation in the timing of nuclear and prenuclear pitch accents, even in closely related languages and dialects. Atterer and Ladd (2004) found small differences of the timing of both the low and the high targets of prenuclear rising accents in Northern and Southern German, with later peaks attested for the Southern speakers (see also Gilles 2005).

In addition, both Northern and Southern speakers of German were found to align the prenuclear rises later than suggested by data for English and Dutch. Mücke et al. (2008, 2009) found a similar difference between speakers of (Northern) Düsseldorf German and (South Eastern) Vienna German, the latter aligning the beginning and end of prenuclear rises and the peaks of rising-falling accents later. Kleber and Rathcke (2008) supplemented these studies with data from East Central German, which showed earlier alignments of the beginnings of prenuclear rises than both Northern and Southern German. Cross-linguistic variation has also been found for nuclear $F_0$ peaks. Ladd et al. (2009) found both prenuclear and nuclear peaks to be aligned later in Scottish Standard English than in RP English and nuclear peaks to be aligned earlier in RP English than in Dutch. Van Leyden and van Heuven (2006) observed a similar variation of peak timing in rising accents of the Lowland Scots dialects of Orkney and Shetland, with later peaks being found in the Orkney dialect.\footnote{There may also be differences in variability, as suggested by Dalton and Ní Chasaide (2007), who showed for Cois Fharráige Irish that the timing of the prenuclear and nuclear peak was unaffected by the number of adjacent unstressed syllables. For the closely related Inis Oírr Irish, however, the peak timing was found to be more variable.}

Grabe (1998) reports nuclear peaks aligning at 66% of the accented syllable in Southern British English, but at 98% by speakers of Braunschweig German. Peters (1999) found similar differences between northern German varieties, with speakers of Hamburg German aligning the nuclear peak at 57% of the accented syllable and Berlin speakers at 79%. This difference was increased when alignment is expressed as a proportion of the duration of the vowel (Hamburg 35%, Berlin 71%). Interestingly, in both varieties the timing of $F_0$ was found to be affected by the distance to the upcoming IP boundary, but only in Berlin German was there an effect of the end of the accented word, which may indicate that the timing of earlier peaks may be less affected by the end of the
accented word. As for the alignment of the target of the L-tone after H*, no cross-linguistic variation has to our knowledge been reported for West Germanic.

In the investigation reported here, we were particularly interested in the extent to which the dependence of the timing of the post-peak low target on nearby prosodic boundaries varies across varieties and languages that are geographically close. We report a reading experiment that was designed to examine the variation in the timing of the rising-falling accent in seven varieties spoken along the coast of the Netherlands and North-Western Germany. While representing a section of a traditional dialect continuum, this geographic cline includes three standard language areas, Dutch, Frisian, and German. A further opportunity to show that effects of geographic proximity may be independent of the sociolinguistic allegiance of the dialects was provided by the bidialectal speakers from North-West Germany, who read comparable materials twice, once in their dialect and once in their regiolectal standard German.

There is no unique autosegmental representation of the High Fall. The ToBI framework originally developed for English (Beckman and Ayers Elam 1997; Beckman et al. 2005) represents it as (L+)H* L-L%, where the end of the fall is accounted for by the ‘phrase accent’. In the analysis of Gussenhoven (2005), the High Fall is represented as H*L L%, in which the end of the falling movement is represented as a trailing tone. The same analysis was applied to English (Gussenhoven 1991, 2004) and Standard German (Peters 2009, 2014). In our investigation, we will put the two hypotheses of the alignment of L-, i.e., an alignment with the upcoming post-nuclear stressed syllable and an alignment with the end of the nuclear word, to the test. In Section 4, we will suggest that the assumptions underlying the existence of a trailing tone in those analyses do not in fact imply the phonetic alignment characteristics that have been assumed in the ToBI framework.

1.3 Timing strategies

Assuming an accented word with initial stress, there are theoretically three strategies speakers may adopt in the alignment of an accentual peak when accommodating to variation in word length. First, increasing the number of word-internal postnuclear syllables may affect the distance of a given target in the LHL-contour to the end of the nuclear word, but not its distance to the beginning of the word. In this strategy, the tonal target stays at a fixed distance from the beginning of the nuclear word, but not its duration (Strategy 1). Second, increasing the number of word-internal postnuclear syllables may affect the distance of the tonal target to the beginning of the nuclear word, but not its distance to the word end. This strategy represents Pierrehumbert’s (1980) word-based hypothesis for the alignment of L-. In
this case, the tonal target occurs at a fixed distance from the end of the nuclear word irrespective of its duration (Strategy 2). Third, increasing the size of the nuclear word may affect the target’s distance from both the beginning of the nuclear word and from its end. In this case, increasing word size results in a proportional adjustment of the alignment (Strategy 3). The same three options will apply to the alignment of the rise-fall with respect to the interval between the beginning of the accented syllable and the first post-nuclear word-stressed syllable. We will refer to this interval as the ‘nuclear foot’, in the spirit of Abercrombie (1964). Note that the nuclear foot may be longer than the nuclear word. This is the case if the first post-nuclear word begins with an unstressed syllable, which is included in the nuclear foot. As will be clear, Strategy 2 now represents the widely entertained hypothesis that the upcoming stress provides the alignment point for L-.

Figure 1 illustrates the three strategies for the alignment of H* relative to either the nuclear word or nuclear foot. C1 marks the beginning of the nuclear word or nuclear foot, where the initial syllable is accented. Wb and St mark the final boundary of the nuclear word and foot, respectively, which may or may not coincide. t1 and t2 mark the intervals from H* to the beginning of the nuclear word or foot and from H* to the end of the nuclear word or nuclear foot, respectively. The interval \( t_1 + t_2 \) corresponds to the overall duration of the nuclear word or foot. The three panels of Figure 1 illustrate the three strategies for accommodating an increased length of the nuclear word or foot on \( t_1 \) and \( t_2 \). A correlation analysis will be used to detect these possible effects of increasing length. Specifically, a positive correlation of \( t_1 + t_2 \) with \( t_2 \) and no correlation with \( t_1 \) will indicate that Strategy 1 is adopted (left panel), while a positive correlation of \( t_1 + t_2 \) with \( t_1 \) and no correlation with \( t_2 \) suggests that Strategy 2 is adopted (central panel). Finally, positive correlations between \( t_1 + t_2 \) and both \( t_1 \) and \( t_2 \) suggest that Strategy 3 is adopted (right panel).

Figure 1: Three strategies for the timing of H* relative to the beginning (C1) and end (Wb/St) of the nuclear word or nuclear foot.
Figure 2 illustrates the three strategies for the timing of L1 relative to the accentual peak (H) and the end of the nuclear word or foot (Wb/St).

Figure 2 illustrates the three strategies for the timing of the tonal low target at the end of the nuclear falling movement, here referred to as L1. We defined \( t_1 \) as the distance of L1 from the target of H*, because the location of H* more so than that of C1 determines the time available for it. \( t_2 \) is the distance of L1 from the end of the nuclear word or nuclear foot (Wb/St).

The aim of the present paper is to answer the following questions:

- **The timing strategy for H.** Does H occur at a fixed distance from the beginning of the nuclear word/foot? (Strategy 1) Or is it affected by the location of the final boundary of the nuclear word or foot (Wb or St)? And if so, does H stay at a fixed distance from Wb or St (Strategy 2), or does it move proportionally to the duration of the nuclear word or foot (Strategy 3)?

- **The timing strategy for L1.** Does L1 occur at a fixed distance from the preceding H? (Strategy 1) Or is it affected by the location of Wb or St? And if so, does H stay at a fixed distance from Wb or St (Strategy 2), or does it move proportionally to the duration of the nuclear word or foot (Strategy 3)?

- Are there overall differences in the timing of H and L1 in the varieties examined? And if so, do these differences have any relevance for the strategies adopted for the timing of H and L1?

---

3 It is important to see that our research question cannot be addressed by using target delay with reference to a landmark in the accented rime, like its end or beginning. Increases in the number of syllables in a word with initial stress will shorten all syllables (Nooteboom 1972; Rietveld et al. 2004), and rime duration will affect those measures regardless of word length (Schepman et al. 2006). Thus, L1 delay or H delay with reference to the rime end will be an inevitable consequence of increasing word length, but this effect cannot be interpreted as shifts of these targets to the right word boundary. To see if H moves right with increasing word length, the relevant landmarks therefore are the beginning and the end of the word.
2 Method

2.1 Speech materials

We used ordinary statements as carrier sentences. In order to elicit High Falls, these carrier sentences were presented as answers to a preceding question, as in (1).

(1)  
   a. Waarom kijk je zo geschrokken?
       ‘Why are you looking so frightened?’

   b. De buren gaan iets ergs over tante Mol beweren.
       ‘The neighbours are about to say something nasty about Aunt Mol.’

Each carrier sentence contained a fictitious proper name that was intended to bear the nuclear accent, *Mol* in (1b), followed by an infinitive verb form in postnuclear position, *beweren* in (1b). To prevent stress clash effects on the timing of the nuclear pitch accent gesture, we made sure that in each carrier sentence the proper name was preceded by an unstressed syllable. As proper names we used three oxytones, three paroxytones, and three proparoxytones, all of which start with the lexical stressed syllable. As verb forms we used three pairs of paroxytones which differed by the presence or absence of an initial unstressed syllable corresponding to the prefix *be*- in all our languages. Table 1 lists all proper names and verb forms used in the Standard Dutch version of the test sentences classified according to their position in the carrier sentence (nuclear vs. postnuclear) and their stress patterns.

| Nuclear | Postnuclear |
|---------|-------------|
| s       | sw          | sww | sw | wsw |
| Mol     | Molber      | Molberen | weren | beweren |
| Lum     | Lumber      | Lumberen | landen | belanden |
| Mel     | Melber      | Melberen | noemen | benoemen |

Proper names and verb forms were combined such that six types of carrier sentences were created, differing by the distance between the lexically stressed syllable of the nuclear word and the final word boundary of the nuclear word (=WbDist) and by the distance between the lexically stressed syllable of the nuclear word and the first postnuclear lexical stress (=StDist), both counted in terms of the number of intervening syllables. In *Mol beweren* in (1), for example,
WbDist = 0, as no syllable intervenes between the accented syllable and the final word boundary of Mol, while StDist = 1, as one syllable intervenes between the nuclear and the first postnuclear stress. Table 2 lists one sample dialogue for each experimental condition of the Standard Dutch version. WbDist was varied from 0 to 2, whereas StDist was varied from 0 to 3.

Table 2: Combinations of the factors WbDist and StDist of the Standard Dutch version. Nuclear and postnuclear stressed syllables are marked in bold.

| Experim. condition | WbDist | StDist | Dialogue                                      |
|---------------------|--------|--------|----------------------------------------------|
| 1                   | 0      | 0      | A. Wat heb je tegen je kind gezegd?          |
|                     |        |        | B. Het moet zich tegen buurman Mol weren.     |
| 2                   | 0      | 1      | A. Waarom kijk je zo geschrokken?            |
|                     |        |        | B. De buren gaan iets ergs over tante Mol beweren. |
| 3                   | 1      | 1      | A. Ik heb echt altijd hoofdpijn.             |
|                     |        |        | B. Dan moet je niet tegen de pillen van dokter Molber weren. |
| 4                   | 1      | 2      | A. Ik hoorde dat de school is afgebrand.     |
|                     |        |        | B. Ja, dat gaat burgermeester Molber beweren.|
| 5                   | 2      | 2      | A. Wat gaat er gebeuren?                    |
|                     |        |        | B. Ze willen de vossen uit het bos bij Molberen weren. |
| 6                   | 2      | 3      | A. Wat is het probleem?                     |
|                     |        |        | B. Iedereen gelooft wat ze over opa Molberen beweren. |

Note that WbDist and StDist are not independent. StDist corresponds to WbDist or WbDist + 1, but it is never smaller than WbDist. Different levels of WbDist can be compared while keeping StDist constant if we check condition 2 against 3 for StDist = 1 and condition 4 against 5 for StDist = 2. Likewise, different levels of StDist can be compared if we check condition 1 against 2 for WbDist = 0, condition 3 against 4 for WbDist = 1, and condition 5 against 6 for WbDist = 2.

