Chapter 6

Conclusions

6.1 Introduction

The primary goal of the work reported in this dissertation is to test the existence of domain-edge and domain-span processes in suprasyllabic English speech timing—specifically, at the word level and the utterance level—and to determine the domain and locus of each process for which empirical support is found. Experiment 1 shows that the word is a domain of durational processes; Experiment 2 examines the *type* of processes which occur at the word level and the utterance level, and finds support for domain-edge processes, but little support for domain-span processes.

A number of domain-edge processes are identified in Experiment 2, including word-initial lengthening and utterance-final lengthening. There is also some evidence of phrase-initial lengthening and an absence of hierarchical lengthening utterance-initially. Chapter 1 asks whether domain-edge processes at different levels affect different loci: the evidence from domain-initial processes suggests that the locus is the syllable onset in each case; as noted below, there is also evidence of different levels of domain-final lengthening affecting different subconstituents of the word-rhyme.

There is little evidence for domain-span processes at either word level or utterance level. Word-span compression (polysyllabic shortening) is only found in pitch-accented words: what is termed the polysyllabic accent effect arises because the locus of accentual lengthening is the word, with lengthening greatest at word edges; variation in the distribution of accentual lengthening according to word length results in shortening of subconstituents in words of more syllables. There is little evidence for utterance-span compression: in certain cases, there are differences in subsyllabic constituent duration in utterances of different length, but these appear to be better interpreted as domain-edge processes. There is evidence of word-rhyme compression, a sub-word domain-span process. As word-rhyme compression is the only suprasyllabic domain-span process supported by empirical evidence, a more theoretically sat-

isfactory approach is to interpret it as a domain-edge process at the word-level: as discussed in Chapter 5, both utterance-final lengthening and word-rhyme compression (or word-final lengthening) appear to have a word-rhyme locus, although different subconstituents are affected in each case.

A model of suprasyllabic English speech timing incorporating these processes is outlined in Section 6.2, suggesting that the primary processes associated with prosodic structure are localised lengthening effects at domain edges and within pitch-accented words. The domains and loci of these processes are described in Section 6.2.1. Two important principles of the model, which distinguish it from some other approaches to speech timing, are discussed in Section 6.2.2. Firstly, the model proposes that suprasyllabic durational effects are directly related to linguistic structure rather than derived from factors which are essentially non-durational, such as biomechanical constraints on articulation. Secondly, in contrast with rhythmical approaches, the consequences of suprasyllabic linguistic structure are localised, rather than distributed throughout the speech string. Suggestions for future research which could test the predictions of the timing model are presented in Section 6.3.

The model presented below is concerned with suprasyllabic speech timing, but the scope of the term "suprasyllabic" intended here is slightly different from in the descriptive framework presented in Chapter 1. There the durational effects of lexical stress distribution are included among the suprasyllabic factors, which determine the durational consequences of the organisation of syllables into higher-level constituents; however, the evidence reviewed in Chapter 2 indicates little support for prominence-delimited constituents, such as feet headed by lexical stresses. Rhythmical processes, defined in Chapter 1 as the durational effects of the arrangement of lexical stress, appear to be localised—the stress adjacency effect—rather than reflecting a diffuse adjustment across stress feet according to the number of syllables they contain. Furthermore, the stress adjacency effect appears to be independent of prosodic constituent structure: the lengthening of a stressed syllable when followed by another stressed syllable occurs both within prosodic constituents and across constituent boundaries.

For the purposes of the discussion here, rhythmical processes may be reclassified as belonging to the set of syllabic factors outlined in Chapter 1. These arise from the organisation of segments into a string of stressed and unstressed syllables. This classification is not intended to suggest that speech timing is necessarily derived from syllabic and suprasyllabic representations thus formulated, but simply to indicate that the durational consequences of the organisation of syllables into higher-level constituents appear distinct from the durational effects of the distribution of lexical stress.

6.2 A model of English suprasyllabic speech timing

The suprasyllabic level of speech timing determines the durational consequences of the organisation of syllables into prosodic constituents. The effect of these structural processes is lengthening, localised at the edges of hierarchical constituents and within phrasally-prominent words.

This model adopts the view, suggested in Chapter 5, that word-rhyme compression may be best interpreted as word-final lengthening. Thus, there are no domain-span effects at the suprasyllabic level: that is, there is no inverse relationship between constituent length and subconstituent duration, either localised on the head of the constituent or diffuse throughout its length. Because of this, segments not contained within a locus of domain-edge lengthening or accentual lengthening are not subject to any durational process at the suprasyllabic stage which directly results from linguistic structure. The only suprasyllabic durational effects outside of the loci of lengthening may be some readjustment in segmental durations as a result of durational processes within the loci: in particular, as discussed in Chapter 5, there may be some compensatory shortening of the vowel following a syllable onset lengthened word-initially and possibly some residual lengthening following an accented word[1].

