English Speech Timing: A Domain and Locus Approach

Laurence White

PhD The University of Edinburgh 2002

Abstract

This dissertation presents a descriptive framework for suprasyllabic processes in speech timing, and describes speech production experiments that investigate durational processes at the word level and the utterance level and the interaction of these processes with the effects of pitch accent. The experimental evidence suggests a model of a speech timing comprised of localised effects, in contrast with the diffuse processes typical of accounts that focus on the rhythmic organisation of speech.

Within the descriptive framework, two types of process are associated with the domain, a familiar concept in prosodic phonology. Domain-edge processes lengthen segments near the initial and final boundaries of constituents: for example, word-initial lengthening and utterance-final lengthening. Domain-span processes are hypothesised to arise from an inverse relationship between the size of some constituent and the duration of some subconstituent: for example, word-span compression (polysyllabic shortening) and utterance-span compression.

The particular segments affected by each domain-edge or domain-span process are termed the ”locus”: for example, the word is a domain of initial lengthening and the locus is the word-initial syllable onset. It is hypothesised that each process is associated with a locus defined in phonological terms, and that processes may be distinguished by their distinct loci. The experimental work examines the loci of durational effects, indicating support for domain-edge processes—but not domain-span processes—at the word level and the utterance level.

Utterance-final lengthening is found to be progressive, affecting syllable codas and the final syllable nucleus within a word-rhyme locus. These results contradict the idea of a gradual deceleration in speech at the end of utterances. Utterance-initial shortening suggests that where the boundary cue is the termination of the preceding silence there is an absence of the hierarchical lengthening demonstrated word-initially and phrase-initially. There is no evidence of an utterance-span effect.

Word-initial lengthening is supported, with a syllable onset locus, as indicated by previous results. Word-initial lengthening is found not to interact with accentual lengthening, and may be attenuated in polysyllables.

Polysyllabic shortening, a domain-span process at the word level, is not supported. The previously-observed effect arises from variation in the distribution of accentual lengthening between monosyllables, disyllables and trisyllables. The locus of accentual lengthening is shown to be the word, with the greatest lengthening tending to be found at word edges. Because total lengthening is no greater in polysyllables than in monosyllables, the effect on particular subconstituents is attenuated when the word contains more syllables.

Word-rhyme compression is proposed to account for variation in nucleus duration according to the number of subsequent syllables in accented and unaccented words. Because it is the only domain-span process supported, it may be theoretically preferable to interpret word-rhyme compression as a domain-edge effect at the word level, similar to utterance-final lengthening but affecting nuclei rather than codas within the locus.

This dissertation is my own work, and describes research that I have carried out. No part of this dissertation has been submitted for any other degree or professional qualification.

Laurence White

To my Mum and my Dad
with all my love.

Contents

Acknowledgements

I have greatly benefited from the help of many people during the production of this dissertation, and I would like to thank the following colleagues and friends for their particular contributions:

—my supervisors Alice Turk and Steve Isard, for sharing their knowledge and experience, for their encouragement and their attention to detail and, latterly, for their long-distance draft-reading.

—my examiners Bob Ladd and Jan van Santen, for their helpful and encouraging comments and for a very interesting discussion.

—present and former computing and technical staff of the Linguistics Department, including Mike Bennett, Morag Brown, Norman Dryden, Eddie Dubourg, Irene Macleod, Cedric Macmartin and Stewart Smith, for their friendly and highly effective assistance; also Rob Clark, for his unerring technical advice.

—present and former administrative staff of the Linguistics Department, including Fiona Gilbert, Lynda Hutchison and Ethel Jack, for ever-willing assistance.

—present and former staff and students of the Linguistics Department, for contributing to a friendly and stimulating working environment.

—Louise Kelly, Simon King, Ineke Mennen and Astrid Schepman, for their tremendous help, advice, kindness and encouragement during my time in the Linguistics Department.

—Marcus Ackroyd and Lucy Ellis, for blazing the Sisyphean trail.

—the Monday lunchtime five-a-siders, for getting the week started in my Edinburgh days.

I would also like to thank my family and my friends, for their encouragement and for all their support in recent years.

To do justice to the positive contribution of Sarah Davies to my life over the time that I have been writing this dissertation would require more pages than follow this one. My thanks are unbounded.

Chapter 1

The domain and locus in speech timing

1.1 Introduction

Speech timing is a potentially ambiguous guide to speech structure. There are a large number of influences on the duration of segments within a speech string, some of which are a result of factors which are not fundamentally durational, such as biomechanical constraints on speech production. Evidence presented in this dissertation suggests, however, that certain suprasyllabic timing processes are directly related to linguistic structure within particular domains; furthermore, each process has a characteristic locus within its domain. Because the loci are distinct, the apparent ambiguity in speech timing is reduced.

The descriptive framework incorporating the concepts of the domain and the locus, representing cause and effect in speech timing, is described below, where a characterisation of domain-edge and domain-span processes is abstracted from a description of lengthening and shortening effects. The experimental work reported here focuses on domain-edge and domain-span processes at the word level and the utterance level. The results suggest a model of suprasyllabic speech timing based upon localised lengthening effects at domain edges and in pitch-accented words, each process having a particular phonologically-defined locus. There is little empirical support for domain-span processes, in contrast with approaches to speech timing that suggest that suprasyllabic linguistic structure has durational consequences throughout the speech string.

1.1.1 Lengthening and shortening

Much of the systematic variation in English speech timing can be described either in terms of constituent structure or in terms of the distribution of prominences. Within both frames of reference, relatively large increases in duration tend to be associated with intonational events: constituents that have an intonational marking at their final edge—a boundary tone—also tend to show lengthening of segments near the boundary (for example: Price *et al.* 1991); likewise, the phrase-level prominence indicated by a pitch accent is associated with significant lengthening of the accented syllable and neighbouring syllables within the word (Sluijter 1995; Turk & Sawusch 1997; Turk & White 1999).

The timing consequences of constituent structure and the distribution of prominences are not restricted to the locations of intonational events, however: at lower levels, both prominence and constituent boundaries may be marked primarily by duration. With regard to prominence, segments are longer in lexically-stressed syllables than in unstressed syllables; furthermore, a stressed syllable when followed immediately by an unstressed syllable may be shorter than when followed by another stressed syllable (for example: Van Lancker *et al.* 1988). With regard to constituents, both initial and final boundaries may be indicated by lengthening: syllable onset consonants are longer word-initially than word-medially (for example: Oller 1973; Cooper 1991); the final syllable rhyme of words and phrases may be lengthened even in the absence of an intonational phrase boundary (for example, Beckman & Edwards 1990; Wightman *et al.* 1992). Additionally, the duration of the constituents of a lexical word may be influenced by its length: "polysyllabic shortening" refers to the hypothesised inverse relationship between primary stressed syllable duration and the number of syllables in the word (for example: Lehiste 1972; Port 1981).

The timing consequences of such effects when investigated experimentally are described by comparing segmental durations at two or more levels of some condition, with the many other potential sources of durational variation kept constant as far as possible. As there is no procedure for fixing an objective reference duration for segments of any given type, any comparison between segments of different durations may be described either as "lengthening" or "shortening", but the conventional usage of these terms reflects an apparent distinction of processes. Typically, processes associated with constituent boundaries are described as lengthening, relative to the duration of segments placed constituent-medially, for example:

Word-initial lengthening: /p/ is longer in *porter* than in *report*.

Phrase-final lengthening: /æk/ is longer in

The big cat was black, but the small one was grey.

than in
The black cat was big, but the grey one was small.

The effects of prominence, particularly phrasal stress, are also described in terms of lengthening relative to the absence of prominence. For example:

Accentual lengthening: /sæmən/ is longer in
I said salmon sandwich, not tuna sandwich.
than in
I said salmon sandwich, not salmon salad.

In contrast, durational variation resulting from the phonological length (typically, the number of syllables) of the constituent in which measured segments are placed is generally described as shortening: thus, the duration in the phonologically-shortest constituent is the reference point[1]. A stressed syllable is held to have greatest duration in a monosyllabic word or utterance, and to become shorter as more syllables are added within the constituent. For example,

Polysyllabic shortening: /stɪk/ is longest in *stick*, shorter in *sticky* and shorter still in *stickiness*.

Similarly, variation in stressed syllable duration due to the number of following unstressed syllables is described as foot-level shortening (for example: Rakerd *et al.* 1987): the duration of the first of a pair of adjacent monosyllables is greater than when one or more unstressed syllables intervene.

Foot-level shortening: /pitʃ/ is longer in *peach light* than in *peach delight*.

Both constituents and prominences are thus associated with "lengthening" and with "shortening".

1.1.2 Cause and effect in speech timing

There is a large body of experimental work examining durational processes, reviewed in detail in Chapter 2; however, due to the apparently hierarchical organisation of speech, precise qualitative characterisation of the cause and the effect of timing processes can be experimentally elusive. For example, in order to ascertain that word

[1]Here and subsequently, the term "phonological length" is used where clarification is required to mean the size of some constituent measured in terms of some phonological unit, generally the syllable. This is in distinction to what might be termed "phonetic length", that is, acoustic duration. As far as possible the terms "length" and "constituent size" are used in this "phonological" sense, and the term "duration" refers to the acoustically-observed quantity; the use of terms such as "lengthening" and "shortening" to refer to contrasts in acoustic duration is unavoidable, however. The additional use of the term "phonological length" to refer to the distinction between tense and lax vowels is acknowledged, but not intended in the following discussion, unless explicitly indicated.

length influences segment duration, it is necessary to keep constant the length of other constituents which include that word. If the lengths of a number of nested constituents vary at the same time, it cannot be said for certain where the influence lies. Furthermore, the problem of qualitative characterisation applies not only to the constituent which causes the durational variation, but also to the section of speech which manifests the effect: for instance, if the effect of phrase-finality is reported in terms of lengthening of a syllable or word or other arbitrary unit, information about the distribution of the effect within or beyond that unit will be missed.

This imprecision in the characterisation of timing processes, particularly in the description of the location of the effect, would matter less if such processes were the result of diffuse changes in speech rate: for example, Klatt (1976) and Cummins (1999) suggest that utterance-final lengthening is a gradient process, reflecting "global deceleration and a reduction in articulatory effort over a number of syllables" Cummins (1999:476). Similarly, lengthening associated with the presence of a pitch accent could be interpreted as a by-product of the fundamental frequency movement, with segments lengthened in order to provide sufficient "ground" for the realisation of a perceptible pitch excursion (for example: Beckman & Edwards 1992). In this type of analysis, the distribution of lengthening amongst segments in the vicinity of an utterance boundary or pitch accent depends upon factors such as the expandability of individual segments and the relative sharpness of deceleration or fundamental frequency movement.

In contrast, the working hypothesis adopted here is that many speech timing effects above the level of syllabic organisation are directly related to linguistic structure rather than indirect effects of other processes. A systematic description of durational variation in speech is proposed here in terms of two structures defined in phonological terms: the "domain" and the "locus", the former characterising the cause and the latter describing the effect. For example, lengthening is observed at the end of intonational phrases, relative to utterance-medial position: thus, the domain of lengthening is the intonational phrase and the locus—that is, the stretch of speech which undergoes lengthening—is the rhyme of the phrase-final syllable.

The domain-and-locus approach is applied in the present research to durational variation arising from structural relationships within speech above the level of syllabic organisation. The suprasyllabic level of description is distinguished from the segmental level and the syllabic level in Section 1.2, where the treatment of speech rate, and the distinction between global and local rate variation is also discussed. The rationale for the domain-and-locus description of durational variation in speech is discussed further in Section 1.3, where a classification of durational effects in domain-and-locus terms is presented. The specific research questions addressed in this dissertation and an overview of the theoretical and empirical work undertaken are presented in Section 1.4.

1.2 Sources of durational variation

This dissertation examines the durational consequences for segments of the organisation of syllables into words and larger constituents: that is, suprasyllabic factors rather than segmental and syllabic factors. These distinctions are illustrated by classifying Klatt's (1976) list of durational factors in English speech—which he organises according to their possible origin: syntax, semantics, physiology, etc.—into segmental, syllabic and suprasyllabic factors[2]. The other factor mentioned by Klatt, speech rate, is also discussed.

1.2.1 Segmental factors

Segmental durational factors are those which are intrinsic to particular phonemes or classes of phonemes. Klatt identifies a number of such factors: the difference between long and short vowels; the greater duration of voiceless fricatives than voiced fricatives; the greater duration of bilabial stops than alveolar and velar stops. There are likely to be articulatory explanations for the origin of these distinctions: for example, short vowels are more centralised and thus their articulation takes less time; likewise, a bilabial stop creates a larger supralargyngeal cavity than alveolar or velar stops, thus more time is required for sufficient pressure to build-up behind the point of constriction to create an audible release. Although influenced by such articulatory considerations, these distinctions may in many cases be part of the phonology of the language.

Klatt also suggests another aspect of segmental duration which must be considered. He observes that, when durational factors are expressed as shortening rules, the co-application of two shortening rules is less than the sum of the two effects separately. To account for this, he proposes that a certain minimum duration is "required to execute a satisfactory articulatory gesture", thus the segment is incompressible beyond a certain amount. The mathematical details of models of timing incorporating the concept of incompressibility are not relevant to the present discussion, but there are perceptual and articulatory reasons for believing that segments will indeed exhibit some degree of incompressibility; furthermore, this is likely to vary between phonemes of different types. As discussed in Chapter 2, not all phonemes may manifest a particular durational effect to the same degree.

[2]This classification is intended to illustrate the distinctions between segmental, syllabic and subsyllabic factors rather than being an exhaustive survey of the types of processes at each level of description.

1.2.2 Syllabic factors

Syllabic durational factors are those which result from the organisation of segments into a string of syllables. One of the most perceptually salient of the syllabic factors identified by Klatt is the lengthening of vowels in stressed syllables compared with unstressed syllables; onset consonants are also longer in stressed syllables. Klatt also suggests that there is a small durational difference between vowels in syllables carrying primary and secondary lexical stress, although, as discussed below, this may only be present in phrasally-stressed words. The durational difference between stressed and unstressed syllables is clearly a linguistically-motivated distinction: greater duration is conferred as a result of the stress value of the syllable as determined by morphology rather than by articulatory considerations. Klatt also notes that the durational difference between stressed and unstressed syllables is greatest in phrase-final position, indicating an interaction between syllabic and suprasyllabic levels.

Within and between syllables, there are durational effects arising from the number and type of adjacent segments. Consonants are shorter in clusters than when they occur singly, this difference being more marked in stressed syllables: whether consonants must be tautosyllabic to manifest shortening in clusters is uncertain. Vowels are longer when followed by voiceless coda consonants than when followed by voiced coda consonants[3], a difference that Klatt claims is, like the stressed vs unstressed syllable durational difference, much greater in phrase-final syllables than in phrase-medial syllables. Indeed, he says that the duration of the preceding vowel can serve in phrase-final position as a primary perceptual cue to the distinction between voiced and voiceless consonants in the syllable coda.

1.2.3 Suprasyllabic factors

Suprasyllabic durational factors are those that arise from the linguistic structure of a syllabified string. Klatt identifies a number of effects which may be classified as suprasyllabic, and which can be placed into three categories: boundary-related lengthening, lengthening due to prominence, and shortening due to the phonological size of the constituent.

There are two effects that Klatt identifies as primary perceptual cues, on the basis of their perceptual salience and linguistic significance: phrase-final lengthening, and lengthening due to emphatic or contrastive stress. The type of phrase which occasions final lengthening is discussed in detail in Chapter 2; as Klatt observes there is a lack of consistent correspondence between syntactic phrasing and phrase-final lengthening. Additional boundary effects proposed include word-initial and word-final lengthen-

[3]This difference is less marked in some dialects, particular in Scottish English, where it may be obscured by larger dialect-specific allophonic durational effects.

ing, utterance-final lengthening extending over several syllables, and paragraph-final lengthening, whereby the final sentence of a read paragraph tends to be longer than if it were non-final. As well as the effects of emphatic or contrastive stress, which may be related to the presence of a pitch accent within the phrasally-stressed word, Klatt also suggests that entire phrases may be emphasised by lengthening: for example, at a dramatic point in the relation of a story, a speaker may reduce speech rate across a whole phrase or sentence.

There are two constituents for which Klatt proposes an inverse relation between constituent size and subconstituent duration, the word and the utterance. Initial stressed syllables are shorter in polysyllabic words: for example, /spid/ is shorter in *speedy* than in *speed*, and shorter still in *speediness*. This can either be interpreted as word-final lengthening or as a result of what he calls: "a desire on the part of a talker to maintain a nearly constant duration for each word, independent of its phonetic composition" Klatt (1976:1214). Klatt appears not to favour the latter explanation, continuing: "it is hard to believe that constant word duration is a very important constraint in English, since duration has already been implicated as functioning to differentiate inherently long from short vowels, voiced from voiceless fricatives, phrase-final from non-final syllables, etc." With regard to the effect of utterance length, Klatt claims that there is a general trend for speech rate to become faster in longer sentences, thus shortening individual segments within the sentence.

Previous experimental studies of suprasyllabic durational processes related to boundaries, constituent length and prominence are reviewed in Chapter 2, including the durational effects associated with variation in the size of hypothesised prominence-delimited constituents. Experimental investigations of such processes at the word-level and the utterance-level are presented in Chapter 3 and Chapter 4.

1.2.4 Speech rate

In addition to the durational factors classified here as segmental, syllabic and suprasyllabic, Klatt also observes that changes in overall speech rate, which may be affected by a number of non-linguistic factors, "exert a complex influence on the durational patterns of a sentence. When speakers slow down, a good fraction of the extra duration goes into pauses (Goldman-Eisler 1968). On the other hand, increases in speaking rate are accompanied by phonological and phonetic simplifications as well as differential shortening of vowels and consonants" Klatt (1976:1210).

It is worth considering the concept of speech rate further, as it is an unstated factor in much of the experimental work reviewed in Chapter 2. The term "speech rate" may be used in two senses. Most commonly, it is used to mean *global* speech rate: the rate of production of speech segments over a stretch of speech of arbitrary length as influenced by factors such as physiology, dialect and emotional arousal. Occasionally it refers to *local* speech rate: the influence of linguistic structure, such as the distribution of constituents and prominences, on speech timing (it is used in this sense by, for example, Dankovičová 1997). Clearly, the latter usage relates to the linguistically-motivated durational patterns that are the present objects of study, so to prevent confusion, the term "speech rate" is only used here in the global sense.

In the discussion of previous research and in the experiments presented here, speech rate is taken to be normal except where otherwise stated, where "normal" implies a self-selected rate by the speaker, possibly under explicit instruction to speak normally. In a small study, Gaitenby (1965) compares segmental duration between subjects talking at their own "normal" rate and finds that most segments are expanded or compressed in proportion to the length of the whole utterance; voiced consonants, however, are very close in absolute duration between talkers. She also finds that the widest variations in length are observed on stressed syllables.

It is not certain whether patterns observed between speakers at their individual natural speaking rates also apply to the different rates of a single speaker. In particular, it is possible that the results obtained at normal speech rate with respect to suprasegmental effects may not generalise across the range of possible rates.

Global speech rate variation does have two well-attested consequences. Firstly, as described by Gee & Grosjean (1983), at faster speech rates utterances tend to be broken up into fewer and larger phrases, a process discussed further in Chapter 2. Secondly, researchers who look at the effect of global speech rate variation find that some of the durational effects evident at normal speech rates may be more marked at slower rates and may be attenuated or disappear at faster rates. Rakerd *et al.* (1987), for example, find significant interactions between speech rate and syntactic structure, and between speech rate and foot structure: preboundary lengthening is attenuated in fast speech, as is the shortening of a stressed syllable followed by an unstressed syllable compared with when followed by another stressed syllable. Inspection of their graphical data suggests that phrase-final lengthening is preserved to a reasonable extent at higher speech rate, whereas the difference in stressed syllable duration between monosyllabic and disyllabic feet is greatly reduced in fast speech. The evidence that Beckman & Edwards (1990) find for a phrase-medial word-final lengthening effect is stronger at slower speech rates. In contrast, Port (1981) finds that the rhyme of a pitch-accented syllable shows less shortening in response to an increase in global speech rate than does the overall sentence: at a faster rate the accented syllable rhyme is 76% of its normal duration whereas the sentence as a whole shortens to 72% of its normal duration. Port reports that this result is in agreement with an earlier study by Peterson & Lehiste (1960).

It appears that some durational effects are more robust under increases in speech rate than others. It is possible that rephrasing as observed by Gee & Grosjean (1983) could underlie some of these observations: results that appear to reflect a diminution of final lengthening may actually be a result of some boundaries not being realised at all at faster rates. This explanation could not account for the attenuation of shortening in disyllabic feet at fast rate, however, unless the second of a pair of adjacent stressed syllables became unstressed—and thus refooted—at fast rate. A speculative interpretation is that the durational effects associated with intonational events are more robust in fast speech than other timing processes: thus, intonational-phrase-final lengthening and accentual lengthening are preserved whereas final lengthening of minor constituents and the effects of stress adjacency are attenuated.

1.3 A descriptive framework

1.3.1 The concept of the domain

The concept of the domain is a familiar one in prosodic phonology—the study of connected speech processes—as presented, for example, by Selkirk (1986) and Nespor & Vogel (1986). Within the hierarchies of prosodic constituents proposed by these and other researchers, each prosodic constituent is proposed to have a certain relation to syntax, which is in general more straightforward for lower-level constituents: at higher levels, a number of arrangements of prosodic constituents are available for the same syntactic structure. The evidence for constituents is provided by segmental processes occurring within and between words, such as flapping in American English and r-insertion in non-rhotic dialectics of British English. There are three types of rules proposed by Nespor & Vogel to explain the occurrence of segmental variations in connected speech: domain-span rules, domain-juncture rules and domain-limit rules. Domain-span rules describe processes which may occur anywhere within a particular constituent but not at its edge; domain-limit and domain-juncture rules occur only at the edges of constituents, with the latter further requiring particular segmental conditions on the other side of the constituent boundary.

These prosodic constituents and segmental processes are described in more detail in Chapter 2; for the purposes of the present discussion there are two important points to be made about the distinction between prosodic constituents and the concept of the domain as used here.

Firstly, the constituents described by researchers such as Selkirk and Nespor & Vogel are not necessarily co-extensive with the domains of durational effect, although they may prove to be so. Nespor & Vogel, for example, claim that some of the constituents that they identify as domains of segmental processes may also be the domains of durational variations, but do not present empirical evidence to support these claims. Furthermore, although it seems clear that syntax is not wholly reliable in predicting the occurrence of durational effects such as final lengthening, as discussed in Chapter 2, it remains to be seen whether the prosodic constituents evinced by segmental processes are in all cases isomorphic with the domains of durational processes. There is indeed, no intrinsic reason why a domain need be a constituent at all, although the fact that a constituent serves as the domain of some effect in speech may be used as evidence of its existence. An illustration of a domain which is not co-extensive with a constituent might be taken from intonational phonology: Faure *et al.* (1980), for example, suggest that the phonetic realisation of stress contrasts is different following the nuclear accent in an intonational phrase . One might thus assert that the stretch of speech between the nuclear accent and the intonational phrase boundary is the "domain of stress neutralisation". Such a statement may be theoretically sufficient without the requirement that this domain be identified as a constituent of some kind.

The second important distinction to be made between the concept of the domain as exemplified in the prosodic phonology of Nespor & Vogel (1986) and its use here is in the type of processes manifested within the domain. As described above, durational variation due to the position of measured segments within some constituent tend to be described as lengthening, whereas the effect constituent length tends to be described as shortening. An alternative to the subjective standpoint inherent in a lengthening and shortening description is to describe the relationship between the domain and the measured segments as either a domain-edge process or a domain-span process. These terms and their relation to the domain-limit, domain-juncture and domain-span processes of prosodic phonology are described further in Section 1.3.3.

1.3.2 The concept of the locus

Many early studies of durational effects such as final lengthening report the results in terms of the amount of lengthening of some unit, typically the word or syllable, in the experimental condition compared with some control condition, for example, constituent-final syllables as compared with constituent-medial syllables. As discussed above, this could in some cases reflect an implicit assumption that durational variation is not related to linguistic structure, but is instead a consequence of the articulatory process: for example, final lengthening may be seen as a process of deceleration at the end of a speech unit and thus not targeted at any particular segments of speech. More recent evidence suggests that the distribution of durational variation is, in some cases at least, determined by linguistic structure: thus, for example, phrase-final lengthening is manifest on the final syllable rhyme (for example: Wightman *et al.* 1992) and accentual lengthening is distributed throughout the word (Sluijter

1995; Turk & White 1999) and attenuated by word boundaries.

If durational variation is related to linguistic structure, then it becomes necessary to know what the constituents of this structure are. The concept of the domain represents one part of the process: segments may be lengthened when they stand in a particular relation to a particular domain, such as the constituents of syntactic or prosodic structure. An account of linguistically-derived timing also requires a description of where the effect associated with some domain is manifest: this is the locus of durational variation.

One reason for the proposal here of the term locus is the resolution of an ambiguity present in some recent investigations of the distribution of durational variation associated with certain processes. For example, studies of final lengthening, such as Cambier-Langeveld (2000) for Dutch, report that it primarily affects the rhyme of the final syllable before the constituent boundary, citing this as evidence that the "domain" of final lengthening is the syllable rhyme. In customary recent usage, however, the term "domain" is applied to the constituent causing the lengthening effect (for example, Fougeron & Keating 1997). As the domain of final lengthening cannot be at the same time the intonational phrase and the rhyme of the final syllable of that phrase, the additional term "locus" is proposed to describe the phonologically-defined section of speech which is directly subject to a durational process, whether characterised as lengthening or shortening.

A further reason for emphasising the importance of the locus in the description of durational variation is that its identification may serve to clarify the nature of the effect. There are many factors which may influence the duration of a speech segment, outlined in Section 1.2, and as Port (1981:262) observes: "this large and impressive literature [on the durational realisation of linguistic features] has not resulted in an integrated view of the role of linguistically motivated timing in speech perception or speech production [. . .] Not only may several different phonological features affect the duration of a single phonetic interval, such as a vowel duration, but many other factors that are marginally phonological or nonlinguistic may also influence its duration [such as word length, syntactic boundary location, speech rate . . .] How can timing be an effective source of phonological information when it is subject to such a variety of overlapping distortions?" The question of whether morphosyntactic facts such as word length and boundary location are integrated into phonological structure may be left aside for the moment; the important point here is that, given the diversity of influences, speech segment duration may appear an intrinsically ambiguous cue to linguistic structure.

Port observes—like Klatt mentioned above—that limits on the compressibility of speech segments mean that a combination of "shortening factors" have less effect than the factors individually, he then concludes that: "despite massive evidence that relatively fine-grained articulatory timing comprises an ubiquitous aspect of the acoustic specification of English words, there is, nevertheless, also evidence that in normally complex speech situations very little of this potential information may actually be available for a listener." Of course, listeners do not experience segmental duration in isolation. As well as the variations in acoustic energy across time and frequency which allow determination of segmental identity, there are variations in the fundamental frequency and amplitude of the speech signal which are known to correlate with durational variation. Thus, greater phrasal prominence is conferred by a combination of duration, fundamental frequency and amplitude variation. Similarly, major boundaries are marked by both duration and fundamental frequency.

Experimental studies show, however, that durational variation alone may serve to indicate the presence of a boundary to the listener in some cases (for example: Price *et al.* 1991; White 1993). Furthermore, there must be sufficient information about linguistically-motivated patterns of speech timing accessible to the listener or the systematic variations which do not arise as a direct consequence of articulation could not be learned. Indeed, it is not clear why such systematic durational variation would be present in speech if it were not accessible to listeners.

Klatt (1976:1220) expresses the apparently intrinsic ambiguity of speech timing thus:

> A paradox seems to exist which might prevent listeners from making effective use of durational cues in forming hypotheses about sentence structure. In order, e.g., to perceive the lengthening at a phrase boundary as a cue to the end of a constituent, it seems necessary to know the identity of the lengthened segments. Durations are lengthened relative to the inherent durations for the segments in question. But a listener can't know the inherent duration until he has identified the vowel, so we have a version of "the chicken or the egg" paradox[4].

At the segmental level, however, the intrinsic durational properties of different phonemes are in most cases not a primary cue to segmental identity in English: like microprosodic variations in fundamental frequency, they may be regarded as a predictable perturbation which can be factored out by listeners when attending to suprasyllabic variation. For example, although different vowels in English have different durations in citation form, differences in vowel quality are more important for identification. Similarly, voiced obstruents tend to be shorter than voiceless obstruents, but duration itself is not a significant cue to voicing in most cases. Likewise, at the syllabic level, much of the variation is likely to be predictable if segmental identities are known.

[4]Klatt's paradox has been resolved in modern approaches to speech recognition where multiple hypotheses are processed in parallel. The quotation serves to illustrate potential problems for listeners, however these may be solved in practice.

In some contexts, certain features may be neutralised and duration may play a more important role: for example, Klatt suggests that where voiced fricatives are devoiced syllable-finally, duration may be the primary cue to voicing. This is not, however, simply a matter of judging absolute duration, but of comparing the relative durations of adjacent segments. For example, the interpretation of the word /meɪs/ may depend on both vowel duration and consonant duration: in "mace", in many dialects of English, the vowel will be relatively short and the coda relatively long; in "maize" the vowel will be relatively long and the coda relatively short. Thus, the distribution of durational variation may sometimes play a role in identifying of segments in connected speech.

At the suprasyllabic level, it is argued in this dissertation that the distribution of variation is of primary importance in the interpretation of speech timing as a cue to structure. Klatt identifies duration as a primary cue in two structural distinctions: phrase-finality vs non-phrase-finality, and the presence or absence of emphatic stress: both phrase-final and emphatically-stressed syllables show substantial lengthening. As observed by Shattuck-Hufnagel & Turk (1996:238): "Apparently, listeners are able to distinguish prominence-related from boundary-related duration lengthening, perhaps because the two factors affect different parts of the syllable (Edwards *et al.* 1991)." As discussed above, the lengthening associated with phrase boundaries appears to affect primarily the syllable rhyme (for example: Wightman *et al.* 1992); as shown by Turk & Sawusch (1997) and Turk & White (1999), the lengthening associated with emphatic accent also affects the accented syllable onset and adjacent syllables within the same word, and may extend further.

Clearly, fundamental frequency variation will also be associated with final lengthening, at least at the end of intonational phrases, and with emphatic stress. Fundamental frequency variation is not always unambiguous, however: as discussed in Chapter 2, pitch accents are not always realised with a large fundamental frequency movement, particularly towards the end of a phrase. Furthermore, some structurally-related durational variation, such as word-initial lengthening and polysyllabic shortening, is apparently not associated with fundamental frequency variation. It is hypothesised here that the ambiguity in speech timing as a cue to linguistic structure is greatly reduced if the locus of durational variation is considered in each case.

1.3.3 A domain-and-locus description

The concepts of the domain and the locus provide a descriptive framework for speech timing processes which allows established effects to be treated more systematically. In particular, structurally-determined lengthening may, in many cases, be seen as a domain-edge process and shortening may often be seen as a domain-span process.

As described now, this classification implies processes with different characteristics, particularly with regard to the locus.

Domain-span processes

The durational effect of constituent length is generally held to be manifest in an inverse relationship between the phonological length of some constituent and the duration of some subconstituent: in other words, there is an inverse relationship between the length of the domain and the duration of the locus. As indicated in Section 1.1, a number of units have been held to be domains of such processes, both syntactic/prosodic constituents such as the word and the sentence/utterance[5], and prominence-delimited constituents, such as the cross-word foot: evidence for such processes is described in detail in Chapters 2 and 3.

As outlined above, this effect is generally described as "shortening"; the alternative adopted here is to adapt the terminology of prosodic phonology (for example: (Nespor & Vogel 1986)), and describe it as a "domain-span" process. This expresses the fact that the duration of the locus is related to the domain's length and not to the position of the locus within the domain. This usage of "domain-span" is, however, distinct from its usage by Nespor & Vogel: in prosodic phonology, a domain-span applies within a domain, but not at its edge. As used here with regard to durational processes, the effect may in principle apply to the locus *anywhere* within a domain; the crucial factor is not the position of the locus but the length of the domain.

The primary characteristic of a domain-span process is the inverse relationship between domain length and subconstituent duration. This may be depicted as a compression of the measured subconstituent proportional to the amount of additional material within the domain. Thus, the term "domain-span compression" is sometimes used to clarify the nature of the effect: the term is pleonastic, however, as all domain-span processes are hypothesised to manifest such compression.

There are a number of candidates for the subconstituents that may comprise the locus of a domain-span process. One possibility is that the locus is some privileged unit, such as the head of the constituent that comprises the domain. In a polysyllabic word, for example, the head would be the syllable carrying the primary lexical stress, or the nucleus of that syllable. In a larger prosodic constituent, for example, the intonational phrase, the head would be the syllable or word carrying the main stress accent (nuclear accent). In prominence-delimited constituents, the heads are part of the definition; for example, a foot of almost any type comprises a strong syllable followed by a number of weak syllables.

[5]Differences in syntactic and prosodic definitions of the word are discussed in Chapter 2, as is the prosodic concept of the utterance.

A head-based account of domain-span shortening is suggested by studies which report the effect of constituent length on the duration of the constituent head, but this is frequently the *only* unit whose duration is reported, sometimes for practical reasons: for example, in a comparison of monosyllabic, disyllabic and trisyllabic words, the primary stressed syllable is the only unit measurable in all conditions. In such cases, it is impossible to know if the locus extends beyond the constituent head.

The alternative hypothesis in domain-span shortening is that the locus is co-extensive with the domain: thus, all segments within a domain are shortened if the length of the domain increases. This account is suggestive of accounts of speech timing, such as the theories of isochrony discussed in Chapter 2, which suggest a unit with a global timing function.

The distinction between these two characterisations of the locus of domain-span processes may in practice be a matter of degree. Within the word, for example, the greatest polysyllabic shortening effect may be seen within the primary stressed syllable, either because this is the phonological head or because it is likely to be longer and thus more compressible than other syllables. Likewise, the syllable nucleus may show the greatest effect, either as the head of the syllable, or, because of the intrinsic durational properties of vowels and their relative expandability compared to other phone classes such as stops, as the longest part of the syllable.

The experimental investigation of domain-span effects is complicated by the fact that an increase in the size of some lower-level constituent such as the word also increases the size of higher-level constituents which dominate the word, such as syntactic or prosodic phrases. This potential ambiguity of interpretation may resolvable if the locus of the effect is different between two domains, as discussed further in relation to domain-edge processes.

Domain-edge processes

Durational processes associated with constituent boundaries are typically characterised as "lengthening", where segments at the boundary have greater duration than similar segments placed constituent-medially: for example, word-initial lengthening or phrase-final lengthening. In the domain-and-locus description, such effects are called "domain-edge" processes.

Nespor & Vogel (1986) discuss two types of segmental process at constituent boundaries in prosodic phonology: domain-limit processes and domain-juncture processes. The former occur at the initial or final edges of constituents, regardless of context; the latter additionally require a certain segmental context on the other side of the boundary. Durational processes at constituent boundaries appear to have more in common with domain-limit segmental processes, as position in the constituent is apparently the determining factor rather than any preceding or following context. The term domain-edge is preferred to domain-limit here, however, and subsumes both domain-initial and domain-final processes.

The locus of domain-edge processes is, of course, adjacent to the domain edge itself. Studies discussed in Chapter 2 suggest that the locus may be structurally determined rather than simply being a matter of proximity. In particular, the position of the final stressed syllable in a phrase or utterance may influence the extent of the locus of final lengthening: where the stressed syllable is in absolute-final position, the stressed syllable rhyme appears to manifest the lengthening effect; where there is an unstressed syllable between the final stressed syllable and the boundary, the unstressed syllable shows the greater lengthening, but the stressed syllable may also show some lengthening (for example: Nakatani *et al.* 1981; Turk 1999, Cambier-Langeveld 2000). This suggests that domain-edge lengthening is progressive within the locus: that is, the subconstituents of the locus nearer to the boundary show a greater effect. As discussed in Chapter 2, within the rhyme of an absolute-phrase-final syllable, there may be more lengthening on the coda than on the nucleus.

The terms "lengthening" and "shortening" suggest that domain-edge and domain-span processes are in conflict, but this is not necessarily the case, as shown by considering a lexical word in citation form: that is, spoken in isolation. For example, the syllable /dɛt/ is predicted to shorten in the sequence

debt . . . cadet . . . cadetship

by both domain-edge and domain-span processes. The domain-span hypothesis predicts that stressed syllable will have the longest duration in the monosyllable because the domain is phonologically shortest; in the phonologically-longest domain, the trisyllable, the stressed syllable will have the shortest duration. The domain-edge hypothesis predicts that stressed syllable have the longest duration in the monosyllable because it is both domain-initial and domain-final; it will have the shortest duration in the trisyllable, because it is neither domain-initial nor domain-final. Thus, determining which process is responsible for observed effects is not always straightforward. Furthermore, in this case, the domain of either type of process could be the word, or a phrase of some kind, or the utterance, as these constituents are co-extensive in citation form.

A similar dichotomy of interpretation exists for durational processes within prominence-delimited constituents, as pointed out by Fowler (1990:201): "Stressed syllables in monosyllabic stress feet that are identified as unshortened in a stress-timed account can just as well be identified as lengthened; so the absence of stress-timed shortening can instead be described as a lengthening of a stressed syllable at the right edge of a foot. As such it may count as a sort of preboundary lengthening." The use of the terms "lengthening" and "shortening" may be simply a matter of the viewpoint from which the effects are considered; however, although domain-edge and domain-span processes may make the same predictions about durational variation at the syllable level, they are theoretically distinguishable if patterns of durational variation in subsyllabic constituents are taken into account. This is the approach taken in the experimental work reported in Chapter 4 and in other recent work (Turk & Shattuck-Hufnagel 2000).

A domain-edge process is suggested if the locus is aligned with the boundary and the magnitude of the effect within the locus diminishes with distance from the boundary; in contrast, in a domain-span process, the variation is not expected to be progressive within the locus and may be greatest on the phonological head of the constituent.

As noted above, there is another potential ambiguity of interpretation of observed domain-edge effects, particularly effects in constituents within nested hierarchies, such as syntactic or prosodic constituents. The association of a domain-edge effect with a particular constituent is made more difficult by the fact that the edge of that constituent also corresponds with the edge of subordinate constituents: in Nespor & Vogel (1986)'s prosodic phonology, for example, the final boundary of an intonational phrase is necessarily the final boundary of a phonological phrase and all other subordinate constituents. This is a consequence of the "strict layer rule", described in Chapter 2. Thus, unless appropriate experimental controls are exercised, it cannot be guaranteed that the effect is associated with the intonational phrase rather than with a subordinate constituent. Similarly, the presence of higher-level boundaries should be taken into account when examining domain-edge processes in lower-level constituents. As when distinguishing domain-edge and domain-span processes, consideration of the locus may be helpful here: it is possible that higher-level constituents may have a locus which extends further from the domain-edge, as well as manifesting a greater magnitude of lengthening than lower-level constituents. Evidence relating to possible patterns of hierarchical lengthening is considered in later chapters.

Prominence

The arrangement of prominences, such as stressed syllables, may have durational consequences, because as domain-span processes are associated with prominence-delimited units, such as cross-word feet, as well as with syntactic or prosodic constituents. In addition, the direct lengthening effects of prominence—stressed syllables are longer than unstressed syllables; accented syllables are longer still—should be accommodated within a complete descriptive framework of speech timing. These observations do not appear to be classifiable as domain-edge or domain-span effects, however. Indeed, although the locus of lengthening may be observable, it is less certain how the domain should be characterised.

The lengthening associated with lexical stress appears to have a syllable locus: all parts of the syllable are likely to be longer in a stressed syllable than an unstressed syllable, although as noted above, Klatt (1976) only alludes to lengthening in the stressed syllable onset and coda. If lengthening were only associated with primary stress, then it could be said that the lexical word is the domain of stress-related lengthening; within polysyllabic words, however, there may be more than one syllable which has a lexical stress, with the difference between the two only becoming apparent when the word carries a pitch accent. This suggests a domain with more than one locus, which whilst not inconceivable, is less theoretically satisfactory. As discussed in Chapter 2, the within-word foot may be defined as beginning with a stressed syllable and continuing to the next stressed syllable or to a word boundary, whichever is the sooner. Such a unit might be said to be the domain of stress-related lengthening, which could then be viewed as a domain-edge effect, specifically a domain-initial effect with a syllable locus. There is, however, a circularity in the definition of such a unit, unless independent evidence for its existence can be found. Furthermore, stress-related lengthening seen as a domain-edge effect would lack the progressive quality of lengthening associated with other domain-edge effects.

With regard to the phrasal level of prominence, studies such as Turk & Sawusch (1997) and Turk & White (1999) refer to the domain of accentual lengthening to mean the stretch of speech which undergoes lengthening. It is clearly more consistent within the current framework to use the term locus for this purpose, although, as for lexical stress, it is not clear how the domain of the process ought to be characterised. Both the experimental evidence regarding the locus of accentual lengthening and the question of its domain are discussed in more detail in Chapter 2. Also raised in Chapter 2 is the possibility that many previous studies of durational effects in speech—including domain-edge and domain-span effects—use materials in which the words being measured are likely to contain a pitch accent. Thus, the findings of such studies do not necessarily apply to words which do not carry an accent.

The fact that both lexical and phrasal stress are associated with greater duration may be seen as a reason for treating degrees of prominence as variations within a single dimension. This is the approach often taken in phonological accounts of prominence: Hayes (1983) and Selkirk (1984), for example, represent different levels of stress as columns of different heights on a metrical grid, in an attempt to account for the distribution of lexical and phrasal stresses. The approach taken here, however, is to treat lexical and phrasal stress as largely distinct from the point of view of speech timing. This separation reflects the identification by Bolinger (1981) of two different types of prominence relations: full vs reduced vowels (in effect, the presence or absence of lexical stress) and accented full vowels vs unaccented full vowels (in effect, the presence or absence of phrasal stress). There are phonetic and theoretical reasons to favour this approach in the analysis of durational effects.

In the context of a discussion about the distribution of stress and the phenomenon of stress shift, Fourakis & Monahan (1988:284) say that metrical phonology is "a theoretical construct based on impressionistic transcriptions and very little acoustic research has been carried out to test its predictions." Certainly, the acoustic correlates of the different levels of prominence represented in the grid do not support a unified approach: differences in the acoustic marking of different degrees of stress are qualitative as much as quantitative. In particular, phrasal stress is indicated primarily by pitch accent and its durational consequences may extend beyond the stressed syllable; in contrast, the primary cues to lexical stress are vowel quality and duration, with the most significant durational effects on the vowel nucleus of the stressed syllable. Beckman & Edwards (1994) examine the articulatory differences underlying the accented vs unaccented distinction and the stressed vs unstressed distinction, finding that the two prominence contrasts are associated with variation in different articulatory parameters. In the terms of the task-dynamic model they apply to their data, accented syllables are articulated with greater gestural magnitude than unaccented syllables, whereas stressed syllables are articulated with less gestural stiffness than unstressed syllables. This effectively means that increased duration associated with phrasal stress is, at least in part, a reflection of the more extreme displacement of articulators compared with unaccented stressed syllables, whereas the increased duration associated with lexical stress reflects slower movement of articulators towards their target[6].

There are also levels of prominence distinguished in grid-based approaches to metrical phonology which do not have specific durational consequences. Sluijter & van Heuven (1995) suggest that the distinction between primary and secondary lexical stress is not realised durationally, but by variation in high frequency energy levels. Indeed, Bolinger's two-tier approach to prominence suggests that "primary and secondary word stress differ not in the degree or type of articulatory or acoustic prominence, but in the instructions they provide for the placement of pitch accent" (Shattuck-Hufnagel & Turk 1996:22); likewise Ladefoged (1993) suggests that the distinction is only phonetically realised when one of the stresses carries a pitch accent. Nuclear pitch accents and non-nuclear pitch accents are also placed on different levels in metrical grid representations: Silverman & Pierrehumbert (1990), while not directly addressing the durational consequences of this distinction, make the claim based on their results regarding peak alignment that "the distinction exists in prosodic organisation, but [we] do not find evidence for it in the inventory of English pitch accents or the

[6]The vocalic sections of unstressed syllables will also have a less extreme articulation than stressed syllables in most cases, by the nature of the centralised reduced vowel.

phonetic rules for pronouncing them" (Silverman & Pierrehumbert 1990:105).

The theoretical reasons for treating lexical stress and phrasal stress as distinct relate to the observation of Beckman & Edwards (1990:165) that "accent patterns belong to the intonation and thus are part of the post-lexical phrasal phonology, whereas stress patterns are largely specified in the lexicon." The distribution of lexical stresses is predictable to a great extent from the text of an utterance and stands in arbitrary relation to meaning[7]. In contrast, there are a number of possible accentual patterns for any given utterance, and the accentual choice made by a speaker provides information regarding the pragmatic interpretation of the utterance (Ladd 1996). Whether the unified approach to prominence taken in metrical phonology has utility in other areas is beyond the scope of this discussion; as the accounts of the durational consequences of lexical and phrasal stress in Chapter 2 indicate, for timing purposes the two-tier approach advocated by Bolinger appears more appropriate.

A further, terminological distinction should be made at this point: the term "rhythm" is used widely in discussion of speech timing, sometimes apparently to encompass all systematic durational variation, sometimes more narrowly to refer to the durational effects of all levels of prominence, and sometimes to refer only to the distribution of lexical stresses. The latter approach is taken here: rhythm means the pattern of alternation of strong syllables (carrying primary or secondary lexical stress) and weak (unstressed) syllables. The term "pitch accent" as used here relates solely to the occurrence of phrasal stress, which is intonationally marked. To avoid any confusion, the term "phrasal stress" is generally avoided henceforth.

1.4 Research questions

This dissertation examines a number of durational effects at the suprasyllabic level and attempts to characterise the domain and the locus in each case. Constituents of syntactic and prosodic hierarchies and prominence-delimited units are considered as potential domains of edge and span processes in the review of research presented in Chapter 2. The new experimental work described in Chapters 3 and 4 examines domain-edge and domain-span processes at two structural levels: the word and the utterance. Specific questions raised in the research review and in the experimental work include:

- Do all constituents serve as domains of edge and span processes or are such processes specific to particular levels of some putative hierarchy?

[7] There are, of course, a small number of minimal pairs such as ˈ*accent* and *acˈcent* in which the noun carries primary stress on the first syllable and the verb carries primary stress on the second syllable. In the great majority of cases, however, the placement of stress within a word is of no syntactic, semantic or pragmatic significance.

- If domain-edge or domain-span processes are associated with more than one constituent, is the locus the same in each case?

The durational effects of pitch accent are also considered: firstly, the locus of accentual lengthening is examined; secondly, given the likely presence of pitch accent in many previous studies of durational effects, the interaction between accentual lengthening and domain-edge and domain-span processes is examined.

1.4.1 The structure of the dissertation

These research questions are applied in Chapter 2 to previous experimental studies. The work discussed—like the experimental work presented in later chapters—is that which arises from what may be characterised as the "laboratory phonology" approach to the study of speech timing: particular hypotheses are tested about specific durational effects by examining recordings of subjects reading prepared materials, wherein certain factors are controlled whilst others are experimentally varied. Evidence for domain-edge and domain-span processes from such work is discussed with regard to syntactic constituents and prosodic constituents; evidence for the existence of such processes in prominence-delimited units is also examined. Previous studies of the durational effects of pitch accent are reviewed.

Two particular aspects of this previous experimental work are examined further in Experiment 1, presented in Chapter 3. This experiment is designed to determine: firstly, whether the word is a domain of durational processes or whether previous experimental findings regarding the word are a result of domain-edge or domain-span processes at other levels of linguistic structure; secondly, whether the presence or absence of pitch accent has an influence on observed effects at the word level. The findings of Experiment 1 indicate that the word is a domain of durational variation and that some of this variation may be affected by the presence of pitch accent.

Experiment 2, presented in Chapter 4, is designed to determine the type of processes operating at the word level, whether domain-edge or domain-span processes, or some combination; the existence of domain-edge and domain-span processes at the utterance level is also tested. Examination of the loci of variation indicates that there are domain-edge processes, but not domain-span processes, at the word level and the utterance level; there is some evidence of a domain-span effect within a sub-word constituent, but this may also be interpreted as a word-final effect. The distribution of accentual lengthening within monosyllables, disyllables and trisyllables is found to be responsible for apparent word-span compression. Domain-edge processes do not appear to interact with the durational effects of pitch accent.

The conclusions drawn from Experiments 1 and 2 are summarised in Chapter 5, and compared with other experimental findings regarding domain-edge and domain-span processes. A model of English suprasyllabic speech timing is outlined in Chapter 6, based upon localised lengthening effects at domain edges and within pitch-accented words.

Conventions used in this dissertation

The terms "polysyllable" and "polysyllabic" are used here to refer to a word of two or more syllables, typically in the contrast "monosyllables vs polysyllables", rather than in the standard sense of a word of many, or more than three, syllables.

Where results are given as percentage of lengthening or shortening, this is calculated by a standard formula. For lengthening, the difference between the longer and shorter duration is expressed as a proportion of the duration in the shorter context. For shortening, the difference between the longer and shorter duration is expressed as a proportion of the duration in the longer context.

Where comparisons of segmental duration between two contexts are made in the text, underlining is used to indicate the measured segments where these are not otherwise identified, thus: "a stressed syllable is longer in an accented monosyllable than in an accented disyllable: for example, *thank fulfil* vs *thankful Phil*."

Chapter 2

Suprasyllabic factors in durational variation

2.1 Introduction

Suprasyllabic timing processes associated with constituents and prominences are described here in terms of the domain and locus framework outlined in Chapter 1. Each process suggested by experimental findings is classified as a domain-edge or a domain-span process; cases where such classification is problematic are discussed. The most appropriate characterisation of the domain of each process is considered, whether this might be in terms of constituents—syntactic or prosodic—or in terms of units defined by lexical stresses or phrasal stresses. An attempt is also made to characterise the locus in each case, and it is considered whether similar processes—particularly domain-initial and domain-final processes—with different domains have similar loci. In some cases the evidence currently available is insufficient to resolve these issues and the experimental work presented in Chapters 3 and 4 is intended to shed further light on them: the particular questions to be addressed experimentally are indicated in such cases.

The influence of syntactic constituents upon speech timing is considered in Section 2.2, and reasons are discussed for the occasional failure of syntactic constituency adequately to predict the occurrence of domain-edge durational effects. A theory of prosodic constituency is presented in Section 2.3, which is motivated in part by the shortcomings of syntactic descriptions of domain-edge effects. The timing effects of prominence are considered under two distinct categories, for reasons outlined in Chapter 1: the effect of the distribution of lexical stress is discussed in Section 2.4, and the effect of pitch accent is discussed in Section 2.5.

In some cases, the designs of previous experiments suggest alternative interpretations from those advanced by the authors. Where questions of experimental design are directly relevant to the experiments presented in later chapters, previous results are reviewed here but critical analysis of experimental design is presented in the introduction to the relevant new experiment. Thus, studies of polysyllabic shortening are evaluated in Chapter 3 and studies of hierarchical domain-edge effects are evaluated in Chapter 4. Other findings are critically evaluated in the present chapter. In particular, the timing consequences of the distribution of lexical stress—although relevant to the general discussion and to experimental design—are not examined experimentally in this dissertation, and so previous work is analysed in detail here.

As the subject of the dissertation is speech timing in English, experiments relating to other languages are not reported as a matter of course, particularly where results between English and other languages are in agreement. Such experiments are discussed where the results pertain to questions that have not been satisfactorily resolved in English. How far the durational effects of constituents and prominences may be universal is beyond the scope of this dissertation, but where such effects may be seen to be phonologically determined, rather than a by-product of articulation, it is likely that cross-linguistic differences will be observed.

2.2 Syntactic constituents

Durational and intonational variation can influence listeners' judgements about syntactic structure, principally through the delimitation of phrases and clauses (for example: Scott 1982; Beach 1991; Price *et al.* 1991; Marslen-Wilson *et al.* 1992; White 1993). Suprasegmental information may serve to indicate the location of a boundary where syntactic structure is ambiguous. White (1993) plays subjects sentence fragments ending in verbs of ambiguous transitivity, for example:

- *When the tractor pulls ...*

and asks them to indicate which of a pair of sentences represents the most likely complete sentence, for example:

1. *When the tractor pulls, the rope is stretched tight.*
2. *When the tractor pulls the rope, the tree crashes down.*

A number of versions of each sentence fragment are prepared by resynthesising real speech, and the only differences between the versions are the duration of the final verb and the fundamental frequency contour on the voiced section of the verb. Subjects are significantly more likely to select Sentence 1, where the verb is clause-final, when the duration of the verb is greater and when the fundamental frequency contour is a sharp fall-rise rather than a shallow fall.

These experimental manipulations correspond to suprasegmental variations that have been observed at the end of syntactic phrases and clauses in natural speech by, for example, Cooper & Paccia-Cooper (1980) for duration, and Cooper & Sorenson (1981) for intonation. Speakers are not wholly consistent in the production of such boundary cues, however. As Klatt (1976) observes, boundaries such as that between a noun phrase and a verb phrase may or may not be accompanied by lengthening of the pre-boundary syllable: extra-syntactic factors such as the length of the noun phrase appear to influence the occurrence of suprasegmental boundary cues. Shattuck-Hufnagel & Turk (1996) observe that the same sentence may be realised on different occasions as a single prosodically-delimited phrase or as multiple phrases. Speech rate is a contributory factor: more prosodic phrases are realised at slower rates. In some cases syntactic constituent structure is actually violated: suprasegmental boundary cues may indicate that the main boundary in a NP-VP sentence comes after the verb but before the rest of the verb phrase, apparently influenced by a tendency for prosodic phrases to divide utterances symmetrically. Similarly, function words may be prosodically separated from the content words to which they are most closely bound syntactically. An example of the tendencies to ignore the major syntactic boundary in a sentence and to divide an utterance into approximately equal prosodic phrases is provided by the following television announcement[1] , where "|" indicates a phrase boundary):

Newsnight is over |on BBC2.

The major syntactic phrase boundary, which follows the first word, is not realised prosodically, but the compound prepositional phrase *over on BBC2* is prosodically divided, so as to split the whole utterance into two five-syllable phrases.

Segmental phonological processes such as assimilations or elisions have a similarly indirect relationship to syntax, prompting researchers such as Selkirk (1986) and Nespor & Vogel (1986) to postulate hierarchies of prosodic constituents, influenced by syntax but not isomorphic with syntactic constituents. Such prosodic constituents are hypothesised to be the domains of both segmental and suprasegmental processes.

The relevance of theories of prosodic constituency to speech timing effects such as final lengthening is discussed in Section 2.3. It is not certain, however, that all durational processes have an indirect relationship to syntactic structure. In particular, initial lengthening[2], a domain-edge process, and polysyllabic shortening, a domain-span process, may prove to be closely associated with the lexical word. Domain-span

[1]This example was provided by John Wells on the *phonet* mailing list.

[2]This effect is sometimes called "onset lengthening" but this term conflates the position of the effect, the start of some constituent, and its apparent locus, the onset of the initially-placed syllable.

processes have also been associated with higher-level syntactic constituents such as the sentence. In the remainder of this section, the relationship between syntactic constituents and the domains and loci of durational processes is examined, to assess the effectiveness of syntactically-defined domains in predicting the occurrence of such processes.

2.2.1 Domain-edge effects: final lengthening

Speech segments have been proposed to lengthen before a pause, at the end of different types of syntactic phrases, and at the end of lexical words. The experimental evidence for each effect is considered here in terms of a domain-and-locus description of speech timing. As the boundaries of higher-level and lower-level constituents frequently coincide, the domain responsible for observed lengthening is not always unambiguously identifiable; the locus of the effect in each case may serve to distinguish between domains. The limitations of a characterisation of domains in purely syntactic terms are discussed.

Utterance-final lengthening

Klatt (1976:1211) says that a word spoken at the end of an utterance is about the same duration as it would be when spoken in isolation and "perhaps as much as twice as long as it would be at the beginning of a sentence." Researchers finding evidence for this effect generally measure the duration of segments at the end of single sentences read in isolation (for example: Gaitenby 1965; Oller 1973; Klatt 1975), thus the effect is often called prepausal lengthening or sentence-final lengthening. The domain of the effect under these conditions may be loosely termed the "utterance"—a pause-delimited and relatively self-contained stretch of speech—which may frequently correspond with the similarly-named constituent of certain prosodic hierarchies (see Section 2.3 for discussion). The essence of utterance-final lengthening is that it occurs before a significant break in speech, and that it is a large and reliable effect.

This broad definition of the domain of utterance-final lengthening reflects agreement within the literature about the circumstances in which it is observed, but the locus of the effect is less well established. Klatt (1976:1212) says that lengthening "often extends over several syllables," although he does not cite evidence in support of this claim. Similarly, Cummins (1999:476) states, without citing evidence, that utterance-final lengthening is "characterised by global deceleration and a reduction in articulatory effort over several syllables." Klatt, in common with other researchers, identifies the rhyme of the utterance-final syllable as being the position of the most substantial durational increment. Oller (1973) examines the effect using the reiterant nonsense words *babab*, *bababab*, *babababab* and *bababababab*, in which the position of primary lexical stress is systematically varied. Both the nucleus and coda of utterance-final syllables show some lengthening compared with syllables earlier in the word: the nucleus has a mean duration of 240 ms compared with 140 ms utterance-medially, and the coda has a mean duration of 80 ms compared with 60 ms utterance-medially. There is only a small lengthening effect on the final syllable onset, and only in the case that the syllable is unstressed. Campbell & Isard (1991) support this finding: syllables lengthened sentence-finally show most of the effect in the syllable rhyme. Campbell & Isard's results also suggest that the composition of the rhyme affects the degree of final lengthening of its constituents: open syllables (those without coda consonants) show more lengthening on the vowel nucleus than do closed syllables.

Examination of the reported duration in Oller's study of vowels and consonants in non-final position within reiterant polysyllables does not provide consistent evidence of the progressive lengthening within the word which would indicate a locus of final lengthening extending to the left of the final syllable; some, but not all, results indicate a slight trend in this direction, and the author makes no comment about such results.

In contrast with many studies of speech timing which use reiterant speech, Oller uses a number of different reiterant templates, such as /baɪb/, /baɪbaɪb/ etc.; /sæs/, /sæsæs/ etc.; /stæst/, /stæstæst/ etc.; /bæ/, /bæbæ/ etc.; /pæp/, /pæpæp/ etc. He finds that the results are largely comparable between these different reiterant materials. Utterance-final fricatives, however, show a particularly large amount of final lengthening, greater that found on the vowel: a mean duration of around 210 ms compared with 70 ms non-utterance-finally. The distribution of lengthening within the locus is often not reported in other studies, with durations simply given for the final syllable or the final syllable rhyme. Cambier-Langeveld's (2000) study of Dutch does report subsyllabic durations, indicating that the distribution of phrase-final and utterance-final lengthening tends to be progressive: that is, the final segment shows the greatest effect—both in absolute terms and in proportion to its non-final duration—and the lengthening of preceding segments diminishes with distance from the boundary. Berkovits (1993b) reports a similar finding for the distribution of utterance-final lengthening in Hebrew. For English, Campbell & Isard (1991) report the distribution of sentence-final lengthening for subsyllabic units in terms of z-scores[3] and find that the coda is lengthened slightly more than the nucleus: the former having a z-score of 1.14 and the latter having a z-score of 1.09. More data are required to determine how far this small difference represents a real trend in English.

Certainly, the discrepancy in Oller's results for the various coda consonants in-

[3]Z-scores here represent the difference between the means for particular sentence-final segments and the overall means for the same segments, in terms of the number of standard deviations of the overall mean for each segment.

dicates that the distribution of lengthening within the locus may vary according to segmental identity, an observation which should be considered when assessing other studies which utilise reiterant speech. As Oller points out, the result may be influenced by the difficulty in comparing certain segment durations in medial and absolute-final position, where they may have "differential mechanisms of termination".

A further point may be made about this study which also applies to other experiments investigating final lengthening. The design of the materials makes it very likely that the target words are pitch accented. The durational effects of pitch accent are discussed in Section 2.5, but it should be noted that the presence of an accent could have significant consequences for the interpretation of such results: firstly, the effect may have greater or lesser magnitude than would otherwise have been observed; secondly, the locus of final lengthening may be different in the presence of pitch accent. In Oller's experiment on utterance-medial, phrase-final words discussed below, it may even be the case—in the absence of independent evidence that a phrase boundary is realised—that the observed durational variation is entirely the effect of pitch accent rather than phrase-finality.

Cambier-Langeveld (2000) provides some data for English and Dutch regarding the interaction between utterance position and accent and the distribution of lengthening beyond the utterance-final syllable. She measures the duration of both syllables of the names *Johnny*, *Michael*, *Macy* and *Joseph* in frame sentences of the type:

I think that Joseph told Johnny about Macy.

where the position of each name and the location of focus—and hence nuclear accent—is systematically varied. Table 2.1 shows mean duration of both syllables in utterance-medial and utterance-final words, when accented and unaccented[4]. The data suggest that there is not a strong interaction between utterance-final lengthening and accentual lengthening, particularly for word-final syllables, which show more final lengthening: there the absolute effect is the same for unaccented and accented words, and the proportional effect is less in the accented case because syllables are longer in accented words.

The locus of utterance-final lengthening appears to extend beyond the final syllable according to the data presented in Table 2.1, although the effect is clearly greater closer to the end of the phrase. It may be noted that the utterance-medial words in this comparison are potentially phrase-final and so the magnitude of utterance-final lengthening may be an underestimate of what it would be relative to a context with no final lengthening influence. Furthermore, if phrase-final lengthening is present on the

[4]Cambier-Langeveld (2000) examines the data in terms of the effect of utterance position on the magnitude of accentual lengthening, but the effect of accent on the magnitude of utterance-final lengthening may be derived from the reported means.

	Stressed syllable, word-initial	Unstressed syllable word-final
Unaccented	19.4 (12%)	122.1 (83%)
Accented	5.4 (3%)	121.5 (68%)

Table 2.1: Data taken from Cambier-Langeveld (2000). Mean utterance-final lengthening (ms) of initial and final syllables in words such as *Joseph* and *Macy*, when accented and unaccented. The proportion of lengthening is shown in parentheses.

utterance-medial words and has a less extensive locus than the utterance-final effect, the relative magnitude of the utterance-final effect on the word-initial and word-final syllables may be distorted[5]. It seems clear, however, that some final lengthening may extend to the left of the final syllable where it is unstressed; furthermore, the locus and magnitude of lengthening are not greatly affected by the presence or absence of pitch accent.

Experiment 2, presented in Chapter 4, examines utterance-final lengthening further, to determine how it is distributed within the locus, and to attempt to corroborate Cambier-Langeveld's (2000) finding that lengthening extends beyond the final syllable where the penultimate syllable carries the primary lexical stress, also testing the assertion that lengthening may extend over several syllables. In addition, Experiment 2 looks again for evidence of an interaction between utterance position and accent.

Phrase-final lengthening

Oller (1973) reports an experiment which he says provides evidence of phrase-final lengthening. Using reiterant nonsense words in the template:

The [bab|babab|bababab|. . .] is on the table

he finds that the vocalic nucleus is substantially longer in the word-final syllable, which is also the final syllable in the noun phrase. In contrast with the utterance-final measurements however, he fails to find much lengthening of coda consonants. In the light of subsequent research, Oller's original interpretation may be re-evaluated, questioning his identification of phrase-final lengthening in this case.

Klatt (1975) identifies the boundary between a noun phrase and a verb phrase as one of the positions at which lengthening of the preboundary syllable may occur, even in the absence of a pause. He also identifies a number of other syntactically-defined locations at which syllables may be lengthened: before conjunctions, between nouns

[5]Data from Cambier-Langeveld (2000) indicate, however, that for Dutch the locus of final lengthening is not more extensive at utterance edges than at intonational phrase edges.

and prepositional phrases and before embedded clauses. Klatt (1976), however, makes it clear that such locations may not always elicit final lengthening effects:

> Goldhor (1976) found that a noun in subject position in a simple sentence is lengthened more when preceded by an adjective than when the noun appears by itself [. . .] Perhaps in a short phrase consisting of a single noun, the speaker does not bother to lengthen the last syllable of the noun to set off the short noun phrase, but rather groups the subject and verb phrase into a single phonological phrase. The addition of an adjective makes it less likely that he will ignore the phrase boundary.(Klatt 1976:1212)

As noted above, Gee & Grosjean (1983) find evidence supporting this view: the boundaries of certain syntactic phrases are not marked by suprasegmental features, and in some cases, the phrasing suggested by such features may contradict the conventional syntactic interpretation. As well as phrase length, Gee & Grosjean suggest that speech rate and symmetry influence the occurrence of boundaries: fewer phrases are manifest at faster speech rates, and there is a tendency to divide utterances up into phrases of approximately equal size.

The approach taken by Cooper & Paccia-Cooper (1980) to this lack of isomorphism between syntactic phrases and suprasegmental features is to postulate an algorithm which takes performance factors such as speech rate and constituent length into account in mediating between syntax and phonetics. A more thoroughgoing approach is to propose a separate phonological structure—the prosodic hierarchy—which is influenced by syntax but distinct from it, as outlined in Section 2.3.

The question of whether a hierarchy of prosodic constituents accurately predicts the occurrence of final lengthening aside, it is clear that syntactic constituents alone do not. This is one reason for questioning the identification by Oller (1973) of phrase-final lengthening in his experiment, where the syntactic phrase comprises only a determiner and a noun: it is uncertain that such a phrase would be delimited by suprasegmental features in normal speech.

Oller's identification of phrase-final lengthening may also be questioned on the grounds that he describes most of the lengthening as occurring on the syllable nucleus, with little consistent evidence of lengthening on the coda. This result contrasts with other studies of phrase-final lengthening which suggest that, as for utterance-final lengthening, the locus of durational variation is the rhyme of the preboundary syllable. Wightman *et al.* (1992) report that both the vocalic nucleus and coda consonants undergo significant preboundary lengthening, with the strongest utterance-medial preboundary effects being comparable to those found utterance-finally[6]. They further suggest that segments preceding the final nucleus do not show preboundary

[6] Wightman *et al.* (1992) also claim that the amount of utterance-medial final lengthening is correlated with the strength of the prosodic boundary, discussed further in Chapter 4.

lengthening, although their method of analysis may mask some smaller variation: in cases where the final syllable is unstressed, they report the normalised duration of the preceding stressed syllable nucleus and the mean normalised duration of all the material intervening between the stressed syllable nucleus and the final syllable nucleus. They found no evidence of preboundary lengthening in this "nucleus-to-nucleus" stretch, and evidence of a slight tendency towards lengthening in the final stressed syllable nucleus when followed by one or more unstressed preboundary syllables.

Two points may be made about these findings. Firstly, the use of the nucleus-to-nucleus measure could mask some localised lengthening within this stretch, such as on the final syllable onset. Secondly, in common with other studies of final lengthening, the effects reported here may be confounded with the effects of pitch accent. This is particularly true of studies of large corpora of speech where the materials have not been designed to separate the two factors. As nuclear pitch accents tend to occur late in an intonational phrase, there will frequently be cases where the preboundary syllable is subject to both final lengthening and accentual lengthening.

As mentioned above, Cambier-Langeveld (2000) finds that accentual lengthening and utterance-final lengthening combine approximately additively in English; in Dutch, she finds that accentual lengthening is markedly attenuated in utterance-final position. With regard to the interaction between accent and utterance-medial phrase-final lengthening, Price *et al.* (1991), analysing the materials used by Wightman *et al.*, report that the lengthening of the vowel nucleus associated with prominence is less if the prominent syllable is part of the last word in an intonational phrase or utterance (the two types of boundaries are grouped together). This suggests a sub-additive interaction between accentual lengthening and final lengthening, such that the effect of both together is less than the sum of the individual effects, although the graphical data presented by Price *et al.* appear support this conclusion only for syllables that are not in word-final position: word-final syllables appear to show at least as large an effect of accent phrase-finality as when phrase-medially, although Price *et al.* make no mention of this apparent trend.

Nakatani *et al.* (1981) record reiterant versions of two-word phrases within frame sentences of normal English words, for example:

We fished in a mama ma we found in the woods

where the reiterant stretch is spoken with the stress pattern of the phrase *remote stream*. They vary the number of syllables and the stress distribution of the two-word reiterant phrases. There is no control over pitch accent placement and examination of the example materials suggests that lexical stress and pitch accent are often confounded. Among the conditions examined is the effect on reiterant syllable duration of position in syntactically-defined phrases. They find comparable final lengthening effects for stressed and unstressed syllables: both are lengthened in absolute phrase-final position, by at least 100ms for three out of four speakers (subsyllabic durations are not reported). They further state: "Compared to the elongation of phrase-final syllables, the effects of other phrase positions on syllable duration were very small indeed" (Nakatani *et al.* 1981:97), but report a small lengthening effect on the phrase-penultimate stressed syllable in the case that the phrase-final syllable is unstressed. None of the observed effects are analysed statistically.

It may be assumed that many of the phrase-penultimate stressed syllables are also pitch accented in Nakatani *et al.*'s study, thus the small observed effect may not generalise to syllables that do not receive pitch accent. Of course, it is very likely in normal speech that phrase-final lexical words will carry pitch accent. In a study using real rather than reiterant speech, Turk (1999) reports lengthening of the rhyme, but not the onset, of the primary lexically stressed syllable in phrase-penultimate position, whether the final syllable carries secondary stress (for example, in *Frankfurt*) or is unstressed (for example, in *Boston*). Turk acknowledges that the words measured are pitch accented in both phrase-final and non-final positions: thus, it seems that the locus of phrase-final lengthening may extend to the left of the final syllable, at least in accented context, although lengthening on the absolute final syllable is greater in all cases. Where the primary lexical stress is in absolute final position, there is no evidence of lengthening on the preceding utterance-penultimate syllable: for example, the first syllable in *Tibet* is not lengthened when the word is phrase-final.

The evidence discussed here suggests the locus of final lengthening extends from the rhyme of the final primary stressed syllable to the phrase boundary. The distribution of the effect may be progressive and is slight in all cases on the onset of the final syllable[7]. It remains to be seen whether this description generalises to non-pitch-accented words and to the case where the final primary stressed syllable in the phrase is followed by more than one syllable before the phrase boundary. Although the experiment presented in Chapter 4 is largely concerned with durational effects at the word-level and at the utterance-level, it does provide some data which relate to lengthening at utterance-internal phrase boundaries.

As discussed above, although final lengthening clearly delimits certain utterance-internal phrases, the phrases themselves are not always syntactically defined. The types of phrases that constitute the domain of final lengthening and their relation to syntactic phrases are discussed further in Section 2.3.

[7] As pointed out by Klatt (1976:1213), the distribution of lengthening within the locus also depends on segmental identity (as observed by Oller regarding utterance-final lengthening): "The lengthening seen at a phrase boundary makes phrase-final fricative and sonorant consonants as much as 40–100ms longer than in non-phrase-final word-final positions. Plosive consonants do not seem to lengthen as much at a phrase boundary."

Word-final lengthening

Word-final lengthening is the extra duration that segments receive because of their proximity to the end of a word in the absence of any larger boundary. Research on word-final lengthening has tended to focus on durational variation in stressed syllables in lexical words. Given the uncertainty discussed earlier about the factors that contribute to prosodic boundaries in speech, separation of word boundaries and phrase boundaries has not proved straightforward for researchers. Klatt (1976:1213) observes: "Early investigators reported large word-final lengthening effects (Barnwell 1971; Lehiste 1972; Klatt 1973), but they did not always control for phrase-final lengthening effects. Word-final lengthening has not always been observed by all investigators (Harris & Umeda 1974) and is probably too small an effect to contribute significantly to the decoding of word boundary locations."

There are other durational influences which affect studies of word-final lengthening. The influence of pitch accent must be considered, as outlined above. Furthermore, the effect of within-word position may be confounded with the effect of word size: as discussed in Section 2.2.3, the hypothesised polysyllabic shortening effect proposes an inverse relationship between word size in syllables and stressed syllable duration.

These multiple influences upon stressed syllable duration within words may be one reason for the continuing uncertainty about the existence of a word-final lengthening effect, with commentators even disagreeing on the interpretation of the same experimental results. Discussing a paper by Beckman & Edwards (1990), Fowler (1990:203) says: "Findings in these experiments and elsewhere (Oller 1973; Klatt 1975) suggest to me that word-final lengthening is fairly ubiquitous," whereas Cutler (1990:208) in the same volume and commentating on the same Beckman & Edwards's paper, says: "In contrast to phrase-final lengthening, word-final lengthening occurs inconsistently. It is more evident at a slow rate of speech; even then, not all speakers show it with all sentences."

The experiments of Beckman & Edwards (1990) which show most evidence of word-final lengthening involve the recording by American English speakers of the phrases *pop opposed* and *poppa posed* in carrier sentences designed to elicit appropriate pitch accent and phrase boundaries at three self-selected speech rates. The vowel durations of the first two syllables in the phrase are reported, the results generally suggesting a word-final lengthening effect even in the absence of a following intonational phrase boundary: the first vowel /a/ tends to be longer in the monosyllabic context and the second vowel /ə/ tends to be longer in the disyllabic context. The data are presented subject-by-subject, which makes comparisons between conditions difficult, but the effect appears to be relatively consistent in both the conditions where the word *pop* was accented and where it was unaccented. Word-final lengthening appears greater for the full vowel than for the schwa and broadly speaking seems more marked at slow speech rate. It is reported that "[the effect] was not significant for all subjects at all rates, but [...] was consistently in the same direction" (Beckman & Edwards 1990:167). A statistical analysis across subjects is not available, so the significance of the mostly small differences observed is uncertain.

Aside from the difficulties in assessing the data due to the method of presentation of the results, there is a specific problem of reliability, alluded to by Cutler (1990:213): the experiment only uses a single phonetic sequence which is manifestly ambiguous in its lexical structure and so there exists "the possibility that the temporal distinctions [the subjects] produced arose from their desire to emphasise the perceived contrast.[8]" Cutler further observes: "Speakers were required to articulate three /p/ phonemes, each followed by a different vowel, in the crucial region to be measured. It is not particularly far-fetched to suggest that the difficulty of articulating such a sequence with clarity might have been a contributory factor in rendering speakers more prone to produce lengthening effects."

These objections are somewhat assuaged in a further experiment reported by Beckman & Edwards (1990), using the phrases *super station*, *superstition* and *Sioux perspective* and measuring the duration of the first two vowels in various accent conditions. The results of this experiment suggest, however, that much of the variation apparent in the other recordings may be artifactual: for four out of six subjects, the only major effect is the greater duration of /u/ in the monosyllable *Sioux*, in the comparison between pitch-accented words. This could be interpreted as evidence of word-final lengthening; alternatively, it could be seen as polysyllabic shortening amplified by pitch accent. Beckman & Edwards do not report onset or coda durations: data about the distribution of the effect within the syllable might help to distinguish word-edge and word-span interpretations, an approach taken in Experiment 2 in Chapter 4. A study by Silverman & Pierrehumbert (1990) of the alignment of accent peaks also reports word-final lengthening effects: for example, the rhyme of the first syllable in <u>*Mom*</u> *Le Mann* is shorter than in <u>*Mam*</u>*alie Le Mann*[9], where the syllable carries a pitch accent in each case. Clearly, the polysyllabic shortening interpretation is also available here, and once again the existence of the effect in the absence of pitch accent is uncertain.

Turk & White (1999) look at the effect of word affiliation on the distribution of lengthening due to accent. An unaccented condition suggests no evidence of word-final lengthening or a balance of word-initial and word-final lengthening: in pairs of

[8]As discussed in Chapter 3, Turk & White (1999) use materials with a similar ambiguity of lexical affiliation, and a control condition without this ambiguity reveals no difference in subjects' production of durational variation. In that experiment, however, a large number of different sentences are used and the recordings are made in blocks designed to deflect attention from possible ambiguity.

[9]This is a study of American English, in which the first syllable is homophonous in these two phrases.

phrases such as *bake enforce* and *bacon force*, there is no significant difference in the duration of the central syllable /ən/ where there is no pitch accent in the phrase. A similar lack of difference is observed for central syllables with secondary lexical stress, such as in the phrase pair *there foreclose* and *therefore close*. If the first word in these phrases carries a pitch accent, then the central syllable is longer when affiliated to that word, but the simplest explanation appears to be the blocking of accentual lengthening at a word boundary.

Turk & Shattuck-Hufnagel (2000) look at a number of hypothesised durational effects related to word structure, and suggest that the most consistent interpretation of their data rules out word-final lengthening, particularly in unaccented words. This research is discussed in detail in Chapter 3 in relation to the results of Experiment 1, which tests the existence of word-level timing processes. Experiment 2, presented in Chapter 4, is designed to distinguish the various interpretations of word-level effects, including word-initial and word-final lengthening, accentual lengthening and polysyllabic shortening.

2.2.2 Domain-edge effects: initial lengthening and shortening

Word-initial lengthening

Oller (1973), as outlined in Section 2.2.1, records reiterant nonsense words in sentences such as *Say a bababab*. In none of the experimental contexts are the initial consonants of the target words phrase-initial or utterance-initial. As stated above, it seems likely that the target words are pitch accented although the placement of accent is not mentioned. In his conclusions, he reports that "the initial consonant effect is even more conclusively demonstrated than most final-syllable effects" (Oller 1973:1244), the effect in question being the lengthening of syllable onset consonants in word-initial syllables when compared with word-medial syllables. He finds the effect for a range of consonants: for /b/, based on recordings of seven subjects; for /s/, /p/ and for both consonants in the cluster /st/, based in single-subject recordings. The fact that both the consonants in the onset cluster show lengthening suggests that the locus of the effect is at least as extensive as the syllable onset, rather than localised on the initial consonant of the syllable.

The magnitude of the effect varies according to the lexical stress of the syllable: for example, in unstressed syllables /b/ is reported to be about 30 ms longer in word-initial position than in word-medial position, where its mean duration is about 40–60 ms; in stressed syllables /b/ is about 20 ms longer in word-initial position than in word-medial position, where its mean duration is about 80–90 ms. Clearly, initial lengthening is proportionally much greater for onset consonants in unstressed syllables, as these are shorter but show more lengthening even in absolute terms. This could be because consonants in stressed syllables are already quite long and do not have much scope for further expansion, for perceptual or for physiological reasons[10].

There is further effect of stress which is not discussed by Oller, but can be inferred from the graphically-presented data. For the consonant /p/, he reports both the duration of the consonant closure and the duration of the whole closure and aspiration interval. From these results, it can be seen that the lengthening of consonant closure in word-initial position is greater for unstressed syllables than for stressed syllables—as for /b/—and also that aspiration duration is greater word-initially for unstressed syllables but greater word-medially for stressed syllables. Estimated from the graph, with a margin of error of about 3 ms either way, aspiration duration is 25 ms word-initially and 15 ms word-medially for unstressed syllables; for stressed syllables it is 25 ms word-initially and 40 ms word-medially.

Oller does not discuss the possibility of the word-initial lengthening effect extending beyond the syllable onset, but there is evidence from the graphical data of a slight trend towards lengthening of the vowel in word-initial syllables in some experiments. Once again, there appears to be an interaction with lexical stress: vowels in word-initial unstressed syllables appear to be longer in many instances than vowels in word-medial unstressed syllables, particularly by comparison with vowels in the second syllable of words. There is no evidence that the locus of word-initial lengthening extends beyond the syllable onset in stressed syllables.

Further evidence for an interaction between stress and word-initial lengthening may be inferred from the results of a study by Cooper (1991), which is primarily concerned with examining the relationship between the place of articulation of voiceless stops and their duration. Two speakers say the nonsense words *pipip*, *titit* and *kikik* in the carrier sentence *Say . . . again*. The primary stress is either on the first or the second syllable, and it may be surmised that the target word is accompanied by a pitch accent, although this is not stated.

The results for the consonant /p/ are shown in Table 2.2. These exemplify the trend for all three consonants, which is similar to the pattern inferred from Oller's study. Closure duration for all the voiceless stops is longer word-initially than word-medially, for both stressed and unstressed syllables. Once again, aspiration duration is longer word-initially only in unstressed syllables: in stressed syllables, aspiration duration is greater word-medially[11]. Comparison between stressed and unstressed syllable onset durations suggests that the source of this difference may be word-initial shortening of aspiration duration in stressed syllables, rather than a word-medial

[10] Voiced stops can only be sustained for a short time before the build up of supraglottal pressure during closure causes voicing to stop.

[11] As in Oller's study, these trends are not discussed but may be estimated from the graphical data: aspiration duration is greater word-medially in all three consonants for one subject, and in two out of three for the other subject.

lengthening. This conclusion may be drawn because the difference in aspiration duration between stressed and unstressed syllables is comparable word-medially to the stressed vs unstressed differences seen for closure duration; word-initially, however, the difference in aspiration duration between stressed and unstressed context is reduced for subject KM and effectively absent for subject ES.

Syllable context		Subject KM		Subject ES	
		Closure	Aspiration	Closure	Aspiration
Stressed syllable	Word-initial	105	54	98	27
	Word-medial	79	65	82	31
Unstressed syllable	Word-initial	83	44	77	33
	Word-medial	54	28	71	29

Table 2.2: Durations (ms) of closure and aspiration for the onset consonant /p/ in the study by Cooper (1991). The durations shown here have been estimated from graphical data and have a margin of error of 1 ms either way.

Given the likely presence of a pitch accent on the target words in Oller's and Cooper's studies, this interaction between syllable stress and word position for aspiration duration may in fact represent an interaction with pitch accent, because the syllable with primary lexical stress in the accented word will carry the accent, at least for nuclear-accented syllables. A study by Fougeron & Keating (1997) using reiterant speech also finds word-initial lengthening effects on onset consonants, but as they only report the duration of the onset consonant /n/ in unstressed syllable context, their study does not shed further light on the possible interaction between syllable stress and word position, particularly its effect on aspiration duration in voiceless stops. Experiment 2, reported in Chapter 4, examines this interaction further, by looking at the duration of a range of onset consonants in stressed syllables which are initial or medial in accented and unaccented words.

The studies reported here do not discuss the possibility that the locus of word-initial lengthening extends beyond the onset consonant, but Turk & Shattuck-Hufnagel (2000) present evidence—discussed in more detail in Chapter 3—that suggests the effect is largely confined to the onset, and this conclusion is supported by evidence from Fougeron & Keating (1997) and Byrd (2000). The locus of initial lengthening is examined further in Experiment 2.

One further point should be made here: studies of word-initial lengthening tend to use reiterant versions of lexical words. It is not clear whether these results generalise to function words, and how they relate to prosodically-defined words. In the latter regard, if the domain of initial lengthening is the prosodic word rather than the lexical word, lengthening may be seen lexical-word-internally in compounds if they comprise more than one prosodic word. The question of lexical and prosodic word structure in relation to speech timing is considered in more detail in Section 2.3.

Utterance-initial shortening

Another example of a domain-initial shortening effect has occasionally been observed: the shortening of certain onset consonants utterance-initially; that is, immediately following a break in speech. Fourakis & Monahan (1988) find that the unstressed syllable /mə/ has a mean duration of 75ms in an utterance beginning *Machines will* and a mean duration of 107ms in an utterance beginning *Your machines will*. They consider the possibility that this is due to a measurement error, suggesting that the utterance-initial amplitude at the beginning of /mə/ might have been too weak to appear on the spectrogram, but think this explanation is unlikely because the effect is consistent across speakers and because they "would expect that there should be enough energy in the signal four to six [glottal] pulses into the utterance to register on the spectrograph" (Fourakis & Monahan 1988:293).

Fourakis & Monahan do not report subsyllabic durations, so it is impossible to know what the locus of the effect is, but Fougeron & Keating (1997) report onset consonant duration in a study using reiterant speech. For two of their three speakers, utterance-initial /n/ is shorter than word-initial /n/ in utterance-medial position: for example, for one speaker, /n/ is about 55ms long utterance-initially and about 70ms longer utterance-medially. The magnitude of the effect, where observed, is less than that reported for the whole syllable in Fourakis & Monahan's (1988) study, which might indicate that the locus of utterance-initial shortening encompasses the vocalic nucleus as well, but given the small number of subjects used by Fougeron & Keating, this difference may simply be a matter of between-speaker variation. As with the Fourakis & Monahan study, the syllable in question is not the primary stress: in the Fougeron & Keating study, it is the first syllable in a reiterant version of the number *89*. Furthermore, the syllable onsets in both studies are nasal consonants, so it is uncertain whether the effect applies to all types of onset consonants.

Fougeron and Keating's study is concerned in part with discovering whether the magnitude of initial lengthening is correlated with the strength of prosodic boundaries: that is, whether constituents at higher levels of the prosodic hierarchy show greater initial lengthening effects. The design of this and a similar study by Wightman *et al.* (1992) are analysed in detail in Chapter 4. Experiment 2 presented there includes a comparison of word-initial syllable duration in utterance-initial and utterance-medial position, to determine if the utterance-initial shortening of nasal onset consonants in unstressed syllables is observed in stressed syllables as well, and if so, if the effect also applies to other types of onset consonants. The duration of subsequent vowels is also reported, to determine if the locus of the effect is the syllable onset or something larger.

