Effects of vowel length and "right context" on the alignment of Dutch nuclear accents

Astrid Schepman*

Robin Lickley**

D. Robert Ladd

Edinburgh University

Final re-revised version January 2005

* Division of Psychology, University of Abertay Dundee, Bell Street, Dundee, DD1 1HG, UK, email: a.schepman@tay.ac.uk

** Department of Speech and Language Sciences, Queen Margaret University College, Clerwood Terrace, Edinburgh, EH12 8TS, email: rlickley@qmuc.ac.uk

Department of Theoretical and Applied Linguistics, University of Edinburgh, George Square, Edinburgh EH8 9LL, email: bob@ling.ed.ac.uk

Running title: Alignment of Dutch nuclear accents

Abstract

We measured the alignment of f0 landmarks with segmental landmarks in nuclear "pointed hat" accents in controlled speech materials in Dutch. We varied the phonological length of the stressed vowel and the "right context" (syllable membership of following consonant, presence or absence of stress clash). The nuclear accented word was always followed by an unaccented content word. Based on previous work we expected that the alignment would be substantially affected by vowel length, stress clash and syllable membership, but the only important effect was that of vowel length. We believe this can be explained by the fact that most previous studies have dealt with prenuclear accents and/or with nuclear accents in utterance-final position, whereas we are dealing with nuclear accents that are not in utterance-final position. We also explored the effects of using different quantitative definitions of our dependent and independent variables, and of using Multiple Regression rather than ANOVA, and conclude that our findings are robust regardless of the variables or analysis technique used. An important methodological conclusion from our comparative analyses is that tonal alignment is best expressed relative to a nearby segmental landmark. Proportional measures may also be useful, but need further investigation.

Key words: PROSODY, INTONATION, ALIGNMENT, f0, DUTCH

1. Introduction

1.1. Background

Since the late 1970s, when Bruce (1977) demonstrated that the Swedish word accent distinction is based on different patterns of temporal coordination between f0 movements and stressed vowels, many studies have investigated the temporal alignment of fundamental frequency features with the segmental string. This work has consistently shown that alignment is highly lawful, and that the principles are at least partly language-specific. Two examples of the kind of finding produced by this research are the following:

Silverman & Pierrehumbert 1990 studied the alignment of the f0 peak on pitch-accented words in English as a function of their length and syllable structure and as a function of the distance between the test pitch accent and a following one. Specifically, they measured "peak delay" in the first stressed syllable of names like *Mom Lemm, Mamalie LeMann, Mama LeMonick*, and so on, where peak delay was defined as the distance from the beginning of the stressed vowel to the f0 peak. They found consistent effects on peak delay depending on the duration of the stressed vowel, and concluded that their data could be best accounted for if peak delay was expressed as a *proportion* of the duration of the syllable rhyme. With peak delay so expressed, they found consistent independent effects on peak delay when (a) the stressed syllable was word-final, and (b) there was a stress clash, i.e. the stressed syllable in the first name was immediately followed by the stressed syllable of the last name. There were also effects of speech rate which we will not consider further here. Results comparable to Silverman & Pierrehumbert's have been obtained by Caspers & van Heuven 1993 for Dutch and by Prieto, van Santen & Hirschberg 1995 for Mexican Spanish; cf. also the related perceptual study by Rietveld and Gussenhoven 1995.

Arvaniti, Ladd & Mennen 1998 studied the alignment of the beginning and end of rising f0 movements accompanying prenuclear or non-final accents in Greek (i.e. the local f0 minimum and maximum in the vicinity of the accented syllable). Unlike Silverman and Pierrehumbert, they were concerned to avoid any effects of stress clash and word boundary: their accented syllables were always antepenultimate in the word, and always separated from the following pitch accent by three unstressed syllables. Instead, they manipulated the actual duration of the accented syllable and the following syllable onset (e.g. by using fricatives, clusters, and low vowels to create longer syllables and nasals and high vowels to create shorter syllables). They found that the duration of the f0 rise was highly correlated with the duration of the accented syllable and following onset consonant. They interpreted this as indicating a consistent alignment of the f0 minimum with the beginning of the syllable and of the f0 maximum with the beginning of the following unstressed vowel. Results comparable to Arvaniti et al.'s have been obtained by Xu 1997, 1998, 1999 for Mandarin Chinese, by Ladd, Faulkner, Faulkner & Schepman 1999 for British English, by Ladd, Mennen & Schepman 2000 for Dutch, and by Ishihara 2003 for Japanese.

Findings of the sort exemplified by the two studies just discussed differ in their emphasis – the one looks for systematic sources of variation, the other for invariance when certain sources of variation are held constant – but they are by no means incompatible. It is entirely plausible to regard f0/segmental coordination as involving an ideal target alignment (of the sort investigated by Arvaniti et al.) which can be perturbed or modified under certain conditions (of the sort investigated by Silverman & Pierrehumbert). Indeed, Bruce's original work on the realisation of the Swedish word accent distinction suggested just such an approach to certain cases (1977: sec. 5.4.). However, even if such a synthesis of the two lines of research is possible, the recent accumulation of data from studies of various languages makes it clear that we are still some way from being able to identify the "ideal" alignment patterns in any given case, and some way from understanding the many factors that affect the realisation of the ideal.

1.2. The present study

This paper reports a study of the alignment of the f0 peak of nuclear (rising-)falling accents in ordinary declarative sentences in Standard Dutch. In the terminology of the IPO (Institute for Perception Research) tradition (e.g. ’t Hart, Cohen & Collier 1990), these accents are known as “pointed hats” and are analysed as a combination of a “Type 1 Rise” and a “Type A Fall” on a single prominent syllable. A ToBI-style (e.g. Silverman et al., 1992) transcription of these accents would be L+H* L- (L%) or H* L- (L%); in the ToDI system (a ToBI-based transcription system for Dutch which dispenses with the notion of the “phrase accent”; see Gussenhoven et al. 1999) they would be transcribed H*L L%. Whatever the details of the analysis, these accents are intended to be ordinary declarative accents, with no special emphasis and no substantial downstep or contextual lowering.

We were specifically concerned to investigate two main empirical questions:

- Are the “right context” effects that have been reported in the literature primarily due to the phonetic pressures of “tonal crowding” or to structural factors like “word boundary” and “stress clash”?

- Can the effect of phonological vowel length (or tenseness) on alignment reported by Ladd et al. 2000 be attributed to differences of syllable structure (as Ladd et al. proposed), or is vowel length itself directly relevant?

However, by investigating nuclear accents we also hoped to extend the range of cases for which good alignment data are available, and as we shall see, our results are also relevant to a third question:

- do nuclear and prenuclear accents behave the same way with respect to f0 alignment?

Furthermore, we have also used our data to address essentially methodological questions that arise from differences of practice among those who have studied alignment:

- is the most appropriate quantitative characterisation of alignment a *difference* measure, such as the distance between an F0 peak and a segmental landmark like the end of a stressed vowel (e.g. Arvaniti et al. 1998), or a *proportion* measure, such as the distance between the beginning of the stressed syllable and the F0 peak expressed as a fraction of the duration of the stressed syllable (e.g. Silverman & Pierrehumbert 1990)?
- What independent variables should be examined to gain a full understanding of alignment, and are these independent variables best characterised as categorical or continuous? For example, if vowel length is relevant, is it a matter of phonological vowel quantity – categorically short or long – or of phonetic vowel duration?

The background to the three substantive questions is sketched in the next subsections. We return to the methodological questions in section 5.

1.2.1. "Right context" and stress clash

Several studies (e.g. Steele 1986; Silverman & Pierrehumbert 1990; Prieto et al. 1995; Arvaniti, Ladd & Mennen 1998, forthcoming) report that alignment is substantially affected by the position of the accented syllable in the word and/or the "Abercrombian foot" (the interval between one stressed syllable and the next; cf. Abercrombie 1964). Specifically, it has generally been found that the alignment of f0 targets with an accented syllable is later as the number of following unstressed syllables increases. Looked at the other way, the closer an upcoming boundary or stressed syllable is, the earlier the alignment of the f0 targets will be; the "right context" may push the f0 movements "leftward". As a statement of general tendencies, this summary is uncontroversial.

However, at least two things remain unclear. One is the size of the phonological domain within which the summary statement is valid: do right-context effects operate at the level of the foot, the (prosodic) word, or some larger prosodic unit like the intonation phrase? The other is more fundamental: is the effect of a following boundary or a following stressed syllable "phonological or phonetic", i.e. is it primarily structurally conditioned, or is it attributable to physical constraints on f0 realisation? For example, suppose tonal targets are aligned earlier in syllables that are closely followed by a boundary or by another stressed syllable. This could be because the following boundary or stress marks the end of a phonological domain that specifies the realisation of the f0 movement – this would be a case of "structural" conditioning. Alternatively, the influence could be at the level of physical constraints on speech production, e.g. the earlier alignment could result from the need to complete a f0 movement before the end of an utterance (this could explain the results of Steele 1986 and Arvaniti et al. forthcoming) or before the beginning of a subsequent f0 movement (this could explain Silverman & Pierrehumbert 1990, Prieto et al. 1995, and Arvaniti et al. 1998).

It is not straightforward to investigate these questions empirically, in part because of the number of potentially confounding issues involved in designing materials to test hypotheses. For example, to test the effects of "stress clash", we first have to decide whether (putting it in terms of metrical phonology) clashes at all levels in the grid are to be considered, or whether by definition "stress clash" is present only between pitch-accented syllables. If we do not consider lower-level clashes, we prejudge the "phonological vs. phonetic" question, because, by definition, we will be considering only cases where there clearly exists a plausible phonetic explanation in terms of f0 movement. On the other hand, if we do consider lower-level clashes, we need to make decisions about a large number of disagreements of detail over what counts as a clash at what level. A specific issue here is whether clashes can be defined in terms of Abercrombian feet that potentially cross word boundaries, or whether, as assumed by Selkirk 1984 and others, the foot is part of a strictly-layered prosodic structure and is by definition a sub-domain of the prosodic word.