We used three carrier sentences per condition, which amounts to a total of 18 sentences, each of which was prompted by a question. The Standard Dutch version of the sentences was used for speakers from Zuid-Beveland, Rotterdam, and Amsterdam. For the other speakers, we used translations into the local language, which was West Frisian, Dutch Low Saxon, German Low Saxon, or High German (for more information on our speakers see Section 2.3). In the translations, the proper names were kept constant, whereas the postnuclear words were adapted according to the respective language. In some cases, particularly in the Low Saxon and High German language versions, modified versions rather than translations of the Standard Dutch version had to be used in order to keep the metrical structure constant across languages and dialects. An overview of the sentence in all language versions is given in the Appendix.
2.2 Recording procedure

The mini-dialogues were presented in a booklet, one dialogue per page. To prevent order effects, the dialogues were presented in pseudo-randomized order, which was reversed for half of the subjects per variety, and 87 mini-dialogues from other experiments were added as fillers. To reduce effects of the experimenter’s presence on the speakers’ dialect level, our speakers were recorded in pairs, with one speaker producing the context sentence and the other the carrier sentence. The participants switched roles at the end of the task after they had repeated any mispronounced sentences. The German Low Saxon and High German test sentences were recorded by the same speakers on two visits which were at least four weeks apart. Recordings were made in a quiet room either in the homes of our speakers or in a public building. We used a portable digital recorder (Zoom H4) with a 48 kHz sampling rate, 16 bit resolution, and stereo format. The participants wore head-mounted Shure WH30XLR wired condenser microphones.4

2.3 Speakers

Recordings were made in six places along the coastal line stretching from Zeeland in the South-West to Weener in the North-East (see Figure 3). These recordings covered six local varieties, Zeelandic Dutch in Zuid-Beveland (ZB),

![Figure 3: Recording locations in the Netherlands and North-West Germany.](image)

4 For the recordings of Low Saxon and High German in Weener, which were added at a later time, we used the Tascam HD P2 recorder and Sennheiser MKE 2 wired condenser microphones.
Rotterdam Dutch (RO), Amsterdam Dutch (AM), West Frisian in Grou (GR), Dutch Low Saxon in Winschoten (WI), and German Low Saxon in Weener (WL). The Weener speakers were recorded on a second visit to obtain comparable data of the local standard variety, which is Weener High German (WH).

We recorded a total of 125 speakers, aged between 16 and 45. The speakers from Zuid-Beveland, Grou, and Winschoten were bilingual with Standard Dutch and their local language. All regional speakers and at least one of their parents spoke the indigenous variety fluently, and they were raised in the location concerned. The speakers from Weener were bilingual with German Low Saxon and High German. Except for the speakers of West Frisian, our speakers were less familiar with their local language as a written language, which may have had a negative influence on the fluency of the speech in the reading task of some speakers. This is particularly true of our Weener speakers when speaking German Low Saxon.

Eleven speakers were excluded because their speech was disfluent or because their linguistic background later turned out not to be representative of the local variety. Table 3 gives an overview of the speakers recorded per variety. The participants were naïve as to the purpose of the experiment and were paid for their participation.

Table 3: Number of speakers from Zuid-Beveland (ZB), Rotterdam (RO), Amsterdam (AM), Grou (GR), Winschoten (WI), and Weener, with WL for Weener Low German and WH for Weener High German.

|          | ZB | RO | AM | GR | WI | WL/WH | Total |
|----------|----|----|----|----|----|--------|-------|
| total    | 18 | 20 | 24 | 23 | 20 | 20     | 125   |
| selected | 18 | 19 | 18 | 23 | 18 | 18     | 114   |
| male     | 10 | 12 | 12 | 3  | 4  | 7      | 48    |
| female   | 8  | 7  | 6  | 20 | 14 | 11     | 66    |

Note: Note that in the remainder of this paper the place labels will likewise be used for the places and for the varieties spoken in these places.

2.4 Acoustic analysis

All recorded carrier sentences were converted to monaural files and stored on computer disk as separate wave files. Utterances were excluded from further analysis if they showed deviant pitch patterns due to accent position, choice of pitch accent, or choice of final boundary tone. In particular, we excluded utterances with a downstepped nuclear accent. As it happens, a relatively small number of the utterances in our experiment had downstep. We also excluded utterances with hesitation pauses on and around the target word. After exclusion
of all irregular items, about 81% of the recorded utterances were left for acoustic analysis, which were evenly distributed across speaker groups.

Acoustic and auditory analyses of the data were done with the help of the speech processing software package Praat (Boersma and Weenink 2008). We inserted the labels listed in Table 4.

Table 4: Acoustic measurement labels.

1. Segmental labels
   - C1: beginning of the onset of the nuclear syllable
   - WB: end of the nuclear word
   - ST: end of the nuclear foot (= beginning of the first postnuclear lexically stressed syllable)

2. Pitch labels
   - L0: time stamp of the minimum F₀ at or around the beginning of the nuclear syllable (= begin of rise)
   - H: time stamp of the maximum F₀ on the nuclear word (= nuclear peak)
   - L1: time stamp of the ‘elbow’ after the nuclear peak

In general, syllable boundaries were determined on the basis of visual inspection of the waveform and the broadband spectrogram, aided by auditory information. We placed all labels at negative-to-positive zero-crossings of the sound wave. L₀ and H were determined semi-automatically using a Praat function to locate the F₀ minimum or maximum in a selected region. Semi-automatic determination of the elbow after the nuclear peak (L₁) was found to yield less inter-rater agreement for our data set. Therefore, we determined L₁ visually by looking for the location of the highest rate of F₀ change near the bottom line of the nuclear contour. For a comparison of manual with automatic detection methods, see del Giudice et al. (2007). Pitch tracking errors such as octave jumps were corrected by hand. Using the labels in Table 4, we calculated the acoustic variables given in Table 5.

From the above measures, all variables can be derived that are necessary to identify the timing intervals t₁ and t₂ that we used to define the three timing

Table 5: Acoustic variables (all in ms).

1. Durational variables
   - WordDur: the duration of the nuclear word
   - FootDur: the duration of the nuclear foot

2. Latency variables
   - L₀ Delay: the distance of L₀ from the beginning of the nuclear word
   - H Delay: the distance of H from the beginning of the nuclear word
   - L₁ Delay: the distance of L₁ from H
strategies in Section 1. Table 6 gives an overview of variants of $t1$ and $t2$ for the timing of H and L with respect to the nuclear word and foot.

**Table 6:** Derived variables $t1$ and $t2$ for the alignment of H and L1 with respect to the nuclear word and foot.

| Timing of H          |                                                                 |
|----------------------|-----------------------------------------------------------------|
| **Nuclear word**     |                                                                 |
| $t1_{H/w}$           | the distance of H from the beginning of the nuclear word         |
| $t2_{H/w}$           | the distance of H from the end of the nuclear word               |
| $t1_{H/w} + t2_{H/w}$| duration of nuclear word                                       |
| **Nuclear foot**     |                                                                 |
| $t1_{H/f}$           | the distance of H from the beginning of the nuclear foot         |
| $t2_{H/f}$           | the distance of H from the end of the nuclear foot               |
| $t1_{H/f} + t2_{H/f}$| duration of nuclear foot                                       |

| Timing of L1         |                                                                 |
|----------------------|-----------------------------------------------------------------|
| **Nuclear word**     |                                                                 |
| $t1_{L/w}$           | the distance of L1 from H                                       |
| $t2_{L/w}$           | the distance of L1 from the end of the nuclear word             |
| $t1_{L/w} + t2_{L/w}$| duration of nuclear word                                       |
| **Nuclear foot**     |                                                                 |
| $t1_{L/f}$           | the distance of L1 from H                                       |
| $t2_{L/f}$           | the distance of L1 from the end of the nuclear foot             |
| $t1_{L/f} + t2_{L/f}$| duration of nuclear foot                                       |

To check the reliability of measurements, one sentence per experimental condition was quasi-randomly selected from one male and one female speaker per variety and labelled independently by the first author and a trained phonetician ($6 \times 2 \times 7 = 84$ sentences). Table 7 gives the averaged absolute differences between the two measurements for each of the acoustic variables listed in Table 5. For most variables inter-rater agreement was acceptable. The largest mean absolute difference was found for L1 Delay, which can mainly be attributed to a lower agreement about the location of L1.

**Table 7:** Comparison of two independent measurements ($N = 2 \times 84$ measurements per variable).

| Variable   | Mean absolute difference (ms) | Variable   | Mean absolute difference (ms) |
|------------|-------------------------------|------------|-------------------------------|
| WordDur    | 12                            | L0 Delay   | 16                            |
| FootDur    | 17                            | H Delay    | 5                             |
|            |                               | L1 Delay   | 28                            |
2.4.1 Statistical analysis

To account for inter-speaker differences and for the interdependence of observations per speaker, we built multi-level regression models using Maximum Likelihood Estimation of Linear Mixed Effects Models (LME Models) in SPSS including Speaker and Sentence (sentences per condition) as random factors. Pairwise comparisons reported for the levels of the fixed factors were carried out as part of the Mixed Effect procedure in SPSS and adapted using the Bonferroni correction.

To estimate the amount of variation explained by WbDist or StDist when choosing either t1 or t2 as the dependent variable, we compared the fit of each model before and after adding WbDist or StDist as a fixed factor. For comparison of model fit we used $\Omega^2$, according to Xu (2003), which can be estimated by

$$r^2 = 1 - \frac{\hat{\sigma}^2}{\sigma_0^2}$$

where $\hat{\sigma}^2$ is the estimate of the residual variance of the final model and $\sigma_0^2$ is the estimate of the residual variance of the null model, which in our case only contained the random factors. The values of $\Omega^2$ range between 0 and 1.

3 Results

3.1 Duration of nuclear word and nuclear foot

To ensure that the experimental manipulation of the number of syllables of the nuclear word and the nuclear foot was not compensated for by a change in speech rate, we first checked whether adding syllables indeed increased the duration of the nuclear word (WordDur) and foot (FootDur).

Figure 4 reveals that increasing the number of syllables between the accented syllable and the right word boundary (=WbDist) from 0 to 2 had the desired effect. Every additional syllable increased WordDur. For statistical analysis, we built LME Models including Speaker and Sentence as random factors and WbDist (0, 1, 2) and Dialect (ZB, RO, AM, GR, WI, WL, WH) as fixed factors. Significant main effects on WordDur were found for both WbDist, $F(2,1863) = 3353.58$, $p < 0.001$, and Dialect, $F(6,139) = 7.59$, $p < 0.001$. In addition, there was a significant interaction between WbDist and Dialect, $F(12,1862) = 10.85$, $p < 0.001$. 


Pairwise comparisons of levels of WbDist revealed significant differences between all levels of WbDist at the 0.1% level. Levels of Dialect revealed that WL differed significantly from ZB ($p < 0.01$), RO ($p < 0.05$), and WH ($p < 0.001$).

Figure 5 shows analogous measurements of the effect of varying the distance between the nuclear syllable and the first postnuclear stress (StDist) from 0 to 3 syllables on the duration of the nuclear foot (FootDur). Again, increasing the number of intervening syllables had the desired effect. For statistical analysis we included Speaker and Sentence as random factors and StDist (0, 1, 2, 3) and Dialect as fixed factors. Significant main effects on FootDur were found for StDist, $F(3,1859) = 2813.01$, $p < 0.001$, and Dialect, $F(6,141) = 6.37$, $p < 0.001$. There was also a significant interaction between StDist and Dialect, $F(18,1858) = 4.28$, $p < 0.001$.

Pairwise comparisons of levels of StDist revealed significant differences between all levels of WbDist at the 0.1% level. Comparisons of levels of Dialect revealed that WL differed significantly from WH ($p < 0.001$). We conclude that all
experimental manipulations had the desired effects on the location of the final word and foot boundary in all varieties.

In Sections 3.2 and 3.3 we examine the effects of WbDist and StDist on the timing of H within the accented word and the nuclear foot and of L between the preceding tonal target, H, and the end of the nuclear word and foot. As we were particularly interested in differences in timing strategies as a function of language, we fitted separate regression models for each variety.

3.2 Alignment of H

3.2.1 Effects of the distance of the final word boundary

Figure 6 illustrates for each variety the effects of manipulating the number of postnuclear syllables within the accented word (WbDist) on the locations of L0, H, and L1 while keeping the distance of the accented syllable to the next postnuclear stress (StDist) constant.

In the lower parts of the panels WbDist varies between 0 and 1 syllable, while StDist is restricted to 1 syllable. Examples are *Mol beweren* (WbDist = 0) and *Molber weren* (WbDist = 1). In the upper parts of the panels, WbDist varies between 1 and 2 syllables, while StDist is restricted to 2 syllables, as in *Molber beweren* (WbDist = 1) and *Molberen weren* (WbDist = 2).

Even though we were investigating the timing of H and L1, we have added locations of L0 (the beginning of the rise). The diagrams show that in all varieties, except WL and WH, L0 occurs close to the beginning of the nuclear word so that we do not expect substantial differences between H-latencies from L0 and latencies from the beginning of the nuclear word. Since WL and WH show an earlier alignment of L0, we will consider this case in Section 3.4.

Figure 6 shows that increasing WbDist from 0 to 2 syllables results in an expected increase of the duration of the accented word, which is indicated by the overall length of the horizontal bars. On the other hand, the H-latency from the beginning of the nuclear word (at 0 ms) remains fairly constant. In RO and AM, the timing of H for StDist = 2 is more variable, but only in AM does word duration correlate positively with the location of H.

