Temporal Marking of Accents and Boundaries

Cambier - Langeveld, G.M.

Publication date
2000

2. THE DOMAIN OF FINAL LENGTHENING IN DUTCH[1]

2.1. Introduction

Speakers have numerous ways at their disposal to mark the prosodic structure of their speech. One way for speakers to signal prosodic structure is by indicating the grouping of words into higher level constituents, using a variety of prosodic features, so that the structure within the typically continous speech flow is made clear to the listener. From the literature it is clear that the prosodic features pause, pitch and/or lengthening of the preboundary segments are used to mark prosodic boundaries (e.g. Cooper & Paccia-Cooper, 1980; Nooteboom, Brokx & de Rooij, 1976; de Rooij, 1979). Which prosodic features are used to mark a certain boundary may depend on the language and on the type of boundary involved (de Pijper & Sanderman, 1994), but is also to some degree variable, e.g. across speakers and speech rates (Caspers, 1994; Sanderman, 1996). The lengthening of preboundary segments, which will be our main concern in this chapter, is referred to as **final lengthening**.

In the past, final lengthening was thought to indicate syntactic boundaries (Klatt, 1975). With the development of prosodic phonology, final lengthening is now assumed to occur at the edges of prosodic boundaries (just like other phonetic and phonological processes; cf. §1.1). A final lengthening effect marking the end of an utterance or intonational phrase has been firmly established by numerous phonetic studies (e.g. Crystal & House, 1988). Other prosodic boundaries may be marked by lengthening of the preboundary segment(s) as well (Beckman & Edwards, 1990). Moreover, the amount of final lengthening has been shown to be related to the depth of the following boundary, both as produced by the speaker and as required by the listener (Nooteboom & Doodeman, 1980; Gussenhoven & Rietveld, 1992; cf. §1.4.1). Thus, a deeper prosodic boundary is, and should be, marked by more final lengthening.

Although final lengthening is a well documented phenomenon, relatively little attention has been paid to the question of which segments are affected. In Cooper & Paccia-Cooper (1980) it is said that most of the syllable lengthening before utterance boundaries in English is due to lengthening of the vowel, and that only sonorant and

---

[1] Parts of this chapter have also been published as Cambier-Langeveld (1997) and Cambier-Langeveld, Nespor & van Heuven (1997).

continuant segments can be lengthened. This was questioned in a study on Dutch by Hofhuis, Gussenhoven & Rietveld (1995), in which all types of consonants (liquids, nasals, fricatives and stops) were found to be lengthened due to a following boundary. Moreover, these consonants, which were in coda position, were lengthened more than the preceding vowel, showing that final lengthening is strongest in the final segment (even if this is a consonant) and gets weaker as the segment is further to the left, i.e. away from the boundary. This is in agreement with Berkovits (1993, 1994), who found that in Hebrew the final lengthening effect is progressively distributed from left to right across the segments of the final syllable. She notes that studies of final lengthening which focus only on the amount of lengthening in the final vowel or syllable overlook a significant aspect of the lengthening effect. The same point was made earlier in Edwards & Beckman (1988), where the sonority profile of English final syllables is shown to change as a result of a disproportionate lengthening of the second part of this syllable (in contrast to the sonority profile of an accented syllable, which remains unaffected; cf. §1.2). They note that for an adequate account of lengthening effects and rhythm, it is not sufficient to state the lengthening in terms of millisecond values or durational ratios for acoustic segments; we also need a description of its distribution.

The studies discussed so far have only been concerned with lengthening within the final syllable, be it the final vowel (e.g. Nooteboom & Doodeman, 1980), the final rhyme (e.g. Gussenhoven & Rietveld, 1992) or the whole final syllable (e.g. Edwards & Beckman, 1988; Berkovits, 1994). Only a few studies have investigated the possibility of an effect extending to the left beyond the final syllable. Some of these will be discussed below. As far as we are aware, there have been no perception experiments on final lengthening in anything other than the final syllable, or on the importance of the specific distribution of final lengthening over the preboundary segments for boundary perception.

One study which is concerned with the domain of final lengthening in production is an American English database study by Wightman et al. (1992). Their results seem to indicate that in speech production final lengthening is confined to the rhyme of the final syllable, as is assumed in several other studies. However, as already mentioned in Chapter 1 (§1.4.1), their analysis considers only four pre-boundary units: the final coda consonants (if any), the final vowel, any segments between the last stressed vowel and the final vowel, and the last stressed vowel. It is important to note that their definition of a stressed vowel implies that any number of syllables may intervene between the last stressed syllable and the relevant boundary: " [...] a vowel was marked as stressed if two criteria were satisfied: (1) the vowel was marked in the lexicon (dictionary) as receiving lexical stress, and (2) the match between the acoustic waveform and the stressed vowel model in the SRI recognizer was better than the match between the acoustic waveform and the unstressed vowel model." (Wightman et al., 1992; p. 1713). The third increment (from 'final vowel' to 'any segments between the last stressed vowel and the final vowel'), therefore, is a highly variable one, and may be too crude to reveal small effects in more than the final rhyme, e.g. in the final syllable, or in the final (disyllabic) foot. Moreover, factors which may influence the effect under observation can be overlooked in database studies, because the materials are not designed to find such effects. For example, word structure, (final) syllable weight, word stress, etc. could influence the size of the domain of final lengthening, as would be the case if the unit affected by final lengthening is the final foot (which can consist of one or two syllables, depending mainly on the segmental structure of the final syllable; see §1.3.2.4).

There is evidence that, at least in some languages, final lengthening can be found in more than the final syllable. Kohler (1983), using the German words *eine* 'a (fem.)' and *einige* 'some' (both with initial stress), found that the effect spreads across the whole word, decreasing towards the beginning of the word. The largest effect is always found in the final syllable. The results of a production experiment on Dutch by Hofhuis (in prep.) indicate that under certain conditions final lengthening may spread across more than the final syllable in this language, too. To recall (cf. §1.4.1), her data consist of monosyllabic and disyllabic words, where the latter differ from the former only in the addition of a word-final schwa, occurring at five different boundaries. In the disyllabic words, an effect of boundary depth on segment duration was found not only in the final syllable, but also in the penultimate rhyme (a full vowel), and sometimes even in the penultimate onset (i.e., in the whole word). If this larger domain of lengthening is due to the final schwa, then apparently the segmental content of the final syllable influences the size of the lengthened unit in this language. However, since Hofhuis' material contains only schwa-final disyllabic words, it is impossible to differentiate between foot/syllable structure, syllable weight or vowel quality as the relevant factor determining the domain of final lengthening. Furthermore, the lengthening of the penultimate syllable in Hofhuis' disyllabic words can be expressed as lengthening of the final foot, the whole word, or even some unit that has no status in the standard prosodic hierarchy, e.g. a unit starting with the last primary stressed syllable and including all following syllables up to the boundary (cf. the so-called 'Abercrombian foot'; Abercrombie, 1965).

In the next section, we will formulate the specific research questions for this chapter, which follow in part from the discussion of previous work on final lengthening above.

2.1.1. Research questions

In this chapter, we will report on a series of experiments on the production and perception of the domain of final lengthening in Dutch. In §2.2, possible factors influencing the domain of final lengthening in speech production are investigated. As discussed above, most studies on final lengthening have focused on the lengthening effect in the final syllable. In this syllable, the amount of lengthening has been shown to be related to the depth of the following boundary, yet very little is known about the effect of boundary depth on the domain that is lengthened. One could imagine that a deeper prosodic boundary not only triggers a larger amount of lengthening, but also increases the size of the unit affected.

To recall, in a study on English by Wightman et al. (1992), a final lengthening effect was found only in the final rhyme, yet evidence from a production experiment on Dutch (Hofhuis, in prep.) suggests that penultimate syllables can also be affected. The conditions under which final lengthening may extend further back than the final syllable are not quite clear, since in the latter study only disyllabic words with a final schwa were used. Word stress, foot structure, vowel quality or syllable weight could all have played a role here. Prior to coming to a proper description of the domain(s) of final lengthening, we need to understand the factors involved in determining the size of the lengthened unit and their exact influence. This leads us to the research questions in (1):

(1a) What is the relationship between boundary depth and the domain of final lengthening?

(1b) Do stress and/or the structure of the final syllable/word have an effect on the domain of final lengthening?

Following the production experiments, two perception experiments were run. The first was designed to check if the durational differences found in the production data suffice for boundary perception at the phonological phrase and the intonational phrase levels. This was done in the form of an identification experiment in which listeners were asked to indicate which type of utterance they thought the speaker had intended. The second was an acceptability experiment, in which listeners were presented with utterances and asked to judge the acceptability of the durational build-up of the test words. The perception experiments were designed to answer the following questions:

(2a) To what extent are the durational differences found in production used in boundary perception?

(2b) Are listeners sensitive to differences in the size of the domain over which a certain amount of lengthening is spread?

In §2.2, the production experiments will be discussed. The perception experiments will be described in §2.3.

2.2. Production experiments

2.2.1. Approach

To answer the research question in (1a), repeated below for convenience, material is needed in which a certain string of segments is followed by prosodic boundaries of different depths.

(1a) What is the relationship between boundary depth and the domain of final lengthening?

Any other factors, i.e. apart from the following boundary, which influence the duration of segments must be held as constant as possible. By constructing material such that the same string of segments is placed before the crucial boundary every time, it is possible to abstract away from inherent differences in duration between segments and from the influence of adjacent segments. Furthermore, prominence relations and overall speech rate must also be kept as constant as possible. In studies where one carrier phrase is used, only replacing one target word by another, it is likely that speakers will produce a consistent prominence pattern at a steady rate; however, the sentences in the present study could very well have different default patterns, which means that extra measures should be taken in order to ensure the elicitation of a constant prominence pattern. Since such prosodic features are not represented in spelling, one cannot make them 'given' when presenting the material to the speakers only on paper. It was therefore decided to present the material both visually (i.e. printed on paper) and auditorily. The auditory stimuli to be repeated by the speakers are meant to convey the desired (i.e. constant) intonation contour, prominence relations, speech rate etc. to the speakers, in the hope that when speakers are required to repeat an utterance they will copy these features. However, one should not allow the prosodic feature under investigation, i.e. (final) lengthening, to be influenced by the input which is given to the speakers. Rather, the input speech should contain no temporal markers at all, so that any temporal boundary markers in the speakers' reproduction must have been 'implemented' by the speakers themselves, thus reflecting natural lengthening effects in speech. It should be noted that this approach assumes that the correct prosodic features in the input speech will be copied, whereas the incorrect temporal structure will be automatically corrected by the speakers.

Human speech is not suitable as input for this experiment, since phonetic effects like lengthening presumably occur naturally and cannot be 'turned off' by a human speaker. We are therefore obliged to use synthetic speech for the input, since that allows us to produce speech without lengthening effects.

Diphone synthesis was chosen to be used for the input speech. A diphone consists of a transition between two phonemes, with roughly half a phoneme on either side (Dixon & Maxey, 1968). Thus, the Dutch word *prosodie* 'prosody' will be built up from the following diphones, where # indicates silence (at the beginning and end):

#p pr ro: o:s so: o:d di i#

The idea behind diphone synthesis is that different realisations of one phoneme differ least at a point halfway through that phoneme. In a sequence of diphones, the first half of one phoneme is adjoined to the second half of that same phoneme. Any discrepancies are then smaller and easier to adjust than those between two different phonemes. Elsendoorn & 't Hart (1982) were the first to make a set of (~2000 accented) Dutch diphones; Drullman & Collier (1991) made a new set which - unlike its predecessor - also included a set of (~1500) unaccented diphones. As each diphone is excerpted from the same position in the same type of (nonsense) word, there are no temporal effects in the original set. Rules for lengthening have been added to the synthesis program (DS; Rijnsoever, 1988), but can be turned on or off as the user of the program pleases.