6.2.1 Sources of lengthening

Prosodic constituents

The constituents which are manifest by segmental and suprasegmental processes in connected speech are not always isomorphic with syntactic constituents, as discussed in Chapter 2. The experiments described in this dissertation are not designed explicitly to test the number of levels of prosodic constituency, but examine two particular levels, the word and the utterance; however, consideration of the results of Experiment 2 and of previous studies, as discussed in Chapter 5, suggests that there may be four levels of prosodic constituency: word, phonological (or intermediate) phrase, intonational phrase, utterance. The word, phonological phrase and intonational phrase are domains of initial lengthening, with a syllable onset locus; hierarchical lengthening effects may be absent utterance-initially, at least in isolated sentences. There appear to be at least two levels of final lengthening, word-final and intonational-phrase-final; whether the latter is distinct from utterance-final lengthening is discussed below. The locus of final lengthening is the word-rhyme, although different subconstituents are

[1] As suggested in Chapter 5, it remains to be seen whether residual lengthening, if confirmed, may be accommodated as an adjustment outside the locus (reflecting a difference in the articulatory strategies for final lengthening and accentual lengthening) or whether residual lengthening represents a problem that requires some adjustment in the picture of the locus of accentual lengthening presented here.

affected at different levels of constituency. Durational evidence for each of these constituents is reviewed now.

Prosodic word

Experiment 2 supports previous findings that the lexical word is a domain of initial lengthening, which has a syllable onset locus. Experiment 2 also suggests the existence of word-rhyme compression, an inverse relation between word-rhyme length and subconstituent duration: within the stressed syllable only the nucleus is affected, and a similar distribution is hypothesised with unstressed syllables. As discussed in Chapter 5, the fact that the word-rhyme—the domain of this effect—is also the locus of utterance-final lengthening, suggests that word-rhyme compression could be interpreted as a domain-final process at the word level. This explanation is preferred, as there is no support for other domain-span processes found in Experiment 2. Thus, according to this model, word-final lengthening affects syllable nuclei within a word-rhyme locus. The word is also the locus of accentual lengthening, as discussed below.

As discussed in Chapter 2, the mapping from lexical and function words to prosodic words—in particular, the prosodic status of compound lexical words and of function words—is a matter of unresolved debate, although there is some evidence that compounds, as well as monomorphemic lexical words, have a one-to-one mapping with prosodic words. Whether function words can stand alone as prosodic words is a less important issue for speech timing, because most function words are monosyllabic and are unlikely to receive phrasal stress. The prosodic level at which function words may form co-ordinate structures with adjacent lexical words remains to be established, as do the possible differences between proclitics and enclitics.

Phonological phrase

The evidence from initial lengthening in Experiment 2 suggests at least one distinct level of phrasing between the word and the intonational phrase. Experiment 2 was not designed explicitly to test for such levels of constituency, but the greater duration of word-initial syllables in utterance-medial context compared with near-utterance-initial context suggests that, in some cases, the utterance-medial words are preceded by phrase boundaries; these boundaries must be below the level of the intonational phrase, as intonational phrase boundaries are not present adjacent to utterance-medial test syllables. This finding agrees with Fougeron & Keating (1997), who find that onset duration is greater word-initially than word-medially and increases further phonological-phrase-initially and intonational-phrase-initially[2].

[2]The findings of Experiment 2 and those of Fougeron & Keating (1997) do not rule out the possibility that there are more levels of phrasing between the word and the intonational phrase, but it is more

Wightman *et al.* (1992) suggest that at least two levels of phrasing between the word and the intonational phrase are distinguished by final lengthening; however, in comparisons in Experiment 2 between utterance-medial and near-utterance-final test syllables, the utterance-medial syllables did not show greater subconstituent duration compatible with domain-final lengthening. This could mean either that the utterance-medial syllables are rarely followed by phrase boundaries or that such boundaries exist, but they are not marked by final lengthening. This finding should be treated with caution, as Experiment 2 is not designed to test for such an effect; as noted above, however, analogous comparisons found evidence of phrase-initial lengthening in Experiment 2. Furthermore, as discussed in Chapter 4, there are a number of reasons for believing that the results of Wightman *et al.* may not be replicable in normal speech.

Intonational phrase

Previous studies reviewed in Chapter 2 indicate: that both syllables within a disyllabic word-rhyme may be subject to lengthening; that lengthening tends to be progressive within this locus; that the nucleus and coda, but not the onset, of absolute-final syllables are lengthened. Thus, in Experiment 2, it is taken as established that the intonational phrase is a domain of final lengthening, and durational effects at intonational phrase boundaries are not investigated. There also appears to be an intonational-phrase-initial effect: results such as Fougeron & Keating (1997) suggest that syllable onsets are longer intonational-phrase-initially than when initial in lower-level constituents.

