5. FINAL DISCUSSION AND CONCLUSIONS

5.1. Introduction

In this thesis, some experiments on final lengthening (marking the edges of prosodic constituents) and accentual lengthening (marking the head of an Intonational Phrase) have been described. Although both lengthening effects are well-established in the literature, the precise specification of the domain that is lengthened, the distribution of lengthening across the segments within the lengthened domain and factors influencing any aspect (amount, domain, and/or distribution) of these lengthening effects is elusive. While we do not claim to have answered all questions regarding final and accentual lengthening, or to have a full-scale specification of either of these effects ready, we hope to contribute to a better understanding of how these lengthening effects work in languages such as Dutch and English. A better understanding of these durational prosodic markers may be useful for automatic speech recognition, as well as for the improvement of synthetic speech, which still suffers from an unnatural temporal organisation.

In the following section (§5.2), we will summarize the main findings of each chapter. In §5.3, we place the topics of this thesis in a broader perspective, in a discussion of the practical and theoretical implications of our work. Finally, in §5.4, we will mention some limitations of the present research and give suggestions for further research.

5.2. Summary of main findings

5.2.1. Final lengthening

There is a vast literature on final lengthening as a boundary cue. Both in production and perception, boundary depth has been shown to systematically affect the amount of final lengthening: a deeper boundary is marked by more final lengthening in production, and listeners use this duration cue to assess boundary depth (for Dutch: Nooteboom & Doodeman, 1980; for English: Gussenhoven & Rietveld, 1992; Wightman et al., 1992). The domain which is affected by final lengthening is often assumed to be the final syllable (e.g. Edwards & Beckman, 1988; Berkvits, 1994) or rhyme (e.g. Allen et al., 1987; Gussenhoven & Rietveld, 1992). Wightman et al. (1992) also found that the final lengthening effect in (American) English is confined
to the final rhyme, although they do not consider the effect in the onset of the final syllable separately, and therefore cannot distinguish between the final rhyme or syllable as the domain of final lengthening. Hofhuis' (in prep.) data, on the other hand, indicate that the domain of final lengthening in Dutch may exceed the final syllable at least in some cases. The experiments described in Chapter 2 were meant to determine under which conditions final lengthening may extend further into the word than the final syllable/rhyme. During this search for the factor(s) influencing the domain of final lengthening, other aspects of final lengthening can be ascertained or verified in the process, such as the relation between boundary depth and the amount of final lengthening, and the distribution of final lengthening across the constituent-final segments.

The production experiments described in §2.2 included four types of boundaries: a Prosodic Word (PW) boundary, a Phonological Phrase (PhP) boundary, an Intonational Phrase (IP) boundary, and an Utterance (U) boundary. The IP boundary was consistently marked by more final lengthening than the PhP boundary (by some 65 ms), which in turn differed only marginally from the PW boundary (8 ms). Occasionnally (for some words and some speakers), more lengthening was found at the U boundary than at the IP boundary. The final segment's duration is most informative for boundary cueing, since in this segment the different types of boundaries are optimally distinguished from one another.

In a perception experiment (§2.3.1), it appeared that the (small) durational difference between PW and PhP boundaries could still be used by listeners to distinguish between the two. Even though these two boundary types were often confused, they were correctly identified above chance, regardless of whether the original intonation contour was preserved or had been reduced to only a declination line. The high correlation between final syllable durations and the listeners' responses suggest that listeners were indeed led by the input's durations, rather than by other possible (prosodic or segmental) cues. In all, the results are in line with the observations by others that boundary depth is cued by different amounts of lengthening, and that particularly Intonational Phrase boundaries are clearly distinguished from lower boundaries by lengthening of the preboundary segments. Also, the distribution of final lengthening in Dutch shows a consistent progressive pattern: most of the lengthening is found in the final syllable, and within the final syllable, most of the lengthening is found in the final rhyme/segment.

As for the domain which is affected by final lengthening, the final rhyme indeed proves to be the unit that is consistently affected. The onset of the final syllable is sometimes also lengthened, but not in all cases. The main production experiment (§2.2.3) revealed that the penultimate syllable may be included in the domain of final
lengthening, but only when the final syllable contains a schwa, i.e., when the expandability of the final syllable is low. It is not the case that final lengthening starts in the last stressed syllable, or that it has the final foot as its domain, since this would have given lengthening in more than the final syllable in other words as well (e.g. in *yucca*), yet the effect of boundary depth is clearly restricted to the segments in the final syllable when this syllable contains a full vowel. A first attempt to address the perceptual relevance of these differences between words (§2.3.2), however, did not reveal any effect of specific word type on the preferred domain size across which a fixed amount of lengthening is spread. Only total word duration was found to affect the acceptability of a word’s durational structure as a function of boundary depth. The finding that the domain over which final lengthening is spread does not affect the acceptability of the words’ durational structure suggests that it has little communicative function. On the basis of the results from Chapter 2, we concluded that it is the limited expandability of a schwa, which is a very short vowel and by definition unstressed, in combination with the importance of realizing the appropriate amount of final lengthening, which leads to a final lengthening effect extending further into the word than the final syllable. So, it seems that the domain of final lengthening is not a fixed (prosodic) constituent, but is primarily determined by phonetic restrictions.