To examine which strategy our speakers used (Section 1.3), we proceeded in two steps. First, we examined the effect of the categorical variable WbDist on the location of H (Model 1). Second, we examined the effect of the continuous variable word duration (WordDur) on the location of H (Model 2). In the first case, we used only data with StDist = 1 and StDist = 2 in order to control for
Figure 6: Mean durations of the nuclear words and mean locations of L0, H, and L1 as a function of the experimental conditions WbDist = 0 and 1 for StDist = 1, and for WbDist = 1 and 2 for StDist = 2. The beginning of the nuclear word is set to 0 ms, and the end is indicated by WB.
stress position when comparing different levels of WbDist, as illustrated in Figure 4. In the second case we included all levels of WbDist and StDist.

To account for the effect of WbDist on the location of H (Model 1), we fitted two LME Models, each with a different dependent variable. According to Strategy 1, increasing WbDist should increase the distance of H from the final boundary of the accented word, which we used as our first dependent variable, \( t^2 \) (Model 1a). According to Strategy 2, we expected that increasing WbDist would increase the latency of H from the beginning of the accented word, which was our second dependent variable, \( t^1 \) (Model 1b). In both models, we used Speaker and Sentence as random factors and WbDist as a fixed factor. As WbDist and StDist do not vary independently of each other (see Section 2.1), we carried out separate tests for the two relevant levels of distance of stress, StDist = 1 and StDist = 2.

Table 8 shows significant effects of WbDist on \( t^2_{H/w} \) in all varieties and for both levels of StDist, whereas significant effects on \( t^1_{H/w} \) were restricted to GR and WH for StDist = 1 and AM and GR for StDist = 2.

**Strategy 1** was taken to be revealed when WbDist significantly predicted \( t^2_{H/w} \) but not \( t^1_{H/w} \), **Strategy 2** when WbDist significantly predicted \( t^1_{H/w} \) but not \( t^2_{H/w} \), and **Strategy 3** when WbDist significantly predicted both \( t^1_{H/w} \) and \( t^2_{H/w} \) (see Section 1.3). Accordingly, we conclude that **Strategy 1** was used by speakers of ZB, RO, WI, and WL regardless of the value of StDist, and by speakers of AM

---

### Table 8: Effects of WbDist on \( t^1_{H/w} \) and \( t^2_{H/w} \)

|         | \( t^1_{H/w} \) | \( t^2_{H/w} \) |
|---------|-----------------|-----------------|
|         | \( df \) | \( F \) | \( \Omega^2 \) | \( df \) | \( F \) | \( \Omega^2 \) |
| ZB      | 1       | 1,76  | 3.82  | 0.04  | 1,80  | 326.52*** | 0.78       |
| RO      | 1       | 1,82  | 0.27  | 0.00  | 1,84  | 341.57*** | 0.79       |
| AM      | 1       | 1,56  | 1.10  | 0.02  | 1,58  | 159.75*** | 0.73       |
| GR      | 1       | 1,100 | 4.88* | 0.04  | 1,101 | 140.03*** | 0.58       |
| WI      | 1       | 1,80  | 0.97  | 0.01  | 1,81  | 210.62*** | 0.72       |
| WL      | 1       | 1,76  | 1.43  | 0.03  | 1,75  | 248.92*** | 0.77       |
| WH      | 1       | 1,73  | 7.32** | 0.09 | 1,76  | 263.02*** | 0.77       |
| ZB      | 2       | 1,64  | 0.06  | 0.00  | 1,68  | 31.52***  | 0.33       |
| RO      | 2       | 1,69  | 3.32  | 0.04  | 1,67  | 79.23***  | 0.55       |
| AM      | 2       | 1,55  | 4.36* | 0.06  | 1,55  | 52.34***  | 0.49       |
| GR      | 2       | 1,106 | 4.67* | 0.04  | 1,103 | 84.40***  | 0.43       |
| WI      | 2       | 1,73  | 2.47  | 0.03  | 1,72  | 42.85***  | 0.37       |
| WL      | 2       | 1,73  | 1.54  | 0.02  | 1,70  | 187.60*** | 0.73       |
| WH      | 2       | 1,81  | 3.39  | 0.04  | 1,81  | 375.88*** | 0.82       |
when StDist was 1. **Strategy 3** was used by speakers of GR and WH when StDist was 1 and by speakers of AM and GR when StDist was 2.

The estimates of $\Omega^2$ added in Table 8 suggest that WbDist improves the model fit substantially if the model is designed to predict the location of H relative to the final word boundary ($t_{2H/w}$) (last column). On the other hand, WbDist does not substantially help in predicting the location of H relative to the beginning of the accented word ($t_{1H/w}$) (5th column). In particular, the estimates of $\Omega^2$ suggest that in those cases where **Strategy 3** was used, moving the final word boundary rightwards by increasing the number of syllables was accompanied by a much smaller amount of rightward-movement of H.

As a second step, we used the duration of the nuclear word (WordDur) to predict $t_{1H/w}$ and $t_{2H/w}$ (Model 2). The scatter plots in Figure 7 show for each variety the relation between WordDur (x-axis) and the distance of H from the beginning of the word ($t_{1H/w}$, y-axis, first and second columns) and the distance of H from the final word boundary ($t_{2H/w}$, y-axis, third and fourth columns). The graphs reveal for all varieties strong positive correlations between WordDur and the distance of H from the final word boundary (WordDur vs. $t_{2H/w}$), but not between word duration and the distance of H from the beginning of the word (WordDur vs. $t_{1H/w}$).

To determine the significance of the effect of WordDur on $t_{1H/w}$ and $t_{2H/w}$, we fitted a LME Model including Speaker and Sentence as random factors and WordDur as a covariate using either $t_{1H/w}$ as a dependent variable (Model 2a) or $t_{2H/w}$ (Model 2b). Table 9 lists $F$ values and estimates of $\Omega^2$. The results show significant effects of WordDur on both dependent variables in all varieties, suggesting **Strategy 3** for speakers of all varieties. The estimates of $\Omega^2$, however, reveal that WordDur predicts $t_{2H/w}$ much better than $t_{1H/w}$. In the latter case, WordDur only marginally improves model fit. These results suggest for all varieties that if there is a proportional movement of H relative to the size of the nuclear word, the distance from the final word boundary increases more than the distance from the beginning of the nuclear word. We conclude that, similarly to the findings for Model 1 using WbDist as a predictor variable, increasing WordDur may cause some peak delay, but the effect is small.

### 3.2.2 Effects of the distance of the following lexical stress

In the previous section we examined the effect of word length (WbDist) and word duration on H–distances to the boundaries of the nuclear word ($t_{1H/w}$ and $t_{2H/w}$). We now turn to the effect of the distance to the following lexical stress (StDist) and the duration of the nuclear foot (FootDur) on the location of H with respect to the boundaries of the nuclear foot.
Figure 7: Scatterplots of word duration vs. $t_{1H/w}$ (first and second columns) and vs. $t_{2H/w}$ (third and fourth columns).
Again, we start with examining effects of the categorical variable, which in this case is StDist. Figure 8 illustrates for each variety the effects of manipulating the number of postnuclear syllables within the nuclear foot (StDist) on the locations of L0, H, and L1 while keeping the distance of the accented syllable to the end of the word (WbDist) constant.

In the lower parts of the panels, StDist varies between 0 and 1 syllable, while WbDist is 0. Examples are Mol weren (StDist = 0) and Mol beweren (StDist = 1). In the middle parts of the panels, StDist varies between 1 and 2 syllables, while WbDist is restricted to 1 syllable. Examples are Molber weren (StDist = 1) and Molber beweren (StDist = 2). In the upper parts of the panels, StDist varies between 2 and 3 syllables, while WbDist is restricted to 2 syllables, as in Molberen weren (StDist = 2) and Molberen beweren (StDist = 3).

The graphs show that increasing StDist from 0 to 3 syllables results in an expected increase of the duration of the nuclear foot, which is indicated by the overall length of the horizontal bars. Increasing StDist substantially affects the location of H with respect to the end of the nuclear foot, but not with respect to its beginning, which coincides with the beginning of the nuclear word at 0 ms.

To account for the effect of StDist on the location of H, we fitted two LME Models differing by the dependent variable. According to Strategy 1, increasing StDist is expected to increase the distance of H from a nearby segmental landmark, in this case the end of the nuclear foot, which we used as our first dependent variable, $t_2H/f$ (Model 1a). According to Strategy 2, we expected that increasing StDist increases the distance of H from the beginning of the nuclear foot, which we used as our second dependent variable, $t_1H/f$ (Model 1b). In both models, we used Speaker and Sentence as random factors and StDist as a fixed factor. We carried out three separate tests for the relevant levels of distance of word boundary, WbDist = 0, WbDist = 1, and WbDist = 2.

### Table 9: Effects of WordDur on $t_{1H/w}$ and $t_{2H/w}$

|       | $t_{1H/w}$ |         | $t_{2H/w}$ |         |
|-------|------------|---------|------------|---------|
|       | df         | F       | $\Omega^2$ | df      | F       | $\Omega^2$ |
| ZB    | 1,237      | 4.24*   | 0.02       | 1,237   | 2,814.67*** | 0.92      |
| RO    | 1,266      | 7.71**  | 0.02       | 1,266   | 3,122.64*** | 0.92      |
| AM    | 1,205      | 26.06***| 0.10       | 1,205   | 1,342.53*** | 0.87      |
| GR    | 1,372      | 80.00***| 0.16       | 1,371   | 1,750.12*** | 0.83      |
| WI    | 1,270      | 7.05**  | 0.02       | 1,270   | 1,595.40*** | 0.86      |
| WL    | 1,267      | 25.79***| 0.08       | 1,267   | 2,580.42*** | 0.91      |
| WH    | 1,281      | 22.38***| 0.06       | 1,281   | 2,731.08*** | 0.91      |
Figure 8: Mean durations of nuclear feet and mean locations of L0, H, and L1 as a function of the experimental conditions StDist = 0 and 1 for WbDist = 1, StDist = 1 and 2 for WbDist = 1, and StDist = 2 and 3 for WbDist = 2. The beginning of the nuclear foot is set to 0 ms and the end is indicated by ST.
Table 10 shows significant effects of StDist on $t_{1H/ft}$ for all levels of WbDist, whereas significant effects on $t_{1H/ft}$ were restricted to GR for WbDist = 0, RO and WI for WbDist = 1, and ZB, AM, and WL for StDist = 2. Estimates of $\Omega^2$ reveal a large or medium effect of StDist on $t_{2H/ft}$, but no or a small effect on $t_{1H/ft}$. We conclude that our speakers used either Strategy 1 or Strategy 3 for aligning H with respect to the nuclear foot, just as they did with respect to the nuclear word. Moving the following postnuclear stress away from the accented syllable did not substantially affect the location of H relative to the beginning of the nuclear foot. This result is not unexpected in view of the previous finding that H occurs at a rather fixed distance from the beginning of the nuclear word, as in our test sentences the beginning of the nuclear foot always coincided with the beginning of the nuclear word.

Figure 9 shows scatter plots for the relation between foot duration (FootDur, x-axis) and the H-latency from the beginning of the foot ($t_{1H/ft}$, y-axis, first and
Figure 9: Scatterplots of foot duration vs. $t1_{H/f}$ (first and second columns) and vs. $t2_{H/f}$ (third and fourth columns).
second columns) as well as the H-latency from the final foot boundary \((t2_{H/f})\) y-axis, third and fourth columns). The graphs show strong positive correlations between foot duration and the distance of H from the final foot boundary (FootDur vs. \(t2_{H/f}\)) but not between foot duration and the distance of H from the beginning of the foot (FootDur vs. \(t1_{H/f}\)).

To determine the significance of the effect of FootDur on \(t1_{H/f}\) and \(t2_{H/f}\), we fitted LME Models including Speaker and Sentence as random factors and FootDur as a covariate using either \(t1_{H/f}\) (Model 2a) or \(t2_{H/f}\) (Model 2b) as a dependent variable. Table 11 lists \(F\) values and estimates of \(\Omega^2\). Significant effects of FootDur on both dependent variables are found in all varieties except \(t1_{H/f}\) in ZB. As in the case of WordDur, estimates of \(\Omega^2\) reveal that FootDur predicts \(t2_{H/f}\) much better than \(t1_{H/f}\). In the latter case, FootDur only marginally improves model fit (all \(\Omega^2 \leq 0.1\)). These results suggest that our speakers used **Strategy 3**, except for ZB speakers who used **Strategy 1**. To account for the significant effects of StDist in both Model 2a and 2b, we may concede that there are effects of StDist on the location of H relative to both the beginning and end of the nuclear foot, such that increasing the duration of the foot moves H proportionally rightwards. As was observed for word duration in Section 3.2.1, however, the rightward movement of H is much smaller than the movement of the final foot boundary.