The choice for a repetition task using synthetic speech as input may be well motivated, but also raises some new questions. Primarily, the assumption must be checked that speakers do indeed correct the temporal structure of the input speech, such that in their reproductions a final lengthening effect is found which they have implemented themselves. Apart from questions concerning the validity of the assumptions underlying the present approach, there are also a few practical issues which need to be resolved. The most important are:

- What kinds of different structures are needed in preboundary position in order to answer research question (1b) in some detail?
- Do the speakers require another feature, e.g. a boundary-marking pitch movement, at the major (prosodic) boundaries in order to grasp the structure of the utterances? In other words, what should be the intonation contour of the input?
- Does final lengthening interact in any way with lengthening due to prominence?

To answer these questions, a pilot study was carried out prior to the main experiment, which served both to check whether a repetition task is appropriate for the present research, and to give us some ideas for the set-up of the main experiment. The latter will include more target words, which will have segmentally varying final syllables in order to investigate whether the prosodic structure of these syllables has any influence on the domain of final lengthening.

In §2.2.2, a report is given of the preliminary experiment forming the pilot study; in §2.2.3 the main experiment will be described.

2.2.2. Pilot study

2.2.2.1. Introduction

The preliminary experiment described in this section not only serves as a try-out for the main experiment described in the following section, but is also necessary to answer some specific questions concerning the required form of the input for the repetition task.

As this research involves lengthening in speech production, one would like possible lengthening effects not to be influenced by the input utterances that speakers were asked to repeat. Consequently, the prosodic boundaries in the input are not to be marked by lengthening, so that if speakers do indeed produce speech with final lengthening when they are asked to repeat speech which does not contain any lengthening effects, such lengthening effects cannot have been copied, but are inserted by the speakers, and can therefore be assumed to reflect general (natural) lengthening effects in human speech.[2]

One question which will be answered in the present study is whether speakers need some other prosodic feature to mark the major boundaries in the input speech. Two

[2] This methodology may have a restraining influence on the production of final lengthening, since speakers may try to stay close to the temporal organisation of the input utterances. However, this does not enervate the idea that whatever lengthening *is* found reflects natural lengthening effects. We will come back to this issue in Chapter 5.

possible candidates come to mind; pauses and boundary-marking pitch movements. The former, however, is also a temporal marker. Therefore, intonational marking is the only option left when speech without temporal effects of any kind is preferred. Each stimulus sentence will be synthesized with three different intonation contours, which will be described below in §2.2.2.2.

Another variable in the pilot study is focus distribution. The target word is put either in or out of focus using a precursor question, and the target word is realized with or without a pitch accent on the primary stressed syllable accordingly. This variable is added to see if final lengthening interacts in any way with accentual lengthening, and to determine if in the main experiment the target words should be placed in or out of focus (or both).

2.2.2.2. Material

The one target word which is used for this experiment is a fairly long one, with four syllables: *rododendron* /roːdoːˈdɛndrɔn/ 'rhododendron' (kind of plant). Such a long word allows us to distinguish between possible domains of final lengthening, such as the whole word, the final syllable *dron*, or the syllables starting with the primary stressed syllable, i.e. *dendron*. The results should give us some idea of the relevant domain(s), and of how the target words for the main experiment should be varied in order to answer research question (1b) in detail.

Four sentences containing the target word were composed such that they contained the same number of syllables in total, as well as the same segmental material. Each sentence consists of an intonational phrase plus a parenthetical, i.e. an additional remark. The parenthetical was realized without a pitch accent, so that every utterance contained only one pitch accent. In the case where the target word is utterance-final, the parenthetical precedes the intonational phrase that contains the target word. In the other cases the parenthetical follows the intonational phrase containing the target word. The four sentences, given in (3), differ in the type of boundary that follows the target word (underlined):

(3)	PW-boundary:	Piet wil die rare <u>rododendron</u>planten, gek als hij is. 'Piet wants those strange rhododendron plants, crazy as he is'
	PhP-boundary:	Piet wil die rare <u>rododendron</u> planten, gek als hij is. 'Piet wants to plant that strange rhododendron, crazy as he is'
	IP-boundary:	Piet wil die rare <u>rododendron</u>, plantengek als hij is. 'Piet wants that strange rhododendron, plantcrazy as he is'
	U-boundary:	Plantengek als hij is wil Piet die rare <u>rododendron</u>. 'Plantcrazy as he is, Piet wants that strange rhododendron'

The shallowest boundary is that between the two members of a compound; syntactically, they form one word, but prosodically they form two words (cf. §1.3.2.3). This will be called the PW-boundary (*p*rosodic *w*ord-boundary). In the second sentence, the same elements that constitute the two members of a compound in the first sentence now form two syntactic words, namely a noun (the target word itself) and a verb. Since the noun phrase of which the target word is the head also includes a determiner and an adjective, it is branching and thus cannot be restructured with the verb into one phonological phrase (cf. §1.3.2.1). Consequently, according to the phonological phrase formation rules discussed in §1.3.2.1, there is a phonological phrase-boundary between the target word and the verb; it will therefore be called the PhP-boundary (*ph*onological *p*hrase-boundary).

In the third sentence the *i*ntonational *p*hrase-boundary (or IP-boundary), indicated by an orthographic comma, is in a different place in the utterance, since it must now follow immediately after the target word. Segmentally, however, the target word is still followed by the same word (sequence) as in the other two sentences. As the total length of the *utterance* was kept constant, this implies that the intonational phrase in which the target word occurs is shorter in the IP-boundary condition than in the PW- and PhP-boundary condition (and the following parenthetical is longer, but this is not relevant since lengthening effects in the target word only are studied). The length of an IP may affect the rate of speech, in terms of syllables per second, in which it is uttered; we will come back to this issue in §2.2.3.2.

The final sentence places the parenthetical in front of the intonational phrase, so that the target word concludes the utterance. This also necessitates turning the order of the subject (*Piet*) and the main verb (*wil*) around. This will be referred to as the U-boundary (*u*tterance-boundary). In all, the four prosodic boundaries are: PW, PhP, IP and U.

To answer the question whether final lengthening interacts in any way with lengthening due to prominence, each of the above four sentences occurs with two different focus distributions. Using precursor questions, either the noun phrase of which the target word is the head or the proper name *Piet* is put in focus, so that the pitch accent falls on the stressed syllable of *rodo*den*dron* or on *Piet*, as exemplified in (4) (capitals indicate the location of the pitch accent):

(4) Q: Wat wil Piet?
'What does Piet want?'
A: Piet wil die rare rodoDENdron, plantengek als hij is.

Q: Wie wil die rare rododendron?
'Who wants that strange rhododendron?'
A: PIET wil die rare rododendron, plantengek als hij is.

These precursor questions were spoken by the author and recorded directly onto a computer disk. (As only the answers, i.e. the test sentences, were to be repeated, the temporal structure of the preceding questions is not relevant, so that human speech could be used; the questions serve only to determine the place of the accent.)

In accordance with the focus distribution imposed on the sentence by the preceding question, an accent-lending pitch movement was realized either on the primary stressed syllable of the target word (*den*) or on *Piet*. The IP-boundary in each sentence was either not marked at all, marked by a single pitch movement (either a rise or a fall) or by a double pitch movement (i.e. a rise-fall combination). The timing, size and excursion of these movements are determined by the synthesis program accompanying the diphone set used to generate the utterances. The diphone set used is named bloem30_nostand.ind, which was made in Eindhoven by Drullman and Collier, and spoken by the native Dutch professional speaker Bloemendal. This diphone set was found to afford the highest intelligibility for Dutch by van Bezooijen & Pols (1993).[3] The declination of the pitch contour was also implemented automatically by the diphone synthesis program. Only accented diphones were used, so that there were no durational differences between diphones set in in accented and unaccented syllables, and all temporal effects were turned off.

When the IP-boundary is not intonationally marked, the intonation contour consists only of a 'pointed hat' on the accented syllable, i.e. an accent-lending early rise immediately followed by a late fall in the same syllable. This contour will be referred to as 1&A (cf. 't Hart, Collier & Cohen, 1990; all following transcriptions of intonation will also be in the tradition of this work. See also Appendix A). The double boundary-marking pitch movement also consists of a rise immediately followed by a fall, yet these are timed later in the syllable, and are referred to as 2&B. In this contour, the accented syllable is still marked by the same pointed hat. In the PW-, PhP- and IP-boundary sentences, the accent precedes the IP boundary, so that the

[3] Superior synthesis systems may have appeared more recently, but these were not available at the time the experiment was run.

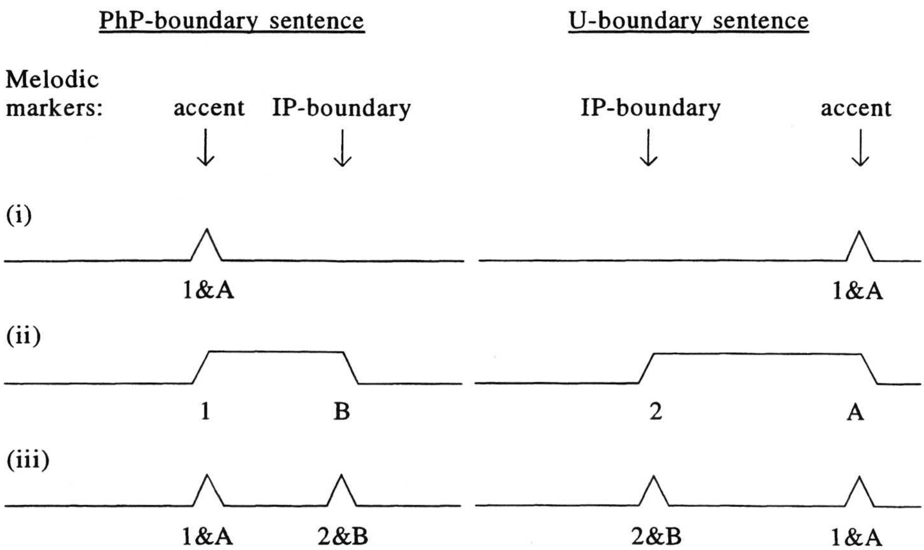
total intonation contour is 1&A 2&B. In the U-boundary sentence, the IP-boundary precedes the accent, so that the order in which the pitch movements occur is switched around into 2&B 1&A.

For the last intonation contour, the rise of the first rise-fall combination of the previous contour is joined to the fall of the second rise-fall combination to form a so-called 'flat-hat' contour. When the accent precedes the IP-boundary, the result is a contour called 1B (with an accent-lending early rise, 1, followed by a boundary-marking fall, B, in another syllable). When the accent follows the IP-boundary, the rise is boundary-marking (2) and the fall is accent-lending (A). The total intonation contour for the single movement boundary marker is therefore 1B or 2A.

In sum, the boundary between the intonational phrase and the parenthetical is melodically marked in three ways, giving rise to three different intonation contours:

(5) (i) no melodic boundary marker (total intonation contour: 1&A)
(ii) light single-movement melodic boundary marker (total intonation contour: 1B/2A)
(iii) heavy double-movement melodic boundary marker (total intonation contour: 1&A, 2&B)

Examples:



Piet wil die rare rododendron planten, gek als hij is. Plantengek als hij is wil Piet die rare rododendron.

Finally, every sentence was synthesized once with the target word *rododendron* in it, and once again with the target word replaced by the nonsense word /pa:pa:'pa:pa:/, i.e. a reiterant version of the lexical target word. This word has the same number of syllables and was made to have the same melodic structure as the target word. Reiterant speech allows us to study prosodic phenomena while abstracting from segmental influences (Liberman & Streeter, 1978; Nakatani & Schaffer, 1978).

The complete material for the pilot study thus comprizes 48 utterances: 1 target word x 4 boundary depths x 2 focus distributions x 3 intonation contours x 2 versions (lexical and reiterant) of the target word.

2.2.2.3. Subjects and procedure

Three male and two female speakers participated in the experiment. Three of them (two male and one female) may be called phonetically trained, while the other two had no experience in taking part in an experiment such as this and no phonetic background. Both trained and untrained speakers are included because for the main experiment it is important to know whether speakers should have some (phonetic) experience or not: the task could be quite difficult for non-trained speakers because the synthesized stimulus sentences sound unnatural due to the lack of any temporal effects, and due to the poor quality of diphone speech in general.