The above conclusion is supported by the findings in Chapter 4, when compared to those in Chapter 2. In Chapter 4, we investigated final lengthening in combination with accentual lengthening. The disyllabic names included in the test materials for that experiment have a final unstressed syllable ending in either a full vowel (*Mina, Jannie*) or a reduced vowel/schwa plus consonant (*Kobus, Peter*). Looking at the effect of final position in these names, a total lengthening of some 130 ms was found, while the final two syllables of the target words from Chapter 2 were lengthened by only 70 ms (on average).\(^1\) Of these 130 ms, some 25 ms (range: 20-31 ms) are found in the initial (penultimate) syllable. So, while the final lengthening effect in Chapter 2 only extended back into the penultimate syllable when the final syllable contained a schwa, apparently the much larger effect in Chapter 4 exceeds the expansion possibilities of other types of unstressed syllables too, and hence part of the effect spills over onto the penultimate syllable. This finding contradicts the claim that only a final schwa leads to a larger domain of final lengthening, but does provide support for

---

\(^1\) Recall that subjects were asked to repeat synthetic utterances without any temporal effects in Chapter 2, but not in Chapter 4. Although in Chapter 2 speakers indeed added final lengthening of their own accord, hence reflecting a natural occurrence of this lengthening effect, this methodology may have influenced the amount of final lengthening. However, (part of) the difference in amount of final lengthening between Chapter 2 and Chapter 4 could also be due to differences in utterance length, utterance-internal structure, target word lengths and inter-speaker differences (see also §5.4 below).
the conclusion that this domain is larger than the final syllable only when this syllable cannot be lengthened to the desired degree, suggesting that the domain of final lengthening is predominantly determined by phonetic rather than phonological considerations. So, both the required amount of final lengthening (in order to properly signal the following boundary) and the expandability of the segments in the final syllable determine whether the domain of final lengthening is restricted to the final syllable or includes the penultimate syllable as well. As far as we can tell, differences in boundary depth are best reflected in the duration of the final segment, but a deeper boundary does not appear to lead to a larger domain of lengthening.

5.2.2. Accentual lengthening

The research concerning accentual lengthening in this thesis (Chapter 3) involved an experimental paradigm designed to answer very specific research questions. It does not address the question of the structural conditions under which accentual lengthening is found, but instead concentrates on the question of which segments are affected. In Chapter 3, we used contrastive focus to elicit a pitch accent on one of two accentable positions, while our main interest went to units directly adjacent to the accented syllable. Still, a general comment on the overall size of the accentual lengthening effects in our data may be in order. In non-final positions, we found an average lengthening effect of some 50 ms (range: 40-60 ms). Very similar amounts were also found in Chapter 4. In our data, roughly the same absolute amount of lengthening was found in disyllabic words as in monosyllabic words. So, the accented syllable of a disyllabic word is lengthened less than an accented monosyllabic word (on average, by 35 ms instead of 50 ms), since the former receives only part of the total amount of lengthening. For monosyllabic words, an absolute lengthening of 50 ms means a lengthening of some 20-25% with respect to unaccented durations, and for disyllabic words, this is a lengthening of some 15%. These amounts are comparable to those reported by others (e.g. Nooteboom, 1972; Sluijter & van Heuven, 1995, 1996).

Our primary goal was to investigate whether claims made in the past about different domains of accentual lengthening in Dutch versus English could actually be the result of different experimental designs. Research on Dutch seems to show that accentual lengthening is spread across all syllables in the accented word (e.g. Eefting, 1991; Sluijter & van Heuven, 1995), while data from Turk & Sawusch (1997) suggest that the domain of accentual lengthening in English begins with the accented syllable, i.e., does not include unstressed syllables to the left of the accented syllable. Yet the experimental design employed by Turk & Sawusch differs from that used in studies
on Dutch, and in such a way that the differences in experimental design could explain the different results. We therefore applied the Turk & Sawusch paradigm to Dutch, to see if comparable conditions would give the same results for the two languages or if a cross-linguistic difference could be confirmed.

Our secondary goal was to specify what type of boundary (or boundaries) blocks or attenuates the spreading of accentual lengthening from the accented syllable to neighbouring syllables. Although previous work on Dutch has addressed the question whether accentual lengthening affects only the accented word and/or the whole of the accented word (Eefting, 1991; van Heuven, 1993), these previous studies did not examine the actual blocking effect of a word boundary, since the presence vs. absence of a word boundary was not included as a variable. Also, while accentual lengthening in Dutch has been claimed to spread throughout the word, it is unclear how (whether) this word-sized domain can be defined more accurately (as a prosodic word, syntactic word, clitic group?)