In the present study we considered clash at the level of the prosodic word, i.e. "below" the level of adjacent pitch accents, but constructed our materials in such a way that they might shed light on the question of whether a construct like the Abercrombian foot is relevant in describing speech timing. As will be explained in more detail in the Method section, our materials take advantage of the fact that in Dutch, as in German, the non-finite verb form at the end of a clause is very often unaccented even though it is a content word with a lexically stressed syllable. Specifically, this feature of Dutch made it easy to embed nouns as test words into sentences in which they bear the final pitch accent but are not utterance-final, such as *Hij kon de* ***maan*** *zien* /ηEI κ□ν δ↔ ∪μα⌉ν ∩ζι⌉ν/ 'He could see the moon'. Here the test word *maan* normally bears the nuclear accent but is followed by another (post-nuclear) content word and is therefore not in absolute sentence-final position. Given these test materials, if we find that nuclear accented syllables show different alignment characteristics from those reported in the literature for prenuclear syllables, we can be sure that the difference is not the result of the accented syllables being in absolute utterance-final position. Moreover, the occurrence of post-nuclear content words allows us to manipulate the proximity of following lexically stressed syllables without also affecting the position of pitch accents. For example, in *Hij kon de **maan** zien* we have a "stress clash" involving the syllables /μα⌉ν/ and /ζι⌉ν/, whereas in *Hij had de **maan** gezien* /ηEI ηAδ δ↔ ∪μα⌉ν ξ↔∩ζι⌉ν/ 'He had seen the moon' we do not, but the syllable /ζι⌉ν/ has no pitch accent in either case. With regard to the question of whether right context effects are "phonological or phonetic", we can therefore rule out any obvious contribution of "time pressure" if we find right context effects of the sort described above.

1.2.2. Vowel length and syllable structure

Ladd et al. 2000 found a significant effect of phonological vowel length on the alignment of prenuclear accent peaks in Dutch: with long vowels, the peak was aligned near the end of the vowel, but with short vowels the peak was aligned in the following consonant. By comparing test words containing the long vowel /ι⌉/ and the short vowel /I/ (which are virtually identical in duration despite having different phonological quantity or "tenseness"), Ladd et al. were able to show that the effect is not merely due to phonetic differences of vowel duration, but to the phonological distinction between "long" and "short". However, they speculated that the relevant factor is not the phonological length distinction itself, but rather the differences in syllable structure that follow from vowel length. This speculation is based on the following well-established facts about Dutch syllable structure: In syllables containing a long vowel, a single following consonant in the same word is unambiguously the onset of the next syllable, whereas in syllables with a short vowel a single following consonant to some extent fills the role of coda consonant and is arguably ambisyllabic[1]. This means that the alignment difference could be interpreted as a reflex of aligning the F0 peak *with the end of the syllable*: in an open syllable (the long vowel cases) the end of the syllable (and the alignment point of the accentual peak) is the end of the vowel, whereas in a closed syllable (the short vowel cases, assuming ambisyllabicity) the end of the syllable and the alignment point of the accentual peak are located somewhere in the coda consonant.

In the present study, we can test this by having one and two syllable words with long and short vowels, such as *maan* /μα⌉ν/ 'moon', *manen* /∪μα⌉ν↔/ 'moons', *man* /μΑν/ 'man', *mannen* /∪μΑν↔/ 'men'. If the effect of vowel length on alignment is primarily due to differences of syllable structure, then the addition of the suffix should have a large effect on alignment in the long vowel cases (because in e.g. /μα⌉ν/ the /ν/ is part of the test syllable but it /∪μα⌉ν↔/ it is not), but should have a smaller effect, or no effect at all, in the short vowel cases (because in both e.g. /∪μΑν↔/ and /μΑν/ the /ν/ is presumably part of the test syllable). If on the other hand the effect of vowel length is directly due to the vowel's phonological status as long or short and not to syllable structure, then any effect of adding a suffix should be similar for both long and short vowel cases.

1.2.3. Nuclear vs. prenuclear accents

Finally, we should briefly sketch the background to the distinction we have made between nuclear and prenuclear accents. For a brief period in the mid-1980s, there was considerable interest in the question of whether nuclear and prenuclear accents are the same sort of entity. (The terms "nuclear" and "prenuclear" are used here in roughly the sense that they had in the British descriptive tradition of e.g. Palmer, 1922 and O'Connor & Arnold, 1973; see Ladd

1996: sec. 6.1 for discussion). Early attempts at speech synthesis in English discovered empirically that the f0 peaks in nuclear accents need to be aligned earlier than in prenuclear accents in order to sound right (Silverman and Pierrehumbert 1990: 74). This finding took on theoretical significance in the context of Pierrehumbert's analysis of English intonation (1980), which posited a single inventory of "pitch accents" for English and ascribed any difference between nuclear and prenuclear accents either to mere position (the nuclear accent being the last accented word in a phrase) or to the following "phrase accent". Silverman & Pierrehumbert's 1990 paper was specifically aimed at showing that alignment in prenuclear accents is subject to the same sort of right context effects that Steele (1986) found for nuclear accents, thereby indirectly validating Pierrehumbert's decision to collapse the distinction between the two types. To our knowledge, this issue has not been investigated further since 1990, and most studies of alignment have been concerned with prenuclear accents.

Our study did not attempt any direct comparison between nuclear and prenuclear accents, but (as will be seen) our results are sufficiently different from the results of our previous study of prenuclear accents in Dutch that a comparison between the two studies seems valuable. We return to this question in Section 4.

2. Method

As in most other recent studies of alignment, we recorded controlled speech materials read aloud under laboratory conditions, and then measured the alignment of specific f0 points at preplanned locations. We have elsewhere (Lickley, Schepman & Ladd, forthcoming) defended the validity of this approach from the criticisms of those who believe that intonational phenomena should be studied primarily on the basis of spontaneous speech.

2.1. Speech materials

We designed a corpus of 120 short sentences intended to contain only a single pitch accent when read aloud. The phonological properties of the test word (the word on which the accent was expected to occur) and of its “right context” (i.e. the following word) were systematically varied. The sentences were mostly of the following general form:

[subject pronoun] - [auxiliary verb] - [determiner] - [TEST WORD] - [non-finite lexical verb]

where the test word was most commonly a lexical noun, normally the grammatical object of the verb. The lexical verb was either an infinitive (with the stress on the first syllable) or a past participle (with an unstressed prefix *ge-* /ξ↔/)[2]. The test words had either a phonologically short or a phonologically long stressed vowel, and were either monosyllabic or disyllabic. Approximately half of the disyllabic test words were monosyllabic nouns with the plural suffix *–en* (phonetically /-↔/ or /-↔ν/); the remainder either had other suffixes with a schwa vowel (*-lijk* /-λ↔κ/, *-ig* /-↔ξ/, *-er* /-↔ρ/) or were monomorphemic (e.g. *honger* /∪η□N↔ρ/ ‘hunger’ or *Rome* /∪ρο⌉μ↔/ ‘Rome’). The materials were thus based on a 2 x 2 x 2 design, with the variables VOWEL LENGTH (long or short), SUFFIX (absence or presence of “suffix” or weak second syllable), and PREFIX (absence or presence of prefix on following verb form, i.e. infinitive or past participle). As far as possible the consonants in the test words were sonorants, to permit easy extraction and measurement of f0. The stressed vowels in the test words were /α⌉/, /A/, /ο⌉/, /□/, /ε⌉/ and /E/; we avoided high vowels because the phonologically long high vowels in Dutch are phonetically quite short.

Examples of the eight resulting types of test sentences are given in the following list. Note that only two of these examples were actually used in the experiment, and the other six have been deliberately constructed to be as parallel as possible so that the reader can more easily appreciate the structure of the materials. In the actual experiment we varied the content of the sentences as much as possible within the strict limits imposed by our design, precisely so that our speakers would *not* appreciate the structure of our materials! We use the following examples throughout the paper to refer to specific combinations of independent variables. A full list of materials is given in the Appendix.

- (VOWEL LENGTH long, SUFFIX absent, PREFIX absent) *Hij kon de maan zien* /ηEI κ□ν δ↔ ∪μα⌉ν ∩ζι⌉ν/ 'He could see the moon'

- (VOWEL LENGTH long, SUFFIX absent, PREFIX present) *Hij had de maan gezien* / ηEI ηAδ δ↔ ∪μα⌉ν ξ↔∩ζι⌉ν/ 'He had seen the moon'

- (VOWEL LENGTH long, SUFFIX present, PREFIX absent) *Hij kon de manen zien* /ηEI κ□ν δ↔ ∪μα⌉ν↔ ∩ζι⌉ν/ 'He could see the moons'

- (VOWEL LENGTH long, SUFFIX present, PREFIX present) *Hij had de manen gezien* /ηEI ηAδ δ↔ ∪μα⌉ν↔ ξ↔∩ζι⌉ν/ 'He had seen the moons'

- (VOWEL LENGTH short, SUFFIX absent, PREFIX absent) *Hij kon de man zien* /ηEI κ□ν δ↔ ∪μAν ∩ζι⌉ν/ 'He could see the man'

- (VOWEL LENGTH short, SUFFIX absent, PREFIX present) *Hij had de man gezien* /ηEI ηAδ δ↔ ∪μAν ξ↔∩ζι⌉ν/ 'He had seen the man'

- (VOWEL LENGTH short, SUFFIX present, PREFIX absent) *Hij kon de mannen zien* /ηEI κ□ν δ↔ ∪μΑν↔ ∩ζι⌉ν/ 'He could see the men'

- (VOWEL LENGTH short, SUFFIX present, PREFIX present) *Hij had de mannen gezien* /ηEI ηΑδ δ↔ ∪μΑν↔ ξ↔∩ζι⌉ν/ 'He had seen the men'

The materials just described were intended to shed light on the two main empirical questions outlined in the introduction: whether stress clash (variously defined) leads to right context effects (section 1.2.1.), and whether phonological vowel length affects alignment directly, or indirectly through its effects on syllable structure (section 1.2.2.).

We can clearly test for effects of stress clash by looking for differences among the four conditions that result from combinations of the SUFFIX and PREFIX variables: there should be big differences of alignment between sentences like *Hij kon de man zien* ("SUFFIX absent, PREFIX absent"), where the accented syllable is immediately followed by another stressed syllable, and sentences like *Hij had de mannen gezien* ("SUFFIX present, PREFIX present"), where the accented syllable is separated from the following stressed syllable by two unstressed syllables. More subtly, our materials were also designed to shed light on whether the prosodic effects of foot structure are best expressed in terms of the Selkirkian "within-word" foot or the Abercrombian foot that ignores word boundaries. We can do this by comparing the "SUFFIX present, PREFIX absent" condition (e.g. *Hij kon de mannen zien*) with the "SUFFIX absent, PREFIX present" condition (e.g. *Hij had de man gezien*). If the accented syllable forms a foot with the following unstressed syllable regardless of whether it is part of the same word (as suggested by Abercrombie's definition), then these two cases should show the same alignment behaviour. If, on the other hand, a monosyllabic accented word forms some sort of prosodic unit on its own, without the unstressed prefix of the following word (as suggested by Selkirk and much other work in metrical phonology), then the two conditions should yield different results.

As for the effects of vowel length, as already explained, the "SUFFIX present" condition creates different syllable structures from the "SUFFIX absent" condition in the case of long vowels but presumably not in the case of short vowels. If the effects of phonological vowel length on alignment are actually due to syllable structure, as proposed by Ladd et al. 2000, then the VOWEL LENGTH variable should interact with the SUFFIX variable: specifically, there should be a large effect of SUFFIX with long vowels but not with short vowels.