Utterance

Experiment 2 shows that utterance-final lengthening is progressive within a word-rhyme locus, but does not affect all of the subconstituents: the most consistent interpretation is that utterance-final lengthening affects syllable codas within the word-rhyme, plus the absolute-final syllable nucleus. It may be that, as Cambier-Langeveld (2000) finds for Dutch, utterance-final lengthening is no greater in magnitude or more extensive in locus than intonational-phrase-final lengthening, in which case the effect can be identified with just a single domain, the intonational phrase, because all utterance boundaries are necessarily intonational phrase boundaries. It may be, therefore, that of the three levels of constituency suggested by initial lengthening, only the word and the intonational phrase are marked by final lengthening; this hypothesis requires further investigation, however.

Experiment 2, and previous studies reviewed in Chapter 2, suggest that there is an

parsimonious to predicate the minimum number of levels required by the data.

utterance-initial shortening effect, at least for syllable onsets that are wholly measurable utterance-initially. As discussed in Chapter 4, this effect may arise because there is no reason for utterance-initial boundaries to be signalled suprasegmentally: the first syllable spoken is preceded by silence, and this is a sufficient cue to its position. If this interpretation of utterance-initial shortening is correct, then it simply represents an absence of hierarchical lengthening and is not logically associated with any particular domain: all constituents have their onsets aligned with the start of a new stretch of speech, except where a break is caused by interruption or hesitation pausing. As such, the effect does not represent evidence for the utterance as a constituent; thus, if the intonational phrase were shown to be adequate to account for patterns of pre-pausal lengthening, then there would be no durational evidence requiring that the utterance be proposed as a distinct prosodic constituent. As discussed in Chapter 2, however, there may be segmental evidence that the utterance may dominate two or more intonational phrases, although the intonational characteristics of the utterance are not established.

Phrasal prominence

It is argued in Chapter 1 that lexical stress and phrasal stress should be treated as distinct for speech timing purposes. This proposal appears to be justified: the durational effects of the distribution of lexical stress are independent of constituent structure, as discussed in Chapter 2; in contrast, Experiment 2 shows that the distribution of lengthening due to pitch accent is determined by constituent structure at the word level. Thus, accentual lengthening is included within the suprasyllabic level of speech timing, together with the domain-initial and domain-final processes also associated with suprasyllabic constituents.

Experiment 2 indicates that accentual lengthening has a specific locus—the word—and that lengthening tends to be greatest at domain edges. The total amount of accentual lengthening is no greater in polysyllables than in monosyllables and thus, where the word is longer, the accentual lengthening on any given constituent is reduced. This reduction is greatest for subsyllabic constituents that are word-initial or word-final in shorter words and word-medial in longer words.

Experiment 2, together with previous results, suggests that the durational effects of pitch accent do not strongly interact with domain-initial or domain-final lengthening. There is some evidence, however, that utterance-final (or intonational-phrase-final) lengthening may be slightly attenuated in accented words. This interaction may be a ceiling effect, reflecting limits to the expandability of segments.

Differences between Experiment 2 and previous studies regarding the distribution of accentual lengthening within the word are discussed in Chapter 5. One pos-

sibility is that different levels of phrasal prominence—contrastive/emphatic, nuclear, prenuclear—may be distinguished by variation in the magnitude and/or distribution of accentual lengthening.

6.2.2 The nature of suprasyllabic timing processes

The most important features of the timing model outlined above may be summarised thus:

- The durational consequence of suprasyllabic linguistic structure is lengthening within phonologically-defined loci.
- These localised lengthening effects only occur at domain-edges and within pitch accented words; the loci are different in each case.
- There are no domain-span processes; thus, there are no direct durational consequences of linguistic structure outside the loci of lengthening effects.

The following sections outline how these features indicate an approach to suprasyllabic speech timing which is different from some others which have been proposed: firstly, in this timing model, observed effects are directly related to linguistic structure; secondly, these durational consequences of linguistic structure are localised rather than distributed throughout the utterance.

Durational effects and linguistic structure

The timing model supports the view that structural processes—domain-initial lengthening, domain-final lengthening, accentual lengthening—function as cues to speech structure for listeners, and further proposes that these cues are distinctive because there are different loci for each type of process. These loci are defined by reference to phonological constituents, suggesting, for reasons outlined below, that the effects result from processes that are intrinsically durational. An alternative is to regard the location of durational effects as contingent upon the occurrence of linguistic events that are not durational in their abstract representation. Indeed, some statistically consistent variations in duration are not the result of processes which are fundamentally durational, but rather arise as a by-product of articulatory planning or biomechanical constraints: at the segmental stage, for example, an articulatory explanation may be proposed to account for the greater duration of labial plosives or the shorter steady state (and higher fundamental frequency) of high vowels.