2.2.3 Domain-span effects: polysyllabic shortening

The polysyllabic shortening effect may be described thus:

The more syllables that there are in a word, the shorter will be the segments within those syllables.

By this definition, it is strictly a domain-span compression effect, as characterised in Chapter 1, affecting the duration of segments within the locus regardless of position. Most investigators (for example: Lehiste 1972; Port 1981) examine the duration of the primary stressed syllable in relation to word length, as it is the only syllable available for comparison between monosyllabic and disyllabic contexts; such studies might be taken to imply that the locus of polysyllabic shortening is the primary stressed syllable. In contrast, Barnwell (1971:87) concludes from his data that "when the number of syllables for words in a given (constant) structural location is increased [. . .] both the stressed and unstressed vowels decrease in size [and] the consonants also decrease in size although not as much." This assertion suggests that the locus of the effect is co-extensive with the domain: all parts of the word are subject to durational variation due to word length, with other factors—such as, for example, articulatory constraints and perceptual salience—determining how the timing adjustments are worked out.

The question of the locus of polysyllabic shortening may be left aside for the moment, however, as the results of previous studies do not conclusively demonstrate that the effect is attributable to a domain-span process at all. Indeed, studies often cited as evidence for polysyllabic shortening, such as Barnwell (1971), Lehiste (1972) and Klatt (1973), are elsewhere held to demonstrate another process: Klatt (1976), for example, cites these studies as evidence of word-final lengthening, possibly confounded by phrase-final lengthening. As a typical experimental design compares the duration of the primary stressed syllable in words such as *sleep*, *sleepy* and *sleepiness*, the increase in word length coincides with an increase in the distance of the measured syllable from the end of the word, and possibly from the end of a phrase, depending on the structure of the carrier sentence.

There are other potential confounds in previous studies of polysyllabic shortening, such as rhythmical effects and the influence of pitch accent, discussed below. The interpretation of such studies, in particular Lehiste (1972) and Port (1981), is analysed in Chapter 3 and Experiment 1 is designed to eliminate potential confounds to determine if previous observations may indeed be attributed to a word-level process. The results of Experiment 1 are considered in relation to those of Turk & Shattuck-

Hufnagel (2000); Experiment 2, in Chapter 4, attempts to determine the *type* of process occurring at the word-level.

Two other issues should be mentioned at this point. Firstly, as with word-initial lengthening, it is uncertain whether polysyllabic shortening, if it can be demonstrated unambiguously, operates over a syntactic domain—the lexical word—or a phonological domain—the prosodic word or a similar near-word-sized unit. The nature of the prosodic word is discussed further in Section 2.3.2.

Secondly, the study by Huggins (1975) is sometimes cited as evidence for polysyllabic shortening, but this is not his conclusion. He reports the durations of the syllables /tʃiz/ and /baʊnd/ in the context of sentences formed from the template

Cheese(s) (a)bound(ed) (ab)out.

where the presence of the unstressed syllables in brackets is systematically varied. He finds that the duration of /tʃiz/ is reduced by an unstressed syllable within the same word, but not by an unstressed syllable following the word boundary. From this, the existence of a word-level effect might be inferred. He also finds, however, that /baʊnd/ is shortened by an unstressed syllable both within the word and across a word boundary. His conclusion is that the shortening is a rhythmical effect—as discussed in Section 2.5—which is blocked at the major syntactic boundary between *cheese* and *bound*. Furthermore, Huggins (1975:462) states that this conclusion is "speculative, being based on very little data," and Rakerd *et al.* (1987:149) report that Huggins failed to replicate the shortening effect with added sentences and subjects.

2.2.4 Domain-span effects: sentence-level shortening

Lehiste (1974), cited in Klatt (1976:1212), finds "a general tendency for readers to talk faster when phrases or sentences are longer." This effect, also alluded to by Jones (1942–43), suggests a domain-span compression effect, analogous to polysyllabic shortening, but over a larger domain such as the sentence. The domain may also be the phonological utterance, which, as discussed in Section 2.2.1, may generally be equated with the sentence in studies requiring subjects to read isolated sentences. As rate of speech is held to increase over the whole sentence, the locus of such an effect would be co-extensive with the domain. As with polysyllabic shortening, however, it remains uncertain whether this interpretation represents the most accurate account.

In a single subject study, Gaitenby (1965) measures the duration of phrases in sentences of varying lengths.

1. *Why don't you?*

2. *Why don't you get tickets?*

3. *Why don't you get tickets for tomorrow?*

4. *Why don't you get tickets for tomorrow night?*

5. *Why don't you get tickets for tomorrow night's programme?*

She finds that each added phrase is longer in sentence-final position than otherwise, but that, other than this effect, there is no clear pattern of variation due to sentence length. Thus *get tickets* is 740 ms in Sentence 2 and 590 ms in Sentence 3, but Sentence 4 and Sentence 5 cause little further variation in the duration of *get tickets*, which is 590 ms in the former and 580 ms in the latter. Similar results are reported for the other phrases in the sentence.

In contrast, Rakerd *et al.* (1987) find that the duration of a sentence-medial syllable is greater when the sentence in which it is uttered is shorter. They measure the duration of monosyllabic lexical words in a pair of sentences of different lengths, for example, *date* in the context of:

1. *His first date aroused some anxiety.*

2. *His first date aroused some anxiety for obvious reasons.*

The mean duration of the monosyllabic targets is 11 ms greater in shorter sentences such as Sentence 1. This may be seen as evidence of a sentence-span effect, but other interpretations remain available. Firstly, if a domain-span effect is responsible for the observation, the domain may be better characterised as a prosodic entity such as the utterance rather than a syntactic entity such as the sentence. Furthermore, the locus of the effect remains to be established: is it just the duration of lexically-stressed or pitch-accented syllables that is affected by domain length, or does the adjustment take place throughout the domain? If the former, is this because such syllables are structurally significant or simply because they are longer and so have more scope for durational variation?

Secondly, there may be an alternative, indirect explanation of apparent sentence-span shortening: the length of the sentence may influence its phrasing, such that the measured word or syllable stands in a different position with respect to boundaries in the two versions of the experimental sentence. For example, Sentence 1 might be realised as two phrases with a boundary after the subject noun phrase: thus *date* would be subject to phrase-final lengthening. If Sentence 2 elicited a different phrasing—perhaps two phrases with a boundary after the object noun phrase—*date* would be shorter because of its phrase-internal position. Thus an apparent domain-span effect would have a domain-edge explanation.

The experiment described in Chapter 4 is designed to test the existence of domain-span processes in the word and the sentence/utterance.

2.3 Prosodic constituents

As the evidence discussed above demonstrates, syntax does not reliably predict the occurrence of all durational processes associated with constituents of spoken English: phrase-final lengthening, in particular, appears to take place in domains that are only indirectly related to syntactic structure. Prosodic hierarchies have been proposed by researchers such as Nespor & Vogel (1986), Hayes (1989) and Selkirk (1978; 1986; 1996) to account for the domains of occurrence of segmental phonological processes, and in certain cases, intonational phenomena (see Shattuck-Hufnagel & Turk 1996 for a review). It has been proposed by researchers such as Wightman *et al.* (1992) that it is these prosodic constituents which constitute the domains of final lengthening effects. Other researchers, such as Fougeron & Keating (1997), have linked the strength of prosodic constituent boundaries to the magnitude of initial lengthening effects.

Specific details of prosodic hierarchies vary between theories: Nespor & Vogel's (1986) account is presented in Section 2.3.1 as an illustration of how prosodic constituents may be derived. Constituents which may influence durational processes, and which may have different names in different theories, are discussed in Section 2.3.2.

2.3.1 A prosodic hierarchy

Nespor & Vogel's framework is proposed as one of a number of interacting phonological subsystems, "each governed by its own principles, such as the theories of the metrical grid, lexical phonology, autosegmental phonology and prosodic phonology" (Nespor & Vogel 1986:1). This separation reflects one possible standpoint on the disparate influences on speech timing, distinguishing those aspects of durational variation that are directly predictable from syntax; those that can be accounted for within the domains provided by prosodic phonology; those that are better described within a different phonological subsystem, such as metrical phonology; and those that are "better relegated to phonetic interpretation along with other details of phonetic timing such as the different durations of vowels before voiced and unvoiced obstruents" (Beckman & Edwards 1990:158)[12]. Nespor & Vogel's distinction is not always clear-cut, however: they include as evidence for their prosodic constituents metrical effects such as stress shift, and intonational phenomena that might otherwise be regarded as belonging to autosegmental phonology. Some researchers go further and explicitly attempt to link apparently distinct subsystems: for example, Beckman & Edwards (1990,

[12]The influence of syntax on speech timing, originally considered to be direct, is often now regarded as necessarily mediated by prosodic structure. As this discussion in this chapter illustrates, the evidence for this view is stronger for some timing processes, such as phrase-final lengthening, than it is for others, such as the various hypothesised word-level effects.

1994) attempt to identify each level of prominence as the head of a particular prosodic constituent. Whether different influences should be separated or whether they reflect an undiscovered unified representation of speech timing remains to be seen, but some separation appears at present to be necessary: as discussed in Section 2.4, some rhythmical aspects of speech timing do not appear to respect the boundaries of syntactic or prosodic constituents.

Nespor & Vogel present a hierarchy of prosodic constituents—syllable, foot, phonological word, clitic group, phonological phrase, intonational phrase, phonological utterance—which is observable because "each prosodic constituent serves as the domain of application of specific phonological rules and phonetic processes" (Nespor & Vogel 1986:1). The prosodic constituents above the level of the word are constructed according to rules which refer to the output of the morphosyntactic component of the grammar, but also take into account factors such constituent length, branching, and prominence relations.

The phenomena for which prosodic constituents serve as domains include effects such as liaison in French, *raddoppiamento sintattico* in dialects of Italian, nasal assimilation in Spanish, flapping in American English, and linking-r and intrusive-r in dialects of British English. Such processes are often referred to as external sandhi rules, meaning word-external: because the word does not have a privileged position within their framework, Nespor & Vogel do not use the qualifier "external". As mentioned above, they also invoke certain non-segmental phenomena such as iambic reversal (also known as stress shift) in American English as evidence for certain constituents. The application of these processes within the domains delimited by the constituents of the prosodic hierarchy is governed by three types of rules: domain-span rules, domain-limit rules and domain-juncture rules.

Domain-span rules are those that apply within a domain but not at its edge[13]. For instance, the phenomena of linking-r and intrusive-r in certain non-rhotic dialects of British English are accounted for by a phonological utterance-span rule of r-insertion. Thus if any of the vowels /ɔ/, /ɑ/ or /ə/ is followed immediately by another vowel within a phonological utterance, an /ɹ/ may be articulated between them. At utterance boundaries, however, this process is blocked. Iambic reversal is also cast as a domain-span rule of the phonological phrase: an iambic word may be realised as a trochee when followed by a stressed syllable within the same phonological phrase. Thus, although the word *thirteen* is iambic in citation form, in the phrase *thirteen men* it may have its main stress realised word-initially. As a domain-span rule, iambic reversal is blocked by the intervention of a phonological phrase boundary.

[13]The usage of the term "domain-span" is rather different in Nespor & Vogel's account of segmental phonology from its usage in this dissertation, where it refers to durational variation related to the size of a particular constituent and affecting subconstituents regardless of position.

Domain-limit rules apply at the edges of domains but not within domains. An example is the glottalisation of /t/ in English, which may occur when /t/ is syllable-final. The aspiration of /t/ is also described as domain-limit rule, in this case at the level of the foot: /t/ is aspirated in certain dialects of American English when it is foot-initial but not elsewhere.

Domain-juncture rules, like domain-limit rules, occur at the edges of domains but not within them, further requiring that the appropriate conditions are met on the other side of the domain-boundary. The trilling of orthographic *r* in Spanish is described as a domain-juncture rule at the level of the syllable: a syllable-initial *r* may be trilled rather than tapped if the previous syllable ends with a sonorant.

The principles which govern the hierarchical construction of prosodic constituents into trees are in important respects different from those which apply in syntax. In particular, prosodic trees are n-ary branching rather than binary branching, forming an essentially flat structure, rather than binary trees which have in principle unlimited depth. There may be any number of sister nodes at a given level of the prosodic tree: one of them is designated the strong node according to specific principles applying to each level, the others are all designated as weak. The other important principle of construction in Nespor & Vogel's theory of the prosodic hierarchy is the "strict layer rule". This means that any constituent of a given level is exhaustively contained within the superordinate constituent of which it is a part, and conversely, any non-terminal constituent is composed exclusively of one or more immediately lower units. Not all theories of the prosodic hierarchy require that strict layering holds in all circumstances, however: following an optimality theory approach taken by Ito & Mester (1992), Selkirk (1996) analyses the strict layering rule into four constraints called "headedness", "layeredness", "exhaustivity" and "nonrecursivity". In Selkirk's prosodic hierarchy, the latter two constraints are violated in some circumstances: where exhaustivity is violated, a constituent may dominate another constituent of more than one level down; where nonrecursivity is violated, a constituent may dominate another constituent of the same level.

An important characteristic of Nespor & Vogel's approach, particularly in relation to speech timing, is the restructuring of constituents allowed at higher levels of the hierarchy, from the phonological phrase upwards. The length of constituents is important in determining whether larger constituents are formed from the merger of two or more constituents at the same level. In reference to intonational phrase (IP) restructuring, they say that "the faster the rate of speech, the longer the [IPs] of a given utterance tend to be; conversely, the slower the rate of speech, the shorter the [IPs] tend to be. In this way, some more abstract notion of length in terms of timing or rhythm may be respected"[14] (Nespor & Vogel 1986:195). They further state that "for both [IP] and [phonological utterance] restructuring, what seems to be crucial is some average length in terms of timing" (Nespor & Vogel 1986:240).

Nespor & Vogel's prosodic hierarchy is intended to be universal, with the same set of constituents combining in similar fashion in all languages, although certain parameters—such as the direction of affiliation of lower-level constituents into higher-level ones—may vary between languages. They claim that although a particular language may lack segmental domain rules for some phonological constituents, these constituents may still be necessary in order to define relative prominence relations, or to account for other phenomena such as constituent-final lengthening. This suggestion may be intended to imply that all constituents within the hierarchy, with the likely exception of subword units, manifest final lengthening to some extent. Specific experimental evidence for a hierarchy of constituents delimited by durational effects of increasing magnitude at higher levels—what may be termed "hierarchical lengthening"—is discussed in Chapter 4, where Experiment 2 explores two potential levels of hierarchical variation: the word and the utterance.

2.3.2 Prosodic constituents and speech timing

The general principles of the prosodic hierarchy having been outlined, brief descriptions of particular prosodic constituents are presented in this section. Particular emphasis is given to those constituents that may serve as domains of suprasyllabic durational processes. Points of major divergence between the hierarchy of Nespor & Vogel and other theories, such as that of Selkirk (1996), are also indicated. In accordance with common practice, what Nespor & Vogel call the "phonological word" is referred to here as the "prosodic word". Also, Nespor & Vogel refer to a sub-word constituent which they call the "foot"; as discussed in Section 2.4, there are a number of possible types of feet, and so the term "within-word foot" is used here to avoid ambiguity.

The Syllable

Subsyllabic units such as the onset, rhyme and coda are excluded from Nespor & Vogel's prosodic hierarchy. Firstly, they do not obey the principle of construction which requires each level to be exclusively composed of one type of constituent. Secondly, a number of segmental phonological rules which can be expressed in terms of subsyllabic units can also be expressed in terms of a domain defined by the syllable itself. Syllables are thus the lowest unit in the prosodic hierarchy. Syllabification is according

[14]It may be noted that domain-span compression effects, of the type discussed in Section 2.2.4, make a similar link between constituent size and *local* speech rate, although there the direction of causality is reversed: the rate of speech is faster because the constituent is longer.

the maximal onset principle, over the domain of the phonological word. Resyllabification is possible over larger domains in some languages, such as many European Romance languages, but not in English.

As mentioned in Chapter 1, there are a number of durational processes which may be related to the domain of the syllable. Firstly, the distribution of lexical stress is generally described in terms of syllables: the primary acoustic indicator of lexical stress, together with vowel quality, is lengthening of the stressed vowel and, to a lesser extent, of the onset and coda consonants. Other durational processes relating to the distribution of stressed and unstressed syllables are described in Section 2.4. There are also a number of processes which may be related to the composition of syllables or subsyllabic units, such as the shortening of consonants in clusters and the lengthening of a vowel before a voiced coda.

The Within-Word Foot

Each within-word foot is composed of one relatively strong syllable and a number of relatively weak syllables, up to a prosodic word boundary. In Nespor & Vogel's account, some languages allow only binary feet, but in English any number of weak syllables may be included. Other parameters which vary between languages are the position of the strong syllable—it is the leftmost syllable in English—and the possibility of refooting across words. Nespor and Vogel say that this is not allowed in English: the domain over which feet must be well-formed is the prosodic word.

Studies discussed in Section 2.2.1 have suggested that the within-word foot may be the locus of phrase-final lengthening (for example: Nakatani *et al.* 1981), although such lengthening seems to exclude the onset consonant of the first syllable of the foot, and possibly the other onset consonants. Turk (1999) finds, however, that words such as *Woodstock*, containing two within-word feet, showed phrase-final lengthening on the word-initial syllable as well as on the word-final syllable. As described in Section 2.5, the results of Turk & Sawusch (1997) suggest the possibility that the within-word foot may be the locus of accentual lengthening, although subsequent research such as Turk & White (1999) shows that accentual lengthening can extend beyond the within-word foot. Thus, it now appears that the within-word foot is not the locus of final lengthening or accentual lengthening.

The Prosodic Word

Nespor & Vogel's (1986) rules for the construction of the prosodic word differ between languages—indeed, they say that it is the constituent which varies most between languages—but there are no instances in which a prosodic word may be larger than a terminal node of the syntactic tree. In some languages, such as Greek and Latin, there is exact identity between these terminal nodes and prosodic words, but for most languages each morphological stem must belong to a separate word, with affixes generally attaching by adjacency to the stem within the same terminal syntactic node. In languages such as Hungarian and Italian, morphological or phonological criteria identify certain affixes which must comprise separate prosodic words. The status of English is not explicitly discussed in Nespor & Vogel's account, although it may be surmised that it belongs to the set of languages identifying each morphological stem with a separate prosodic word.

The prosodic word is constructed from all the feet within its domain. The relative prominence rules require that, within a given language, in the unmarked case the strong foot will consistently be either the leftmost one or the rightmost one in the prosodic word of a given language. In English, it is the rightmost foot which is strong in the unmarked case.

There are two main points of contention regarding the mapping between syntactic and prosodic words: one is the prosodic status of function words, the other is the prosodic nature of compounds.

In Nespor & Vogel's formulation, every syntactic word—lexical and functional—forms a single prosodic word (except where the syntactic word contains more than one morphological stem). This significantly differs from other theories, notably that of Selkirk (1996), which accord prosodic word status to lexical words as a matter of course, but to function words only in certain cases. According to Selkirk, function words form prosodic words in citation form and when in focus (that is, when pitch accented); most also form prosodic words when phonological-phrase-final. In other circumstances, monosyllabic function words are regarded as clitics—unstressed syllables which are not dominated by either a within-word foot or a prosodic word of their own. Disyllabic function words, and a few monosyllabic function words such as *up*, *too* and *off* are regarded as footed by virtue of containing a stressed syllable, but do not form prosodic words of their own. The structures within which function words coordinate with lexical words depend whether they precede a lexical word (proclitics) or follow it (enclitics).

Selkirk (1996) argues that an enclitic function word—for example, the pronoun *him* in *need him*—may join prosodic structure within a nested prosodic word (PW): $((need)_{PW} him)_{PW}$. The verb in *need him* is directly dominated by a prosodic word, which is itself dominated by another prosodic word which also dominates the pronoun; thus, the strict layering constraint of nonrecursivity—described in Section 2.3.1—is violated. A proclitic function word may join prosodic structure at a higher level, such as the phonological phrase (PP); for example, the auxiliary *can* in *can paint*: $(can\ (paint)_{PW})_{PP}$; here the strict layering constraint of exhaustivity is violated. In contrast, Nespor & Vogel (1986) claim both lexical words and function words form prosodic words in their own right; "functional" prosodic words may then combine with "lexical" prosodic words within clitic groups.

Clearly, Nespor & Vogel's theory would predict that prosodic-word-level timing processes should apply to both content words and function words, but Selkirk's account would not, at least for proclitics. The resolution of this issue is not straightforward: given the closed-class nature of function words, and the fact that most are monosyllabic, the experimental examination of processes such as initial lengthening and polysyllabic shortening is rendered highly problematic and to some extent irrelevant. Furthermore, in practical terms, function words tend to be much shorter than lexical words, and thus any systematic durational variation would be slight in, for example, disyllabic function words compared with monosyllabic function words[15]. The issue of the nesting of prosodic words may be more tractable, however: Selkirk says that "words with 'level 2' suffixes [...] in English have exactly the same prosodic structure as verbs with object clitics." Thus, if durational processes apply across a prosodic word domain, they should apply equally to *saw it* and to *sawing*.

The theoretical standpoints regarding compounds are harder to determine, but the main point of contention appears to be whether compounds—words containing two morphological stems, such as *kneecap* and *shakedown*—form single prosodic words, pairs of prosodic words or nested prosodic words. For other languages, apparently including English, Nespor & Vogel suggest that each morphological stem forms its own prosodic word. Selkirk (1984) claims that, syntactically, compounds have a nested structure, for example: $((\text{knee})_W(\text{cap})_W)_W$, and thus both parts of compound are joined within a single word. It is unclear how such a structure is to be interpreted prosodically: as noted above with regard to function words, Selkirk proposes that prosodic words may be nested, so the two prosodic words in a compound may be dominated by a higher prosodic word, but does not explicitly discuss the case of syntactic compounds. She does say, however, that "a sequence of lexical words [...] in morphosyntactic representation [...] is characteristically prosodised as a sequence of prosodic words [...] in phonological representation" (Selkirk 1996:188), which might be taken to imply a one-to-one mapping. As syntactic compound words are nested, they would also be nested prosodically.

If the prosodic word, rather than the lexical word, is the domain word-level processes, then, if the above interpretation of Selkirk's theory is correct, such processes would operate recursively. For example, each prosodic word individually and the

[15]In certain circumstances, such as phrase-finally and when in focus, function words are not reduced: the most obvious timing consequence of this is lengthening. It may be expected that the distribution of such lengthening would reflect that seen in lexical words when phrase-final or accented, but this issue is beyond the scope of this dissertation.

higher prosodic word would exhibit word-initial lengthening and polysyllabic shortening. As mentioned above in relation to function words, the fact that many compounds are formed of two monosyllabic stems makes testing the predictions of the nested and non-nested hypotheses difficult for polysyllabic shortening; for word-initial lengthening, one prediction would be that would be greater in compounds, as a result of the operation of two recursive word-initial processes, than in monomorphemic words: thus, /k/ would be longer on *crowbar* than in *crowing*.

What durational evidence there is on this issue suggests that compounds do exhibit similar durational properties to noncompounds. Barnwell (1971) observes that compounds comprising two monosyllabic words behave for timing purposes as a single disyllabic word. Klatt (1975), based upon a small amount of data, finds that vowel duration is better predicted if compounds are treated as "single phonological words" over which a word-span compression process operates. Turk & White (1999) find similar patterns of accentual lengthening across a range of disyllabic words containing two stressed syllables, including some noncompounds—such as *skimpy* and *capsize*—and compounds—such as *kneecap* and *downstairs*.

The issue of the nesting of prosodic words is not tested directly in Experiment 1 or Experiment 2, and it may be assumed that the term "word" is used to mean a lexical word, whether noncompound or compound.

The Clitic Group

As observed above, clitics are monosyllabic function words that have some phonological dependence on non-clitic constituents. Clitics may not, for example, receive contrastive stress or stand alone as utterances. In Nespor & Vogel's theory, the clitic group dominates single prosodic words and adjacent clitics: the construction of the clitic group makes reference to the syntactic structure of a language, so that in cases where clitics may either encliticise or procliticise, the direction of attachment is often determined by syntactic structure. In most languages, including English, the strong element within the clitic group is the prosodic word that contains the non-clitic element.

The behaviour of clitics in Romance languages such as Italian strongly motivates the hypothesised clitic group, as illustrated by Nespor & Vogel. It is less clear whether languages such as English have phonological processes which support a single level at which all clitics are associated with adjacent lexical words. Selkirk's (1996) treatment of clitics is different: as discussed above, enclitics attach to a prosodic word recursively dominating the preceding prosodic word. Proclitics are not dominated by a prosodic word, but attach to the phonological phrase which dominates the following prosodic word. Thus, in the phrases *a conversion*, *for a massage*, *that you could ask*, all the function words are unfooted syllables attached to the phonological phrase which dominates the phrase-final prosodic word. Selkirk's phonological phrase may thus be said to share some characteristics of Nespor & Vogel's clitic group[16].

Due, perhaps, to the uncertain status of clitics in English, the timing consequences of this constituent have not been explored, although the formulations of Nespor & Vogel and Selkirk differ in their predictions: if there is a domain-span timing process which applies to the clitic group, proclitics and enclitics should be equivalent in their effect on the size of the constituent according to Nespor & Vogel. In Selkirk's theory however, enclitics should influence prosodic word timing processes, whereas proclitics only affect lexical words though processes which may hold at the phonological phrase level.

The Phonological Phrase

Nespor & Vogel (1986:168) construct the phonological phrase from: "a clitic group which contains a lexical head (X) and all clitic groups on its non-recursive side up to the maximal projection of X." (In languages, such as English, whose syntactic trees are right-branching, the rightmost node of the phonological phrase is the strong element; in languages whose syntactic trees are left-branching, it is the leftmost node which is strong.) For example, the utterance:

The indolent donkey| masticated| the tasty thistle| languorously.

divides as indicated into four phonological phrases

Although the phonological phrase (Φ) has a precise syntactic definition, it is the lowest constituent in the prosodic hierarchy for which optional restructuring is possible, and this operates according to non-syntactic criteria: "A nonbranching Φ which is the first complement of X on its recursive side is joined into the Φ that contains X" (Nespor & Vogel 1986:168). The requirement that the adjacent phonological phrase be nonbranching reflects the intuition that the length of prosodically-defined phrases tends to cluster around some mean, with very short or very longer phrases being less likely. For example, the *indolent donkey* sentence above could be restructured into two phonological phrases by combining the first two phrases and the last two phrases, but this is unlikely because the phrases are already relatively long. A similarly-structured utterance with shorter phrases, such as:

The mouse| chewed| the cheese| fast.

is more likely to undergo restructuring to form two phonological phrases with a boundary following *chewed*.

[16] Selkirk (1996) does not make the distinction between major and minor phonological phrases that she did in previous work such as Selkirk (1986).

It is possible that there would be no acoustic boundary within this sentence: such an observation suggests either that Φ may restructure more than Nespor & Vogel propose or that Φ is not a domain of final lengthening. The phonological phrase is proposed by Nespor & Vogel, however, as the domain of final lengthening in English: "Syntactic phrasing plays a role in lengthening only in the sense that the phonological phrase is built on the basis of syntactic notions. The reason a distinction [between syntactic and phonological phrases] was not made earlier is most likely due to the fact that the two phrasings often do coincide" (Nespor & Vogel 1986:1990). They relate the blocking of iambic reversal at phonological phrase boundaries to the fact that lengthening in this position increases the interval between adjacent stresses.

Nespor and Vogel do not, however, put forward specific experimental evidence in support of their claim that the phonological phrase is the domain of final lengthening, and the status of the assertion is not clear[17]. As mentioned above, Nespor & Vogel suggest that (unidentified) constituents which are not supported by segmental phonological evidence within a particular language may prove necessary as domains of final lengthening, implying that more than one constituent may show this effect. They allude to prominence relations influencing the relative magnitude of final lengthening between intonational phrases[18], but they make no specific claims for constituents below the level of the phonological phrase, such as the prosodic word. (As discussed in Section 2.2.1, the evidence supporting final lengthening at the word-level is uncompelling, so Nespor & Vogel may be justified in ignoring it.) Thus, the extent to which they regard final lengthening as specifically associated with a phonological phrase domain or as a property of some or all prosodic constituents is uncertain.

If the phonological phrase does manifest final lengthening, it is at first sight sufficient to account for all final lengthening at or above this level: according to the strict layering hypothesis, final edges of higher constituents must necessarily coincide with final edges of lower constituents. The end of every utterance is also the end of an intonational phrase, and the end of every intonational phrase is also the end of a phonological phrase. If lengthening of greater magnitude is observed at the end of the higher constituent, this may be related to the fact that the final phonological phrase in the intonational phrase or the utterance is the stronger. For example, the sentence

There was a large black dog and a tiny brown dog

may be realised as three phonological phrases, with a boundary after the verb and an-

[17]Selkirk (1996:202) also suggests a link between phonological phrases and speech timing: "Independent evidence for or against this right-edge alignment of phonological and syntactic phrases in English will most likely come from phenomena involving intonation or durational patterns, which are argued to be characterised in terms of phrase-sized units of prosodic structure."

[18]Nespor & Vogel (1986:223) say that sentence-final intonation and lengthening seem to indicate that "the last [intonational phrase] of [an utterance] is the strong one."

other before *and*. In this case, Nespor & Vogel would predict final lengthening on the word *dog* in both instances. If the second instance of the word is found to be longer than the first, it is not necessary to postulate intonational-phrase-final or utterance-final lengthening; rather, it is more parsimonious simply to state that the final phonological phrase in the intonational phrase is the strong one and thus shows more final lengthening.

If final lengthening is shown to occur consistently at the end of phonological phrases, there are two possible observations which might require the postulation of lengthening at the end of higher-level domains. Firstly, if the phonological phrase is the only domain of final lengthening, the strong phrase should still show the greatest degree of lengthening even when it is not the rightmost phrase of a set of daughters. (This requires some independent test of phonological phrase strength, which Nespor and Vogel suggest may relate to the location of new information within the dominating intonational phrase.) If, in this case, the final phonological phrase still shows the greatest magnitude of final lengthening, this would suggest that the effect relates to a higher-level domain. Secondly, the locus of the effect may provide information about the possible domain: higher-level constituents may not only show lengthening of greater magnitude but may also show a more extensive locus of lengthening. If final lengthening effects are observed with consistently different loci, this suggests the existence of different domains. A Dutch study by Cambier-Langeveld (2000) examining this is discussed below in relation to the phonological utterance.

Studies which have attempted to find phonological-phrase-final lengthening in English, in the context of hierarchical lengthening effects, are discussed in Chapter 4; the experiment reported there examines the locus of final lengthening in English in relation to utterance-medial and utterance-edge durational effects.

The Intonational Phrase

Although Nespor & Vogel identify the phonological phrase with one particular suprasegmental effect—final lengthening—they do not suggest any particular correspondence with intonational phenomena. Major and minor phonological phrases as defined by Selkirk (1986) are identified as domains of intonational phenomena in other languages: the major phrase is required to characterise tonal phrasing in Xiamen and Shanghai Chinese (Lin 1993); the minor phrase (called the accentual phrase by Beckman & Pierrehumbert 1986) appears to correspond to the domain of occurrence of pitch accent in Japanese. It is not clear, however, whether both of these levels are required in English. As Wightman *et al.* (1992:1709) observe: "In Japanese, [Beckman & Pierrehumbert 1986] find clear evidence for an accentual phrase as a simple grouping of words, but in English, they find the evidence for justifying it much less compelling, although it is clearly possible to define such a unit."

As its name suggests, the intonational phrase is, in contrast, based upon the domains over which intonation contours extend, where a complete intonational contour includes, at least, a nuclear pitch accent, a phrase tone and a boundary tone, with optional prenuclear accents (see, for example: Ladd 1996). Nespor & Vogel also identify the ends of intonational phrases with the position of potential rule-governed pauses in English, although Hayes (1989) says that the utterance is the domain of non-hesitation pauses. The syntactic information used to define IPs is very general, although certain constructions form obligatory intonational phrases, including parentheticals, nonrestrictive relative clauses, tag questions, vocatives and expletives (Nespor & Vogel say that these are all strings that are in some way external to the root sentence). Beyond this, any number of phonological phrases within the same utterance may be dominated by a single intonational phrase, or a number of intonational phrases may each dominate smaller groups of phonological phrases: "The use of n-ary branching trees avoids the problems [associated with] binary branching trees. That is, since there is no additional constituent structure between [IP] and Φ, restructuring may group any sequence of Φs into smaller [IPs], as long as the division respects the syntactic and argument structure conditions as well as the general timing conditions..." (Nespor & Vogel 1986:205). These timing conditions are identifiable with the factors discussed in Section 2.2: the length of phrases, speech rate and style; additionally, certain "semantic prominence relations" may also affect restructuring at this level.

Some researchers have questioned the inclusion of the intonational phrase within prosodic hierarchies largely motivated by segmental phonological phenomena. Selkirk (1990:195) refers to it as an "honorary constituent of syntax", because it is "more closely tied in to the semantic properties of sentences" than other prosodic constituents, and states that she is "reluctant to accept the existence of a prosodic structure in general solely on the basis of the existence of intonational phrases." Gussenhoven (1992) goes further and suggests that intonational constituents—which he calls "association domains"—actually form a distinct class from prosodic constituents, and do not consistently map to any particular prosodic constituent. A relevant question is whether Nespor & Vogel are correct to make the connection, as discussed above, between segmental phonological effects and suprasegmental "phonetic" effects at a lower level, by identifying the phonological phrase as the domain of final lengthening. If this is justified, then given that there is no doubt that the ends of intonational phrases—through whatever mechanism—show final lengthening, all that would be required to make a compelling case for including both constituents within the same representation would be to show that the former are exhaustively dominated by the latter—that the boundaries of intonational phrases only occur at the boundaries of phonological phrases.

The Phonological Utterance

The utterance is discussed above in relation to final lengthening at the end of syntactically-defined constituents: the definition of the utterance given there as a pause-delimited and relatively self-contained stretch of speech is also a reasonable working definition of the phonological utterance. In practice, the utterance very often coincides with the syntactic sentence: as defined by Nespor & Vogel, the utterance comprises all the intonational phrases dominated by the highest node in the syntactic tree. Utterance-restructuring is possible, however: two short, temporally-adjacent sentences may comprise a single utterance when there exists "a syntactic relation (ellipsis, anaphora) and/or a positive semantic relation (and, therefore, because) between the [utterances] in question" (Nespor & Vogel 1986:244). This process is optional, depending on style and rate of speech amongst other factors.

As mentioned in Section 2.3.1, Nespor and Vogel say that the utterance is the domain, in certain dialects of English, of r-insertion. This may provide evidence of utterance restructuring between sentences; for example, in:

Let's take the car. I'm exhausted.

an /ɹ/ may be inserted between the two sentences if they comprise a single utterance. Shattuck-Hufnagel & Turk (1996) discuss such observations in relation to flapping of /t/ in American English, proposed by Nespor & Vogel to be another utterance-span effect. They say that information is lacking about the intonational contours of utterances, which makes it difficult to identify their occurrence. Between-sentence effects such as flapping, and by the same argument r-insertion, may be a result of two closely-linked sentences being realised as one intonational phrase and thus associated with that domain rather than the utterance.

In experimental investigation of speech timing, sentences are very often spoken in isolation, and by definition comprise a single utterance. Thus sentence-final lengthening may be readily equated with utterance-final lengthening. The problem of identifying final lengthening effects with a particular domain is discussed above in relation to the phonological phrase; here it may be asked whether sentence-final lengthening actually relates to a phonological constituent identifiable as the utterance, or to subordinate constituents such as the intonational phrase or the phonological phrase. As stated above, the locus of lengthening at the end of different constituents may help to resolve this question. Cambier-Langeveld (2000) does not find strong evidence in Dutch of lengthening at the end of phonological phrases, or at least, no greater lengthening than at the end of prosodic words. She does, however, find final lengthening at the end of intonational phrases and utterances, where the locus of the effect is consistent between the two types of boundary: the final syllable rhyme, plus the penultimate syllable rhyme only in the case that the final syllable contains a reduced vowel. For their part, Nespor & Vogel suggest that patterns of sentence-final intonation and lengthening indicate that the last intonational phrase of an utterance is the strong one, but what these distinctive suprasegmental patterns are is not specified. As discussed in Section 2.2.1, variations in final lengthening between domains in English speech remain uncertain; this issue is discussed further in Chapter 4 and Experiment 2 looks for evidence of final lengthening at utterance-medial and utterance-final word boundaries.

2.4 Lexical stress

2.4.1 The concept of speech rhythm

As mentioned in Chapter 1, the term "rhythm" is used here to mean the pattern of alternation of strong syllables (carrying primary or secondary lexical stress) and weak (unstressed) syllables. The concept of "speech rhythm", whether under this restricted definition or when more widely applied to speech timing phenomena, is both prevalent and elusive. It worthwhile considering briefly what it means to assert that speech is a rhythmical activity, before examining the durational consequences of such an assertion.

Rhythm implies the repeated occurrence of events marked by common features. It is an essentially temporal phenomenon, but not exclusively auditory. Visually-observed movement may be said to possess rhythm through the repetition of some sequence of events in time: the sight of a person walking or running and of the wings of a bird in flight, convey a clear sense of rhythm. Visually-observed rhythms in many instances have a particular quality of flow; that is, a sequence of events may be seen to recur without there being any sense of a privileged point in the sequence. There is no position which may be described as the onset, although one may be arbitrarily assigned. This is similarly true from the point of view of the person observed at a rhythmic activity: one may run, with a greater or lesser degree of rhythm experienced proprioceptively, without any awareness of a point in the sequence of movements which mark the beginning of the next repetition. In other cases, rhythm may arise from a sequence of manifestly discrete events, each with a clear onset in time: a heartbeat proprioceptively experienced is rhythmic and discrete.

In the auditory domain, rhythm seems almost to require this quality of discreteness: it is difficult, if not impossible, introspectively to characterise an auditorily-experienced rhythm without making reference to the onset, or some other specific point, of each event in a sequence. Below a certain rate of occurrence, each repetition of a regular acoustic event is experienced as an equal contributor to the sequence. A heartbeat heard at a normal resting frequency is a discrete sequence of equally salient or prominent thuds. Above a certain frequency however, the auditory system imposes a structure on a sequence of acoustic events such that the events are perceived as discrete groups, with the first in each group possessing the greatest degree of salience. A very fast heartbeat may be heard as a sequence of paired beats, with the first in each pair being the more prominent.

This grouping tendency, and the conferring of salience on the first event in a group, is clearly a particular feature of the auditory information processing system, discussed by Woodrow (1951:1232):

> By rhythm, in the psychological sense, is meant the perception of a series of stimuli as a series of groups. The successive groups are ordinarily of similar pattern and experienced as repetitive. Each group is perceived as a whole and therefore has a length lying within the psychological present.

The size and perceptual experience of these groups depends to some extent on the rate at which successive stimuli occur. When acoustic events occur in rapid succession, it would seem that the brain devotes most of its processing power to a certain fraction of the events, every second or third or fourth occurrence, the others being perceived largely as inferred and degraded copies of the salient instances. There is an indisputable sense in which the auditory sensation of the prominent events is different from that of the others: the sound of a rapidly ringing bell creates the ineluctable percept that certain of the chimes are different from the others, and introspection is required to demonstrate to oneself that they must in fact be identical. Auditory sampling is reflected in onomatopoeic descriptions of regular acoustic events: the chime of a slowly tolling bell is either *ding* or *dong* according to pitch; at a faster rate it is *ding-dong*; the chime of a rapid alarm bell is *ding-a-ling-a-ling*. Auditory sampling clearly has an upper limit as well: above a certain rate of occurrence, processing is no longer devoted to individual events and the sound is experienced as a continuum, although at relatively low frequencies individual events may be inferred if not perceptually isolated.

This sampling property of the human auditory processing system seems to have influenced the evolution and development of acoustic production systems. The concept of prominence-headed groups of sounds is clearly fundamental in music. Furthermore, the regularity of occurrence of these prominences and disruptions to this regularity are planned by the composer and can be interpreted by the listener. Similarly in speech, certain parts of an utterance are more prominent than others, with salience conferred in production and, it seems likely, in perception as well. Given the existence of these patterns of prominence, it is tempting to believe that, as in music, prominences are produced at regular intervals of time, and disruption from this regularity can be interpreted by the listener. The empirical evidence for such regularity is lacking, however. In particular, the rhythmical foot, comprising a initial syllable of high prominence and a number of subsequent less prominent syllables has often been asserted as a phonological unit with consequences for the timing of speech production (for example: Lehiste 1977). The evidence reviewed below indicates that, as traditionally defined, there is remarkably little support for the concept of the foot that cannot be better accounted for by reference to other linguistically-motivated structures.

Different conceptions of the metrical foot are examined in Section 2.4.2; durational evidence for the foot as a linguistic unit is discussed in Section 2.4.3; the evidence for a localised durational effect of prominence is presented in Section 2.4.4; the non-durational evidence for the rhythmical foot as a linguistic unit is considered in Section 2.4.5.

2.4.2 Isochrony and the rhythmical foot

Prominence-headed constituents

There are at least as many conceptions of prominence-headed constituents as there are postulated degrees of prominence. Firstly, strong syllables may be defined as those that contain a full vowel and carry some degree of lexical stress, and weak syllables as unstressed and containing a reduced vowel. By these definitions, the most basic prominence constituent into which an utterance can be (almost) exhaustively divided is the cross-word foot: this begins at the onset of a strong syllable and contains all subsequent weak syllables up to the onset of the next strong syllable[19]. Where an utterance begins with an anacrusis—one or more weak syllables before the first strong syllable—under this definition those syllables are extra-metrical, that is, not contained within any cross-word foot.[20]

The cross-word foot so defined corresponds closely to poetical concepts of feet as composed of strong and weak syllables, in most cases one strong syllable and a number of weak syllables. Poetical feet may be left-headed or right-headed and are classified according to the number of syllables they contain: for instance, trochees and dactyls are left-headed disyllabic and trisyllabic feet; iambs and anapaests are right-headed disyllabic and trisyllabic feet. The types of poetical feet and their use in verse are essentially matters of convention, however: to postulate both left-headed and right-headed cross-word feet in the description of natural speech would seem to confound description of an utterance's rhythmical structure rather than assist it. A single word may be described as an iamb or an anapaest, but in describing an utterance,

[19]There are other prominence-delimited constituents which have been called cross-word feet, as discussed below.

[20]Some metrical theories posit silent stresses to deal with anacrusis. For instance, in Abercrombie's (1965) account of the phonetics of verse structure, there are held to be silent stresses as the head of feet before line-initial—but not stanza-initial—unstressed syllables.

it suffices to identify the strong syllables and count the number of unstressed syllables headed by each. Most researchers choose to use a left-headed foot in metrical description, sometimes by convention. Left-headed feet have also been hypothesised to have specific consequences for timing, as discussed below.

The cross-word foot may be regarded as the most basic rhythmical unit, as it reflects the lowest level of prominence contrasts (leaving aside the difference between primary and secondary lexical stress, which, as discussed in Chapter 1, does not seem to have significant timing consequences). Thus, the cross-word foot so defined is referred to here as simply the foot, except where clarification is required, with the number of syllables it contains specified where appropriate. There is another proposed constituent of speech, the within-word foot, also composed of a strong syllables and a number of subsequent weak syllables; as its name implies does not extend beyond word boundaries. It is fundamentally non-exhaustive, as it excludes word-initial unstressed syllables and most function words, and as such is not appropriate for a complete description of speech rhythm. It has been considered as the locus of durational effects such as final lengthening and accentual lengthening, as discussed elsewhere in this chapter.

Certain strong syllables in speech are pitch accented, and perceived as being more prominent than other strong syllables by virtue of that fact. The specific durational consequences of pitch accent are considered in Section 2.5, but is appropriate here to consider the possible existence of a prominence constituent headed by a pitch-accented syllable. This may be defined as beginning with the onset of a pitch-accented syllable and continuing until the onset of the next pitch-accented syllable, encompassing all intervening syllables. This unit—another type of cross-word foot—is sometimes referred to as the Abercrombian foot (after David Abercrombie; for example: Abercrombie 1965[21]). As in the case of the cross-word foot, there may be extra-metrical syllables before the onset of the first Abercrombian foot in an utterance. A prominence-headed constituent may also be postulated from the onset of one nuclear accented syllable up to the onset of the next; the durational evidence for this will be discussed briefly below.

The isochrony hypothesis

All the prominence-headed constituents discussed above have at some time been considered as potentially isochronous units in speech. Early formulations of isochrony, such as that by Classe (1939), are applied by Pike (1945) to a typological distinction between languages such as English, called "stress-timed", and languages such as French,

[21] An alternative interpretation of Abercrombie's formulation is cross-word feet headed by primary lexical stresses. Clearly, primary lexical stress is often, but not always, accompanied by pitch accent.

called "syllable-timed". In stress-timed languages, it is asserted, lexically-stressed syllables occur at regular intervals of time; by contrast, in syllable-timed languages, syllables themselves occur at equal time intervals. Thus, in English the durations of cross-word feet will tend to be equal, however many syllables they contain.

A formal statement of the isochrony hypothesis could be:

Rhythmical feet are of equal duration at constant speech rate.

The mechanism Pike proposes for the timing adjustment required to compensate for the variation in the number of syllables within feet is a compression of unstressed syllables in proportion to their number. Others consider the effect in terms of the shortening of a lexically-stressed syllable as unstressed syllables are added (for example: Jones 1942–43). According to either mechanism, isochrony is a domain-span compression process: an inverse relation between the number of syllables in a foot and the duration of constituents of the foot, independent of their position. Attempts have been made to demonstrate isochrony for different types of rhythmical feet, and as outlined below, have shown the hypothesis to be false.

2.4.3 Experimental studies of isochrony

In the widely-cited paper *Isochrony Reconsidered*, Lehiste (1977) presents a review of the evidence for isochrony of rhythmical feet in speech. As discussed above, prominence-headed constituents may be postulated for each distinct level of prominence in speech. The effect on speech timing of a number of these constituents has been experimentally examined.

Lea (1974) is cited by Lehiste as looking at the relationship between the number of unstressed syllables between lexical stresses and the duration of the interstress intervals. Lea does not find evidence for foot-level isochrony: the average interstress interval appears to increase almost linearly with the number of unstressed syllables in the foot. Lehiste also discusses her own studies of interstress intervals, which use "relatively more homogenous material" (Lehiste 1973; Lehiste 1975): sentences comprising four cross-word feet of varying lengths. As in Lea's study, there is a regularity of duration for feet of the same numbers of syllables, but there are marked differences in the duration of feet of different length, even controlling for position of the foot within the sentence. Lehiste indicates one instance in this study where the same speaker has a difference in mean durations of monosyllabic and disyllabic feet of 133 ms, despite the feet occurring in the same sentence-medial position.

Bolinger (1965) measures intervals between pitch accents, corresponding to the length of the Abercrombian foot as defined in Section 2.4.2 above. As cited by Lehiste, he fails to find evidence for isochrony in the timing of this putative constituent. Of the 53 Abercrombian foot durations measured, 13 are approximately twice as long as shortest such foot in the study. Lehiste also cites the study by Shen & Peterson (1962) of intervals between "primary stresses"—meaning the primary stress in each sentence—effectively corresponding to a prominence constituent headed by nuclear pitch accents. Their materials, containing sentences of varying length, elicit from all speakers very large ranges in the intervals between primary sentence stresses, with no detectable central tendency.

The review of evidence presented by Lehiste makes it clear that prominence-headed constituents are not produced isochronously in English speech. She nevertheless makes the claim that foot-level isochrony exists (that is, isochrony in the cross-word foot headed by lexical stresses), and suggests that deviations from isochrony are not perceived in the majority of cases. When large enough to be perceived, the deviation in foot-level isochrony may interpreted by the listener as evidence of the presence of a syntactic boundary.

Evidence for her assertion that deviations from isochrony are not perceived is presented from an experiment (Lehiste 1975) which is the perceptual analogue of her studies which fails to demonstrate isochrony in production. She plays recorded utterances composed of four cross-word feet to listeners, and finds that they have "considerable difficulty" in identifying which of the feet is the longest or shortest in any sentence. In contrast, the equivalent task using synthesised materials (clicks separated by noise with the interstress intervals of the spoken sentences) is found to be much easier.

Lehiste proposes this finding as evidence of "perceptual isochrony". She says that "if listeners cannot identify the actually longest or shortest measures in spoken English sentences, the measures must seem to them to have equal duration; if you cannot tell them apart they must be alike. Isochrony would then be a perceptual phenomenon" (Lehiste 1977:256). It is not clear that this conclusion can be drawn: it is one thing to find that a task is difficult, and another to prove that its terms of reference have some perceptual reality. The result could equally well be taken to indicate that listeners do not attend to interstress intervals because they are not linguistically important. If the outcome can be taken as support for two contradictory conclusions, their validity must be questioned.

Perceptual isochrony is questioned by Scott *et al.* (1985), who find that both French and English listeners tend to regularise interstress intervals found in speech more than the beats of a simple non-speech stimulus when they imitate the rhythm of the stimulus. If stress-timing were a property of English, and French were syllable-timed, one would expect to find differences between French and English performance on this task, as French speakers should not have a bias towards perceptual isochrony in speech. Furthermore, both sets of speakers also regularise the beats on a third version of the task, using non-speech materials with the acoustic complexity of speech. This outcome raises "the possibility that subjects are not actually doing anything very interesting at all—that they are simply exhibiting a response bias toward evenly spaced taps when the task becomes difficult" (Scott *et al.* 1985:161).

Lehiste (1977) acknowledges that the difference in performance between speech and non-speech stimuli might be a gradient distinction; she further suggests that the determination of the just-noticeable-difference (JND) for non-speech on her experimental task might indicate a minimum value for the durational difference required between feet for deviations from isochrony to be observed. Using the same design, with four click-delimited intervals of noise, she systematically varies their duration so that three of the four are isochronous and the fourth, variously located, is longer or shorter. Her finding is that the amount of durational variation required to distinguish the non-isochronous foot from the others ranges from 30 ms to 100 ms. This is sometimes cited as evidence that the JND for duration in speech perception ranges from 30–100 ms (for example: Couper-Kuhlen 1986:54); this is not claimed by Lehiste, however, and it is clear from the task described that this conclusion should not be drawn. Firstly, the task uses non-speech stimuli. Because the task is likely to be easier with sounds of less complexity than speech, it could be argued that the JND represents a minimum threshold for speech. Listeners, however, have expectations based on linguistic experience about the relative durations of speech sounds, and may actually be more sensitive to durational variation which is linguistically interpretable: for instance, Klatt & Cooper (1975) suggest JNDs for individual speech sounds vary according to position within syntactic structures such as words and sentences. The durational range for the non-speech task in Lehiste's experiment only relates to extended stretches of sound designed to correspond to feet. There is no strong evidence that listeners actually pay attention to the duration of feet, or even, as discussed in Section 2.4.5, that the concept of foot has any psychological validity at all.

Lehiste contends that listeners do pay attention to variations in foot duration, provided they are above the threshold of perception: "In English, lengthening of interstress intervals is frequently used to signal the presence of a syntactic boundary" (Lehiste 1977:253). She cites O' Malley *et al.* (1973) as suggesting that boundaries in speech are marked by acoustic cues including lengthening, but takes issue with them because they locate "their boundary signals at given points in a linear sequence, without relating them to the general rhythmic structure of the utterances". This comment implies that it is not the precise location of the lengthened segments that is important but the fact of lengthening of the interstress interval. If the domain-and-locus description of systematic durational variation is valid, however, it is precisely the "given point" of the lengthened segments that is important, rather than differences in foot duration. As Lehiste herself points out, perceptible lengthening of an interstress interval could be used to signal anything and, if it used in speech to signal a syntactic boundary, lengthening alone cannot indicate anything else, at least where the possibility of a boundary exists however, lengthening of the rhyme of a word-final syllable is less ambiguous as a cue, even in the absence of other suprasegmental cues such as fundamental frequency variation, and allows durational variation in a different locus to perform a different function. Lehiste further proposes that polysyllabic shortening and the shortening of consonants in clusters are evidence of isochrony at work, using a foot-span compression process to explain phenomena which might be better accounted for as localised durational adjustments within the word or syllable.

Lehiste puts forward perceptual evidence in support of the view that lengthening of the interstress interval signals a boundary, referring to the study by Lehiste *et al.* (1976) thus: "disambiguation was produced solely by increasing the length of the interstress interval, and the results of our study show that this is indeed a sufficient cue for signalling the presence of a boundary" (Lehiste 1977:262). It could be argued that although such lengthening is *sufficient*, specifically-located lengthening is a better cue. Furthermore, it is far from clear that the study she cites does, in fact, demonstrate that interstress lengthening is a boundary cue. In particular, she says that in Lehiste *et al.* (1976) the "interstress interval was increased by increasing the duration of each sampling period by the same factor; the durational relationships of the segments to each other remained the same." In fact, in the 1976 study, the durational manipulation is not reliably placed in the interstress interval. Lehiste *et al.* identify nine sentences in which the durational manipulations produced a significant response bias; in only one instance is the only durational manipulation a variation in the length of the cross-word foot (underlined):

The president of the university's committee on education policies came.

In five cases, two adjacent feet are lengthened, as in:

The old men and women stayed at home.

In three cases, part of the sentence is lengthened, whilst another part is shortened, such as:

I know more beautiful women than Mary.

where the underlined words were both and alternately lengthened and shortened.

It is possible to reinterpret these results in a variety of ways: for instance, in the *old men and women* example, as a phrasing effect obtained by delimiting a constituent with a local speech rate variation. The main point here is that the results of Lehiste *et al.* (1976) do not support a theory of perceptual isochrony in which perceived deviations from isochrony indicate the presence of a syntactic boundary (or indeed, a prosodic boundary).

Finally, to suggest that isochrony is nevertheless present as an "underlying tendency" is equally invalid, implying as it does that stressed syllables would be regularly spaced if not for all the other durational factors clouding the picture. Lehiste offers a quotation which expresses this view: "The implication is that speech attempts to be very rhythmic, but fails because of durational constraints" (Barnwell 1971:88). This view does seem insupportable under the weight of subsequent evidence, both the continual failure to observe isochrony and the replication of experiments which demonstrate other durational effects such as phrase-final lengthening, word-initial lengthening and accentual lengthening. Firstly, although the question of the precise characterisation—within, for instance, a domain-and-locus description—of these well-attested patterns remains open, the patterns themselves have been found to be statistically significant under a variety of conditions, in contrast to isochrony. Secondly, stress-adjacent lengthening is a prominence-related durational effect, discussed in the following section, which is be observed even in the presence of another structural effect—final lengthening—indicating that the presence of one type of timing process need not preclude the observation of another. Finally, segmental restrictions on durational variation are very minor in comparison with structural influences; most phones may be produced over wide range of durations, and so segmental variation in the composition of feet cannot be held responsible for failures to observe isochrony.

2.4.4 The stress adjacency effect

The relative prominence of two adjacent syllables has consequences for speech timing. Bolinger (1965) observes that a full vowel syllable followed by another full vowel syllable is longer than when followed by a reduced vowel syllable. The presence of a full vowel in a syllable almost always indicates the presence of (primary or secondary) lexical stress; likewise, a reduced vowel in a syllable is a reliable indicator that the syllable is unstressed. The observation may be restated thus:

The duration of a stressed syllable is greater when
the following syllable is stressed than when it is unstressed.

The effect is in the direction predicted by the isochrony hypothesis, where the interpretation would be that a disyllabic foot is compressed (or the monosyllabic foot expanded) in order to equalise foot durations. Although, as demonstrated in the previous section, there is very little empirical support for the isochrony hypothesis, there is strong evidence for what may be termed the "stress adjacency effect". As the evidence reviewed now suggests, this seems to be a strictly rhythmical effect, apparently independent of the boundaries of words and higher-level syntactic or prosodic constituents.

The study of Van Lancker *et al.* (1988) is designed to explore the influence of word boundary and pitch accent upon the occurrence of the effect described by Bolinger. Example sentences designed to test the effect of accent include the following, where the measured syllable (underlined) is followed in the first of each pair by another full vowel syllable and in the second by a reduced vowel syllable. These may be designated the full-full and full-reduced conditions respectively.

Accented condition

They <u>rot</u>ate both ways... vs ... *They <u>wrote</u> it both ways.*[22]
<u>John</u> cracked the walnuts... vs ... *<u>John</u> corrected the papers*

Unaccented condition

They describe the elec<u>tron</u> lineup... vs ... *They describe the elec<u>tron</u> alignment.*

There are a number of potential problems with the design as illustrated by these sentences. Firstly, it is apparent from the description of the experimental design that pitch accent is equated with primary lexical stress in their materials, which raises the possibility that the presence of accent was assumed rather than detected in their recordings. Secondly, the relationship of the measured syllables to syntactic boundaries is not always constant within sentence pairs: if a boundary of some kind is present after the target syllable in one instance—such as after /roʊt/ in *wrote it* but not in *rotate*—the stress adjacency effect may be masked or exaggerated by domain-final lengthening. Thirdly, the immediate phonetic environment of the measured syllable is often quite different within sentence pairs, as acknowledged by Van Lancker *et al.* Finally, in common with many speech timing studies, the durational effects of syllable number within particular constituents presents a potential confound. In the first sentence pair above, the measured syllable is in a disyllabic word in the full-full condition and in a monosyllabic word in the full-reduced condition; polysyllabic shortening would predict a durational change in the opposite direction to stress adjacency. In the

[22]This an American study, hence the first syllable in *rotate* has primary stress. It may be noted that the syllabification implied here by comparison with *wrote it* is contrary to the principle of maximal onset: that is, the authors are presumably intending to compare syllables with the same structure in the two sentences, but /ɹoʊt/, which is clearly a syllable in *wrote it*, may not be a syllable in *rotate*. Consultation with an American speaker suggests that the syllabification is also contrary to the evidence from flapping: the medial /t/ in *rotate* is unlikely to be flapped, unlike the /t/ in *wrote it*.

other sentence pairs, the sentence in the full-reduced condition has an extra syllable: a domain-span effect operating over some constituent which includes these syllables would predict a durational difference in the same direction as the stress adjacency hypothesis. This point is explored further in relation to studies of polysyllabic shortening in Chapter 3.

Although some of these possible confounds are present in many of the sentence comparisons, there is sufficient diversity of materials to give the results some validity, particularly as confounding factors would favour the null hypothesis in some cases: for example, any effect of word-final lengthening or polysyllabic shortening in the *wrote it* vs *rotate* sentence pair would tend to mask the effect predicted by the experimental hypothesis.

Van Lancker *et al.* (1988:343) report that "a target syllable followed by a full syllable is longer than one followed by a reduced syllable, whether that target is accented or unaccented and regardless of word boundaries which may intervene between the two syllables." The target syllable durations shown in Table 2.3 are for the accented and unaccented conditions, pooled across the word boundary conditions (thus, in some cases the measured syllable is followed by a word boundary and in others not).

	Full-full condition	Full-reduced condition
Accented	177	175
Unaccented	155	133

Table 2.3: Reported durations (ms) of the measured syllables in the study by Van Lancker *et al.* (1988). The durations shown here have been converted from mm of spectrogram, the unit reported by the authors.

Comparison of the two rows of the table shows the lengthening effect of pitch accent, discussed in Section 2.5 below; given that many of the "unaccented" syllables are likely to be found in words containing a previous pitch accent, it is probable that these will also be subject to some degree of lengthening, as suggested by studies such as Turk & White (1999), so the magnitude of the effect may be underestimated here. Comparison of the two columns of the table shows the effect of the following syllable type upon measured syllable duration: it is immediately apparent that the magnitude of stress-adjacent lengthening for unaccented syllables is much greater than that observed for the accented syllables, with 17% lengthening in the former condition and only 1% lengthening in the latter. There is, however, no suggestion in the reported statistical analysis of an interaction between the accent condition and the syllable type condition. Unless due to a misprinted figure, this difference—even if not significant in this analysis—seems worthy of comment.

The findings of Rakerd *et al.* (1987) agree with those of Van Lancker *et al.* in finding that the stress adjacency effect is not blocked by word boundaries. This study goes further in looking systematically at the influence of larger boundaries, with the effect examined across major syntactic boundaries (MSB, corresponding to a noun phrase/verb phrase boundary) and within the same noun phrase (WPH). Rakerd *et al.* record subjects reading aloud sentence pairs such as:

WPH *The strong peach light was unpleasant.* vs *The strong peach delight was unpleasant.*

MSB *The young duke armed his subjects.* vs *The young duke disarmed his subjects.*

The measured syllable (underlined) is a stressed syllable followed by either another stressed syllable or an unstressed syllable; between the two syllables there is either a minor or a major syntactic boundary. Rakerd *et al.* report substantial final lengthening in preboundary syllables, suggesting that the major syntactic boundaries are prosodically realised in many cases, but there are some potential problems with the experimental design as outlined. Firstly, as in the study by Van Lancker *et al.*, the phonetic environment of the measured syllable is not kept constant: segmental effects on duration are likely to be small and unsystematic, however. Secondly and more importantly, there is no control of the distribution of pitch accents and no account given of their occurrence: inspection of the materials suggests that the measured syllables are likely targets for pitch accent in many cases.

The results show that "shortening of a stressed target syllable by an unstressed syllable across a word boundary was not significantly less at an NP/VP boundary than elsewhere, even though the target word shows substantial domain-final lengthening at the boundary" (Rakerd *et al.* 1987:152), supporting and extending the findings of Van Lancker *et al.* It may be noted, however, that the observed stress-adjacent lengthening is very small: a mean of 7ms, where the duration of the measured syllable in the non-adjacent condition was in the range 135ms–190ms. In comparison, Van Lancker *et al.* find a mean of 22ms stress-adjacent lengthening in their unaccented condition, with a non-adjacent mean duration of 133ms.

Taking both studies into account, the simplest explanation of the smaller stress adjacency effect observed by Rakerd *et al.* is that the measured syllables in their materials are, in many cases, pitch accented; furthermore, that the attenuation of stress-adjacent lengthening in the presence of accentual lengthening suggested by Van Lancker *et al.*'s data is a real effect. This conclusion is, of course, speculative without access to the original recordings, so the hypothesised interaction between stress adjacency and pitch accent remains unproven at present.

Fant *et al.* (1991) present the stress adjacency effect as the sole major timing consequence of the rhythmical organisation of speech, summarising the situation thus:

> A foot's duration increases on the whole proportionally to the number of unstressed phonemes or syllables contained within it. The increment per segment is close to the average duration of unstressed segments. Foot shortening effects as a trend towards isochrony have been much discussed in the literature [...] The main effect appears to be in the step from none to one following unstressed syllables in the foot. However, in our experience, these effects are marginal and not sufficient as a basis for a theory of "stress timing" (Fant *et al.* 1991:84).

If this analysis is correct, then one may ask whether the concept of the foot is required to characterise the stress adjacency effect. Fowler (1990) suggests that it might be, saying that "the lengthening of a stressed syllable at the right edge of a foot [...] may count as a sort of preboundary lengthening." Likewise, Rakerd *et al.* (1987) investigate what they call "foot-level shortening", but the experimental evidence they provide relates only to the monosyllabic vs disyllabic foot effect highlighted by Fant *et al.* (1991).

An alternative characterisation would eliminate the foot altogether and describe the effect as arising purely from the adjacency of stressed syllables. In order to distinguish this "stress adjacency" account from the "foot-level" account, certain predictions might be made. Firstly, only the stress adjacency effect could predict any lengthening of the second of two adjacent stressed syllables. Secondly, only the foot-level effect could predict any shortening in the disyllabic vs trisyllabic foot comparison, for instance, in the case that the effects of foot-finality diminished gradually with distance from the end of the foot.

As seen above, Fant *et al.* (1991) suggest that the latter effect is not observed. In contrast, Rakerd *et al.* (1987:154) cite Fowler (1981) as providing evidence that "Two unstressed syllables shorten a stressed syllable more than does one", and furthermore that the shortening effect is confined (or largely confined) to the first stressed syllable. Examination of the method used in Fowler's study shows, however, that it cannot be used to support the assertion made by Rakerd *et al.* The experiment uses reiterant words composed of stressed syllables (/si/, /sʌ/) and unstressed syllables (/sʌ/) in a fixed frame sentence. The durational effect on stressed syllables of the number of preceding and following unstressed syllables is reported. The duration of stressed syllables is found to be inversely related to the number of unstressed syllables which follow it, an effect Fowler calls "anticipatory shortening"; limited evidence is found for an inverse relationship between stressed syllable duration and the number of unstressed syllables preceding it (a hypothesised effect Fowler calls "backward shortening"). The problem with the interpretation of these results is that the recording procedure requires subjects to produce the sentence "in time" with a metronome; that is, stressed syllables are placed on regularly-occurring beats. The procedure is used so that the coarticulatory correlates of a large and reliable anticipatory lengthening effect can be examined, and Fowler notes that "metronome pacing may strongly inflate the anticipatory shortening parameter and may even decrease the backward shortening parameter" (Fowler 1981:49). Given this, Rakerd *et al.* are mistaken in drawing any conclusions from this experiment about the durational effect on the second of two stressed syllables of intervening unstressed syllables, and about the consequences for both stressed syllables of more than one unstressed syllable intervening between them.

The experimental design used by Fourakis & Monahan (1988) apparently allows direct consideration of these questions. They measure the duration of pairs of stressed syllables when separated by either one or two unstressed syllables. The sentences are constructed with the template:

(your) ma<u>chines</u> will (soon) <u>pass</u> (at) my <u>desk</u> today

where the measured stressed syllables are underlined and the presence of the bracketed words is systematically varied.

At first sight, the results seem contradictory, offering apparent support for both the foot-level hypothesis and the stress-adjacent hypothesis: for the former because there is a shortening effect in the disyllabic vs trisyllabic foot context; for the latter because this shortening affects the following stressed syllable as well as the preceding stressed syllable.

There are, however, a number of confounding factors which may contribute to this puzzling state of affairs. Firstly, the presence or absence of additional stressed syllables affects not only foot structure but the relationship of the measured syllables to phrase boundaries—and may affect the placement of those boundaries—as well as the length of the utterance and other constituents. Secondly, lexical stress may once again be confounded with pitch accent: at least some of the measured stressed syllables are likely to be pitch accented. Indeed, it is possible that pitch accents are what Fourakis & Monahan intend to delimit their feet: certainly, words such as *will* and *soon* are not always lexically unstressed.

Finally deciding between the competing stress adjacency and foot-level hypotheses is beyond the scope of this dissertation, although some of the possible confounding factors in previous studies, such as the effect of utterance length on stressed syllable duration, are addressed experimentally. What seems clear, however, is that the adjacency of lexically-stressed syllables lengthens the first syllable compared with when one or more unstressed syllables are interposed. It seems a strong possibility that this stress adjacency effect is greater when the first syllable is unaccented. There is little unambiguous evidence for the application of foot-based timing rules beyond this localised effect however.

It appears rather awkward to define the stress adjacency effect in the domain-and-locus terms that may be applied to durational effects within syntactic or prosodic constituents. Perhaps the most precise characterisation would be to say that the domain is the spondee—a foot of two strong syllables—and the locus is the first of these two syllables. Such a domain would seem to belong to a separate representation of speech timing from the constituent-based structures discussed in Sections 2.2 and 2.3. Indeed, as suggested by Rakerd *et al.* (1987), the stress adjacency effect appears to operate across constituent boundaries and independently of the timing effect of these boundaries.

2.4.5 The rhythmical foot as a linguistic unit

Given the lack of evidence for the timing consequences of rhythmical foot structure, it is worth considering briefly whether there is any other evidence for the foot's linguistic existence: that is, do rhythmical feet constitute part of the grammar of English, or are they merely a descriptive tool best applied to restricted domains such as poetics?

Within the area of segmental phonology, there does not seem to be evidence of processes that require a domain delimited by the boundaries of the cross-word foot. Nespor & Vogel (1986) describe certain rules of English phonology as operating in a within-word foot domain, but they argue that segmental processes are regulated by a prosodic structure that is partially determined by syntax. It seems unlikely that a constituent which by definition ignores syntactic boundaries, such as the cross-word foot, could be accommodated within a framework utilising prosodic constituents, which are constructed with specific reference to syntax.

With regard to suprasegmental phenomena, within-word feet—but not cross-word feet—have been proposed as relevant to the description of final lengthening, but recent evidence suggests the locus of the effect may sometimes extend over more than one within-word foot (see Section 2.2.1). In intonational research, recent work on tonal alignment in Greek, Dutch and English, such as Arvaniti *et al.* (1998) and Ladd *et al.* (1999), has indicated the primary importance of the structure of the accented syllable—rather than larger constituents such as the foot or the word—in determining the location of intonation peaks and valleys. Rhythmical constituents do not appear to be relevant, although more research remains to be undertaken in this area[23].

[23] In studies carried out so far, the design of experiments has been influenced by the need to have adjacent pitch accents some distance apart, to avoid distortion of the intonation contour due to "tonal crowding". It would be interesting to vary the degree of tonal crowding in such experiments, in order to examine the hypothesis that there is some correlation between the distance from one accented syllable to the next—this corresponds to the Abercrombian foot, by one definition—and the position of the accent maxima and minima. An alternative hypothesis would be that tonal crowding only has consequences for alignment under conditions of adjacency; that is, when one pitch-accented syllable immediately follows another.

Although segmental and suprasegmental evidence for the rhythmical foot is lacking, there may be some usefulness in conceptions of prominence-headed constituents to describe the *distribution* of lexical stresses in English speech. In the prescriptive domain of poetic verse, for which conceptions of rhythmical feet were originated, there is rule-based alternation of strong and weak syllables and the choice of lexical items is in part conditioned by the demands of the particular foot structure of the type of verse. Furthermore, there are constraints upon the number of feet per line. Abercrombie (1965) alludes to this when he says that prose (and by implication, speech) is rhythmic but not metrical, but verse is both rhythmic and metrical: prose exhibits organisation at the level of the foot, whereas verse is organised in terms of feet, and also in terms of a superordinate metrical unit, the line. This is evident, for instance, in verse of iambic pentameters, such as[24]:

Was this the face that launched a thousand ships,
And burnt the topless towers of Illium?

in which the choice of words has been influenced by the requirement to alternate weak and strong syllables so as to form iambuses—right-headed disyllabic feet—and furthermore by the requirement that there be five such feet per line.

Abercrombie's assertion that prose (and also speech) are not metrical in the sense that he intends for poetry seems uncontroversial. Neither syntactic nor prosodic constituents require a certain number of feet: for example, noun phrases, verb phrases, and intonational phrases may all be well-formed as monosyllables. Syntactic phrases furthermore have no theoretical upper limit on their length. Prosodic constituents—especially higher-level constituents such as the intonational phrase—do tend to be limited in length, as outlined in Section 2.1: two small phonological phrases may restructure as a single larger phrase; potentially lengthy intonational phrases may be realised as two smaller phrases of approximately equal size. There is no evidence, however, that these processes make reference to the number of feet within constituents: although Nespor & Vogel (1986:203) refer to rules governing the restructuring of phrases as reflecting "a general abstract rhythmic organisation of speech into chunks of a more or less fixed temporal length", it seems clear that here the term "rhythmic organisation" is being used in the much broader sense of timing in general. The use of the term "abstract" seems questionable in this context, as surely the restructuring of phrases is a matter of speech performance rather than underlying representation, as suggested by their later assertion: "for both [intonational phrase] and [utterance] restructuring, what seems to be crucial is some average length in terms of timing, most probably due ultimately to physiological considerations" (Nespor & Vogel 1986:240).

[24]Christopher Marlowe (1604), *Dr Faustus*, Act 5, Scene 1.

Speech is therefore not metrical in the Abercrombian sense; furthermore, there seems no compelling reason to regard it as rhythmical in the sense that he intends. That is, there is no statistically-based support for the idea that feet themselves have some organisational significance in speech. The rhythmical pattern of speech is very largely determined by the interaction of lexical phonology and syntax: the words in an utterance are chosen on the basis of meaning and not for their pattern of strong and weak syllables[25], and how those words are ordered within a sentence is a matter of syntactic organisation. What arises from these processes has a manifest rhythm in the sense that there are alternations of stressed and unstressed syllables, but this rhythm is almost entirely an emergent property, contingent upon other processes that make no reference to lexical stress.

There are observable regularities in the rhythm that emerges from these lexical and syntactic processes, regularities that are to some extent language-specific. Scott *et al.* (1985) say that English "conspires" to place stresses at regular intervals. At the simplest level, lexical words—containing at least one stress—and unstressed function words tend to alternate: long collocations of function words or lexical words are rare; furthermore, lexical words of several syllables contain secondary stresses. Thus, as observed by Dauer (1983:58), "in a text of conversational English, the vast majority of interstress intervals contain from one to four syllables, with a limit of five [...] Counterexamples are difficult to construct in English." She contrasts this situation with Spanish, which allows only one stressed syllable per word, and where interstress intervals may contain as many as seven syllables.

The maintenance of this emergent rhythmical regularity is occasionally subject to processes which may be seen as entirely rhythmical. The occurrence of a long string of weak syllables often induces the placement of a stress upon one of the syllables, with the stress likely to be towards the centre of the "underlying" interstress interval. The phenomenon of stress retraction may also be seen an purely rhythmical process: when a word-final stressed syllable precedes a word-initial stressed syllable, the primary stress in the first word may be retracted, thus *thirˈteen men* may become *ˈthirteen men*. The existence of this effect is not uncontroversial, however. Cooper & Eady (1986) fail to find acoustic correlates of stress retraction in cases like the *thirteen men* example. Grabe *et al.* (1994) and Grabe & Warren (1995) likewise find that the two stressed syllables in contexts such as *thirteen* are not different in fundamental frequency range, duration and amplitude, except where the second syllable precedes a phrase boundary and/or carries a nuclear accent and is thus marked by greater lengthening and frequency range; furthermore, where the two syllables have equal acoustic prominence,

[25]There are statistical regularities of rhythm within the lexicon, such as the predominance of word-initial stressed syllables (Cutler & Carter 1987), but there is no evidence that word selection takes account of whether these regularities are or are not observed.

listeners cannot reliably judge which is stressed when they are excised from context. In related research, Shattuck-Hufnagel *et al.* (1994) analyse stress shift in terms of the likely placement of pitch accent within a word rather than variation in the acoustic correlates of lexical stress.

In their analysis of stress retraction, Shattuck-Hufnagel *et al.* refer to the tendency for the first pitch accent in a phrase to be placed as early as possible, with the final accent tending to be placed late. In an intonational phrase with, by default, an early accent and a late accent, the length of the phrase may favour the placement of additional accents: long phrases will have three, four or more accents, which will tend to be evenly spaced throughout the phrase, subject to the availability of suitable accent sites. Such an observation could be restated in terms of feet—"Abercrombian feet[26] tend to be of equal size"—but taking this step may imply timing consequences that are not intended in a statement about distribution.

For both lexical stresses and pitch accents, there is no need to invoke prominence-delimited constituents to capture statistical regularity in the occurrence of prominences. The further step of postulating prominence-delimited constituents as the units underlying the control of speech timing in English speech is not only unnecessary, but incorrect, as the evidence discussed in Section 2.4.3 indicates. Domain-span timing processes do not operate over constituents delimited by prominences.

In early speech research, where instrumental evidence was difficult to obtain, rhythmical concepts were arrogated from the prescriptive domain of poetics, but the relevance of such concepts to the science of speech is not supported by empirical techniques now available. Despite this lack of evidence, however, research is still pursued within frameworks that, at least implicitly, presuppose the existence of prominence-delimited constituents. Conceptions of the metrical grid that bracket groups of syllables according to levels of prominence are thus positing a unit which is not empirically motivated. Likewise, the view that "rhythm is manifested as the temporal binding of events to specific and predictable phases of a superordinate cycle" (Cummins & Port 1998:147), is not only implicitly invoking the rhythmical foot as an organisation principle of speech, but is going beyond Abercrombie to suggest that speech, like poetical verse, has an higher level of metrical coordination.

Cummins & Port (1998) cite evidence to support this view, based upon their experimental paradigm of "speech cycling": they require subjects to repeat short phrases, such as *big for a duck*, containing two stressed syllables which are to be aligned with the low and high tones of a metronome for which the timing intervals are varied. They find that English speakers tend to place the onsets of stressed syllables more regularly than the occurrence of the tones: that is, stresses align with points within the over-

[26] Abercrombian feet are sometimes delimited by pitch accents, as in this instance, and elsewhere by primary lexical stresses.

all phrase repetition cycle which divide the cycle into regular intervals. They suggest that the patterns can be understood as the nesting of one unit (the stress foot) within a larger unit (the phrase repetition cycle).

As the authors repeatedly state that the observed co-ordination between stress placement and the higher-level cycle is task-specific, it is unclear why the results should be thought to relate to the performance of normal speech. They state themselves that: "This dynamical interpretation of the speech cycling task suggests why isochrony should be more readily observable within the confines of a repetition task than in normal conversational speech. Repetition generates a stable cycle to which nested processes can (or must) entrain. The changing demands of unconstrained speech production do not allow this stability to persist, though it would emerge occasionally as the speech content permitted. In its simplest form, the speech cycling task provides a stable period [the phrase repetition cycle] within which nested and entrained periods can be seen to emerge" (Cummins & Port 1998:150). They later claim, however, that "By structuring an utterance so that prominences (stresses, beats) lie at privileged phases of a higher-level prosodic unit, rhythm is seen as an organisational principle which has its roots in the coordination of complex action and its effect in the realm of prosodic structure" (Cummins & Port 1998:167).

There appear no grounds for the latter assertion: Cummins & Port admit that the superordinate cycle within which stresses are aligned is a product of the experimental task and make no suggestion about which prosodic constituent might fulfil such a function in normal speech. Moreover, it is one thing to show that subjects are capable of performing a task and quite another to claim that this behaviour is habitual: it would be spurious to infer, from the fact that people can learn to juggle, that the timing of hand movements when eating with a knife and fork is entrained to some superordinate cycle.

Finally, it is interesting to note the lack of intuitive evidence for isochrony even in the rhythmically and metrically-ordered domain of poetic verse. Iambic pentameters should be ideal material for the observation of isochronous performance, with each foot in a regular line being disyllabic. In performance, however, the stresses frequently occur with sufficient irregularity to obscure the fact that verse rather than prose is being spoken. Indeed, a performer attempting to place each stress "on the beat" would sound unnatural and possibly comical.

2.5 Pitch accent

2.5.1 The distribution of pitch accents

All intonational phrases include at least one pitch accent. The primary acoustic correlate of pitch accent occurs in the fundamental frequency contour: commonly, there is a rise in F0, but there may be a fall instead, or even a level stretch corresponding to the absence of normal declination. All types of prominence-lending variation in fundamental frequency may be collectively termed "pitch accents".

Ladd (1996) reviews a phonological classification of pitch accents according to high and low pitch targets, based upon the work of Pierrehumbert and others. He also presents an account of the placement of phrase-level prominence in terms of focus; for current purposes, the important facts are that each pitch accent within a phrase is located on a lexically-stressed syllable (with occasional exceptions due, for example, to contrastive focus) and the final pitch accent—known as the nuclear accent—is perceptually the most prominent. Pitch accents tend to be associated with primary lexical stresses, but tonal crowding may cause movement of prenuclear accents onto secondary lexical stresses (for example: Shattuck-Hufnagel *et al.* 1994).

Although pitch accents are characterised primarily by fundamental frequency variation, they are also associated with significant increases in duration in the vicinity of the pitch accent. Klatt (1976:1210) cites Coker *et al.* (1973) as finding that "[one] acoustic correlate of emphatic or contrastive stress is an increase in the duration of a word by 10%–20% or more." Later researchers such as Sluijter (1995), Turk & Sawusch (1997) and Turk & White (1999) have attempted a more precise characterisation of the distribution of such lengthening. Before reviewing this experimental work, it worth considering again a terminological point raised in Chapter 1. Turk & Sawusch (1996) present evidence relating to what they call the "domain" of accentual lengthening. In the current work, the term "domain" is reserved for the unit which occasions the durational variation—such as the intonational phrase or the utterance in the case of final lengthening—and the term "locus" is used for the unit which actually manifests the effect: in phrase-final lengthening, typically the rhyme of the final syllable. This use of the term "locus" seems to be applicable, at least to a first approximation, to durational processes associated with pitch accent; as discussed below, however, the term "domain", which seems valid and useful in the description of both domain-edge effects and domain-span effects such as polysyllabic shortening, may be rather less susceptible to precise characterisation in relation to pitch accent.

2.5.2 The locus of accentual lengthening

Turk & Sawusch (1997) show that the locus of accentual lengthening extends to all parts of the accented syllable and beyond. Firstly, they measure the durations of vowels and consonants in the context of near-homophonous word pairs such as *beef arm* and *bee farm*, where the central consonant—/f/ in this case—could be either word-initial or word-final. As expected, the vocalic nuclei of accented syllables show a large amount of lengthening, having a mean duration of 206 ms compared with 163 ms in the unaccented context. The effect on neighbouring consonants is analysed by measuring the duration of the central consonant when word-initial or word-final, in accented or unaccented words: the word-initial consonants show a large lengthening effect due to pitch accent, being approximately 34 ms longer compared with an unaccented mean of about 101 ms; the word-final consonants also show significant lengthening in an accented word, but the magnitude of the effect is smaller, approximately 10 ms greater than the unaccented mean of around 83 ms.

Turk & Sawusch also examine the durational effect of pitch accent upon unstressed syllables adjoining the accented syllable. Again they use pairs of near-homophonous two-word phrases, this time with a central unstressed syllable which could be either word-initial or word-final, for example: *thank fulfil* and *thankful Phil*. They find that word-final unstressed syllables show a large amount of lengthening in an accented word—approximately 20 ms compared with the unaccented context of 150 ms—whereas word-initial unstressed syllables show no significant effect of membership of an accented word. From this they conclude that the locus of accentual lengthening is most likely to be a unit beginning with a lexically-stressed syllable and including, at least, a subsequent unstressed syllable within the same word. This unit possibly corresponds to the within-word foot, although other within-word candidates are not ruled out by the available evidence.

This finding contrasts with that of Sluijter (1995) who examines the durational effects of pitch accent in English disyllabic words. She measures both syllables of words contrasting only in lexical stress placement—such as *ˈcompact* and *comˈpact*—when accented and unaccented. As found by Turk & Sawusch, word-final unstressed syllables show large amounts of accentual lengthening, 38 ms longer than the unaccented mean of 236 ms. Sluijter, however, finds that word-initial unstressed syllables are also lengthened when the word is accented, by 23 ms compared with 147 ms mean duration in the unaccented condition.

Turk & Sawusch consider an alternative explanation of their findings which suggests a possible reason for this discrepancy. In their materials, word-initial unstressed syllables, when in an unaccented word, are preceded by a pitch-accented syllable, albeit with an intervening word boundary. In the accented word condition, word-initial unstressed syllables are followed by an accented syllable without an intervening word boundary. For example,

Unaccented word condition *THANK fulfil*

Accented word condition *thank fulFIL*

where capitals indicate the placement of the accent. If there were a rightward lengthening effect of pitch accent, attenuated but not blocked by a word boundary, which matched a leftward within-word effect in the accented word condition, then this could explain the failure to observe any accentual lengthening on word-initial unstressed syllables.

Turk & White (1999) replicate and extend this experiment by Turk & Sawusch, including a "baseline" condition in which there is no pitch accent on either word in the near-homophonous phrase pairs:

Baseline condition *thank fulfil*

Comparison of word-initial unstressed syllables in this condition with the accented word condition shows that there is indeed a small lengthening effect to the left of the primary stressed syllable within accented words, and also a rightwards lengthening effect attenuated across a word boundary. These results are shown in the top line of Table 2.4. Another experiment reported by Turk & White indicates that this cross-word-boundary effect is not influenced by the strength of syntactic affiliation of the words on either side of the boundary, at least within the context of a verb phrase.

Example phrase type	Unaccented baseline duration	Lengthening:			
		Phrase-initial accent		Phrase-final accent	
thank fulfil	165 ms	7 ms	4%	8 ms	5%
thankful Phil	166 ms	21 ms	13%	4 ms	2%
knee capsize	207 ms	10 ms	5%	7 ms	3%
kneecap size	206 ms	26 ms	13%	-1 ms	0%

Table 2.4: Data taken from Turk and White (1999). The size of the accentual lengthening effect for word-initial and word-final unstressed syllables and secondary stressed syllables compared with the unaccented baseline condition. Example phrase pairs are shown for each context, with the measured syllable underlined.

Notwithstanding the small cross-word-boundary effect to the right of a pitch-accented syllable, further results reported by Turk & White and shown in Table 2.4 strengthen the conclusion that the distribution of accentual lengthening is primarily related to word structure. Firstly, they show that stressed as well as unstressed syllables adjacent to a pitch-accented syllable undergo lengthening: in phrases such as *knee capsize* and *knee capsize* the central syllable, which carries secondary lexical stress, shows a very similar pattern of accentual lengthening to the unstressed syllable in phrases such as *thankful Phil* and *thank fulfil*. This result rules out the within-word foot as the locus of accentual lengthening, as a secondary stressed syllable is separately footed from the primary stressed syllable within a word. Secondly, Turk & White find lengthening on both syllables which follow the pitch-accented antepenultimate syllable in words such as *property* and *alternative*. The magnitude of lengthening, is greatest on the pitch-accented syllable itself (30 ms, 23 %), and is greater on the word-final syllable (22 ms, 14%) than on the penultimate syllable (12 ms, 11%), although direct comparisons are only suggestive as the phonetic composition of the syllables is different in each case, and the penultimate syllables are substantially shorter even in the unaccented comparison.