2.2. Speakers

The speakers were six students at the University of Nijmegen in their early to mid twenties. Three were female (SK, SH, ND) and three male (AC, XC and JP). (A seventh speaker (MF, female) was recorded but her recordings were not analysed because she frequently placed the main accent on a word other than the intended test word). As would be expected from the Nijmegen student population, some of the speakers (especially XC, JP and AC) had identifiably southern Dutch accents, while the others had general Dutch ("Algemeen Nederlands") accents. The speakers were paid a small sum for their participation.

2.3. Recording and analysis procedures

The test sentences were printed on sheets of paper, ten sentences per A4 page, in a randomized order. The test sentences were interspersed with materials for a separate experiment not reported here. The interspersed sentences were short questions. By including roughly equal numbers of declarative sentences and questions we hoped to prevent the speakers from settling into a monotonous reading style. We gave the speakers no explicit instructions about intonation; they were simply asked to read the sentences naturally and to read each sentence individually rather than treating the whole series as a list. They were asked to repeat the whole sentence if they made an error, but this was a rare event.

Recordings were made on DAT tape using professional equipment in a sound-treated recording booth in the Nijmegen University Phonetics Laboratory. The digital recordings were transferred to a Sun workstation at the University of Edinburgh for acoustic analysis using ESPS Waves+ software. F0 was extracted with the get_f0 programme using its default settings, i.e. a 7.5 ms correlation window and a 10 ms frame shift. Acoustic measurements were made by the first and second authors on the basis of simultaneous displays of waveform, f0 contour, and wide-band spectrogram.

We labelled the following segmental landmarks in the speech files:

- **C0**: the beginning of the onset consonant of the accented syllable of the test word
- **V0**: the beginning of the vowel of the accented syllable of the test word
- **C1**: the beginning of the postvocalic consonant immediately following the stressed vowel of the test word[3]
- **V1**: the end of the consonant starting at **C1** (if a suffix was present, this was the beginning of the unstressed vowel of the suffix of the test word, hence the label "**V1**")
- **C2**: the beginning of the first consonant of the word following the test word
- **V2**: the beginning of the stressed vowel of the word following the test word
- **C3**: the end of the stressed vowel of the word following the test word

The details of how these labels match up to actual segments in the various experimental conditions are exemplified in Figure 1.

////////// FIGURE 1 ABOUT HERE ///////

In addition to the segmental boundaries, we marked the location of the f0 valley at the onset of the accented syllable 'L', and we marked the location of the f0 peak towards the end of the accented syllable 'H'. Mostly, these f0 values were easy to identify visually, especially for the peaks, as they tended to occur on steady-state portions of vowels. For the valleys, there were sometimes possible microprosodic effects of the segments on the f0 contour. If such effects were clear, we chose the next lowest f0 point as our L, although we did not make this adjustment if we were in doubt. If two candidate points had the same f0 values, then we consistently chose the one that occurred first. Note that in some instances, the L could not be identified, as it was too close to unvoiced preceding consonants (this was true for 38 of 681 cases, or 5.6%, compared to 1.7% of missing data for H).

We used the labels of segment boundaries and f0 landmarks just described to derive dependent variables for expressing alignment. Our choices here were based on past work, in particular on Arvaniti et al. 1998 and Ladd et al. 2000. We calculated two alignment measures, which express time intervals in milliseconds. The first, the alignment of the H, is expressed by the dependent variable *H-C1,* which is time at which C1 occurs minus the time at which H occurs, i.e. a difference score. This variable has negative values when H precedes C1, i.e. when the f0 maximum is aligned during the accented vowel, and positive values when the f0 maximum occurs after the end of the vowel. The second, the alignment of L, is expressed by the dependent variable *L-C0*. This has negative values when the f0 minimum is aligned before the beginning of the accented syllable and positive values when it is aligned during the onset consonant. In section 5 we explore whether the choice of dependent variable makes an important difference to our general conclusions about how alignment works. In addition to the two alignment measures we also calculated segment durations, as control measures. These are described in section 3.3.

3. Results

As a general first approach, we analysed the measures of interest with a 2 x 2 x 2 x 6 ANOVA. As noted above, the factors reflecting the manipulations in the speech materials were VOWEL LENGTH (Long /Short); SUFFIX (a shorthand for the presence / absence of an unstressed second syllable in the target word – e.g. *man, mannen*); and PREFIX (a shorthand for the presence / absence of an unstressed post-nuclear past participle prefix on the next word – e.g. *zien, gezien*). In addition, we took SPEAKER as a factor. In all analyses, "items" was the random factor, and all fixed factors were treated as between-items. The data are illustrated in Figure 2.

/////// FIGURE 2 ABOUT HERE ///

3.1. Alignment of H

3.1.1. Effects of experimental manipulations

The most substantial finding of the ANOVA is a main effect of VOWEL LENGTH, $F(1, 659) = 297.74$, $p < 0.0001$, with alignment for long vowels being earlier (56 ms before the end of the vowel) than for short vowels (16 ms before the end of the vowel). To some extent this replicates experiment 1 of Ladd et al. (2000), which showed a similar difference in alignment for long-vowel versus short-vowel syllables. However, the peaks are earlier here than in Ladd et al. 2000, as one might expect given that these are nuclear rather than prenuclear accents (see sec. 1.2.3 above). We return to this issue in the discussion.

The main effect of SUFFIX was not significant ($F < 1$), but there was a significant interaction between VOWEL LENGTH and SUFFIX, $F(1,659) = 7.24$, $p < 0.01$. In short vowels H aligned 12 ms before the end of the vowel when there was a suffix, and slightly earlier, at 19 ms before the vowel offset, when the suffix was absent. The pattern of means was reversed in long vowels (59 ms vs 52 ms before the vowel offset for SUFFIX absent vs present, respectively). Neither of these differences was significant on a Bonferroni t-test ($p > 0.6$). Note that, even though we found an interaction, the pattern of means is different from the pattern one would expect if peaks aligned with the edge of the syllable (see section 1.2.2.), which would have shown there to be no difference between the suffix conditions in short vowels, but a very substantial difference in long vowels.

There was a small (7 ms) but significant main effect of the factor PREFIX, $F(1, 659) = 10.42$, $p < 0.001$, with the H aligning later when the prefix was present than when it was absent. In this case, the effect did not interact with that of VOWEL LENGTH.

One final analysis is relevant to the general question of stress clash and right context effects. We analysed the combined effect of the four combinations of the SUFFIX and PREFIX using the Bonferroni t-test. This test did not reach significance when long and short vowels were combined. However, separating the long and short vowels revealed a significant difference, in short vowels only, between the "SUFFIX absent, PREFIX absent" condition (i.e. the "stress clash" condition) and the "SUFFIX present, PREFIX present" condition; alignment was 16 ms earlier in the former than the latter condition. This was the stress clash effect we expected. However, no other contrasts reached significance.

3.1.2. Speaker variation

The alignment of H relative to C1 is shown separately for each speaker in Figure 3. Although this figure shows rather large individual differences in overall alignment, it can be seen that the pattern was relatively stable across conditions. In the following paragraph we briefly summarise the individual differences we found, but we emphasise that the overall patterns of alignment are very similar for all speakers.

////// FIGURE 3 ABOUT HERE /////

The ANOVA revealed that the effect of SPEAKER was significant, $F(5,3295) = 109.29$, $p < 0.0001$. Speakers showed a strong variation in their mean individual alignment points, with AC aligning 75 ms before the end of the vowel, SH 71 ms, SK 23 ms, ND 22 ms, JP 13 ms and XC 8 ms. The effect of VOWEL LENGTH interacted with that of SPEAKER, $F(5, 3295) = 2.274$, $p < 0.05$. An inspection of the means revealed small variations in effect size for the VOWEL LENGTH effect across different speakers as a source of the interaction. In addition, there were three-way interactions between VOWEL LENGTH, SPEAKER and SUFFIX and also between VOWEL LENGTH, SPEAKER and PREFIX, again due to minor differences in effect sizes.

3.2. Alignment of L

The dependent variable here is the duration of the time interval (in ms) between the local f0 minimum at the start of the rise (L) and the start of the onset consonant (C0). Overall, L was aligned 2 ms after C0. However, there was a main effect of VOWEL LENGTH, $F(1,633) = 11.69$, $p < 0.001$, with L alignment in short vowels averaging at 3 ms before C0, while in long vowels alignment was later, at 5 ms after C0. That is, the effect of VOWEL LENGTH on H alignment is clearly mirrored in L alignment as well, though the size of the effect on L alignment is very small[4]. Note that this pattern of L alignment was also found by Ladd et al. 2000. This suggests that the anchoring of L to the onset of the syllable, found by Arvaniti et al. (1998) and others, can be subtly influenced in one direction or the other by time pressure; cf. also Prieto et al. (1995).

There was some variability between speakers. There was a main effect of SPEAKER, $F(5, 633) = 6.02$, $p < 0.0001$, with overall L alignment ranging from 5 ms before C0 (ND) to 15 ms after C0 (SK). (JP aligned L on average 4 ms before C0 and the other three speakers aligned L on average exactly at C0). There was a significant interaction between PREFIX and

SPEAKER, $F(5, 633) = 2.85$, $p < 0.05$. Inspection of the means revealed that five speakers showed earlier alignment of the L in the no prefix condition, with differences ranging from 5 to 12 ms earlier for the no prefix condition, while one speaker (SK) showed a reversed pattern.

3.3. Segment Duration Measures

We analysed the effect of our main factors on the duration of the three segments (onset consonant (C0 to V0), vowel (V0 to C1) and postvocalic consonant (C1 to V1)) that formed part of or flanked the target syllable. It is important to set out these results in some detail because segment durations may affect the alignment patterns indirectly, as we shall see later in this section.

The most striking effects were those of VOWEL LENGTH and SUFFIX. The relevant means are presented in Table I.

///// TABLE I ABOUT HERE /////

For the factor VOWEL LENGTH, the vowel duration and the onset consonant duration were longer ($p < 0.001$) when vowels were phonologically long, but the effect was reversed for the offset consonant ($p < 0.01$). While an effect was obviously expected for the duration of the vowel itself, it is not obvious that the consonants flanking the vowel should be affected by VOWEL LENGTH, too, nor that this effect should be in opposite directions. As for the factor SUFFIX, consonant segments were shorter ($p < 0.001$) when a suffix was present, as would be expected from the fact that the suffix increases the number of syllables in the accented word ("polysyllabic shortening"; cf. Lehiste, 1970; see also Turk & White 1999). The effect of

SUFFIX was most pronounced for the durations of the postvocalic consonant, whose syllable membership may also be affected by the presence of a suffix. For the duration of the vowel, there was no significant main effect of SUFFIX, but an interaction ($p < 0.0001$): the short vowels were shorter when a suffix was present, in line with the "polysyllabic shortening" seen above, while the long vowels were longer when a suffix was present. Both contrasts were significant on Bonferroni t-tests ($p < 0.001$). The lengthening of long vowels when a suffix was present is presumably due to the fact that the accented syllable in the long vowel words is unambiguously an open syllable when the suffix is present, i.e. the final consonant is syllabified with the next syllable. Thus, although we did not find evidence of differences in alignment based on syllable structure, we have some evidence here that syllable structure did have an effect on a different aspect of the data.