At the suprasyllabic stage, utterance-final lengthening, for example, may be interpreted as a gradual deceleration in speech rate. Klatt (1976:1212) says: “It is not

known whether a speaker learns to lengthen segments at the ends of phrase boundaries in order to help the listener decode the message, or if there is simply a natural tendency to slow down at the ends of all motor sequences or planning units. Since utterance-final lengthening often extends over several syllables, it is probably related to the general deceleration of motor activity at the ends of speaking acts." Likewise, Cummins (1999:476) says that utterance-final lengthening "is characterised by global deceleration and a reduction in articulatory effort distributed over several syllables". Fowler (1990:205) goes further, asking whether the occurrence of pre-boundary lengthening effects in general "should be predictable from a description of the grammatical or phonological structure, from the syntactic structure or whether it shouldn't be predictable from the *grammar* at all? [...] The lengthening may reflect the braking the inertial systems show generally as they stop gently[3]."

The evidence from Experiment 2 shows, however, that the extent of lengthening before an utterance boundary is structurally determined, rather than a gradual effect over several syllables. When the final primary stressed syllable immediately precedes the boundary, the nucleus and coda are lengthened, but not the onset; where the final primary stressed syllable is the penultimate or antepenultimate syllable in the utterance, only the coda is lengthened, together with the following unstressed syllables. The distribution of lengthening is progressive, as would be predicted by the "gradual deceleration" interpretation, but the locus of the effect is defined in phonological terms[4]. If final lengthening has a phonologically-defined locus in English, as this evidence suggests, then it must be perceptible within the speech string or speakers of English would not be able to acquire the pattern in childhood.

A similar argument may be made with regard to accentual lengthening, which has been suggested to arise as a by-product of intonation, providing space where required for a full pitch excursion (for example: Beckman & Edwards 1992). Evidence discussed in Chapter 4 indicates, however, that pitch excursions—at least for rising prenuclear accents—are aligned according to accented syllable structure, and, more importantly, that it is the position of the fundamental frequency valley and peak, rather than the

[3]It may be questioned whether the vocal apparatus constitute an inertial system at the level alluded to by Fowler (1990:205). Individual articulators, such as the jaw and tongue, have inertia, and show patterns of acceleration and deceleration in articulatory gestures. There is no source of inertia within an ongoing stream of gestures, however, and it is perfectly possible to stop speaking suddenly and without lengthening of the final sequence of segments, for example, when interrupted mid-phrase.

[4]With regard to intonational-phrase-final lengthening, Turk (1999) observes that the effect extends to the primary stressed syllable in phrase-penultimate position, when it is followed by a secondary stressed syllable as well as by an unstressed syllable. The rhyme of both syllables is lengthened where the primary stress is penultimate, but the onset of the final unstressed or secondary stressed syllable, between the two lengthened rhymes, does not show a lengthening effect. If phrase-final lengthening is due to deceleration, it would not manifest such discontinuity within the locus. As described in Chapter 4, a discontinuous distribution of lengthening may also be inferred in utterance-final lengthening, although subconstituent durations are not measured for unstressed syllables in Experiment 2.

slope or duration of the rise, that is preserved in syllables of different durations (for example: Ladd *et al.* 1999). Thus, there is no reason for the accented syllable or word to be lengthened to accommodate the pitch excursion. If the distribution of accentual lengthening cannot be explained as an indirect effect of some other process, then it must be an intrinsically durational process; given that it has a phonologically-defined locus, the distribution of lengthening must be perceptible to listeners, or it could not be acquired.

Snow (1994) has evidence that suprasyllabic durational patterns are learned rather than a secondary effect of intrinsic speech production constraints or a by-product of intonation. In a developmental study, he finds that young children acquire phrase-final intonation patterns earlier than the lengthening of phrase-final syllables. Furthermore, a consistent pattern of final lengthening begins to emerge alongside "the transition to combinatorial speech", suggesting a developmental relationship between suprasyllabic speech timing and syntax.

The suprasyllabic durational processes in this model thus could be included amongst the set of timing rules which are part of the phonology of English, as proposed by Port (1981:272):

> But the more closely the temporal structures of language are examined, the greater the variety of patterns of implementation that are found across languages [...] and the greater is the extent to which phonological variables are found to control timing. This makes it more plausible to argue with Klatt (1974, 1976, 1979) that some timing rules, at least, should be viewed as part of the phonology of a language, even though it is still unclear to what extent the phonology itself must be modified to incorporate such rules [...] The proposal made here is that certain aspects of the timing of phonetic intervals in speech, which may be shaped over time by mechanical factors, are incorporated into the phonology of the language itself, and thus should not be viewed as simply part of the peripheral production mechanism.

It is not necessary, however, to propose that durational processes are phonological. An alternative could be to regard them as part of language-specific phonetic implementation. For example, Nolan (1998) distinguishes quantitative phonetic interpretation and linguistic-phonetic interpretation. Quantitative phonetic interpretation describes the processes whereby phonological entities are translated into acoustic realisation, according to the constraints imposed by the physiology of the vocal apparatus and the physical laws of aerodynamics: for example, the shorter steady state and higher fundamental frequency of high vowels, as mentioned above. In contrast, Nolan states that the essential characteristic of linguistic-phonetic interpretation "is that it draws upon, and manipulates, linguistic abstractions rather than parameters definable in the vocabulary of the physical sciences." This is the nature of suprasyllabic durational

processes suggested by the results of the experiments reported here and by other studies discussed throughout. The distribution of certain effects—initial lengthening, final lengthening, accentual lengthening—appears to be describable in terms of linguistic abstractions—such as the syllable onset, nucleus and coda, and the word—rather than in terms of physical processes such as deceleration of articulators.