Turk & White (1999:171) conclude that "both the left edge of a pitch accented syllable and the left and right edges of a word-sized unit thus appear to attenuate the spread of accentual lengthening." It is interesting to note the contrast between these results and reported studies of domain-edge effects such as initial lengthening and final lengthening. In the latter, for the most part, the effect appears to be well constrained within the boundaries of its manifest locus. For accentual lengthening, however, no single unit fully characterises the locus, because lesser degrees of lengthening extend beyond its apparent boundaries. The best available description of the locus would be "a unit beginning with a primary stressed syllable and extending rightwards to a word boundary", with additional small amounts of lengthening on syllables adjacent to the primary stress, rightwards across a word boundary but leftwards only within a word. An alternative possibility, particularly suggested by the results of the Turk & White study of trisyllabic words, is that the locus of accentual lengthening is actually bimodal: firstly, there is a large lengthening effect on the accented syllable itself, which is attenuated but not blocked by syllable-initial and syllable-final boundaries, causing a small amount of lengthening on adjacent syllables; secondly, there is a fairly large lengthening effect on the final syllable of the accented word. Where the word-final syllable is adjacent to the accented syllable, as in *THANKful*, the accented-word-final effect combines with a smaller lengthening effect due to adjacency to the accented syllable. Where the syllable is word-initial and thus precedes the accented syllable, as in *fulFIL*, there is only the small amount of lengthening due to the adjacency effect. One of the purposes of the experiment presented in Chapter 4 is to determine whether the bimodal locus or the rightward locus or some other interpretation best characterises the distribution of accentual lengthening.

2.5.3 The domain of accentual lengthening

The findings described in Section 2.4 about the timing consequences of lexical stress lend support to the notion that the constituents that undergo domain-edge durational processes arise from a different type of representation from that in which the timing effects of prominence might be derived (an issue discussed in more detail with regard to Dutch by Sluijter & van Heuven 1995). For pitch accent, however, constituent boundaries appear to exert some influence on the precise distribution of its durational effects. The interaction between accentual lengthening and syntactic/prosodic constituents, particularly the lexical word, will be considered in the experiments presented in Chapters 3 and 4.

A related issue which remains unresolved is the place of accentual lengthening within a domain-and-locus description of speech timing. As just seen, the concept of the locus seems less clear-cut than in discussions of domain-edge effects. It is even less clear how the domain of accentual lengthening might be characterised. The natural answer would be a constituent which contains a single pitch-accented syllable, but there does not appear to be much evidence for this unit in English, where even a single word may in some cases contain more than one pitch accent: Beckman & Pierrehumbert (1986) propose such a constituent—the accentual phrase—but this appears to be better motivated for Japanese than for English. Alternatively, if accentual lengthening was only associated with nuclear pitch accent, then its domain would be the intonational phrase. Studies of the durational effects of accent have tended to elicit it through contrast (for example: *I said "PROPERTY sale", not "LEGACY sale"*), where the target accent is intended to be the last or only accent in the intonational phrase. The durational effects of prenuclear pitch accents, and of accents not elicited contrastively, have not been much studied. Given the parallels between prenuclear and nuclear accents in terms of peak alignment discussed by Silverman & Pierrehumbert (1990), and also the apparent ubiquity of durational phenomena in relation to other aspects of prosodic and metrical structure, it seems unlikely however that prenuclear accents should not show durational correlates similar to those seen for nuclear accents. Thus, there is not an obvious candidate for the domain of accentual lengthening within a constituent hierarchy.

It remains an open question whether attempts to unify prosodic constituents and prominences within a single type of representation will ultimately succeed. The durational evidence presented in Section 2.4 suggests that the effects of lexical stress in particular may require a separate type of representation; for accent, however, the association of the locus of lengthening with word structure suggests an accommodation within a constituent hierarchy, although the concept of the domain appears less useful for accentual lengthening than for processes at constituent boundaries.

2.6 Summary

In this chapter, the evidence for suprasyllabic timing effects in English speech is reviewed and placed, wherever possible, into a domain-and-locus framework, which incorporates domain-span processes, where the length of some constituent and duration of some subconstituent are inversely related, and domain-edge processes, where the duration of subconstituents is greater at the initial or final edge of a domain than domain-medially. Another durational process, accentual lengthening, is less straightforwardly accommodated within the domain-and-locus framework; furthermore, the durational effects of pitch accent appear often to be overlooked in previous research relating to domain-edge and domain-span processes.

Possible domain-span processes considered include foot-level isochrony, utterance-level shortening and polysyllabic shortening. The available evidence strongly suggests that foot-level isochrony does not exist in English speech. There is, however, strong evidence for a localised rhythmical process, the stress adjacency effect, which appears to operate in a domain which is distinct from the constituents defining the occurrence of domain-edge effects such as final lengthening and initial lengthening. The stress adjacency effect has to be controlled in experiments relating to other durational processes, such as those presented in Chapter 3 and Chapter 4.

There is insufficient evidence available regarding the other domain-span processes. It is possible that utterance-span shortening reflects a general tendency for domain-span processes to operate within all prosodic domains; alternatively, reported results may reflect indirect effects of constituent length, such as restructuring. Experiment 2 in Chapter 4 is, in part, designed to test claims about utterance-span shortening. The discussion of polysyllabic shortening indicates that some of the previous experiments relating to this effect have contained confounding factors such as stress-adjacent lengthening, phrase-final lengthening and accentual lengthening. The design of these experiments is examined in detail in Chapter 3; Experiment 1 is designed to eliminate some of these potential confounds. Word-level durational processes and their interactions are further examined in Experiment 2 presented in Chapter 4.

The existence of domain-edge processes is better attested, although the precise nature of the domains involved is not always clear. Final lengthening occurs at the end of utterances and at the end of certain utterance-medial phrases: whether these observations reflect a single underlying domain or a hierarchy of domains is uncertain. The evidence for initial lengthening indicates quite strongly that it relates to a word-type domain, which may be the lexical word or the prosodic word. An apparent interaction between initial lengthening and lexical stress or pitch accent has been observed, as has a possible utterance-initial shortening effect. The design of experiments which suggest hierarchical domain-edge effects is critically evaluated in Chapter 4; differences between utterance-edge and utterance-medial processes are addressed in Experiment 2, particularly with regard to the locus of the effect in each case, and to its interaction with the durational effects of pitch accent.

There are a number of different questions to be addressed by the experimental work described in this dissertation. These questions have been raised in this chapter and are presented in the following chapters as formal experimental hypotheses. Testing these hypotheses will contribute to answering three general questions:

- Are previously-observed domain-edge and domain-span effects specific to certain constituents, or do they reflect tendencies at all levels?
- Does each domain-edge and domain-span process have a phonologically-defined locus? Does the locus vary between different levels of hierarchical processes?
- Is the presence of pitch accent a prerequisite for the occurrence of certain durational effects, or does accentual lengthening simply serve to amplify existing effects and possibly modify their loci?

Experiment 1, presented in Chapter 3, looks for evidence of word-level durational processes. Experiment 2, presented in Chapter 4, is designed to determine the *type* of processes—whether domain-edge or domain-span—which operate at the word level and the utterance level.

Chapter 3

Experiment 1: Evidence for a word-level durational effect

3.1 Introduction

Experiment 1 addresses the question: "Is the word a domain of durational variation in English speech?". Previous experiments testing word-level durational effects contain potential confounding factors, as indicated in Chapter 2; the design of two such experiments—Lehiste (1972) and Port (1981)—is examined below. The problems of interpretation of these experiments demonstrate that, in order to test for word-level effects, durational variation which may arise from domain-edge and domain-span processes at other levels of constituent structure must be eliminated; in addition, the durational effects of the distribution of lexical stresses and pitch accents must be controlled. The design of Experiment 1 incorporates these considerations and the results—discussed with reference to other studies of word-level processes, in particular Turk & Shattuck-Hufnagel (2000)—indicate that the word itself, together with a subword constituent, may be domains of durational processes. Experiment 2, in Chapter 4, examines whether word-level effects are best characterised as arising from domain-edge or domain-span processes, or from some combination of such processes.

3.1.1 Polysyllabic shortening: a word-level effect

It has often been asserted that there is an inverse relationship between the size of the word and the duration of its constituents. For example, Ladefoged (1993:95) says that: "[in] sets of words such as *speed, speedy, speedily* [. . .] the vowel in the stressed syllable gets progressively shorter as a result of adding extra syllables in the same word". As in this example, investigation of the effect largely focuses on the durational consequences for the primary stressed syllable in lexical words[1]. The formal description of the polysyllabic shortening effect suggested in Chapter 2 may be revised slightly to reflect this:

The duration of the primary stressed syllable in a lexical word is inversely related to the number of syllables in that word.

This putative relationship has long been attested for English speech[2]. Jones (1942–43) discusses the phenomenon, illustrating both the terms in which the effect has commonly be described and the mechanism by which it has been explained.

> In Southern English the lengths of 'long' vowels in *strongly* stressed syllables depends not only on the type of sound following but also to a large extent on the number of weakly stressed syllables following them. The long [iː], for instance, is very long in words like *me* or *mead* when they are said by themselves. But the same vowel is not nearly so long in *immediate* [ɪˈmiːdjət], and it is still less long in *immediately* [ɪˈmiːdjətli]. The same applies to *muse* [mjuːz], *music* [mjuːzɪk], *musical* [mjuːzɪkl], *morn*, *morning*, and so on. The principle also applies to the short chroneme[3], though it is of course less noticeable: compare *wed, wedding, leg, legacy, rob, robber, robbery*.
>
> The same thing may be noticed in sentences. *Days* said by itself has a very long [eɪ]. In the expression *for days on end* the [eɪ] is less long, and in *the days of the week* it is still less long.
>
> It is easy to see why this is. There is a tendency in English to make the strong stresses follow each other as nearly as possible at equal intervals of time. Consequently, when there is a sequence consisting of a strongly stressed syllable followed by one or more weakly stressed syllables, English speakers instinctively try to cram this sequence into the same space of time as a single strongly stressed syllable. For instance when we count *one, two, three* and *thirteen, fourteen, fifteen*, our rhythm remains approximately the same. The [ɔː] of *fourteen* is therefore necessarily much shorter than that of *four*, and so on. (Jones 1942–43:4)

It is interesting to observe that Jones, in common with later researchers, chooses to frame his description of the effect in terms of the lexical word, but frames his explanation in rhythmical terms. In the first paragraph, he describes polysyllabic shortening as defined above—a domain-span process at the word level—but in the third paragraph he explains the observation as a domain-span process at the cross-word foot

[1]Experiment 2 in Chapter 4 also examines the effect of word length on unstressed syllables.

[2]This process has a morpho-phonological analogue: for instance, Chomsky & Halle (1968) describe a process called *trisyllabic laxing* through which tense vowels in monosyllabic or disyllabic root morphemes may be realised as lax vowels in derived forms with more syllables, for example: *serene* vs *serenity*; *sane* vs *sanity*.

[3]By "short chroneme", Jones means any lax vowel

level[4]. Similarly, Ladefoged (1993:119) observes that "[Polysyllabic shortening] can be interpreted as a tendency to minimise the variation in the length of words containing only a single stress, so that adjacent stresses remain much the same distance apart."

If the observation of polysyllabic shortening may be attributed to the size of a constituent other than the word itself, then it is misleading to discuss the phenomena in terms of words at all. If the rhythmic account of the effect is sufficient, then the shorter duration of *four* in *fourteen, fifteen, sixteen* compared with *four, five, six* should be equivalent to the duration of the same syllable in *four and five and six*. The word-level conception of polysyllabic shortening would be superfluous. As discussed in Chapter 2, there is strong evidence for the existence of a particular rhythmical effect: "stress-adjacent lengthening" is the greater duration of a stressed syllable when immediately followed by another stressed syllable compared with its duration when followed by one or more unstressed syllables. One of the questions addressed in this chapter is whether this explanation is sufficient to explain the observed word-level timing patterns.

There are other potential confounds in previous research on polysyllabic shortening. Experiments on English speech find that the duration of a stressed syllable is greater when it comprises a monosyllabic word than when the same stressed syllable is *followed* by one or more unstressed syllables within a word. For example, Lehiste (1972) finds /stɪk/ is shorter in *sticky* and *stickiness* than in the monosyllable *stick*. Clearly, if the important parameter is word size rather than the size of some other constituent, such as the within-word foot[5], a similar relationship should pertain between stressed syllable duration and the number of syllables which *precede* it within the word.

Researchers into languages other then English have investigated polysyllabic shortening in terms of the effect of syllables preceding and following the stressed syllable. For example, Nooteboom (1972) for Dutch and Lindblom (1968) for Swedish find evidence for polysyllabic shortening in both cases, although the shortening effect of following syllables appears to be the greater. This evidence suggests that there is indeed an influence of word-level structure on stressed syllable duration in these languages, but the asymmetry of the effect suggests the operation of domain-edge processes rather than a domain-span process.

[4] In the second paragraph, he describes a domain-span process at the sentence level, which might be equated with the phonological utterance level. Experiment 2 described in Chapter 4 considers the existence of domain-span processes at the utterance level as well as the word level.

[5] The within-word foot begins at the onset of a stressed syllable and continues to the next stressed syllable onset or to a word boundary, whichever is the sooner.

3.1.2 Word-edge and word-span processes

The distinction between domain-edge and domain-span processes is discussed in Chapter 1. Polysyllabic shortening, as defined above, is a word-span durational process: the inverse relationship between word length and stressed syllable duration holds regardless of the position of the syllable within the word. In contrast, domain-edge processes are dependent upon position: a segment of speech is longer at the edge of some domain rather than when placed domain-medially.

Evidence for domain-edge processes has been put forward for the initial and final edges of a number of domains within English speech. As indicated in Chapter 2, phrase-final lengthening is widely observed, often with a syllable rhyme locus: it seems clear that final lengthening occurs at the ends of intonational phrases and utterances; it is less clear whether such processes apply at the edge of lower-level prosodic constituents like the word. There is also strong evidence for the operation of a word-initial lengthening process, possibly with a syllable onset locus; there may also be an utterance-initial shortening process. Experiments investigating domain-edge processes at the edges of hierarchical prosodic constituents, including the word, are analysed in more detail in Chapter 4.

With regard to word-level processes in English, is first necessary to determine whether the domain of the effect described by Jones (1942–43) and Ladefoged (1993) is indeed the word. As well as the lack of evidence about the effect of syllables preceding the primary stressed syllable, and the potential confound of stress-adjacent lengthening, previous studies have not controlled pitch accent placement, itself a source of lengthening within words, as discussed in Chapter 2. Studies of polysyllabic shortening by Lehiste (1972) and Port (1981) and their potential confounds are discussed here, and then an experiment is presented which is designed to eliminate these confounds in order to test the existence of a word-level effect.

There are a number of studies, other than those by Lehiste and Port, which examine the durational effects of word length. Some, such as those by Barnwell (1971) and Nakatani *et al.* (1981), have designs which contain some of the confounds discussed below. Two other studies sometimes cited as indicating word-level effects have been discussed in Chapter 2. Firstly, the findings of Huggins (1975) were shown to be more strongly indicative of a rhythmical effect independent of word boundaries. Secondly, the study by Beckman & Edwards (1990), which apparently provides evidence of word-final lengthening, was shown to be inconclusive, and possibly confounded by a word-span effect. A study by Turk & Shattuck-Hufnagel (2000), which considers both word-edge and word-span processes, and the influence of pitch accent, will be discussed in more detail following the presentation of the results of Experiment 1.

3.1.3 Experimental evidence for a word-level effect

Two experimental studies which have been cited as evidence for a word-level durational effect are discussed here: Lehiste (1972) and Port (1981). The design and results of each experiment are given separately, followed by a general discussion of confounding factors.

Lehiste 1972

This study is often cited as evidence for a word-level durational effect, for example, by Klatt (1976) and Port (1981). Although Lehiste's own interpretation of the results is somewhat different, the experimental design, as outlined below, suggests a word-level effect. Indeed, Lehiste appears to come to this conclusion in a later paper, where she cites her 1972 study as evidence of the "reduction in duration of monosyllabic stems when various suffixes are added" (Lehiste 1977:260).

The acoustic duration of four lexically-stressed syllables—/stɪk/, /slip/, /spid/, /ʃeɪd/—referred to here as "test syllables", is measured in various speech contexts:

- As isolated monosyllabic words.
- In isolated disyllabic and trisyllabic words derived from the monosyllable by the addition of various suffixes. The same set of suffixes is used for each of the four test syllables. The full set of isolated words is shown in Table 3.1.
- As monosyllabic words in three different short sentences, which have zero, one or two unstressed syllables between the test syllable and the next stressed syllable, for example: *the stick fell*, *the stick is broken*, *the stick was discarded*.

Monosyllables	Disyllables			Trisyllables	
stick	sticky	sticker	sticking	stickily	stickiness
sleep	sleepy	sleeper	sleeping	sleepily	sleepiness
shade	shady	shader	shading	shadily	shadiness
speed	speedy	speeder	speeding	speedily	speediness

Table 3.1: The full set of monosyllables and polysyllables used in the isolated word conditions in Lehiste's (1972) study.

Thus, there are nine different tokens for each of the four test syllables: one isolated monosyllable, five isolated polysyllables and three monosyllables in sentential context. There are two different methods of presentation of the tokens: in one the subject reads each of the tokens ten times; in the other the subjects reads all of the tokens in a single list, the whole list being repeated ten times. Thus there are 180 repetitions of each of the four test syllables, and the full set of materials is read by two subjects, graduate students of Ohio State University.

For the analysis, the durations are pooled between the two different reading methods, as variability between the two is found in general to be non-systematic. Evidence is found which appears to demonstrate a word-level effect. To consider initially just the isolated words: the test syllable spoken as a monosyllabic word is considerably longer than when in a disyllable, and is shorter still in a trisyllable. Results are quoted for each speaker and for each test syllable in each experimental context: for example, for one speaker, /stɪk/ has a mean duration of 432 ms as a monosyllable, 348 ms in the disyllable *sticking* and 272 ms in the trisyllable *stickiness*. The overall means can be calculated by pooling across the two speakers and the four test syllables[6]: in the isolated word conditions, the mean duration of the test syllable in monosyllables is 461 ms, in disyllables it is 342 ms, and in trisyllables it is 300 ms. This may be expressed as a polysyllabic shortening ratio, the amount of shortening as a fraction of the test syllable duration in the monosyllable:

$$(duration_{monosyl} - duration_{polysyl})/duration_{monosyl}$$

Polysyllabic shortening is 26% in the disyllable and 35% in the trisyllable.

While the monosyllabic suffixes *–y*, *–er* and *–ing* are equivalent in their shortening effect, the disyllabic suffixes are not. The suffix *–ily* produces more shortening than the monosyllabic suffixes, but less than the suffix *–iness*. The mean test syllable duration for the suffix *–ily* is 312 ms and for the suffix *–iness* it is 287 ms. This could be attributable to the much greater acoustic length of the suffix *–iness*, which has a mean duration of 445 ms compared with 313 ms for the *–ily* suffix. The acoustic duration of the suffix is not the sole determiner of the amount of polysyllabic shortening, however: Table 3.2 (derived from Lehiste's data) shows that the suffix *–ing* has much greater duration than other monosyllabic suffixes, but difference in the shortening effect is minimal.

Measurement of the vowel which comprises the syllable nucleus reveals that it is this part of the syllable that undergoes the most shortening, although onset and coda are also shortened to some extent (the distribution of shortening between onset and coda is not reported). The acoustic length of the vowel is important: the long vowels in /ʃeɪd/ and /spid/ are more compressible (that is, undergo proportionately more shortening relative to the rest of the syllable) than the short vowel in /stɪk/. Also, the vowel in /slip/, followed by an voiceless stop, is shorter and less compressible than the vowels in /ʃeɪd/ and /spid/.

[6] All overall means reported here have been calculated from the by-speaker and by-test-syllable means presented by Lehiste (1972).

Suffix	Test syllable duration	Suffix duration
-y	345	183
-er	342	194
-ing	338	251
-ily	312	313
-iness	287	445

Table 3.2: The mean durations (ms) of the test syllables and suffixes for the isolated word condition in Lehiste's (1972) study.

The results for test syllable duration in sentential contexts are rather harder to interpret, given that comparisons are to be made with isolated words. As Klatt (1976) indicates, a word spoken in isolation may be as much as twice as long as the same word spoken in a sentential context. In Lehiste's data, for example, the mean duration of the test word in isolation is in all conditions more than 100 ms longer than when it is in a two-word phrase followed by a monosyllabic stressed word, for example: *speed* vs *speed kills*, or *sleep* vs *sleep heals*. This fact alone makes it hard to compare the effects of adding unstressed syllables within a word and across a syntactic boundary, and indeed no clear picture emerges.[7] A more reliable comparison would control overall sentence length, as well as other factors such as pitch accent placement. These methodological points will be expanded upon after examination of another experiment which provides evidence of a word-level durational effect.

Port 1981

The first experiment presented in Port's (1981) paper may be seen as a replication of Lehiste's (1972) study, extended to sentential contexts. The test words are constructed using four "English or English-like" monosyllables in combination with two suffixes.

[7] Lehiste says, however, "The figures show that, by and large, the ratios [of test syllable duration in the isolated monosyllabic word to duration in the longer word or sentence context] are the same for disyllabic words consisting of the base plus suffix, and for disyllabic sequences taken from sentences in which the base word is followed by an unstressed syllable" (page 2021). Examination of the figures shows that this is not the case. In fact, the test syllable in isolated word contexts such as *shading* is consistently longer than in sentential contexts such as *The shade increased*: the mean duration of the test syllable in all disyllabic word contexts is 342 ms; in all sentential contexts where the word boundary is followed by an unstressed-stressed syllable sequence, the mean duration is 292 ms. Furthermore, the difference in duration between monosyllables in sentential context followed by zero or one unstressed syllables—for example, *the stick fell* vs *the stick is broken*—is much less, in most cases, than the difference between test syllable duration in isolated monosyllables compared with disyllables—for example, *stick* vs *sticker*. Lehiste's own conclusion is that "there is no way in which morpheme boundaries and word boundaries could be distinguished on the basis of the temporal patterns" (page 2021). Given the pattern of mean durations, however, it seems safer to draw the limited conclusion about monosyllables vs polysyllables in citation conditions, and refrain from inference which relies on comparison between citation and sentential context. As mentioned above, this appears to be Lehiste's revised verdict in later discussion of these results.

The full set of test words is:

- *deeb, deeber, deeberly*
- *dib, dibber, dibberly*
- *deep, deeper, deeperly*
- *dip, dipper, dipperly*

As can be seen from the materials, there are two other experimental conditions in the experiment, the tensity of the vowel (tense /i/ vs lax /ɪ/) and the voicing of the bilabial stop in the coda (voiced /b/ vs voiceless /p/). The test words are spoken in the sentence frame:

"I say testword *again every Monday."*

The subjects are ten adult speakers of American English, who read each of the 12 test words in the sentence frame five times, with the materials presented in pseudo-random order.

Port reports the durations of the onset, nucleus and coda for each test syllable in each condition[8]. As in Lehiste's study, there is evidence of a word-level effect. The mean duration of the test syllable is 294 ms in monosyllables, 260 ms in disyllables, and 250 ms in trisyllables. This represents 12% polysyllabic shortening in the disyllabic context and 15% in the trisyllabic context, compared with 26% and 35% respectively in Lehiste's study of isolated words. The lesser shortening effect seen by Port, both in proportional and absolute terms, probably reflects the fact that Lehiste's comparisons are derived from words in citation form rather than in sentential contexts.

All parts of the syllable show significant effects of word length upon duration. Table 3.3 shows the means—pooled across the vowel tensity and coda voicing conditions—of the onset, nucleus and coda by word length, calculated from the reported data. The nucleus shows the most polysyllabic shortening between monosyllables and disyllables words, 18% compared with 12% over the whole syllable. Between disyllables and trisyllables, however, the shortening effect is comparable between the nucleus and the coda, about 4% in each case, and is about 2% in the onset.

This differential distribution of polysyllabic shortening in the monosyllable vs disyllable case as compared with the disyllable vs trisyllable case suggests that longer segments within the locus of shortening are more compressible. It is as if, because the nucleus has already undergone a large amount of shortening in the disyllable compared with the monosyllable, it is relatively resistant to further compression. To com-

[8] The overall means reported here are derived from the by-test-word results given by Port (1981).

Word length	Onset	Nucleus	Coda	Total
Monosyllable	96	114	84	294
Disyllable	93	93	74	260
Trisyllable	91	89	70	250

Table 3.3: Mean durations (ms) by word length, calculated from the data reported in Port's (1981) study.

pensate, other segments within the locus of shortening are subject to greater compression in the disyllable vs trisyllable comparison[9].

The other experimental conditions do not have their durational effects distributed throughout the syllable: the nucleus shows significant effects of vowel tensity—the tense vowel is longer—and postvocalic voicing—the vowel is longer when the following consonant is voiced; in the coda, the voiced stop is shorter than the voiceless stop, but there is no durational effect of the tensity of the previous vowel; onset duration is unaffected by either vowel tensity or postvocalic voicing.

Port notes that the shortening effect of the first added syllable is much greater than that of the second. He suggests two, possibly parallel, explanations. Firstly, he refers to Klatt's (1976) concept of incompressibility of vowels beyond a certain minimum duration, such that shortening due to combinations of factors will tend to be asymptotic towards this minimum. (That this is not an absolute physiological minimum is indicated by a much shorter duration for the same stressed syllable reported by Port under conditions of rapid speech.) Secondly, he points out that it is only the first added syllable which moves the stressed syllable from its original word-final position. This observation is discussed further below.

3.1.4 Alternative interpretations of the observed effect

Huggins (1975:459), with reference to two other studies of polysyllabic shortening observes that: "Neither Lindblom [1964] nor Barnwell [1971] established that the unstressed syllable *had* to fall in the same word ..." Similarly, neither Lehiste (1972) nor Port (1981) demonstrate conclusively that domain of the observed effect is the word itself. Some alternative interpretations are considered now.

Stress-adjacent lengthening

The temporal influence of the distribution of lexical stresses has to be controlled: the first of a pair of adjacent stressed syllables is longer than when it is followed by one or

[9] An alternative interpretation of this effect is discussed in Chapter 4.

more unstressed syllables. The possibility of stress-adjacent lengthening is confounding factor in Lehiste's (1972) experimental design, where in one presentation of the materials the words are read out as lists of each item.

Series 1 *"stick, stick, stick . . . "*

Series 2 *"sticky, sticky, sticky . . . "*

Series 3 *"stickiness, stickiness, stickiness . . . "*

Stress-adjacent lengthening could cause the measured syllable to be longer in the list of monosyllables than in the lists of polysyllables, although the magnitude of the effect may not be sufficient to account for all of the observed variation. Van Lancker *et al.* (1988) finds that the presence of an unstressed syllable between two stressed syllables causes shortening of the first by 22 ms when it is unaccented (and much less when it is accented), compared with a mean disyllabic shortening effect observed by Lehiste of 119 ms. The syllables in the latter experiment are much larger, however, and the proportional shortening effects are more similar: 26 % in the Lehiste experiment and 14 % in the Van Lancker *et al.* experiment. The possible presence of pitch accent in Lehiste's materials is discussed further below.

Whether a durational difference would also be expected due to stress distribution between Series 2 and Series 3 above is less clear; as discussed in Chapter 2, evidence for a shortening effect of two unstressed syllables compared with one is inconclusive, but if such an effect exists, is likely to be quite small. As such it would not represent a major confound in the experiment of Port (1981). Stress-adjacency is not present in his design. The measured syllable is followed by an unstressed syllable in the monosyllabic test-word sentence: for example,

"I say dib again every Monday."

Both experiments contain some other potentially confounding factors, however.

The within-word foot

In the above example from Port (1981), the words *dib*, *dibber* and *dibberly* all begin with a stressed syllable and contain no other stressed syllables; thus they are co-extensive with another proposed prosodic constituent, the "within-word foot". As defined by Halle & Vergnaud (1987), the within-word foot begins at the onset of a stressed syllable and extends to the onset of the next stressed syllable or to a word boundary, whichever is the sooner. The evidence for the within-word foot as locus of durational processes is discussed in Chapter 2. The experimental evidence of Lehiste and Port cannot distinguish between the word and the within-word foot as domains of the observed effect.

Phrase length

Some previous studies of the relationship between word length and stressed syllable duration, such as those of Lehiste and Port, leave open the possibility that it is not the length of the word which is important, but the length of some other constituent which dominates the word. Domain-span compression effects have primarily been associated with low-level constituents of speech, such as words (polysyllabic shortening) or cross-word feet (isochrony), but higher-level constituents may exhibit domain-span compression: Jones (1942–43)—as quoted in Section 3.1.1—and Lehiste (1974) claim that the durations of words in an utterance may be shorter if the utterance is longer.

The evidence for domain-span compression processes and their application across domains at different levels of speech structure is inconclusive. Experiment 2 is, in part, designed to address this question, but where domain-span processes cannot be ruled out they remain a potential confound. For example, in the experiment of Port (1981), the duration of a primary syllable is measured in monosyllabic, disyllabic and trisyllabic contexts, within a constant sentence frame: *"I say [dib|dibber|dibberly] again every Monday"*. In this design, the size of any constituent which wholly dominates the word is covariant with the size of the word, thus any compression effect cannot be definitively ascribed to the word level. The results of Lehiste (1972) and Port (1981) do not rule out the possibility that, for example, the phonological phrase or the intonational phrase or the utterance may be the domains of the observed shortening effect.

Position in phrase

It is well-known that the final segments of intonational phrases undergo lengthening relative to equivalent phrase-medial segments. This effect is generally thought to be greatest on the rhyme of the final syllable in the phrase (see Chapter 2 for discussion). If the duration of a stressed syllable in a monosyllabic word is measured immediately preceding an intonational phrase boundary, then the test syllable would have greater duration due to adjacency to the phrase boundary than in a disyllabic word. For example, measurement of the syllable /stɪk/ in:

1. *"Valerie threw a stick, and Jim threw a stone."*
2. *"Valerie had a sticker, and Jim had a flag"*

would show—assuming the sentence is uttered as two intonational phrases in accordance with punctuation—the test syllable to be longer in the monosyllable than in the disyllable. Because of the presence of the phrase boundary, it would not be possible to attribute this effect to word size.

Unlike domain-span compression effects, the existence of intonational-phrase-final lengthening effects is well-attested and the magnitude of the effects is quite large. As outlined in Chapter 2 and discussed in more detail in Chapter 4, it is not clear how far domain-final lengthening is a tendency at all levels of some hierarchical prosodic structure. An additional complication is that speakers may give the same sentence different phrasings according both to their pragmatic interpretation and to variables such as phrase length and speech rate: thus manifest prosodic boundaries are hard to control and even to identify in experimental materials.

In Lehiste's experiment, it is very difficult to determine the influence of phrasing on the observed durations: the materials are read in lists, which tend to induce particular patterns of phrasing, such as the placement of a phrase boundary after every three or four items. In Port's experiment, the most likely place for a phrase boundary in the carrier sentence

"I say testword *again every Monday."*

would seem to be after the word "again", which would remove the test word from the locus of significant domain-final lengthening. Given, however, the likely presence of a pitch accent on the test word, it is possible that a phrase boundary of some kind may occur immediately afterwards. As the details of manifest phrasing are not made explicit, the presence of this possible confound cannot be ruled out.

Pitch accent

As described in Chapter 2, there is evidence that the lengthening associated with pitch accent is not confined to the accented syllable itself, but may also affect adjacent syllables (for example: Turk & Sawusch 1997; Sluijter 1995). This lengthening appears to be attenuated, but not entirely blocked, by word boundaries, and appears to be greater on syllables following the accented syllable than on those preceding it (Turk & White 1999). Because word structure is a factor in determining the distribution of lengthening due to pitch accent, the presence or absence of accent should be taken into account when determining the relationship between word length and the subconstituent duration.

The use of a fixed frame sentence in Port (1981), as shown above, would be expected to elicit a nuclear pitch accent on the target word, as it is the new information in successive sentences. It is possible that the observed durational effect of word structure may not generalise to words without a nuclear pitch accent. One possibility is that the lengthening effect of pitch accent may be less on the primary stressed syllable in polysyllables than in monosyllables (Turk & White 1999). Alternatively, the polysyllabic shortening effect observed in pitch-accented words may actually depend on the presence of pitch accent, and not occur at all when words are unaccented. The effect of word length on stressed syllable duration needs to be investigated for both accented and unaccented words.

3.2 Experimental design

Lehiste (1972) and Port (1981) suggest the existence of a word-level durational effect, but both experiments, and others such as Barnwell (1971) and Nakatani *et al.* (1981), include some of the confounding influences upon speech segment duration outlined in Section 3.1.4. The purpose Experiment 1 is to observe the effect of word size on primary stressed syllable duration, whilst eliminating or controlling these other durational influences.

3.2.1 Experimental purpose

The question addressed in Experiment 1 may be restated: "Is there a word-level durational effect or is the effect previously observed due to other factors which influence speech timing?" The other factors considered are:

The within-word foot Is the within-word foot, rather than the word, the domain of the observed effect?

Stress-adjacent lengthening Is the lengthening of a stressed syllable due to the presence of an immediately-following stressed syllable a contributory factor in the observed effect?

Pitch accent Is the relationship between word size and stressed syllable duration different for accented and unaccented words?

Higher constituent length Does adding syllables to some domain which dominates the word, such as a phrase or utterance, rather than to the word itself, cause the observed effect?

Phrase-final lengthening Does the observed shortening occur because the additional syllables move the stressed syllable from phrase-final position, where it would undergo lengthening?

The experiment is designed to address this question by explicitly testing the word, within-word foot and pitch accent hypotheses, whilst controlling the influences of the other factors.

3.2.2 Experimental materials

The experimental materials are constructed from pairs of near-homophonous[10] two-word three-syllable phrases[11], shown in Table 3.4. In one phrase of the pair, the word boundary is placed between the first and second syllables; in the other phrase, the word boundary is placed between the second and third syllables. The initial and final syllables in the phrase pairs carry the primary lexical stress in each word: it is the duration of these syllables, referred to as "test syllables", that is the dependent variable. In half of the phrase pairs, the central syllable contains a reduced vowel and is unstressed; in the other half, the central syllable contains a full vowel and has secondary lexical stress.

	Word boundary after first syllable	Word boundary after second syllable
Reduced (unstressed) central syllable	bake enforce can inspire thank fulfil cube explain toe content pay perform Dan surprise day today *A*	bacon force cannon spire thankful Phil cubics plane token tent paper form dancer prize data day *B*
Full (stressed) central syllable	knee capsize near bisect there foreclose skim Peking shake downstairs there foursquare crow barbette hard whereby *C*	kneecap size nearby sect therefore close skimpy king shakedown stairs therefore square crowbar bet hardware buy *D*

Table 3.4: Phrase pairs used in the experimental materials. The letters A–D at the foot of each cell are used in Section 3.3.1 to indicate how the sentences are grouped into blocks for the recording.

The phrases are presented in carrier sentences designed to elicit no pitch accent

[10]The term "near-homophonous" is intended to indicate that there are few phonetic differences between the phrases within each pair. Certain specific differences may be observed for some Scottish speakers, in particular: the /w/ may differ in voicing between *hard whereby* and *hardware buy*; the final vowels in *four* and *therefore* may differ, being /o/ and /ɔ/ respectively; some vowels realised in Standard Southern British English as /ə/ may be distinct for some Scottish speakers.

[11]The recordings of these materials were also used for one of the experiments reported in Turk & White (1999).

within the phrase, or pitch accent on the initial syllable of the phrase, or pitch accent on the final syllable of the phrase (capitals indicate emphatic stress):

No Accent SAY "thank fulfil", don't SHOUT "thank fulfil".

Initial Accent Say "THANK fulfil", don't say "BOAT fulfil".

Final Accent Say "thank FULFIL", don't say "thank SURREAL".

The phrases and carrier sentences are designed to vary the size of the within-word foot and the word containing the test syllable, whilst keeping other potential influences on duration constant. These controlled factors are:

- segmental composition of phrases, within phrase pairs.
- phrase length and utterance length.
- position of test syllables with respect to phrase boundaries.

The placement of nuclear pitch accent with respect to the test syllables and the number of unstressed syllables between stressed syllables are systematically varied.

3.2.3 Predictions of the experimental hypotheses

The combination of two locations of the test syllable (phrase-initial and phrase-final) and two types of central syllable generates four subsets of materials, within which direct comparison of measured durations is possible.

In "left-headed" words, the test syllable is word-initial in the disyllabic word of the minimal pair, for example: *knee* and *kneecap*. In "right-headed" words, the test syllable is word-final in the disyllable, for example: *size* and *capsize*. Within these two subsets, the materials may be grouped according to whether the other syllable in the disyllable is stressed or unstressed. For example, in *thankful Phil* vs *thank fulfil*, the central syllable is lexically unstressed; in contrast, the central syllable in *kneecap size* vs *knee capsize* carries secondary lexical stress in both contexts.

The relevant constituent structure of the materials is shown in Figure 3.1. Each subset of the materials is treated as a separate experiment for the purposes of statistical analysis. In all four experiments, the word-level hypothesis predicts that the duration of the test syllable, carrying the primary lexical stress, will be less in the disyllable than in the monosyllable. The predictions of the within-word-foot hypothesis are different according to the constituent structure in each experiment, as outlined below.

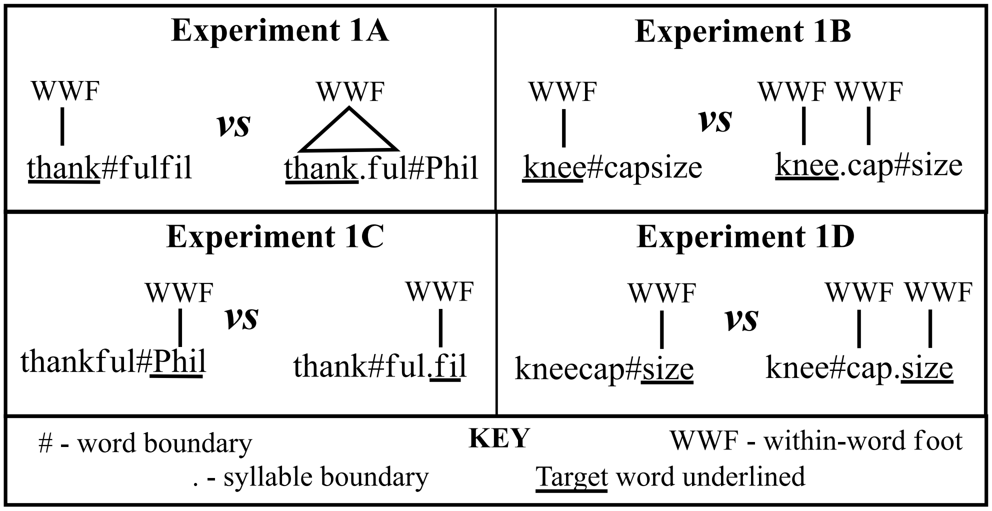

Figure 3.1: The constituent structure of the materials. Example phrases are shown for each experiment and the test syllables are underlined in each case.

Experiment 1A: left-headed words, reduced central syllable

The materials in Experiment 1A are shown in Table 3.5[12]. This experiment represents a partial replication of the experiments of Lehiste (1972) and Port (1981); partial because the comparison here is between monosyllables and disyllables, whereas both Lehiste and Port had a trisyllabic condition as well. Both the word-level hypothesis and the within-word-foot hypothesis predict that the test syllable should be longer in the monosyllabic word than in the disyllabic word, because here the word and the within-word foot containing the test syllable are co-extensive.

Test syllable	Monosyllabic context	Disyllabic context
/bek/	bake enforce	bacon force
/kan/	can inspire	cannon spire
/θaŋk/	thank fulfil	thankful Phil
/kjub/	cube explain	cubics plane
/to/	toe content	token tent
/pe/	pay perform	paper form
/dan/	Dan surprise	dancer prize
/de/	day today	data day

Table 3.5: Experiment 1A materials. The phonetic transcriptions represent the test syllable in Scottish Standard English.

[12] All the test syllables in this experiment are transcribed in Scottish Standard English, as the experimental subjects were all speakers of Edinburgh dialects.

Experiment 1B: left-headed words, full central syllable

The materials in Experiment 1B are shown in Table 3.6. Unlike the experiments of Lehiste (1972) and Port (1981), the additional syllable in the disyllable carries secondary lexical stress and so comprises a separate within-word foot. In this experiment, the size of the within-word foot containing the test syllable remains constant—it is monosyllabic whether the word itself is monosyllabic or disyllabic—and thus any variation in the duration of the test syllable cannot be attributed to processes at the within-word-foot level.

Test syllable	Monosyllabic context	Disyllabic context
/ni/	knee capsize	kneecap size
/niɹ/	near bisect	nearby sect
/ðeɹ/	there foreclose	therefore close
/skɪm/	skim Peking	skimpy king
/ʃek/	shake downstairs	shakedown stairs
/ðeɹ/	there foursquare	therefore square
/kɹo/	crow barbette	crowbar bet
/haɹd/	hard whereby	hardware buy

Table 3.6: Experiment 1B materials. The phonetic transcriptions represent the test syllable in Scottish Standard English.

Experiments 1C and 1D

The materials in Experiment 1C, shown in Table 3.7, and in Experiment 1D, shown in Table 3.8, also allow a distinction to be made between the within-word-foot and word-level hypotheses. In the right-headed materials, the test syllable is in a monosyllabic within-word foot, regardless of the number or type of syllables which precede it within the word. If the durational variation observed by Lehiste (1972) and Port (1981) has a within-word-foot domain, then there should be no difference in the duration of the test syllable between monosyllables and disyllables. If the previous observations really do reflect a word-level process, however, then a similar pattern of polysyllabic shortening should be observed in the right-headed materials and the left-headed materials.

Pitch accent hypotheses

All four experiments have two accent conditions: the word containing the test syllable is either accented or unaccented. If a relationship is observed between either within-word-foot size or word size and stressed syllable duration, two hypotheses exist about

Test syllable	Monosyllabic context	Disyllabic context
/fɔɹs/	bacon force	bake enforce
/spaɪɹ/	cannon spire	can inspire
/fɪl/	thankful Phil	thank fulfil
/plen/	cubics plane	cube explain
/tɛnt/	token tent	toe content
/fɔɹm/	paper form	pay perform
/pɹaɪz/	dancer prize	Dan surprise
/de/	data day	day today

Table 3.7: Experiment 1C materials. The phonetic transcriptions represent the test syllable in Scottish Standard English.

Test syllable	Monosyllabic context	Disyllabic context
/saɪz/	kneecap size	knee capsize
/sɛkt/	nearby sect	near bisect
/kloz/	therefore close	there foreclose
/kɪŋ/	skimpy king	skim Peking
/steɹz/	shakedown stairs	shake downstairs
/skweɹ/	therefore square	there foursquare
/bɛt/	crowbar bet	crow barbette
/baɪ/	hardware buy	hard whereby

Table 3.8: Experiment 1D materials. The phonetic transcriptions represent the test syllable in Scottish Standard English.

the influence of pitch accent: firstly, the effect is only be observed when the test syllable is in a pitch-accented word; secondly, pitch accent amplifies the effect observed in unaccented words. The null hypothesis is that there is no interaction between word length or within-word foot length and pitch accent.

Other factors

The other factors considered in Section 3.1.3 as explanations of the previously-observed word-level effect are not directly tested in these experiments. The length in syllables of the phrase and the utterance containing the test syllable is kept constant. The adjacency of the test syllable to phrase boundaries is kept constant, assuming no phrase boundary is inserted between the two words of the test phrase[13]. Evidence of stress-adjacent lengthening may be observed in the comparison between the test syllable

[13] As described in Section 3.3.1, precautions were taken to prevent the insertion of a phrase boundary, and recorded tokens with a perceptible boundary were not included in the analysis.

durations in Experiment 1A and Experiment 1B, with the test syllable followed immediately by another stressed syllable in the latter but not the former: the comparison is necessarily indirect, however, because of the different segmental composition of the phrases, but the effect is well attested by previous studies and likely to be observed.

If no differences are observed due to the experimental conditions within the experiments, then the conclusion must be that one or more of the other factors considered—phrase/utterance length, phrase boundaries, stress-adjacency—is responsible for the effect observed by Lehiste, Port and other.

3.3 Experimental procedure

There are 16 phrase pairs in the experimental materials, as shown in Table 3.4. The phrases are placed in carrier sentences, and the appropriate emphasis for each of the three pitch accent conditions is indicated with block capitals. The dependent variables are the durations of the first and last syllables in the first phrase—the test phrase—of each carrier sentence. In the initial and final accent conditions, as described in Section 3.3.3, only the syllable carrying the pitch accent is actually measured. The three accent conditions are, for example:

Initial accent Say "BAKE enforce", don't say "TANK enforce".

Final accent Say "bake ENFORCE", don't say "bake REMOVE".

No accent SAY "bake enforce", don't SHOUT "bake enforce".

where the dependent variables are the durations of the underlined syllables.

3.3.1 Recording

The recordings for this experiment took place as part of another experiment, reported in Turk & White (1999), which is concerned with structural influences on accentual lengthening (see Chapter 2 for discussion). In that experiment, the dependent variable is the duration of the central syllable of the test phrase. In addition to the 96 sentences described above, there are another 48 sentences in the full set of materials.

The test phrases in these additional sentences are unpaired trisyllabic phrases, included as a test of the validity of the experimental method. A criticism which may be directed at previous experiments of a similar design, such as Turk & Sawusch (1997), is that the use of phonetically-similar phrase pairs with ambiguous lexical structure, such as *bake enforce* vs *bacon force*, might lead speakers to indicate the correct interpretation by consciously or unconsciously signalling the word boundary more markedly than they otherwise would. Thus accentual lengthening might be blocked in the experiment, where it might otherwise cross a word boundary.

The additional phrases are included to test this hypothesis. As shown in Table 3.9, these trisyllabic phrases all have a word boundary after the first syllable, and do not have a phonetically-similar analogue with a word boundary after the second syllable. There are two types of unpaired phrases: half were judged nonsensical and half were judged meaningful. This condition is included as a test of the hypothesis that the prosodic realisation of the experimental phrases might be different where they are perceived as meaningless. There are sixteen unpaired phrases, which have three pitch accent conditions like the paired phrases.

	Reduced Central Syllable	Full Central Syllable
Unpaired Words: Nonsense	joke enforce plane inspire plank fulfil tube explain *E*	tree capsize pier bisect hair foreclose swim Peking *F*
Unpaired Words: Meaningful	play today big surprise please perform stay content *G*	stay downstairs new Peking don't capsize stand foursquare *H*

Table 3.9: Unpaired test phrases also included in Experiment 1 recordings. The letters E–H at the foot of each cell are used to indicate how the sentences are grouped into blocks for the recording.

The complete set of materials was presented to subjects in three blocks, which kept apart the pairs of phonetically-similar phrases. This was done to deflect the subjects' attention away from the potential ambiguity of the phrases. The blocks are made up as follows, with the letters indicating from which cell of Tables 3.4 and Table 3.9 the phrases are taken. An example phrase from each set is shown:

Block 1 A F H (*bake enforce*; *tree capsize*; *stay downstairs*)

Block 2 B D (*bacon force*; *kneecap size*)

Block 3 C E G (*knee capsize*; *joke enforce*; *play today*)

There are 16 different phrases in each block, with each phrase presented in all three frame sentence types, corresponding to the three pitch accent conditions. Thus, there are 48 test sentences within each block.

The sentences were printed onto 6″ × 4″ record cards. Two sets were prepared, sorted into the above blocks, and the order randomised within each block. The order of presentation of the blocks was counterbalanced between subjects, with subjects reading both randomised sets of sentences for a particular block before moving on to the next block.

Before reading the sentences, subjects were given written instructions to read each sentence aloud naturally and to emphasise the words in capital letters as though, for example, they were correcting someone. The instructions told the subjects that they should read the sentence again if they made a mistake or did not emphasise the sentence in the right way.

Each subject read 12 practice sentences before beginning the recording, these sentences being randomly selected from the full set of sentences. Once the recording began, subjects controlled the rate at which they read the sentences. The experimenter asked subjects to repeat a sentence where it was judged that it had not been read satisfactorily. This was done if:

- the lexical content of the sentence was misread.
- the words in capitals were not emphasised.
- other words were emphasised.
- a pause was perceived within the test phrase.

Recordings were made direct to disk in ESPS format at a sample rate of 16kHz.

3.3.2 Experimental subjects

The experimental subjects were six female speakers of Scottish English from the Edinburgh area, who reported no speech or hearing problems. They each were paid five pounds for the recordings. None of the subjects were given any specific information about the purpose of the recordings until after they had completed them.

3.3.3 Measurement of syllable duration

Each of the 96 experimental sentences was read twice by each of the six subjects. Thus there were 1152 sentences recorded for this experiment, each containing two test syllables.

For all the experiments, the unaccented test syllables are taken from the No Accent condition, in which speakers were directed to emphasise words in the carrier sentence outside the test phrase; for example: *SAY "thank fulfil", don't SHOUT "thank fulfil"*. For Experiments 1A and 1B, where the test syllable is phrase-initial, the accented test syllables are taken from the Initial Accent condition, where the emphasis is on the first word of the test phrase; for example: *Say "THANK fulfil", don't say "BOAT fulfil"*. For Experiments 1C and 1D, where the test syllable is phrase-final, the accented test syllables are taken from the Final Accent condition, where the emphasis is on the last word of the test phrase, for example: *Say "thank FULFIL", don't say "thank SURREAL"*. In all cases, it is the phrase in the first half of the sentence (for example, *SAY "thank fulfil"*) from which the measurements are taken; the second half of the sentence (for example, *don't SHOUT "thank fulfil"*) being used to assist the speakers in their placement of emphasis in the first phrase.

1536 syllables were measured for this experiment: two-thirds of the phrase-initial test syllables and two-thirds of the phrase-final test syllables. The start and end points of these syllables, shown in Tables 3.10 and 3.11, were hand-labelled by analysis of the waveforms and spectrograms using XWaves. Details of the criteria used to determine the beginning and end of each test syllable are given in Appendix A.

Experiment 1A		Experiment 1B	
Test syllable	Following context	Test syllable	Following context
/bek/	/(ə)n/	/ni/	/k/
/kan/	/(ə)n/	/niɹ/	/b/
/θaŋk/	/f/	/ðeɹ/	/f/
/kjub/	/ɪ/	/skɪm/	/p/
/to/	/k/	/ʃek/	/d/
/pe/	/p/	/ðeɹ/	/f/
/dan/	/s/	/kɹo/	/b/
/de/	/t/	/haɹd/	/w/

Table 3.10: The test syllables in Experiments 1A and 1B. The following context is given for the purposes of illustrating segmentation criteria. The preceding context is /se/ in all cases.

About 6% of the total sentences recorded were discarded because the reading was judged incorrect. The main reasons for discarding sentences were incorrect placement of emphatic stress, or pausing between the words of the test phrase. These sentences were not re-recorded. This meant that durations were not available for 104 of the 1536 test syllables, but because each subject read two repetitions of each sentence, there were only 8 data points for which both tokens were missing. There are 192 different data points in each experiment, with each data point comprising two measurements. The missing data are distributed as follows:

Experiment 1A 10 missing measurements; no missing data points.

Experiment 1B 31 missing measurements; 3 missing data points.

Experiment 1C		Experiment 1D	
Preceding context	Test syllable	Preceding context	Test syllable
/n/	/fɔɹs/	/p/	/saɪz/
/n/	/spaɪɹ/	/aɪ/	/sɛkt/
/(ə)l/	/fɪl/	/ɔɪ/	/kloz/
/s/	/plen/	/i/	/kɪŋ/
/n/	/tɛnt/	/n/	/steɪz/
/pə/	/fɔɹm/	/ɔɪ/ (or /oɪ/)	/skweɪ/
/s(ə)/	/pɹaɪz/	/aɪ/	/bɛt/
/t(ə)/	/de/	/eɪ/	/baɪ/

Table 3.11: The test syllables in Experiments 1C and 1D. The preceding context is given for the purposes of illustrating segmentation criteria. The following context is /nɔt/ in all cases.

Experiment 1C 21 missing measurements; 1 missing data point.

Experiment 1D 42 missing measurements; 4 missing data points.

In no case is the amount of missing data considered large enough to warrant further action being taken prior to statistical analysis.

3.3.4 Statistical analysis: procedure

The Analyses of Variance presented here for each experiment have fixed factors of Word Length and Pitch Accent, with the duration of the test syllable as the dependent variable. The main ANOVA in each case has subjects as a random factor: this is a By-Subjects or F_1 analysis, recommended by Raaijmakers *et al.* (1999) as the appropriate analysis for experiments where items are matched between conditions. Recent practice in linguistic research has been also to treat items as a random factor—the By-Items or F_2 analysis—and to regard differences as significant if both the By-Subjects and the By-Items analysis attain a certain level, such as $p < .05$. Raaijmakers *et al.* argue that this practice arises from a misinterpretation of the recommendations of Clark (1973). Clark advocates the use of the *minF′* statistic, calculated using F_1 and F_2, in language experiments of particular designs, for instance, where items are a nested factor within each experimental condition.

In accordance with current practice, results are also reported here for the analysis with Items as a random factor. Given the persuasive arguments of Raaijmakers *et al.* on this subject, however, the significance level of the F_2 statistic will not be regarded as indicative of the presence or absence of an experimental effect. Differences are described as significant if $p < .05$ on the By-Subjects analysis and highly significant if $p < .01$.

Planned comparisons between two levels of particular experimental conditions are carried using one-tailed t-tests.

3.4 Results

The mean durations for each experiment according to the conditions of Word Length and Pitch Accent are shown in Table 3.12.

	Word Length	
	Monosyllable	Disyllable
Experiment 1A	*thank fulfil*	*thankful Phil*
Unaccented	201	172
Accented	242	192
Experiment 1B	*knee capsize*	*kneecap size*
Unaccented	247	220
Accented	314	267
Experiment 1C	*thankful Phil*	*thank fulfil*
Unaccented	324	317
Accented	415	386
Experiment 1D	*kneecap size*	*knee capsize*
Unaccented	351	339
Accented	447	422

Table 3.12: Mean test syllable duration (ms) in Experiment 1. Example phrases are included for illustration. The test syllable in each phrase is underlined.

Two trends can be clearly observed in the data. Firstly, in each experiment, the test syllable is shorter in disyllables than in monosyllables for both accent conditions. This suggests that the word-level hypothesis—that test syllable duration is inversely related to word length—may be correct. Secondly, in each experiment, for a given word length, the test syllable is longer in an accented word than in an unaccented word. The latter result is expected given previous studies of accentual lengthening such as Turk & Sawusch (1997).

Statistical analysis is presented below for each experiment, indicating that it is correct to accept the word-level hypothesis and reject the within-word foot hypothesis. The relationship between the word-level effect and the presence or absence of pitch accent is also examined.

As outlined in Section 3.3.1, additional unpaired phrases were recorded to test the validity of the experimental methodology. The results indicate that the methodology is sound. As reported in Turk & White (1999), there is no evidence from the durational variation of central syllable in the test phrases that subjects differentiated prosodically between paired and unpaired stimuli. From this it may be inferred that the use of phonetically-similar phrase pairs is not a confounding factor in this experiment. There was also "little positive evidence that the meaningful *vs* nonsense nature of the test stimuli influenced the [observed] durations" (Turk & White 1999:189).

3.4.1 Experiment 1A: left-headed words; reduced central syllable

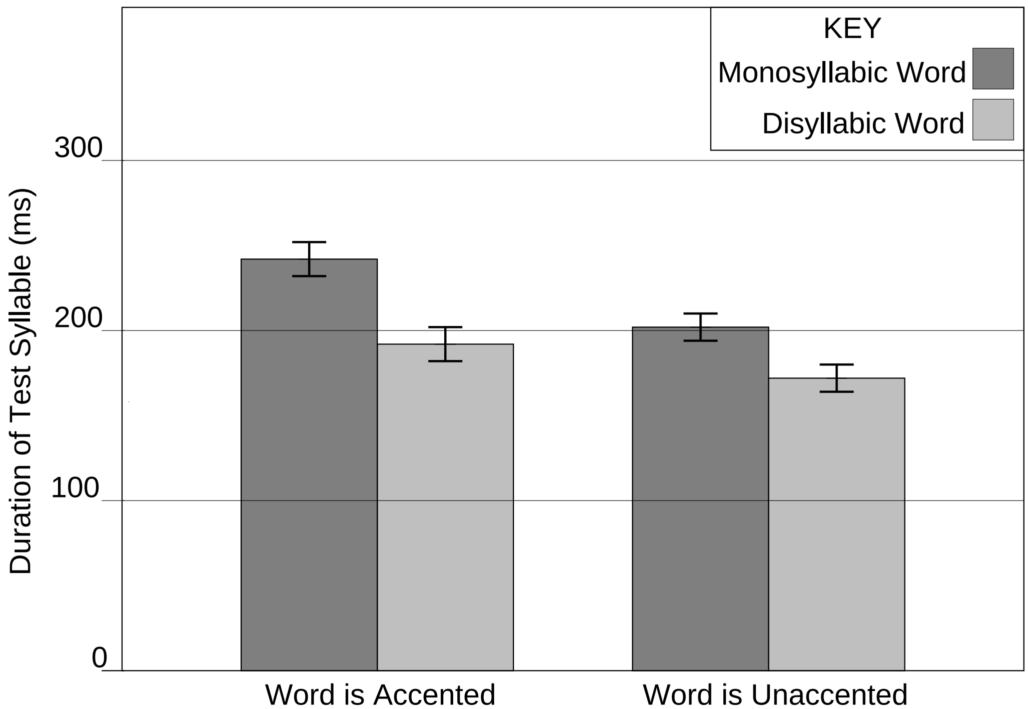

Figure 3.2: Mean test syllable duration (ms) in Experiment 1A: left headed words; reduced central syllable.

Figure 3.2 shows mean test syllable duration for Experiment 1A according to the experimental conditions of Word Length and Pitch Accent. (Here and throughout, the error bars represent two standard errors either side of the mean. This interval contains the population mean with 95% confidence.) The test syllable is longer in monosyllables than in disyllables: for example, /θaŋk/ is longer in *thank fulfil* than in *thankful Phil*. As shown in Table 3.13, the main effect of Word Length is highly significant. This represents a partial replication of the results of Lehiste (1972) and Port (1981), and provides support for both the word-level hypothesis and the within-word-foot hypothesis, as the two units are co-extensive in this experiment.

The effect of Word Length on test syllable duration is consistent across all the materials. Mean test syllable duration for each phrase pair, averaged over both accent conditions, is given in Appendix B, where it can be seen that shortening effect from in

	Degrees of freedom	F Ratio	Significance level
By-Subjects analysis			
Word Length	1,5	89.44	$p < .001$
Accent	1,5	120.22	$p < .001$
Word Length × Accent	1,5	47.50	$p < .001$
By-Items analysis			
Word Length	1,7	55.04	$p < .001$
Accent	1,7	131.95	$p < .001$
Word Length × Accent	1,7	12.38	$p < .01$

Table 3.13: ANOVA summary table for Experiment 1A.

disyllables compared with monosyllables applies in all cases.

The lengthening effect of Pitch Accent on test syllable duration can also be seen in Figure 3.2. The test syllable is longer when the word is accented, both in the comparison between accented and unaccented monosyllables and in the comparison between accented and unaccented disyllables. Table 3.13 shows that the main effect of Pitch Accent is highly significant.

Table 3.13 also shows a highly significant interaction between Word Length and Pitch Accent: the lengthening effect of Pitch Accent is greater in monosyllables (41 ms, 20%) than in disyllables (20 ms, 12%); this also means that the shortening effect of Word Length (in disyllables compared to monosyllables) is greater in accented words than in unaccented words. Planned comparisons show, however, that the main effects remain highly significant when analysed separately according to the levels of the other factor: thus, the effect of Pitch Accent is highly significant ($p < .001$) for both monosyllables and disyllables, and the effect of Word Length is highly significant ($p < .001$) for both accented and unaccented words.

3.4.2 Experiment 1B: left-headed words; full central syllable

Figure 3.3 shows mean test syllable duration for Experiment 1B according to the experimental conditions of Word Length and Pitch Accent. The test syllable is longer in monosyllables than in disyllables: for example, /ni/ is longer in *knee capsize* than in *kneecap size*. Table 3.14 shows that the main effect of Word Length is significant. This result is similar to that of Experiment 1A in supporting the word-level hypothesis; however, the within-word-foot interpretation is not available here, as the within-word foot containing the test syllable is monosyllabic in all cases.

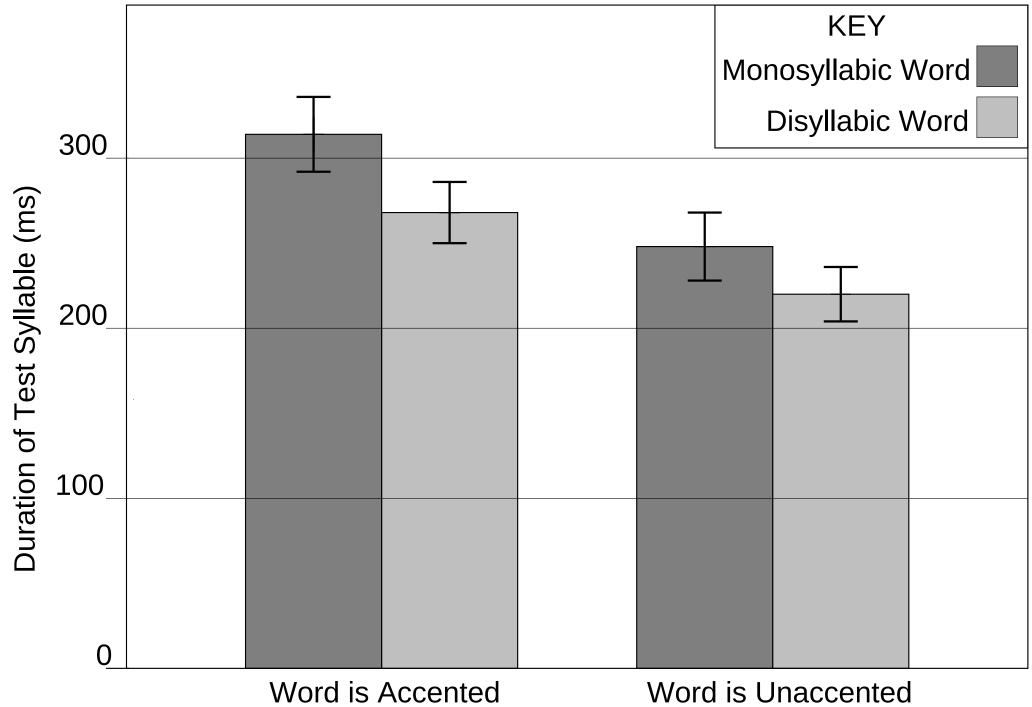

Figure 3.3: Mean test syllable duration (ms) in Experiment 1B: left-headed words; full central syllable.

As for Experiment 1A, the effect of Word Length on test syllable duration is consistent across all the materials: mean test syllable duration for each phrase pair (averaged over both accent conditions) are given in Appendix B.

The lengthening effects of Pitch Accent on the test syllable are evident in this experiment: the main effect of Pitch Accent is highly significant. Unlike Experiment 1A, however, there is no significant interaction between Word Length and Pitch Accent in the By-Subjects analysis. Planned comparisons confirm that the effect of Pitch Accent is highly significant ($p < .001$) on the test syllable in both monosyllables and disyllables. The effect of Word Length on the test syllable is significant ($p < .05$) in the unaccented condition and highly significant ($p < .001$) in the accented condition. It may be noted that the interaction between Word Length and Pitch Accent is significant ($p < .01$) in the By-Items analysis shown in Table 3.14.

3.4.3 Experiment 1C: right-headed words; reduced central syllable

Figure 3.4 shows mean test syllable duration for Experiment 1C according to the experimental conditions of Word Length and Pitch Accent. As for the left-headed words, the test syllable is longer in monosyllables than in disyllables: for example, /fıl/ is longer in *Phil* than in *fulfil*. The main effect of Word Length is significant, as shown

	Degrees of freedom	F Ratio	Significance level
By-Subjects analysis			
Word Length	1,5	13.18	$p < .025$
Accent	1,5	90.45	$p < .001$
Word Length × Accent	1,5	1.79	NS ($p > .05$)
By-Items analysis			
Word Length	1,7	24.82	$p < .01$
Accent	1,7	118.28	$p < .001$
Word Length × Accent	1,7	21.00	$p < .01$

Table 3.14: ANOVA summary table for Experiment 1B.

in Table 3.15. This provides further support for the word-level hypothesis. It may be noted that both in absolute and proportional terms, the effect of word length is smaller than in the left-headed materials, as discussed in Section 3.4.5 below.

Mean stressed syllable duration for each phrase pair are given in Appendix B, which shows that although the effect of Word Length is small, it is consistent across almost all of the materials. For example, the mean duration of /fɔɹm/ in "paper form" is 343 ms compared with 326 ms in "pay perform". The one exception to this trend is /fɔɹs/, which has a mean duration of 307 ms in "bacon force" and 311 ms in "bake enforce".

The lengthening effect of Pitch Accent on test syllables is evident in Figure 3.4: this effect is highly significant, as shown in Table 3.15. The interaction between Word Length and Pitch Accent is not significant, although the value of the F Ratio (5.04, *df* = 1,5) approaches significance. Planned comparisons indicate that the effect of Pitch Accent is highly significant ($p < .001$) in both monosyllables and disyllables. The effect of Word Length is not significant in unaccented words, but is significant in accented words ($p < .005$). This suggests the interaction between Pitch Accent and Word Length is real, but that there is insufficient data to attain significance (the interaction is significant in the By-Items analysis).

3.4.4 Experiment 1D: right-headed words; full central syllable

Figure 3.5 shows mean test syllable duration for Experiment 1D according to the experimental conditions of Word Length and Pitch Accent. The pattern of results is very similar to that for Experiment 1C, with the test syllable longer in monosyllables than in disyllables: for example, /saɪz/ is longer in "kneecap size" than in "knee capsize".

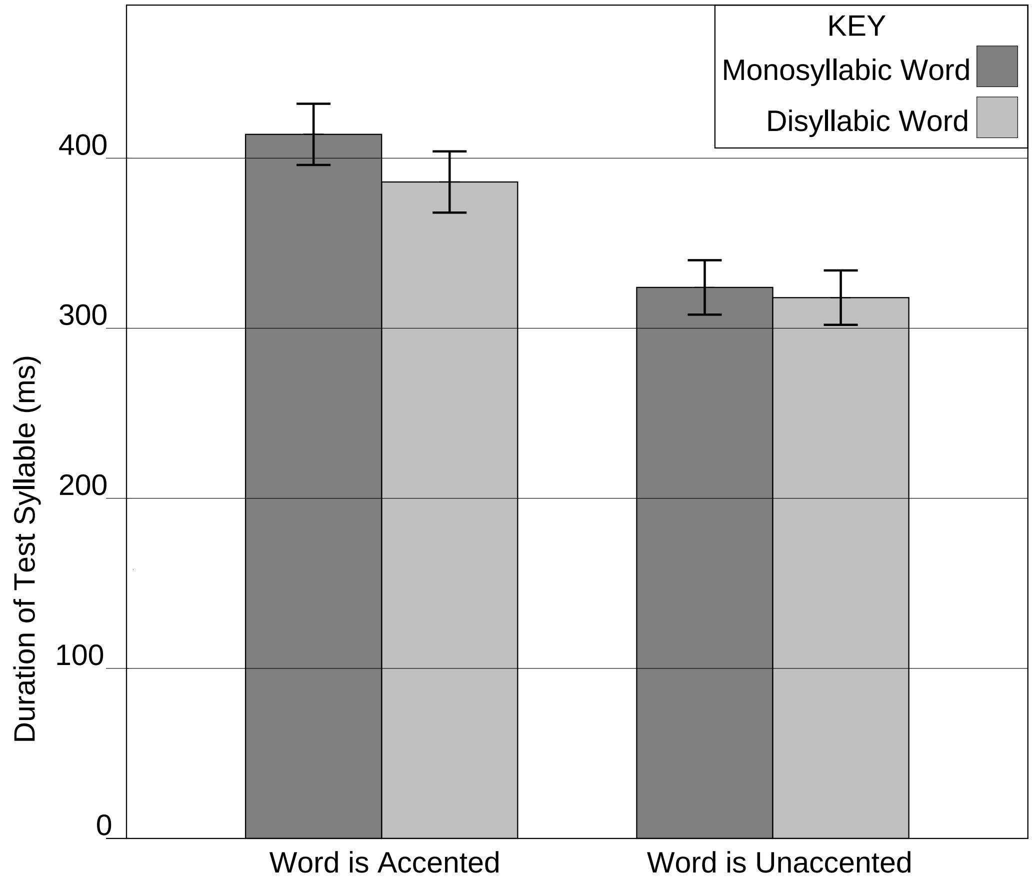

Figure 3.4: Mean test syllable duration (ms) in Experiment 1C: right-headed words; reduced central syllable.

As in Experiment 1C, the effect is smaller than in the left-headed materials, but is consistent across phrases, as shown in Appendix B. The main effect of Word Length is significant, as shown Table 3.16.

Once again, there is a large lengthening effect of Pitch Accent on the test syllable, which is highly significant, as shown in Table 3.16. Unlike Experiment 1C, the interaction between Word Length and Pitch Accent is also significant. The nature of this interaction is illustrated by planned comparisons. The effect of Pitch Accent is highly significant in monosyllables and disyllables. The effect of Word Length is significant in accented words ($p < .05$), but not in unaccented words. This is the pattern also observed in Experiment 1C.

	Degrees of freedom	F Ratio	Significance level
By-Subjects analysis			
Word Length	1,5	7.08	$p < .05$
Accent	1,5	165.21	$p < .001$
Word Length × Accent	1,5	5.04	NS ($p > .05$)
By-Items analysis			
Word Length	1,7	11.31	$p < .025$
Accent	1,7	285.39	$p < .001$
Word Length × Accent	1,7	8.49	$p < .025$

Table 3.15: ANOVA summary table for Experiment 1C.

	Degrees of freedom	F Ratio	Significance level
By-Subjects analysis			
Word Length	1,5	9.48	$p < .05$
Accent	1,5	257.21	$p < .001$
Word Length × Accent	1,5	6.93	$p < .05$
By-Items analysis			
Word Length	1,7	16.94	$p < .01$
Accent	1,7	334.97	$p < .001$
Word Length × Accent	1,7	1.23	NS ($p > .05$)

Table 3.16: ANOVA summary table for Experiment 1D.

3.4.5 General discussion of the experimental results

Evidence for a word-level effect

The results of the experiments support the word-level hypothesis. In all four sets of materials, the duration of the test syllable is greater in monosyllables than in disyllables. The magnitude of the effect is not constant, however. Figure 3.6 shows mean test syllable duration for all four experiments in the accented condition. In all these comparisons between monosyllables and disyllables, the effect of word length is significant, but is greater in the left-headed materials (Experiments 1A and 1B) than in the right-headed materials (Experiments 1C and 1D), in both absolute and proportional terms.

In the accented condition of Experiment 1A, with materials such as *thank* vs *thank-*

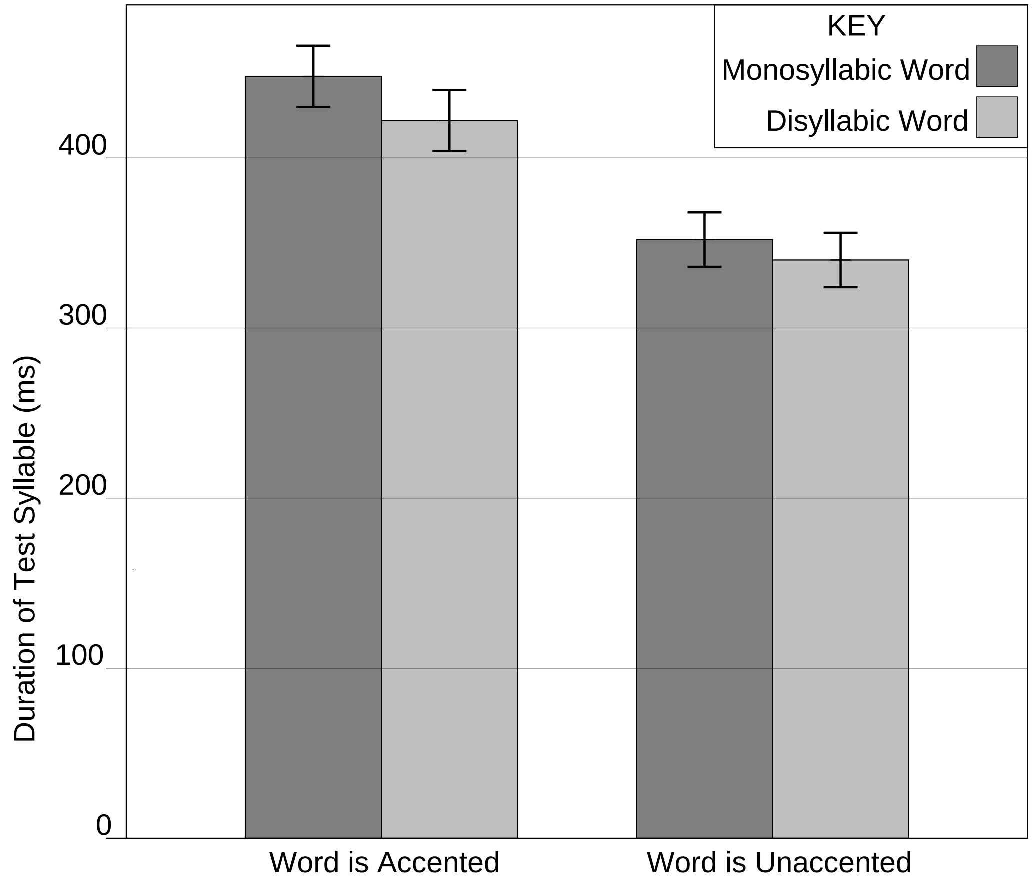

Figure 3.5: Mean test syllable duration (ms) in Experiment 1D: right-headed words; full central syllable.

ful, mean test syllable duration is 50 ms less in disyllables than in monosyllables, a "polysyllabic shortening" effect of 21%. Mean test syllable duration in the accented condition of Experiment 1B is slightly greater, possibly influenced by stress-adjacent lengthening, as discussed below. The 47 ms difference between monosyllables and disyllables—for example, *shake* vs *shakedown*—is very similar to the difference in Experiment 1A and represents a 15% "polysyllabic shortening" effect.

The test syllables in the right-headed materials are substantially longer than those in the left-headed materials, probably due to phrase-final lengthening, as discussed below. The size of the word-length effect is smaller than for the left-headed materials. In the accented condition of Experiment 1C, mean test syllable duration is 29 ms greater in monosyllables such as *Phil* than disyllables such as *fulfil*. This represents a 7% "polysyllabic shortening" effect. In the accented condition of Experiment 1D, mean test syllable duration is 25 ms greater in monosyllables such as *stairs* than in disyllables such as *downstairs*, a 6% "polysyllabic shortening" effect.

This quantitative variation in the effect of word length could be interpreted in

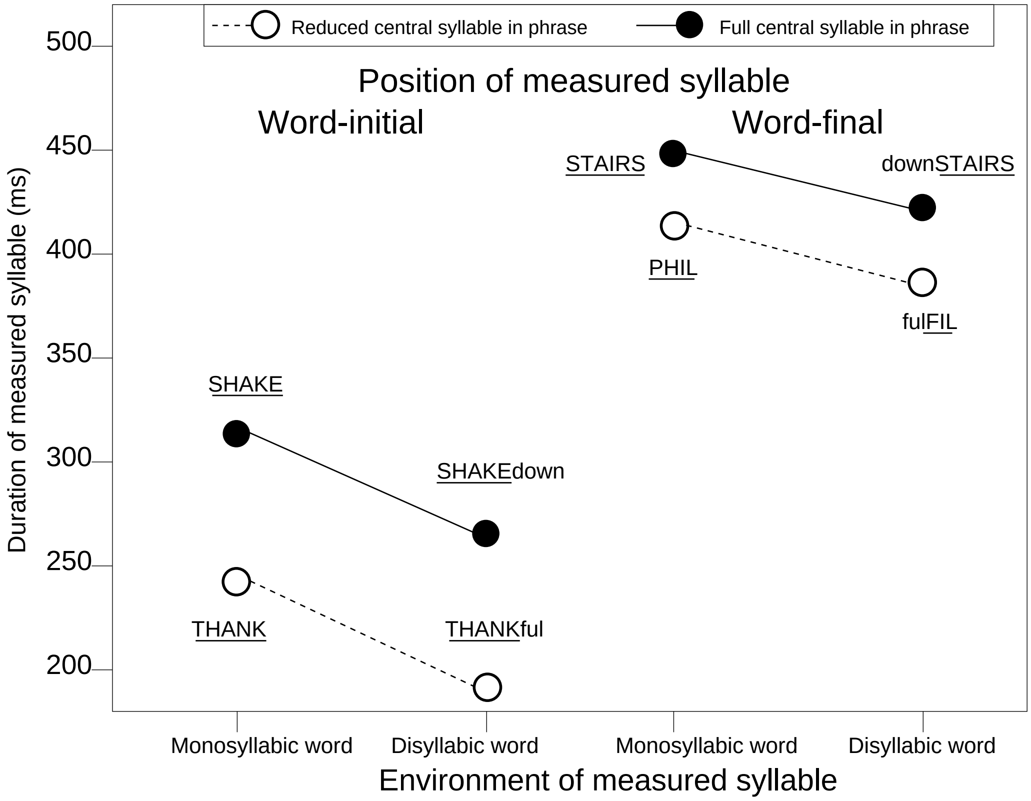

Figure 3.6: Mean test syllable duration (ms) in the accented condition. Example words are given for each mean duration and the test syllable is underlined.

at least two ways. Possibly syllables that are lengthened due to phrase-finality are less susceptible to word-span compression. Alternatively, the differences between left-headed and right-headed words may arise from different underlying processes: specifically, the observed word-level effect may be a result of the position of the test syllable within the word rather than the size of the word. These two hypotheses are discussed further in Section 3.5.

Figure 3.7 shows mean test syllable durations for all four experiments in the unaccented condition. For the left-headed materials, the pattern resembles that for the accented condition, although the effect of word length is smaller in absolute and proportional terms. In the unaccented condition of Experiment 1A, mean test syllable duration is 29 ms greater in monosyllables such as *thank* than in disyllables such as *thankful*, a 14% "polysyllabic shortening" effect. In the unaccented condition of Experiment 1B, mean test syllable duration is 27 ms greater in monosyllables such as *shake* than in disyllables such as *shakedown*, an 11% "polysyllabic shortening" effect.

For the right-headed materials, the effect of word length is very small in the unaccented condition, and comparisons of test syllable duration between monosyllables and disyllables show that the differences are not significant in either experiment. In

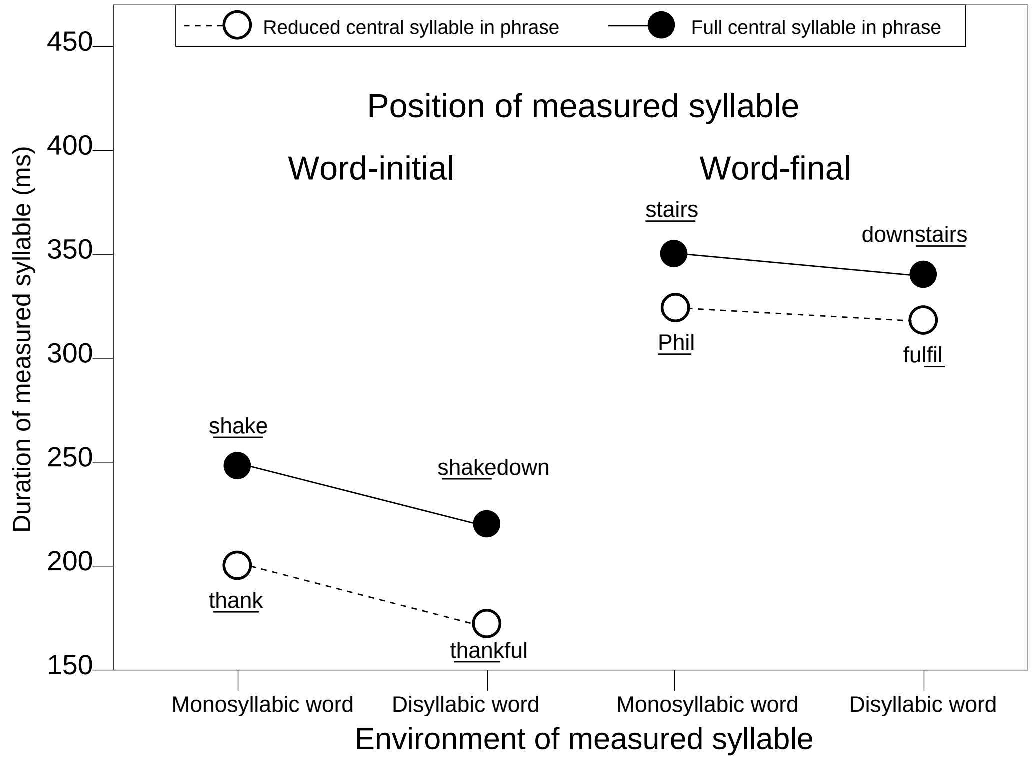

Figure 3.7: Mean test syllable duration (ms) in the unaccented condition. Example words are given for each mean duration and the test syllable is underlined.

the unaccented condition of Experiment 1C, mean test syllable duration is 7 ms greater in monosyllables such as *Phil* than in disyllables such as *fulfil*, a 2% "polysyllabic shortening" effect. In the unaccented condition of Experiment 1D, mean test syllable duration is 12 ms greater in monosyllables such as *stairs* than in disyllables such as *downstairs*, a 3% "polysyllabic shortening" effect.

One interpretation of this finding that accenting simply amplifies the word-length effect, by between 13 ms (Experiment 1D) and 21 ms (Experiment 1A): by this reasoning, the lack of a significant effect in the unaccented condition for the right-headed materials is due to the size of the sample being insufficient to demonstrate an effect which is present but very small. An alternative interpretation is that the word-length effect only occurs when the word carries a pitch accent: this is consistent with the data from the right-headed materials but not from the left-headed materials. These hypotheses will be discussed further in Section 3.5.

Evidence for a within-word-foot effect

Taken alone, the result of Experiment 1A supports both the word-level hypothesis and the within-word-foot hypothesis. In the other three experiments, however, the word-level hypothesis can account for the pattern of durational variation, subject to the qualifications just discussed, but the within-word-foot hypothesis cannot. Thus it is parsimonious to reject the within-word-foot hypothesis.

The effect of pitch accent

The lengthening effect of pitch accent is clear: in all four experiments, for both monosyllables and disyllables, the test syllable is longer in a word receiving contrastive phrasal stress, the primary prosodic marker of which is a pitch accent. This result is in line with previous findings; what is more interesting is the interaction between pitch accent and word length.

The nature of this interaction is discussed above in terms of the effect of word length in the presence or absence of pitch accent. An alternative way to analyse the interaction is by considering the amount of lengthening due to pitch accent in monosyllables and disyllables. Table 3.17 shows the amount of accentual lengthening in these contexts for each experiment, clearly indicating that test syllables in disyllables are lengthened less due to pitch accent than the same syllables in monosyllables, although the effect of accent is significant in all cases. The proportional lengthening effects for monosyllables and disyllables are comparable between the experimental materials in Experiments 1B, 1C and 1D, with 27–28% accentual lengthening in monosyllables and 21–24% accentual lengthening in disyllables. In Experiment 1A, the lengthening effect of accent is smaller in both contexts, although the difference between monosyllables and disyllables remains.

		Unaccented duration (ms)	Accentual lengthening: Magnitude (ms)	Proportion
Experiment 1A				
Monosyllable	*bake enforce*	201	41	20%
Disyllable	*bacon force*	172	20	12%
Experiment 1B				
Monosyllable	*knee capsize*	247	67	27%
Disyllable	*kneecap size*	220	47	21%
Experiment 1C				
Monosyllable	*bacon force*	324	91	28%
Disyllable	*bake enforce*	317	69	22%
Experiment 1D				
Monosyllable	*kneecap size*	351	96	27%
Disyllable	*knee capsize*	339	83	24%

Table 3.17: Accentual lengthening effect of test syllables in monosyllables and disyllables in each experiment. Example phrase pairs are shown for each context, with the test syllable underlined.