By including test units of several types, we were able to investigate the effects of the boundary between onset/rhyme and nucleus/coda, between two syllables, two stems and two content words and between a content word and a function word. Within the accented syllable, both onset and coda consonants were found to be affected, confirming that at least the whole accented syllable is lengthened. Within the accented (syntactic) word, we found that unstressed syllables to the right of the pitch accent are lengthened more than unstressed syllables to the left. This contradicts Sluijter & van Heuven’s (1995) claim that accentual lengthening is spread evenly across all syllables in the word, yet is still compatible with their findings, and those of Nooteboom (1972) and Eefting (1991), in that all syllables within the word are lengthened to some degree. Crucially, our findings for Dutch are also compatible with those of Turk & Sawusch for English, in that there is an asymmetry between leftward and rightward lengthening. Moreover, Turk & White (1999) provide a more complete picture of accentual lengthening in English, in which syllables to the left of the pitch accent also show a certain amount of lengthening for some subjects (but less than syllables to the right), a picture highly consistent with the picture we now have of accentual lengthening in Dutch. So, these data show that once experimental differences are eliminated, accentual lengthening in Dutch behaves similarly to accentual lengthening in English.

A morphosyntactic word boundary was found to have a strong attenuating effect on the spreading of accentual lengthening, in that an unstressed syllable neighbouring the accented syllable is consistently lengthened less when there is an intervening word boundary than when it belongs to the same word (comparing e.g. da in PAN # damasten vs. PANDa # masten). The prosodic word boundary between two stems of a
compound, however, has no more effect than a normal syllable boundary (comparing the results for [REIS]ₚw(doel)ₚw vs. [PANda]ₚw). Rather, our Dutch data seem to indicate that the left and right edge of a morphological compound (i.e., a syntactic word, rather than a prosodic word) attenuate the spreading of accentual lengthening. We will come back to this point in §5.3.3 below.

In addition to the compound, we also seem to need the syllable to properly describe the distribution of accentual lengthening. In Turk & White (1999), the weaker leftward effect within the word is interpreted as an attenuating effect of the left edge of an accented syllable. Also, the accented syllable is generally lengthened most, and the distribution of lengthening within the accented syllable seems to be governed by different principles than within the (syntactic) word. Summarizing, we find that the distribution of accentual lengthening is influenced by syllable boundaries and syntactic word boundaries, but not by the prosodic word boundary between two stems of a compound. We can refute the claim that Dutch and English have different domains of accentual lengthening; rather, the different conclusions for Dutch and English drawn in the past can be attributed to differences in experimental design.

5.2.3. The interaction between final and accentual lengthening

In the literature, final lengthening and accentual lengthening are generally studied separately, in order to avoid any confounding of effects. In Chapter 4, an experiment is described which investigates specifically whether there is an interaction between final lengthening and accentual lengthening in Dutch and English. Since, in more spontaneous speech, the nuclear pitch accent often (by default) falls on the last accetable position, it may occur quite often that the word which undergoes final lengthening is also accented. It is therefore very important for speech synthesis (and speech recognition) to know what happens when final position is combined with accent.

In Chapter 4, the effect of a pitch accent on the duration of the Dutch target words is significantly smaller in final position than in initial or medial position. In non-final positions, an average accentual lengthening effect of 21% (47 ms) was found, an amount similar to that found in Chapter 3 and reported by others (e.g. Nooteboom, 1972; Sluijter & van Heuven, 1995, 1996), yet in final position the effect is reduced to a mere 4% (13 ms). A similar reduction of accentual lengthening in final position was found in Chapter 3. Initially, the frame sentence for the experiment reported in Chapter 3 was translated literally from the one used by Turk & Sawusch (1997), with the test material in intonational-phrase final position. The accented syllables in intona-
tional-phrase final or penultimate position were on average only lengthened by 6%,
while the accented syllables in non-final position were lengthened by 25%. (These results led to a revision of the frame sentence, see §3.3.3).

Results for the English target words from Chapter 4, on the other hand, show that in this language the absolute amount of accentual lengthening remains fairly constant across sentence positions. Indeed, Turk & Sawusch (1997) also found comparable amounts of lengthening across positions in their English study. Only in Dutch, then, do we find a significant interaction between final and accentual lengthening, in that the presence of final lengthening seems to suppress the realization of accentual lengthening. In English, there seems to be more 'room' within the durational limitations of the language to add accentual lengthening on top of final lengthening. This is found not only at the phrasal level, but also at the phonemic level (English vowels exhibit a larger range of durations than Dutch vowels; Elsendoorn, 1984). In §4.3, we argued that durational flexibility may have to be postulated as an independent property, which shows some variation across languages. A difference in durational flexibility may have phonological consequences: higher ceiling durations make duration a more reliable cue in English than in Dutch, and hence a feature more readily available for signalling multiple linguistic functions.

5.3. Practical and theoretical implications

The experiments reported in each chapter aimed to answer rather specific research questions. The results from these experiments also have some wider implications. In §5.3.1, we will call attention to several observed properties of lengthening effects that ought to be considered in any comprehensive durational model, e.g. the influence of linguistic structure, language-specific phonetic implementation, and the interdependence of final and accentual lengthening. In §5.3.2, we will reflect on difficulties that may arise when comparing lengthening effects across studies and languages, and in §5.3.3 we will discuss the implications of our results for prosodic constituency in Dutch.