Localised durational effects

The timing model outlined above suggests that suprasyllabic durational effects are localised at certain important points in prosodic constituent structure: boundaries and phrasal prominences. There is little evidence for the diffuse durational effects associated with the interpretation of domain-span processes—outlined in Chapter 1—which states that the duration of segments across the whole domain is inversely related to the phonological length of the domain. The alternative conception of domain-span processes, that the duration of certain subconstituents within the domain—in particular, the domain head—is inversely related to domain length, has some support at the word-rhyme level. The word-rhyme, however, is not a constituent into which an utterance can be exhaustively parsed. Thus, there is no support for the idea, typical of rhythmical approaches to duration, that segmental duration is directly determined by constituent structure throughout the speech string: for example, according to the isochrony hypothesis, one parameter which must be known for the determination of the duration of each segment is the length of the cross-word foot which dominates it. The approach taken here, in contrast, implies that parameters such as word length are only important within the loci of durational processes and do not affect segmental durations elsewhere.

An observation by Nolan (1998) makes an analogy between speech timing and intonation: "The representation of rhythm-related phenomena[5] has been handicapped because we don't have a 'rhythm contour' parallel to the F0 contour in which to identify significant events and trends." The analogy with fundamental frequency is interesting, because according to accounts of intonation within the autosegmental-metrical framework, as described in Ladd (1996), the significant events in fundamental frequency variation to which Nolan alludes do not themselves comprise a contour at the most abstract level of representation: the primitives of intonational representation are simply high and low pitch targets. Fundamental frequency is not specified throughout the utterance, and a contour becomes apparent through a process of interpolation between the abstract pitch targets when they are related to the segmental string.

The approach to suprasyllabic speech timing taken here is similar to the approach

[5]It is clear from the context that Nolan is using the term "rhythm" in its broadest sense—as outlined in Chapter 1—to embrace structural durational processes in general.

to intonation described by Ladd (1996), insofar as the important durational events are localised and between the loci the only durational processes may be adjustments to accommodate the localised effects. This approach contrasts with "rhythmic" approaches to speech timing, which have in common the identification of some constituent into which the speech string may exhaustively be parsed which imposes durational constraints on its subconstituents. The most obvious example is a timing model based on stress-delimited feet. As discussed in Chapter 2, it has very frequently been observed that stressed syllables in English tend to recur at regular intervals, and although this statement has distributional validity—it is indeed rare to encounter a long unbroken string of unstressed syllables—there is no support for its application within a model of speech timing. The time between successive stressed syllables is approximately proportional to the number of intervening unstressed syllables, and variations in the interstress interval reflect syllable composition as well as localised effects such as domain-edge lengthening.

The timing model presented here proposes that not only do stress-delimited units have no validity as primitives of a timing model, but also that there is *no* unit into which an utterance may be exhaustively parsed that consistently imposes timing constraints upon its subconstituents. This is in contradiction to many theoretical accounts of speech timing which propose that there are units which mediate between linguistic structure and segmental duration.

As discussed in Chapter 5, van Santen describes the "syllabic mediation" assumption behind some models of speech timing thus: "The duration of a segment depends mostly on the (pre-computed) syllable duration and the segment's identity" (van Santen 1997:237). Mediation may also apply to other constituents, implying the precalculation of that constituent's duration prior to the determination of the duration of subconstituents. It is a feature of rhythmical theories of timing, but is absent from the current model: here the determination of duration is primarily a bottom-up process—based upon segmental identity, syllable composition and stress distribution—with the effects of superordinate constituents being localised rather than diffuse.

Descriptions of speech timing by Couper-Kuhlen (1986), Local & Ogden (1996) and Cummins & Port (1998) propose a mediative function for stress-delimited units in speech timing, despite the lack of support for isochrony. Cummins & Port (1998) cite evidence to support this approach, based upon their experimental paradigm of "speech cycling", but, as discussed in Chapter 2, the finding that prominences tend to be placed at certain phases of the phrase repetition cycle appears to be highly task-dependent.

Most approaches to speech timing which suggest durational meditation by some suprasyllabic unit appeal to rhythmical units: that is, units delimited by lexical stresses.