The effect of word length on accentual lengthening of the test syllable is very likely a reflection of the finding, in studies such as Turk & White (1999), that other syllables in polysyllabic words also undergo accentual lengthening. An unstressed or secondary stressed syllable which follows the primary stress in a disyllable is lengthened by about 13% when the word is accented. A syllable preceding the primary stress in a disyllable undergoes about 4% accentual lengthening. Thus, accentual lengthening is greatest on the primary stressed syllable, but also affects other syllables within the word, and more strongly those following the primary stress.

Given such results, it might be expected that the attenuation of accentual lengthening in disyllables would be greater for left-headed words, where there is more lengthening on the other syllable, than for right-headed words. That this does not appear to be the case argues against a representation of accentual lengthening which posits a fixed amount of additional duration to be distributed among eligible syllables.

Accentual lengthening of the test syllable in Experiment 1A is less, in both word-length contexts, as shown in Table 3.17. A reason for this is that the test syllable in Experiment 1A is relatively short (being unaffected by stress-adjacent lengthening or phrase-final lengthening, as discussed below). Thus, in line with Weber's law of psychophysics, a smaller difference in duration would be required to make a perceptible change[14]. If this explanation is sufficient, then there is no need to regard the materials in Experiment 1A as distinctive, further undermining the within-word-foot hypothesis.

Another explanation of the smaller accentual lengthening effect in Experiment 1A is suggested by examination of the experimental materials: the carrier sentence for the unaccented condition in Experiment 1A has a pitch-accented word preceding the test syllable, for example:

SAY "bake enforce", don't SHOUT "bake enforce".

Turk & White (1999) suggest that a small amount of accentual lengthening can be found on the syllable immediately following an accented syllable, even where a word boundary intervenes. Thus in Experiment 1A, the unaccented target syllable might be slightly longer, due to the foregoing accent, than it would be in the absence of any durational influence of accent. Two observations suggest that, if true, this may only be a partial account of the attenuated accentual lengthening seen in Experiment 1A. Firstly, the materials are comparable in Experiment 1B, where no such attenuation is observed: accentual lengthening of the test syllable there is equivalent to that seen in

[14] In psychophysical experiments, the concept of the just noticeable difference (JND) is based upon consecutively-presented stimuli. For such a notion to operate in on-line prosodic processing, there must be an assumption that listeners can compare perceptions with expectations derived from foregoing context. According to this assumption, the listener generates an expectation of the duration of a given syllable using knowledge of speech rate and speaker characteristics. If the observed duration is longer than the expectation by more than the JND, the listener may infer some kind of suprasegmental marking.

Experiments 1C and 1D. Secondly, the boundary between *SAY* and the test syllable in Experiments 1A and 1B is likely to be stronger than that between the first two words of phrases such as *SEND Mark now*, the experimental context used by Turk & White (1999) to find the small cross-boundary effect. A stronger boundary may more greatly attenuate the cross-boundary effect, rendering an already small effect nugatory.

Evidence for the stress adjacency effect

The results relating to the stress adjacency effect are only indicative of a possible trend, as the segmental material is different between the phrases containing an unstressed central syllable and those containing a secondary stressed central syllable. The difference in mean test syllable duration between Experiments 1A and 1B is quite large, however, and suggests support for studies such as Van Lancker *et al.* (1988) and Rakerd *et al.* (1987). Mean test syllable duration in Experiment 1A, when it is followed by an unstressed syllable, is 187 ms in the unaccented condition and 216 ms in the accented condition. Mean test syllable duration in Experiment 1B, when it is followed by a secondary stressed syllable, is 234 ms in the unaccented condition and 291 ms in the accented condition. Furthermore, previous studies suggest the possibility that stress-adjacent lengthening is attenuated where the test syllable is pitch accented. In the current data, the magnitude of the effect appears to be greater for pitch-accented syllables. This finding is particularly surprising: the absolute magnitude of the effect cannot be gauged from these unmatched materials, but the direction of the trends is contrary to that previously observed.

It may be noted, with the foregoing qualification regarding segmental differences, that phrase-final test syllables are longer in the phrases containing a secondary stressed central syllable (for example, *knee capsize*) than in those containing an unstressed central syllable (for example, *thank fulfil*). Mean test syllable duration in Experiment 1C, when it is preceded by an unstressed syllable, is 321 ms in the unaccented condition and 400 ms in the accented condition. Mean test syllable duration in Experiment 1D, when it is preceded by a secondary stressed syllable, is 345 ms in the unaccented condition and 434 ms in the accented condition. This supports the interpretation discussed in Chapter 2 that previous observations indicate stress-adjacent lengthening, rather than foot-final lengthening: only the first stressed syllable would be affected by a foot-final effect, whereas both stressed syllables could be affected by an adjacency effect. All the results pertaining to this effect are merely suggestive, however; no firm conclusions can be drawn in the absence of phonetically-matched materials.

Other sources of durational variation

As outlined in Section 3.2.1, this experiment is designed to test the word-level and within-word-foot hypotheses, whilst controlling the influences of other factors thought to have been potential confounds in previous experiments. The effects of pitch accent and stress-adjacency have already been discussed; two further influences are also considered—higher constituent length and phrase-final lengthening.

In this experiment, the influence of these factors upon test syllable duration is considered constant. This is predicated on the assumption that subjects realise the two-word three-syllable phrase containing the test syllables as a single prosodic phrase, such as a phonological phrase. Thus, in sentences such as:

- Say "BAKE enforce", don't say "TANK enforce".
- Say "bake ENFORCE", don't say "bake REMOVE".
- SAY "bake enforce", don't SHOUT "bake enforce".

it is assumed that no prosodic boundary larger than a word boundary is realised between the words in the first phrase in quotation marks. Subjects were instructed not to pause at this point, and utterances containing perceptual evidence of a prosodic break were rejected from the analysis.

This assumption does not mean that the test syllables were not influenced by higher prosodic boundaries, simply that such influences were considered to be fixed. Indeed, it is noticeable that the duration of test syllables in phrase-final position is considerably greater than in phrase-initial position. Although a direct comparison cannot be made between the two positions because the segmental composition of the syllables is different, it seems very likely that much of this difference may be attributed to phrase-final lengthening. As discussed in Chapter 2, the lengthening that occurs at the end of major prosodic constituents is well established. The structure of the experimental sentence materials used here would suggest the placement of an intonational phrase boundary at that point in the sentence. The possibility that phrase-final lengthening may be a factor in the variable magnitude of the word-level effect is discussed in Section 3.5 below.

In this experiment, assuming no major prosodic break within the test phrase, the length of any phrasal constituent dominating the word containing the test syllable is kept constant. The effect of variation in the length of the utterance is examined in Experiment 2.

3.5 Discussion

3.5.1 The mechanism of the word-level effect

The main finding from this experiment is that the location of word boundaries influences primary stressed syllable duration. This represents confirmation of the word-level hypothesis, but the mechanism underlying the observation remains unclear. The result in this experiment resembles a previous study of Swedish, in which "adding syllables after the main stress causes more drastic shortening than attaching them before" (Lindblom 1968:20); Nooteboom (1972) finds a similar pattern for Dutch.

Such results argue against the description of "polysyllabic shortening" presented in Section 3.1, which frames it as a domain-span process at the word-level: in that account a temporal compression is exerted on the primary stressed syllable which is proportional to the number of additional syllables in the word, whether before or after the primary stress. In abstract terms, one could imagine an algorithm which allots each word in a sentence a fixed duration based upon parameters determined by factors such as its position, its prominence, its structural relationships with the other words in the sentence, and the overall rate of speaking. Once the word has been given its temporal allotment, another algorithm determines how this time is shared out between the sub-word constituents. Clearly a syllable in a monosyllable would take all the time allotted the word, whereas in a disyllable there would be less time available to any single syllable, and in a trisyllable even less. Constraints on compressibility would mean that polysyllables would tend to be longer than temporally-equivalent monosyllables, despite their isochrony in the underlying representation.

This idealisation serves to illustrate one theoretical account of the experimental observations. An additional or alternative explanation is required to explain the difference between left-headed and right-headed words in the magnitude of the word-level effect. Possible interpretations are:

An interaction with final lengthening. Phrase-final lengthening of the test syllable in the right-headed words may attenuate the word-span compression effect: evidence from Turk & Shattuck-Hufnagel (2000) suggest, however, that this cannot be the full explanation.

Two domain-span processes. Word-span compression (polysyllabic shortening), plus an additional domain-span compression process in a sub-word constituent, the "word-rhyme", beginning with a primary stress and continuing to the word boundary: studies suggest that this constituent may be important in phrase-final lengthening and accentual lengthening.

Two domain-edge processes. A small word-initial lengthening effect plus a larger word-final lengthening effect: other studies indicate that evidence for word-initial lengthening is stronger than for word-final lengthening.

A combination of edge and span processes. Word-rhyme-span compression, with a syllable nucleus locus and word-initial lengthening with a syllable onset locus: Turk & Shattuck-Hufnagel (2000)'s results suggest these processes may be accompanied by a word-span compression process with a stressed syllable locus.

Arguments for and against these interpretations are considered now.

An interaction with final lengthening

The difference in the size of the word-level effect in right-headed and left-headed words may be an experimental artefact: the syllables in the right-headed materials are phrase-final, and their relatively large duration suggests phrase-final lengthening; it may be that syllables lengthened due to phrase-finality are less susceptible to a word-span compression effect. Price *et al.* (1991) suggest that phrase-final syllables do not gain as much duration due to pitch accent as non-phrase-final syllables. It may be that syllables in phrase-final position are more resistant to both "lengthening", such as due to accent, and "shortening", such as due to word length.

The design of Turk & Shattuck-Hufnagel's (2000) experiment is similar to that of Experiment 1, and investigates various word-level durational mechanisms including word-initial and word-final lengthening, and polysyllabic shortening. They measure the durations of all three syllables in pairs of test phrases such as *tune acquire* and *tuna choir*, in three accent conditions analogous to those in Experiment 1, but the carrier sentence they use makes it less likely that the final syllable in the test phrase will be followed by a phrase boundary, for example:

SHOUT "tuna choir" again; don't SAY "tuna choir" again.

The test syllable (underlined) is taken from the second phrase in quotation marks and is followed by *again* before the utterance-final boundary. In this context, it is unlikely that there will be a major boundary such as an intonational phrase boundary following the test-phrase-final syllable. Their data suggest this interpretation: although phrase-initial and phrase-final syllables are not phonetically matched, the mean durations are taken from 11 differently-composed phrase pairs and show no evidence of the large difference between phrase-initial and phrase-final syllables found in Experiment 1.

Despite the apparent lack of phrase-final lengthening on the test-phrase-final syllable in Turk & Shattuck-Hufnagel's experiment, they also find some discrepancy in the word-length effect between left-headed and right-headed words, at least in the conditions where there is a pitch accent within the test phrase: thus /tun/ is almost

10% longer in *TUNE acquire* than in *TUNA choir*, but /kwaɪr/ is only 5% longer in *tuna CHOIR* than in *tuna ACQUIRE*, indicating that the possible interaction with final lengthening in Experiment 1 is not the full explanation of the observation.

Two domain-span processes

Another explanation of the difference between left-headed and right-headed words is that the difference reflects the operation of two domain-span processes: a word-level effect evident in all the materials, and an effect in a subword unit which only occurs in the left-headed materials. As the compression effect on the primary stress in left-headed disyllables occurs both when the additional syllable is unstressed (for example: *thankful* vs *thank*) and when it has secondary stress (for example: *kneecap* vs *knee*), this unit could not be the within-word foot, but rather a unit that begins with a primary stress and continues to a word boundary. There is some evidence of such a unit having durational consequences: accentual lengthening appears to affect syllables preceding and following the primary stressed syllable within a word, but the greater amount of lengthening is manifest on following syllables, whether unstressed or carrying secondary stress (for example: Turk & White 1999). Furthermore, Turk (1999) finds phrase-final lengthening in accented words with a locus comprising a primary stressed syllable and another syllable to its right, either unstressed or with secondary stress. If such a unit, which may be termed the "word-rhyme", is important as a locus of accentual lengthening and phrase-final lengthening, it may also demonstrate a domain-span compression effect[15].

Turk & Shattuck-Hufnagel (2000) consider the possibility that the asymmetry in their results, comparable to that found by Lindblom (1968) for Swedish, is due to an asymmetrical polysyllabic shortening effect, which may be seen as functionally equivalent to the two-process domain-span model just discussed, which proposes a word-span compression effect and a word-rhyme-span compression effect. They reject this hypothesis, however, because they do not find strong evidence of variation in the duration of the unstressed central syllable in phrases such as *tune acquire* and *tuna choir*. This syllable shows a significant durational difference between these contexts only when the preceding syllable is accented, seemingly the result of the rightward spread of accentual lengthening being attenuated at the word boundary (Turk & White 1999). No other word-level mechanism appears to influence the duration of the central unstressed syllable between these two contexts, which could either indicate the absence

[15] It should be noted that not all parts of the word-rhyme are affected by accentual lengthening and final lengthening to the same degree. In particular, studies discussed in Chapter 2 indicate that lengthening of a phrase-final stressed syllable—corresponding to a monosyllabic word-rhyme—does not affect the syllable onset. In contrast, the onset shows a large accentual lengthening effect, at least in monosyllabic words.

of a word-level effect or the operation of a symmetrical effect, contrary to the evidence of stressed syllable duration. One possible solution is to postulate that the locus of the domain-span processes at the word level or the word-rhyme level (in the two-process model outlined above) is the primary stressed syllable. Two hypotheses are suggested in Chapter 1 regarding the locus of domain-span processes: firstly, the locus is co-extensive with the domain, thus all the subconstituents are subject to compression proportional to domain-length; secondly, the locus is the phonological head of the domain. These findings appear to favour the latter interpretation.

Turk & Shattuck-Hufnagel propose a different solution to the asymmetry problem. They suggest that a symmetrical word-span effect may be accompanied by what they call "syllable ratio equalisation". This relates to the observation by Abercrombie (1965) of durational variation within the near-homophonous phrases (in Standard Southern English):

1. *Take Greater London*

2. *Take Grey to London*

Abercrombie says that within the cross-word foot /ˈɡɹeɪtə/ the "quantity" of the two syllables is more similar in (1) than in (2); in terms of poetical feet, *Greater* is a spondee (two equal syllables) and *Grey to* is a trochee (a long syllable followed by a short syllable). Albrow (1968) relates what Abercrombie calls syllable quantity to duration: thus, the syllables in *Greater* should have more similar durations than those in *Grey to*. While this pattern is likely to be observed, other mechanisms can account for the observation without requiring a mechanism that explicitly attempts to equalise syllable durations, a goal which seems doomed to failure given that one syllable contains a full vowel and the other contains a reduced vowel. Polysyllabic shortening or word-final lengthening predict the main durational difference: /ɡɹeɪ/ is shorter in *Greater* than in *Grey*. This fact alone would cause the two syllables in *Greater* to have more similar durations than those in *Grey to*, and this process may be supplemented by the unstressed /tə/ gaining some duration as a result of belonging to a content word, possibly accented, rather than a function word. In Turk & Shattuck-Hufnagel's experiment, the lack of a word-level effect on the unstressed central syllable leads them to posit syllable ratio equalisation and suggest that it only affects the first syllable of the sequence: thus /tun/ is shortened in *tuna* compared with *tune*, but the duration of /ə/ is unaffected by constituency in *tuna* or *acquire*. As stated above, proposing a within-word domain-span process with a primary stressed syllable locus makes similar predictions about syllable durations without requiring an additional *type* of process. Thus parsimony suggests that syllable ratio equalisation should be discounted unless it can account for results that are not explicable in terms of existing processes—for example, domain-edge or domain-span processes—for which independent evidence exists. Some further evidence for this process put forward by Turk & Shattuck-Hufnagel (2000) is discussed below.

It may be noted that, although the durational variation reported by Turk & Shattuck-Hufnagel is in the direction suggested by syllable ratio equalisation, the mean duration of the unstressed central syllable is never more than a third of that of the preceding stressed syllable. Abercrombie's original observation, while having some basis in observable durational variation, somewhat overstates the significance of the process: the phonetic facts strongly indicate that variation in the placement of the word boundary does not cause /ˈɡɹeɪtə/, a trochee in *Grey to*, to become a spondee in *Greater*[16]. The unstressed syllable undoubtedly remains unstressed and contains a short, reduced vowel in both contexts.

Two domain-edge processes

A third interpretation of the discrepancy between left-headed and right-headed words observed in the present experiment is that the durational differences reflect distinct *domain-edge* processes: thus word size is significant because additional syllables change the alignment of the test syllable with word boundaries. According to this interpretation of Experiment 1, there is a small word-initial lengthening effect—for example, in *Phil* compared with *fulfil*—and a large word-final lengthening effect—for example, in *thank* compared with *thankful*.

As discussed in Chapter 2, there is good evidence for a word-initial lengthening effect, probably with a syllable-onset locus, but evidence for the existence of a word-final lengthening effect in the absence of phrase boundaries is more equivocal. Turk & Shattuck-Hufnagel (2000) conclude that their data do not provide support for word-final lengthening. Firstly, the duration of the unstressed central syllable is not consistently longer word-finally than word-initially; however, this could also be interpreted as a balance of word-initial lengthening and word-final lengthening. Secondly, the distribution of lengthening in /tun/ in *tune* compared with *tuna* is not progressive, as might be expected with final lengthening: the greatest difference is found on the syllable centre with less in absolute and proportional terms on the syllable-final consonant[17].

[16]If *Greater* were really best classified as spondee, this would be expected to elicit stress-adjacent lengthening. Evidence reviewed in Chapter 2 suggests that this process affects the first stressed syllable more than the second and so, paradoxically, the equalisation would then entail a subsequent adjustment away from equality.

[17]In order to segment their materials reliably, Turk & Shattuck-Hufnagel (2000) divide their measured stressed syllables into the first consonant, the last consonant, and the vocalic nucleus plus remaining consonants, the latter being termed the "syllable centre". Thus, if the onset or coda contain consonant clusters, the syllable centre is not isomorphic with the nucleus.

It seems then that the small word-level effect observed in the right-headed materials in this experiment might be explicable in terms of word-initial lengthening of the syllable onset; however, explaining the much larger word-level effect in the left-headed materials in terms of word-final lengthening is not well supported by the evidence of previous research. As discussed in Section 3.4.5, it is assumed that no boundary larger than a word boundary intervenes between the two words of the test phrases in this experiment. If this assumption is valid, the explanation of the durational difference between the first syllables of *thank fulfil* and *thankful Phil* remains an open question. The most likely account appears to be some form of domain-span compression effect, possibly over the word rhyme.

An alternative candidate is syllable ratio equalisation, as discussed above: Turk & Shattuck-Hufnagel have a further reason for proposing that this mechanism, in conjunction with other processes, accounts for their results. They find that the subsyllabic distribution of word-level shortening is different in left-headed and right-headed words, with most of the variation in comparisons such as *tune* vs *tuna* occurring on the stressed syllable centre, and in comparisons such as *choir* vs *acquire* on the stressed syllable onset—due to word-initial lengthening—with not much difference in variation between the centre and the final consonant. They propose that syllable ratio equalisation, affecting only the left-headed words, might be focused on the stressed syllable nucleus, whereas the symmetrical polysyllabic shortening effect might be more evenly distributed: thus, the distribution of the word-level effects in pairs such as *tune* vs *tuna* would be a result of polysyllabic shortening and syllable ratio equalisation, and pairs such as *choir* and *acquire* would manifest the effects of word-initial lengthening and polysyllabic shortening, with the lengthening effects of accent also to be considered where present. The two-process model proposed above could accommodate this result if the locus of polysyllabic shortening were the primary stressed syllable and the locus of word-rhyme compression were the nucleus of the primary stressed syllable.

A combination of edge and span processes

As just noted, Turk & Shattuck-Hufnagel (2000) propose a combination of symmetrical polysyllabic shortening, word-initial lengthening and syllable ratio equalisation to account for the variations in the distribution of durational effects that they observe in their materials, combined with the fairly well-established effects of pitch accent. Polysyllabic shortening is assumed to be relatively evenly distributed within the stressed syllable, whereas syllable ratio equalisation is proposed to affect primarily the syllable nucleus.

An alternative, which could account for the results of Experiment 1, is a combination of word-rhyme-span compression and word-initial lengthening: the former accounts for the difference in *thank* vs *thankful* and the latter for the difference in *Phil* vs *fulfil*. This combination of processes could also account for the main trends in Turk & Shattuck-Hufnagel's data, although the subsyllabic distribution of effects suggests that a small additional polysyllabic shortening effect is required, which appears to affect all parts of the stressed syllable. Word-rhyme compression appears to have a syllable nucleus locus: as noted above, the greatest difference in *tune* vs *tuna* is on the stressed syllable nucleus. Word-initial lengthening appears to have a syllable onset locus: thus the greatest difference in *choir* vs *acquire* is on the stressed syllable onset.

A third alternative is asymmetric polysyllabic shortening combined with word-initial lengthening. The choice between these alternative combinations of processes is partly a question of the relative theoretical simplicity of the competing hypotheses. As mentioned above, syllable ratio equalisation is a different type of process to domain-edge and domain-span processes, thus it is parsimonious not to propose its existence if established processes are available as alternatives. The choice between asymmetric polysyllabic shortening and a two-process domain-span model may be made on similar grounds. Although parsimony suggests that one process is better than two, the two-process model proposes two effects that have a locus which is within the head of the domain—the locus appears to be the stressed syllable for polysyllabic shortening and the stressed syllable nucleus for word-rhyme-span compression—and have similar relationships between locus duration and domain-size. The asymmetrical process, by definition, lacks internal consistency, with a change in word-length to the left of the locus affecting it differently from a change in word-length to its right. Furthermore, both the domains in the two-process model have independent support: the word, whether syntactically or prosodically defined; and the word rhyme, which may be a locus of both accentual lengthening and phrase-final lengthening.

The number and type of word-level processes are examined further in Experiment 2, which looks at how subsyllabic durations vary in monosyllables, disyllables and trisyllables.

3.5.2 The interaction between pitch accent and word-level effects

Another difference between the left-headed and right-headed materials lies in the interaction with pitch accent. For left-headed materials, such as /ni/ in *knee capsize* vs *kneecap size*, the effect of word length is significant for syllables in both accented and unaccented contexts, although the magnitude of the effect is less in the unaccented case. For right-headed materials, such as /saɪz/ in *kneecap size* vs *knee capsize*, the effect of word length is only significant in the accented context.

As mentioned in Section 3.4.5, this difference may be due to there being insufficient data for the small effect in Experiments 1C and 1D to be statistically significant in the unaccented condition. If, for example, the process underlying the observation is word-initial lengthening with a syllable-onset locus, the measured duration of the whole syllable would be consistently affected, but the proportion of lengthening would be less than on the onset considered alone. In this case, the interaction suggests that word-initial lengthening of the syllable onset is proportionately greater in accented words.

The alternative account of the interaction in the right-headed materials is that the word-level effect only operates in the presence of pitch accent. This would seem to require a different process underlying the observation in the left-headed materials. One conclusion could be that word-rhyme compression is amplified by pitch accent, but exists in the absence of accent, whereas word-initial lengthening is dependent upon the presence of pitch accent.

Previous studies indicate that the word-rhyme may be the locus of the greatest durational effects of pitch accent, as noted above; thus, an amplification of the domain-span effect might reasonably be expected in the presence of pitch accent. One interpretation of domain-span processes is as adjustments within the locus of lengthening effects—accentual lengthening and final lengthening—such that the additional duration is distributed in accordance with the amount of segmental materials within the locus. This experiment suggests, however, that the word-rhyme-span effect is not dependent on the presence of pitch accent or phrase-finality, as might be predicted from such a model.

The evidence for existence of word-initial lengthening is clear—syllable onsets are longer word-initially than word-medially—but the effect of pitch accent has not been explicitly examined. Both Oller (1973) and Cooper (1991) find that unstressed syllable onsets undergo proportionally more word-initial lengthening than stressed syllable onsets. For instance, Oller (1973) finds that the closure duration of /b/ is subject to about 30 ms onset lengthening in unstressed syllables and 20 ms in stressed syllables. Given the greater duration of the /b/ in stressed syllable context, the difference in the proportion of lengthening is even more marked. The effect of pitch accent may be inferred from the materials, which utilise reiterant-style nonsense words in sentence-final position; for example: *Say a bababab*, *Was it a bababab?* and *It was a bababab*. It is very likely that the nonsense word would receive a pitch accent here and thus the stressed syllable would also be a pitch-accented syllable.

Cooper (1991) suggests a similar effect: the closure durations of voiceless stops appear to undergo word-initial lengthening of a similar magnitude in stressed and unstressed syllables, which would mean a proportionally greater effect in unstressed syllables. The aspiration duration of voiceless stops is also longer word-initially in unstressed syllables; in stressed syllables, however, the aspiration duration is in most cases longer word-medially. The presence of a pitch accent on the stressed syllables may also be inferred from the materials in Cooper's study as well.

These results offer a slightly ambiguous picture of the interaction between word-initial lengthening and pitch accent: apparently, the measured syllable onsets are taken from accented words, but accented syllables themselves show less word-initial lengthening than do unstressed syllables in an accented word. Turk & Shattuck-Hufnagel (2000) look explicitly at the effect of accent and word position on stressed syllable onsets and find word-initial lengthening in all accent conditions, but most reliable and of greatest magnitude (about 10%) when the word is accented. In contrast, the aspiration duration of syllable-onset voiceless stops is reliably longer word-initially only where the word is unaccented. This finding is similar to those inferred from the studies of Oller and Cooper, where stressed syllable aspiration duration, measured in apparently pitch-accented words, does not show a word-initial lengthening effect, being in fact larger word-medially in most cases.

Thus a rather complicated picture emerges of interactions between lexical and/or phrasal stress and word position differentially affecting syllable-onset closure duration and aspiration duration. It does seem, however, that stressed syllable onsets may show greater word-initial lengthening of closure duration in pitch-accented words, a finding which appears to be consistent with the results of the present experiment. These issues are examined further in Experiment 2.

3.6 Summary

The word-level hypothesis, that there is a relationship between word length and stressed syllable duration, is supported by the results of this experiment, but the best description of this effect does not appear to be a simple word-span compression effect. One interpretation of the pattern of results in Experiment 1 is that two effects are present. Firstly, a domain-span compression effect over the word-rhyme, which extends from the onset of a stressed syllable up to the word-final boundary; comparison with the results of Turk & Shattuck-Hufnagel 2000 suggests that the locus of this effect may be the nucleus of the primary stressed syllable. Secondly, a domain-initial lengthening effect at the word-level, with a locus suggested by other studies to be the syllable onset. The present experiment suggests that this effect may be greatly amplified by, or possibly dependent on, the presence of pitch accent. The results of Turk & Shattuck-Hufnagel (2000) indicate that word-rhyme-span compression and word-initial lengthening may be accompanied by a small domain-span effect at the word-level (polysyllabic shortening).

Other explanations of the observations are available, and other studies present different conclusions about the precise nature of word-level durational processes. Experiment 2 is designed to test for domain-edge and domain-span processes at the word level and the utterance level.

Chapter 4

Experiment 2: Domain-edge and domain-span processes

4.1 Introduction

Experiment 2 tests the existence of domain-edge and domain-span processes at two levels of linguistic structure, the word and the utterance; the locus of accentual lengthening and its interaction with domain-edge and domain-span processes are also examined. The results—described in detail in later sections of this chapter—show evidence for domain-edge processes, at both word level and utterance level: word-initial lengthening, utterance-initial shortening and utterance-final lengthening, effects which do not appear affected by the presence of pitch accent. The results do not support domain-span processes at either word level or utterance level. The previously-observed word-span effect (polysyllabic shortening) is shown to be a result of accentual lengthening: the locus of lengthening is the word, and the distribution of the effect varies between monosyllables, disyllables and trisyllables, such that there is less accentual lengthening of the primary stressed syllable when there are more additional syllables in the word. There is evidence for only one domain-span effect, at the word-rhyme level, and this may be interpretable as a word-final effect.

4.1.1 Background

Experiment 1, in Chapter 3, shows that stressed syllables have greater duration in monosyllabic words than in disyllabic words. This result is observed when the word is unaccented or accented, and whether the additional syllable in the disyllable is unstressed or secondary stressed. The shortening effect is much greater for stressed syllables which are word-initial in the disyllable than for those which are word-final: for example, /pe/ is 25 % shorter in *paper form* than in *pay perform*; /fɔɹm/ is 5% shorter in *pay perform* than in *paper form*[1].

The asymmetry in these observations suggests that they may arise from a small word-initial lengthening process and a large word-final lengthening process, but other explanations are possible: for example, polysyllabic shortening, a word-span compression process. As discussed in Chapter 2, domain-span processes are also linked with constituents other than the word: in particular, an utterance-span compression process has been proposed, such that words are longer in shorter sentences.

Several syntactic and prosodic constituents are identified with domain-edge processes: studies discussed in Chapter 2 link final lengthening with words, various syntactic and prosodic phrases, and utterances; others find initial lengthening in words and sometimes in higher-level phrases. Some studies suggest that domain-edge processes are associated with a number of levels of a prosodic constituent hierarchy, with greater lengthening at higher-level constituent boundaries, and there is also evidence for an utterance-initial *shortening* effect.

The durational effects of pitch accent, discussed in Chapters 2 and 3, are also examined in Experiment 2. The interactions of accent with edge and span processes are considered, to determine whether accent simply amplifies existing effects, or alters how they are manifest by, for example, changing the locus.

Research on domain-edge and domain-span processes at the word level and the utterance level and on the influence of pitch accent is reviewed[2] in Section 4.1.2, and experimental hypotheses are formulated. Questions of experimental design raised by studies which present evidence for a hierarchy of domain-edge processes are discussed in Section 4.1.3.

4.1.2 Experimental hypotheses

The experimental hypotheses to be tested in Experiment 2 are presented here, and the previous research which suggests each hypothesis is briefly reviewed. The hypotheses relate to the concept of the locus in general, and to:

- Word-edge and word-span processes.
- Utterance-span processes.
- Utterance-initial processes.
- Utterance-final processes.
- Accentual lengthening, and its interaction with edge and span processes.

[1] The transcriptions are for the experimental subjects, Scottish English speakers from Edinburgh.

[2] Previous research is discussed in detail in Chapters 2 and 3.

For each set of hypotheses, the findings of Experiment 2 are briefly noted at the end of the subsection (indicated thus $\Longrightarrow$); full details are given in the appropriate results section, as indicated.

The locus

Domain-edge locus hypothesis: each domain-edge process has a characteristic phonologically-defined locus.

- The alternative hypothesis is that lengthening is greatest at the boundary, and decreases gradually with distance from the edge, without being confined to particular phonological constituents.

As defined in Chapter 1, the locus is the stretch of speech which undergoes the durational effect associated with a domain: for example, the word is a domain of initial lengthening and the locus is the onset of the word-initial syllable. Studies of speech timing often look at the effect of some condition upon the duration of a particular preselected unit: for example, in Experiment 1 and elsewhere, the effect of word length on stressed syllable duration is reported. This approach could mask the fact that certain durational effects may be focused at specific loci: domain-final lengthening may be restricted to the final syllable rhyme (for example: Oller 1973; Campbell & Isard 1991; Wightman *et al.* 1992); domain-initial lengthening may be restricted to the initial syllable onset (for example: Oller 1973; Fougeron & Keating 1997). Lengthening may increase within the locus towards the domain boundary (for example: Cambier-Langeveld 2000) but this distribution may vary according to segmental identity (Oller 1973). To measure the duration of a syllable or a foot or a word ignores the possible differential distribution within this unit, with some constituents affected strongly and others not at all.

Domain-span locus hypothesis: a domain-span compression process shortens only the phonological head of the domain.

- This may be interpreted either as an all-or-nothing distinction, or as a gradient effect, with lesser compression seen outside the phonological head of the domain. The alternative hypothesis is that all parts of the domain are equally affected by domain-span shortening.

The position of the locus of a domain-edge process is, by definition, the domain boundary. For domain-span processes, two possibilities for the identity of the locus are suggested by previous studies. Firstly, studies such as Barnwell (1971) indicate that all the subconstituents of the domain may be affected: for example, all the syllables in the word are shortened as the length of the word increases; in this case, the domain and the locus are co-extensive. Secondly, studies such as Lehiste (1972) may be taken to imply that the domain head may be affected—for example, the primary stressed syllable in the word is shortened as the word increases in length—but the other syllables are unaffected; similarly, as found by Lehiste (1972) and Port (1981), the vowel nucleus within the stressed syllable may show a greater effect of word length than the onset or coda.

If the greatest effect of a domain-span process is found to occur on the domain head, this could either be because it is the locus of the process, or because the locus is co-extensive with the domain but the most prominent subconstituent of the domain is longer and more compressible. In the latter case, unstressed syllables, being shorter, may manifest smaller domain-span compression effects than a stressed syllable. Likewise, the syllable nucleus will often be the longest part of the stressed syllable, and even when it is no longer than the onset or coda, it may be more compressible than onset or coda consonants such as stops.

$\Longrightarrow$ The locus hypotheses pertain to all parts of Experiment 2. The domain-edge locus hypothesis is supported: in particular, word-initial lengthening (Section 4.6) and utterance-final lengthening (Section 4.5) are shown to be confined to phonologically-defined loci. The domain-span locus hypothesis is partially supported: the only domain-span effect indicated by Experiment 2—word-rhyme-span compression as described in Section 4.6—is localised within the stressed syllable on the nucleus; however, unstressed syllables within the word-rhyme also appear to be affected. The experimental outcomes relating to the locus hypotheses are discussed in detail in Section 4.10.

Word-edge and word-span processes

Word-span hypothesis: primary stressed syllable duration is inversely related to word length in syllables.

Word-rhyme-span hypothesis: primary stressed syllable duration is inversely related to word-rhyme length in syllables.

Word-initial hypothesis: the duration of the primary stressed syllable onset is greater word-initially than word-medially in utterance-medial words.

Word-final hypothesis: the duration of the primary stressed syllable rhyme is greater word-finally than word-medially in utterance-medial words.

The asymmetry in the word-level effect in Experiment 1 could be an experimental artefact: the word-final test syllables (underlined) are also phrase-final (for example: *SAY "knee capsize", don't SHOUT "knee capsize"*) and phrase-final lengthening may mask a symmetrical underlying word-span compression process: polysyllabic shortening as defined in Chapter 2. Turk & Shattuck-Hufnagel (2000) suggest this process accounts for their results, in conjunction with word-initial lengthening and what they call "syllable ratio equalisation"[3].

Alternatively, the asymmetry in Experiment 1 may be caused by a combination of polysyllabic shortening and another domain-span compression process, at the word-rhyme-level: as defined in Chapter 3, the word-rhyme begins with a primary stressed syllable and continues to a word boundary. The results of Lehiste (1972) and Port (1981) suggest a word-rhyme-span process, but the effect of adding syllables before the primary stress is not examined, and so comparison of word-rhyme-span and word-span hypotheses is not possible. Turk & Shattuck-Hufnagel (2000) consider the existence of "an asymmetric polysyllabic shortening effect", functionally equivalent to an additive combination of word-span and word-rhyme-span effects, but reject it for reasons outlined in Chapter 3. This combination of domain-span processes could account, however, for the observations of Lindblom (1968) for Swedish and Nooteboom (1972) for Dutch.

A combination of word-edge processes could also explain the results of Experiment 1: smaller word-initial lengthening and larger word-final lengthening. Oller (1973), Cooper (1991), Fougeron & Keating (1997) and Turk & Shattuck-Hufnagel (2000) all find evidence for word-initial lengthening; Beckman & Edwards (1990) present evidence which could indicate word-final lengthening, but Turk & Shattuck-Hufnagel (2000) do not find strong evidence for it.

Experiment 1 could also be interpreted as a combination of domain-edge and domain-span processes: in particular, word-initial lengthening and word-rhyme-span compression. The word-level experimental hypotheses presented above are thus expressed separately for each process. The hypotheses are formulated with reference to primary stressed syllable duration; the effect of word length on unstressed syllable duration is also examined.

Word-initial accent hypothesis: word-initial lengthening of stressed syllable onsets only occurs, or is significantly greater, in accented context.

[3] It is argued in Chapter 3 that syllable ratio equalisation is superfluous to the interpretation of word-level results and thus no explicit hypothesis regarding it is formulated here. Where relevant, it is considered in relation to the results of Experiment 2.

Experiment 1 also suggests that word-initial lengthening is only significant in the presence of accent; similarly, Turk & Shattuck-Hufnagel (2000) find that word-initial lengthening is greatest and most statistically reliable when the syllable is accented.

Word-initial aspiration accent hypothesis: word-initial lengthening of aspiration in stressed-syllable-onset voiceless stops only occurs in unaccented syllables; in accented syllables, aspiration duration may be greater word-medially.

The duration of aspiration of syllable-onset voiceless stops may show a particular interaction between word position and lexical or phrasal stress. The results of Oller (1973) and Cooper (1991) suggest that in syllable-onset voiceless stops, closure duration is longer word-initially in all contexts, but aspiration duration is longer word-initially in unstressed syllables and word-medially in stressed syllables. If, as it appears, test words are accented in those studies, lexical stress and pitch accent are confounded. Turk & Shattuck-Hufnagel's (2000) results suggest that the presence of pitch accent may be the determining factor in this interaction: measuring the onsets of stressed syllables only, they find that aspiration duration of syllable-onset voiceless stops shows reliable word-initial lengthening only in unaccented context.

$\Longrightarrow$ The results for word-edge and word-span processes in Experiment 2 are described in Section 4.6. The word-span hypothesis is not supported, as an inverse relationship between primary stressed syllable duration and word length is only found for accented words. The word-rhyme-span hypothesis and the word-final hypothesis are both partially supported: there is an effect of word-rhyme length localised on the stressed syllable nucleus, which also affects unstressed syllables; this also could be interpreted as a word-final effect, as discussed in Section 4.6 and further in Chapter 5. The word-initial hypothesis is clearly supported, and the effect is comparable in accented and unaccented words, hence the word-initial accent hypothesis is not supported. The word-initial aspiration accent hypothesis is also not supported: for syllable-onset voiceless stops, closure and aspiration duration are similarly affected by word position, and there appears to be no interaction with accent.

Utterance-span processes

Utterance-span hypothesis: primary stressed syllable duration is inversely related to the number of syllables in an utterance.

Some studies claim that a domain-span process operates at the sentence/utterance level: Lehiste (1974), cited in Klatt (1976), finds that constituents of the sentence are shorter when the sentence is longer; a study by Rakerd *et al.* (1987) of monosyllabic content words supports this conclusion. In contrast, Gaitenby (1965) finds little effect of sentence length, beyond what could be explained as final lengthening. The syntactic sentence and the prosodic utterance are often co-extensive, particularly where, as in Experiment 2, single sentences are read in isolation.

⟹ The utterance-span hypothesis is not supported in Experiment 2, as described in Section 4.5.

Utterance-initial processes

Utterance-initial hypothesis: word-initial stressed syllable onsets are shorter utterance-initially than utterance-medially.

Hierarchical lengthening studies indicate greater domain-edge lengthening at higher-level boundaries. Both Fourakis & Monahan (1988) and Fougeron & Keating (1997), however, find utterance-initial *shortening* of unstressed syllable onsets. Experiment 2 tests this effect in primary stressed syllables.

⟹ The utterance-initial hypothesis is supported in Experiment 2 for certain stressed syllable onsets, as described in Section 4.5

Utterance-final processes

Utterance-final locus hypothesis: the locus of utterance-final lengthening is the word-rhyme.

Klatt (1976) and Cummins (1999) claim that utterance-final lengthening extends over several syllables, although Oller (1973) and Campbell & Isard (1991) indicate that the largest effect occurs on the utterance-final syllable rhyme. Cambier-Langeveld (2000) finds a similar locus for utterance-final lengthening to that suggested by Nakatani *et al.* (1981) and Turk (1999) for phrase-final lengthening: both the utterance-penultimate stressed syllable, if the final syllable is unstressed, and the absolute final syllable are lengthened. This result suggests that the locus could be the word-rhyme, possibly excluding the onset of the primary stressed syllable.

Utterance-final locus distribution hypothesis: the distribution of utterance-final lengthening is progressive within the locus.

Where both syllables in the word-rhyme show phrase-final or utterance-final lengthening, previous studies suggest that the effect appears to be progressive: that is, there is greater lengthening closer to the boundary. Cambier-Langeveld (2000) for Dutch and Berkovits (1993a, 1993b) for Hebrew suggest that lengthening is similarly progressive within the final syllable rhyme. Oller (1973) for English, however, indicates variations in the distribution of lengthening according to the segmental identity of subsyllabic constituents.

Utterance-final accent hypothesis: the magnitude of utterance-final lengthening is not affected by the presence of a pitch accent on the utterance-final word.

Some studies may confound final lengthening with the effects of pitch accent: Oller (1973), for example, measures utterance-final lengthening in words that appear likely to be accented. The locus and magnitude of final lengthening may be different when the utterance-final word is not accented. Price *et al.* (1991) suggest an interaction between accentual lengthening and phrase-final lengthening such that neither one is as great in the presence of the other as in its absence. Cambier-Langeveld (2000) finds, however, that the magnitude of lengthening on utterance-final syllables is very similar in accented and unaccented context: the proportion of final lengthening is greater in unaccented context, because the syllables are shorter, but the interaction is not significant. For Dutch, in contrast, she finds a large attenuation of final lengthening in accented context.

⟹ As described in Section 4.8, the utterance-final locus hypothesis is supported, although lengthening does not occur on all parts of the word-rhyme: the stressed syllable coda is lengthened in utterance-antepenultimate position and utterance-penultimate position, as are the following unstressed syllables within the word; where the stressed syllable is in utterance-final position, the nucleus and coda are lengthened. The utterance-final locus distribution hypothesis and the utterance-final accent hypothesis are supported, although there is a non-significant trend towards attenuation of final lengthening in accented words.

Pitch accent

Word-rhyme accent hypothesis: the locus of accentual lengthening is the word-rhyme.

Bimodal accent hypothesis: the locus of accentual lengthening is the primary stressed syllable and the word-final syllable.

Sluijter (1995), Turk & Sawusch (1997) and Turk & White (1999) find that pitch accent causes lengthening of all parts of the accented syllable and also affects other syllables in the same word. In disyllables, a syllable following the accented syllable (for example: /kæp/ in *kneecap*) is lengthened more than a syllable preceding an accented syllable (for example: /kæp/ in *capsize*). In left-headed trisyllables such as *property*, both unstressed syllables show some lengthening. Two interpretations of the locus of accentual lengthening are suggested, as presented above. In one case, the bimodal hypothesis, the locus is suggested to comprise two discrete subconstituents. In either case, it appears that syllables adjacent to the locus within the word also manifest small lengthening effects.

Word-span accent hypothesis: the magnitude of accentual lengthening on the primary stressed syllable is inversely related to word length.

In Experiment 1, accentual lengthening of the primary stressed syllable is greater, in absolute and proportional terms, in monosyllables than in disyllables. Experiment 2 examines this apparent interaction between pitch accent and word length.

$\Longrightarrow$ As described in Section 4.9, the word-rhyme accent hypothesis and the bimodal accent hypothesis are not supported in Experiment 2: although there is accentual lengthening within the word rhyme, there is also lengthening in the unstressed syllable preceding the word-rhyme, suggesting that the locus of accentual lengthening is the word; furthermore, lengthening appears greatest at word edges. The word-span accent hypothesis is supported: accentual lengthening on the stressed syllable is greatest in monosyllables and progressively less in disyllables and trisyllables.

4.1.3 Experimental design considerations

The design of Experiment 2, which tests the hypotheses outlined above, is described in Section 4.2. Some of the issues which inform the design and interpretation of the experiment are discussed now, with particular reference to two previous studies of domain-edge processes in hierarchical constituents.

When measuring the magnitude of a domain-edge durational effect in a particular constituent, it is necessary to reliably identify the constituent itself. Syntactic constituents have well-established discovery procedures, but as illustrated in Chapter 2, it seems clear that syntactic constituents do not consistently comprise the domains of durational processes. In many cases, there is a process of re-analysis of syntactic constituents into the prosodic constituents which appear to be the domains of durational processes. Important aspects of this mapping remain uncertain, however:

- The number of prosodic constituents required to account for the segmental and suprasegmental processes of connected speech.

- Whether prosodic hierarchies comprise a single structure or partially-independent structures, such as prominence and constituent hierarchies.

- The balance between syntax and other factors such as speech rate, constituent length and pragmatic interpretation in the construction of prosodic constituents.

At the highest and lowest levels of structure, prosodic constituents are quite well motivated. Each lexical word appears to comprise a prosodic word, although, as discussed in Chapter 2, the prosodic status of grammatical words and the existence of nested prosodic words remain matters of debate. Intonational phrases are readily identifiable in speech through characteristic events—one or more pitch accents, declination of fundamental frequency, and a boundary tone—although the mapping of intonational phrases onto syntax is variable, with a number of possibilities available for most sentences. Utterances are fairly well-defined in most contexts, generally identifiable with a single sentence, encompassing one or more intonational phrases, and very likely to be pause-delimited.

Between the levels of the word and the intonational phrase, however, the number of prosodic constituents and their segmental and suprasegmental consequences are less clear. The evidence for certain constituents, such as the clitic group and various types of phonological phrase, is discussed in Chapter 2. In the current context, the important questions are:

- How are the boundaries of these putative constituents identified?

- What are the durational processes associated with these boundaries?

Experiments on hierarchical lengthening processes generally adopt one of two approaches to identifying prosodic boundaries. In corpus-based studies, such as Ladd & Campbell (1991) and Wightman *et al.* (1992), a set of recorded sentences is examined by a phonetician, who locates boundaries of different types according to acoustic criteria. In experimental studies, such as Ferreira (1991) and Fougeron & Keating (1997), sentences are designed to elicit specific prosodic boundaries at particular points in the recorded utterance. One study of each type is examined here, to illustrate the difficulties inherent in such experiments, and thereby inform the design and interpretation of Experiment 2.

Wightman, Shattuck-Hufnagel, Ostendorf and Price (1992)

Four professional radio announcers read 35 pairs of phonetically-similar but syntactically-ambiguous sentences, which are designed to be fully voiced. Speech segment labels are aligned automatically and the resulting segment durations are normalised for phoneme identity and speech rate.

Three researchers assess boundary strength between each pair of words, using a system of seven "break indices", the precursor to that described in Silverman *et al.* (1992) as part of the ToBI ("Tone and Break Indices") system for the transcription of prosody. The number of indices reflects the maximum number of prosodic constituents at and above the word level proposed in the literature, except for proposals such as Ladd (1986) allowing recursion of constituents[4]. An index of 0 indicates no perceived prosodic break and corresponds to a clitic group boundary; an index of 4 corresponds to an intonational phrase boundary, the lowest level with a boundary tone; an index of 6, marking the end of a sentence, effectively equates to an utterance boundary.

Wightman *et al.* test the hypothesis that boundaries induce lengthening in the stress-delimited foot that contains them (see for example: Lehiste 1977). They measure the duration of four stretches of speech before each boundary: the final stressed syllable vowel and the segments between the final stressed vowel and the final vowel (where the preboundary syllable is not stressed), the final syllable vowel and the final syllable coda. They also measure the duration from the boundary to the start of the next stressed vowel. They find lengthening correlated with break index is confined to the vowel and coda of the preboundary syllable. As the preboundary vowel shows the highest correlation, they compare vowel duration, normalised for phoneme identity and speech rate, at boundaries marked with different break indices and find significant differences between vowels preceding boundaries marked with break indices 1, 2, 3 and 4. The durational differences between indices 0 and 1, and between indices 4, 5 and 6 are not significant, the latter meaning that utterance-final lengthening is no greater than intonational-phrase-final lengthening, at least on the final vowel. They conclude that preboundary lengthening can distinguish four levels of prosodic constituents: prosodic word; "a grouping of words within a larger unit", possibly the accentual phrase (Beckman & Pierrehumbert 1986); intermediate intonational phrase; intonational phrase.

[4]The seven break indices used by Wightman *et al.* (1992) do not correspond exactly to the seven levels, discussed in Chapter 2, described by Nespor & Vogel (1986): two of Nespor & Vogel's constituents—the syllable and the within-word foot—are below the word level, whereas the break indices used by Wightman *et al.* refer to perceived degrees of juncture *between* (but not within) orthographic words. Clearly, the seven break indices thus imply more prosodic constituents above the word level than the proposal of Nespor and Vogel, although it may be noted that the number of break indices is reduced to five in the subsequent development of the ToBI system.

There are several reasons to doubt the ubiquitous occurrence in normal speech of all these constituents and their distinctive degrees of preboundary lengthening. The first is the method of allocating break indices to boundaries: labels are assigned by ear, with no visual display, and another paper on the same corpus states that: "care was taken in the discussion [of discrepancies between labellers] to point out possible biases from syntax in order to avoid such influences insofar as possible" (Price *et al.* 1991:2963). This implies that the assignment of labels for constituent boundaries below the level of the intonational phrase—for which independent cues to juncture are available—is based largely on perceptions of degrees of lengthening, as segmental effects will only be manifest sporadically.

It seems unlikely that all biases from syntax could be eliminated: for example, would an intermediate intonational phrase boundary be placed within a prepositional phrase, contravening the theoretical accounts of its construction, if the degree of acoustic juncture demanded it? Moreover, it seems that some account *should* be taken of syntax if the experiment is not merely to demonstrate labellers' skill in perceiving degrees of durational variation. The later development of the ToBI system reflects this: as Wightman (2002) observes that: "the subjective opinion of the labeller was de-emphasised in favour of the identification of a specified set of prosodic phrasal constituents." In the prototype system exploited by Wightman *et al.* (1992), however, the determining characteristic of break indices below the intonational phrase appears to be perceived final lengthening.

Another problem is that the durational effects observed are not necessarily associated with boundaries, given the absence of control for the effects of pitch accent. Some of the boundaries associated with break indices will be preceded by pitch-accented words, which will be lengthened as a result of accent. Indeed, such lengthening could induce in the listener the perception of a lower-level boundary: a boundary thus identified would show relative lengthening, but it would be wrong to identify this as final lengthening. Accentual lengthening and final lengthening might be distinguished by their locus: final lengthening mainly affects the final syllable rhyme; accentual lengthening also affects the syllable onset, and preceding syllables within the word. Wightman *et al.* do find lengthening correlated with break index only on the final vowel and coda, but the manner of their segmental groupings may mask lengthening elsewhere. For instance, stressed and unstressed preboundary syllables are grouped together, and the onset duration of preboundary stressed syllables is not reported. Where this syllable is accented, it should show lengthening on the onset as well[5]. In addition, final

[5]Wightman *et al.*, furthermore, find no hierarchical lengthening effect in the post-boundary stretch, contradicting Fougeron & Keating (1997), discussed below, who find hierarchical syllable-onset lengthening. This discrepancy may arise because Wightman *et al.* do not report onset duration separately, but group it with all the post-boundary material up to the first stressed vowel. Similar masking of a localised

lengthening appears to be progressive, but here the greatest effect is evident on the final syllable nucleus; inspection of the confidence intervals for the normalised means suggests that differences in coda duration between different break indices are not significant. This pattern of durational variation may mean that final lengthening and accentual lengthening are somewhat confounded in the experimental materials.

Apart from the problem of interpretation of Wightman *et al.*'s study, there is also the question of its applicability to normal speech. Firstly, significant differences in preboundary lengthening are only obtained with data that is twice manipulated to reduce variance, using normalisation procedures for phoneme identity and speech rate; in a related study using the same data-set, Price *et al.* (1991) report mean durations normalised only for phone identity, and find the only significant difference in preboundary lengthening is between break index 4 (the intonational phrase) and the lower level indices. Secondly, as Price *et al.* (1991) indicate, the corpus is produced by professional announcers using their "standard radio style of speaking". They acknowledge that the manner of speech is not what would be observed in spontaneous conversation: "we found the FM radio style to have more clearly and consistently marked prosodic cues than a nonprofessional speaking style (Price *et al.* 1988), while sounding acceptably natural. Our hope was that this style would be easier to label prosodically, and therefore the contributions of specific phonological cues would be easier to identify [...] However, care must be taken in generalising our results to spontaneous speech, where prosodic cues may be less clearly marked" (Price *et al.* 1991:2959). One may go further, and suggest that this style of speech is qualitatively distinct from normal speech, and even from the reading of printed text in a laboratory by untrained speakers. Even if it is safe to conclude that trained speakers can make a certain number of durational distinctions that become significant in rate-normalised data, it does not follow that these distinctions, and the constituents that they imply, are customarily used by normal speakers or perceivable by normal listeners.

Fougeron and Keating (1997)

Three phoneticians, including the second author, read reiterant versions of algebraic expressions, for example:

- 89 + 89 + 89 + 89 = a lot
- (89 + 89) × (89 + 89) = a lot
- 89 × (89 + 89 + 89) = a lot

are all spoken as the sequence:

effect may occur in the preboundary vowel-to-vowel stretch.

- nonono no nonono no nonono no nonono = a lot

with brackets indicating the intended phrasing.

Five levels of prosodic constituency are identified: utterance; intonational phrase, generally ending at the close of a bracketed expression; phonological phrase, generally encompassing a "plus" or "times" and a subsequent numeral (a phonological phrase boundary is identified acoustically as an "intermediate break"); word, each numeral, such as "89", being regarded as a single word; syllable.

Durational measurements are reported for the recordings based on the "89" paradigm, with word-final lexical stress; other numerals are used to vary the position of lexical stress for a larger articulatory study. Mean durations of syllable onset /n/ are reported according to the highest constituent boundary they follow. Two of the three speakers distinguish four constituents according to onset duration: greatest intonational-phrase-initially: less phonological-phrase-initially; less again word-initially; shortest word-medially. One speaker shows a marked utterance-initial shortening effect, with duration comparable to word-medial position; the other has utterance-initial /n/ comparable to IP-initial /n/. The third speaker also shows a large utterance-initial shortening effect; intonational phrases and phonological phrases are not distinguished, but show initial lengthening relative to word-initial position, which shows lengthening relative to word-medial position.

Results for vowel duration preceding constituent boundaries are less clear-cut: "final vowels are generally poor indices of the hierarchical level of prosodic domains both in their spatial (linguopalatal) and temporal characteristics, though all three speakers have longer vowels [intonational and phonological] phrase-finally than [word-] finally" (Fougeron & Keating 1997:3736). The use of reiterant open syllables may contribute to this, because syllable codas may show greater final lengthening than nuclei. The relative lack of differentiation of boundary strength by preboundary vowels contrasts with Wightman *et al.*, where vowels actually show stronger hierarchical effects than coda consonants: the failure to replicate this result supports the view that Wightman *et al.*'s observations might in part be due to the presence of accent rather than phrase boundaries.

Another comparison demonstrates the limited applicability of these results: Wightman *et al.* look for seven levels of phrasing and find final lengthening evidence for most, including two levels between word and intonational phrase. Fougeron & Keating use materials designed to induce three levels of phrasing above the word-level, and find evidence for these levels from onset consonant durations. They do not, however, find reliable evidence for a single level of hierarchical final lengthening between the word and the intonational phrase, whereas Wightman *et al.* find two such levels. Furthermore, Wightman *et al.* do not find evidence for the hierarchical lengthening of syllable onsets observed by Fougeron & Keating. Clearly, these results are incompatible.

One reason for this difference may be the rather specialised nature of the experimental tasks. Wightman *et al.* use professional radio announcers and admit that their phrasing of the experimental materials may be rather different from that of ordinary speakers in everyday context. Fougeron & Keating use trained phoneticians, including one of the authors and two others who are "naive about the present study". Firstly, if it relevant to mention ignorance of the purpose of the recordings, it may be asked why one of the authors is nevertheless used as a subject: Vogel (1992), amongst others, have questioned the validity of data produced by experimenters themselves in phonological studies. Secondly, the tasks of disambiguating contrived bracketed expressions, like the reading of texts by professional radio announcers, clearly requires a special type of linguistic performance. Apparently, the strategy adopted in the disambiguation task involves lengthening of syllable onsets at phrase boundaries focused at phrase onsets; however, finding that trained subjects are capable of performing a particular task does not demonstrate that normal speakers habitually exhibit such behaviour.

Factors in the design of Experiment 2

This above discussion illustrates difficulties inherent in experimental examinations of durational variation in speech, where not all potential influences can be reliably controlled, and some factors, such as the number of levels of prosodic constituents, are neither theoretically agreed nor empirically established. A possible criticism of Experiment 1 is that, in restricting potential durational influences, the experimental materials are too tightly constrained. For example, the linguistic status of the test materials (underlined) is uncertain in phrases such as:

- SAY " <u>thank fulfil</u>", don't SHOUT "thank fulfil".

This could be interpreted as the phrasal equivalent of citation form; certainly, the prosodic phrasing is not subject to the degree of variation found in normal speech.

The sentences in Experiment 2 are designed to allow subjects a more normal range of prosodic choices, whilst at the same controlling the environment of the measured segments in certain important ways. Firstly, as in Experiment 1, the intended placement of pitch accent is specified in the sentence by writing words to be emphasised in block capitals. As there are a range of possible locations for accents within utterances, this should not require subjects to deviate further from a normal speaking style than is inherent in a sentence-reading task.

The question of constituent boundary location is more problematic. Two domain-edge conditions are examined: word-edge and utterance-edge. As the utterance is the highest constituent, certainly within single, short sentences, controlling the placement of utterance boundaries is straightforward. At the word-level, however, the task is more complicated given the number of putative higher-level constituents whose boundaries could coincide with the boundaries of experimentally-significant words. Furthermore, because the relationship between syntax and prosody is not deterministic, sentences cannot be written which elicit a particular phrasing consistently.

The approach taken in Experiment 2 is to eliminate the well-established durational effects of intonational phrase boundaries, whilst allowing variation in the placement of boundaries of such lower-level constituents as may exist, and which may have durational consequences. As the comparison of Wightman *et al.* (1992) and Fougeron & Keating (1997) shows, the durational consequences of constituents between word-level and intonational-phrase-level are far from clearly established. As discussed in Chapter 2, at least one constituent, the phonological phrase, may be required to account for segmental and suprasegmental effects in connected speech: whether it manifests distinctive domain-initial and/or domain-final lengthening remains uncertain[6].

Intonational phrase boundaries are controlled in Experiment 2 in a number of ways. Firstly, the experimental sentences are relatively short, and free of any syntactic structures associated with obligatory intonational phrase construction. Secondly, as utterance-internal pauses are most strongly associated with intonational phrase boundaries, subjects are instructed to read each experimental sentence without pausing. Thirdly, each recorded utterance is listened to, and those with perceptible intonational phrase boundaries adjacent to the experimentally-significant words are excluded from analysis.

The lack of specific measures to control the placement of phrase boundaries below the level of the intonational phrase has consequences for the interpretation of the experimental findings with regard to word-edge effects. The edges of some experimentally-significant words may coincide with boundaries of, for example, phonological phrases; therefore, if word-initial or word-final lengthening effects are found, these may be attributed, at least in part, to phrasal boundaries rather than word boundaries. This proviso is considered further in the discussion of the experimental results. The possible occurrence of phonological phrase boundaries is less relevant to hypotheses about word-span, utterance-edge and utterance-span effects, and about pitch accent and its interactions.

[6] A study by Cambier-Langeveld (2000) of Dutch, discussed in Chapter 2, found no difference in speech segment duration word-finally and phonological-phrase-finally.

4.2 Experimental design

Experiment 2 is designed to examine the experimental hypotheses outlined in Section 4.1.2. The durations of the onset, nucleus and coda of a primary stressed syllable are measured when spoken in words and utterances of varying lengths, where the syllable's position with respect to word and utterance boundaries is varied, and where the syllable is accented and unaccented. The durations of unstressed syllables in polysyllabic words are also measured.

4.2.1 Experimental materials

Stressed syllable durations in Experiment 1 are measured in monosyllabic words and in left-headed and right-headed disyllabic words. Word-level contexts are extended in Experiment 2 to include left-headed and right-headed trisyllabic words, because the word-edge, word-span and word-rhyme-span hypotheses make different predictions about stressed syllable duration in these trisyllables compared with disyllables.

For both left-headed pairs like *captain* and *captaincy* and right-headed pairs like *commend* and *recommend*, the word-span hypothesis predicts shorter primary stressed syllable duration in trisyllables than disyllables. The word-rhyme-span hypothesis predicts such shortening in left-headed pairs, but not in right-headed pairs.

The word-edge hypotheses predict no difference in either left-headed or right-headed pairs, because the alignment of the stressed syllable with the word boundary is unchanged: in both *captain* and *captaincy*, the measured syllable /kæp/ is word-initial and not word-final; in both *commend* and *recommend*, the measured syllable /mɛnd/ is word-final and not word-initial.

When the primary stress is word-medial in the trisyllable, comparisons with either left-headed or right-headed disyllables do not allow distinction of the word-edge and word-span hypotheses: both the word-initial and word-span hypotheses predict shortening of /sɛns/ in *consensus* compared with *census*; likewise, the word-final, word-span and word-rhyme-span hypotheses all predict shortening of /tɛnt/ in *contentment* compared with *content*. Such trisyllables do not usefully distinguish the hypotheses and so are not used here.

It is possible that edge and span processes may both operate. For example, a large amount of shortening of /mɛnd/ between *mend* and *commend* and a smaller shortening effect between *commend* and *recommend* could indicate a word-span effect together with a word-initial effect. In order to test this, it is necessary to examine subsyllabic durations: the locus of word-initial lengthening, which affects only the *mend* vs *commend* comparison, is likely to be the syllable onset, whereas the word-span shortening effect is likely to affect all parts of the syllable and possibly the nucleus more than the onset and coda.

Keywords

The dependent variables in this experiment are the durations of the onset, nucleus and coda of syllables carrying primary lexical stress, referred to as "test syllables". The word containing the test syllable is referred to as the "keyword".

The experiment uses sixteen different test syllables, each spoken in three different lexical contexts: monosyllable, disyllable and trisyllable. The three keywords for each test syllable are referred to as "keyword triads". There are eight keyword triads in which the test syllable is always word-initial, the left-headed keywords, and eight in which the test syllable is always word-final, the right-headed keywords. The full set of keywords[7] is given in Table 4.1.

Left-headed triads				**Right-headed triads**			
Test syllable	σ	σ*σ*	σ*σσ*	Test syllable	σ	*σ*σ	*σσ*σ
/kæp/	**cap**	captain	captaincy	/mɛnd/	**mend**	commend	recommend
/sɛns/	**sense**	censor	censorship	/pəʊz/	pose	**dispose**	indispose
/dɒg/	**dog**	dogma	dogmatist	/pɔt/	**port**	report	misreport
/fɪʃ/	**fish**	fissure	fisherman	/dʒus/	**juice**	produce	reproduce
/meɪs/	**mace**	mason	masonry	/pəʊz/	pose	**suppose**	presuppose
/pɑt/	**part**	partner	partnership	/pəʊz/	pose	**compose**	decompose
/spɛk/	**speck**	spectre	spectacle	/sɛnd/	**send**	descend	condescend
/tɛn/	**ten**	tendon	tendency	/meɪn/	**main**	humane	inhumane

Table 4.1: Keywords used in Experiment 2. Test syllables are shown schematically as σ, other syllables as *σ*. The keywords used to refer to each triad are shown in bold.

The syllables immediately adjacent to the test syllable are unstressed, to control stress-adjacent lengthening, and are identical in the disyllables and trisyllables in each keyword triads, to minimise any effect of local phonetic variation on test syllable duration. Stress and phonetic environment are also controlled in the carrier sentence.

Carrier sentences

In Experiment 1, experimental phrases are placed in what could be described as a metalinguistic context within the complete sentence, for example:

- Say "BAKE enforce", don't say "TANK enforce".

[7] Transcriptions are given for Standard Southern British English. Unlike American English, *juice* and the final syllable of *produce* and *reproduce* are homophonous, being pronounced /dʒus/.

In Experiment 2, keywords are placed in sentences designed to be read as meaningful utterances, with a full normal prosodic realisation. At the same time, the length of the utterances and the immediate environment of the test syllables is controlled[8].

The full design of the experimental materials is shown schematically in Table 4.2 for right-headed keywords. The design of the materials for left-headed keywords is the mirror-image of the syllable configurations shown.

	Number of added syllables: **. . . in utterance**	**. . . in word** 0	1	2
Utterance-medial	0	$\|\ldots\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	$\|\ldots\sigma\sharp\sigma\boldsymbol{\sigma}\sharp\ldots\|$	$\|\ldots\sharp\sigma\sigma\boldsymbol{\sigma}\sharp\ldots\|$
	1	$\|\ldots\sigma\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	$\|\ldots\sigma\sigma\sharp\sigma\boldsymbol{\sigma}\sharp\ldots\|$	
	2	$\|\ldots\sigma\sigma\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$		$\|\ldots\sigma\sigma\sharp\sigma\sigma\boldsymbol{\sigma}\sharp\ldots\|$
Utterance-edge	0	$\|\boldsymbol{\sigma}\sharp\ldots\|$		
	1	$\|\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	$\|\sigma\boldsymbol{\sigma}\sharp\ldots\|$	
	2	$\|\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$		$\|\sigma\sigma\boldsymbol{\sigma}\sharp\ldots\|$

Table 4.2: Schematic representation of the experimental keywords and carrier sentences. The design shown is for right-headed words; the design for left-headed words is the mirror image of this representation. KEY: $\boldsymbol{\sigma}$- target syllable; σ- other syllable; $\|$ - utterance boundary; $\sharp$ - word boundary; ... symbolises the initial and final parts of the utterance, of fixed length within each triad, and at least two syllables long.

The design shown in Table 4.2 systematically varies word length and utterance length. Where the test syllable is utterance-medial, these factors are varied both independently and concurrently. Where the test syllable is close to the utterance edge, utterance length covaries with word length and varies whilst word length is fixed. The resultant five series of test syllable environments are outlined in turn now.

Keyword Series A: utterance-medial; fixed utterance-length; variable word-length

Keyword Series A primarily tests the word-edge, word-span and word-rhyme-span hypotheses, and is illustrated schematically for right-headed keywords in the top line of the utterance-medial section of Table 4.2. Utterance-level factors are kept constant, so that word-level effects of length and position can be examined. To eliminate utterance-length effects, the number of syllables within the keyword is counterbalanced in the rest of the sentence, maintaining a constant number of syllables in the

[8] The requirement for phonetic balance in the environment of the test syllable, together with sentence length constraints, mean that some of the sentences are semantically rather strange. Subjects appeared, however, to have no difficulty in reading these sentences with the same fluency as with the others.

complete sentence for each keyword of a triad. To minimise utterance-edge effects, the distance of test syllables from utterance edges is kept constant within each keyword triad and keywords are separated from the ends of the carrier sentence by at least two syllables.

In left-headed triads the monosyllabic keyword has one more syllable to its right than the disyllabic keyword and two more than the trisyllabic keyword. For the right-headed triads, the counterbalancing syllables are to the left of the keyword. Examples of counterbalanced sentences are shown in Table 4.3: in that table and subsequently, keywords are in bold, and words emphasised in the unaccented condition—as outlined below—are in capitals. The complete set of sentences is shown in Table C.1 in Appendix C.

Right-headed keywords	
$\|\ldots\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jessica **mend** it AGAIN.
$\|\ldots\sigma\sharp\sigma\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jessie **commend** it AGAIN.
$\|\ldots\sharp\sigma\sigma\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jess **recommend** it AGAIN.

Left-headed keywords	
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed AGAIN.
$\|\ldots\sharp\boldsymbol{\sigma}\sigma\sharp\sigma\ldots\|$	I SAW the **mason** reclaimed it ALL.
$\|\ldots\sharp\boldsymbol{\sigma}\sigma\sigma\sharp\ldots\|$	I SAW the **masonry** cleaned AGAIN.

Table 4.3: Experimental sentences for Keyword Series A.

The syllable immediately following the test syllable in the left-headed triads and immediately preceding the test syllable in the right-headed triads is unstressed, whether it is within the same word or within an adjacent content word, as in Table C.1, or comprises an adjacent function word, for example:

Series A.1 Kate GAVE the **sense** of the SCRIPT away. (cf. Kate GAVE the **censor** the SCRIPT again.)

Series A.1 John THREW the **cap** to the BED again. (cf. John THREW the **captain** the BADGE again.)

This is to control the effect of stress adjacency on the test syllable.

The left-headed keywords *dog* and *part* are followed in Series A monosyllables by closed-class words containing syllables (underlined) that are reducible but could be realised with full vowels:

Series A.1 Tim KNEW the **dog** may decline AGAIN. (cf. Tim KNEW the **dogma** declined AGAIN.)

Series A.1 Jim LIKES his **part** <u>no</u> more than MOST. (cf. Jim LIKES his **partner** more than MOST.)

As described in Chapter 2, a stressed syllable may be lengthened when followed immediately by another stressed syllable, relative to when followed by an unstressed syllable. It is considered unlikely that merely full-vowel unstressed syllables will cause this effect, but because of this possibility, the keywords *dog* and *part* are excluded from those analyses which make a comparison which could be biased by stress-adjacent lengthening if it did occur: for example, shortening of the test syllable between *dog may* and *dogma* could be regarded either as a word-level effect or a consequence of adjacent full vowel syllables in the former but not the latter.

The immediate phonetic environment of the test syllable is kept constant, as far as possible. As shown in Table 4.3, /mɛnd/ is always preceded by /kə/ and followed by /ɪt, and /meɪs/ is always preceded by /ðə/ and followed by /ən/. In the examples above for keywords *sense* and *cap*, the reduction of function words in connected speech means that the test-word-adjacent syllables are similar to the within-word equivalents: thus, *sense of* vs *censor*; *cap to* vs *captain*. In these sentences, in contrast with those shown in Table 4.3, the requirement to maintain a consistent phonetic environment and the lack of segmentally-appropriate counterbalancing words means that the syntactic structure of the sentences differs within the keyword triads. All of the left-headed triads except *mace* show some variation in syntactic structure in Series A. For the right-headed triads, three keywords—*port*, *juice* and *pose*—have structural variation between the three sentences; for the latter two, this variation is minimal.

There are two pitch accent conditions: in the accented condition, subjects emphasise the keyword (shown, in the examples given here, in bold); in the unaccented condition, they emphasise two other words in the sentence (shown, in the examples given here, in block capitals), one before and one after the test syllable. Words to be emphasised in the unaccented condition never immediately precede the test syllable, as the word following an accented word may be lengthened somewhat. The required placement of emphasis is indicated to subjects in all cases by the use of block capitals.

The two accent conditions apply in all keyword series, likewise the consistent environment of the test syllable in terms of lexical stress and segmental identity, except where otherwise stated.

Keyword Series B: utterance-medial; variable utterance-length; fixed word-length

Keyword Series B primarily tests the utterance-span hypothesis, and is illustrated schematically for right-headed keywords in the left-hand column of the utterance-medial section of Table 4.2. The length of the keyword is fixed whilst the length of the whole utterance varies. Example sentences are shown in Table 4.4 and the full set of sentences in Series B is shown in Table C.2 in Appendix C.

Right-headed keywords	
$\|\ldots\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jessica **mend** it AGAIN.
$\|\ldots\sigma\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JOHNNY saw Jessica **mend** it AGAIN.
$\|\ldots\sigma\sigma\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JONATHAN saw Jessica **mend** it AGAIN.
Left-headed keywords	
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed AGAIN.
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed once AGAIN.
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed by them AGAIN.

Table 4.4: Example experimental sentences for Keyword Series B.

As indicated in Table 4.2, the shortest sentence in Series B is also the monosyllabic keyword sentence in Series A. In order to keep the phonetic environment of the test syllable constant, additional syllables in the other two sentences are not placed immediately adjacent to the keyword, but elsewhere in the utterance. Due to an oversight in the preparation of the materials for the keyword *part*, the additional syllables are adjacent to the keyword; because of the resulting unbalanced phonetic environment, *part* is not analysed with the other keywords in Series B.

Keyword Series C: utterance-medial; variable utterance-length; variable word-length

Keyword Series C primarily tests the utterance-edge hypotheses by comparison with Series E, as outlined below; it also tests the word-edge, word-span and word-rhyme-span hypotheses examined in Series A. Series C is illustrated schematically for right-headed keywords in the diagonal line of the utterance-medial section of Table 4.2. As in Series A, the number of syllables in the keyword is varied; in contrast with Series A, the length of the carrier sentence is not counterbalanced with the length of the keyword. As the length of the keyword increases from monosyllable to disyllable to trisyllable, the length of the utterance also increases. Example sentences are shown in Table 4.5 and the full set of sentences is shown in Table C.3 in Appendix C.

As indicated in Table 4.2, the sentence containing the monosyllabic keyword is common also to Series A and Series B[9]. In Series A the requirement to counterbalance word length and sentence length whilst maintaining the phonetic environment of the test syllable often means that the syntactic structure of the sentences varies within

[9] As for Series A, the left-headed keywords *dog* and *part* are excluded from analyses where the potential stress-adjacent lengthening in the monosyllables could confound examination of the experimental hypotheses.

Right-headed keywords	
$\|\ldots\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jessica **mend** it AGAIN.
$\|\ldots\sigma\sigma\sharp\sigma\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Alison **commend** it AGAIN.
$\|\ldots\sigma\sigma\sharp\sigma\sigma\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Alison **recommend** it AGAIN.
Left-headed keywords	
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed AGAIN.
$\|\ldots\sharp\boldsymbol{\sigma}\sigma\sharp\sigma\sigma\ldots\|$	I SAW the **mason** disinclined it ALL.
$\|\ldots\sharp\boldsymbol{\sigma}\sigma\sigma\sharp\sigma\sigma\ldots\|$	I SAW the **masonry** disinterred AGAIN.

Table 4.5: Example experimental sentences for Keyword Series C.

the keyword triad. In Series C, however, this variability is eliminated. The carrier sentences for the monosyllabic, disyllabic and trisyllabic keywords are structurally equivalent except for very minor differences for the keywords *mace* and *port*.

The carrier sentences are often identical within a keyword triad; elsewhere, small non-structural changes occur where a repeated sequence of syllables is potentially awkward to pronounce. For example, for the keyword *mend*, the repeated /kəkə/ sequence in *Jessica commend* was avoided by substituting *Alison* for *Jessica*.

Keyword Series D: utterance-edge; variable utterance-length; fixed word-length

Keyword Series D tests the utterance-edge and utterance-span hypotheses, and is illustrated schematically for right-headed keywords in the left-hand column of the utterance-edge section of Table 4.2. This is the utterance-edge equivalent of Series B, both having fixed word-length and variable utterance-length. The keyword is a monosyllable in all three sentences and adjacent to the sentence edge in the shortest. As the number of syllables in the sentence increases, the distance between the test syllable and the utterance edge also increases. Example sentences are shown in Table 4.6 and the full set of sentences is shown in Table C.4 in Appendix C. There is only one emphasis in the carrier sentence in the unaccented condition; this is similarly true for Series E, the other utterance-edge series.

In contrast with the other keyword series, it is not possible in Series D always to keep the immediate stress and phonetic environment of the test syllable constant. In particular, syntactic restrictions on the occurrence of function words and phonological restrictions on their realisation—phrase-final function words are generally pronounced in their strong form—mean that sentences cannot be constructed in which the keyword is separated from the boundary by a single unstressed syllable. Thus *mend* is preceded by the adverb *now* in this condition and *mace* is followed by the preposition

Right-headed keywords	
$\|\sigma\sharp\ldots\|$	**Mend** it AGAIN for me please.
$\|\sigma\sharp\sigma\sharp\ldots\|$	Now **mend** it AGAIN for me please.
$\|\sigma\sigma\sharp\sigma\sharp\ldots\|$	Will you **mend** it AGAIN for me please.
Left-headed keywords	
$\|\ldots\sharp\sigma\|$	Albert THREW the **mace**.
$\|\ldots\sharp\sigma\sharp\sigma\|$	Albert THREW the **mace** up.
$\|\ldots\sharp\sigma\sharp\sigma\sigma\|$	Albert THREW the **mace** again.

Table 4.6: Example sentences for Keyword Series D.

up, a function word which receives stress by virtue of its phrase-final position. These restrictions on sentence construction place limitations on the interpretation of results in this part of the keyword series; but when the keyword is separated from the boundary by two syllables there is scope for placing it adjacent to an unstressed syllable. All the left-headed keywords in this condition are either followed by *again* or by *today*, which have initial unstressed syllables. The right-headed keywords are all preceded by either *Will you*, *Can you* or *All the*.

Given the restrictions on the lexical items available in sentence-initial and sentence-final position, it is furthermore very difficult to keep the phonetic environment of the test syllable the same within the series, and also by comparison with the disyllables and trisyllables in other keyword series. It is possible in some cases: *speck* is followed by /tə/ in both *spectacle* and in *speck today*; in some other left-headed triads the test syllable is followed by /ə/ both in the polysyllabic words and in the Series D condition. In general, however, the variations in the phonetic environment of the test syllable provide further reason for caution, although any durational differences due to phonetic environment are unlikely to be systematic with respect to the experimental hypotheses.

Keyword Series E: utterance-edge; variable utterance-length; variable word-length

Keyword Series E primarily tests the utterance-edge hypotheses, by comparison with the utterance-medial Series C: in both series, word-length and utterance-length co-vary and so differences in durational patterns between the series reflect the presence of the utterance boundary in Series E, where the distance between the test syllable and the utterance-edge increases as the length of the keyword increases from monosyllable to disyllable to trisyllable. The series is illustrated schematically for right-headed keywords in the diagonal line of the utterance-edge section of Table 4.2. Example sentences are shown in Table 4.7 and the full set of sentences is shown in Table C.5 in Appendix C.

Right-headed keywords	
$\|\sigma\sharp\ldots\|$	**Mend** it AGAIN for me please.
$\|\sigma\sigma\sharp\ldots\|$	**Commend** it AGAIN for me please.
$\|\sigma\sigma\sigma\sharp\ldots\|$	**Recommend** it AGAIN for me please.
Left-headed keywords	
$\|\ldots\sharp\sigma\|$	Albert THREW the **mace**.
$\|\ldots\sharp\sigma\sigma\|$	Albert THREW the **mason**.
$\|\ldots\sharp\sigma\sigma\sigma\|$	Albert THREW the **masonry**.

Table 4.7: Example materials for Keyword Series E.

The sentence containing the monosyllabic keyword is also the shortest sentence in Series D; additional sentences are constructed by replacing the monosyllabic keyword with the disyllabic or trisyllabic versions. Direct comparisons may be made with Series A and C as the stress and phonetic environment of the test syllable is constant between these series[10].

4.3 Experimental procedure

There are five keyword series for each test syllable, each containing three sentences. As shown in Table 4.2, the three utterance-medial series have one sentence in common, as do the two utterance-edge series, making a total of twelve sentences for each keyword, each to be read in two accent conditions. There are eight right-headed keywords and eight left-headed keywords, making 384 experimental sentences in total, listed according to keyword series in Tables C.1 to C.5 in Appendix C.

4.3.1 Recording

The experimental sentences were divided into three blocks for the purposes of the recording: Block A contains 112 sentences; Blocks B and C contain 136 sentences each. The composition of each block is shown in Table C.6 in Appendix C.

For each of the six subjects, two full sets of sentences from each block were prepared in separate random orders. For Block A, each randomised set was split into two

[10]For the left-headed keywords, the preceding phonetic environment is the same in Series A, C and E except for the keyword *speck*; for the right-headed keywords, it is the same except for *port*, *send* and *main*. The durational effect of these differences is, however, likely to be small and non-systematic with respect to the experimental hypotheses.

groups of 56 sentences each; for Blocks B and C, each set was split into three groups, two of 45 sentences and one of 46 sentences.

Each group of sentences was converted into a Postscript file, with one sentence in the centre of each page. These files were displayed on a Sun workstation monitor. Each subject took part in three recording sessions, reading through all of the six or eight groups of sentences for one block in each session. The order of presentation of the blocks was counterbalanced between subjects.

Before the first recording, subjects were given an instruction sheet telling them to read aloud the sentences as they appeared on the screen, and to emphasise the words in capital letters, speaking naturally and without pausing mid-sentence. Subjects were told to re-read a sentence if they made a mistake or did not place the emphasis in the correct place. They were told they could control the rate of presentation of the sentences, pressing the space bar on the computer keyboard to bring up the next. Before each session they were given a short practice, using a random sample of ten sentences from the block of sentences to be read.

During a recording session, subjects read the two or three groups of sentences from the first randomised set of sentences from a block, followed by the two or three groups from the second randomised set, pausing if they wished between groups of sentences. Subjects were not made aware of any systematic grouping or repetition of sentences. Each recording session lasted around half an hour and subjects attended three sessions, thereby reading all of the experimental sentences twice.

Subjects were asked to repeat sentences during the recording if:

- the lexical content of the sentence was misread.
- the words in capitals were not emphasised.
- words not in capitals were emphasised.
- a pause was perceived within the utterance.

Recordings were made direct to disk in ESPS format at a sample rate of 16 kHz.

4.3.2 Experimental subjects

Three females and three males took part in the recordings, all undergraduate or post-graduate students of Edinburgh University. Four of the subjects were speakers of Standard Southern British English; two were speakers of Northern English dialects. Subjects received five pounds after each recording session and an additional five pounds for completing all three sessions. Subjects were not given any specific information about the purpose of the recordings until after they had completed them.

4.3.3 Measurement of syllable duration

Each subject recorded 768 experimental sentences, excluding corrections—that is, two repetitions of each sentence—making a total of 4608 recorded sentences. UNIX shell scripts were used on a SparcStation for file processing; the recordings were analysed using XWaves.

Utterances were extracted from the original recordings and merged into sound files with other sentences for the same speaker and keyword, to facilitate the application of consistent measurement criteria for each test syllable. There were four sound files for each keyword for each subject:

- 2 repetitions × 7 sentences, utterance-medial, unaccented.
- 2 repetitions × 7 sentences, utterance-medial, accented.
- 2 repetitions × 5 sentences, utterance-edge, accented.
- 2 repetitions × 5 sentences, utterance-edge, accented.

Identifying labels were aligned with each utterance and a three-stage labelling procedure took place. Firstly, utterance production labels—for example, identifying incorrect placement of emphasis—were applied as appropriate. Secondly, the start and end points of the onset, nucleus and coda of each test syllable were identified by inspection of waveforms and spectrograms. Finally, the start and end points of the additional syllables within polysyllabic keywords were identified for a large subset of the utterances. Each of these stages is now described in detail.

Utterance production labelling

Each utterance was listened to in order to determine whether it had been produced correctly. This judgement involved the application of a number of criteria regarding pronunciation, placement of emphasis and the presence of audible boundaries within the utterance.

An utterance was labelled as excluded due to pronunciation if the speaker misread the words in the sentence, or pronounced the keyword differently from their other productions of the same keyword. In some cases, this label was applied during test syllable labelling, when it was determined that some acoustic feature of the keyword meant that segmentation criteria could not be reliably applied.

An utterance was labelled as excluded due to accent placement if the keyword was perceived as unaccented in the accented condition or accented in the unaccented condition. Often this judgement was straightforward; where it was more difficult, the decision was based on the relative prominence of the keyword and other words in the utterance. On a few occasions, reference was made to the fundamental frequency contour of the utterance, but this often proved inconclusive. If the keyword correctly carried or did not carry an accent, other misplacement of accent did not exclude the utterance from analysis: in the unaccented condition, where an accent was not placed on a word emphasised with block capitals but elsewhere in the utterance; in the accented condition, where there was an additional accent in the utterance to that on the keyword.

Another source of accent misplacement was the shifting of primary stress in polysyllables. For example, the right-headed word *reproduce* should in this experiment have the primary stress on the final syllable, but subjects occasionally placed the primary stress—thus also the accent if present—on the initial syllable of the word. Utterances in which this occurred were excluded.

Utterance production labels were also applied regarding the occurrence of boundaries within the utterance. The judgement of the presence of a boundary was perceptual, and was made where a juncture was perceived corresponding to an intonational phrase boundary: the primary criterion was the presence of an audible boundary tone, such as a steep fall or a fall-rise; additional criteria were the presence of an audible pause or perceptible pre-boundary lengthening, or the existence of two complete intonational phrases within the utterance. Visual inspection of fundamental frequency contours, which were often discontinuous, was seldom used. Utterances were excluded if an intonational phrase boundary was perceived adjacent to the keyword, but were not excluded if the boundary was elsewhere in the utterance.

The number of utterances excluded from analysis according to each criterion was:

- Utterance misread: 35 (0.8%)
- Keyword accented in the unaccented condition: 60 (1.3%)
- Keyword unaccented in the accented condition: 43 (0.9%).
- Primary lexical stress misplaced within the keyword: 29 (0.6%).
- Intonational phrase boundary adjacent to the keyword: 187 (4.1%)

Utterances for which both repetitions are excluded represent missing data points. There are 72 such utterances—3.1% of the total—listed in Tables D.1–D.3 in Appendix D. The number of missing data points is given in the results for each analysis; in most cases, less than 5% of the total data points are missing.

Test syllable labelling

The start and end points of subsyllabic constituents of each test syllable were labelled by visual inspection of the waveform and spectrogram. For consistency, labelling was performed keyword-by-keyword. To eliminate bias due to the order of labelling, each speaker's four files for each keyword were labelled in random order. The order of speakers for each keyword was also random.