The main effect of PREFIX on the duration of the postvocalic consonant was significant, $F(1,659) = 17.79$, $p < 0.0001$, with the postvocalic consonants being shorter when a prefix was present (65 ms) than when it was absent (71 ms). However, this main effect was moderated by an interaction with SUFFIX, $F(1,659) = 5.01$, $p < 0.05$. The effect of PREFIX was non-significant when there was also a suffix present, $F(1,327) = 2.68$, $p > 0.1$, (means: 54 vs 57 ms. for PREFIX present vs. PREFIX absent), while the effect was significant when there was no suffix, $F(1,332) = 16.54$, $p < 0.0001$, (means: 77 vs 86 ms respectively). In other words, the prefix has an effect on the duration of the postvocalic consonant only when the latter is word-final.

The main effects of SPEAKER and some interactions with SPEAKER reached significance. Main effects of SPEAKER tended to be minor, and could be due to minor differences in speech rate, while the inspections of the means related to the interactions tended to reveal very minor numerical differences in effect sizes as likely sources.

Before leaving the topic of segment durations, we should note the possibility that duration differences may have an indirect effect on our alignment measures. For example, a contributory factor to the lack of a significant effect of SUFFIX on alignment may the substantial main effect of the factor SUFFIX on the duration of the postvocalic consonant. We have seen that the lack of suffix is associated with a longer postvocalic consonant. While there might be some leftward pressure on alignment from the lack of a suffix, this effect may be cancelled by the extra time made available by the longer postvocalic consonant. Note that similar observations were made by Silverman and Pierrehumbert (1990) and Prieto et al. (1995). We return to this question in section 5 below.

4. Discussion

In the preceding section we have presented a straightforward report of the effects of our experimental manipulations. Where possible, we have related our findings to the notions of stress clash and prosodic structure that we originally set out to explore. However, the effects we have been discussing, though perhaps ultimately of interest to a full account of f0 - segmental alignment, are in general extremely small. At this point we need to step back from the details of what we did find and consider the significance of what we did not find.

On the basis of past studies, we had good reason to expect that our experiment would reveal substantial effects of both syllable structure and stress clash. Specifically, on the basis of Silverman and Pierrehumbert's findings for English and Prieto et al.'s findings for Spanish, we expected substantially earlier peak alignment in the stress clash cases (SUFFIX absent, PREFIX absent) than in the other combinations of SUFFIX and PREFIX conditions. Statistically, that is, we expected a large interaction between the factors SUFFIX and PREFIX. Similarly, on the basis of Ladd et al.'s interpretation of the effects of phonological vowel length on alignment, we expected substantially earlier peak alignment when the consonant following the accented vowel is unambiguously syllabified as the onset of the following syllable (SUFFIX present, VOWEL LENGTH long) than in the other combinations of SUFFIX and VOWEL LENGTH conditions - again, a large interaction.

Neither of these expected interactions was found. The overwhelming impression of our overall findings is that there is a fairly fixed pattern of alignment for long vowels and another for short vowels, and that these patterns are only minimally affected by the manipulations of SUFFIX and PREFIX. The lack of large interactions, like Sherlock Holmes's dog that did nothing in the night-time, requires comment.

In the case of the anticipated interaction of VOWEL LENGTH and SUFFIX, it is true that the expectation of an interaction was based on an interpretation rather than a clear empirical finding. That is, it could be that Ladd et al.'s analysis of the difference in alignment between long and short vowels in Dutch is simply wrong: the difference has nothing to do with syllable structure, and no interaction should have been expected.

The absence of any large stress clash effect (the expected interaction between SUFFIX and PREFIX) is somewhat harder to explain away. It is true that the previous reports were based on English and Spanish, and it is conceivable that English and Spanish exhibit stress clash effects on alignment and Dutch does not. This, however, seems most unlikely; as noted in section 3.3, we do find similar durational effects of right context to those found by Silverman & Pierrehumbert and by Prieto et al. in English and Spanish respectively. Another possible explanation for our failure to find the expected interaction between SUFFIX and PREFIX is that "stress clash" is restricted to cases of clash between *pitch accented* stressed syllables (an assumption implicit in the work of e.g. Shattuck-Hufnagel, Ostendorf and Ross, 1994). On this view, the clashes we have set up between a nuclear accented syllable and an immediately following *unaccented* (postnuclear) stressed syllable somehow "don't count" as clash.

However, we think this is unlikely as well, because in the case of the short vowels there is evidence for a small stress clash effect. As seen in section 3.1.1, the "SUFFIX absent, PREFIX absent" condition (i.e. the "stress clash" condition) showed slightly but significantly earlier alignment with short vowels than the "SUFFIX present, PREFIX present" condition. This finding is consistent with the expectation that stress clash pushes alignment back (see section 1.2.1.); the restriction to short vowels could be due their relatively late peak alignment.

In any case, we believe there may be a unifying explanation for our failure to find both interactions, namely that the peak of a nuclear accent is aligned earlier than that of a prenuclear accent. There are striking differences between the data in the present paper and the data reported by Ladd et al. 2000, on which our expectations were based, and similar differences appear in other European languages as well. The Dutch prenuclear accents described by Ladd et al. 2000 have a peak late in the vowel for long vowels and in the middle of the following consonant for short vowels, whereas in the present study we find a peak midway through the vowel for long vowels and late in the vowel for short vowels. For English, recall that early attempts at speech synthesis in English found empirically that nuclear peaks need to be aligned earlier than prenuclear ones; our own work on English (in progress) is putting this finding on a more secure experimentally-based footing. Similar differences have turned up in Greek: Arvaniti et al. 1998 reported prenuclear peaks aligning early in the unstressed vowel following the accented syllable, while Arvaniti et al. (forthcoming) report nuclear high peaks about two-thirds of the way through the accented vowel. The same effects have also been reported for Spanish (Nibert 2000, Face 2002; for a review see Hualde 2002).

Although our experiment was not, of course, set up to compare nuclear and prenuclear alignment directly, the prenuclear data from Ladd et al. 2000 were based on methodologically very similar experiments and we feel it is appropriate to compare the results across the two experiments. Specifically, we compared *H-C1* in our data to *H-C1* in Experiment 1 from Ladd et al (2000), which involved prenuclear Dutch H* tones, in a range of phonologically long and short vowels. While in the nuclear data, H aligned 55 and 15 ms before C1 for long and short vowels respectively, the means were 25 ms before C1 and 12 ms after C1, respectively, in the prenuclear data. The difference between the two types of alignment is illustrated in figure 4, which shows earlier alignment in nuclear position (top panel), than in prenuclear position (bottom panel). A Bonferroni t-test showed that both the nuclear vs. prenuclear contrasts were significant ($p < 0.0001$). There were no significant differences in vowel duration as a function of intonational status to explain this difference in alignment, with vowel durations being 77 vs 84 ms for prenuclear vs nuclear, respectively, for short vowels, and 133 vs 139 ms, respectively, for long vowels ($p > 0.1$ in both cases).

/////FIGURE 4 ABOUT HERE/////

A plausible explanation for the difference of alignment can be sought in the idea that the pitch movement in a nuclear accent is more complex than in a prenuclear accent. Specifically, in "pointed hat" nuclear accents, the f0 must not only rise to the peak whose alignment we are measuring, but must also fall immediately afterwards. As we noted in the introduction, this immediately following fall is analysed as a separate accentual pitch movement in the traditional IPO analysis, and as the reflex of a "phrase accent" in some autosegmental-metrical descriptions of intonational phonology[5]. In both analyses, that is, the fall is a *separate phonological event,* distinct from the rise to the peak; only the rise is shared by both prenuclear and nuclear accents. When conceived of in this way, it seems reasonable to expect that this additional phonological event would make its own demands on the temporal organisation of the utterance. Indeed, just such an explanation has been proposed by Nibert (2000, cited in Hualde 2002) for the difference in alignment between prenuclear and nuclear accent peaks in Spanish; Nibert rejects Face's (2002) proposal that there are two categorically different accent types, H*+L and H+L*, and treats them all as H*, with or without a following phrase accent L. In any case, *something* appears to be causing nuclear accents to align earlier than prenuclear ones. The autosegmental analysis suggests that that "something" is the phrase accent, but if we prefer to call it a "Type A Fall" the essence of the explanation is unaffected.

We believe that this "something", in turn, provides the germ of an explanation for our failure to find right substantial context effects as a result of our experimental manipulations. In ways that are not quite clear, the right context established by the separate intonational event (phrase accent, Type A Fall, etc.) is the most important factor determining the alignment of the peak in nuclear accents, important enough that it overrides (at least in the statistical analysis, and ignoring the small stress clash effect seen with short vowels) the subtle effects of syllable structure and stress clash. In prenuclear accents, where the fall from the f0 peak to the beginning of the next accent is merely a transition, there is no rightward pressure from the string of intonational events themselves, and the effects of syllable structure and stress clash manifest themselves. In this connection we note that Caspers and van Heuven (1993) found greater effects of "time pressure" on prenuclear accents (their "Type 1" pitch rises) than on nuclear accents (their "Type A" pitch falls). Though they interpreted their results as involving a difference between falls and rises, their findings can readily be recast in the terms suggested here.

5. On defining the dependent variable for alignment

In the previous section we have suggested a possible explanation for the absence of the expected interactions, based on an analysis of the phonology of nuclear accents. There is, of course, a much more prosaic possible explanation, namely that our choice of *H-C1* as the principal dependent variable simply masks the effect we are looking for. Recall that some other researchers have used a “peak delay” measure of alignment that characterises alignment relative to the beginning of the vowel or the beginning of the syllable, and that some researchers have argued for defining “peak delay” as a proportion of the duration of the syllable rhyme. It is possible that one of these alternative dependent variables is the “right” way of characterising alignment quantitatively, and that only the right variable will give us the expected pattern of results.

The possibility of drawing misleading conclusions from poorly chosen dependent variables is not a trivial concern. For example, Prieto et al.’s observation (1995: 437) that “onset [consonant] duration is closely correlated with peak delay” leads them to reject Silverman and Pierrehumbert’s conclusion that the location of the f0 peak is best characterised as a proportion of the syllable rhyme. Yet a moment’s thought shows how the correlation of peak delay and onset duration could be a meaningless artefact of the quantitative definition of peak delay.