In contrast, Nolan (1998) suggests a description of speech timing based upon speech rate within prosodic words. This is based upon the work of Dankovičová (1997) in Czech, who reports "rallentando" within an intonational phrase: thus, the local speech rate becomes slower in each successive prosodic word as the end of the phrase approaches. It may be that Czech differs from English in exhibiting diffuse durational characteristics; however, the only statistically significant effect of phrase position is the greater duration of the final prosodic word in the phrase, suggesting that there may in fact be similarly localised domain-edge effects. In any case, it does not seem a justified extrapolation to suggest that prosodic words should be regarded as having a privileged or mediative function in speech timing. This, however, is the implication of Nolan's proposal that the speech rate variation within successive prosodic words could be used in a description of the rhythm contour discussed above. Even if the suggestion is not intended to imply that prosodic words have a privileged or mediative function in speech timing, the gradual slowing down of speech rate over the intonational phrase implied by such a description has no place in English speech timing, as indicated by the result for utterance-final lengthening in Chapter 4.

Domain-span processes in speech timing, such as polysyllabic shortening and utterance-span compression, imply a mediative function for certain constituents, particularly where the locus is co-extensive with the domain. The experimental evidence reported in this dissertation suggests that such processes are not part of suprasyllabic speech timing. Furthermore, the evidence reviewed in Chapter 2 indicates that there is no mediative function for prominence-delimited units either.

6.3 Directions for future research

The model of suprasyllabic speech timing described in Section 6.2 proposes that there are two important types of process—domain-edge lengthening and accentual lengthening—and identifies the locus in each case. For both types of process, the number of levels of structure—constituents or prominences—which manifest distinct effects is still an open question, particularly with regard to final lengthening and accentual lengthening: comparisons within the experiments reported here and with other results suggests variation in the locus of lengthening which may be associated with distinct levels.

Experimental work which could resolve these questions is suggested in the following sections. In addition, because suprasyllabic processes are observed to be localised, it is assumed that they are perceptible by listeners, but apart from phrase-final lengthening, the use of such effects as cues to structure, particularly when confined to a specific locus, has not been examined.

Levels of initial lengthening

As outlined in Chapter 5, the evidence relating to initial lengthening supports previous findings, providing evidence for at least two levels of domain-initial lengthening below the intonational phrase—the word and the phonological/intermediate phrase—and indicating a syllable onset locus that is not more extensive at higher levels. There are three new findings associated with initial lengthening suggested by the results of Experiment 2: attenuation of word-initial lengthening in polysyllables; the absence of hierarchical lengthening utterance-initially; and compensatory shortening of subsequent segments.

Experiment 2 provides evidence of a greater magnitude of word-initial lengthening in monosyllables than in disyllables or trisyllables. As studies have often examined the effect of position in words of fixed length, this has not previously been noted, although a re-assessment of the data of Turk & Shattuck-Hufnagel (2000) provides some support for the finding, which is the only direct link between word length and subconstituent duration indicated by the experimental evidence reported here. An attempt to replicate this result should examine unaccented words—in accented words the observation is attributable to the polysyllabic accent effect—and place the words in contexts where they are not preceded by a higher-level phrase boundary. The duration of, for example, of /m/ in *man* ... *manage* ... *manager* should be measured, to see whether it is shorter in the longer words and if so, whether it is a binary effect—monosyllable vs polysyllable—as found in Experiment 2, or gradient, so that syllable onsets are shorter still in trisyllables than in disyllables. If evidence were found for such the latter effect, it would also be interesting to see if it were observed in unstressed syllables; for example, whether /k/ is longer in *con'tent* than in *contented*. If shown to be a reliable, the attenuation of word-initial lengthening in polysyllables may be examined as a cue to word structure.

Utterance-initial shortening of a syllable onset consonant relative to its word-initial duration was observed for the nasal /m/ and the fricative /s/ in Experiment 2, and proposed to indicate an absence of hierarchical lengthening when the utterance onset is cued by the termination of the preceding silence. Evidence of such an effect in a greater variety of onset consonants would support this interpretation. It would also to be interesting to see whether the distinct effect in partially-measurable onsets such as affricates and voiceless stops—no utterance-initial variation in stressed syllables and some utterance-initial lengthening in unstressed syllables—is replicable. If so, articulatory studies may be required to determine if this is a result of a more forceful stop release being produced utterance-initially.

The experiments presented here, in common with previous studies, have shown that the locus of word-initial lengthening is the syllable onset and does not extend to

the following vowel. Little work has been done, however, to determine if there is a lengthening effect on the vowel when in absolute word-initial position, although work has been done on glottalisation at higher-level boundaries (for example: Dilley *et al.* 1996), which will tend to increase the duration of the whole vowel measured from the onset of glottalisation. One problem with studying the duration of word-initial vowels is that, in word-medial position, they may not be regarded as syllable-initial according to the maximal onset principal when preceded by a consonant: for example, /ə/ in *unless* compared with *mason* could be regarded as non-syllable-initial in the latter context, according to the maximal onset principle. Furthermore, a syllable-initial vowel would be difficult to measure when preceded by another vowel: for example, the segmentation of /aɪ/ and /ə/ would be difficult in *iron* compared with *unless*. Some data on this question would, however, be useful, particularly in resolving the potential influences on the central unstressed syllable in pairs such as *bake enforce* and *bacon force*, as discussed below.