Labels are placed for all keywords at the start point of each of subsyllabic constituent: onset, nucleus, and coda. The start of the nucleus marks the end of the onset and the start of the coda marks the end of the nucleus; the end of the coda is also labelled. It should be noted that singleton test syllable coda consonants could also be regarded as ambisyllabic or as unstressed syllable onsets in left-headed disyllables and trisyllables: for example *mason* and *masonry*. For consistency, such consonants are always labelled as codas in Experiment 2, a decision which appears to be supported by evidence from final lengthening reported in Section 4.8.

The right-headed *port*, *compose*, *dispose* and *suppose*, and the left-headed *ten*, *cap* and *part* all have voiceless stops as their syllable onset. In this experiment, the onset measurement includes both the duration of stop closure and the duration of aspiration following the release of the stop. There are arguments for treating aspiration as part of the nucleus, because the vocalic gesture begin in the oral cavity at the stop release; however, the laryngeal vocalic gesture begins with the onset of voicing, and this is taken as the start of the nucleus in Experiment 2. There are hypotheses which predict distinct effects in closure and aspiration duration, as discussed in Section 4.1.2, and so these are measured separately within the overall onset measurement. The decision to include aspiration as part of the onset appears to be justified by the results reported in Section 4.6 regarding the effect of word position on onset duration: the pattern of variation is very similar for stop closure duration and aspiration duration.

The words containing voiceless stops as their syllable onset thus all have an additional label placed at the point of stop release, corresponding to the start of aspiration[11]. Similarly, the test syllable onset of the keyword *juice* has a label at the position of plosive release in the affricate onset /dʒ/, corresponding to the start of frication: the duration of frication is wholly contained within the onset measure, but the extra label provide a partial-onset measure utterance-initially, where stop closure duration cannot be determined acoustically. The aspiration of voiceless stops similarly provides a partial-onset measure in utterance-initial position.

The placement of the labels for each keyword is illustrated in Tables E.1 and E.2 in Appendix E. The criteria used for segmentation are described in Appendix A. Labels were placed by reference to a speech waveform and a colour wideband spectrogram. For consistent application of visual criteria, judgements were made using a window size of not less than 400ms. As spectrogram colour indicates intensity, the appearance of a particular colour at a particular frequency was sometimes used to aid segmenta-

[11]Being preceded by an /s/, the /p in *speck* is unaspirated, and thus did not receive this additional label.

tion, for example, for the start of certain fricative consonants.

Not all constituents of all test syllables are reliably measurable in all contexts. As mentioned above, the period of stop closure in the syllable onsets for the right-headed keywords *compose*, *dispose*, *port*, *suppose* and *juice* does not have an acoustic correlate in utterance-initial position, and aspiration or frication duration is used where necessary as a partial-onset measure; full syllable duration is therefore available in utterance-edge conditions only for *main*, *mend* and *send*.

Coda stop consonants are sometimes difficult to measure reliably: the codas of the right-headed *port* and the left-headed *cap*, *dog*, *part* and *speck* are all occasionally elided or glottalised. For the right-headed keywords, *port* is excluded from measurements relating to coda duration, including whole syllable duration. For the left-headed keywords, measurements relating to coda duration, including whole syllable duration, are consistently available only for the keywords *fish*, *mace*, *sense* and *ten*. Where measurements are reliably available for all keywords, such as syllable onsets of left-headed keywords, or syllable nuclei of all keywords, then data are combined across the full set.

Additional syllable labelling

The additional syllables in disyllabic and trisyllabic keywords were measured for Keyword Series C and Keyword Series E (Series B and D feature only monosyllabic keywords). Because the primary comparison required for these syllables is between accented and unaccented keywords of the same length and in the same utterance position, it was not necessary to include Keyword Series A: that is, the fact that Series A has fixed utterance length is not relevant for the analysis of the effects of accent on additional syllable duration. Series C was used rather than Series A because the sentences in Series C have consistent structure within each keyword triad.

The additional syllables are labelled syllable-2 and syllable-3. For the left-headed keywords, syllable-2 immediately follows the primary stressed syllable in the disyllable and trisyllable, and syllable-3 follows syllable-2 in the trisyllable: for example, in *cap(tain)(cy)*, /tən/ is syllable-2 and /sɪ/ is syllable-3. For the right-headed keywords, syllable-2 immediately precedes the primary stressed syllable in the disyllable and trisyllable, and syllable-3 precedes syllable-2 in the trisyllable: for example, in *(mis)(re)port*, /ɹə/ is syllable-2 and /mɪs/ is syllable-3.

The labelling procedure was the same as for the test syllable, with labelling done keyword-by-keyword, and the order of speech files and speakers randomised. For the unstressed syllables, however, only full syllable durations were measured, for two main reasons. Firstly, due to the constraints on the experimental materials described in Section 4.2.1, the unstressed syllables have variable composition; thus, onset, nucleus and coda measures are not available for all syllables. Secondly, unstressed syllables are likely to manifest reduction and elision of segments; the variability in these processes place severe restrictions on the application of consistent segmentation criteria.

The placement of labels for unstressed syllables is illustrated in Tables E.3 and E.4 in Appendix E. The criteria used for segmentation are described in Appendix A.

For the right-headed keywords, syllable-initial stop closure durations could not be measured utterance-initially. Thus, for the disyllables *commend*, *dispose*, *produce*, *compose* and *descend*, syllable-2 duration is taken from the onset stop release. Similarly, for the trisyllables *presuppose*, *decompose* and *condescend*, syllable-3 duration is taken from the onset stop release. In addition, syllable-2 of the keyword *humane* is excluded from comparisons between disyllabic and trisyllabic contexts: the syllable-initial /h/ is realised quite differently in the two contexts (*develop humane* vs *inhumane*), and thus its durations are not considered comparable. It is included in other comparisons, such as those regarding the durational effects of accent, examined separately in disyllables and trisyllables.

Data processing

Test syllable constituent durations and additional syllable durations were derived from XWaves label files using UNIX shell scripts and extracted together with utterance production labels. Checks were carried out on this data. Firstly, missing or duplicated tokens were accounted for, and relabelled where necessary. Secondly, the range of values was inspected for each subsyllabic unit in each type of target syllable and outlying values were identified. The speech files corresponding to these tokens were inspected again: in some cases, misplaced labels were realigned; in other cases, unusually large or small subsyllabic durations were seen to have been correctly measured. The full set of durations was then recalculated prior to statistical analysis.

4.3.4 Statistical analysis

In the analyses reported in this chapter, items (that is, the test syllables being measured) are matched between experimental conditions. Therefore, as discussed in Chapter 3, the By-Subjects (F_1) analysis of variance is considered the appropriate statistical test. Following widespread practice, By-Items (F_2) analyses are also reported, but only By-Subjects statistics are taken as the benchmark of significance. Analysis of variance are performed using SPSS statistical software.

There are 72 missing data points, where utterances have both tokens excluded from analysis. Where less than about 5% of data are missing in a particular analysis, no action is taken; where a greater proportion are missing, the keyword with the most missing data points is excluded from the analysis—indicated in the text as "missing data"—this procedure being repeated until the proportion of missing data points is less than 5%. This step was only necessary in a few of the reported analyses. The total number of data points is reported below each table reporting mean duration in the following results sections; also noted is the proportion of the total data points that are missing.

Values of F, degrees of freedom[12] and significance levels are reported for the main analyses relating to each experimental hypothesis, and for subsidiary analyses where new results are obtained. Where subsidiary analyses simply corroborate established results, the results of ANOVAs are reported as "not significant", "approaching significance", "significant" ($p < .05$) or "highly significant" ($p < .01$). Except where otherwise stated, such reporting implies that By-Subjects and By-Items analyses have the same outcome.

4.4 Results overview

The 4608 recorded sentences represent a large data-set addressing the hypotheses outlined in Section 4.1.2 regarding domain-edge and domain-span processes and the durational effects of pitch accent. These hypotheses are tested by statistical analysis of durational variation in test syllable onset, nucleus and coda, and in the unstressed syllables in polysyllables. The analyses utilise different subsets of the full data-set to address different hypotheses, and are presented in later sections of this chapter, after an overview of the main results.

The most important findings in Experiment 2 are:

- There are no domain-span processes at word level (polysyllabic shortening) or utterance level, although there is evidence for a sub-word level process, word-rhyme-span compression.

- The word is the locus of accentual lengthening and the variation in distribution of accentual lengthening between monosyllables, disyllables and trisyllables accounts for previously-observed polysyllabic shortening: there is less lengthening of the stressed syllables in a disyllable than in a monosyllable, and less still in a trisyllable.

- Word-initial lengthening has a syllable onset locus; there may be additional lengthening of the onset phrase-initially.

- There is inverse relationship—word-rhyme-span compression—between stressed

[12]Where data are missing in a particular By-Subjects or By-Items analysis, the degrees of freedom reported may be non-integral.

syllable nucleus duration and word-rhyme length; the duration of unstressed syllables shows a similar inverse relationship with word-rhyme length[13].

- Word-initial stressed syllable onsets are, in some cases, shorter utterance-initially than utterance-medially.

- Utterance-final lengthening affects stressed syllable codas in penultimate and antepenultimate position, plus the following unstressed syllable(s); in absolute-final position, the stressed syllable nucleus and coda are lengthened. For all constituents, the lengthening effect is greater closer to the boundary.

- There is some evidence of compensatory shortening[14] in segments near to the loci of domain-edge lengthening processes.

These results are now described according to the experimental hypotheses formulated in Section 4.1.2.

Word-edge and word-span hypotheses

Word-edge and word-span results are described in detail in Section 4.6, where Figure 4.4 shows the effect of word length on onset, nucleus and coda duration. The results are outlined here according to the experimental hypotheses.

Word-span hypothesis: primary stressed syllable duration is inversely related to word length.

This hypothesis is not supported, as there is an inverse relationship between primary stressed syllable duration and word length only in accented words. For example, in the accented condition: /mɛnd/ is shorter in *commend* than in *mend*, and shorter still in *recommend*; likewise, /meɪs/ is shorter in *mason* than in *mace* and shorter still in *masonry*. These effects arise from variations in accentual lengthening, as described below.

Word-rhyme-span hypothesis: primary stressed syllable duration is inversely related to word-rhyme length.

Word-final hypothesis: the duration of the primary stressed syllable rhyme is greater word-finally than word-medially in utterance-medial words.

---

13 Subsyllabic durations are not measured for unstressed syllables, as noted in Section 4.3.3.

14 As described in Section 4.5.2, the term "compensatory" is used here to characterise durational effects in which some constituent shortens or lengthens apparently as a result of an opposite durational variation elsewhere, whatever the underlying interpretation of the effect.

These hypotheses are both supported, although the locus of the effect is not the whole syllable or the syllable rhyme: in accented and unaccented words, there is an inverse relationship between word-rhyme length and stressed syllable nucleus duration; this could be interpreted as a word-final effect, although the coda is not affected. Thus, /eɪ/ is shorter in *mason* than in *mace* and shorter still in *masonry*, even when unaccented. Unstressed syllables also show an effect of word-rhyme length in accented and unaccented left-headed keywords: for example, /ən/ is shorter in *masonry* than in *mason*.

Word-initial hypothesis: the duration of the primary stressed syllable onset is greater word-initially than word-medially in utterance-medial words.

Word-initial accent hypothesis: word-initial lengthening of stressed syllable onsets only occurs, or is significantly greater, in accented context.

The word-initial hypothesis is clearly supported, and the effect is comparable in accented and unaccented words, hence the word-initial accent hypothesis is not supported. Thus, /m/ is much longer in *mend* than in *commend* both when accented and when unaccented; there is, however, no difference in the duration of /m/ between *commend* and *recommend* when unaccented.

Word-initial aspiration accent hypothesis: word-initial lengthening of aspiration in stressed-syllable-onset voiceless stops only occurs in unaccented syllables; in accented syllables, aspiration duration may be greater word-medially.

This hypothesis is not supported: the effects of word position are similar for closure and aspiration duration of syllable-onset voiceless stops, and there appears to be no interaction with accent. Thus, the closure and aspiration of /p/ are longer in *pose* than in *suppose* or *presuppose*, both when accented and when unaccented.

Utterance-initial hypothesis

Utterance-initial results are described in detail in Section 4.7; Figure 4.5 shows the effect of utterance position on test syllable onsets in right-headed keywords, where these are fully-measurable utterance-initially.

Utterance-initial hypothesis: word-initial syllable onsets are shorter utterance-initially than utterance-medially.

This hypothesis is supported for certain stressed syllable onsets. The right-headed keywords *main*, *mend* and *send* have show utterance-initial shortening of stressed syllable onset duration: thus, /m is shorter in ‖*mend*... than in ‖...*mend* .... The right-headed keywords *port*, *compose*, *dispose*, *suppose* and *juice* show no effect of utterance position on aspiration/frication duration when accented and a slight lengthening of aspiration/frication duration utterance-initially when unaccented.

Utterance-final hypotheses

Utterance-final results are described in detail in Section 4.8, where Figure 4.9 shows the distribution of utterance-final lengthening in left-headed keywords.

Utterance-final locus hypothesis: the locus of utterance-final lengthening is the word-rhyme.

This hypothesis is supported, although lengthening does not occur on all parts of the word-rhyme. Where the stressed syllable is in utterance-final position, the nucleus and coda are lengthened: thus, /eɪs/ is longer in *... mace*‖ then in *... mace ...* ‖. The stressed syllable coda is lengthened in utterance-antepenultimate position and utterance-penultimate position, as are the following unstressed syllables within the word (subsyllabic durations are not obtained for unstressed syllables): thus, /s/ and /ən/ are longer in *... mason*‖ than in *... mason ...* ‖; likewise, /s/, /ən/ and /ɹɪ/ are longer in *... masonry*‖ then in *... masonry ...* ‖.

Utterance-final locus distribution hypothesis: the distribution of utterance-final lengthening is progressive within the locus.

This hypothesis is supported: lengthening is greater on the coda than on the nucleus of an absolute-utterance-final stressed syllable; and lengthening is greater on an absolute-final unstressed syllable than on an utterance-penultimate unstressed syllable.

Utterance-final accent hypothesis: the magnitude of utterance-final lengthening is not affected by the presence of a pitch accent on the utterance-final word.

This hypotheses is supported: there is no interaction between accent and utterance position (medial vs final), although there is a non-significant trend towards attenuation of utterance-final lengthening in accented words.

Utterance-span hypothesis

Utterance-span results are described in Section 4.5, where Figure 4.1 shows the lack of a systematic effect of utterance length on test syllable duration.

Utterance-span hypothesis: primary stressed syllable duration is inversely related to the number of syllables in an utterance.

This hypothesis is not supported: there is no systematic shortening of stressed syllable duration in longer utterances, in either accent condition.

Pitch accent hypotheses

Pitch accent results are described in detail in Section 4.9, where Figure 4.11 shows the distribution of accentual lengthening in left-headed and right-headed keywords.

Word-rhyme accent hypothesis: the locus of accentual lengthening is the word-rhyme.

Bimodal accent hypothesis: the locus of accentual lengthening is the primary stressed syllable and the word-final syllable.

These hypotheses are not supported: although there is accentual lengthening within the word-rhyme—on all parts of the accented syllable and on subsequent unstressed syllables—there is also lengthening in unstressed syllables preceding the word-rhyme, suggesting that the locus of accentual lengthening is the word. Thus, the syllables /ən/ and /ɹɪ/ are lengthened when *masonry* is accented (although this effect does not attain significance); similarly, the syllables /ɹɛ/ and /kəm/ are lengthened when *recommend* is accented. The magnitude of lengthening tends to be greater at word edges than word-medially.

Word-span accent hypothesis: the magnitude of accentual lengthening on the primary stressed syllable is inversely related to word-length.

The word-span accent hypothesis is supported: accentual lengthening on the stressed syllable is greatest in monosyllables and progressively less in disyllables and trisyllables. Total accentual lengthening is no greater in polysyllables than in monosyllables but is distributed between more syllables. The effect of this is what has previously been described as polysyllabic shortening, and may be re-interpreted as the polysyllabic accent effect: an inverse relationship between stressed syllable duration and word length in accented words only.

Locus hypotheses

The results of all parts of Experiment 2 inform the concept of the locus, which is discussed in Section 4.10.

Domain-edge locus hypothesis: each domain-edge process has a characteristic phonologically-defined locus.

The domain-edge locus hypothesis is supported: in particular, word-initial lengthening has a syllable onset locus and utterance-final lengthening affects certain constituents of the word-rhyme.

Domain-span locus hypothesis: a domain-span compression process shortens only the phonological head of the domain.

The domain-span locus hypothesis is partially supported. There is no support for word-span compression or utterance-span compression, but word-rhyme-span compression is supported and found to be localised within the stressed syllable on the nucleus. Unstressed syllables within the word-rhyme also appear to be affected, however, and it may be that the effect should be interpreted as word-final lengthening.

Summary

Mean test syllable durations[15] are reported in Table F.1 in Appendix F for the right-headed keywords *main*, *mend* and *send*, and in Table F.2 in Appendix F for the left-headed keywords *fish*, *mace*, *sense* and *ten*. The trends discussed above are discernible in the patterns of syllable durations according to length of the word and the utterance, and according to position within these domains.

Detailed results, statistical analysis and discussion are presented now for each set of experimental hypotheses. Utterance-span results are presented first, because the absence of an utterance-span effect affects the interpretation of the data for other analyses: in particular, it allows comparisons between different keyword series to be made without any allowance for utterance length. The discussion of results in each section focuses primarily on the interpretation of the present experiment; comparisons with previous results, particularly where these suggest different conclusions, are made in Chapter 5.

4.5 Utterance-span results

The utterance-span hypothesis predicts that primary stressed syllable duration is inversely related to the number of syllables in an utterance. As stated above, this hypothesis is not supported in the three sets of analyses presented in Section 4.5.1. The primary test of the hypotheses is provided by Keyword Series B, where keywords remain monosyllabic whilst utterance length varies. The durations of test syllable subconstituents in Series B are also compared with their durations in Series D, where monosyllabic keywords are in shorter utterances with different syntactic structures: evidence from these comparisons indicates some durational differences that appear to arise from domain-edge effects rather than utterance-span effects. The duration of test syllable subconstituents are also compared between Series A and Series C where disyllables and trisyllables are in utterances of varying length: evidence indicates no differences between the shorter Series A and the longer Series C utterances, allowing

[15]Total syllable durations cannot be reported for all keywords in all experimental conditions for reasons outlined in Section 4.3.3. The means in the summary tables are calculated without tokens for which both repetitions are missing, as also described in Section 4.3.3.

both to be used for the analysis of word-level effects in Section 4.6.

Section 4.5.2 discusses the lack of evidence for an utterance-span effect and the processes underlying the apparent domain-edge effects observed in the Series B vs Series D comparison.

4.5.1 Results and analysis

Keyword Series B

Example sentences in Keyword Series B are illustrated in Table 4.8. For the analysis of the effects of utterance length, left-headed and right-headed keywords are analysed together because there is no systematic difference between the two sets of monosyllables with respect to the utterance-span hypotheses[16].

Right-headed keywords	
$\|\ldots\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jessica **mend** it AGAIN.
$\|\ldots\sigma\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JOHNNY saw Jessica **mend** it AGAIN.
$\|\ldots\sigma\sigma\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JONATHAN saw Jessica **mend** it AGAIN.
Left-headed keywords	
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed AGAIN.
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed once AGAIN.
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed by them AGAIN.

Table 4.8: Example experimental sentences for Keyword Series B.

Figure 4.1 shows mean syllable duration in Keyword Series B. There is no effect of Utterance Length on syllable duration by Subjects: $F(2,10.2) = 0.62$, $p = .556$ [by Items: $F(2,22.4) = 1.11$, $p = .348$]. The lengthening effect of Pitch Accent is highly significant by Subjects: $F(1,5) = 30.91$, $p < .005$ [by Items: $F(1,5) = 264.81$, $p < .001$]. There is no interaction between Utterance Length and Pitch Accent by Subjects: $F(2,10.2) = 0.21$, $p = .817$ [by Items: $F(2,24.3) = 0.10$, $p = .901$].

Although the lack of support for the utterance-span hypothesis is clear at the syllable level, it may be that some subsyllabic constituent shows an effect which is masked by variation in the other subsyllabic constituents and so the results for onset, nucleus and coda are shown in Table 4.9. This confirms that there is no relationship between the number of syllables in the utterance and the duration of test syllable onset, nucleus or coda: ANOVAs show no significant effect of Utterance Length, and no sig-

[16] As already stated, the left-headed keyword *part* is excluded from all analyses here because its phonetic environment is not consistent throughout Series B; other keywords are excluded for certain analyses because of inconsistently measurable codas.

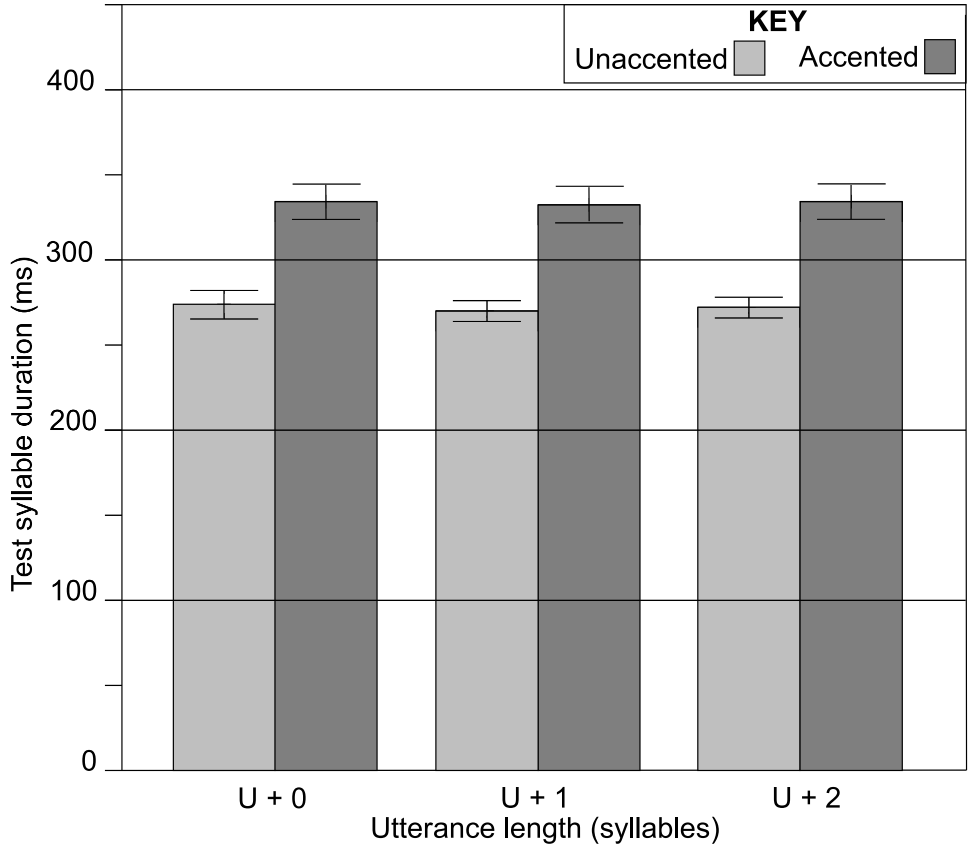

Figure 4.1: Mean test syllable duration (ms) in Keyword Series B for all keywords, excluding *part*, *cap*, *dog*, *speck* and *port*. Data points = 396; missing = 3.8%. Here and throughout, error bars represent two standard errors either side of the mean: this interval contains the population mean with 95% confidence.

nificant interaction between Utterance Length and Pitch Accent, for all subsyllabic constituents.

Therefore, all analyses of Series B data indicate no support for the utterance-span hypothesis.

Keyword Series B vs Keyword Series D

One possible reason for the lack of an utterance-span effect in Keyword Series B is that the difference in utterance length examined, only one or two syllables, might be insufficiently large relative to utterance length to cause a significant difference in test syllable duration. There are, however, other comparisons within the full set of experimental materials which provide a greater difference in utterance length: in particular, for some monosyllabic keywords, the longest sentences in Series B are four syllables

	Syllables added in utterance		
	0	1	2
Onset			
Unaccented	102	101	103
Accented	128	128	129
Nucleus			
Unaccented	91	90	90
Accented	105	107	105
Coda			
Unaccented	82	81	81
Accented	101	97	99

Table 4.9: Mean duration (ms) of subsyllabic constituents of the test syllable in Keyword Series B for right-headed and left-headed keywords. Means for onset and nucleus exclude keyword *part*: data points = 540; missing = 3.5%. Means for coda exclude keywords *part*, *cap*, *dog*, *speck* and *port*: data points = 396; missing = 3.8%.

longer than the longest sentences in Series D[17], as discussed in Section 4.2.1.

There are three such left-headed keywords examined, shown in Table 4.10. The utterance-span hypothesis predicts that the duration of the test syllable should be greater in the shorter Series D utterance than in the longer Series B utterance. It is possible, however, that proximity to the utterance boundary could also cause lengthening in the Series D test syllables: thus the locus of any effect should be examined to determine if the durational difference is greater closer to the boundary.

Series B.3 sentence	Series D.3 sentence	Difference in syllables
John THREW the **cap** to the big red BED again.	John DESIGNED the **cap** today	4
Kate GAVE the **sense** of the latest SCRIPT away.	Kate EXPLAINED the **sense** again.	4
I SAW the **mace** unreclaimed by them AGAIN.	Albert THREW the **mace** again.	4

Table 4.10: Sentences from Series B and Series D, containing left-headed keywords and differing in utterance length by four syllables.

Table 4.11 shows mean duration of the onset, nucleus and coda[18] for these three

[17] Although the shortest D.1 sentences from Series D provide a greater contrast in utterance length, they have keywords utterance-finally, where they are likely to be subject to final lengthening, and the D.2 sentences of intermediate length have keywords in utterance-penultimate position: as well as the possibility of lengthening due to proximity to the utterance boundary, the keywords are followed by full vowel syllables which could also cause lengthening of the test syllable.

[18] Although *cap* has an inconsistently measurable coda, in this data-set there are only two tokens for which neither repetition has a measurable coda.

keywords. There is little evidence of an utterance-span effect, even though the difference in utterance length is four syllables. Indeed, the direction of the durational difference in the syllable onset and nucleus is the opposite to that predicted by the utterance-span hypothesis.

	Keyword Series	
	B	D
Onset		
Unaccented	95	91
Accented	120	108
Nucleus		
Unaccented	88	83
Accented	98	96
Coda		
Unaccented	118	130
Accented	136	144

Table 4.11: Mean duration (ms) of subsyllabic constituents of the test syllable for left-headed keywords in Series B.3 and Series D.3 utterances: utterance length is four syllables greater in Series B than in Series D. Keywords used are *cap*, *sense* and *mace*: data points = 72; missing = 3 for onset and nucleus; missing = 5 for coda.

For syllable onsets, the effect of Utterance Length is significant by Subjects: $F(1,5) = 17.51$, $p < .01$ [by Items: $F(1,2.2) = 67.57$, $p < .05$]. For syllable nuclei, the effect of Utterance Length is not significant by Subjects: $F(1,5) = 0.34$, $p = .584$ [by Items: $F(1,2) = 16.91$, $p = .053$]. For syllable codas, the effect of Utterance Length approaches significance by Subjects: $F(1,5) = 4.43$, $p = .089$ [by Items: $F(1,2.1) = 32.17$, $p = .026$]. For onsets and nuclei, the effect of Accent is significant or highly significant; for codas, the effect is almost is significant by Subjects ($p = .051$) [by Items: not significant]. For neither onsets, nuclei nor codas is there a significant interaction between Utterance Length and Pitch Accent.

This pattern of results is somewhat surprising, with only coda durations showing any evidence of the durational effect predicted by the utterance-span hypothesis. It seems likely, however, that none of the results are actually related to utterance length as such. This contention is supported by a comparison between the Series B.1 utterances for the same keywords and the Series D.3 utterances used in the above comparison, as shown in Table 4.12: although the difference in utterance length is only two syllables, for onset, nucleus and coda, a very similar pattern of results emerges, the only difference being that the greater duration of test syllables codas in Series D is significant by Subjects ($p < .005$). If these effects were directly due to utterance length, a similar pattern should be seen in the comparison within Series B, where there is also an utterance length difference of two syllables.

Series B.1 sentence	Series D.3 sentence	Difference in syllables
John THREW the **cap** to the BED again.	John DESIGNED the **cap** today	2
Kate GAVE the **sense** of the SCRIPT away.	Kate EXPLAINED the **sense** again.	2
I SAW the **mace** unreclaimed AGAIN.	Albert THREW the **mace** again.	2

Table 4.12: Sentences from Series B and Series D, containing left-headed keywords and differing in utterance length by two syllables.

The result for test syllable onsets is the clearest and most unexpected: they have greater duration in the longer utterances. The difference may be related to the structure of the utterance, with a prosodic phrase boundary being more likely to precede the test syllables in the longer utterances. Examination of the Series B sentences in Table 4.10 suggests that this is unlikely, as the keywords are all preceded by *the*: the existence of a phrase boundary between a determiner and a noun would not be expected according to current theoretical accounts of prosodic structure. An alternative explanation for this observation, as what might be called a compensatory effect, is suggested in Section 4.5.2.

The nucleus shows least evidence of any utterance length effect, despite the fact that, as discussed in Section 4.1.2, domain-span effects might be expected to affect the nucleus maximally, as the phonological head of the syllable. There is a very slight lengthening of the nucleus in Series B compared with Series D, but this is not significant. Certainly, there is no evidence of an utterance-span compression effect.

Syllable codas, however, show a tendency, approaching significance, to be longer in the shorter utterance. Given the absence of any other support for the utterance-span hypothesis, and the fact that the same effect is obtained in the comparison between Series B.1 and D.3 utterances, with only two syllables difference in length, an interpretation in terms of utterance-final lengthening seems more plausible. In the Series D sentences, the test syllable is antepenultimate in the utterance: it is possible that the syllable coda may be lengthened in this position as well as in utterance-penultimate and utterance-final position: evidence for utterance-final lengthening is discussed further in Section 4.8.

Comparison of the right-headed keywords in Series B and Series D provides a further test of the effects of utterance length. Here, however, there are only two keywords—*suppose* and *send*—for which the longest sentences in Series B are four syllables longer than those in Series D; this comparison is therefore expanded to include keywords with three syllables difference in utterance length. Of these further four keywords, two—*dispose* and *juice*—have keywords in utterance-antepenultimate position in the Series B utterance: given the evidence just found for lengthening of the syllable coda in this position, these keywords are excluded from the analysis.

Table 4.13 shows the longest sentences from Series B and Series D for the keywords *suppose*, *send*, *main* and *port*. It should be noted that phonetic environment of the test syllables differs between these Series B and Series D utterances, raising the possibility of microprosodic effects influencing test syllable duration: such effects are unlikely, however, to induce a systematic bias in the results over a range of test syllables.

Series B.3 sentence	Series D.3 sentence	Difference in syllables
I LET the ballet dancer **pose** it TODAY.	Can you **pose** it TODAY please.	4
I MADE Peter Burgundy **send** to them ALL.	Will you **send** it TODAY please.	4
You MUST really continue **main** treatment NOW.	All the **main** roads SCARE me a lot.	3
I CHECKED each one in every **port** for TOM.	Can you **port** some QUICKLY please.	3

Table 4.13: Sentences containing right-headed keywords from Keyword Series B and Keyword Series D that differ in utterance length by three or four syllables.

Table 4.14 shows mean subsyllabic durations for keywords *suppose*, *send*, *main* and *port*. Both the nucleus and coda show some lengthening in the shorter Series D utterances compared with the longer Series B utterances. As for the left-headed keywords, syllable onset duration is greater in the longer utterances.

	Keyword Series	
	B	D
Onset		
Unaccented	108	99
Accented	134	124
Nucleus		
Unaccented	92	96
Accented	114	118
Coda		
Unaccented	60	65
Accented	74	85

Table 4.14: Mean duration (ms) of subsyllabic constituents of the test syllable for right-headed keywords in Series B.3 and D.3. Utterances are three or four syllables longer in Series B than Series D. Keywords used are *suppose*, *send*, *main* and *port*, the latter being excluded from the syllable coda duration means: data points for onset and nucleus = 96, missing = 1; data points for coda = 72, missing = 1.

For syllable onset, the effect of Utterance Length is significant by Subjects: $F(1,5) = 19.20$, $p < .01$ [by Items: $F(1,3) = 9.34$, $p = .055$]. For syllable nucleus, the effect of Utterance Length is not significant by Subjects: $F(1,5) = 1.58$, $p = .264$ [by Items: $F(1,3.1) = 23.32$, $p < .05$]. For syllable coda, the effect of Utterance Length is significant by Subjects: $F(1,5.1) = 15.85$, $p < .05$ [by Items: $F(1,2) = 4.98$, $p = .155$]. For onset and nucleus, the effect of Accent is significant or highly significant; for coda, the effect is significant by Subjects but not by Items. For neither onset, nucleus nor coda is there a significant interaction between Utterance Length and Pitch Accent.

As in the left-headed keyword results, there is little evidence of a direct relationship between utterance length and test syllable duration. This is supported by a comparison between the Series B.1 utterances and the Series D.3 utterances used in the above comparison, shown in Table 4.15, where the difference in utterance length between the two series is only one or two syllables: the pattern of results is very similar to that shown in Table 4.14, and the differences in onset and coda duration remain statistically significant. Because the durational effects are preserved, despite the smaller difference in syllable number, it must be concluded that the effects relate to the different structures in the two sets of utterances, rather than to their different lengths.

Series B.1 sentence	Series D.3 sentence	Difference in syllables
I LET the dancer **pose** it TODAY.	Can you **pose** it TODAY please.	2
I MADE Burgundy **send** to them ALL.	Will you **send** it TODAY please.	2
You MUST continue **main** treatment NOW.	All the **main** roads SCARE me a lot.	1
I CHECKED each one in every **port** for TOM.	Can you **port** some QUICKLY please.	1

Table 4.15: Sentences containing right-headed keywords from Keyword Series B and Keyword Series D that differ in utterance length by one or two syllables. Keywords are in bold.

The syllable onset, as for left-headed keywords, has greater duration in the longer utterance, but the phrase-initial lengthening explanation seems to be better motivated for right-headed keywords: examination of the Series B sentences in Tables 4.13 and 4.15 suggests that a prosodic boundary, such as a phonological phrase boundary, could occur preceding the keyword in some of the utterances. This seems most likely for the keywords *send* and *suppose*, where the test syllable is preceded by a noun phrase/verb phrase boundary; possible also for *main*, which follows a verb/noun phrase boundary; but unlikely for *port*, which is preceded by a noun-phrase-medial boundary. In the Series D sentences, syntactic structure and the fact that only two syllables precede the test syllables would tend to preclude the placement of phrase boundaries. The patterns of results for each keyword tends to support this analysis: *send* and *suppose* have at onsets at least 10 ms longer in Series B than in Series D, for both accented and unaccented keywords; *main* has 7 ms difference when accented and no difference when unaccented (where a preceding phrase boundary might be considered less likely); *port* has 10 ms difference when unaccented, but only 2 ms difference when accented. Thus, apart from *port* in the unaccented condition, the interpretation of the results as due to phrase-initial lengthening is supported. The relation between syntax and the occurrence of prosodic phrase boundaries is apparently probabilistic: that is, phrase boundaries are more likely preceding keywords in Series B than in Series D. The magnitude of the observed durational difference thus does not represent the true magnitude of phrase-initial lengthening, because only a proportion of the keywords in Series B will be phrase-initial. The evidence for phrase-edge processes is discussed further in Section 4.5.2.

For right-headed keywords, there is a small lengthening of the syllable nucleus in Series D compared with Series B, contrasting with the left-headed keywords, for which the trend was reversed. In neither case does the difference approach significance, however, and it must be concluded that there is no strong evidence of an utterance-span effect on test syllable nucleus duration.

The results for test syllable coda for right-headed keywords, as for left-headed keywords, are in the direction to support the utterance-span hypothesis, with greater duration in the shorter Series D utterances, but the lack of other evidence in favour of this, and the fact that the effect is comparable when the utterance length difference is one or two syllables to when it is three or four syllables, suggest that the explanation must lie elsewhere. For left-headed keywords, the utterance-antepenultimate position of the longer syllable codas suggests an utterance-final effect, but this explanation is not available for the right-headed syllables. Another possibility is that the effect is caused by utterance-medial phrase boundaries following the test syllables in the Series D utterances but not in the Series B utterances. The slight—but not significant—lengthening of the test syllable nucleus in Series D compared with Series B supports this explanation, as phrase-final lengthening would be expected to affect the syllable rhyme, but probably the coda more than the nucleus. Examination of the Series B sentences in Table 4.13 does not, however, provide compelling reasons to favour this interpretation: *pose* and *send* are followed by object pronouns which should show close prosodic attachment to the preceding verb; *main* is an adjective qualifying the following noun. A possible interpretation of this result as a compensatory effect is discussed in Section 4.5.2.

Keyword Series A vs Keyword Series C

As there appears to be no direct effect of utterance length on test syllable duration, Series A and Series C ought to show equivalent patterns of durational variation. In both series, word length is varied in utterance-medial position; the primary difference is that in Series A utterance length is fixed and in Series C utterance length covaries with word length.

For disyllabic keywords, the Series C utterance contains one more syllable than the Series A utterance, for example:

Series A.2 I saw Jessie commend it again.

Series C.2 I saw Alison commend it again.

For trisyllabic keywords, the Series C utterance contains two more syllables than the Series A utterance, for example:

Series A.3 I saw Jess recommend it again.

Series C.3 I saw Alison recommend it again.

Table 4.16 shows mean syllable durations for polysyllables in Series A and C, where no systematic difference is evident. For both disyllabic and trisyllabic comparisons in left-headed and right-headed contexts, analyses of variance confirm that there is no significant effect of Utterance Length and no interaction between Utterance Length and Accent. Furthermore, there are no such effects for the durations of onset, nucleus and coda analysed separately, except for onset duration in right-headed trisyllables, where there is a significant interaction between Utterance Length and Accent[19].

Given the lack of an utterance-span effect, Series C is used in addition to Series A in Section 4.6.1 to examine the effects of word length on test syllable duration.

4.5.2 Discussion

The utterance-span hypothesis predicts that primary stressed syllable duration is inversely related to utterance length. In the comparisons within Keyword Series B and between Keyword Series A and C, there is no significant effect of utterance length on test syllable duration. In the comparison between Series B and D, different parts of the syllable are affected in different ways by utterance length, suggesting that the observations may best be interpreted as arising from phrase-initial and utterance-final lengthening: furthermore, the structure of the sentences in question, rather than their length as such, appears to be the important factor.

This may underlie the observations by Lehiste (1974) and Rakerd *et al.* (1987) of a link between utterance length and durational variation: as observed in Chapter 2,

[19]Interaction by Subjects: $F(1,6.3) = 10.20$, $p < .05$ [by Items: $F(1,5.2) = 0.47$, $p = .521$]. In the unaccented condition, test syllable onsets from Series C right-headed trisyllables are 1 ms longer than those from Series A; in the accented condition they are 2 ms shorter: the magnitude of this interaction, together with the fact of it being entirely exceptional, indicate an anomalous result rather than an effect demanding further consideration.

	Keyword Series	
	A	C
Right-headed disyllables		
Unaccented	244	248
Accented	292	295
Right-headed trisyllables		
Unaccented	238	238
Accented	267	266
Left-headed disyllables		
Unaccented	286	286
Accented	331	327
Left-headed trisyllables		
Unaccented	279	276
Accented	306	302

Table 4.16: Mean test syllable duration (ms) in Series A and Series C. Series C utterances are longer by one syllable in the disyllable comparison, and by two syllables in the trisyllable comparison. All keywords except *port* are included for the right-headed disyllables: data points = 168; missing = 2.4%. All keywords except *port* and *dispose* (missing data) are included for the right-headed trisyllables: data points = 144; missing = 4.2%. Keywords *fish*, *mace*, *sense* and *ten* are included for left-headed disyllables and trisyllables: data points = 192; missing, for disyllables = 6.3%, for trisyllables = 2.1 %.

prosodic phrasing is influenced by the length of sentences and syntactic phrases. Thus sentence length may indirectly influence duration by affecting the number and placement of boundaries at which lengthening processes may occur. In general, longer sentences are likely to contain more prosodic boundaries rather than fewer, thus causing more aggregate lengthening, so the general trend would be in the opposite direction to that predicted by the utterance-span hypothesis.

The loci of variation, in the comparison between Series B and D, suggest domain-edge rather than domain-span interpretations: there is no significant durational variation on the test syllable nucleus—the phonological head of the syllable—but durational effects are seen on the onset and coda. The variations in test syllable onset and coda durations are considered in across the full set of keywords in Appendix G.

Three types of effects are suggested by this analysis: phrase-initial lengthening, utterance-final lengthening and compensatory shortening. Compensatory shortening is suggested to occur in association with both of the lengthening effects: the term "compensatory" is used here to characterise durational effects in which some constituent shortens or lengthens apparently as a result of an opposite durational variation elsewhere, whatever the underlying interpretation for the effect.

In Appendix G, there is a comparison of test syllable subconstituent duration between Series B.3 and Series D.3 utterances for right-headed keywords, such as:

Series B.3 I MADE Peter Burgundy **send** to them ALL.

Series D.3 Will you **send** it TODAY please.

There is evidence of lengthening of test syllable onsets in Series B compared with Series D, suggesting that in some cases the test syllable may be preceded by a phrase boundary[20]. There is also some evidence of lengthening of the coda for certain keywords in the Series D sentences, although this result is not significant across all keywords. The Series D sentences do not appear likely to have a phrase boundary following the test syllable, thus final lengthening cannot explain the greater length of the coda; rather, the observation of phrase-initial lengthening in the Series B onset and shortening of the Series B coda suggest that there may be a compensatory relationship between onset duration and coda duration.

In Appendix G, there is also a comparison between Series B.3 and Series D.3 for left-headed keywords, such as:

Series B.3 I SAW the **mace** unreclaimed by them AGAIN.

Series D.3 Albert THREW the **mace** again.

There is evidence of lengthening of the coda in Series D, which may be due to the utterance-antepenultimate position of the syllable. There is also evidence of lengthening of the onset in Series B; the structure of the sentences suggests that this is unlikely to be phrase-initial lengthening. The explanation may be compensatory shortening of the onset in Series D due to lengthening of the coda.

There is also some evidence of a compensatory effect associated with word-initial lengthening reported in 4.6.2. Possible interpretations of apparent compensatory effects are discussed in Chapter 5, as is the absence of evidence for utterance-span compression in relation to previous studies.

[20] As noted in Appendix G, the occurrence of a phrase boundary at a given point in a particular utterance is a matter of probability: syntactically more important boundaries are more likely to be realised prosodically. As keyword-adjacent intonational phrase boundaries are excluded from the final data-set, the phrases associated with these boundaries will be below the intonational phrase: for example, the phonological phrase. The question of (phonological) phrase-initial lengthening is considered further in Section 4.6 in relation to word-level results.

4.6 Word-edge and word-span results

As an overview of the durational effects of word length, primary stressed syllable durations are presented in Section 4.6.1 for monosyllables, disyllables and trisyllables. These results indicate that the word-span hypothesis—primary stressed syllable duration is inversely related to word length—is strongly supported only for accented keywords. Word-level effects are then examined in detail for subsyllabic constituents.

Results are reported first for right-headed keywords. Onset durations indicate support for the word-initial hypothesis: the duration of the primary stressed syllable onset is greater word-initially than word-medially in utterance-medial words. The hypotheses regarding the interaction of word-initial lengthening and accent are then examined. The word-initial accent hypothesis is not supported: thus, once the effects of word length in accented words are taken into account, word-initial lengthening of stressed syllable onsets is not significantly greater in accented syllables than unaccented syllables. The word-initial aspiration accent hypothesis is also not supported: onset aspiration duration is longer word-initially in both unaccented and unaccented syllables. The possible influence of phrase-initial position on syllable onset duration, suggested in Section 4.5, is also examined. The results for right-headed keywords for nucleus and coda duration are reported, in general indicating support for the word-span hypothesis in accented keywords only.

Results for left-headed keywords are then reported. Onset durations suggest that word-initial lengthening may be greater in monosyllables than in polysyllables. Nucleus duration suggest support for the word-span hypothesis and also for word-rhyme-span hypothesis: primary stressed syllable duration is inversely related to word-rhyme length. Coda duration indicates support for the word-span hypothesis in accented words only, and does not offer support for the word-final hypothesis—that the duration of the primary stressed syllable rhyme is greater word-finally than word-medially in utterance-medial words—because in unaccented words, there is no effect of word position on the coda.

Unstressed syllable durations are also reported. The results for right-headed keywords offer some support for word-initial lengthening, plus a word-span effect in accented keywords only. The results for left-headed keywords could be taken as support for word-rhyme-span compression or for word-final lengthening .

Section 4.6.2 discusses the implications of the results for the four types of process considered at the word-level: word-span compression; word-rhyme-span compression; word-final lengthening; and word-initial lengthening.

4.6.1 Results and analysis

Primary stressed syllable duration

Example sentences in Series A, where the number of syllables in the word varies but the total number of syllables in the utterance is fixed, are shown in Table 4.17.

Right-headed keywords	
$\|\ldots\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jessica **mend** it AGAIN.
$\|\ldots\sigma\sharp\sigma\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jessie **commend** it AGAIN.
$\|\ldots\sharp\sigma\sigma\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jess **recommend** it AGAIN.
Left-headed keywords	
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed AGAIN.
$\|\ldots\sharp\boldsymbol{\sigma}\sigma\sharp\sigma\ldots\|$	I SAW the **mason** reclaimed it ALL.
$\|\ldots\sharp\boldsymbol{\sigma}\sigma\sigma\sharp\ldots\|$	I SAW the **masonry** cleaned AGAIN.

Table 4.17: Example sentences for Keyword Series A.

Mean test syllable duration for right-headed keywords in Series A is shown in Figure 4.2. The effect of Word Length is highly significant by Subjects: F(2,10.1) = 37.43, p < .001 [by Items: F(2,12.4) = 71.12, p < .001]. The effect of Accent is highly significant by Subjects: F(1,5) = 29.38, p< .005 [by Items: F(1,6.2) = 308.99, p< .001]. There is a highly significant interaction between Word Length and Accent by Subjects: F(2,10.2) = 23.99, p < .001 [by Items: F(2.62) = 11.90, p < .01]. Planned comparisons indicate that the difference in test syllable duration between monosyllables and disyllables is highly significant for both accented and unaccented keywords; the difference between disyllables and trisyllables is highly significant for accented keywords but not significant for unaccented keywords. Two processes are suggested by these results: a word-initial lengthening effect in accented and unaccented keywords, contributing to the difference in, for example, *mend* vs *commend*; and a word-span effect in accented keywords only, contributing to the difference in *mend* vs *commend* and *commend* vs *recommend*. Examination of subsyllabic durations below supports these conclusions.

Mean stressed syllable duration for left-headed keywords in Series A is shown in Figure 4.3. The effect of Word Length is highly significant by Subjects: F(2,10.1)= 12.84, p < .005 [by Items: F(2,6) = 9.97, p < .05]. The effect of Accent is significant by Subjects: F(1,5) = 15.81, p < .05 [by Items: F(1,3) = 57.06, p = .005]. The interaction between Word Length and Accent is also highly significant by Subjects: F(2,10.2) = 7.71, p < .01 [by Items: F(2,6.2) = 11.90, p < .01]. Planned comparisons indicate that the difference in test syllable duration between monosyllables and disyllables approaches significance for accented keywords (p = .057) but not significant for unaccented keywords;

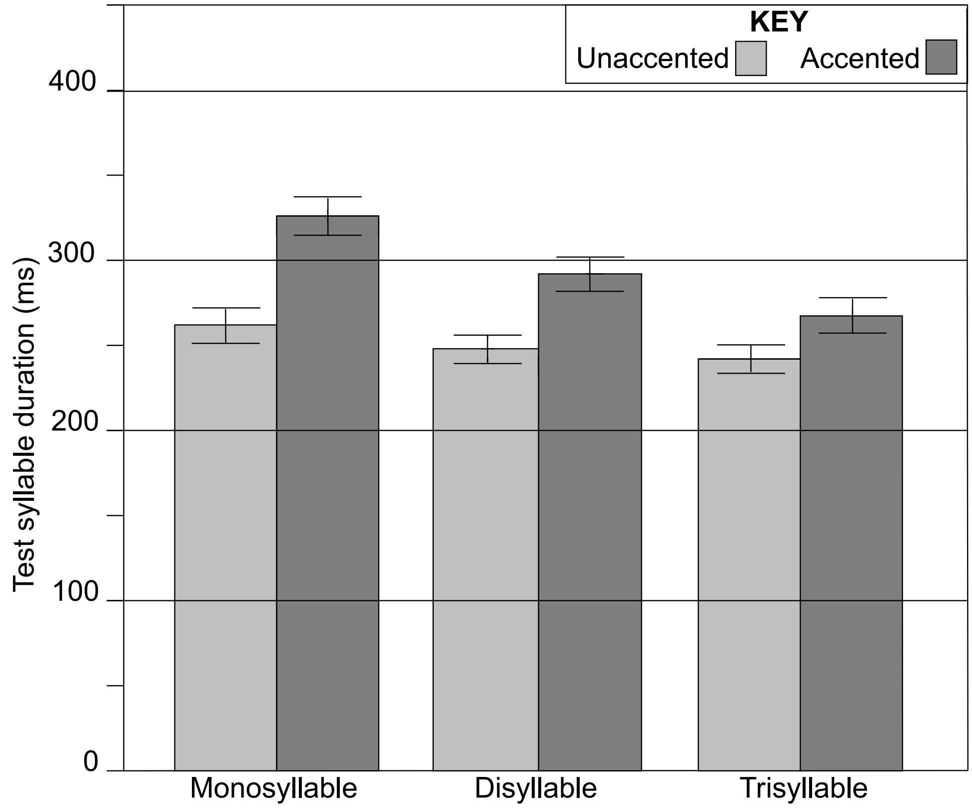

Figure 4.2: Mean test syllable duration (ms) in Series A for right-headed keywords excluding *port*: data points = 252; missing = 3.8%.

the difference between disyllables and trisyllables is highly significant for accented keywords but not significant for unaccented keywords. The difference between monosyllables and trisyllables is, however, significant for unaccented keywords as well as for accented keywords. As for right-headed keywords, there is evidence of a word-span effect in accented keywords only. The slight shortening effect in unaccented keywords, significant in, for example, *mace* vs *masonry* suggests a word-rhyme-span effect, because there is no comparable difference in right-headed keywords. Analysis of subsyllabic durations below indicates that the locus of this effect is the syllable nucleus.

The interaction between word length and accent for left-headed and right-headed keywords arises because the effect of word length is greater in accented keywords than in unaccented keywords. This is illustrated in Table 4.18: the absolute and proportional shortening of the test syllable due to an additional syllable in the word is substantially greater in the accented condition. An alternative view of this effect is that accentual lengthening of the test syllable is greater in words of fewer syllables, as discussed in more detail later. Apart from the monosyllable vs disyllable (and monosyllable vs trisyllable) comparison for right-headed keywords and the monosyllable vs

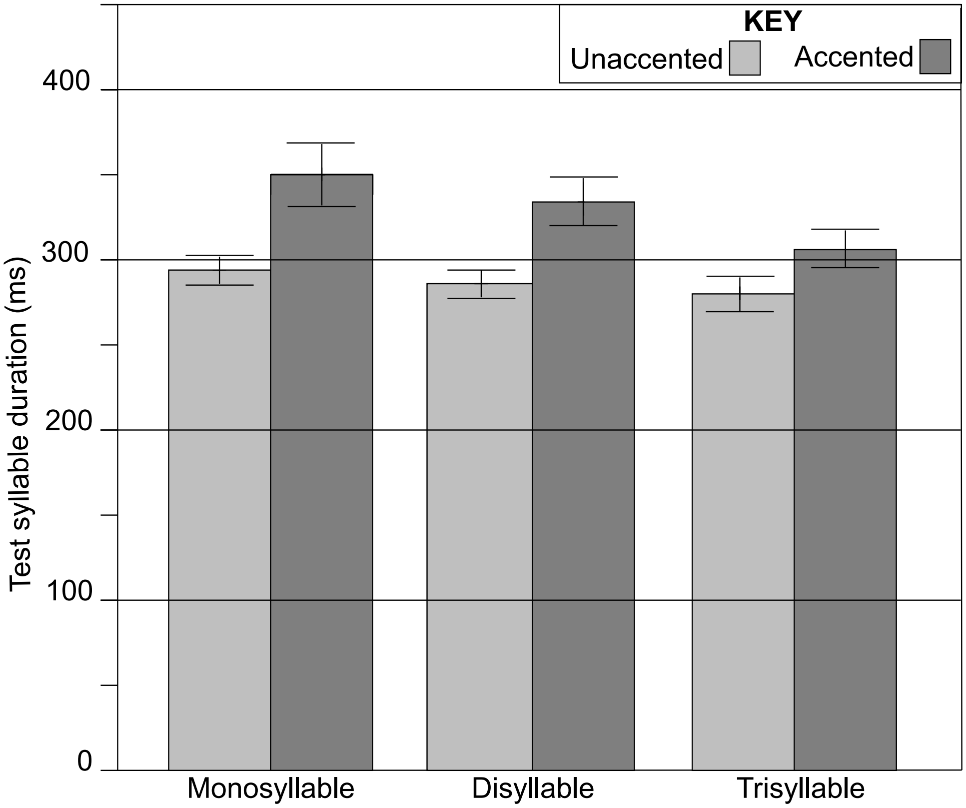

Figure 4.3: Mean test syllable duration (ms) in Series A for left-headed keywords *fish*, *mace*, *mend* and *send*: data points = 144; missing = 2.0%.

trisyllable for left-headed keywords, none of the differences in unaccented keywords are significant.

Table 4.18 also shows that for the comparison between monosyllables and disyllables, the effect of word length is much greater in right-headed keywords than in left-headed keywords, due to the large word-initial lengthening effect: analysis of sub-syllabic durations, reported below, indicates that the locus of this effect is the syllable onset. Thus, /m/ is longer in *mend* than in *commend* or *recommend*.

The magnitude of the effect between disyllabic and trisyllabic keywords is comparable in Table 4.18 in left-headed and right-headed accented keywords, suggesting a word-span effect. Right-headed unaccented keywords show very little effect of word length, beyond that explicable as word-initial lengthening. Left-headed unaccented keywords show a slightly larger effect, comparable between monosyllables vs disyllables and disyllables vs trisyllables. This suggests that the word-final hypothesis is not supported: the loss of word-final lengthening in, for example, *fissure* vs *fish*, should cause a large shortening effect in /fɪʃ/, whereas there should be little or no difference in syllable duration between *fisherman* vs *fissure*.

	Disyllable vs Monosyllable		Trisyllable vs Disyllable	
Right-headed keywords	18 ms 6.9%	**34 ms** **10.4%**	3 ms 1.2%	**24 ms** **8.2%**
Left-headed keywords	7 ms 2.4%	**18 ms** **5.2%**	7 ms 2.4%	**25 ms** **7.6%**

Table 4.18: The absolute and proportional shortening of the test syllable in disyllables compared with monosyllables, and trisyllables compared with disyllables. Figures on the left in each cell are for the unaccented condition; figures on the right in bold are for the accented condition.

All of the hypotheses are now considered further by examining the distribution of durational variation in the subsyllabic constituents, for right-headed keywords and for left-headed keywords. The primary test of these hypotheses is Keyword Series A, where word length varies as utterance length remains fixed. As the results presented in Section 4.5.1 indicate no direct effect of utterance length on test syllable duration, Keyword Series C, in which word length and utterance length covary, is used as an additional test of the word-level hypotheses. Series C is illustrated in Table 4.19: the first sentence in each set, which contains the monosyllabic keyword, is also in to Series A, and so results reported for monosyllabic keywords are identical, except where a different subset of the data is used.

Right-headed keywords	
$\|\ldots\sigma\sigma\sharp\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Jessica **mend** it AGAIN.
$\|\ldots\sigma\sigma\sharp\sigma\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Alison **commend** it AGAIN.
$\|\ldots\sigma\sigma\sharp\sigma\sigma\boldsymbol{\sigma}\sharp\ldots\|$	JOHN saw Alison **recommend** it AGAIN.
Left-headed keywords	
$\|\ldots\sharp\boldsymbol{\sigma}\sharp\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed AGAIN.
$\|\ldots\sharp\boldsymbol{\sigma}\sigma\sharp\sigma\sigma\ldots\|$	I SAW the **mason** disinclined it ALL.
$\|\ldots\sharp\boldsymbol{\sigma}\sigma\sigma\sharp\sigma\sigma\ldots\|$	I SAW the **masonry** disinterred AGAIN.

Table 4.19: Example experimental sentences for Keyword Series C.

In the analysis of subsyllabic durational variation, results are presented for Series A. The results for Series C are in Appendix H: subsyllabic mean durations are given in Table H.1 for right-headed keywords and in Table H.3 for left-headed keywords; ANOVA results for Series C are shown in Table H.2 for right-headed keywords and Table H.4 for left-headed keywords. Results for Series A and C are compared in the following discussion, but the details of statistical tests for Series C are not given where the pattern of results is equivalent to that for Series A (this includes cases where the difference is merely in the level of significance).

Right-headed keywords: onset duration

Table 4.20 shows mean test syllable onset duration for right-headed keywords in Series A. There is a highly significant effect of Word Length by Subjects: $F(2,10.1) = 64.49$, $p < .001$ [by Items: $F(2,14.1) = 44.93$, $p < .001$]. The effect of Accent is highly significant by Subjects: $F(1,5) = 30.13$, $p < .005$ [by Items: $F(1,7.2) = 231.03$, $p < .001$]. There is also a highly significant interaction between Word Length and Accent by Subjects: $F(2,10.1) = 21.62$, $p < .001$ [by Items: $F(2,14.7) = 47.00$, $p < .001$].

Planned comparisons reveal that the difference in onset duration between monosyllables and disyllables is highly significant for both accented and unaccented contexts. Thus, for example, /m/ is shorter in *commend* than in *mend*, and the effect is greater in accented keywords (31 ms, 24%) than in unaccented keywords (17 ms, 17%). There is no significant difference in onset duration between unaccented disyllables and trisyllables, but there is a highly significant shortening effect (13 ms, 13%) between accented disyllables and trisyllables: thus, for example, /s/ is shorter in *condescend* than *descend*.

	Number of syllables in word		
	1	2	3
Unaccented	99	82	81
Accented	128	97	84

Table 4.20: Mean test syllable onset duration (ms) for all right-headed keywords in Series A: data points = 288; missing = 4.5%.

The data-set for Series C for syllable onset excludes the keyword *port* (missing data), but the pattern of results (Appendix H) is very similar to that for Series A: a large shortening effect between monosyllables and disyllables for both accent conditions and a further shortening effect between disyllables and trisyllables for accented keywords only. As in Series A, the effects of Word Length, Accent and the interaction between Word Length and Accent are all highly significant.

The results for onset duration support the word-initial lengthening hypothesis: the duration of the syllable onset is increased by adjacency to the initial boundary of an utterance-medial word. The further shortening, in accented keywords only, between disyllables and trisyllables indicates a word-span effect which appears dependent on the presence of pitch accent. An alternative view of this interaction is that accentual lengthening on the syllable onset is greater in words of fewer syllables.

Interaction between word-initial lengthening and accent: onset duration

The word-initial accent hypothesis predicts that word-initial lengthening of stressed syllable onsets is greater in accented syllables. This hypothesis is apparently supported: word-initial lengthening occurs in the primary stressed syllable onset of both accented and unaccented keywords—thus, /m/ is longer in *mend* than in *commend*—but the effect is greater in accented keywords. Not all of the durational difference between accented monosyllables and disyllables may be due to the word-initial effect, however: a proportion may result from the word-span compression effect. The evidence from the disyllable vs trisyllable comparison suggests a word-span effect on test syllable onsets only in the accented context: thus, /m/ is longer in *commend* than in *recommend* in the comparison between accented keywords but not between unaccented keywords. This word-span effect in accented words could account for the apparently greater word-initial effect in the presence of accent. If so, word-initial lengthening as such is independent of pitch accent: in an accented word, an onset consonant seems to gain extra duration due to word-initial position which is actually attributable to its being in a monosyllable rather than a disyllable (for example: /m/ in *mend* rather than *commend*). If this interpretation is correct, then the word-initial accent hypothesis is not in fact supported, and there is no interaction between word-initial lengthening and accent.

Interaction between word-initial lengthening and accent: voiceless stop closure and aspiration duration

The word-initial aspiration accent hypothesis states that word-initial lengthening of aspiration in stressed-syllable-onset voiceless stops only occurs in unaccented syllables; in accented syllables, onset aspiration duration may be longer word-medially. Table 4.21 shows the mean test syllable closure and aspiration duration in Series A for the right-headed keywords which have voiceless stops as their test syllable onset: *compose*, *port* and *suppose*[21]. The evidence for both a word-initial effect and a word-length effect is comparable to that found for syllable onset duration.

For closure duration, the effect of Word Length is highly significant by Subjects: $F(2,10.5) = 45.45$, $p < .001$ [by Items: $F(2,4) = 7.20$, $p = .047$]. The effect of Accent is significant by Subjects: $F(1,5) = 6.70$, $p < .05$ [by Items: $F(1,2.1) = 139.06$, $p < .01$]. The interaction between Word Length and Accent is highly significant by Subjects: $F(2,10.9) = 23.08$, $p < .001$ [by Items: $F(2,4.5) = 32.93$, $p < .005$]. Planned comparisons indicate that the shortening in closure duration between monosyllables and disyllables is significant for unaccented words (9 ms, 13%) and highly significant for accented words

[21]The keyword *dispose* is excluded because the test syllable onset is unaspirated in disyllabic and trisyllabic contexts, indicating the syllabification /dɪ.spəʊz/.

	Number of syllables in word		
	1	2	3
Closure duration			
Unaccented	69	60	61
Accented	85	69	58
Aspiration duration			
Unaccented	43	36	33
Accented	56	44	34

Table 4.21: Mean test syllable onset closure and aspiration duration (ms) in Keyword Series A for keywords *compose*, *suppose* and *port*: data points = 108; missing = 6.

(16 ms, 19%). Between disyllables and trisyllables, unaccented keywords show no significant difference and accented keywords show a highly significant shortening effect (11 ms, 16%).

The pattern of results for aspiration duration is very similar. The effect of Word Length is highly significant by Subjects: $F(2,10.1) = 20.19$, $p < .001$ [by Items: $F(2,4.1) = 18.63$, $p < .01$]. The effect of Accent is significant by Subjects: $F(1,5) = 9.25$, $p < .05$ [by Items: $F(1,2.1) = 38.62$, $p < .05$]. The interaction between Word Length and Accent is significant by Subjects: $F(2,10.2) = 6.50$, $p < .05$ [by Items: $F(2,4.2) = 5.98$, $p = .059$]. Planned comparisons indicate that the shortening in aspiration duration between monosyllables and disyllables is significant for unaccented words (7 ms, 16%) and highly significant for accented words (12 ms, 21%). Between disyllables and trisyllables, unaccented keywords show no significant difference and the accented keywords show a highly significant shortening effect (10 ms, 23%).

For both onset closure and aspiration duration for *compose*, *suppose* and *port*, the pattern of results shown in Table 4.21 closely resembles that seen for onset duration in the full data-set: a shortening effect between monosyllables and disyllables which is somewhat greater in accented keywords and a shortening effect between disyllables and trisyllables only evident in accented keywords. This suggests that, as discussed above, the word-initial effect is complemented in the accented context by a word-span compression effect, thus increasing manifest size of word-initial lengthening. The present results suggest that either the previously-observed word-initial shortening of aspiration duration in accented syllables is not robust, or it is not observed here because word-span shortening between monosyllables and disyllables in the accented condition serves to mask the effect of word position.

Word-initial effect: influence of phrase boundaries.

Results reported in Section 4.5.1 suggest that monosyllabic keywords in utterance-medial position may in some cases be preceded by phrase boundaries which are not present when the keywords are near the start of the utterance; thus, the apparent word-initial lengthening effect could actually be a result of phrase-initial lengthening.

To test the effect of word-initial position in the absence of phrase boundaries, test syllable onsets are examined in Series D, where they are near the start of the utterance and unlikely to be immediately preceded by a phrase boundary, and compared with Series E, which has polysyllables in similar utterances. Series B utterances, where test syllables in monosyllabic keywords appear to have a higher probability of occurring in phrase-initial position, are included for reference. Example utterances in these experimental conditions are shown in Table 4.22.

B.2	JOHNNY saw Jessica **mend** it AGAIN.
D.2	Now **mend** it AGAIN for me please.
E.2	**Commend** it AGAIN for me please.
B.3	JONATHAN saw Jessica **mend** it AGAIN.
D.3	Will you **mend** it AGAIN for me please.
E.3	**Recommend** it AGAIN for me please.

Table 4.22: Example utterances in Keyword Series B, D and E.

Table 4.23 shows the mean test syllable onset duration (ms) for right-headed keywords in Series B, D and E, excluding those in the shortest utterance length conditions (B.1, D.1, E.1). A number of trends may be observed, apart from the clear lengthening effect of pitch accent throughout. Firstly, the comparison within Series E illustrates the effect of word length discussed above: onsets are shorter in trisyllables than in disyllables, and this effect is greater in accented keywords. Secondly, the comparison between Series B.3 and Series D.3 shows the influence of phrase boundaries indicated in Section 4.5.1: test syllable onsets in monosyllables are slightly longer utterance-medially, where they seem more likely to be preceded by a phrase boundary, such as that between phonological phrases; this effect does not depend on accent[22]. Thirdly, the comparisons between Series D and Series E show that there is a large difference in test syllable onset duration between monosyllables and disyllables (D.3 vs E.2), and between monosyllables and trisyllables (D.3 vs E.3), even when the monosyllabic keyword is unlikely to be phrase-initial; this effect is greater in accented keywords but still large in unaccented keywords.

[22]The difference between word-initial and potentially phrase-initial syllable onsets is not evident in the B.2 vs D.2 comparison in Table 4.23. Comparison within Series D suggests that onsets are slightly lengthened in the D.2 utterances, possibly because they are preceded by the full vowel syllable *now* in all

	Keyword Series		
	B	D	E
	B.2	*D.2*	*E.2*
Unaccented	97	98	79
Accented	129	127	90
	B.3	*D.3*	*E.3*
Unaccented	100	95	76
Accented	129	122	83

Table 4.23: Mean test syllable onset duration (ms) for right-headed keywords in Series D and Series E. All keywords are included for the U+1 comparison: data points = 192; missing = 2.1%. All keywords except *dispose* (missing data) are included for the U+2 comparison: data points = 168; missing = 4.2%.Test syllable onset duration (ms) for right-headed keywords in Series B is included for reference, for the same keywords in each comparison.

The difference in onset duration between Series D.2 and Series E.2, an effect of Word Length, is highly significant by Subjects: $F(1,5) = 128.12$, $p < .005$ [by Items: $F(1,7) = 26.41$, $p < .005$]. The effect of Accent is highly significant by Subjects: $F(1,5) = 34.61$, $p < .005$ [by Items: $F(1,7) = 149.95$, $p < .001$]. The interaction between Word Length and Accent is highly significant by Subjects: $F(1,5) = 40.87$, $p < .005$ [by Items: $F(1,7.1) = 66.53$, $p < .001$]. The effect of Word Length evident between Series D.3 and Series E.3 is highly significant by Subjects: $F(1,5) = 112.08$, $p < .001$ [by Items: $F(1,7) = 30.17$, $p < .005$]. The effect of Accent is highly significant by Subjects: $F(1,5) = 22.57$, $p < .01$ [by Items: $F(1,7.5) = 443.59$, $p < .001$]. The interaction between Word Length and Accent is highly significant by Subjects: $F(1,5) = 19.25$, $p < .01$ [by Items: $F(1,7.3) = 59.17$, $p < .001$].

The evidence strongly indicates word-initial lengthening as well as phrase-initial lengthening. The word-initial effect is greater in accented keywords; as discussed in Section 4.5.1, the phrase-initial effect does not show such an interaction. Given that word-initial and phrase-initial processes might be expected to manifest similar interactions, this observation supports the interpretation of the word-level effects discussed above: word-initial lengthening comparable in accented and unaccented keywords, plus word-span compression in accented keywords only.

Right-headed keywords: nucleus duration

Table 4.24 shows mean test syllable nucleus duration for the right-headed keywords in Series A, where there is no strong evidence for either a word-initial effect or a word-

cases, except for *main*, preceded by *the*.

span effect.

	Number of syllables in word		
	1	2	3
Unaccented	97	97	97
Accented	114	118	111

Table 4.24: Mean nucleus duration (ms) in the test syllable for right-headed keywords in Series A: data points = 288; missing = 4.5%.

The effect of Word Length is not significant by Subjects: $F(2,10.2) = 2.22$, $p = .159$ [by Items: $F(2,14.6) = 1.89$, $p = .187$]. The effect of Accent is significant by Subjects: $F(1,5) = 13.08$, $p < .05$ [by Items: $F(1,7) = 23.82$, $p < .005$]. The interaction between Word Length and Accent is not significant by Subjects: $F(2,10.2) = 2.63$, $p = .119$ [by Items: $F(2,14.5) = 3.72$, $p = .05$]. Planned comparisons indicate that none of the differences in vowel duration are significant for either accented or unaccented keywords, although in the accented condition the difference in vowel length between disyllabic and trisyllabic context approaches significance ($p = .077$).

These results strongly suggest that the word-initial lengthening seen for onset duration does not extend to the vowel of the word-initial syllable. Furthermore, there is little evidence of the word-span effect suggested by previous results for accented words, except for a non-significant difference between disyllables and trisyllables. Nucleus duration is actually slightly shorter in accented monosyllables than accented disyllables : a possibly compensatory relationship between this effect and word-initial lengthening is discussed in Section 4.6.2.

The results for nucleus duration in Series C (Appendix H) are similar, with little clear evidence of a word-span effect: nucleus duration is slightly longer in disyllables than monosyllables, and shorter in trisyllables than disyllables, particularly for accented keywords. The effect of Word Length is, however, significant in a By-Subjects analysis: $F(2,10.2) = 5.79$, $p < .05$ [by Items: $F(2,12.2) = 3.70$, $p = .055$]. The effect of Accent is significant and the interaction between Word Length and Accent not significant. The results of planned comparisons are equivalent to those for Series A: for both accented and unaccented keywords, there is no significant durational difference in pairwise comparisons of Word Length; once again, only the difference between accented disyllables and trisyllables approaches significance ($p = .067$).

Right-headed keywords: coda duration

Table 4.25 shows mean test syllable coda duration for right-headed keywords in Series A. Although differences in coda duration according to word length are small, particularly for the unaccented keywords, they are in the direction predicted by the word-span hypothesis.

	Number of syllables in word		
	1	2	3
Unaccented	69	68	66
Accented	85	80	75

Table 4.25: Mean test syllable coda duration (ms) for right-headed keywords in Series A excluding keyword *port*: data points = 252; missing = 3.8%.

The effect of Word Length illustrated in Table 4.25 is significant by Subjects: $F(2, 10.1) = 4.92$, $p < .05$ [by Items: $F(2,13.1) = 11.75$, $p = .001$]. The effect of Accent is highly significant by Subjects: $F(1,5) = 27.07$, $p < .005$ [by Items: $F(1,6) = 8.01$, $p < .05$]. The interaction between Word Length and Accent is highly significant by Subjects: $F(2,10.8) = 7.55$, $p < .01$ [by Items: $F(2,14) = 4.92$, $p < .05$]. Planned comparisons indicate that none of the differences in coda duration due to Word Length are significant in the unaccented condition, but coda duration is significantly greater in monosyllables than trisyllables in the accented condition.

The pattern of variation of coda duration is similar in Series C (Appendix H), with evidence of a small word-span effect, greater in accented keywords. The effect of Word Length is highly significant in Series C, as is the effect of Accent. For Series C, however, the interaction between Word Length and Accent is not found to be significant by Subjects: $F(2,10.2) = 2.40$, $p = .131$.

Right-headed keywords: summary

The results for the subsyllabic constituents of test syllables in right-headed keywords suggest two durational processes. Firstly, a large word-initial lengthening effect, with a syllable onset locus. Secondly, a smaller word-span effect only present, or only of significant magnitude, in accented words. The locus of the word-span effect appears to be the whole syllable, but durational variation is more evident on the onset and coda than on the nucleus, where evidence for a word-span effect is slight. These results are discussed further in Section 4.6.2.

Left-headed keywords: onset duration

Table 4.26 shows mean test syllable onset duration for left-headed keywords in Series A. There is a significant effect of Word Length by Subjects: $F(2,10.1) = 5.97$, $p < .05$ [by Items: $F(2,10.1) = 7.71$, $p < .01$]. The effect of Accent is also significant by Subjects: $F(1,5) = 10.8$, $p < .05$ [by Items: $F(1,5) = 188.85$, $p < .001$]. The interaction between

Word Length and Accent is not significant by Subjects: $F(2,10.4) = 1.48$, $p = .271$ [by Items: $F(2,10.5) = 7.02$, $p < .05$]. Planned comparisons indicate that for both unaccented and accented keywords, the difference between monosyllables and disyllables is significant.

	Number of syllables in word		
	1	2	3
Unaccented	112	104	106
Accented	133	124	121

Table 4.26: Mean duration (ms) of test syllable onset for all left-headed keywords in Series A except *dog* and *part*: data points = 216; missing = 1.4%.

For left-headed keywords in Series C (Appendix H), mean test syllable onset durations are very similar to those for Series A: the effects of Word Length and Accent are significant, and the interaction between Word Length and Accent is not significant. The difference between monosyllables and disyllables approaches significance for unaccented keywords ($p = .063$) and is significant for accented keywords.

The results for onset duration offer some support for the word-final hypothesis, because the source of the significant effect of word length appears to be the difference between monosyllables and disyllables/trisyllables, in both accented and unaccented keywords: for example, /m/ is longer in *mace* than in *mason* or *masonry*. There is no evidence of a shortening effect between disyllables and trisyllables in unaccented keywords and only a very slight, non-significant effect in accented keywords. Paradoxically, although this result offers some support for a binary distinction between word-final and non-final stressed syllables, the syllable onset is not generally believed to manifest domain-final lengthening. An alternative interpretation, discussed further in Section 4.6.2, is that the results indicate that the magnitude of word-initial lengthening is greater in monosyllables than in polysyllables.

Left-headed keywords: nucleus duration

Table 4.27 shows mean test syllable nucleus duration for left-headed keywords in Series A. The effect of Word Length is highly significant by Subjects: $F(2,10.3) = 27.27$, $p < .001$ [by Items: $F(2,10.1) = 32.81$, $p < .001$]. The effect of Accent approaches significance by Subjects: $F(1,5) = 4.87$, $p = .078$ [by Items: $F(1,5.1) = 68.94$, $p < .001$]. The interaction between Word Length and Accent is highly significant by Subjects: $F(2,11.8) = 10.92$, $p < .005$ [by Items: $F(2,10.2) = 3.19$, $p = .084$]. Planned comparisons show that the durational difference between monosyllables and disyllables is only significant in accented keywords, but the difference between disyllables and trisyllables is significant in both accented and unaccented keywords. The difference between monosyllables and trisyllables is highly significant in both accent conditions.

	Number of syllables in word		
	1	2	3
Unaccented	79	76	67
Accented	89	79	72

Table 4.27: Mean test syllable nucleus duration (ms) for left-headed keywords in Series A except *dog* and *part*: data points = 216; missing = 1.4%.

The results for nucleus duration are very similar for Series C (Appendix H): for both accented and unaccented words, there is shortening between monosyllables and disyllables and further shortening between disyllables and trisyllables. The effect of Word Length is highly significant, the effect of Accent is significant and the interaction between Word Length and Accent is significant. The results of planned comparisons are very similar, although for Series C the difference between disyllables and trisyllables only approaches significance in the accented condition.

The results for nucleus duration provide support for the word-span hypothesis. The interaction indicates that the word-span compression effect on the nucleus is differently distributed in unaccented and accented keywords. For Series A, disyllabic shortening compared with the monosyllabic context is 3 ms (4%) for unaccented words and 10 ms (11%) for accented words; trisyllabic shortening compared with the disyllabic context is 9 ms (12%) for unaccented words and 7 ms (9%) for accented words. The fact that the differences between disyllables and trisyllables are comparable with those between monosyllables and disyllables favours the word-span hypotheses rather than the word-final hypothesis, as the latter predicts no difference in stressed syllable duration in disyllables vs trisyllables, or a much smaller difference than between monosyllables and disyllables. The results show that, for example, the vowel /eɪ/ in /meɪs/ is shorter in *mason* than in *mace* and shorter still in *masonry*. This constitutes good evidence for a word-span effect; however, unlike most of the foregoing results, the evidence suggests an effect in both accented and unaccented keywords.

If the word-span effect does only affect accented words, then the presence of a difference in the unaccented left-headed keywords for nucleus duration could be taken to indicate support for the word-rhyme-span hypothesis: this states that primary stressed syllable duration is inversely related to the word-rhyme length, and the results here suggest that the effect is localised on the syllable nucleus, as discussed further in Section 4.6.2.

Left-headed keywords: coda duration

Table 4.28 shows mean test syllable coda duration for left-headed keywords in Series A. The effect of Word Length is not significant by Subjects: $F(2,10.3) = 2.00$, $p = .184$ [by Items: $F(2,6.1) = 0.65$, $p = .557$]. The effect of Accent is significant by Subjects: $F(1,5) = 9.60$, $p < .05$ [by Items: $F(1,3) = 10.48$, $p < .05$]. There is a significant interaction between Word Length and Accent according to a By Subjects analysis: $F(2,10.5) = 5.31$, $p < .05$ [by Items: $F(2,6.2) = 5.66$, $p < .05$]. Planned comparisons reveal that the only significant difference in coda duration according to Word Length for either Accent condition is between monosyllables and trisyllables in the accented condition ($p < .05$).

	Number of syllables in word		
	1	2	3
Unaccented	104	105	105
Accented	131	126	116

Table 4.28: Mean test syllable coda duration (ms) for the left-headed keywords *fish*, *mace*, *sense* and *ten* in Series A: data points = 144; missing = 3%.

The results for coda duration are very similar for Series C (Appendix H): there is no evidence of a word-span effect in the unaccented context, but good evidence for a word-span effect in the accented context. The effect of Word Length approaches significance in Series C, as does the effect of Accent, and the interaction between Word Length and Accent is significant. The difference between monosyllables and trisyllables in the accented condition is significant in planned comparisons.

Because of the significant interaction and the apparent word-span effect which it indicates in accented keywords only, the data from the accented condition are analysed separately: for Series A, the effect of Word Length approaches significance by Subjects: $F(2,11.1) = 3.70$, $p = .059$ [by Items: $F(2,6.2) = 2.16$, $p = .194$]; For Series C, the effect of Word Length is significant by Subjects: $F(2,14.3) = 5.08$, $p < .05$ [by Items: $F(2,6.1) = 1.83$, $p = .239$]. These analyses support the conclusion that there is a word-span effect in accented keywords only. The word-final hypothesis is not supported as this predicts as shortening of the syllable rhyme—including the coda—in left-headed monosyllables vs disyllables: however, /s/ is no longer in *mace* than in *mason* in the unaccented condition.

Left-headed keywords: summary

In general, the results for left-headed keywords support the word-span hypothesis, but the nature of the effect apparently varies throughout the syllable. Coda duration suggests that the word-span effect only operates in the presence of pitch accent, as appeared to be the case for right-headed keywords, but nucleus duration indicates a word-span effect both in the presence and the absence of pitch accent. The results for onset duration apparently favour a word-final lengthening interpretation, both in the presence and absence of pitch accent, but previous research suggests that the syllable onset is less likely than the nucleus or coda to manifest final lengthening.

The processes underlying these results are discussed further in Section 4.6.2.

Unstressed syllable duration

As described in Section 4.3.3, the additional syllables in polysyllabic keywords are measured for Series C, but not for Series A: because there is no evidence of an utterance-span effect, the decision to use Series C was made because it has more consistent syntactic structure between sentences. The effect of word length on unstressed syllables may be examined by looking at the duration of syllable-2. For right-headed keywords, this is to the immediate left of the stressed syllable in utterances such as:

Series C.2 JOHN saw Alison **commend** it AGAIN.

Series C.3 JOHN saw Alison **recommend** it AGAIN.

For left-headed keywords, syllable-2 is to the immediate right of the stressed syllable in utterances such as:

Series C.2 I SAW the **mason** disinclined it ALL.

Series C.3 I SAW the **masonry** disinterred AGAIN.

As these examples illustrate, the phonetic environment of syllable-2 is not consistent between disyllables and trisyllables; this may cause microprosodic variations in duration, which are unlikely to be systematic with respect to the experimental hypotheses.

Table 4.29 shows syllable-2 duration for right-headed keywords[23]. The effect of Word Length is significant by Subjects: $F(1,5.1) = 10.98$, $p < .05$ [by Items: $F(1,5) = 3.09$, $p = .139$]. The effect of Accent is significant by Subjects: $F(1,5) = 12.10$, $p < .05$ [by Items: $F(1,5.1) = 36.69$, $p < .005$]. The interaction between Word Length and Accent approaches significance by Subjects: $F(1,5.1) = 5.30$, $p = .069$ [by Items: $F(1,5.1) = 12.47$, $p < .05$].

Unstressed syllables in accented keywords are 17ms (13%) shorter in trisyllables than in disyllables—thus, /kə/ is shorter in *recommend* than in *commend*—but the equivalent difference in unaccented keywords is only 4 ms (3%). There is clearly a

[23] As described in Section 4.3.3, the keyword *main* is excluded because the onset of /hju/ is not comparably articulated in the two contexts *develop humane* and *inhumane*.

	Number of syllables in word	
	2	3
Unaccented	132	128
Accented	151	134

Table 4.29: Mean syllable-2 duration (ms) for all right-headed keywords except *port* (missing data) and *main* in Series C: data points = 144; missing = 4.2%.

word-level effect: one interpretation, in line with the results for test syllable duration, is that there is a word-initial lengthening effect in accented and unaccented keywords, together with a word-span compression effect in accented keywords only. The small magnitude of apparent word-initial lengthening may reflect the fact that it has a syllable onset locus, and variation in the rest of the syllable may mask the effect in the onset.

Table 4.30 shows syllable-2 duration for left-headed keywords[24]. The effect of Word Length is significant by Subjects: $F(1,5.3) = 50.37$, $p < .005$ [by Items: $F(1,5) = 15.23$, $p < .05$]. The effect of Accent approaches significance by Subjects: $F(1,5.3) = 5.48$, $p = .063$ [by Items: $F(1,5.1) = 5.79$, $p = .061$]. The interaction between Word Length and Accent is not significant by Subjects: $F(1,5) = 3.48$, $p = .115$ [by Items: $F(1,5.5) = 0.16$, $p = .707$].

	Number of syllables in word	
	2	3
Unaccented	95	70
Accented	108	75

Table 4.30: Mean syllable-2 duration (ms) for all left-headed keywords in Series C except *dog* and *part*: data points = 144; missing = 5.6%.

Unstressed syllables in left-headed keywords show a large word-level effect: 25 ms (26%) shorter in trisyllables than disyllables in the unaccented condition, and 33 ms (31 %) shorter in trisyllables than disyllables in the accented condition. There is no interaction with pitch accent, unlike the results for test syllable duration, most of which indicate a word-span effect only in accented keywords. The comparison for unstressed syllables confounds the effects of word-length and position-in-word: for example, /ən/ is word-final in the disyllable *mason* but not word-final in the trisyllable *masonry*. Thus, these results could be interpreted as evidence for word-span or

[24]As noted in Section 4.2.1, the syllable following the keyword in Series C for *dog* and *part* is likely to contain a full vowel and thus may have durational influence on the preceding syllable. Although this seems particularly unlikely where the preceding syllable is unstressed, these keywords are excluded from the following analysis.

word-rhyme-span compression or word-final lengthening, although the latter is not supported by most of the results for stressed syllable duration, and word-span compression of stressed syllables appears only to be significant in accented words.

4.6.2 Discussion

Results for whole test syllable duration, illustrated in Figures 4.2 and 4.3, indicate two main effects:

- word-initial lengthening;
- word-span compression, much greater in pitch-accented words.

The proportion of shortening between monosyllables and disyllables and between disyllables and trisyllables is illustrated in Figure 4.4 for subsyllabic constituents in Series A; for the purposes of comparison, results are shown for the same data-sets for all constituents.

The experimental hypotheses are here considered with regard to the patterns of durational variation of subsyllabic constituents.

Word-span hypothesis

Figure 4.4 shows that all subsyllabic constituents in accented keywords manifest some evidence of word-span compression except for nuclei in the right-headed monosyllable vs disyllable comparison. In contrast, constituents in unaccented keywords do not show strong evidence of word-span compression except for the onset in the monosyllable vs disyllable comparison for right-headed and left-headed keywords and the nucleus in both comparisons for left-headed keywords. These apparently anomalous results for unaccented keywords are discussed below.