Table 11: FootDur as a predictor of \(t1_{H/f}\) and \(t2_{H/f}\).

|       | \(\Omega^2\) | \(\Omega^2\) |
|-------|--------------|--------------|
| \(t1_{H/f}\) | \(t2_{H/f}\) | \(t1_{H/f}\) | \(t2_{H/f}\) |
| df   | \(F\) | \(\Omega^2\) | df   | \(F\) | \(\Omega^2\) |
| ZB   | 1,239 | 1.53 | 0.00 | 1,239 | 3,459.61** | 0.94 |
| RO   | 1,268 | 12.08** | 0.03 | 1,268 | 3,783.04** | 0.94 |
| AM   | 1,202 | 14.48*** | 0.05 | 1,202 | 1,747.77*** | 0.90 |
| GR   | 1,371 | 41.45*** | 0.09 | 1,371 | 2,549.39*** | 0.88 |
| WI   | 1,268 | 15.32*** | 0.05 | 1,268 | 2,336.91*** | 0.90 |
| WL   | 1,267 | 33.48*** | 0.10 | 1,267 | 3,119.58*** | 0.92 |
| WH   | 1,281 | 30.74*** | 0.08 | 1,281 | 3,115.64*** | 0.92 |

3.3 Alignment of L

3.3.1 Effects of the distance of the final word boundary

Figure 6 shows that the alignment of L1 is very similar to the alignment of H when WbDist is varied. Adding one more syllable while keeping the distance to
the following lexical stress constant hardly changes the L1-latency either from the preceding H or from the beginning of the nuclear word. One exception is ZB, where L1 occurs about 70 ms earlier in monosyllabic words (WbDist = 0) than in disyllabic words (WbDist = 1) for StDist = 1. We also note that in AM and GR the timing of L1 is rather variable. In no variety does L1 move with the final word boundary, with the exception of AM when StDist = 2.

To account for the effect of WbDist on the location of L1, we fitted two LME Models. According to Strategy 1, we should expect that increasing WbDist increases the distance of L1 from the final word boundary, which we used as our first dependent variable ($t_{2L/w}$) (Model 1a). According to Strategy 2, we expected that increasing WbDist increases the latency from a preceding landmark. We chose H as our landmark, such that the duration of the nuclear fall was the dependent variable of the second model ($t_{1L/w}$) (Model 1b). In both models, we used Speaker and Sentence as random factors and WbDist as a fixed factor. Again, separate tests were performed for StDist = 1 and StDist = 2.

Table 12 shows significant effects of WbDist on $t_{2L/w}$ in all varieties and for both levels of StDist, except for AM when StDist was 2. Significant effects of WbDist on $t_{1L/w}$ were found for ZB and GR when StDist was 1 and for RO, AM, and GR when StDist was 2. Taking Strategy 1 to be used when WbDist significantly predicts $t_{2L/w}$ but not $t_{1L/w}$, Strategy 2 when WbDist significantly predicts $t_{1L/w}$ but not $t_{2L/w}$, and Strategy 3 when WbDist significantly predicts both $t_{1L/w}$ and $t_{2L/w}$

|     | $t_{1L/w}$ |   | $t_{2L/w}$ |   |
|-----|-----------|---|-----------|---|
|     | df | F  | $Q^2$ | df | F   | $Q^2$ |
| ZB  | 1  | 1,96 | 27.74*** | 0.23 | 1,82 | 86.76*** | 0.50 |
| RO  | 1  | 1,83 | 0.27  | 0.00 | 1,84 | 105.13*** | 0.53 |
| AM  | 1  | 1,56 | 2.5   | 0.05 | 1,56 | 50.76***  | 0.48 |
| GR  | 1  | 1,101 | 9.20** | 0.09 | 1,101 | 77.26***  | 0.44 |
| WI  | 1  | 1,81 | 3.91  | 0.04 | 1,80 | 132.86*** | 0.61 |
| WL  | 1  | 1,78 | 0.07  | 0.00 | 1,76 | 134.55*** | 0.64 |
| WH  | 1  | 1,78 | 0.29  | 0.00 | 1,74 | 188.42*** | 0.71 |
| ZB  | 2  | 1,63 | 0.35  | 0.00 | 1,65 | 6.47*     | 0.09 |
| RO  | 2  | 1,65 | 4.70* | 0.04 | 1,64 | 15.68***  | 0.22 |
| AM  | 2  | 1,59 | 23.26*** | 0.30 | 1,59 | 0.37     | 0.01 |
| GR  | 2  | 1,106 | 14.07*** | 0.12 | 1,106 | 53.57***  | 0.33 |
| WI  | 2  | 1,76 | 0.07  | 0.00 | 1,71 | 26.38***  | 0.26 |
| WL  | 2  | 1,72 | 0.10  | 0.00 | 1,76 | 88.84***  | 0.52 |
| WH  | 2  | 1,83 | 0.32  | 0.00 | 1,81 | 205.41*** | 0.71 |
and $t_2L/w$ (see Section 1), we may conclude that **Strategy 1** was used by speakers of RO, AM, WI, WL, and WH when StDist was 1, and by speakers of ZB, WI, WL, and WH when StDist was 2; **Strategy 2** was used by speakers of AM when StDist was 2; and **Strategy 3** was used by speakers of ZB and GR when StDist was 1 and by speakers of RO and GR when StDist was 2. Estimates of $\Omega^2$ show that in those cases where **Strategy 3** was used, variation of WbDist was accompanied by some proportional movement of L1, but L1 moved rightwards to a lesser extent than the final word boundary when WbDist was increased.

The scatter plots in Figure 10 show for each variety the relation between WordDur (x-axis) and the distance of L1 from H ($t_1L/w$, y-axis, first and second columns) and the distance of L1 from the final word boundary ($t_2L/w$, y-axis, third and fourth columns). The graphs reveal strong positive correlations between word duration and distance of L1 from the final word boundary (WordDur vs. $t_2L/w$) but not between word duration and distance of L1 from H (WordDur vs. $t_1L/w$).

To determine the significance of the effect of WordDur on $t_1L/w$ and $t_2L/w$ we fitted LME Models including Speaker and Sentence as random factors and WordDur as a covariate using either $t_1L/w$ (Model 2a) or $t_2L/w$ (Model 2b) as a dependent variable. Table 13 lists $F$ values and estimates of $\Omega^2$. Significant effects of WordDur on $t_2L/w$ were found for all varieties and on $t_1L/w$ for ZB, RO, WI, WL, and WH. These results suggest that speakers of AM and GR used **Strategy 1** and that speakers of all other varieties used **Strategy 3**. For **Strategy 3** the estimates of $\Omega^2$ suggest that moving the final word boundary rightwards increased the distance of L1 from the final word boundary much more than the distance from the beginning of the word.

### 3.3.2 Effects of the distance to the following lexical stress

Figure 8 shows that overall the alignment of L1 is similar to the alignment of H under different StDist conditions. Adding one more syllable to the nuclear foot while keeping the distance of the accented syllable to the final word boundary constant does not have a large effect on the distance of L1 from both the preceding H and the beginning of the nuclear foot. An exception is the monosyllabic word in ZB, where L1 occurs earlier when StDist = 1. Also, AM and, to a lesser extent, GR show substantial variation of the timing of L1.

To account for the effect of StDist on the location of L1, we fitted two LME Models each for the three levels of distance of word boundary, WbDist = 0, WbDist = 1, and WbDist = 2. Table 14 shows significant effects of StDist on $t_2L/f$ in all varieties, except for AM when WbDist was 0. A significant effect of StDist
Figure 10: Scatterplots of word duration vs. $t_{1w}$ (first and second columns) and vs. $t_{2w}$ (third and fourth columns).
Table 13: Effects of WordDur on $t_{1/L/w}$ and $t_{2/L/w}$.

|   | $t_{1/L/w}$ | $t_{2/L/w}$ |
|---|-------------|-------------|
|   | df | $F$ | $\Omega^2$ | df | $F$ | $\Omega^2$ |
| ZB | 1,247 | 33.81*** | 0.12 | 1,247 | 425.98*** | 0.64 |
| RO | 1,265 | 8.41** | 0.03 | 1,265 | 539.47*** | 0.68 |
| AM | 1,211 | 2.27 | 0.00 | 1,211 | 173.78*** | 0.47 |
| GR | 1,367 | 1.87 | 0.00 | 1,367 | 532.37*** | 0.59 |
| WI | 1,273 | 20.56*** | 0.06 | 1,273 | 773.32*** | 0.74 |
| WL | 1,270 | 20.58*** | 0.06 | 1,270 | 927.77*** | 0.78 |
| WH | 1,277 | 15.20*** | 0.05 | 1,277 | 1,676.06*** | 0.86 |

Table 14: Effects of StDist on $t_{1/L/f}$ and $t_{2/L/f}$.

|   | $t_{1/L/f}$ | $t_{2/L/f}$ |
|---|-------------|-------------|
|   | WbDist | df | $F$ | $\Omega^2$ | df | $F$ | $\Omega^2$ |
| ZB | 0 | 1,70 | 69.06*** | 0.50 | 1,70 | 215.47*** | 0.74 |
| RO | 0 | 1,73 | 7.45** | 0.08 | 1,73 | 4.03* | 0.05 |
| AM | 0 | 1,56 | 8.67** | 0.14 | 1,56 | 0.00 | 0.00 |
| GR | 0 | 1,110 | 1.38 | 0.01 | 1,109 | 43.07*** | 0.28 |
| WI | 0 | 1,84 | 6.93** | 0.07 | 1,86 | 130.51*** | 0.60 |
| WL | 0 | 1,75 | 0.20 | 0.00 | 1,70 | 87.36*** | 0.54 |
| WH | 0 | 1,78 | 7.51** | 0.09 | 1,75 | 249.98*** | 0.74 |
| ZB | 1 | 1,69 | 1.09 | 0.01 | 1,69 | 83.95*** | 0.55 |
| RO | 1 | 1,74 | 3.49 | 0.04 | 1,86 | 66.71*** | 0.44 |
| AM | 1 | 1,61 | 3.25 | 0.06 | 1,57 | 62.24*** | 0.53 |
| GR | 1 | 1,98 | 15.85*** | 0.13 | 1,97 | 23.04*** | 0.19 |
| WI | 1 | 1,81 | 1.64 | 0.01 | 1,75 | 128.35*** | 0.63 |
| WL | 1 | 1,72 | 1.08 | 0.01 | 1,73 | 97.85*** | 0.57 |
| WH | 1 | 1,81 | 2.02 | 0.02 | 1,78 | 108.83*** | 0.58 |
| ZB | 2 | 1,65 | 0.35 | 0.00 | 1,62 | 51.01*** | 0.47 |
| RO | 2 | 1,74 | 2.50 | 0.03 | 1,75 | 30.80*** | 0.30 |
| AM | 2 | 1,59 | 12.72** | 0.17 | 1,60 | 59.86*** | 0.50 |
| GR | 2 | 1,104 | 8.50** | 0.07 | 1,104 | 23.60*** | 0.19 |
| WI | 2 | 1,70 | 0.08 | 0.00 | 1,66 | 96.82*** | 0.60 |
| WL | 2 | 1,78 | 0.14 | 0.00 | 1,76 | 45.70*** | 0.37 |
| WH | 2 | 1,80 | 0.85 | 0.01 | 1,78 | 43.03*** | 0.36 |
on $t_{1L/f}$ was found for ZB, RO, AM, WI, and WH when WbDist was 0, for GR when WbDist was 1, and for AM and GR when WbDist was 2. These results suggest that **Strategy 1** was used by speakers of GR and WL when WbDist was 0, by speakers of ZB, RO, AM, WI, WL, and WH when WbDist was 1, and by speakers of ZB, RO, WI, WL, and WH when WbDist was 2; **Strategy 2** was used by speakers of AM when WbDist was 0; and **Strategy 3** was used by speakers of ZB, RO, WI, and WH when WbDist was 0, by speakers of GR when WbDist was 1, and by speakers of AM and GR when WbDist was 2. Estimates of $\Omega^2$ suggest that $t_{2L/f}$ increased much more than $t_{1L/f}$ when **Strategy 3** was used. As in the case of using WbDist as a predictor variable (see Section 3.3.1), AM speakers were found to use **Strategy 2** when WbDist was 0, but the amount of variance reduced by StDist was small ($\Omega^2 = 0.14$).

We conclude that all varieties examined, except AM for WbDist = 0, used **Strategy 1** or **Strategy 3**. Both strategies resulted in timing patterns where L1 was not, or not substantially, delayed when the distance of the first postnuclear stress from the nuclear stress was increased.

Figure 11 shows scatter plots for foot duration (x-axis) and the distance of L1 from the beginning of the foot ($t_{1L/f}$, y-axis, first and second columns) as well as the distance of L1 from the final foot boundary ($t_{2L/f}$, y-axis, third and fourth columns). The graphs reveal strong positive correlations between foot duration and the distance of L1 from the final foot boundary (FootDur vs. $t_{2L/f}$), but not between foot duration and the distance of L1 from the preceding H (FootDur vs. $t_{1L/f}$).