5.3.1. Lengthening effects: ingredients for a comprehensive model

In this thesis, two durational phenomena were studied, which were assumed to mark aspects of prosodic structure (cf. Chapter 1). Within a stretch of speech, lengthening may signal the approach of a constituent boundary and/or the relative prominence of a syllable. In a very simple production model, the speaker may be thought of as planning the locations of accents and boundaries, while both accentual lengthening
and final lengthening follow as part of the phonetic realization of these accents and boundaries. In such a model, accentual lengthening may be viewed as a by-product of the extra time needed to articulate an important syllable clearly, and to make it salient (through the realization of a pitch movement). As van Heuven & Sluijter (1996) note, "the general claim is that an accented syllable is pronounced more elaborately, therefore more slowly, than an unaccented syllable." (p. 246). Perceptually, this also gives the listener more time to recognize the more essential parts of the speech flow. As for final lengthening, a parallel has been suggested with slowing down at the end characteristic of any mechanical movement (cf. Vaissière, 1983). Fowler (1990), for example, suggests that “[t]he lengthening may reflect the braking that inertial systems show generally as they stop gently.” (p. 206). Years earlier, Cooper & Paccia-Cooper (1980) claimed that final lengthening is produced “in order to allow the speaker an extra fraction of time for planning upcoming material in the next phrase”, in addition to “a general relaxation response” at the ends of major constituents (p. 199).

There are, however, some findings in this thesis and elsewhere which indicate that a comprehensive account of lengthening effects cannot dismiss accentual lengthening as purely a consequence of the time needed to make a syllable salient, or final lengthening as purely a natural slowing-down phenomenon. Here, we will concentrate on relevant observations based on our own data, and on supporting evidence from the literature bearing on the same issues.

First of all, in the case of accentual lengthening, we have shown that its distribution is strongly influenced by linguistic structure, which is unexpected if it is merely a by-product of accentuation. In Chapter 3, a (syntactic) word boundary was found to attenuate the spreading of accentual lengthening. Furthermore, in all of our data on accentual lengthening (from Chapters 3 and 4) we found a fairly consistent absolute amount of accentual lengthening (approximately 50 ms) in all words, which is realized in the accented syllable when this is the only syllable in the word, but is spread out across both syllables in disyllabic words. These observations show that the distribution of accentual lengthening is sensitive to word boundaries, and that accentual lengthening spreads over more than just the accented syllable in Dutch (provided that the accented word is polysyllabic). In contrast, pitch movements in Dutch are associated with a particular syllable (‘t Hart & Collier, 1975), and the size, slope and timing of a particular type of pitch movement is relatively fixed (cf. ‘t Hart et al., 1990). Thus, the distribution of accentual lengthening seems to be (partially) regulated independently of the actual pitch accent. Also, our finding that neither accentual lengthening nor final lengthening are restricted to one syllable imply that models which consider durational effects within the syllable only (as in Edwards & Beckman, 1988; Beckman, Edwards & Fletcher, 1992; see §1.2) may overlook an
important aspect of these lengthening effects.

Secondly, languages show various kinds of differences in the way that final and accentual lengthening are phonetically implemented. While some observed differences may be accounted for by referring to other structural differences between languages, this is not always the case. In Chapter 4 (§4.3), we noted that English seems to allow for more variability in duration at the segmental level and at the phrasal level than Dutch. This led us to the proposition that a certain degree of durational flexibility may be an independent, language-specific property. Indeed, Delattre (1966) also found differences in the extent to which different languages lengthen segments under the influence of prominence and boundaries: in a comparative study, he found the largest lengthening effects in English, smaller effects in German and French, and the smallest effects in Spanish. It may be that the perception of languages as 'stress-timed' or 'syllable-timed' (terms introduced by Pike, 1945) correlates in some way with the extent to which stressed and unstressed syllables are differentiated durationally (and otherwise); for discussion see e.g. Nespor (1990), Fletcher (1991). This traditional dichotomy can only account for some of the cross-linguistic differences, since it makes no claims about differences within either the stress-timed or the syllable-timed set of languages. It certainly cannot account for the difference we found between English and Dutch, since these are both held to be stress-timed languages. In so far that cross-linguistic temporal differences cannot be independently motivated by other (phonological or phonetic) differences between the relevant languages, they seem to call for a language-specific account of temporal phenomena, as they have also been developed for intonation (or pitch).

In Chapter 3, we refuted the claim that Dutch and English have different domains of accentual lengthening. However, other studies indicate that there are language-specific ways in which accentual lengthening is distributed. Fant, Kruckenber & Nord (1991), comparing data from English, Swedish and French, find a relatively symmetrical distribution of accentual lengthening within the syllable in English, a concentration of the lengthening in the vowel and coda consonants in Swedish, and a reverse asymmetry (lengthening in the vowel and onset consonants) in French. They observe that “[t]his finding is coherent with the relevance of VC units in Swedish and of CV units in French” (p. 361). The dominance (high frequency) of CV-syllables is indeed larger in French than in Swedish or English (an observation also linked to the 'stress-timed' vs. 'syllable-timed' dichotomy, see above). The relevance of a VC unit in Swedish is based on the relation between the phonological length of the vowel and that of the coda consonant; they are inversely related, i.e., a long vowel is followed by a short consonant and a short vowel is followed by a long consonant or a consonant
cluster (Elert, 1964). The segment quantity distinction in Swedish has also been found to interact with the distribution of accentual lengthening (Bannert, 1979). In Swedish, accentual lengthening primarily affects the long unit, whether this is the vowel or the consonant, and hence the quantity distinction is enhanced in accented environment (Heldner & Strangert, submitted).