Regarding the word-span effect, the comparison between disyllables and trisyllables is the more transparent because the monosyllable vs disyllable comparison may be influenced by word-edge effects. Word-span compression is quite clear in accented keywords between disyllables and trisyllables: thus /mɛnd/ is shorter in *recommend* than *commend* and /meɪs/ is shorter in *masonry* than *mason*. The word-span compression interpretation is supported because the effect is evident in left-headed and right-headed keywords, and in all subsyllabic constituents, rather than being confined to a specific locus.

Most of the evidence suggests that word-span compression is confined to accented words. In unaccented keywords, according to this interpretation, primary stressed syllable duration is constant, all other factors being equal. When the word is accented, all

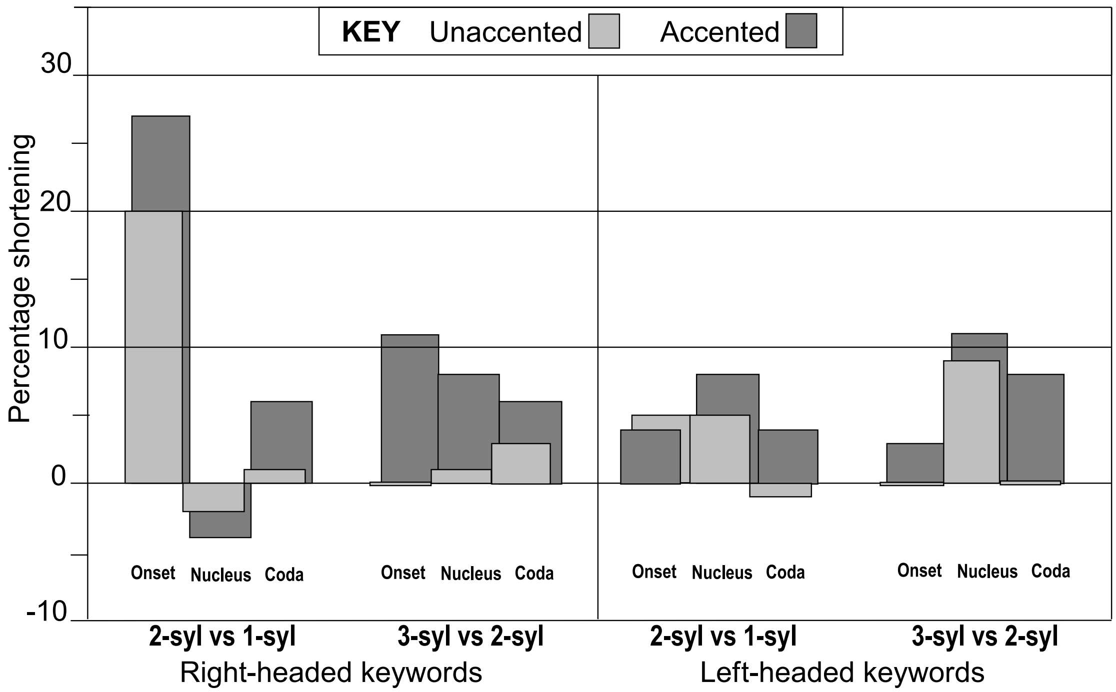

Figure 4.4: Proportional word-level shortening in monosyllabic, disyllabic and trisyllabic keywords in Series A. Proportional shortening is the difference in duration expressed as a percentage of syllable duration in the longer context. All right-headed keywords are included except *port*: data points = 252; missing = 3.8%. Left-headed keywords are *fish*, *mace*, *mend* and *send*: data points = 144; missing = 2.0%.

parts of the primary stressed syllable are lengthened, by a large amount in the monosyllable, less in the disyllable and less still in the trisyllable. This produces an apparent word-span compression effect in the comparison of constituent duration between accented keywords. Given its origin, however, it is more straightforward to describe it as an effect of word length on the magnitude of accentual lengthening. What may be called the "polysyllabic accent effect" simply encapsulates the fact that:

Accentual lengthening of the primary stressed syllable is greater in words of fewer syllables.

Because accentual lengthening affects all parts of the syllable, the locus of the polysyllabic accent effect is the whole syllable. The relative shortening effect on the different constituents should be a function of the amount of accentual lengthening they receive. If this is the case, the results shown in Figure 4.4 here suggest that the distribution of accentual lengthening is different in right-headed and left-headed words: this is examined further in Section 4.9. How previous evidence for polysyllabic shortening may be interpreted in the light of these results is discussed in Chapter 5.

Word-rhyme-span hypothesis

The word-rhyme-span hypothesis predicts an effect of word length in left-headed words, but not in right-headed words. This is not supported by onset and coda duration, where the effect of word length is at least as great in right-headed keywords. The pattern of results for nucleus duration is somewhat different, however. In both comparisons for unaccented keywords, monosyllable vs disyllable and disyllable vs trisyllable , the nucleus shows evidence of a word-span compression effect in left-headed keywords but not in right-headed keywords. Furthermore, in the accented condition, the nucleus shows the greatest effect of word-length for left-headed keywords, whereas for right-headed keywords the onset shows the greatest effect of word-length. These results suggest a process, which may be called "word-rhyme compression", independent of pitch accent and with a syllable nucleus locus.

Results for unstressed syllable duration indicate that the locus may extend to subsequent syllables within the word-rhyme. For left-headed keywords, the word-final syllable in the disyllable is shorter when word-medial in the trisyllable: for example, /ən/ is shorter in *masonry* than in *mason*. This effect appears to be independent of accent, although the proportional shortening is slightly greater in accented keywords.

The polysyllabic accent effect could explain the unstressed syllable result in left-headed accented keywords. The fact that there is an almost comparable effect in unaccented left-headed keywords requires a further mechanism which could be word-rhyme compression. Subsyllabic durations are not available for unstressed syllables, but it may be hypothesised that the locus of the effect may be the syllable nucleus, as for stressed syllables. Thus, a full statement of word-rhyme compression is:

There is an inverse relationship between the number of syllables in the word-rhyme and the duration of the nuclei of those syllables, both stressed and unstressed.

In accented keywords, of course, the unstressed syllable is likely to be influenced also by the polysyllabic accent effect, thus it appears that the two effects combine sub-additively, as the shortening in the disyllable vs trisyllable comparison is little greater than in the unaccented condition.

Two alternative interpretations of the effect observed in left-headed keywords are available: word-final lengthening, as discussed below, and syllable ratio equalisation, proposed by Turk & Shattuck-Hufnagel (2000). Syllable ratio equalisation is discussed with regard to Experiment 1 in Chapter 3, where it is suggested to be supererogatory; it is reassessed in Chapter 5 with regard to the results of Experiment 2.

Word-final lengthening

As shown in Figure 4.4, the results for left-headed accented keywords indicate that all parts of the syllable are affected by word-length, with the syllable nucleus showing the greatest effect. This supports the polysyllabic accent effect discussed above: a word-final process would be expected to affect the nucleus and coda, and probably the latter more than the former. The results for unstressed syllables in left-headed keywords could be interpreted as support for word-final lengthening: in both accented and unaccented contexts, the unstressed word-final syllable in the disyllable is longer then when placed word-medially in the trisyllable: thus, for example, /ən/ is longer in *mason* than in *masonry*.

The results for stressed syllables in unaccented keywords seem, however, to suggest a domain-span compression interpretation, for a number of reasons. The word-length effect in unaccented left-headed keywords is primarily manifest on the nucleus, and the coda is unaffected, contrary to previous observations of domain-edge effects. Furthermore, the magnitude of the word-length effect on the unaccented nucleus is actually greater in the disyllable vs trisyllable comparison than in the monosyllable vs disyllable comparison: there is greater shortening of /eɪ/ in *mason* vs *masonry* than in *mace* vs *mason*. A word-final effect would be expected to cause a large difference in test syllable duration in the latter comparison, but little or no difference in the former.

Word-rhyme compression is not, therefore, simply an alternative statement of a word-final effect. The locus of domain-final lengthening at higher-level boundaries is known to include the syllable rhyme, and tends to be progressive; that is, greater closer to the boundary. The effect observed here has a syllable nucleus locus and is not progressive. For theoretical consistency, it does not appear appropriate to describe it as a word-final effect.

Because of the lack of support for the utterance-span hypotheses, and the interpretation of the word-span effect as resulting from variations in accentual lengthening—the polysyllabic accent effect—the word-rhyme compression is the only domain-span effect for which Experiment 2 provides evidence. For reasons of parsimony, an interpretation which posits a demonstrated process in a new domain is to be preferred to an interpretation which requires an new type of process. It is clear that final lengthening exists at the intonational-phrase-level and/or the utterance-level; to extend this process to the word-level would be theoretically more satisfying than suggesting a single domain-span process, that is, word-rhyme compression.

Results presented in Section 4.8.1 indicate that utterance-final lengthening is localised on certain subconstituents of the word-rhyme; this suggests parallels with the effects observed here in left-headed keywords which leave open the possibility of a word-final characterisation of the process. This is explored further in Chapter 5.

Finally, there does not appear in Experiment 2 to be any evidence of phrase-final lengthening below the level of the intonational phrase. Lower-level phrase boundaries were not controlled, and as seen below, there is some evidence of phrase-initial lengthening, suggesting some right-headed keywords may be preceded by boundaries such as those between phonological phrases. For left-headed keywords, however, the lack of support for the word-final hypothesis also indicates an absence of evidence of phrase-final lengthening in utterance-medial context, a conclusion supported by the comparisons between utterance-medial and near-utterance-final words in Section 4.5. Possible levels of final lengthening are discussed further in Chapter 5.

Word-initial lengthening

The results for word-initial lengthening are fairly clear: the syllable onset, including aspiration duration in onset voiceless stops, has greater duration word-initially than word-medially. The apparently greater lengthening in accented words is probably due to the polysyllabic accent effect, as described above. This interpretation is supported by the evidence relating to phrase-initial lengthening, which does not indicate an interaction with accent.

The results for unstressed syllables provide some support for the word-initial lengthening hypothesis: for example, /kə/ is longer in *commend* than in *recommend*, although the effect is very small in the unaccented condition. Thus, if word-initial lengthening does not interact with accent, then the results for unstressed syllables must also be interpreted as a combination of word-initial lengthening, with a syllable onset locus, and the polysyllabic accent effect, affecting the whole syllable.

The finding, observed in the results of Oller (1973) and Cooper (1991), of differential interactions with pitch accent (or lexical stress) in the closure and aspiration duration of syllable onset voiceless stops is not supported here. Such effect cannot be ruled out by this evidence, however, because in Experiment 2 the polysyllabic accent effect may prove a confounding factor: for example, the duration of aspiration of /p/ is compared in *pose* and *suppose*, where possible word-medial lengthening of aspiration in accented syllables may be masked by shortening due to word length.

The results for phrase-initial lengthening are only suggestive, given that boundaries below the intonational phrase have not been controlled: it appears that, for example, /m/ in *mend* is slightly longer in utterance-medial context (for example: *Johnny saw Jessica mend it again*), where it may be preceded by a phrase boundary, than near the utterance-edge (for example: *Will you mend it again for me please*), where a preceding phrase boundary is less likely. Experiment 2 thus supports Fougeron & Keating's (1997) finding of three levels of syllable onset durational variation below the intonational phrase: word-medial; word-initial; phrase-initial.

There are two outstanding anomalies in the results illustrated in Figure 4.4, both of which may relate to word-initial lengthening. Firstly, the syllable onset in unaccented left-headed keywords shows an effect of word-length in the monosyllable vs disyllable comparison; thus, /m/ is shorter in *mason* than in *mace*: this may be interpreted as indicating that word-initial lengthening has lesser magnitude in polysyllables than in monosyllables. This is an analogue of the polysyllabic accent effect: the degree of lengthening is inversely proportional to word length. For word-initial lengthening, however, this seems to be a binary distinction between monosyllables and polysyllables, as there is no evidence of shortening of /m/ in *masonry* compared with *mason*, except in the accented condition, for which the polysyllabic accent effect already provides an explanation.

The second outstanding result is the slight lengthening of the syllable nucleus in the monosyllable vs disyllable comparison for right-headed accented and unaccented keywords. The durational difference between, for example, /ɛ/ in *mend* and *commend* is not significant, but does run contrary to the trend of the other results. If the interpretations of Experiment 2 are correct, then the syllable nucleus in this comparison ought to show no durational variation in unaccented keywords and show word-span compression in accented keywords; in fact, the nucleus is slightly longer in the disyllable. This could be interpreted as a compensatory effect: as can be seen in Figure 4.4, word-initial lengthening is very large in comparison with the other durational effects, including accentual lengthening. The magnitude of word-initial lengthening may induce an opposite adjustment in nucleus duration in the direction of maintaining a constant syllable duration; this could also underlie the relatively small word-initial lengthening effect apparent in the durations for whole unstressed syllables. This interpretation of the small observed effect is speculative, but evidence presented in Section 4.5 suggests similar effects may be associated with phrase-initial syllables and utterance-antepenultimate syllables. Possible compensatory effects are discussed further in Chapter 5.

4.7 Utterance-initial results

The utterance-initial hypothesis predicts that word-initial syllable onsets are shorter utterance-initially than utterance-medially. Only right-headed keywords are used in the analyses presented in Section 4.7.1—left-headed keywords are not placed utterance-initially in the experimental materials—and there are two different groups of right-headed keywords analysed, as described below. The utterance-initial hypothesis is supported for those keywords in which the test syllable onset is fully measurable utterance-initially.

The experimental comparison for the utterance-initial hypothesis is between Series B (utterance-medial) and Series D (utterance-initial). For example, the utterance-initial hypothesis predicts that /m/ is shorter in D.1 than B.1, thus:

Series B.1 JOHN saw Jessica **mend** it AGAIN.

Series D.1 **Mend** it AGAIN for me please.

Results are reported first for test syllable onset duration in Series B and Series D: although fully-measurable onsets show utterance-initial shortening, partial-onset (aspiration/frication) duration shows slight utterance-initial lengthening in unaccented context. Nucleus and coda durations are also reported, indicating some evidence of compensatory durational effects, tending to counter the utterance-initial effects on the syllable onset.

The interpretation of observed utterance-initial effects is discussed in Section 4.7.2, as are explanations for the apparent compensatory effects.

4.7.1 Results and analysis

Keyword grouping

Utterance-initial shortening has only been observed in nasal onset consonants; thus, keywords with different test syllable onset consonants are treated separately in the first analysis. In the right-headed monosyllables, as discussed in Section 4.3.3, only *main*, *mend* and *send* have onsets which are entirely measurable utterance-initially. The monosyllable *juice* begins with the voiced affricate /dʒ/ for which the duration of closure cannot be measured utterance-initially; thus, the duration of frication is compared utterance-initially and utterance-medially. The keywords *port*, *compose*, *dispose* and *suppose* have the voiceless stop /p/ as their test syllable onset, for which closure duration cannot be measured in utterance-initial position; thus, aspiration duration is examined to test the utterance-initial hypothesis.

Table 4.31 shows mean onset duration for *mend* and *main*, and separately for *send*; mean onset frication duration for *juice*; mean test syllable onset aspiration duration (ms) for *compose*, *dispose*, *port* and *suppose*. The keywords with fully-measurable onsets show the same trend: *send*, and *main* and *mend* together, have greater onset duration utterance-medially than utterance-initially. In contrast, *juice* has onset frication duration longer utterance-initially than utterance-medially, although the difference is small compared with the utterance-initial shortening shown by *main*, *mend* and *send*. The keywords *compose*, *dispose*, *port* and *suppose* show no effect of utterance position on aspiration duration in accented keywords, but aspiration duration is slightly longer utterance-initially in unaccented keywords.

	Utterance position	
	Initial	Medial
/m/ duration: *main* and *mend*		
Unaccented	43	69
Accented	66	103
/s/ duration: *send*		
Unaccented	97	115
Accented	124	143
/dʒ/ frication duration: *juice*		
Unaccented	55	50
Accented	62	58
/p/ aspiration duration: *compose, dispose, port* and *suppose*		
Unaccented	46	41
Accented	54	55

Table 4.31: Mean test syllable onset duration (ms) for right-headed keywords in Series B and Series D: data points for each keyword = 24; missing = 0 for *mend* and *main*; missing = 1 for keyword *send*; missing = 0 for *juice*; missing = 4 for *compose*, *dispose*, *port* and *suppose*.

Thus, *main*, *mend* and *send* are analysed together, as all three show utterance-initial shortening in the onset. For the remaining keywords, *juice* may reasonably be grouped with the keywords *compose*, *dispose*, *port* and *suppose* as all have partially-measurable onsets, with closure duration unmeasurable utterance-initially; furthermore, these keywords do not show utterance-initial shortening, but rather manifest slight utterance-initial lengthening, at least in the unaccented condition[25].

Test syllable onset duration

Figure 4.5 shows mean test syllable onset duration for *main*, *mend* and *send*. The effect of Utterance Position is highly significant by Subjects: $F(1,5.1) = 150.61$, $p < .001$ [by Items: $F(1,2) = 40.38$, $p < .05$]. The effect of Accent is significant by Subjects: $F(1,5) = 13.88$, $p < .05$ [by Items: $F(1,2) = 2228.57$, $p < .001$]. There is no significant interaction between Utterance Position and Accent by Subjects: $F(1,5) = 0.91$, $p = .384$ [by Items: $F(1,2) = 4.19$, $p = .177$]. For unaccented monosyllables, the onset is 25 ms (30%) shorter utterance-initially; for accented monosyllables, the onset is 31 ms (27 %)

[25] Because *port* has an inconsistently measurable coda, it is excluded from further analysis to provide a consistent keyword set between all subsyllabic constituents.

shorter utterance-initially.

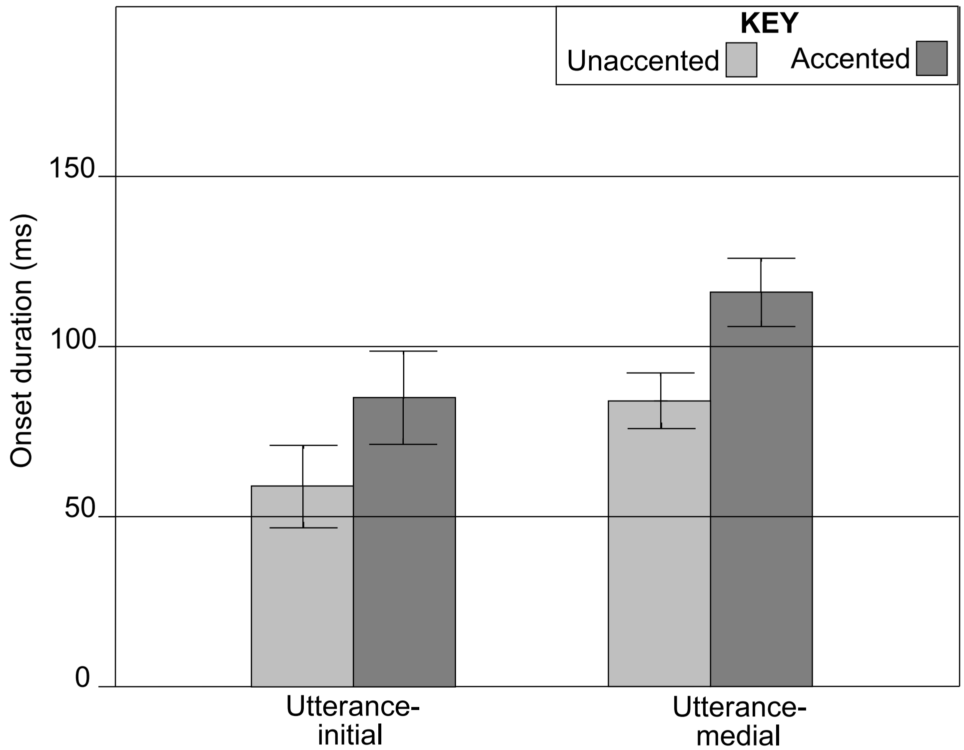

Figure 4.5: Mean test syllable onset duration (ms) for monosyllables *main*, *mend* and *send* in Series B and Series D: data points = 72; missing = 1.

Figure 4.6 shows mean duration of test syllable onset aspiration/frication for *compose*, *dispose*, *suppose* and *juice*. The effect of Utterance Position is not significant by Subjects: $F(1,5) = 0.67$, $p = .449$ [by Items: $F(1,3) = 1.69$, p .283]. The effect of Accent is significant by Subjects: $F(1,5) = 9.36$, $p < .05$ [by Items: $F(1,3.2) = 105.89$, $p = .002$]. The interaction between Utterance Position and Accent is significant by Subjects: $F(1,5.1) = 8.03$, $p < .05$ [by Items: $F(1,3.1) = 5.09$, $p = .108$]. The nature of this interaction is indicated by two-tailed t-tests: the effect of Utterance Position is significant for unaccented keywords, where aspiration/frication duration is 6 ms (13%) shorter utterance-medially, but not for accented keywords.

The results for *main*, *mend* and *send* clearly support the utterance-initial hypothesis, for both accented and unaccented keywords: onset duration is shorter utterance-initially. The results for partial-onset duration for *compose*, *dispose*, *suppose* and *juice* do not support the utterance-initial hypothesis: unaccented keywords manifest a small utterance-initial lengthening effect and accented keywords show no effect of utterance-position. Because of these differences, nucleus and coda durations are analysed according to the keyword groupings used for onset and partial-onset duration.

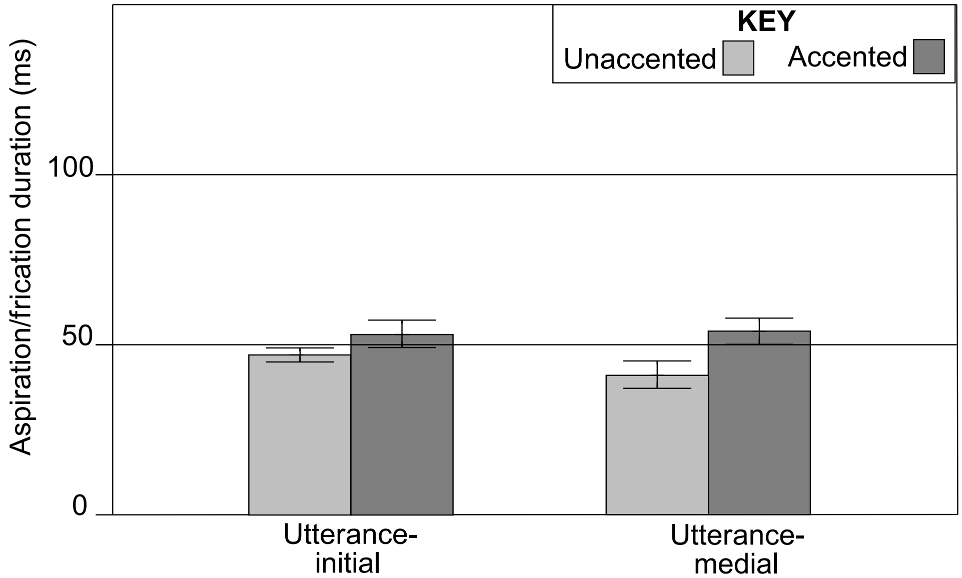

Figure 4.6: Mean test syllable onset aspiration/frication duration (ms) for monosyllables from keywords *compose*, *dispose*, *suppose* and *juice* in Series B and Series D: data points = 96; missing = 4.

Test syllable nucleus duration

Table 4.32 shows mean nucleus duration for *main*, *mend* and *send*, which demonstrate a large utterance-initial shortening effect in their onsets. For nucleus duration there is no effect of Utterance Position by Subjects: F(1,5) = 0.03, p = .875 [by Items: F(1,2) = 0.02, p = .910]. There is a highly significant effect of Accent by Subjects: F(1,5) = 20.43, $p < .01$ [by Items: F(1,2) = 4.31, p = .173]. There is no significant interaction between Utterance Position and Accent by Subjects: F(1,5.1) = 0.51, p = .505 [by Items: F(1,2) = 0.26, p = .660]. Thus, for these keywords, the effect of utterance-initial position does not extend beyond the test syllable onset.

	Utterance position	
	Initial	Medial
Unaccented	82	82
Accented	99	97

Table 4.32: Mean test syllable nucleus duration (ms) for monosyllables *main*, *mend* and *send* in Series B and Series D: data points = 72; missing = 1.

Table 4.33 shows the mean test syllable nucleus duration for the keywords *compose*, *dispose*, *suppose* and *juice*. The effect of Utterance Position is not significant by Subjects: F(1,5) = 3.72, p = .111 [by Items: F(1,3) = 5.31, p = .104]. The effect of Accent is highly significant by Subjects: F(1,5) = 18.15, $p < .01$ [by Items: F(1,3) = 39.33, $p < .01$]. The interaction between Utterance Position and Accent is significant by Subjects: $F(1,5.4) = 23.13$, $p < .005$ [by Items: $F(1,3) = 3.34$, $p = .164$]. The nature of this interaction is indicated by two-tailed t-tests: in the unaccented context, the effect of Utterance Position is highly significant, with 13 ms (12%) shortening utterance-initially; in accented context, the effect of Utterance Position is not significant.

	Utterance position	
	Initial	Medial
Unaccented	98	111
Accented	133	134

Table 4.33: Mean test syllable nucleus duration (ms) for monosyllables from keywords *compose*, *dispose*, *port* and *suppose* in Series B and Series D: data points = 96; missing = 4.

The utterance-initial shortening of nucleus duration for unaccented monosyllables contrasts with the utterance-initial lengthening of onset aspiration/frication duration shown in Figure 4.6 for the same unaccented monosyllables. The interpretation of this compensatory relationship between aspiration duration and vowel duration is discussed in Section 4.7.2.

Test syllable coda duration

Table 4.34 shows mean test syllable coda durations for *main*, *mend* and *send*. The effect of Utterance Position is highly significant by Subjects: $F(1,5.1) = 20.89$, $p < .01$ [by Items: $F(1,2) = 3.18$, $p = .216$]. The effect of Accent is also highly significant by Subjects: $F(1,5) = 20.57$, $p < .01$ [by Items: $F(1,2) = 13.67$, $p = .066$]. The interaction between Utterance Position and Accent is not significant by Subjects: $F(1,5.2) = 1.91$, $p = .224$ [by Items: $F(1,2) = 0.54$, $p = .541$]. Thus, coda duration is shorter utterance-medially, by 10 ms (13%) for unaccented keywords and by 7ms (7%) for accented keywords,

	Utterance position	
	Initial	Medial
Unaccented	75	65
Accented	94	87

Table 4.34: Mean test syllable coda duration (ms) for monosyllables *main*, *mend* and *send* in Series B and Series D: data points = 72; missing = 1.

The finding of an utterance-initial lengthening effect on coda duration for keywords *main*, *mend* and *send* is surprising, particularly as the syllable nucleus for these keywords, which is closer to the start of the utterance, shows no sign of any durational effect; however, the utterance-initial codas are compared with utterance-medial codas which show evidence of a possible compensatory shortening effect, as reported in Section 4.5.1. Such shortening would be sufficient explanation of the durational difference, without proposing a further utterance-initial effect.

Table 4.35 shows the mean test syllable coda duration for *compose*, *dispose*, *suppose* and *juice*. There is no significant effect of Utterance Position by Subjects: $F(1,51) = 0.14$, $p = .723$ [by Items: $F(1,3.2) = 1.90$, $p = .258$]. The effect of Accent is significant by Subjects: $F(1,5.1) = 8.78$, $p < .05$ [by Items: $F(1,3) = 3.55$, $p = .156$]. The interaction between Utterance Position and Accent approaches significance by Subjects: $F(1,5.3) = 5.31$, $p = .067$ [by Items: $F(1,3.1) = 0.70$, $p = .462$]. The tendency towards an interaction reflects the slight utterance-initial lengthening of the coda in unaccented keywords and the even smaller utterance-initial shortening of the coda in accented keywords.

	Utterance position	
	Initial	Medial
Unaccented	76	72
Accented	81	83

Table 4.35: Mean test syllable coda duration (ms) for monosyllables from keywords *compose*, *dispose*, *suppose* and *juice* in Series B and Series D: data points = 96; missing = 4.

4.7.2 Discussion

The comparison between utterance-initial and utterance-medial syllables suggests two trends: firstly, utterance-initial durational variation in the stressed syllable onset, either shortening or lengthening according to the composition of the onset; secondly, compensatory shortening, whereby lengthening in the onset is balanced by the opposite effect in the nucleus.

Utterance-initial durational processes

The utterance-initial shortening found by Fougeron & Keating (1997) for nasals in unstressed syllables is found here in stressed syllables and a similar effect is found for the fricative /s/ in stressed syllable onsets. Stressed syllable onsets containing stops or affricates do not demonstrate utterance-initial shortening, however: the duration of aspiration or frication which follows the stop release is longer utterance-initially in unaccented words, and shows no effect of position in accented words.

Evidence presented in Section 4.6.1 shows that syllable onsets are longer word-initially than word-medially, and also indicates further lengthening (phonological-) phrase-initially. These results agree with Fougeron & Keating's finding of hierarchical lengthening of syllable onsets, which they find to be further lengthened intonational-phrase-initially. This raises the question of why utterance-initial boundaries, at the highest level of putative prosodic hierarchies, do not show consistent evidence for further domain-initial lengthening.

A possible explanation for the absence of an utterance-initial lengthening effect is that there is no reason for utterance-initial boundaries to be signalled suprasegmentally, at least not when the utterance is spoken in isolation. If hierarchical initial-lengthening effects are cues to juncture, they are superfluous utterance-initially: the listener will know that the first syllable spoken is utterance-initial because it is preceded by silence. This interpretation predicts that utterance-initial onset duration should be comparable to that found word-medially, as the lowest level of the hierarchy of onset lengthening. The results of Fougeron & Keating (1997) for one speaker show exactly this pattern; for the second speaker, onset duration is slightly longer utterance-initially than word-medially, but shorter than word-initially[26].

This interpretation of utterance-initial shortening may be tested by examining test syllable onset duration in Keyword Series E, where the onset is utterance-initial in the monosyllable and word-medial in the disyllable and trisyllable, for example:

E.1 **Mend** it AGAIN for me please.

E.2 **Commend** it AGAIN for me please.

E.3 **Recommend** it AGAIN for me please.

Table 4.36 shows mean test syllable onset duration for *main*, *mend* and *send* in Series E. The effect of Word Length is not significant by Subjects: $F(2,10.1) = 0.07$, $p = .931$ [by Items: $F(2,4) = 0.10$, $p = .904$]. The effect of Accent is significant by Subjects and by Items. The interaction between Word Length and Accent approaches significance by Subjects: $F(2,10.2) = 3.64$, $p = .064$ [by Items: $F(2,4) = 9.75$, $p < .05$].

The usual word-initial lengthening, as described in Section 4.6.1 is not evident, as the small word-length effect in the accented condition may be attributed to the polysyllabic accent effect. The onset is actually slightly shorter word-initially than word-medially in the unaccented condition. The almost significant interaction may be attributable to one or both of these trends, although none of the differences due to Word Length are significant in t-tests for either accented or unaccented keywords.

[26] As mentioned in Section 4.1.3, the third speaker did not show an utterance-initial shortening effect: onset duration utterance-initially was comparable to that found intonational-phrase-initially.

	Word length		
	Monosyllable	Disyllable	Trisyllable
Unaccented	59	68	67
Accented	85	80	77

Table 4.36: Mean test onset syllable duration (ms) for the keywords *main*, *mend* and *send* in Series E: data points = 108; missing = 3.

The pattern of onset duration according to word length may be contrasted with that found utterance-medially; in Series C, for example:

C.1 JOHN saw Jessica **mend** it AGAIN.

C.2 JOHN saw Alison **commend** it AGAIN.

C.3 JOHN saw Alison **recommend** it AGAIN.

Table 4.37 shows mean test syllable onset duration for *main*, *mend* and *send* in Series C. Word-initial lengthening is evident, as is the additional word-span compression in accented keywords attributable to the polysyllabic accent effect. In a comparison of Series C and Series E, the interaction between Utterance Position and Word Length is highly significant by Subjects: $F(2,10.1) = 20.98$, $p < .001$ [by Items: $F(2,4) = 15.59$, $p < .05$]. This supports the conclusion that the word-initial lengthening evident in the utterance-medial Series C is absent in the utterance-initial Series E.

	Word length		
	Monosyllable	Disyllable	Trisyllable
Unaccented	84	69	70
Accented	116	90	76

Table 4.37: Mean test onset syllable duration (ms) for the right-headed keywords *main*, *mend* and *send* in Series C: data points = 108; missing = 1.

Therefore, the hypothesis that utterance-initial shortening represents an absence of any domain-initial hierarchical lengthening is quite well supported by the data for keywords *main*, *mend* and *send*, which have fully-measurable test syllable onsets utterance-initially.

The results for the other keywords, where stop closure cannot be measured at the start of the utterance, show a somewhat different pattern. There is no utterance-initial shortening of aspiration or frication following the stop release in accented keywords, and in unaccented keywords the partial onset—for example, the aspiration of /p/ in /pəʊz/—is slightly longer utterance-initially than utterance-medially. This could be an articulatory effect: aspiration or frication following a stop release may be more forceful and thus longer where this is initiating phonation than in the middle of the speech stream; the difference may be masked in accented keywords because of their generally greater gestural magnitude and duration. Utterance-initial processes are discussed further in Chapter 5.

Compensatory effects

A compensatory effect, similar to those discussed in Section 4.7.1, is suggested by one of the results in Section 4.7. In the test syllables for the keywords *compose*, *dispose*, *suppose* and *juice* in the unaccented condition, the partial-onset—aspiration or frication following the stop release—shows a small utterance-initial lengthening effect, and the nucleus shows a slightly larger utterance-initial shortening effect; the coda is also slightly longer utterance-initially, although the effect is not significant. The sum of these effects produces a whole syllable duration which is not significantly different utterance-initially and utterance-medially. In the accented condition, there is no significant effect of utterance position in either the whole syllable or in the subsyllabic constituents.

Compensatory shortening of the nucleus following lengthening of the partial-onset is predictable if the articulatory gestures involved are considered. The vowel durations reported in Experiment 2 are taken from the onset of voicing, because, as discussed in Section 4.3, aspiration duration shows a similar pattern of word-initial lengthening to consonant closure duration: thus, onset duration for voiceless stops includes closure and aspiration duration. As shown in Figure 4.7, however, the vocalic gesture begins in the oral cavity with the release of the stop, at which point the articulator begin moving towards its steady state position for the vowel. In voiceless stops, the vocalic gesture does not begin in the larynx until the onset of voicing. Thus, if the onset of voicing is delayed, aspiration duration increases and vowel duration decreases. Conversely, an earlier onset of voicing will mean shorter aspiration duration and longer vowel duration.

It is, of course, possible for aspiration duration and vowel duration to increase concurrently; such an effect would be expected, for example, in a pitch-accented syllable. The inverse relationship between vowel duration and aspiration duration observed here would arise when there is a durational process that affects only the onset, such as initial lengthening, or only the nucleus, such as word-rhyme compression or phrase-final lengthening (the latter affecting also the syllable coda).

Thus, the evidence of a compensatory relationship within utterance-initial stressed syllables appears to be explicable as an effect of articulatory overlap in the case of *compose*, *dispose* and *suppose*. There is, however, some direct evidence of compen-

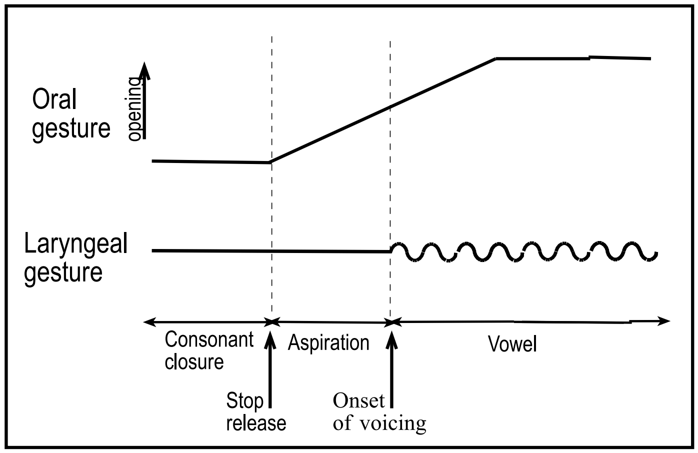

Figure 4.7: Schematic diagram of the relationship between aspiration duration and vowel duration.

satory relationships at three levels of prosodic structure: word-initial, phrase-initial and utterance-final. The interpretation of these effects is discussed further in Chapter 5.

4.8 Utterance-final results

The principal test of utterance-final lengthening is the comparison between the two left-headed keyword series in which the length of the keyword covaries with the length of the utterance: Series C, where the keyword is utterance-medial, and Series E, where the keyword is utterance-final. This is the utterance-final within-word analysis: results for the primary stressed syllable and for additional syllables in polysyllabic keywords support the utterance-final hypothesis that the locus of lengthening is the word-rhyme, although not all subconstituents of the word-rhyme are lengthened. The utterance-final locus distribution hypothesis, that lengthening is progressive within this locus, is supported. The utterance-final accent hypothesis, that the magnitude of lengthening is not affected by the presence of a pitch accent on the utterance-final word, is also supported, although there is non-significant attenuation of lengthening in accented words. Results are reported first for whole test syllable duration, then separately for the onset, nucleus and coda of the test syllable, and then for unstressed syllables in disyllables and trisyllables.

Series D contains test syllables in monosyllabic keywords, separated from the utterance-final boundary by zero, one or two syllables. The utterance-final cross-word analysis compares Series D with Series B, which contains utterance-medial monosyllabic keywords, to determine the pattern of utterance-final durational variation where word boundaries intervene between the test syllable and the utterance edge. These results are only suggestive, however, because there are potentially confounding influences in the experimental materials; in particular, full-vowel syllables immediately follow the test syllables when in utterance-penultimate position[27].

The experimental support for the three utterance-final hypotheses is discussed in Section 4.8.2.

4.8.1 Results and analysis

Within-word results: full stressed syllable duration

Table 4.38 shows example sentences from Series C and Series E; the experimentally-important difference between the two series being the proximity of the utterance-final boundary in Series E. Differences in test syllable duration should be directly comparable between the two series because, as seen in Section 4.5.1, there is no direct effect of utterance length on test syllable duration.

Series C: utterance-medial keywords	
$\|\ldots\sharp\sigma\sharp\sigma\sigma\ldots\|$	I SAW the **mace** unreclaimed AGAIN.
$\|\ldots\sharp\sigma\sigma\sharp\sigma\sigma\ldots\|$	I SAW the **mason** disinclined it ALL.
$\|\ldots\sharp\sigma\sigma\sigma\sharp\sigma\sigma\ldots\|$	I SAW the **masonry** disinterred AGAIN.
Series E: utterance-final keywords	
$\|\ldots\sharp\sigma\|$	Albert THREW the **mace**.
$\|\ldots\sharp\sigma\sigma\|$	Albert THREW the **mason**.
$\|\ldots\sharp\sigma\sigma\sigma\|$	Albert THREW the **masonry**.

Table 4.38: Example experimental sentences for Series C and Series E.

Figure 4.8 shows mean duration of utterance-final and utterance-medial test syllables in monosyllables, disyllables and trisyllables. The effect of Utterance Position is highly significant by Subjects: $F(1,5) = 75.10$, $p < .001$ [by Items: $F(1,3) = 28.34$, $p < .05$. This indicates that test syllables are longer utterance-finally than utterance-medially.

The effect of Word Length is highly significant by Subjects: $F(2,10.1) = 153.84$, $p < .001$ [by Items: $F(2,6) = 91.80$, $p < .001$]; this is attributable to the polysyllabic accent effect and the word-rhyme-span compression effect, as discussed in Section 4.6.1. The effect of Accent is significant by Subjects: $F(1,5) = 14.83$, $p < .05$ [by Items: $F(1,3) = 98.84$, $p < .005$].

The interaction between Utterance Position and Word Length is highly significant by Subjects: $F(2,10,5) = 230.70$, $p < .001$ [by Items: $F(2,6) = 13.34$, $p < .01$]. This indi-

[27] In the reporting of these results, the phrase “utterance-final” generally refers to the experimental condition, wherein the measured syllable itself could be “absolute-final”, “utterance-penultimate” or “utterance-antepenultimate”.

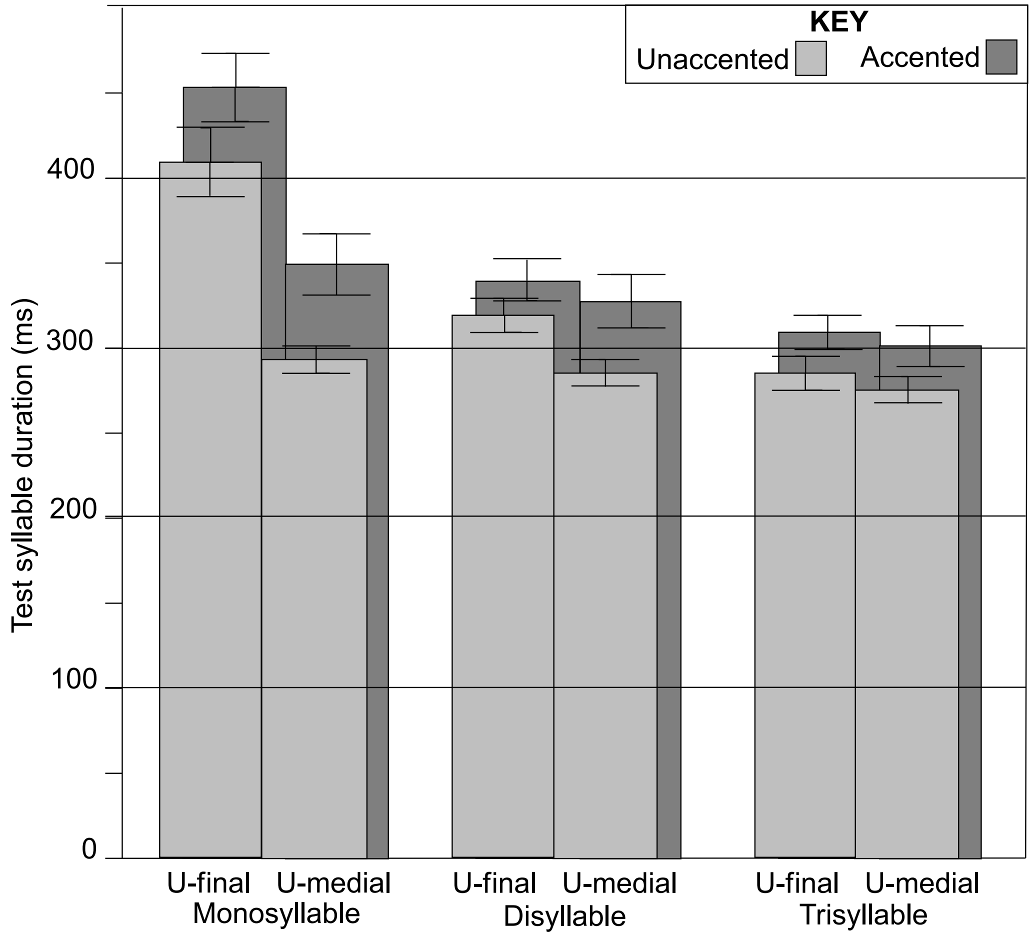

Figure 4.8: Mean test syllable duration (ms) for the left-headed keywords *fish*, *mace*, *sense* and *ten* in Series C and Series E: data points = 288; missing = 2.8%.

cates that the magnitude of utterance-final lengthening is greatest in monosyllables, where the stressed syllable is in absolute-final position; utterance-final lengthening is less when there are one or two syllables in the keyword between the stressed syllable and the end of the utterance. The magnitude and locus of the effect in each case is examined in more detail below.

The interaction between Accent and Utterance Position is not significant by Subjects: F(1,5) = 2.43, p = .179 [by Items: F(1,3) = 4.54, p = .122], indicating support for the utterance-final accent hypothesis. The interaction between Word Length and Accent is highly significant by Subjects: F(2,11.1) = 24.00, $p < .001$ [by Items: F(2,6.1) = 18.90, $p < .005$], indicating that word-span shortening is greater for accented keywords. As proposed in Section 4.6.1, this is due to the polysyllabic accent effect, which affects all parts of the test syllable; in addition, both accented and unaccented left-headed keywords are affected by word-rhyme-span shortening, which has a syllable nucleus locus. There is no significant three-way interaction between Accent, Utterance Position and Word Length by Subjects: $F(2,10.6) = 2.03$, $p = .179$ [by Items: $F(2,6.1) = 1.47$, $p = .301$].

Figure 4.8 shows clearly that test syllable duration is greater utterance-finally than utterance-medially, particularly when it immediately precedes the utterance boundary. The distribution of the effect within the test syllable is examined now.

Within-word results: stressed syllable onset duration

Table 4.39 shows mean test syllable onset duration for left-headed keywords in Series C and Series E[28]. The effect of Word Length is highly significant by Subjects: $F(2,10.1) = 22.22$, $p < .001$ [by Items: $F(2,14) = 17.31$, $p < .001$]. Both utterance-final and utterance-medial keywords show small but consistent word-level shortening effects; as discussed in Section 4.6.2, the evidence from Experiment 2 indicates that word-initial lengthening may be attenuated in polysyllables, both when accented and unaccented, in addition to the shortening due to the polysyllabic accent effect .

	Number of syllables in word					
	1	2	3	1	2	3
	Utterance-final			*Utterance-medial*		
Unaccented	109	101	99	105	101	98
Accented	128	117	115	128	119	115

Table 4.39: Mean test syllable onset duration (ms) for all left-headed keywords in Series C and Series E: data points = 576; missing = 2.1%.

There is no significant effect of Utterance Position by Subjects: $F(1,5) = 0.01$, $p = .931$ [by Items: $F(1,7) = 0.10$, $p = .757$]. For unaccented and accented keywords, onset duration at each word length is comparable utterance-finally and utterance-medially. This is true even in monosyllables, where the test syllable in Series E is in absolute-final position in the utterance: in unaccented keywords, the onset is 4 ms longer utterance-finally; in accented keywords, the onset is the same duration utterance-finally and utterance-medially.

There is a significant effect of Accent by Subjects: $F(1,5) = 9.65$, $p < .05$ [by Items: $F(1,7) = 151.46$, $p < .001$]. The interaction between Utterance Position and Word

[28] As discussed in Section 4.2.1, the left-headed keywords *dog* and *part* are followed in monosyllabic utterance-medial context by potentially full-vowel syllables; thus they are excluded from certain analyses, because of the possibility of stress-adjacent lengthening. There are included in the utterance-final analysis, however: firstly, any stress-adjacent lengthening in the utterance-medial series would bias the results towards the null hypothesis; secondly, any such effect is likely to be small relative to utterance-final lengthening, and distributed throughout the syllable rather than having a particular subsyllabic locus; finally, this possible confound only relates to the comparison between monosyllables. The keywords *dog* and *part* are not, however, included in the above analysis of full syllable duration, because of their inconsistently-measurable codas.

Length is not significant by Subjects: F(2,10.1) = 0.49, p = .628 [by Items: F(2,14) = 3.48, p = .058]; nor is the interaction between Utterance Position and Accent: F(1,5.1) = 1.18, p = .326 [by Items: F(1,7) = 1.21, p = .308]. The interaction between Word Length and Accent does approach significance by Subjects: F(2,10.3) = 3.37, p = .075 [by Items: F(2,14.2) = 5.81, p = .014], indicating that the greatest effect of Word Length is between accented monosyllables and disyllables/trisyllables. There is no evidence of any three-way interaction between Utterance Position, Word Length and Accent by Subjects: F(2,10.3) = 0.17, p = .849 [by Items: F(2,14.4) = 3.06, p = .078].

These results indicate that the test syllable onset is not lengthened utterance-finally. Previous results generally observe final lengthening on the rhyme of the preboundary syllable, so this result is in accordance with expectations.

Within-word results: stressed syllable nucleus duration

Table 4.40 shows mean test syllable nucleus duration in Series C and Series E. The effect of Word Length is highly significant by Subjects: F(2,10.1) = 88.63, $p < .001$ [by Items: F(2,14) = 100.21, $p < .001$], indicating the polysyllabic accent effect and the word-rhyme span effect.

	Number of syllables in word					
	1	2	3	1	2	3
	Utterance-final			*Utterance-medial*		
Unaccented	126	87	75	88	83	75
Accented	133	90	81	100	90	82

Table 4.40: Mean test syllable nucleus duration (ms) for all left-headed keywords in Series C and Series E: data points = 576; missing = 2.1%.

The effect of Utterance Position is significant by Subjects: F(1,5) = 16.48, $p < .05$ [by Items: F(1,7) = 33.88, $p < .005$]. There is a highly significant interaction between Utterance Position and Word Length by Subjects: F(2,10.1) = 62.65, $p < .001$ [by Items: F(2,14) = 51.49, $p < .001$], indicating that utterance-final lengthening is greatest for monosyllables: where the test syllable is in absolute-final position in the utterance, the nucleus is 38 ms (43%) longer in the unaccented condition and 33 ms (33%) longer in the accented condition.

Because of the interaction between Utterance Position and Word Length, the effect of Utterance Position on nucleus duration is analysed separately for monosyllables, disyllables and trisyllables, as shown in Table 4.41: this indicates that utterance-final lengthening only occurs when the test syllable immediately precedes the utterance boundary. In disyllables and trisyllables, where there are one or two unstressed syllables between the test syllable and the utterance boundary, there is no evidence of final lengthening of the test syllable nucleus. It may also be noted that there is no evidence of an interaction between Utterance Position and Accent, even in monosyllables, suggesting that the effects of final lengthening are similar whether the keyword is unaccented or accented.

	By-Subjects analysis			By-Items analysis		
Source of variation	Degrees of freedom	*F Ratio*	Significance level	Degrees of freedom	*F Ratio*	Significance level
Monosyllables						
Position	1,5	44.11	< .005	1,7	49.41	< .001
Accent	1,5.1	18.95	< .01	1,7	15.86	< .01
Interaction	1,5.1	0.90	.386	1,7	2.32	.171
Disyllables						
Position	1,5.1	0.11	.750	1,7	2.76	.141
Accent	1,5	6.24	.054	1,7	3.52	.103
Interaction	1,5.1	2.68	.161	1,7.2	1.34	.284
Trisyllables						
Position	1,5.1	0.64	.460	1,7.1	0.37	.564
Accent	1,5.1	12.51	< .05	1,7	10.92	< .05
Interaction	1,5.1	0.12	.746	1,7.2	2.46	.160

Table 4.41: Analyses of variance of test syllable nucleus duration according to word length for all left-headed keywords in Series C and Series E. There are 192 data points in each analysis, with 1.0% missing data for monosyllables and disyllables and 4.2% missing data for trisyllables.

In the analysis of the full data set for all three levels of Word Length, the interaction between Utterance Position and Accent is not significant by Subjects: $F(1,5) = 1.07$, $p = .348$ [by Items: $F(1,7) = 5.32$, $p = .054$], nor is the interaction between Accent and Word Length: $F(2,10.2) = 1.14$, $p = .358$ [by Items: $F(2,14.1) = 2.65$, $p = .105$]. The three-way interaction between Utterance Position, Accent and Word Length is not significant by Subjects: $F(2,10.6) = 1.77$, $p = .218$.

Within-word results: stressed syllable coda duration

Table 4.42 shows mean test syllable coda duration in Series C and Series E for keywords *fish*, *mace*, *sense* and *ten*. The effect of Word Length is highly significant by Subjects: $F(2,10.4) = 64.81$, $p < .001$ [by Items: $F(2,6) = 8.54$, $p < .05$]. The effect of Utterance Position is highly significant by Subjects: $F(1,5.1) = 100.18$, $p < .001$ [by Items: $F(1,3) = 13.15$, $p < .05$]. The interaction between Utterance Position and Word Length is also highly significant by Subjects: $F(2,11.6) = 185.41$, $p < .001$ [by Items: $F(2,6) = 5.55$, $p < .05$]. In contrast with nucleus duration, utterance-final lengthening of the coda extends to disyllables and trisyllables, although the effect is greatest in monosyllables, as indicated by the interaction: in monosyllables, final lengthening is 73 ms (70%) in unaccented keywords and 65 ms (50%) in accented keywords; in disyllables, final lengthening is 24 ms (22%) in unaccented keywords and 15 ms (12%) in accented keywords; in trisyllables, final lengthening is 8 ms (8%) in unaccented keywords and 7 ms (6%) in accented keywords.

	Number of syllables in word					
	1	2	3	1	2	3
	Utterance-final			*Utterance-medial*		
Unaccented	177	131	113	104	107	105
Accented	196	138	120	131	123	113

Table 4.42: Mean test syllable coda duration (ms) for left-headed keywords *fish*, *mace*, *sense* and *ten* in Series C and Series E: data points = 288; missing = 2.8%.

Because of the interaction between Utterance Position and Word Length, the effect of Utterance Position on coda duration is analysed separately for monosyllables, disyllables and trisyllables, as shown in Table 4.43[29].

For monosyllables and disyllables, Table 4.43 shows that the effect of Utterance Position is highly significant by Subjects. For trisyllables, the effect of Utterance Position on coda duration is almost significant by Subjects ($p = .052$). For neither monosyllables nor disyllables nor trisyllables is there a significant interaction between Utterance Position and Accent, indicating that—as for nucleus duration—the utterance-final lengthening effect is similar in unaccented and accented context. As discussed in Section 4.3.3, test syllable codas could be regarded as ambisyllabic or as onset consonants in left-headed polysyllables, such as *mason* or *masonry*: treating them here as coda consonants allows the results to be interpreted as a consistent pattern of lengthening on the coda in monosyllables, disyllables and trisyllables.

In the analysis of the full data-set for all three levels of Word Length, the interaction between Utterance Position and Accent is not significant by Subjects: $F(1,5.1) = 1.11$, $p = .338$ [by Items: $F(1,3) = 2.02$, $p = .249$]. There is, however, a significant interaction between Accent and Word Length by Subjects: $F(2,12.3) = 14.23$, $p < .005$ [by Items:

[29]These analyses use different subsets of the data according to the number of missing data points of the keywords *cap*, *dog* and *speck* due to coda glottalisation: they are included if the missing data is not unevenly distributed, and the proportion of missing data in the full data-set is not greater than about 5%.

	By-Subjects analysis			By-Items analysis		
Source of variation	Degrees of freedom	*F Ratio*	Significance level	Degrees of freedom	*F Ratio*	Significance level
Monosyllables						
Position	1,5.2	205.53	p < .001	1.6	10.45	p < .05
Accent	1,5.1	11.74	p < .05	1,6	19.50	p < .005
Interaction	1,5.6	1.71	.242	1,6.1	0.44	p = .533
Disyllables						
Position	1,5.1	16.85	p < .01	1,5	26.28	p < .005
Accent	1,5.1	5.28	p < .01	1,5.1	6.93	p < .05
Interaction	1,5.2	0.27	.624	1,5.2	7.00	p < .05
Trisyllables						
Position	1,5.1	6.40	.052	1,3.1	16.24	p < .05
Accent	1,5	3.12	.137	1,3.6	87.39	p < .005
Interaction	1,5.1	0.00	.950	1,3.1	0.00	.984

Table 4.43: Analyses of variance of test syllable coda duration according to word length for all left-headed keywords in Series C and Series E. For monosyllables, all keywords are included except *part*: data points = 168; missing data = 3.0%. For disyllables, all keywords are included except *part* and *speck* : data points = 144, missing data = 4.2%. For trisyllables, the keywords used are *fish*, *mace*, *sense* and *ten*: data points = 96, missing data = 2.

$F(2,6.1) = 8.17$, $p < .05$], in contrast with the findings for nucleus duration. This reflects the finding, discussed in Section 4.6.1, of a word-span effect on the coda in accented keywords only.

The three-way interaction between Utterance Position, Accent and Word Length is not significant by Subjects: $F(2,10.6) = 1.77$, $p = .218$.

Within-word results: final unstressed syllable duration in disyllables

In disyllables such as *captain* and *partner*, the effect of utterance position on the word-final unstressed syllable (syllable-2) may be examined by comparing Series C and Series E utterances, as for stressed syllable duration. The full data-set is not used for this analysis, however, for a reason indicated in Table I.1 in Appendix I, which gives mean syllable-2 duration keyword-by-keyword. Five keywords show some utterance-final lengthening in accented and unaccented keywords; for *captain*, the effect is only evident in unaccented keywords. Keywords *dogma* and *tendon*, in contrast, both show an utterance-final shortening effect, of between 15 ms and 25 ms for *tendon* and between 4 ms and 26 ms for *dogma*.

The anomalous behaviour of these two keywords may be due to the context in which they are spoken utterance-medially:

Series C.2 Tim knew the **dogma** may decline again

Series C.2 I heard the **tendon** denied again.

In both cases, syllable-2 is followed by another very similar syllable (both underlined), which may cause speakers to take particular care over enunciation in this iterative sequence[30]. Given this account of the anomalous results for *dogma* and *tendon*, they are excluded from the analysis of syllable-2 duration in disyllables.

Mean syllable-2 duration is shown in Table 4.44. The effect of Utterance Position is significant by Subjects: $F(1,5) = 12.55$, $p < .05$ [by Items: $F(1,5) = 13.65$, $p < .05$]. The utterance-final lengthening is 29 ms (35%) in unaccented disyllables and 20 ms (21%) in accented disyllables. The effect of Accent is not significant by Subjects: $F(1,5.1) = 2.99$, $p = .144$ [by Items: $F(1,5.2) = 9.31$, $p < .05$], and the interaction between Utterance Position and Accent is not significant by Subjects: $F(1,5.1) = 2.23$, $p = .195$ [by Items: $F(1,5.9) = 4.33$, $p = .083$].

	Utterance position	
	Final	Medial
Unaccented	112	83
Accented	114	94

Table 4.44: Mean syllable-2 duration (ms) in Series C and Series E for left-headed disyllables excluding *dogma* and *tendon*: data points = 144; missing = 3.5%.

Thus, both syllables in trochaic words such as *captain*, *fissure*, *mason* and *censor* manifest significant utterance-final lengthening. The absolute magnitude of the effect is comparable between the two syllables, but the proportional effect is greater on the shorter unstressed syllable. The distribution of final lengthening within polysyllables is examined further below.

[30]Cutler (1990:213) makes a similar point concerning Beckman & Edwards (1990): "It is not particularly far-fetched to suggest that the difficulty of articulating such [an iterative] sequence with clarity might have been a contributory factor in rendering speakers more prone to produce lengthening effects." Elsewhere in the Experiment 2, slight modifications to materials within particular keyword series have been made to avoid such clashes, as outlined in Section 4.2.

Within-word results: penultimate unstressed syllable duration in trisyllables

Syllable-2 is word-medial in trisyllables, for example: /tɪn/in *captaincy* and /ən/in *masonry*. In utterance-final trisyllables, syllable-2 is in utterance-penultimate position.

Table 4.45 shows mean syllable-2 duration for the trisyllables of left-headed keywords. The effect of Utterance Position is significant by Subjects: $F(1,5.2) = 10.71$, $p < .05$ [by Items: $F(1,7) = 4.61$, $p = .069$]. The utterance-final lengthening of syllable-2 is 6 ms (9%) in unaccented dss and 7 ms (10%) in accented disyllables. The effect of Accent approaches significance by Subjects: $F(1,5.2) = 4.70$, $p = .081$ [by Items: $F(1,7.1) = 9.02$, $p < .05$]. There is no significant interaction between Utterance Position and Accent by Subjects: $F(1,5.2) = 0.16$, $p = .708$ [by Items: $F(1,7.1) = 0.10$, $p = .761$].

	Utterance position	
	Final	Medial
Unaccented	74	68
Accented	79	72

Table 4.45: Mean syllable-2 duration (ms) in Series C and Series E for all left-headed trisyllables: data points = 192; missing = 2.6%.

Utterance-penultimate unstressed syllables, like stressed syllables, show significant lengthening relative to utterance-medial syllables, but the effect is small for unstressed syllables. This may be due to the distribution of the effect: in the test syllable only the coda undergoes significant lengthening in utterance-penultimate position. Subsyllabic constituent durations were not obtained for unstressed syllables, but the importance of subsyllabic structure may be inferred from the pattern of the results for each keyword, shown in Table I.2 in Appendix I. Where syllable-2 is an open syllable, there is little consistent evidence of utterance-final lengthening: only *spectacle* of open syllable-2 keywords shows utterance-final lengthening for both unaccented and accented context, 8 ms (16%) for the former and 4ms (7%) for the latter. By contrast, three keywords have large and consistent utterance-final lengthening effects: *captaincy*, *masonry* and *tendency*, and for these syllable-2 is closed (that is, contains a coda consonant). Keyword *captaincy* has 13 ms (11%) lengthening in unaccented context and 13 ms (10%) in accented context; keyword *masonry* has 18ms (27%) in unaccented context and 21 ms (31%) in accented context; keyword *tendency* has 34ms (31%) in unaccented context and 4 ms (3%) in accented context. This result supports the conclusion from the results for stressed syllables, that the locus of final lengthening in utterance-penultimate position is the coda, rather than some arbitrary set of segments at the end of the syllable.

Within-word results: final unstressed syllable duration in trisyllables

Syllable-3 is word-final in trisyllables, for example: /sɪ/in *captaincy* and /ɹɪ/in *masonry*. In utterance-final trisyllables, syllable-3 is in absolute-final position.

Table 4.46 shows mean duration of syllable-3 in left-headed trisyllables[31]. The effect of Utterance Position is highly significant by Subjects: $F(1,5) = 52.00$, $p < .005$ [by Items: $F(1,7) = 18.53$, $p < .005$]. There is a large utterance-final lengthening effect in word-final unstressed syllables: 40 ms (25%) in unaccented keywords and 35 ms (20%) in accented keywords. The effect of Accent approaches significance by Subjects: $F(1,5) = 5.74$, $p = .062$ [by Items: $F(1,7.5) = 89.44$, $p < .001$]. There is no interaction between Utterance Position and Accent by Subjects: $F(1,5.1) = 0.38$, $p = .563$ [by Items: $F(1,7.3) = 2.32$, $p = .170$].

	Utterance position	
	Final	Medial
Unaccented	203	163
Accented	212	177

Table 4.46: Mean final unstressed syllable duration (ms) for the trisyllables of all left-headed keywords in Series C and Series E: data points = 192; missing = 2.6%.

Thus, as for the test syllable in monosyllabic context, and for the word-final unstressed syllable in disyllabic context, there is a large lengthening effect in the absolute-final syllable relative to utterance-medial position. Also, for syllable-3 in trisyllables shown in Table 4.46, as for syllable-2 in disyllables shown in Table 4.44, the lengthening effect of accent is attenuated utterance-finally and does not attain significance; however, in neither instance is the interaction between utterance position and accent statistically significant.

Cross-word results: utterance-penultimate position

In utterance-penultimate position when followed by an unstressed syllable within the same word, the test syllable coda is lengthened relative to utterance-medial position . A comparison of the U+1 conditions in Series B and Series D, for example:

Series B.2 I SAW the **fish** again here TODAY.

Series D.2 Bob said he SAW the **fish** off.

[31] The keywords *dogmatist* and *partnership* have coda stop consonants in syllable-3 that are not always released; for these keywords syllable-3 duration is measured up to the point of closure of the coda stop consonant.

may indicate whether this effect also holds when the utterance-penultimate syllable is followed by another syllable which comprises a separate word.

As described in Section 4.2.1, the test syllables in Series D.2 sentences are necessarily followed by full-vowel syllables, because function words are not reduced phrase-finally. This could result in stress-adjacent lengthening of the test syllables, so this comparison contains a possible confound; as already noted, the utterances for the keywords *dog* and *part* also contain full vowel syllables following the test syllable in the Series B.2 utterances. The locus of any effect within the syllable may indicate its origin, however.

Table 4.47 shows mean onset, nucleus and coda durations in Series B.2 and D.2 utterances for left-headed keywords.

	Utterance position	
	Final	Medial
Onset		
Unaccented	107	105
Accented	127	130
Nucleus		
Unaccented	90	88
Accented	104	102
Coda		
Unaccented	123	105
Accented	135	131

Table 4.47: Mean duration (ms) of subsyllabic constituents of the test syllable in Series B.2 and Series D.2 utterances (U+1 condition) for left-headed keywords. Means for onset and nucleus include all keywords: data points = 196; missing = 6. Means for coda include keywords *fish*, *mace*, *sense* and *ten*: data points = 96; missing = 5.

The syllable onset shows no effect of Utterance Position by Subjects: $F(1,5.2) = 0.02$, $p = .894$ [by Items: $F(1,7.1) = 0.00$, $p = .992$] and no interaction between Utterance Position and Accent by Subjects: $F(1,5.3) = 1.24$, $p = .314$ [by Items: $F(1,7.3) = 6.37$, $p = .038$]. Thus, there is no evidence of either final lengthening or of the tendency towards utterance-medial lengthening of test syllable onsets in Series B noted in Section 4.5.1.

The syllable nucleus shows no effect of Utterance Position by Subjects: $F(1,5.1) = 0.06$, $p = .812$ [by Items: $F(1,7) = 0.82$, $p = .397$] and no interaction between Utterance Position and Accent by Subjects: $F(1,5.1) = 0.12$, $p = .747$ [by Items: $F(1,7.2) = 1.08$, $p = .334$]. The lack of a difference between Series B and Series D in nucleus duration suggests that the effect of stress-adjacency is not a factor in this comparison. The result mirrors that found for utterance-penultimate test syllables in trochaic disyllables such as *mason*, which show no effect of utterance position on the nucleus.

The syllable coda shows no effect of Utterance Position by Subjects: $F(1,5.2) = 2.59$, $p = .166$ [by Items: $F(1,3) = 3.63$, $p = .153$]; the interaction between Utterance Position and Accent approaches significance by Subjects, however: $F(1,5.2) = 4.74$, $p = .079$ [by Items: $F(1,3.1) = 3.16$, $p = .170$]. Because of this, unaccented and accented keywords are also analysed separately. Unaccented keywords show a significant effect of Utterance Position on coda duration by Subjects: $F(1,5.1) = 7.47$, $p < .05$ [by Items: $F(1,3) = 3.34$, $p = .165$]. Accented keywords show no significant effect of Utterance Position on coda duration by Subjects: $F(1,6.2) = 0.78$, $p = .409$ [by Items: $F(1,3.1) = 6.73$, $p = .079$].

The lengthening of the coda in unaccented keywords seems likely to be attributable to utterance-finality, rather than stress-adjacency, given the lack of effect of utterance position for onset and nucleus. If so, this result is somewhat anomalous, as none of the previous results for utterance-final lengthening indicate an interaction between utterance position and accent.

Cross-word results: utterance-antepenultimate position

In utterance-antepenultimate position, there is evidence of lengthening of the test syllable coda when followed by two unstressed syllables within the word. A comparison of Series B.3 and Series D.3 utterances, for example:

Series B.3 I SAW the **fish** again with it TODAY.

Series D.3 Bob said he SAW the **fish** again.

may indicate if this effect also holds when the utterance-antepenultimate syllable is followed by two syllables which comprise a separate word.

Table 4.48 shows mean onset, nucleus and coda durations in Series B.3 and D.3 utterances for left-headed keywords.

The syllable onset shows a significant effect of Utterance Position by Subjects: $F(1,5.1) = 10.72$, $p < .05$ [by Items: $F(1,7.2) = 5.94$, $p < .05$] and no interaction between Utterance Position and Accent by Subjects: $F(1,5.1) = 0.58$, $p = .481$ [by Items: $F(1,7.1) = 0.04$, $p = .854$]. This results indicates utterance-medial lengthening of test syllable onsets in Series B, as already noted in Section 4.5.1. It is unlikely to be influenced by stress-adjacent lengthening in Series B, as the effect is very similar when *dog* and *part* are excluded.

The syllable nucleus shows no effect of Utterance Position by Subjects: $F(1,5) = 0.76$, $p = .423$ [by Items: $F(1,7.1) = 14.07$, $p < .01$] and no interaction between Utterance Position and Accent by Subjects: $F(1,5.1) = 0.62$, $p = .467$ [by Items: $F(1,7.1) = 0.04$, $p = .844$] (and the results are very similar when *dog* and *part* are excluded). Thus, as for

	Utterance position	
	Final	Medial
Onset		
Unaccented	101	105
Accented	122	129
Nucleus		
Unaccented	84	90
Accented	97	100
Coda		
Unaccented	113	104
Accented	128	124

Table 4.48: Mean duration (ms) of subsyllabic constituents of the test syllable in Series B and Series D U+1 condition for left-headed keywords. Means for onset and nucleus include all keywords: data points = 192; missing = 4.7%. Means for coda include keywords *fish*, *mace*, *sense* and *ten*: data points = 120; missing = 7.5%.

the within-word analysis, utterance-antepenultimate test syllables show no evidence of final lengthening on the nucleus.

The effect of Utterance Position on the syllable coda approaches significance by Subjects: $F(1,5.1) = 4.22$, $p = .094$ [by Items: $F(1,4) = 2.85$, $p = .166$]. There is no interaction between Utterance Position and Accent by Subjects: $F(1,5.1) = 2.94$, $p = .145$ [by Items: $F(1,4.1) = 0.62$, $p = .473$]. Here there is no possibility of a confound due to stress-adjacency, as all test syllables are followed by reduced syllables in Series B and Series D, but the analysis should be treated with caution due to the relatively high proportion of missing data (7.5%).

This caveat notwithstanding, the pattern of apparent utterance-final lengthening in test syllables followed by another monosyllabic or disyllabic word before the utterance boundary resembles that for within-word final lengthening effects: the coda, but not the onset or the nucleus, may be lengthened in utterance-penultimate and utterance-antepenultimate position, but by much less than when in absolute-final position. The within-word results regarding this effect are fairly clear; the cross-word results are more equivocal, with the possibility that other durational processes may, in some cases, influence the observed durations.

4.8.2 Discussion

Experiment 2 tests three specific hypotheses relating to utterance-final lengthening:

Utterance-final locus hypothesis: the locus of utterance-final lengthening is the word-rhyme.

Utterance-final locus distribution hypothesis: the distribution of utterance-final lengthening is progressive within the locus.

Utterance-final accent hypothesis: the magnitude of utterance-final lengthening is not affected by the presence of a pitch accent on the utterance-final word.

The evidence relating to each of these hypotheses is considered now.

The locus of final lengthening

Experiment 2 finds that an utterance-final stressed syllable undergoes a large amount of lengthening, and an utterance-penultimate stressed syllable undergoes a significant but substantially smaller amount of lengthening relative to utterance-medial position. Furthermore, there is some evidence of lengthening of a stressed syllable in utterance-antepenultimate position when followed by two unstressed syllables. The unstressed syllables following the final stressed syllable also show lengthening; in all cases, the syllable in absolute-final position, whether stressed or unstressed, shows the greatest amount of lengthening. Thus, for example, all three syllables in *fisherman* are longer utterance-finally than utterance-medially, with the greatest effect seen on the absolute-final syllable /mən/.

These results support the utterance-final hypothesis, indicating that the locus of lengthening is the word-rhyme. This constituent begins with the primary stressed syllable within a word and extends to the word boundary. Unstressed syllables preceding the word-rhyme are not measured in this study; however, Turk (1999) shows that the unstressed first syllable of words such as *Tibet* is not lengthened intonational-phrase-finally. The present results could also be taken to indicate that the locus of final lengthening is the within-word foot, defined in Chapter 2 as beginning with a stressed syllable and continuing to a word boundary or another stressed syllable onset, whichever is the sooner; however, Turk (1999) also shows intonational-phrase-final lengthening on both syllables in words such as *Woodstock*, which comprises two within-word feet. For this reason, the word-rhyme interpretation is preferred.

Not all parts of the word-rhyme are affected by final lengthening, however: where the stressed syllable is in absolute-final position, the syllable rhyme shows significant lengthening but the onset does not. For utterance-penultimate and utterance-antepenultimate stressed syllables, only the syllable coda shows significant utterance-final lengthening[32]. The results for unstressed syllables also suggest this distribu-

[32] The lengthening of the coda in utterance-antepenultimate stressed syllables is appraoches significance (p = .052) in the analyses for Experiment 2.

tion: only those unstressed syllables which have codas show consistent lengthening in utterance-penultimate position, although, as subsyllabic durations were not obtained for unstressed syllables, the actual distribution of lengthening may only be inferred.

There is some evidence that the stressed syllable coda may also be lengthened in utterance-penultimate and utterance-antepenultimate position where there is a separate word intervening before the boundary. Thus, for example, /s/ in /meɪs/ may be longer in *mace up*|| and *mace again*|| than utterance-medially, although the effect of utterance position does not attain significance in these cases. Furthermore, stress-adjacent lengthening may be a confounding factor in the utterance-penultimate analysis.

Thus, proposing the word-rhyme as the locus of utterance-final lengthening requires two qualifications. Firstly, lengthening occurs only on the syllable rhyme in absolute-final position, and only on the syllable coda in utterance-penultimate and utterance-antepenultimate position. There is no constituent which maps precisely to this locus, so its description necessarily entails two levels: at the syllabic level, the locus is the word rhyme; at the subsyllabic level, the locus is the syllable rhyme or the syllable coda, according to position. Secondly, the word-rhyme is by definition a within-word constituent; future research may yet reveal that the locus of utterance-final lengthening is a cross-word constituent.

The distribution of final lengthening within the locus

The utterance-final locus distribution hypothesis predicts that lengthening is progressive within the locus. As neither stressed nor unstressed syllables show an interaction between the durational effects of utterance position and pitch accent in Experiment 2, the mean durations may be pooled across accent conditions for the purposes of illustrating the distribution of lengthening effects. Table 4.9 shows the magnitude of utterance-final lengthening on all constituents which show any evidence of such an effect[33]: it is clear that greater lengthening is found nearer to the end of the utterance and the trend is particularly clear when the proportion of lengthening is considered. Thus, the utterance-final locus distribution hypothesis is strongly supported in the pooled data for a number of keywords.

Influence of pitch accent on utterance-final lengthening

The utterance-final accent hypothesis predicts that the magnitude of utterance-final lengthening is not affected by the presence of a pitch accent on the utterance-final word. Experiment 2 supports this hypothesis: in none of the statistical analyses of

[33]Lengthening is shown for illustrative purposes on the test syllable nucleus in all conditions, although, as noted above, the effect is only significant on the nucleus in absolute-final test syllables.

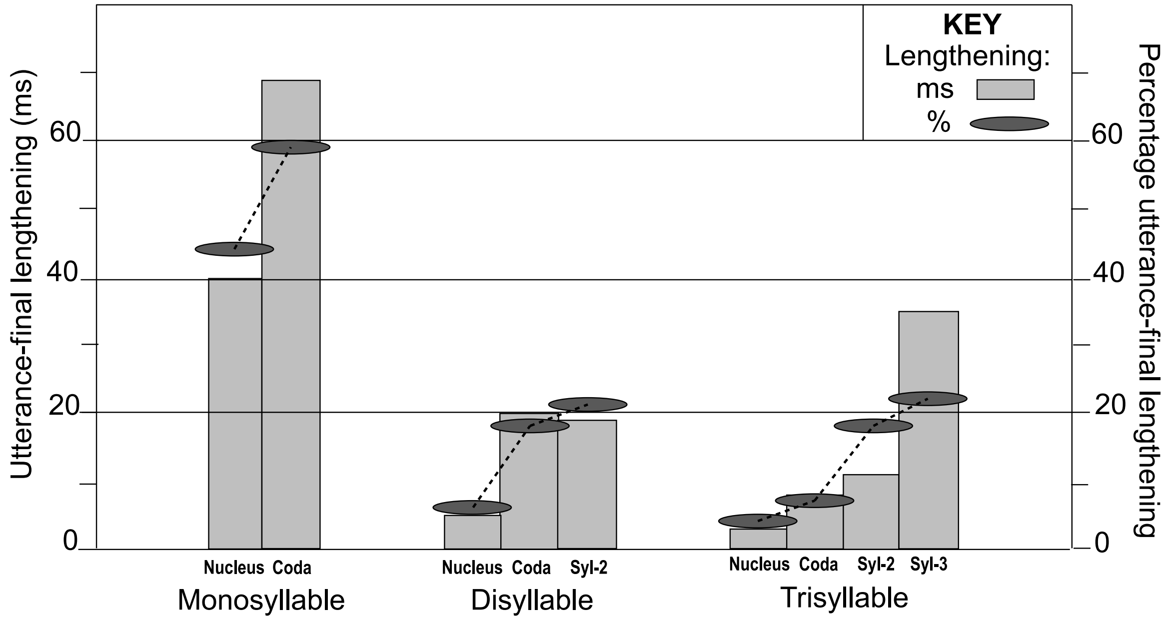

Figure 4.9: The magnitude and proportion of final lengthening (ms) for the left-headed keywords *fish*, *mace*, *sense* and *ten* in Series C and Series E on the test syllable nucleus and coda, and for the whole unstressed syllable(s) in disyllables and trisyllables.

utterance-final lengthening effects, for either stressed or unstressed syllables, is a significant interaction found between utterance position and pitch accent. Table 4.49 shows the mean and proportion of lengthening for stressed syllable subsyllabic durations and whole unstressed syllable durations.

Although there are no statistical interactions found between final lengthening and accentual lengthening, Table 4.49 clearly shows that for almost all comparisons, both the magnitude and proportion of final lengthening is greater in unaccented words. The difference in the proportion of lengthening is more marked because constituents are shorter in unaccented words, and thus a similar magnitude of final lengthening will be proportionally greater in the unaccented case. This result is quite similar to that reported in Cambier-Langeveld (2000) and may be interpreted as a ceiling effect, as discussed in Chapter 5.

This slight attenuation of final lengthening in accented keywords notwithstanding, it seems that the two effects are largely independent. This conclusion is supported by the observation that the locus of final lengthening is consistent between accented and unaccented words.

The results of Experiment 2 regarding utterance-final lengthening are discussed further in Chapter 5, and compared with previous findings.

	Stressed syllable				Unstressed syllable			
	Nucleus		Coda		Syllable-2		Syllable-3	
Monosyllables								
Unaccented	38 ms	43%	73 ms	70%				
Accented	33 ms	33%	65 ms	50%				
Disyllables								
Unaccented			24 ms	22%	29 ms	35%		
Accented			15 ms	15%	20 ms	21%		
Trisyllables								
Unaccented			8 ms	8%	6 ms	9%	40 ms	25%
Accented			7 ms	6%	7 ms	10%	35 ms	20%

Table 4.49: Magnitude and proportion of utterance-final lengthening in accented and unaccented keywords.

4.9 Accentual lengthening results

The lengthening effect of pitch accent is observed in many of the foregoing analyses of syllabic and subsyllabic durations, and is statistically significant in most instances. Two particular aspects of accentual lengthening are of interest here: the interaction between lengthening and the number of syllables in the word; and the distribution of lengthening in the stressed syllable and the rest of the accented word.

The word-span accent hypothesis, that the magnitude of accentual lengthening on the primary stressed syllable is inversely related to word length, is examined first in Section 4.9.1 and found to be supported.

The distribution of accentual lengthening is examined in left-headed and right-headed monosyllables, disyllables and trisyllables. The results for left-headed keywords provide some support for both the word-rhyme accent hypothesis, which states the locus is the word-rhyme, and the bimodal accent hypothesis, which states the locus is the primary stressed syllable and the word-final syllable. The results for right-headed keywords, however, show that the locus of accentual lengthening is the whole word.

The interpretation of these results is discussed further in Section 4.9.2.

4.9.1 Results and analysis

Accentual lengthening and word length

The word-span accent hypothesis predicts that the magnitude of accentual lengthening on the primary stressed syllable is inversely related to word length. The magnitude and proportion of accentual lengthening is shown in Figure 4.10 for keywords in Series A.

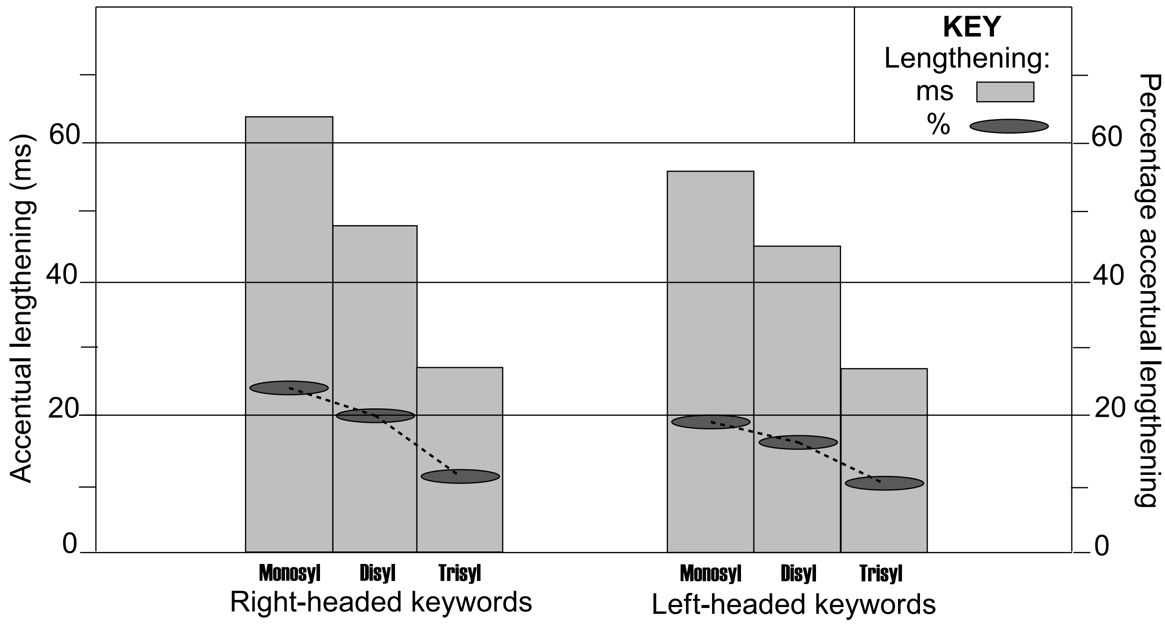

Figure 4.10: The magnitude and proportion of accentual lengthening (ms) of the test syllable in Series A. All right-headed keywords are included except *port* : data points = 252, missing = 3.8%. Left-headed keywords included are *fish*, *mace*, *sense* and *ten* : data points = 144, missing = 2.1%.

The analyses in Section 4.6.1 of test syllable duration show a significant interaction between word length and pitch accent for both right-headed and left-headed keywords: accentual lengthening of the stressed syllable is greatest in monosyllables and least in trisyllables. For right-headed keywords in Series A, the effect of Accent is highly significant by Subjects: $F(1,5) = 29.38$, $p < .005$ [by Items: $F(1,6.2) = 308.99$, $p < .001$], and the interaction between Word Length and Accent is highly significant by Subjects: $F(2,10.2) = 23.99$, $p < .001$ [by Items: $F(2,13) = 36.49$, $p < .001$]. For left-headed keywords in Series A, the effect of Accent is significant by Subjects: $F(1,5) = 15.81$, $p < .05$ [by Items: $F(1,3) = 57.06$, $p < .01$], and the interaction between Word Length and Accent is highly significant by Subjects: $F(2,10.2) = 7.71$, $p < .01$ [by Items: $F(2,6.2) = 11.90$, $p < .01$].

Details of the effect of accent and the interaction between accentual lengthening and word length for test syllable onset, nucleus and coda are given in Section 4.6.1 for Series A and in Appendix H for Series C. In most cases, subsyllabic constituents show the same pattern of durational variation as the full test syllable: significant accentual lengthening, and a significant interaction with word length such that accentual lengthening is greatest in monosyllables and least in trisyllables.

This pattern, termed the polysyllabic accent effect in Section 4.6.2, apparently underlies the finding here and in previous experiments of polysyllabic shortening that, in accented words, primary stressed syllables are longer in monosyllables, shorter in disyllables and shorter still in trisyllables; according to the polysyllabic accent hypothesis this is due to greater accentual lengthening in words of fewer syllables. The lack of strong evidence for a similar effect in unaccented keywords supports this finding: localised effects such as word-initial lengthening on the syllable onset and word-rhyme-span shortening on the syllable nucleus appear to arise from separate processes.

The reason that accentual lengthening is greater in words of fewer syllables appears to be that the additional syllables in polysyllables undergo some accentual lengthening themselves, as though there is a fixed amount of lengthening to be shared out amongst the syllables of the accented word. Because the interaction between accentual and word length is similar in right-headed and left-headed words, as shown in Figure 4.10, it is expected that the lengthening of unstressed syllables in polysyllables should show a similar degree of symmetry: that is, the effects should be similar on the additional syllables in left-headed and right-headed keywords. This is not always observed, however: studies such as Turk & White (1999) suggest there is more accentual lengthening on syllables following the primary stress than on those preceding it. The distribution of accentual lengthening in Experiment 2 is examined now.

The distribution of accentual lengthening

As just discussed, there may be a relationship between the magnitude of accentual lengthening in the test syllable and in the additional syllables in polysyllables. The durational effects of accent are examined for the subsyllabic constituents of the test syllable in monosyllables, disyllables and trisyllables, and for the additional syllables in polysyllabic keywords.