To determine the significance of the effect of FootDur on $t_{1L/f}$ and $t_{2L/f}$, we fitted LME models including Speaker and Sentence as random factors and FootDur as a covariate using either $t_{1L/f}$ (Model 2a) or $t_{2L/f}$ (Model 2b) as a dependent variable. Table 15 shows significant effects of FootDur on both dependent variables in all varieties except AM when $t_{1L/f}$ was the dependent variable. Estimates of $\Omega^2$ reveal that FootDur predicts $t_{2L/f}$ much better than $t_{1L/f}$. In the latter case, FootDur only marginally improved model fit. These results suggest that speakers of all varieties used **Strategy 3**, except the speakers of AM who used **Strategy 1**. Again, estimates of $\Omega^2$ suggest that even when **Strategy 3** was adopted, L1 was rather stably aligned relative to the preceding H. Using FootDur to predict $t_{1L/f}$ removed only a small part of the variance of residuals ($\Omega^2 \leq 0.06$).

### 3.4 Cross-linguistic variation

In Sections 3.2 and 3.3 we carried out separate analyses for each level of Dialect, as we were interested in the timing strategies of individual varieties. In this section, we add some information on this cross-linguistic variation based on direct comparisons of the varieties included in our corpus. The comparisons of the nuclear words and feet produced in different experimental conditions in Section
Figure 11: Scatterplots of foot duration vs. $t_{1L/f}$ (first and second columns) and vs. $t_{2L/f}$ (third and fourth columns).
Table 15: Effects of FootDur on $t_{L/f}$ and $t_{2L/f}$

|        | $t_{L/f}$ |        | $t_{2L/f}$ |
|--------|-----------|--------|------------|
|        | $df$      | $F$    | $Ω^2$      | $df$      | $F$    | $Ω^2$ |
| ZB     | 1,248     | 16.63***| 0.06       | 1,248     | 599.10***| 0.71 |
| RO     | 1,268     | 10.01** | 0.04       | 1,268     | 643.65***| 0.71 |
| AM     | 1,202     | 0.69    | 0.00       | 1,202     | 259.54***| 0.56 |
| GR     | 1,367     | 5.97*   | 0.02       | 1,367     | 590.52***| 0.62 |
| WI     | 1,280     | 7.99**  | 0.02       | 1,280     | 1,140.16***| 0.81 |
| WL     | 1,269     | 18.19***| 0.03       | 1,269     | 1,106.83***| 0.81 |
| WH     | 1,277     | 9.80**  | 0.03       | 1,277     | 1,897.66***| 0.87 |

3.1 have shown that word and foot duration differed only marginally across varieties. The picture changes when we look at the timing of the accentual gesture. Figure 12 shows the alignment of L0 and H relative to the beginning of the nuclear word and foot (L0 Delay and H Delay) and of L1 relative to H (L1 Delay). For all variables we observe inverted U-shaped patterns, with a later occurrence of the three tonal targets in the more ‘central’ varieties, AM, GR, RO, and, to a lesser extent, WI than in the more ‘peripheral’ varieties, ZB, WL, and WH. These data suggest that the whole rising-falling $F_0$-gesture occurs later when we approach the geographical midpoint of our research area.5

To assess the effect of the linguistic background of our speakers on the alignment of L0, H, and L1, we fitted LME models for the three dependent variables L0 Delay, H Delay, and L1 Delay, including Speaker and Sentence as random factors and Dialect as a fixed factor in each case. Dialect had a significant effect on L0 Delay, $F$

---

Figure 12: L0 Delay, H Delay, and L1 Delay in each variety.

---

5 A similar pattern was found for the size of the rising movement suggesting that larger latencies of nuclear peaks correlate with larger rise sizes. In Weener speakers, higher beginnings of the rises were found to contribute to smaller rise sizes.
Table 16: Pairwise comparisons of levels of Dialect for the dependent variables L0 Delay, H Delay, and L1 Delay (levels of significance: 0.05, 0.01, and 0.001).

|       | 1 vs. 2 | 2 vs. 3 | 3 vs. 4 | 4 vs. 5 | 5 vs. 6/7 | 6 vs. 7 |
|-------|---------|---------|---------|---------|-----------|---------|
| L0 Delay |         |         | ***/*** |         |           | ***/*** |
| H Delay  | *       | ***     |         | ***     | ***/***   | ***/*** |
| L1 Delay | ***     | ***     | ***     | ***     |           |         |

(6,163) = 14.67, p < 0.001; H Delay, F(6,135) = 21.34, p < 0.001; and L1 Delay, F(6,138) = 15.70, p < 0.001. Table 16 shows pairwise comparisons of levels of Dialect. The distribution of significant differences between levels is in line with the overall inverted U-shaped patterns observed in Figure 12. Most significant differences were found between the more ‘peripheral’ varieties, ZB, WL, and WH, on the one hand and the more ‘central’ varieties on the other. For L0 Delay only those differences between levels that involved either WL or WH\(^6\) reached significance.

---

6 Figure 6 showed that L0 was aligned earlier in WL and WH than in the other varieties. An analysis of data from the same speakers in another experiment (Peters et al. 2014) has shown that the earlier alignment of L0 (and H\(^*\)) in the Weener speakers correlates with a higher beginning of the nuclear rising movement. While the early beginnings of the rise can partly be explained by a more frequent occurrence of prenuclear high and rising accents in Weener data, earlier occurrences also occur when no prenuclear accents are present.
4 Discussion

4.1 Alignment of H

The nuclear $F_0$ peak, $H$, kept a fairly constant latency from the beginning of the nuclear word and foot when the distance between the nuclear syllable and the final word boundary ($WbDist$) or the following postnuclear stress ($StDist$) was varied. Similar results were obtained when continuous variables such as the overall word duration ($WordDur$) or duration of the nuclear foot ($FootDur$) were used as predictors of the location of $H$.

Statistical analyses revealed that speakers of all varieties either used Strategy 1 or Strategy 3. Strategy 1 was defined as exclusively increasing the $H$-distance from the final word or foot boundary when word or foot size is increased, respectively. Strategy 3 was defined as increasing both the $H$-distance from the final word or foot boundary and the initial word or foot boundary when word or foot size is increased, respectively. Under Strategy 3, delaying the final word or foot boundary is accompanied by a proportional delay of $H$ in the same direction. Estimates of the variance of residuals explained, $\Omega^2$, however, showed that even in varieties using Strategy 3, this effect is fairly small. A comparison of the $\Omega^2$ estimates for $t_{1H/w}$ and $t_{2H/w}$ in Tables 8 and 9 and of $t_{1H/f}$ and $t_{2H/f}$ in Tables 10 and 11 shows that the amount of variance of residuals explained in predicting $t_{2H/w}$ or $t_{2H/f}$ was at least 5 times larger than when predicting $t_{1H/w}$ or $t_{1H/f}$, respectively. We conclude that in the varieties examined here, the size of the nuclear word or foot has little or no effect on the timing of the nuclear $F_0$ peak relative to the beginning of the nuclear word and foot.

This finding is largely in line with the findings of Schepman et al. (2006) on Standard Dutch, who found no general effect of the size of the nuclear word for $WbDist = 0$ and $WbDist = 1$, and only a small effect for $StDist = 1$ when compared with $StDist = 0$. Ladd et al. (2009) found an effect of stress distance on nuclear accentual peaks in speakers of English (RP and Scottish Standard English). In this case, however, the target words occurred in final or prefinal position of the utterance, such that tonal crowding effects arising from the presence of IP-final boundary tones may have come into play.

4.2 Alignment of L1

L1 was found to be aligned at a fairly fixed distance from the preceding $H$ when $WbDist$ or $StDist$ was increased stepwise. Again, statistical analyses suggested that either Strategy 1 or Strategy 3 was adopted, and again the outcome in adopting
these strategies was very similar. Even when Strategy 3 was adopted, increasing the size of the word or foot had only a small effect on the distance of L1 from the preceding high target. One exception was AM, which was found to adopt Strategy 2 when WbDist was increased from 1 to 2 syllables (for StDist = 2; cf. Table 12) and when StDist was increased from 1 to 2 syllables (for WbDist = 0; cf. Table 14). In these cases, L1 was kept at a rather fixed distance from the end of the nuclear word or foot rather than from the beginning of the nuclear word or foot, as illustrated by the upper half of the panel for AM in Figure 6 and the lower third of the panel for AM in Figure 8. Using the absolute durations of the nuclear word and foot as predictor variables and including data of all levels of WbDist and StDist, all varieties except AM and GR for WordDur and AM for FootDur were found to adopt Strategy 3 (see Tables 13 and 15). Again, however, the effects on the distance of L1 from the beginning of the word and foot were small. Estimates of $\Omega^2$ for $t_{2L/w}$ and $t_{2L/f}$ were more than five times larger than for $t_{1L/w}$ and $t_{1L/f}$, respectively. Moreover, with foot duration as the dependent variable, speakers were no longer found to adopt Strategy 2.

Overall, the timing patterns of L1 were very similar to those of H in most varieties. L1, like H, was fairly stably aligned with the beginning of the preceding landmark, with a tendency to proportional timing in some varieties. These findings are fully in line with those by Barnes et al. (2010) for American English, which, unlike Barnes et al. 2006, found the alignment of L1 relative to H to be unaffected by the distance of the word boundary and first postnuclear stress. A comparison of estimates of $\Omega^2$ for $t_{2H/w}$, $t_{2H/f}$, $t_{2L/w}$, and $t_{2L/f}$ for H and L1 shows that the amount of variance explained when predicting the alignment of L1 was smaller than when predicting the alignment of H. One possible explanation may be that in the case of L1 a tonal target rather than a segmental landmark was used as point of reference. Therefore, we carried out additional analyses not reported here, in which we used segmental landmarks like the beginning of the nuclear word and foot or the end of the accented syllable as points of reference for the alignment of L1. These analyses provided the same overall patterns.7

---

7 Note that the correlation coefficients reported by Barnes et al. (2010) are not directly comparable to our $r^2$ values used as estimates of $\Omega^2$. In contrast to our estimates of $\Omega^2$, Pearson’s correlation coefficients are based on the data distributions presented in Figures 5, 7, 8, and 9 without taking account of inter-speaker variation and interdependencies between recorded sentences, which in our null model are accounted for by including Speaker and Sentence as random factors.
4.3 Cross-linguistic variation

Whereas comparisons of the overall length of the nuclear word and foot in Section 3.1 did not yield substantial cross-linguistic variation, except for WL (see below), comparisons of the timing of L0, H, and L1 in Section 3.4 have shown that in the more ‘central’ varieties, RO, AM, GR, and WI, the overall accentual gesture is delayed when compared to the ‘peripheral’ varieties, ZB, WL, and WH. Accordingly, we observed inverted U-shaped patterns in Figure 12, with the largest values for L0 Delay, H Delay, and L1 Delay being found in AM and GR. Cross-linguistic differences in the timing of the accentual gesture have been found for a number of languages and dialects for both nuclear and prenuclear accents (e.g., Atterer and Ladd 2004; Ladd et al. 2009), but we did not expect gradual variation as it is manifested by the observed inverted U-shaped pattern, which suggests that geographical distance matters more in this area than historical-grammatical origin of the languages involved (for similar findings see Peters et al. 2014).

We also observed that AM and GR most often adopted Strategy 3 for the timing of both H and L1, and AM was the only variety that adopted Strategy 2 for the timing of L1 (see Table 12). Strategy 2 and Strategy 3 have in common that increasing word or foot size increases H Delay or L1 Delay. The word and foot boundaries in these varieties might conceivably have acted as stronger attractors for H and L1 because of smaller absolute distances between the prosodic boundaries and the tonal targets when compared to other varieties (cf. Schepman et al. 2006). In that case, however, we should expect stronger effects of WbDist and StDist for smaller values of WbDist and StDist, where H and L1 occur closer to the prosodic boundaries, which was not the case.

We also observed that the timing of L1 was generally more variable in AM than in all other varieties, which may be explained by the higher number of $F_0$ contours that could not easily be classified as either ordinary high falls or as ‘late peak’ modification of the rising-falling accent (Ladd 1983), the late C-fall instead of the early A-fall of ‘t Hart et al. (1990). The latter is typically realized in AM with a high plateau, ranging from the nuclear syllable to the last stress in the intonational phrase. In Figure 13, we illustrate the ordinary fall, the ‘late

Figure 13: Stylized $F_0$ contours attested for AM: (a) ordinary nuclear fall, (b) regional variant of the ‘late peak’ modification of the nuclear accent, and (c) instance of an $F_0$ contour that could not easily be classified as either (a) or (b).
peak’ modification with a high plateau, and an $F_0$ contour that could not easily be classified as either one of these. Note that the classification of these contours as instances of ordinary falls yielded very large values for L1 Delay.