The speakers were native speakers of standard Dutch, as judged by a two-member panel of Dutch phoneticians. They varied in age from 36 to 55 years old.

The experiment was run in individual sessions. The material was presented to the speakers both auditorily using a Sony SS-E34 loudspeaker and visually on paper. Each precursor question was followed by two answers, first with the lexical target word in it, and then with the reiterant target word *papapapa* in it. There were 24 questions, which were randomized and presented to three speakers first in the order 1-24; after a 10 minute break they were presented again in the reversed order (24-1). The other two speakers started with the sentences in the order 24-1 and continued with the order 1-24 after the break.

The speakers were seated in a sound insulated booth and could take their time to read the instructions carefully. They were instructed to repeat the answers only, with the same melody as they had just heard. Both the questions (as recorded by the author) and the two following answers (made by diphone synthesis) were played back on-line by a computer, and the speech produced by the speakers was recorded on-line (44.1 kHz, 16 bits). The computer was placed outside the recording booth, so that the speakers were in the booth only with the loudspeaker, the sheets of paper with the instructions and the written stimuli, and a Sennheiser MKH-416 directional condenser microphone.

After the first answer to a given question was heard, i.e. the version with the lexical word *rododendron* in it, a warning tone indicated the beginning of the recording time in which this answer was to be repeated. After five seconds, a second tone (different from the first) indicated the end of this recording time. One second after this end-tone, the second answer (with the reiterant target word in it) was presented, also followed by a recording time of five seconds signalled by a beginning and end tone. After another second the next question was heard, followed by two answers etc. This steady timing pattern enhances a constant speech rate.

Before the experiment began, four practice examples were given. This was done so that the speakers could get used to the type of speech and contrasts involved in the experiment and to the speed at which the material was presented. During the experiment, if either the experimenter or the speaker was dissatisfied with the speaker's performance, the experiment could be interrupted at any time by the experimenter (the author) who monitored the recording from outside the booth. The experiment would then be continued starting from the preceding precursor question.

2.2.2.4. Results

Every stimulus sentence was repeated twice by every speaker. Only the first recordings were segmented, while the second recording served as back-up in case a mistake or other problem had been overheard during the experiment. In these cases (2 out of 240), the second recording replaced the first. The results are thus based on 24 sentences x 5 speakers = 120 measurements, both for the lexical and the reiterant version of the target word. The target words were segmented into phonemes according to the guidelines given for Dutch by van Zanten et al. (1991).

An analysis of variance (ANOVA) was performed on the data, with boundary type, focus distribution and intonation contour as fixed factors, with repeated measures over speakers and total word duration as the dependent variable. Separate ANOVA's were run for the lexical version of the target word (*rododendron*) and for the reiterant (*papapapa*) version. The results for the lexical target word will be discussed first.

Neither intonation contour nor focus distribution had any significant effect on the results ($F[2,8]=1.49$, n.s., and $F[1,4]=2.08$, n.s., respectively). Only the type of boundary at which the target word occurs has a significant effect on the duration of the target word ($F[3,12]=30.59$; $p<.001$). On the basis of these findings, a series of oneway ANOVA's were run with segment duration as the dependent variable, and only boundary type as a fixed factor, to find out where the interaction lies between duration and boundary type, i.e., which segments of the word are lengthened in relation to the boundary at which the word occurs. The results of this analysis form the preliminary answer to our research question in (1a) concerning the influence of boundary depth on the domain of final lengthening.

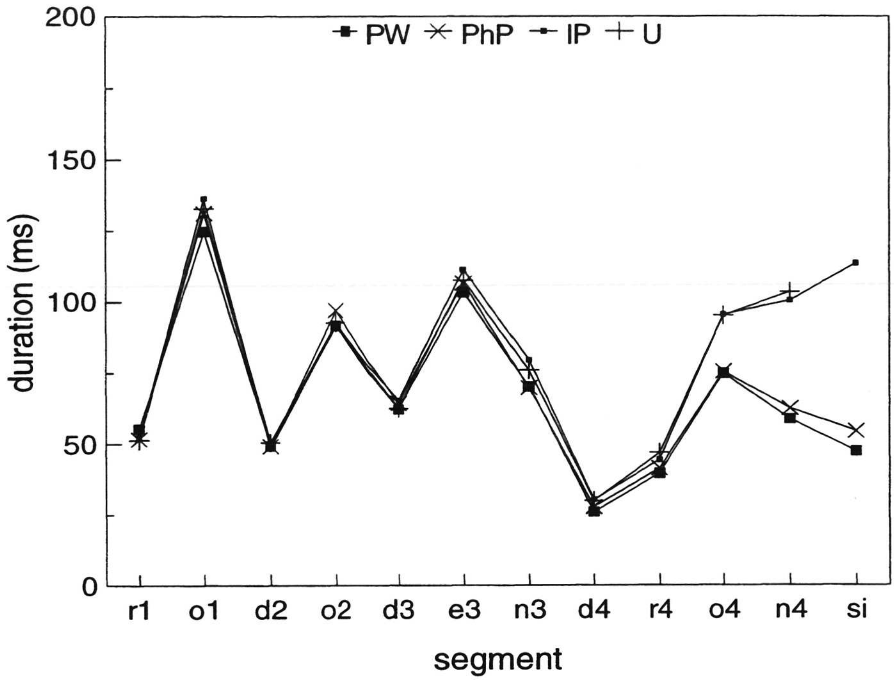


Figure 2.1. Segment durations for each boundary depth, for the lexical target word 'rododendron'.

In Figure 2.1, the duration of each segment in the word *rododendron* (denoted on the x-axis in conventional orthography) for each of the four prosodic boundaries is given (means and standard deviations are given in Appendix B). On the x-axis, the number next to the segment indicates the syllable in which this segment occurs; for instance, o2 stands for the /oː/ in the second syllable. The silent interval between the target word and the following word is indicated by 'si'; this interval is undefined for the U-boundary, since no word follows the target word in that condition.

It is clear from Figure 2.1 that there is no effect of final lengthening until the final syllable of the word is reached; a difference between the boundary types is found only in the final syllable. In o4 and n4, i.e. in the rhyme of the final syllable, there is a significant effect of boundary type on segment duration ($F[3,12]=16.20$, $p<.001$ for o4; $F[3,12]=33.60$, $p<.001$ for n4). Not all the boundaries in the material are signalled by different degrees of lengthening. Only the IP boundary is clearly and consistently lengthened with respect to the next-shallower PhP boundary. In o4 as well as in n4, two homogeneous subsets are formed (Newman-Keuls procedure with α=.05): PW and PhP do not differ from each other, nor does IP differ from U. The silent interval after the IP-boundary is also significantly longer than after the shallower boundaries (F[2,8]=10.76, p=.005, followed by Newman-Keuls post-hoc comparison). The effect of boundary type in the onset of the last syllable is not significant (F[3,12]=1.80, p=.20 for d4; F[3,12]=2.34, p=.13 for r4). In Figure 2.1 we can see that the gap between the boundaries of different depths gets larger as the target phoneme occurs closer to the word edge. This indicates that final lengthening is indeed progressive, which agrees with the results of Berkovits (1994) and Hofhuis et al. (1995). In the final segment, there is also a small tendency for the PhP-boundary to be marked more by final lengthening than the PW-boundary, and for the U-boundary to be marked more than the IP-boundary, but this tendency does not reach significance. Still, the results agree with the observation by others in that there is more final lengthening at deeper boundaries.

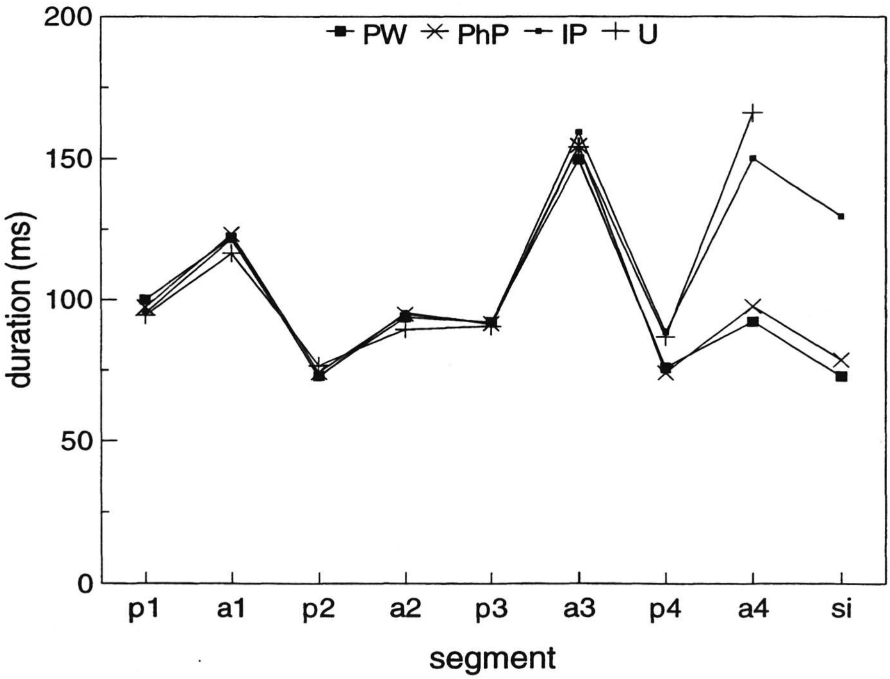


Figure 2.2. Segment durations for each boundary depth, for the reiterant target word 'papapapa'.

Similar results are obtained for the reiterant target word *papapapa*. An ANOVA with boundary type, focus distribution and intonation contour as fixed factors, with repeated measures over speakers and total word duration as the dependent variable, shows that again only the type of boundary at which the target word occurs has a significant effect on the duration of the target word ($F[3,12]=10.44$, $p=.001$). Figure 2.2 shows the same information as Figure 2.1, but is based on the reiterant version of the target word (means and standard deviations are given in Appendix B). Again, the IP-boundary has a clear lengthening effect on the preboundary segments. Only the segments in the final syllable and the silent interval following the word are affected ($F[3,12]=11.95$, $p=.001$ for p4; $F[3,12]=21.88$, $p<.001$ for a4; $F[2,8]=12.61$, $p=.003$ for the silent interval). In p4, two homogeneous subsets are formed (Newman-Keuls procedure with $\alpha=.05$). In a4, *three* subsets are formed: not only is there a significant amount of lengthening at the IP-boundary, distinguishing it from the PW- and PhP-boundary, but there is also significantly more lengthening at the U-boundary than at the IP-boundary. However, this effect is found only for some speakers, as Figure 2.3 shows.

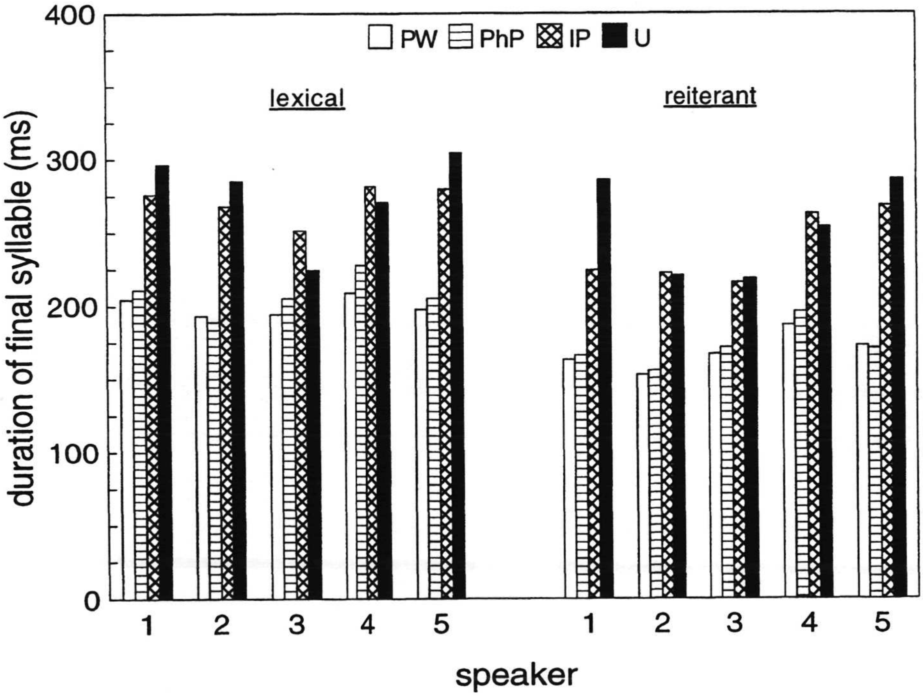


Figure 2.3. Final syllable duration per speaker and boundary depth, for the lexical target word (left) and the reiterant target word (right).