Finally, the view that final lengthening and accentual lengthening should be analyzed as two separate, independent phenomena may not be appropriate for all languages. In Dutch, these two effects were clearly shown to interact with one another (see Chapter 4). Moreover, the extent to which these effects interact was shown to be language-specific. In the literature, we found evidence for an interdependence of final and accentual lengthening of quite another nature. There are at least two languages which, besides independent occurrences of final and accentual lengthening, also exhibit a lengthening that is both boundary-related and stress-related.² These are European Portuguese (Frota, 1998) and French (Fant, Kruckenberg & Nord, 1991; Hirst, Astésano & Di Cristo, 1998). Fant et al. (1991) compared stress-related lengthening in French, Swedish and English. They let subjects read (translations of) a piece of text from a Swedish novel, and stresses were assigned through listening. In French, 25 of the 64 stresses occurred in prepausal position (English: 14 out of 77, Swedish: 6 out of 72), and the distribution of stress-related lengthening was found to be different in prepausal positions versus other positions in French only. In non-prepausal locations, they find stress-related lengthening only in the first part of the syllable (onset and vowel), but in prepausal positions a larger difference between stressed and unstressed syllables is found, which is primarily located in the second part of the syllable (vowel and coda). This is shown in Figure 5.1, which is a copy of Fant et al.’s Figure 2. Note that the extra lengthening in panel (a) vs. (b) is not final lengthening; the figure shows stress-related lengthening, which is larger and distributed differently in prepause vs. other locations. Since the figure shows the increase in duration due to stress (not absolute durations), no information on final lengthening can be extracted from this figure (nor from the rest of Fant et al.’s paper).

² Contrary to our earlier definition of ‘stress’ as a word-level phenomenon, but in agreement with the literature discussed here, the term ‘stress’ in this discussion of French and Portuguese refers to any degree of prominence, as indicated by listeners.
**Figure 2.** Stress-induced increase in segment durations in French. (a) Prepause locations. (b) Other locations.

**Figure 5.1.** Taken from Fant, Kruckenberg & Nord (1991).

Fant et al.'s results for French are supported by Hirst, Astésano & Di Cristo (1998). They studied accentual lengthening in French, for various types of accent, in the onset, nucleus and coda of the accented syllable. They found accentual lengthening in the coda only when an accent is 'final in a terminal Intonation Unit' (as opposed to initial accents, and accents final in a non-terminal Intonation Unit). This finding is consistent with the results of Fant et al., depicted in Figure 5.1, if we can assume that 'final in a terminal Intonation Unit' corresponds to Fant et al.'s 'prepause locations'. In any case, these findings show that in French, it is not always possible to keep the effects of stress and final position on segment durations apart. This may have to do with the observation that stress typically falls on the final syllable both at the lexical and at the phrasal level in French (Fletcher, 1991), which results in an inherent confounding of prominence and final position. This confounding is evident in the paper by Hirst et al. (1998), in which 'final in a terminal Intonation Unit' is literally used as an accent category.

Frota (1998) runs into a lengthening in European Portuguese that is also related to both stress and final position. While Fant et al. studied stress-related lengthening and found that they had to take position into account, Frota studied final lengthening and found that it was sensitive to stress position. In her study of boundary phenomena, Frota found two distinct effects: (non-prepausal) IP-final lengthening and prepausal lengthening. The former is the interesting case: at the end of non-prepausal IP's, final
lengthening was primarily found in the last stressed syllable (also the head of the IP), which was either the penultimate or the final syllable in the IP. Note that the similarity between these findings and Fant et al.'s (1991) findings for French lies in the ambiguity of the lengthening as stress- or boundary-related, since this particular lengthening requires both stress and final position. They are different in that Frota reports a shift of lengthening from one syllable to another, while Fant et al. report on durational effects of stress and prepausal position within the syllable. Also, Portuguese does not have fixed final word stress, and hence does not confound stress and final position inherently (as French arguably does). Frota points out that the confounding in Portuguese could be another manifestation of the important role of heads in that language, and suggests that in this language the last stressed syllable absorbs both the head-related function of accentual lengthening and the boundary-related function of IP-final lengthening.

Summarizing, we argue that a comprehensive durational model should:

- look beyond the level of the accented/final syllable
- take effects of linguistic structure into account
- allow for language-specific phonetic implementation
- consider possible interactions between final and accentual lengthening; they are not always independent, separable effects.

5.3.2. Lengthening effects: making cross-linguistic comparisons

The experiment described in Chapter 3 was motivated by a careful comparison of the experimental designs of some Dutch and English studies concerned with accentual lengthening. The experiments on Dutch by Eefting (1991) and Sluijter & van Heuven (1995, 1996) placed words in accented and unaccented environments, and their materials seem suitable to address the question whether unstressed (as well as stressed) syllables in an accented word are durationally affected by a pitch accent. The experiments on English (Turk & Sawusch, 1997; Turk & White, 1999) were specifically designed to investigate whether unstressed syllables within and outside the accented word are lengthened due to accent, and can therefore address the same question. Still, the results revealed that the different conclusions which have been drawn for these languages on the basis of these studies were an artifact of experimental differences. The issues raised in Chapter 3 thus form a striking illustration of the special care we should take when making claims about cross-linguistic differences based on different studies. Even when studies may in principle be addressing the same questions, the precise set-up of materials may differ in ways that can influence the results, and hence the conclusions.
In the previous section, we tried to restrict the discussion of cross-linguistic durational differences to those found in comparative studies, such as in Delattre (1966) and Fant et al. (1991). A comparison of lengthening effects in various languages across studies is not only a precarious venture because of differences in experimental designs, but also because the literature on both final and accentual lengthening suffers from an inconsistent use of terms. In what follows, we discuss some illustrative discrepancies, contradictions and confusions that we came across in the literature.