Finally, the question arises of why the nuclear word was found to be longer in WL than in ZB and RO in Section 3.1, and why both the nuclear word and foot were found to be longer in WL than in WH, which data were recorded by the same speakers. As for WL and WH, there are two factors which may have contributed to these differences. First, Weener speakers were less familiar with Low Saxon as a written language than with High German. As a result, they may have read the Low Saxon sentences less fluently than the High German ones, which in turn may have led to overall differences in speech rate. Second, closer inspection of Figure 4 reveals that the longer overall durations of the nuclear word in Weener Low German may be particularly due to longer durations of target words with WbDist = 2, that is Molberen, Lumberen, and Melberen. In unpublished research the same speakers were found to realize target words of the Molberen type as Mol. be:rn ([mʊl.ɛːɐn]) rather than as Mol.be.ren ([mʊl.ə.ɾɛn]) in nearly 50% of the cases when they read the Low Saxon versions of the test sentences, but in less than 25% of the cases when they read the High German versions. Even if word forms like Mol.be:rn contain one syllable less than words like Mol.be.ren, the longer duration of the full-voweled second syllable increases the overall duration of both the nuclear word and the nuclear foot.

4.4 Implications for phonological theory

According to classical ToBI, trailing tones are tones that occur at a fixed distance from the preceding starred tone, whereas ‘phrase accents’ are timed independently. Grice et al. (2000) defined the ‘phrase accent’ as a boundary tone that seeks an association with a postnuclear stressed syllable. If L1 were in fact attracted to a postnuclear stress in languages like English, German, or Dutch, it would be reasonable to adopt the ToBI representation of the nuclear fall as H* L-L%, originating from Pierrehumbert (1980).

The attraction of L1 to a postnuclear stress was apparently supported by smaller-scale studies like Benzmüller and Grice (1998) for German and Barnes et al. (2006) for American English. In the data reported here, the position of the first postnuclear lexical stress had no appreciable effect on the timing of either H or the end of the fall in seven varieties of coastal continental West Germanic.8

8 One well-known issue in research on the timing of nuclear falls is the determination of L1 in the $F_0$ trajectory. As reported in Section 2.4, the reliability test provided a mean absolute
Together with the findings by Barnes et al. (2010), our results provide evidence that the timing of L1 is not substantially affected by the position of the following lexical stress, which from the view of classical ToBI and its further developments in Grice et al. (2000) argues in favour of representing L1 as a trailing tone rather than a ‘phrase accent’. While this representation is consistent with that proposed in Gussenhoven (1991, 2005), which is a bitonal pitch accent H*L, his analysis of a tone as either trailing or phrasal was not primarily guided by constant latencies of tonal targets within the pitch accent, but by contextual and language-specific implementation rules. Specifically, for Dutch, English, and German (cf. Peters 2014), L in a nuclear H*L is timed with reference to H*, but in a prenuclear H*L, it may be timed with reference to the first tone in a following pitch accent. Rather, L1 is regarded as part of the nuclear accent for structural and semantic reasons. For instance, in Dutch, the nuclear and pre-nuclear realizations of H*L appear functionally identical, both being typical of citation pronunciation. Moreover, the allophonic right-alignment of the trailing tone of prenuclear pitch accents is a general feature of the analysis and equally applies to trailing H. The decision to analyse the H-tone that associates with the last stressed syllable in the tonal dialect of Roermond Dutch as a boundary tone by Gussenhoven (2000) is based on the fact that the tone complex H_L consistently expresses either non-finality or interrogativity as well as on the fact that they appear without the pitch accent L* in unaccented syllables. Additionally, in this case their status as a complex boundary melody is indicated by the failure of the two tones making up the complex to be separated by a lexical tone on the final mora of the IP.

An alternative interpretation of these findings is presented by Barnes et al. (2010). While observing that the target of L is stably timed with the preceding H* target in American English, they suggest that this is due to the need for low F0 to occur immediately after the nuclear peak, in order for the contour to be distinct from one where prenuclear H* is followed by L*, where the slope down from H* deviation of 28 ms in the measurements of L1 by two independent transcribers. This amount of variation is comparable to that found in other studies using either manual or semi-automatic methods. Barnes et al. (2010) argue that determining the “relativised area under the curve” (R-AUC) rather than attempting to locate a single tonal target specified by L1 may be a more adequate approach in assessing differences in the timing of the nuclear falls. Note, however, that Barnes et al. (2010) applied their analysis to nuclear falling-rising contours (H* L-H% or H*L H%) rather than to nuclear falls (H* L-L% or H*L L%). In falling-rising contours, dispensing with the location of the end of the nuclear fall (L1) results in the calculation of the area covered by the falling F0 trajectory between H and the second low target, which corresponds to the beginning of the final rise. It is not clear how this analysis could be transferred to simple falling contours where no phonetic evidence for a low target is available that corresponds to the low target at the beginning of the final rise in falling-rising contours.
may be more gradual. They suggest that, given their findings, L may still represent a phrase accent L- that finds an association somewhere in the post-nuclear stretch, but that the preservation of the contrast between H* L- and H* L* requires speakers to produce a sharp fall. There are two considerations that may be relevant here. First, there is no evidence for any syllabic association of L at all (Pierrehumbert 1980: 57); rather, the constant timing of L after H* suggests the left-alignment of an unassociated tone, where ‘alignment’ has the meaning of Prince and Smolensky (1993; cf. Gussenhoven 2000). In that meaning, a tone’s alignment specifies its location with respect to some other element in the phonological structure, independent of whether the tone associates in that location. Second, the existence of contrastive patterns cannot generally be used to argue for a non-phonological interpretation of the phonetic characteristics that are unique to one or the other member of a contrast. A perhaps somewhat crude analogy may illustrate this second point. If we were to argue that the short duration of the tongue glide in *roil* serves to distinguish it from *raw eel*, this would in no way support any independent, earlier position that *roil* was a disyllabic structure. At a minimum, we can maintain, therefore, that an interpretation of the post-H* low target as a phrase accent is not supported by the timing data.

One of the advantages of defining tone class without reference to detailed timing patterns becomes evident when comparing the phonetic realisation of closely related varieties that differ with respect to the timing of L1. Reference to timing patterns in distinguishing between H*+L L% and H* L-L% contours would imply that the nuclear falls produced by the same speaker are represented differently. As Table 14 shows, AM speakers adopted **Strategy 2** for the timing of L1 in sentences containing monosyllabic nuclear words (WbDist = 0), but **Strategy 1** in sentences containing disyllabic nuclear words (WbDist = 1). According to the ToBI-Grice et al. view, we would have to conclude that in the first case they used H* L-L% and in the second H*+L L%. There is, however, no other motivation for assuming that AM speakers used categorically different contours depending on whether WbDist was 0 or 1. They clearly did not express different meanings in the two contexts. The simpler assumption is that H*L L% was used in all cases, and that phonetic implementation rules show some sensitivity for the timing of L1 to an upcoming stress in words with final main stress.

## 5 Conclusions

The primary motivation for measuring the time alignments of the main turning points of nuclear rising-falling melodies in the speech of our speakers was to establish the extent to which the timing of this rising-falling gesture covaried...
with the distance of the nuclear syllable from the end of the accented word and from the stressed syllable in a following word. We were particularly interested in an evaluation of the claim in Grice et al. (2000) that the timing of the low target after the peak is determined by the location of the first post-nuclear stressed syllable. The theoretical interest within the ToBI framework lies in the interpretation of the low target as either a pronunciation of a trailing tone in a H*+L pitch accent or the pronunciation of a separate tonal category, a ‘phrase accent’ (L-). Because the earlier literature had suggested that the phrase accent L- aligns with the end of the nuclear word, we investigated L-distances to the righthand word boundary as well as to the beginning of the first post-nuclear stressed syllable. In none of the seven varieties did we find a general effect of the location of the upcoming word boundary or the upcoming stressed syllable that could support an interpretation of the low target, L1, as a pronunciation of a L- that associates with the post-nuclear stress. The critical measures for detecting alignment shifts were target latencies from the beginning and the end of the nuclear word, which were tested in separate analyses for the final word boundary and the location of the upcoming stressed syllable, independently for the separate levels of the other variable.

The investigation included the alignment of the peak before L1, notated as H. Significant rightward shifts in the alignment of H as a function of word length were detected in 4 out of the 14 cases (7 varieties × 2 levels for distance to upcoming stress), but the effects were negligibly small and occurred in the same dialect for both levels of the distance to the upcoming stress only in the case of GR. As for the effect of the upcoming stress, out of 21 cases (7 varieties × 3 levels for word length), 6 showed significant but negligible effects, in no case occurring in the same variety under more than one condition for word length. Five significant but small rightward shifts for L1 as a function of word length were found in the 14 relevant cases, 2 of them in GR. As for the effect of the distance to the upcoming stress, 8 out of 21 such cases were found. One of these represented a non-negligible leftward shift in the case of ZB in the monosyllabic word condition. Also, in two conditions there were effects in AM, in which data are the most variable.

These data do not support the conclusion that there is an L- in any of the varieties we have investigated. Within the philosophy of the ToBI framework, in which alignments are generally taken to be criterial features for morphological tone type, this means that the low target after the peak must be interpreted as a trailing L in a bitonal H*+L pitch accent. This conclusion may equally be drawn from the results obtained by Barnes et al. (2010) from 15 speakers of American English. The analysis is independently supported by earlier analyses of the West Germanic languages (Gussenhoven 1991, Gussenhoven 2004, Gussenhoven 2005; Féry 1993; Grabe 1998). In those analyses, the phonological nature of
intonational tones is, however, not generally determined by alignment properties, which may vary contextually for the same tone.

A reviewer challenged us to speculate on the causes of the small effects of word length and of the distance to the upcoming postnuclear stress on the alignment of H and L, however unsystematic these were. We agree that they require an explanation. Both effects have been entertained as plausible categorical features of West Germanic intonation. Also, we never found the reverse effects in our data. In general, we think they may be understood as reflexes of the general ergonomic context within which the phonologies of languages are embedded. The effect of increasing word length might be related to a tendency to assign melodic patterns to words, as opposed to phrases or syllables, a tendency that acquired grammatical status in many tone languages and more generally in languages in which word-like constituents are acceptable units and relatively short prosodic constituents like the accentual phrase have developed. On that assumption, the word-length effect is due to the variable availability of time for the articulation of tonal targets, as evidenced in research by Silverman and Pierrehumbert (1990) and Prieto et al. (1995). The tendency for tonal targets to gravitate towards the upcoming stressed syllable is less easily understood in this way. If we were to argue that stressed syllables are a natural locus for hyperarticulation, to be held responsible for their phonological complexity and faithfulness, we can explain why languages like Athenian Greek and Roermond Dutch (see Section 1) associate boundary tones with post-nuclear stressed syllables. An ergonomic benefit may be afforded by the synchronization of segmental and laryngeal features, but this would not explain why a gradual approach towards such favourable alignment points should bring any benefit at all. Leaving this topic for now, we would like to stress that the timing of the end of the fall suggests that this tone is left-aligned with the preceding H*-tone, where by ‘aligned’ we mean the edge-to-edge alignment of Optimality Theory (Prince and Smolensky 1993; Gussenhoven 2000). In this structurally specified location, however, the tone does not associate, since its detailed phonetic timing is in no way governed by syllable structure (Pierrehumbert 1980). As will be true for any phonological segment or tone, its duration and timing will depend on a variety of factors, of which distance between the accented syllable and the word end as well as an upcoming stressed syllable may play a small role, depending on the variety concerned.

Another reviewer challenge concerned the extent to which AM-theory can deal with pitch configurations in which tones are not represented by identifiable targets. By way of preamble, we first withdraw from a position that AM-theory constrains languages to having tone systems in which each tone is realized as a turning point in a pitch contour made up of linear interpolations between these points. A representation consisting of LHL may in one language be realized as a
rise-plateau-fall and in another as a slow-rise-plus-fall, and yet another as a rise-
plus-slow-fall. ToBI treats both aspects as language-specific features of phonetic
implementation. It rejects generalized linear interpolation, as in the case of L-,
which is realized as low-level pitch between H* and H%, or in the case of an H-
realized as mid-level pitch between L* and H%. It assumes unidentifiable tonal
targets in the high level nuclear tone, as in tomatoes, bananas, cucumbers..., where cucumbers may have no turning points, but is described with three tones,
H*, H-, and L%. Here, H- has the same pitch as H*, and L% has the same pitch
as H-, by convention. That is, the relation between the pitch contour and the
tonal analysis is mediated by conventions that specify how tones are realized in
specific contexts, much as allophonic rules specify pronunciations of segments.
If we assume that linear interpolations are part of the grammars of languages,
not universal phonetic realization rules, we may for a given language define the
contexts in which linear interpolations apply and the contexts in which tones
are realized continuously so as to fill up empty space between the targets of tones
they are flanked by. For Dutch, Gussenhoven (2005: 127) restricts linear interpola-
tions to tones within tonal morphemes, i.e., multitional pitch accents and the
initial boundary tone %HL. Between morphemes, the last tone continues its
pronunciation until the end of the intonational phrase or until the first tone of
the next morpheme, as in the case of trailing L in a nuclear H*L, which continues
until a final boundary H% is reached, or in the case of the level high contour
referred to above for cucumbers, which is described as H* without a following
boundary tone. The three contours in Figure 13 can in principle be described by
the same LHL tone sequence, but with different assumptions about the phonolo-
gical location of the tones (Gussenhoven 2000, Gussenhoven 2004: Ch 8). In
contour (a), the H left-aligns with L on its left and final L left-aligns with H. In
contour (b), final L right-aligns with the boundary, producing a linear interpola-
tion between the target of H and that of final L, while in contour (c), H aligns both
left and right, creating a level high stretch between the two high targets.