In Figure 2.3, the durations of the final syllable in the lexical and the reiterant reproductions are given, broken down by boundary depth and speaker. Speakers 1, 2 and 3 may be called phonetically trained. On the right, the results for the final syllable of the reiterant target word are given; here we see that especially Speaker 1 distinguishes the utterance boundary from the IP-boundary as much as the IP-boundary is distinguished from the lower boundaries. This is a phonetically trained female speaker. On the left, where the final syllable durations for the lexical target word are given, we see again that this speaker distinguishes U from IP, but here the effect is smaller and cancelled out by the durations of the other speakers.

In general, the speakers differ in the extent to which they mark different prosodic boundaries by different degrees of final lengthening, but they all lengthen more at an IP-boundary than at a PhP- or PW-boundary. The phonetically naive speakers (4 and 5) do not differ in any systematic way from the phonetically trained speakers.

2.2.2.5. Discussion

As for the general questions addressing the appropriateness of the present approach for the type of experiments discussed here, one may conclude that the speakers were quite capable of doing the task. Generally, no difference could be detected between phonetically trained and phonetically naive speakers. Regardless of which intonation contour accompanied the sentences, the speakers had no problem interpreting the structure of the sentences and adding the temporal organisation appropriate for such a structure. This is exactly what is required for the set-up of this experiment.

A four-syllable word with penultimate stress was used in the pilot study so that a distinction could be made between hypothetically possible domains of lengthening such as the final syllable (*dron*), the part starting with the primary stressed syllable (*dendron*), and the whole word (*rododendron*). In the present study, the effect of final lengthening was restricted to the final syllable, and for the lexical target word it was only significant in the final rhyme. This implies that neither the whole word nor the last stressed syllable are affected by final lengthening. As noted in §1.3.2.4 (p. 15), the foot structure of words such as *rododendron* (with a final VC syllable) is somewhat problematic. In the main experiment, words will be included that undisputedly have a final disyllabic foot, so that we can differentiate between the final syllable and the final foot as the domain of final lengthening. The main experiment will also contain words which are longer than two syllables, since at this point we do not wish to exclude the possibility of final lengthening effects exceeding the final foot.

Only one of the prosodic boundaries used in the experiment is clearly marked by more final lengthening than the next lower boundary. As soon as final lengthening starts, there is a significant difference between an intonational phrase boundary and a phonological phrase boundary. Other levels are not consistently marked; in fact, the PhP-boundary never differed significantly from the PW-boundary, while the distinction made between the intonational phrase and the utterance depends largely on the speaker. If the following experiment yields similar results, the conclusion may be drawn that only intonational phrase boundaries are obligatorily marked by final lengthening, and that the choice for marking finer distinctions in boundary depth is up to the speaker.

Focus did not have a significant effect on the absolute duration of the segments. This is somewhat surprising, since [+focus] leads to accentuation, and accented words in Dutch are normally lengthened by some 15% relative to unaccented words (Sluijter & van Heuven, 1995). However, our speakers did not pronounce the target word more quickly in the [-focus] condition, even though in this condition it expresses information already given in the precursor question. In hindsight, the [-focus] conditions are still quite likely to elicit some prominence on the target word due to rhythmical principles. Generally, when a certain minimal amount of material follows or precedes the nuclear accent in an utterance, another word will be made prominent, either by a non-nuclear accent and/or by other prosodic means (like duration). In the [-focus] sentences, *rododendron* is a very good candidate for a secondary phrasal prominence. Informal listening to the material confirms that both the stressed syllable of the target word and of the first word of the parenthetical were often realized with a certain amount of prominence (in italics; nuclear accent in capitals):

PIET wil die rare rodo*den*dron( )planten, *gek* als hij is.
PIET wil die rare rodo*den*dron, *plan*tengek als hij is.
*Plan*tengek als hij is wil PIET die rare rodo*den*dron.

The difference in focus distribution, as determined by the precursor question, in this case did not lead to the large difference in prominence we had anticipated, because due to rhythmical factors the [-focus] condition also bears some prominence. Indeed, [-focus] does not necessarily imply [-accent], and even if the word is unaccented (i.e., not marked by an accent-lending pitch movement), prominence laid on by rhythmical principles may well be realized by means of other prosodic features, such as lengthening.

The lack of the expected effect of focus distribution on the duration of the word implies that this experiment cannot answer the question whether final lengthening interacts with accentual lengthening, since no accentual lengthening was found. The issue of the interaction between final lengthening and accentual lengthening will be taken up again in Chapter 4. Focus will not be included as a variable in the main experiment described in the next section, since it is the same type of experiment and material, so there is no reason to assume that this time an effect of accent will be found. Instead, the target words will only occur in [+focus] condition.

Despite some complaints from the speakers concerning the unnaturalness of the stimulus sentences, which can only be expected when diphone synthesis without temporal effects is used, the speakers were still able to grasp the structure of the utterances without any prosodic cues as to this structure in the input speech. Judging by the speakers' imitations, the punctuation in the written text together with the precursor questions gave enough information to elicit speech with the required prosodic structure from the speakers. Since it does not appear to make the repetition task any easier, it is not necessary to mark the major (IP-) boundaries in the material by a boundary-marking pitch movement (if anything, the inclusion of three different intonation contours made the task harder). As a result, the main experiment described in the next section will only contain utterances with just a pitch movement combination 1&A on the word in focus. This contour is standardly produced by the diphone synthesis program and, more importantly, was most accurately and easily repeated by the speakers.

2.2.3. Main production experiment

2.2.3.1. Introduction

The pilot study described in §2.2.2 has shown that a repetition task using diphone synthesis without any temporal effects elicits speech that contains the desired temporal effects. This means that the speakers did not copy the temporal structure of the input speech, but implemented their own temporal structure in their speech. This temporal structure can thus be said to reflect natural temporal effects, which speakers always display in their speech.

The results of the pilot study indicate that the domain of final lengthening in the word *rododendron* is restricted to the final syllable. In the following production experiment, a number of other target words will be included which differ in the structure of their final two syllables. These materials will allow us to answer research question (1b), repeated below:

(1b) Do stress and/or the structure of the final syllable/word have an effect on the domain of final lengthening?

2.2.3.2. Material

Five words are used in this experiment which have different types of final syllables, resulting in different prosodic structures in the final two syllables. These words are given in (6) (primary word stress is marked with ´):

(6)	márathon	harmónika	yúcca	tándem	móde
	/ˈmaːraːtɔn/	/hɑrˈmoːnikaː/	/ˈjukaː/	/ˈtɛndəm/	/ˈmoːdə/
	(id.)	'concertina'	(id.)	(id.)	'fashion'

These words are used to answer the research question in (1b) concerning the role of final syllable structure and/or stress on the domain of final lengthening. The word *marathon* is included to confirm the results obtained for the word *rododendron*, since both have a final closed (VC) syllable that does not bear primary stress. The word *tandem* also has a closed final syllable, yet this syllable contains a schwa and is therefore not a separate foot under any analysis (cf. §1.3.2.4). The words *yucca*, *mode* and *tandem* have a final unstressed syllable, and each form one disyllabic foot.

The word *harmónika* does not follow the dominant stress pattern for VV-final words (stress on the penult if this does not contain schwa; Kager, 1989), but has antepenultimate stress instead (as more VV-final words do with only /i/ in the rhyme of the penult; Kager, 1989; Booij, 1995). In a stress-based foot structure theory, this pattern can be derived by marking the final syllable as extrametrical (*harmóni<ka>*), just as in words with a final VC-syllable (*mára<thon>*). However, a final extrametrical VV-syllable differs from an extrametrical VC-syllable in that it is not heavy (cf. §1.3.2.4). Rhythmically, though, they both receive secondary stress (marked with `): *harmónikà, máratòn*. Crucially, the word *harmonika* makes a distinction possible between effects of secondary stress on the final syllable (*harmonika* patterning with *marathon*) and final syllable structure or weight (*harmonika* patterning with *yucca*).

The words also differ in final vowel quality (only *mode* and *tandem* do not have a final full vowel). Finally, the word-final rhymes can be ordered along a weight scale from heavy to light.

secondary vs. no stress (~disyllabic foot)	márathòn harmónikà | móde tándem yúcca
final syl. structure: VC vs. VV vs. schwa	marathon tandem | harmonika yucca | mode
final vowel quality: full vs. reduced	marathon harmonika yucca | tandem mode
final syl. weight: from heavy to light	marathon | harmonika yucca | tandem | mode

The boundaries included in the present experiment are the same as those in the pilot study (PW, PhP, IP, U). The sentences for each word differ minimally in order to form grammatical sentences with the different boundary types immediately following the target word. Each sentence consists of an intonational phrase and a parenthetical, as in the pilot study. In the U-boundary conditions the parenthetical precedes the intonational phrase that contains the target word. In the other cases the parenthetical follows the intonational phrase.

In the pilot study, the test sentences were all equally long. The choice for an equal number of syllables in each utterance together with the other requirements on the (differences between the) utterances implies that the number of syllables in the intonational phrase in which the target word occurs is smaller in the IP-boundary and U-boundary conditions than in the PW-boundary and PhP-boundary conditions (cf. (3)). The problem was already noted in §2.2.2.2. This difference between the PW- and PhP-boundary sentences on the one hand and the IP- and U-boundary sentences on the other could have been the cause of at least some of the difference in the amount of lengthening between these two groups; a long intonational phrase (PW and PhP) is spoken at a faster speaking rate than a shorter intonational phrase (IP and U), thus giving rise to shorter durations in the former case. Therefore the requirement of an equal number of syllables in each utterance should perhaps be given up in favour of an equal number of syllables in the intonational phrase in which the target word occurs. To make sure that the length of the intonational phrase does not jeopardize our conclusions, two versions of the IP- and U-boundary sentences were included in the experiment, one constructed similarly to the way it was done in the pilot study (i.e. with the same number of syllables in the *utterance* as in the PW- and PhP-boundaries) and one in which the utterance is lengthened to obtain the same number of syllables in the *intonational phrase* as in the PW- and PhP-boundary conditions. An example is given in (7), where the sentences adapted to the length of the intonational phrase in which the target word occurs are named IP' and U' (see Appendix C for the complete set of sentences):

(7)

IP	[Zij zijn op weg met een snelle <u>tandem</u>]$_{IP}$, rijdend door de mooie duinen.
IP'	[Zij zijn vandaag op weg met een snelle <u>tandem</u>]$_{IP}$, rijdend door de mooie duinen.
	'They are (today) away with a fast tandem, riding through the beautiful dunes.'
U	Rijdend door de mooie duinen [zijn zij op weg met een snelle <u>tandem</u>]$_{IP}$.
U'	Rijdend door de mooie duinen [zijn zij vandaag op weg met een snelle <u>tandem</u>]$_{IP}$.
	'Riding through the beautiful dunes are they (today) away with a fast tandem.'
cf. PW	[Zij gaat op weg met een snelle <u>tandem</u>rijder]$_{IP}$, door de mooie duinen.
	'She goes away with a fast tandem rider, through the beautiful dunes.'

All sentences were synthesized using the same diphone set as in the pilot study (see §2.2.2.2). Each utterance was realized with a 'pointed hat' (1&A) on the primary stressed syllable of the target word, without any melodic boundary markers. The material for the present experiment thus consists of 30 utterances (5 target words x 6 sentences).