The possible existence of compensatory shortening following word-initial or phrase-initial lengthening of the onset, discussed in Chapter 4, could also be investigated. For compensatory shortening of the nucleus following word-initial lengthening, there are difficulties in the design of appropriate materials due to the possible confounding factors of position-in-word—syllable nuclei may be subject to word-final lengthening, depending on word structure—and word length in accented words.

Levels of final lengthening

An assumption behind the experimental work reported here is that the intonational phrase is a domain of final lengthening, and this is not directly tested. There is little evidence in Experiment 2 of phrase-final lengthening at lower levels, although this could be interpreted as an absence of phrase boundaries following the measured syllables in the experimental materials or the lack of final lengthening at phrase boundaries. Utterance-final lengthening is shown to have a distinct, phonologically-defined locus—the word-rhyme—but only syllable codas and the final syllable nucleus appear to be lengthened within this locus: subsyllabic durations are required to confirm the distribution in unstressed or secondary stressed syllables following the final primary stressed syllable. There is also evidence that the word-rhyme may be the locus of word-final lengthening, with syllable nuclei lengthened within the locus, although this could also be interpreted as a word-rhyme span effect.

An important question is whether the utterance-final effect should actually be associated with the intonational phrase, given that utterance-final boundaries are necessarily also intonational phrase boundaries. A speech production experiment similar to that carried out by Cambier-Langeveld (2000) for Dutch could determine whether

the locus of lengthening is more extensive utterance-finally than intonational-phrase-finally, and whether the magnitude of the effect differs between the two levels: it seems likely that, as for Dutch, no difference might be found, thus supporting the idea of a single effect at the intonational phrase level. Cambier-Langeveld also examines durational effects word-finally and phonological-phrase-finally and finds no difference between the two; the evidence from Experiment 2 suggests that this result might also be found for English, if materials can be devised to ensure the reliable placement of lower-level phrase boundaries.

Distinguishing between the word-rhyme compression and word-final lengthening interpretations of the effect seen in Experiment 2 on stressed syllables in left-headed words is difficult because position-in-word and word-rhyme length are necessarily confounded. As outlined in Chapter 5, however, the durational effects on unstressed syllables of word-rhyme length and position-in-word may be distinguished in order to choose between the word-rhyme compression and word-final lengthening hypotheses. The results for Experiment 2 show a durational difference in unstressed syllables in left-headed disyllables and trisyllables: for example, /ən/ is shorter in *masonry* than in *mason*, a result which is compatible with word-rhyme compression and with word-final lengthening. If the additional syllable in the trisyllable is placed before the common unstressed syllable, however, only word-rhyme compression predicts a durational difference: for example, /m(ə)n/ should be shorter in *posterman* than in *postman*; likewise /ɪŋ/ should be shorter in *surfing* then in *surfeiting*. An absence of a durational difference would be strong support for the word-final hypothesis.

In addition, the word-rhyme compression and word-final hypotheses make different predictions about the distribution of durational differences in unstressed syllables in word-initial and word-final position. As noted in Chapter 5, studies such as Turk & White (1999) and Turk & Shattuck-Hufnagel (2000) do not find a difference in unstressed syllable duration according to word membership in phrases such as *thankful Phil* and *thank fulfil*, but subsyllabic durations are not examined for unstressed syllables and the possibility remains that there is a balance of word-initial and word-final processes. The word-final hypothesis predicts that the nucleus (and possibly the coda in unstressed syllables) should be longer word-finally whilst the onset should be longer word-initially. The word-rhyme hypothesis is neutral as regards this difference: word-finally, the unstressed syllable is in a maximally long word-rhyme[6]; word-initially, it does not belong to a word-rhyme.

[6] A monosyllabic word-rhyme would not contain any unstressed syllables; thus a disyllabic word-rhyme is maximally short as regards the unstressed syllable.

Levels of phrasal prominence

Experiment 2 strongly suggests that the word is the locus of accentual lengthening. As discussed in Chapter 5, there are differences in the distribution of lengthening between these results and some previous studies. These differences may reflect variation in accentual lengthening according to the level of phrasal prominence. A direct experimental comparison of the magnitude and distribution of lengthening between emphatic/contrastive stress, nuclear pitch accent and prenuclear accent could resolve this issue: such as result would provide some support for the idea of the linking of hierarchies of constituents and prominences, as it would suggest that both are associated with a number of levels of distinctive durational effects.

The existence of small residual lengthening effects on the syllable following an accented word presents a potential problem for the view of the word as the locus of accentual lengthening. It is possible that the observation of some lengthening following an accented monosyllable (Turk & White 1999) may not be replicated with accented polysyllables, particularly if they are uttered in normal sentences rather than metalinguistic context, where the whole phrase may be in focus to some extent.