An important additional finding is that there appears to be a reflection of a
geographical cline in the alignment of peaks. Varieties spoken in the central
zone of the coastal arc have later alignments than those spoken at the periphery.
This finding concurs with the results of an investigation of contrastive focus
realization in these dialects reported in Peters et al. (2014). They found that
relative to the peripheral varieties, the central varieties lengthened nuclear
onsets and rimes in corrective and narrow focus conditions, while also delaying
the accent peak more. Since the central zone represents the prestigious heart-
land of varieties spoken in the Netherlands, the interpretation of this finding
would appear to be that the central zone is innovative. This assumption is
supported by the claim that, by and large, there are rightward shifts in language
development and by the preponderance of rightward H-spreading in tone languages (Hyman 2007). Zuid-Beveland and, to a lesser extent, Winschoten thus represent an older phase of the language.

Acknowledgments: This study was carried out as part of the project *Intonation in Varieties of Dutch*, funded by the *Netherlands Organisation of Scientific Research* (NWO), grant number 360-7-180. We thank Marron C. Fort and Garrelt van Borssum for translating the test sentences into Dutch and German Low Saxon, respectively. We also thank Rachel Fournier and Joop Kerkhoff for valuable practical and technical support, Wilbert Heeringa for statistical advice, and our research assistants Marjel van Dijk, Lian van Hoof, Jan Michalsky, and Renske Teeuw for their help with data collection and annotation. We are grateful to the two anonymous reviewers and Associate Editor Jon Barnes for their valuable comments.

Appendix

A.1 Speech materials

| Code | Language                |
|------|-------------------------|
| SD   | Standard Dutch (used for ZB, RO, and AM) |
| WF   | West Frisian (GR)       |
| DLS  | Dutch Low Saxon (WI)    |
| GLS  | German Low Saxon (WL)   |
| HG   | High German (WH)        |
| E    | English translation     |

As we did not plan to make recordings in Weener right from the beginning, the dialogues were not optimized for translations into Weener Low Saxon and High German, and some had to be modified to keep the metrical structure constant across languages. Translations of those alternative dialogues into English are added after slashes.

WbDist = 0, StDist = 0

1. SD A *Wat heb je tegen je kind gezegd?*  
B *Het moet zich tegen buurman Mol weren.*

WF A *Wat hast tsjin dyn bern sein?*  
B *It moat moarn nei buorman Mol harkje.*

DLS A *Wat hest doe tegen dokter van Holt zegd?*  
B *Hai mog in ‘t hoes van noaber Mol wonen.*
1. SD A Waarom kijk je zo geschrokken?
   B De buren gaan iets ergs over tante Mol beweren.

WF A Wêrom sjochto no sa skroekken?
   B De buorlju wolde it relaas oer tante Mol beharkje.

2. SD A Valt er hier nog wat te zien?
   B Er zal zo een parachutist in de wei van Boer de Lum landen.

WF A Krijt men hjir noch wat te sjen?
   B Der sil sa in parasjutist yn it lân fan Boer de Lum lânje.

DLS A Vaalt ter hier nog wat te zain?
   B Der zel zo n parachutist in de swoag van Boer de Lum lânne.

GLS A Gifft dat hier noch wat to sehn?
   B Daar sall so en malle Undögd up de Süster van Opa Lümm luaren.

HG A Gibt es noch was zu sehen?
   B Da soll so ein übler Kerl auf die Schwester von Opa Lümm lauern.

E A Is there anything to see here?
   B Soon a parachutist will be landing in the field of farmer Lum./
      They say that a bad guy ambushes the sister of grandpa Lum.

3. SD A De jonge ouders denken nog na over een naam.
   B Misschien kunnen ze hun dochter naar oma Mel noemen.

WF A De jonge âlden tinke noch nei oer in namme.
   B Wa wit kinne se har dochter nei beppe Mel neame.

DLS A De jonge ollers dinken nog noa over ‘n noam.
   B Meschain willen ze heur dochter Mel nuinen.

GLS A De junge Ollen denken noch over en Naam na.
   B Villicht können se hör Dochter na Oma Mell nömen.

HG A Die jungen Eltern denken noch über einen Namen nach.
   B Vielleicht können sie ihre Tochter nach Oma Mell nennen.

E A The new parents are still thinking about a name for their child.
   B They may want to call their daughter after grandmother Mel.

WbDist = 0, StDist = 1
1. **SD** A Waarom kickst du so verfehrt?
   B Umdat frömde Lü dat Huus van Oma Moll bewohnen.

**E** A Why are you looking so frightened?
   B The neighbours are about to say something nasty about Aunt Mol./
   The neighbours live in the house of Aunt Mol./
   Because strangers live in the house of grandma Moll.

2. **SD** A Waar ben je zo boos over?
   B De bal zal wel in de heg van buurman Lum belanden.

**WF** A Wêr bisto sa likk oer?
   B De bal sil wol op it hiem fan buorman Lum belânjie.

**DLS** A Waar bist doe zo kwoad over?
   B De baal zel wel in de heeg van noaber Lum belannen.

**GLS** A Waarover büst du so düll?
   B He wull uns achter d’ Heeg van Nahber Lümm beluren.

**HG** A Worüber bist du so böse?
   B Er wollte uns hinter der Hecke von Nachbar Lümm belauern.

**E** A What are you so angry about?
   B The ball will end up in the hedge of our neighbour Lum.

3. **SD** A In Hoevelaken zoeken ze een nieuwe burgemeester.
   B Ja, ze willen Rob van Mel benoemen.

**WF** A Yn Hoevelaken sykje se in nije boargemaster.
   B Ja, se wolle Rob van Mel beneame.

**DLS** A In Hoevelaken zuiken ze n nije börgermeester.
   B Joa, ze willen Rob van Mel benuimen.

**GLS** A In Esterwegen söken se en neje Buurmester.
   B Ja, as Kandidat willen se Jan van Mell benömën.

**HG** A In Esterwegen suchen sie einen neuen Bürgermeister.
   B Ja, als Kandidaten wollen sie Jan van Mell benennen.

**E** A They are looking for a new mayor in Hoevelaken/Esterwegen.
   B Yes, they want to appoint Rob/Jan van Mel.

---

**WbDist = 1, StDist = 1**

1. **SD** A Ik heb echt altijd hoofdpijn.
   B Dan moet je je niet tegen de pillen van dokter Molber weren.

**WF** A Ik haw no echt pineholle.
B Dan moatsto no goed nei it advys fan dokter Molber harkje.

DLS A Wat is ter?
B Johan wil in ’t hoes van dokter Molber wonen.

GLS A Wat is d’r?
B Johan will in ’t Huus van Dokter Molber wohnen.

HG A Was ist los?
B Johann will im Haus von Doktor Molber wohnen.

E A I really always have a headache./What is going on?
B Then you will have to pay close attention to what Doctor Molber will tell you./
Johann would like to live in Doctor Molber’s house.

2. SD A Wat is er gebeurd?
B Een vliegtuig wilde achter het huis van slachter Lumber landen.

WF A Wat is der no bard?
B In fleannasine woe achter it hûs fan slachter Lumber lânje.

DLS A Wat is ter gebeurd?
B ‘N vlaiger wol achter ‘t hoes van slachter Lumber lannen.

GLS A Wat is geböhrt?
B En Frömde wull achter dat Huus van Slachter Lümber luren.

HG A Was ist passiert?
B Ein Fremder wollte hinter dem Haus von Schlachter Lümber lauern.

E A What happened?
B An aircraft wanted to land behind butcher Lumbers’ house./
A stranger wanted to snoop around at the back of butcher Lumber’s house.

3. SD A Zijn er nog leuke roddels?
B Ja, mijn zus laat zich tegenwoordig mevrouw de Melber noemen.

WF A Binne der noch moaie nijtsjes?
B Ja, ús mem lit har tsjintwurich mefrou de Melber neame.

DLS A Binnen der nog orizinele proaterijen?
B Joa, mien zus let zok tegensworeg mevrouw de Melber nuimen.

GLS A Gifft dat noch nette Protere?
B Ja, mien Süster lett sük nu Frollein van Melber nömen.

HG A Gibt es noch nette Gerüchte?
B Ja, meine Schwester lässt sich jetzt Fräulein von Melber nennen.

E A Any fun gossip?
B Yes, my sister now wants to be known as Mrs. de Melber.
1. SD A Ik hoorde dat de school is afgebrand.
   B Ja, dat gaat burgermeester Molber beweren.
WF A Ik hearde dat de skoalle ôfbaarnd is.
   B De ferhalen sil boargemaster Molber beharkje.
DLS A Wat gaat ter gebeuren?
   B De bürgermeester gaat ’t hoes van meester Molber bewonen.
GLS A Wat is d’r?
   B Se laten mi neet dat Huus van Mester Molber bewohnen.
HG A Was ist los?
   B Sie lassen mich nicht das Haus von Lehrer Molber bewohnen.
E A I heard that the school burnt down./
   What is going on?
   B Yes, that is what Mayor Molber is going to claim./
   They do not allow me live in the house of teacher Molber.

2. SD A Waar maak je je druk om?
   B Ik wil niet nog eens bij meester Lumber belanden.
WF A Wêr makkest dy drok om?
   B Ik wol net noch ris by master Lumber belânje.
DLS A Woar moakst doe die drok om?
   B Ik wil nait nog ains bie meester Lumber belannen.
GLS A Wat wullt du van hum?
   B He sall neet noch maal Mester Lümber beluren.
HG A Was willst du von ihm?
   B Er soll nicht noch mal Lehrer Lümber belauern.
E A What are you worried about?/What do you want from him?
   B I don’t want to end up with teacher Lumber again./
   He mustn’t spy on teacher Lumber ever again.

3. SD A Nog nieuws uit het bestuur?
   B Ja, als nieuwe secretaris wil men Johan Melber benoemen.
WF A Noch nijs ût it bestjoer?
   B Ja, als nije sekretaris wol’ se Johan Melber beneame.
DLS A Nog nijs oet ’t bestuur?
   B Ja, als nije secretoaris wil men Johan Melber benuimen.
GLS A De een of anner Neeligheid ut de Vörstand?
   B Ja, as neje Sekretär willen se Johann Melber benömen.
HG A Irgendwelche Neuigkeiten aus dem Vorstand?
   B Ja, als neuen Sekretär will man Johann Melber benennen.
E A Is there any news from the board?
   B Yes, they want to appoint Johan Melber as the new secretary
1. SD  A  Wat gaat er gebeuren?  
      B  Ze willen de vossen uit het bos bij Molberen *welen*.  
WF  A  Wat is no dochs barre?  
      B  Se wolle nei de foksen út it bosk by Molberen *harkje*.  
DLS A  Wat is het probleem?  
      B  Oma Thijs mag nait in ‘t huus van tante Molberen *wonen*.  
GLS A  Wat is dat Problem?  
      B  Opa Theis dürt neet in ‘t Huus van Oma Molberen *wohnen*.  
HG  A  Was ist das Problem?  
      B  Opa Theis darf nicht im Haus von Oma Molberen *wohnen*.  
E   A  What’s going on?  
      B  They want to chase the foxes from the woods near Molberen.

2. SD  A  Wat is er gebeurd?  
      B  De piloot moest plotseling in ‘t dorpje Lumberen *landen*.  
WF  A  Wat is der bard?  
      B  De piloot moast ynienen yn ‘t doarpke Lumberen *länje*.  
DLS A  Wat is ter gebeurd?  
      B  De piloot mos pal in ‘t dörpke Lumberen *lannen*.  
GLS A  Waar büst du bang vör?  
      B  Daar kunn en Deev achter de Döör van Oma Lümberen *luren*.  
HG  A  Wovor hast du Angst?  
      B  Da könnte ein Dieb hinter der Tür von Oma Lümeren *lauern*.  
E   A  What happened?/What are you afraid of?  
      B  The pilot had to make a sudden landing in the village of Lumberen./  
          There might be a thief lurking behind Granny Lumberen’s door.

