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 <u>dib</u> 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.

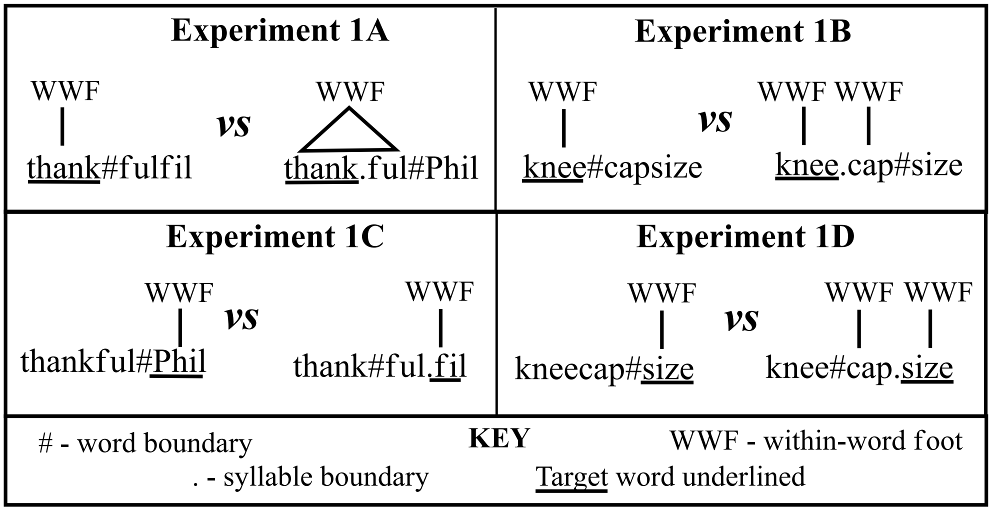

Figure 3.1: The constituent structure of the materials. Example phrases are shown for each experiment and the test syllables are underlined in each case.

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 "<u>BAKE</u> enforce", don't say "TANK enforce".

Final accent Say "bake EN<u>FORCE</u>", don't say "bake REMOVE".

No accent SAY "<u>bake</u> en<u>force</u>", 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

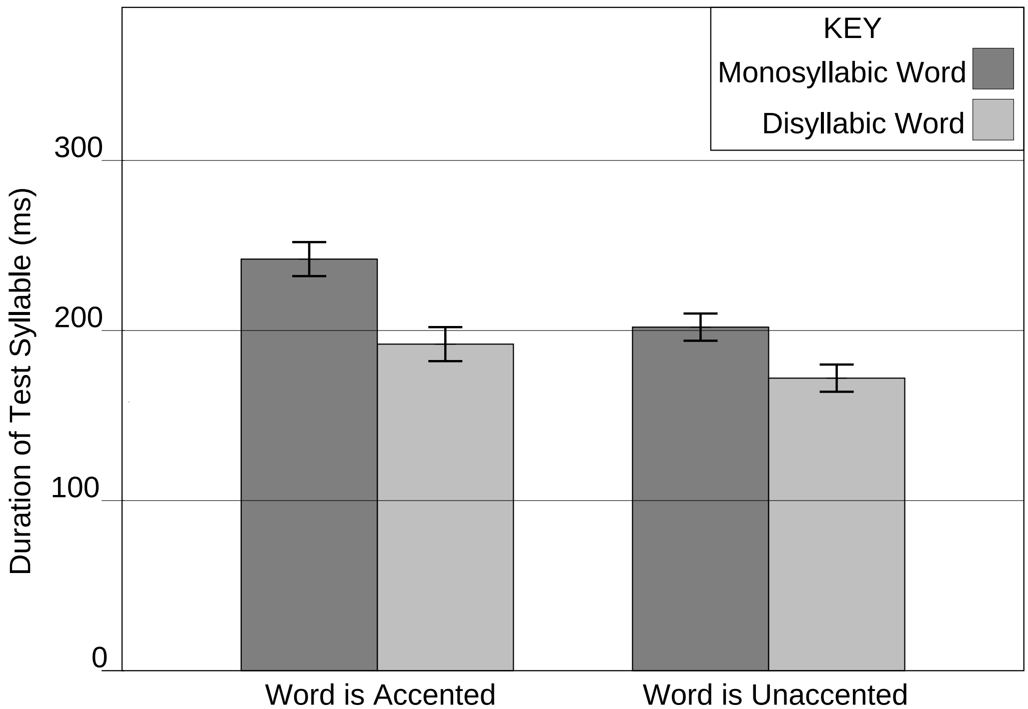

Figure 3.2: Mean test syllable duration (ms) in Experiment 1A: left headed words; reduced central syllable.

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.