2.2.3.3. Subjects and procedure

Three male and three female speakers participated in the experiment. Four of these are phonetically trained. Two of the speakers had also taken part in the pilot study. All of the speakers were native speakers of standard Dutch, as judged by two native Dutch phoneticians. The speakers were between 34 and 55 years of age.

All utterances were preceded by a precursor question, which put focus on the target word. These 30 question-and-answer pairs were randomized and presented to half of the speakers first in the order 1-30 and then in the order 30-1, while the other half began with the order 30-1 followed by the order 1-30. Speakers could take a short break halfway through the experiment.

The recording procedures and instructions for the present experiment were exactly the same as those for the pilot study, except that the auditory stimuli were given over headphones instead of using a loudspeaker.

2.2.3.4. Results

Both recordings of each utterance by every speaker were segmented according to the guidelines given in van Zanten et al. (1991). To begin with, the two versions (non-adapted and adapted) of the IP-boundary condition and the U-boundary condition were compared. The IP and U conditions do not differ from the IP' and U' conditions respectively ($F[1,118]<1$ for both IP and U). As expected, the means for the adapted versions are a little shorter than for the non-adapted versions, i.e. they come closer to the PW- and PhP-durations (IP: 563 ms vs. 569 ms; U: 559 ms vs. 565 ms). In order not to bias the results towards a large distinction between PW and PhP on the one hand and IP and U on the other, only the adapted versions will be used in the analyses. The results are thus based on 5 words x 4 boundaries x 6 speakers x 2 repetitions = 240 measurements.

An analysis of variance (ANOVA) with total word duration as the dependent variable, boundary type as fixed factor and random factors of Speakers and Words show that boundary type has a significant effect on word duration (by Words: $F[3,12] = 74.68$; $p<.001$; by Speakers: $F[3,15] = 43.15$, $p<.001$). In order to find out which segments are affected, analyses of variance were run as above, with the following dependent variables:

final coda	–	*marathon*	–	*tandem*	–
final nucleus	*harmonika*	*marathon*	*mode*	*tandem*	*yucca*
final onset	*harmonika*	*marathon*	*mode*	*tandem*	*yucca*
penult. rhyme	*harmonika*	*marathon*	*mode*	*tandem*	*yucca*

Taking the five target words together, the effect of boundary type is significant in the whole of the final syllable, i.e., in the final coda (by Words: $F[3,3] = 32.75$, $p=.009$; by Speakers: $F[3,15] = 20.13$, $p<.001$), the final nucleus (by Words: $F[3,12] = 12.07$, $p=.001$; by Speakers: $F[3,15] = 25.92$, $p<.001$) and the final onset (by Words: $F[3,12] = 8.50$, $p=.003$; by Speakers: $F[3,15] = 18.81$, $p<.001$). In each case, post-hoc comparisons (Newman-Keuls, $\alpha=.05$) indicate that two homogeneous subsets are formed: {PW, PhP} and {IP, U}. Interestingly, the effect in the penultimate rhyme is significant only by Speakers, but not by Words (by Words: $F[3,12] = 2.69$, $p=.09$; by Speakers: $F[3,15] = 10.70$, $p=.001$). The fact that the effect of boundary depth on the duration of the penultimate rhyme is not consistent across words could mean that final lengthening reaches back into the penultimate rhyme only in some words, but not in others, i.e., that some words have a larger lengthened domain than others.

In order to investigate how far final lengthening reaches back into each of the target words, ANOVA's were run with segment duration as the dependent variable, boundary type as fixed factor and repeated measures over speakers, separately for each target word. The results of these analyses are given in Table 2.1. The segment durations for each of the five words are depicted in Figure 2.4. The segments of each word (denoted in conventional Dutch orthography) are used as labels on the x-axis (means and standard deviations are given in Appendix D).

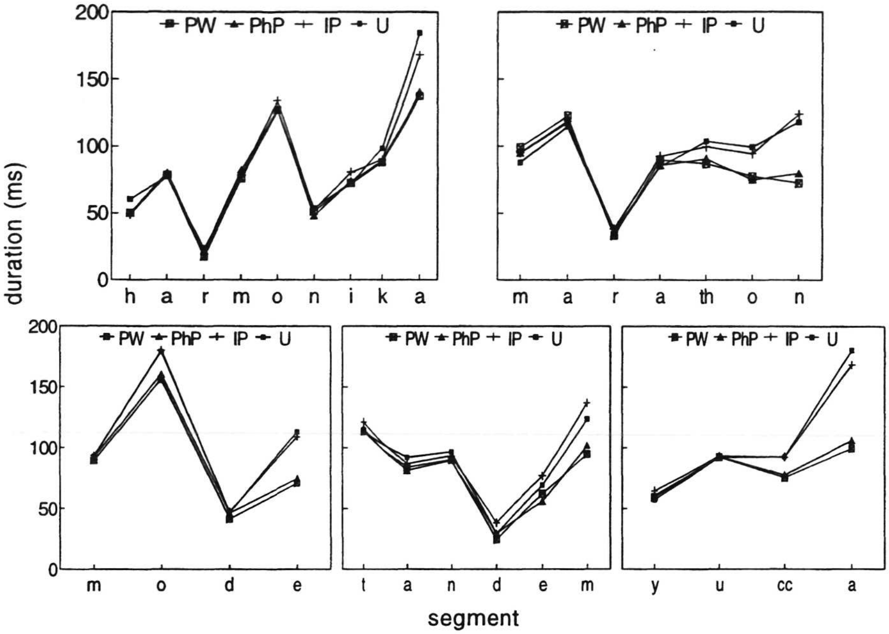


Figure 2.4. Segment durations for each boundary type; one diagram for each word.

Table 2.1. Effect of boundary type on segment durations; F-ratios and significance (bold face indicates a significant effect at the .01 level).

	h	a	r	m	o	n	i	k	a
F(3,12)	1.29	<1	<1	2.35	1.46	1.95	3.91	3.33	7.83
p	n.s.			n.s.	n.s.	n.s.	.03	.05	**.002**
	m	a	r	a	th	o	n		
F(3,12)	3.58	<1	<1	<1	10.00	6.00	16.17		
p	.04				**.001**	**.007**	**<.001**		
	t	a	n	d	e	m			
F(3,12)	2.23	6.91	<1	10.24	6.66	13.88			
p	n.s.	**.004**		**.001**	**.004**	**<.001**			
	m	o	d	e					
F(3,12)	<1	15.32	<1	12.10					
p		**<.001**		**<.001**					
	y	u	cc	a					
F(3,12)	1.04	<1	10.52	65.19					
p	n.s.		**.001**	**<.001**					

As can be expected, there is a significant effect of boundary type on the duration of at least the final rhyme of each word. Post-hoc analyses (Newman-Keuls, α=.05) show that the PW and PhP boundaries are never significantly different, but form one homogeneous subset for all words. The IP and U boundaries form a second homogeneous subset, except in the final rhyme of *harmonika* and *yucca*, where three subsets are formed: {PW, PhP}, {IP} and {U}.

As the distance to the right word edge increases, the effect of boundary depth becomes increasingly opaque. The two-way distinction in boundary depth consistently made in the final rhyme is also found in the final onsets of *marathon* and *yucca* (as can be seen in Figure 2.4, and by the results of post-hoc analyses), while in *tandem* only the IP boundary differs from the rest. In *harmonika*, the effect in the penultimate rhyme and the final onset is significant only at a .05 level, and is not very consistent: in *i*, there is only lengthening for the IP boundary, while in *k*, there is only lengthening for the U boundary.[4]

As for the penultimate rhyme, we see a very clear effect of boundary depth only in the word *mode*, where we find the same homogeneous subsets as in final rhymes. In the word *tandem* too, we find a significant effect in the expected direction, yet here the subsets given by a Newman-Keuls post-hoc procedure are {PhP, PW, IP} and {IP, U}, i.e., only the U boundary differs significantly from the PW and PhP boundaries.

The data for the two final syllables of the target words are presented in a different way in Figure 2.5; the lengthening of the penultimate and of the final syllables is shown for each target word, with the duration at the shallowest boundary (PW) taken as 'zero duration'. Although there is no statistical difference between PW and PhP, the final syllables consistently show a lengthening for PhP of some 8 ms (≈4%). On the left of the figure, where the lengthening of the penultimate syllable is shown, it can be seen that this syllable is lengthened by some 30 ms (≈11%) in *mode*, and that *tandem* lies somewhere in between *mode* and the other words.

[4] Note that in the initial segment of *marathon*, there is also a significant effect of boundary type at the .05 level. However, this can have nothing to do with final lengthening, since a) it is too far from the right word edge, b) the effect is not in the expected direction; the segment is **shorter** for U than for the other boundaries. We therefore hesitate to accept effects that are only significant at a .05 level, and concentrate on those that are significant at a .01 level.

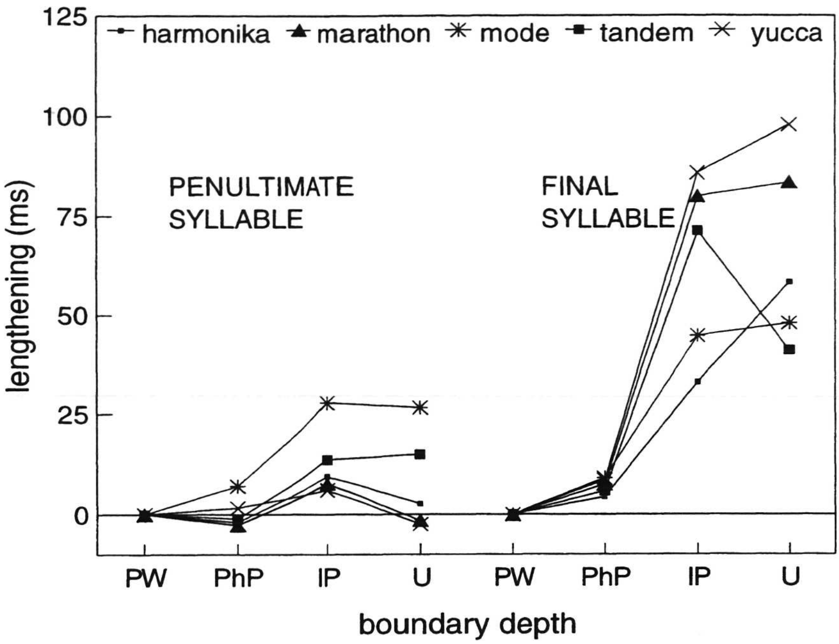


Figure 2.5. Lengthening (re. PW-boundary) of the final two syllables for the five target words.

In the reiterant data obtained in the pilot study, some speakers were found to distinguish U from IP, whereas others did not. In Figure 2.6, the durations of the final syllables from the present data per boundary depth are given for each speaker, for all words taken together. Speakers 5 and 6 are phonetically naive. As was found in the pilot study, some speakers (particularly Speaker 4) lengthen more at a U boundary than at an IP boundary, while others (particularly Speaker 5) seem to do the opposite. The variation between IP and U is larger than that between PW and PhP. The large lengthening effect at an IP boundary is found for every speaker.

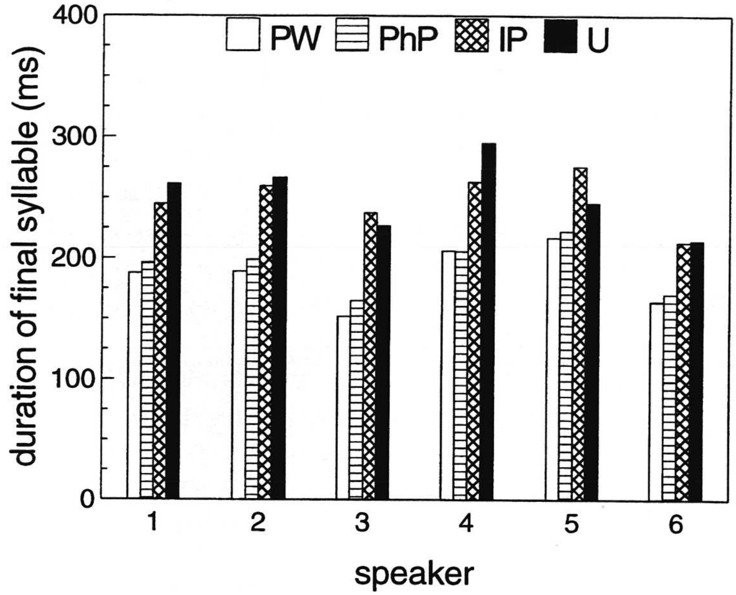


Figure 2.6. Final syllable durations, broken down by speaker and boundary depth, for all words taken together.