Suppose, for example, that the “true” production target for accent peaks is a point halfway through the rhyme of the accented syllable. Let us then define three variables, C (the duration of the onset consonant), R (the duration of the rhyme), and D (“peak delay” *as defined by Silverman and Pierrehumbert,* i.e. the distance of the peak from the beginning of the vowel). Given these definitions and the “true” basis of our empirical data, we will find a high correlation between D and R, such that D = 0.5(R), but we will not expect any correlation between D and C. Now let us define a further variable P (“peak delay” as defined by Prieto et al., i.e. distance of the peak from the beginning of the accented *syllable*). We will now find a good correlation between C and P, but this is merely an inevitable consequence of the fact that P = C + D and reveals nothing about the production target halfway through the rhyme.

In this section we investigate whether our conclusions would be changed by a different choice of variables. With regard to the dependent variable, we consider two possible issues: the choice of reference point for the definition of the dependent variable, and the choice of an absolute or a proportional definition for the dependent variable. We reran the ANOVA for the main experimental factors using a variety of different dependent variables, most of them taken from or adapted from previous research. This permitted us to compare our measure (*H-C1*) to other possible measures. We first examined other dependent variables that, like ours, express alignment as the absolute time interval between the F0 peak and some other acoustic landmark. We then also investigated the effect of expressing alignment as a proportion of a segmentally defined interval. We find from these analyses that our pattern of results is robust, and that our overall conclusions do not change.

A further issue is whether to define the *independent* variables as categorical or continuous. To investigate this question, we explored our data using multiple regression, as has been done by some previous researchers including Silverman and Pierrehumbert, 1990 and Prieto et al., 1995. Again, these analyses show that the overall pattern of our results is robust, and that there are no important advantages to be gained from multiple regression analysis. Note that Silverman and Pierrehumbert used categorical variables in their multiple regression analysis, which in any case makes their analysis effectively equivalent to ANOVA, but without interactions. The multiple regression analyses in Prieto et al., however, used continuous measures (e.g. the duration of other segments) as independent variables or predictors – for example, in place of our factor VOWEL LENGTH (long or short) they used variables like vowel duration (in ms.) – and this might uncover effects that would be obscured by our ANOVA technique. We redid a number of our analyses using such independent variables in a multiple regression model. One entirely predictable effect is that the overall variance on the dependent variable explained by the model (i.e. the R^2) is somewhat higher for the model with continuous variables than for the one with categorical variables, as there simply is more variation on the independent variables. But like the correlation of “peak delay” with onset consonant duration in Prieto et al.’s work, this there were no obvious ways in which the “improvement” from the multiple regression analysis suggested the existence of effects that were not revealed by our original ANOVA.

In the interests of brevity we omit further discussion of the comparison of ANOVA with multiple regression techniques. However, in the next two subsections we report some details of the comparison of different dependent variables. We believe that this comparison is important and timely given the considerable methodological variation in current work in this field.

5.1 Alignment defined as the duration of a time interval from an acoustic landmark

Up till now we have defined the alignment of the f0 peak under investigation as the duration, in ms., of the interval between the peak and the *end of the stressed vowel* (C1 in our terms). This choice was based on previous research (e.g. Arvaniti et al. 1998, Ladd et al. 2000), which suggested that C1 is a likely phonological target position. However, there are other possible acoustic landmarks relative to which alignment could be measured. For example, Silverman and Pierrehumbert (1990) define their variable “peak delay” as the duration, in ms, of the interval between the peak and the *beginning* of the stressed vowel (V0 in our terms). We will refer to this measure as “*Silverman peak delay*”. Prieto et al. (1995), who also call their dependent variable “peak delay”, define it (as we just saw) as the duration of the time interval between the peak and the beginning of the stressed *syllable* (C0, in our terms). We refer to this measure as "*Prieto peak delay*". One might also measure the alignment of the peak relative to the preceding F0 minimum; this is equivalent to measuring the duration of the F0 rise, and would be an obvious choice of dependent variable for a traditional model (e.g. the IPO model of 't Hart et al. 1990) that treats f0 changes rather than f0 targets as the fundamental building blocks of intonation. We refer to this variable as "*Rise time*".

In order to establish whether expressing the dependent variable differently affects the peak alignment results, we repeated the ANOVA described in section 3.2 for all of the alternative dependent variables mentioned in the previous paragraph. We found that three main effects were unchanged irrespective of how the dependent variable is defined:

- First, a significant effect of PREFIX was found using all measures (except for *Rise time)*, $F(1,633) = 7.78$ for *Silverman peak delay*; F(1,633 = 4.57 for *Prieto peak delay*, (for both $p < 0.05$) and $F < 1$, n.s., for *Rise time*. The pattern was stable, and in the same direction for all measures on which it reached significance: alignment is slightly later when there is a prefix on the next word. This suggests a consistent small right-context effect of "time pressure" (Caspers and van Heuven 1993) on the realisation of the pitch accent, time pressure that is less acute when there is an additional syllable between the accented syllable and the following lexically stressed syllable.
- Second, the main effect of VOWEL LENGTH was statistically robust for all dependent variables, $F(1,633) = 56.23$; $F(1,633) = 92.63$; and $F(1,633) = 27.67$, for *Silverman peak delay*, *Prieto peak delay* and *Rise time*, respectively; for all, $p < 0.0001$. Interestingly, the apparent direction of the VOWEL LENGTH effect reverses depending on the location landmark relative to which the location of the peak is expressed. For example, *Prieto peak delay* (H-C0) is 143 ms when the vowel is short, and 165 ms when the vowel is long; that is, H aligns "later" when the vowel is long than when it is short. (This can be seen clearly in Figure 2.) This also applies to all other measures

that express alignment relative to a landmark occurring before the peak. Despite the change in sign, however, the effect is significant whichever way it is expressed, so it does not seem to be an artefact of the measure chosen.

- Finally, the main effect of SPEAKER was significant, with very similar patterns, for all dependent variables, $F(5,633) = 112.47$ with *Silverman peak delay*; $F(5,633) = 101.61$ with *Prieto peak delay*; and $F(5,633) = 70.07$ with *Rise time*; for all $p < 0.0001$. The interactions between SPEAKER and the other factors did not give drastically different patterns of results from the ones already reported in section 3.2.

Two effects were somewhat more variable in their patterns of significance. The main effect of SUFFIX, which as already reported did not reach significance with *H-C1*, is also non-significant with *Silverman Peak Delay*, $F < 1$, n.s. Both these measures express alignment relative to a *nearby* segmental landmark. The effect did reach significance with *Prieto peak delay* and *Risetime*, $F(1,633) = 3.93$; and $F(1,633) = 6.209$, respectively; for both, $p < 0.05$, both of which use a landmark that is further removed from H. However, even when statistically significant, the effect of SUFFIX was numerically very small (i.e. for *Prieto peak delay* alignment is 5 ms earlier with suffix than without; for *Risetime* the difference is 7 ms). We suspect that by defining alignment relative to a more distant landmark, the variables *Prieto peak delay* and *Risetime* are more influenced by variations in segment duration than *Silverman Peak Delay* and our *H-C1* (again, recall our comments above on the correlation between *Prieto peak delay* and onset consonant duration). We therefore conclude that measuring alignment relative to a nearby variable is preferable. However, we acknowledge that other interpretations of the difference are possible.

The interaction between VOWEL LENGTH and SUFFIX also showed a dichotomy for nearby versus more distant landmarks, but in this case the interaction was significant for the measures which use the nearby segmental landmarks: *Silverman Peak Delay*, $F(1,633) =$

7.53, $p < 0.005$, and, as previously reported, *H-C1*. As already noted (section 1.2.2 and section 4), the interaction with *H-C1* was not of the type predicted on the basis of Ladd et al. 2000. The interaction with *Silverman peak delay* was similarly small numerically, and did not support the original hypothesis either. The interaction was not significant with the other two dependent variables, $F(1,633) = 2.54$, $p > 0.1$ with *Prieto peak delay*; $F < 1$, n.s., with *Rise time*. There is no basis for abandoning our principal empirical conclusion, namely that the expected effect of syllable structure is absent.

All other effects and interactions not explicitly mentioned in the foregoing paragraphs remained stable regardless of the choice of dependent variable.

5.2. Alignment defined as a proportion of a time interval

In the preceding subsection we expressed alignment as the duration of an interval between the peak and one of four different acoustic landmarks. However, it is also possible to express this interval as a proportion of some other interval, such as the duration of the vowel, the syllable rhyme, or the whole syllable. In some of their analyses, for example, Silverman and Pierrehumbert (1990), express "peak delay" not as a time interval with a duration in ms. but as a proportion of the syllable rhyme. They report that this way of expressing alignment gives more consistent results in their analyses. (For example, they found that in monosyllabic test words not immediately followed by a stressed syllable both of their speakers aligned their peaks roughly two-thirds of the way through the syllable rhyme.) Given that the main factors in our experiment had quite substantial effects on segment durations (see section 3.3), we felt it was important to explore whether expressing alignment proportionally makes any difference to our conclusions. As we have already suggested, we think that Prieto et al.'s rejection of a proportional measures of alignment is premature.

There are many possible candidate time intervals as a proportion of which alignment can be expressed, but we limited our exploration to the alignment as a proportion of the syllable rhyme, following Silverman and Pierrehumbert (1990). We chose the rhyme in part because it is widely recognised as a phonological unit, but in part because the analyses just reported in section 5.1 suggest that alignment is best expressed relative to nearby segmental landmarks, and the syllable rhyme is bounded by two landmarks that are both relatively near the alignment point. However, Silverman and Pierrehumbert's definition of "syllable rhyme" does not transfer readily to Dutch, as it excludes any intervocalic consonant following the stressed vowel (e.g. the second *m* in *Mamalie*). In Dutch, as noted above (section 1.2.2. and footnote 1), it is usual to distinguish between intervocalic consonants after long vowels (which are excluded from the rhyme) and those after short vowels (which are frequently included in the rhyme). Consequently, our dependent variable "*Proportional alignment*" is based on the Dutch treatment of intervocalic consonants. Specifically, it is defined as:

$$\frac{H - V0}{Bs - V0}$$

where Bs is the time of occurrence of the syllable boundary (i.e. the end of the rhyme), V0 is the time of occurrence of the vowel onset (i.e. the beginning of the rhyme), and H is the time of occurrence of the peak.