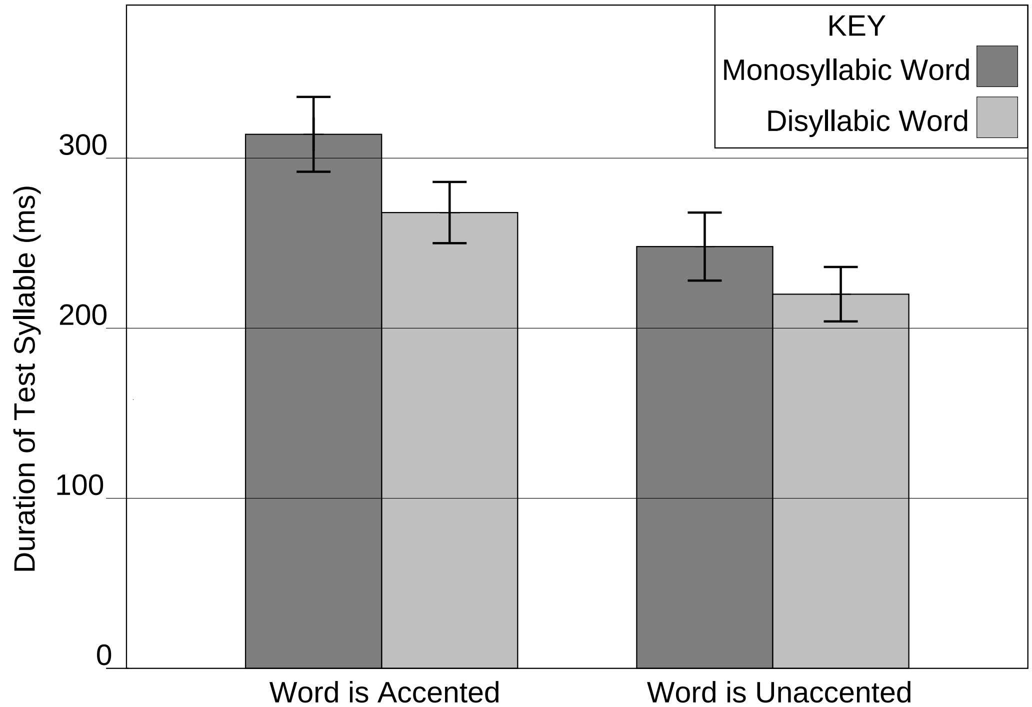

Figure 3.3: Mean test syllable duration (ms) in Experiment 1B: left-headed words; full central syllable.

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".