2.2.3.5. Discussion

The results of the main production experiment confirm the results of the pilot study. Both experiments converge in showing that, in Dutch:

- the amount of lengthening is largest in the final segment and decreases in the preceding segments,
- a phonological phrase boundary is not marked by significantly more lengthening than a prosodic word boundary,
- the end of an intonational phrase and of an utterance are clearly lengthened,
- a deeper boundary does not lead to a larger domain of lengthening, and
- neither the whole word nor the last stressed syllable are necessarily affected by final lengthening.

The results of the main experiment also give additional information. Words ending in a long vowel do not behave any differently than words ending in a closed syllable, in the sense that in both cases final lengthening is restricted to the final syllable (consistently reaching significance in the final rhyme, and sometimes affecting the onset as well). This implies that the domain of final lengthening is not the final foot, since a final VV-word like *yucca* forms one disyllabic foot yet is only partly lengthened. However, words with a final schwa have a larger domain of lengthening than words with a full vowel in the final syllable (be it open or closed); the effect in *mode* reaches back into the penultimate syllable. The word *tandem*, having a final schwa + coda consonant, seems to lie in between *mode* and the other words. This leads to the conclusion that when the rhyme of the final syllable contains only a schwa (i.e., with final ultra-light syllables), final lengthening begins in the penultimate syllable. This effect can also be found, to a lesser extent, in words ending with a schwa plus a coda consonant, i.e., in the second lightest type of rhyme.

One could argue that these conclusions are based on data with very few words. Still, they are strengthened by other data in the literature. In general, the bulk of the final lengthening effect is found in the final rhyme, agreeing with the results of Wightman et al. (1992). The finding that the penultimate syllable is lengthened when the final rhyme contains only a schwa is supported by Hofhuis' data (in prep.), which involved eight such schwa-final words. The finding that *only* schwa-final words have a domain of final lengthening larger than the final syllable explains why this larger domain is usually not found in more general final lengthening studies, which do not differentiate between different types of final syllables; the relatively small effect in the penultimate syllable (remember that final lengthening is progressively distributed) in only a small subset of the total number of words will be obscured by other words.

The effect of syllable weight on the domain of final lengthening may be explained by the fact that in general, light syllables contain less segments, and are phonetically shorter, than heavy syllables. Shorter and fewer segments result in a smaller expandability of the syllable. A schwa in particular, which is by definition a reduced syllable (spectrally **and** temporally), may not be able to be lengthened to the degree that is required by the boundary depth. Thus, in such cases, segments preceding a final schwa will have to participate in the final lengthening, resulting in a larger domain which is lengthened. The domain of final lengthening is then determined by purely phonetic considerations (such as inherent duration of segments and their expandability; cf. Klatt, 1976; Allen et al., 1987). This may also clarify the non-systematic behaviour of the final onset; the variation in the absolute duration of the final rhyme (due to factors such as segment identity, stress, speaking rate, etc.), together with (inherent) differences in expandability, may result in a non-systematic effect in the final onset.

2.3. Perception experiments

2.3.1. Identification experiment

2.3.1.1. Introduction

The previous section was concerned with the domain of final lengthening in speech production. Following these production experiments, two perception experiments were run: an identification experiment (this section) and an acceptability experiment (§2.3.2). The experiment described in this section was devised to find out if listeners can establish which boundary is intended by the speaker, when the sentence is structurally ambiguous. Thus, the present experiment should answer research question (2a), repeated below:

(2a) To what extent are the durational differences found in production used in boundary perception?

The segments of the target words in the production experiments were significantly longer only at the Intonational Phrase (IP) boundary and the Utterance (U) boundary. The Prosodic Word (PW) and Phonological Phrase (PhP) boundaries were not significantly differentiated, nor was the U-boundary consistently differentiated from the IP-boundary. However, other prosodic features may have been used to signal the difference between these boundaries. From the literature it is clear that durational and melodic cues may be used together as boundary markers (Sanderman, 1996, and references given therein). The present identification experiment was run with material obtained in the main production experiment (§2.2.3) to see if listeners can determine the type of boundary intended by the speaker. Both the role of the durational structure and of intonational cues on the recognition of boundary depth will be investigated.

2.3.1.2. Material

Only one set of sentences from the main production experiment has exactly the same sequence of words for the PW, PhP and IP sentences.[5] These are given in (8):

[5] The word order in the U-sentences always differed radically from those in the other sentences, since the IP containing the target word, which precedes the parenthetical in the other sentences, necessarily follows this parenthetical in the U-sentences.

(8) PW Piet wil die rare yuccaplanten, gek als hij is.
'Piet wants those strange yucca plants, crazy as he is.'
PhP Piet wil die rare yucca planten, gek als hij is.
'Piet wants to plant that strange yucca, crazy as he is.'
IP Piet wil die rare yucca, plantengek als hij is.
'Piet wants that strange yucca, plant crazy as he is.'

All sentences had been repeated twice by the speakers in the production experiment. Both recordings of the sentences in (8) of every speaker were used in the present study.

To examine the role of intonation in the identification of the boundary types, two versions of each utterance were used; one with the original intonation as realized by the speakers, and one in which all melodic markers are removed, i.e. in which the intonation consists only of a declination line (made with PSOLA[6]). This declination line approximates the lower declination line in the original utterance.

All versions of the utterances were presented twice to the listeners, giving 3 intended boundaries x 6 speakers x 2 recordings x 2 intonation contours x 2 repetitions = 144 utterances.

2.3.1.3. Subjects and procedure

Sixteen native Dutch listeners, aged 18 to 53, participated in the experiment. They had no known hearing impairments.

The utterances were recorded on tape and presented to the listeners in group sessions over high-quality headphones in quasi-random order. An utterance by one speaker was never directly followed by an utterance of that same speaker. The listeners received instructions which included an explanation of the different structures of the sentences in (8). For each utterance, the three sentences were given on paper, and the listeners had to indicate which of these three they thought was the one meant by the speaker (a forced choice task).

2.3.1.4. Results

The results are based on a total of 144 utterances x 16 listeners = 2304 judgments. As expected, listeners recognized the IP boundary very well: the number of PW and PhP boundaries mistaken for an IP boundary and vice versa are very small. This can be

[6] PSOLA (Pitch-Synchronous OverLap-Add) is a resynthesis technique which modulates pitch through a time domain manipulation; it does not affect spectral quality (Moulines & Verhelst, 1995).

seen in Table 2.2, where the responses for each intended boundary type are given (in total number of responses and in percentages). The results for the utterances with intonation and for those with only declination are given separately. Grey shading indicates correct results. The percentages in each row add up to 100%.

Table 2.2. Given responses per intended boundary and intonation condition.

output ▸ input ▾		responses PW	responses PhP	responses IP
with intonation	PW	179 (46.6%)	202 (52.6%)	3 (0.8%)
	PhP	132 (34.4%)	243 (63.3%)	9 (2.3%)
	IP	4 (1.0%	4 (1.0%	376 (97.9%)
declination only	PW	209 (54.4%)	164 (42.7%)	11 (2.9%)
	PhP	167 (43.5%)	200 (52.1%)	17 (4.4%)
	IP	5 (1.3%)	4 (1.0%)	375 (97.7%)

When the speaker produces an IP boundary, it is recognized as such in nearly all cases (98%). The results for the other two boundary types are more evenly distributed. Only the PhP boundary with intonation seems to do fairly well (63%); however, this is partly due to a bias toward PhP responses when intonation is present (this interaction between responses and intonation condition will be discussed below). Still, this bias is stronger when it is indeed a PhP boundary, indicating a weak ability on the part of the listeners to hear which boundary is intended. Even when the intended IP boundaries and IP responses are disregarded, a significant effect of intended boundary on the responses remains in the expected direction, both with intonation (χ^2=10.8, df=1, p=.001) and in the absence of intonation (χ^2 =8.2, df=1, p= .004): an intended PW boundary indeed leads to more PW responses than PhP responses, and an intended PhP boundary favours PhP responses.

The absence of intonation does not lead to a large loss of information: the intonation contour has no main effect on the percentage of correct responses (F[1,70]<1). Within the PW and PhP data, however, there is a significant **interaction** between intonation contour and given responses (χ^2=14.1, df=1, p<.001), in that PhP responses prevail when intonation is present and more PW responses are given when intonation is absent. This interaction is presumably the result of a strategy followed by the listeners, rather than having any theoretical relevance. This strategy is likely to be the result of the relatively poor ability to hear the difference between a PW and a PhP boundary, in combination with the presence of one binary variable in the material which is clearly audible, namely the presence or absence of the original intonation contour. I assume that during the experiment, some listeners began to link this audible binary distinction to the binary choice which they had to make when the utterance was clearly not the IP sentence. There are several observations which support this assumption. First of all, some listeners explicitly declared that they had adopted such a strategy, since they had little else to go by. Secondly, the correlation is stronger in the second half of the experiment, indicating that the strategy is developed in the course of the experiment. Finally, and most persuasively, the listeners differ in the way they connected the melodic variable to their response; some of them linked the presence of intonation to a PhP response and the absence of intonation to a PW response, but others have done the exact opposite. In all, this indicates that the interaction has no theoretical implications.

Table 2.3. Given responses according to a strategy linking PW and PhP responses to the presence or absence of intonation, per listener and per repetition. Grey shading indicates the direction of the bias for two subsets of listeners.

strategy ▸	with intonation=PW, only declination=PhP			with intonation=PhP, only declination=PW		
listeners ▾	1st time	2nd time	total	1st time	2nd time	total
4	9 (20%)	6 (13%)	15 (16%)	37 (80%)	42 (87%)	79 (84%)
7	17 (35%)	15 (36%)	32 (36%)	31 (65%)	27 (64%)	58 (64%)
9	8 (17%)	0 (0%)	8 (8%)	40 (83%)	47 (100%)	87 (92%)
11	10 (21%)	8 (17%)	18 (19%)	37 (79%)	38 (83%)	75 (81%)
total	44 (23%)	29 (16%)	73 (20%)	145 (77%)	154 (84%)	299 (80%)
2	32 (68%)	32 (70%)	64 (69%)	15 (32%)	14 (30%)	29 (31%)
8	25 (52%)	33 (69%)	58 (60%)	23 (48%)	15 (31%)	38 (40%)
12	23 (51%)	33 (69%)	56 (60%)	22 (49%)	15 (31%)	37 (40%)
14	26 (55%)	33 (69%)	59 (62%)	21 (45%)	15 (31%)	36 (38%)
total	106 (57%)	131 (69%)	237 (63%)	81 (43%)	59 (31%)	140 (37%)

Data supporting the latter two arguments are given in Table 2.3. In this table, the number of responses according to the strategy linking the presence of intonation to PW boundaries and the absence of intonation to PhP boundaries or vice versa is given per repetition for the eight listeners responding in accordance with such a bias in more than 60% of the cases (excluding IP responses). Listener 4, for example, has given a total of (15+79=) 94 non-IP responses. The number of PhP responses in reaction to an input utterance with intonation plus the number of PW responses in reaction to an input utterance with only declination equals 79 (=84% of 94). As the table shows, Listeners 4, 7, 9 and 11 have linked the presence of intonation to a PhP boundary and the absence of intonation to a PW boundary (grey shading). Since their bias is stronger than that of the listeners doing the opposite (2, 8, 12 and 14; grey shading), the overall results show an interaction in the same direction.

The lack of an effect of intonation contour on the percentage of correct responses shows that the material did not contain any melodic markers that could help the listener in determining what kind of boundary the speakers had intended. Still, listeners recognized even PW and PhP boundaries above chance. In order to investigate whether the small durational differences in the input could have provided some cue as to the intended boundary, we calculated the correlation between the percentage of PW responses and the duration of the final syllable of the target word, since this is the syllable which is affected by final lengthening, and its duration may therefore be a cue to boundary depth. Shorter final syllable durations are expected to elicit more PW responses, and longer durations should elicit fewer PW responses, so we expect a negative correlation.