If we use a proportional definition of alignment, it is important to note that our predictions change, especially with regard to vowel length and syllable structure. Recall that in section 1.2.2 we predicted an interaction between VOWEL LENGTH and SUFFIX. This prediction is the statistical operationalisation of the theoretical claim that the peak is conditioned by the (Dutch) syllable boundary: because the end of the vowel is equivalent to the syllable boundary for long vowels but not for short vowels, we predicted that the peak alignment, defined relative to the end of the vowel, would be affected by vowel length. However, if alignment is expressed as a proportion of the rhyme, the same substantive prediction now translates statistically to predicting a null effect. That is, if the peak is somehow tracking the syllable boundary, then alignment defined as a proportion of the syllable rhyme should be roughly constant.

Strictly speaking, such constancy is not observed in our reanalysis. Indeed, the ANOVA using *Proportional alignment* as the dependent variable reveals a large and significant interaction between VOWEL LENGTH and SUFFIX, $F(1,659) = 26.91$, $p < 0.0001$, which shows that alignment is not conditioned in a constant way by the syllable boundary. However, our results suggest that a proportional definition of alignment may be worth further consideration. Most obviously, in the SUFFIX ABSENT condition (i.e. with monosyllabic test words like *maan* and *man*), the alignment expressed proportionally is indeed constant for short vowels and long vowels (0.40 and 0.39 respectively), despite substantial differences in the duration of the rhyme (178 vs 207 ms., respectively, both including the postvocalic consonant). In the SUFFIX PRESENT condition, the proportions are somewhat larger: alignment is at 0.50 of the rhyme in words with short vowels and 0.61 of the rhyme in words with long vowels, with rhyme durations of 132 and 145 ms. respectively. However, since these are the cases where the treatment of the intervocalic consonant following the stressed vowel is dependent on one's analysis of Dutch syllable structure, it is possible that a different definition of the rhyme might make these proportions appear more constant. We do not pursue this matter further here, but we cannot rule out the possibility that a proportional definition of alignment may ultimately prove to be "correct". We also suggest that a carefully constructed experiment comparing the predictions of proportional and absolute-interval definitions of alignment would be a valuable contribution to future research.

In other respects, the adoption of a proportional definition of alignment does not affect our previous conclusions. There is still no interaction between PREFIX and SUFFIX ($F < 1$), corroborating the conclusion that effects of stress clash are either absent or very small. The main effect of PREFIX remains significant ($p < 0.0001$), with alignment being later when there is a prefix (0.50) than when there is not (0.45); again, this suggests that small right-context effects due to "time pressure" are real. The main effect of SPEAKER also remains significant, and individual speakers' alignment differences persist when alignment is expressed proportionally; this is a useful finding, as it suggests that the main effect of SPEAKER observed in all previous analyses was not an artefact of speech rate. The previously reported interactions of other factors or interactions with SPEAKER remained stable and small.

5.3. Summary: does the definition of the variables matter?

In this section of the paper we have considered the possibility that we failed to find the expected interactions in our main experiment because of the way we defined our dependent variable of alignment. We did this by running the same analyses as in section 3 of the paper using different quantitative definitions of alignment, including two with different segmental reference points (*Prieto peak delay* and *Silverman peak delay*), one with a pitch-related reference point (*Rise time*), and one in which alignment is considered as a proportion of the syllable rhyme (*Proportional alignment*). Although we naturally found differences of detail, the overall pattern of our results is not affected.

In one sense it therefore makes little difference how alignment is quantitatively defined. However, we found reason to draw a few methodological conclusions from our reanalyses.

First, we believe that it is better to express alignment relative to a nearby acoustic landmark than a more distant one. The more distant the landmark, the greater the variance, and the greater the likelihood of uninformative correlations. The same point has been made on a somewhat different basis by Atterer and Ladd (2004). Specifically in the present case, we believe that the results using either our *H-C1* variable or *Silverman peak delay* are more dependable than results based on defining alignment relative to the beginning of the accented syllable, either explicitly (*Prieto peak delay*) or implictly *(Rise time).*

Second, we believe that it is better to express alignment relative to a segmental landmark rather than to another f0 landmark: the conclusions from analyses based on *Rise time* were sometimes at odds with those based on the other three alignment measures, and in the multiple regression analyses *Rise time* consistently led to a poorer model fit. Since, as we noted above, *Rise time* is an obvious choice of dependent variable for any theory that treats f0 changes as the building block of intonation, this finding represents further evidence against such theories, and in favour of the idea that f0 changes are best treated simply as transitions between segmentally anchored f0 targets.

Finally, we leave the issue of proportional vs. absolute-duration measures of alignment unresolved. We note that we did not find any strict constancy of proportional alignment relative to the syllable rhyme, which is what the strongest version of a proportional alignment theory would lead us to expect. However, we also note the existence of cases (notably the case of monosyllabic test words, i.e. the SUFFIX ABSENT condition) where absolute measures suggest that alignment is affected by our experimental manipulations but the proportional measure suggests that there is no effect. This may be coincidence, or it may suggest that proportional measures more accurately reflect the factors underlying alignment.

6. Conclusions

We may summarise the paper's principal findings as follows:

- There is a difference between Dutch nuclear and prenuclear accents with respect to the alignment of f0 peaks: in nuclear accents the peak is aligned earlier and is less susceptible to right context effects such as the effect of stress clash. The fact that nuclear accents exhibit earlier alignment confirms reports on other languages.
- There is an effect of vowel length on alignment in Dutch which – contrary to our earlier findings (Ladd et al. 2000) – may be at least partially independent of syllable structure. We say that this "may be" independent of syllable structure, because we acknowledge the possibility that an appropriately chosen proportional definition of alignment would reveal that there is indeed a relation between alignment and syllable structure. On the basis of our results we are unable to rule out such a possibility.
- In our present state of understanding, the most appropriate quantitative variables for expressing f0/segmental alignment are those that define alignment as the time interval between the f0 target in question and a *nearby* segmental landmark; the more distant the landmark, the greater the variance.

We hope that the first two of these conclusions will stimulate further research, and we believe that the third will provide subsequent researchers with important methodological guidance.

Footnotes:

1. This analysis of Dutch syllable structure is more or less universally taken for granted; see e.g. Booij (1995) and van der Hulst (1984). Among other things, this analysis is manifested in various conventions of Dutch orthography, notably (a) the convention that most long vowels are spelled with a single letter in open syllables and with a double letter in closed

syllables, and (b) the convention that intervocalic consonants are spelled with a single letter following a long vowel and with a double letter following a short vowel. There is also experimental evidence that Dutch native speakers are sensitive to syllable structure, defined according to this standard analysis, when they perform various psycholinguistic word-manipulation tasks (e.g. Schiller, Meyer and Levelt, 1997).

2. Due to an oversight, the materials included two items (13 and 24) that begin with an unstressed syllable (a lexical prefix) even in the 'PREFIX absent' condition. The data reported here include these items. Analyses and means for data excluding those items were virtually identical, showing only very small numerical differences.

3. Readers, especially those who speak a non-rhotic variety of English, may wish to know that coda /r/ is pronounced as a consonant in most varieties of Dutch and that this was the case with all our speakers.

4. The effect is particularly small in relation to the frame shift of 10 ms used in the f0 extraction. Nevertheless, this small numerical difference probably reflects a genuine difference, because the means represent a large number of data points, which has given the experiment a high level of statistical power.

5. We acknowledge the existence of autosegmental analyses without the phrase accent (e.g. Grabe 1998 on German and English; ToDI Gussenhoven et al., 1999 on Dutch; Frota 2002 on Portuguese), but we feel that the evidence for phrase accents presented in Grice et al. 2000 is compelling, and that our findings here represent further support for the general view that nuclear accents are tonally distinct from prenuclear accents. However, further discussion of this issue is well beyond the scope of this paper.

Acknowledgements:

This research was funded by the UK Economic and Research Council (ESRC) under grant no. R000-23-7447 to Edinburgh University, Principal Investigator D. R. Ladd. We are grateful to Carlos Gussenhoven, for providing access to the Nijmegen phonetics lab and for recruiting the speakers; Jan van Santen for useful discussions; Wladyslaw Cichocki for advising us on statistical analysis; and Michael Bennett, Eddie Dubourg, Cedric Macmartin and Stewart Smith for technical support in Edinburgh. The paper was considerably improved by comments from Jonathan Harrington and an anonymous referee.

References:

Abercrombie, David (1964). Syllable quantity and enclitics in English. In D. Abercrombie et al. (eds.), *In Honour of Daniel Jones,* London: Longman, pp. 216-222.

Arvaniti, A., Ladd, D. R. and Mennen, I. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics,* 26, 3-25.

Arvaniti, A.. Ladd, D. R. and Mennen, I. (forthcoming). The intonation of Greek polar questions and emphatic statements: fundamental frequency and intonational phonology. In preparation.

Booij, G. (1995). *The Phonology of Dutch*. Oxford: Clarendon Press.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund, Sweden: Gleerup

Caspers, J. and van Heuven (1993). Effects of Time Pressure on the Phonetic Realization of the Dutch Accent-Lending Pitch Rise and Fall. *Phonetica,* 50, 161-171.

Face, T. (2002) Local intonational marking of Spanish contrastive focus. *Probus* 14(1): 71-92.

Frota, S. (2000*) Prosody and focus in European Portuguese: Phonological phrasing and intonation.* New York: Garland Publishing.

Frota, S. (2002) Tonal association and target alignment in European Portuguese nuclear falls. In Gussenhoven, Carlos; Warner, Natasha (eds). *Laboratory Phonology 7*, Berlin/New York: Mouton de Gruyter, 387-418.

Grabe, E. (1998). *Comparative intonational phonology: English and German.* Ph.D. thesis, Max Planck Institute for Psycholinguistics, Nijmegen.

Grice, M., Ladd, D. R. and Arvaniti, A. (2000). On the place of phrase accents in intonational phonology. *Phonology,* 17, 143-185.

Gussenhoven, C., Rietveld, T. and Terken, J. (1999). TODI: Transcription of Dutch Intonation. http://lands.let.kun.nl/todi/todi/home.htm

Gussenhoven, C., Rietveld, T. and Terken, J. (1999). ToDI: Transcription of Dutch Intonation. Available through: http://lands.let.kun.nl/todi/todi/home.htm

't Hart, J., Collier, R. and Cohen, A. (1990). *A perceptual study of intonation: An experimental-phonetic approach*. Cambridge: Cambridge University Press.

Hualde, J. I. (2002). Intonation in Spanish and other Ibero-Romance languages: Overview and status quaestionis. In *Romance Phonology and Variation. Selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000*, ed. by Caroline Wiltshire & Joaquim Camps, 101-115. Amsterdam: Benjamins.

Hulst, H. van der (1984). Syllable Structure and Stress in Dutch. Dordrecht: Foris.

Ishihara, T. (2003). A phonological effect on tonal alignment in Tokyo Japanese. *Proceedings of ICPhS*, Barcelona, pp. 615-618.

Ladd, D. R. (1996). *Intonational Phonology.* Cambridge: Cambridge University Press.