A good example of disagreement across studies comes from the discussion on the universality of final lengthening. Vaissièr e (1983) refers to final lengthening as a prosodic feature common to a very large number of languages, yet mentions Finnish (Lehiste, 1965), Estonian (Lehiste, 1965) and Japanese (Han, 1961) as languages with little or no final lengthening. Notably, other researchers did find final lengthening in two of these so-called counterexamples. Krull (1997) refutes the claim that Estonian has no final lengthening. The literature on Japanese is an especially interesting example of confusion. Besides Vaissièr e’s reference to Japanese as a language with little or no final lengthening, Japanese has even been claimed to have sentence-final shortening (Kaiki & Sagisaka, 1992). Campbell (1992) questions these findings, and showed that Kaiki & Sagisaka’s ‘sentence-final shortening’ was at least partly due to a significant shortening effect of an onset plosive /t/ on a following vowel /a/, in combination with a predominance of the past-tense marker -ta in sentence-final position in a written-text-based corpora such as theirs. Indeed, in a controlled experiment Ueyama (1999) found clear lengthening at IP boundaries, but no consistent difference between IP and U in Japanese (that is, results highly similar to those for Dutch, cf. Chapter 2).

The claim made in Berkovits (1984) that Hebrew has no utterance-final lengthening is also later refuted, by Berkovits herself. Berkovits (1991) argues that the different findings for English and Hebrew with respect to utterance-final lengthening in her earlier study were due to differences in speaking rate.

Maddieson (1997), in his work on language universals, claims that lengthening in preboundary position is a ‘universal tendency’. Understandably, he hesitates to call it a true universal. The above examples, however, must have made clear that any apparent counterexamples need to be considered very carefully.

As pointed out earlier, final lengthening is not the only term used for boundary-related lengthening. Unfortunately, it is far from clear whether what is referred to as clause-final lengthening by Allen et al. (1987) is the same as Berkovits’ (1991) utterance-final lengthening, or as any of the following terms: phrase-final lengthening (e.g. Ueyama, 1999), preboundary lengthening (e.g. Gussenhoven & Rietveld, 1992),
prepausal lengthening (e.g. Frota, 1998), and/or domain-final lengthening (e.g. Fougeron & Keating, 1997). Although there might be real differences between languages in the types of boundaries that are marked by lengthening, without a unified terminology for prosodic constituents that is widely used and agreed upon cross-linguistically, it is virtually impossible to draw explicit conclusions on this subject from such a diverse set of studies.

Note that besides a non-homogeneous terminology, the relationship between boundary depth and the amount of final lengthening (e.g. Ladd, 1986; Wightman et al., 1992) may be an additional source of problems when comparing final lengthening effects reported in different studies. Results in terms of amounts of lengthening depend not only on the type of higher constituent that is expected or reported to exhibit final lengthening, but also on the type of material that is used as reference condition.

The literature on stress-related lengthening suffers from a serious confounding of the terms ‘stress’ and ‘accent’. As with final lengthening, the levels or types of prominence that are compared are likely to influence the results. Problems may also arise when comparing findings from different groups of languages, e.g. from stress-accent languages vs. tone languages, since the interpretation of a certain term (e.g. stress) depends highly on the type of language under investigation. Thus, comparisons across studies on accentual lengthening are not only hindered by a confounding of terms, but should also take into account the language-specific ways in which prominence distinctions are made, and in which F0 is employed.

As a case in point, let us consider the claim by Potisuk, Gandour & Harper (1996) that the role of F0 as an accent cue is diminished in tone languages (in this case, in Thai), since F0 is already employed in making lexical tone contrasts. Each of their target syllables occurs as the first member of either a noun-verb sequence, displaying a strong-strong stress pattern, or of a noun-noun compound, which forms a weak-strong stress pattern. Crucially, it is not clear from their paper whether it involves a stress distinction at the phrasal level, or only at some lower level (e.g. the word). The contribution of duration in making the strong-weak stress distinction was found to be much larger than that of F0 and intensity. The question is, should this hierarchy of cues be compared to that for lexical stress or for pitch accents in stress-accent languages? Potisuk et al. (1996), apparently unaware of the fact that only pitch accents, but not lexical stresses, are marked by F0 in stress accent languages, simply compare their results with the reported cues to accent. Considering the results, however, it may be more appropriate to compare this prominence distinction in Thai to lexical stress in other languages.
5.3.3. Lengthening effects: evidence for prosodic constituents

Prosodic phonology provides the constituents which serve as the domain of application for phonological rules, and which account for some of the acoustic patterns found in the phonetic realization of language. The definitions of some prosodic constituents in Dutch, given and discussed in Chapter 1, are primarily based on phonological evidence, and subsequently used as a basis for setting up our materials. Conversely, the durational data obtained in this thesis potentially provide evidence for the physical (and psychological) reality of certain constituents.