In Figure 2.7, the percentage of PW responses (collapsed over intonation contours) are plotted against the final syllable durations. The figure does not explicitly differentiate between the three possible intended boundaries, but at least the intended IP boundaries can be inferred; all data points with a final syllable duration of more than 220 ms (and eliciting close to 0% PW responses) are intended IP boundaries.

When all cases are included, a correlation of $r=-.942$ ($p<.001$) is found. When only the intended PW and PhP boundaries are included, the correlation is still highly significant ($r=-.873$, $p<.001$). This implies that the duration of the final syllable of the target word, irrespective of the type of boundary actually intended by the speaker, greatly influences the listeners' responses.

Figure 2.7 also shows which data points are related to which speaker. Since some speakers have long final syllable durations and others have shorter durations, and given the strong relationship between final syllable duration and responses of the listeners, the responses also depend on the speaker. In Table 2.4, the percentage of correct responses per intended boundary are given for each speaker.

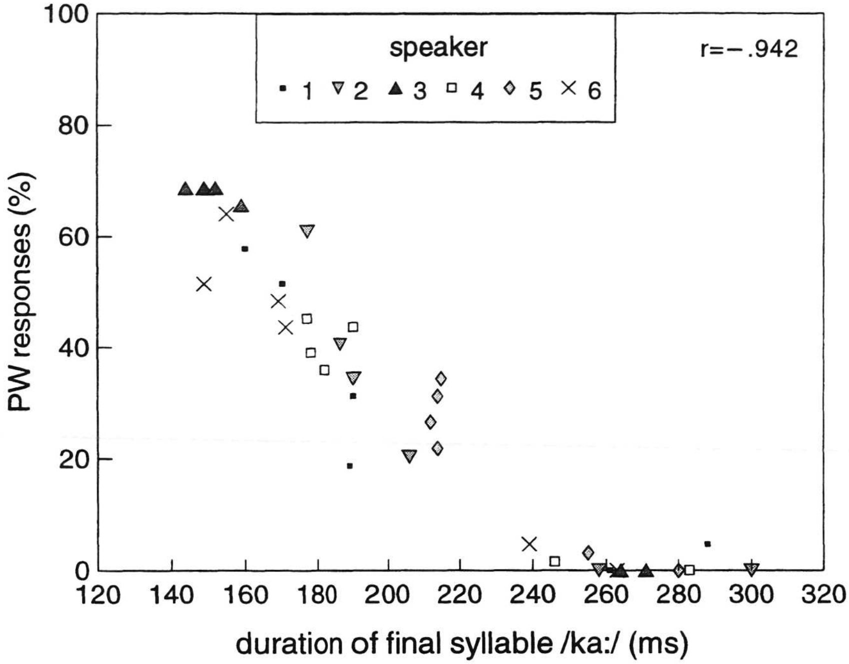


Figure 2.7. Percentage of PW responses plotted against the duration of the final syllable of the target word yucca*, broken down by speakers.*

Since the errors involving IP boundaries are minimal, we can assume that a PW or PhP boundary which is not correctly recognized will generally have been mistaken for a PhP or PW boundary, respectively. Speaker 3 has a high percentage of correctly recognized PW boundaries, but a low percentage of correctly recognized PhP boundaries, indicating a clear bias toward PW for this speaker. As we can see in Figure 2.7, this speaker has the shortest final syllables durations (within the set of PW/PhP boundaries). Similarly, Speaker 5 has the longest durations within this set, and is indeed most often judged to produce a PhP boundary (low % correct for PW, high % correct for PhP). Speakers 1 and 2 have the largest differences in their final syllable durations, and their utterances also yield the best results in the identification task.

Table 2.4. Percentage of correct responses per speaker and boundary type.

% correct	Speaker 1	Speaker 2	Speaker 3	Speaker 4	Speaker 5	Speaker 6
PW	54.7	50.8	68.8	42.2	32.8	53.9
PhP	75.0	68.8	31.3	57.8	66.4	46.9
IP	96.9	100	98.4	96.9	97.7	96.9
total	75.5	73.2	66.2	65.6	65.6	65.9

2.3.1.5. Discussion

The extent to which listeners can establish what type of boundary was intended by the speaker seems to be strongly related to the extent to which these boundaries are durationally marked. So, the IP boundary is easily distinguished from the two shallower boundaries, whereas the distinction between the PW and PhP boundaries is much harder to perceive. Although PW and PhP were internally differentiated above chance, the fact that several of the listeners felt they had to fall back on the strategy described above, linking their PhP/PW responses to the only clearly audible binary variable of intonation, shows that they were very uncertain in choosing between the two.

The loss of information contained in the intonation contour does not lead to systematically poorer results. In the case of IP boundaries, final lengthening is apparently a sufficient cue for boundary perception, since even without intonation the IP boundaries were correctly recognized in 98% of the cases. With such a high score even without intonation, the presence or absence of melodic markers in the intonation contour can hardly make a difference. In the case of PhP boundaries, the data indicate that there were no melodic markers to distinguish this boundary from a PW boundary; if there were, the presence of intonation would have had a positive effect on the results. Recall, however, that the speech material was obtained through a repetition task: the speakers were asked to repeat utterances which contained only a pointed hat on the accented syllable, and no melodic boundary markers. When speakers have more freedom with respect to their intonation, they may well make more use of melodic boundary markers.

Even though the PhP boundary was not marked by significantly more final lengthening than the PW boundary in the production, still listeners seem to have based their decisions on the duration of the final syllable of the target word. Indeed, both PW and PhP boundaries were correctly recognized above chance, although the difference between the two is quite hard to tell, judging by the listeners' uncertainty. In all, the data indicate that the durational differences found in production are quite efficiently used in boundary perception at all levels investigated.

2.3.2. Acceptability experiment

2.3.2.1. Introduction

Perception experiments have shown that the correlation between boundary depth and the *amount* of final lengthening is perceptually relevant (Nooteboom & Doodeman, 1980; Gussenhoven & Rietveld, 1992), but as yet, no attention has been paid to the desired *distribution* of this lengthening over the domain-final segments. In this section, an acceptability experiment will be described investigating whether listeners are sensitive to differences in the size of the domain that is lengthened. In the experiment, some original aspects of the final lengthening in the production material, such as its progressive distribution, are maintained. This experiment is therefore not designed to see if progressiveness is a first requisite for the perception of final lengthening, since all distributions used here are progressive. Rather, the objective of the present experiment is to answer research question (2b), repeated below:

(2b) Are listeners sensitive to differences in the size of the domain over which a certain amount of lengthening is spread?

If listeners are sensitive to such differences, we would expect them to require a larger domain of final lengthening when the final syllable has only a schwa in its rhyme, which is what we found in production.

2.3.2.2. Material

The material was taken from the production experiment. To reduce the amount of material, only the utterances produced by Speaker 1 were used. This speaker obtained the best results in the identification experiment described in §2.3.1, implying that this speaker made the clearest distinctions between the different boundaries. Only the PhP and IP boundaries are included, since an IP boundary is most clearly marked by final lengthening and the PhP boundary counts as the baseline condition. The PhP rather than the PW sentences were used because the PhP sentences structurally resemble the IP sentences more closely. For the same reason, the IP versions *not* corrected for the length of the IP in which the target word occurs were used (see §2.2.3.2). The sentences used are given in (9), with the target words underlined (for translations, see Appendix C):

(9) PhP Jan heeft een mooie harmonika gehoord, hier ver vandaan.
IP Jan hoort een mooie harmonika, gespeeld hier ver vandaan.

PhP Jan kon een goede marathon lopen, in zijn jonge jaren.
IP Jan liep een goede marathon, lopend in zijn jonge jaren.

PhP Henk wil de laatste mode volgen, zijn vrienden daarbij meeslepend.
IP Henk volgt de laatste mode, volgens zijn vrienden daar bij meetkunde.

PhP Zij gaan vandaag met een snelle tandem rijden, door de mooie duinen.
IP Zij zijn op weg met een snelle tandem, rijdend door de mooie duinen.

PhP Piet wil die rare yucca planten, gek als hij is.
IP Piet wil die rare yucca, plantengek als hij is.

For each sentence, two versions spoken by Speaker 1 were available. For the PhP boundary, the utterance was used in which the target word was shortest, and for the IP boundary, the utterance with the longest target word was used, so that the durational distinction between the two versions of the target word is optimal.

In order to test the acceptability of various duration distributions, the segment durations were synthetically varied (using PSOLA, see footnote 6 on p. 54) so as to conform to the outcome of the calculations given below. The durations meant to convey a PhP boundary are implemented in the original PhP utterance, while the longer durations meant to indicate an IP boundary are implemented in the original IP utterance.

As a starting point, the mean segment duration was calculated from the two segment durations (PhP/IP) of the target words in the sentences used. These mean segment durations give the first temporal word structure presented both in the PhP and the IP utterance:

Example: *yucca* /ˈjuka:/

	original PhP	original IP	mean segment durations	
j	58 ms	64 ms	(58+64)/2	= **61 ms**
u	79 ms	83 ms	(79+83)/2	= **81 ms**
k	88 ms	92 ms	(88+92)/2	= **90 ms**
a:	103 ms	196 ms	(103+196)/2	= **149.5 ms**

Thus, a segment which is lengthened at an IP boundary has a starting point which is longer than its PhP-duration, whereas the starting point for a non-lengthened segment is approximately its original duration, and the size of the lengthening effect determines the deviation of the starting point from the actual durations. In this way, the progressive distribution in the original data is incorporated in the determination of the segment durations for this experiment.

Starting from these mean segment durations, the target word in the original PhP utterance is shortened and the target word in the original IP utterance is lengthened. The lengthening (for IP) or shortening (for PhP) is distributed over either the final rhyme, the final syllable, or the penultimate rhyme plus the final syllable. The share that each segment within these domains receives depends on the share that segment's duration has in the total duration of the domain over which the lengthening is spread (on the basis of the mean segment durations):

Example: domain /uka:/, i.e. penultimate rhyme plus final syllable.

Duration of domain /uka:/ = 81 ms + 90 ms + 149.5 ms = 320.5 ms

share for u	= 81 / 320.5	= **25.27%**
share for k	= 90 / 320.5	= **28.08%**
share for a:	= 149.5 / 320.5	= **46.65%**

Thus, in absolute terms, a short segment is lengthened less than a long segment, but relatively speaking they are lengthened in equal proportions.

The total word duration (with respect to the mean durations) was increased (for IP) or decreased (for PhP) by either 50 ms or 100 ms, which corresponds roughly to the minimal and maximal amount of lengthening found in the production data.

Example: +/- 50 ms

share for u	= 50 ms * 25.27% = **12.64 ms**
share for k	= 50 ms * 28.08% = **14.04 ms**
share for a:	= 50 ms * 46.65% = **23.32 ms**

	+50 ms. for IP boundary:		-50 ms. for PhP boundary:	
u	(81 + 12.64 =)	**93.64 ms**	(81 - 12.64 =)	**68.36 ms**
k	(90 + 14.04 =)	**104.04 ms**	(90 - 14.04 =)	**75.96 ms**
a:	(149.5 + 23.32 =)	**172.82 ms**	(149.5 - 23.32 =)	**126.18 ms**

The durational structures for *yucca* are illustrated in Figure 2.8. The lines above the mean will be presented to the listeners in the IP boundary context; below the mean, the durational structures for PhP are shown. The condition '100 ms, final rhyme

(/a:/)' does not occur, since this would give unnaturally long or short durations. This condition is only included when the final rhyme consists of a vowel and a coda consonant (in *tandem* and in *marathon*). In the case of *mode*, the whole final syllable was not shortened by 100 ms either, since the result would be an unrecognizable /d/ of 13 ms.

All utterances were presented twice to the listeners. In all, the total material consists of 126 utterances (*marathon* and *tandem*: 28 times each, *yucca* and *harmonika*: 24 times each, and *mode*: 22 times).