Ladd, D. R., Faulkner, D., Faulkner, H. and Schepman, A. (1999). Constant “segmental anchoring” of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America,* 106, 1543-1554.

Ladd, D. R., Mennen, I. and Schepman, A. (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America,* 107, 2685-2696.

Lehiste, I. (1970). Suprasegmentals. Cambridge: MIT Press.

Lickley, R., Schepman, A. and Ladd, D. R. (forthcoming). Alignment of "Phrase Accent" Lows in Dutch Falling-Rising Questions: Theoretical and Methodological Implications. Under revision, *Language and Speech.*

Nibert, H. J. (2000). Phonetic and phonological evidence for intermediate phrasing in Spanish intonation. Doctoral dissertation, University of Illinois at Urbana-Champaign, ms.

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

Palmer, H. (1922). *English Intonation with Systematic Exercises*. Cambridge: Heffer.

Pierrehumbert, J. B. (1980). *The Phonology and Phonetics of English Intonation.* PhD thesis, MIT.

Prieto, P., van Santen, J. and Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics,* 23, 429-451.

Rietveld, T. and Gussenhoven, C. (1995). Aligning pitch targets in speech synthesis: effects of syllable structure. *Journal of Phonetics*, 23, 375-385

Selkirk, E. (1984) *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Schiller, N. O., Meyer, A. S., & Levelt, W. J. M. (1997). The Syllabic Structure of Spoken Words: Evidence from the Syllabification of Intervocalic Consonants. *Language and Speech*, 40, 103-140.

Shattuck-Hufnagel, S., Ostendorf, M., and Ross, K. (1994). Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics*, 22, 357-388.

Silverman, K. and Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. In J. Kingston and M. Beckman (eds.), *Papers in Laboratory Phonology I*, Cambridge: Cambridge University Press, pp. 72-106.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert J. and Hirschberg, J. (1992). TOBI: A standard for labeling English prosody. *Proceedings of ICSLP* 2: 867-870. Banff, Alberta, Canada.

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

Turk, A. & White, L.S. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics* 27.171-206.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25: 61-83

Xu, Y. (1998). Consistency of tone-syllable alignment across different syllable and speaking rates. *Phonetica,* 55, 179-203.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27, 55-105.

Appendix:

Items used in the experiment. The eight conditions are marked in the items table.

	Long VOWEL LENGTH, SUFFIX absent, PREFIX absent		Long VOWEL LENGTH, SUFFIX absent, PREFIX present
1	Hij kon de maan zien.	61	Ze had de maan gezien.
2	Hij wil de laan zien.	62	Hij had de laan gezien.
3	Ze wil haar naam geven.	63	Ze had haar naam gegeven.
4	Hij wil een haan kopen.	64	Hij had een haan gekocht.
5	Ze zou z'n haar knippen.	65	Ze had z'n haar geknipt.
6	Hij zou z'n loon krijgen.	66	Hij had z'n loon gekregen.
7	Ze zou loom worden.	67	Ze was loom geworden.
8	Hij kon een moor zien.	68	Hij had een moor gezien.
9	Ze wil met Noor spreken.	69	Ze had met Noor gesproken.

10 Hij zou room kopen.

11 Ze zou leer kopen.

12 Hij zou meel kopen.

13 Ze zou leem verzamelen.

14 Hij zou met Leen spreken.

15 Ze wil het meer zien.

Long VOWEL LENGTH,

SUFFIX present, PREFIX absent

16 Hij wou de manen zien.

17 Ze zou de namen krijgen.

18 Hij zou hanig worden.

19 Ze kon de hanen zien.

20 Hij wil een hamer kopen.

21 Ze kon de molen zien.

22 Hij kon de holen zien.

23 Het zou romig worden.

24 Hij zou de lonen betalen.

25 Ze zou Rome zien.

26 Hij wil Wenen zien.

27 Ze wou de Here zien.

28 Hij kan de hemel zien.

29 Ze zou lenig worden.

30 Hij zou lelijk worden.

Short VOWEL LENGTH,

SUFFIX absent, PREFIX absent

31 Ze kon een man zien.

32 Hij kon een ram zien.

33 Ze wil een lam zien.

70 Hij had room gekocht.

71 Ze had leer gekocht.

72 Hij had meel gekocht.

73 Ze had leem verzameld.

74 Hij had met Leen gesproken.

75 Ze had het meer gezien.

Long VOWEL LENGTH,

SUFFIX present, PREFIX present

76 Hij had de manen gezien.

77 Ze had de namen gekregen.

78 Hij was hanig geworden.

79 Ze had de hanen gezien.

80 Hij had een hamer gekocht.

81 Ze had de molen gezien.

82 Hij had de holen gezien.

83 Het was romig geworden.

84 Hij had de lonen betaald.

85 Ze had Rome gezien.

86 Hij had Wenen gezien.

87 Ze had de Here gezien.

88 Hij had de hemel gezien.

89 Ze was lenig geworden.

90 Hij was lelijk geworden.

Short VOWEL LENGTH,

SUFFIX absent, PREFIX present

91 Ze had een man gezien.

92 Hij had een ram gezien.

93 Ze had een lam gezien.

34 Hij zou een ham krijgen.

35 Ze wou in z'n wang knijpen.

36 Er zou een non komen.

37 Ze zou wol kopen.

38 Hij zou een rol kopen.

39 Ze kon een mol zien.

40 Hij kon een hol zien.

41 Ze wou een ren kopen.

42 Hij zou een hel krijgen.

43 Ze zou met Len spreken.

44 Hij kon een hen zien.

45 Ze wou de rem kiezen.

Short VOWEL LENGTH,

SUFFIX present, PREFIX present

46 Er zouden mannen komen.

47 Ze zou een hanger kopen.

48 Hij zou wallen bouwen.

49 Ze mag z'n wangen voelen.

50 Hij wou een mangel kopen.

51 Ze konden mollen zien.

52 Hij wil een roller kopen.

53 Ze kon een hommel zien.

54 Hij wil z'n longen testen.

55 Ze zou honger krijgen.

56 Hij kon een renner zien.

57 Ze zou een menger spreken.

58 Hij kon de rellen zien.

59 Ze kon de hennen zien.

60 Hij zou een hengel krijgen.

94 Hij had een ham gekregen.

95 Ze had in z'n wang geknepen

96 Er was een non gekomen.

97 Ze had wol gekocht.

98 Hij had een rol gekocht.

99 Ze had een mol gezien.

100 Hij had een hol gezien.

101 Ze had een ren gekocht.

102 Hij had een hel gehad.

103 Ze had met Len gesproken.

104 Hij had een hen gezien.

105 Ze had de rem gekozen.

Short VOWEL LENGTH,

SUFFIX present, PREFIX present

106 Er waren mannen gekomen.

107 Ze had een hanger gekocht.

108 Hij had wallen gebouwd.

109 Ze had z'n wangen gevoeld.

110 Hij had een mangel gekocht.

111 Er waren mollen gezien.

112 Hij had een roller gekocht.

113 Ze had een hommel gezien.

114 Hij had z'n longen getest.

115 Ze had honger gekregen.

116 Hij had een renner gezien.

117 Ze had een menger gesproken.

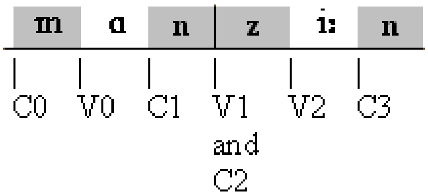
118 Hij had de rellen gezien.

119 Ze had de hennen gezien.

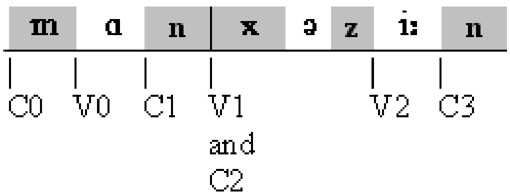
120 Hij had een hengel gekregen.

Figure 1

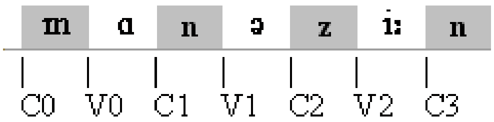
SUFFIX absent, PREFIX absent:



SUFFIX absent, PREFIX present:



SUFFIX present, PREFIX absent:



SUFFIX present, PREFIX present:

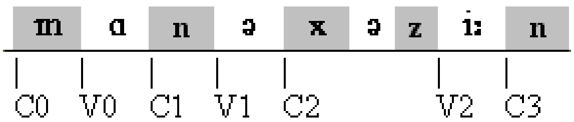


Figure 1 Caption:

Measurement points for the four combinations of SUFFIX and PREFIX absent vs present. Consonants are shaded grey, while vowels are white. Note that the measurement points were the onsets of the relevant segments.

Figure 2:

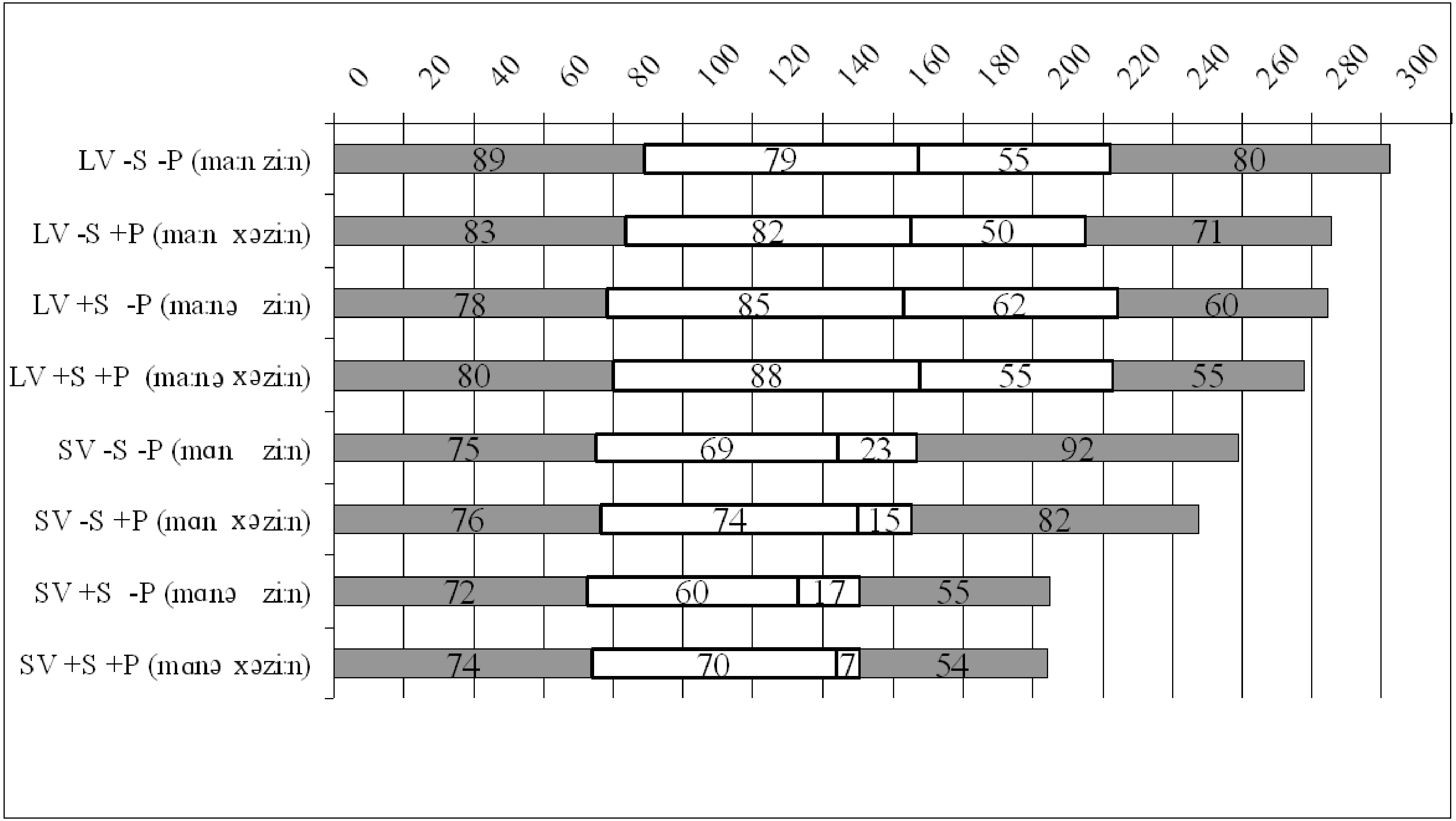


Figure 2 Caption on separate sheet.

Figure 2 Caption:

Mean durations of the CVC segments of the target word and location of H (the f0 peak) for the target syllable, as a function of the eight experimental conditions. The labels LV and SV refer to Long vs Short vowels, + or – refer to the absence and presence of the Suffix (S) or Prefix (P). Grey-shaded parts of the horizontal bars are the durations of consonants (e.g. /m/ and /n/ in *man*), while the white portion is the vowel. The vowel portions are divided into two parts, separated by vertical lines that indicate the location of H. The portion of white bar *after* the vertical line represents H to C1 (our dependent variable), while the portion of the white bar *before* the vertical line is the pre-peak remainder of the vowel.

Figure 3:

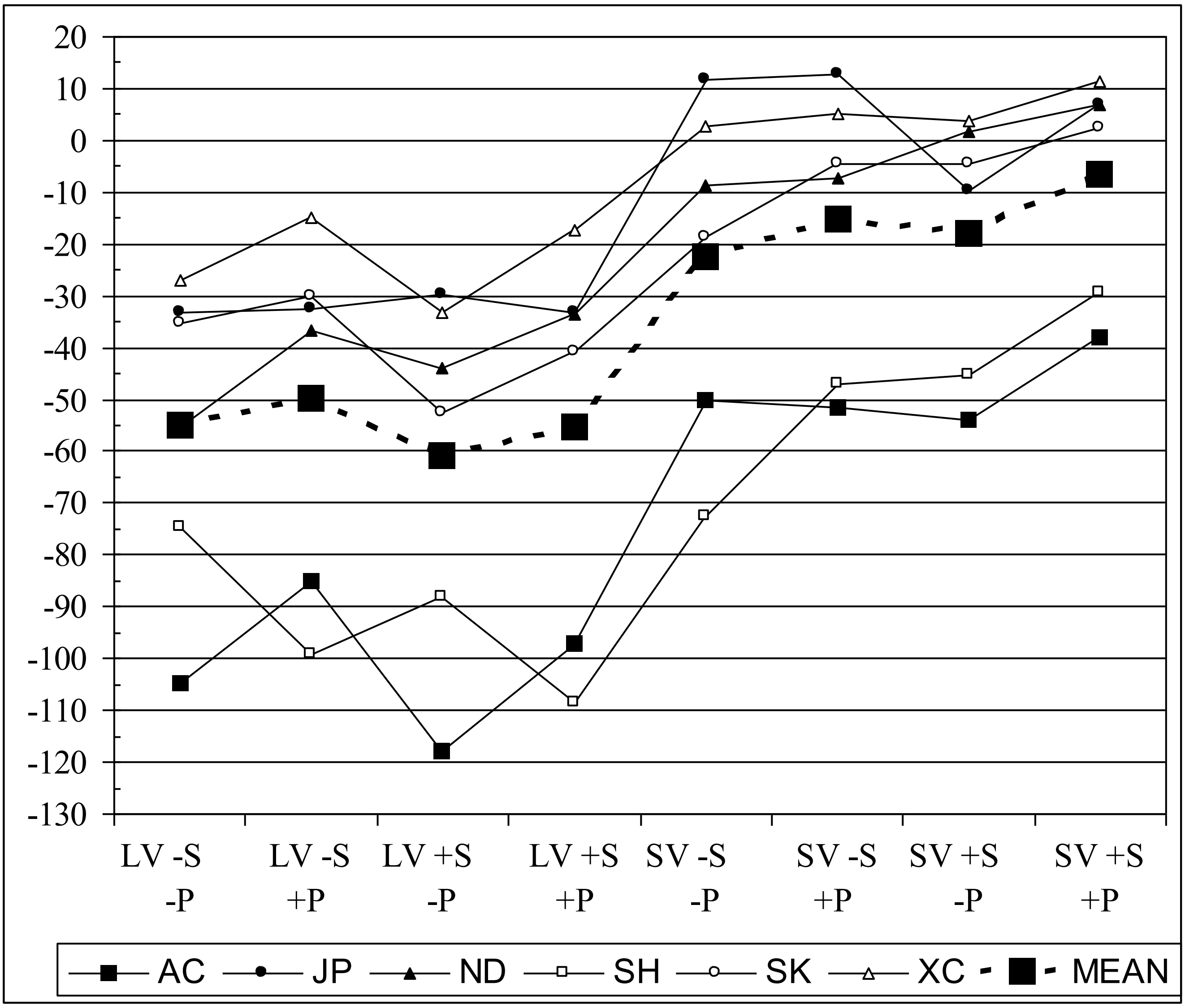
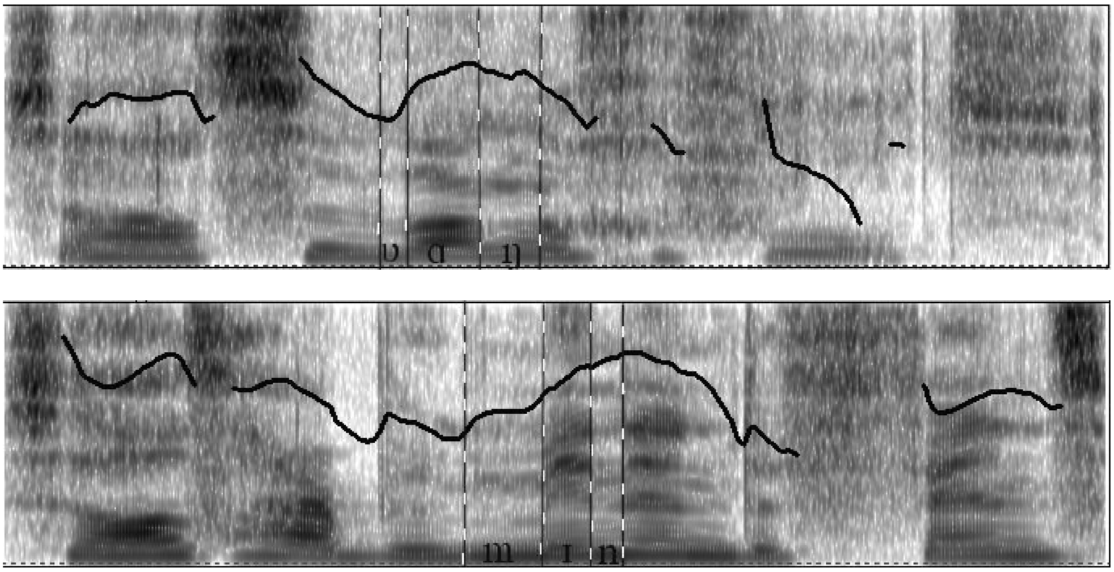


Figure 3 Caption:

Mean values of H – C1 for the eight experimental conditions, separated by speaker. LV and SV refer to Long vs Short vowels, + or – refer to the absence and presence of the Suffix (S) or Prefix (P). Negative numbers indicate alignment before C1, while positive numbers indicate alignment after C1.

Figure 4



Caption figure 4

Nuclear (top panel) and pre-nuclear (bottom panel) alignment: Spectrograms (0 – 8000 Hz) with superimposed f0 tracks (100-300 Hz), both with durations of 1.32 seconds. Added vertical lines represent, from left to right, C0, V0, C1 and V1, with relevant target segments indicated between those vertical lines.

The figures show speaker SK reading *Ze had z'n <u>wang</u>en gevoeld*. /z↔ hAt s↔n ∪ςAN↔ ξ↔∩vυ⎤lt/ 'She had felt his cheeks' in the top panel and *Ze had haar be<u>min</u>nende fans*... /z↔ hAt hα⎤ρ b↔∪mIn↔nd↔ ∪fEns/ in the bottom panel. (The target syllables are underlined.) The latter fragment was excerpted from a longer sentence which continues *...achter hun rug uitgelachen* (Lit. 'She had her adoring fans…behind their backs laughed-at', i.e. 'She had laughed at her adoring fans behind their backs'), and formed part of Experiment 2 in Ladd et al. 2000. The f0 peak in the top panel (256 Hz) occurs towards the end of the /A/, while that in the bottom panel (262 Hz) occurs at the beginning of the /↔/ which follows the /n/.

Table I

		Duration of Onset Consonant (C0 to V0) (in ms)	**Duration of Vowel (Vo to C1) (in ms)**	**Duration of Postvocalic Consonant (C1 to V1) (in ms)**
VOWEL LENGTH	Long	**82**	**139**	**66**
	Short	**74**	**84**	**71**
SUFFIX	Absent	**81**	**111 overall** Long: 133 Short: 91	**81 overall** Long: 75 Short: 87
	Present	**76**	**111 overall** Long: 145 Short: 77	**56 overall** Long: 57 Short: 54

Table I Caption:

Mean durations of the segments making up the CVC string of the target syllable, as a function of VOWEL LENGTH and SUFFIX. Note that these some of these main effects and two-way interactions were subject to minor interactions with SPEAKER, as reported in section 3.3.