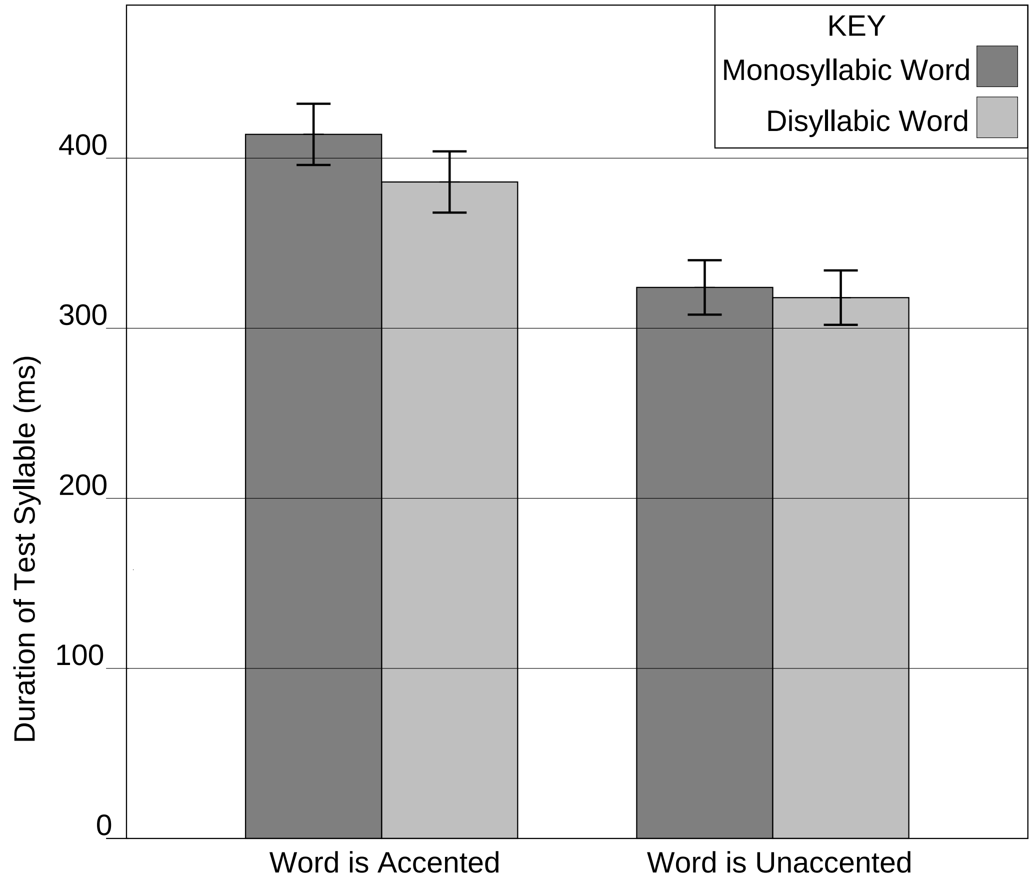

Figure 3.4: Mean test syllable duration (ms) in Experiment 1C: right-headed words; reduced central syllable.

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-*

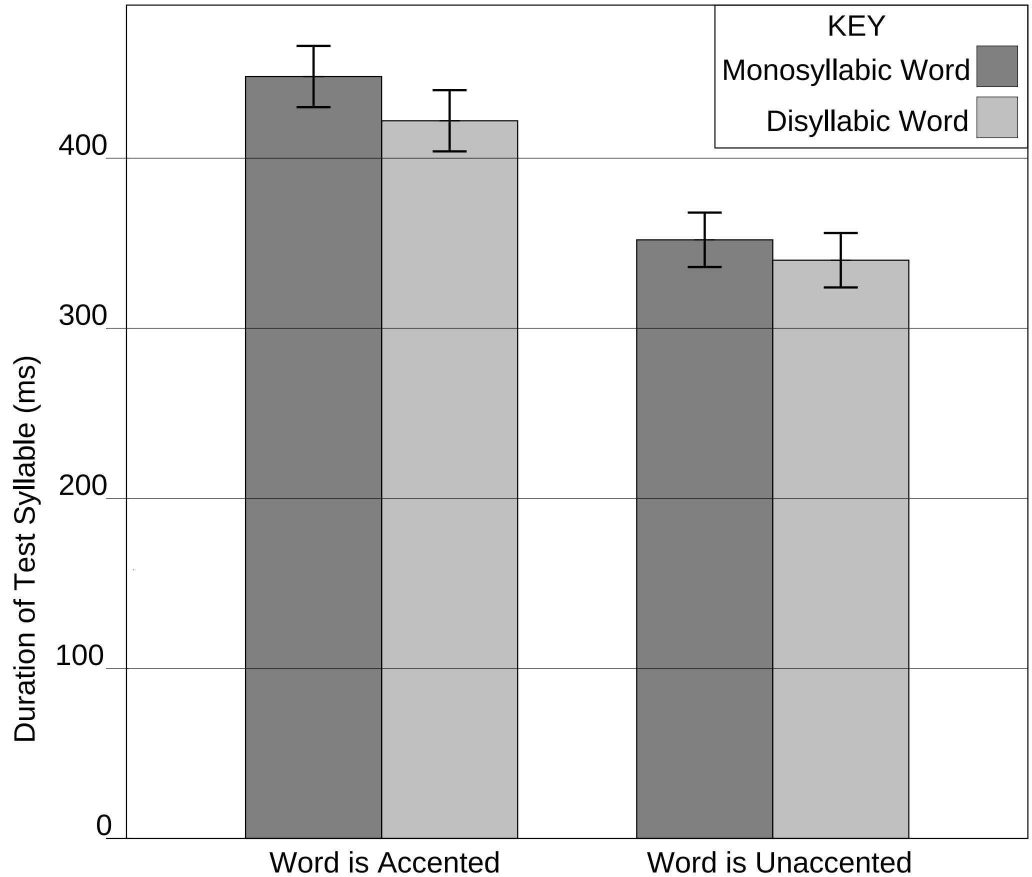

Figure 3.5: Mean test syllable duration (ms) in Experiment 1D: right-headed words; full central syllable.