The final lengthening data confirm that the right edge of intonational phrases is clearly marked by lengthening. The domain which is lengthened is arguably not a prosodic constituent, but is determined by phonetic considerations. The domain of accentual lengthening, on the other hand, does seem to be constrained by linguistic structure. In Chapter 3, we found that the unit most suitable for describing the patterns in the data obtained so far is the compound, i.e. the terminal element of the syntactic tree. While the prosodic word may correspond to this unit in some languages, this is not the option proposed for Dutch: in Dutch, each member of a compound is said to form a separate prosodic word (Nespor & Vogel, 1986). The current Dutch prosodic hierarchy therefore does not provide us with the constituent apparently most appropriate to describe the distribution of accentual lengthening.

Is there any other evidence for the compound as a relevant phonological or prosodic constituent in Dutch? As a matter of fact, there might be, if we take a closer look at some phonological processes which have formerly been claimed to have the phonological phrase as their domain. Dutch has two phonological rules of voice assimilation, traditionally called Progressive Assimilation and Regressive Assimilation (Booij, 1995). The rule of Progressive Assimilation states that a fricative is devoiced after a voiceless obstruent, and the rule of Regressive Assimilation states that voiceless obstruents become voiced before a voiced stop, as exemplified in (1) (taken from Booij, 1995):

(1) Progressive Assimilation
- opvallend 'remarkable' \( pv \rightarrow [pf] \)
- afval 'trash' \( fv \rightarrow [f] \)
- stoepzout ‘pavement salt’ \( pz \rightarrow [ps] \)

Regressive Assimilation
- klapband ‘flat tyre’ \( pb \rightarrow [bz] \)
- kookboek ‘cooking book’ \( kb \rightarrow [gb] \)
- misdad ‘crime’ \( sd \rightarrow [zd] \)
Menert (1994) argues that since the application of these rules is sensitive to speech rate and style, they must have a constituent at least the size of the phonological phrase as their domain, on the basis of the claim that below this level 'performance factors' such as speech rate and style do not influence prosodic structure. Booij (1995), instead, claims that "Phonological rules like Progressive Assimilation, which apply obligatorily within prosodic words and compounds, become optional in larger domains: for instance, in phonological phrases." (Booij, 1995: p. 59). He then goes on to use the term 'prosodic compounds' to refer to the complex words within which these rules apply. If his claim about the environments in which these phonological rules apply is correct, it seems to provide phonological evidence for the compound as a relevant unit for Dutch phonology. So, while the prosodic word in Dutch is said to be as large as the stem because the stem serves as the domain of syllabification, r-colouring and coordination reduction, the compound also seems to be a relevant unit in Dutch, serving as the domain of voice assimilation and of accentual lengthening.

Notice that the issue cannot simply be solved by proposing a recursive prosodic word node (the compound), which joins together lower prosodic words (stems), since in principle phonological rules cannot refer to these finer distinctions within existing levels (Dresher, 1994). Rather, recursive structure is proposed precisely when both the higher and the lower constituents in such a structure within one level display the same behaviour with respect to a certain rule or phenomenon. For instance, Booij (1995) incorporates clitics into the prosodic word because a clitic+host combination forms one domain of syllabification (and the domain of syllabification is the prosodic word). Similarly, Ladd & Campbell (1991) propose that domains within a recursive structure allow boundaries of the same type to be realized with different strengths (not different means!). If a phonological rule or phonetic process needs to refer specifically to the compound as its domain of application, while other rules refer to the stem, this typically constitutes evidence for a separate level of constituency (cf. Nespor & Vogel, 1986).

We leave the issue of how the syntactic compound should be integrated into the prosodic hierarchy, and the analysis of multiple compounds, for future research, since our knowledge on this subject is too limited, and it is beyond the scope of this study.

5.4. Limitations and suggestions for further research

The largest part of this thesis is concerned with the production of final lengthening and accentual lengthening. In Chapter 2, two perception experiments (an identification experiment and an acceptability experiment) involving final length-
enning were described. The identification experiment (§2.3.1) served to demonstrate the importance of the final lengthening effect found in the production for boundary perception. The acceptability experiment (§2.3.2) was a preliminary attempt to address the perceptual importance of the domain across which final lengthening is spread, yet the distributional alterations made in the stimuli did not affect the acceptability of the target words’ durational structure. Clearly, less conservative variations are needed in order to find significant acceptability effects. In our experience, however, despite the ever-improving speech resynthesis possibilities, altering segment durations synthetically for experimental purposes still raises problems, probably because in natural speech the duration of a segment affects the extent to which it is spectrally reduced, and because its duration is evaluated with respect to the duration of surrounding segments. Thus, the interaction of the duration of one particular segment with that of other segments and with its spectral quality makes it difficult to study the effects of independently changing segment durations. Still, the progressive distribution of final lengthening versus the relatively linear distribution of accentual lengthening is bound to play a role in perception, yet the perceptual relevance of these different distributions remains an almost unexplored topic.