The distribution of accentual lengthening in monosyllables

Table 4.50 shows the durational effect of accent for test syllable onset, nucleus and coda of monosyllables from the right-headed keyword set. ANOVA details are shown in Table J.1 in Appendix J: in By-Subjects analyses, all subsyllabic constituents show highly significant effects of Accent. The onset shows the greatest amount of lengthening and the nucleus shows the least, although this is still a large effect.

	Onset	Nucleus	Coda
Unaccented	99	96	68
Accented	129	115	84
Lengthening	30%	20%	24%

Table 4.50: Duration (ms) of subsyllabic constituents of test syllables in right-headed monosyllables in Series B. For onset and nucleus duration, all keywords are included: data points = 288, missing = 1.4%; for coda duration, all keywords except *port* are included: data points = 252, missing = 1.6%.

Table 4.51 shows the durational effect of accent in test syllable onset, nucleus and coda for monosyllables from the left-headed keyword set. ANOVA details are shown in Table J.1 in Appendix J: the effect of Accent is significant by Subjects for onset and nucleus, and approaches significance for coda duration ($p = .065$)[34]. The pattern of lengthening resembles that for right-headed keywords, as would be expected given that for monosyllables there is no structural difference between the two groups of keywords (they are analysed separately for the purposes of comparison with the two groups of disyllables and trisyllables). Although the proportion of lengthening is slightly smaller in this set—presumably due to differences in segmental composition—the onset once again shows the largest effect of accent and the nucleus shows the smallest effect.

	Onset	Nucleus	Coda
Unaccented	105	89	114
Accented	129	100	138
Lengthening	23%	12%	21%

Table 4.51: Duration (ms) of subsyllabic constituents of test syllables in left-headed monosyllables in Series B. For onset and nucleus duration, all keywords are included: data points = 288, missing = 5.2%; for coda duration, keywords *fish*, *mace* and *sense* are included (*ten* is excluded because of missing data): data points = 108, missing = 4.6%.

The lesser lengthening of the nucleus is not attributable to the decision to include the duration aspiration of voiceless stop onset consonants within onset duration rather than within nucleus duration. Applying the alternative segmentation—that is, including aspiration within the nucleus—shows a slight lessening of the proportion of lengthening in the onset and a slight increase in the proportion in the nucleus, but

[34]The lack of a significant effect for coda duration may be due to the smaller data set used: *ten* is excluded because its inclusion creates a data-set with 7.6% missing data, greater than the conventionally-allowable proportion of 5%. If *ten* is included, however, the pattern of results is similar to that shown in Table 4.51, with unaccented duration of 105 ms and accented duration of 130 ms, a 24% accentual lengthening effect which is significant by Subjects: $F(1,5) = 9.64$, $p < .05$.

the trend remains the same. For right-headed keywords: accentual lengthening of the onset excluding aspiration is 29%; accentual lengthening of the nucleus including aspiration is 23%. For left-headed keywords: accentual lengthening of the onset excluding aspiration is 22%; accentual lengthening of the nucleus including aspiration is 16%.

The distribution of accentual lengthening in disyllables

Table 4.52 shows the durational effect of accent for test syllable constituents of right-headed disyllables. ANOVA details are shown in Table J.2 in Appendix J: the effect of Accent is highly significant by Subjects for onset and is significant for nucleus and coda. Thus the subsyllabic constituents of the stressed final syllables of words such as *compose*, *commend* and *humane* are all lengthened when the word is accented, but by less than in the equivalent monosyllables, such as *pose*, *mend* and *main*. Furthermore, the distribution of lengthening is different: in right-headed monosyllables, the nucleus shows less accentual lengthening than the onset or coda; in the right-headed disyllables, the proportion of lengthening is similar for onset, nucleus and coda.

	Syllable-2	Onset	Nucleus	Coda
Unaccented	120	83	98	68
Accented	138	98	117	81
Lengthening	15%	18%	19%	19%

Table 4.52: Duration (ms) of test syllable onset, nucleus and coda, and of syllable-2, in right-headed disyllables in Series A and C. For onset and nucleus duration, all keywords are included: data points = 192, missing = 2.1%; for coda duration, all keywords except *port* are included: data points = 168, missing = 2.4%; for syllable-2 duration, all Series C keywords are included: data points = 96, missing = 3.

Table 4.52 also shows mean duration of syllable-2, the initial unstressed syllable in right-headed disyllables: for example, /kə/ in *commend*. The effect of Accent is significant by Subjects: $F(1,5.1) = 14.28$, $p < .05$ [by Items: $F(1,7) = 29.62$, $p < .005$].

Table 4.53 shows the durational effect of accent for test syllable constituents of left-headed disyllables such as *fissure*, *mason* and *censor*. ANOVA details are shown in Table J.2 in Appendix J: the effect of Accent is significant by Subjects for onset and coda, and approaches significance for nucleus duration ($p = .077$). The lengthening effects are distributed differently compared to right-headed disyllables, and resemble the results for left-headed and right-headed monosyllables, where the onset and coda show greater proportional lengthening than the nucleus. The amount of lengthening is in all cases less than that found in equivalent monosyllables such as *fish*, *mace* and *sense*.

	Onset	Nucleus	Coda	Syllable-2
Unaccented	100	83	106	99
Accented	119	90	124	109
Lengthening	19%	8%	17%	10%

Table 4.53: Duration (ms) of test syllable onset, nucleus and coda, and of syllable-2, in left-headed disyllables in Series A and C. For onset and nucleus duration, all keywords are included: data points = 192, missing = 3.6%; for coda duration, keywords *fish*, *mace*, *sense* and *ten* are included: data points = 96, missing = 6.3%; for syllable-2 duration, all Series C keywords are included: data points = 96, missing = 4.

Table 4.53 also shows mean duration of syllable-2, the final unstressed syllable in left-headed disyllables: for example, /mə/ in *dogma*. The effect of Accent is not significant by Subjects: $F(1,5.2) = 3.91$, $p = .103$ [by Items: $F(1,7.3) = 3.26$, $p = .113$]. The lack of a significant effect of pitch accent for word-final unstressed syllables is somewhat surprising, given the magnitude of the effect and the results of previous experiments, and may reflect the relatively small size of the data-set. As discussed below, the effect of accent on syllable-2 more nearly attains significance when disyllables and trisyllables are analysed together.

The distribution of accentual lengthening in trisyllables

Table 4.54 shows the durational effects of accent for test syllable constituents of right-headed trisyllables, such as *recommend*, *inhumane* and *reproduce*. ANOVA details are shown in Table J.3 in Appendix J: for onset duration and for nucleus duration, the effect of Accent approaches significance by Subjects ($p = .060$ and $p = .054$ respectively); for coda duration, there is a highly significant effect of Accent by Subjects. The distribution of lengthening is different from that found in right-headed monosyllables and disyllables: here, the greatest amount of lengthening is on the coda; there is little lengthening on the syllable onset and an intermediate amount on the nucleus.

Table 4.54 also shows the effects of pitch accent on syllable-2 and syllable-3 in right-headed trisyllables: for example, in *recommend*, /kə/ is syllable-2 and /rə/ is syllable-3. For syllable-2, the effect of Accent is not significant by Subjects: $F(1,5.1) = 3.88$, $p = .105$ [by Items: $F(1,6.3) = 10.14$, $p < .05$], and the apparent accentual lengthening effect is smaller than for syllable-2 in disyllables. For syllable-3, the effect of Accent is significant by Subjects: $F(1,5.1) = 9.31$, $p < .05$ [by Items: $F(1,6.1) = 7.66$, $p < .05$].

Table 4.55 shows the durational effects of accent for test syllable constituents of left-headed trisyllables such as *captaincy*, *masonry* and *fisherman*. ANOVA details are shown in Table J.3 in Appendix J: the effect of Accent on onset duration is significant by Subjects; the effect of Accent approaches significance by Subjects for nucleus dura-

	Syllable-3	Syllable-2	Onset	Nucleus	Coda
Unaccented	140	129	76	98	66
Accented	161	135	80	110	76
Lengthening	15%	5%	5%	12%	15%

Table 4.54: Duration (ms) of test syllable onset, nucleus and coda, and of syllable-2 and syllable-3, in right-headed trisyllables in Series A and C. For the onset, nucleus and coda duration, all keywords except *port* are included: data points = 168, missing = 5.9%; for syllable-2 duration, all keywords in Series C except *port* are used, data points = 84, missing = 3; for syllable-3 duration, all keywords in Series C except *port* are used, data points = 84, missing = 3.

tion (p = .069) and coda duration (p = .088). The distribution of lengthening contrasts with the pattern seen for monosyllabic keywords and disyllabic left-headed keywords, where onset and coda showed more lengthening than nucleus. It also contrasts with the pattern seen for right-headed trisyllabic words, where the coda shows the greatest degree of lengthening. Here, the onset shows the greatest degree of lengthening, with lesser effects seen on the nucleus and coda.

	Onset	Nucleus	Coda	Syllable-2	Syllable-3
Unaccented	100	75	105	68	163
Accented	116	82	114	72	177
Lengthening	16%	9%	9%	6%	9%

Table 4.55: Duration (ms) of test syllable onset, nucleus and coda, and of syllable-2 and syllable-3, in left-headed trisyllables in Series A and C. For the onset and nucleus duration, all keywords are included: data points = 192, missing = 2.1%; for coda duration, keywords *fish*, *mace*, *sense* and *ten* are included: data points = 96, missing = 2; for syllable-2 and syllable-3 duration, all keywords in Series C are used, data points = 96, missing = 4.

Table 4.55 also shows the effects of pitch accent on syllable-2 and syllable-3 in left-headed trisyllables: for example, in *masonry*, /ən/ is syllable-2 and /rɪ/ is syllable-3. For syllable-2, the effect of Accent is not significant by Subjects: $F(1,5.7) = 3.48$, $p = .114$ [by Items: $F(1,7.2) = 3.04$, $p = .124$]. For syllable-3, the effect of Accent, although relatively large in terms of the mean durations, is not significant by Subjects: $F(1,5) = 3.89$, $p = .105$ [by Items: $F(1.7.6) = 46.47$, $p < .001$]. (The By-Items analysis shows a highly significant effect, which could indicate that it might also be significant by Subjects with additional data.)

Accentual lengthening of syllable-2 is more reliable when data for left-headed disyllables and trisyllables are pooled together. The full analysis is reported in Section

4.6.1, where the effect of Accent approaches significance by Subjects (p = .088) and there is no significant interaction between Word Length and Accent (p = .178). This suggests that the lack of a significant effect for either disyllables or trisyllables alone may be attributable to the relatively small data-set, combined with the relatively small size of the effect.

The effects of accent on subsyllabic durations of test syllables clearly vary in magnitude and distribution according to both word length and the position of the stressed syllable within the word. The results for syllable-2 and syllable-3 duration indicate that the effects of accent may extend beyond the primary lexically-stressed syllable, as previously found by Sluijter (1995), Turk & Sawusch (1997) and Turk & White (1999). The pattern of results does not suggest strong support for either the rightward accent hypothesis or the bimodal accent hypothesis. Both hypotheses predict more lengthening on syllables following the stressed syllable than on those preceding it, which was not the case: word-initial unstressed syllables, in particular, show large and reliable lengthening due to pitch accent.

4.9.2 Discussion

Accentual lengthening and word length

The word-span accent hypothesis predicts that the magnitude of accentual lengthening on the primary stressed syllable is inversely related to word length: the results for onset, nucleus and coda duration clearly demonstrate this in left-headed and in right-headed keywords. A similar effect is observed on the common unstressed syllable in disyllables and trisyllables: in right-headed keywords, syllable-2—for example, /kə/ in *commend* and *recommend*—shows 19ms (14%) accentual lengthening in disyllables and 6 ms (5%) in trisyllables; in left-headed keywords, syllable-2—for example, /ən/ in *mason* and *masonry*—shows 11ms (13%) accentual lengthening in disyllables and 6 ms (10%) in trisyllables.

Figure 4.11 shows the mean accentual lengthening for the whole word in monosyllables, disyllables and trisyllables: the absolute magnitude of accentual lengthening is, to a first approximation, comparable between words of different lengths. Certainly, there is no suggestion of greater accentual lengthening in polysyllables in proportion to their greater length. Accentual lengthening as a proportion of total word length clearly decreases from monosyllables to disyllables to trisyllables.

The data presented in Figure 4.11 suggest that accentual lengthening is a property of the word: thus, a certain amount of additional length due to accent is assigned to the word as a whole, and distributed amongst the constituent syllables. An alternative is to link the amount of accentual lengthening to fundamental frequency variation: for example, Beckman & Edwards (1992:374) suggest that accentual lengthening may

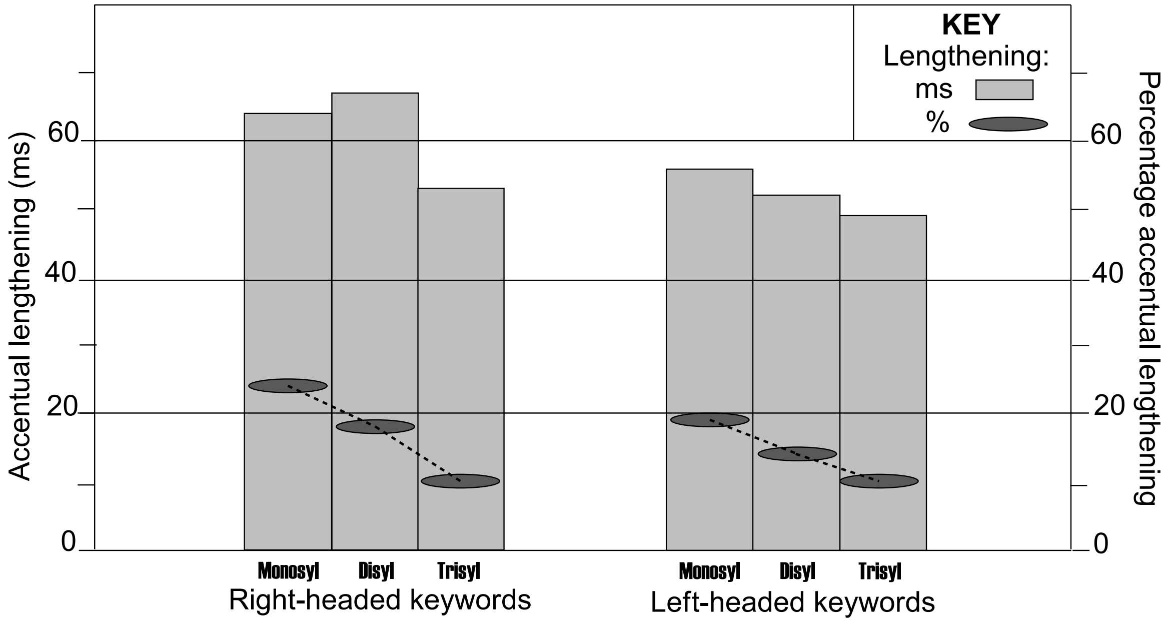

Figure 4.11: The magnitude and proportion of accentual lengthening (ms) of the whole keyword in Series C. All right-headed keywords are included except *port* : data points = 252, missing = 3.6%. Left-headed keywords included are *fish*, *mace*, *sense* and *ten* : data points = 144, missing = 4.9%.

occur "simply to make room for realising the associated pitch shapes in the intonation contour." This would suggest that words of fewer syllables require more lengthening, because in polysyllables the same pitch excursion could be accommodated without the need for much accentual lengthening; as Figure 4.11 indicates, however, there is not much attenuation of accentual lengthening across the word as a whole in polysyllables compared with monosyllables.

Research on the alignment of pitch accents also suggests a decoupling of accentual lengthening and the pitch excursion itself. Ladd *et al.* (1999) indicate that the slope and duration of a rising prenuclear pitch accent can vary to accommodate changes in speech rate or duration of the accented syllable. What is preserved under these manipulations is the alignment and F_0 level of the start and the end of the rise: the start is aligned with beginning of the stressed syllable onset and the end is aligned with the post-vocalic consonant; the F_0 levels of the start and end points differ between speakers but are fixed for individual speakers. This data is obtained for prenuclear rising accents, and it is possible that the alignment targets are different for nuclear accents—at least 94% of the accents on keywords in Experiment 2 are nuclear—although Silverman & Pierrehumbert (1990) suggest that prenuclear and nuclear accents are not distinguished phonetically. It seems highly probable that other types of accent do have fixed alignment points, however: thus even if tonal targets are differently aligned in nuclear accents—for example, with the edges of the accented word rather then the accented syllable—variation in the duration of the word would only be required to accommodate the excursion if it had a fixed slope or fixed duration. Arvaniti *et al.* (1998), Ladd *et al.* (1999) and Ladd *et al.* (2000), for a number of languages, indicate that the slope and duration of the pitch excursion are subject to variation according to the structure of the accented syllable, and in some cases, adjacent syllables.

The evidence thus suggests that the amount of accentual lengthening within a word is not an accommodation of F0 movement, but rather a fundamentally durational property of the word as a whole; as discussed below, the structure of the word determines how accentual lengthening is distributed amongst its constituents.

The distribution of accentual lengthening

Figure 4.12 shows the mean accentual lengthening in the constituents of monosyllables and polysyllables[35], clearly indicating that neither the word-rhyme accent hypothesis nor the bimodal accent hypothesis are supported. The former predicts that the locus of accentual lengthening is the word-rhyme; the latter that the locus is the primary stressed syllable and the word-final syllable. Both these hypotheses have some support from the data for left-headed keywords; however, the unstressed syllables preceding the primary stress in right-headed keywords also show lengthening effects. Indeed, the only lengthening effects which attain significance for unstressed syllables are when syllable-2 or syllable-3 are word-initial in right-headed disyllables and trisyllables.

These results strongly suggest that the locus of accentual lengthening is the word itself. Variations in the distribution of lengthening within the word appear to derive from two main sources: segmental composition and word structure. The effect of segmental composition can be seen, for example, in the difference between left-headed and right-headed monosyllables, which have no structural differences: in the former, the coda shows the greatest accentual lengthening effect and in the latter, it is the onset. Similarly, Figure 4.12 may be compared with the results in Section 4.9.1, which have somewhat different data-sets: for left-headed keywords in the larger data-set used in the statistical analysis, the syllable onset shows the greatest degree of lengthening in monosyllables (24ms, 23%), whereas for the smaller data-set used in Figure 4.12 it is lengthened by less than the syllable coda. These differences may be attributed to segmental variation, in particular, the relative expandability of different phonemes.

[35]The mean durations from which Figure 4.12 is derived are taken, for the purposes of comparison, from the same subset of the data for all constituents; thus the proportion of accentual lengthening differs slightly for some constituents from that reported in Section 4.9.1, where the largest available data-set was used for each constituent.

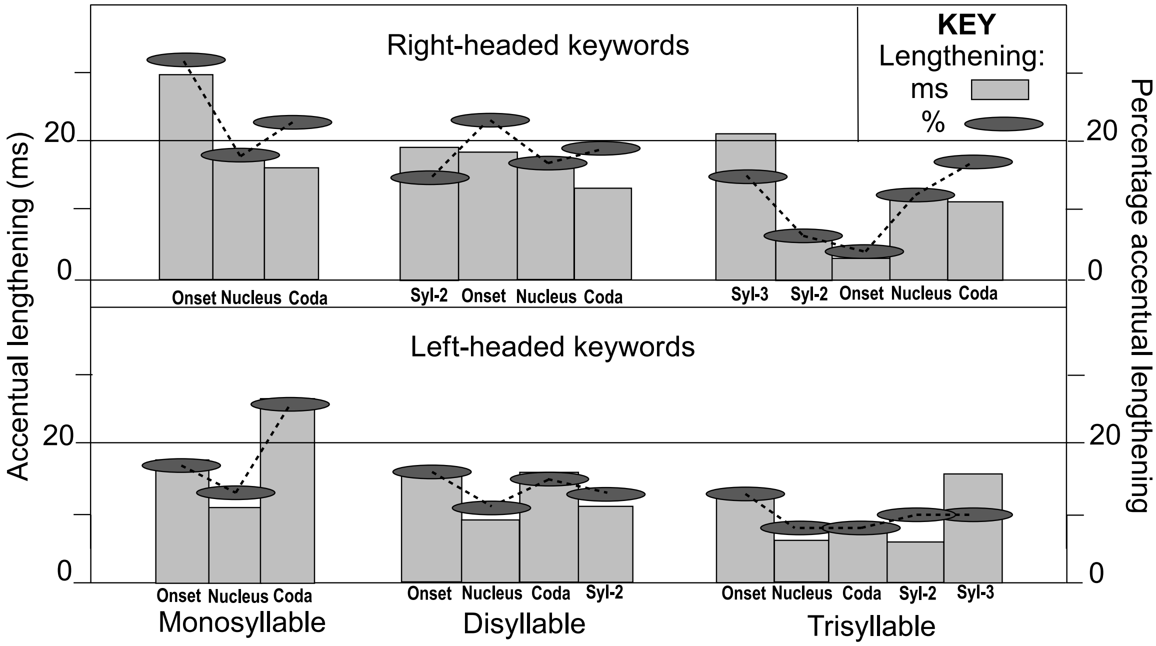

Figure 4.12: The magnitude and proportion of accentual lengthening (ms) for test syllable onset, nucleus and coda, and for syllable-2 and syllable-3 in Series C keywords. All right-headed keywords are included except *port* : data points = 252, missing = 3.6%. Left-headed keywords included are *fish*, *mace*, *sense* and *ten* : data points = 144, missing = 4.9%.

The clearest effect of word structure is the tendency for the most accentual lengthening to be evident at word edges, in both left-headed and right-headed keywords. Thus, in monosyllables the onset and coda show more lengthening than the nucleus. This pattern is also evident in disyllables, but the additional syllable reduces the accentual lengthening of the adjacent constituent by more than lengthening is reduced in the syllable as a whole: in right-headed disyllables, the accentual lengthening of the onset is 12 ms less than in the monosyllable, whereas the differences for nucleus and coda are only 1 ms and 3 ms respectively; in left-headed disyllables, the accentual lengthening of the coda is 11 ms less than in the monosyllable, whereas the onset and nucleus show only 2 ms less lengthening each. In trisyllables, the test syllable constituent on the same side as the additional syllables shows an even more marked reduction in accentual lengthening: in right-headed trisyllables, the accentual lengthening of the onset is 15 ms less than in the disyllable, whereas the differences for nucleus and coda are only 5 ms and 2 ms respectively; in left-headed trisyllables, the accentual lengthening of the coda is 8 ms less than in the disyllable, and the onset and nucleus both show 3 ms reduction in accentual lengthening.

As discussed in Section 4.6.2, the polysyllabic accent hypothesis predicts that durational differences between constituents of monosyllables, disyllables and trisyllables reflect the amount of accentual lengthening that the constituents receive in each context. Thus, according to the pattern in Figure 4.12 just described, right-headed keywords should manifest large polysyllabic shortening effects on the syllable onset and left-headed keywords on the syllable coda. Figure 4.4 in Section 4.6.2 shows word-level shortening effects in Series A[36], where the results for the disyllable vs trisyllable comparison particularly bear out this analysis: in right-headed keywords, the onset shows the greatest polysyllabic accent effect; in left-headed keywords, the coda shows a large effect. The fact that the nucleus shows the greatest polysyllabic accent effect in left-headed keywords may be attributed to the additional influence of word-rhyme-span shortening.

In the monosyllable vs disyllable comparison (see again Figure 4.4 in Section 4.6.2), the pattern of results is complicated by additional factors. For right-headed keywords, the onset does show the greatest word-level shortening effect, but word-initial lengthening clearly plays a large part in this: the effect is significantly greater in the accented condition, however, indicating the polysyllabic accent effect. For left-headed keywords, the difference in the coda is no greater than that in the onset and nucleus, contrary to the predictions based on the amount of accentual lengthening, as shown in Figure 4.12. As discussed in Section 4.6.2, however, both the other constituents may be subject to additional influences: firstly, the word-level shortening of the onset in the left-headed monosyllable vs disyllable comparison shown in Figure 4.4 may reflect an attenuation of word-initial shortening in polysyllabic words; secondly, the word-level shortening of the nucleus in the same comparison is influenced by word-rhyme-span shortening. Thus, processes which were hypothesised largely on the basis of word-level durational variation in unaccented keywords appear to have some corroboration in the pattern of word-level variation in accented keywords, insofar as they can explain why the magnitude of word-span compression for each constituent does not correspond to the predictions based on accentual lengthening.

The evidence thus appears strongly to indicate that the word itself is the locus of accentual lengthening. Possible reasons are discussed in Chapter 5 for differences between these results and those of previous studies which suggest a slightly different picture of the distribution of lengthening within the locus.

4.10 Discussion: the concept of the locus

There are two hypotheses which specifically relate to the concept of the locus.

Domain-edge locus hypothesis: each domain-edge process has a characteristic phonologically-defined locus.

[36]Figure 4.12 above shows accentual lengthening results for the same set of keywords in Series C.

Domain-span locus hypothesis: a domain-span compression process shortens only the phonological head of the domain.

How far the results of Experiment 2 support these hypotheses is discussed now.

4.10.1 Domain-edge processes

Experiment 2 finds evidence for domain-initial processes at the word-level and the phrase-level. Stressed syllable onsets are lengthened word-initially compared to word-medially, an effect which may be greater in monosyllables than in polysyllables. There also appears to be further lengthening word-initially when there is a phrase boundary preceding the word. Utterance-initially, the stressed syllable onset appears not to be subject to any lengthening, although it is also word-initial and phrase-initial: its duration is comparable to that found word-medially, at least when the whole onset is acoustically measurable.

In these cases, significant lengthening is restricted to the stressed syllable onset, strongly suggesting that this is a phonologically-defined locus. If the process was simply the lengthening of the initial part of the word by a certain amount, then the nucleus should also show some lengthening where the onset is intrinsically shorter or less expandable. This is not observed; indeed, the nucleus or coda of the word-initial syllable may be somewhat shorter following a word-initial or phrase-initial consonant: such compensatory effects are discussed further in Chapter 5.

The locus of utterance-final lengthening appears to extend from the final stressed syllable to the utterance boundary, at least where no word boundary intervenes: the absolute-final syllable, whether stressed or unstressed, shows lengthening in both the nucleus and coda; utterance-penultimate and utterance-antepenultimate syllables are only lengthened in the coda. Furthermore, the effect is progressive, with constituents closer to the boundary showing more lengthening. These observations strongly support the concept of a phonologically-defined locus, in this case the word-rhyme, although not all parts of the word-rhyme are affected. If utterance-final lengthening were a result of general deceleration over several syllables, then all constituents following the antepenultimate coda—the earliest to show lengthening—should be lengthened. Where the stressed syllable is in penultimate position, however, only the coda is lengthened: a general deceleration should also affect the onset and nucleus. Likewise in final position, the stressed syllable rhyme is lengthened but the onset is unaffected.

The concept of the locus manifested in these observations is less straightforward than that observed in domain-initial effects, as the description of the locus is not simply a matter of identifying a phonological constituent. Furthermore, it is not sufficient to say that all codas in the word-rhyme are lengthened, as the nucleus is also lengthened absolute-finally.

It is clear that the utterance-final lengthening observed in this experiment has a distinct locus; it may be questioned, however, whether the domain of the effect is actually the utterance. Wightman *et al.* (1992) for English and Cambier-Langeveld (2000) for Dutch find that intonational phrase boundaries and utterance boundaries are not distinguished by the magnitude of final lengthening on the preboundary syllable, a result which could be most simply interpreted by positing just one level of final lengthening, as all utterance boundaries are necessarily also intonational phrase boundaries. There are some problems with the interpretation of the results of Wightman *et al.* (1992), as discussed in Section 4.1.3, so it remains to be seen if the magnitude of final lengthening in English is greater utterance-finally than intonational-phrase-finally.

Whether or not the magnitude of the effect on the final syllable rhyme differs between the two constituents, the locus could distinguish the two domains if it is more extensive utterance-finally. Turk (1999) suggests that a penultimate stressed syllable may be lengthened intonational-phrase-finally, at least when accented, although she does not report which subsyllabic constituents are affected. Cambier-Langeveld (2000) finds no difference in the locus of final lengthening for Dutch between the intonational phrase and the utterance, at least within the final and penultimate syllables. It remains to be seen if the antepenultimate stressed syllable coda is also lengthened intonational-phrase-finally. If this is the case, there appears no reason for proposing two levels of domain-final lengthening, at least at and above the intonational-phrase.

In contrast, it seems clear that the utterance manifests a distinct durational effect at its initial boundary: lower-level boundaries show degrees of lengthening, but all lengthening appears to be absent utterance-initially, despite this boundary necessarily corresponding with the start of intonational and lower-level phrases. As discussed in Section 4.7.2, it may be that this effect is only observed when the utterance is spoken in isolation; where there are successive utterances separated by only short pauses, the need to indicate the onset of a new utterance may mean that an appropriate degree of lengthening of the utterance-initial onset is observed.

4.10.2 Domain-span processes

Two levels of linguistic structure are considered in Experiment 2 as possible domains of span compression processes: the utterance and the word. There is no strong evidence to suggest that the constituents of the utterance are shorter when the utterance contains more syllables; durational variations between utterances of different lengths appear much more likely to be due to timing processes at utterance-medial constituent boundaries. Similarly, most of the durational differences in stressed syllables between monosyllables, disyllables and trisyllables appear to relate either to word-initial lengthening, as discussed above, or to the effects of pitch accent: stressed syllables are longer in accented monosyllables than accented disyllables, and longer in accented disyllables than in accented trisyllables because accentual lengthening is greater in words of fewer syllables. Studies such as Turk & Shattuck-Hufnagel (2000), which find some evidence of polysyllabic shortening in unaccented words, are discussed in Chapter 5.

The observed word-span effect in accented words reflects the clear finding of Experiment 2 that the locus of accentual lengthening is the whole word, with word edges generally subject to more lengthening than word-medial constituents. The total magnitude of accentual lengthening within the whole word is, to a first approximation, similar in monosyllables, disyllables and trisyllables: the trend in most cases is for the total lengthening to be slightly less in longer words, the opposite of what would be expected if all constituents were subjected to a similar proportion of accentual lengthening. Because aggregate accentual lengthening does not increase in longer words, stressed syllable subconstituents show less lengthening in disyllables than in monosyllables, and less again in trisyllables compared with disyllables. This results in the polysyllabic accent effect, an inverse relation between stressed syllable duration and the number of syllables in accented words. This inverse relation does not appear to hold in unaccented words, however, and it is not necessary to postulate a distinct domain-span mechanism to explain the observation, because the facts about the distribution of accentual lengthening, in conjunction with word-initial lengthening and word-rhyme compression, account for the observed durational variation at the word-level.

The only clear evidence for a domain-span process in Experiment 2 is found in left-headed keywords such as *mace*, *mason* and *masonry*: the stressed syllable nucleus /eɪ/ is shorter in words of more syllables even when unaccented; the unstressed syllable /ən/ is similarly shorter in trisyllables than in disyllables, although the subsyllabic distribution of the effect is not reported for unstressed syllables. One explanation for these observations is that there is a domain-span process at the word-rhyme-level: thus, as the length of the word-rhyme increases, the duration of subconstituent duration decreases. The locus of word-rhyme compression appears to be the syllable nucleus in stressed syllables; it may be that the locus is likewise the nucleus in unstressed syllables. There is little evidence that other subsyllabic constituents are affected in stressed syllables.

It is considered in Section 4.1.2 that the locus of a domain-span process might either be the domain itself, with all constituents subject to compression, or the head of the domain. The head of the word-rhyme, if correctly identified as a prosodic constituent, would logically be the stressed syllable or the stressed syllable nucleus. The latter certainly appears to be the locus of word-rhyme compression within the stressed syllable; however, unstressed syllables within the word-rhyme also seems to be subject to word-rhyme compression. If the locus of the effect is the nuclei of all constituent syllables, then the picture somewhat resembles that found for utterance-final lengthening: in both cases, subsets of constituents within the word-rhyme are affected. Utterance-finally, all the codas within the word-rhyme, plus the final nucleus, are lengthened. Word-rhyme compression could likewise be seen as a domain-final effect: according to this interpretation, the stressed syllable nucleus is shorter in a word-rhyme of more syllables because it is further from the final edge of the word. Likewise, unstressed syllables are shorter word-medially than word-finally because of their adjacency in the latter position to the word boundary, rather than because of a direct effect of word-rhyme length. This has some support from Oller (1973) and Wightman *et al.* (1992), who find most lengthening at lower-level boundaries on the stressed syllable nucleus rather than the coda. The word-final lengthening interpretation of word-rhyme compression is discussed further in Chapter 5 and suggestions for resolving the issue are presented in Chapter 6.

There is one other source of evidence for domain-span processes in Experiment 2: the finding of compensatory shortening in syllables in which the onset has been lengthened word-initially or phrase-initially, or the coda has been lengthened utterance-finally. The adjustments in the syllable are in the direction of maintaining constant syllable duration despite the lengthening of the onset or coda, thus one interpretation of the observations could be that there is a domain-span effect at the syllable level. This is discussed in Chapter 5, where alternative explanations is also considered.

4.11 Summary

Word-level processes

The results of Experiment 2 suggest evidence for three durational processes relating to word structure:

Accentual lengthening: all constituents are lengthened in an accented word.

Word-initial lengthening: the primary stressed syllable onset is longer word-initially than word-medially.

Word-rhyme compression: the duration of the primary stressed syllable nucleus is inversely related to the number of syllables in the word-rhyme.

The finding that the word is the locus of accentual lengthening agrees with research such as Sluijter (1995) and Turk & White (1999), although the magnitude of lengthening is distributed slightly differently in Experiment 2: unstressed syllables preceding the primary stress show greater lengthening than previously observed. The magnitude of lengthening on the primary stressed syllable is greatest in monosyllables, less in disyllables and less still in trisyllables. Because there is a trend towards greater accentual lengthening at the initial and final edges of words, the attenuation of accentual lengthening is particularly marked on stressed syllable constituents adjoining the added syllables: the stressed syllable onset in right-headed keywords such as *mend*, *commend* and *recommend*; the stressed syllable coda in left-headed keywords such as *mace*, *mason* and *masonry*.

The polysyllabic accent effect emerges because of this link between word length and the magnitude of accentual lengthening: in accented words only, stressed syllable duration is inversely related to word length. Previous evidence for polysyllabic shortening is presented in research such as Lehiste (1972) and Port (1981), where the presence of a pitch accent on the measured syllable seems highly probable, although Turk & Shattuck-Hufnagel (2000) do find some evidence of polysyllabic shortening in unaccented words.

Word-initial lengthening, which is noted by Oller (1973) and Cooper (1991) among others, does not appear to be dependent upon the presence of pitch accent. There are two other observations made in Experiment 2 effects relating to word-initial lengthening. Firstly, there is evidence that the magnitude of lengthening of the word-initial syllable onset may be greater in monosyllables than in disyllables and trisyllables. Secondly, there may be some compensatory shortening within the rest of the word-initial syllable, although the magnitude of this is not as large as the word-initial lengthening itself. There is also some evidence in Experiment 2 that word-initial onset consonants are further lengthened phonological-phrase-initially, and the remainder of the initial syllable may manifest further compensatory shortening.

Word-rhyme compression has not previously been proposed, although its effect in left-headed disyllables is similar to syllable ratio equalisation, for which Turk & Shattuck-Hufnagel (2000) put forward evidence. It seems likely the combination of the polysyllabic accent effect and word-rhyme compression is sub-additive, because shortening of syllable nucleus duration in left-headed sequences such as *mace*, *mason* and *masonry* is not markedly greater in accented words than in unaccented words. It may be that word-rhyme compression, the only domain-span effect for which evidence is found in Experiment 2, is better understood as word-final lengthening affecting syllable nuclei within a word-rhyme locus.

Experiment 1 in Chapter 3 examines the effect of word length and accent on the duration of initial and final stressed syllables in phrases such as *thank fulfil* and *thankful Phil*. Initial syllables, such as /θaŋk/, are significantly longer in monosyllables than in disyllables, when accented and unaccented; final syllables, such as /fɪl/, as significantly longer in monosyllables than in disyllables only when accented, there being a slight but non-significant difference in unaccented words. These results are broadly compatible with the three word-level processes proposed here. The phrase-initial syllables are subject to both the polysyllabic accent effect and word-rhyme compression, the latter operating both in accented and unaccented words. Phrase-final syllables are subject to the polysyllabic accent effect and word-initial lengthening. Although word-initial lengthening ought to affect the stressed syllable onset in accented and unaccented words—thus causing the /f/ of /fɪl/ to be shorter in *fulfil* than in *Phil*—only full syllable durations are reported and the magnitude of the effect may be masked by variation in the rest of the syllable, particularly if there is a compensatory shortening effect in the nucleus.

Utterance-level processes

Experiment 2 finds evidence for two effects relating to utterance structure:

Utterance-initial shortening: the onset of word-initial primary stressed syllables is shorter utterance-initially than utterance-medially.

Utterance-final lengthening: all codas within the word-rhyme are lengthened utterance-finally; the nucleus is also lengthened in absolute-utterance-final position.

Fourakis & Monahan (1988) and Fougeron & Keating (1997) report utterance-initial shortening in unstressed syllable onsets; here it is observed in stressed syllable onsets, and explained as an absence of any hierarchical initial lengthening effect, because, being preceded by silence, the utterance-initial boundary requires no such cue. According to this interpretation, it would not be correct to describe utterance-initial shortening as a process, but rather as the absence of a process observed at other hierarchical levels.

Many studies, as discussed in Chapter 2, have reported that the greatest final lengthening in intonational phrases or utterances is found on the final syllable rhyme. Cambier-Langeveld (2000) finds that stressed syllables in utterance-penultimate position are lengthened, an effect also suggested by Nakatani *et al.* (1981) and Turk (1999) for intonational phrases. Experiment 2 shows that the coda of a stressed syllable in utterance-penultimate or utterance-antepenultimate position is lengthened, although the trend is for greater lengthening nearer the boundary; both the nucleus and coda of the utterance-final stressed syllable show large amounts of lengthening.

Unstressed syllables in utterance-final position also show a large amount of lengthening relative to utterance-medial position; utterance-penultimate syllables show a small lengthening effect, which indirect evidence suggests may be localised on the syllable coda.

The hypothesis that stressed syllable duration is inversely related to the number of syllables in the utterance, proposed by Jones (1942–43) and Lehiste (1974), is not supported. There is some evidence of utterance-medial phrase-initial lengthening of stressed syllable onsets, but little evidence of utterance-medial phrase-final lengthening below the intonational phrase level.

The results of Experiment 2, and the durational processes they suggest, are discussed further in Chapter 5, particularly where results differ from those of previous studies. An model of English suprasyllabic speech timing incorporating these processes is presented in Chapter 6, and ideas are put forward for research which could resolve some of the outstanding theoretical issues.

Chapter 5

General discussion

5.1 Introduction

Domain-edge and domain-span timing processes at the word level and the utterance level are examined in Experiment 2, presented in Chapter 4, as is the durational effect of pitch accent and the interaction of pitch accent with edge and span processes. There is support for domain-edge processes: in particular, word-initial lengthening and utterance-final lengthening. Results also suggest that the syllable onset—the locus of word-initial lengthening—may be further lengthened phrase initially (for example, phonological-phrase-initially); however, there is an absence, for certain onset consonants, of any domain-initial lengthening effect utterance-initially. There is little support for domain-span processes at the word level and the utterance level, although there is evidence of a domain-span process at the word-rhyme level. The locus of accentual lengthening is shown to be the word, with the greatest lengthening effects observed word-initially and word-finally. The variation in accentual lengthening of the primary stressed syllable according to word length accounts for previously-observed polysyllabic shortening, an apparent domain-span process at the word level.

The evidence for these processes is considered further in this chapter: in particular, possible interpretations are suggested where the findings of Experiment 2 are different from those of previous studies. Domain-span processes are discussed in Section 5.2, domain-edge processes in Section 5.3 and the effects of accent in Section 5.4.

There is one other class of observations arising from Experiment 2, compensatory shortening, where domain-edge lengthening within some locus appears to be accompanied by an opposite, though smaller, effect outside the locus. Interpretations of compensatory effects are considered in Section 5.5.

5.2 Domain-span processes

There are three domain-span processes examined in Experiment 2: word-rhyme-span compression, which is supported; word-span compression (polysyllabic shortening) and utterance-span compression, which are not supported. Additional experimental evidence for and against these processes is considered in the following sections.

5.2.1 Polysyllabic shortening

In studies of polysyllabic shortening such as Lehiste (1972) and Port (1981), the measured word is likely to be pitch accented, as discussed in Chapter 3. If so, such results are compatible with the conclusion of Experiment 2 that observed polysyllabic shortening is a result of the variable distribution of accentual lengthening according to word length. In addition, both Lehiste and Port examine series of left-headed words only, such as *speed* ... *speedy* ... *speedily*. As word-rhyme compression appears to affect syllable nuclei within the word-rhyme, the results of Lehiste and Port, which show the greatest word-level effect on the nucleus and smaller effects on the onset and coda, are compatible with a combination of word-rhyme compression and the polysyllabic accent effect[1].

Turk & Shattuck-Hufnagel (2000) look for evidence of polysyllabic shortening in accented and unaccented words, using two unaccented conditions: one where emphatic stress precedes the test phrase, and one where the emphatic stress is on the other word in the test phrase. The full set of conditions is exemplified by utterances for the keyword *tuna* (underlined):

Accented Say "DOLPHIN choir" again, don't say "<u>TUNA</u> choir" again.

Accent-on-say SHOUT "tuna choir" again, don't SAY "<u>tuna</u> choir" again.

Accent-in-phrase Say "tuna SCHOOL" again, don't say "<u>tuna</u> CHOIR" again.

The results for left-headed words are largely compatible with those of Experiment 2: thus, in comparisons such as *tune* vs *tuna*, the "syllable centre"[2] shows a word-length effect in all accent conditions, as predicted by word-rhyme compression (and also by the polysyllabic accent effect when the word is accented).

[1]As discussed in Chapter 2, experiments for Swedish (for example: Lindblom & Rapp 1972) and Dutch (Nooteboom 1972) suggest a small word-span compression effect on stressed vowel duration due to syllables preceding the stressed syllable and a large effect due to syllables following the main stress. This is compatible with a combination of the polysyllabic accent effect and word-rhyme compression, if the words are accented. Examination of the materials suggests this to be the case: Lindblom & Rapp, like Port (1981), use words in a fixed frame sentence, and Nooteboom reports stressed vowel duration in isolated words; in both cases, the measured words are highly likely to carry phrasal stress.

[2]Details of the subsyllabic constituents measured by Turk & Shattuck-Hufnagel are given Chapter 3.

Word-initial syllable onset aspiration is significantly longer in *tune* than in *tuna* in the accented condition, which could be due to the polysyllabic accent effect, and also to the attenuation of initial lengthening in polysyllables. The latter effect is suggested in Experiment 2 by the shortening of word-initial syllable onsets in unaccented left-headed disyllables compared with monosyllables (for example, /m/ is shorter in *mason* than in *mace*) and could also be responsible for the observation by Turk & Shattuck-Hufnagel of shortening of stressed syllable onset closure duration in disyllables in the accent-in-phrase condition. No evidence of this effect is observed in the accent-on-say condition, however, possibly influenced by the presence of an accent immediately preceding the measured syllable onset (for example: *. . . don't SAY "t̲una choir" again*): Turk & White (1999) show some lengthening in words following accented monosyllables, which could mask any effect of word length on the magnitude of initial lengthening.

The results for the stressed syllable final consonant in left-headed words are in line with those for Experiment 2. The difference in duration of, for example, /n/ in *tune* vs *tuna*, only approaches significance in the accented condition, in line with the polysyllabic accent effect and the assertion that the locus of word-rhyme compression is the syllable nucleus.

Some of the results for right-headed keywords reported by Turk and Shattuck-Hufnagel are less in agreement with those of Experiment 2. The first consonant, however, clearly shows the word-initial lengthening effect, although not as consistently as in Experiment 2: for example, in all three accent conditions either the closure or the aspiration of /k/ is longer in *choir* compared with *acquire*.

According to the polysyllabic accent hypothesis, the syllable centre should show a word-length effect in the accented condition, but not otherwise. This is not what is observed, however: the only significant difference in a By-Subjects analysis in, for example, /waɪ/ between *choir* vs *acquire*, is in the accent-on-say condition. In the other two conditions, the greater duration of /waɪ/ in the monosyllable than in the disyllable is significant only in a By-Items analysis. It may be noted, however, that the syllable nucleus in right-headed keywords in Experiment 2 actually shows a small lengthening effect in the disyllable compared with the monosyllable: for example, /ɛ/ is slightly longer in *commend* than in *mend*. This effect, found in both accented and unaccented keywords, is the only positive correlation found between word-length and subconstituent duration in Experiment 2 and is attributed in Section 5.5 to a compensatory effect prompted by the large word-initial lengthening effect on the preceding segment in the monosyllable.

This may explain the lack of a word-length effect for the syllable centre in accented right-headed keywords in Turk & Shattuck-Hufnagel's study, which show the largest word-initial lengthening of the syllable onset. The compensatory shortening of the syllable centre following the lengthened onset in *choir* compared with *acquire* may serve to mask the underlying polysyllabic accent effect on the nucleus. The reason for the observation of a word-length effect in right-headed words in the accent-on-say condition is unclear, however: possible explanations are discussed further below.

The results for final consonant duration indicate no significant differences according to word length in right-headed keywords. This is in line with the results of Experiment 2 except for the accented condition, where /ɹ/ should be longer in *choir* than in *acquire*, due to attenuation of accentual lengthening in the disyllable. The word-length effect on the test syllable coda in right-headed accented keywords is relatively small, however: for example, in Experiment 2, the shortening of /z/ in *presuppose* compared with *suppose* when accented is smaller than that found on the onset and nucleus. The reason given in Chapter 4 is that accentual lengthening appears greater at the edges of accented words than at the centre, and thus the polysyllabic accent effect is greatest on segments which have different alignment with word edges in monosyllables, disyllables and trisyllables: /p/ receives a lot of accentual lengthening in *pose*, less in *suppose* and less still in *presuppose*, whereas the word-final consonant remains word-final throughout the sequence, and thus accentual lengthening is attenuated only slightly. The evidence for the distribution of accentual lengthening which may underlie these observations is compared with previous results in Section 5.4 below.

A further contributing factor to the absence of a polysyllabic accent effect on the word-final consonant could be the unusual syllable segmentation used by Turk & Shattuck-Hufnagel: in four out of 11 cases, part of the syllable coda is measured with the syllable centre. Furthermore, the final consonant in the phrase is also antepenultimate in the utterance, being followed only by the word *again*, and thus may be subject to some durational influence of utterance-finality: the actual magnitude of accentual lengthening appears to be slightly, if not significantly, reduced utterance-finally, as found in Experiment 2 and by Cambier-Langeveld (2000); a corollary of reduced accentual lengthening would be a smaller difference in accentual lengthening between monosyllables and disyllables.

Given these factors which may affect final consonant results, the most problematic discrepancy between the results of Turk & Shattuck-Hufnagel and the processes suggested by Experiment 2 is the finding, mentioned above, of a word-length effect on the syllable centre in the unaccented, accent-on-say condition (for example: *. . . don't SAY "tuna* <u>*choir*</u>*" again*), and also a significant effect (By-Items only) in the unaccented, accent-in-phrase condition (for example: *. . . don't say "TUNA* <u>*choir*</u>*" again*): thus, in both cases /waɪ/ is longer in *choir* than *acquire*. This is not predicted by any of the word-level mechanisms proposed in Chapter 4: word-initial lengthening has a syllable onset locus; the polysyllabic accent effect only pertains to accented words; word-rhyme compression, although having a syllable nucleus locus, is proportional to the length of the word-rhyme, which is constant in right-headed words with a final primary stressed syllable. One factor which may contribute somewhat to the observed effect is the segmentation used by Turk & Shattuck-Hufnagel: where the syllable contains a complex onset, only the first consonant is measured separately. This is the case for three out of 11 right-headed keywords—*quit* vs *acquit*, *press* vs *oppress*, *choir* vs *acquire*—where the syllable centre contains a consonant that would normally be regarded as part of the onset. As it seems very clear that the onset is subject to lengthening in monosyllabic context, there is likely to be lengthening within the syllable centre in these three tokens. How much effect this has on the magnitude of the observed effect across all the materials is uncertain.

It is possible that the structure of the experimental materials and the placement of accents in Turk & Shattuck-Hufnagel's study may also have some bearing on the outstanding disparity between these findings and those of Experiment 2. Firstly, the test words, like those in Experiment 1, are placed in a metalinguistic context, with the phrase containing these words spoken in something like a citation form, for example (test phrase underlined):

SHOUT "tuna choir" again, don't SAY "<u>tuna choir</u>" again.

It may be that this structure causes the whole phrase to be focused, so that words within the phrase are subject to a certain amount of lengthening even when unaccented. This could account for the evidence of polysyllabic shortening in the "unaccented" conditions, although it is not clear why this evidence should be largely restricted to the stressed syllable centre of right-headed keywords.

Secondly, the pattern of results differs between the two "unaccented" conditions: as Turk & Shattuck-Hufnagel (2000:427) point out, "the magnitude of word-boundary-related effects appears to be tied to the occurrence of a pitch accent nearby". If the effects of accent placement were entirely restricted to the accented word, there should be no difference between the unaccented accent-on-say condition and the unaccented accent-in-phrase condition. As noted in Chapter 2, Turk & White (1999) find a small amount of lengthening on the syllable following an accented monosyllable despite the intervening word boundary; Herment-Dujardin & Hirst (2002) suggest that "words preceding a focused word are usually shorter"[3]. These factors could play a part in Turk & Shattuck-Hufnagel's study; in contrast, in the unaccented condition in Experiment 2, pitch accents are placed some distance from the measured syllables.

[3]The interpretation of durational effects of accent placement beyond the accented word is discussed further in Section 5.4.

This discussion suggests that the results of previous studies regarding polysyllabic shortening are, to a first approximation, in line with the findings of Experiment 2. Most of the differences between Experiment 2 and Turk & Shattuck-Hufnagel (2000) may be related to methodological factors, in particular the phrasing and placement of accent in the materials in the latter study, and possibly the segmentation into subsyllabic constituents used there. If it is correct, then, to conclude that polysyllabic shortening arises as a result of variations in accentual lengthening, it is important to know how the distribution of lengthening varies according to word length. The picture presented in Chapter 4 is compared with the results of the studies by Sluijter (1995), Turk & Sawusch (1997) and Turk & White (1999) in Section 5.4.

5.2.2 Word-rhyme compression

The results of Experiment 2 suggest a word-rhyme-span process, independent of pitch accent, with a syllable nucleus locus. This finding agrees with previous results for left-headed keywords, such as those of Lehiste (1972) and Port (1981), which indicate that the syllable nucleus shows the greatest compression effect due to word length. Similarly, Oller (1973) finds a word-final lengthening effect which appears greater on the syllable nucleus than on the coda. If it is correct to infer that the test words in these previous experiments are accented, than two processes would be present. Firstly, the polysyllabic accent effect , causing all parts of the syllable to be shorter in accented words of more syllables. Secondly, a word-rhyme compression effect, further shortening only the syllable nucleus when it is separated from the end of the word by one or more syllables. The evidence from the current experiment suggests that the polysyllabic accent effect and word-rhyme compression combine sub-additively: the effect of two shortening processes being less than the sum of the two processes acting separately suggests limits to compressibility of the sort proposed by Klatt (1976).

Word-rhyme compression is similar in effect to what Turk & Shattuck-Hufnagel (2000) call "syllable ratio equalisation". As discussed in Chapter 3, this is proposed to account for Abercrombie's (1965) observation that the syllables in *greater* are more similar in duration than those in *Grey to*. Turk & Shattuck-Hufnagel find evidence to support such a process with a syllable nucleus locus: in comparisons such as *tune* vs *tuna*, most of the shortening in the disyllable is manifest on the syllable centre, in all accent conditions. In addition, the central unstressed syllable in comparisons such as *tune acquire* vs *tuna choir* does not manifest durational variation due to word-membership such as would be consistent with an "asymmetrical polysyllabic shortening" process; thus, they argue, if polysyllabic shortening is symmetrical, the greater word-length effect on stressed syllable centres in left-headed keywords must be due to an additional process.

The results of Experiment 2 suggest that Turk & Shattuck-Hufnagel are correct to identify the syllable centre/nucleus as the locus of the effect within the primary stressed syllable; there is little evidence in onset and coda duration of word-rhyme compression. The results of Experiment 2 indicate, however, that the locus extends to other syllables within the word rhyme: for example, /ən/ is shorter in *masonry* than in *mason*, and it is hypothesised that the effect is localised on the nucleus in unstressed syllables as well. The effect on unstressed syllables in Turk & Shattuck-Hufnagel's experiment is difficult to determine, as they may be subject to different durational influences according to the different position-in-word and pitch accent conditions. In comparisons such as *tune acquire* vs *tuna choir*, the central unstressed syllable is either word-initial or word-final, and the main durational influence on this syllable is lengthening due to membership of an accented word, whether preceding or following the primary stressed syllable. In the accent-on-say condition, however, there is no accent within the test phrase and there appears to be either an absence of word-initial and word-final effects, or a balance between the two, as there is no difference in the duration of, for example, /ə/ in *tuna* vs *acquire*. Word-initial lengthening is well-established, as discussed in Section 5.3, but the locus appears to the syllable onset, and in Turk & Shattuck-Hufnagel's materials there is only one central syllable out of 11 which contains an onset. As the existence of a durational effect on absolute word-initial vowels in not certain, the balance between different processes on the central unstressed vowel cannot be determined[4]: if there is no effect of absolute-word-initial position on vowels, however, then there appears to be no evidence for a word-final lengthening effect. The word-rhyme compression interpretation may be still available, however: for the unstressed syllable, a disyllabic word-rhyme is the shortest possible, thus a domain-span effect might only exert a compression in a longer word-rhyme. In word-initial position, the unstressed syllable is not dominated by a word-rhyme and would likewise not be subject to word-rhyme compression. This absence of a compression effect in both cases could explain the lack of durational difference in, for example, /ə/ in *tuna* vs *acquire*.

It is argued in Chapter 3 that the syllable ratio equalisation mechanism is unnecessary, because a combination of domain-span processes could account for the results equally adequately without requiring the postulation of a different *type* of process. Experiment 2 does not, however, provide support for the existence of any suprasyllabic domain-span process other than word-rhyme compression. Thus, the argument put forward in Chapter 3 against syllable ratio equalisation on the grounds of parsimony is not valid here, because both processes appear unique of their kind; that is, both

[4]A word-initial vowel may be glottalised, which is likely to lengthen its total duration. Studies such as Dilley *et al.* (1996) indicate that glottalisation is more likely at higher-level prosodic phrase boundaries and when the word is accented.

word-rhyme compression and syllable ratio equalisation are processes predicated to explain results within a single domain.

Theoretical and empirical arguments are available for maintaining the preference for the domain-span interpretation. Firstly, the greater shortening of the syllable nucleus for left-headed keywords which Turk & Shattuck-Hufnagel observe in monosyllables vs disyllables is also observed in disyllables vs trisyllables in Experiment 2: thus, /eɪ/ is shorter in *masonry* than in *mason*. It is not clear how syllable ratio equalisation would accommodate this result. Secondly, syllable ratio equalisation appears to belong more naturally to the sphere of rhythmical processes, defined in Chapter 1 as those arising from the distribution of lexical stress; as indicated in Chapter 2, evidence suggests that the primary rhythmical process is stress-adjacent lengthening, which does not respect the boundaries of syntactic or prosodic constituents and appears to arise from a separate subsystem in the representation of speech timing. The effect on stressed syllable nucleus duration under discussion here appears, in contrast, to be dependent on word structure.

An account in terms of existing processes, for which independent evidence exists at other levels of linguistic structure, would be preferable to either word-rhyme compression or syllable ratio equalisation. The interpretation of word-rhyme compression as word-final lengthening is examined further in Section 5.3.2.

5.2.3 Utterance-span compression

Jones (1942–43) claims that domain-span compression processes may be observed in sentences; Lehiste (1974; cited in Klatt 1976) and Rakerd *et al.* (1987) find experimental evidence to support this. The results of Experiment 2, however, are line with those of Gaitenby (1965), who finds little effect of sentence length that cannot be better explained as a domain-edge process. Segments near the utterance-final boundary are subject to lengthening, and thus will be shorter if utterance size is increased by placing additional words following the measured segments. For example, the phrase *for tomorrow* is measured in:

1. Why don't you get tickets for tomorrow?

2. Why don't you get tickets for tomorrow night's programme?

and found to be longer in Sentence 1 than in Sentence 2. The results of Experiment 2 suggest that the difference is likely to be localised within the word-rhyme /mɒɹəʊ/ and progressive within this locus.

In addition, as suggested by Gee & Grosjean (1983), longer utterances tend to comprise more constituent phrases than shorter utterances. Where the alignment of a measured subconstituent with respect to phrase boundaries is different in longer and shorter utterances, durational consequences will be observed. For example, in:

1. The frog leapt from the lily pad.

2. Bill the happy bullfrog leapt from the lily pad.

the durational of /l/ in *leapt* may be greater in Sentence 2, where it is more likely to be preceded by a phrase boundary than in Sentence 1. Similarly, the duration of /ɒg/ in *frog* may be greater in Sentence 2 than in Sentence 1, although, as pointed out in Chapter 4, the evidence for phrase-final lengthening below the intonational phrase remains inconclusive.

Such observations are in the opposite direction to those suggested by Jones, Lehiste and Rakerd *et al.*, who predict shortening in longer utterances. In addition to the absence of utterance-final lengthening, as illustrated in the Gaitenby example above, there are two reasons why an inverse relationship between utterance size and subconstituent duration might be observed. Firstly, as discussed in Chapter 2, Gee & Grosjean suggest that there is a tendency for the phrases into which an utterance is divided to be of similar size. In some cases, this could mean that words are adjacent to a phrase boundary in a shorter sentence, but not in a longer sentence, the boundary being placed elsewhere for reasons of balance between subconstituents. Secondly, in certain time-restricted circumstances, there is likely be an inverse correlation between the size of the utterance and the rate at which it is spoken. In an experimental task, for example, where speakers are required to read successive sentences presented at a fixed rate, it may be necessary to read longer sentences more quickly to finish them within the time available. Such an effect may also be observed in non-experimental speech tasks when time is restricted: for example, a radio continuity announcer with a short time between programmes will talk more quickly when the content of links is longer. In most normal situations, however, speech is not restricted to a fixed time frame, although emotional factors such as a feeling of pressure to complete one's turn could cause an increase in speech rate in some circumstances.

Given the lack of evidence for an utterance-span effect, it may be asked whether other constituents manifest domain-span compression. The evidence for word-span compression suggests this is an indirect effect related to accent. Word-rhyme compression is supported but may be interpreted as a word-final effect. Furthermore, the word-rhyme has not been proposed as a constituent of prosodic hierarchies, and violates the constraint of exhaustivity—as discussed in Chapter 2—required by some accounts. The possibility that constituents between the word and the utterance, such as various types of prosodic phrase, could be domains of compression processes has not been examined. In Experiment 2, the intonational phrase is, in most cases, co-extensive with the utterance, so it seems unlikely that there exists an intonational-phrase-span compression process. The possibility remains that other levels of the prosodic hierarchy, such as phonological phrases, may manifest an inverse relationship between length and subconstituent duration, but the most significant consequences of variations in phrasing appear to be localised and associated with boundaries.

5.3 Domain-edge processes

There is evidence in Experiment 2 of domain-edge lengthening word-initially, phrase-initially and utterance-finally; there is also some evidence of an utterance-initial shortening effect. The nature of these processes and the additional experimental evidence for them are discussed in the following sections. As noted above, word-rhyme compression may be interpreted as a word-final lengthening process; arguments for this interpretation are presented below.

5.3.1 Initial lengthening

Experiment 2 suggests two levels of initial lengthening below the level of the intonational phrase, word-initial and (phonological) phrase-initial, and also suggests that there may be an absence of hierarchical lengthening effects utterance-initially.

Word-initial lengthening

Word-initial lengthening is a large effect, with a syllable onset locus: this supports findings, discussed in Chapter 2, of Oller (1973) and Cooper (1991), who find lengthening of the word-initial syllable onset, and of Fougeron & Keating (1997), Turk & Shattuck-Hufnagel (2000) and Byrd (2000) who suggest that lengthening does not extend beyond the word-initial syllable onset. The greater magnitude of lengthening in accented words is probably due to the additional influence of the polysyllabic accent effect: the stressed syllable onset is word-initial in monosyllables, such as *mend*, and word-medial in disyllables and trisyllables, such as *commend* and *recommend*; thus, both word length and position-in-word should cause shortening in polysyllables compared with monosyllables when the word is accented. Confirmation of this would require comparison of syllable onsets in word-initial and word-medial position in accented and unaccented words of equal length: for example, /d/ in *debtor* vs *cadet* would be expected to show no interaction between word-initial lengthening and pitch accent[5].

[5] Even in this comparison, there may be a slightly larger word-initial lengthening effect in accented words: /d/ is likely to receive greater accentual lengthening in *debtor* than in *cadet*, as the magnitude of lengthening appears to be greatest at the edges of accented words.

Where the stressed syllable onset is a voiceless stop, the effect of position-in-word appears to be similar both for closure duration and for aspiration duration in accented and unaccented syllables; in contrast, Oller (1973) and Cooper (1991) indicate that word-initial lengthening of closure duration is greater in unaccented syllables and that aspiration duration is longer word-medially than word-initially in accented syllables. Those studies do not conflate word length and position-in-word because they examine the effect on onset duration of position in words of fixed length; in Experiment 2, however, possible word-medial lengthening of aspiration could be masked by the polysyllabic accent effect: thus, the aspiration of /p/ is longer in *pose* than in *suppose*, at least partly due to the greater accentual lengthening it receives in the monosyllable[6].

There are two features associated with word-initial lengthening which have not previously been identified. One is the possible compensatory shortening of the following vowel, as discussed in Section 5.5. The other is the attenuation of word-initial lengthening in polysyllables: onset duration is shorter in monosyllables such as *mace* than in left-headed disyllables and trisyllables such as *mason* and *masonry*, even in unaccented words, where no effect of word length is predicted on the syllable onset. This shortening effect appears to be a binary distinction, as there is no durational difference in test syllable onset duration between unaccented left-headed disyllables and trisyllables. Examination of the data of Turk & Shattuck-Hufnagel (2000) provides some support for this finding, as discussed in Section 5.2.1 above.

There is evidence in Experiment 2 of word-initial lengthening in unstressed syllables, as found by Oller, Cooper and Fougeron & Keating: here the effect is found to be small in unaccented words. This resembles the pattern found in Experiment 1 for stressed syllables: in both cases, subsyllabic constituents are not measured and it may be that initial lengthening of the onset is masked by compensatory shortening of the nucleus. In the accented condition, lengthening due to word-initial position is reinforced by the polysyllabic accent effect, as outlined above.

Phrase-initial lengthening

The evidence for phrase-initial lengthening in Experiment 2 is that word-initial stressed syllable onsets have greater mean duration in utterance-medial context than in near-utterance-initial context: apparently the measured onset is more likely to be preceded by a phrase boundary in utterance-medial position, and where there is such a bound-

[6] An alternative interpretation of the results of Oller (1973) and Cooper (1991) is that word-initial lengthening of closure and aspiration is greater in *unstressed* syllables. For example, Cooper finds that aspiration duration of /k/ is greater in *k̲iˈkik* than in *ˈkik̲ik*, whereas the effect of position is reversed in stressed syllables: thus, the aspiration of /k/ is shorter in *ˈk̲ikik* than in *kiˈk̲ik*. In Experiment 2, only stressed syllables are examined, and aspiration duration is longer word-initially in both accented and unaccented words. As pointed out above, this lengthening in accented words is also predicted by the polysyllabic accent effect.

ary, this is associated with lengthening of the following syllable onset: for example, /m/ is longer in

B.3 JONATHAN saw Jessica **mend** it AGAIN.

than in

D.3 Will you **mend** it AGAIN for me please.

The type of phrase must be between the word level and the intonational phrase level, probably what is commonly called the phonological phrase. This result supports that of Fougeron & Keating (1997), who find word-initial and phonological-phrase-initial lengthening of unstressed syllable onsets, as well as intonational-phrase-initial lengthening. In contrast, Wightman *et al.* (1992) do not find a correlation between prosodic boundary strength and syllable onset duration: possible confounding factors in their experiment are discussed in Chapter 4.

Utterance-initial effects

There is evidence in Experiment 2 of an absence of hierarchical lengthening utterance-initially, so that syllable onsets have comparable duration to those found word-medially. Fourakis & Monahan find that the syllable /mə/ is 30 % shorter utterance-initially than when preceded by another syllable; they do not, however, report subsyllabic durations, so the locus of the effect cannot be determined precisely. Fougeron & Keating find a similar effect in reiterant speech with a precise locus: the syllable onset /n/ is shorter for two out of three speakers utterance-initially than utterance-medially; for one speaker, for example, /n/ is 21% shorter utterance-initially. Both of those studies use nasal onset consonants in unstressed syllables; in Experiment 2 the utterance-initial shortening effect is found for the nasal /m/ and the fricative /s/ in stressed syllable onsets. The interpretation of this effect given in Chapter 4 is that the utterance boundary is not signalled durationally because the end of the foregoing silence provides a sufficient cue.

Stressed syllable onsets containing stops or affricates do not demonstrate utterance-initial shortening, however: closure duration cannot be measured acoustically in utterance-initial position and the duration of aspiration or frication which follows the stop release is longer utterance-initially in unaccented words, and shows no effect of position in accented words. The reason for this is not clear. It is unlikely to be because aspiration/frication should be regarded as part of the nucleus, because these phases show the same pattern of word-initial lengthening as the preceding closure duration. As suggested in Chapter 4, it may reflect a more forceful—and thus longer—stop release where this is initiating phonation; the lack of effect in accented syllables may be attributable to the fact that they have more forceful articulation in any position.

Articulatory data could indicate the pattern of consonant closure duration in utterance-initial and utterance-medial position, and it may be that the whole onset gesture shows the same durational pattern seen for *main*, *mend* and *send*, with apparent absence of any hierarchical lengthening effect; alternatively, the whole gesture for /m/ or /s/ could show lengthening utterance-initially. Such findings would run counter to the idea of degrees of hierarchical lengthening as cues for listeners, however: there is no point in varying the duration of a gesture with no acoustic correlate[7].

More research is required using a range of onset phonemes to determine the relative contributions of prosodic hierarchical structure and of constraints upon articulation due to initiation of phonation. The evidence presented here suggests that for some onset consonants at least, the utterance-initial shortening effect previously noted by Fourakis & Monahan (1988) and Fougeron & Keating (1997) may be the result of degrees of domain-initial lengthening associated with levels of a prosodic hierarchy and the special status of the utterance-edge as a boundary that does not require cues other than the fact of speech being initiated.

Different results may be obtained in a long stretch of continuous speech, containing a number of phonological utterances. Here, the breaks in speech between utterances may not be greater than planned mid-utterance pauses at, for example, intonational phrase boundaries. In this case, the utterance should manifest initial lengthening relative to lower-level phrases, because the primary cue to the utterance-initial boundary in isolated sentences, the breaking of hitherto ongoing silence, is not present in continuous speech. The utterance boundary may, of course, be marked by pre-boundary lengthening, but as discussed in the following section, the magnitude of this effect may be no greater utterance-finally than intonational-phrase-finally.

5.3.2 Final lengthening

Utterance-final lengthening

The results of Experiment 2 indicate a number of facts about utterance-final lengthening:

- The locus of lengthening is the word rhyme, although not all subconstituents are affected.

- Lengthening is progressive within the locus.

[7]It may be noted that Fougeron & Keating (1997) find that the articulatory strengthening (in terms of the degree of maximum linguopalatal contact for /n/) of syllable onsets at higher domain edges does not show the same pattern as for duration, where two of the three speakers show utterance-initial shortening relative to lower domain edges. For linguopalatal contact, the maximum at the domain edge is at least as great utterance-initially as at lower levels.

- There is no significant interaction between final lengthening and accentual lengthening.

The question of the number of levels of final lengthening remains unresolved, however.

Researchers such as Oller (1973) and Campbell & Isard (1991) indicate that lengthening in absolute-utterance-final position is localised on the syllable nucleus and syllable coda. Studies such as Nakatani *et al.* (1981) and Turk (1999) find that the locus of *phrase-final* lengthening may extend to the penultimate syllable where the phrase-final syllable is unstressed. Cambier-Langeveld (2000) finds a similar locus for utterance-final lengthening in English, an expected result given that the locus of utterance-final lengthening ought to be at least as extensive as that for phrase-final lengthening. Experiment 2 supports the finding that the locus of lengthening in an absolute-utterance-final stressed syllable is the syllable rhyme; it further shows that when the stressed syllable is in utterance-penultimate position and followed by an unstressed syllable within the word (for example, ... *mason*‖) the lengthening on the stressed syllable is less than absolute-finally, and restricted to the syllable coda[8], and the following unstressed syllable undergoes a large amount of lengthening. Furthermore, a small amount of lengthening is found on the stressed syllable coda in antepenultimate position, when followed by two unstressed syllables within the word (for example, ... *masonry*‖) and both the unstressed syllables show lengthening, the greatest effect being on the syllable in absolute-final position.

The distribution of utterance-final lengthening within unstressed syllables was not measured directly, although from the differences between materials it may be inferred that lengthening is localised on the syllable coda in the penultimate unstressed syllable: thus, all the words that have a medial syllable with a coda consonant (*masonry*, *tendency*, *captaincy*) show an utterance-penultimate lengthening effect on that syllable; the other unstressed syllables are open and do not show utterance-penultimate lengthening except for /tə/ in *spectacle*. It might be assumed that lengthening in absolute-utterance-final unstressed syllables is localised on the syllable rhyme, as for stressed syllables. Subsyllabic measurements from a range of unstressed syllables would be required to resolve this issue.

Experiment 2 also suggests there may be some lengthening of the stressed syllable coda when followed by a word boundary and one or two unstressed syllables before the utterance boundary (for example ... *mace up*‖ and ... *mace again*‖). More research needs to be done to determine if this effect is consistent and is related to utterance

[8] As noted in Chapter 4, the stressed syllable coda /s/ in *mason* could also be regarded as ambisyllabic or as the onset of the following test syllable. The decision to treat it as a coda consonant in Experiment 2 allows consistency in interpretation of the results: final lengthening of syllable onsets is not widely reported.

position, rather than to possible confounds such as stress-adjacent lengthening and segmental effects: these may arise because the phonetic and stress environment of the test syllable is not always matched between utterance-medial and utterance-final contexts. Given that utterance-final function words are not reduced, however, it may be impossible fully to distinguish the potential durational influences in this case. If utterance-final lengthening were consistently found across word boundaries, as just discussed, then the identification of the locus as the word-rhyme would have to be altered to a constituent headed by a primary lexical stress and continuing to the start of the next primary lexical stress, regardless of word boundaries: as discussed in Chapter 2, this corresponds to one interpretation of the Abercrombian foot.

In Experiment 2, the duration of syllables preceding the final stressed syllable is not measured; however, where the stressed syllable is in absolute-final-position, Turk (1999) finds no evidence of phrase-final lengthening on the unstressed syllable that precedes it. This and other studies, such as Oller (1973), strongly suggest that the locus of final lengthening begins with the final stressed syllable. The word-rhyme, as defined in Chapter 3, is the constituent which best describes the locus of utterance-final lengthening; as noted above, not all subconstituents of the word-rhyme are affected by lengthening.

The distribution of lengthening on syllables within the word-rhyme clearly tends to be progressive. Cambier-Langeveld (2000) finds that in words such as *Joseph* and *Macy*, absolute-final unstressed syllables show greater lengthening than the preceding stressed syllables. Similar trends are also apparent within syllables: for example, for intonational-phrase-final lengthening, Turk (1999) reports a progressive lengthening effect within the final syllable rhyme, as do others, such as Berkovits (1994) for Hebrew. In Experiment 2, lengthening is shown to be progressive both within the absolute-final syllable rhyme and within the word-rhyme where the final stressed syllable is in utterance-penultimate or utterance-antepenultimate position. As Oller (1973) indicates, however, there may be variations from this trend due to segmental differences in expandability: for example, final voiced stops may manifest less lengthening than other coda consonants.

Price *et al.* (1991) suggest that final lengthening and accentual lengthening combine subadditively. Cambier-Langeveld (2000) finds this effect for Dutch, but only slight evidence of such an interaction for English: as shown in Table 2.1 in Chapter 2, both syllables in disyllabic words such as *Joseph* and *Macy* show a greater proportion of final lengthening in unaccented keywords and the initial stressed syllable also shows a greater absolute magnitude of lengthening in unaccented words, but the interaction between accent and utterance position is not statistically significant for either syllable. The result of Experiment 2 is quite similar to that reported by Cambier-Langeveld:

there is no significant interaction between pitch accent and utterance position, but there is a tendency towards a slightly greater effect of utterance position in the comparison between unaccented words.

The lack of statistical support indicates that this is not a major effect, such as Cambier-Langeveld finds for Dutch, but the consistency with which is observed in Experiment 2 and by Cambier-Langeveld suggests that it may be genuine. The observation is probably explicable as a ceiling effect: if segments are lengthened by a large amount due to accent, there may be less scope for lengthening due to utterance-final position. This is supported by the evidence from Dutch. Cambier-Langeveld argues that Dutch has less variability in vowel duration than English, and thus the combination of two lengthening effects is greatly sub-additive. The wider range of possible durations for English mean that the durational ceiling is only reached in extreme cases and there is some evidence that there is no ceiling effect phrase-finally. In Experiment 1, the accentual lengthening on the phrase-final syllable in phrases such as *thank ful<u>fil</u>* and *thankful <u>Phil</u>* is greater in absolute magnitude and proportionally comparable to that found on phrase-initial syllables, with the caveat that the syllables compared between the two positions are segmentally different. Similarly, Cummins (1999), looking at a range of speech rates, finds that accentual lengthening and phrase-final lengthening combine approximately additively for most rates; at the fastest rate, there is little evidence of either effect, a fact which may reflect a general change in the nature of the task under unusual conditions rather than segmental limitations on duration as such. Thus, the subadditivity of final lengthening and accentual lengthening may be a particular feature of utterance-final lengthening, which would suggest that it has greater magnitude in English than phrase-final lengthening.

It is not demonstrated that utterance-final lengthening is distinct from intonational-phrase-final lengthening, however. Wightman *et al.* (1992) report no significant difference in the magnitude of final lengthening of the final syllable rhyme between the intonational phrase and the utterance. It is possible that the locus of lengthening may be more extensive utterance-finally; however, Cambier-Langeveld (2000) finds a similar locus for intonational-phrase-final and utterance-final lengthening in Dutch.

There is a lack of evidence in Experiment 2 for phrase-final lengthening below the intonational phrase. There are two alternative interpretations of this observation: firstly, the left-headed keywords in the experimental materials are never followed by phrase boundaries, or so rarely as to have no effect on the mean duration of the pre-boundary nucleus and coda; secondly, phrase boundaries are present following left-headed keywords—as there appear to be preceding right-headed keywords in some cases—but these boundaries do not have durational consequences for the preboundary word. The latter interpretation contradicts studies such as Wightman *et al.* (1992), discussed in Chapter 4, that find evidence of a number of levels of phrasing distinguished by the magnitude of final lengthening; however, Cambier-Langeveld (2000) finds no durational difference in Dutch between word-final syllables and phonological-phrase-final syllables.

Word-final lengthening

One reason given in Chapter 4 for not favouring the interpretation of word-rhyme compression as a domain-final effect is that the nucleus but not the coda is affected: previous studies suggest that domain-final lengthening is progressive—and Experiment 2 finds progressive lengthening utterance-finally—thus at least as much domain-final lengthening would be expected on the coda as on the nucleus. The results for utterance-final lengthening indicate, however, that final lengthening need not be continuous within the locus, but may be localised on discrete subconstituents. For example, the utterance-antepenultimate stressed syllable coda is lengthened, but it is likely that not all subconstituents of the following unstressed syllable undergo lengthening. If the locus of utterance-final lengthening is discrete—affecting certain subconstituents of the word-rhyme—then word-rhyme compression, which affects a different set of subconstituents of the word-rhyme, may also be interpretable as a final lengthening effect.

The strong theoretical reason for favouring the word-final interpretation is that the results of Experiment 2 do not support either word-span or utterance-span compression, so a reinterpretation of word-rhyme compression as a domain-final process—affecting only the nuclei in a word-rhyme locus—would eliminate one type of durational process altogether. This is preferable in theoretical terms to proposing the existence of domain-span compression processes to explain only one set of observations.

Against this domain-final interpretation is the fact that the magnitude of the word-rhyme compression effect may be greater between disyllables and trisyllables than between monosyllables and disyllables: for example, in unaccented words, where there is no influence of the polysyllabic accent effect, the /eɪ/ of /meɪs/ is shortened by 3 ms (4%) between *mace* and *mason* and by 9 ms (12%) between *mason* and *masonry*. The largest difference in a domain-final effect might be expected to be between absolute-final and penultimate position, as is seen for utterance-final lengthening; however, it could be argued that the greatest domain-span compression effect might also be expected between monosyllabic and disyllabic contexts, so the data are not conclusive on this point.

As discussed in Chapter 2, Oller (1973) finds evidence for what he identifies as phrase-final lengthening on the syllable nucleus of the final syllable in reiterant words.

He intends a syntactic definition of "phrase", but the noun phrases containing the measured syllables and the carrier sentences themselves are both rather short (for example, *The* *bababab* *is on the table*) and so a major prosodic boundary may be unlikely in these circumstances. In this case, the finding that the durational effect of word position is localised on the syllable nucleus may be seen as support for the results found in Experiment 2, either as word-rhyme compression or word-final lengthening.

As noted above, the word-final interpretation is preferable theoretically because it does not require the postulation of another type of process. It would be desirable to have an empirically-testable difference in the predictions of the two processes, but the effects of position-in-word and word-rhyme length are necessarily confounded for stressed syllables. For an unstressed syllable, however, it is possible to vary the length of the word-rhyme whilst maintaining the alignment of the measured syllable with the word boundary; similarly, its position with respect to the word boundary can be altered whilst the effects of word-rhyme length are kept constant. These possibilities for experimentally distinguishing the two hypotheses are described in Chapter 6.

5.4 Pitch accent

Experiment 2 provides strong evidence that the locus of accentual lengthening is the word, as suggested by Sluijter (1995) and Turk & White (1999); furthermore, the evidence from Experiment 2 suggest that lengthening is greatest at word edges, and that the variation in the distribution of lengthening between monosyllables, disyllables and trisyllables is responsible for previously-observed polysyllabic shortening. This is the polysyllabic accent effect, as defined in Chapter 4. The distribution of accentual lengthening found in Experiment 2 is here compared with that observed in previous studies, and possible reasons for the differences are discussed.

5.4.1 Accentual lengthening in stressed syllables

Turk & Sawusch (1997) report the durational effect of accent on the constituents of monosyllabic words. Their results are shown in Table 5.1: onset and coda durations are estimated from their graphical data. The mean durations for onset, nucleus and coda from Experiment 2 are reported for comparison: these are the means from the data-set used in Figure 4.12 in Chapter 4, pooled between left-headed and right-headed keywords.

In both data-sets shown in Table 5.1, the onset shows the largest proportional effect and the nucleus shows a somewhat smaller proportional effect[9]. The most notable

[9]The largest data-set available for the same Experiment 2 keywords as used in Table 5.1 includes all the monosyllables in Series B, the fixed word-length and variable utterance-length condition. In that

	Onset		Nucleus		Coda	
Experiment 2	26 ms	27%	16 ms	17%	19 ms	23%
Turk & Sawusch (1997)	34 ms	34 %	43 ms	21%	10 ms	12 %

Table 5.1: Comparison of accentual lengthening in subsyllabic constituents of monosyllables. Mean lengthening is shown on the left in each cell; the proportion of lengthening compared with the unaccented duration is on the right. The Experiment 2 data are from Series C: all keywords except *cap*, *dog*, *part*, *speck* and *port* are included.

difference between two data-sets is that the effect on the coda is rather less in Turk & Sawusch's data. This cannot be attributed to variation in the expandability of segments, because the same phonemes are used in the onset and coda measurements: Turk & Sawusch use phrases such as *bee farm* and *beef arm* and measure the vowel in the first word and the central consonant, word-initially and word-finally. The accent is always within the phrase, either on the first word or the second word, and it is possible that the presence of an accent on an adjacent word could cause some lengthening in the unaccented condition, thereby reducing the apparent accentual lengthening effect. This seems more likely to occur in onset consonants, however: Turk & White (1999) demonstrate that some lengthening may spread rightwards across a word boundary following an accented monosyllable but not leftwards. In Turk & Sawusch's materials, /f/ could be lengthened in *BEE farm*, for example, in which case the magnitude of the lengthening in *bee FARM* would be underestimated. It seems much less likely that /f/ would receive any lengthening in *beef ARM*, however, so the reason for the lower estimate of coda accentual lengthening in Turk & Sawusch remains unclear.

The pattern of stressed syllable accentual lengthening that emerges from consideration of the results of the two experiments together is that: firstly, all constituents show significant lengthening; secondly, the onset usually shows the greatest proportional effect; thirdly, the relative magnitude of lengthening in nucleus and coda is variable.

Turk & Sawusch only report stressed syllable accentual lengthening for monosyllables. As described in Chapter 4, Experiment 2 shows that accentual lengthening in the stressed syllable is attenuated in disyllables and trisyllables, with the most marked reduction being in the onset for right-headed keywords—*port*, *report*, *misreport*—and the coda for left-headed keywords—*fish*, *fissure*, *fisherman*. It appears that, in general, accentual lengthening is greater on constituents close to word-initial or word-final boundaries than on word-medial constituents.

data-set, the mean lengthening is are slightly different. In particular, the nucleus effect is slightly larger and the coda effect is slightly smaller than in Table 5.1: onset 27 ms, 28%; nucleus 18 ms, 20%; coda 17 ms, 21%. As the same keywords are involved in both comparisons, this is apparently a result of random variation.

5.4.2 Accentual lengthening in unstressed syllables

Table 5.2 compares the accentual lengthening found on unstressed syllables in disyllables in Experiment 2 with the results of Sluijter (1995) and Turk & White (1999). In Experiment 2, there is more accentual lengthening on word-initial unstressed syllables, such as /kə/ in *commend*, than on word-final unstressed syllables, such as /ən/ in *mason*, and the effect only attains significance for word-initial syllables[10]. The results of Turk & White and Sluijter suggest that the effect on word-final syllables in Experiment 2 would be significant with more data, but the reasons for the difference in magnitude of the word-initial and word-final effects is not clear.

	Word-initial		Word-final	
Experiment 2	18 ms	15%	10 ms	10 %
Sluijter (1995)	13 ms	9%	38 ms	16 %
Turk & White (1999)	8 ms	5%	21 ms	13%

Table 5.2: Comparison of accentual lengthening in unstressed syllables in disyllables. Mean lengthening is shown on the left in each cell; the proportion of lengthening compared with the unaccented condition is on the right. The Experiment 2 data are from all Series C keywords.

In both previous experiments, the measured syllables are in words spoken in metalinguistic context, for example:

Turk & White Say "THANKFUL Phil", don't say "SHAMELESS Phil".

Sluijter Please produce COMPACT for him again.

where measured unstressed syllables are underlined. In Turk & White's materials, the emphasis in the target word is explicitly contrastive; in Sluijter's materials, it is likely to be contrastive given that the target word is the only difference between successive sentences. It may be that the different distribution of accentual lengthening within the word found in Experiment 2 reflects a difference in realisation between contrastive stress and normal phrasal stress. Certainly, the smaller and less reliable accentual lengthening effect observed in word-final syllables in Experiment 2, compared with the previous studies, could reflect a lesser degree of phrasal stress. The results for word-initial syllables, however, do not correspond with the idea of contrastive stress as a simple amplification of non-contrastive accentual lengthening: there is more lengthening in the non-contrastive case in Experiment 2 than in the contrastive case in

[10]The right-headed keywords—listed in Chapter 4—may suggest the possibility that they could be realised differently in accented and unaccented words, in particular, that the first syllable may be spoken with a full vowel. Such variation was not observed in recordings, however.

Turk & White (1999). Selkirk (2002) suggests that are differences in the prosodic properties of contrastive and presentational focus[11]: in particular, she identifies different types of pitch accent that are typically associated with the two types of focus; it may also be that they manifest different patterns of accentual lengthening.

At least three other factors may affect these differences. Firstly, there may be dialectal differences in accentual lengthening distribution: Turk & White, for example, use Scottish English speakers, who show a similar pattern of accentual lengthening to the American English speakers used by Turk & Sawusch (1997), whereas Experiment 2 uses speakers from northern and southern England. Secondly, there may be segmental differences in expandability, both between the materials in the different studies and between the word-initial and word-final syllables in Experiment 2[12]. Thirdly, the right-headed keywords in Turk & White's study are placed phrase-finally in the carrier sentence (for example: *SAY "thank fulfil", don't SHOUT "thank fulfil"*): as discussed in Section 5.2.1, there is some evidence of a slight attenuation of accentual lengthening utterance-finally. Although there does not appear to be an attenuation of accentual lengthening for the phrase-final stressed syllables, such as /fɪl/ in *thank fulfil*, these tokens would have to be recorded phrase-medially to ensure that there is no attenuation of accentual lengthening in a phrase-final word on syllables such as /fʊl/ in *thank fulfil*.