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

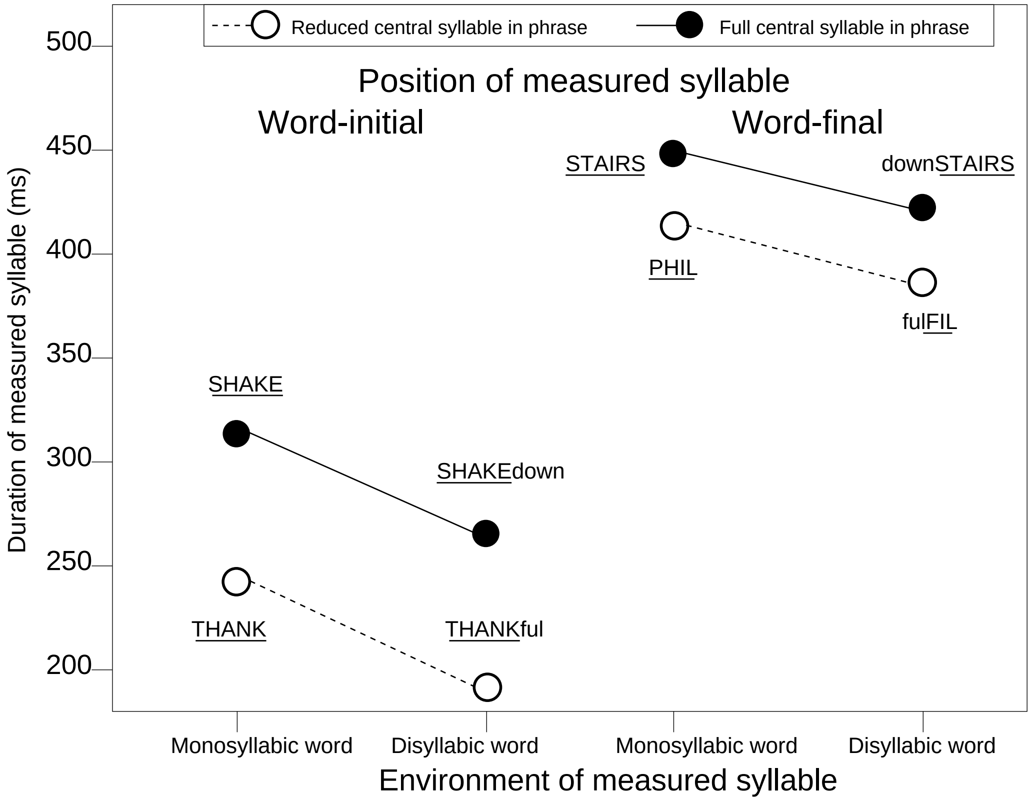

Figure 3.6: Mean test syllable duration (ms) in the accented condition. Example words are given for each mean duration and the test syllable is underlined.

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

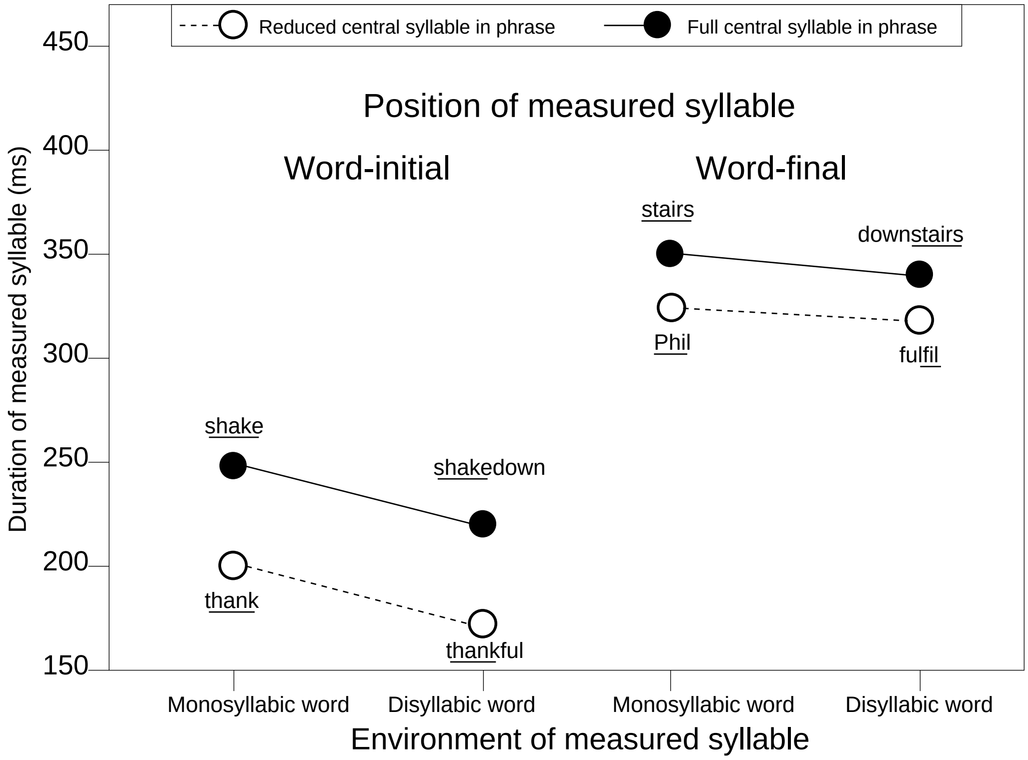

Figure 3.7: Mean test syllable duration (ms) in the unaccented condition. Example words are given for each mean duration and the test syllable is underlined.

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.