Cues to structure

As discussed in Chapter 2, it is well established that phrase-final lengthening can act as a cue for listeners to a boundary in speech; however, the perceptual impact of the lengthening of word-initial and phrase-initial syllable onsets has not been greatly investigated. It is hypothesised here that listeners are capable of using syllable onset duration as a cue to word juncture and to phrasing.

The effect of word-initial lengthening could be investigated experimentally using materials with lexically ambiguous word segmentation. Speech synthesis techniques allow materials to be constructed which lack spectral variation and other possible cues to structure. For example, perception of the boundary in homophonous phrase pairs such as *bake enforce* and *bacon force* could be investigated by varying only the duration of /f/; the prediction being that longer durations of this segment should be associated with a higher likelihood of subjects choosing the *bacon force* interpretation of the materials. If compensatory effects are robust, then a certain degree of shortening of the /ɔ/vowel following the /f/ should further increase the rate of identification of this as a word-initial consonant. In addition, lengthening of the word-initial /b/ may, other things being equal, incline subjects to the *bake enforce* interpretation, because initial lengthening appears to be attenuated in polysyllables.

Similar techniques could be applied to investigate initial lengthening at the phrasal level. The types of structural ambiguities discussed in Chapter 2 with regard to the investigation of final lengthening could also be used in this instance, with the ex-

perimental hypothesis that higher degrees of initial lengthening following a phrasal boundary facilitate perception of that boundary, either by making it more likely or by shortening reaction times in on-line perception tasks. It may well be that subjects are more sensitive to degrees of durational variation in the potential loci of suprasyllabic effects than in the speech string as a whole, and psychophysical experiments using materials lacking a complete prosodic structure may over-estimate the just noticeable differences required at important positions in speech.

Investigation of the perception of phrasal stress could also be informed by the results reported here, which suggest that lengthening is distributed across the accented word, but is greater at word edges. If this is correct, then the identification of words in lexically-ambiguous pairs of the *bake enforce* vs *bacon force* type should be facilitated, when one is in focus, if the distribution of lengthening more closely resembles the pattern established in Experiment 2. Whether such a mechanism is used to identify words in normal speech perception is not certain. It seems more likely that degrees of accentual lengthening, and possibly variation in its distribution within the word, may serve as cues for the listener in determining the level of phrasal prominence.

6.4 Summary

The purpose of this dissertation has been to investigate durational processes associated with suprasyllabic speech structure utilising the domain and locus framework outlined in Chapter 1. Two types of processes were suggested by previous results—domain-edge processes and domain-span processes—and the evidence was examined for their existence at the word level and the utterance level.

The experimental methodology has been supported: examining the loci of durational effects serves to distinguish between different underlying processes, distinguishing, for example, word-edge from word-span processes and phrase-initial from utterance-span processes. A picture of suprasyllabic speech timing is suggested by these results in which the important processes are:

- Domain-initial lengthening.
- Domain-final lengthening.
- Accentual lengthening.

The existence of domain-span processes is not supported at the word level or the utterance level, and it is suggested that the diffuse durational processes that domain-span processes imply have no place in suprasyllabic speech timing, in contrast with the claims of accounts based upon rhythmical constituents.

There is evidence for at least two hierarchical levels of initial lengthening below the intonational phrase, including the word, and also evidence of a distinctive absence of lengthening utterance-initially. In all cases, the locus of initial lengthening is the initial syllable onset; the extent of lengthening does not alter as the magnitude of lengthening increases at higher levels.

Utterance-final lengthening is shown to have a word-rhyme locus, and it appears that syllable codas and the final syllable nucleus are lengthened within this locus. Further investigation is required to determine if utterance-final lengthening is distinguished from intonational-phrase-final lengthening by the magnitude or the locus of the effect. The word-rhyme is also important at the word level, as it appears to be the domain of a span compression effect, affecting syllable nuclei within the locus. Given the lack of support for other domain-span processes and the structural similarities between this effect and utterance-final lengthening, an interpretation of the observation as word-final lengthening is preferred; further investigation is required to resolve this issue.

The locus of accentual lengthening is shown to be the word. The distribution of lengthening varies between monosyllables, disyllables and trisyllables, with the edges of words tending to manifest the greatest effect. The durational pattern that this variation produces appears to account for what has previously been described as polysyllabic shortening. Turk & Shattuck-Hufnagel (2000:428) allude to the apparent complexity of durational processes at the word level: "We look forward to the day when the patterns we account for with a complicated set of mechanisms can be explained by a more parsimonious model." This picture presented in this dissertation is a move towards that goal, by proposing only one durational mechanism—accentual lengthening—that is unique to the word level, together with initial and final lengthening, that apply at various levels of constituency.

The constituents to which suprasyllabic timing processes apply are taken to be prosodically defined. Much work remains to be done to determine the combination of factors which contribute to the construction of prosodic structure within speech. The evidence presented here strongly suggests, however, that the durational consequences of prosodic structure are not distributed throughout the speech string, but localised at significant points. In this way, the extraction of useful information from the mass of influences on speech timing appears a more tractable problem for the listener.