Table 5.3 compares the accentual lengthening found on unaccented syllables in left-headed trisyllabic words in Experiment 2 with the results of Turk & White (1999). In both experiments, there is less accentual lengthening on word-medial unstressed syllables, such as /ən/ in *masonry*, than on word-final unstressed syllables, such as /rɪ/ in *masonry*. In Experiment 2, however, the effects do attain not significance.

	Word-medial		Word-final	
Experiment 2	4 ms	6%	14 ms	9%
Turk & White (1999)	12 ms	11%	22 ms	14%

Table 5.3: Comparison of accentual lengthening in unstressed syllables in left-headed trisyllables. Mean lengthening is shown on the left in each cell; the proportion of lengthening compared with the unaccented condition is on the right. The Experiment 2 data are from all left-headed keywords in Series C.

[11] What Selkirk calls "presentational" focus may be interpreted as non-contrastive phrasal stress.

[12] The magnitude of accentual lengthening of both unstressed syllables in the left-headed keywords may be underestimated slightly, as noted in Table E.3 in Appendix E. For the unstressed syllable immediately following the test syllable, the test-syllable-final stops in *spec.tre*/*spec.ta.cle* and *cap.tain*/*cap.tain.cy* are not always released, thus the closure durations of the unstressed syllable onset stops are not reliably measurable and so the syllables are measured from the stop release. For the word-final unstressed syllables in trisyllables, the final stops are not always released in *dogmatist* and *partnership* and so the syllable is measured up to the onset of stop closure in these cases.

As just discussed with regard to disyllables, the smaller effect in Experiment 2 could be because the phrasal stress is not contrastive. However, the effects of accent in right-headed trisyllables, not reported previously in English, are large and reliable word-initially: syllables such as /ɹɛ/ in *recommend*, for example, show 21 ms (15%) accentual lengthening. Word-medial syllables, such as /kə/ in *recommend*, show 6 ms (5%) lengthening, comparable with the effect seen for word-medial syllables in left-headed trisyllables.

There may be a particular reason for the strong effect seen on the word-initial syllable in right-headed trisyllables: in most cases, these syllables contain full vowels, and may receive secondary lexical stress. In some cases, as noted in Section 4.3.3, speakers realise these words with the primary stress on the initial syllable; although such cases are discarded from the analysis where detected, it may be that the secondary stressed syllables manifest more accentual lengthening than unstressed syllables. Against this hypothesis is the fact that Turk & White (1999) find that the pattern of accentual lengthening in disyllables containing a secondary stress, such as *kneecap* and *capsize*, is very similar to that found in disyllables containing only one stress, such as *thankful* and *fulfil*.

The pattern of accentual lengthening within the word in Experiment 2 may be seen as mirroring the pattern within the primary stressed syllable in both Experiment 2 and Turk & Sawusch (1997): in most cases, the onset shows the greater degree of lengthening within the primary stressed syllable; similarly, within the word, initial unstressed syllables showing more lengthening than final unstressed syllables.

A picture of accentual lengthening emerges thus: the locus is the whole lexical word, and the greatest lengthening effects are to be found word-initially and word-finally. This reflects the bimodal accent hypothesis presented in Chapter 4—that the locus is the primary stressed syllable and the word-final syllable, and syllables adjacent to the locus also manifest small lengthening effects—but extends it to word-initial syllables also. If this picture is accurate, then the relatively small word-initial effects found in Sluijter (1995) and particularly in Turk & White (1999) need to be accounted for. A number of factors may influence the observed difference: variation in distribution of accentual lengthening between contrastive and non-contrastive accent; dialectal differences in the distribution of accentual lengthening; variation within the locus due to segmental differences; and differences in the pattern of accentual lengthening phrase-finally and phrase-medially.

Accentual lengthening beyond the word

It is not clear how the finding reported in Turk & White (1999) of a small amount of lengthening on the word following an accented monosyllable should be accommodated within the view of the word as the locus of accentual lengthening. Turk & White show, for example, that the duration of /fʊl/ is about 4% greater in *THANK fulfil* than in *thank fulfil*. This effect is much smaller than most of the accentual lengthening effects observed within the word: for example, /fʊl/ is about 13% longer in *THANKful Phil* than in *thankful Phil*. There is no evidence of a similar effect in words preceding an accented monosyllable: for example, /fʊl/ is not significantly longer in *thankful PHIL* than in *thankful Phil*.

Turk & White suggest that the "residual" lengthening in an unaccented word following an accented monosyllable indicates that word boundaries attenuate lengthening, rather than block it altogether. They further suggest that the effect is not observed in a word preceding an accented monosyllable because the left edge of the accented syllable also attenuates lengthening: the combined effects of the syllable boundary and the word boundary effectively eliminate any lengthening on the preceding word. The results discussed above for Experiment 2 suggest, however, that accentual lengthening is not attenuated at the left edge of the accented syllable, at least for non-contrastive pitch accent.

The existence of residual lengthening beyond the boundaries of the accented words needs further investigation, as the effect has only been observed following accented monosyllables in metalinguistic context. If confirmed—for example, by measuring the duration of syllables adjacent to accented monosyllables and polysyllables in more normal speech contexts—residual lengthening would indicate that the picture, described above, of the word as the locus of accentual lengthening may need some adjustment[13]

[13]There does not appear to be any residual lengthening effect beyond the locus associated with domain-edge processes. Indeed, there is some evidence, discussed in Section 5.5, that domain-edge lengthening may occasion *shortening* in segments near the locus. There is also a general difference in the nature of the two lengthening processes: accentual lengthening appears to have a gradient distribution within the word, with all subconstituents affected and greater effects nearer the constituent boundaries; utterance-final lengthening, as an example of a domain-edge process, is confined to certain phonologically-defined subconstituents of the word-rhyme, and there is no evidence of significant lengthening beyond those subconstituents. Thus, accentual lengthening is diffuse and gradient within the locus, whereas final lengthening, and also initial lengthening, are localised. It is possible that this difference may relate to the different articulatory mechanisms through which lengthening is implemented. Using the task-dynamic approach, Beckman *et al.* (1992) say that the closing gesture in an accented syllable is phased later, relative to the opening gesture: this means that the articulators move further in the opening gesture, before this is truncated by the onset of the closing gesture. In contrast, final lengthening is associated with an increase in gestural stiffness of the closing gesture of the syllable; thus, the articulators move to the same target as in a non-final syllable, but take a longer time to get there. Details of how the articulatory implementation of lengthening may relate to its distribution are beyond the scope of this dissertation; if there is a link between articulatory strategy and distribution, however, it may be that apparently different durational effects beyond the locus—compensatory shortening for domain-edge processes, residual lengthening for pitch accent—are a reflection of these distinct articulatory strategies for lengthening within the locus.

5.4.3 The domain of accentual lengthening

The question of the domain of accentual lengthening, discussed in Chapter 2, remains unclear. In Experiment 2, the accent on the keyword is usually—though not always—the last in the utterance[14]; in such cases, it is, by definition, the nuclear accent. If accentual lengthening only affected nuclear-accented words, then the domain of accentual lengthening could be said to be the intonational phrase. It seems highly probable, however, that prenuclear pitch accents also cause lengthening: for example, Silverman & Pierrehumbert (1990:103) provide data on tonal alignment which support a "parallel phonological and phonetic treatment of nuclear and prenuclear pitch accents", also suggesting that the magnitude of lengthening of the accented syllable may be greater for nuclear accents than prenuclear accents[15]. As discussed above, the distribution of lengthening within the accented word may be different for contrastive and non-contrastive nuclear accent. Thus, it may be that there are a number of levels of pitch accent distinguished by the magnitude of accentual lengthening and its distribution within the locus. If there is a hierarchy of accentual lengthening, the different levels might be associated with different domains.

One theoretically attractive approach would be to identify each prosodic constituent with a particular level of prominence: thus, each constituent would be associated with a certain degree of initial and, at least in some cases, final lengthening, and also with a particular degree of lengthening within the accented word, which is the head of the constituent. Beckman & Edwards (1990, 1994) suggest a link between hierarchies of constituents and prominences; however, although the word may be associated with lexical stress and the intonational phrase with nuclear pitch accent, it is not clear what prominence should be associated with other levels such as the utterance or the phonological phrase, and as discussed in Chapter 2, no constituent has been identified which contains only a single prenuclear pitch accent.

Selkirk (2002) suggests that domain of contrastive focus may in fact be the intonational phrase and the domain of what she calls "presentational focus"—which may be equated with non-contrastive pitch accent—is the major phonological phrase. This may be useful from the point of view of timing if it is shown, as discussed above, that contrastive and non-contrastive accent are associated with differing degrees and distribution of lengthening; however, not all intonational phrases contain a contrastive focus.

[14] Utterances in which subjects realised an additional accent following the accented keyword were not excluded from the analysis in this experiment. Such utterances made up 5.8% of the total. In some of these utterances, the accent on the keyword is consequently prenuclear; in other cases, the following accent is in a separate intonational phrase.

[15] Silverman & Pierrehumbert (1990) do not provide data on the magnitude of prenuclear and nuclear accentual lengthening.

To say that a phrase is a domain of pitch accent, of whatever type, does not localise the durational effect of accent in the way that a similar statement regarding domain-initial or domain-final processes allows the locus of lengthening to be identified: the placement of accents within an intonational phrase is influenced by non-structural factors, in particular, the pragmatic interpretation according to context. Thus, the theoretical utility of the concept of the domain is more evident with regard to the description of domain-edge processes than to accentual lengthening. The results of Experiment 2 indicate that there are at least two or three levels of domain-initial durational effects, all having a syllable onset locus; there may also be at least two levels of domain-final lengthening, affecting different constituents of the word-rhyme. In these cases, to specify the domain of the initial or final effect provides important information as to its relative magnitude and, for final lengthening, its locus. For the purposes of predicting durational variation associated with pitch accent, however, the important information appears to be that there is an accent on a word of particular length and syllable structure. It remains to be seen whether the type of accent should also be specified, for example, if prenuclear accentual lengthening is of lesser magnitude than nuclear accentual lengthening, or if the distribution of contrastive accentual lengthening is different within the locus to that of nuclear accent.

5.5 Compensatory effects

The term "compensatory" is defined in Chapter 4 as characterising durational effects in which some constituent shortens or lengthens, apparently as a result of an opposite durational variation elsewhere, whatever the underlying interpretation for the effect. In Experiment 2, there are three observations which may be interpreted as compensatory shortening: following word-initial lengthening; following phrase-initial lengthening; and preceding utterance-final lengthening. There is no strong evidence for relating these observations directly to constituent edges or to constituent length, or to any other source of suprasyllabic durational variation. In two out of three cases, they arise from comparisons between segmental durations in contexts that are not phonetically balanced, which means that durational effects arising from the adjacency of segments cannot be ruled out.

The observations suggesting compensatory shortening are reviewed here and possible interpretations are suggested. The existence of compensatory processes remains to be confirmed by experiment, however: suggestions for such research are advanced in Chapter 6.

5.5.1 Evidence for compensatory shortening

Word-initial lengthening and compensatory shortening

In the word-edge and word-span analysis presented in Chapter 4, the test syllable nucleus for right-headed keywords is slightly shorter in monosyllables compared with disyllables. For example, in:

Series A.1 JOHN saw Jessica **mend** it AGAIN.

Series A.2 JOHN saw Jessie **commend** it AGAIN.

the /ɛ/ in /mɛnd/ is slightly shorter in *mend* than in *commend*, particularly in the accented condition. This effect, which may be seen in Figure 4.4 in Chapter 4, is not significant, but runs counter to the trend for the nucleus in the accented disyllable vs trisyllable comparison, and also for other subsyllabic constituents in the accented condition: in all other cases in right-headed and left-headed keywords, the duration of test syllable subconstituents is greater in accented monosyllables than in accented disyllables. The syllable onset undergoes a large amount of lengthening in the monosyllable compared with the disyllable—31ms (32%) in accented words, 17ms (21%) in unaccented words—and the shortening of the nucleus in the same comparison suggests a compensatory relationship.

Compensatory shortening of the syllable nucleus following word-initial lengthening has not previously been noted. This is to some extent because some studies finding evidence for word-initial lengthening report only onset duration. Turk & Shattuck-Hufnagel (2000), however, report the duration for American English of the syllable centre as well and, as discussed in Section 5.2, they do not observe a consistent word-length effect in the syllable centre in right-headed keywords: in particular, in the accented condition, the /oʊ/ in *pose* is only slightly longer than in *oppose*, and this difference is not significant by Subjects. The polysyllabic accent hypothesis would predict a larger difference in this case, arising from the attenuation of accentual lengthening of /oʊ/ in the disyllable. It may be that the compensatory shortening following the large word-initial lengthening of /p/ in *pose* serves to mask the greater accentual lengthening in the monosyllable[16].

[16]In some of their materials, Turk & Shattuck-Hufnagel (2000) have complex stressed syllable onsets, such as /kw/ in *choir* vs *acquire*. As discussed in Chapter 3, only the first consonant in such onsets are included in the initial consonant measure, and the other consonant is measured as part of the "syllable centre", in order to facilitate accurate segmentation. It is to be expected that both parts of a complex onset would show initial lengthening; thus, in three out of 11 cases, the durational effect on the syllable centre may overestimate the lengthening of the nucleus in monosyllables compared with disyllables, because the syllable centre includes an onset consonant. This could make any underlying compensatory shortening less apparent.

Phrase-initial lengthening and compensatory shortening

In the utterance-span analysis presented in Chapter 4, there is a comparison of test syllable subconstituent duration between Series B.3 and Series D.3 utterances for right-headed keywords, such as:

Series B.3 I MADE Peter Burgundy **send** to them ALL.

Series D.3 Will you **send** it TODAY please.

There is evidence of lengthening of test syllable onsets in Series B compared with Series D, suggesting that in some cases the test syllable may be preceded by a phrase boundary (such as a phonological phrase boundary). There is also some evidence of shortening of the coda for certain keywords in the utterance-medial (Series B) sentences, although this result is not significant across all keywords. These two observations suggest that there may be a compensatory relationship between onset duration and coda duration[17].

Utterance-final lengthening and compensatory shortening

In Chapter 4, there is also a comparison between Series B.3 and Series D.3 for left-headed keywords, such as:

Series B.3 I SAW the **mace** unreclaimed by them AGAIN.

Series D.3 Albert THREW the **mace** again.

There is evidence of lengthening of the coda in Series D.3, which may be due to the utterance-antepenultimate position of the syllable. There is also evidence of lengthening of the onset in Series B.3; the structure of the sentences suggests that this is unlikely to be phrase-initial lengthening. The explanation may be compensatory shortening of the onset in Series D.3 due to lengthening of the coda. Compensatory effects are not observed in Experiment 2 in other analyses of utterance-final lengthening, for example, where the test syllable is in absolute-final position, the rhyme is lengthened and there is no durational effect on the onset.

Compensatory shortening of the onset in response to final lengthening of the syllable rhyme has not been reported previously; Turk (1999), however, reports shortening

[17] An alternative interpretation is that the coda is shortened due to domain-span compression in the longer Series B utterances. This is unlikely, however, because the effect on the coda is preserved in a comparison using Series B.1 sentences, which are structurally similar to the Series B.3 sentences, but are two syllables shorter and little longer than the Series D.3 sentences. There is a possibility of segmental effects confounding the comparison, however. For example, in the *send* sentences illustrated above, the coda is followed by a consonant in the B.3 sentence, but by a vowel in the D.3 sentence: it is not clear if shortening in consonant clusters occurs between, as well as within, syllables; if it does, then this represents an alternative explanation for this particular comparison.

for one of two speakers on the rhyme of the unstressed syllable preceding the phrase-final stressed syllable.

5.5.2 Interpretations of compensatory shortening

One possible interpretation of compensatory effects is described by de Jong (1991:7): "given an actual situation in which the ratio between consonant and vowel duration is held constant, variation in the placement of the boundary mark between consonant and vowel could, by itself, generate a negative correlation." Thus, a compensatory effect could arise from random fluctuation in measurement, but this is unlikely to underlie the observations in Chapter 4: the variation in onset duration associated with word-initial lengthening is large and is systematic rather than random; furthermore, the compensatory effects associated with phrase-initial lengthening and utterance-final lengthening are between non-adjacent segments.

It is possible, however, that compensatory shortening may have a similar explanation in terms of articulation: structurally-determined lengthening, such as that observed at domain edges, may cause a change in the articulation of the segments within the locus of lengthening which affects the acoustic realisation of the boundary between these segments and adjoining segments, and thus affects the identification of that boundary when segmental durations are measured. This could only explain the compensatory shortening associated with word-initial lengthening, where the lengthened and shortened segments are adjacent. This effect is, however, the most robust of the three possible compensatory effects, as the measured segments have a consistent phonetic environment[18].

Alternatively, the articulatory change within the locus may cause changes in the articulation of adjoining segments which could have durational consequences for those segments. In their task-dynamic approach discussed in Section 5.4, Beckman & Edwards (1994) suggest that accentual lengthening is associated with changes in gestural phasing and that final lengthening is associated with changes in gestural stiffness. It is possible that these mechanisms could have consequences for the articulation and duration of adjacent segments. It not clear whether this alternative articulatory interpretation could explain compensatory effects between segments which are not adjacent—for example, variation in articulation of the onset could have articulatory consequences on the coda, which could affect its duration—and details of the mechanisms which could underlie such an effect are beyond the scope of this dissertation. One question for future investigation is whether such effects would be expected to

[18] Although not statistically significant, the effect associated with word-initial lengthening represents, for accented words at least, a small shortening effect where there are good grounds for expecting a lengthening effect.

be symmetrical: that is, whether changes in the articulation of a lengthened segment could affect how segments preceding it are articulated, as well as those following it.

An alternative interpretation of compensatory processes is that they arise from a domain-span process at the syllable level. This suggests that syllable durations are derived initially—taking into account prominence, position, stress distribution etc.—and the durations of subsyllabic segments are adjusted subsequently. This is termed "syllabic mediation" by van Santen (1997), who expresses the concept thus: "The duration of a segment depends mostly on the (pre-computed) syllable duration and the segment's identity [. . . When] two contexts produce the same overall duration of a given syllable, then all segments should also have the same duration in the same contexts" (van Santen 1997:237)[19]. Thus, if syllables have pre-calculated durations, lengthening in one part of the syllable—for example, word-initial lengthening of the onset—would cause a similar amount of shortening elsewhere; segmental durations outside the syllable should not be affected by lengthening effects within the syllable.

All of the observations of compensatory shortening reported above are compatible with the syllable-span hypothesis in terms of direction, but not in terms of magnitude. In particular, the compensatory shortening of the nucleus is not commensurate with the word-initial lengthening of the onset: total syllable duration is greater in word-initial position than in word-medial position. Furthermore, if they exist, syllable-span effects should be ubiquitous; as noted above, however, the onset of the utterance-absolute-final syllable shows no compensatory shortening in Experiment 2, despite the very large lengthening effect in the syllable rhyme.

Another interpretation of compensatory shortening is predicated on the existence of domain-span processes in higher-level constituents: thus shortening would occur in response to a lengthening effect in order to maintain the duration of some superordinate unit. The evidence from Experiment 2 indicates that such processes do not occur at the word level or the utterance level; in addition, evidence discussed in Chapter 2 suggests that such processes do not occur in prominence-delimited units such as the cross-word foot.

Compensatory shortening: summary

The most likely interpretation of compensatory effects appears to be that there is some articulatory adjustment in the locus of lengthening that affects how the boundary between the locus and the following segment is realised: this could account for shortening of the nucleus following word-initial lengthening. Whether a more complex articulatory explanation is required to explain the other observations of compensatory

[19]Van Santen rejects the syllabic mediation hypothesis, partly because of the localisation within the syllable of factors such as phrase-finality, word-initial position and lexical stress.

shortening is uncertain: the effects associated with apparent phrase-initial lengthening and utterance-final lengthening arise from comparisons between materials which are not phonetically balanced, and may simply be anomalous results.

5.6 Summary

The evidence reviewed in this chapter suggests that the primary processes associated with suprasyllabic linguistic structure are domain-edge lengthening and accentual lengthening, both of which are associated with phonologically-defined loci. The number of hierarchical levels of constituent structure associated with domain-initial and domain-final lengthening processes remains to be established, as does the number of levels of phrasal stress differentiated by variations in the magnitude and distribution of accentual lengthening.

There is evidence for one type of domain-span process—word-rhyme compression—but this may also be interpreted as a word-final process. The existence of domain-span processes at the word level and the utterance level is not supported in Experiment 2, and examination of previous experiments suggests that support for these domain-span processes may be reinterpreted, although in a few cases results suggest a different picture from that indicated by Experiment 2.

In Chapter 6, a model of suprasyllabic speech timing is outlined which incorporates these processes, and the principles encapsulated by this model are contrasted with other views of suprasyllabic speech timing processes. Suggestions are made for further research which explores questions about the nature of domain-edge lengthening and accentual lengthening raised in the present chapter.

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 satisfactory 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 possibility 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 pre-calculation 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 *conˈtent* 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 experimental 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.

Bibliography

ABERCROMBIE, D. 1965. A phonetician's view of verse structure. In *Studies in Phonetics and Linguistics*, 16–25. London: Oxford University Press.

ALBROW, K.H., 1968. The rhythm and intonation of spoken English. Programme in Linguistics and English Teaching Paper No. 9, University College London.

ARVANITI, A., D.R. LADD, & I. MENNEN. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26.3–25.

BARNWELL, T.P. 1971. An algorithm for segment durations in a reading machine context. Technical Report 479, Research Laboratory of Electronics, Massachusetts Institute of Technology.

BEACH, C.M. 1991. The interpretation of prosodic patterns at points of syntactic ambiguity: Evidence for cue trading relations. *Journal of Memory and Language* 30.644–663.

BECKMAN, M.E., & J. EDWARDS. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In (Kingston & Beckman 1990), 152–178.

——, & ——. 1992. Intonational categories and the articulatory control of duration. In *Speech Perception, Production and Linguistic Structure*, ed. by Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sapisaka, 356–375. Tokyo: Ohmsha.

——, & ——. 1994. Articulatory evidence for differentiating stress categories. In (Keating 1994), 7–33.

——, ——, & J. FLETCHER. 1992. Prosodic structure and tempo in a sonority model of articulatory dynamics. In (Docherty & Ladd 1992), 68–86.

——, & J.B. PIERREHUMBERT. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3.255–309.

BEL, B., & I. MARLIEN (eds.) 2002. *Speech Prosody 2002: Aix-en-Provence*. Conference proceedings: internet publication.

BERKOVITS, R. 1993a. Progressive utterance-final-lengthening in syllables with final fricatives. *Language and Speech* 36.89–98.

—— 1993b. Utterance-final lengthening and the duration of final-stop closures. *Journal of Phonetics* 21.479–489.

—— 1994. Durational effects in final lengthening, gapping, and contrastive stress. *Language and Speech* 37.237–250.

BOLINGER, D. 1965. *Forms of English: Accent, Morpheme, Order*. Cambridge, Massachusetts: Harvard University Press.

—— 1981. Two kinds of vowels, two kinds of rhythm. Technical report, Indiana University Linguistics Club.

BYRD, D. 2000. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57.3–16.

CAMBIER-LANGEVELD, T., 2000. *Temporal Marking of Accents and Boundaries*. University of Amsterdam dissertation.

CAMPBELL, W.N., & S.D. ISARD. 1991. Segment durations in a syllable frame. *Journal of Phonetics* 19.37–47.

CHOMSKY, N., & M. HALLE. 1968. *The Sound Pattern of English*. New York: Harper and Row.

CLARK, HERBERT H. 1973. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behaviour* 12.335–359.

CLASSE, A. 1939. *The Rhythm of English Prose*. Oxford: Blackwell.

COKER, C.H., N. UMEDA, & C.P. BROWMAN. 1973. Automatic synthesis from ordinary English text. *Institute of Electrical and Electronics Engineers: Transactions of Audio Electroacoustics AU-21.* 293–297.

CONNELL, B., & A. ARVANITI (eds.) 1995. *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. Cambridge: Cambridge University Press.

COOPER, A.M. 1991. Laryngeal and oral gestures in English /p, t, k/. In *Proceedings of the XIIth International Congress of Phonetic Sciences, Aix-en-Provence*, volume 2, 50–53.

COOPER, W.E., & S.J. EADY. 1986. Metrical phonology in speech production. *Journal of Memory and Language* 25.369–384.

——, & J. PACCIA-COOPER. 1980. *Syntax and Speech*. Cambridge, Massachusetts: Harvard University Press.

——, & J.M. SORENSON. 1981. *Fundamental Frequency in Sentence Production*. New York: Springer.

COUPER-KUHLEN, E. 1986. *An Introduction to English Prosody*. London: Edward Arnold.

CUMMINS, F. 1999. Some lengthening factors in English speech combine additively at most rates. *Journal of the Acoustical Society of America* 105.476–480.

——, & R. PORT. 1998. Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26.145–171.

CUTLER, A. 1990. From performance to phonology: Comments on Beckman and Edwards's paper. In (Kingston & Beckman 1990), 208–214.

——, & D.M. CARTER. 1987. The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language* 2.133–142.

DANKOVIČOVÁ, J. 1997. The domain of articulation rate variation in Czech. *Journal of Phonetics* 25.287–312.

DAUER, R.M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11.51–62.

DE JONG, K.J. 1991. An articulatory study of consonant-induced vowel duration changes in English. *Phonetica* 48.1–17.

DILLEY, L., S. SHATTUCK-HUFNAGEL, & M. OSTENDORF. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24.423–444.

DOCHERTY, G., & D.R. LADD (eds.) 1992. *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press.

EDWARDS, J., M.E. BECKMAN, & J. FLETCHER. 1991. The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America* 89.369–382.

FANT, G., A. KRUCKENBERG, & L. NORD. 1991. Durational correlates of stress in Swedish, French and English. *Journal of Phonetics* 19.351–365.

FAURE, G., D. J. HIRST, & M. CHAFCOULOFF. 1980. Rhythm in English: Isochronism, pitch, and perceived stress. In *The Melody of Language*, ed. by L. R. Waugh & C. H. van Schooneveld, 71–79. Baltimore: University Park Press.

FERREIRA, F. 1991. Creation of prosody during sentence production. *Psychological Review* 100.233–253.

FOUGERON, C., & P.A. KEATING. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101.3728–3740.

FOURAKIS, M., & C.B. MONAHAN. 1988. Effects of metrical foot structure on syllable timing. *Language and Speech* 31.283–306.

FOWLER, C. 1981. A relationship between coarticulation and compensatory shortening. *Phonetica* 38.35–50.

FOWLER, C.A. 1990. Lengthenings and the nature of prosodic constituency: Comments on Beckman and Edwards's paper. In (Kingston & Beckman 1990), 201–207.

GAITENBY, JANE H. 1965. The elastic word. Technical Report SR–2, Haskins Laboratories, New York.

GEE, J.P., & F. GROSJEAN. 1983. Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology* 15.411–458.

GOLDHOR, R., 1976. Sentential determinates of duration in speech. Massachussetts Institute of Technology Masters dissertation.

GOLDMAN-EISLER, F. 1968. *Psycholinguistics: Experiments in Spontaneous Speech*. New York: Academic.

GRABE, E., & P. WARREN. 1995. Stress shift: Do speakers do it or do listeners hear it? In (Connell & Arvaniti 1995), 95–110.

——, ——, & F. NOLAN. 1994. Resolving category ambiguities—evidence from stress shift. *Speech Communication* 15.101–114.

GUSSENHOVEN, C. 1992. Intonational phrasing and the prosodic hierarchy. In *Phonologica 1988*, ed. by W. Dressler, H.C. Luschützky, O.E. Pfeiffer, & J.R. Rennison, 89–99. Cambridge: Cambridge University Press.

HALLE, M., & J.-R. VERGNAUD. 1987. *An Essay on Stress*. Cambridge: Massachussetts Institute of Technology Press.

HARRIS, M.S., & N. UMEDA. 1974. Effect of speaking mode on temporal factors in speech. *Journal of the Acoustical Society of America* 56.1016–1018.

HAYES, B. 1983. A grid-based theory of English meter. *Linguistic Inquiry* 14.357–393.

—— 1989. The prosodic hierarchy in meter. In *Phonetics and Phonology Volume 1: Rhythm and Meter*, ed. by P. Kiparsky & G. Youmans, 201–260. San Diego: Academic Press.

HERMENT-DUJARDIN, S., & D. HIRST. 2002. Emphasis in English: a perceptual study based on modified synthetic speech. In (Bel & Marlien 2002). Conference proceedings: internet publication.

HUGGINS, A.W.F. 1975. On isochrony and syntax. In *Auditory Analysis and Perception of Speech*, ed. by G. Fant & M.A.A. Tatham, 455–464. Orlando: Academic Press.

ITO, J., & R.-A. MESTER, 1992. Weak layering and word binarity. University of Santa Cruz manuscript.

JONES, D. 1942–43. Chronemes and tonemes. *Acta Linguistica* 3.1–10.

KEATING, P.A. (ed.) 1994. *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*. Cambridge: Cambridge University Press.

KINGSTON, J., & M.E. BECKMAN (eds.) 1990. *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press.

KLATT, D.H. 1973. Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America* 54.1102–1104.

—— 1974. On predicting the duration of the phonetic segment [s] in English. *Journal of Speech and Hearing Research* 17.51–63.

—— 1975. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3.129–140.

—— 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59.1208–1220.

—— 1979. Synthesis by rule of segmental durations in English sentences. In *Frontiers in Speech Communication Research*, ed. by B. Lindblom & S. Ohman, 287–300. New York: Academic.

——, & W.E. COOPER. 1975. Perception of segment duration in sentence context. In *Structure and Process in Speech Perception*, ed. by A. Cohen & S. Nooteboom. Heidelberg: Springer-Verlag.

LADD, D.R. 1986. Intonational phrasing: The case for recursive structure. *Phonology Yearbook* 3.311–340.

—— 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

——, & N. CAMPBELL. 1991. Theories of prosodic structure: Evidence from syllable duration. In *Proceedings of the XIIth International Congress of Phonetic Sciences, Aix-en-Provence*, volume 2, 290–293.

——, D. FAULKNER, H. FAULKNER, & A. SCHEPMAN. 1999. Constant ”segmental anchoring” of f-0 movements under changes in speech rate. *Journal of the Acoustical Society of America* 106.1543–1554.

——, I. MENNEN, & A. SCHEPMAN. 2000. Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America* 107.2685–2696.

LADEFOGED, P. 1993. *A Course in Phonetics*. Fort Worth: Harcourt Brace.

LEA, W.A. 1974. Prosodic aids to speech recognition: IV. A general stragey for prosodically-guided speech understanding. Technical Report PX10791, Sperry Univac, DSD, St Paul, Minnesota.

LEHISTE, I. 1972. The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America* 51.2018–2024.

—— 1973. Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America* 54.1228–1234.

——, 1974. Interaction between test word duration and the length of utterance. Ohio State University Working Papers in Linguistics 17.

—— 1975. The role of temporal factors in the establishment of linguistic units and boundaries. In *Phonologica 1972*, ed. by W.U. Dressler & F.V. Mares, 115–122. Munich-Salzburg: Verlag.

—— 1977. Isochrony reconsidered. *Journal of Phonetics* 5.253–263.

——, J.P. OLIVE, & L.A. STREETER. 1976. The role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America* 60.1199–1202.

LIN, J.-W., 1993. Lexical government and prosodic domains. University of Massachusetts at Amherst unpublished manuscript.

LINDBLOM, B., & K. RAPP. 1972. Re-examining the compensatory adjustment of vowel duration in swedish words. Technical report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm.

LINDBLOM, B.E.F. 1964. A note on segment duration in Swedish polysyllables. Quarterly Progress and Status Report 1/1964, Speech Transmission Laboratory, Stockholm.

—— 1968. Temporal organisation of syllable production. Quarterly Progress and Status Report 2–3/1968, Speech Transmission Laboratory, Stockholm.

LOCAL, J., & R. OGDEN. 1996. A model of timing for nonsegmental phonological structure. In *Progress in Speech Synthesis*, ed. by J.P.H. van Santen, J.P. Olive, R.W. Sproat, & J. Hirschberg, chapter 9, 109–121. New York: Springer-Verlag.

MARSLEN-WILSON, W.D., L.K. TYLER, P. WARREN, P. GRENIER, & C.S. LEE. 1992. Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology* 45.73–87.

NAKATANI, L.H., K.D. O'CONNOR, & C.H. ASTON. 1981. Prosodic aspects of American English speech rhythm. *Phonetica* 38.84–106.

NESPOR, M., & I. VOGEL. 1986. *Prosodic Phonology*. Dordrecht: Foris Publications.

NOLAN, F., 1998. Phonological representation and phonetic interpretation in intonation analysis. Unpublished manuscript of conference paper presented at Laboratory Phonology VI, York.

NOOTEBOOM, S.G., 1972. *Production and Perception of Vowel Duration: A Study of Durational Properties in Dutch*. University of Utrecht dissertation.

O' MALLEY, M.H., D.R. KLOKER, & B. DARA-ABRAMS. 1973. Recovering parentheses from spoken algebriac expressions. *Institute of Electrical and Electronics Engineers: Transactions on Audio and Electro-acoustics AU-21.* 217–220.

OLLER, D.K. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54.1235–1247.

PETERSON, G., & I. LEHISTE. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32.693–703.

PIKE, K.L. 1945. *The Intonation of American English*. Ann Arbor: University of Michigan Press.

PORT, R.F. 1981. Linguistic timing factors in combination. *Journal of the Acoustical Society of America* 69.262–274.

PRICE, P.J., M. OSTENDORF, S. SHATTUCK-HUFNAGEL, & C. FONG. 1991. The use of prosody in synactic disambiguation. *Journal of the Acoustical Society of America* 90.2956–2970.

——, M. OSTENDORF, S. SHATTUCK-HUFNAGEL, & N. VEILLEUX. 1988. A methodology for analyzing prosody. *Journal of the Acoustical Society of America* 84.Supplement 1 S99.

RAAIJMAKERS, J.G.W., J.M.C SCHRIJNEMAKERS, & F. GREMMEN. 1999. How to deal with 'the language-as-fixed-effect fallacy': Common misconceptions and alternative solutions. *Journal of Memory and Language* 41.416–426.

RAKERD, B., W. SENNETT, & C.A. FOWLER. 1987. Domain-final lengthening and foot-level shortening in spoken English. *Phonetica* 44.147–155.

SCOTT, D.R. 1982. Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America* 71.996–1007.

——, S.D. ISARD, & B. DE BOYSSON-BARDIES. 1985. Perceptual isochrony in English and French. *Journal of Phonetics* 13.155–162.

SELKIRK, E.O. 1978. On prosodic structure and its relation to syntactic structure. In *Nordic Prosody II*, ed. by T. Fretheim, 111–140. Trondheim: Tapir.

—— 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Massachusetts: Massachussetts Institute of Technology Press.

—— 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3.371–405.

—— 1990. On the nature of prosodic constituency: Comments on Beckman and Edwards's paper. In (Kingston & Beckman 1990), 179–200.

—— 1996. The prosodic structure of function words. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, ed. by J.L. Morgan & K. Demuth, 187–213. Mahwah, New Jersey: Lawrence Erlbaum.

—— 2002. Contrastive FOCUS vs. presentational *focus*: Prosodic evidence from right node raising in English. In (Bel & Marlien 2002). Conference proceedings: internet publication.

SHATTUCK-HUFNAGEL, S., M. OSTENDORF, & K. ROSS. 1994. Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics* 22.357–388.

——, & A.E. TURK. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25.193–247.

SHEN, Y., & G.G. PETERSON. 1962. Isochronism in English. Studies in Linguistics, Occasional Papers 9, University of Buffalo.

SILVERMAN, K., M. BECKMAN, J. PITRELLI, M. OSTENDORF, C. WIGHTMAN, P. PRICE, J. PIERREHUMBERT, & J. HIRSCHBERG. 1992. ToBI: A standard for labeling English prosody. In *Proceedings of the International Conference on Spoken Language Processing, Banff, II*, volume 2, 867–870.

SILVERMAN, K.E.A., & J.B. PIERREHUMBERT. 1990. The timing of prenuclear high accents in English. In (Kingston & Beckman 1990), 72–113.

SLUIJTER, A.M.C., 1995. *Phonetic Correlates of Stress and Accent*. University of Leiden dissertation.

——, & V.J. VAN HEUVEN. 1995. Effects of focus distribution, pitch accent and lexical stress on the temporal organisation of syllables in Dutch. *Phonetica* 52.71–89.

SNOW, D. 1994. Phrase-final syllable lengthening and intonation in early child speech. *Journal of Speech and Hearing Research* 37.831–840.

TURK, A.E. 1999. Structural influences on boundary-related lengthening in English. In *Proceedings of the 14th International Congress of Phonetic Sciences*, volume 2, 1117–1120.

——, & J.R. SAWUSCH. 1996. The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America* 99.3782–3790.

——, & ——. 1997. The domain of accentual lengthening in American English. *Journal of Phonetics* 25.25–41.

——, & S. SHATTUCK-HUFNAGEL. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics* 28.397–440.

——, & L.S. WHITE. 1999. Structural influences on accentual lengthening in English. *Journal of Phonetics* 27.171–206.

VAN LANCKER, D., J. KREIMAN, & D. BOLINGER. 1988. Anticipatory lengthening. *Journal of Phonetics* 16.339–347.

VAN SANTEN, J.P.H. 1997. Segmental duration and speech timing. In *Computing Prosody: Computational Models for Processing Spontaneous Speech*, ed. by Y. Sagisaka, N. Campbell, & N. Higuchi, 225–249. New York: Springer-Verlag.

VOGEL, I. 1992. Comments on chapters 3 and 4. In (Docherty & Ladd 1992), 124–127.

WHITE, L.S., 1993. The contribution of prosodic cues to boundaries in sentence processing by listeners. University of Cambridge MPhil dissertation.

WIGHTMAN, C.W. 2002. ToBI or not ToBI? In (Bel & Marlien 2002). Conference proceedings: internet publication.

——, S. SHATTUCK-HUFNAGEL, M. OSTENDORF, & P. PRICE. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91.1707–1717.

WOODROW, H. 1951. Time perception. In *Handbook of Experimental Psychology*, ed. by S.S. Stevens, 1224–1236. New York: Wiley.

Appendix A

Speech segmentation criteria

The location of segment boundaries in Experiment 1 and Experiment 2o is determined by inspection of speech waveforms and spectrograms. The criteria used in making such judgements are given here. For consistency, where alternatives are available, a single criterion is applied to all examples of a particular test item for a particular speaker. The parameters of application of certain criteria, for example, the choice of which formant break to associate with a given stop closure, is applied with similar consistency. Where labels are associated with the start or end of pitch periods, they are placed at the point of zero crossing on the waveform.

Stop closure: The end of pitch period before a significant drop in waveform amplitude or at a break in a particular formant.

- Following a fricative, the drop in spectrogram energy intensity in a particular frequency range may be used.

Glottalised stop closure: The end of pitch period preceding a break in formant structure.

- A change in the shape of successive pitch periods may be taken as a guide: for example, lengthening or doubling.

Stop release: The start of waveform burst; the first used if multiple bursts are present.

Vowel start/end: The start/end of a pitch period corresponding to appearance/disappearance of a particular formant.

- The shape and magnitude of successive pitch periods may be also used as a guide.

Nasal start/end: The start/end of a pitch period corresponding to the appearance/disappearance of nasal formants on spectrogram.

- Where a nasal adjoins a vowel, the waveform amplitude minimum is usually a reliable criterion.

Fricative start/end: The start/end of continuous spectrogram energy in a particular frequency range. Additional criteria for particular speakers and phonemes include:

- the end/start of a particular preceding or following formant;
- the end/start of spectrogram voicing (for voiceless fricatives);
- a change in amplitude of waveform periodicity.

Certain sequences, such vowel-approximant or approximant-vowel, are particularly difficult to label reliably and are avoided. Where approximants are present in the experimental materials, their start and end points are determined with reference to the adjacent consonant.

Appendix B

Experiment 1: Measured syllable duration by phrase

Test phrase (target syllable underlined)	Target syllable duration (ms)	Number of tokens
bake enforce	254.2	23
bacon force	224.1	22
can inspire	242.5	24
cannon spire	224.0	24
thank fulfil	227.8	21
thankful Phil	199.9	24
cube explain	249.0	23
cubics plane	197.2	24
toe content	186.0	24
token tent	124.9	24
pay perform	186.8	23
paper form	139.2	24
dan surprise	248.5	24
dancer prize	217.4	23
day today	176.3	23
data day	135.1	24

Table B.1: Durations by phrase type: left-headed words; reduced central syllable

Test phrase (target syllable underlined)	Target syllable duration (ms)	Number of tokens
knee capsize	228.4	23
kneecap size	192.7	21
near bisect	269.9	22
nearby sect	249.6	16
there foreclose	194.4	24
therefore close	155.0	22
skim Peking	408.0	22
skimpy king	336.6	23
shake downstairs	334.2	22
shakedown stairs	314.9	22
there foursquare	188.8	23
therefore square	161.6	22
crow barbette	232.8	24
crowbar bet	210.1	24
hard whereby	411.7	21
hardware buy	335.4	22

Table B.2: Durations by phrase type: left-headed words; full central syllable

Test phrase (target syllable underlined)	Target syllable duration (ms)	Number of tokens
bake enforce	311.2	23
bacon force	307.1	21
can inspire	461.9	24
cannon spire	466.3	22
thank fulfil	345.5	21
thankful Phil	351.2	23
cube explain	289.2	24
cubics plane	325.0	21
toe content	371.0	23
token tent	391.6	24
pay perform	326.2	24
paper form	343.0	22
dan surprise	421.8	23
dancer prize	429.0	22
day today	289.0	24
data day	312.6	22

Table B.3: Durations by phrase type: right-headed words; reduced central syllable.

Test phrase (target syllable underlined)	Target syllable duration (ms)	Number of tokens
knee capsize	378.0	22
kneecap size	412.0	19
near bisect	373.5	21
nearby sect	409.1	17
there foreclose	400.5	24
therefore close	408.0	20
skim Peking	355.5	22
skimpy king	372.7	22
shake downstairs	401.7	20
shakedown stairs	409.9	20
there foursquare	434.1	23
therefore square	453.0	23
crow barbette	329.3	23
crowbar bet	344.1	22
hard whereby	360.1	23
hardware buy	379.9	21

Table B.4: Durations by phrase type: right-headed words; full central syllable

Appendix C

Experiment 2: Sentence materials

Left-headed keywords	Right-headed keywords
John THREW the **cap** to the BED again. John THREW the **captain** the BADGE again. John THREW the **captaincy** BADGE again.	JOHN saw Jessica **mend** it AGAIN. JOHN saw Jessie **commend** it AGAIN. JOHN saw Jess **recommend** it AGAIN.
Kate GAVE the **sense** of the SCRIPT away. Kate GAVE the **censor** the SCRIPT again. Kate GAVE the **censorship** SCRIPT away.	BETH saw Clematis **pose** it ALL. BETH saw Clemmie **dispose** it ALL. BETH saw Clem **indispose** it ALL.
Tim KNEW the **dog** may decline AGAIN. Tim KNEW the **dogma** declined AGAIN. Tim KNEW the **dogmatist** line AGAIN.	I CHECKED in every **port** for TOM. I CHECKED the old **report** for TOM. I CHECKED the **misreport** for TOM.
I SAW the **fish** again TODAY. I SAW the **fissure** crack TODAY. I SAW the **fisherman** TODAY.	GREG let big Oprah **juice** it ALL. GREG let Bobbie **produce** it ALL. GREG let Bob **reproduce** it ALL.
I SAW the **mace** unreclaimed AGAIN. I SAW the **mason** reclaimed it ALL. I SAW the **masonry** cleaned AGAIN.	I LET the dancer **pose** it TODAY. I LET the man **suppose** it TODAY. I LET him **presuppose** it TODAY.
Jim LIKES his **part** no more than MOST. Jim LIKES his **partner** more than MOST. Jim LIKES his **partnership** the MOST.	I SAW Widdicombe **pose** it AGAIN. I SAW Jodie **compose** it AGAIN. I SAW Joe **decompose** it AGAIN.
I MADE the **spec** to collect WOOD. I MADE the **spectre** collect WOOD. I MADE the **spectacle** from WOOD.	I MADE Burgundy **send** to them ALL. I MADE Megan **descend** to them ALL. I MADE May **condescend** to them ALL.
I HEARD the **ten** denied AGAIN. I HEARD the **tendon** go AGAIN. I HEARD the **tendency** TODAY.	You MUST continue **main** treatment NOW. You MUST maintain **humane** treatment NOW. You MUST cease **inhumane** treatment NOW.

Table C.1: The full set of experimental sentences for Keyword Series A. Keywords are shown in bold; the words to be emphasized in the unaccented keyword condition are shown in block capitals.

Left-headed keywords	Right-headed keywords
John THREW the **cap** to the BED again. John THREW the **cap** to the big BED again. John THREW the **cap** to the big red BED again.	JOHN saw Jessica **mend** it AGAIN. JOHNNY saw Jessica **mend** it AGAIN. JONATHAN saw Jessica **mend** it AGAIN.
Kate GAVE the **sense** of the SCRIPT away. Kate GAVE the **sense** of the new SCRIPT away. Kate GAVE the **sense** of the latest SCRIPT away.	BETH saw Clematis **pose** it ALL. BETHAN saw Clematis **pose** it ALL. BETHANY saw Clematis **pose** it ALL.
Tim KNEW the **dog** may decline AGAIN. Tim KNEW the **dog** may decline it AGAIN. Tim KNEW the **dog** may decline it all AGAIN.	I CHECKED in every **port** for TOM. I CHECKED one in every **port** for TOM. I CHECKED each one in every **port** for TOM.
I SAW the **fish** again TODAY. I SAW the **fish** again here TODAY. I SAW the **fish** again with it TODAY.	GREG let big Oprah **juice** it ALL. GREGOR let big Oprah **juice** it ALL. GREGORY let big Oprah **juice** it ALL.
I SAW the **mace** unreclaimed AGAIN. I SAW the **mace** unreclaimed once AGAIN. I SAW the **mace** unreclaimed by them AGAIN.	I LET the dancer **pose** it TODAY. I LET the tap dancer **pose** it TODAY. I LET the ballet dancer **pose** it TODAY.
Jim LIKES his **part** no more than MOST. Jim LIKES his **part** there no more than MOST. Jim LIKES his **part** in it no more than MOST.	I SAW Widdicombe **pose** it AGAIN. I SAW Ann Widdicombe **pose** it AGAIN. I SAW Anna Widdicombe **pose** it AGAIN.
I MADE the **spec** to collect WOOD. I MADE the **spec** to collect the WOOD. I MADE the **spec** to collect all the WOOD.	I MADE Burgundy **send** to them ALL. I MADE Pete Burgundy **send** to them ALL. I MADE Peter Burgundy **send** to them ALL.
I HEARD the **ten** denied AGAIN. I HEARD the **ten** denied it AGAIN. I HEARD the **ten** denied it all AGAIN.	You MUST continue **main** treatment NOW. You MUST not continue **main** treatment NOW. You MUST really continue **main** treatment NOW.

Table C.2: The full set of experimental sentences for Keyword Series B. Keywords are shown in bold; the words to be emphasized in the unaccented keyword condition are shown in block capitals. Due to an oversight in the preparation of the experimental materials, the phonetic environment immediately following the test syllable **part** is not constant within the keyword triad, as discussed in Chapter 4, and so this keyword is excluded from these analyses.

Left-headed keywords	Right-headed keywords
John THREW the **cap** to the BED again. John THREW the **captain** to the BED again. John THREW the **captaincy** to the BOARD again.	JOHN saw Jessica **mend** it AGAIN. JOHN saw Alison **commend** it AGAIN. JOHN saw Alison **recommend** it AGAIN.
Kate GAVE the **sense** of the SCRIPT away. Kate GAVE the **censor** for the SCRIPT again. Kate GAVE the **censorship** of the SCRIPT away.	BETH saw Clematis **pose** it ALL. BETH saw Clematis **dispose** it ALL. BETH saw Clematis **indispose** it ALL.
Tim KNEW the **dog** may decline AGAIN. Tim KNEW the **dogma** may decline AGAIN. Tim KNEW the **dogmatist** may decline AGAIN.	I CHECKED in every **port** for TOM. I CHECKED in the old **report** for TOM. I CHECKED in every **misreport** for TOM.
I SAW the **fish** again TODAY. I SAW the **fissure** behind TODAY. I SAW the **fisherman** again TODAY.	GREG let big Oprah **juice** it ALL. GREG let big Oprah **produce** it ALL. GREG let big Olive **reproduce** it ALL.
I SAW the **mace** unreclaimed AGAIN. I SAW the **mason** disinclined it ALL. I SAW the **masonry** disinterred AGAIN.	I LET the dancer **pose** it TODAY. I LET the postman **suppose** it TODAY. I LET the dancer **presuppose** it TODAY.
Jim LIKES his **part** no more than MOST. Jim LIKES his **partner** no more than MOST. Jim LIKES his **partnership** no more than MOST.	I SAW Widdicombe **pose** it AGAIN. I SAW Robinson **compose** it AGAIN. I SAW Robinson **decompose** it AGAIN.
I MADE the **spec** to collect WOOD. I MADE the **spectre** to collect WOOD. I MADE the **spectacle** to collect WOOD.	I MADE Burgundy **send** to them ALL. I MADE Robinson **descend** to them ALL. I MADE Burgundy **condescend** to them ALL.
I HEARD the **ten** denied AGAIN. I HEARD the **tendon** denied AGAIN. I HEARD the **tendency** denied AGAIN.	You MUST continue **main** treatment NOW. You MUST develop **humane** treatment NOW. You MUST abandon **inhumane** treatment NOW.

Table C.3: The full set of experimental sentences for Keyword Series C. Keywords are shown in bold; the words to be emphasized in the unaccented keyword condition are shown in block capitals.

Left-headed keywords	Right-headed keywords
John DESIGNED the **cap**. John DESIGNED the**cap** then. John DESIGNED the **cap** today.	**Mend** it AGAIN for me please. Now **mend** it AGAIN for me please. Will you **mend** it AGAIN for me please.
Kate EXPLAINED the **sense**. Kate EXPLAINED the **sense** well. Kate EXPLAINED the **sense** again.	**Pose** it ALL today. Now **pose** it ALL today. Can you **pose** it ALL today.
Tim thought I KNEW the **dog**. Tim thought I KNEW the **dog** then. Tim thought I KNEW the **dog** again.	**Port** some QUICKLY please. Now **port** some QUICKLY please. Can you **port** some QUICKLY please.
Bob said he SAW the **fish**. Bob said he SAW the **fish** off. Bob said he SAW the **fish** again.	**Juice** it TODAY please. Now **juice** it TODAY please. Can you **juice** it TODAY please.
Albert THREW the **mace**. Albert THREW the **mace** up. Albert THREW the **mace** again.	**Pose** it TODAY please. Now **pose** it TODAY please. Can you **pose** it TODAY please.
I hope Jim FINDS his **part**. I hope Jim FINDS his **part** soon. I hope Jim FINDS his **part** again.	**Pose** it AGAIN for me please. Now **pose** it AGAIN for me please. Will you **pose** it AGAIN for me please.
I FOUND an awful **speck**. I FOUND an awful **speck** there. I FOUND an awful **speck** today.	**Send** it TODAY please. Now **send** it TODAY please. Will you **send** it TODAY please.
Jill OBSERVED the **ten**. Jill OBSERVED the **ten** then. Jill OBSERVED the **ten** today.	**Main** roads SCARE me a lot. The **main** roads SCARE me a lot. All the **main** roads SCARE me a lot.

Table C.4: The full set of experimental sentences for Keyword Series D. Keywords are shown in bold; the words to be emphasized in the unaccented keyword condition are shown in block capitals.

Left-headed keywords	**Right-headed keywords**
John DESIGNED the **cap**. John MALIGNED the **captain**. John RESIGNED the **captaincy**.	**Mend** it AGAIN for me please. **Commend** it AGAIN for me please. **recommend** it AGAIN for me please.
Kate EXPLAINED the **sense**. Kate EXPLAINED the **censor**. Kate EXPLAINED the **censorship**.	**Pose** it ALL today. **Dispose** it ALL today. **Indispose** it ALL today.
Tim thought I KNEW the **dog**. Tim thought I KNEW the **dogma**. Tim thought I KNEW the **dogmatist**.	**Port** some QUICKLY please. **Report** some QUICKLY please. **Misreport** some QUICKLY please.
Bob said he SAW the **fish**. Bob said he SAW the **fissure**. Bob said he SAW the **fisherman**.	**Juice** it TODAY please. **Produce** it TODAY please. **Reproduce** it TODAY please.
Albert THREW the **mace**. Albert THREW the **mason**. Albert THREW the **masonry**.	**Pose** it TODAY please. **Suppose** it TODAY please. **Presuppose** it TODAY please.
I hope Jim FINDS his **part**. I hope Jim FINDS his **partner**. I hope Jim FINDS his **partnership**.	**Pose** it AGAIN for me please. **Compose** it AGAIN for me please. **Decompose** it AGAIN for me please.
I FOUND an awful **spec**. I FOUND an awful **spectre**. I FOUND an awful **spectacle**.	**Send** it TODAY please. **Descend** it TODAY please. **Condescend** it TODAY please.
Jill OBSERVED the **ten**. Jill OBSERVED the **tendon**. Jill OBSERVED the **tendency**.	**Main** roads SCARE me a lot. **Humane** roads SCARE me a lot. **Inhumane** roads SCARE me a lot.

Table C.5: The full set of experimental sentences for Keyword Series E. Keywords are shown in bold; the words to be emphasized in the unaccented keyword condition are shown in block capitals.

Block	Composition	Number of sentences
A	left-headed, utterance-medial, unaccented right-headed, utterance-medial, unaccented	56 56
B	left-headed, utterance-edge, unaccented right-headed, utterance-edge, accented right-headed, utterance-medial, accented	40 40 56
C	left-headed, utterance-edge, accented left-headed, utterance-medial, accented right-headed, utterance-edge, unaccented	40 56 40

Table C.6: Division of experimental sentences in Experiment 2 for the purposes of recording.

Appendix D

Experiment 2: Excluded utterances

Utterances were excluded from Experiment 2 analysis because: the lexical content of the utterance was misread; the keyword received an accent in the unaccented condition; the keyword did not receive an accent in the accented condition; primary lexical stress was misplaced within the keyword; there was an intonational phrase boundary adjacent to the keyword. The excluded utterances are listed in Tables D.1–D.3.

Series	Keyword	Subject	Sentence
A	*compose*	1	⋆ I saw Widdicombe POSE it again
		5	⋆ I SAW Widdicombe pose it AGAIN
		5	I SAW Joe decompose it AGAIN
		5	I saw Joe DECOMPOSE it again
	dispose	1	Beth saw Clem INDISPOSE it all
		3	Beth saw Clem INDISPOSE it all
		6	⋆ Beth saw Clematis POSE it all
	juice	4	Greg let Bob REPRODUCE it all
		6	Greg let Bob REPRODUCE it all
	main	3	You must cease INHUMANE treatment now
	mend	6	JOHN saw Jessie commend it AGAIN
	port	2	I checked the MISREPORT for Tom
		4	I checked the MISREPORT for Tom
B	*compose*	1	⋆ I saw Widdicombe POSE it again
		5	⋆ I SAW Widdicombe pose it AGAIN
	dispose	3	BETHAN saw Clematis pose it ALL
		6	⋆ Beth saw Clematis POSE it all

Table D.1: Experiment 2 utterances for right-headed keywords (part 1) for which pairs of tokens are excluded from analysis. ⋆ indicates that the sentence is common to Series A, B and C.

Series	Keyword	Subject	Sentence
C	*compose*	1	⋆ I saw Widdicombe POSE it again
		5	⋆ I SAW Widdicombe pose it AGAIN
		6	I saw Robinson COMPOSE it again
	dispose	1	Beth saw Clematis INDISPOSE it all
		3	Beth saw Clematis INDISPOSE it all
		6	⋆ Beth saw Clematis POSE it all
		6	BETH saw Clematis dispose it ALL
	juice	4	Greg let big Olive REPRODUCE it all
	port	1	I checked in every MISREPORT for Tom
		2	I checked in every MISREPORT for Tom
		3	I checked in every MISREPORT for Tom
		4	I checked in every MISREPORT for Tom
		6	I checked in every MISREPORT for Tom
	send	6	I MADE Robinson descend to them ALL
D	*dispose*	3	⋆ Pose it ALL today
		3	Now pose it ALL today
		5	Now pose it ALL today
		5	Can you pose it ALL today
	send	6	⋆ Send it TODAY please
		6	Now send it TODAY please
		6	Will you send it TODAY please
E	*dispose*	1	Indispose it ALL today
		5	Indispose it ALL today
		6	Indispose it ALL today
		3	⋆ Pose it ALL today
	juice	4	REPRODUCE it today please
		6	REPRODUCE it today please
	port	4	MISREPORT some quickly please
		2	MISREPORT some quickly please
	send	6	⋆ Send it TODAY please
		6	Descend it TODAY please
		6	Condescend it TODAY please
	suppose	2	PRESUPPOSE it today please

Table D.2: Experiment 2 utterances for right-headed keywords (part 2) for which pairs of tokens are excluded from analysis. ⋆ indicates that the sentence is common to Series A, B and C or common to Series D and E

Series	Keyword	Subject	Sentence
A	*dog*	1	Tim knew the DOGMA declined again
		5	⋆ Tim knew the DOG may decline again
	fish	1	I saw the FISSURE crack today
		5	I saw the FISSURE crack today
	ten	5	⋆ I heard the TEN denied again
B	*dog*	4	Tim knew the DOG may decline it all again
		5	⋆ Tim knew the DOG may decline again
		5	Tim knew the DOG may decline it all again
	fish	5	I saw the FISH again here today
	mace	1	I saw the MACE unreclaimed by them again
		4	I saw the MACE unreclaimed by them again
		5	I SAW the mace unreclaimed once AGAIN
		5	I SAW the mace unreclaimed by them AGAIN
	speck	5	I made the SPEC to collect the wood
	ten	1	I heard the TEN denied it again
		4	I heard the TEN denied it again
		4	I heard the TEN denied it all again
		5	⋆ I heard the TEN denied again
		5	I heard the TEN denied it again
		5	I heard the TEN denied it all again
C	*dog*	5	⋆ Tim knew the DOG may decline again
	fish	1	I saw the FISSURE behind today
		4	I saw the FISSURE behind today
		5	I saw the FISSURE behind today
	mace	1	I saw the MASON disinclined again
		1	I saw the MASONRY disinterred again
		6	I SAW the masonry disinterred AGAIN
	speck	3	I MADE the spectacle to collect WOOD
		5	I made the SPECTACLE to collect wood
	ten	5	⋆ I heard the TEN denied again
D	*dog*	1	Tim thought I KNEW the dog again
	ten	1	Jill observed the TEN today
E	*cap*	1	John RESIGNED the captaincy
	mace	5	Albert THREW the mason

Table D.3: Experiment 2 utterances for left-headed keywords for which pairs of tokens are excluded from analysis. ⋆ indicates that the sentence is common to Series A, B and C.

Appendix E

Experiment 2: Keyword labelling

The placement of labels for test syllable constituents is shown in Tables E.1 and E.2. The placement of labels for other syllables in keywords is shown in Tables E.3 and E.4.

Keyword	Test syllable	**onset**	**aspiration**	**nucleus**	**coda**	**end**
cap	/kæp/	/k/ closure	/k/ release	/æ/ start	/p/ closure	/p/ release
sense	/sɛns/	/s/ start		/ɛ/ start	/n/ start	/s/ end
dog	/dɒg/	/d/ closure		/d/ release	/g/ closure	/g/ release
fish	/fɪʃ/	/f/ closure		/ɪ/ start	/ʃ/ start	/ʃ/ end
mace	/meɪs/	/m/ start		/eɪ/ start	/s/ start	/s/ end
part	/pɑt/	/p/ closure	/p/ release	/ɑ/ start	/t/ closure	/t/ release
speck	/spɛk/	/s/ start		/p/ release	/k/ closure	/k/ release
ten	/tɛn/	/t/ closure	/t/ release	/ɛ/ start	/n/ start	/n/ end

Table E.1: Placement of test syllable labels for the left-headed keywords in Experiment 2. Each subsyllabic constituent is identified by the label (in bold) placed at its start.

Keyword	Test syllable	onset	aspiration	nucleus	coda	end
mend	/mɛnd/	/m/ start		/e/ start	/n/ start	/d/ release
dispose	/pəʊz/	/p/ closure	/p/ release	/əʊ/ start	/z/ start	/z/ end
port	/pɔt/	/p/ closure	/p/ release	/ɔ/ start	/t/ closure	/t/ release
juice	/dʒus/	/dʒ/ closure	/d/ release	/u/ start	/s/ start	/s/ end
suppose	/pəʊz/	/p/ closure	/p/ release	/əʊ/ start	/s/ start	/s/ end
compose	/pəʊz/	/p/ closure	/p/ release	/əʊ/ start	/s/ start	/s/ end
send	/sɛnd/	/s/ start		/ɛ/ start	/n/ start	/d/ release
main	/meɪn/	/m/ start		/eɪ/ start	/n/ start	/n/ end

Table E.2: Placement of test syllable labels for right-headed keywords in Experiment 2. Each subsyllabic constituent is identified by the label (in bold) placed at its start, except for *juice* where the "aspiration" label indicates the start of the fricated part of the affricate /dʒ/.

Keyword	Syllables	Position of label			
		Syllable-2 start	**Syllable-2 end**	**Syllable-3 start**	**Syllable-3 end**
cap	/kæp.tən.sɪ/	/t/ release	/n/ end		/ɪ/ end
sense	/sɛns.ə.ʃɪp/	/ə/ start	/ə/ end		/p/ release
dog	/dɒg.mə.tɪst/	/m/ start	/ə/ end		/t/ closure
fish	/fɪʃ.ə.mən/	/ʃ/ end	/ə/ end		/n/ end
mace	/meɪs.ən.ɹɪ/	/s/ end	/n/ end		/ɪ/ end
part	/pɑt.nə.ʃɪp/	/n/ start	/ə/ end		/p/ closure
spec	/spɛk.tə.kəl/	/t/ release	/ə/ end		/l/ end
ten	/tɛn.dən.sɪ/	/d/ closure	/n/ end		/ɪ/ end

Table E.3: Placement of additional syllable labels for the left-headed keywords in Experiment 2: syllable-2 immediately follows the primary stressed syllable in the disyllable and trisyllable; syllable-3 follows syllable-2 in the trisyllable. The test-syllable-final stops in *spec.ta.cle* and *cap.tain.cy* are not always released, thus the closure durations of the syllable-2 onsets are not reliably measurable and so the syllables are measured from the syllable-2 stop release. The syllable-3 final stops are not always released in *dogmatist* and *partnership* and so the syllable is measured up to the onset of stop closure.

<table>
<tr><th rowspan="2">Keyword</th><th rowspan="2">Syllables</th><th colspan="4">Position of label</th></tr>
<tr><th>Syllable-3 start</th><th>Syllable-3 end</th><th>Syllable-2 start</th><th>Syllable-2 end</th></tr>
<tr><td>mend</td><td>/ɹɛ.kə.mɛnd/</td><td>/ɹ/ start</td><td colspan="2">/k/ closure</td><td>/m/ start</td></tr>
<tr><td>dispose</td><td>/ɪn.dɪs.pəʊz/</td><td>/ɪ/ start</td><td colspan="2">/d/ closure</td><td>/s/ end</td></tr>
<tr><td>port</td><td>/mɪs.ɹə.pɔt/</td><td>/m/ start</td><td colspan="2">/ɹ/ start</td><td>/p/ closure</td></tr>
<tr><td>juice</td><td>/ɹi.pɹə.dʒus/</td><td>/ɹ/ start</td><td colspan="2">/p/ closure</td><td>/d/ closure</td></tr>
<tr><td>suppose</td><td>/pɹi.sə.pəʊz/</td><td>/p/ closure</td><td colspan="2">/s/ start</td><td>/p/ closure</td></tr>
<tr><td>compose</td><td>/di.kəm.pəʊz/</td><td>/d/ closure</td><td colspan="2">/k/ closure</td><td>/p/ closure</td></tr>
<tr><td>send</td><td>/kɒn.də.sɛnd/</td><td>/k/ closure</td><td colspan="2">/d/ closure</td><td>/s/ start</td></tr>
<tr><td>main</td><td>/ɪn.hju.meɪn/</td><td>/ɪ/ start</td><td colspan="2">/n/ end</td><td>/m/ start</td></tr>
</table>

Table E.4: Placement of additional syllable labels for the right-headed keywords in Experiment 2: syllable-2 immediately precedes the primary stressed syllable in the disyllable and trisyllable; syllable-3 precedes syllable-2 in the trisyllable.

Appendix F

Experiment 2: Results overview

	Number of added syllables: … in utterance	… in word: 0		1		2	
Utterance-medial	0	232	**300**	219	**269**	212	**251**
	1	231	**299**	219	**270**		
	2	234	**298**			211	**242**
Utterance-edge	0	216	**278**				
	1	232	**296**	235	**274**		
	2	233	**300**			225	**265**

Table F.1: Mean test syllable duration (ms) in Experiment 2 for the right-headed keywords *main*, *mend* and *send*, on the left in each cell for unaccented keywords and in bold on the right in each cell for accented keywords.

	Number of added syllables: ... in utterance	... in word: 0		1		2	
Utterance-medial	0	293	**349**	286	**331**	279	**306**
	1	291	**349**	286	**327**		
	2	290	**344**			276	**302**
Utterance-edge	0	410	**453**				
	1	318	**357**	319	**339**		
	2	294	**338**			286	**309**

Table F.2: Mean test syllable duration (ms) in Experiment 2 for the left-headed keywords *fish*, *mace*, *sense* and *ten*, on the left in each cell for unaccented keywords and in bold on the right in each cell for accented keywords.

Appendix G

Experiment 2: Additional utterance-span results

There is a comparison in Experiment 2, reported in Section 4.5.1, of test syllable duration between those Series B and Series D sentences which differ in length by three or four syllables, intended to determine the effect of utterance length on stressed syllable duration. The results suggest domain-edge rather than domain-span interpretations: there is no significant durational variation on the test syllable nucleus, but the onset and coda manifest contrasting durational effects. The variations in test syllable onset and coda durations are considered here across the full set of keywords, regardless of utterance length.

Test syllable onset duration

In order to see if the pattern of differences in onset duration between Series B vs Series D relates to the likelihood of phrase boundary occurrence, the full set of keywords are examined. Table G.1 shows the mean durational difference which, as there is no interaction between utterance length and pitch accent, is pooled across accented and unaccented conditions.

The most noticeable trend in Table G.1 is that substantial lengthening of the onset in Series B is more widespread for right-headed keywords. A By-Subjects analysis finds a significant effect of Utterance Position for all right-headed keywords: $F(1,5) = 10.78$, $p < .05$ [by Items: $F(1,7) = 6.34$, $p < .05$]. Where right-headed keywords are preceded by a noun phrase/verb phrase boundary in the Series B sentence, a prosodic phrase boundaries may sometimes be realised. A prosodic phrase boundary seems less likely in the two Series B utterances which have a different syntactic environment for the keyword—*main* is within the verb phrase headed by the preceding verb; *port* is preceded by a noun-phrase-internal word boundary—and these two keywords do

Series B.3 utterance	Series D.3 utterance	Lengthening
Right-headed keywords		
I LET the ballet dancer **pose** it TODAY.	Can you **pose** it TODAY please.	17%
GREGORY let big Oprah **juice** it ALL.	Can you **juice** it TODAY please.	11%
JONATHAN saw Jessica **mend** it AGAIN.	Will you **mend** it AGAIN for me please.	10%
BETHANY saw Clematis **pose** it ALL.	Can you **pose** it ALL today.	10%
I MADE Peter Burgundy **send** to them ALL.	Will you **send** it TODAY please.	8%
You MUST really continue **main** treatment NOW.	All the **main** roads SCARE me a lot.	5%
I CHECKED each one in every **port** for TOM.	Can you **port** some QUICKLY please.	4%
I SAW Anna Widdicombe **pose** it AGAIN.	Will you **pose** it AGAIN for me please.	-10%
Left-headed keywords		
Tim KNEW the **dog** may decline it all AGAIN.	Tim thought I KNEW the **dog** again.	10%
I SAW the **mace** unreclaimed by them AGAIN.	Albert THREW the **mace** again.	9%
John THREW the **cap** to the big red BED again.	John DESIGNED the **cap** today.	4%
Kate GAVE the **sense** of the latest SCRIPT away.	Kate EXPLAINED the **sense** again.	4%
Jim LIKES his **part** in it no more than MOST.	I hope Jim FINDS his **part** again.	4%
I SAW the **fish** again with it TODAY.	Bob said he SAW the **fish** again.	2%
I HEARD the **ten** denied it all AGAIN.	Jill OBSERVED the **ten** today.	2%
I MADE the **spec** to collect all the WOOD.	I FOUND an awful **speck** today.	-4%

Table G.1: Percentage lengthening of test syllable onset in Keyword Series B.3 utterances compared with Keyword Series D.3 utterances. The test syllable is in bold; words emphasised in the unaccented condition are in capitals.

show less lengthening of the onset in Series B.

Variations in segmental environment are also a factor in these results. For most right-headed keywords, the test syllable is preceded by a vowel in both Series B and Series D utterances; for *compose*, however, the test syllable is preceded by a nasal consonant in Series B (*Widdicombe pose*) and a vowel in Series D (*Will you pose*), and it has a longer syllable onset in Series D. One possibility is that the /p/ is shortened in Series B because it is preceded by /m/: it is not clear whether consonants must be tautosyllabic for shortening in consonant clusters to occur. *Clematis pose* in Series B similarly has a test syllable onset /p/ preceded by another consonant, in contrast with Series D where it is preceded by a vowel. In this case, however, the shortening seen in *Widdicombe pose* is not observed.

For the left-headed keywords in Table G.1, all the test syllable onsets are noun-phrase-internal, and thus the likelihood of a phrase boundary seems low, although possibly more likely in the longer utterance. The small differences between Series B and Series D reflect this, except for the utterances for *dog* and *mace*, which do not appear to have any structural reasons for the more likely placement of phrase boundaries preceding the test syllable in Series B[1]. A By-Subjects analysis finds a significant effect of Utterance Position on onset duration: $F(1,5.1) = 10.72$, $p < .05$ [by Items: $F(1,7.2) = 5.94$, $p < .05$] for all left-headed keywords shown in Table Table G.1.

[1]The likely full vowel syllable in *dog may* could, however, cause lengthening in comparison to *dog again*.

One possible factor regarding the placement of boundaries in read speech is that subjects may pay less attention to syntax in their prosodic phrasing than when generating utterances spontaneously: factors such as utterance length and the relative lengths of potential phrases may have more weight in the reading task, where deep levels of linguistic processing may have less input into prosodic planning. In the extreme case, subjects reading lists of items tend to break the list into phrases containing equal numbers of words (Gee & Grosjean 1983). The reading of linguistically-meaningful written sentences may lie, in performance terms, between list-reading and natural speech: in the latter, where non-hesitation breaks are highly likely to delimit structurally-significant units and factors such as length and speech rate may affect how many such units are realised phrasally. In all cases, the occurrence of a phrase boundary at a given point in a particular utterance is a matter of probability: syntactically more important boundaries are more likely to be realised prosodically.

Overall, it seems that most right-headed keywords in Series B have syntactic structures which may induce prosodic boundaries and associated lengthening preceding the keyword, but such boundaries are less likely in Series D. As keyword-adjacent intonational phrase boundaries are excluded from the final data-set, the phrases associated with these boundaries will be below the intonational phrase: for example, the phonological phrase. The results for left-headed keywords are considered below in relation to variations in coda duration.

Test syllable coda duration

Test syllable codas show the opposite pattern to onsets in the comparison between Series B and D, being longer in the shorter Series D utterances; as discussed in Section 4.5.1, however, this effect does not seem likely to be interpretable as an utterance-span effect. To consider the interpretation of these results further, mean coda durations are shown in Table G.2 for all keywords, excluding those with frequently glottalised codas. Because there is no interaction between utterance length and pitch accent, the mean durational difference is pooled across accented and unaccented conditions.

A By-Subjects analysis finds that the effect of Utterance Position approaches significance for left-headed keywords: $F(1,5.1) = 4.22$, $p = .094$ [by Items: $F(1,4) = 2.85$, $p = .166$]. The explanation for the greater duration of test syllable codas in left-headed keywords in Series D seems likely to be their utterance-antepenultimate position, particularly as no other parts of the syllable show such lengthening. The only keyword which does not show this pattern is *ten*, which does not have a consistent phonetic environment between the two series.

In accordance with the word-level results presented in Section 4.6, it may be noted that there is no evidence of left-headed keywords in utterance-medial position (Series

Series B utterance	Series D utterance	Lengthening
Right-headed keywords		
I MADE Peter Burgundy **send** to them ALL.	Will you **send** it TODAY please.	22%
You MUST really continue **main** treatment NOW.	All the **main** roads SCARE me a lot.	14%
I SAW Anna Widdicombe **pose** it AGAIN.	Will you **pose** it AGAIN for me please.	4%
GREGORY let big Oprah **juice** it ALL.	Can you **juice** it TODAY please.	4%
BETHANY saw Clematis **pose** it ALL.	Can you **pose** it ALL today.	3%
I LET the ballet dancer **pose** it TODAY.	Can you **pose** it TODAY please.	2%
JONATHAN saw Jessica **mend** it AGAIN.	Will you **mend** it AGAIN for me please.	-8%
Left-headed keywords		
I SAW the **mace** unreclaimed by them AGAIN.	Albert THREW the **mace** again.	17%
Kate GAVE the **sense** of the latest SCRIPT away.	Kate EXPLAINED the **sense** again.	11%
John THREW the **cap** to the big red BED again.	John DESIGNED the **cap** today.	9%
I SAW the **fish** again with it TODAY.	Bob said he SAW the **fish** again.	3%
I HEARD the **ten** denied it all AGAIN.	Jill OBSERVED the **ten** today.	-7%

Table G.2: Percentage lengthening of test syllable coda in Keyword Series D.3 utterances compared with Keyword Series B.3 utterances, excluding keywords *port*, *speck*, *dog* and *part*. The test syllable is in bold; words emphasised in the unaccented condition are in capitals.

B) manifesting phrase-final lengthening effects: this includes those in the Series B.1 utterances, for example in Table 4.12, which are also used in Series A and Series C to examine word-level effects.

The interpretation of the results for right-headed keywords shown in Table G.2 as phrase-final lengthening is not compelling, as the syntactic structures of the Series D utterances do not appear likely to result in prosodic boundaries immediately following the test syllable[2]. The results for onset duration discussed above suggest, however, that there is sometimes a phrase boundary preceding the right-headed keyword in Series B. The shortening of the codas in Series B could be interpreted as a compensatory adjustment to the phrase-initial lengthening of their onsets. A similar effect could account for the apparently anomalous lengthening of onset duration for Series B left-headed keywords compared with Series D: the codas appear to be lengthening due to their utterance-antepenultimate position in Series D, and thus the onsets may undergo some compensatory shortening. The tentative explanation is supported by the apparent occurrence of another compensatory effect in Experiment 2: the possible shortening of the syllable nucleus following word-initial lengthening. Compensatory effects are discussed in more detail in Chapter 5.

The effect in right-headed codas is only small, however, and a By-Subjects analysis of coda duration across all keywords except *port* finds a no significant difference

[2]It may be noted that the difference between coda duration in Series B and Series D may be affected for the keywords *dispose* (as in *Clematis pose*) and *juice* by the utterance-antepenultimate position of the keywords in Series B. As just noted for left-headed keywords, this position appears to cause some lengthening of the coda: thus, a tendency for some other reason towards lengthening of the coda in Series D may be masked.

between Series B and Series D: $F(1,5) = 2.43$, $p = .180$ [by Items: $F(1,6) = 2.27$, $p = .182$]. It may be noted, furthermore, that only two keywords, *send* and *main*, show a strong effect in this comparison, and both have differences in the following phonetic environment between Series B and Series D: there may be shortening of consonants in clusters in Series B for *send* and *main*, although it is not clear if this effect is reliable across syllable boundaries. (In contrast, as noted above, the Series D lengthening effect may be underestimated in the comparison for *dispose* and *juice*.). Indeed, the variability in the pattern of results for both onsets and codas in the Series B vs Series D comparisons indicates the importance of experimentally controlling influences on segmental duration other than those pertaining to the experimental conditions.

Appendix H

Experiment 2: Additional word-edge and word-span results

	Number of syllables in word		
	1	2	3
Onset			
Unaccented	95	77	76
Accented	125	95	79
Vowel nucleus			
Unaccented	98	102	98
Accented	116	119	110
Coda			
Unaccented	69	69	66
Accented	85	82	77

Table H.1: Mean duration (ms) of the subsyllabic constituents of the test syllable for right-headed keywords in Series C. The data-set for onset and nucleus is smaller than that for Series A because here the keyword *port* is excluded due to a cluster of missing measurements. The data-set for coda is equivalent to that for Series A, where *port* was also excluded. The means for onset, nucleus and coda duration are taken from all the keywords except *port*: data points = 252; missing = 3.6%.

	By Subjects analysis			By Items analysis		
Source of variation	Degrees of freedom	*F Ratio*	Significance level	Degrees of freedom	*F Ratio*	Significance level
Onset						
Word Length	2,10	52.69	$p < .001$	2,12.1	64.39	$p < .001$
Accent	1,5	24.27	$p < .005$	1,6.1	80.56	$p < .001$
Interaction	2,10.1	21.64	$p < .001$	2,12.7	64.67	$p < .001$
Nucleus						
Word Length	2,10.2	5.79	$p < .05$	2,12.2	3.70	$p = .055$
Accent	1,5	10.60	$p < .05$	1,6	19.98	$p < .005$
Interaction	2,10.2	1.63	$p = .244$	2,12.4	1.13	$p = .354$
Coda						
Word Length	2,10.2	10.01	$p < .005$	2,12.8	10.98	$p < .005$
Accent	1,5	28.03	$p < .005$	1,6	10.34	$p < .05$
Interaction	2,10.2	2.49	$p = .131$	2,12.85	2.05	$p = .169$

Table H.2: Results of analyses of variance for subsyllabic durations in right-headed keyword test syllables in Series C.

	Number of syllables in word		
	1	2	3
Onset			
Unaccented	112	106	104
Accented	133	123	120
Nucleus			
Unaccented	79	74	66
Accented	89	78	73
Coda			
Unaccented	104	107	105
Accented	131	123	113

Table H.3: Mean duration (ms) of the subsyllabic constituents of the test syllables for left-headed keywords in Series C. The means for onset and nucleus duration are taken from all the keywords excluding *dog* and *part*: data points = 216; missing = 4.2%. The means for coda duration are taken from the keywords *fish*, *mace*, *sense* and *ten*: data points = 144; missing = 4.9%.

	By Subjects analysis			By Items analysis		
Source of variation	Degrees of freedom	*F Ratio*	Significance level	Degrees of freedom	*F Ratio*	Significance level
Onset						
Word Length	2,10.4	7.95	$p < .01$	2,10.1	10.53	$p < .005$
Accent	1,5	8.54	$p < .05$	1,5.1	145.59	$p < .001$
Interaction	2,10.5	0.30	$p = .744$	2,10.6	5.88	$p < .05$
Nucleus						
Word Length	2,10.8	36.25	$p < .001$	2,10.1	27.24	$p < .001$
Accent	1,5.1	12.57	$p < .05$	1,5.1	58.08	$p < .005$
Interaction	2,12.5	5.65	$p < .05$	2,10.2	1.51	$p = .267$
Coda						
Word Length	2,11.2	3.23	$p = .078$	2,6	0.97	$p = .433$
Accent	1,5.1	5.33	$p = .068$	1,3	16.67	$p < .05$
Interaction	2,12.2	6.19	$p < .05$	2,6.1	3.00	$p = .124$

Table H.4: Results of analyses of variance for subsyllabic durations in left-headed key-word test syllables in Series C.

Appendix I

Experiment 2: Additional utterance-final results

Keyword	Utterance position			
	Final		Medial	
captain	160	**166**	153	**167**
dogma	106	**118**	132	**122**
fissure	88	**85**	44	**44**
mason	109	**111**	94	**102**
partner	113	**115**	94	**97**
censor	77	**83**	35	**39**
spectre	119	**123**	78	**83**
tendon	151	**157**	166	**182**

Table I.1: Mean final unstressed syllable duration (ms) for left-headed disyllabic keywords Series C and Series E: data points = 24 for each keyword; missing = 0 for all keywords, except *fissure* (3 missing) and *mace* (2 missing). In each cell, durations on the left are from unaccented context; durations in bold on the right are from accented context.

Keyword	Utterance position			
	Final		Medial	
captaincy	135	**148**	122	**135**
dogmatist	61	**69**	69	**69**
fisherman	29	**35**	31	**31**
masonry	84	**89**	66	**68**
partnership	51	**62**	56	**58**
censorship	34	**30**	31	**31**
spectacle	59	**59**	51	**55**
tendency	143	**137**	109	**133**

Table I.2: Mean penultimate unstressed syllable duration (ms) for left-headed trisyllabic keywords Series C and Series E: data points = 24 for each keyword; missing = 0 for all keywords, except *captaincy* (1 missing), *mace* (2 missing) and *speck* (2 missing). In each cell, durations on the left are from unaccented context; durations in bold on the right are from accented context.

Appendix J

Experiment 2: Additional pitch accent results

	Degrees of freedom		F Ratio		Significance level	
Right-headed						
Onset	1,5	(1,7.1)	26.64	(398.34)	$p < .005$	($p < .001$)
Nucleus	1,5	(1,7)	16.38	(19.34)	$p < .01$	($p < .005$)
Coda	1,5	(1,6)	35.98	(15.21)	$p < .005$	($p < .01$)
Left-headed						
Onset	1,5	(1,7.1)	11.92	(223.06)	$p < .05$	($p < .001$)
Nucleus	1,5.1	(1,7)	13.32	(22.96)	$p < .05$	($p < .005$)
Coda	1,5	(1,2)	5.57	(5.16)	$p = .065$	(NS)

Table J.1: Results of analyses of variance for the effect of Accent on the duration of subsyllabic constituents of the test syllable for monosyllabic keywords in utterance-medial keyword series. In each cell, the results of By-Subjects analyses are on the left and the results of By-Items analyses are in parentheses on the right. The data-sets used are described in the tables of mean duration in Section 4.9.1. NS = not significant, thus $p > .05$.

	Degrees of freedom	F Ratio	Significance level
Right-headed			
Onset	1,5 (1,7)	21.25 (50.07)	$p < .01$ ($p < .001$)
Nucleus	1,5 (1,7)	12.67 (24.12)	$p < .05$ ($p < .005$)
Coda	1,5 (1,6)	14.70 (8.81)	$p < .05$ ($p < .05$)
Left-headed			
Onset	1,5 (1,7.5)	10.37 (340.53)	$p < .05$ ($p < .001$)
Nucleus	1,5.1 (1,7.1)	4.90 (6.39)	$p = .077$ ($p < .05$)
Coda	1,5.2 (1,3)	7.68 (11.39)	$p < .05$ ($p < .05$)

Table J.2: Results of analyses of variance for the effect of Accent on the duration of sub-syllabic constituents of the test syllable for disyllabic keywords in utterance-medial keyword series. In each cell, the results of By-Subjects analyses are on the left and the results of By-Items analyses are in parentheses on the right. The data-sets used are described in the tables of mean duration in Section 4.9.1. NS = not significant, thus $p > .05$.

	Degrees of freedom	F Ratio	Significance level
Right-headed			
Onset	1,5 (1,6.1)	6.28 (5.18)	$p = .054$ ($p = .062$)
Nucleus	1,5 (1,6)	5.84 (14.55)	$p = .060$ ($p < .01$)
Coda	1,5 (1,6)	37.39 (5.84)	$p < .005$ ($p = .052$)
Left-headed			
Onset	1,5 (1,7.1)	9.99 (113.90)	$p < .05$ ($p < .001$)
Nucleus	1,5 (1,7.1)	5.31 (16.23)	$p = .069$ ($p < .01$)
Coda	1,5 (1,3.1)	4.47 (13.93)	$p = .088$ ($p < .05$)

Table J.3: Results of analyses of variance for the effect of Accent on the duration of sub-syllabic constituents of the test syllable for trisyllabic keywords in utterance-medial keyword series. In each cell, the results of By Subjects analyses are on the left and the results of By Items analyses are in parentheses on the right. The data-sets used are described in the tables of mean duration in Section 4.9.1. NS = not significant, thus $p > .05$.