3. SD  A  Waarom wacht je bij de telefoon?  
      B  Simon zou mijn naam bij de Boer de Melberen*noemen*.  
WF  A  Wêrom wachtest by de telefoon?  
      B  Simon soe myn namme by de Boer de Melberen *neame*.  
DLS A  Woarom wachst doe bie de telefoon?  
      B  Simon zol mien noam bie de Boer de Melberen *nuinen*.  
GLS A  Waar liggt dat Problem?  
      B  Simon sall mi neet mehr Unkel Melberen *nömen*.  
HG  A  Wo liegt das Problem?  
      B  Simon soll mich nicht mehr Onkel Melberen *nennen*.  
E   A  Why are you waiting by the phone?/What's the problem?  
      B  Simon was going to mention my name to farmer de Melberen./  
      Simon shall no longer call me Uncle Melberen.
1. SD  A  Wat is het probleem?
     B  Iedereen gelooft wat ze over opa Molberen beweren.

WF  A  Wat is it probleem?
     B  Hast elkenien wit dat se pake Molberen beharkje.

DLS  A  Woar bist doe zo kwoad over?
      B  Ik mag nait ‘t huus van opa Molberen bewonen.

GLS  A  Waar liggt dat Problem?
      B  Ik will neet dat Huus van Opa Molberen bewohnen.

HG  A  Wo liegt das Problem?
      B  Ich will nicht das Haus von Opa Molberen bewohnen.

E  A  What’s the problem?/What are you so angry about?
     B  Everybody believes what they are saying about grandfather Molberen./
       I’m not allowed to live in grandpa Molberen’s house./
       I don’t want to live in grandpa Molberen’s house.

2. SD  A  Wat ben je onrustig! Wat is er toch?
     B  Ik wil niet meer in Café de Lumberen belandén.

WF  A  Wat bisto ünridejch! Wat is der oan?
     B  Ik wol net mear yn Café de Lumberen belânje.

DLS  A  Wat bist doe ongedureg! Wat is ter toch?
      B  Ik wil nait meer in Café de Lumberen belannen.

GLS  A  Waarum büst du so düll? Wat is d’r denn?
      B  He sall neet noch eenmaal Förster Lümberen beluren.

HG  A  Warum bist du so böse? Was ist los?
      B  Er soll nicht noch einmal Förster Lümberen belauern.

E  A  Why are you so restless? What’s going on?/Why are you so angry?
     B  I don’t want to end up in the Lumberen pub again./
       He mustn’t spy on forester Lumberen ever again.

3. SD  A  Wat lees je daar in de krant?
     B  De commissaris zal Joop de Melberen benoemen.

WF  A  Wat lësto dër yn de krante?
     B  De kommissaris sil Joop de Melberen beneam.

DLS  A  Wat leest doe doar in t nijsblad?
      B  De commissaris zel Joop de Melberen benuimen.

GLS  A  Wat giff dat Neeis?
      B  De Kommission sall Karl van Melberen benömen.

HG  A  Was gibt es Neues?
      B  Die Kommission soll Karl von Melberen benennen.
E  A  What is it you are reading in the paper?/What’s the news?
   B  The governor is about to appoint Joop de Melberen./
      The committee is about appoint Karl von Melberen.

References

Abercrombie, David. 1964. Syllable quantity and enclitics in English. In David Abercrombie, D. Fry, P. MacCarty, N. Scott, & J. Trim (eds.), In honour of daniel jones. Papers contributed on the occasion of his eightieth birthday 12 September 1961, 216–222. London: Longmans.

Arvaniti, Amalia. 2002. The intonation of yes-no questions in Greek. In Marianthi Makri-Tsiliipakou (ed.), Selected papers on theoretical and applied linguistics, 71–83. Thessaloniki: Department of Theoretical and Applied Linguistics, School of English, Aristotle University.

Arvaniti, Amalia, D. Robert Ladd & Ineke Mennen. 1998. Stability of tonal alignment: The case of Greek prenuclear accents. Journal of Phonetics 26. 3–25.

Atterer, Michaela & D. Robert Ladd. 2004. On the phonetics and phonology of ‘segmental anchoring' of F0: Evidence from German. Journal of Phonetics 32. 177–197.

Barnes, Jonathan, Stefanie Shattuck-Hufnagel, Alejna Brugos & Nanette Veilleux. 2006. The domain of realization of the L- phrase tone in American English. Proceedings of Speech Prosody 2006, Dresden, May 2–5, 2006.

Barnes, Jonathan, Nanette Veilleux, Alejna Brugos & Stefanie Shattuck-Hufnagel. 2010. Turning points, tonal targets, and the English L- phrase accent. Language and Cognitive Processes 25. 982–1023.

Beckman, Mary E. & Gayle Ayers Elam. 1997. Guidelines for ToBI labelling (version 3.0, March 1997). The Ohio State University Foundation. [http://ling.ohio-state.edu/Phonetics/ToB1/tobi_homepage.htm]

Beckman, Mary E., Julia Hirschberg & Stefanie Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In Sun-Ah Jun (ed.), Prosodic typology: The phonology of intonation and phrasing, 9–54. Oxford: Oxford University Press.

Benzmüller, Ralf & Martine Grice. 1998. The nuclear accentual fall in the intonation of Standard German. ZAS Papers in Linguistics: Papers on the conference “The word as a phonetic unit”, 79–89. Berlin, Germany.

Boersma, Paul & David Weenink. 2008. Praat: doing phonetics by computer (Version 5.0.25) [computer program]. [http://www.praat.org] (accessed 31 May 2008).

Caspers, Janneke & Vincent J. van Heuven. 1993. Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall. Phonetica 50. 161–171.

Dalton, Martha & Ailbhe Ní Chasaide. 2007. Melodic alignment and micro-dialect variation in Connemara Irish. In Carlos Gussenhoven & Tomas Riad (eds.), Tones and tunes, vol. 2: Experimental studies in word and sentence prosody, 293–315. Berlin and New York: Mouton de Gruyter.

del Giudice, Alex, Ryan Shosted, Kathryn Davidson, Mohammad Salihie & Amalia Arvaniti. 2007. Comparing methods for locating pitch ‘elbows'. Proceedings of ICPHS XVI. 117–1120.

Féry, Caroline. 1993. German intonational patterns. Tübingen: Niemeyer.
Frota, Sonja. 2002. Tonal association and target alignment in European Portuguese nuclear falls. In Carlos Gussenhoven & Natasha Warner (eds.), Papers in laboratory phonology 7, 387–418. Berlin and New York: Mouton de Gruyter.

Gili Fivela, Barbara. 2002. Tonal alignment in two Pisa Italian peak accents. In Bernard Bel & Isabelle Marlien (eds.), Proceedings of Speech Prosody 2002, Aix-en-Provence, 11–13 April 2002, 339–342.

Gilles, Peter. 2005. Regionale Prosodie im Deutschen. Variabilität in der Intonation von Abschluss und Weiterweisung. Berlin and New York: de Gruyter.

Grabe, Esther. 1998. Comparative intonational phonology: English and German (MPI series in psycholinguistics 7). Ponsen en Looien: Wageningen.

Grice, Martine, D., Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. Phonology 17. 143–185.

Gussenhoven, Carlos. 1991. Tone segments in the intonation of Dutch. In Thomas F. Shannon & Johan P. Snapper (eds.), The Berkeley conference on Dutch linguistics 1989, 139–155. Lanham, MD: University Press of America.

Gussenhoven, Carlos. 1993. The Dutch foot and the chanted call. Journal of Linguistics 29. 37–63.

Gussenhoven, Carlos. 2000. The boundary tones are coming: On the non-peripheral realization of boundary tones. In Michael B. Broe & Janet B. Pierrehumbert (eds.), Papers in laboratory phonology V: Acquisition and the Lexicon, 132–151. Cambridge: Cambridge University Press.

Gussenhoven, Carlos. 2004. The phonology of tone and intonation. Cambridge: Cambridge University Press.

Gussenhoven, Carlos. 2005. Transcription of Dutch intonation. In Sun-Ah Jun (ed.), Prosodic typology: The phonology of intonation and phrasing, 118–145. Oxford: Oxford University Press.

Hayes, Bruce & Aditi Lahiri. 1992. Durationally specified intonation in English and Bengali. In Rolf Carlson, Lennart Nord, & Johan Sundberg (eds.), Proceedings of the 1990 Wenner-Gren Center Conference on music, language, speech, and brain, 78–91. New York: Stockton Press.

Hyman, Larry M. 2007. Universals of tone rules: 30 years later. In Tomas Riad & Carlos Gussenhoven (eds.), Tones and tunes I: Typological studies in word and sentence prosody, 1–34. Berlin and New York: Mouton de Gruyter.

Kleber, Felicitas & Tamara Rathcke. 2008. More on the “segmental anchoring” of prenuclear rises: Evidence from East Middle German. In Proceedings of Speech Prosody 2008, Campinas, Brazil, May 6–9, 2008.

Ladd, D. Robert. 1978. Stylized intonation. Language 54. 517–540.

Ladd, D. Robert. 1983. Phonological features of intonational peaks. Language 59. 721–759.

Ladd, D. Robert, Ineke Mennen & Astrid Schepman. 2000. Phonological conditioning of peak alignment in rising pitch accents in Dutch. Journal of the Acoustical Society of America 107. 2685–2696.

Ladd, D. Robert, Astrid Schepman, Laurence White, Louise M. Quarmby & Rebekah Stackhouse. 2009. Structural and dialectal effects on pitch peak alignment in two varieties of British English. Journal of Phonetics 37. 145–161.

Liberman, Mark. 1975. The intonation system of English. Cambridge, MA: MIT Ph.D. dissertation.

Mücke, Doris, Martine Grice, Johannes Becker & Anne Hermes. 2009. Sources of variation in tonal alignment: Evidence from acoustic and kinematic data. Journal of Phonetics 37. 321–338.
Mücke, Doris, Martine Grice, Anne Hermes & Johannes Becker. 2008. Prenuclear rises in Northern and Southern German. In *Proceedings of Speech Prosody 2008*, Campinas, Brazil, May 6–9, 2008.

Niebuhr, Oliver & Gilbert Ambrazaitis. 2006. Alignment of medial and late peaks in German spontaneous speech. In *Proceedings of Speech Prosody 2006*, Dresden, May 2–5, 2006.

Nooteboom, Sieb G. 1972. *Production and perception of vowel duration. A study of durational properties of vowels in Dutch*. University of Utrecht Ph.D. dissertation.

O’Connor, J. D. & G. F. Arnold. 1973. *Intonation of colloquial English* (2nd ed.). London: Longman.

Peters, Jörg. 1999. The timing of nuclear high accents in German dialects. In *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, August 1–7, 1999, Vol. 3, 1877–1880.

Peters, Jörg. 2009. Intonation. In *Duden bd. 4: Die Grammatik*, (8th rev. ed.), 95–128. Mannheim: Bibliographisches Institut.

Peters, Jörg. 2014. *Intonation*. Heidelberg: Universitätsverlag Winter.

Peters, Jörg, Judith Hanssen & Carlos Gussenhoven. 2014. The phonetic realization of focus in West Frisian, Low Saxon, High German, and three varieties of Dutch. *Journal of Phonetics* 46. 185–209.

Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. Cambridge, MA: MIT Ph.D. dissertation.

Pierrehumbert, Janet B. & Mary E. Beckman. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.

Prieto, Pilar, Jan van Santen & Julia Hirschberg. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23. 429–451.

Prince, Alan & Paul Smolensky. 1993. Optimality theory: Constraint interaction in Generative grammar. *Rutgers University Center for Cognitive Science Technical Report* 2.

Rietveld, Toni, Joop Kerkmoff & Carlos Gussenhoven. 2004. Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics* 32. 349–371.

Schepman, Astrid, Robin Lickley & D. Robert Ladd. 2006. Effects of vowel length and ‘right context’ on the alignment of Dutch nuclear accents. *Journal of Phonetics* 34. 1–28.

Silverman, Kim E. & Janet B. Pierrehumbert. 1990. The timing of prenuclear high accents in English. In John Kingston & Mary E. Beckman (eds.), *Papers in laboratory phonology I*, 71–106. Cambridge: Cambridge University Press.

Steele, Shirley A. 1986. Nuclear accent F0 peak location: Effects of rate, vowel, and number of following syllables. *Journal of the Acoustical Society of America* 80(Supplement 1). s51.

’t Hart, Johan, René Collier & Antonie Cohen. 1990. *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.

van de Ven, Marco & Carlos Gussenhoven. 2011. The timing of the final rise in falling-rising intonation contours in Dutch. *Journal of Phonetics* 39. 225–236.

van Leyden, Klaske & Vincent J. van Heuven. 2006. On the prosody of Orkney and Shetland dialects. *Phonetica* 63. 149–174.

Xu, Ronghui. 2003. Measuring explained variation in linear mixed effects models. *Statistics in Medicine* 22. 3527–3541.