As far as the production of final lengthening is concerned, the largest gap in our knowledge is an explanation of the variation in the amount of final lengthening. While boundary depth has a well-established effect on this amount, there also seem to be differences between words, speakers, and effects of constituent length (e.g. utterance length) and/or constituent-internal structure, and maybe more. Little is known about how to account more precisely for the variability in amounts of final lengthening. It is a relevant issue from the perspective of speech synthesis and automatic speech recognition; as yet, many synthesis systems have no more than a three-level distinction for lengthening at boundaries, since it is unclear whether listeners would profit from any finer distinctions.

Within the experimental design employed in Chapter 3 to investigate accentual lengthening, some logical extensions would be to include longer words (i.e. with more unstressed syllables between the two stressed syllables), to add a ‘baseline’ condition with no accent on either of the stressed syllables (as is done in Turk & White, 1999), and to continue the search for the precise unit within which accentual lengthening is not attenuated. Future research will be aimed at the latter objective, investigating in more detail the behaviour of function words within the Turk & Sawusch (1997) design.

The paradigm used in Turk & Sawusch (1997), and replicated in Chapter 3 for Dutch, necessitates the use of nonsense-word combinations. The paradigm is pre-
eminently usable to tease apart the effects of position in the word and directional effects, and to have the same segmental material (the so-called test unit) across conditions, but generalizations to normal speech should be made with caution. In particular, we know very little about the realization of contrastive accents (as elicited in Chapter 3) as compared to non-contrastive accents. We are not too concerned about our use of metalinguistic contrasts: Sluijter & van Heuven (1995), comparing words in sentences with a metalinguistic interpretation (Did you say ‘COFFin’ or ‘MUFFin’?) with normally focused words (Did you buy a COFFin or a MUFFin?), showed that there are no durational differences between the two. Also, in Heldner & Strangert’s (submitted) paper on Swedish accentual lengthening, no differences were found with respect to lengthening between accents elicited through emphatic focus (Q: ‘Who is draining the pond?’ A: The MAN is draining the pond) versus contrastive focus (Q: ‘Is it the man or the woman that is draining the pond?’ A: The MAN is draining the pond). However, we think it likely that accents systematically elicited (often through focus manipulation, with the use of precursor questions as above) in an experimental task and environment, being more emphatic than accents found in spontaneous speech, exhibit much larger and clearer cues than normal accents. Therefore, it remains to be seen to what extent our results can be generalized to other types of accent occurring more often in normal speech. Work on the realization of contrastive accents versus spontaneous pitch accents would therefore be welcomed.

The finding in Chapter 4 that accentual lengthening is strongly reduced in final position in Dutch raises a number of related questions. First of all, it makes one wonder what happens to other accent cues in final position, such as pitch movement, amplitude, spectral quality and spectral tilt. In Chapter 4, a preliminary attempt was made to investigate the effect of final position on pitch excursion size. The data indicate that pitch excursions are slightly smaller in final position than in non-final positions, yet the reduction in pitch excursion size is not nearly as drastic as that of the amount of lengthening. A more full-scale study of the strength of each of the accent cues in final position is called for to confirm these findings. As a matter of fact, the hierarchy of cues to accent (i.e., the relative strength of each acoustic correlate of accent) needs to be re-established for final positions, since the durational cue in final position is reduced at least in production. While accentual lengthening is apparently suppressed by final lengthening, other acoustic parameters marking accent are not also used to mark boundaries, and so these accent cues (especially amplitude, spectral quality and spectral tilt) may remain unaffected in final position because they need not compete with boundary cues. One could even imagine that other accent cues are strengthened in final position, in order to compensate for the reduction of the durational cue.
Since a reduced durational effect in production is likely to affect the role of duration in perception, new perception experiments are also in order. In particular, we would like to know whether or not the reduction of accentual lengthening in final position affects prominence perception, or if other accent cues can fill up the gap left by the (at this point speculative) reduced role of duration.

In disyllabic names with initial stress, still only a very small amount of accentual lengthening was found in final position, even though it is then the penultimate syllable which is accented. Apparently, the interaction between final lengthening and accentual lengthening is not only found when the final syllable is accented, but also when the penultimate syllable is accented. It could be the case that accentual lengthening is reduced when the accented word is in final position, or that there have to be at least two syllables between the accented syllable and the boundary invoking final lengthening. In other words, more research is needed to establish what the minimal distance of an accent to the boundary should be, and in what terms this distance should be stated, in order to find a ‘normal’ amount of accentual lengthening (comparable to that found in initial and medial positions).

Finally, in contrast to the abundance of literature on the acoustics of boundary markers and accent cues, there is very little work on articulatory correlates. In the present context, articulatory studies may prove useful for disentangling the effects of various factors inducing lengthening. At the word level, on the basis of duration alone it is hard to distinguish between a stronger rightward effect of accentual lengthening, leading to longer durations for word-final unstressed syllables, and a combined effect of accentual lengthening and word-final lengthening (cf. Turk & White, 1999). At a higher level, it would be interesting to know more about the articulatory features of accent in final position, so that we can specify more precisely how final lengthening interacts with accentual lengthening.