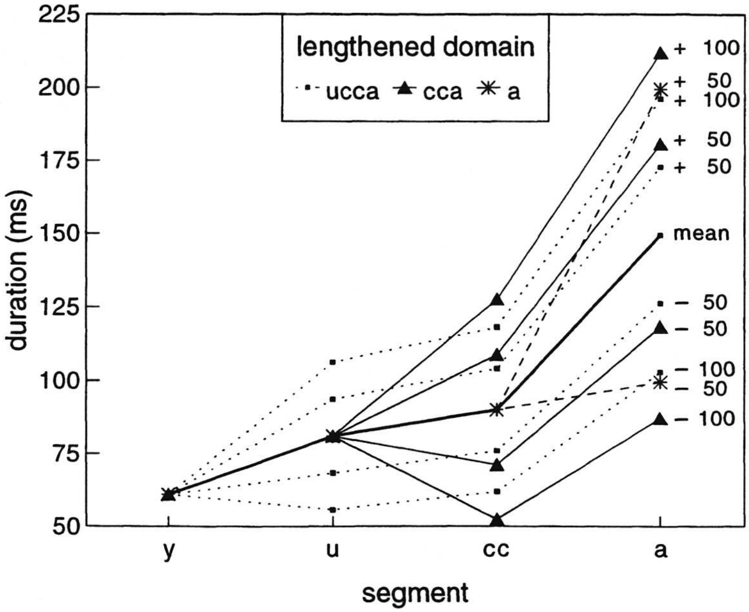


Figure 2.8. Durational structures for the word yucca *as material for the acceptability experiment.*

2.3.2.3. Subjects and procedure

28 native Dutch listeners, aged 18 to 67, participated in the experiment. They had no known hearing impairments.

The utterances were presented in group sessions in quasi-random order over good-quality headphones. Two successive sentences never had the same target word. Each utterance was presented on paper, with the target word underlined. Listeners were instructed to pay close attention to the durational build-up of that word, and to the position of that word in the sentence. They then had to indicate the acceptability of this durational structure on a scale from 1 (unacceptable) to 10 (perfect).

2.3.2.4. Results

An ANOVA was run with acceptability scores as the dependent variable, fixed factors of boundary type, total word duration and domain, with repeated measures over words and the results collapsed over listeners and repetitions. Since the 'domain' variable is not defined when the total word duration equals the starting point durations, this duration is excluded from this statistical analysis (leaving 2968 cases).

In Figure 2.9, the mean acceptability for each durational structure are given at a PhP boundary (left) and at an IP boundary (right). On the x-axis, the total word durations are given (mean, -/+ 50 ms, -/+ 100 ms), while the three bars give the results for the three domains over which the 50 or 100 ms are spread (domain i = final rhyme, domain ii = final syllable, domain iii = penultimate rhyme plus the final syllable).

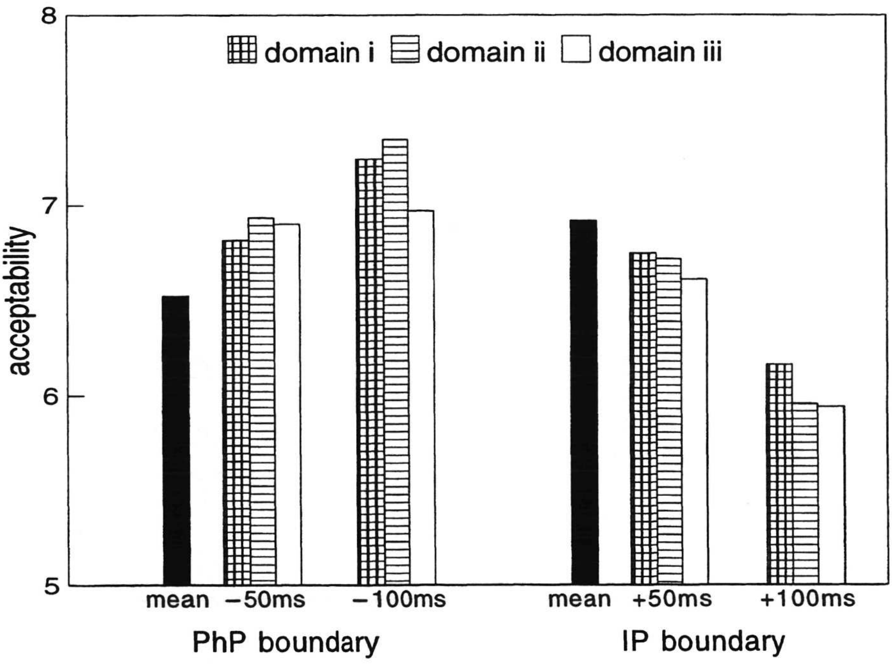


Figure 2.9. Mean acceptability of each durational structure, for all words taken together (domain i = final rhyme, domain ii = final syllable, domain iii = penultimate rhyme plus the final syllable).

In general, only the type of boundary (PhP or IP) and total word length (-/+ 50 ms, -/+ 100 ms) have an effect on the acceptability ($F[1,4]=10.44$, $p=.03$ and $F[1,4]=9.78$, $p=.03$ respectively). These factors interact too, as shown in Figure 2.9. ($F[1,4]=50.83$, $p=.002$). The domain over which the 50 or 100 ms is spread has not the least bit of influence on the acceptability ($F[2,8]<1$).

As can be concluded from Figure 2.9, the mean durations calculated from Speaker 1's original durations were actually fairly long, thus giving a high score when implemented at an IP boundary, and lower scores when placed at a PhP boundary. In fact, any further lengthening at an IP boundary sounds worse than the mean duration, while shortening at a PhP boundary increases acceptability (giving an interaction between boundary type and total word duration). Still, it is clear that only the total word length affects acceptability, but not the way this duration is built up; in fact, the difference between any of the three domains is generally as little as 0.2 on a 10-point scale.

To see if the acceptability of the spreading of final lengthening over the three domains is different for each of the target words, the results are given per domain and boundary type for each word in Table 2.5. Crucially, for the words included in this experiment (Words included in an ANOVA as a fixed factor), there is no interaction between the lengthened domain and the specific words ($F[8,2953]=1.6$, n.s.). If anything, we would expect that spreading across the penultimate rhyme and the final syllable in the IP boundary condition would be more acceptable for *mode* than for other words. This does not seem to be the case; in fact, this combination has the lowest acceptability score (5.6). Thus, the results are not even in the direction we would expect them to be on the basis of the results of the production experiment (§2.2.3).

Table 2.5. Mean acceptability per word, boundary type and domain.

word ▸ distribution ▾		marathon	harmonika	yucca	tandem	mode	total
PhP	domain i	6.9	7.3	6.8	7.2	6.2	6.9
	domain ii	7.0	7.3	7.3	7.3	6.4	7.1
	domain iii	6.5	6.9	7.2	7.5	6.6	6.9
IP	domain i	6.8	6.6	6.5	6.4	6.6	6.6
	domain ii	6.9	5.7	6.5	6.4	6.2	6.3
	domain iii	6.7	6.5	6.3	6.2	5.6	6.3

2.3.2.5. Discussion

In the production, a significant effect of final syllable weight on the domain of final lengthening was found. However, the present results indicate that the exact domain over which final lengthening is distributed does not influence the acceptability of the durational build-up of a word. Listeners only react in a systematic way to the total duration of the IP-final word. Looking at the results per word, no consistent effects could be detected; those conditions which agree with the results from the production do not sound more acceptable than those that do not. This indicates that the increased domain of final lengthening for schwa-final words is not perceptually relevant.

2.4. Conclusions

In this chapter, two production experiments and two perception experiments were described. The production experiments were designed to investigate whether boundary depth, stress and/or the structure of the final syllable/foot have any influence on the size of the domain of final lengthening (and on the magnitude of the lengthening effect). The pilot study described in §2.2.2 mainly addresses the appropriateness of the methodology using a repetition task with input speech without any temporal effects; this was considered desirable in order to elicit a constant intonational and rhythmical pattern and at the same time to make sure that the temporal effects found in the speakers' speech cannot have been copied from this input speech, and must therefore reflect natural lengthening effects added by the speakers of their own accord. A highly significant final lengthening effect was found at the IP boundary, consistently across speakers, so that we can conclude that the methodology is sound. The same methodology was therefore applied in the main production experiment described in §2.2.3.

The data from the main production experiment show that a large amount of final lengthening is found at IP-boundaries, and that it is progressively distributed across the segments of the final syllable or rhyme. These results agree with the general characteristics of final lengthening observed by other researchers (Edwards & Beckman, 1988; Berkovits, 1994; Wightman et al., 1992). Since in our data final lengthening could distinguish only between two levels (namely the intonational phrase level or above and the phonological phrase level or below), no relation could be established between boundary depth and the domain of final lengthening. Other studies have shown that a deeper prosodic boundary triggers a larger amount of final lengthening (Cooper & Paccia-Cooper, 1980; Nooteboom & Doodeman, 1980; Ladd & Campbell, 1991; Wightman et al., 1992); in our data, finer distinctions in the amount of final lengthening are only found in the final segment of some words (see, for example, Figures 2.1 and 2.2, and the *yucca* data in Figure 2.4). We take this to indicate that the duration of the final segment, which is affected most, is the most informative with respect to the depth of the following boundary, and that boundary depth does not influence the size of the domain which is lengthened.

The only case in which final lengthening exceeds the final syllable is with words with a final schwa, particularly with the word *mode* (and to a lesser extent with the word *tandem*). Similar results were obtained by Hofhuis (in prep.). The fact that this word forms one disyllabic foot cannot be the determining factor, since the same holds for other words included in the experiment in which final lengthening was restricted to the final syllable. We therefore argue that it is final syllable weight, or the inherent duration and expandability of segments closely tied to syllable weight, which influences the domain of final lengthening. Only when the final syllable is ultralight may final lengthening extend beyond this syllable to affect the penultimate syllable as well.

The identification experiment described in §2.3.1 showed that there is some correlation between the durational differences found in production and the type of boundary perceived by listeners. The IP boundary, clearly marked in production, is recognized very well, and even PhP and PW boundaries, which are durationally only marginally distinguished, are correctly recognized above chance level, with or without the presence of intonational cues. There was a significant correlation between the duration of the final syllable and listeners' responses. Moreover, speakers having long final syllable durations are often thought to produce Phonological Phrase boundaries whereas speakers having short final syllables are more often judged to have produced a Prosodic Word boundary. Durational aspects of the input therefore greatly influence boundary perception. The acceptability experiment described in §2.3.2, however, revealed that listeners are not sensitive to the size of the domain which is lengthened, but only to the amount of final lengthening. Thus, it is of no perceptual importance whether the lengthening is concentrated in the final rhyme or is spread out across the penultimate rhyme and the final syllable. Apparently, the effect of syllable weight on the domain of final lengthening found in production does not have any communicative function, since such differences are not picked up by the listener; if this effect was important to the listener, we should have found that a spreading of final lengthening over more than the final syllable was more acceptable for words with a light final syllable, and a concentration of final lengthening in the final rhyme should have been more acceptable for words with a heavy final syllable. Since no such tendencies could be detected, the differences in the size of the domain of final lengthening in the production must be due to factors only relevant for production. We suggest that it is the inherent short duration of a schwa, and its limited expandability (cf. Klatt, 1976), which necessitates lengthening of preceding segments in order to reach the desired amount of final lengthening which is required to indicate the depth of the following boundary. The final segment of *mode* is indeed only lengthened by 38 ms (=53%), while the final segment of the word *yucca* (also a CVcv word, and placed in a highly similar sentence) is lengthened by 72 ms (=70%).

As for implementation of final lengthening in speech synthesis programs, it seems clear that the determination of the appropriate *amount* of final lengthening is the most crucial factor. Presumably, this lengthening should primarily affect the final segment, and should at least extend across the final rhyme. If the required amount of final lengthening can be accomplished in these segments within the limits of their possible durations, no further measures should be necessary; if not, final lengthening may be spread across the final two syllables.